



RESEARCH ARTICLE

Heart Disease Prediction with Machine Learning Algorithm Using Feature Selection by Genetic Algorithm

¹ Berna Vatansver, ² Hakan Aydın, * Ali Çetinkaya

^{1,2} Istanbul Gelisim University, Computer Eng., Istanbul, Turkey
 bernavatansver1999@hotmail.com orcid.0000-0001-6441-6660
 haaydin@gelisim.edu.tr orcid.0000-0002-0122-8512

*Istanbul Gelisim University, TTO UAM R&D Center, Istanbul, Turkey
 alcetinkaya@gelisim.edu.tr orcid.0000-0003-4535-3953

HIGHLIGHTS

- Predicting heart disease by analyzing the factors that affect it.
- Utilizing Machine Learning (ML) algorithms by selecting dataset features for prediction processes
- Demonstrating that the prediction results obtained after feature selection using the Genetic Algorithm (GA) method and machine learning methods have a higher success rate.

Keywords:

- Machine Learning Algorithms
- Genetic Algorithms
- Heart Disease Prediction
- Feature Selection

Article Info:

Received : 07.10.2021
 Accepted : 06.11.2021
 Published : 21.12.2021

DOI:

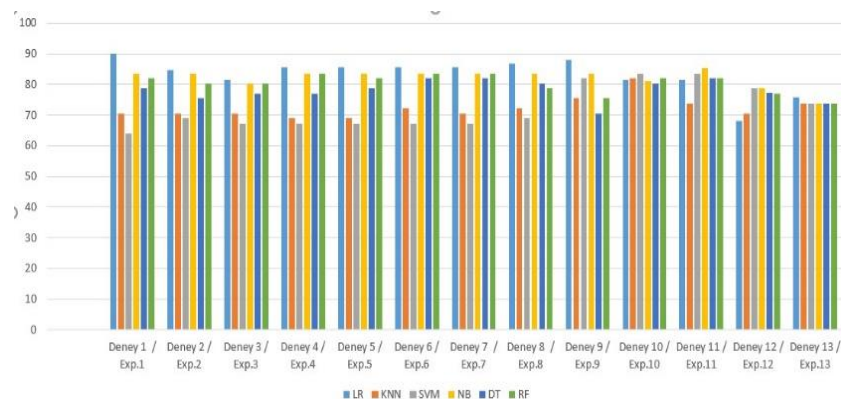
10.53525/jster.1005934

*Correspondence:

Ali Çetinkaya
alacetinkaya@gelisim.edu.tr
 +90 212 422 70 00 / 7187

GRAPHICAL ABSTRACT

The aim of this study was determined as the prediction of heart disease by analyzing the factors that affect this disease with Machine Learning (ML) algorithms using the feature selection with the Genetic Algorithm (GA) approach. In this context, K-Nearest Neighbor (KNN), Logistic Regression (LR), Decision Tree (DT), Random Forest (RF), Naive Bayes (NB) and Support Vector Machine (SVM) algorithms were used in the study. The graph of the success rates obtained in the experiments conducted in the first group using these ML algorithms within the scope of the study is presented below.



Aim of Article : The purpose of this study is to predict heart disease with ML algorithms using feature selection with GA approach.

Theory and Methodology : In this study, ML methods and GA are used.

Findings and Results: In the study, the importance of feature selection with GA was emphasized by proving that the accuracy of heart disease prediction was higher after feature selection was made with GA in the data set.

Conclusion : As a result of the study, it has been shown that heart disease can be predicted with higher success rate and accuracy with ML algorithms using feature selection with the GA approach.



Genetik Algoritma Yaklaşımıyla Öznitelik Seçimi Kullanılarak Makine Öğrenmesi Algoritmaları ile Kalp Hastalığı Tahmini

¹ Berna Vatansver, ² Hakan Aydın, * Ali Çetinkaya

^{1,2} İstanbul Gelişim Üniversitesi, Bilgisayar Müh. Bölümü, İstanbul, Türkiye
bernavatansver1999@hotmail.com, orcid.0000-0001-6441-6660,
haaydin@gelisim.edu.tr orcid.0000-0002-0122-8512

*İstanbul Gelişim Üniversitesi, TTO UAM R&D Center, İstanbul, Türkiye
alacetinkaya@gelisim.edu.tr orcid.0000-0003-4535-3953

Alıntı / Citation :

Vatansver, B., Aydın, H., Çetinkaya, A. (2021). Genetik Algoritma Yaklaşımıyla Öznitelik Seçimi Kullanılarak Makine Öğrenmesi Algoritmaları ile Kalp Hastalığı Tahmini, *Journal of Scientific Technology and Engineering Research*, 2(2):67-80. DOI: 10.53525/jster.1005934

ÖNE ÇIKANLAR / HIGHLIGHTS

- Kalp hastalığını etkileyen faktör ve özelliklerin analiz edilerek bu hastalığın tahmin edilmesi
- Geleceğe dönük tahmin işlemleri için veri seti öznitelik seçimi yapılarak Makine Öğrenmesi (MÖ) algoritmalarından yararlanılması
- Genetik Algoritma (GA) metodu ve makine öğrenmesi yöntemleri kullanılarak öznitelik seçimi yapıldıktan sonra elde edilen tahmin sonuçlarının daha yüksek başarı oranına sahip olduğunun gösterilmesi

Makale Bilgileri

Geliş Tarihi : 07.10.2021
Kabul Tarihi : 06.11.2021
Yayın Tarihi : 21.12.2021

DOI:

10.53525/jster.1005934

*Sorumlu Yazar:

Ali Çetinkaya
alacetinkaya@gelisim.edu.tr
+90 212 422 70 00 / 7187

ÖZET / ABSTRACT

Günümüzde Yapay Zeka (YZ) uygulamalarından sağlık sektöründe pek çok hastalığın teşhis ve tedavisinde giderek artan oranlarda faydalanılmaktadır. Bu çalışmanın amacı Genetik Algoritma (GA) yaklaşımı ile öznitelik seçimi yaparak Makine Öğrenmesi (MÖ) algoritmalarıyla kalp hastalığını tahmin etmek olarak belirlenmiştir. Bu kapsamda çalışmada MÖ algoritmalarından K-En Yakın Komşu (K-EYK), Lojistik Regresyon (LR), Karar Ağacı (KA), Rastgele Orman (RO), Naive Bayes (NB) ve Destek Vektör Makinesi (DVM) algoritmaları ile 3 (üç) farklı grupta toplamda 28 (yirmi sekiz) deney gerçekleştirilmiştir. Çalışma sonucunda en yüksek başarı oranı GA ile yapılan deneylerde %93.44 ile elde edilmiştir. Çalışmamızda yapılan diğer deneylerde ise LR ile %90.16, K-EYK ile %81.96, DVM ile %83.60, NB ile %85.24, KA ile %81.96 ve RO ile %83.60 oranında başarı oranı elde edilmiştir. Çalışmada ABD Cleveland eyaletindeki bireylerin verisini içeren veri seti kullanılmıştır. Araştırmamızda elde edilen sonuçlarda kalp hastalığı tahmininin GA yaklaşımı ile öznitelik seçimi yapılması durumunda daha yüksek doğruluk oranının elde edildiği görülmüştür. Çalışmamızın GA ile öznitelik seçimi yapılarak MÖ ile kalp hastalığının tahmin edilmesinde sağlık alanındaki literatüre katkı sağlayacağı değerlendirilmektedir.

Anahtar Kelimeler / Keywords: Makine Öğrenmesi Algoritmaları, Genetik Algoritmalar, Kalp Hastalığı Tahmini, Öznitelik Seçimi

I. GİRİŞ [INTRODUCTION]

Günümüzde Yapay Zekâ (YZ) uygulamaları sağlık sektöründe pek çok hastalığın teşhis ve tedavisinde giderek artan oranlarda kullanılmaktadır. Son yıllarda sağlık sektöründe pek çok hastalığın teşhis ve tedavisinde YZ tabanlı sistemlerin kullanıldığı görülmektedir. YZ

teknik ve yöntemleri ile sağlık alanında birçok hastalığın tespiti, erken teşhisi ve tedavisi mümkündür. YZ yöntemlerinin sağlık alanında kullanımı hastalıkların başta hastalık tanılarının insan hataları olmak daha yüksek doğruluk ve oranda yapılabilmesi olmak üzere hastalıkların tanı ve teşhisinde pek çok imkân sunmaktadır. Kalp hastalıkları da bunlardan bir tanesidir.



Bu hastalık dünyanın dört bir tarafında yaşanan ölümlerin en büyük sebeplerinden birisi olarak öne çıkmaktadır. Kalp hastalığı, Dünya’da ve Türkiye’de ölüm sebebi olarak birinci sırada gösterilmektedir [1,2]. Kalp hastalıkları kadın ve erkeklerde dünyanın en yaygın hastalıkları arasındadır. Gelecekte bu hastalık nedeniyle ölüm sayısının daha da artacağı öngörülmektedir. Bu hastalık aynı zamanda pek çok başka hastalığın ortaya çıkmasına da sebep olmaktadır. Son yıllarda yapılan araştırmalar ani olarak hayatını kaybeden bireylerin yüzde 46,2’sinin kalp ve damar hastalıkları nedeniyle hayatını kaybettiklerini, bu nedenle bu hastalığın erken teşhisinin ve tedavisinin çok önemli olduğunu ortaya koymaktadır [3]. Bulanık Mantık, Doğal Dil İşleme, Konuşma Tanıma, Örüntü Tanıma, Yapay Sinir Ağları (YSA), Makine Öğrenmesi (MÖ), Genetik Algoritmaları (GA) YZ kapsamında sağlık sektöründe uygulama imkânı bulmaktadır. MÖ öğrenmesi ile örnek veri setlerine dayanan sınıflandırma veya tahmin işlemleri yapılabilmektedir. MÖ öğrenmesi ile tahmin işlemlerinde veri setinin başarısını etkileyen birçok öznelik bulunmaktadır. Veri setinde yer alan öznelikler tahmin başarı oranlarını olumlu veya olumsuz düzeyde etkileyebilmektedir. Bu nedenle veri setindeki tüm özneliklerin hepsini birebir alıp kullanmak yerine, veri setini daha belirgin olarak betimleyebilecek, veri setini en doğru temsil edebilecek nitelikli öznelik alanlarının belirlenmesi önem kazanmaktadır. Bu şekilde belirlenecek olan öznelik verileri sınıflandırma yöntemlerinin tahmin başarı oranlarını artırabilecektir. GA yöntemi de bu bağlamda özneliklerin belirlenmesinde kullanılan yöntemlerden bir tanesidir. GA ile gerçekleştirilen uygulamalarda analiz yapılmasında yüksek doğruluk oranına sahip başarılı sonuçlar elde edilebilmektedir. [4].

MÖ çeşitli görevlerin öğrenilmesi, mantıksal ve ikili çıkarımlar yoluyla otomatik hesaplama yöntemlerini kapsayan bir süreç olarak kullanılmaktadır. MÖ ile pek çok hastalığın tanısı yüksek doğruluk ile yapılabilmektedir. Literatürde hastalık tanı tahminlerinin en yüksek doğruluk oranı ile alınabilmesi ile MÖ kullanılarak çeşitli çalışmalar yapıldığı görülmektedir. En yüksek doğruluk oranı ile YZ teknikleriyle hastalıkların tahmin edilmesinin hastalıklarının tedavisine büyük katkılar sağlayacağı bir gerçektir. Bu çalışmanın amacı GA yaklaşımıyla öznelik seçimi kullanılarak MÖ algoritmaları ile kalp hastalığını tahmin etmek olarak belirlenmiştir. Bu bağlamda çalışmada YZ’nin bir alt dalı olan MÖ öğrenmesi ile daha yüksek doğruluk oranına sahip tahmin sonuçları elde edilmesinde GA ile öznelik

seçiminin önemini belirlenmesi hedeflenmiştir. Bu bağlamda çalışmada öznelik seçimi için GA’dan faydalanılmış ve bu yöntemin başarı oranına etkisi ortaya konmaya çalışılmıştır. Çalışmada kullanılan veri seti üzerinde farklı deneyler yapılmış ve bu deneyler neticesinde elde edilen başarı oranları değerlendirilmiştir. Deneylerde veri setinde öznelik seçim yöntemlerinden YZ tabanlı GA kullanılarak daha belirgin öznelikler seçilmiş ve MÖ sınıflandırma algoritmaları ile kalp hastalığının tahmini gerçekleştirilmiştir.

Bu makale şu şekilde düzenlenmiştir: Çalışmanın ikinci bölümünde sağlık alanında hastalıkların tanımlanmasında literatürde yer alan çalışmalar incelenmiştir. Üçüncü bölümünde yapılan deneyler anlatılmış ve elde edilen sonuçlar belirtilmiştir. Son bölümde ise çalışma neticesinde ulaşılan sonuçlar değerlendirilmiştir.

II. İLGİLİ ÇALIŞMALAR [RELATED STUDIES]

Sağlık alanında YZ yönteminin kullanılması ile ilgili çalışmalar, son yıllarda giderek artan bir hızla yaygınlaşmaya başlamıştır. Sağlık alanına yönelik olarak MÖ olmak üzere YZ teknik ve yöntemleri kullanılarak pek çok çalışmanın bulunduğu, ayrıca bu çalışmaların giderek arttığı görülmektedir. Biyomedikal veri kümesi üzerinde YSA algoritması dahil edilerek yapılan çalışmada elde edilen başarı oranları karşılaştırıldığında en yüksek doğruluk oranının K-EYK algoritması ile olduğu görülmüştür [5]. Karaciğer hastalığı üzerine denetimli makine öğrenimi algoritmalarını ile tahmin elde eden bir çalışmada LR, RO, KA, DVM, K-EYK ve NB için %75, %74, %69, %64, %62 ve %53 doğruluk değeri bulunmuştur [6]. MÖ ve Veri Madenciliği (VM) teknikleriyle öğretmenlerin psikososyal risk düzeylerinin tahmin edilmesi amacıyla yapılan çalışmada kullanılan veri seti üzerinde Elektrokardiyografi (EKG), Elektromiyografi (EMG) ve Elektroensefalografi (EEG) gibi fizyolojik faktörler MATLAB ortamında test edilmiş ve MÖ algoritmalarından KA, K-EYK, DVM karşılaştırılmış ve %99,9 doğruluk oranına ulaşılmıştır [7]. MÖ algoritmaları kullanılarak klinik tabanlı veriler üzerinden gerçekleştirilen çalışmada radyoloji verileri ile bütünlük sağlanarak kanser hastaları için kişisel tedavi aracı geliştirilmesi amaçlanmıştır [8]. Kalp hastalığı tahminin tespiti amacıyla yapılan diğer bir çalışmada ise RO yöntemi ve DVM kullanılmış ve çalışmalar neticesinde %87,87 tahmin doğruluk oranına ulaşılmıştır [9]. Sağlık hizmetleri alanında yorumlanabilir makine öğrenmesi modellerinin kullanımı üzerine yapılan çalışmada ihtiyaç duyulan MÖ algoritmasının sisteme nasıl ekleneceği araştırılmıştır [10]. Diyabet teşhisi



konmasında hastalığın tahmin edilebilmesine yönelik olarak YZ teknikleri kullanılarak yapılan çalışmada optimizasyon algoritmaları kullanılmış ve optimizasyon algoritmalarının kullanımının tahmin sonucu başarı oranını artırabileceği gösterilmiştir [11]. Araç algılama sistemi geliştiren bir çalışmada YZ temelli çalışan bir video algılayıcı sistem kullanılarak ortalama hız, araç sınıflandırma vb. araç takibi bilgileri üretilir [12].

Öz nitelik seçimiyle kanser sınıflandırması yapılan bir çalışmada sinyal görüntü oranı, K-EYK, DVM yararlanılmıştır. K-EYK ile %99.3 doğruluk oranı elde edilmiştir [13]. DVM algoritması ile tahmin işlemleri amacıyla yapılan çalışmada Veri Madenciliği (VM) ve R programlama dili kullanılmıştır [14]. Literatürde içinde bulunduğumuz Pandemi dönemine ilişkin olarak da çalışmalar yapıldığı görülmektedir. Yüksek risk altındaki hastaları, ölüm oranlarını ve diğer anormallikleri belirlemek için yapılan çalışmada MÖ öğrenmesi ile risk faktörlerini yaş, sosyal alışkanlıklar, konum ve iklim faktörlerinin hastalığa olan etkisi araştırılmıştır [15]. Ortaöğretim 6, 7 ve 8. sınıf öğrencilerine, belirli ölçütlere göre 24 sorudan oluşan anket uygulanarak yapılan bir diğer çalışmada anket sonuçları dikkate alınarak, MÖ öğrenmesiyle Türkçe, Matematik dersleri ve dönem sonu genel başarı puan ortalamalarının regresyon / çok sınıflı MÖ modelleri oluşturulmuştur [16]. Cilt kanserinin YZ teknik ve yöntemleri ile tahmin edilmesi konusunda yapılan çalışmada %94 ile doğruluk oranına ulaşılmıştır [17]. Kalp hastalıklarının tahmin edilmesi amacıyla yapılan bir diğer çalışmada VM sınıflandırma teknikleri kullanılmış ve çalışma neticesinde kalp hastalıkları tahminlerinde risk faktörlerinin tahmin oranlarının yüksek veya düşük olarak belirlenmesine etki ettiği ortaya konmuştur [18]. Kalp ve meme kanseri hastaları için tahmini için NB, KA, Simple Logistic, J48 algoritmalarının kullanıldığı çalışmada çapraz doğrulama yöntemleriyle algoritmaların performansları karşılaştırılmış ve Naive Bayes'in en başarılı sonuç verdiği anlatılmıştır [19]. DVM, NB, J48, RO, Adaboost, LR, KA, Tek ve Çok Katmanlı Perceptron ile yapılan çalışmada kalp hastalığı tahmini yapılması amaçlanmıştır [20]. MÖ öğrenmesi yöntemleriyle kalp hastalıklarının tahmin edilmesi için Anfis (Adaptive Neuro-Fuzzy Inference System) ve Bulanık K Ortalamalar yöntemleri kullanılmış ve K kat çaprazlama yöntemi ile %98,35 başarı oranı elde edilmiştir [21]. VM, YSA ve GA kullanılarak Matlab ortamında kalp hastalığı riskinin tahmini amacıyla yapılan çalışmada %89 başarı oranı elde edilmiştir [22]. Diğer bir çalışmada Koroner Arter Hastalığı riski tespitinde bir sınıflama modeli %83,498 başarıyla Çok katmanlı YSA sınıflandırma yöntemi ile geliştirilmiştir [23]. Kalp hastalarının yaşam kalitesini artırmak

hastaneye yatışını azaltmak ve iyileştirmek için kanıta dayalı tedavi kapsamında yapılan çalışmada bilimsel teknikler ve ölçümler kullanılmıştır [24]. "Kardiyovasküler Hastalık Risk Faktörleri Bilgi Düzeyi (KARRİF-BD)" ölçeği ile verilerin toplandığı çalışmada ise t testi, Kruskal Wallis Varyans Analizi, Mann Whitney U testi ve Spearman korelasyon analizi kullanılmış, çalışma neticesinde Göğüs ağrısı tipi, egzersize bağlı anjin ve egzersize bağlı ST depresyonu nitelikleri yüksek oranda birbiriyle ilişkili olduğu gösterilmiştir [25]. Kalp hızı ve kalp hızı değişkenliğinin kardiyovasküler risk faktörleri ile ilişkisinin araştırıldığı çalışmada başta kalp hızı ölçümü olmak üzere farklı değerler kullanılmıştır [26]. Yaşa bağlı olarak yapısal değişimlerin gen ifade düzeyleriyle olan ilişkisinin araştırıldığı çalışmada da yaş arttıkça gen ifade düzeylerindeki azalmanın görüldüğü ortaya onmuştur [27]. Obezite ile birçok hastalığın meydana geldiği düşüncesinden hareketle yapılan bir diğer çalışmada obezitenin kalp hastalığı oluşumunda etken faktör olduğu gösterilmiştir [28]. Amerikan Kalp Derneği (AHA) tarafından yapılan çalışma kalp hastalığı, felç ve kardiyovasküler risk faktörleri ile ilgili en güncel istatistiklere katkı sağlamak amacıyla yapıldığı görülmektedir [29]. Temel bileşen analizi ile hipertansiyon hastalığının tahmininin amaçlandığı çalışmada Katmanlı Algılayıcı Ağ (ÇKA), Karar Tablosu ve C4.5 sınıflandırma algoritmaları kullanılmış ve çalışma neticesinde en başarılı sonucun %88 ile NB algoritması ile alındığı görülmektedir [30]. Kalp hastalarının sosyo-demografik hastalığa ilişkin özelliklerin yaşam kalitesi ve sağlıklı yaşam biçimi davranışlarına etkisinin araştırıldığı çalışmada; erkek, 49-60 yaş arası, evli, lisans ve üstü mezunu, sosyo-ekonomik durumu iyi, çalışan, ek hastalığı olmayan bireylerin hastalıklarıyla ilgili eğitim alanlarının risk faktörü araştırılmıştır [31]. Kardiyovasküler hastalıkların yaşla ilişkisinin araştırıldığı çalışmada Kardiyovasküler hastalıklardan ölümlerin ve majör komplikasyonların %78'i, 75 yaş üzerindeki hastalarda görüldüğü çalışma sonucu olarak belirtilmiştir [32]. NB, Destek Vektör Makineleri, Lineer Diskriminant Analiz Algoritması, Yapay Sinir Ağları, C4.5, C5.0, K-Nearest Neighbor algoritmaları ile yapılan çalışmada bu algoritmaların başarı oranları ortaya konmuştur [33]. Diğer bir çalışmada MÖ algoritmalarının sağlık alanında hastalık biyolojisi, ilaç keşfi, mikrobiyom araştırmaları ve sentetik biyoloji gibi birçok alanda kullanılabileceği araştırılmıştır [34]. Benzer şekilde başka bir çalışmada aynı araştırmanın ZeroR, OneR, NB, J48 Karar Ağacı, RO, Çoklu Algılayıcılar, K-EYK, LR, DVM gibi sınıflandırma algoritmaları ile WEKA programında yapıldığı görülmektedir [35]. MÖ yöntemleriyle telekomünikasyon sektöründeki müşteri kaybı tahmini üzerine çalışmada Yeniden Örnekleme, Sentetik Azınlık



Aşırı Örneklemeye Yöntemi (SMOTE), Uyarlanabilir Sentetik Örneklemeye (ADASYN), Rastgele Aşırı Örneklemeye (ROSE) yöntemleri uygulanmış, sınıflandırma problemi için NB, KA, RO, Yapay Sinir Ağları, LR, DVM ve K-EYK yöntemleri 10 kat Çapraz Geçerlemeyle araştırılmıştır [36]. Futbol takımların hücumu dayalı ve savunmaya dayalı yetenekleri dikkate alınarak maçın sonucunu ne kadar başarılı tahmin edebildiği konusunda yapılan çalışmada 16 futbol liginde yapılan testler sonucu Karar Tablosu çoğunlukla en başarılı algoritma olduğu anlatılmaktadır [37]. Fiber optik dağıtılmış akustik algılama yapan bir çalışmada sınıflandırma bölümünde iki farklı derin öğrenme algoritmasından yararlanılmıştır [38]. MÖ ve VM tahmin ve modelleme ile ilişkisinin araştırıldığı çalışma, tahmine dayalı modelleme, finansal hizmetler, sağlık hizmetleri, telekomünikasyon vb. alanları üzerine yapılmış olan bir çalışmadır [39]. Koroner arter hastalığı teşhisi için MÖ ve VM teknikleri konusunda yapılan çalışmada, çalışma kapsamında geliştirilen veri tabanı sunulmaktadır [40].

Biyomedikal veri kümeleri ile MÖ algoritmaları üzerine yapılan bir çalışmada YSA, K-EYK ve Lineer Diskriminant Analiz (LDA) algoritmaları ile EEG veri kümesi üzerinde sınıflandırma işlemleri sonucunda epileptik veriler başarı ile sınıflandırılmıştır [41]. MÖ yöntemleriyle kredi kartı dolandırıcılığına karşı yapılan bir çalışmada elde edilen en yüksek başarı sonucu Naive Bayes algoritmasında %97.92 olarak bulunmuştur [42]. Kredi kartı dolandırıcılığını MÖ yöntemleri kullanılarak tahmin eden bir başka çalışmada ise en yüksek başarı oranı çok katmanlı yapay sinir ağları ile %99,943'tür. Naive Bayes algoritması ile %98,207 başarı sonucu elde edilmiştir [43]. Değişken önem değerlendirmesine dayalı MÖ algoritmalarını karşılaştırmalı analizini yapan bir çalışmada, Random Forest, Cubist ve MARS, hem tahmin performansı hem de değişken önem değerlendirmesi için etkilidir sonucu çıkarılmıştır [44].

Çok katmanlı algılayıcı (ÇKA) sinir ağı modeli ile DVM yöntemleri kullanılarak yapılan çalışmada buzlanmadan kaynaklı kazaların önlenmesine yönelik %87.26 doğruluk oranına sahip bir buzlanma tahmin algoritması ve mobil uygulama geliştirilmiştir [45]. EEG sinyallerinden epileptik nöbet sınıflandırılmasının yapıldığı çalışmada DVM kullanıldığı görülmektedir [46]. Bölgelere göre itfaiye istasyonu ihtiyacının MÖ kullanılarak tahmin sınıflandırılmasında en başarılı algoritmanın %93,84 ile RO algoritması olduğu yer almaktadır [47]. Sağlık hizmeti üzerine veri madenciliği ile yapılan çalışmada NB, KA ve DVM algoritmaları karşılaştırılmıştır [48]. Kayısı çekirdeklerinin tatlı veya acı

olarak sınıflandırılması üzerine yapılan çalışmada K-EYK, DVM, KA, Rasgele Orman, NB ve Çok Katmanlı Algılayıcı algoritmaları kullanılmıştır [49]. Sosyal medya paylaşımları ile markaların itibarını analiz etmesini ve tüketiciler ile ilişkisini yöneten bilgi sistemi geliştirilmesinin hedeflendiği çalışmada MÖ yöntemleriyle sosyal medya marka itibar analizi alanında yazılım yapılmış ve çalışma neticesinde modeller arasından %90 F-ölçütü başarı skoruna sahip NB seçilmiştir [50]. Diğer bir çalışmada ise veri kümelerinin MÖ ile analiz edilmesi araştırılmıştır [51]. Duygu analizi konusunda yapılan çalışmada kümeleme, sınıflandırma, YSA ve VM teknikleri kullanılarak araştırma yapılmıştır [52]. Doğal dil işleme ve MÖ kullanılarak Twitter üzerinden iletilen mesajların hastalık bilgisinin içerme durumuna göre hastalık yayılım hızı analizi gerçekleştirilmiştir [53].

Geliştirilen MÖ uygulamasında yapılan deneylerin üzerinde GA ile öznitelik seçimi yapılarak başarı oranının arttığı belirtilmiştir [54]. Genetik algoritma kullanarak MRI tabanlı beyin tümörü Sınıflandırması yapan bir çalışmada %92,03 doğruluk elde edilmiştir [55]. GA ile çözüme ulaşılabilecek evrim süreci bilgisayar ortamında taklit edilmektedir. Karmaşık çok boyutlu arama uzayında bütünsel en iyi çözümü aramaktadır [56]. GA, çok sayıda sınırlama içeren ve karmaşık optimizasyon sorunlarını yazılım yöntemleriyle çözerler. GA ve sınıflandırıcı yöntemler kullanarak kanseri tahmin eden bir çalışmada en başarılı sonuç akciğer kanserine ait verilerde %94.09 olarak DVM algoritmasında elde edilmiştir [57]. Amacı tek bir çözüm değil çözümler kümesi ortaya çıkarmaktır. Hatta çözümsüz problemlere çözüm üretebilmektir. Bir problemin çözümü birden fazla olabilir. Bu çözüm kümesine nüfus adı verilir. Bireye ait her bir elemana gen denir. Algoritmada yapılan işlemler genetik operatör olarak adlandırılır. Genetik Operatörler seçme (selection), çaprazlama (crossover) ve mutasyon (mutation)'dur. Mutasyon sayesinde genetik çeşitlilik artar [58]. Bu sayede daha çok çeşit ortaya çıkar. Elde edilen bu nüfus ileriki nesillerde kullanılır. GA'da Mutasyon ile, kayıp veya keşfedilmemiş genetik materyaller bulunur ve erken yakınsamayı önler, yerel optimum seviyede kalır [59]. Her nesil için bu adımlar tekrarlanır.

Çalışmamızda elde edilen başarı oranı ile literatürde yer alan benzer çalışmaların başarı oranları Tablo 1'de sunulmuştur. Çalışmamızın sonuçlarının araştırılan diğer çalışmalarla benzer sonuçlar içerdiği ve çalışmalarda belirtilen referans aralığında başarı sonucu gösterdiği görülmektedir.



Tablo 1. Literatür taramasının değerlendirilmesi

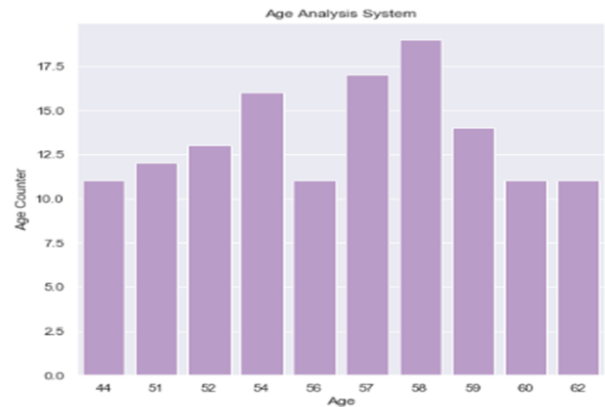
Literatür Adı	Kullanılan Algoritma Tipi	Veri Seti	Başarı Sonucu	Hata Sonucu
Polatgil, 2020	Anfis ve Bulanık K Ortalamalar	Kalp	%98.35	%1.65
Ekrem, 2020	Rastgele Orman	Kalp	%87.87	%12.13
Ünver, 2020	Naive Bayes	Hipertansiyon	%88	%12
Kabaoğlu, 2021	ÇKA	2017 yılı kış ayının hava tahmin verileri	%87.26	%12.74
Selimoğlu, 2021	Çok katmanlı yapay sinir ağları	Kredi Kartı Dolandırıcılık Teşhis	%99.943	%0.057
Mishra, 2011	K-EYK Algoritması	Lösemi	%99.3	%0.7
Rahman, 2019	Logistic Regresyon	Hint karaciğer hastaları	%75	%25
Çalışmamız	Genetik Algoritma	Kalp	%93.44	%6.56

III. TEORİ VE YÖNTEM [THEORY AND METHODOLOGY]

A. Veri Seti (Dataset)

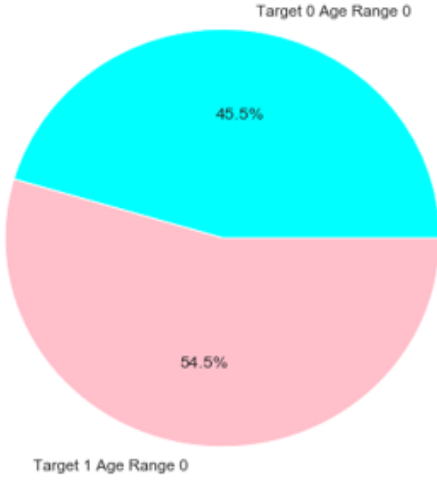
Bu çalışmada Kaggle'dan elde edilen kalp veri seti kullanılmıştır. Bu veri seti ABD'nin Cleveland eyaletindeki bireylerin verisidir. Kalp veri seti; Andras JANOSI, William STEINBRUNN, Matthias PFISTER ve Robert DETRANO tarafından oluşturulmuştur [60]. Kullanılan veri setinde yaralanma öznelik bilgileri 14 sütun ve 303 satırdan oluşmaktadır. Veri setinde hasta bireyler 1 rakamı, sağlıklı bireyler ise 0 rakamı ile ifade edilmiştir. Sütunlarda yer alan tüm veriler sayısal olarak gösterilmiştir. Veri setinde yer alan 14 farklı öznelik değerleri şu şekildedir: Yaş, cinsiyet (0=kadın,1=erkek), göğüs ağrısı tipi (0=tipik angina, 1=atipik angina, 2=anginal olmayan ağrı, 3=asemptomatik angina), dinlenme kan basıncı (tansiyon), kolesterol, açlık kan şekeri (120'den büyükse 1,değilse 0), dinlenme EKG ölçümü (0=normal,1=ST-T dalga anormalliği, 2=Sol ventrikül hipertrofisi), maksimum kalp atış hızı, egzersiz kaynaklı ağrı (1=evet,0=hayır), Egzersizin neden olduğu ST depresyonu, ST segmentinin eğimi (1=yukarı eğimli,2=düz,3=aşağı eğimli), ana damar sayısı(0-3), Talesimi (0=talasemi hastalığı yok, 1=sabit kusur, 2=normal kan akışı, 3=geri dönüştürülebilir kusur). Veri setinde boş değer (null) bulunmamaktadır. Hedef sütunu y, diğer tüm sütunlar x olarak ayrılmıştır. Hold-out yöntemiyle veriler %20 test ve %80 eğitim (train) verisi olarak ayrılmıştır. Bu kapsamda çalışmada kullanılan toplamda 303 adet veriden 242 adedi eğitim verisi ve 61 adedi ise test verisi olarak kullanılmıştır. Verilerin kalitesini artırmanın yolu verileri ön işlemden geçirmek suretiyle veri kalitesinin artırılması yoluyla

gerçekleşebilir. Bu maksatla veri setimiz üzerinde veri ön işleme adımları gerçekleştirilmiştir. Veri setimizdeki verilerin tamamlanmamış, tutarsız ya da güncelliğini kaybetmiş olup olmadığı kontrol edilmiş ve kaliteli veri için veri ön işleme adımları uygulanmıştır. Bu bağlamda null değere sahip veriler sonuca dahil edilmeden önce veri setinden çıkarılmıştır. Veri ön işleme aşamasında eksik, gürültülü ve tutarsız olan veriler iyileştirilmiştir. Veri setimizde yer alan bireylerin yaşlarını gösteren sütun grafik Şekil 1'de gösterilmektedir. Bu grafiğe göre, örneğin 58 yaşında olan birey sayısı 19 kişidir.



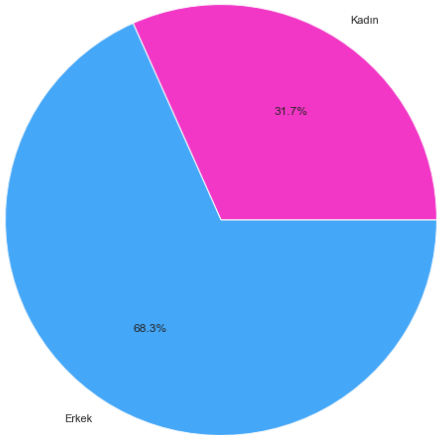
Şekil 1. Yaşa göre sütun grafiği

Şekil 2'de genç yaşta bulunan bireylerin hasta ve sağlıklı olma durum oranları sunulmuştur. Söz konusu şekilde hasta bireyler 1 rakamı ile ve sağlıklı bireylerde 0 rakamı ile gösterilmiştir.



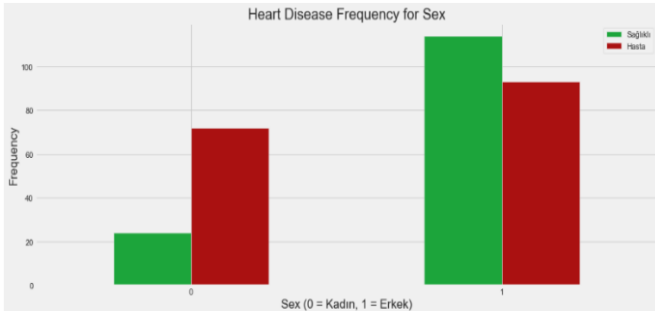
Şekil 2. Gençlerin kalp hastası/sağlıklı oranı

Şekil 3'te ise veri setinde bulunan bireylerin cinsiyet dağılımına göre pasta grafiği oluşturulmuştur. Grafiğe göre veri setindeki bireylerin %31.7' si kadın, %68.3' ü erkeklerden oluşmaktadır.



Şekil 3. Cinsiyet dağılımına göre pasta grafiği

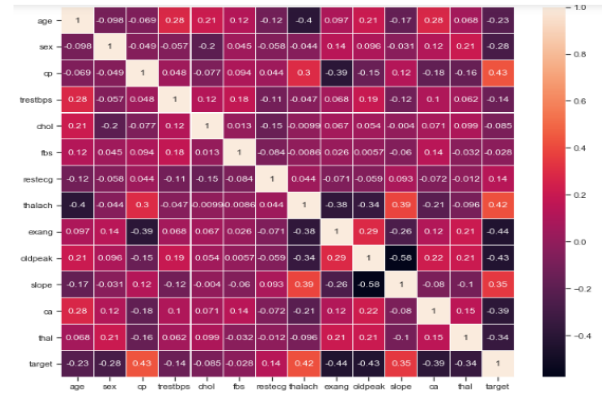
Şekil 4'te kalp hastalığının cinsiyete bağlı görülme sıklığı sunulmuştur. Bu grafiğe göre cinsiyeti kadın olan 23 sağlıklı ve 70 hasta birey yer alırken cinsiyeti erkek olan toplamda 100 sağlıklı ve 90 hasta birey yer almaktadır.



Şekil 4. Kalp hastası-cinsiyet dağılımı

Şekil 5'de veri setindeki özneliklerin korelasyon matrisi sunulmuştur. Bu matriste kullanılan korelasyon (r) kavramı ile özelliklerin birbiriyle ilişkisini gösterilmiştir. Bu bağlamda -1'e olan yakınlık değişkenler arasındaki negatif ilişkiyi, +1'e olan yakınlık ise pozitif ilişkiyi ve 0'a yakınlık da birbiriyle ilişkisi olmadığını göstermektedir. Bu duruma ilişkin durum Denklem (1)'de formülize edilmiştir.

$$-1 \leq r \leq +1 \quad (1)$$



Şekil 5. Korelasyon matrisi

Şekil 5 incelendiğinde bir bireyin kalp hastası olmasına göğüs ağrısı tipi, egzersize kaynaklı ağrı ve egzersiz neden olduğu, ayrıca ST depresyonu arasında pozitif ilişki bulunduğu görülmektedir. Bu durum kalp hastalığı tanısının konmasında etkindir. Ayrıca cinsiyet ve yaş öznelikleri de kalp hastalığı tanısında etkilidir. Cinsiyeti erkek olan bireylerin ve yaşlıların kalp hastası olma ihtimalinin daha yüksek olduğu görülmektedir.

B. Makine Öğrenmesi (MÖ) Algoritmaları

MÖ algoritma yöntemleri denetimli öğrenme, yarı denetimli öğrenme ve denimsiz öğrenme olmak üzere üç başlık altında gösterilebilir. K-EYK, NB, KA, RO, LR ve DVM algoritmaları bu kapsamda en çok kullanılan algoritmalar arasında sayılabilir.

LR, çeşitli tahmin değişkenleri ile ikili sonuç arasındaki ilişkiyi değerlendiren istatistiksel bir tekniktir [61, 62]. LR analizinin sağlık hizmetleri alanında sıkça kullanıldığı görülmektedir [63]. Bu algoritma bir sınıflandırma algoritmasıdır. Bu algoritmanın formülü Denklem (2)'de sunulmuştur.

$$\text{logic}(p) = \ln\left(\frac{p}{1-p}\right) \quad p = \text{probability} \quad (2)$$

K-EYK algoritması da sık kullanılan sınıflandırma algoritmaları arasındadır. Bu algoritma parametrik bir algoritma değildir. Tahmin işlemlerinde veri seti içinde belirlenen K değeri hesaplanacak en yakın komşuların eleman sayısını ifade eder. Örneğin; K değeri 5 ise en yakın 5 komşunun elde edilen değer ile arasındaki uzaklık hesaplanır. Uzaklık hesaplar Kenan Denklem (3) Öklid, Denklem (4) Manhattan ve Denklem (5) Minkowski standart mesafe ölçümü denklemlerinden yararlanır.

$$\text{Öklid} = \sqrt{\sum_{i=1}^k (x_i - y_i)^2} \quad (3)$$

$$\text{Manhattan} = \sum_{i=1}^k |x_i - y_i| \quad (4)$$

$$\text{Minkowski} = \left(\sum_{i=1}^k (|x_i - y_i|^q) \right)^{1/q} \quad (5)$$

MÖ yöntemlerinden DVM algoritması sınıflandırma ve regresyon için verilerin analizinde kullanılan denetimli öğrenme modelidir. Bu algortmada amaç, iki sınıf arasındaki uzaklığın maksimum olduğu durumun bulunmasıdır. Vektör uzayı tabanlı olan bu makine öğrenmesi algoritması parametre almaz. DVM algoritması Denklem (6)'da formül olarak gösterilmiştir.

$$\{x_i, y_i\} | i = 1 \dots L, y_i \in \{-1, 1\}, x \in R^D \quad (6)$$

NB algoritması, sınıflandırma algoritmalarındandır. Bu algoritmanın çalışma mantığı her bir eleman için tek tek her durumun olasılığını hesaplayarak en yüksek olasılık değerini temel alarak sınıflandırma yapmaktadır. NB algoritması Denklem (7)'de göre formül olarak gösterilmiştir.

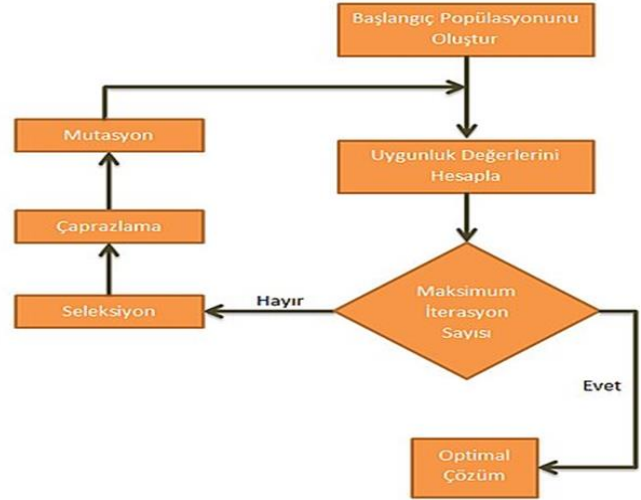
$$P\left(\frac{\text{class}}{\text{data}}\right) = \frac{p(\text{data/class}) \times p(\text{class})}{p(\text{data})} \quad (7)$$

KA algoritması, veri kümesindeki çok büyük verileri belirli kurallar ile daha küçük kümelerle ayıran sınıflandırma algoritmasıdır. RO algoritması denetimli sınıflandırma algoritmalarındandır. Büyük verileri kümeleyerek küçük ağaçlar oluşturur. Bu algortmada ağaç sayısı artarsa kesinlik artar. Bu algoritmanın KA algoritmasından farkı kök düğüm bulma ve düğümlere bölme işleminin rastgele çalışmasıdır.

C. Genetik Algoritmalar (GA)

Teknolojinin gelişmiş algoritmalarından biri olan GA çalışma mantığı insan genlerinin nesiller arası geçiş sürecine benzetilmektedir. GA, doğal genetik ve biyolojik evrimin mekaniklerine dayanan fonksiyon optimizasyonu için kullanılan stokastik bir yöntemdir. Genetik

algoritmanın çalışma prensibi Şekil 6'da gösterilmiştir [64].



Şekil 6. Genetik algoritma akış diyagramı

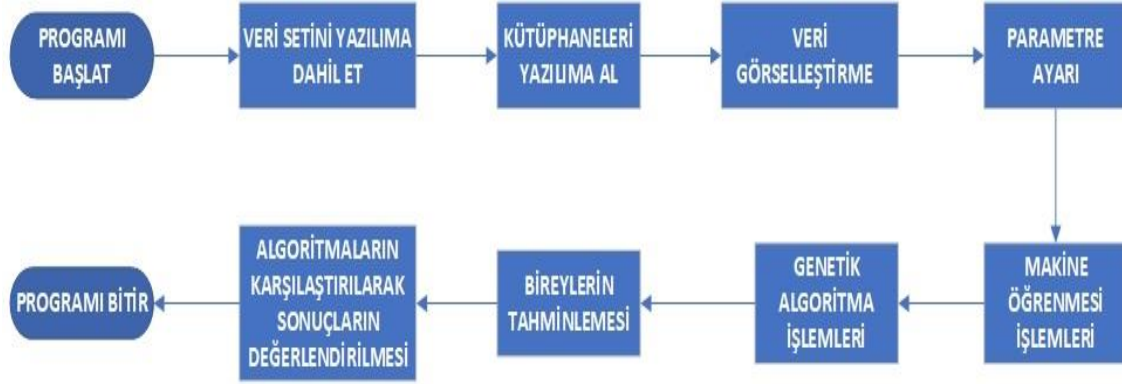
GA çalışma adımlarında ilk olarak “nüfus” bilgisi bulunmaktadır. Bu bilgi ile verilere uygun popülasyon rastgele olarak başlatılır. İkinci aşama “uygunluk” değeridir. Sonuca ulaştıran her bir çözümün (kromozom) uygunluk değeri bulunur. Üçüncü aşamada gelecek nesil ile gen geçişi arasında yeni bir popülasyon olması için ebeveyn olarak en uygun kromozomlar seçilir. Dördüncü aşamada çaprazlama işlemi yapılmaktadır. Ebeveynler birleştirilip yeni kromozom seti oluşturulur. Daha sonra yeni oluşan popülasyona eklenir. Beşinci aşamada oluşan popülasyon setindeki kromozomun gen değişimleri yapılır. Bu işleme Mutasyon denilmektedir.

D. Sistemin Akış Diyagramı

Çalışmada Şekil 7 üzerinde belirtilen akış diyagramı kullanılmıştır. Akış diyagramına uygun olarak öncelikle Jupyter notebook yazılım programı başlatılmaktadır. Tahmini ve analizi yapılacak olan veri seti çalışmaya eklenmiştir. Kullanılacak algoritmalar, istatistikler, görselleştirme-animasyonlar için gerekli olan kütüphaneler yazılıma eklenmiştir. Elde edilen istatistiklerin görsel şekiller çıkarılmıştır. Veri setindeki bireylerin yaş dağılımı, cinsiyet ayrımı, öznitelik gruplandırma vb. istatistiksel sonuçları daire, satır, sütun şekilleri ile gösterilmiştir. Gerekli analizler yapılır. Daha sonrasında öznitelik ayrımı yapılır. Bu ayrım sonucu hedef dışındaki tüm veriler giriş(x), hedef özniteliği ise çıkış(y) olarak ayrılmıştır. Hold out yöntemiyle %20 test ve %80 eğitim verisi olarak veri seti ayrımı yapılır. LR, KA, NB, K-EYK, RO, DVM algoritmaları olan 6 makine öğrenme algoritması sırasıyla programa eklenir.

Sonrasında başarı ve hata oranları hesaplanır. Bu temel MÖ algoritmaları sütun grafiği ile karşılaştırması yapılır. Öznitelik seçimi olarak adlandırılan 24 farklı özniteliklerin algoritmalarındaki başarı sonuçları hesaplanır. Yapılan tahminlerin her algoritma için sonucu karşılaştırılmıştır. Çalışma kapsamında toplamda 28 adet farklı deney yapılmıştır. Çalışmada kümesi içinde hangi

özelliklerin sonuç üzerinde ne kadar etkili olduğunu belirlemek amacıyla öznitelik seçimi (feature selection) yöntem, özellik seçimi işlemi sistemin daha hızlı çalışmasını sağladığından dolayı uygulanmıştır.



Şekil 7. Çalışmanın Akış Diyagramı

IV. BULGULAR VE TARTIŞMA [RESULT AND FINDINGS]

Çalışma kapsamında 3 farklı grupta toplamda 28 adet deney gerçekleştirilmiştir. İlk grupta 13 adet deney yapılmıştır. Bu grupta yapılan deneylerin amacı veri setindeki öznitelik sayısının her bir deneyde farklılaştırılarak başarı oranının ortaya komasıdır. İkinci grupta 10 adet deney yapılmıştır. Bu grupta yapılan deneylerin amacı, birinci grup deneyler sonucunda ortaya çıkan modellerin test verisi ile test edilmesidir. Üçüncü grup deneyde ise 5 adet deney yapılmıştır. Bu grupta yapılan deneylerin amacı GA ile öznitelik seçimi yapılarak başarı oranını ortaya koymaktır. Deneylerde veri setinde bulunan bireylerin hasta veya sağlıklı olduğunun tahmini, çalışmadaki en başarılı sonuç %93.44 ile GA olmuştur.

Diğer algoritmalar için en başarılı sonuçlar; LR=%90.16, K-EYK=%81.96, DVM=%83.60, NB=%85.24, KA=%81.96, RO=83.60 olarak hesaplanmıştır. Çalışmanın başarı sonuçları %68.13 ile %93.44 arasındadır. İlk grup 13 deneyde giriş verilerinin ve MÖ algoritmalarının değişikliğine göre başarı sonuçları değişmiştir. Bu deneylerde elde edilen başarı sonuçları Tablo 2’de, hata sonuçları ise Tablo 3’de gösterilmiş, ayrıca algoritmalarından elde edilen başarı sonuçları karşılaştırılmıştır.

Tablo 2. Öznitelik seçimi ile MÖ algoritmalarının başarı sonuçları

Deney No.	LR (%)	K-EYK (%)	DVM (%)	NB (%)	KA (%)	RO (%)
1.	90.16	70.49	63.93	83.60	78.68	81.96
2.	84.61	70.49	68.85	83.60	75.40	80.32
3.	81.31	70.49	67.21	80.32	77.049	80.32
4.	85.71	68.85	67.21	83.60	77.049	83.60
5.	85.71	68.85	67.21	83.60	78.68	81.96
6.	85.71	72.13	67.21	83.60	81.96	83.60
7.	85.71	70.49	67.21	83.60	81.96	83.60
8.	86.81	72.13	68.85	83.60	80.32	78.68
9.	87.91	75.40	81.96	83.60	70.49	75.40
10.	81.31	81.96	83.60	81.24	80.32	81.96
11.	81.31	73.77	83.60	85.24	81.96	81.96
12.	68.13	70.49	78.68	78.68	77.049	77.049
13.	75.82	73.77	73.77	73.77	73.77	73.77

Tablo 2 ve 3 üzerinde yapılan 13 deney kapsamında yaş, cinsiyet, göğüs ağrısı tipi, dinlenme kan basıncı, kolesterol, açlık kan şekeri, dinlenme ekg ölçümü, en fazla kalp atış hızı, egzersize bağlı ağrı, dinlenmeye göre egzersizin neden olduğu ST depresyonu, ST segmentinin eğimi, ana damarların sayısı, talasemi sütunları x, hedef sütunu y olarak temel MÖ algoritmalarının öznitelik seçimi yapılarak başarı oranı değerleri hesaplanmıştır.



Tablo 3. Öznitelik seçimi ile MÖ algoritmalarının hata sonuçları

Deney No.	LR (%)	K-EYK (%)	DVM (%)	NB (%)	KA (%)	RO (%)
1.	09.84	29.51	36.07	16.40	21,32	18.04
2.	15.39	29.51	31.15	16.40	24.60	19.68
3.	18.69	29.51	32.79	19.68	22.951	19.68
4.	14.29	31.15	32.79	16.40	22.951	16.40
5.	14.29	31.15	32.79	16.40	21.32	18.04
6.	14.29	27.87	32.79	16.40	18.04	16.40
7.	14.29	29.51	32.79	16.40	18.04	16.40
8.	13.19	27.87	31.15	16.40	19.68	21.32
9.	12.09	24.60	18.04	16.40	29.51	24.60
10.	18.69	18.04	16.40	18.76	19.68	18.04
11.	18.69	26.23	16.40	14.76	18.04	18.04
12.	31.17	29.51	21.32	21.32	22.951	22.951
13.	24.18	26.23	26.23	26.23	26.23	26.23

Bu deneylere ilişkin elde edilen sonuçlar aşağıda yer almaktadır. Yapılan deneylerde genel olarak her algoritma için veriler %20 test ve %80 eğitim verisi olarak ayrılmıştır.

- LR algoritması kullanılarak yapılan deneylerde LR aşamasında algoritma 109 defa eğitilerek %90.16 başarı sonucuna ulaşılmıştır. Hata oranı %9.84'tür.
- K-EYK algoritmasında 3 komşu dikkate alınmıştır. Komşu sayısının tek sayı verilmesi optimum sonuç olmasına yardımcı olacağından hareketle komşu sayısı belirlenmiştir. Bu algoritma 41 defa eğitilerek %70.49 başarı sonucuna ulaşılmıştır. Hata oranı %29.51'dir.
- DVM algoritması 100 defa eğitilerek %63.93 başarı sonucuna ulaşılmıştır. Hata oranı da %36.07'dir.
- NB algoritması 41 defa eğitilerek %83.60 başarı sonucuna ulaşılmıştır. Hata oranı %16.40'tır.
- KA algoritması 109 defa eğitilerek %78.68 başarı sonucuna ulaşılmıştır. Hata oranı da 21.32'dir.
- RO algoritması 41 defa eğitilerek %81.96 başarı sonucuna ulaşılmıştır. Hata oranı %18.04'tür.

Tablo 2 ve 3 üzerindeki deneylerden 7 numaralı deneyde; dinlenme EKG ölçümü, en fazla kalp atış hızı,

egzersize bağlı ağrı, dinlenmeye göre egzersizin neden olduğu ST depresyonu, ST segmentinin eğimi, ana damarların sayısı, Talasemi sütunları x, hedef sütunu y olarak temel MÖ algoritmalarının başarı oranı değerleri hesaplanmıştır. Veri setinde bulunan yaş, cinsiyet, göğüs ağrısı tipi ve dinlenme kan basıncı, kolesterol, açlık kan şekeri hariç tüm sütunlar dahil edilmiştir.

Bu deneylere ilişkin elde edilen sonuçlar aşağıda yer almaktadır:

- LR algoritmasında %30 test ve %70 train olarak ayrılmıştır. Bu algoritma 41 defa eğitilerek %85.71 başarı sonucuna ulaşılmıştır. Hata oranı %14.29'dur.
- K-EYK algoritmasında 3 komşu dikkate alınmıştır. Komşu sayısının tek sayı verilmesi optimum sonuç olmasına yardımcı olur. Bu algoritmada %20 test ve %80 train olarak ayrılmıştır. Algoritma 41 defa eğitilerek %70.49 başarı sonucuna ulaşılmıştır. Hata oranı %29.51'dir.
- DVM algoritması 100 defa eğitilerek %67.21 başarı sonucuna ulaşılmıştır. Hata oranı da %32.79'dur.
- NB algoritması 41 defa eğitilerek %83.60 başarı sonucuna ulaşılmıştır. Hata oranı %16.40'tır.
- KA algoritması 109 defa eğitilerek %81.96 başarı sonucuna ulaşılmıştır. Hata oranı da 18.04'tür.
- RO algoritmasında %20 test ve %80 train olarak verisi ayrılmıştır. Bu algoritma 41 defa eğitilerek %83.60 başarı sonucuna ulaşılmıştır. Hata oranı %16.40'tır.

Çalışmada yapılan ikinci grup deneylerde veri setinde yer alan hasta ve sağlıklı olan toplam 10 bireyin verileri MÖ yöntemleri ile tahminleri yapılarak başarı ve hata sonuçları listelenmiştir. Örneğin; Deney 1'de tahmin edilen birey hastadır ve değeri "1"dir. LR algoritması ile tahmin edilmesi istendiğinde %90.16 başarı oranıyla bireyin hasta olduğu tahmin edilmiştir.

Tablo 4'de görüldüğü gibi bu tahminin hata oranı 0.1764'dür. Hasta birey: Hedef 1 ve Sağlıklı birey: Hedef 0 olarak tanımlanmıştır. Tahminler; %20 test, %80 train olarak ayrılarak 100 defa eğitilerek sonuçlandırılmıştır.

Tablo 4. Veri setinde bulunan 10 bireyin tahmini ve algoritmaların başarı sonuçları

Deney No.	Gerçek Veri	LR (%)	K-EYK (%)	DVM (%)	NB (%)	KA (%)	RO (%)
1.	1	%90,16 (1)	%70,49 (0)	%63,93 (1)	%83,60 (1)	%75,40 (0)	%81,96 (0)
2.	0	%90,16 (0)	%70,49 (1)	%63,63 (0)	%83,60 (0)	%75,40 (0)	%81,96 (0)
3.	1	%90,16 (1)	%70,49 (0)	%63,93 (1)	%83,60 (1)	%75,40 (1)	%81,96 (1)
4.	0	%90,16 (0)	%70,49 (1)	%63,93 (1)	%83,60 (1)	%75,40 (0)	%81,96 (0)
5.	0	%90,16 (0)	%70,49 (0)	%63,93 (0)	%83,60 (0)	%75,74 (0)	%81,96 (0)
6.	1	%90,16 (1)	%70,49 (1)	%63,93 (1)	%83,60 (1)	%75,40 (1)	%81,96 (1)
7.	1	%90,16 (1)	%70,49 (1)	%63,93 (1)	%83,60 (1)	%75,40 (1)	%81,96 (1)
8.	0	%90,16 (0)	%70,49 (0)	%63,93 (0)	%83,60 (1)	%75,40 (0)	%81,96 (0)
9.	0	%90,16 (0)	%70,49 (0)	%63,93 (0)	%83,60 (0)	%72,13 (0)	%81,96 (0)
10.	1	%90,16 (0)	%70,49 (1)	%63,93 (1)	%83,60 (0)	%73,77 (1)	%81,96 (1)

Tablo 5. Veri setinde bulunan bireylerin GA ile tahmini

Deney No.	Gerçek Veri	GA ile Tahmin Veri
1.	1	1
2.	1	1
3.	0	0
4.	0	1
5.	0	0

Yapılan üçüncü aşamadaki deneyde ise GA ile veri setindeki bir bireyin değerleri girilerek hasta veya sağlıklı olup olmadığı tahmin edilmiştir. Tablo 5 üzerinde veri seti içerisinde bulunan 5 adet bireyin GA ile hasta veya sağlıklı olma tahmini yapılmıştır. Bu çalışma kapsamında GA kullanılarak %93.44 başarı ve %6.56 hata sonucu elde edilmiştir. Tablo 5 üzerinde “0” bilgisi sağlıklı bireyi, “1” bilgisi ise hasta bireyi temsil etmektedir.

Tablo 6. Genetik algoritma parametreleri

Parametreler	Değerler
Birey Sayısı	303
Birey Seçim Yöntemi	Sıralama
Çaprazlama Olasılığı	1
Mutasyon Olasılığı	0.1

Tablo 6 üzerinde GA ile analizi yapılan işlemlerde kullanılan parametreler verilmiştir. Bu parametrelerden “Birey Sayısı” ile GA’da hesaplanacak çalışmanın popülasyonuna ait birey sayısını belirlenmiştir. “Birey Seçim Yöntemi” parametresi ile bireylerin çaprazlaması için hangi yöntemin seçeceğini belirlenmiştir. “Çaprazlama Olasılığı” parametresi ile yeni nesli oluştururken seçilen bireylerin ne oranda çaprazlama yapacağını belirten sayıdır. “Mutasyon Olasılığı” parametresi ise mutasyon yapılırken hangi oranda yapılacağını belirleyen olasılık parametresidir.

Tablo 7 üzerindeki veriler ise Tablo 5 üzerindeki 1 numaralı deneydeki bireyin detaylı bilgileri ve GA sonucu tahmin verisi verilmiştir.

Tablo 7. Veri setinde bulunan bir bireyin detay bilgileri

Tahmini yapılacak bireyin özellikleri	Değerler
Yaş	37
Cinsiyet	Erkek
Göğüs ağrısı tipi	2
Dinlenme kan basıncı	130
Kolesterol	250
Açlık kan şekeri	0
Dinlenme ekg ölçümü	1
Maksimum kalp atış hızı	187
Egzersize bağlı ağrı	0
Dinlenmeye göre egzersizin neden olduğu ST depresyonu	3.5
ST segmentinin eğimi	0
Ana damar sayısı	0
Talasemi	2
Hedef	1(Hasta)

Şekil 8’de GA ile bireyin tahmin sonucunu gösteren kod gösterilmiştir.

```
data=data[data["Ca"]==0]
data=data[data["Thal"]==2]
tahmin=logmodel.predict([[37,1,2,130,250,0,1,187,0,3.5,0]])
print(tahmin,accuracy_score(y_test,predictions))
#target 1
```

[1] 0.9344262295081968

Şekil 8. GA ile hastalık tahmini kod parçası

Genetik algoritma (GA), çok boyutlu uzayda maliyet fonksiyonuna uygun olacak şekilde iterasyonlar yaparak optimum çözümü ararlar. LR ise bağımlı ve bağımsız



değişken arasındaki ilişkiyi en az sayıda değişken ile en iyi uyuma sahip olacak şekilde tahmin etmeyi amaçlar. Farklı sayıda değişken içeren regresyon modellerinden başarı sonucu yüksek olacağından genellikle Genetik Algoritma kullanılması tavsiye edilir. Bu sebeple LR başarısı yüksek çalışmaların GA başarısı da yüksektir. Bu çalışmada da LR yöntemi başarı sonucu % 90.16 olarak elde edilmiştir. Çalışmada GA temel MÖ yöntemleri ile karşılaştırılmış ve daha yüksek başarı sonucunu göstermiştir.

V. SONUÇ VE DEĞERLENDİRME [CONCLUSION]

Günümüzde YZ'nin bir alt dalı olan MÖ algoritmaları sağlık alanında çeşitli hastalıklara ilişkin teşhislerde uygun maliyet, hızlı ve daha doğru sonuçlar elde ederek tahminler yapmak mümkündür. Dünyada kalp hastalığı sonucu ölümlerin ilk sırada olması, aynı zamanda her geçen gün kalp hastalığı nedeniyle ölüm sayısının artması bu hastalığa ilişkin teşhislerde doğru tahminler yapılmasını gerektirmektedir. MÖ algoritmaları ile eğer bu hastalık yüksek başarı oranları ile tahmin edilebilirse erken teşhis ve önlemler sayesinde ölüm oranları azaltılabilir. Bu çalışmada kalp hastalığının GA yaklaşımıyla öznitelik seçimi kullanılarak MÖ algoritmaları ile tahmini gerçekleştirilmiştir. Bu kapsamda MÖ yöntemleri ve GA'dan faydalanılmıştır. Çalışmada sınıflama işlemi için hangi özniteliklerin önemli hangilerinin ise daha az önemli olduklarının belirlenmesi aşamasında öznitelik seçimi işlemi için GA metodundan faydalanılmıştır. Çalışmada ABD'nin Cleveland eyaletindeki bireylerin verisinden yararlanılarak Andras Janosi, William Steinbrunn, Matthias Pfister ile Robert Detrano tarafından oluşturulan ve Kaggle'dan [47] alınan veri seti kullanılmıştır. Çalışmada öncelikle kullanılan veri setinde GA'dan faydalanılmış ve MÖ yöntemlerinde öznitelik seçimi yapılmış, daha sonra belirlenen MÖ algoritmaları kalp hastalığı tahmininin yüksek doğruluk oranı ile yapılabileceği ortaya konmuş, böylelikle veri setinde GA ve MÖ yöntemlerinde öznitelik seçiminin tahmin doğruluk oranını artırdığı gösterilmiştir. Çalışma kapsamında kullanılan veri seti üzerinde K-EYK, LR, KA, RO, NB, DVM ve GA algoritmaları ile 3 grup deney gerçekleştirilmiştir. Farklı aşamada toplam 28 (yirmi sekiz) deney gerçekleştirilmiştir.

Çalışma kapsamında yapılan deneylerde alınan başarı sonuçları sırasıyla LR=90.16, K-EYK=81.96, DVM=83.60, NB=85.24, KA=81.96, RO=83.60 olarak ölçülmüştür. Elde edilen hata oranlarına bakıldığında en başarılı sonucun GA yaklaşımıyla yapılan tahminde %93.44 başarı oranı elde edildiği görülmektedir. Çalışmada GA ile tahmin yapıldıktan sonra elde edilen başarı oranının daha yüksek olduğu gözlemlenmektedir. Çalışmada elde edilen sonuçlar GA yaklaşımı ve MÖ

yöntemlerinde öznitelik seçimi kullanılarak kalp hastalığının daha yüksek başarı oranı ve doğrulukla tahmin edilebileceğini göstermektedir. Bu çalışma bize göstermiştir ki ilgili algoritma yapıları üzerinde veri sayısı ve öznitelik sayısı değişimi algoritmalar arasında başarı oranını etkilemektedir.

GA ve MÖ yöntemleri ile geliştirilen bu çalışmanın sağlık alanındaki literatüre katkı sağlaması amaçlanmıştır.

ÇIKAR ÇATIŞMASI [CONFLICTS OF INTEREST]

Yazarlar arasında ve ilgili kurumları arasında herhangi çıkar çatışması olmadığını bildirmişlerdir.

ETİK KURALLARA UYGUNLUK [RESEARCH AND PUBLICATION ETHICS]

Yazarlar bu makalenin etik kurul onayı veya herhangi bir özel izin gerektirmediğini beyan ederler.

KAYNAKLAR [REFERENCES]

- [1] M. Pavithra, A. M. Sindhana, T. Subajanaki, and S. Mahalakshmi, 2021. "Effective Heart Disease Prediction Systems Using Data Mining Techniques", *Annals of the Romanian Society for Cell Biology*, 6566-6571.
- [2] L. Tokgozoglu, M. Kayikcioglu, and B. Ekinici, 2021. "The Landscape of Preventive Cardiology in Turkey: Challenges and successes", *American Journal of Preventive Cardiology*, 100184.
- [3] İ. Ozcan, B. Tasar, A. B. Tatar, and O. Yakut, 2019. "Destek Vektör Makinesi Algoritması ile Kalp Hastalıklarının Tahmini", *Bilgisayar Bilimleri*, 4(2), 74-79.
- [4] A. Coşkun, and N. Arıcı, 2006. "Genetik Algoritma Kullanılarak Deriden Geçiş Özelliklerine Göre Kimyasal Maddelerin Molekül Yapılarının Bulunması", *Politeknik Dergisi*, 9(4), 255-261.
- [5] M. Karakoyun, and M. Hacıbeyoğlu, 2014. "Biyomedikal Veri Kümeleri ile Makine Öğrenmesi Sınıflandırma Algoritmalarının İstatistiksel Olarak Karşılaştırılması", *Dokuz Eylül Üniversitesi Mühendislik Fakültesi Fen ve Mühendislik Dergisi*, 16(48), 30-42.
- [6] A. K. M. Rahman, F. M. J. Shamrat, Z. Tasnim, J. Roy, and S. A. Hossain, 2019. "A Comparative Study on Liver Disease Prediction Using Supervised Machine Learning Algorithms", *International Journal of Scientific & Technology Research*, 8(11), 419-422.
- [7] Şengür, D., 2021. "Eeg, Emg And Ecg Based Determination of Psychosocial Risk Levels in Teachers Based on Wavelet Extreme Learning Machine Autoencoders", *Politeknik Dergisi*, 1-1.
- [8] M. Nazari, I. Shiri, and H. Zaidi, 2021. "Radiomics-based machine learning model to predict risk of death within 5-years in clear cell renal cell carcinoma patients", *Computers in Biology and Medicine*, 129, 104135.



- [9] Ö. Ekrem, O. K. M. Salman, B. Aksoy, and S. A. İnan, 2020. "Yapay Zekâ Yöntemleri Kullanılarak Kalp Hastalığının Tespiti", *Mühendislik Bilimleri ve Tasarım Dergisi*, 8(5), 241-254.
- [10] M. A. Ahmad, C. Eckert, and A. Teredesai, 2018. "Interpretable Machine Learning In Healthcare", *In Proceedings Of The 2018 Acm International Conference On Bioinformatics, Computational Biology, And Health Informatics*, 559-560.
- [11] U. Köse, 2019. "Zeki Optimizasyon Tabanlı Destek Vektör Makineleri ile Diyabet Teşhisi", *Politeknik Dergisi*, 22(3), 557-566.
- [12] Ş. Pazar, M. Bulut, and C. Uysal, 2020. "Yapay Zeka Tabanlı Araç Algılama Sistemi Geliştirilmesi", *Journal of Scientific, Technology and Engineering Research*, 1(1), 31-37.
- [13] D. Mishra, and B. Sahu, 2011. "Feature Selection for Cancer Classification: a Signal-To-Noise Ratio Approach", *International Journal of Scientific & Engineering Research*, 2(4), 1-7.
- [14] E. Çağlayan Akay, 2020. "Ekonometride Büyük Veri ve Makine Öğrenmesi: Temel Kavramlar", *Der Yayınları*.
- [15] S. Kushwaha, S. Bahl, A. K. Bagha, K. S. Parmar, M. Javaid, A. Haleem, and R. P. Singh, 2020. "Significant Applications Of Machine Learning For Covid-19 Pandemic", *Journal of Industrial Integration and Management*, 5(4).
- [16] M. Gök, 2017. "Makine Öğrenmesi Yöntemleri ile Akademik Başarının Tahmin Edilmesi", *Gazi Üniversitesi Fen Bilimleri Dergisi Part C: Tasarım ve Teknoloji*, 5(3), 139-148.
- [17] C. Akyel, and N. Arıcı, 2020. "Cilt Kanserinde Kıl Temizliği ve Lezyon Bölütlemesinde Yeni Bir Yaklaşım", *Politeknik Dergisi*, 23(3), 821-828.
- [18] A. Abdullah Sheik, and R. R. Rajalaxmi, 2012. "A Data Mining Model For Predicting The Coronary Heart Disease Using Random Forest Classifier", *In International Conference in Recent Trends in Computational Methods, Communication And Controls*, 22-25.
- [19] V. Chaurasia, and S. Pal, 2014. "Performance Analysis Of Data Mining Algorithms For Diagnosis And Prediction Of Heart And Breast Cancer Disease", *Review of Research*, 3(8).
- [20] Ö. Özmen, A. Khdr, E. Avcı, 2018. "Sınıflandırıcıların Kalp Hastalığı Verileri Üzerine Performans Karşılaştırması", *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, 30(3), 153-159.
- [21] M. Polatgil, 2020. "Anfis ve Bulanık K Ortalamalar ile Kalp Hastalığının Tespit Edilmesi", *Bilişim Teknolojileri Dergisi*, 13(4), 443-449.
- [22] S. Amin Umar, K. Agarwal, and R. Beg, 2013. "Genetic Neural Network Based Data Mining In Prediction Of Heart Disease Using Risk Factors", *In 2013 IEEE Conference On Information and Communication Technologies*, 1227-1231.
- [23] Ş. Cihan, B. Karabulut, G. Arslan, and G. Cihan, 2018. "Koronar Arter Hastalığı Riskinin Veri Madenciliği Yöntemleri ile İncelenmesi", *Uluslararası Mühendislik Araştırma ve Geliştirme Dergisi*, 10(1), 85-93.
- [24] M. Pagnesi, M. Adamo, and M. Metra, 2021. "March 2021 at a glance: focus on epidemiology, prevention and COVID-19", *European Journal of Heart Failure*, 23(3): 347-349.
- [25] E. M. Göktepe, and M. Yağanoğlu, 2020. "Veri Bilimi Uygulamalarının Hastalık Teşhisinde Kullanılması: Kalp Krizi Örneği", *Journal Of Information Systems And Management Research*, 2(2), 26-32.
- [26] A. Çil, E. Edem, and Ö. Göldedi, 2018. "Hipertansif Hastalarda Kalp Hızı Değişkenliği ile Kardiyovasküler Risk Faktörlerinin İlişkisi", *Sakarya Tıp Dergisi*, 8(2), 303-310.
- [27] Y. Olgar, D. Billur, and B. Turan, 2021. "Yaşa Bağlı Kalp Fonksiyon Değişiklikleri ve miRNA'lar", *Ankara Üniversitesi Tıp Fakültesi Mecmuası*, 74(2):239-244.
- [28] I. Kalan, and Y. Yeşil, 2010. "Obezite ile İlişkili Kronik Hastalıklar", *Diyabet ve Obezite*, 78.
- [29] S. S. Virani, A. Alonso, H. J. Aparicio, E. J. Benjamin, M. S. Bittencourt, C. W. Callaway, ... and American Heart Association Council on Epidemiology and Prevention Statistics Committee and Stroke Statistics Subcommittee, 2021. "Heart disease and stroke statistics—2021 update: a report from the American Heart Association", *Circulation*, 143(8), e254-e743.
- [30] M. H. Ünver, Y. Kökver, and A. Çifci, 2020. "Hipertansiyon Tahmini İçin Temel Bileşen Analizinin Kullanımı", *Uluslararası Mühendislik Araştırma ve Geliştirme Dergisi*, 12(3), 42-51.
- [31] N. Küçükberber, K. Özdu, and H. Yorulmaz, 2011. "Kalp Hastalarında Sağlıklı Yaşam Biçimi Davranışları ve Yaşam Kalitesine Etki Eden Faktörlerin Değerlendirilmesi", *Anadolu Kardiyol Derg*, 11, 619-26.
- [32] M. İyigün, 2021. "Kardiyovasküler ve Torasik Cerrahi", *In Geriatrik Anestezi*, 283-294. *Ankara: Güneş Kitabevi*.
- [33] A. Dar Masroof, and G. Atul, 2019. "Performance Analysis Of Data Mining Algorithms", *Journal Of Computational And Theoretical Nanoscience*, 16(9), 3849-3853.
- [34] M. D. Camacho, M. K. Collins, K. R. Powers, C. J. Costello, and J. J. Collins, 2018. "Next-Generation Machine Learning For Biological Networks", *Cell*, 173(7), 1581-1592.
- [35] M. E. Taşçı, and R. Şamlı, 2020. "Veri Madenciliği ile Kalp Hastalığı Teşhisi", *Avrupa Bilim ve Teknoloji Dergisi*, 88-95.
- [36] A. M. Aydın, 2021. "Müşteri Kaybı Tahmininde Sınıf Dengesizliği Problemi", *Politeknik Dergisi*, 1-1.
- [37] B. Karaoğlu, 2015. "Makine Öğrenmesi ile Spor Karşılaşmalarının Modellenmesi", *Emo Bilimsel Dergi*, 5(9), 1-5.
- [38] E. Uzundurukan, and A. Kara, 2020. "Deep Learning Based Threat Classification for Fiber Optic Distributed Acoustic Sensing Using SNR Dependent Data Generation", *Journal of Scientific, Technology and Engineering Research*, 1(2), 4-12.



- [39] D. Sarkar, R. Bali, and T. Sharma, 2018. "Forecasting Stock And Commodity Prices in Practical Machine Learning With Python, *Apress*, 467-497.
- [40] R. Alizadehsani, M. Roshanzamir, M. Abdar, A. Beykikhoshk, A. Khosravi, M. Panahiazar, A. Koohestani, F. Khozeimeh, S. Nahavandi, and N. Sarrafzadegan, 2019. "A Database For Using Machine Learning And Data Mining Techniques For Coronary Artery Disease Diagnosis", *Scientific Data*, 6(1), 1-13.
- [41] B. Çağlıyan and U. Köse, "Epilepsi EEG Verilerinin Makine Öğrenmesi Teknikleriyle Sınıflandırılması", *Avrupa Bilim ve Teknoloji Dergisi*, no. 23, pp. 163-172.
- [42] J. O. Awoyemi, A. O. Adetunmbi, and S. A. Oluwadare, 2017. "Credit card fraud detection using machine learning techniques: A comparative analysis", *In 2017 International Conference on Computing Networking and Informatics (ICCNi)*, *IEEE*, 2017, 1-9.
- [43] M. Selimoglu, and A. Yılmaz, 2021. "Kredi Kartı Dolandırıcılık Tespitinin Makine Öğrenmesi Yöntemleri ile Tahmin Edilmesi", *Beykent Üniversitesi Fen Ve Mühendislik Bilimleri Dergisi*, 13(2), 28-33.
- [44] H. Yıldırım, 2021. "Comparative Analysis of Machine Learning Algorithms Based on Variable Importance Evaluation", *Journal of Scientific, Technology and Engineering Research*, 2(2), 46-53.
- [45] H. Kabaoglu, E. Uçar, and F. Duran, 2021. "Buzlanma Tahmini Yapan Mobil Uygulama Geliştirilmesi", *Politeknik Dergisi*, 1-1.
- [46] E. Tuncer, and D. E. Bolat, 2021. "Destek Vektör Makinaları ile EEG Sinyallerinden Epileptik Nöbet Sınıflandırması", *Politeknik Dergisi*, 1-1.
- [47] C. Aydın, 2018. "Makine Öğrenmesi Algoritmaları Kullanılarak İtfaiye İstasyonu İhtiyacının Sınıflandırılması", *Avrupa Bilim ve Teknoloji Dergisi*, (14), 169-175.
- [48] K. Arumugam, M. Naved, P. P. Shinde, O. Leiva-Chauca, A. Huaman-Osorio, and T. Gonzales-Yanac, 2021. "Multiple disease prediction using Machine learning algorithms", *Materials Today: Proceedings*.
- [49] A. F. Şenel, 2020. "Makine Öğrenmesi Algoritmaları Kullanılarak Kayısı İç Çekirdeklerinin Sınıflandırılması", *Bitlis Eren Üniversitesi Fen Bilimleri Dergisi*, 9(2), 807-815.
- [50] Ş. B. Bozyiğit, and Ç. Tarhan, 2020. "Makine Öğrenmesi Yöntemleri Kullanılarak Sosyal Medyada Marka İtibarı Analizi", *Yönetim Bilişim Sistemleri Dergisi*, 6(2), 57-76.
- [51] M. Libbrecht, W. Noble, 2015. "Machine Learning Applications in Genetics And Genomics", *Nature Reviews Genetics*, 16(6), 321-332.
- [52] M. Atalay, and E. Çelik, 2017. "Büyük Veri Analizinde Yapay Zekâ ve Makine Öğrenmesi Uygulamaları", *Mehmet Akif Ersoy Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 9(22), 155-172.
- [53] A. Öztürk, Ü. Durak, and F. Badıllı, 2020. "Twitter Verilerinden Doğal Dil İşleme ve Makine Öğrenmesi ile Hastalık Tespiti", *Konya Mühendislik Bilimleri Dergisi*, 8(4), 839-852.
- [54] A. E. Keleş and M. Kaya Keleş, "Genetik Algoritma ile Öznitelik Seçimi Kullanılarak Yapım Yönetiminde Çalışan Liderlik Algısının Tahmini", *Çukurova Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi*, vol. 33, no. 4, pp. 97-110.
- [55] N.B. Bahadure, A. K. Ray, and H. P. Thethi, 2018. "Comparative Approach of MRI-Based Brain Tumor Segmentation and Classification Using Genetic Algorithm", *Journal of digital imaging*, 31(4), 477-489.
- [56] S. Katoch, S. S. Chauhan, and V. Kumar, 2021. "A review on genetic algorithm: past, present, and future", *Multimedia Tools and Applications*, 80(5), 8091-8126.
- [57] H. Candan, A. Durmuş, and G. Harman, 2019. "Genetik Algoritma ve Sınıflandırıcı Yöntemler ile Kanser Tahmini", *Veri Bilimi*, 2(1), 30-34.
- [58] S. Karasu, and Z. Saraç, 2020. "Güç Kalitesi Bozulmalarının Hilbert-Huang Dönüşümü, Genetik Algoritma ve Yapay Zeka/Makine Öğrenmesi Yöntemleri İle Sınıflandırılması", *Politeknik Dergisi*, 23(4), 1219-1229.
- [59] B. C. Kalaycı, O. Ertenlice, H. Akyer, and H. Aygoren, 2017. "A Review On The Current Applications Of Genetic Algorithms in Mean-Variance Portfolio Optimization", *Pamukkale University Journal of Engineering Sciences*, 23(4), 470-476.
- [60] Heart Disease Data Set, Available, <https://archive.ics.uci.edu/ml/datasets/Heart+Disease>, Accessed Time: 15.06.2021
- [61] M. Koklu, S. Sarigil, and O. Ozbek, 2021. "The use of machine learning methods in classification of pumpkin seeds (*Cucurbita pepo* L.)", *Genetic Resources and Crop Evolution*, 1-14.
- [62] P. Ranganathan, S. C. Pramesh, and R. Aggarwal, 2017. "Common Pitfalls In Statistical Analysis: Logistic Regression", *Perspectives in Clinical Research*, 8(3), 148.
- [63] C. P. Austin, and J. Merlo, 2017. "Intermediate And Advanced Topics In Multilevel Logistic Regression Analysis", *Statistics in Medicine*, 36(20), 3257-3277.
- [64] M. Öztürk, 2018. Paksoy T., and Öztürk M., "Genetik Algoritmalar (GA) ile Nüfus Artış Tahmini", *Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi*, 11(1), 40-51.