



# Geleneksel Makine Öğrenmesi Yöntemleri ve Metasezgisel Yöntemlerle Öznitelik Seçim Yöntemlerinin Karşılaştırılması

## Comparison of Traditional Machine Learning Methods and Metaheuristic Methods for Attribute Selection

İsmail AÇAR<sup>1\*</sup>, İbrahim Berkan AYDİLEK<sup>2</sup>

<sup>1</sup> Harran Üniversitesi, Bilgisayar Mühendisliği Bölümü, [ismailacr63@gmail.com](mailto:ismailacr63@gmail.com)

ORCID: <https://orcid.org/0009-0007-9888-7429>

<sup>2</sup> Harran Üniversitesi, Bilgisayar Mühendisliği Bölümü, [berkanaydilek@harran.edu.tr](mailto:berkanaydilek@harran.edu.tr)

ORCID: <https://orcid.org/0000-0001-8037-8625>

### MAKALE BİLGİLERİ

#### Makale Geçmişi:

Geliş 13 Mart 2024  
Revizyon 24 Nisan 2024  
Kabul 17 Mayıs 2024  
Online 30 Haziran 2024

Anahtar Kelimeler:  
Öznitelik Seçimi,  
Metasezgisel algoritmalar,  
Makine öğrenmesi,  
Genetik algoritmalar,  
Parçacık sürü optimizasyon  
algoritması

### ARTICLE INFO

#### Article history:

Received 13 March 2024  
Received in revised form 24 April 2024  
Accepted 17 May 2024  
Available online 30 June 2024

#### Keywords:

Feature selection,  
Metaheuristic Algorithms,  
Machine Learning,  
Genetic Algorithms,  
Particle Swarm Optimization  
Algorithm

Doi: 10.24012/dumf.1451945

\* Sorumlu Yazar

### ÖZ

Öznitelik seçim yöntemleri, makine öğrenmesi için çok önemli bir yere sahiptir. Veri kümesinin boyutu arttıkça makine öğrenmesi modelinin performansı düşmektedir. Öznitelik seçim yöntemi aynı zamanda bir optimizasyon süreci olarak düşünülebilmektedir. Gürültülü olan ya da model için alakasız olan öznitelikler elenince başarı artabilmektedir. Bu çalışma kapsamında UCI veri deposundan 3 farklı veri kümesi kullanılmış ve 5 farklı makine öğrenmesi algoritması kullanılarak 10 katlı çapraz doğrulama yöntemiyle algoritmaların başarımları ve çalışma süreleri ölçülmüştür. Öznitelik seçim yöntemi olarak geleneksel makine öğrenmesi ile öznitelik seçim yöntemleri ve meta sezgisel algoritmalar kullanılarak öznitelik seçim işlemleri gerçekleştirilmiştir. Kullanılan öznitelik seçim yöntemleri şu şekildedir; Korelasyon Analizi, Lasso, Hipotez testi ve metasezgisel algoritmalarından Genetik Algoritma ve Parçacık Sürü Optimizasyon Algoritması yöntemleri kullanılmıştır. Yapılan deneyler sonucunda öznitelik seçim yöntemlerinin uygulanması ile birlikte başarımın arttığı görülmektedir. Çalışma süreleri bakımından algoritmalar karşılaştırıldığında en hızlı algoritmanın Karar ağacı algoritması ve en yavaş çalışan algoritmanın Gradyan Artırma Makineleri (GBM) algoritması olduğu gözlemlenmiştir. Başarım oranına göre değerlendirildiğinde Hafif Gradyan Artırma Makineleri (LighGBM) algoritmasının daha başarılı sonuçlar verdiği görülmüştür. Metasezgisel algoritmalarla yapılan öznitelik seçiminin başarım ve çalışma süresi bakımından daha uygun olduğu gözlemlenmiştir.

### ABSTRACT

Feature selection methods are very important for machine learning. As the size of the dataset increases, the performance of the machine learning model may decrease. The feature selection method can also be considered as an optimization process. Success can be improved by eliminating noisy or irrelevant attributes for the model. In this study, 3 different datasets from the UCI repository were used and 5 different machine learning algorithms were used to measure the performance and runtime of the algorithms using 10-fold cross-validation. As feature selection methods, traditional machine learning feature selection methods and meta-heuristic algorithms were used for feature selection. The feature selection methods used are as follows; Correlation Analysis, Lasso, Hypothesis testing and Genetic Algorithm and Particle Swarm Optimization metaheuristic algorithms. As a result of the experiments, it is seen that the performance increases with the application of feature selection methods. When the algorithms are compared in terms of running times, it is observed that the fastest algorithm is the Decision tree algorithm and the slowest algorithm is the Gradient Boosting Machines (GBM) algorithm. When evaluated according to the success rate, it was observed that the Lightweight Gradient Boosting Machines (LighGBM) algorithm gave more successful results. It was observed that feature selection with metaheuristic algorithms is more suitable in terms of performance and runtime.

## Giriş

Veri kümelerinde boyut sayısı arttıkça eldeki makinelerin işlem yükü artmakta bununla birlikte işlem hızı azalmaktadır. Makine öğrenmesi yöntemlerinde bazen eldeki veri kümesinde çok sayıda öznitelik durumu ile karşı karşıya kalılabilmektedir. Öznitelik seçiminin yapılması makine öğrenmesi algoritmalarının performansı üzerinde doğrudan olumlu sonuçlar doğurabilmektedir. Öznitelik seçim yöntemleri bir optimizasyon süreci olarak düşünülebilir. Model için uygun özniteliklerin seçimi, gereksiz özniteliklerin ortadan kaldırılması algoritmanın başarı oranını artırabilmektedir. Gereksiz olan öznitelikler veri kümesinden atıldığında makine öğrenmesi modellerinde aşırı öğrenme probleminin önüne geçilebilir ve böylelikle daha genellebilir sonuçlar ortaya çıkmış olur. [1]

Bununla birlikte aşırı öğrenme problemleriyle karşılaşılacak modellerde öznitelik seçim yöntemlerinin uygulanması sonucunda modellerin performanslarında iyileşme olduğu görülmüştür. [2]

Bu çalışmada geleneksel makine öğrenmesi yöntemleri ve doğadaki davranışlardan esinlenilerek yapılmış algoritmalar olan metasezgisel algoritmaları ile öznitelik seçim yöntemleri üzerine deneysel çalışmalar yapılmıştır. Geleneksel makine öğrenmesi yöntemleri ile 3 farklı öznitelik seçim yöntemi, metasezgisel algoritmalar ile öznitelik seçimi için 2 farklı algoritma kullanılmıştır. Öznitelik seçim yöntemleri UCI veri deposundan temin edilen 3 farklı veri kümesi üzerinde uygulanarak sonuçlar karşılaştırılmıştır. 3 farklı veri kümesinin kullanılmasının amacı; kullanılan öznitelik seçim yöntemlerinin genel etkisi ve diğer öznitelik seçim yöntemlerine göre başarı durumunu gözlemleyebilmektir.

Makalenin organizasyonu şu şekilde yapılmıştır;

Birinci bölümde, metasezgisel algoritmalar ve geleneksel makine öğrenmesi yöntemleri kullanılarak yapılmış öznitelik seçim yöntemleri ile ilgili çalışmalar incelenerek ilgili yöntemlerin başarılı yönleri gözlemlenmiştir. İkinci bölümde, çalışma kapsamında kullanılan beş farklı öznitelik seçim yöntemleri ve kullanılan beş farklı makine öğrenmesi sınıflandırma yöntemleri hakkında detaylı bilgi verilmiştir. Üçüncü bölümde kullanılan veri kümeleri hakkında bilgi verilerek , veri kümeleri üzerinde beş farklı öznitelik seçim yöntemi uygulanarak öznitelik sayısı indirgenmiş daha sonra beş makine öğrenmesi sınıflandırma yöntemi 10 katlı çapraz doğrulama kullanılarak deneyler yapılmış, çalışma süreleri ve 10 katlı çapraz doğruluk başarımları tablolarda gösterilmiş ve metasezgisel yöntemler ile geleneksel makine öğrenmesi yöntemlerinin öznitelik seçim yönündeki başarımları 3 farklı grafikte gösterilmiştir. Dördüncü bölümde ise öznitelik seçim yöntemlerinin öneminden ve

ilgili çalışmanın genel olarak başarılı ve başarısız yönleri ele alınmıştır.

## İlgili Çalışmalar

Literatürde öznitelik seçim yöntemleri yaygın kullanılan bir yöntemdir. Öznitelik seçim yöntemi, ilk defa 1970'li yılların başlarında kullanılmaya başlanmıştır. Öncelikli olarak istatistiksel modelleme ve desen tanıma problemlerinde kullanılmıştır. Bu dönemde öznitelik seçimiyle ilgili bir takım erken çalışmalar bulunuyordu, "Entropi Kavramının Örüntü Tanımadaki Kullanımı" adlı çalışma bu konuda yapılmış ilk çalışmadır. Bu çalışmayla birlikte öznitelik seçim yöntemlerinin gelişim ve yaygınlığı zamanla daha da artmış ve gittikçe karmaşıklaşmıştır. Öznitelik seçim yöntemleri makine öğrenmesi için çokça başvurulan bir konu olarak ele alınmaktadır. [3]

Ele alınan konu ile ilgili literatürde yer alan belli başlı çalışmalar incelenmiş ve kısaca bahsedilmiştir.

*Öznitelik Seçimi Problemleri İçin İkili Beyaz Köpekbalığı Optimizasyon Algoritması* adlı çalışmada öznitelik seçim problemi için beyaz köpek balığı optimizasyon algoritmasını ele almıştır. Bu meta sezgisel algoritma öznitelik seçimi için başarılı sonuçlar vermektedir. (Beyaz Köpek Balığı Optimizasyon Algoritmasında S, V ve U şekilli transfer fonksiyonlarının ikili forma dönüştürülmesi sonucunda öznitelik seçimi kullanılmıştır). [4]

Genetik Algoritma ile Özellik Seçimi Kullanılarak Makine Öğrenimi Algoritması ile Kalp Hastalığı Tahmini adlı çalışmada kalp veri kümesi üzerinde 7 farklı algoritma kullanılarak 3 deney yapılmıştır. 3 grup deneyler sonucunda toplamda 28 deney gerçekleştirilmiştir. Çalışmalardan elde edilen sonuçlar incelendiğinde en başarılı sonucun Genetik Algoritma ile %93.44 başarımlarının elde edildiği görülmüştür. Çalışma kapsamında Genetik Algoritma ve Makine Öğrenmesi yöntemleri kullanılarak öznitelik seçim yöntemleri yapıldığında kalp hastalığının daha başarılı bir şekilde tahmin edildiği görülmüştür. [5]

Öğretmenlerde Etkili Psikososyal Risk Düzeylerinin Belirlenmesi için Parçacık Sürü Optimizasyonu Tabanlı Öznitelik Seçme adlı çalışmada Kolombiyadaki devlet okullarında görev yapan öğretmenlerden toplanmış veriler kullanılmıştır. Veri kümesi 123 psikososyal öznitelikten ve 4 sınıftan oluşmaktadır. Bu çalışmada parçacık sürü optimizasyon algoritması kullanılarak öznitelik seçimi yapılmış ve 5 katlı doğrulama yöntemi ile karar ağacı, destek vektör makineleri , k-en yakın komşu , yapay sinir ağı sınıflandırıcıları kullanılarak deneyler yapılmıştır. Elde edilen sonuçlar incelendiğinde destek vektör makinelerinin %99.9 başarımları ile diğer yöntemlere göre daha başarılı sonuçlar verdiği görülmüştür. [6]

*Ağırlıklandırılmış, Koşullu Karşılıklı Bilgi İle Öznitelik Seçimi* adlı çalışmada öznitelik seçiminde daha etkin bir yöntem önerisi yapılmıştır. Önerilen yöntemde, sezgisel bir fonksiyon kullanılarak özniteliklerin seçimi üzerindeki fazlalık etkisi, hem sınıfla olan hem de diğer özniteliklerle olan karşılıklı bilgi ilişkileri göz önünde bulundurularak değerlendirilmiştir. Böylelikle, her problem için dinamik bir çözüm sunarak öznitelik seçiminde daha başarılı sonuçlar elde edilmesi sağlanmıştır. [7]

*Yüz İfadesi Tanıma için Mesafe Oranlarına Dayalı Öznitelik Çıkarımı ve Genetik Algoritmalar ile Seçimi* adlı çalışmada 4 erkek ve 3 kadın katılımcıdan toplanan görüntüler ile yüz işaretçileri arasındaki mesafe oranları baz alınarak öznitelik çıkarımı yapılmıştır. Öznitelik çıkarımı sonucunda 120 tane öznitelik elde edilmiştir. Daha sonra 4 farklı öznitelik seçimi yöntemi ile öznitelik sayısı azaltılarak sonuçlar karşılaştırılmıştır. Genetik algoritmanın öznitelik seçiminde diğer yöntemlere göre daha başarılı sonuçlar verdiği görülmüştür. Bununla birlikte sınıflandırıcı olarak rastgele orman algoritması kullanılmıştır. [8]

*EEG Sinyallerinin Sınıflandırılmasında Evrimsel Öznitelik Seçim Metotlarının Kullanılması* çalışmasında Ayrık Dalgacık dönüşümü ve Bir Boyutlu Yerel İkili Örüntü yöntemleri kullanılarak öznitelik çıkarımı ve differansiyel evrim tabanlı yöntemi ile öznitelik seçimi yapılarak elde edilen sonuçlar değerlendirilmiştir. Elde edilen sonuçlar incelendiğinde başarı oranlarının literatürdeki çalışmalara göre daha başarılı olduğu gözlemlenmiştir. [9]

*Siber Saldırıları için Rastgele Orman Algoritması Kullanılarak Öznitelik Seçimi* çalışmasında Dört farklı öznitelik seçimi yöntemi kullanılarak sonuçlar karşılaştırılmıştır. Öznitelik sayıları 13, 6, 5 ve 4 yapılarak deneyler yapılmıştır. Rastgele Orman algoritması sınıflandırıcısı kullanılarak 4 öznitelik ile %97.24 başarı oranı ile diğer yöntemlere göre daha başarılı sonuçlar verdiği gözlemlenmiştir. Böylece öznitelik sayısının azalması ile birlikte çalışma süresinin azaldığı görülmüştür. [10]

*Filtre Tabanlı Öznitelik Seçim Yöntemleri Kullanılarak Metinlerde Duygu Sınıflandırması Üzerine Karşılaştırmalı Bir Çalışma* adlı çalışmada duygu sınıflandırma problemi için filtre tabanlı öznitelik seçim yöntemleri kullanılarak sonuçlar karşılaştırılmıştır. Çalışma kapsamında 9 veri kümesi, 6 öznitelik seçim yöntemi ve 3 sınıflandırma algoritması kullanılarak 4 farklı değerlendirme metriği ile elde edilen sonuçlar değerlendirilmiştir. Elde edilen sonuçlar incelendiğinde en başarılı sonucun Naive Bayes Algoritması kullanılarak ortalama %94.34 doğruluk oranına ulaştığı gözlemlenmiştir. [11]

*Rüzgâr Gücü Tahmininde Genetik Algoritma ile Öznitelik Seçimi* adlı çalışmada SCADA programı ile elde edilen veri kümesi ile NASA'nın meteorolojik veri kümesi kullanılarak rüzgâr gücü tahmini yapılmıştır. Çalışma kapsamında 5 makine öğrenmesi algoritması kullanılarak 3 farklı

değerlendirme metriği ile sonuçlar karşılaştırılmıştır. Genetik algoritma ile öznitelik seçimi yapılarak öznitelik sayısı 47'den 9'a azaltılarak modelin başarımı artırılmıştır. R2 metriği kullanılarak 0.98 doğruluk oranına ulaşılmıştır. [12]

*Parkinson Hastalarının Tespitinde Karınca Koloni Algoritması ile Seçilen Özniteliklerin Performansa Etkisi* adlı çalışmada Parkinson hastalarına ait ses verileri kullanılarak öznitelik çıkarımı yapılmıştır. Daha sonra gereksiz olan öznitelikleri tespit etmek için karınca koloni algoritması kullanılarak öznitelik sayısı azaltılmıştır. Elde edilen sonuçlar incelendiğinde seçilen özniteliklerin hastalık tespitinde oldukça önemli olduğu görülmüştür. [13]

*Salp Sürü Algoritması ile Öznitelik Seçimi ve Sınıflandırıcı Performans Değerlendirmesi* adlı çalışmada 3 veri kümesi ve 3 sınıflandırma algoritması kullanılmıştır. Bu çalışmada 3 veri kümesi üzerinde salp sürü algoritması ile öznitelik seçimi yapılarak sınıflandırma algoritmaları ile sonuçlar karşılaştırılmıştır. Modellerin başarımlarını ölçmek için 7 farklı değerlendirme metriği kullanılmıştır. Elde edilen sonuçlar göz önüne alındığında çalıştırma süresi bakımından k-en yakın komşu algoritmasının diğer algoritmalara göre daha hızlı çalıştığı gözlemlenmiştir. Seçilen özniteliklerin sayılarına göre destek vektör makineleri ve rastgele orman algoritmasının daha başarılı sonuçlar verdiği görülmüştür. [14]

## Materyal ve Yöntem

Bu çalışmada Öznitelik seçim yöntemleri için geleneksel makine öğrenmesi yöntemlerinden Korelasyon Analizi, Hipotez testi ve Lasso yöntemleri kullanılmıştır. Meta sezgisel algoritmalarla öznitelik seçimi için Genetik Algoritma ve Parçacık Sürü Optimizasyon yöntemleri kullanılmıştır. Seçilen özniteliklerin sınıflandırma başarımı için Karar Ağacı, Rastgele Orman, Gradyan Artırma Makineleri (GBM), Hafif Gradyan Artırma Makineleri (LightGBM) ve Aşırı Gradyan Artırma Makineleri (XGBoost) algoritmaları kullanılmış ve 10 katlı çapraz doğrulama yöntemi ile model başarısı ölçülmüştür. Kullanılan yöntemler alt başlıklarda açıklanmıştır.

## Öznitelik seçimi Yöntemleri

Çalışma kapsamında 5 farklı öznitelik seçim yöntemi kullanılmıştır.

### Korelasyon Analizi:

Korelasyon analizi -1 ile 1 arasında değerler alır. +1 değerine doğru yaklaştıkça pozitif yönlü ilişki artarken, -1 değerine doğru yaklaştıkça negatif yönlü ilişki artmaktadır. Korelasyon iki değişken arasındaki ilişkiyi ifade eder. Öznitelik seçimi yöntemi için sıkça kullanılan bir yöntemdir. [15]

## Hipotez Testi:

Hipotez testi oluşturulurken özniteliklerin katsayılarına yönelik anlamlılık değerleri belirlenir. Anlamlılık değeri istatistikte p ile ifade edilir, genellikle 0.01 ve 0.05 seçilir. P değeri belirlenmiş olan anlamlılık değerinden düşük olursa öznitelik model için istatistiksel olarak anlamlıdır ve modelde tutulabilirken büyük olduğu durumda istatistiksel olarak anlamsız ve modelden çıkartılır. [16]

## Lasso Yöntemi

Lasso yöntemi katsayılara ceza terimi uygulayarak katsayıları sıfıra kadar indirgemiş olur. Lasso yöntemi katsayıları sıfıra indirgemesi sayesinde hem aşırı öğrenme problemlerinin önleme hem de öznitelik seçim işlemleri için oldukça kullanışlı bir yöntemdir. [17]

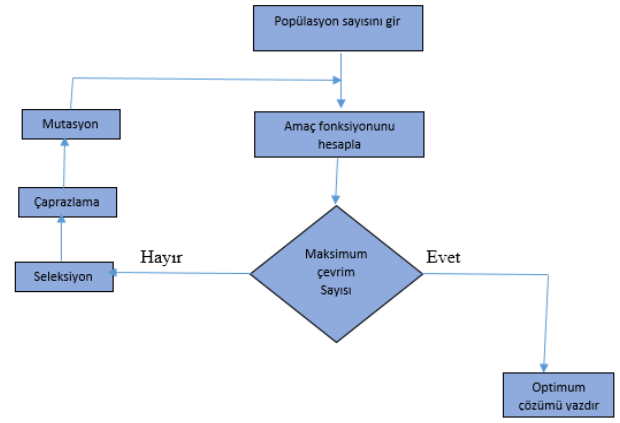
## Genetik Algoritma

Genetik algoritma Darwin'in doğal seçim yönteminden esinlenilerek geliştirilmiş optimizasyon yöntemidir. Bu yöntem 1975 yılında John Holland tarafından ortaya atılmıştır. [18]

Genetik algoritma doğada en iyinin hayatta kalarak yeni nesiller oluşmasına, güçsüz olan bireylerin elenmesi prensibine dayanmaktadır. Genetik algoritma öznitelik seçim yönteminde sıklıkla kullanılan bir yöntemdir.

Genetik algoritmada öncelikli olarak parametre kodlaması işlemi gerçekleştirilir. Parametre kodlamasında eldeki veriler bilgisayarın anlayabileceği şekilde kodlanır. Sıklıkla kullanılan parametre kodlaması binary kodlaması, permütasyon kodlama, değer kodlama ve ağaç kodlama yöntemleri kullanılmaktadır. Kullanılan yöntem mevcut probleme göre belirlenir. İkinci adımda başlangıç popülasyonu da belirlenerek başlangıç popülasyonu sayısı probleme göre şekillendirilir. Başlangıç popülasyonu çok büyük veya az büyük seçilmemeli, ikisi arasında optimum sayı belirlenmelidir. Üçüncü aşamada uygunluk değeri hesaplanır. Burada her bir çözümün ne kadar iyi olduğu hesaplanır. Dördüncü aşamada kromozom seçimi işlemi gerçekleştirilir. Kromozom seçimi için rulet tekeri, sıralı seçim, sabit durum ve turnuva yöntemi seçimleri kullanılabilir. Hangi yöntemin daha iyi olduğu probleme göre değişmektedir. Beşinci adımda çaprazlama ve mutasyon yöntemleri uygulanır. Çaprazlama işleminde ebeveynlerin genleri karşılıklı değiştirilir böylelikle anne ve babalarından daha farklı bireyler meydana gelmiş olur. Çaprazlama için; tek noktalı, çift noktalı, uniform ve aritmetik gen takası yöntemleri bulunmaktadır. Mutasyon işlemini nesillerdeki kromozom çeşitliliğini artırmak için kullanılmakta böylelikle farklı kromozom yapısına sahip bireyler meydana gelmiş olmaktadır. Son aşamada durdurma kriteri sağlanıyorsa algoritma durdurulur. [19]

Genetik Algoritmanın akış şeması Şekil 1'de verilmiştir.



Şekil 1. Genetik Algoritma akış şeması

## Parçacık Sürü Optimizasyon Algoritması:

Parçacık sürü optimizasyon algoritması doğada birlikte hareket eden kuş ve balık sürülerinin yiyecek arama ve dış tehlikelere karşı kenetlenmelerinden esinlenilerek 1995 yılında James Kennedy ve Russell Eberhart tarafından geliştirilmiş bir meta sezgisel algoritmadır. [20]

Parçacık sürü optimizasyon algoritmasında her bir parçacık bir kuşu veya balığı temsil eder. Parçacıklar rastgele değerle başlatılır ve hız ve konum değerleri güncellenir. Tüm parçacıkların en iyi çözümü genel en iyi, parçacığın en iyi çözümü kişisel en iyi olarak ifade edilir. Parçacık sürü optimizasyon algoritmasında parametre sayısı az olduğu için oldukça kullanışlı bir yöntemdir. [21]

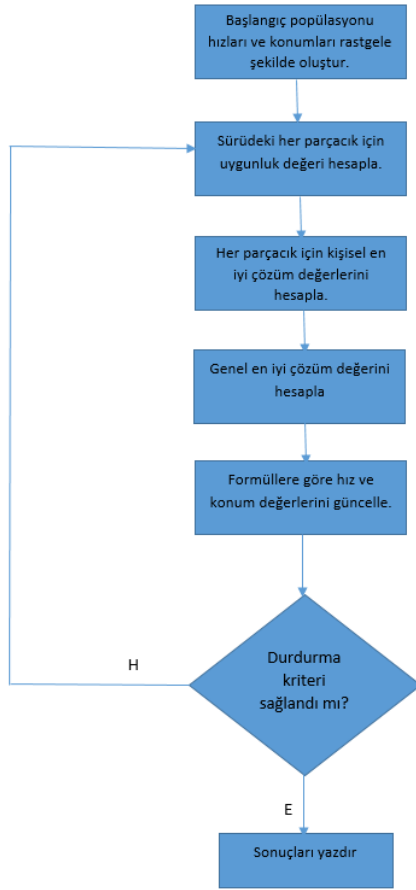
Parçacık sürü optimizasyon algoritmasının denklemleri (1) ve (2)'de verilmiştir.

$$vi(t + 1) = w.vi(t) + c1.r1.(Pbest - xi(t)) + c2.r2.(Gbest - xi(t)) \quad (1)$$

$$xi(t + 1) = xi(t) + vi(t + 1) \quad (2)$$

Pbest kişisel en iyi çözümü ve Gbest genel en iyi çözümü ifade ederken, v hızı, w atalet ağırlığını, x konumu, c1 ve c2 pozitif ivme sabitlerini, r1 ve r2 ise 0 ile 1 arasında rastgele sayısal değerleri ifade etmektedir.

Parçacık Sürü Optimizasyon Algoritmasının akış şeması Şekil 2'de verilmiştir.



Şekil 2. Parçacık Sürü Optimizasyon Algoritması Akış Şeması

### Sınıflandırma Yöntemleri:

Çalışma kapsamında 5 farklı sınıflandırma algoritması kullanılmıştır.

### Karar Ağacı

Karar ağaçları, sınıflandırma ve regresyon problemleri için sıklıkla kullanılan bir makine öğrenmesi algoritmasıdır. Karar ağaçlarında kökten, dal ve yapraklara doğru hiyerarşik olarak bölünme prensibi temel alınır. Karar Ağaçlarında hangi öznelikten itibaren bölüneceğine bilgi kazancı ve gini gibi yöntemler kullanılarak karar verilir. [22]

Karar Ağaçları algoritması eğitimi hızlı, ayrıca uzaklık temelli bir algoritma olmadığı için normalizasyon işlemine gerek duyulmaz. Öte yandan bu yöntemlerde aşırı öğrenme problemi ile sıkça karşılaşmaktadır. Bunun önüne geçilmek için çok sayıda ağaçtan oluşan topluluk öğrenme yöntemleri geliştirilmiştir.

### Rastgele Orman

Rastgele orman algoritması, birden fazla karar ağacının bir araya gelerek toplu bir şekilde öğrenme yöntemine dayanan bir sınıflandırma ve regresyon algoritmasıdır. Karar ağaçları algoritmasında aşırı öğrenme problemi ile sıklıkla karşılaşmaktadır. Bu algoritma, aşırı öğrenme problemine

bir çözüm olarak ortaya atılmış bir makine öğrenmesi algoritmasıdır. [23]

Rastgele orman algoritması topluluk öğrenme yöntemlerinden biridir. Rastgele orman, karar ağaçları algoritmasının aşırı öğrenme problemi önüne geçilmesi için geliştirilmiştir. Öte yandan bu yöntemde hiperparametre sayısı çok sayıda olduğu için modelin ayarlanması zorlaşabilmektedir. Çalışma süresi bakımından incelendiğinde karar ağaçlarına göre daha yavaş çalıştığı gözlemlenmektedir.

### Gradyan Artırma Makineleri (GBM)

Gradyan Artırma Makineleri, zayıf sınıflandırıcıları baz alarak geliştirilmiş bir gradyan artırma algoritmasıdır. Zayıf sınıflandırıcıları bir araya getirerek güçlü sınıflandırıcılar oluşturma esasına dayanan bir yöntemdir. [24]

GBM algoritması, zayıf sınıflandırıcıların hatalardan ders çıkararak güçlü sınıflandırıcılar oluşturması fikrine dayalı çalıştığından başarılı sonuçlar verdiği görülmektedir. Ayrıca topluluk öğrenme yöntemine dayalı olduğu için aşırı öğrenme problemine karşı güçlü olduğu gözlemlenmektedir. Öte yandan model eğitim süreci çok uzun sürmektedir.

### Hafif Gradyan Artırma Makineleri (LightGBM)

LightGBM algoritması, GBM algoritmasının hız ve tahmin performansını iyileştirmek amacıyla geliştirilmiş bir algoritmadır. Yüksek işlem hızı, Gpu desteği ve büyük veri kümelerini işleyebilme yeteneği sayesinde oldukça kullanışlı bir yöntemdir. [25]

Lightgbm algoritması, diğer algoritmalara göre daha hızlı tahminler üretmekte, aşırı öğrenme problemine karşı güçlü olduğu gözlemlenmektedir. Öte yandan hiperparametre sayısının fazla olmasından dolayı hiperparametre ayarının yapılması zorlaşabilmektedir.

### Aşırı Gradyan Artırma Makineleri (XGBoost)

XGBoost algoritması, GBM algoritmasını optimize ederek daha kullanışlı bir yöntem ortaya çıkarmaktadır. Aşırı öğrenme probleminin önüne geçmek, eksik verilerle çalışabilmesi ve hızlı işlem yapması yönleriyle kullanışlı bir algoritmadır. Diğer makine öğrenmesi algoritmalarına göre birçok problem için daha çok öne çıkmaktadır. [26]

XGBoost algoritması, büyük veri kümeleri üzerinde hızlı ve başarılı tahminler vermektedir. Ayrıca diğer topluluk öğrenme yöntemlerinde olduğu gibi aşırı öğrenme problemine karşı dirençli olduğu görülmektedir. Öte yandan hiperparametre sayısının çokluğundan dolayı ayarlanması uzun sürmektedir.

### DeneySEL Çalışmalar

Bu bölümde deney kapsamında kullanılan veri kümeleri ve deneylerden elde edilen sonuçlardan bahsedilecektir.

**Kullanılan Veri kümeleri:****Kullanılan Veri kümeleri:**

Bu çalışma kapsamında UCI veri deposundan temin edilen 3 farklı veri kümesi üzerinde deneyler yapılmıştır. Tablo 1'de veri kümeleri hakkında bilgi verilmiştir.

Sonar: 61 öznitelikten ve 208 gözlemden oluşmaktadır. Denizaltıların ve su altı nesnelerin sınıflandırılması amacıyla hazırlanmıştır. Bu veri kümesinde kaya ve maden olmak üzere 2 farklı sınıf kullanılmıştır.

Parkinson: 23 öznitelik ve 195 gözlemden oluşmaktadır. Parkinson hastalığı beyinde dopamin üreten hücrelerin azalması sonucunda ortaya çıkan bir hastalıktır. Veri kümesi, hastalığa etki eden 23 özellik belirlenerek

hastalığın erken teşhisini kolaylaştırmak için çeşitli hastalardan toplanmıştır. Veri kümesi hasta olma ve hasta olmama durumlarını ifade eden 2 sınıftan oluşmaktadır.

Breast Cancer: 32 öznitelik ve 569 gözlemden oluşmaktadır. Meme kanseri, genellikle kadınlarda belli bir yaştan sonra meme dokularının kontrolsüz büyümesi sonucu ortaya çıkan bir hastalıktır. Veri kümesi, meme kanserine etki eden 31 özellik belirlenerek hastalığın erken teşhisini kolaylaştırmak için oluşturulmuştur. Veri kümesi hasta olma ve hasta olmama durumlarını ifade eden 2 sınıftan oluşmaktadır.

Tablo 1. Kullanılan veri kümeleri

Veri kümesi	Öznitelik Sayısı	Gözlem Sayısı
Sonar	61	208
Parkinson	23	195
Breast Cancer	32	569

**Bulgular ve Tartışma:**

Çalışma kapsamında 3 farklı veri kümesi üzerinde 5 farklı öznitelik seçim yöntemi uygulanmış ve orijinal durumdaki hali ile karşılaştırılmıştır. Deneyler kapsamında 5 farklı makine öğrenmesi algoritması 10 katlı çapraz doğrulama yöntemiyle uygulanmıştır.

Her bir veri kümesi için sınıflandırma doğruluk başarımları ve milisaniye cinsinden çalışma süreleri hesaplanmış, öznitelik seçim yöntemi için en yüksek doğruluk başarımları değerleri koyu olarak gösterilmiştir.

Her bir veri kümesi için öznitelik seçim yöntemleri uygulandıktan sonra ortaya çıkan 10 katlı çapraz doğrulama sonuçları ve çalışma süreleri üç farklı veri kümesi için elde edilen sonuçlar değerlendirilmiştir.

Sonar veri kümesi için 5 farklı öznitelik seçim yöntemi ve 5 makine öğrenmesi algoritması kullanılarak elde edilen sonuçlar Tablo 2'de verilmiştir. Deney sonuçları, 10 katlı çapraz doğrulama ve çalışma süreleri bakımından öznitelik seçim yöntemi uygulandıktan sonra sonuçlar veri kümesinin orijinal hali ile karşılaştırılmıştır.

Sonar veri kümesi için deneylerden elde edilen sonuçlar Tablo 2'de verilmiştir. Elde edilen sonuçlar gözlemlendiğinde öznitelik seçim yöntemleri uygulandıktan sonra doğruluk başarımları oranları artmıştır.

Veri kümesinde öznitelik seçimi yapılmadan algoritmaların başarımları oranları incelendiğinde LightGBM algoritması daha başarılı sonuçlar vermiştir. Öznitelik seçim yöntemleri uygulandıktan sonra Parçacık sürü optimizasyon algoritması 39 öznitelik ile 0.8434 doğruluk başarımları oranını, Genetik algoritma 31 öznitelik ile 0.8382 doğruluk başarımları oranını vermiştir. Deney sonuçları incelendiğinde metasezgisel algoritmaların öznitelik seçiminde oldukça başarılı sonuçlar verdiği gözlemlenmiştir.

Tablo 2. Sonar Veri kümesi Deney sonuçları (10 cv doğruluk sonuçları [çalışma süresi (ms)])

Kullanılan öznelik seçim yöntemleri	Öznelik Sayısı	Karar Ağacı için 10-cv ms süreleri	Rastgele Orman için 10-cv ms süreleri	Gradyan Artırma Makineleri için 10-cv ms süreleri	Hafif Gradyan Artırma Makineleri için 10-cv ms süreleri	Aşırı Gradyan Artırma Makineleri için 10-cv ms süreleri
Veri kümesinin orijinal hali	61	0.7063[52.71]	0.7835[1219.56]	0.8070[3234.79]	<b>0.8312</b> [472.07]	0.8140[1819.97]
Korelasyon Analizi yöntemi (thres:0.8)	38	0.6761[43.82]	0.7526[1335.06]	<b>0.8125</b> [2435.82]	0.7834[324.22]	0.7408[441.61]
Lasso yöntemi	35	0.6816[40.74]	0.8316[1350.40]	0.7949[2455.34]	<b>0.8319</b> [331.93]	0.8257[422.69]
Hipotez testi yöntemi	36	0.7349[75.52]	<b>0.8011</b> [1316.84]	0.7710[2652.90]	0.7709[343.23]	0.7772[440.01]
PSO yöntemi	39	0.7118[42.71]	<b>0.8434</b> [1458.02]	0.8074[2168.17]	0.8136[336.69]	0.8132[462.86]
Genetik Algoritması yöntemi	31	0.7360[60.58]	0.7651[1273.89]	0.8250[2090.06]	<b>0.8382</b> [334.80]	0.8313[424.39]

Parkinson veri kümesi için deneylerden elde edilen sonuçlar Tablo 3'te verilmiştir. Elde edilen sonuçlar incelendiğinde öznelik seçimi yapıldıktan sonra başarı oranının arttığı gözlemlenmiştir. Veri kümesinin orijinal hali incelendiğinde lightgbm algoritması daha başarılı sonuçlar vermiştir.

seçiminde 0.8971 doğruluk oranı ile daha başarılı sonuçlar vermiştir.

Öznelik seçim yöntemi uygulandıktan sonra Genetik algoritma ile öznelik sayısı 23 ten 12'ye düşmüş ve başarı oranının arttığı görülmüştür. Genetik algoritma öznelik

Tablo 3. Parkinson Veri kümesi Deney sonuçları (10 cv doğruluk sonuçları [çalışma süresi (ms)])

Kullanılan öznelik seçim yöntemleri	Öznelik Sayısı	Karar Ağacı için 10-cv ms süreleri	Rastgele Orman için 10-cv ms süreleri	Gradyan Artırma Makineleri için 10-cv ms süreleri	Hafif Gradyan Artırma Makineleri için 10-cv ms süreleri	Aşırı Gradyan Artırma Makineleri için 10-cv ms süreleri
Veri kümesinin orijinal hali	23	0.7881[66.65]	0.8624[2161.30]	0.8762[2512.22]	<b>0.8823</b> [469.69]	0.8690[417.57]
Korelasyon Analizi yöntemi (thres:0.8)	10	0.7933[44.59]	0.8638[1883.91]	0.8700[1426.75]	0.8771[349.45]	<b>0.8848</b> [369.73]
Lasso yöntemi	12	0.8148[69.20]	0.8767[2060.26]	0.8690[1586.41]	<b>0.8828</b> [399.42]	0.8548[381.95]
Hipotez testi yöntemi	15	0.8219[70.05]	0.8424[1848.64]	<b>0.8690</b> [1680.00]	0.8623[402.22]	0.8629[394.16]

<b>PSO yöntemi</b>	19	0.8500[81.49]	0.8552[1789.30]	0.8767[1851.08]	0.8828[399.97]	<b>0.8838</b> [399.68]
<b>Genetik Algoritması yöntemi</b>	12	0.8576[63.58]	0.8762[1879.56]	0.8771[1446.14]	<b>0.8971</b> [387.69]	0.8938[389.71]

Breast Cancer veri kümesi kullanılarak elde edilen deney sonuçları Tablo 4' te verilmiştir. Deney sonuçları incelendiğinde öznelik seçim yöntemleri uygulandıktan sonra başarı oranının azaldığı gözlemlenmiştir. Veri kümesi orijinal hali incelendiğinde lightgbm algoritması ile 0.9766

başarı oranı elde edilirken, öznelik seçimi yapıldıktan sonra doğruluk başarı oranı 0.8969' a düşmüştür. Breast Cancer veri kümesinde, metasezgisel algoritmalar öznelik seçiminde diğer yöntemlere göre daha başarısız sonuçlar vermiştir

Tablo 4. Breast Cancer Veri kümesi Deney sonuçları (10 cv doğruluk sonuçları[çalışma süresi(ms)])

Kullanılan öznelik seçim yöntemleri	Öznelik Sayısı	Karar Ağacı için 10-cv ms süreleri	Rastgele Orman için 10-cv ms süreleri	Gradyan Artırma Makineleri için 10-cv ms süreleri	Hafif Gradyan Artırma Makineleri için 10-cv ms süreleri	Aşırı Gradyan Artırma Makineleri için 10-cv ms süreleri
<b>Veri kümesinin orijinal hali</b>	32	0.9225[169.16]	0.9508[4871.24]	0.9602[12285.11]	<b>0.9766</b> [3838.36]	0.9648[1334.73]
<b>Korelasyon Analizi yöntemi (thres:0.8)</b>	15	0.8969[134.11]	0.9461[4088.72]	0.9414[6482.11]	<b>0.9484</b> [3035.88]	0.9483[1154.30]
<b>Lasso yöntemi</b>	15	0.9155[120.44]	0.9554[3862.01]	0.9530[6312.27]	<b>0.9578</b> [3199.66]	0.9601[381.95]
<b>Hipotez testi yöntemi</b>	20	0.9342[120.82]	0.9555[4497.90]	0.9508[8179.60]	<b>0.9648</b> [3215.57]	0.9532[1311.23]
<b>PSO yöntemi</b>	14	0.9131[106.65]	0.9438[4285.74]	0.9484[6228.71]	<b>0.9650</b> [2865.00]	0.9454[1136.01]
<b>Genetik Algoritması yöntemi</b>	16	0.9083[108.91]	0.9508[4192.99]	0.9508[7122.63]	<b>0.9555</b> [387.69]	0.9530[1104.11]

Öznelik seçim yöntemlerinin başarımlarını sağlıklı bir şekilde karşılaştırabilmek için tüm veri kümeleri için aynı parametre değerleri kullanılmıştır. Elde edilen sonuçlar ve çalışma süreleri incelendiğinde kullanılmış olan 3 veri kümesinin 2 tanesinde başarı oranı artmıştır. Ayrıca elde edilen sonuçların başarımlarını ölçmek için diğer algoritmalara göre daha başarılı sonuçlar veren Hafif gradyan artırma makineleri algoritması ile 3 farklı veri kümesinin ortalama doğruluk, fl-skor, duyarlılık ve hassasiyet sonuçları Şekil 6 üzerinde gösterilerek karşılaştırılmıştır.

5 farklı algoritmanın 3 veri kümesi üzerindeki çalışma sürelerini incelendiğinde çalışma sürelerinin bazı algoritmalarda azalırken bazı algoritmalarda arttığı görülmüştür. 5 farklı algoritma özelinde tüm veri kümeleri üzerinde çalışma süreleri incelendiğinde en hızlı çalışan algoritmanın Karar Ağacı algoritması olduğu, en yavaş çalışan algoritmanın Gradyan Artırma Makineleri algoritması olduğu görülmüştür. Algoritmalar bazında

öznelik seçim yöntemleri uygulandıktan sonra çalışma süreleri incelendiğinde; Karar Ağacı algoritması için Breast Cancer veri kümesinde kullanılan tüm öznelik seçim yöntemleri için çalışma süreleri azalmıştır. Rastgele Orman algoritmasında Parkinson ve Breast Cancer veri kümelerinde kullanılan tüm öznelik seçim yöntemleri için çalışma süreleri azalmıştır. Gradyan Artırma Makineleri (GBM) algoritmasında 3 veri kümesinde de kullanılan tüm öznelik seçim yöntemleri için çalışma süreleri azalmıştır. Hafif Gradyan Artırma Makineleri (LightGBM) algoritmasında 3 veri kümesinde de kullanılan tüm öznelik seçim yöntemleri için çalışma süreleri azalmıştır.

Aşırı Gradyan Artırma Makineleri (XGBoost) algoritmasında kullanılan 3 veri kümesinde de tüm öznelik seçim yöntemleri için çalışma süreleri azalmıştır. Kullanılmış olan 3 veri kümesi içinden Breast Cancer veri kümesinde tüm algoritmalar için kullanılan tüm yöntemlerde çalışma süresinin azaldığı gözlemlenmiştir.



Parkinson veri kümesinde hem başarı oranının arttığı hem de çalışma sürelerinin azaldığı ortaya konmuştur.

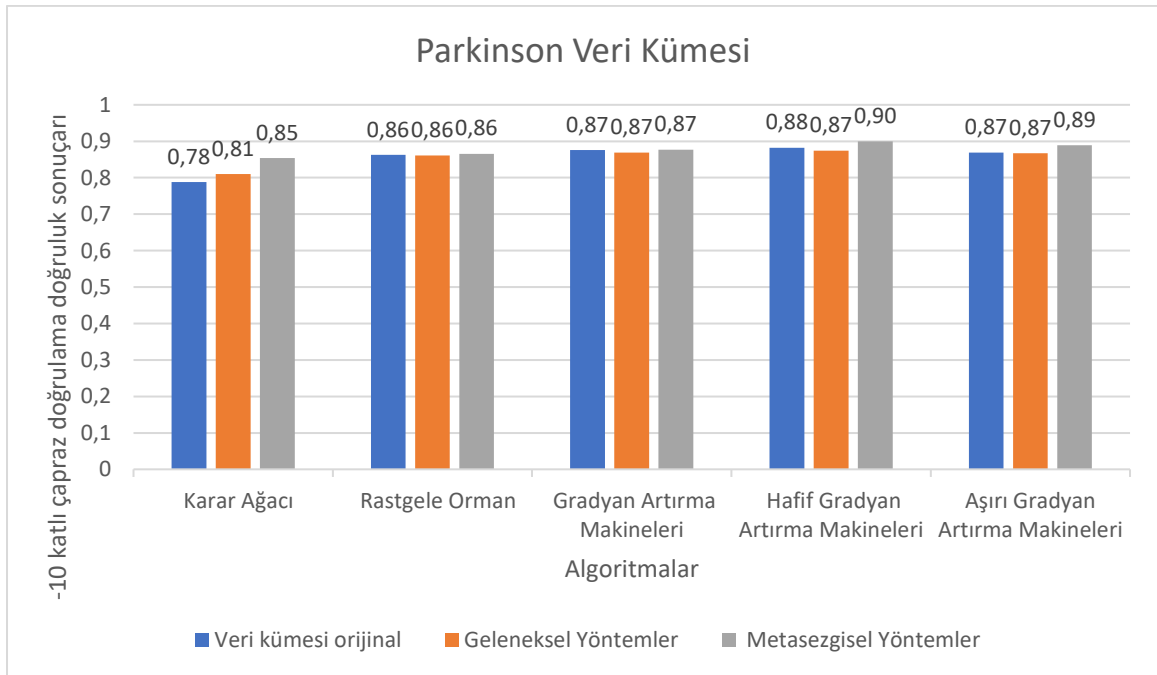
Deneylerden elde edilen sonuçlar incelendiğinde meta sezgisel algoritmalarla parçacık sürü optimizasyon ve genetik algoritmanın öznelik seçiminde geleneksel makine öğrenmesi yöntemlerinden korelasyon analizi, lasso ve hipotez testine göre daha başarılı sonuçlar verdiği görülmüştür.

Genel olarak incelenen 3 veri kümesi üzerinde uygulanmış olan öznelik seçim yöntemlerinde başarılı sonuçlar elde edildiği görülmüştür.

Metasezgisel yöntemler ve geleneksel yöntemlerin öznelik seçimindeki etkilerini gözlemleyebilmek için deneylerde kullanılmış olan metasezgisel yöntemlerin ortalamaları ve geleneksel yöntemlerin ortalamaları alınıp her bir algoritma için veri kümesinin orijinal hali ile karşılaştırılmıştır.

Şekil 3'te Parkinson veri kümesi için kullanılan yöntemlerin ortalama doğruluk başarımları ve veri kümesinin orijinal halinin doğruluk başarımları gösterilmiştir. Parkinson veri kümesi için genel olarak sonuçlar incelendiğinde, Karar Ağacı algoritması veri

kümesinin orijinal halinde 0.78 ile en düşük başarımlarını verirken, Hafif Gradyan Artırma Makineleri(LightGBM) algoritması, metasezgisel algoritmalar ile öznelik seçimi yapıldığında 0.9 ile en yüksek doğruluk başarımlarını vermiştir. Şekil 3'te görüldüğü gibi metasezgisel yöntemlerin 5 makine öğrenmesi algoritması üzerinde genel olarak geleneksel yöntemlere göre daha başarılı sonuçlar verdiği görülmüştür. LightGBM algoritmasının diğer algoritmalara göre daha başarılı olduğu gözlemlenmiştir.

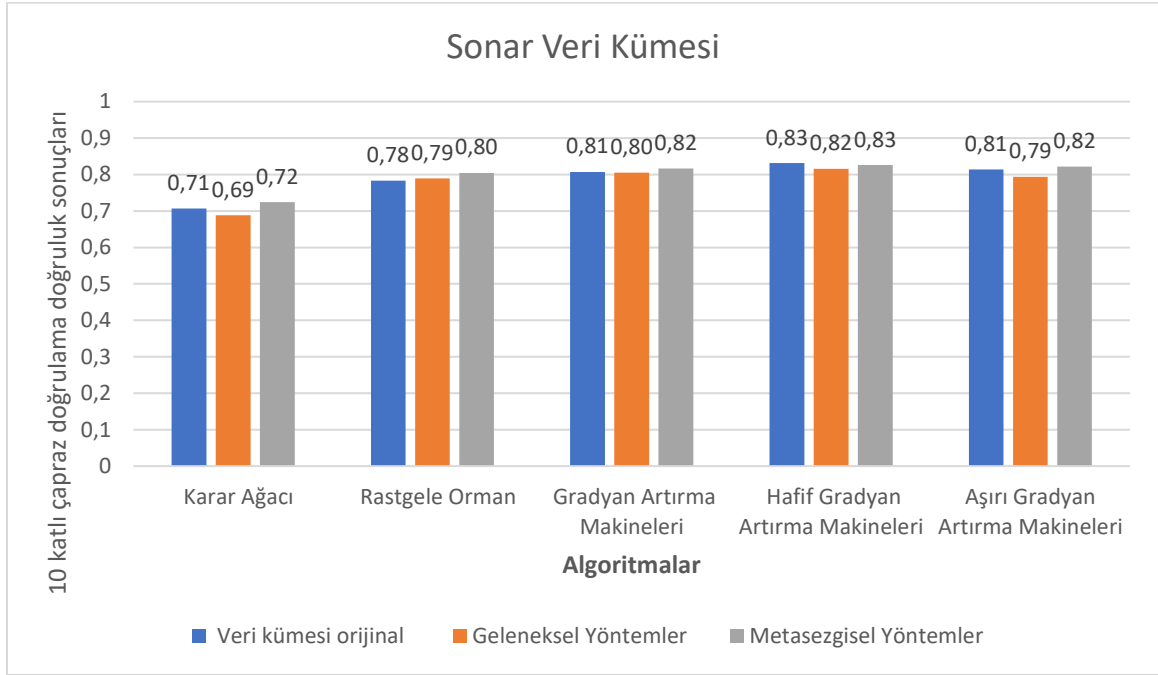


Şekil 3. Parkinson veri kümesi ortalama doğruluk başarımları

Şekil 4'te Sonar veri kümesi için kullanılan yöntemlerin ortalama doğruluk başarımları ve veri kümesinin orijinal halinin doğruluk başarımları gösterilmiştir. Sonar veri kümesi için genel olarak sonuçlar incelendiğinde, Karar Ağacı algoritması geleneksel makine öğrenmesi yöntemleri kullanılarak öznelik seçimi yapıldığında 0.69 başarımları ile en düşük doğruluk başarımlarını verirken,

LightGBM algoritması metasezgisel algoritmalar ile öznelik seçimi yapıldığında 0.83 ile en yüksek doğruluk başarımlarını vermiştir. Şekil 4'teki sonuçlar incelendiğinde LightGBM algoritması dışında metasezgisel algoritmaların daha başarılı sonuçlar verdiği görülmüştür. Algoritmalar bazında incelendiğinde LightGBM

algoritmasının daha başarılı sonuçlar verdiği gözlemlenmiştir.

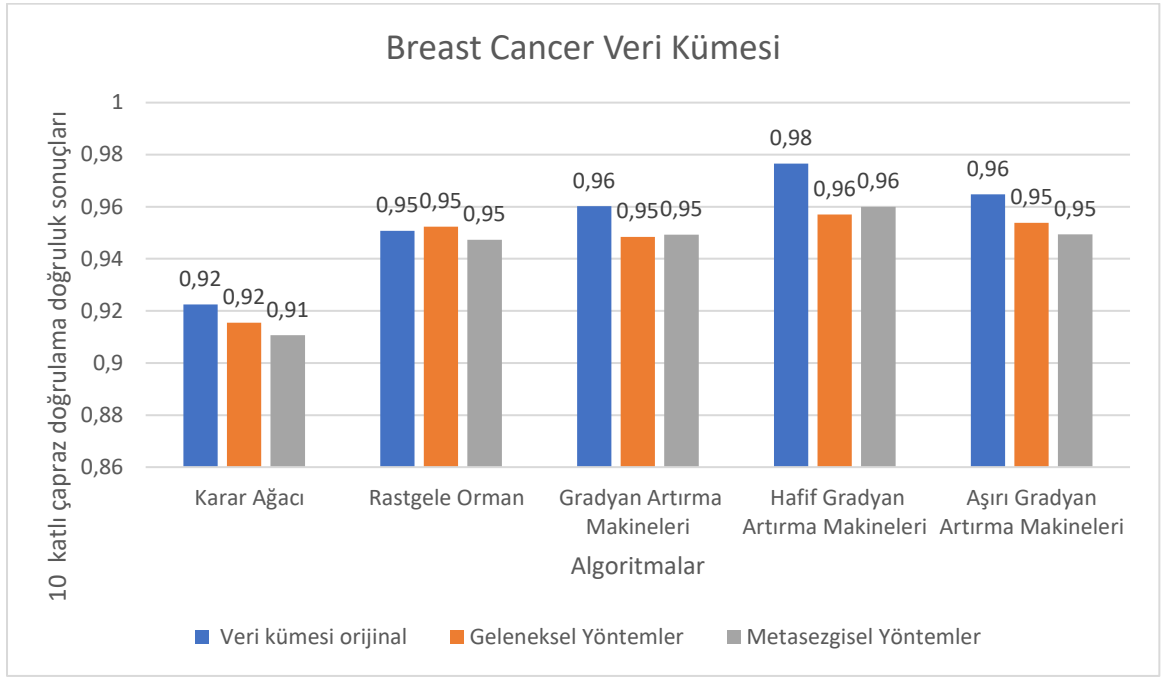


Şekil 4. Sonar veri kümesi ortalama doğruluk başarımları

Şekil 5'de Breast Cancer veri kümesi için kullanılan yöntemlerin ortalama doğruluk başarımları ve veri kümesinin orijinal halinin başarımları gösterilmiştir. Breast Cancer veri kümesi için genel olarak sonuçlar incelendiğinde, Karar Ağacı algoritması metasezgisel algoritmalar ile öznelik seçimi yapıldığında 0,9108 doğruluk başarımları ile en düşük doğruluk başarımlarını verirken, LightGBM algoritması, veri kümesinin orijinal hali kullanılarak öznelik seçimi yapıldığında 0,9766 doğruluk başarımları ile en yüksek doğruluk başarımlarını vermiştir. Ortalama doğruluk başarımları incelendiğinde metasezgisel yöntemlerin tüm

algoritmalar için geleneksel yöntemlere ve veri kümesinin orijinal haline göre daha başarısız sonuçlar verdiği görülmüştür. Bu veri kümesi üzerinde metasezgisel yöntemlerin başarısız sonuçlar vermesi; veri kümesinin dengesiz bir şekilde dağılması, veri kümesinde aykırı değerlerin olması ve parametrelerin uygun bir şekilde belirlenememesinden dolayı başarısız sonuçlar verdiği düşünülmektedir.

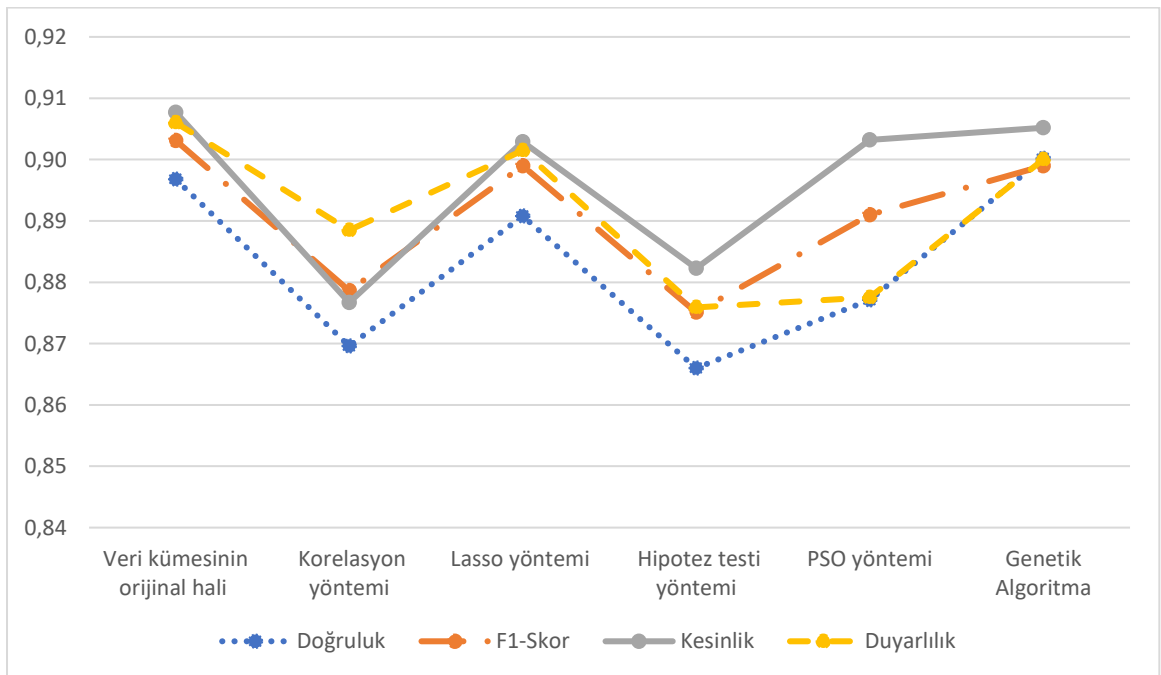
Genel olarak öznelik seçimi yapılmadan kurulan modellerin daha başarılı sonuçlar verdiği gözlemlenmiştir.



Şekil 5. Breat cancer veri kümesi ortalama doğruluk başarımları

Elde edilen deneysel çalışmaların doğruluğunu ölçmek için Hafif Gradyan Artırma Makineleri Algoritması kullanılarak 3 veri kümesinin Doğruluk, F1-Skor, Kesinlik ve Duyarlılık başarımlarının ortalamaları alınarak Şekil 6'da gösterilmiştir. Genel olarak sonuçlar incelendiğinde Genetik Algoritmanın öznelik seçiminde diğer yöntemlere

göre daha başarılı sonuçlar verirken, korelasyon analizi en düşük başarımları verdiği görülmüştür. Kullanılan değerlendirme metriklerine göre Doğruluk değeri genel olarak daha düşük sonuçlar verirken, duyarlılık değerinin daha başarılı sonuçlar verdiği görülmüştür



Şekil 6. Hafif Gradyan Artırma Makineleri ile 3 veri kümesi ortalama

## Doğruluk , F1-Skor, Kesinlik , Duyarlılık sonuçları

**Sonuçlar**

Bu çalışmada UCI veri kümesi deposundan literatürde en sık kullanılan 3 veri kümesi üzerinde 3 farklı geleneksel makine öğrenmesi yöntemi ile öznelik seçimi ve doğadan ilham alarak ortaya çıkmış 2 farklı meta sezgisel algoritma kullanıldı. Yapılan deneysel çalışmalar sonucunda öznelik seçim yöntemleri uygulandıktan sonra genel olarak başarı oranlarının arttığı gözlemlenmiştir.

Veri kümesinde öznelik sayısı çok olduğunda öznelik seçim yöntemlerinin kullanılması önem arz etmektedir. Bunun nedeni veri kümesinde öznelik sayısının çok olması durumunda bazı öznelikler model performansını azaltmakta ve başarı oranlarının düşmesine neden olabilmektedir.

Veri kümeleri üzerinde öznelik seçim yöntemleri kullanılarak gürültülü ya da model için ilgisiz öznelikler elenince başarı artabilir.

Sonuç olarak kullanılan öznelik seçimi yöntemlerinden parçacık sürü optimizasyon ve genetik algoritmaların öznelik seçiminde daha az sayıda öznelik ile oldukça başarılı sonuçlar elde edildiği söylenebilir.

**Kaynaklar**

- [1] Koşan, Muhammed Ali, Coşkun, Aysun, Karacan, Hacer. 2019. "Yapay Zekâ Yöntemlerinde Entropi", *Bilişim Sistemleri ve Yönetim Araştırmaları Dergisi*, Cilt 1, S.1, sf. 15-21
- [2] Kaynar, Oğuz, Arslan, Halil, Görmez, Yasin, Işık, Yunus Emre. 2018. "Makine Öğrenmesi ve Öznelik Seçim Yöntemleriyle Saldırı Tespiti", *Bilişim Teknolojileri Dergisi*, Cilt 11, S. 2, sf.175-185
- [3] Kaynar, Oğuz, Arslan, Halil, Görmez, Yasin, Işık, Yunus Emre. 2018. "Makine Öğrenmesi ve Öznelik Seçim Yöntemleriyle Saldırı Tespiti", *Bilişim Teknolojileri Dergisi*, Cilt 11, S. 2, sf.175-185
- [4] Onay, Funda Kutlu. 2023. "Öznelik Seçimi Problemleri İçin İkili Beyaz Köpekbalığı Optimizasyon Algoritması", *Gümüşhane Üniversitesi Fen Bilimleri Dergisi*, Cilt 13, S.2, sf. 281-298
- [5] Vatanserver, Berna, Aydın, Hakan, Çetinkaya, Ali. 2021. "Heart Disease Prediction with Machine Learning Algorithm Using Feature Selection by Genetic Algorithm", *Bilim, Teknoloji ve Mühendislik Araştırmaları Dergisi*, Cilt 2, S.2, sf. 67-80
- [6] Şengür, Dönüş. 2021. "Öğretmenlerde Etkili Psikososyal Risk Düzeylerinin Belirlenmesi için Parçacık Sürü

Elde edilen sonuçlar incelendiğinde doğadaki varlıklardan esinlenerek ortaya çıkarılmış metasezgisel algoritmaların öznelik seçiminde başarılı olduğu gözlenmiştir. Bu yöntemler öznelik seçimi dışında endüstriyel tasarım projelerinde de oldukça başarılı sonuçlar vermektedir. Bir uçağın kanat tasarımı, masa için uygun genişlikte yükseklikte malzemelerin seçimi, baz istasyonlarının yerleşim yerlerinin belirlenmesi, elektronik devre tasarımı gibi endüstriyel ve mühendislik çalışmalarında yaygın olarak kullanılmaktadır. Öte yandan bir yapay sinir ağının eğitimi ve makine öğrenmesi algoritmalarında hiperparametrelerin optimize edilmesi gibi alanlarda kullanılmaktadır.

Farklı metasezgisel algoritmaların iyi yönlerinin birleştirilerek geliştirilen hibrit metasezgisel algoritmalar başarılı sonuçlar vermektedir. Metasezgisel algoritmalar ile ilgili çok sayıda yöntem geliştirilmiş olmasına rağmen, çözümü aranan problemlerle ilgili çalışmalar gün geçtikçe farklılaştığından dolayı, bu konu popüleritesini ve güncelliğini korumaktadır.

Metasezgisel algoritmaların kısıtları gözönüne alındığında ise, her zaman en iyi çözümü garanti etmedikleri, Yüksek bellek kullanımı gereksimi ve parametrelerinin ayarlanmasının uzun zaman gerektirdiği gözlemlenmiştir.

Optimizasyonu Tabanlı Öznelik Seçme", *Türkiye Eğitim Dergisi*, Cilt 6, S. 2, sf. 488-496

[7] Çelik, Ceyhan, Bilge Hasan Şakir. 2015. "Ağırlıklandırılmış Koşullu Karşılıklı Bilgi İle Öznelik Seçimi", *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*, Cilt 30, S. 4, sf.585-596

[8] Sağbaş, Ensar Arif, Gökalp Osman, Uğur Aybars. 2019. "Yüz İfadesi Tanıma için Mesafe Oranlarına Dayalı Öznelik Çıkarımı ve Genetik Algoritmalar İle Seçimi", *Veri Bilim Dergisi*, Cilt 2, S.1, sf.19-29

[9] Abbasoğlu, Ferda, Gülcü, Ayla, Ulvi, Başpınar. 2021. "EEG Sinyallerinin Sınıflandırılmasında Evrimsel Öznelik Seçim Metotlarının Kullanılması", *International Journal of Advances in Engineering and Pure Sciences*, C.33, S.2, sf.171-179

[10] Bilen, Abdulkadir, Özer, Ahmet Bedri. 2021. "Siber Saldırıları için Rastgele Orman Algoritması Kullanılarak Öznelik Seçimi", *Fırat Üniversitesi Fen Bilimler Dergisi*, C.34, S.1, sf.31-37

[11] Sağbaş, Ensar Arif. 2023. "Filtre Tabanlı Öznelik Seçim Yöntemleri Kullanılarak Metinlerde Duygu Sınıflandırılması Üzerine Karşılaştırmalı Bir Çalışma", *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, C.35, S.1, sf.239-250

- [12] Yağmur, Ece, Yağmur, Sercan. 2022. “Rüzgâr Gücü Tahmininde Genetik Algoritma ile Öznitelik Seçimi”, *Afyon Kocatepe Üniversitesi Fen ve Mühendislik Bilimleri Dergisi*, C.22, S.5, sf.1028-1040
- [13] Narin, Ali. 2020. “Parkinson Hastalarının Tespitinde Karınca Koloni Algoritması ile Seçilen Özniteliklerin Performansa Etkisi”, *Düzce Üniversitesi Bilim ve Teknoloji Dergisi*, C.8, S.4, sf. 2443-2454
- [14] Can, Celal, Kaya, Yasin, Kılıç, Fatih. 2021. “Salp Sürü Algoritması ile Öznitelik Seçimi ve Sınıflandırıcı Performans Değerlendirmesi”, *Avrupa Bilim ve Teknoloji Dergisi*, C.30, sf.12-16
- [15] Xue Wen, Chen and Michael, Wasikowski. “Fast: a roc-based feature selection metric for small samples and imbalanced data classification problems”, *In Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining, Las Vegas, USA, 2008*, pp.124-132
- [16] Cengiz, Mehmet Ali, Terzi, Yüksel. 2018. *Hipotez Testleri Ders Notları*, On Dokuz Mayıs Üniversitesi Fen-Edebiyat Fakültesi İstatistik Bölümü.
- [17] Orhan, Hikmet, Vergili, Merve. 2022. “Genomik Veri Setlerinin LASSO ve Elastik Net Regresyon Yöntemleri ile Analizi”, *Süleyman Demirel Üniversitesi Sağlık Bilimleri Dergisi*, Cilt 13, Sayı 3, sf.485 – 496
- [18] Karaboğa, Derviş. 2020. “Yapay Zeka Optimizasyon Algoritmaları”, Nobel Akademik Yayıncılık, sf.75
- [19] Karaboğa, Derviş. 2020. “Yapay Zeka Optimizasyon Algoritmaları”, Nobel Akademik Yayıncılık, sf.76-82
- [20] Karaboğa, Derviş. 2020. “Yapay Zeka Optimizasyon Algoritmaları”, Nobel Akademik Yayıncılık, sf.182
- [21] Karaboğa, Derviş. 2020. “Yapay Zeka Optimizasyon Algoritmaları”, Nobel Akademik Yayıncılık, sf.183-189
- [22] Ozkan, Yalçın. 2020. “Veri Madenciliği Yöntemleri”, Papatya Yayıncılık, sf. 39-40
- [23] Breiman Leo, “Consistency For A Simple Model of Random Forests”, Berkeley, USA. Sep.9,2004.  
<https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=2a42f39add8332a7139d44a6e77496c0571e4f24>
- [24] Friedman, Jerome Harold. 2001. “Greedy Function Approximation: A Gradient Boosting Machine”, *The Annals of Statistics*, C. 29, S. 5, sf.1189-1232
- [25] Alsharı, Haithm, Saleh, Abdulrazak Yahya, Odaş, Alper. 2021. “Comparison of Gradient Boosting Decision Tree Algorithms for CPU Performance”, *Erciyes Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, Cilt 37, S. 1, sf. 157-168
- [26] Chen, Tianqi and Guestrin, Carlos. “XGBoost: A Scalable Tree Boosting System”, *Uluslararası Bilgi Keşfi ve Veri Madenciliği Konferansı Bildirileri, Washington, USA,2016*, sf. 785-794