



T.C.
KONYA TEKNİK ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ



VERİ BİLİMİ VE MÜHENDİSLİK
OPTİMİZASYON PROBLEMLERİNİN
ÇÖZÜMÜ İÇİN YENİ BİR YAKLAŞIM:
KAOTİK YAPAY ALG ALGORİTMASI

Bahaeddin TÜRKOĞLU

DOKTORA TEZİ

Bilgisayar Mühendisliği Anabilim Dalı

Haziran-2022
KONYA
Her Hakkı Saklıdır

TEZ KABUL VE ONAYI

Bahaeddin TÜRKOĞLU tarafından hazırlanan “VERİ BİLİMİ VE MÜHENDİSLİK OPTİMİZASYON PROBLEMLERİNİN ÇÖZÜMÜ İÇİN YENİ BİR YAKLAŞIM: KAOTİK YAPAY ALG ALGORİTMASI” adlı tez çalışması 03/06/2022 tarihinde aşağıdaki jüri tarafından oy birliği ile Konya Teknik Üniversitesi Lisansüstü Eğitim Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı’nda DOKTORA TEZİ olarak kabul edilmiştir.

Jüri Üyeleri

İmza

Başkan

Prof.Dr. Harun UĞUZ

.....

Danışman

Dr.Öğr.Üyesi Ersin KAYA

.....

Üye

Prof.Dr. Gülay TEZEL

.....

Üye

Doç.Dr. Mehmet Akif ŞAHMAN

.....

Üye

Dr.Öğr.Üyesi Ahmet ÖZKİŞ

.....

Yukarıdaki sonucu onaylarım.

Prof. Dr. Saadettin Erhan KESEN
Enstitü Müdürü

Dr.Öğr.Üyesi Sait Ali UYMAZ bu tez çalışmasının ikinci danışmanıdır.

TEZ BİLDİRİMİ

Bu tezdeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edildiğini ve tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her türlü ifade ve bilginin kaynağına eksiksiz atıf yapıldığını bildiririm.

DECLARATION PAGE

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Bahaeddin TÜRKOĞLU

Tarih:

ÖZET

DOKTORA TEZİ

VERİ BİLİMİ VE MÜHENDİSLİK OPTİMİZASYON PROBLEMLERİNİN ÇÖZÜMÜ İÇİN YENİ BİR YAKLAŞIM: KAOTİK YAPAY ALG ALGORİTMASI

Bahaeddin TÜRKOĞLU

**Konya Teknik Üniversitesi
Lisansüstü Eğitim Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı**

**Danışman: Dr.Öğr.Üyesi Ersin KAYA
İkinci Danışman: Dr.Öğr.Üyesi Sait Ali UYMAZ**

2022, 131 Sayfa

**Jüri
Dr.Öğr.Üyesi Ersin KAYA
Prof.Dr. Harun UĞUZ
Prof.Dr. Gülay TEZEL
Doç.Dr. Mehmet Akif ŞAHMAN
Dr.Öğr.Üyesi Ahmet ÖZKİŞ**

Optimizasyon, bir problemin çözüm uzayındaki en uygun çözümü bulma, verilen kısıtlar altında eniyileme işlemidir. Günümüz dünyasında minimum maliyet ile maksimum verimliliğinin hedeflendiği birçok alanda yaygın olarak kullanılmaktadır. Son yıllarda, gerçek dünya problemlerinin giderek artan karmaşıklığı ve zorluğu, daha güvenilir optimizasyon tekniklerine, özellikle metasezgisel optimizasyon algoritmalarına daha fazla ihtiyaç duyulmasına neden olmuştur. Yapay Alg Algoritması (AAA), mikro alglerin karakteristiklerinden ve yaşam davranışlarından esinlenerek geliştirilmiş metasezgisel bir optimizasyon algoritmasıdır. Çeşitli alanlardaki birçok gerçek dünya problemini başarıyla çözerek popüler metasezgisel algoritmalarından birisi haline gelmiştir. Bununla birlikte diğer metasezgisel optimizasyon algoritmalarına benzer şekilde AAA da erken yakınsama ve yerel minimumlara sıkışma eğilimi göstermektedir. Bu problemleri aşmak için algoritmanın yapısının güçlendirilmesi gerekmektedir.

Metasezgisel optimizasyon algoritmalarının karakteristiklerini belirleyen iki önemli arama stratejisi vardır. Bunlardan birisi keşif/çeşitlendirme diğeri sömürü/yoğunlaştırmadır. Keşif süreci, arama uzayını küresel olarak keşfetme yeteneğidir. Bu yetenek, yerel optimumdan kaçınma ve yerel optimuma takılınca kurtulabilme kabiliyetidir. Sömürü süreci ise çözümün uygunluğunu yerel olarak iyileştirmek için, mevcut çözümün yakınındaki muhtemel çözümleri keşfetme yeteneğidir. Bir metasezgisel algoritmanın performansının mükemmel olması bu iki strateji arasındaki dengeye bağlıdır.

Literatürde keşif ve sömürü süreçlerini güçlendirmek ve aralarındaki dengeyi oluşturmak için levy uçuşu, kuantum davranışı, yerel arama, çoklu ve zeki arama, kaos teorisi gibi çok çeşitli stratejiler geliştirilmiştir. Bu stratejilerden birisi kaos teorisinden ilham alınarak geliştirilen kaotik haritalardır. Kaotik haritalar keşif ve sömürü arasındaki dengeyi güçlendiren çok önemli performans artırma stratejisidir nitekim bu haritalar kullanılarak literatürdeki birçok metasezgisel optimizasyon algoritmasının performansı artırılmıştır.

Bu tez çalışmasında AAA kaotik haritalar ile donatılarak Kaotik Yapay Alg Algoritması isminde yeni bir yaklaşım geliştirilmiş ve dört farklı problem uzayına çözüm getirmiştir. Geliştirilen bu yaklaşımla ilk olarak farklı zorluktaki otuz benchmark test fonksiyonu çözülmüştür. Daha sonra basınçlı tank tasarımı, kaynaklı kiriş tasarımı, germe sıkıştırma yayı tasarımı ve Avrupa Uzay Ajansı'ndan alınan sekiz yörünge

tasarım problemi üzerinde test edilerek performansı doğrulanmıştır. Üçüncü olarak Kaotik AAA yaklaşımı, makine öğrenmesinin üç temel alanından birisi olan gözetimsiz öğrenmedeki paylaştırmalı kümeleme problemine uygulanarak performansı analiz edilmiştir. Dördüncü olarak da önerilen kaotik AAA yaklaşımının, makine öğrenmesi algoritmaları için kaçınılmaz kritik bir önışleme süreci olan öznitelik seçiminde kullanılmak üzere ikili versiyonu geliştirilmiştir. Bu tezde geliştirilen kaotik tabanlı yeni yaklaşım, tüm problem uzaylarında literatürdeki farklı zorluk seviyesine sahip problem setleri üzerinde çeşitli popüler algoritmalarla kıyaslanmış, Wilcoxon işaretli sıralar testi ve Friedman istatistiksel testleri yapılarak güvenilirliği sağlanmış ve kıyaslanan algoritmalardan daha performanslı olduğu doğrulanmıştır.

Anahtar Kelimeler: Kaos Teorisi, Kaotik Haritalar, Kümeleme Analizi, Makine Öğrenmesi, Metasezgisel Algoritmalar, Öznitelik Seçimi, Sürü Zekâsı, Yapay Alg Algoritması



ABSTRACT

PhD THESIS

A NOVEL APPROACH TO SOLUTION OF DATA SCIENCE AND ENGINEERING OPTIMIZATION PROBLEMS: CHAOTIC ARTIFICIAL ALGAE ALGORITHM

Bahaeddin TURKOGLU

**Konya Technical University
Institute of Graduate Studies
Department of Computer Engineering**

**Advisor: Asst.Prof.Dr. Ersin KAYA
Co-Advisor: Asst.Prof.Dr. Sait Ali UYMAZ**

2022, 131 Pages

**Jury
Asst.Prof.Dr. Ersin KAYA
Prof.Dr. Harun UGUZ
Prof.Dr. Gulay TEZEL
Assoc.Prof.Dr. Mehmet Akif SAHMAN
Asst.Prof.Dr. Ahmet OZKIS**

Optimization is the process of finding the optimal solution in the solution space of a problem and determining the best one under the given constraints. In today's world, it is widely used in many areas where maximum efficiency is aimed with minimum cost. In recent years, the increasing complexity and difficulty of real-world problems has led to a greater need for more reliable optimization techniques, especially metaheuristic optimization algorithms. Artificial Algae Algorithm (AAA) is a metaheuristic optimization algorithm inspired by the characteristics and life behavior of microalgae. It has become one of the popular metaheuristic algorithms by successfully solving many real-world problems in various fields. However, similar to other metaheuristic optimization algorithms, AAA tends to early converge and get to stuck in local minima. In order to overcome these problems, the structure of the algorithm needs to be strengthened.

There are two important search strategies that determine the characteristics of metaheuristic optimization algorithms. One is exploration/diversification and the other is exploitation/intensification. The discovery process is the ability to explore the search space globally. This ability is the ability to avoid the local optimum and get rid of the local optimum when stuck. The exploitation process is the ability to discover possible solutions near the current solution in order to locally improve the relevance of the solution. The perfect performance of a metaheuristic optimization algorithm depends on the balance between these two strategies.

Various strategies such as levy flight, quantum behavior, local search, multiple and intelligent search, chaos theory have been developed in the literature to strengthen the exploration and exploitation processes and to create the balance between them. One of these strategies is chaotic maps, which are inspired by chaos theory. Chaotic maps are a very important performance enhancement strategy that strengthens the balance between exploration and exploitation, as the performance of many metaheuristic optimization algorithms in the literature has been improved by using these maps.

In this thesis, a new approach named Chaotic Artificial Algae Algorithm has been developed by equipping AAA with chaotic maps and it has brought a solution to four different problem spaces. With this

developed approach, first of all, thirty benchmark test functions of different difficulty were solved. The performance was then validated by testing on the pressure tank design, welded beam design, tension compression spring design, and eight space design problems from the European Space Agency. Thirdly, the chaotic AAA approach is applied to the distributed clustering problem in unsupervised learning, which is one of the three main areas of machine learning, and its performance is analyzed. Fourth, a binary version of the proposed chaotic AAA approach has been developed to be used in feature selection, which is an inevitable critical preprocessing process for machine learning algorithms. The chaotic-based new approach developed in this thesis has been compared with various popular algorithms on problem sets with different difficulty levels in the literature in all problem spaces, its reliability has been ensured by performing Wilcoxon signed-rank test and Friedman statistical tests and it has been verified that it is more performant than the compared algorithms.

Keywords: Artificial Algae Algorithm, Chaos Theory, Chaotic Maps, Clustering Analysis, Feature Selection, Machine Learning, Metaheuristic Algorithms, Swarm Intelligence



ÖNSÖZ

Bu tez çalışması boyunca hiçbir desteğini esirgemeyen değerli danışmanlarım Dr.Öğr.Üyesi Ersin KAYA ile Dr.Öğr.Üyesi Sait Ali UYMAZ'a ;

Tez izleme komitemde yer alan değerli hocalarım Prof.Dr. Harun UĞUZ ve Doç.Dr. Mehmet Akif ŞAHMAN'a ;

Çalışmalarım boyunca anlayış ve destekleri ile her konuda yardımcı olan aileme; İçtenlikle teşekkür eder şükran ve minnetlerimi sunarım.

Bu tez çalışması babam Ali İhsan TÜRKOĞLU'na ithaf edilmiştir.

Bahaeddin TÜRKOĞLU
KONYA-2022

İÇİNDEKİLER

ÖZET	iv
ABSTRACT.....	vi
ÖNSÖZ	viii
İÇİNDEKİLER.....	ix
SİMGELER VE KISALTMALAR.....	xi
1. GİRİŞ	1
1.1. Tezin Literatüre Katkısı	5
1.2. Tezin Organizasyonu	7
2. KAYNAK ARAŞTIRMASI	8
2.1. Kaos Teorisi ve Kaotik Harita Tabanlı Metasezgisel Yaklaşımlar.....	8
2.2. Kümeleme Analizinde Metasezgisel Algoritmaların Kullanımı	12
2.3. Öznitelik Seçiminde Metasezgisel Algoritmaların Kullanımı.....	17
3. MATERYAL VE YÖNTEM	22
3.1. Yapay Alg Algoritması.....	22
3.1.1. Evrimsel Süreç	24
3.1.2. Helisel Hareket	24
3.1.3. Adaptasyon Süreci	26
3.2. Kaos Teorisi.....	28
3.2.1. Kaotik Haritalar	29
3.3. Kaotik Yapay Alg Algoritması.....	37
3.4. Kümeleme Problemleri için Kaotik Yapay Alg Algoritması.....	39
3.5. Öznitelik Seçim Problemi için Kaotik Yapay Alg Algoritması	42
3.5.1. Çözüm Vektörünün Tasarımı.....	48
3.5.2. Uygunluk Fonksiyonu ve K-En Yakın Komşu Sınıflandırıcısı.....	48
3.6. Algoritmaların Test Edilmesi için Kullanılan Optimizasyon Problemleri	51
3.6.1. Klasik Benchmark Fonksiyon Seti ve Değerlendirme Metrikleri	51
3.6.2. Gerçek Dünya Optimizasyon Problemleri.....	55
4. ARAŞTIRMA SONUÇLARI VE TARTIŞMA.....	64
4.1. Global Optimizasyon Problemlerinde Kaotik Yapay Alg Algoritmasının Performansı.....	64
4.1.1. Kaotik Versiyonlarının Performans Analizi	65
4.1.2. Geliştirilen Yöntemin Literatürdeki Algoritmalarla Performans Kıyaslaması	67
4.2. Gerçek Dünya Optimizasyon Problemlerinde Kaotik Yapay Alg Algoritmasının Performans Analizi	77

4.3. Kümeleme Analizinde Kaotik Yapay Alg Algoritmasının Performans Değerlendirmesi.....	81
4.3.1. Kümeleme Analizindeki Değerlendirme Metrikleri.....	85
4.3.2. Performans Değerlendirmesi	87
4.4. Öznitelik Seçme Probleminde İkili Kaotik Yapay Alg Algoritmasının Performansı.....	99
4.4.1. Değerlendirme Kriterleri.....	101
4.4.2. Performans Değerlendirmesi	102
4.4.3. İkili Versiyonlarının Performans Analizi	102
4.4.4. Geliştirilen Varyasyonun Literatürdeki Algoritmalar ile Performans Kıyaslaması.....	105
5. SONUÇLAR VE ÖNERİLER.....	119
5.1 Sonuçlar	119
5.2 Öneriler	122
KAYNAKLAR	123

SİMGELER VE KISALTMALAR

Simgeler

Ap	: AAA'da adaptasyon parametresi
D	: Problem boyutu
e	: AAA'da enerji kaybı parametresi
f	: Amaç fonksiyonu
G	: AAA'da alg kolonilerinin büyüklükleri
N	: Popülasyondaki alg koloni sayısı
Δ	: AAA'da kesme kuvveti katsayısı

Kısaltmalar

AAA	:Yapay Alg Algoritması (Artificial Algae Algorithm)
ABC	:Yapay Arı Kolonisi Algoritması (Artificial Bee Colony Algorithm)
ACO	:Karıncı Koloni Algoritması (Ant colony optimization algorithm)
ALO	:Karıncı Aslanı Optimizasyonu (Ant Lion Optimization)
BAT	:Yarasa Optimizasyon Algoritması (Bat Optimization Algorithm)
BBO	:Biocoğrafya Tabanlı Optimizasyon (Biogeography based Optimization)
BFO	:Bakteri Yiyecek Arama Optimizasyonu (Bacterial Foraging Optimizer)
BHO	:Kara Delik Optimizasyonu (Black Hole Optimization)
BOA	:Kelebek Optimizasyon Algoritması (Butterfly Optimization Algorithm)
CS	:Guguk Kuşu Arama Algoritması (Cuckoo Search Algorithm)
CRO	:Mercan Resifleri Optimizasyonu (Coral Reefs Optimizer)
CSA	:Karga Arama Algoritması (Crow Search Algorithm)
DE	:Diferansiyel Evrim Algoritması (Differential Evolution Algorithm)
ESA	:Avrupa Uzay Ajansı
FFA	:Ateş Böceği Algoritması (Firefly Algorithm)
FOA	:Meyve Sineği Optimizasyon Algoritması (Fruitfly Optimization Algorithm)
FPA	:Çiçek Tozlaşma Algoritması (Flower Pollination Algorithm)
GA	:Genetik Algoritma (Genetic Algorithm)
GOA	:Çekirge Optimizasyon Algoritması (Grasshopper Optimization Algorithm)
GSA	:Yerçekimi Arama Algoritması (Gravitational Search Algorithm)
GTOP	:Küresel Yörünge Optimizasyon Problemleri Veritabanı (Global Trajectory Optimization Problems Database)
GWO	:Gri Kurt Optimizasyonu (Grey Wolf Optimizer)
HC	:Tepe Tırmanma Algoritması (Hill Climbing Algorithm)
HGSO	:Henry Gaz Çözünürlük Optimizasyonu (Henry Gas Solubility Optimization)
HHO	:Harris Şahini Optimizasyonu (Harris Hawk Optimizer)
HOA	:At Sürüsü Optimizasyon Algoritması (Horse Herd Optimization Algorithm)
HS	:Harmoni Arama Algoritması (Harmony Search Algorithm)
KH	:Krill Sürüsü Optimizasyonu (Krill Herd Optimizasyonu)
MFO	:Güve Optimizasyon Algoritması (Moth Flame Optimization Algorithm)
MSE	:Ortalama Karese Hata (Mean Squared Error)
MVO	:Çoklu Evren Optimizasyon Algoritması (Multi Verse Optimization Algorithm)
PIO	:Güvercin İlhamlı Optimizasyon (Pigeon Inspired Optimizer)
PSO	:Parçacık Sürü Optimizasyonu (Particle Swarm Optimization)

SA	:Benzetimli Tavlama Algoritması (Simulated Annealing Algorithm)
SCA	:Sinüs Kosinüs Algoritması (Sine Cosine Algorithm)
SMO	:Örümcek Maymun Optimizasyonu (Spider Monkey Optimization)
SOS	:Simbiyotik Organizmalar Araması (Symbiotic Organisms Search)
SSA	:Salp Sürüsü Algoritması (Salp Swarm Algorithm)
Std.	:Standart Sapma
TS	:Tabu Arama Algoritması (Tabu Search)
WOA	:Balina Optimizasyon Algoritması (Whale Optimization Algorithm)



1. GİRİŞ

Optimizasyon belirli bir problem için olası çözümler arasından en uygun olanı bulma, verilen kısıtlar altında eniyileme işlemidir. Günümüz dünyasında minimum maliyet ile maksimum verimliliğinin hedeflendiği birçok alanda yaygın olarak kullanılmaktadır (Lange, 2013). Optimizasyon problemlerine gösterilebilecek birçok örnekten bazıları şunlardır: Sabit bir büyüklükteki koninin içine yerleştirilecek en geniş hacimli silindirin tasarımı için yarıçap ve yükseklik parametrelerinin tespit edilmesi, çevresi sabit bir dikdörtgenin alanı en büyük olacak şekilde tasarımı için kenar uzunluklarının belirlenmesi, basınçlı tank tasarımı olarak bilinen belirli bir hacim ve basıncı oluşturacak şekilde minimum maliyetli hava tankının tasarımı için tank yarıçapı, yüksekliği ve kullanılan sacın kalınlığının tespit edilmesi gibi örnekler verilebilir (Houssein ve ark., 2020). Yapay sinir ağlarında verilen girişlere ait doğru çıkışları sağlayacak şekilde ağın uygun parametrelerinin saptanması da bir optimizasyon problemidir (Turkoglu ve Kaya, 2020). Araç ve uçuş rotası planlama (Kechagiopoulos ve Beligiannis, 2014), gezgin satıcı problemi (Gündüz ve ark., 2015), ekonomik talep analizi (Kumar ve Dhillon, 2018), rüzgâr türbini ve tesis yerleşim problemi (Beşkirli ve ark., 2018), arazi toplulaştırma (Ozsari ve ark., 2021), enerji tahmin analizi (Kıran ve ark., 2012), mühendislik tasarım problemleri gibi gerçek dünyada karşılaştığımız pek çok problem, temel olarak bir optimizasyon problemine işaret etmektedir.

Bir optimizasyon probleminin karmaşıklığını etkileyen çeşitli faktörler bulunmaktadır. Problemin boyutu, olası çözüm sayısı ve uzayı, probleme özgü kısıtlar bunlardan bazılarıdır. Optimizasyon problemlerdeki tespit edilmesi gereken değişken sayısı arttıkça çözüm uzayı üstel olarak artmaktadır. Dolayısıyla ideal çözümün hesaplanması için harcanan zaman ve bellek maliyetini artırmaktadır.

Optimizasyon problemlerinin çözümünde kullanılan yöntemler deterministik ve stokastik olmak üzere iki ana kategoriye ayrılmaktadır. Deterministik yöntemler ideal çözüme ulaşmak için genellikle gradyan yaklaşımı esas alan yöntemlerdir. Bu yöntemler aynı başlangıç parametreleri altında işletildiğinde aynı ideal çözüme ulaşırlar. Bununla birlikte deterministik yöntemlerin başlangıç parametrelerine yüksek bağımlılık, erken yakınsama ve yerel minimumlarda sıkışma eğilimleri gibi temel dezavantajları vardır. Diğer taraftan stokastik yöntemler ise gradyan içermeyen (gradient-free) ve rastgelelik barındıran teknikler kullanarak ideal çözüme ulaşmaya çalışmaktadır. Stokastik

yöntemler, sezgisel ve metasezgisel olarak iki sınıfta incelenmektedir (Arora ve Anand, 2019a). Sezgisel yöntemler belirlenen kurala bağlı olarak deneme yanılma yoluyla mevcut çözümün geliştirilmesini amaçlamaktadır (Yildiz, 2009). Metasezgisel yöntemler, sezgisel yöntemlerden esinlenerek geliştirilmiş alan bilgisine (domain knowledge) ihtiyaç duymayan, içerisinde birkaç farklı sezgisel yaklaşım mekanizması ve rastgelelik barındırabilen bir üst çatıdır. Metasezgisel yöntemler optimum çözümü her zaman garanti etmemelerine rağmen makul süre zarfında uygulanabilirlik ve hızlı çözüm sunma açısından öne çıkmaktadırlar (Stork ve ark., 2020). Metasezgisel yöntemler birden çok ajanı, kendine has kuralları ve matematiksel denklemleri ile çoklu yineleme (iterasyon) şeklinde işleterek, belirlenen durma kriterine ulaşana kadar iyileştirmeyi amaçlamaktadır. Nihai çözüm olarak arama sürecinin sonunda elde edilen en iyi çözümü (ajanı) sunmaktadır (Hussain ve ark., 2019). Metasezgisel yöntemler sezgisel yöntemlerden farklı olarak problem bağımlı olmadığından ve alan bilgisine ihtiyaç duymadan daha geniş bir uygulama alanına sahiptirler (Blocho, 2020).

Metasezgisel yöntemler arama sürecine rastgele başlangıç çözümleri ile başlar. İdeal çözümü bulmak için iki temel arama davranışı kullanır. Bunlar keşif (exploration) ve sömürü (exploitation) süreçleridir. Keşif, bir yöntemin çözüm uzayını arama kabiliyetini ifade eder. Sömürü mekanizması ise bir çözümün iyileştirilmesi kapasitesini ifade etmektedir. Bu iki süreç metasezgisel yöntemler için son derece kritiktir ve ideal bir dengeyi oluşturacak şekilde tasarlanması gerekmektedir.

Optimizasyon problemleri için, özellikle son yıllarda, doğadan ilhamlı (nature-inspired) metasezgisel yöntemler geliştirilmiş ve popüler hale gelmiştir. Sürü zekâsı karıncalar, arılar, kurtlar, balinalar, yarasalar, ateş böcekleri gibi sürüdeki bireylerin birbirleriyle bilgi paylaşarak etkileşim içerisinde olduğu sürülerin yaşam davranışlarını modelleyen bir yapay zekâ optimizasyon tekniğidir. Bir sürünün, sürü zekâsı algoritması olarak modellenebilmesi için sürüdeki bireylerin merkezi bir kontrol mekanizması olmadan kendi kendilerine organize olabilmeye yeteneğine sahip olması ve sürüdeki bireylerin tek başına değil sürüyle etkileşerek zeki davranışlar göstermesi gerekir. Örneğin sürü zekâsının en popüler algoritmalarından birisini teşkil eden bal arıları tek başına kaldıkları zaman yaşayamayan ancak sürü şeklinde zeki davranışlar gösteren bir canlı türüdür. Sürü zekâsındaki etkileşime; kuş ve balık sürülerinin hız ve konum bilgilerini birbirleriyle paylaşmaları, karıncaların yuva ve yiyecek kaynağı arasında feromon salgılayarak kolonileri ile bilgi paylaşmaları, arıların buldukları nektar

kaynaklarının konumlarını kovan etrafında arı dansı denilen hareketlerle sürüdeki diğer arılarla paylaşmaları, balinaların okyanusta buldukları besin kaynaklarının konumlarını sürüdeki balinalara spiral şekiller oluşturarak paylaşmaları örnek olarak gösterilebilir.

Doğadan ilhamlı metasezgisel yaklaşımlarının popüler olmasının temel nedeni, karmaşık gerçek dünya problemlerinde etkili performanslar göstermiş olmalarıdır. Başarılı performanslarının arkasında yatan en önemli faktör, çözüm uzayının taranmasında ve mevcut çözümlerin geliştirilmesinde doğal olaylarından ve doğadaki canlılardan esinlenilerek ortaya koyulmuş konum güncelleme mekanizmalarının ve süreçlerinin bulunmasıdır.

Literatürde, Genetik Algoritma (GA) (John, 1992), Diferansiyel Gelişim Algoritması (DE) (Storn ve Price, 1997), Parçacık Sürü Optimizasyonu (PSO) (Kennedy ve Eberhart, 1995), Ateşböceği Algoritması (FA) (Yang, 2009), Yapay Arı Kolonisi (ABC) (Karaboga ve Basturk, 2007), Guguk Kuşu Arama Algoritması (CS) (Yang ve Deb, 2009), Yerçekimi Arama Algoritması (GSA) (Rashedi ve ark., 2009), Çoklu Evren Optimizasyonu (MVO) (Mirjalili ve ark., 2016), Gri Kurt Optimizasyonu (GWO) (Mirjalili ve ark., 2014a), Balina Optimizasyon Algoritması (WOA) (Mirjalili ve Lewis, 2016), Ağaç Tohum Algoritması (TSA) (Kiran, 2015), Yapay Alg Algoritması (AAA) (Uymaz ve ark., 2015a) gibi pek çok doğadan ilhamlı metasezgisel optimizasyon algoritması mevcuttur.

GA, karmaşık optimizasyon problemlerini çözmek için biyolojideki doğal seçimden esinlenerek geliştirilmiş, klasikleşmiş bir evrimsel arama algoritmasıdır. GA seçim, çaprazlama ve mutasyon olmak üzere üç temel evrimsel davranışı modelleyen süreçten oluşur. Popülasyondaki her bir birey bir aday çözümü temsil eder. İterasyonlar boyunca algoritmadaki üç süreç ile popülasyon kendini geliştirir ve popülasyondaki en iyi birey en iyi çözümü temsil eder. Seçim sürecinde çoğunlukla rulet seçimi kullanılır, dolayısıyla yüksek uygunluk değerini sahip olan iyi bireyin seçilme ihtimali daha fazladır. Bu yaklaşım minimuma yakınsama hızını artırmakla beraber yerel minimumlara da kolayca takılmaya yol açar. GA'nın başarısı çaprazlama yöntemi, seçim stratejisi, mutasyon oranı ve popülasyon büyüklüğüne bağlı olarak değişiklik gösterebilir (John, 1992).

PSO kuş ve balık sürülerinin sosyal davranışlarından ilham alarak geliştirilen en popüler doğadan ilhamlı metasezgisel optimizasyon algoritmasıdır. Sürüdeki bireylerin konumu aday çözümü temsil eder. Bireyler konumlarını mevcut iterasyona kadar ki kendi

en iyi konumları ve sürüdeki en iyi bireyin konumunu kullanarak iterasyonlar boyunca güncellerler. Uygunluk fonksiyonu bireylerin konumlarının kalitesini değerlendirir ve iterasyonlar tamamlanınca en iyi bireyin konumu aday çözümü temsil eder. Algoritmadaki bireyler en iyi bireye göre kendilerini güncellediği için yüksek sömürü (exploitation) kapasitesine ve hızlı yakınsamaya sahiptir. Ancak bu durum keşif (exploration) kabiliyetini düşürür ve kolayca yerel minimumlara takılmasına yol açabilir. Ayrıca bu algoritma matematiksel denklemleri gereği kontrol parametrelerine yüksek bağımlılık göstermektedir (Kennedy ve Eberhart, 1995).

ABC bal arılarının akıllı yiyecek arama davranışlarından esinlenerek geliştirilmiş bir doğadan ilhamlı metasezgisel optimizasyon algoritmasıdır. ABC’de işçi arı, gözcü arı ve kâşif arı olmak üzere üç farklı arı grubu süreçte yer alır. İşçi arılar alan bilgisine dayanarak yiyecek arama yaparlar. Bu arama lokal aramayı (sömürü) temsil eder. Gözcü arı çözüm kalitesini temsil eden nektar miktarına göre bir yiyecek kaynağı belirler. Kâşif arı çözüm uzayında rastgele bir yiyecek seçer. Bu üç farklı arı grubu farklı yiyecek arama stratejilerini temsil eder. ABC bu stratejiler ile keşif ve sömürüyü dengeleyecek şekilde iterasyonlar boyunca arı popülasyonu geliştirir ve en iyi bireyi en iyi çözüm olarak sunar. Literatürde yapılan çalışmalar ABC’nin keşif kabiliyetinin yüksek olduğunu göstermiştir. Bununla birlikte ABC bazen belirlenen iterasyon sayısında bulduğu çözüm yetersizken bazen de erken yakınsama yapma eğilimi göstermiştir (Karaboga ve Basturk, 2007).

Guguk kuşlarının bazı türleri yuva yapmaz, yavrusuna bakıp büyütmez. Bunun yerine, başka bir kuşun yuvasındaki yumurtaları dışarı atıp, kendi yumurtasını o yuvaya bırakır. Yuvanın sahibi olan kuş, yumurtaları kendi yumurtası sanıp, yumurtaları sahiplenir ve yavruları büyütür. Guguk kuşlarının bu asalak (parazitsel) üreme ve uçuş stratejisinden esinlenerek, CS algoritması geliştirilmiştir. CS algoritmasının temel konum güncelleme mekanizması rastgele uçuş veya levy uçuşu olmak üzere iki farklı uçuş stratejisidir. Yapılan çalışmalar, karmaşık optimizasyon problemlerinin çözümünde CS’nin erken yakınsama eğiliminde olduğunu ve keşif kabiliyetinin yetersiz kalabildiğini göstermiştir (Yang ve Deb, 2009).

Kambur balinalar okyanuslarda sürü şeklinde hareket eden ve özel avlanma stratejilerine sahip olan en zeki canlı türlerinden birisidir. WOA kambur balinaların avlanma stratejilerinden ve sosyal davranışlarından esinlenen bir doğadan ilhamlı metasezgisel optimizasyon algoritmasıdır. Algoritma temelde balinaların av arama (search for prey), avını çevrelemesi (encircling prey) ve avlanırken kullandığı kabarcık

saldırı (bubble-net attack) süreçlerinden oluşur. İterasyonlar boyunca av arama ile keşif süreci, kabarcık saldırısı ile de sömürü süreci çalıştırılarak popülasyon geliştirilir. İterasyonlar sonlandığında en iyi ajanın (balina) boyutları çözüm kümesinin parametrelerini oluşturur. Çok erken yakınsama yapan algoritmalarından birisidir. Bu problemini aşmak için literatürde birçok hibrit versiyonu geliştirilmiş, farklı yelpazedeki birçok gerçek dünya probleminin çözümünde uygulanmıştır (Mirjalili ve Lewis, 2016).

AAA mikro alglerin yaşam davranışlarından ve büyüme kinetiklerinden esinlenen bir başka metasezgisel optimizasyon algoritmasıdır. Algoritma helisel hareket, evrimsel süreç ve adaptasyon süreci olmak üzere üç temel bileşenden oluşmaktadır. Helisel hareket alglerin su içerisindeki enerji ve besin yoğunluklarına göre gerçekleştirdiği helisel hareketi modellemektedir. Evrimsel süreçte yeterince büyümemiş alglerin en iyi alglere yaklaşmaya çalıştığı bir süreçtir. Adaptasyon süreci ise sadece belirli iterasyonlar da gerçekleşen yeterince beslenememiş alglerin en iyi alglere benzemeye çalıştığı bir süreçtir. Bu algoritma geniş bir yelpazedeki optimizasyon problemlerini başarıyla çözmektedir. Bu nedenle, geliştirilen yeni optimizasyon modelleri için de ilham kaynağı olmaktadır (Uymaz ve ark., 2015a).

Doğadan ilhamlı metasezgisel algoritmaların çalışma şekilleri birçok yönden birbirlerine benzemektedir. Popülasyon oluşturulması, keşif ve sömürü mekanizmalarını barındırması, en iyi ajanı saklaması, iterasyon başlarında kuvvetli keşif süreçleri, iterasyon sonlarına doğru güçlü sömürü mekanizmalarını çalıştırması benzer yönleridir. Bununla birlikte; her algoritmanın ilham aldığı yapı gereği, farklı konum güncelleme denklemleri, farklı en iyi ajanı takip etme stratejileri, özgün yerel ve global arama süreçleri gibi kendine has değişiklikleri bulunmaktadır. Örneğin; PSO tüm boyutlarda, ABC bir boyutta, AAA üç boyutta, DE rastgele boyutlarda, çözüm uzayında güncelleme (bireyde değişiklik) yapar. Dolayısıyla farklı optimizasyon algoritmalarının çeşitli optimizasyon problemlerinde farklı üstünlükleri vardır.

1.1. Tezin Literatüre Katkısı

Literatürde birçok optimizasyon algoritması olmasına rağmen yeni algoritmalar tasarlanmakta ve var olan algoritmaların geliştirilmesine ihtiyaç duyulmaktadır. “Bedavaya Yemek Yok” (No Free Lunch) teoremine göre bir optimizasyon algoritmasının tüm optimizasyon problemlerini başarıyla çözmesi mümkün değildir

(Wolpert ve Macready, 1997; Ho ve Pepyne, 2002). Tez çalışmamızın temel motivasyonlarından birisi de budur. Bu tez çalışması AAA'nın güçlendirilerek farklı problemler üzerindeki performansını arttırmak ve daha önce uygulanmamış problemlere uygulanacak şekilde geliştirilmesini hedeflemektedir.

AAA'nın diğer metasezgisel algoritmalarla benzer şekilde iki temel problemi mevcuttur. Bunlardan birisi erken yakınsama, diğeri yerel minimumlara takılma eğilimidir. Metasezgisel algoritmaların takıldığı bu iki problemin üstesinden gelmek ve performanslarını artırmak için literatürde çeşitli stratejiler kullanılarak yeni yöntemler önerilmiştir. Bunlardan bazıları; levy uçuşu (Haklı ve Uğuz, 2014), yerel arama (Wu ve ark., 2014), zeki arama (Kiran ve Haklı, 2021), değişken arama (Kiran ve ark., 2015), çoklu ajan kullanımı (Kumar ve ark., 2011), kuantum davranışı (Sun ve ark., 2004) ve kaos teorisi (Atalı ve ark., 2021) gibi stratejilerdir. Bu algoritma geliştirme stratejilerinden birisi de kaos teorisinden esinlenerek geliştirilen kaotik haritalardır ve literatürde gösterdiği başarılı sonuçlarla oldukça popüler olmuştur. Kaotik haritalar algoritmaların verimli bir şekilde keşif ve sömürü süreçlerini dengeleyerek, erken yakınsamadan ve yerel minimumlardan kurtulmasına yardımcı olmaktadır (Demir ve ark., 2020).

Tez çalışması kapsamında ilk olarak AAA, kaotik stratejilerle donatılarak Kaotik Yapay Alg Algoritması geliştirilmiştir. Geliştirilen bu algoritmanın çeşitli benchmark test fonksiyonları ve gerçek dünya problemleri üzerinde test edilerek başarısı doğrulanmıştır.

Geliştirilen kaotik AAA tabanlı yaklaşım, makine öğrenmesinin üç temel alanından birisi olan gözetimsiz öğrenmedeki (unsupervised learning) paylaşılmalı kümeleme (partitional clustering) problemine uygulanarak analiz edilmiştir.

Daha sonra önerilen kaotik AAA yaklaşımı, makine öğrenmesi algoritmaları için kaçınılmaz kritik bir önışleme süreci olan öznitelik seçiminde (feature selection) kullanılmak üzere ikili (binary) versiyonu geliştirilmiştir. Geliştirilen İkili Kaotik Yapay Alg Algoritması literatürdeki farklı zorluk seviyesine sahip veri setleri üzerinde uygulanmış ve popüler algoritmalarla kıyaslanmıştır. İstatistiksel testleriyle birlikte analiz edilmiştir.

Sonuç olarak tez çalışmasının literatüre üç temel konuda katkı sağladığı ifade edilebilir. Bunlardan ilki kaotik yaklaşımların AAA'da ilk kez kullanımıyla güçlü bir metasezgisel algoritmanın geliştirilmesidir. Bir diğeri geliştirilen yöntemin gerçek dünya optimizasyon problemlerinde başarıyla uygulanmasıdır. Üçüncü olarak da öznitelik

seçimi ve kümeleme gibi veri bilimi problemlerinin çözülmesi için güvenilir ve yeni bir sürü zekası yaklaşımı önermektedir.

1.2. Tezin Organizasyonu

Tezin devamı aşağıdaki gibi organize edilmiştir.

Tezin ikinci bölümünde üç başlık altında kaynak araştırması yapılmıştır. İlk başlıkta önerilen algoritmanın geliştirilmesinde kullanılan kaos teorisi ve kaotik haritaların metasezgisel algoritmalarındaki uygulamaları araştırılmıştır. İkinci başlıkta kümeleme analizinde kullanılan yöntemler, özellikle metasezgisel yaklaşımlar incelenmiştir. Üçüncü başlıkta ise öznitelik seçimi için kullanılan algoritmalar hakkında literatür araştırılması araştırmaları yapılmıştır.

Üçüncü bölümde materyal ve yöntem altı başlık altında detaylandırılmıştır. İlk olarak orijinal yapay alg algoritmasının matematiksel denklemleriyle birlikte mekanizmaları ve süreçleri açıklanmıştır. İkinci başlıkta kaos teorisi ve kaotik haritalar, üçüncü başlıkta geliştirilen kaotik yapay alg algoritması akış diyagramlarıyla desteklenerek anlatılmıştır. Dördüncü başlıkta kaotik yapay alg algoritması kümeleme analizinde kullanılmak üzere organize edilmiş, beşinci başlıkta öznitelik seçiminde kullanılmak için geliştirilen ikili kaotik yapay alg algoritması sözde kodlarıyla birlikte açıklanmıştır. Son başlıkta ise geliştirilen yöntemlerin test edilmesi için kullanılan optimizasyon problemleri ve değerlendirme metrikleri detaylandırılmıştır.

Dördüncü kısımda, yapılan deneysel çalışmalar dört başlık altında analiz edilmiştir. İlk başlıkta tezde geliştirilen kaotik yapay algoritmasının çeşitli benchmark fonksiyonlarındaki başarısı, ikinci başlıkta mühendislik optimizasyon problemlerindeki etkinliği, üçüncü başlıkta kümeleme analizindeki başarımı, dördüncü başlıkta öznitelik seçimi problemindeki performansı detaylı bir şekilde analiz edilmiştir.

Beşinci bölümde ise tezde ortaya çıkan sonuçlar açıklanmış ve gelecekte yapılacak çalışmalar için öneriler verilmiştir.

2. KAYNAK ARAŞTIRMASI

Bu tezin amacı kaos teorisinden ilham alarak, AAA'yı geliştirmek ve geliştirilen yeni yaklaşımı, paylaşımlı kümeleme ve öznitelik seçim problemlerine uygulamaktır.

Bu doğrultuda kaynak araştırması üç bölümden oluşturulmuştur. Birinci bölümde kaos teorisi ve kaotik haritaların metasezgisel yaklaşımlarda kullanımı araştırılmıştır. İkinci bölümde paylaşımlı kümeleme analizinde kullanılan metasezgisel yaklaşımlar için kaynak araştırması yapılmıştır. Üçüncü bölümde ise öznitelik seçimi için kullanılan algoritmalar, özellikle metasezgisel tabanlı sarmalayıcı (wrapper) teknikler sunulmuştur.

2.1. Kaos Teorisi ve Kaotik Harita Tabanlı Metasezgisel Yaklaşımlar

Metasezgisel optimizasyon yöntemleri başlangıçta rastgele oluşturulan birden fazla çözümü içeren popülasyon ile başlar. Çözüm uzayında algoritmanın yapısında belirlenmiş davranış modeli çerçevesinde aramalar gerçekleştirir. Algoritmalar yinelemeler boyunca optimum çözüme ulaşmaya çalışırlar.

Yöntemlerin esinlendiği doğal davranış modelleri yonteme ait parametreleri, çözüm uzayının keşfinde kullanılan süreçleri ve mevcut çözümlerin iyileştirilmesini amaçlamaktadır. Çözümlerin iyileştirilmesi için keşif ve sömürü kabiliyetleri kullanılır. Keşif ve sömürü kabiliyetleri birbirleri ile önemli bir değiş-tokuş (trade-off) içerisindedir. Yani sömürü kabiliyetinin fazla olması keşif kabiliyetini kısıtlar, keşif kabiliyetinin fazla olması da sömürü kabiliyetini baskılar. Lokal minimum noktalarından kaçınarak, optimum çözüme hızlı bir yakınsama için metasezgisel yöntemlerin ideal keşif ve sömürü dengesini kurması son derece önemlidir (Tang ve ark., 2018).

Kaos, doğrusal olmayan sistemlerin düzensiz davranışlarıyla ilgilenir. Kaotik sistemler başlangıç koşullarına bağımlı, rastgele davranışlar sergileyen, lineer olmayan (non-linear) deterministik sistemlerdir. Bu özelliklerinden dolayı kaotik sistemler metasezgisel yaklaşımlarda sıklıkla kullanılmıştır. Kaotik yaklaşımların metasezgisel yöntemlerde farklı kullanım amaçları vardır; başlangıç popülasyonunun oluşturulmasında, aday çözümlerin rastgele hareketlerinin gerçekleştirilmesinde, yöntemlerin kritik parametrelerinin belirlenmesinde veya bu modellerin melez (hybrid) kullanılması şeklinde ortaya çıkmaktadır. Literatürde kaotik yaklaşımlara sahip pek çok metasezgisel optimizasyon yöntemi mevcuttur (Atalı ve ark., 2021).

Metasezgisel algoritmaların performansının iyileştirilmesi için önerilen önemli kaotik yaklaşımlardan bazıları bu bölümde sunulmuştur.

WOA balinaların av araması, buldukları avı çevrelemesi ve avlanırken kullandığı kabarcık saldırısı süreçlerini kullanarak optimum çözümü aramaktadırlar. Kabarcık saldırısı ve av arama aşamalarında sırasıyla spiral bir hareketle ve rasgele seçilen bir komşu ile ajanın (balina) konumu güncellenirken, avın çevrelenmesi sürecinde popülasyondaki en iyi ajanın bilgisi kullanılarak mevcut çözümün konumu güncellenmektedir. Ajanın konum güncellemesinin hangi süreç ile gerçekleştirileceğine p parametresi tarafından karar verilir. 2018 yılında Kaur ve Arora tarafından, p parametresi on farklı kaotik harita ile donatılarak Kaotik Balina Optimizasyon Algoritması (CWOA) geliştirilmiştir. Geliştirilen yaklaşımla klasik WOA yönteminin yerel optimuma sıkışma problemini aşması ve çözüm uzayının tarama kapasitesini artırması hedeflenmiştir. CWOA yönteminin performansı yirmi veri seti kullanılarak popüler algoritmalarla kıyaslanmıştır. Yapılan istatistiksel testler bu algoritmanın performansını doğrulamanın yanı sıra, global optimumun bulunmasında da güvenilir bir yaklaşım olduğunu göstermiştir (Kaur ve Arora, 2018).

Kril Sürüsü (KH) algoritmasının ana parametreleri atalet ağırlıklarıdır (inertia weights). Bu parametreler, yakınsama hızına ve ajanların (kril) konumlarını güncelleme stratejilerine karar verme sürecinde kritik bir öneme sahiptir. KH'de, atalet ağırlıkları (ω_n , ω_f), keşif yeteneği için başlangıçta 0,9'a ayarlanır ve arama sürecinde lineer olarak 0.1'e kadar düşürülür. Wang ve ark. tarafından atalet ağırlıkları (ω_n , ω_f) parametreleri kaotik haritalar ile donatılarak kaotik KH algoritması önerilmiştir. Önerilen yöntem on dört popüler benchmark fonksiyonunda literatürdeki güçlü algoritmalarla kıyaslanarak performansı değerlendirilmiştir. Çalışmanın istatistiksel testleri kaotik KH algoritmasının performansını doğrulamıştır (Wang ve ark., 2014).

Çekirge Optimizasyon Algoritması (GOA) $c1$ ve $c2$ olmak üzere iki kritik parametreye sahiptir. Keşif ve sömürü mekanizmaları $c1$ parametresi ile dengelenir. $c2$ parametresi ise çekirgeler arasındaki çekim bölgesini, konfor bölgesini ve itme bölgesini ayarlar. Bu parametreler, yakınsama hızına ve çekirgelerin konumunun güncelleme stratejisine karar verme sürecinde kritik bir öneme sahiptir. GOA'da, bu parametreler, optimizasyon sürecinin ilk aşamasında 1'e ayarlanır ve lineer olarak azaltılır. 2019 yılında Arora ve Anand tarafından kaotik GOA algoritması geliştirilerek, algoritmanın $c1$ ve $c2$ parametreleri çeşitli kaotik haritalar kullanılarak ayarlanmıştır. Circle kaotik haritasına

dayalı kaotik GOA yaklaşımının performansı on üç benchmark setinde çeşitli popüler yöntemlerle kıyaslanarak doğrulanmıştır. (Arora ve Anand, 2019c).

Salp, boyu 1-10 cm arasında olan saydam bir balık türüdür. Planktonlarla beslenirler. Hayalet balık olarak da isimlendirilir. Salp Sürüsü Algoritması (SSA), bu balık sürülerinin okyanuslarda yiyecek arama ve beslenme davranışlarından ilham alan metasezgisel bir optimizasyon algoritmasıdır. SSA'nın performansını etkileyen kritik üç ana parametre mevcuttur. Bu parametreler r_1 , r_2 ve r_3 'tür. r_1 ve r_2 parametreleri yinelemeler boyunca doğrusal olarak azaltılarak salpın konumunun güncellenmesini etkileyen iki önemli parametredir. r_3 parametresi ise salpın bir sonraki konumunun güncellenme stratejisinden sorumlu olan parametredir. 2018 yılında Sayed ve ark. tarafından r_2 parametresini ayarlamak için kaotik haritaların kullanımını öneren kaotik SSA yaklaşımı geliştirilmiştir. Önerilen yaklaşımın performansı yirmi farklı veri seti üzerinde literatürdeki iyi bilinen tekniklerle kıyaslanmıştır. Çalışmanın istatistiksel testleri önerilen yaklaşımın etkinliğini doğrulamıştır (Sayed ve ark., 2018).

Biyocoğrafya Optimizasyon Algoritması (BBO) biyolojik coğrafyadan esinlenilerek ortaya konulmuş, evrimsel bir optimizasyon algoritmadır. Algoritmadaki ajanlar yani aday çözümler emigration, immigration ve mutasyon süreçlerini kullanarak çözüm uzayında hareket etmektedir. Emigration, bir bölgenin yetersiz koşullar nedeniyle terkedilmesidir. Immigration ise bir bölgedeki verimli koşullar nedeniyle, oraya akın edilmesidir. Bir bölgedeki şartlara, ajanların kendilerini değiştirerek uyum göstermelerine ise mutasyon denilmektedir. BBO algoritmasının yerel minimum noktalarına yakalanma ve yavaş yakınsama problemleri bulunmaktadır. Saremi ve ark. 2014 yılında yaptıkları çalışmada BBO yöntemi için önemli olan emigration, immigration ve mutasyon parametrelerinin on farklı kaotik harita ile belirlendiği bir yaklaşımı sunmuştur. Sunulan yaklaşım tek modlu (tek bir global minimumu olan) ve çok modlu (birden çok yerel minimumu sadece bir tane global minimumu olan) nümerik benchmark fonksiyonlarında test edilmiş ve orijinal yöntemin keşif ve sömürü kabiliyetlerini artırdığı gözlemlenmiştir (Saremi ve ark., 2014).

Jothiprakash ve Arunkumar, 2013 yılında yaptıkları çalışmada evrimsel optimizasyon algoritmalarından GA ve DE algoritmasının kaotik versiyonlarını sunmuşlardır. Çalışmada GA ve DE yönteminin başlangıç popülasyonları lojistik kaotik haritasını kullanılarak oluşturulmaktadır. Önerilen kaotik GA ve DE yaklaşımları “maximizing the hydropower production from a reservoir” probleminin çözümünde

kullanılmıştır. Yapılan deneysel çalışma klasik GA ve DE algoritmalarının kaotik versiyonlarının daha başarılı sonuçları olduğunu göstermiştir (Jothiprakash ve Arunkumar, 2013).

2015 yılında yapılan bir çalışmada GA’da yeni bireylerin üretilmesinde önemli bir süreç olan çaprazlama operatörü kaotik yaklaşımlar ile birleştirilerek yeni kaotik çaprazlama operatörü geliştirilmiştir. Yeni kaotik çaprazlama operatörü sayesinde klasik GA yönteminin keşif yeteneği artırılmıştır (Snaselova ve Zboril, 2015). Benzer bir çalışma da DE algoritmasında aday çözümün ve komşularının belirlenmesinde kullanılmıştır. Önerilen stratejide yeni bireylerin yerel minimum noktalarından kaçınması için kaotik mutasyon operatörü kullanılmıştır. Sunulan yeni mutasyon operatöründe lojistik kaotik haritası kullanılmıştır (Zhenyu ve ark., 2006).

Cheng ve Huang tarafından 2010 yılında yapılan çalışmada kaotik GA destekli K-means yaklaşımını sunmuşlardır. Bu yaklaşımda başlangıç popülasyonu kaotik bir yaklaşımla üretilmiş ve klasik GA operatörleri (seçim, çaprazlama ve mutasyon) ile kaotik yaklaşımları melezlenmiştir (Cheng ve Huang, 2010).

Akhtar ve ark. 2019 yılında yaptıkları çalışmada “substitution box design” problemlerinin çözümünde CS yöntemini kaotik bir versiyonunu kullanmıştır. Bu versiyonda başlangıç popülasyonunun oluşturulmasında lojistik ve tent kaotik haritaları kullanılmıştır. Başlangıç popülasyonunun oluşturulmasında kaotik haritaların performansı önemli ölçüde arttırdığını savunmuşlardır (Akhtar ve ark., 2019).

Lui ve ark. yaptıkları bir çalışmada klasik sürü optimizasyon algoritması olan PSO yönteminin performansının geliştirilmesi için iki modifikasyon geliştirilmiştir. Bunlarda ilki “adaptive inertia weight factor” modifikasyonu sayesinde yöntemin keşif kabiliyeti artırılmıştır. İkinci olarak lojistik kaotik haritası kullanılarak en iyi bireyin uygunluk değerini geliştirecek bir kaotik yerel arama modifikasyonu eklenmiştir. Kaotik yerel arama modifikasyonu önerilen yöntemin sömürü kapasitesini arttırmıştır (Liu ve ark., 2005).

GSA, Newton'un kütle çekim yasasından esinlenerek geliştirilmiş bir metasezgisel optimizasyon algoritmasıdır. Bu algorithmada, popülasyondaki bireylerin konumlarının güncellenmesi aşamasında, erken yakınsamanın üstesinden gelmek ve yerel minimumlardan kaçınmak için yerel ve global arama arasında geçişi sağlamak üzere rastgele bir sayı kullanılır. Yapılan çalışmada erken yakınsama ve yerel minimumlardan

kaçma yeteneğini geliştirmek için, kaotik dinamiğe dayalı bir kaotik operatör eklenmiş ve klasik GSA'nın performansı artırılmıştır (Han ve Chang, 2012).

Meyve Sineği Algoritması (FOA) algoritması meyve sineklerinin yiyecek bulma davranışlarından ilham alarak geliştirilmiş metasezgisel bir optimizasyon algoritmasıdır. Algoritmada yeni gıda kaynaklarının üretiminde tekdüze (uniform) dağılım ile rastgele değişkenler kullanır. Mitić ve ark. tarafından önerilen kaotik FOA yaklaşımı gıda kaynaklarının üretimi için kaotik haritalar ile değeri belirlenen yeni bir parametre (alpha) geliştirmişler ve klasik FOA algoritmasının performansını arttırmışlardır (Mitić ve ark., 2015).

2.2. Kümeleme Analizinde Metasezgisel Algoritmaların Kullanımı

Bu bölümde, paylaşımlı kümeleme sürecinde metasezgisel algoritmalarla yapılmış çeşitli çalışmalar sunulmaktadır.

Singh, veri kümeleme problemlerinin verimli bir şekilde çözülmesi için Harris Şahini Optimizasyon (HHO) algoritmasını önermiştir. Bu çalışma, HHO algoritmasının arama modeline rehberlik etmek için kaotik yaklaşımları kullanmış ve HHO algoritmasının rastgele parametrelere bağımlı olma durumunu ortadan kaldırmıştır. HHO algoritmasının performansı, on iki kümeleme veri seti kullanılarak değerlendirilmiş ve altı iyi bilinen teknikle karşılaştırılmıştır. HHO algoritmasının etkinliği, çeşitli performans ölçütleri ve istatistiksel testler kullanılarak doğrulanmıştır. Bu testler, HHO algoritmasının veri kümeleme problemlerini çözme üzerindeki etkinliğini desteklemiştir (Singh, 2020).

Tsal ve ark. paylaşımlı kümeleme tekniklerinin etkinliğini ve verimliliğini artırmak ve büyük miktarda veriyi kümelemek için, Mercan Resifleri Optimizasyonu (CRO) algoritmasından esinlenen Coral Reef Optimization with Substrate Layers (CRO-SL) adı verilen yeni bir metasezgisel algoritma geliştirmişlerdir. Substrate Layers konseptinde sonuçları iyileştirmek için PSO ve Genetik k-means Algoritmasını (GKA) entegre etmişlerdir. CRO-SL algoritmasında amaç büyük veri kümeleme için daha iyi sonuçlar elde etmektir. Ayrıca, veri kümelemenin çalışma süresini azaltmak için önerilen CRO-SL algoritması bulut (cloud) platformunda uygulamışlardır. Performans karşılaştırması için k-means, GKA, PSO, simple coral reef optimization (SCRO) gibi çeşitli kümeleme algoritmalarını kullanmışlardır. CRO-SL algoritmasını dokuz popüler

benchmark veri setinde deęerlendirmişlerdir. Çalışma sonuçlarında, önerilen CRO-SL algoritmasının, bulut platformunu kullanan dięer algoritmalarla kıyasla kümeleme sonuçlarını hızlandırdığı gösterilmiştir (Tsai ve ark., 2019).

Kuwil ve ark. kritik mesafe kümeleme CDC (Critical Distance Clustering) algoritması olarak adlandırılan yeni bir mesafe tabanlı kümeleme algoritması tasarlamışlardır. Bu algoritma, veriler arasındaki benzerliği belirlemek için yeni bir uygunluk fonksiyonu kullanmaktadır. Uygunluk fonksiyonu, Öklid mesafesi ve temel istatistik yöntemleri kullanılarak tasarlanmıştır. Ayrıca önerilen algoritma sadece nitel deęil nicel ve kategorik verilerle de çalışabilmektedir. Geliştirilen algoritmanın performansı yirmi altı kümeleme veri seti üzerinden incelenmiştir. Çalışma sonuçlarında önerilen algoritmanın k-means, DBSCAN ve MST tabanlı kümeleme algoritmalarına kıyasla rekabetçi kümeleme sonuçları sergilediği görülmüştür. Ayrıca bu yeni mesafe tabanlı algoritma ile aykırı deęerlerin başarılı bir şekilde tespit edildiği savunulmuştur (Kuwil ve ark., 2019).

Singh ve ark. kümeleme algoritmalarının yavaş yakınsama ve lokal optimumlara takılma sorunlarını aşmak için Yapay Kimyasal Reaksiyon Optimizasyon (ACRO) kümeleme algoritması adı verilen yeni bir metasezgisel algoritma geliştirmiştir. Yavaş yakınsama ve lokal optimumlara takılma problemleri sırasıyla pozisyon ve komşuluk operatörleriyle çözülmeye çalışılmaktadır. ACRO algoritmasının performansı, beş benchmark ve iki yapay veri seti üzerinde test edilmiştir. ACRO'nun sonuçları klasik kümeleme algoritmalarıyla kıyaslanmıştır. Sonuçlar, ACRO algoritmasının küme içi uzaklık ve f-measure metrikleri açısından daha başarılı olduğunu göstermiştir. Ayrıca, Friedman istatistiksel testi ile ACRO algoritmasının kümeleme alanındaki etkinliğini doğrulanmıştır (Singh ve ark., 2019).

Baalamurugan ve Bhanu, bulut ortamında veri kümeleme analizi için Efficient Stud Krill Herd Clustering (ESKH-C) tekniğini geliştirmiştir. ESKH-C tekniğinin amacı, küme merkezlerinin optimum noktalarının pozisyonlarını hesaplamaktır. Ayrıca, daha verimli hale getirmek için bu algoritmaya Stud seçimi ve Crossover operatörü (SSC) entegre edilmiştir. SSC operatörü, genetik algoritmaların üreme (reproduction) sürecinden esinlenerek tasarlanmıştır ve amacı yakınsama hızını iyileştirmektir. ESKH-C algoritmasının performansı yedi veri seti üzerinde doğrulanmıştır. Geliştirilen algoritma k-means, PSO, ACO ve Bakteri Yiyecek Arama Optimizasyonu (BFO) algoritmaları ile karşılaştırılmıştır. Deneysel sonuçlar, ESKH-C algoritmasının farklı

küme sayısı ve yoğunluğuna sahip çok boyutlu veri setlerini etkin bir şekilde kümelediğini göstermiştir (Baalamurugan ve Bhanu, 2019).

Sharma ve Chhabra, kümeleme algoritmalarının keşif ve sömürü süreçlerinin ideal dengesini oluşturmak için PSO ve çok eşli çaprazlamadan (polygamous crossover) esinlenerek PSOPC adı verilen yeni bir kümeleme algoritması geliştirmiştir. PSOPC algoritmasını yedi klasik kümeleme veri seti üzerinde PSO, GA, DE, FA ve GWO yaklaşımları ile kıyaslamışlardır. PSOPC algoritmasının performansını değerlendirmek için küme mesafesi, küme saflığı ve yakınsama hızı metriklerini kullanmışlardır. Bu metrikler baz alındığında PSOPC algoritmasının kıyaslanan algoritmalarından daha performanslı olduğunu savunmuşlardır (Sharma ve Chhabra, 2021).

Abdulwahab ve ark. kümeleme analizi için Levy Flight Black Hole (LBH) adı verilen etkili bir kümeleme algoritması tasarlamıştır. LBH algoritması, levy uçuşu stratejisini Kara Delik Optimizasyon (BHO) algoritmasına eklemektedir. Yazarlar, BHO algoritmasının bazı veri setlerinde etkin sonuçlar göstermekle birlikte bazı veri setlerinde keşif kabiliyetinin yetersiz olduğunu belirtmişlerdir. BHO algoritmasının keşif sürecindeki eksikliği levy flight stratejisi ile güçlendirerek arama uzayını daha etkin araştırmayı amaçlamışlardır. LBH algoritmasının performansı, altı klasik kümeleme veri seti üzerinde iyi bilinen tekniklerle karşılaştırılmıştır. Çalışma sonuçların da, LBH algoritmasının etkin bir şekilde kümeleme analizi yapabildiği gösterilmiştir (Abdulwahab ve ark., 2019).

Tarkhaneh ve Moser, veri kümeleme için DE algoritmasını geliştirerek Geliştirilmiş Diferansiyel Evrim (IDE) isminde yeni bir yaklaşım önermiştir. IDE algoritmasına küme analizinin etkinliğini arttırmak için Arşimet spirali, Mantegna Levy dağılımı ve komşuluk araması (neighborhood search) süreçleri eklenerek "Adaptive Differential Evolution With Neighborhood Search" (ADENS) yaklaşımını geliştirilmişlerdir. ADENS algoritmasının performansını on iki kümeleme veri seti üzerinde klasik tekniklerle kıyaslanmıştır. Deneysel sonuçlar, ADENS algoritmasının diğer algoritmalara kıyasla daha üstün kümeleme başarısı verdiği göstermişlerdir. Ayrıca, Wilcoxon ve Friedman istatistiksel testleriyle ADENS algoritmasının performansı doğrulanmıştır (Tarkhaneh ve Moser, 2019).

Klasik kümeleme yaklaşımlarında küme merkezlerinin başlangıçta rastgele seçilimi, özellikle çok boyutlu ve karmaşık veri setlerinde erken yakınsamaya neden olmakta ve yaklaşımların başarısını azaltmaktadır. Bu problemin üstesinden gelmek için

Agbaje ve ark. FA ile PSO algoritmasını melezleyerek FAPSO adında yeni bir yaklaşım geliştirmiştir. Önerilen FAPSO yaklaşımında, ilk olarak FA algoritması uygulanmış ve daha sonra optimal çözümü elde etmek için PSO algoritması kullanılmıştır. Algoritmanın performansı, on iki klasik kümeleme veri seti üzerinde dört iyi bilinen kümeleme algoritması ile karşılaştırılmıştır. Çalışmanın deneysel sonuçları, FAPSO'nun DB (davies bouldin) ve CS (clustered index) kümeleme indeksleri açısından diğer kümeleme yöntemlerinden daha avantajlı olduğunu göstermiştir (Agbaje ve ark., 2019).

Zhou ve ark. k-means algoritmasının başlangıç parametrelerine olan bağımlılığını incelemiş, başlangıç parametreleri bazı durumlarda algoritmanın yerel optimumda sıkışmasına sebep olduğunu savunmuşlardır. Bu problemi çözmek için, veri kümelemede kullanılmak üzere Simbiyotik Organizma Aramasından (SOS) esinlenen yeni bir yaklaşımı önermişlerdir. SOS algoritmasının performansı on kümeleme veri seti üzerinde, Çiçek Tozlaşma Algoritması (FPA), DE, PSO, ABC, CS ve k-means ile karşılaştırmışlardır. Çalışmanın deneysel sonuçlarında, SOS algoritmasının daha etkin kümeleme sonuçları gösterdiğini vurgulamışlardır (Zhou ve ark., 2019).

Aljarah ve ark. GWO algoritmasını kümeleme analizi için önermiş ve algoritmanın yerel optimumlara takılma ve erken yakınsama durumunu gözlemlemişlerdir. Yazarlar bu sorunların çok sayıda değişkenden kaynaklandığını belirtmişlerdir. Bu nedenle, lokal optimum ve erken yakınsama konularını ele almak üzere GWO algoritmasını Tabu Arama (TS) algoritması ile melezleyerek TSGWO adı verilen yeni bir yaklaşımı geliştirmişlerdir. TSGWO'nun performansı, on üç klasik kümeleme veri setinde değerlendirilmiş ve literatürde iyi bilinen kümeleme algoritmalarıyla karşılaştırılmıştır. Çalışma sonuçları, TSGWO'nun iyi bilinen metasezgisel algoritmalara göre daha iyi yakınsama oranı elde ettiğini ve lokal optimumlara sıkışma problemini aştığını göstermiştir (Aljarah ve ark., 2020a).

Zhu ve ark. Yarasa Algoritmasının (BAT) lokal optimumlarda takılma ve performans eksikliklerini göz önünde bulundurarak küme analizi için yarasa algoritması tabanlı yeni bir yaklaşım geliştirmişlerdir. Algoritmanın eksikliklerini gidermek, keşif ve sömürü yeteneklerini geliştirmek için algoritmaya iki iyileştirme yapılmıştır. Bunlardan birisi küresel arama yeteneğini geliştirmek için gauss tabanlı yakınsama faktörü eklenmiştir. Diğer yerel arama yeteneğini geliştirmek için, WOA ve sinüs pozisyon güncelleme mekanizmasını kullanan yeni bir süreç geliştirilmiştir. Önerilen bu yeni algoritmanın performansını değerlendirmek için yedi kümeleme veri seti kullanılmıştır.

Çalışma sonuçlarında önerilen iyileştirmelerin BAT algoritmasının performansını önemli ölçüde geliştirildiği gösterilmiştir (Zhu ve ark., 2020).

Senthilnath ve ark. veri kümeleme probleminin çözümünde çiçeklerin tozlaşma sürecinden esinlenen FPA yaklaşımını kullanmışlardır. Bu algoritmanın kümeleme probleminde kullanım amacı optimum küme merkezlerinin konumlarını hesaplamaktır. FPA'nın performansı üç kümeleme veri seti kullanılarak değerlendirilmiş ve GA, PSO, CS, GWO, DE, Örümcek Maymun Optimizasyonu (SMO) ve Harmoni Arama (HS) algoritmaları ile karşılaştırılmıştır. Çalışma sonuçları, diğer algoritmalarla kıyasla FPA algoritmasının minimum kümeleme hatasına sahip olduğunu göstermiştir. Ayrıca, FPA tabanlı kümeleme algoritmasının etkinliği istatistiksel testler kullanılarak doğrulanmıştır (Senthilnath ve ark., 2021).

Mageshkumar ve ark. paylaşımlı kümeleme analizinin performansını arttırmak için Karınca Aslanı Optimizasyonu (ALO) algoritmasını ACO algoritmasıyla birleştirerek ACO-ALO adı verilen melez bir metasezgisel algoritma geliştirmişlerdir. Kümeleme sürecindeki yerel minimumlara sıkışma problemi aşmak için Cauchy mutasyon operatörünü eklemişlerdir. ACO-ALO algoritmasının performansını değerlendirmek için dört kümeleme veri seti üzerinde k-means ve ACO tabanlı kümeleme algoritmaları ile karşılaştırılmışlardır. Çalışma sonuçlarında, ACO-ALO algoritmasının daha etkin kümeleme analizi yaptığını göstermişlerdir (Mageshkumar ve ark., 2019).

Kaur ve ark. özellikle büyük ve karmaşık veri setleri için k-means algoritmasının lokal optimumlara takılma ve yavaş yakınsama problemini incelemişler ve kaos teorisine ve FPA'ya dayalı CFPA adı verilen yeni bir kümeleme algoritmasını geliştirmişlerdir. CFPA algoritmasının performansı on altı kümeleme veri seti üzerinde test edilmiştir. Test sonuçları FPA, CS, BHO, BAT, PSO, FA ve ABC kümeleme algoritmaları ile karşılaştırılmıştır. CFPA'nın etkinliği, küme bütünlüğü, küme saflığı ve çalışma süresi gibi performans metrikleri kullanılarak ölçülmüştür. Çalışma sonuçları, CFPA algoritmasının küme bütünlüğü ve çalışma süresi açısından diğer algoritmalarından daha iyi sonuçlar elde ettiğini göstermiştir (Kaur ve ark., 2020).

Abualigah ve ark. veri kümeleme ve metin kümeleme problemlerinde küresel arama kabiliyetini geliştirmek için KH algoritması ile HS algoritmasını melezleyerek H-KHA adında yeni bir algoritma önermişlerdir. Bu algoritma HS'nin keşif sürecinden ve KH'nin küresel arama mekanizmasından esinlenerek geliştirilmiştir. Önerilen yöntemin performansı yedi klasik ve altı metin kümeleme veri setinde değerlendirilmiştir. Çalışma

sonuçlarında H-KHA'nın kümeleme doğruluğu ve yakınsama oranı açısından klasik kümeleme algoritmalarından daha iyi sonuçları elde ettiğini göstermişlerdir. Yaptıkları istatistiksel testlerle, F-measure metriği açısından H-KHA yönteminin diğer kümeleme algoritmalarından daha etkin olduğunu doğrulamışlardır (Abualigah ve ark., 2017).

2.3. Öznitelik Seçiminde Metasezgisel Algoritmaların Kullanımı

Son yıllarda nesnelerin interneti, sosyal ağlar ve büyük veri gibi teknolojilerin gelişmesiyle devasa miktarlarda veri üretilmektedir. Bu verideki saklı bilgi ve örüntülerin çıkarılması ve işlenmesi ve anlam kazandırılmasında veri madenciliği ve makine öğrenmesi algoritmaları kullanılmaktadır. Ancak böyle yüksek boyutlu veriler gürültü, alakasız ve fazla özelliklerde içerebilmektedir. Bu ilgisiz özellikler arama uzayı genişleterek öğrenme sürecini ve sınıflandırma etkinliğini zorlaştırmaktadır. Bu problem boyutluluk laneti (curse of dimensionality) olarak tanımlanmaktadır (Nouri-Moghaddam ve ark., 2021). Dolayısıyla verilerin analiz edilmesi için yüksek öğrenme zamanına ihtiyaç duyulmakta ve veri madenciliği süreçleri de ancak yoğun hesaplama gerektiren karmaşık modellerle çözülebilmektedir (Mohammadzadeh ve Gharehchopogh, 2021).

Bu problemlerin üstesinden gelmek için geliştirilen en iyi stratejilerden birisi öznitelik seçimidir. Gereksiz ve alakasız özniteliklerin kaldırılması ve azaltılmasıyla, veriler daha anlaşılır ve yorumlanabilir, modeller daha basitleştirilebilir ve daha kısa sürede eğitilebilir hale gelmektedir. Hatta veriler ilgisiz özniteliklerden arındırıldığı için veri madenciliği tekniklerinin ve sınıflandırma algoritmalarının başarısı artabilmektedir.

Öznitelik seçimi problemi geniş arama uzayı ve öznitelikler arasındaki karmaşık ilişkiler nedeniyle zorlu bir süreçtir. Bir öznitelik gereksiz gibi görünürken gerekli, gerekli gibi görünürken gereksiz olabilmektedir. Özniteliklerinin sayısı n kadar olan bir veri kümesinde özniteliklerin seçimi için oluşturulan alt küme sayısı 2^n kadardır. Bu durum n sayısının büyük olduğu durumlarda aramayı imkânsız hale getirmektedir. Öznitelik seçim problemi bu şekilde üstel olarak arttığı için literatürde polinomsal zamanda çözülemeyen (NP-hard) optimizasyon problemi olarak tanımlanmaktadır (Amaldi ve Kann, 1998).

Literatürde öznitelik seçim problemi için kullanılan yaklaşımlar sarmalayıcı (wrapper) yöntemler ve filtre (filter) tabanlı yöntemler olmaz üzere iki kategoriye ayrılmaktadır.

Filtre tabanlı yöntemler, istatistiksel tekniklere ve bilgi teorisine dayanan seçilen öz nitelikler arasındaki en yüksek korelasyonu saptamaya çalışan tekniklerdir. Örnek olarak ki-kare (chi square) (Liu ve Setiono, 1995), bilgi kazanımı (information gain) (Azhagusundari ve Thanamani, 2013), ortak bilgi (mutual information) (Peng ve ark., 2005), kazanç oranı (gain ratio) (Karegowda ve ark., 2010), temel bileşen analizi (principal component analysis) (Song ve ark., 2010), minimum fazlalık maksimum alaka (minimum Redundancy Maximum Relevancy) (Billah ve Waheed, 2020) gibi teknikler gösterilebilir. Bu yöntemler çalışırken herhangi bir öğrenme algoritmasına ihtiyaç duymadıkları için hızlı çalışırlar ve verinin kendi içerisindeki ilişki ve korelasyondan faydalanırlar.

Sarmalayıcı yöntemler üstel bir çözüm kümesinden en optimum çözümü bulmak için makine öğrenmesi tekniklerinden yararlanan ve genellikle metasezgisel yaklaşım kullanan tekniklerdir. Aslında bir veri kümesindeki gereksiz öz nitelikleri çıkartarak en verimli öz niteliklerin saptamak bir optimizasyon problemidir. Optimizasyon algoritmaları bilhassa metasezgisel yaklaşımlar bu işlem için oldukça elverişlidir ve literatürde filtre tabanlı yöntemlerden daha verimli olduğu birçok çalışma tarafından savunulmuştur (Rodrigues ve ark., 2014; Nouri-Moghaddam ve ark., 2021). Metasezgisel yaklaşımlar her zaman en iyi sonucu garanti etmemekle birlikte makul süre zarfında kabul edilebilir çözüm sunmayı amaçlarlar. Öz nitelik seçimi problemindeki yerel minimumlara takılma ve erken yakınsama problemlerinden ötürü literatürde birçok yaklaşım geliştirilmiştir ve yeni yaklaşımlar için de çalışmalar yapılmaktadır (Li ve ark., 2017).

Bu tez kapsamında metasezgisel algoritmalara odaklanıldığı için kaynak araştırması da bu doğrultuda yapılmıştır. Sarmalayıcı öz nitelik seçim işlemi için önerilen önemli metasezgisel yaklaşımlardan bazıları bu bölümde sunulmuştur.

2016 yılında Emary ve ark. öz nitelik seçimi problemini çözmek için GWO'nun iki farklı yaklaşımını içeren bGWO adında ikili versiyonunu geliştirmişlerdir. Bunlardan birisi popülasyondaki en iyi üç bireye yönelik bireysel adımlar için ikili versiyon geliştirilir ve orijinal algoritmanın yapısındaki üç süreç arasında çaprazlama yapar. Diğer sigmoid fonksiyonunu kullanarak sürekli pozisyon güncelleme değerini ikili uzaya sezgisel eşik fonksiyonlarıyla birlikte taşır. Araştırmacıların önerdikleri yaklaşımları on sekiz farklı veri seti üzerinde PSO ve GA ile kıyaslamış ve daha performanslı olduğunu göstermişlerdir (Emary ve ark., 2016).

2017 yılında Mafarja ve Mirjalili, öznitelik seçim probleminde kullanılmak üzere WOA yaklaşımının iki farklı ikili versiyonunu geliştirmiştir. Bunlardan birisi SA algoritmasının sömürü kabiliyetini WOA'da kullanmak üzere geliştirilen melez versiyon diğeri her iterasyondan sonra WOA'nın en iyi çözümünü geliştirmeye yönelik önerilen yerel arama tabanlı yaklaşımdır. Geliştirilen yaklaşımlar on sekiz farklı veri seti üzerinde üç popüler sarmalayıcı öznitelik seçme algoritmasıyla kıyaslanmıştır. Çalışma sonuçları yeni yöntemin kıyaslanan yöntemlerden daha performanslı olduğunu göstermiştir (Mafarja ve Mirjalili, 2017).

2017 yılında Arora ve Anand, öznitelik seçimi problemini çözmek için Kelebek Optimizasyon Algoritmasının (BOA) ikili versiyonunu geliştirmişlerdir. Algoritmanın geliştirilmesinde literatürdeki popüler dört S-şekilli ve dört V-şekilli transfer fonksiyonu sürekli uzayı ikili uzaya dönüştürmek için kullanılmıştır. Yirmi bir farklı veri setinin kullanıldığı çalışmada önerilen yöntem literatürdeki beş klasik ve dört yüksek performanslı öznitelik seçimi yaklaşımı ile kıyaslanmıştır. Çalışma istatistiksel sonuçları BOA tabanlı önerilen yöntemin performansını doğrulamıştır (Arora ve Anand, 2019b).

2018 yılında Mafarja ve ark. tarafından Yusufçuk Kuşu Algoritmasının (DA) ikili versiyonu öznitelik seçim problemini çözmek için geliştirilmiştir. Çalışmadaki sürekli uzayın ikili uzaya dönüştürülmesinde temel motivasyon olarak sekiz adet zaman değişkenli (time varying) transfer fonksiyonu kullanılmıştır. Keşif ve sömürü arasındaki dengeyi kurmak için iterasyon başlarında yüksek boyut sayısında element değişimi yapılırken optimizasyon sonlarına doğru element değişim pozisyonu sayısı azaltılmıştır. Çalışma sonuçları on sekiz veri seti üzerinde GWO, BBA, BGSA, PSO ve GA ile kıyaslanmış ve karışıklık matrisi, ROG eğrisi ve azaltılan öznitelik sayısı üzerinden analiz edilmiştir. Çalışmanın sonuçları ikili DA algoritmasının performans açısından güvenilir olduğunu vurgulamıştır (Mafarja ve ark., 2018).

2019 yılında Sayed ve ark. tarafından Karga Arama Algoritması (CSA) kaotik operatörlerle desteklenerek ikili versiyonu öznitelik seçme probleminde kullanılmak üzere geliştirilmiştir. Yirmi farklı veri setinin kullanıldığı çalışmada on popüler kaotik harita ile parametre analizi yapılmıştır. Analizde tespit edilen sine kaotik haritası yedi iyi bilinen algoritma ile karışıklık matrisi ve azaltılan öznitelik sayısı üzerinden incelenmiştir. Çalışmanın deneysel sonuçların önerilen yöntemin kıyaslanan yöntemlerden daha performanslı olduğu göstermiştir (Sayed ve ark., 2019).

2019 yılında Mafarja ve ark. tarafından öznitelik seçme problemine yönelik ikili Çekirge Optimizasyon Algoritması (GOA) geliştirilmiştir. Algoritmayı ikili uzaya taşımak için dört adet S-şekilli ve dört adet V-şekilli transfer fonksiyonu kullanılmıştır. Bununla birlikte algoritmanın keşif sürecinin başarısını artırmak için mutasyon operatörü eklenmiştir (bGOA-M). Yirmi beş farklı veri seti üzerinde algoritmanın performansı doğrulanmıştır. Sekiz metasezgisel yöntem ve altı filtre tabanlı yöntem ile sonuçlar kıyaslandığında hem bGOA hem de bGOA-M'nin kıyaslanan algoritmalarından daha rekabetçi olduğu gösterilmiştir (Mafarja ve ark., 2019).

2020 yılında Neggaz ve ark. tarafından öznitelik seçiminde kullanılmak üzere Henry Gazı Çözünürlük Optimizasyonun (HGSO) ikili versiyonu geliştirilmiştir. Algoritmanın oluşturduğu ideal öznitelik kümesinin değerlendirilmesi için KNN ve SVM sınıflandırma algoritmaları kullanılmıştır. Çalışma sonuçları GOA, WOA, GWO, SSA ve DA ile kıyaslanmıştır. Çalışma sonuçlarında HGSO'nun daha verimli olduğu istatistiksel testlerle doğrulanmıştır (Neggaz ve ark., 2020).

2020 yılında Baş ve Ülker, öznitelik seçim problemini çözmek için Sosyal Örümcek Algoritmasının (SSA) ikili versiyonunu geliştirmişlerdir. Bu versiyonda önce sekiz adet transfer fonksiyonu ve eşik fonksiyonları ile sürekli çözümler ikili çözümlere transfer edilmiştir. Daha sonra çaprazlama operatörü eklenerek çözümler güçlendirilmiştir ve en iyi modeli bulmak için analizler yapılmıştır. Geliştirilen yöntem yirmi bir farklı büyüklükteki veri seti üzerinde beş popüler yöntem ile kıyaslanmıştır. Çalışmanın sonuçlarına yapılan istatistiksel testler önerilen ikili SSA algoritmasının performansını doğrulamıştır (Emine ve Ülker, 2020).

2021 yılında Pan ve ark. tarafından Güvercin İlhamlı Optimizasyon (PIO) algoritmasının ikili versiyonunu öznitelik seçim probleminde kullanmak üzere geliştirilmiştir. Çalışmada ikili uzaya dönüşüm için dört adet yeni transfer fonksiyonu tanıtılmıştır. Çözüm kalitesini geliştirmek için yeni bir hız güncelleme şeması ve iki aşamalı konum güncelleme denklemi eklenmiştir. Geliştirilen yöntem literatürdeki popüler veri setleri üzerinde, BGWO ve BPSO algoritmaları ile kıyaslanmıştır. Çalışmanın istatistiksel testleri yapılarak önerilen yöntemin başarısı vurgulanmıştır (Pan ve ark., 2021).

2021 yılında Basset ve ark. tarafından Harris Şahini Optimizasyon (HHO) algoritması SA algoritması ile melezlenerek öznitelik seçim probleminde kullanılmak üzere ikili olarak geliştirilmiştir. Algoritmanın ikili versiyonuna ve (and) ile veya (or)

mantıksal operatörleri eklenmiştir. SA algoritması HHO'nun lokal minimumlardan kaçınma performansını arttırmak için kullanılmıştır. Önerilen yöntemin performansı iyi bilinen algoritmalarla yirmi dört adet standart veri seti ve on dokuz adet yapay veri seti kullanılarak karşılaştırılmıştır. Çeşitli istatistiksel testler ile ikili HHO'nun performansı doğrulanmıştır (Abdel-Basset ve ark., 2021a).

2022 yılında Hu ve ark. tarafından Karadul Örümcek Algoritmasının (BWO) performansı arttırılarak ikili versiyonu öznitelik problemini çözmek için geliştirilmiştir. Orijinal algoritmanın performansını arttırmak için eş seçimi, diferansiyel evrim mutasyon operatörü ve dinamik parametre olmak üzere üç yeni iyileştirme yapılmıştır. Yapılan bu geliştirmelerle algoritmanın hem daha hızlı yakınsama yaparak başarısını arttırmak hem de yerel minimumlardan kaçınmasına yardımcı olması hedeflenmiştir. Çalışma sonuçları doğrulamak için on farklı öznitelik seçme problemi çözülmüş, sonuçlar beş iyi bilinen sarmalayıcı öznitelik seçimi tekniği ile kıyaslanmıştır. Yapılan istatistiksel testler önerilen algoritmanın öznitelik seçimi için en umut verici modellerden birisi olduğunu göstermiştir (Hu ve ark., 2022).

2022 yılında Awadallah ve ark. tarafından öznitelik seçimi probleminde kullanılmak üzere At Sürüsü Optimizasyon (HHO) algoritmasının ikili versiyonu geliştirilmiştir. Sürekli uzayda çalışan algoritmanın ikili uzaya dönüştürülmesi için S-şekilli, V-şekilli ve U-şekilli transfer fonksiyonları kullanılmıştır. Her bir transfer fonksiyonunun dört versiyonu da kullanılmıştır. Ayrıca önerilen yöntemin performansını arttırmak için tek noktalı, iki noktalı ve tekdüze çaprazlama operatörleri uygulanmıştır. Geliştirilen on beş farklı versiyonun yirmi dört farklı veri setinde altı metrik üzerinden analizi yapılmıştır. Çalışma sonuçları önerilen yeni yöntemlerin literatürdeki güçlü algoritmalarla kıyasla rekabetçi olduğu göstermektedir (Awadallah ve ark., 2022).

3. MATERYAL VE YÖNTEM

Materyal ve metot bölümü altı alt başlıkta toplanmıştır. İlk olarak orijinal yapay alg algoritmasının süreçleri matematiksel denklemleriyle birlikte açıklanmıştır. İkinci alt başlıkta metasezgisel algoritmalar için kaos teorisi ve kaotik haritalar detaylandırılmıştır. Üçüncü alt başlıkta kaos teorisinden esinlenerek geliştirilen kaotik yapay alg algoritması akış diyagramıyla desteklenerek anlatılmıştır. Dördüncü alt başlıkta kaotik yapay alg algoritması kümele analizinde kullanılmak üzere organize edilmiş, beşinci alt başlıkta öznitelik seçiminde kullanılmak üzere geliştirilen ikili kaotik yapay alg algoritması akış diyagramı ve sözde koduyla birlikte tanıtılmıştır. Altıncı alt başlıkta ise geliştirilen yöntemlerin test edilmesi için kullanılan optimizasyon problemleri ve değerlendirme metrikleri detaylandırılmıştır.

3.1. Yapay Alg Algoritması

Algler CO_2 ve H_2O kullanarak besinlerini üretebilen mikroorganizmalardır. Algler, alg hücrelerinin bir araya gelmesi ile koloni şeklinde yaşayan canlılardır. Yapay Alg Algoritması, alglerin koloni şeklindeki yaşam davranışlarının esinlenerek geliştirilmiş biyolojik ilhamlı metasezgisel bir optimizasyon algoritmasıdır (Uymaz ve ark., 2015a). Bu algorytmada alglerin karakteristik özellikleri de modellenmiştir. Alg kolonileri daha iyi fotosentez yapabilmek için ışık kaynağına doğru hareket etme, değişen ortam şartlarında adapte olma ve mitoz bölünme ile çoğalma özelliklerine sahiptir. Yapay alg algoritması alglerin bu özelliklerini matematiksel olarak modelleyerek optimizasyon problemlerinde daha iyi çözümler üretmek için geliştirilmiştir. Sırasıyla helisel hareket, adaptasyon süreci ve evrimsel süreci gerçekleşir. Her bir alg kolonisi optimizasyon probleminin olası bir aday çözümünü temsil eder. Alg kolonileri bir araya gelerek alg popülasyonunu oluşturur. D boyutlu ve N alg kolonisinden oluşan popülasyon Denklem 3.1’de ve ilgili alg popülasyonun oluşturulması Denklem 3.2’de gösterilmektedir.

$$\text{Populasyon} = \begin{bmatrix} x_{1,1} & x_{1,2} & \dots & x_{1,D} \\ x_{2,1} & x_{2,2} & \dots & x_{2,D} \\ \vdots & \vdots & \dots & \vdots \\ x_{PN,1} & x_{PN,2} & \dots & x_{PN,D} \end{bmatrix} \quad (3.1)$$

$$Populasyon_i^j = x_{min}^j + r_i^j (x_{max}^j - x_{min}^j) \quad i = 1, 2, \dots, N \text{ ve } j = 1, 2, \dots, D \quad (3.2)$$

Denklemlerde D alg kolonisinin boyutunu (bir kolonideki alg hücresi sayısını), PN alg popülasyonu büyüklüğünü (popülasyondaki koloni sayısını), x_{ij} i . alg kolonisinin j . Alg hücresini temsil etmektedir.

Alg popülasyonu alg kolonilerinden, alg kolonileri de alg hücrelerinden oluşur. Her bir alg hücresi aday çözümün bir parametresini temsil eder. Bu hücreler ısı, ışık, besin gibi yeterli ortam şartları sağlandığında büyürler ve daha iyi bir konuma doğru hareket ederler. Algoritmada alglerin büyüme kinetiği (μ) Monod modeli ile Denklem 3.3'teki gibi hesaplanır. Her bir alg kolonisi helisel hareket, adaptasyon ve evrimsel süreç ile bulunduğu konumdan daha iyi bir konuma geçmeye çalışır. Koloni en optimal konuma geldiğinde optimizasyon probleminin optimum çözümü elde edilmiş olur.

$$\mu_i^t = \frac{\mu_{max}^t \times S_i^t}{K_i^t + S_i^t} \quad (3.3)$$

Denklemden μ_{max} maksimum büyüme oranıdır ve 1 olarak kabul edilir. μ spesifik büyüme oranını, S t zamanındaki uygunluk değerini temsil eden besin çözeltisi konsantrasyonudur. K , t zamanında $(G_i^t/2)$ denklemi ile hesaplanan yarı besin koşullarındaki büyüme oranıdır. G , $t + 1$ zamanındaki alg kolonisinin büyüklüğüdür ve Denklem 3.4 ile hesaplanır.

$$G_i^{t+1} = G_i^t + \mu_i^t G_i^t \quad (3.4)$$

Denklemden G_i^t , t zamanındaki i . alg kolonisinin büyüklüğüdür.

AAA üç temel süreçten oluşur. Birincisi besin çözeltisi içerisindeki alglerin üç boyutlu hareketlerini temsil eden helisel hareket sürecidir. İkinci süreç, yeterli ışık ve besin elde eden alg kolonilerinin mitoz bölünme ile çoğaldığı evrimsel süreçtir. Üçüncü süreç, yeterince gelişemeyen alg kolonilerinin en büyük koloniye benzemeye çalıştığı adaptasyon sürecidir. Tezin ilerleyen kısmında bu süreçler detaylı olarak izah edilecektir.

3.1.1. Evrimsel Süreç

Alg kolonileri daha uygun ortam şartlarına ulaştıkça yani daha iyi çözümler buldukça gelişirler ve büyürler. Uygun ortam şartlarını elde edemeyen yani iyi bir uygunluk değerine ulaşamayan koloniler ise gelişemezler ve ölürler. Helisel hareket süreci tamamlandıktan sonra evrimsel süreçte alg kolonileri büyüklüklerine (G) göre sıralanırlar. En büyük alg kolonisinin rastgele seçilmiş bir hücresi en küçük alg kolonisinin aynı hücresine kopyalanır ve evrimsel süreç tamamlanır. Bu işlemler sırasıyla Denklemler 3.5 – 3.7’de gösterilmektedir.

$$\text{biggest}^t = \max (G_i^t) \quad i = 1, 2, \dots, PN \quad (3.5)$$

$$\text{smallest}^t = \min (G_i^t) \quad i = 1, 2, \dots, PN \quad (3.6)$$

$$\text{smallest}_m^t = \text{biggest}_m^t \quad m = 1, 2, \dots, D \quad (3.7)$$

Denklemlerde m rastgele seçilmiş bir alg hücresini, biggest^t ve smallest^t sırasıyla t zamanındaki en büyük ve en küçük alg kolonilerini temsil etmektedirler (Uymaz ve ark., 2015a).

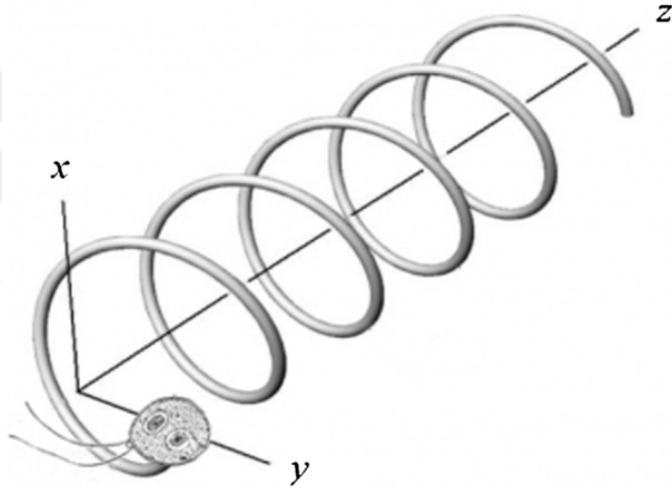
3.1.2. Helisel Hareket

Alg kolonisi daha iyi ortam şartlarına ulaşmak için suda/çözeltide helisel şekilde hareket eder. Alg kolonileri büyüklükleri ile orantılı olarak enerjiye sahiptir. Bu hareketini sahip olduğu enerji miktarınca yapar. Her bir harekette enerji parametresi (e) kadar enerjisinden kaybeder. Yaptığı hareketi sonucunda daha iyi bir ortama ulaşmış ise enerji parametresi değerinin yarısını ($e/2$) tekrar kazanır. Koloninin enerjisi besin konsantrasyonu ile doğrudan orantılıdır. Dolayısıyla daha enerjik olan alg kolonilerinin hareket frekansları daha yüksektir. Bu işlem kolonilerin global arama bir diğer tabirle keşif (exploration) yeteneklerini artırır. Su içerisinde hareketi etkileyen bir diğer faktör sürtünme kuvvetidir. Bu sürtünme kuvveti büyüklüğe bağlı olarak Denklem 3.8 ile hesaplanır. Yüzey alan daha küçük olan koloniler sürtünme yüzeyleri küçük olduğu için su içerisinde daha fazla hareket mesafesine sahiptir. Daha iyi ortam şartlarını elde eden koloniler büyüklüklerinden dolayı daha fazla sürtünme yüzey alanına sahiptirler. Bu

durum iyi bir çözüm yakalamış kolonilerin yerel arama bir diğer tabirle sömürü (explotation) yeteneklerini arttırarak buldukları çözümün etrafında yavaş yavaş arama yaparak çözümü geliştirmelerine daha fazla yardımcı olur. Bu şekilde her bir koloni enerjisi bitene kadar sıvı içerisinde helisel hareketler yaparak daha iyi ortam şartlarına ulaşmaya çalışır.

$$\tau(X_i) = 2\pi \left(\sqrt[3]{\frac{3G_i}{4\pi}} \right)^2 \quad (3.8)$$

Denklemden $\tau(X_i)$, i . alg kolonisinin sürtünme kuvvetini göstermektedir. Her bir alg kolonisi Şekil 3.1’de gösterildiği gibi üç boyutlu helisel hareket yapar.



Şekil 3.1. Alglerin üç boyuttaki helisel hareketi

Bu üç boyutun hangi koloni ile değiştirileceğini belirlerken turnuva seçim yöntemini kullanır. Koloni belirlendikten sonra rastgele seçilecek üç boyut (p , r ve v) belirlenir. Denklemler 3.9 - 3.11 kullanılarak helisel hareket gerçekleştirilir.

$$x_{ip}^{t+1} = x_{ip}^t + (x_{jp}^t - x_{ip}^t)(\Delta - \tau^t(X_i))p \quad (3.9)$$

$$x_{ir}^{t+1} = x_{ir}^t + (x_{jr}^t - x_{ir}^t)(\Delta - \tau^t(X_i)) \cos \alpha \quad (3.10)$$

$$x_{iv}^{t+1} = x_{iv}^t + (x_{jv}^t - x_{iv}^t)(\Delta - \tau^t(X_i)) \sin \theta \quad (3.11)$$

Denklemlerde x_{ip}^t , x_{ir}^t ve x_{iv}^t sırasıyla t zamanındaki i . alg kolonisinin x, y ve z koordinatlarını ve $\tau(X_i)$ sürtünme kuvvetini göstermektedir. p , r ve v birbirinden farklı rastgele seçilmiş üç boyutu, j ise rastgele seçilmiş komşuyu gösterir. X_j turnuva yöntemiyle seçilir ve X_i , X_j ye doğru helisel hareket yapar. $\alpha, \theta \in [0, 2\pi]$ ve $p \in [-1, 1]$ olmak üzere Δ kesme kuvvetini gösterir.

3.1.3. Adaptasyon Süreci

Adaptasyon sürecinde yeterince gelişemeyen alg kolonilerinin en büyük koloniye benzemeye, uyum sağlamaya çalıştığı bir süreçtir. Açlık değeri en yüksek olan alg kolonisi bu işleme tabi tutulur. İterasyon başlangıcında tüm alg kolonilerinin açlık değeri eşit ve sıfırdır. Helisel hareket sürecinde daha iyi bir çözüm bulamayan alg kolonilerin açlık değeri bir artar. Helisel hareket ve evrimsel süreçten sonra açlık değeri en yüksek koloni adaptasyon sürecine girer. Ancak adaptasyon işlemi her iterasyonda gerçekleşmez. Önce 0 ile 1 arasında rastgele bir değer seçilir. Bu değer Adaptasyon parametresinden (Ap) küçükse adaptasyon işlemi gerçekleştirilir. Adaptasyon işlemi Denklemler 3.12 ve 3.13 ile gerçekleştirilir.

$$\text{starving}^t = \max A_i^t \quad (3.12)$$

$$\text{starving}^{t+1} = \text{starving}^t + (\text{biggest}^t - \text{starving}^t) \times \text{rand} \quad (3.13)$$

Denklemlerde A_i^t , t zamanındaki i . alg kolonisinin açlık değerini göstermektedir. starving^t , t zamanındaki açlık değeri en yüksek alg kolonisini temsil etmektedir.

AAA'da alg popülasyonu rastgele başlangıç parametreleri oluşturulur. Adaptasyon parametresi (Ap), kesme kuvveti (Δ) ve enerji kaybı parametresi (e) tanımlanır. Alg popülasyonunu oluşturan her bir alg kolonisi enerjisi bitene kadar helisel hareket yapar. Daha sonra en küçük alg kolonisinin, en büyük alg kolonisine doğru hareket etmeye çalıştığı evrimsel süreç gerçekleşir. En son en aç alg kolonisi (Ap) parametresi doğrultusunda adaptasyon işlemine tabi tutulur ve yeni popülasyon elde edilir. En iyi çözüm her iterasyonda elde edilen çözümlerle karşılaştırılarak seçilip iterasyon boyunca saklanır.

AAA'nın sözde kodu Şekil 3.2'de gösterilmektedir (Uymaz ve ark., 2015a), (Uymaz ve ark., 2015b).

Uygunluk fonksiyonunun tanımlanması $f(x)$, $x=(x_1, x_2, \dots, x_d)$
 N alg kolonisine sahip başlangıç popülasyonunun başlatılması
 Alg kolonilerinin büyüklüklerinin (G) hesaplanması
 Kesme kuvveti parametresi (Δ), Enerji parametresi (e) ve Adaptasyon parametresi (A_p) tanımlanması
While ($t < \text{Maksimum Uygunluk Hesaplama Değeri}$)
 Alg kolonilerin enerjilerinin (E) ve sürtünme yüzeylerinin (τ) hesaplanması
 For $i=1: n$
 Starvation değerlerinin başlatılması
 While ($E(x_i) > 0$)
 Uygunluk fonksiyonlarını değerlerine göre turnuva yöntemi ile j alg kolonisin seçilimi
 Helisel hareket içim rastgele k, l ve m boyutlarının belirlenmesi

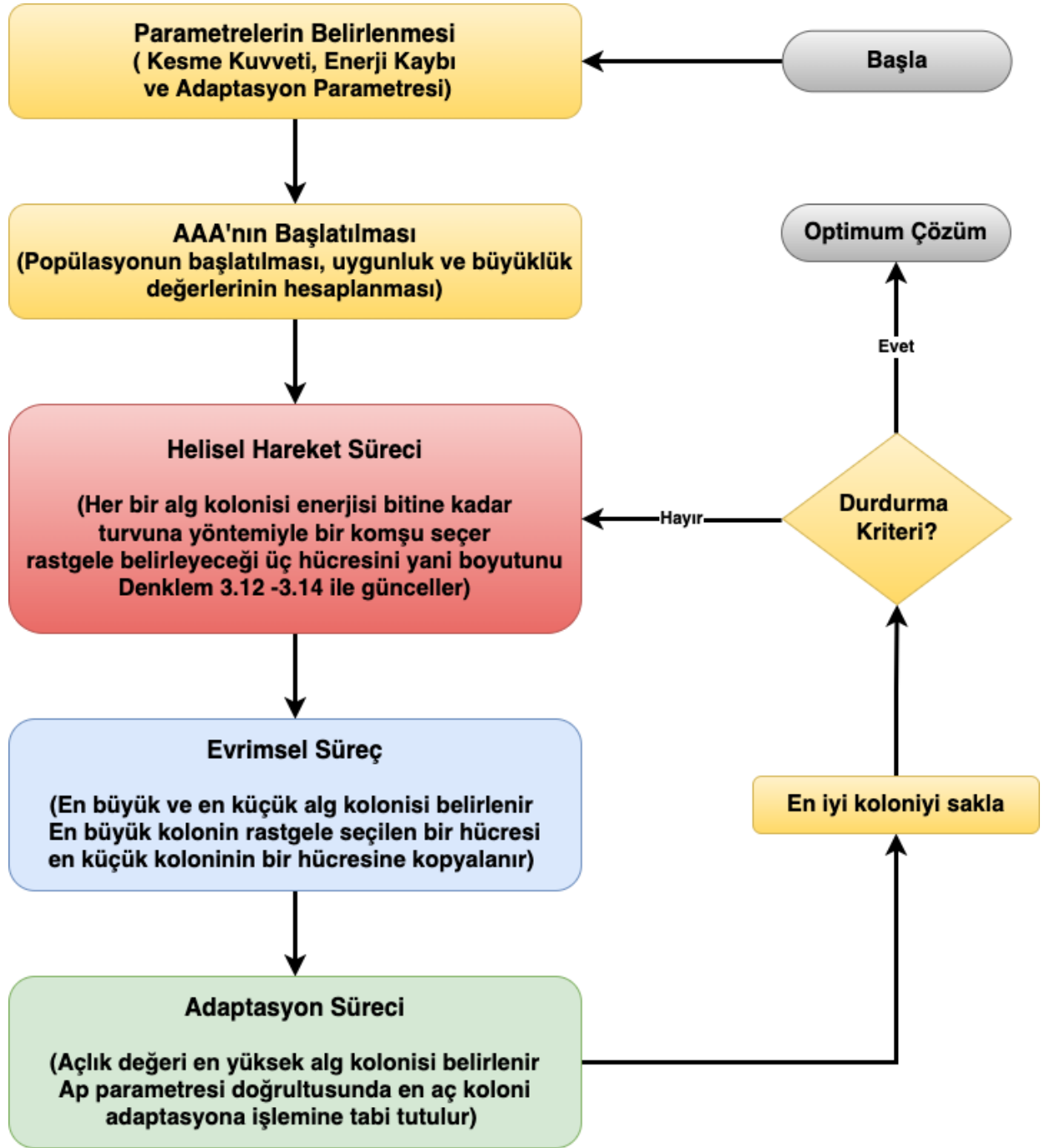
$$x_{ik}^{t+1} = x_{ik}^t + (x_{jk}^t - x_{ik}^t)(\Delta - \tau_i) \cos \alpha$$

$$x_{il}^{t+1} = x_{il}^t + (x_{jl}^t - x_{il}^t)(\Delta - \tau_i) \sin \beta$$

$$x_{im}^{t+1} = x_{im}^t + (x_{jm}^t - x_{im}^t)(\Delta - \tau_i) p$$
 $\alpha, \beta \in [0, 2\pi]; p \in [-1, 1];$
 Yeni çözümün hesaplanması
 $E(x_i) = E(x_i) - (\frac{e}{2})$ Enerji kaybı
 if yeni çözüm daha iyi i. Alg kolonisinin güncellenmesi ve starvation değerinin sıfırlanması
 else $E(x_i) = E(x_i) - (\frac{e}{2})$ enerji kaybı (daha iyi çözüm bulunamadığı için) **end if**
 end While
 if starvation seviyesinin artırılması **end if**
 end For
 Alg kolonilerinin büyüklüklerinin (G) hesaplanması
 Rastgele seçilen r boyutunda en büyük alg kolonisinin hücresinin en küçük alg kolonisine kopyalanması
 $\text{smallest}_r^t = \text{biggest}_r^t$
 En yüksek starvation değerine sahip koloninin belirlenmesi
 if $\text{rand} < A_p$
 $\text{starving}^{t+1} = \text{starving}^t + (\text{biggest}^t - \text{starving}^t) \times \text{rand}$
 end if
 En iyi çözümün saklanması
end While

Şekil 3.2. AAA'nın sözde kodu

AAA'nın akış şeması Şekil 3.3'te gösterilmektedir.



Şekil 3.3. AAA'nın akış şeması

3.2. Kaos Teorisi

Kaos teorisi kaotik, dinamik, doğrusal olmayan (non linear) sistemleri ifade edebilmek için geliştirilmiş bir teodir. Bu sistemler başlangıç parametrelerine yüksek hassasiyet gösterir. Yani başlangıç parametrelerindeki küçük kaotik geliřtirmeler sistemin çıkışında ve performansında büyük değışikliklere yol açabilir. Bununla birlikte

kaos sistemleri rastgelelik (randomness) içerir. Bir diğer ifadeyle kaotik sistemler rastgele davranışlar gösterebilir rastgele parametrelere ihtiyaç duymazlar. Bu sistemlerin bir diğer özelliği ergodik (ergodicity) olmasıdır. Kaosun ergodiklik özelliği belirli bir aralıkta tekrara düşmemeyi garanti eder. Bir diğer tabirle rastgele dizilerde birbirinin aynısı sayılar üretilebilir ancak kaos sistemlerde ergodiklik sayesinde aynı sayıdan tekrar üretilmesi mümkün değildir. Dolayısıyla, optimizasyon problemlerinde tekrara düşerek yerel minimumlara takılmaktan kaçınmak için kaos mekanizması kullanılabilir. Bu özellikleri sayesinde kaos sistemlerden esinlenerek geliştirilmiş kaotik haritalar kullanılabilir. Metasezgisel popülasyon tabanlı optimizasyon algoritmalarının performanslarını arttırmak için, bu haritalar sıklıkla kullanılırlar (Alatas ve ark., 2009; Mirjalili ve Gandomi, 2017).

3.2.1. Kaotik Haritalar

Kaotik haritalar uzun periyotlu rastgele sayılar üretmede sıklıkla kullanılan stokastik, deterministik ve doğrusal olmayan bir stratejidir. Bu haritalar tarafından oluşturulan sayı dizilerindeki sayılar tekrara düşmemesi, bir yerde sıkışmak yerine geniş bir spektruma yayılmış olması, dizi üretim ve saklama maliyetinin düşük olması gibi büyük avantajlara sahiptir. Metasezgisel algoritmaların içerisindeki rastgele üretilen sayıların aynı olması ve belirli bir aralıkta olması gibi nedenlerle bu rastgele sayılar bazen algoritmaların lokal minimumlarda sıkışmasına neden olabilmektedir. Bu gibi sebeplerden ötürü belirli bir sistematigi olan, spektrumun tamamını keşfedebilen ve tekrara düşmeyen kaotik diziler üretilebilir ve rastgele sayı kullanmak yerine bu diziler kullanılabilir. Literatürde Henon (dos Santos Coelho ve Mariani, 2009), Ikeda (Ouannas ve ark., 2019), Zaslavskii (dos Santos Coelho ve Pessôa, 2011), Intermittency (Elaskar ve Del Río, 2017), Liebovitch (Zhao ve Gao, 2021), Tinkerbell (dos Santos Coelho ve Mariani, 2012) gibi çok farklı kaotik harita çeşidi bulunmakla birlikte, çalışmalarda sıklıkla tercih edilen 10 kaotik harita bulunmaktadır. Bunlar Chebyshev (Tang ve ark., 2018), Circle (Jordehi, 2015), Gaussian (Kaur ve Arora, 2018), Iterative (Rather ve Bala, 2020), Logistic (Demir ve ark., 2020), Piecewise (Wang ve ark., 2014), Sine (Saremi ve ark., 2014), Singer (Gandomi ve ark., 2013b), Sinusoidal (Talatahari ve ark., 2012) ve Tent (Gandomi ve ark., 2013a) kaotik haritalarıdır. Kaotik diziler bu kaotik harita denklemlerinden bir başlangıç parametresi ile istenilen büyüklükte oluşturulur. İhtiyaç

duyulan pozisyonda rastgele sayı kullanmak yerine üretilen kaotik diziden sıradaki sayı çekilip kullanılır. Bu dizideki sayılar birbirinin aynı olma olasılığı söz konusu değildir ve geniş bir spektruma yayılır (May, 2004).

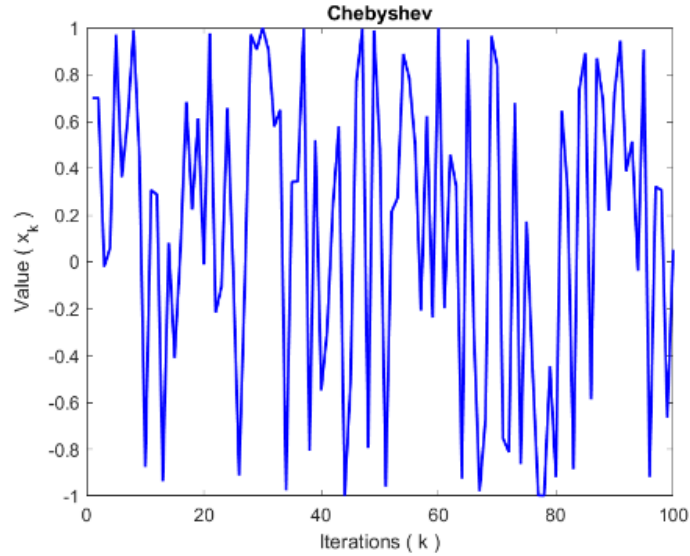
Tezin devamında literatürde yaygın olarak kullanılan kaotik haritalar ve formülasyonları sunulmuştur.

3.2.1.1. Chebyshev Kaotik Haritası

Chebyshev kaotik haritasına ait formülasyon Denklem 3.14'te gösterilmektedir. Denklemde k iterasyon sayısını, x_k k . kaotik sayıyı göstermektedir.

$$x_{k+1} = \cos(k \cos^{-1}(x_k)) \quad (3.14)$$

Chebyshev denklemi kullanılarak başlangıç parametresi $x_1 = 0.7$ ile üretilen 100 sayının spektrumu Şekil 3.4'de gösterilmektedir



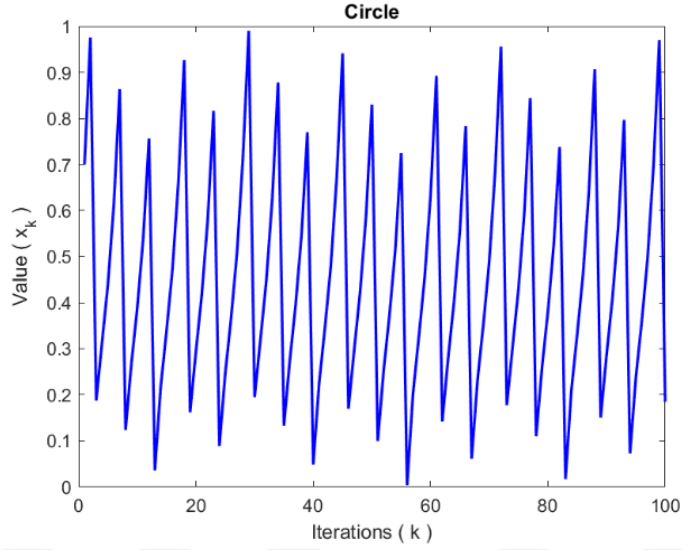
Şekil 3.4. Circle kaotik haritası yayılım spektrumu

3.2.1.2. Circle Kaotik Haritası

Circle kaotik haritasına ait formülasyon Denklem 3.15'te gösterilmektedir. Denklemde k iterasyon sayısını, x_k k . kaotik sayıyı, a ve b denklem sabitlerini temsil etmektedirler.

$$x_{k+1} = x_k + b - \left(\frac{a}{2\pi}\right) \sin(2\pi k) \bmod(1) \quad (3.15)$$

Circle kaotik haritası denklemi kullanılarak başlangıç parametresi $x_1 = 0.7$, denklem sabiti $a = 0.5$ ve $b = 0.2$ ile üretilen 100 sayının spektrumu Şekil 3.5'te gösterilmektedir.



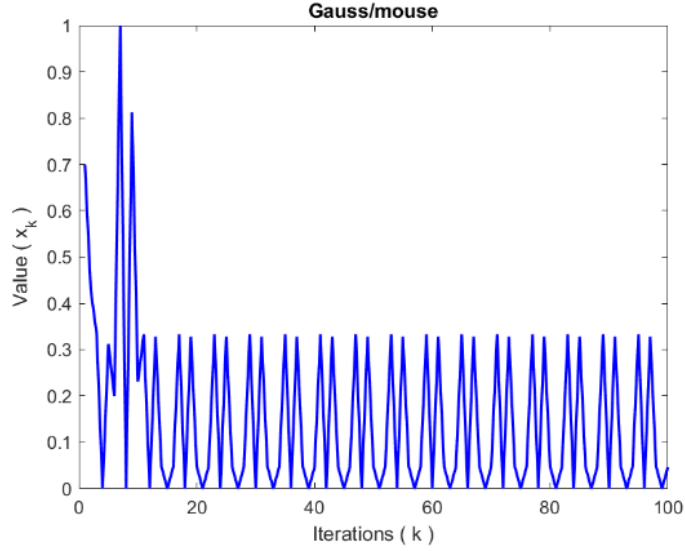
Şekil 3.5. Circle kaotik haritası yayılım spektrumu

3.2.1.3. Gaussian & Gauss/Mouse Kaotik Haritası

Gaussian kaotik haritasına ait Denklem 3.16 aşağıda gösterilmektedir. Denklemde k iterasyon sayısını, x_k k . kaotik sayıyı temsil etmektedir.

$$x_{k+1} = \begin{cases} 0 & x_k = 0, \\ \frac{1}{x_k \bmod(1)} & otherwise, \end{cases} \quad \frac{1}{x_k \bmod(1)} = \frac{1}{x_k} - \left\lfloor \frac{1}{x_k} \right\rfloor \quad x \geq 0 \quad (3.16)$$

Gaussuian map denklemi kullanılarak başlangıç parametresi $x_1 = 0.7$ ile üretilen 100 sayının spektrumu Şekil 3.6'te gösterilmektedir.



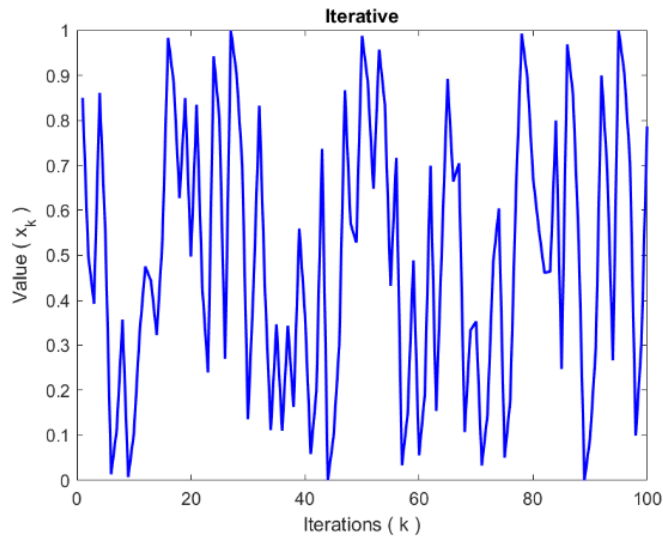
Şekil 3.6. Gaussian kaotik haritasının yayılım spektrumu

3.2.1.4. Iterative Kaotik Haritası

Iterative kaotik haritasına ait Denklem 3.17 aşağıda gösterilmektedir. Denklemde k iterasyon sayısını, x_k k . kaotik sayıyı, a denklem sabitini temsil etmektedir.

$$x_{k+1} = \sin\left(\frac{a\pi}{x_k}\right), a \in (0, 1) \quad (3.17)$$

Iterative kaotik haritası $[-1, 1]$ arasında değerler üretir, bu değerler $[0, 1]$ arasına normalize edilmektedir. Iterative kaotik haritası denklemleri kullanılarak başlangıç parametresi $x_1 = 0.7$ ve kaotik sabiti $a = 0.7$ ile üretilen 100 sayının spektrumu Şekil 3.7’te gösterilmektedir.



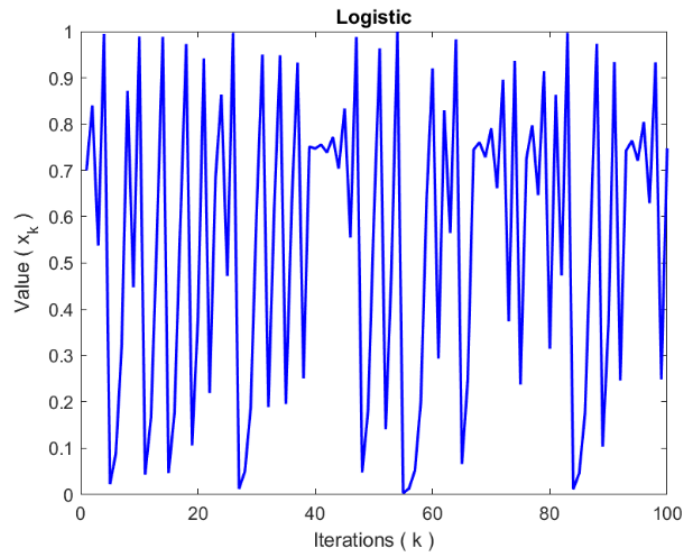
Şekil 3.7. Iterative kaotik haritası yayılım spektrumu

3.2.1.5. Logistic Kaotik Haritası

Logistic kaotik haritasına ait Denklem 3.18 aşağıda gösterilmektedir. Denklemde k iterasyon sayısını, x_k k . kaotik sayıyı göstermektedir. Denklemdeki a değeri lojistik harita sabitidir. Denklem kaotik sayı üretmesi için a sabitinin alabileceği değerler [3.57, 4] aralığındadır.

$$x_{k+1} = ax_k(1 - x_k) \quad (3.18)$$

Logistic kaotik haritası denklemini kullanılarak başlangıç parametresi $x_1 = 0.7$ ve kaotik sabiti $a = 4$ ile üretilen 100 sayının spektrumu Şekil 3.8’te gösterilmektedir.



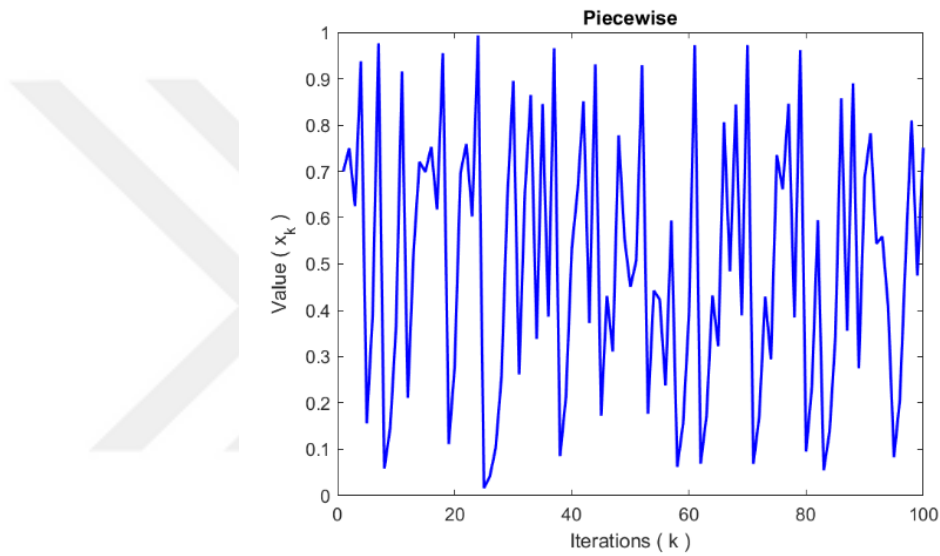
Şekil 3.8. Logistic kaotik haritası yayılım spektrumu

3.2.1.6. Piecewise Kaotik Haritası

Piecewise kaotik haritasına ait Denklem 3.19 aşağıda gösterilmektedir. Denklemlerdeki k iterasyon sayısını, x_k k . kaotik sayıyı, P denklem sabitini göstermektedir.

$$x_{k+1} = \begin{cases} \frac{x_k}{P}, & 0 \leq x_k < P \\ \frac{x_k - P}{0.5 - P}, & P \leq x_k < \frac{1}{2} \\ \frac{1 - P - x_k}{0.5 - P}, & \frac{1}{2} \leq x_k < 1 - P \\ \frac{1 - x_k}{P}, & 1 - P \leq x_k < 1 \end{cases} \quad P = 0.4 \quad (3.19)$$

Piecewise kaotik haritası denklemleri başlangıç parametresi $x_1 = 0.7$ ve denklemler sabiti $P = 4$ ile üretilen 100 sayının spektrumu Şekil 3.9'da gösterilmektedir.



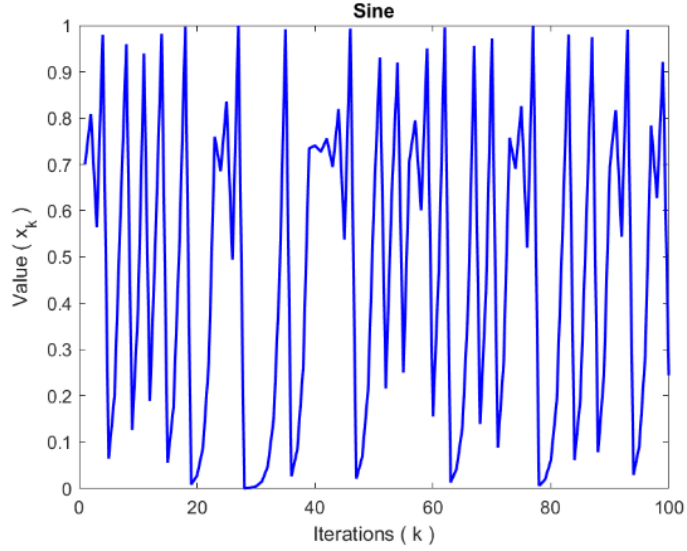
Şekil 3.9. Piecewise kaotik haritası yayılım spektrumu

3.2.1.7. Sine Kaotik Haritası

Sine kaotik haritasına ait Denklem 3.20 aşağıda gösterilmektedir. Denklemde k iterasyon sayısını, x_k k . kaotik sayıyı, a denklemler sabitini göstermektedir.

$$x_{k+1} = \frac{a}{4} \sin(\pi x_k), \quad 0 \leq a < 4 \quad (3.20)$$

Sine kaotik haritası denklemleri kullanılarak başlangıç parametresi $x_1 = 0.7$ ve denklemler sabiti $a = 4$ ile üretilen 100 sayının spektrumu Şekil 3.10'te gösterilmektedir.



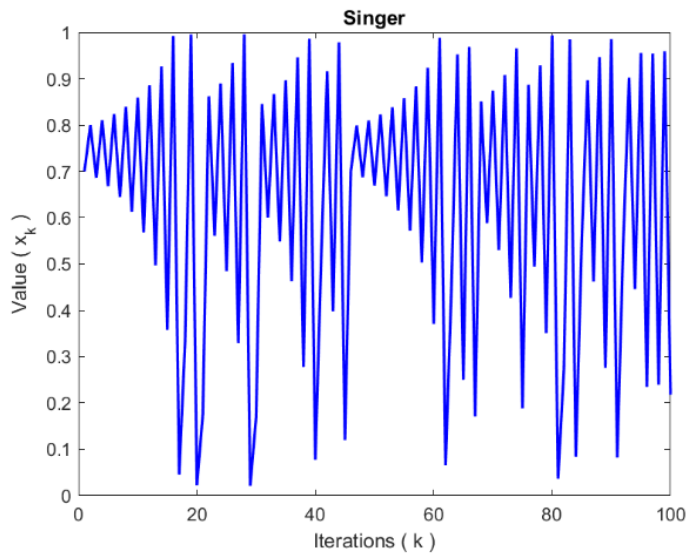
Şekil 3.10. Sine kaotik haritası yayılım spektrumu

3.2.1.8. Singer Kaotik Haritası

Singer kaotik haritasına ait Denklem 3.21 aşağıda gösterilmektedir. Denklemde k iterasyon sayısını, x_k k . kaotik sayıyı, μ denklem sabitini göstermektedir.

$$x_{k+1} = \mu(7.86x_k - 23.31x_k^2 + 28.75x_k^3 - 13.302875x_k^4) \quad (3.21)$$

Singer map denklemi kullanılarak başlangıç parametresi $x_1 = 0.7$ ve denklem sabiti $\mu = 1.07$ ile üretilen 100 sayının spektrumu Şekil 3.11’de gösterilmektedir.



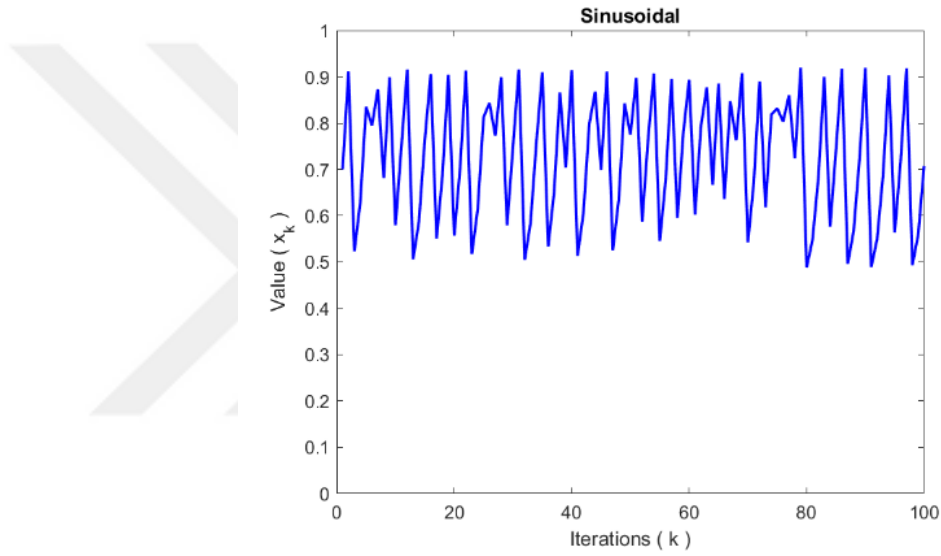
Şekil 3.11. Singer kaotik haritası yayılım spektrumu

3.2.1.9. Sinusoidal Kaotik Haritası

Sinusoidal kaotik haritasına ait Denklem 3.22 aşağıda gösterilmektedir. Denklemde k iterasyon sayısını, x_k k . kaotik sayıyı, a denklem sabitini göstermektedir.

$$x_{k+1} = ax_k^2 \sin(\pi x_k) \quad (3.22)$$

Sinusoidal map denklemleri kullanılarak başlangıç parametresi $x_1 = 0.7$ ve denklem sabiti $a = 2.3$ ile üretilen 100 sayının spektrumu Şekil 3.12’de gösterilmektedir.



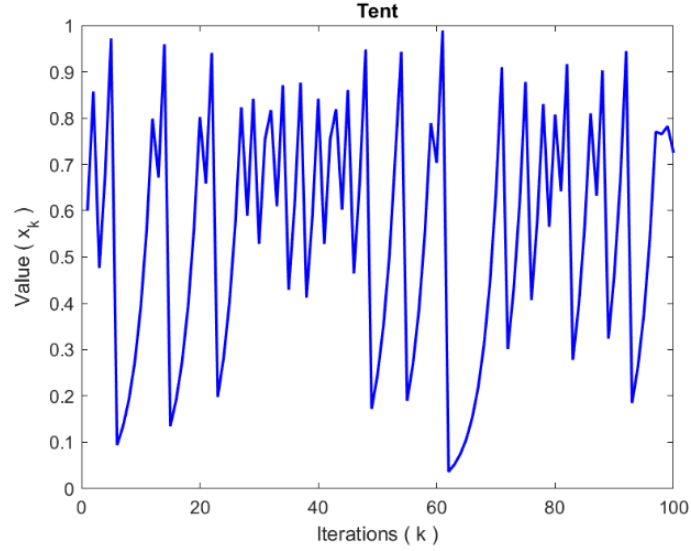
Şekil 3.12. Sinusoidal kaotik haritası yayılım spektrumu

3.2.1.10. Tent Kaotik Haritası

Tent kaotik haritasına ait Denklem 3.23 aşağıda gösterilmektedir. Denklemlerde k iterasyon sayısını, x_k k . kaotik sayıyı temsil etmektedir.

$$x_{k+1} = \begin{cases} \frac{x_k}{0.7}, & x_k < 0.7 \\ \frac{10}{3}(1 - x_k), & x_k \geq 0.7 \end{cases} \quad (3.23)$$

Tent map denklemleri kullanılarak başlangıç parametresi $x_1 = 0.6$ ile üretilen 100 sayının spektrumu Şekil 3.13’te gösterilmektedir.



Şekil 3.13. Tent kaotik haritası yayılım spektrumu

3.3. Kaotik Yapay Alg Algoritması

Metasezgisel optimizasyon yöntemleri başlangıçta rastgele olarak oluşturulan birden fazla çözümün, çözüm uzayında belirlenen davranış modeli çerçevesinde rastgele olarak arama gerçekleştirilerek optimum çözüme ulaşmaya çalışmaktadır. Yöntemlerin esinlendiği doğal davranış modelleri yönteme ait hiper parametreler sayesinde çözüm uzayının keşfi (exploration) ve mevcut çözümlerin iyileştirilmesi (exploitation) kabiliyetlerini dengeli bir şekilde kullanmayı amaçlamaktadır. Keşif ve sömürü kabiliyetleri birbirleri ile önemli bir değiş-tokuş (trade-off) içerisinde. Dolayısıyla, metasezgisel yöntemlerin ideal keşif ve sömürü dengesini kurması lokal minimum noktalarından kaçınarak optimum çözüme hızlı bir yakınsama açısından önemlidir.

Kaos, doğrusal olmayan sistemlerin düzensiz davranışlarıyla ilgilidir. Kaotik sistemler başlangıç koşullarına bağımlı rasgele davranışlar sergileyen non-lineer deterministik sistemlerdir. Bu özelliklerinden dolayı kaotik sistemler metasezgisel yaklaşımlarda uygulanabilir olmuştur. Kaotik yaklaşımların metasezgisel yöntemlerde kullanımı, başlangıç popülasyonunun oluşturulmasında, aday çözümlerin rastgele hareketlerinin gerçekleştirilmesinde, yöntemlerin hiper parametrelerinin belirlenmesinde veya bu modellerin hibrit kullanılması şeklinde ortaya çıkmaktadır. Literatürde kaotik yaklaşımlara sahip pek çok metasezgisel optimizasyon yöntemi mevcuttur.

Yapay Alg algoritmasının en kritik süreci adaptasyon sürecidir (Uymaz ve ark., 2015b). Bu süreç keşif ve sömürü arasındaki dengeyi sağlayan temel süreçtir. Adaptasyon

süreci, bir önceki bölümde tanıtılan on kaotik harita ile zenginleştirilmiştir. Rastgele girilen adaptasyon süreci yerine kaotik seçim operatörü ile girilen bir adaptasyon süreci geliştirilmiştir. Kaotik süreç aşağıdaki Denklem 3.24'te gösterilmektedir.

$$\begin{aligned} & \text{if } C(t) < A_p \text{ then} \\ & \quad \text{starving}^{t+1} = \text{starving}^t + (\text{biggest}^t - \text{starving}^t) \times \text{rand} \\ & \text{end if} \end{aligned} \quad (3.24)$$

Denklemlerde $C(t)$, t . iterasyondaki kaotik haritanın değerini, starving^t en yüksek açıklık seviyesine sahip alg kolonisini, biggest^t en gelişmiş alg kolonisi ve rand 0 ile 1 rastgele seçilmiş bir değeri temsil etmektedir. Kaotik Yapay Alg Algoritması'nın sözde kodu Şekil 3.14'te gösterilmektedir.

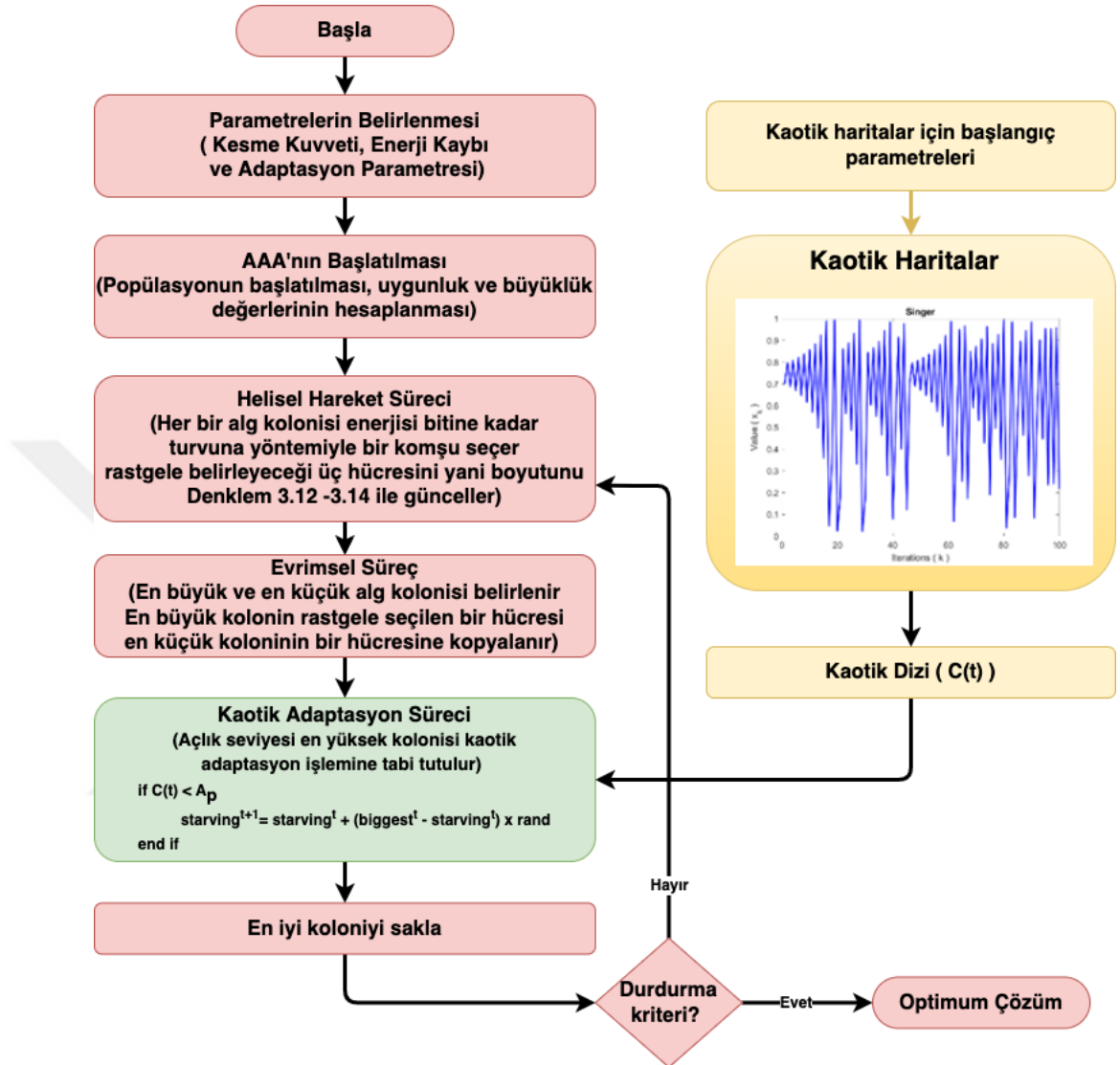
```

Uygunluk fonksiyonunun tanımlanması  $f(x)$ ,  $x=(x_1, x_2, \dots, x_d)$ 
N alg kolonisine sahip başlangıç popülasyonunun başlatılması
Alg kolonilerinin büyüklüklerinin (G) hesaplanması
Kesme kuvveti parametresi ( $\Delta$ ), Enerji parametresi ( $e$ ) ve Adaptasyon parametresi ( $A_p$ ) tanımlanması
While ( $t < \text{Maksimum Uygunluk Hesaplama Değeri}$ )
    Alg kolonilerin enerjilerinin ( $E$ ) ve sürtünme yüzeylerinin ( $\tau$ ) hesaplanması
    For  $i=1: n$ 
        Starvation değerlerinin başlatılması
        While ( $E(x_i) > 0$ )
            Uygunluk fonksiyonlarını değerlerine göre turnuva yöntemi ile  $j$  alg kolonisin seçilimi
            Helisel hareket içim rastgele  $k, l$  ve  $m$  boyutlarının belirlenmesi
             $x_{ik}^{t+1} = x_{ik}^t + (x_{jk}^t - x_{ik}^t)(\Delta - \tau_i) \cos \alpha$ 
             $x_{il}^{t+1} = x_{il}^t + (x_{jl}^t - x_{il}^t)(\Delta - \tau_i) \sin \beta$ 
             $x_{im}^{t+1} = x_{im}^t + (x_{jm}^t - x_{im}^t)(\Delta - \tau_i)p$ 
             $\alpha, \beta \in [0, 2\pi]$ ;  $p \in [-1, 1]$ ;
            Yeni çözümün hesaplanması
             $E(x_i) = E(x_i) - (\frac{e}{2})$  Enerji kaybı
            if yeni çözüm daha iyi,  $i$ . Alg kolonisinin güncellenmesi ve starvation değerinin sıfırlanması
            else  $E(x_i) = E(x_i) - (\frac{e}{2})$  enerji kaybı (daha iyi çözüm bulunamadığı için) end if
        end While
        if starvation seviyesinin artırılması end if
    end For
    Alg kolonilerinin büyüklüklerinin (G) hesaplanması
    Rastgele seçilen  $r$  boyutunda en büyük alg kolonisinin hücresinin en küçük alg kolonisine kopyalanması
     $\text{smallest}_r^t = \text{biggest}_r^t$ 
    En yüksek starvation değerine sahip koloninin belirlenmesi
    if  $\text{Chaotic Map Value}(T) < A_p$ 
         $\text{starving}^{t+1} = \text{starving}^t + (\text{biggest}^t - \text{starving}^t) \times \text{rand}$ 
    end if
    En iyi çözümün saklanması
end While

```

Şekil 3.14. Kaotik Yapay Alg Algoritmasının sözde kodu

Kaotik yapay alg algoritmasının akış diyagramı Şekil 3.15'te gösterilmektedir.

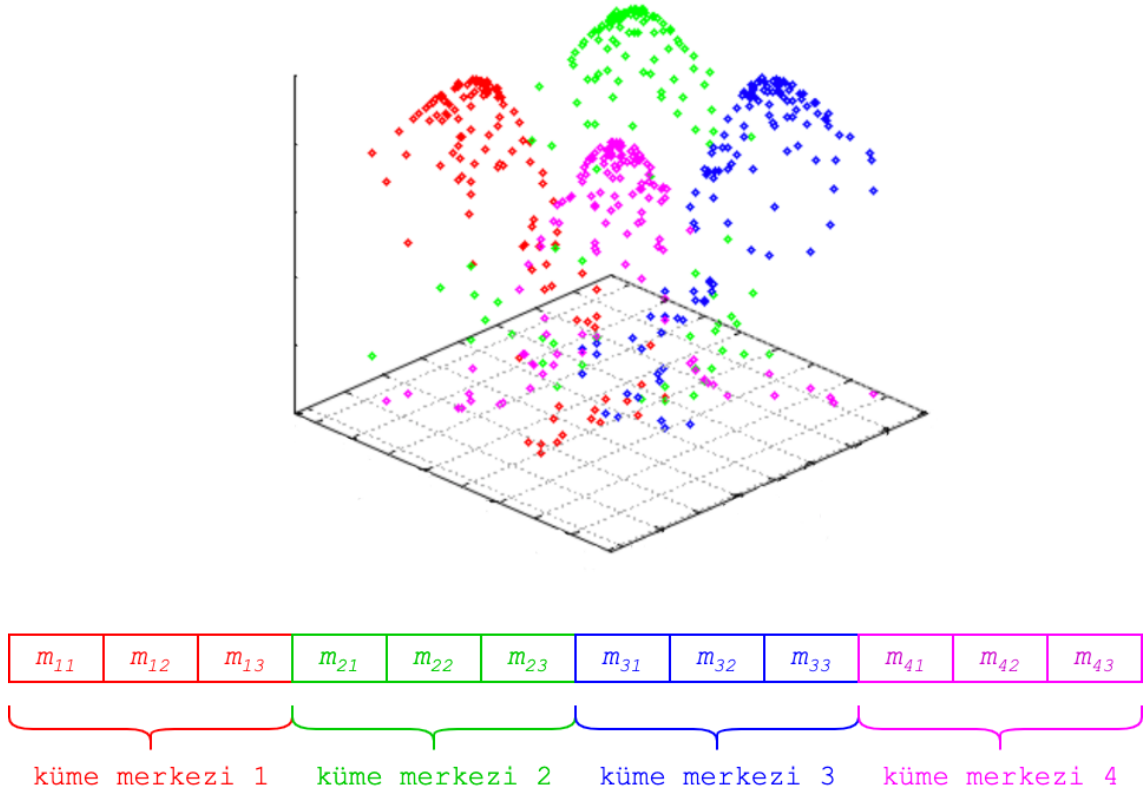


Şekil 3.15. Kaotik Yapay Alg Algoritması'nın akış diyagramı

3.4. Kümeleme Problemleri için Kaotik Yapay Alg Algoritması

AAA tabanlı kümeleme bir paylaşımlı kümeleme tekniğidir. N adet elemandan oluşan bir kümeyi k parçaya böler. Sistemin mantıksal açıdan anlamlı olabilmesi için küme sayısını temsil eden k değerinin eleman sayısını temsil eden N değerinden küçük olmalıdır. Paylaşımlı kümeleme yaklaşımlarında olduğu gibi başlangıçta k parametresinin değerinin belirlenmesi gerekir. AAA tabanlı kümeleme sürecinden benzerlik tabanlı nesnel bir bölümlendirme kriterini optimize etmeye çalışır.

Kümeleme sürecinde alg popülasyonundaki her bir ajan (alg kolonisi) k parametresi sayısında küme merkezlerinin çözüm kombinasyonunu oluşturan bir vektör olarak temsil edilmektedir. Örnek olarak Şekil 3.16’da dört kümenin merkezlerinin üç boyutlu uzayı ve çözüm vektörü gösterilmektedir.



Şekil 3.16. Üç boyutlu noktalara sahip dört kümenin uzaysal ve vektör olarak temsili

Alg kolonilerinin oluşturduğu çözüm vektörünün kalitesi uygunluk fonksiyonu ile belirlenir. Uygunluk fonksiyonu küme içi mesafenin toplamıdır. Bu bir minimizasyon problemidir yani küme içi mesafenin düşürülmesi amaçlanır. Denklem 3.25’te \mathbf{x}_i ile gösterilen her bir alg kolonisi bir aday çözümü temsil eder.

$$\mathbf{x}_i = (\mathbf{m}_{i1}, \mathbf{m}_{i2}, \dots, \mathbf{m}_{ik}, \dots, \mathbf{m}_{iNc}) \quad (3.25)$$

Denklemden \mathbf{m}_{ik} , i^{th} alg kolonisinin k^{th} merkez vektörünü, Nc küme sayısını temsil etmektedir.

Her bir alg kolonisinin uygunluğu Denklem 3.26'daki uygunluk fonksiyonu kullanılarak değerlendirilir. Bu değer veri vektörünün, küme merkezi ile arasındaki minimum mesafelerin toplamıdır.

$$f(x_i) = \sum_{j=1}^N \arg \min dist(x_i, I_j) \quad (3.26)$$

Denklemlerde

x_i : Aday çözümü oluşturan küme merkezlerini,

$f(x_i)$: x_i^{th} alg kolonisinin uygunluk değerini,

$dist(x_i, I_j)$: x_i ile I_j arasındaki Öklid mesafesini,

I_j : j^{th} veri noktasını,

N : veri kümesindeki örnek sayısını temsil etmektedir.

$\arg \min dist(x_i, I_j)$ Fonksiyonu Denklem 3.27'deki gibi her bir veri noktası ile veri noktasının dahil edildiği küme merkezlerinin minimum uzaklarının toplamıdır.

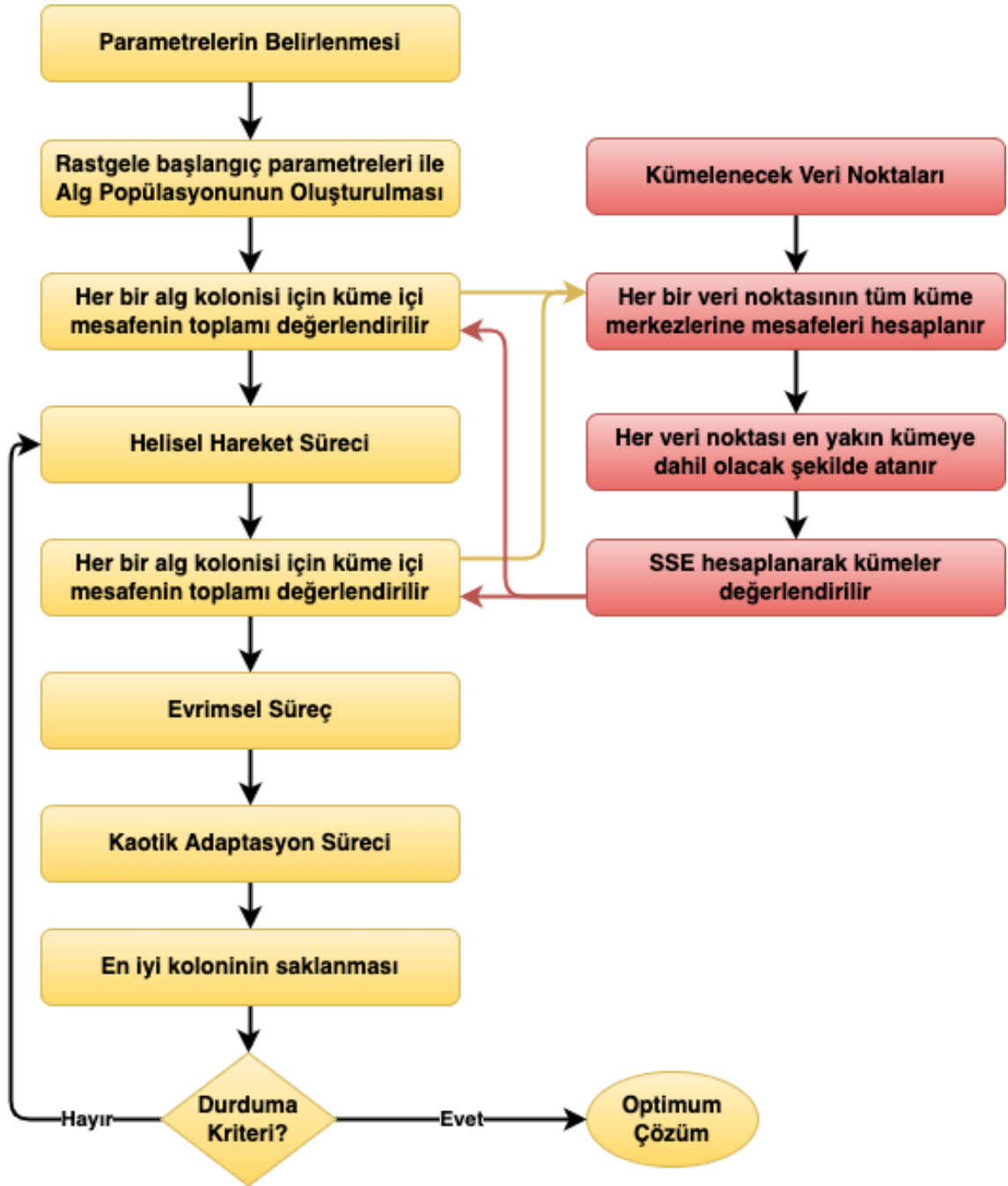
$$\arg \min dist(x_i, I_j) = \min [dist(m_{i1}, I_j), dist(m_{i2}, I_j), \dots, dist(m_{iNc}, I_j)] \quad (3.27)$$

$\arg \min dist(x_i, I_j)$ Fonksiyonu olarak Denklem 3.28'de gösterilen Karesel Hata Toplamı (SSE) kullanılmıştır.

$$Sum\ of\ Squared\ Error = \sum_{n=1}^N d_{nc}^2 \quad (3.28)$$

Denklemden d_{nc}^2 küme merkezi ile veri noktası arasındaki Öklid uzaklığını temsil etmektedir.

SSE değeri her noktanın kendisine en yakın küme merkezine olan uzaklığının karesi alınarak hesaplanan bir metriktir. Bu değer ne kadar küçük olursa hata bir o kadar az anlamına gelir ve başarı o ölçüde yüksek olur. Kaotik AAA tabanlı kümeleme yaklaşımı SSE metriğini bir uygunluk fonksiyonu olarak alır yinelemeler boyunca azaltmaya çalışır. Geliştirilen kaotik AAA tabanlı kümeleme yaklaşımının akış diyagramı Şekil 3.17'de gösterilmektedir.



Şekil 3.17. Kaotik Yapay Alg Algoritması tabanlı kümeleme yaklaşımının akış diyagramı

3.5. Öznitelik Seçim Problemi için Kaotik Yapay Alg Algoritması

Öznitelik seçimi problemlerinde çözüm N boyutlu bir vektör olarak temsil edilir. Burada N veri kümesindeki toplam öznitelik sayısıdır. Bu vektör 0 ve 1'lerden oluşur. Bu vektörün her bir parametresi ilgili özniteliğin yeni veri setine seçilip seçilmeyeceğine

karar verir. Seçilmesi durumunda 1 ile dahil edilmemesi durumunda 0 ile temsil edilir. Tüm olası çözüm öznitelik kombinasyonlarını oluşturma maliyeti 2^N kadardır. Bu sayısı öznitelik sayısını artışına göre üstel (exponential) olarak artacağı için olası tüm kombinasyonlara bakma tekniği olan kaba kuvvet (brute force) yaklaşımı çok maliyetli ve uygulanmaz hale gelir. Metasezgisel yaklaşımlar kaba kuvvet yaklaşımlarının uygulanamadığı yerlerde sıklıkla tercih edilirler. Bu yaklaşımlar kesin çözümü garanti etmemekle birlikte makul zamanda kabul edilebilir bir çözüm sunmaya çalışırlar. Metasezgisel yöntemler öznitelik seçimi için arama uzayında iyi performans gösteren yaklaşımlardan birisidir (Mafarja ve ark., 2017).

Orijinal AAA sürekli uzaydaki çeşitli optimizasyon problemlerini başarıyla çözebilen bir metasezgisel algoritmadır (Uymaz ve ark., 2015a). Tezin bu bölümünde öznitelik seçim problemi bir diğer tabirle ideal öznitelik alt kümesinin seçilmesi için geliştirilen ikili (binary) yapay alg algoritması açıklanacaktır. Önerilen yeni yaklaşımda parametre değerleri öznitelik seçimi problemine uygulanabilmesi için 0 ve 1 (ikili-binary) ile sınırlandırılmıştır. Önerilen yeni yaklaşımda da her bir alg kolonisi helisel hareket, evrimsel süreç ve adaptasyon süreciyle konumunu güncelleyerek çözümü geliştirmeye çalışır. İlk aşama olan helisel hareket sürecinde her bir alg kolonisi turnuva yöntemiyle seçeceği bir alg kolonisi (ışık kaynağı) ile helisel hareketiyle parametrelerini günceller. Diğer iki aşama olan Adaptasyon süreci ve evrimsel süreçte ise sırasıyla en küçük alg kolonisi ve açlık değeri en yüksek alg kolonisi uygun bir adım büyüklüğü ile küresel en iyiye doğru hareket eder. AAA ile aynı stratejiler AAA'nın ikili versiyonunda da uygulanır.

AAA'nın ikili uzayda çalışabilmesi için alglerin konum ve pozisyonlar değerlerinin sürekli değerler yerine ikili değerler formatında olması gerekmektedir. Sürekli uzaydaki değerleri ikili uzaya dönüştürmek için çeşitli stratejiler vardır. Bunlardan birisi transfer fonksiyonlarıdır. Transfer fonksiyonlarının amacı alg kolonilerini sürekli uzaydan ikili uzaya dönüştürmektedir. Diğer bir ifadeyle transfer fonksiyonları, koloninin konum vektöründeki alg hücrelerindeki sürekli değerleri ikili değerlere dönüştürür. Geliştirilen ikili algoritmada uygunluk fonksiyonu, kalitesine ve değerine bağlı olarak alg hücresinin değerinin 0'dan 1'e ya da 1'den 0'a dönüştürüleceğine karar verir (Kennedy ve Eberhart, 1997; Rashedi ve ark., 2010; Mirjalili ve Hashim, 2012; Arora ve Anand, 2019b).

Geliştirilen ikili kaotik yapay alg algoritması için S-şekilli (*S – shape*) ve V-şekilli (*V – shape*) olmak üzere iki tür transfer fonksiyonu grubu kullanılmıştır. Bunlardan dört tanesi S-şekilli (s_1, s_2, s_3, s_4) dört tanesi de V-şekilli (v_1, v_2, v_3, v_4) olmak üzere toplam sekiz tanedir.

S-şekilli transfer fonksiyonları literatürde yaygın olarak kullanılan transfer fonksiyonlarıdır. S-şekilli transfer fonksiyonlarının denklemleri ve detaylı özellikleri Çizelge 3.1’de sunulmaktadır (Mirjalili ve Lewis, 2013; Emary ve ark., 2016; Emine ve Ülker, 2020). S-şekilli transfer fonksiyonlarının denkleme bağlı olarak çıkışı bazen 0 ile 1 arasında sürekli değerlerdir. Bu sürekli değerleri ikili değerlere dönüştürmek için eşik fonksiyonlarının kullanılması gerekir. Eşik fonksiyonları, transfer fonksiyonların sürekli çıkış değerlerini ikili ayrık değerlere (0 ya da 1’e) dönüştürür. Transfer fonksiyonun gradyan değerindeki artış, çözüm vektörünün değerinin 0’dan 1’e ya da 1’den 0’a değişim olasılığı artırmaktadır.

Çizelge 3.1. S-şekilli transfer fonksiyonlarının denklemleri (Beşkirli ve ark., 2018)

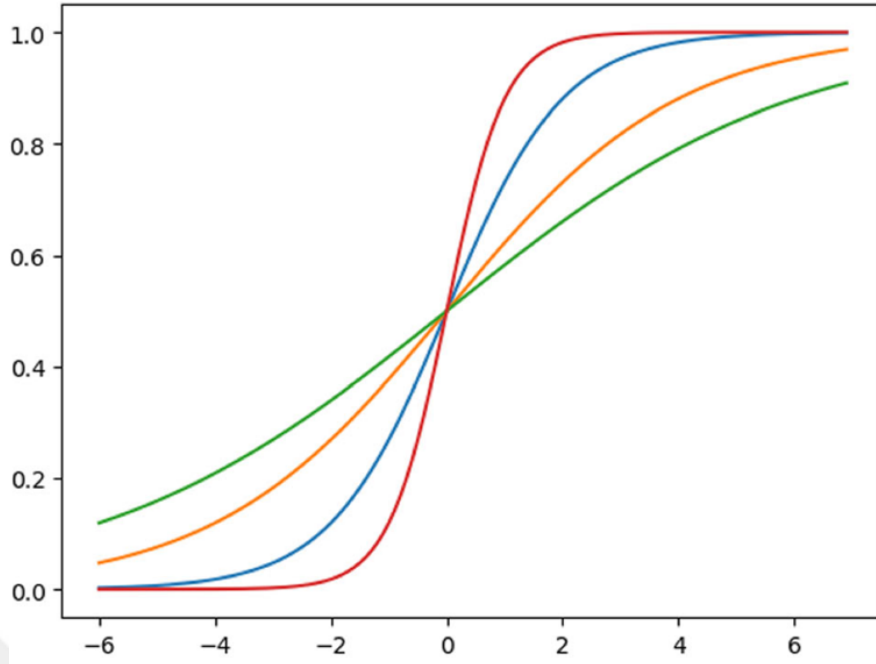
S-şekilli tranfer fonksiyonları	
Kısaltma	Transfer Fonksiyonu Denklemi
S1	$T(x) = \frac{1}{1 + e^{-2x}}$
S2	$T(x) = \frac{1}{1 + e^{-x}}$
S3	$T(x) = \frac{1}{1 + e^{(-x/2)}}$
S4	$T(x) = \frac{1}{1 + e^{(-x/3)}}$

Denklemler 3.30’da S-şekilli fonksiyonlarının dönüşümü için kullanılan eşik değeri fonksiyonu gösterilmektedir.

$$x_i^k(t+1) = \begin{cases} 0, & \text{if } rand < S(T_i^k(t)) \\ 1, & \text{if } rand \geq S(T_i^k(t)) \end{cases} \quad (3.30)$$

Denklemlerde $x_i^k(t)$ ve $T_i^k(t)$, i . alg kolonisinin t . iterasyonundaki k . boyutunu temsil etmektedirler.

S-şekilli transfer fonksiyonlarının denklemlerine ait grafikler denklemdeki aynı rengeyle Şekil 3.18’de gösterilmektedir.



Şekil 3.18. S Şekilli transfer fonksiyonları

V – *şekilli* yaklaşımlarda sürekli değerleri, ikili değerlere dönüştürmek için V – *şekilli* transfer fonksiyonları kullanılır. Alg kolonilerinin ikili arama uzayında hareket ederken kullandıkları transfer fonksiyonlarının denklemleri Çizelge 4.1’de ve grafikleri Şekil 3.13’te

Çizelge 3.2’de V – *şekilli* transfer fonksiyonlarına ait denklemler ve kısaltmaları sunulmaktadır (Mirjalili ve Lewis, 2013; Emary ve ark., 2016; Emine ve Ülker, 2020).

Çizelge 3.2. V-şekilli transfer fonksiyonlarının denklemleri (Emine ve Ülker, 2020)

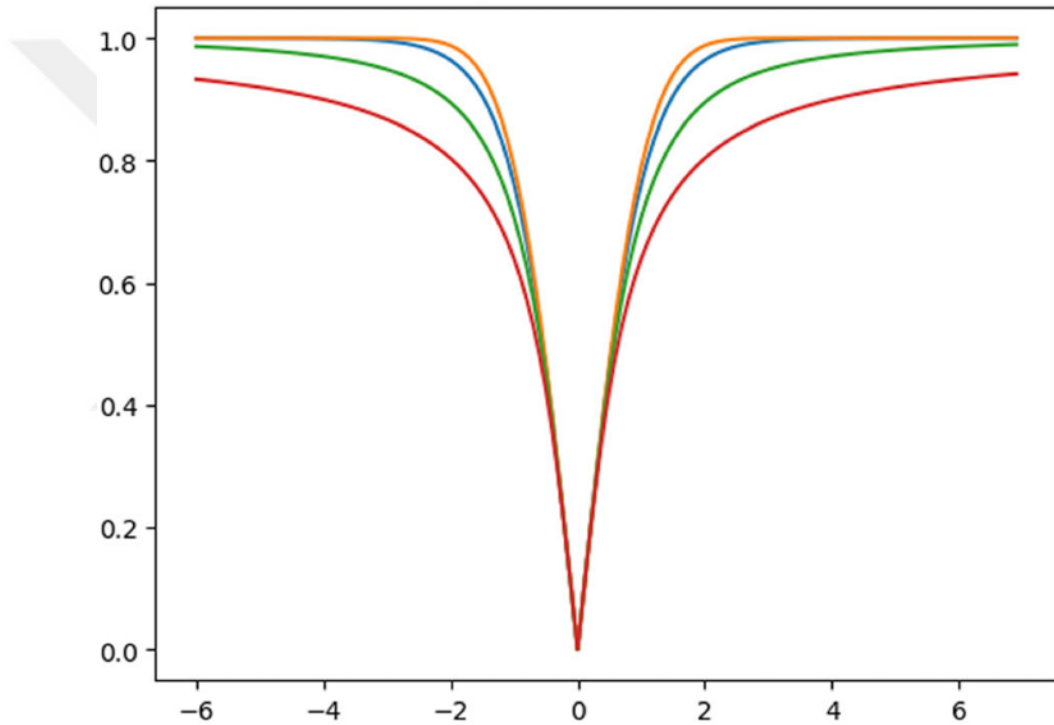
V-şekilli tranfer fonksiyonları	
Kısaltma	Transfer Fonksiyonu Denklemleri
V1	$T(x) = \left \operatorname{erf}\left(\frac{\sqrt{\pi}}{2} x\right) \right $
V2	$T(x) = \tanh(x) $
V3	$T(x) = \left \frac{x}{\sqrt{1+x^2}} \right $
V4	$T(x) = \left \frac{2}{\pi} \arctan\left(\frac{\pi}{2} x\right) \right $

V şekilli transfer fonksiyonlarının eşik değerinin formülasyonları Denklem 3.31’deki tanımlanmıştır.

$$x_i^k(t+1) = \begin{cases} (x_i^k(t))^{-1}, & \text{if } rand < V(T_i^k(t)) \\ x_i^k(t), & \text{if } rand \geq V(T_i^k(t)) \end{cases} \quad (3.31)$$

Denklemlerde $x_i^k(t)$ ve $T_i^k(t)$, i . alg kolonisinin t . iterasyonundaki k . boyutunun pozisyonunu ve $(x_i^k(t))^{-1}$, $x_i^k(t)$. complementini / tamamlayıcısını (0'sa 1'i, 1'se sıfırı) temsil etmektedirler.

V-şekilli transfer fonksiyonlarının denklemlerine ait grafikler denklemdaki aynı rengeyle Şekil 3.19'de gösterilmektedir (Rashedi ve ark., 2010; Mirjalili ve Lewis, 2013; Arora ve Anand, 2019b).



Şekil 3.19. V Şekilli transfer fonksiyonları

V şekilli yaklaşımlarda, alg kolonileri pozisyonları ve uygunluk değerlerini geliştirmek için kullandıkları çözüm vektörlerindeki değişim olasılarını Denklem 3.31'i kullanarak yaparlar. V şekilli fonksiyonların avantajları alg kolonilerinin pozisyonlarını tamamlayıcısını ile değişimini sağlamasıdır. Kısaca alg kolonisi besin yoğunluğunun yüksek olduğu yerlerde ikili değeri değiştirip güncelleyebilir. Ters durumda mevcut pozisyonunu koruma eğilimindedir.

İkili Kaotik Yapay Alg Algoritmasının sözde kodu Şekil 3.20’de gösterilmektedir.

```

Uygunluk fonksiyonunun tanımlanması  $f(x)$ ,  $x=(x_1, x_2, \dots, x_d)$ 
N ikili alg kolonisine sahip başlangıç popülasyonun başlatılması
Alg kolonilerinin büyüklüklerinin ( $G$ ) hesaplanması
Kesme kuvveti parametresi ( $\Delta$ ), Enerji parametresi ( $e$ ) ve Adaptasyon parametresi ( $A_p$ ) tanımlanması
While ( $t < \text{Maksimum Uygunluk Hesaplama Değeri}$ )
    Alg kolonilerin enerjilerinin ( $E$ ) ve sürtünme yüzeylerinin ( $\tau$ ) hesaplanması
    For  $i=1: n$ 
        Starvation değerlerinin başlatılması
        While ( $E(x_i) > 0$ )
            Uygunluk fonksiyonlarını değerlerine göre turnuva yöntemi ile  $j$  alg kolonisin seçilimi
            Helisel hareket içim rastgele  $k, l$  ve  $m$  boyutlarının belirlenmesi
            
$$x_{ik}^{t+1} = x_{ik}^t + (x_{jk}^t - x_{ik}^t)(\Delta - \tau_i) \cos \alpha$$

            
$$x_{il}^{t+1} = x_{il}^t + (x_{jl}^t - x_{il}^t)(\Delta - \tau_i) \sin \beta$$

            
$$x_{im}^{t+1} = x_{im}^t + (x_{jm}^t - x_{im}^t)(\Delta - \tau_i)p$$

             $\alpha, \beta \in [0, 2\pi]; p \in [-1, 1];$ 
            Çizelge 3.1-2’yi kullanarak transfer fonksiyonlarının değerlerinin hesaplanması
            Denklem (3.30- 31)’i kullanarak tranfer fonksiyonlarının çıkışının eşik fonksiyonlarından
            geçirilmesi
            Yeni çözümün hesaplanması
             $E(x_i) = E(x_i) - (\frac{e}{2})$  Enerji kaybı
            if yeni çözüm daha iyi,  $i$ . Alg kolonisinin güncellenmesi ve starvation değerinin sıfırlanması
            else  $E(x_i) = E(x_i) - (\frac{e}{2})$  enerji kaybı (daha iyi çözüm bulunamadığı için) end if
        end While
        if starvation seviyesinin arttırılması end if
    end For
    Alg kolonilerinin büyüklüklerinin ( $G$ ) hesaplanması
    Rastgele seçilen  $r$  boyutunda en büyük alg kolonisinin hücresinin en küçük alg kolonisine kopyalanması
     $\text{smallest}_r^t = \text{biggest}_r^t$ 
    En yüksek starvation değerine sahip koloninin belirlenmesi
    if  $C(t) < A_p$ 
         $\text{starving}^{t+1} = \text{starving}^t + (\text{biggest}^t - \text{starving}^t) \times \text{rand}$ 
        Çizelge 3.1-2’yi kullanarak transfer fonksiyonlarının değerlerinin hesaplanması
        Denklem (3.30- 31)’i kullanarak tranfer fonksiyonlarının çıkışının eşik fonksiyonlarından
        geçirilmesi
    end if
    En iyi çözümün saklanması
end While

```

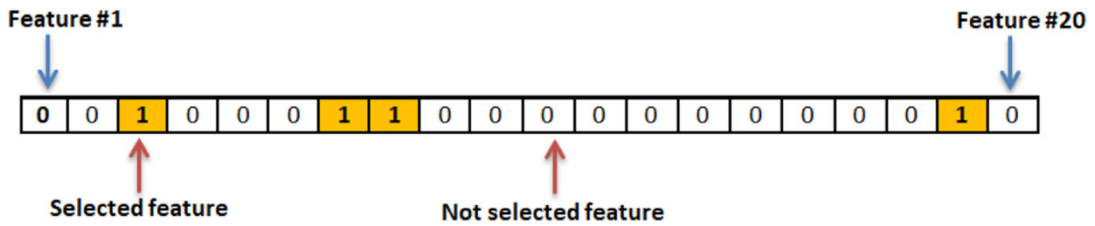
Şekil 3.20. İkili kaotik AAA sözde kodu

Temel olarak en ideal öznitelik alt kümesi seçilmesi sürecinde, her iterasyonda ikili optimizasyon yöntemi ile belirlenen öznitelik alt kümesinden oluşan veri seti oluşturulur. KNN, SVM gibi sınıflandırıcı bir algoritma oluşturulan veri setinin başarısını değerlendirir. İkili çözüm uzayında arama yapan ikili optimizasyon yönteminin temel amacı sınıflandırma başarısını maksimuma çıkaracak en az sayıda öznitelikten oluşan

öznitelik alt kümesini bulmaktır. Öznitelik kümesini belirlerken, ikili operatörle çalışan ikili algoritmalar çözüm uzayındaki her boyut için 0 ve 1'ler oluştuğu için sürekli optimizasyon için çalışan algoritmalarından daha başarılıdırlar.

3.5.1. Çözüm Vektörünün Tasarımı

Öznitelik seçimi problem yapısı gereği arama uzayında sadece ikili değerler alabildiği için problem ikili optimizasyon problemi olarak değerlendirilir. Dolayısıyla çözüm vektörünün her bir bileşeni 0 ya da 1 değerlerinden birisini alır. 1 değerine sahip vektör bileşenin öznitelik alt kümesine seçilip dahil edildiğine, 0 ise ilgili özneliğin kümeye dahil edilmediği anlamına gelir. Örnek bir öznitelik vektörü Şekil 3.21' de gösterilmektedir.



Şekil 3.21. Temsili bir ikili çözüm vektörü (Turabieh ve ark., 2019)

3.5.2. Uygunluk Fonksiyonu ve K-En Yakın Komşu Sınıflandırıcısı

İdeal öznitelik alt kümesinin belirlenmesi iki ayrı amacın karşılaması gereken çok amaçlı bir optimizasyon problemi olarak görülebilir. Bu amaçlardan birisi maksimum sınıflandırma doğruluğuna ulaşmak diğeri minimum öznitelik sayısını seçmektir. Uygunluk fonksiyonu minimum öznitelik sayısı ile maksimum sınıflandırma başarısını oluşturacak şekilde Denklem 3.32'deki gibi birleştirilebilir. Uygunluk fonksiyonunun minimum değerini sağlayacak şekilde oluşturulan öznitelik alt kümesi ideal öznitelik alt kümesidir. Sürüdeki her bir ajan yani her bir çözüm vektörü KNN sınıflandırıcısı kullanılarak uygunluk fonksiyonunda değerlendirilir. Buradaki asıl amaç, sınıflandırma doğruluğu ile öznitelik sayısı arasındaki dengeyi kurmaktır.

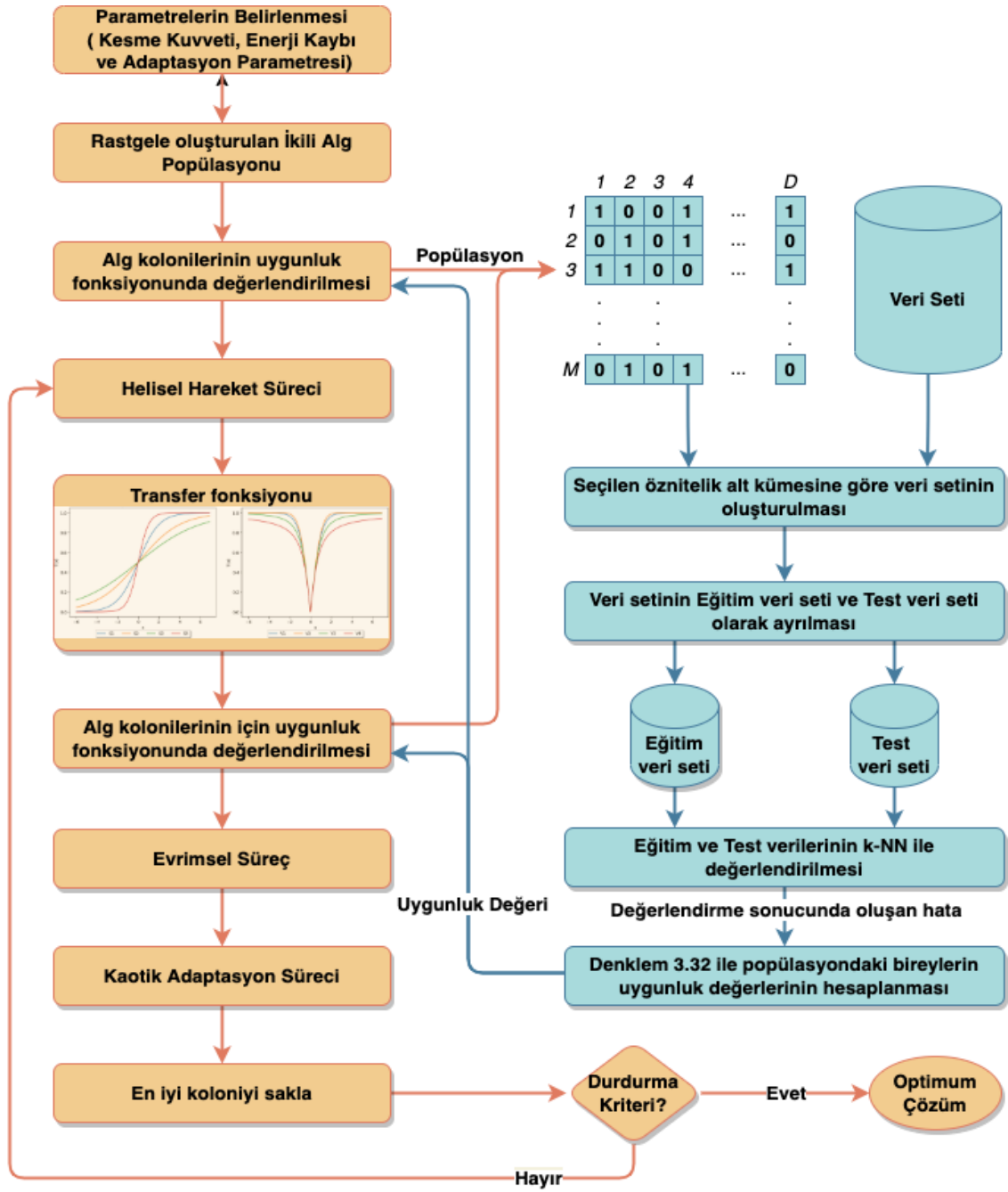
$$\downarrow \text{ Uygunluk Fonksiyonu} = \alpha \gamma_R(D) + \beta \frac{|M|}{|N|} \quad (3.32)$$

Denklemden $\gamma_R(D)$ sınıflandırma hata oranını, $|M|$ seçilen öznitelik sayısını, $|N|$ toplam öznitelik sayısını, α ve β sınıflandırma doğruluğunun ve seçilen öznitelik oranını arasındaki dengeyi kuran, $[0, 1]$ arasında değerler alan ve toplamı 1 olan iki parametreyi temsil etmektedir. Uygunluk fonksiyonu ve parametreler literatüre benzer şekilde seçilmiştir (Emary ve ark., 2016; Mafarja ve Mirjalili, 2017; Emine ve Ülker, 2020).

K- En Yakın Komşu (K-NN) algoritması öznitelik seçimden en sık kullanılan sınıflandırma algoritmasıdır K-NN algoritması genellikle Öklid mesafesini kullanılır. Öklid mesafesinin hesaplama formülü Denklem 3.33'te gösterilmektedir. Literatürdeki çalışmalarda olduğu gibi tez çalışmamızda da K değeri 5 olarak seçilmiştir (Pernkopf, 2005).

$$\text{öklid mesafesi } (x_a - x_b) = \left(\sum_{i=1}^N (x_{ai} - x_{bi})^2 \right)^{0.5} \quad (3.33)$$

Öznitelik seçimi problemini çözmek için geliştirilen kaotik ikili AAA yaklaşımının akış şeması Şekil 3.22'de gösterilmektedir.



Şekil 3.22. Öznitelik seçimi için geliştirilen ikili alg algoritmasının akış diyagramı

3.6. Algoritmaların Test Edilmesi için Kullanılan Optimizasyon Problemleri

Literatürde geliştirilen algoritmaların güvenilirliğini göstermek için farklı karakteristiğe sahip optimizasyon problemleri üzerinde test edilmesi gerekir. Literatürde global optimizasyon problemleri için kabul görmüş klasik benchmark fonksiyon setleri ve gerçek dünya mühendislik optimizasyon problemleri mevcuttur (Mirjalili ve ark., 2014a; Mirjalili ve Lewis, 2016; Mirjalili ve ark., 2017; Faramarzi ve ark., 2020a; Panagant ve ark., 2020; Abualigah ve ark., 2021; Zhao ve ark., 2022). Bu bölüm iki alt başlıktan oluşmaktadır. Birinci alt başlıkta tez kapsamında kullanılan klasik benchmark test fonksiyonları, ikinci alt başlıkta ise geliştirilen yöntem ile çözülen gerçek dünya problemleri detaylandırılmıştır.

3.6.1. Klasik Benchmark Fonksiyon Seti ve Değerlendirme Metrikleri

Geliştirilen yaklaşımın test edilmesi ve doğrulanması için farklı zorluk seviyesindeki otuz iyi bilinen benchmark fonksiyon seti kullanılmıştır. Bu fonksiyonlar tek modlu, (unimodal) çok modlu (multimodal) ve sabit boyutlu (fixed dimension) çok modlu olmak üzere üç farklı karakteristiğe sahiptir. Fonksiyonlardan on bir tanesi tek modlu (Sphere, Schwefel 2.22, Schwefel 1.2, Schwefel 2.21, Rosenbrock, Step, Quartic, Bent Cigar, Sum Power, Zakharov, Elliptic), dokuz tanesi çok modlu (Schwefel, Rastrigin, Ackley, Griewank, Penalized 1, Penalized 2, Levy, Weierstrass, Happy Cat) ve on tanesi sabit boyutlu (Foxholes, Kowalik, Six Hump Camel, Branin, Goldstein-Price, Hartman 3, Hartman 6, Shekel 5, Shekel 7, Shekel 10) çok modlu dur. Tek modlu fonksiyonlar tek bir global optimum noktası dolayısıyla tek bir yerel optimum noktası içerir. Bu fonksiyonlar algoritmanın yakınsama hızını test etmek için kullanılır. Çok modlu fonksiyonlar ise birden fazla yerel minimum ve tek bir global optimum noktası içerir. Bu fonksiyonlar algoritmanın yerel minimumlardan kaçma kabiliyetini test etmek için kullanılır. Diğer grup olan sabit boyutlu çok modlu fonksiyonlar da çok modlu fonksiyonlara benzer şekildedir. Çizelge. 3.3' de bu fonksiyonlardan oluşan benchmark fonksiyon setinin matematiksel ifadeleri gösterilmektedir. Çizelgede tek modlu fonksiyonlar “U” harfi ile, çok modlu fonksiyonlar “M” harfi ile, sabit boyutlu çok modlu fonksiyonlar ise “FD” harfleri ile temsil edilmektedir.

Çizelge 3.3. Deneysel çalışmada kullanılacak benchmark test fonksiyon seti

No.	Fonksiyon Adı	Fonksiyonun Denklemi	Arama Uzağı Aralığı	Boyutu	Minimum Noktası	Tipi
F1	Sphere	$F_1(x) = \sum_{i=1}^{Dim} x_i^2$	[-100,100]	30	0	U
F2	Schwefel 2.22	$F_2(x) = \sum_{i=1}^{Dim} x_i + \prod_{i=1}^{Dim} x_i $	[-10,10]	30	0	U
F3	Schwefel 1.2	$F_3(x) = \sum_{i=1}^{Dim} \left(\sum_{j=i}^i x_j \right)^2$	[-100,100]	30	0	U
F4	Schwefel 2.21	$F_4(x) = \max\{ x_i , 1 \leq i \leq Dim\}$	[-10,10]	30	0	U
F5	Rosenbrock	$F_5(x) = \sum_{i=1}^{Dim} (100(x_{i+1} - x_i^2) + (x_i - 1)^2)$	[-5,10]	30	0	U
F6	Step	$F_6(x) = \sum_{i=1}^{Dim} (x_i + 0.5)^2$	[-10,10]	30	0	U
F7	Quartic	$F_7(x) = \sum_{i=1}^{Dim} Dim \cdot x_i^2 + rand(0,1)$	[-1.28,1.28]	30	0	U
F8	Schwefel	$F_8(x) = \sum_{i=1}^{Dim} x_i \cdot \sin(\sqrt{ x_i }) $	[-100,100]	30	-418.98 $\times dim$	M
F9	Rastrigin	$F_9(x) = \sum_{i=1}^{Dim} [x_i^2 - 10 \cos(2\pi x_i) + 10]$	[-5.12,5.12]	30	0	M
F10	Ackley	$F_{10}(x) = -20 \cdot \exp \left(-0.2 \sqrt{\frac{1}{Dim} \sum_{i=1}^{Dim} x_i^2} \right) + \exp \left(\frac{1}{Dim} \sum_{i=1}^{Dim} \cos(2\pi x_i) \right) + 20 + \exp(1)$	[-50,50]	30	0	M
F11	Griewank	$F_{11}(x) = \frac{1}{4000} \sum_{i=1}^{Dim} x_i^2 - \prod_{i=1}^{Dim} \cos \left(\frac{x_i}{\sqrt{i}} \right) + 1$	[-600,600]	30	0	M
F12	Penalized 1	$F_{12}(x) = \frac{\pi}{Dim} \left\{ \sum_{i=1}^{Dim-1} (y_i - 1)^2 [1 + 10 \sin^2(\pi y_{i+1})] + (y_{Dim-1})^2 + 10 \sin^2(\pi y_1) \right\} + \sum_{i=1}^{Dim} u(x_i, 10, 100, 4)$ $y_i = 1 + \frac{(x_i + 1)}{4}$ $u_{y_i, a, k, m} = \begin{cases} k(x_i - a)^m, & x_i > a \\ 0, & -a \leq x_i \leq a \\ k(-x_i - a)^m, & x_i < a \end{cases}$	[-100,100]	30	0	M

F13	Penalized 2	$F_{13}(x) = \frac{1}{10} \left\{ \sin^2(\pi x_1) + \sum_{i=1}^{Dim-1} (x_i - 1)^2 [1 + \sin^2(3\pi x_{i+1})] + (x_{Dim} - 1)^2 (1 + \sin^2(2\pi x_{i+1})) \right\} + \sum_{i=1}^{Dim} u(x_i, 5, 100, 4)$	[-100,100]	30	0	M
F14	Foxholes	$F_{14}(x) = \left(\frac{1}{500} + \sum_{j=1}^{25} \frac{1}{j + \sum_{i=1}^2 (x_i - a_{ij})^6} \right)^{-1}$	[-65,65]	2	1	FD
F15	Kowalik	$F_{15}(x) = \sum_{i=1}^{11} \left[a_i - \frac{x_1(b_i^2 + b_i x_i)}{b_i^2 + b_i x_3 x_4} \right]^2$	[-5,5]	4	0.00030	FD
F16	Six Hump Camel	$F_{16}(x) = 4x_1^2 - 2.1x_1^4 + \frac{1}{3}x_1^6 + x_1x_2 - 4x_2^2 + 4x_2^4$	[-5,5]	2	-1.0316	FD
F17	Branin	$F_{17}(x) = \left(x_2 - \frac{5.1}{4\pi^2} x_1^2 + \frac{5}{\pi} x_1 - 6 \right)^2 + 10 \left(1 - \frac{1}{8\pi} \right) \cos x_1 + 10$	[-5,5]	2	0.398	FD
F18	Goldstein-Price	$F_{18}(x) = [1 + (x_1 + x_2 + 1)^2 (19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2)] * [30 + (2x_1 - 3x_2)^2 (18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2)]$	[-2,2]	2	3	FD
F19	Hartman 3	$F_{19}(x) = - \sum_{i=1}^4 c_i \exp \left(- \sum_{j=1}^3 a_{ij} (x_j - p_{ij})^2 \right)$	[1,3]	3	-3.86	FD
F20	Hartman 6	$F_{20}(x) = - \sum_{i=1}^4 c_i \exp \left(- \sum_{j=1}^6 a_{ij} (x_j - p_{ij})^2 \right)$	[0,1]	6	-3.32	FD
F21	Shekel 5	$F_{21}(x) = - \sum_{i=1}^5 [(X - a_i)(X - a_i)^T + c_i]^{-1}$	[0,10]	4	-10.1532	FD
F22	Shekel 7	$F_{22}(x) = - \sum_{i=1}^7 [(X - a_i)(X - a_i)^T + c_i]^{-1}$	[0,10]	4	-10.4028	FD
F23	Shekel 10	$F_{23}(x) = - \sum_{i=1}^{10} [(X - a_i)(X - a_i)^T + c_i]^{-1}$	[0,10]	4	-10.5363	FD
F24	Bent Cigar	$F_{24}(x) = x_1^2 + 10^6 \sum_{i=2}^{Dim} x_i^2$	[-100,100]	30	0	U
F25	Sum Power	$F_{25}(x) = \sum_{i=1}^{Dim} x_i ^{(i+1)}$	[-1,1]	30	0	U
F26	Zakharov	$F_{26}(x) = \sum_{i=1}^{Dim} x_i^2 + \left(\sum_{i=1}^{Dim} 0.5ix_i \right)^2 + \left(\sum_{i=1}^{Dim} 0.5ix_i \right)^4$	[-5,10]	30	0	U
F27	Levy	$F_{27}(x) = \sin^2(3\pi x_1) + \sum_{i=1}^{Dim-1} (x_i - 1)^2 [1 + \sin^2(3\pi x_{i+1})] + x_{Dim} - 1 \cdot [1 + \sin^2(2\pi x_{Dim})]$	[-10,10]	30	0	M

F28	Elliptic	$F_{28}(x) = \sum_{i=1}^{Dim} (10^6)^{(i-1)/(Dim-1)} \cdot x_i^2$	[-100,100]	30	0	U
F29	Weierstrass	$F_{29}(x) = \sum_{i=1}^{Dim} \left(\sum_{k=0}^{k_{max}} [a^k \cos(2\pi b^k (x_i + 0.5))] \right) - Dim$ $\cdot \sum_{k=0}^{k_{max}} [a^k \cos(2\pi b^k \cdot 0.5)], a = 0.5, b$ $= 3, k_{max} = 20$	[-1,1]	30	0	M
F30	Happy Cat	$F_{30}(x) = \left \sum_{i=1}^D x_i^2 - D \right ^{\frac{1}{4}}$ $+ 0.5 \left(\sum_{i=1}^D x_i^2 + \sum_{i=1}^D x_i \right) / D + 0.5$	[-100,100]	30	0	M

Algoritmaların performansını değerlendirmek için ortalama ve standart sapma metrikleri kullanılmıştır. Tez çalışmasında bu metriklerden ortalama “mean”, standart sapma “std” olarak kısaltılmıştır. Çizelge 3.4 ‘de ilgili metriklerin matematiksel ifadeleri gösterilmektedir.

Çizelge 3.4. Değerlendirme Metrikleri

Ortalama:	$Ortalama = \frac{1}{n_{run}} \sum_{i=1}^{n_{run}} Fitness_func_i$
Standart Sapma:	$Std.Dev. = \sqrt{\frac{1}{n_{run}} \sum_{i=1}^{n_{run}} (fitness_func_i - Ortalama)^2}$

Algoritmaların genel performans değerlendirmesini yapabilmek için literatürde çeşitli istatistiksel testler bulunmaktadır. Bunlardan birisi bağımlı gözlemlerde uygulanan, parametrik olmayan Friedman testidir (Zimmerman ve Zumbo, 1993). Friedman testi yapılırken sınıflandırma ya da kümeleme başarısı metriğinin ortalama değerleri kullanılır. Friedman testi ikiden fazla örneğin bulunduğu durumlarda uygulanır. Bu tez çalışmasında algoritmaların çeşitli veri setleri ve benchmark problemleri üzerindeki genel performansını değerlendirmek için kullanılmıştır. Karşılaştırma tablolarının en alt satırında Friedman testinden elde edilmiş ortalama sıralama değerleri verilmiştir.

Literatürdeki çalışmalarda kullanılan diğer bir istatistiksel test Wilcoxon işaretli sıralar testidir. Tezde geliştirilen algoritmaların literatürdeki yöntemler ile otuz bağımsız çalışma sonuçlarının doğrulanması için Wilcoxon işaretli sıralar testi yapılmıştır.

Wilcoxon işaretli sıralar testi kıyaslanacak iki algoritmanın, bir fonksiyondaki otuz bağımsız çalışma sonuçlarını değerlendirerek aralarında belirlenen güven seviyesinde anlamlı bir fark olup olmadığını belirlemekte kullanılan parametrik olmayan istatistiksel bir testtir. Bu tez kapsamındaki deneysel çalışmada da birçok çalışmada olduğu gibi Wilcoxon işaretli sıralar testinde güven seviyesi 0,05 olarak belirlenmiştir (Taheri ve Hesamian, 2013). Test sonucunda hesaplanan p-Value değeri 0,05’den büyük bir değer olduğunda sonuçlar arasında istatistiksel açıdan bir farklılık olmadığı anlamına gelmektedir. Ters durumda, p-Value değeri 0,05’ten küçükse sonuçlar istatistiksel açıdan anlamlıdır, yani kıyaslanan algoritmaların başarısı anlamlı olarak birbirinden farklıdır.

3.6.2. Gerçek Dünya Optimizasyon Problemleri

Tez çalışmasında geliştirilen yöntemin test edilmesi için klasik benchmark fonksiyonlarının yanı sıra üç mühendislik tasarım problemi ve sekiz uzay yörüngesi yerleştirme problemi kullanılmıştır. Mühendislik tasarım problemi olarak çalışmalarda en çok kullanılan basınçlı tank (pressure vessel) tasarım problemi, kaynaklı kiriş (welded beam) tasarım problemi ve germe sıkıştırma yay (tension/compression spring) tasarım problemleri tercih edilmiştir. Ayrıca Avrupa Uzay Ajansı’ndan alınan sekiz farklı uzay görevini içeren GTOP (Global Trajectory Optimisation) problem seti kullanılmıştır. Bahsi geçen problemler tezin devamında detaylandırılmıştır.

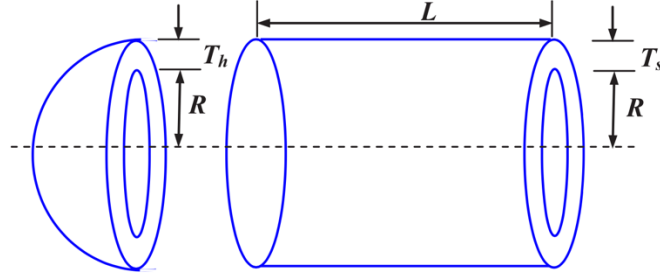
3.6.2.1. Mühendislik Tasarım Problemleri

Bu kısım tezde geliştirilen algoritmanın gerçek dünya problemlerinde test edilmesi için belirlenen popüler üç mühendislik tasarım probleminin tanıtıldığı üç alt başlıktan oluşmaktadır. Bu problemler basınçlı tank tasarım problemi, kaynaklı kiriş tasarım problemi ve germe sıkıştırma yay tasarım problemidir.

3.6.2.1.1. Basınçlı Tank Tasarımı Optimizasyon Problemi

Bu problemin temel amacı, belirli bir hacim ve sabit bir çalışma basıncını oluşturacak ve tasarım kısıtlamalarını karşılayacak şekilde basınçlı bir silindirik hava

tankının toplam maliyetini minimumda tutacak en uygun tasarımı parametrelerini tespit etmektir. Şekil 3.23'te gösterildiği gibi tankın her iki kenarı da kapalıyken başlık kısmı yarım küre şeklindedir (Hassan ve ark., 2014).



Şekil 3.23. Basıncılı tank tasarımı probleminin temsili görünümü (Faramarzi ve ark., 2020a)

Şekil de gösterildiği gibi dört adet tasarım değişkeninin optimize edilmesi gereklidir. İç yarıçap (R), kapağın kalınlığı (T_h), gövdenin kabuk kalınlığı (T_s) ve kapak hariç silindirik parçanın uzunluğu (L). Bu problemin matematiksel kısıtları Denklem 3.34'teki gibidir.

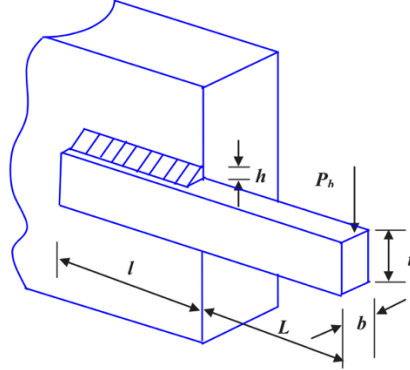
$$\begin{aligned}
 \text{Değişkenler} \quad \vec{x} &= [x_1, x_2, x_3, x_4] = [T_s, T_h, R, L] \\
 \text{Amaç Fonksiyonu} \quad f(\vec{x}) &= 0.6224x_1x_3x_4 + 0.7781x_2x_3^2 + 3.1661x_1^2x_4 + 19.84x_1^2x_3 \\
 \text{Kısıtlar} \quad g_1(\vec{x}) &= -x_1 + 0.0193x_3 \leq 0, \\
 g_2(\vec{x}) &= -x_3 + 0.00954x_3 \leq 0, \\
 g_3(\vec{x}) &= -\pi x_3^2x_4 - \frac{4}{3}\pi x_3^3 + 1296000 \leq 0, \\
 g_4(\vec{x}) &= x_4 - 240 \leq 0, \\
 \text{Değişkenlerin aralıkları} \quad &0 \leq x_1 \leq 99, 0 \leq x_2 \leq 99, 10 \leq x_3 \leq 200, \\
 &10 \leq x_4 \leq 200
 \end{aligned} \tag{3.34}$$

Amaç fonksiyonu tankın imalat maliyetini temsil etmektedir ve ne kadar küçük olursa o kadar verimlidir.

3.6.2.1.2. Kaynaklı Kiriş Tasarımı Optimizasyon Problemi

Bu problemin temel amacı, kaynaklı kirişin üretimini verilen kısıtlar altında en az maliyetli olacak şekilde optimum tasarımı bulmaktır. Şekil 3.24'te gösterildiği gibi dört

parametreyi yani; çubuğun bağlı kısmının uzunluğu (l), kaynak kalınlığı (h), çubuğun yüksekliği (t) ve çubuğun kalınlığının (b) optimum değerini tanımlayarak minimum üretim maliyetini tespit etmektir (Kamil ve ark., 2021).



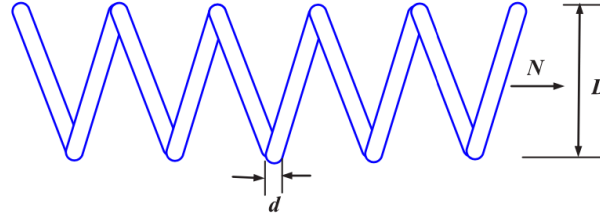
Şekil 3.24. Kaynaklı kiriş tasarım probleminin temsili görünümü (Faramarzi ve ark., 2020a)

Verilen değişkenlerin yedi kısıtlamayı sağlayacak şekilde tespit edilmesi gerekir. Bu problemin matematiksel ifadesi Denklem 3.35'te gösterilmektedir.

$$\begin{aligned}
 \text{Değişkenler} \quad \vec{x} &= [x_1, x_2, x_3, x_4] = [h, l, t, b] \\
 \text{Amaç fonksiyonu} \quad f(\vec{x}) &= 1.10471x_1^2x_2 + 0.04811x_3x_4(14.0 + x_2) \\
 \text{Kısıtlar} \quad g_1(\vec{x}) &= \tau(\vec{x}) - \tau_{max} \leq 0, \\
 g_2(\vec{x}) &= \sigma(\vec{x}) - \sigma_{max} \leq 0 \\
 g_3(\vec{x}) &= \delta(\vec{x}) - \delta_{max} \leq 0 \\
 g_4(\vec{x}) &= x_1 - x_4 \leq 0, \\
 g_5(\vec{x}) &= P - P_c(\vec{x}) \leq 0, \\
 g_6(\vec{x}) &= 0.125 - x_1 \leq 0, \\
 g_7(\vec{x}) &= 1.10471x_1^2 + 0.04811x_3x_4(14.0 + x_2) - 5.0 \leq 0, \\
 \text{Değişken aralıklıkları} \quad &0.1 \leq x_1 \leq 2, \quad 0.1 \leq x_2 \leq 10, \\
 &0.1 \leq x_3 \leq 10, \quad 0.1 \leq x_4 \leq 2
 \end{aligned} \tag{3.35}$$

3.6.2.1.3 Germe Sıkıştırma Yay Tasarım Problemi

Bu problemin temel amacı, belirlenen tasarım kısıtlamalarını karşılayacak şekilde çekme / sıkıştırma yayının minimum ağırlığının oluşturulması için üç tasarım parametresini belirlemektir. Örnek bir germe sıkıştırma yay tasarımı Şekil 3.24'te gösterilmektedir (Çelik ve Kutucu, 2018).



Şekil 3.24. Kaynaklı kiriş tasarım probleminin temsili görünümü (Faramarzi ve ark., 2020a)

Üç tasarım değişkeninin şunlardır: tel çapı (d), ortalama sarım çapı (D) ve aktif sarım sayısı (N). Bu problemin matematiksel ifadesi Denklem 3.36’te gösterilmektedir.

$$\begin{aligned}
 &\text{Değişkenler} \quad \vec{x} = [x_1, x_2, x_3] = [d, D, N] \\
 &\text{Amaç fonksiyonu} \quad f(\vec{x}) = (x_3 + 2)x_2x_1^2 \\
 &\text{Kısıtlar} \quad g_1(\vec{x}) = 1 - \frac{x_2^3x_3}{71785x_1^4} \leq 0, \\
 &\quad g_2(\vec{x}) = \frac{4x_2^2 - x_1x_2}{12566(x_2x_1^3 - x_1^4)} + \frac{1}{5108x_1^2} \leq 0, \\
 &\quad g_3(\vec{x}) = 1 - \frac{140.45x_1}{x_2^2x_3} \leq 0 \\
 &\quad g_4(\vec{x}) = 1 - \frac{x_1 + x_2}{1.5} - 1 \leq 0, \\
 &\text{Değişken Aralıkları} \quad 0.05 \leq x_1 \leq 2.00, \quad 0.25 \leq x_2 \leq 1.30, \\
 &\quad 2.00 \leq x_3 \leq 15.0,
 \end{aligned} \tag{3.36}$$

3.6.2.2. Yörünge Tasarım Problemleri

Küresel Yörünge Optimizasyon Problemleri Veri Tabanı olarak bilinen GTOP, Avrupa Uzay Ajansı (ESA)¹ tarafından yayınlanan uzay görevi tasarım problemlerini içeren bir veri tabanıdır. GTOP veri tabanı, sekiz farklı gezegenler arası uzay yörüngesi tasarım probleminden oluşmaktadır. Bu problemler doğrusal değildir ve çözülmesinin zor olduğu bilinmektedir (Schlueter ve ark., 2021).

GTOP’daki gezegenler arası uzay görevi yörüngelerinin tasarımı problemleri, havacılık biliminin yanı sıra sürü zekâsı ve evrimsel hesaplama araştırmacıları tarafından da aktif olarak çalışılan zorlu bir araştırma alanıdır. 2005 yılından bu yana ESA, GTOP

¹ <https://www.esa.int/gsp/ACT/projects/gtop/>.

veri tabanındaki sayısal kara kutu (black box) gerçek dünya uzay görevi yörünge tasarım optimizasyon problemleri üzerindeki çalışmalarına devam etmektedir. Optimizasyon problemlerinde kara kutu terimi, gerçek problem formülasyonunun bilinmediğini ifade etmektedir. Bu tür kara kutu problemleri, karmaşık gerçek dünya problemlerinin olduğu birçok alanda sıklıkla ortaya çıkmaktadır (Sala ve Müller, 2020). GTOP veri tabanının tüm problemleri tek amaçlı ve süreklidir. Tezin devamında GTOP veri tabanındaki problemler detaylandırılmıştır.

Cassini 1:

Cassini1 benchmark probleminde, Satürn'e yapılan yapılacak gezegenler arası bir uzay görevi modellenmiştir. Cassini uydusu, Satürn'ün yerçekimi tarafından pericenter (gökada enberi) yarıçapı 108.950 km ve dış merkezliği 0.98 olan bir yörüngeye girmektedir. Bu görevde gezegenlerin sırası Dünya–Venüs–Dünya–Jüpiter–Satürn olarak belirlenmiştir. İlk gezegen başlangıç gezegeni ve son gezegen ulaşılacak hedef gezegenidir. Bu problemin amaç fonksiyonu, fırlatma ve yakalama hızı da dahil olmak üzere görev sırasında oluşan toplam hız değişimini (ΔV) en aza indirmektir. Benchmark Çizelge 3.5 'de gösterilen altı karar değişkeni içermektedir.

Çizelge 3.5. Cassini1 için optimizasyon değişkenlerinin açıklaması

Değişken	Açıklaması
1	Başlangıç 1 Ocak 2000 baz alınarak hesaplanan gün sayısı
2-6	Olaylar arasındaki zaman aralığı (örn. Kalkış, geçiş, yakalama)

Bu problem ayrıca dört geçiş manevrası için dört kısıt içermektedir. Bu problemin en iyi bilinen çözümü, $f(x) = 4.9307$ amaç fonksiyonu değerine sahiptir (Schlueter ve ark., 2021).

Cassini 2:

Cassini2 benchmark, Derin Uzay Manevraları (DSM) dahil olmak üzere Satürn'e gerçekleştirilen gezegenler arası bir uzay görevini modellemektedir ve Cassini1'den çok daha zor bir görevdir. Bu görevde gezegenlerin sırası Dünya–Venüs–Venüs–Dünya–Jüpiter–Satürn olarak verilmektedir. İlk gezegen başlangıç gezegeni ve son gezegen ulaşılacak hedeftir. Bu problemin amacı, görev sırasında oluşan toplam ΔV 'yi en aza indirmektir; ancak burada amaç iki uzay aracını bir araya getirme işlemi iken, Cassini1'de

amaç bir yörüngeye girmektir. Cassini 2 problemi Çizelge 3.6 ‘de gösterilen yirmi iki karar değişkenini içermektedir.

Çizelge 3.6. Cassini2 için optimizasyon değişkenlerinin açıklaması

Değişken	Açıklaması
1	Başlangıç 1 Ocak 2000 baz alınarak hesaplanan gün sayısı
2	İlk aşırı hiperbolik hız (km/s)
3	Hiperbolik aşırı hız açıları (kutupsal koordinat çerçevesi)
4	Hiperbolik aşırı hız açıları (kutupsal koordinat çerçevesi)
5~9	Olaylar arasındaki zaman aralığı (örn. Kalkış, geçiş, yakalama)
10~14	DSM’nin gerçekleştiği zaman aralığının fraksiyonu
15~18	Geçiş yarıçapı (gezegende)
19~22	Gezegen yaklaşma vektörünün B düzleminde ölçülen açısı

Bu problem ayrıca dört geçiş manevrası için dört kısıt içermektedir. Bu problemin en iyi bilinen çözümü $f(x) = 8.3030$ amaç fonksiyonu değerine sahiptir (Schlueter ve ark., 2021).

Messenger (Reduced):

Bu problem, Merkür’e yapılan gezegenler arası bir uzay görevini modellemektedir. Bu görevde gezegenlerin sırası Dünya–Dünya–Venüs–Venüs–Merkür olarak verilmiştir. İlk gezegen başlangıç gezegeni ve son gezegen ulaşılacak hedeftir. Bu problemin amacı görev sırasında oluşan toplam hız değişimini ifade eden ΔV ’yi en aza indirmektir. Messenger (Reduced) problemi Çizelge 3.7 ‘de gösterilen on sekiz karar değişkenini içermektedir.

Çizelge 3.7. Messenger (reduced) için optimizasyon değişkenlerinin açıklaması

Değişken	Açıklaması
1	Başlangıç 1 Ocak 2000 baz alınarak hesaplanan gün sayısı
2	İlk aşırı hiperbolik hız (km/s)
3	Hiperbolik aşırı hız açıları (kutupsal koordinat çerçevesi)
4	Hiperbolik aşırı hız açıları (kutupsal koordinat çerçevesi)
5~8	Olaylar arasındaki zaman aralığı (örn. Kalkış, geçiş, yakalama)
9~12	DSM’nin gerçekleştiği zaman aralığının fraksiyonu
13~15	Geçiş yarıçapı (gezegende)
16~18	Gezegen yaklaşma vektörünün B düzleminde ölçülen açısı

Bu problemin en iyi bilinen çözümü, $f(x) = 8.6299$ amaç fonksiyonu değerine sahiptir (Schlueter ve ark., 2021)

Messenger (Full):

Bu problem, Merkür'e gezegenler arası bir uzay görevini modeller ve Messenger (Reduced)'dan farklı olarak Merkür'deki rezonant uçuşlarını da içerir. Bu görevde gezegenlerin sırası Dünya–Venüs–Venüs–Merkür–Merkür–Merkür–Merkür olarak verilmiştir. Bu problemin amacı görev sırasında oluşan toplam ΔV 'yi en aza indirmektir. Messenger (Full) problemi Çizelge 3.8 'de gösterilen yirmi altı karar değişkeni içermektedir.

Çizelge 3.8. Messenger (full) için optimizasyon değişkenlerinin açıklaması

Değişken	Açıklaması
1	Başlangıç 1 Ocak 2000 baz alınarak hesaplanan gün sayısı
2	İlk aşırı hiperbolik hız (km/s)
3	Hiperbolik aşırı hız açıları (kutupsal koordinat çerçevesi)
4	Hiperbolik aşırı hız açıları (kutupsal koordinat çerçevesi)
5~10	Olaylar arasındaki zaman aralığı (örn. Kalkış, geçiş, yakalama)
11~16	DSM'nin gerçekleştiği zaman aralığının fraksiyonu
17~21	Geçiş yarıçapı (gezegende)
12~26	Gezegen yaklaşma vektörünün B düzleminde ölçülen açısı

Bu problemin en iyi bilinen çözümü, $f(x) = 1.9579$ amaç fonksiyonu değerine sahiptir (Schlueter ve ark., 2021)

GTOC 1:

Bu problem, TW229 asteroidine gerçekleştirilen bir uzay görevini modellemektedir. Görev modeli, ESA tarafından 2007'de düzenlenen Küresel Yörünge Optimizasyon Yarışması'nın (GTOC) ilk versiyonundan esinlenmiştir (Izzo, 2007). Görevin amacı, asteroit yörüngesinin eksenindeki değişimi maksimize etmektir. Bu görevde gezegenlerin sırası Dünya–Venüs–Dünya–Venüs–Dünya–Jüpiter–Satürn–TW229 olarak verilmiştir. Bu problem Çizelge 3.9'da gösterilen sekiz karar değişkenini içermektedir.

Çizelge 3.9. GTOC1 için optimizasyon değişkenlerinin açıklaması

Değişken	Açıklaması
1	Başlangıç 1 Ocak 2000 baz alınarak hesaplanan gün sayısı
2~8	Olaylar arasındaki zaman aralığı (örn. kalkış, geçiş, yakalama)

Bu problem ayrıca dört geçiş manevrası için dört kısıt içermektedir. Bu problemin en iyi bilinen çözümü, $f(x) = -1581950$ amaç fonksiyonu değerine sahiptir (Schlueter ve ark., 2021)

Rosetta:

Rosetta problemi derin uzay manevraları da dahil olmak üzere 67P/Churyumov-Gerasimenko kuyruklu yıldızına gerçekleştirilen bir uzay görevini modellemektedir. Bu görevde gezegenlerin sırası Dünya–Dünya–Mars–Dünya–Dünya–67P olarak verilmiştir. İlk gezegen başlangıç gezegeni ve son gezegen ulaşılacak hedeftir. Bu problemin amacı görev sırasında oluşan toplam ΔV 'yi en aza indirmektir. Bu problem Çizelge 3.10'da gösterilen yirmi iki karar değişkenini içermektedir.

Çizelge 3.10. Rosetta için optimizasyon değişkenlerinin açıklaması

Değişken	Açıklaması
1	Başlangıç 1 Ocak 2000 baz alınarak hesaplanan gün sayısı
2	İlk aşırı hiperbolik hız (km/s)
3	Hiperbolik aşırı hız açıları (kutupsal koordinat çerçevesi)
4	Hiperbolik aşırı hız açıları (kutupsal koordinat çerçevesi)
5~9	Olaylar arasındaki zaman aralığı (örn. kalkış, geçiş, yakalama)
10~14	DSM'nin gerçekleştiği zaman aralığının fraksiyonu
15~18	Geçiş yarıçapı (gezegende)
19~22	Gezegen yaklaşma vektörünün B düzleminde ölçülen açısı

Bu problem ayrıca dört geçiş manevrası için dört kısıt içermektedir. Bu problemin en iyi bilinen çözümü, $f(x) = 1.3434$ amaç fonksiyonu değerine sahiptir (Schlueter ve ark., 2021).

Sagas:

Bu problem, Jüpiter'e gerçekleştirilen bir uzay görevini modellemektedir. Bu görevde gezegenlerin sırası Dünya–Dünya–Jüpiter olarak verilmiştir. İlk gezegen

başlangıç gezegeni ve son gezegen ulaşılacak hedeftir. Bu problemin amacı görev sırasında oluşan toplam ΔV 'yi en aza indirmektir. Bu problemde Çizelge 3.11'de gösterilen on iki karar değişkeni bulunmaktadır.

Çizelge 3.11. Sagas için optimizasyon değişkenlerinin açıklaması

Değişken	Açıklaması
1	Başlangıç 1 Ocak 2000 baz alınarak hesaplanan gün sayısı
2	İlk aşırı hiperbolik hız (km/s)
3	Hiperbolik aşırı hız açıları (kutupsal koordinat çerçevesi)
4	Hiperbolik aşırı hız açıları (kutupsal koordinat çerçevesi)
5~6	Olaylar arasındaki zaman aralığı (örn. kalkış, geçiş, yakalama)
7~8	DSM'nin gerçekleştiği zaman aralığının fraksiyonu
9~10	Geçiş yarıçapı (gezegende)
11~12	Gezegen yaklaşma vektörünün B düzleminde ölçülen açısı

Bu problemde ayrıca, yakıt ve fırlatıcı performansına limit getiren iki kısıt da bulunmaktadır. Bu problemin en iyi bilinen çözümü, $f(x) = 18.1877$ amaç fonksiyonu değerine sahiptir (Schlueter ve ark., 2021).

4. ARAŞTIRMA SONUÇLARI VE TARTIŞMA

Bu bölümde geliştirilen kaotik AAA'nın ve ikili versiyonunun performansını değerlendirmek için çeşitli optimizasyon problemleri çözülmüş ve aynı çalışma şartlarında literatürdeki popüler algoritmalarla kıyaslanmıştır. Bu bölüm dört alt başlıktan oluşmaktadır. İlk olarak geliştirilen yöntemin global optimizasyon benchmark testlerindeki performansı değerlendirilmiştir. İkinci alt başlıkta geliştirilen algoritma ve kıyaslanan yöntemlerin gerçek dünya mühendislik optimizasyon problemlerindeki başarımı yorumlanmıştır. Üçüncü alt başlıkta kümeleme analizi için organize edilen kaotik AAA'nın kümeleme veri setleri üzerindeki performansı, dördüncü alt başlıkta ise öznitelik seçimi probleminde geliştirilen ikili kaotik AAA'nın performansı analiz edilmiştir. Tüm deneysel çalışmaların çeşitli istatistiksel testleri yapılarak geliştirilen yöntemlerin performansı doğrulanmıştır.

4.1. Global Optimizasyon Problemlerinde Kaotik Yapay Alg Algoritmasının Performansı

Bu bölümde yapılan deneysel çalışma iki kısımdan oluşmaktadır. Birinci kısımda AAA'nın on farklı kaotik harita (Chebyshev, Circle, Gaussian, Iterative, Logistic, Piecewise, Sine, Singer, Sinusoidal ve Tent) ile donatılmasıyla geliştirilmiş on farklı kaotik AAA versiyonlarının (CAAAv1, CAAAv2, CAAAv3, CAAAv4, CAAAv5, CAAAv6, CAAAv7, CAAAv8, CAAAv9 ve CAAAv10) analizlerini içermektedir. Bu kısımda otuz benchmark fonksiyonu (Sphere, Schwefel 2.22, Schwefel 1.2, Schwefel 2.21, Rosenbrock, Step, Quartic, Bent Cigar, Sum Power, Zakharov, Elliptic, Schwefel, Rastrigin, Ackley, Griewank, Penalized 1, Penalized 2, Levy, Weierstrass, Happy Cat, Foxholes, Kowalik, Six Hump Camel, Branin, Goldstein-Price, Hartman 3, Hartman 6, Shekel 5, Shekel 7, Shekel 10) üzerinde kaotik AAA varyasyonlarının performansı karşılaştırılmıştır. İkinci kısımda ise birinci kısımda yapılan kıyaslama sonucunda elde edilen, en performanslı kaotik AAA varyasyonu literatürdeki iyi bilinen on algoritma (SSA, PSO, FFA, GWO, WOA, MVO, MFO, CS, SCA, JAYA) ve AAA ile kıyaslanmıştır. Yapılan kıyaslama işlemi algoritmaların elde ettiği en iyi sonuç sayısı (Best count) ve Friedman testi değerleri üzerinden analiz edilmiştir. Önerilen kaotik AAA varyasyonunun sonuçları on farklı algoritmaların her biri ile Wilcoxon istatistiksel testi

yapılmış ve sonuçları detaylı olarak analiz edilmiştir. Ayrıca bu bölümde önerilen yaklaşımın ve diğer algoritmaların benchmark fonksiyonları üzerindeki yakınsama eğrileri (convergence curves) verilerek optimum çözüme yakınsama karakteristikleri incelenmiştir.

4.1.1. Kaotik Versiyonlarının Performans Analizi

Bu bölümde, Bölüm 3.1 'de izah edilen kaotik haritalarla donatılan AAA versiyonlarının performansı analizi edilmiştir.

Otuz iyi bilinen benchmark fonksiyonu (Sphere, Schwefel 2.22, Schwefel 1.2, Schwefel 2.21, Rosenbrock, Step, Quartic, Bent Cigar, Sum Power, Zakharov, Elliptic, Schwefel, Rastrigin, Ackley, Griewank, Penalized 1, Penalized 2, Levy, Weierstrass, Happy Cat, Foxholes, Kowalik, Six Hump Camel, Branin, Goldstein-Price, Hartman 3, Hartman 6, Shekel 5, Shekel 7, Shekel 10) üzerinde on kaotik harita (Chebyshev, Circle, Gaussian, Iterative, Logistic, Piecewise, Sine, Singer, Sinusoidal ve Tent) kullanılarak oluşturulan varyasyonların (CAAAv1, CAAAv2, CAAAv3, CAAAv4, CAAAv5, CAAAv6, CAAAv7, CAAAv8, CAAAv9 ve CAAAv10) deneysel çalışma sonuçları Çizelge 4.1'de verilmiştir.

Çizelge 4.1'de varyasyonların benchmark fonksiyonlarda elde ettiği otuz bağımsız çalışmanın ortalaması ve standart sapması sunulmuştur. Her bir fonksiyonda elde edilen en iyi değer koyu stil ile işaretlenmiştir (F1 fonksiyonunda CAAAv3'te olduğu gibi). Çizelge 4.1'in son kısmında elde edilen sonuçların bir özeti verilmektedir. Özet kısmı, elde her bir varyasyonun kaç fonksiyonda en iyi değeri elde ettiği bilgisini ifade eden "Best Count", Friedman sıralama (ranking) değerleri, Friedman sıralama değerine bağlı olarak toplam sıralama (overall ranking) ve Friedman sıralama testinden elde edilen p-Value değerini içermektedir.

Çizelge 4.1'deki sonuçlar analiz edildiğinde CAAAv3 varyasyonu otuz fonksiyondan yirmi tanesinde diğer bir ifadeyle yaklaşık %67'sinde en iyi değeri elde ettiği görülmektedir. Yapılan Friedman sıralama testi sonucunda elde edilen p-Value değeri 0.05 den küçük olduğu için elde edilen sonuçlar arasında istatistiksel açıdan anlamlı bir fark bulunmaktadır. Friedman sıralama testi sonuçları incelendiğinde CAAAv3 varyasyonu 3.05 değeri ile en iyi sonucu elde etmiştir.

Çizelge 4.1. Kaotik Yapay Alg Algoritması versiyonlarının benchmark test fonksiyonlarındaki sonuçları

No.		CAAAv1	CAAAv2	CAAAv3	CAAAv4	CAAAv5	CAAAv6	CAAAv7	CAAAv8	CAAAv9	CAAAv10
		Chebyshev	Circle	Gauss	İterative	Logistic	Piecewise	Sine	Singer	Sinudial	Tent
F1	mean	3.37E-05	3.45E-05	1.49E-05	3.46E-05	7.15E-05	4.20E-05	4.11E-05	6.86E-05	2.07E-04	4.16E-05
	std	2.19E-05	1.73E-05	1.04E-05	1.70E-05	3.33E-05	2.40E-05	2.48E-05	3.28E-05	1.06E-04	1.93E-05
F2	mean	2.73E-04	2.97E-04	2.24E-04	2.88E-04	3.96E-04	2.83E-04	2.62E-04	3.44E-04	6.39E-04	3.23E-04
	std	9.42E-05	9.52E-05	8.18E-05	8.83E-05	9.61E-05	1.02E-04	6.19E-05	7.45E-05	1.81E-04	7.52E-05
F3	mean	3.80E+03	4.02E+03	3.34E+03	4.97E+03	6.11E+03	4.68E+03	3.78E+03	5.49E+03	1.06E+04	5.15E+03
	std	1.41E+03	1.16E+03	1.35E+03	2.14E+03	1.85E+03	1.32E+03	1.23E+03	1.79E+03	2.92E+03	1.92E+03
F4	mean	1.33E+01	1.40E+01	1.33E+01	1.41E+01	1.50E+01	1.45E+01	1.35E+01	1.46E+01	1.91E+01	1.45E+01
	std	2.35E+0	2.24E+00	2.64E+00	2.36E+00	2.01E+00	1.99E+00	3.05E+00	1.79E+00	2.52E+00	1.69E+00
F5	mean	9.13E+01	8.89E+01	7.97E+01	8.96E+01	1.04E+02	8.16E+01	9.88E+01	9.39E+01	1.08E+02	8.25E+01
	std	2.70E+01	3.10E+01	3.55E+01	3.30E+01	2.35E+01	2.99E+01	2.59E+01	3.30E+01	3.20E+01	2.96E+01
F6	mean	2.82E-05	2.63E-05	1.56E-05	4.28E-05	8.63E-05	3.62E-05	3.78E-05	5.92E-05	2.41E-04	4.93E-05
	std	1.27E-05	1.75E-05	1.21E-05	3.36E-05	4.80E-05	1.77E-05	2.41E-05	3.19E-05	1.22E-04	3.53E-05
F7	mean	7.14E-02	7.17E-02	6.60E-02	6.94E-02	7.15E-02	6.31E-02	6.98E-02	7.39E-02	8.06E-02	6.85E-02
	std	1.81E-02	2.63E-02	2.31E-02	1.81E-02	2.20E-02	1.93E-02	1.81E-02	2.00E-02	2.37E-02	2.37E-02
F8	mean	-1.25E+04	-1.24E+04	-1.24E+04	-1.24E+04	-1.24E+04	-1.24E+04	-1.24E+04	-1.24E+04	-1.24E+04	-1.24E+04
	std	1.18E+02	1.32E+02	1.33E+02	1.21E+02	1.39E+02	1.19E+02	1.44E+02	1.17E+02	1.05E+02	1.40E+02
F9	mean	8.39E+00	7.89E+00	4.76E+00	9.08E+00	1.22E+01	8.89E+00	8.74E+00	1.06E+01	1.56E+01	9.41E+00
	std	1.18E+02	2.97E+00	2.01E+00	3.38E+00	3.68E+00	3.25E+00	2.81E+00	3.26E+00	3.43E+00	4.33E+00
F10	mean	2.73E-03	3.03E-03	1.66E-03	3.62E-03	8.04E-03	4.17E-03	3.96E-03	6.73E-03	4.25E-02	4.56E-03
	std	1.05E-03	8.66E-04	6.01E-04	9.74E-04	3.13E-03	1.51E-03	1.31E-03	2.77E-03	3.24E-02	1.37E-03
F11	mean	3.78E-04	3.03E-04	4.45E-04	4.41E-04	7.64E-04	4.79E-04	5.60E-04	6.61E-04	1.48E-03	9.86E-04
	std	3.09E-04	2.41E-04	1.37E-03	4.08E-04	4.42E-04	3.63E-04	7.59E-04	5.52E-04	9.55E-04	1.73E-03
F12	mean	6.06E-07	6.17E-07	2.82E-07	8.42E-07	1.43E-06	1.09E-06	7.61E-07	1.22E-06	4.09E-06	1.07E-06
	std	3.09E-04	3.22E-07	2.23E-07	6.49E-07	9.00E-07	7.27E-07	5.46E-07	6.49E-07	2.38E-06	8.58E-07
F13	mean	6.13E-06	6.21E-06	4.05E-06	5.83E-06	1.60E-05	6.93E-06	5.95E-06	1.28E-05	4.64E-05	8.97E-06
	std	4.32E-06	3.79E-06	3.92E-06	3.13E-06	9.27E-06	4.03E-06	3.28E-06	1.32E-05	2.22E-05	4.49E-06
F14	mean	9.98E-01	9.98E-01	9.98E-01	9.98E-01	9.98E-01	9.98E-01	9.98E-01	9.98E-01	9.98E-01	9.98E-01
	std	1.13E-16	1.13E-16	1.13E-16	1.13E-16	1.13E-16	1.13E-16	1.13E-16	1.13E-16	1.13E-16	1.13E-16
F15	mean	6.87E-04	6.89E-04	6.84E-04	6.51E-04	6.81E-04	6.97E-04	6.70E-04	6.51E-04	6.82E-04	6.74E-04
	std	1.23E-04	8.62E-05	7.31E-05	1.20E-04	6.00E-05	1.72E-04	8.94E-05	9.89E-05	7.94E-05	9.94E-05
F16	mean	-1.03E+00	-1.03E+00	-1.03E+00	-1.03E+00	-1.03E+00	-1.03E+00	-1.03E+00	-1.03E+00	-1.03E+00	-1.03E+00
	std	6.65E-16	6.78E-16	6.71E-16	6.71E-16	6.78E-16	6.65E-16	6.58E-16	6.58E-16	6.71E-16	6.71E-16
F17	mean	3.98E-01	3.98E-01	3.98E-01	3.98E-01	3.98E-01	3.98E-01	3.98E-01	3.98E-01	3.98E-01	3.98E-01
	std	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
F18	mean	3.00E+00	3.00E+00	3.00E+00	3.00E+00	3.00E+00	3.00E+00	3.00E+00	3.00E+00	3.00E+00	2.99E+00
	std	9.90E-16	8.92E-16	1.48E-15	1.39E-15	1.35E-15	7.56E-16	1.37E-15	7.78E-16	8.21E-16	2.20E-15
F19	mean	-3.00E-01	-3.00E-01	-3.00E-01	-3.00E-01	-3.00E-01	-3.00E-01	-3.00E-01	-3.00E-01	-3.00E-01	-3.00E-01
	std	9.90E-16	2.26E-16	2.26E-16	2.26E-16	2.26E-16	2.26E-16	2.26E-16	2.26E-16	2.26E-16	2.26E-16
F20	mean	-3.31E+00	-3.31E+00	-3.30E+00	-3.30E+00	-3.31E+00	-3.32E+00	-3.31E+00	-3.31E+00	-3.32E+00	-3.30E+00
	std	3.02E-02	4.11E-02	4.51E-02	4.84E-02	3.63E-02	1.33E-15	4.11E-02	4.11E-02	1.34E-15	4.84E-02
F21	mean	-8.65E+00	-8.15E+00	-8.56E+00	-8.23E+00	-9.57E+00	-8.47E+00	-8.90E+00	-8.64E+00	-9.57E+00	-8.07E+00
	std	3.02E-02	2.97E+00	2.76E+00	3.06E+00	1.83E+00	2.67E+00	2.61E+00	2.60E+00	1.82E+00	3.08E+00
F22	mean	-9.93E+00	-9.96E+00	-9.56E+00	-9.53E+00	-9.78E+00	-9.57E+00	-9.32E+00	-1.02E+01	-9.99E+00	-9.11E+00
	std	1.82E+00	1.69E+00	2.20E+00	2.29E+00	1.91E+00	2.18E+00	2.50E+00	1.22E+00	1.53E+00	2.63E+00
F23	mean	-9.85E+00	-9.79E+00	-9.47E+00	-1.05E+01	-1.05E+01	-9.57E+00	-1.00E+01	-1.03E+01	-1.01E+01	-1.01E+01
	std	2.13E+00	2.29E+00	2.44E+00	1.98E-15	2.06E-15	2.52E+00	2.00E+00	1.22E+00	1.83E+00	1.68E+00
F24	mean	1.81E+02	1.54E+02	9.67E+01	1.86E+02	4.46E+02	2.92E+02	2.30E+02	3.56E+02	1.15E+03	2.73E+02
	std	1.70E+02	7.57E+01	7.27E+01	7.80E+01	2.18E+02	2.43E+02	1.09E+02	1.46E+02	6.32E+02	1.47E+02
F25	mean	1.24E-18	1.32E-17	3.92E-19	1.51E-18	1.76E-17	5.98E-18	3.98E-18	3.24E-17	5.84E-16	7.05E-18
	std	2.16E-18	5.91E-17	1.59E-18	3.55E-18	3.50E-17	1.38E-1	1.42E-17	7.79E-17	9.87E-16	1.81E-17
F26	mean	6.35E+01	7.51E+01	6.00E+01	7.30E+01	9.19E+01	6.81E+01	7.62E+01	8.30E+01	1.10E+02	6.74E+01
	std	2.42E+01	2.59E+01	1.94E+01	2.99E+01	3.03E+01	2.65E+01	2.75E+01	2.70E+01	3.07E+01	2.58E+01
F27	mean	3.09E-07	3.26E-07	1.57E-07	4.82E-07	9.81E-07	4.20E-07	3.73E-07	6.50E-07	3.65E-06	5.01E-07
	std	1.38E-07	1.93E-07	8.43E-08	6.59E-07	6.51E-07	3.00E-07	2.38E-07	5.00E-07	1.63E-06	2.78E-07
F28	mean	3.61E-02	2.70E-02	2.24E-02	3.70E-02	8.49E-02	4.70E-02	3.41E-02	6.18E-02	1.33E-01	4.65E-02
	std	2.60E-02	1.56E-02	1.01E-02	2.07E-02	4.98E-02	2.69E-02	1.34E-02	3.02E-02	5.33E-02	2.75E-02
F29	mean	9.00E-04	8.40E-04	6.50E-04	9.89E-04	1.33E-03	9.69E-04	9.73E-04	1.26E-03	2.06E-03	1.16E-03
	std	1.84E-04	2.16E-04	1.34E-04	2.15E-04	2.88E-04	1.51E-04	1.82E-04	2.30E-04	3.33E-04	2.32E-04
F30	mean	5.59E-03	5.66E-03	4.32E-03	5.98E-03	6.60E-03	5.45E-03	5.50E-03	6.68E-03	2.99E-02	5.82E-03
	std	1.88E-03	1.78E-03	1.13E-03	2.09E-03	2.70E-03	1.68E-03	1.89E-03	2.12E-03	1.76E-02	1.59E-03

En iyi sonuç sayısı	6	5	20	6	5	6	4	5	6	5
Friedman sıralama değerleri	3.87	4.77	3.05	5.12	7.23	5.22	4.62	6.87	8.02	6.25
Genel sıralama	2	4	1	5	9	6	3	8	10	7
P-Değeri	3.36E-14									

CAAAv3 yaklaşımı on bir tek modlu fonksiyonun sekiz tanesinde, dokuz çok modlu fonksiyonun sekiz tanesinde ve on sabit boyutlu çok modlu fonksiyonun dört tanesinde en iyi sonuçları elde etmiştir. CAAAv3 varyasyonunu tek modlu ve çok modlu fonksiyonların büyük bir kısmında başarılı sonuçlar elde etmesi diğer varyasyonlardan daha iyi keşif ve sömürü dengesini oluşturduğunu göstermektedir.

Tüm benchmark fonksiyonları birlikte değerlendirildiğinde hem en çok birinciliği elde etme (Best Count) hem de Friedman sıralama sonuçları CAAAv3 yaklaşımının diğer varyasyonlardan daha verimli bir keşif sömürü dengesi oluşturduğu ve daha başarılı bir performansı sağladığı görülmektedir.

4.1.2. Geliştirilen Yöntemin Literatürdeki Algoritmalarla Performans Kıyaslaması

Bir önceki bölümde en başarılı kaotik varyasyonunun gauss kaotik haritası ile donatılan CAAAv3 olduğu tespit edilmiştir. Bu bölümde CAAAv3 varyasyonunun literatürdeki popüler yöntemlere karşı olan performansı analiz edilmiştir. Karşılaştırma sürecinde on farklı karakteristik yapıya sahip metasezgisel optimizasyon algoritması kullanılmıştır. Bu algoritmalar PSO (Kennedy ve Eberhart, 1995), GWO (Mirjalili ve ark., 2014a), MVO (Aljarah ve ark., 2020b), Sinüs Kosinüs Algoritması (SCA) (Mirjalili, 2016), FFA (Yang, 2010a), CS (Yang ve Deb, 2009), Jaya Algoritması (JAYA) (Rao, 2016), Güve Alev Optimizasyonu (MFO) (Mirjalili, 2015), SSA (Mirjalili ve ark., 2017) ve WOA (Mirjalili ve Lewis, 2016)'dır. Karşılaştırma sürecinde orijinal AAA ile de kıyaslanmıştır.

Çizelge 4.2'de kıyaslanan algoritmaların kendilerine has çalışma parametreleri gösterilmektedir. Bu tez çalışması kapsamındaki deneysel süreçte Çizelge 4.2'deki parametre değerleri kullanılmıştır. Bu parametreler ilgili algoritma yazarlarının kendi yayınlarında tavsiye ettiği değerlerdir (Faris ve ark., 2016).

Çizelge 4.2. Deneysel çalışmada kıyaslanan algoritmaların parametre değerleri

Algoritmalar	Parametreler	Değerleri
AAA	Δ Share force	2
	e Energy loss	0.3
	A_p Adaptation	0.2
CS	Pa (Discovery rate of alien eggs)	0.25
PSO	c1,c2	2, 2
	$W_{inertia}$ weight	2
	W_{max} , W_{min}	0.9, 0.4
GWO	alpha	2
WOA	alpha	2
FFA	alpha (Randomness)	0.5
	β_{min} (minimum value of beta)	0.2
	Gama (absorption coefficient)	1
MVO	wep_{max}	1
	wep_{min}	0.2
SCA	A	2.0
	r1	[0, 2]
	r2	Random [0, 2π]
	r3	[0, 2]
SSA	Probability of crossover	0.8
	Probability of mutation	0.01
MFO	a (linearly decreases)	[-1, -2]
JAYA	r1	[0, 1]
	r2	[0, 1]

Deneysel çalışmada maksimum uygunluk değerlendirme sayısı (Fitness Evaluations) 50,000 olarak, popülasyondaki birey sayısı 40 olarak belirlenmiştir. Adil bir karşılaştırma olması için otuz birbirinden bağımsız çalışma yapılarak ortalaması kullanılmıştır. Bu parametreler belirlenirken literatürdeki birçok çalışma göz önünde bulundurulmuştur (Gandomi ve ark., 2013a; Kaur ve Arora, 2018; Kohli ve Arora, 2018; Sayed ve ark., 2018).

Çizelge 4.3’ de bütün yöntemlerin her bir fonksiyonda elde ettiği otuz bağımsız çalışma sonuçlarının ortalaması ve standart sapması sunulmuştur. Ayrıca tablonun en alt bölümünde yöntemlerin elde ettiği en iyi sonuç sayısı (Best Count) ve Friedman sıralama testi sonuçlarının bulunduğu bir özet kısmı sunulmaktadır.

Çizelge 4.3. CAAAv3 ile literatürdeki algoritmaların benchmark fonksiyonları üzerindeki sonuçları

		SSA	PSO	FFA	GWO	WOA	MVO	MFO	CS	SCA	JAYA	AAA	CAAAv3
F1	mean	1.44E-08	3.21E-08	3.25E-03	2.38E-53	3.02E-123	4.48E-01	2.67E+03	4.34E-01	4.40E-01	4.71E-09	3.61E-05	1.49E-05
	std	3.85E-09	5.41E-08	1.06E-03	6.62E-53	9.67E-123	1.14E-01	5.83E+03	2.18E-01	1.20E+00	1.90E-08	1.75E-05	1.04E-05
F2	mean	5.18E-01	5.00E+00	8.51E+00	2.24E-04	2.82E-79	3.71E+00	3.77E+01	2.22E+00	7.72E-04	1.38E-07	3.36E-04	2.32E-31
	std	4.92E-01	5.09E+00	1.22E+01	8.18E-05	1.53E-78	1.78E+01	2.05E+01	9.18E-01	2.25E-03	1.09E-07	9.38E-05	2.74E-31
F3	mean	3.05E+02	2.16E+01	2.38E+02	5.74E-11	2.31E+04	5.02E+01	2.59E+04	1.05E+03	4.16E+03	1.31E+04	4.31E+03	3.34E+03
	std	2.72E+02	1.21E+01	2.84E+02	1.32E-10	8.93E+03	2.20E+01	1.62E+04	2.97E+02	3.69E+03	6.20E+03	1.50E+03	1.35E+03
F4	mean	6.61E+00	6.54E-01	1.55E-01	3.36E-11	3.70E+01	8.64E-01	5.97E+01	6.64E+00	2.11E+01	1.25E+01	1.34E+01	1.33E+01
	std	3.08E+00	1.74E-01	5.22E-02	3.32E-11	3.09E+01	3.07E-01	9.28E+00	1.57E+00	9.05E+00	7.18E+00	2.11E+00	2.64E+00
F5	mean	2.54E+02	1.64E+02	4.73E+02	2.70E+01	2.71E+01	3.07E+02	2.68E+06	1.07E+02	8.76E+02	4.36E+02	9.62E+01	7.97E+01
	std	5.03E+02	5.46E+02	1.08E+03	8.41E-01	3.28E-01	6.31E+02	1.46E+07	4.52E+01	1.57E+03	1.99E+03	2.42E+01	3.55E+01
F6	mean	1.57E-08	6.17E-08	3.54E-03	3.70E-01	4.07E-02	4.01E-01	3.02E+03	3.71E-01	4.98E+00	3.60E+00	3.78E-05	1.56E-05
	std	3.92E-09	1.41E-07	1.35E-03	2.86E-01	4.77E-02	1.08E-01	5.99E+03	1.15E-01	1.30E+00	6.16E-01	2.29E-05	1.21E-05
F7	mean	8.39E-02	2.48E+00	4.43E-01	1.03E-03	1.44E-03	1.91E-02	3.71E+00	5.08E-02	5.34E-02	3.28E-02	6.49E-02	6.60E-02
	std	3.92E-02	2.85E+00	1.56E-01	6.50E-04	1.70E-03	7.83E-03	6.35E+00	1.51E-02	5.75E-02	3.87E-02	2.04E-02	2.31E-02
F8	mean	-7.59E+03	-6.13E+03	-2.18E+04	-6.11E+03	-1.06E+04	-7.92E+03	-8.30E+03	-9.79E+03	-3.91E+03	-5.18E+03	-1.24E+04	-1.24E+04
	std	8.26E+02	1.07E+03	3.77E+03	9.04E+02	1.71E+03	7.12E+02	8.17E+02	1.00E+03	3.08E+02	8.89E+02	1.44E+02	1.33E+02
F9	mean	5.37E+01	8.86E+01	7.90E+01	4.94E+00	0.00E+00	1.15E+02	1.82E+02	9.02E+01	3.46E+01	6.91E+01	9.24E+00	4.76E+00
	std	2.24E+01	2.56E+01	2.52E+01	6.52E+00	0.00E+00	3.52E+01	4.48E+01	1.26E+01	4.96E+01	3.24E+01	3.77E+00	2.01E+00
F10	mean	1.64E+00	2.91E-04	1.45E-01	1.66E-03	3.40E-15	1.20E+00	1.66E+01	2.96E+00	1.10E+01	1.91E-01	3.82E-03	2.12E-14
	std	9.86E-01	4.60E-04	3.38E-01	6.01E-04	2.30E-15	6.30E-01	6.48E+00	4.19E-01	9.46E+00	1.05E+00	1.22E-03	3.24E-15
F11	mean	1.29E-02	9.19E-03	1.42E-02	2.42E-03	4.34E-03	5.83E-01	2.11E+01	6.12E-01	4.46E-01	3.32E-02	4.74E-04	4.45E-04
	std	1.34E-02	1.25E-02	9.77E-03	5.14E-03	1.75E-02	9.75E-02	4.55E+01	1.15E-01	3.33E-01	7.24E-02	2.91E-04	1.37E-03
F12	mean	4.63E+00	1.04E-02	3.57E-03	2.85E-02	5.32E-03	1.11E+00	3.12E+01	1.79E+00	4.45E+00	5.86E-01	1.03E-06	2.82E-07
	std	2.03E+00	4.17E-02	1.94E-02	1.18E-02	4.77E-03	8.94E-01	1.69E+02	5.28E-01	6.49E+00	3.12E-0	6.78E-07	2.23E-07
F13	mean	4.16E+00	2.20E-03	3.23E-03	3.70E-01	1.56E-01	9.29E-02	4.59E-01	1.50E+00	4.45E+03	1.84E+04	6.47E-06	4.05E-06
	std	1.25E+01	4.47E-03	5.68E-03	1.86E-01	1.09E-01	7.62E-02	8.99E-01	1.20E+00	1.93E+04	1.01E+05	3.03E-06	3.92E-06
F14	mean	9.98E-01	1.99E+00	9.98E-01	4.32E+00	2.34E+00	9.98E-01	1.85E+00	9.98E-01	1.46E+00	1.10E+00	9.98E-01	9.98E-01
	std	2.39E-16	1.79E+00	2.18E-13	4.35E+00	2.93E+00	1.31E-11	1.61E+00	1.13E-16	8.53E-01	5.42E-01	1.13E-16	1.13E-16
F15	mean	8.71E-04	4.42E-03	1.04E-03	2.35E-03	7.66E-04	5.91E-03	2.63E-03	4.35E-04	1.03E-03	1.08E-03	6.88E-04	6.84E-04
	std	2.42E-04	7.38E-03	2.23E-04	6.11E-03	4.21E-04	1.21E-02	5.01E-03	1.09E-04	3.54E-04	4.86E-04	1.83E-04	7.31E-05
F16	mean	-1.03E+00	-1.03E+00	-1.03E+00	-1.03E+00	-1.03E+00	-1.03E+00	-1.03E+00	-1.03E+00	-1.03E+00	-1.03E+00	-1.03E+00	-1.03E+00
	std	5.60E-15	6.78E-16	1.84E-09	8.00E-09	2.15E-11	2.18E-07	6.78E-16	5.68E-16	2.28E-05	1.88E-05	6.71E-16	6.71E-16
F17	mean	3.98E-01	3.98E-01	3.98E-01	3.98E-01	3.98E-01	3.98E-01	3.98E-01	3.98E-01	3.99E-01	3.98E-01	3.98E-01	3.98E-01
	std	5.21E-15	0.00E+00	1.01E-09	2.51E-07	1.14E-06	5.25E-08	0.00E+00	0.00E+00	7.64E-04	0.00E+00	0.00E+00	0.00E+00
F18	mean	3.00E+00	3.00E+00	3.00E+00	3.00E+00	3.00E+00	3.00E+00	3.00E+00	3.00E+00	3.00E+00	3.07E+00	3.00E+00	2.99E+00
	std	1.08E-13	1.14E-15	1.66E-08	6.11E-06	8.65E-06	9.88E-07	1.61E-15	1.35E-15	2.40E-05	3.36E-01	6.90E-16	1.48E-15
F19	mean	-3.00E-01	-3.00E-01	-3.86E+00	-3.00E-01	-3.00E-01	-3.00E-01	-3.00E-01	-2.94E+00	-3.00E-01	-2.14E-02	-3.00E-01	-3.00E-01
	std	2.26E-16	2.26E-16	3.70E-09	2.26E-16	2.26E-16	2.26E-16	2.26E-16	4.33E-01	2.26E-16	4.51E-02	2.26E-16	2.26E-16
F20	mean	-3.22E+00	-3.26E+00	-3.28E+00	-3.24E+00	-3.26E+00	-3.27E+00	-3.22E+00	-3.32E+00	-2.94E+00	-2.18E+00	-3.30E+00	-3.30E+00
	std	4.22E-02	8.26E-02	5.85E-02	7.87E-02	9.14E-02	6.04E-02	4.82E-02	1.38E-08	3.30E-01	5.07E-01	4.51E-02	4.51E-02
F21	mean	-7.82E+00	-8.30E+00	-8.90E+00	-8.97E+00	-9.73E+00	-7.96E+00	-5.98E+00	-1.02E+01	-2.88E+00	-1.74E+00	-9.06E+00	-8.56E+00
	std	3.41E+00	2.73E+00	2.61E+00	2.18E+00	1.63E+00	2.80E+00	3.37E+00	4.36E-10	2.02E+00	1.01E+00	2.26E+00	2.76E+00
F22	mean	-8.42E+00	-7.88E+00	-1.01E+01	-1.00E+01	-8.72E+00	-8.38E+00	-7.93E+00	-1.04E+01	-3.87E+00	-1.84E+00	-9.83E+00	-9.56E+00
	std	3.15E+00	3.21E+00	1.39E+00	1.34E+00	2.63E+00	2.73E+00	3.36E+00	1.69E-10	2.30E+00	6.19E-01	1.77E+0	2.20E+00
F23	mean	-7.96E+00	-9.82E+00	-1.02E+01	-1.02E+01	-7.72E+00	-9.10E+00	-8.58E+00	-1.05E+01	-4.61E+00	-2.21E+00	-1.05E+01	-9.47E+00
	std	3.51E+00	1.86E+00	1.36E+00	1.58E+00	3.29E+00	2.42E+00	3.33E+00	1.83E-09	1.49E+00	8.44E-01	5.71E-16	2.44E+00
F24	mean	7.74E+02	6.17E-01	4.73E+04	2.45E-46	5.72E-113	3.71E+06	2.33E+10	3.02E+06	5.26E+05	6.46E-03	2.38E+02	9.67E+01
	std	1.56E+03	1.45E+00	3.06E+04	5.17E-46	2.97E-112	1.13E+06	4.30E+10	9.28E+05	1.20E+06	2.25E-02	1.50E+02	7.27E+01
F25	mean	6.20E-04	8.85E-14	3.23E-04	1.44E-133	8.31E-121	3.56E-05	4.14E-15	2.50E-11	5.15E-06	7.46E-15	3.05E-18	3.92E-19
	std	4.07E-04	4.33E-13	1.71E-04	7.83E-133	4.55E-120	2.98E-05	2.22E-14	3.69E-11	1.72E-05	3.49E-14	7.14E-18	1.59E-18
F26	mean	3.93E+00	2.21E+02	2.03E+01	3.91E-18	4.72E+02	1.03E-01	3.71E+02	1.57E+02	1.33E+01	1.20E+02	7.20E+01	6.00E+01
	std	3.13E+00	1.36E+02	5.15E+01	1.25E-17	1.01E+02	3.52E-02	1.03E+02	2.71E+01	9.53E+00	4.83E+01	1.85E+01	1.94E+01
F27	mean	6.62E+00	3.95E-01	2.07E+00	1.02E+00	1.56E-01	1.93E+01	3.03E+01	5.82E+00	2.29E+00	6.41E+00	3.63E-07	1.57E-07
	std	3.18E+00	1.86E+00	3.23E+00	2.98E-01	1.49E-01	7.56E+00	1.40E+01	3.04E+00	3.25E-01	5.67E+00	1.87E-07	8.43E-08
F28	mean	4.02E+06	9.75E-04	1.43E+07	4.53E-50	5.37E-115	6.25E+06	3.85E+07	6.68E+02	7.65E+00	2.76E-06	3.80E-02	2.24E-02
	std	3.59E+06	1.84E-03	7.28E+06	7.23E-50	2.94E-114	2.22E+06	2.85E+07	3.06E+02	1.29E+01	7.79E-06	1.64E-02	1.01E-02
F29	mean	7.11E-01	4.46E-01	9.79E-01	0.00E+00	0.00E+00	4.98E-01	4.75E-01	4.42E-01	5.88E-05	1.33E-05	1.01E-03	6.50E-04
	std	1.35E-01	2.09E-01	2.15E-01	0.00E+00	0.00E+00	1.71E-01	2.96E-01	6.13E-02	1.01E-04	2.55E-05	2.02E-04	1.34E-04
F30	mean	4.29E-03	4.15E-03	1.94E-02	2.97E-02	8.37E-03	2.53E-02	2.66E-03	2.93E-02	6.28E-02	5.65E-02	5.80E-03	4.32E-03
	std	9.72E-04	1.70E-03	4.24E-03	8.05E-03	2.69E-03	7.72E-03	5.23E-04	8.76E-03	1.14E-02	1.67E-02	1.52E-03	1.13E-03

En iyi sonuç sayısı	2	2	2	7	7	0	3	7	0	1	4	8
Freidman sıralama değerleri	7.05	5.88	6.30	4.95	5.15	7.93	9.33	6.07	9.07	8.08	4.40	3.78
Genel sıralama	8	5	7	3	4	9	12	6	11	10	2	1
P-Değeri	1.46E-13											

Çizelge 4.3’ deki sonuçlar analiz edildiğinde, en iyi sonuç sayısı açısından CAAAv3 yöntemi sekiz, CS, WOA ve GWO yöntemleri yedi fonksiyonda en iyi sonucu elde ederek iyi bir performans ortaya koymuşlardır. MVO ve SCA yöntemi ise hiçbir fonksiyonda en iyi sonuç elde edememiştirler. Yöntemlerin elde ettikleri ortalama değerleri Friedman sıralama testine tabi tutulduğunda CAAAv3 yöntemi 3.78 ile en iyi sıralama değerine sahip olurken, orijinal AAA yöntemi de 4.40 sıralama değeri ile ikinci başarılı yöntem olarak görülmektedir. Orijinal AAA yöntemi dört fonksiyonda en iyi değeri elde etmesine rağmen, Friedman sıralama sonuçlarına göre en iyi ikinci skoru elde etmektedir. Bunun nedeni orijinal AAA yönteminin en iyi sonucu elde edemediği fonksiyonlarda da rekabetçi sonuçlar elde etmesidir. CAAAv3 yöntemi sekiz fonksiyonda ulaştığı en iyi değer dışında diğer fonksiyonlarda da rekabetçi sonuçlar elde etmektedir. Yapılan Friedman sıralama test sonucunda elde edilen p-Value (1.46E-13) değerinin 0.05’in altında olması yapılan testin istatistiksel açıdan anlamlı olduğunu göstermektedir.

CAAAv3 yöntemi ile kıyaslanan diğer algoritmaların Wilcoxon işaretli sıralar testi sonuçları Çizelge 4.4 ve 4.5 de verilmiştir. Test sonuçlarının daha anlaşılır ve incelenebilir olması için iki ayrı çizelge şeklinde sunulmuştur. İlgili çizelgelerde yapılan Wilcoxon testine ait p-Value, T ve W değerleri sunulmaktadır. W değeri 0 ise sonuçlar arasında istatistiksel açıdan bir fark yoktur eşitlik söz konusudur. W değeri 1 ise kıyaslanan iki algoritma arasında istatistiksel açıdan anlamlı bir fark vardır ve tez kapsamında geliştirilen CAAAv3 yöntemi kıyaslanan diğer yöntemden daha başarılıdır. W değeri 2 ise sonuçlar arasında istatistiksel açıdan anlamlı bir fark vardır ve kıyaslanan yöntem daha başarılıdır.

Testlerde anlam seviyesi birçok çalışmada olduğu gibi $\alpha = 0,05$ olarak kullanılmıştır. P-değeri (p-value) bu α değeri ile kıyaslanır. $R +$ değeri ilk algoritmanın ikinciden daha iyi olduğu sonuçlardaki sıralama değerlerinin toplamalarını ifade eder ve $R -$ ise ikinci algoritmanın daha iyi olduğu sonuçlardaki sıralama değerlerinin toplamalarını ifade eder. T ise bu toplamaların küçüğünü temsil eder ($T = \min(R+, R-)$) (Derrac ve ark., 2011).

Çizelge 4.4- 4.5'in son satırında Wilcoxon işaretli sıralar testi sonucunda karşılaştırılan iki yöntem için eşitlik, kazanılan ve kaybedilen değerlerin toplamı ayrı ayrı sunulmuştur.

Çizelge 4.4. Geliştirilen CAAAv3 yöntemi ile kıyaslanan yöntemlerin (SSA, PSO, FFA, GWO, WOA, MVO) Wilcoxon işaretli sıralar testi sonuçları

Fonksiyon Numarası	CAAAv3 - SSA			CAAAv3 - PSO			CAAAv3 - FFA			CAAAv3 - GWO			CAAAv3 - WOA			CAAAv3 - MVO		
	p-Value	T	W	p-Value	T	W	p-Value	T	W	p-Value	T	W	p-Value	T	W	p-Value	T	W
F1	1.73E-06	465	2	1.73E-06	465	2	1.73E-06	0	1	1.73E-06	465	2	1.73E-06	465	2	1.73E-06	0	1
F2	1.73E-06	0	1	2.05E-04	52	1	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	465	2	1.73E-06	0	1
F3	1.73E-06	465	2	1.73E-06	465	2	1.73E-06	465	2	1.73E-06	465	2	1.92E-06	1	1	1.73E-06	465	2
F4	6.34E-06	452	2	1.73E-06	465	2	1.73E-06	465	2	1.73E-06	465	2	2.11E-03	83	1	1.73E-06	465	2
F5	5.30E-01	202	0	2.54E-01	288	0	8.22E-02	148	0	3.18E-06	459	2	3.18E-06	459	2	2.18E-02	121	1
F6	1.73E-06	465	2	1.73E-06	465	2	1.73E-06	0	1	2.35E-06	3	1	1.73E-06	0	1	1.73E-06	0	1
F7	4.49E-02	135	1	2.83E-04	56	1	1.73E-06	0	1	1.73E-06	465	2	1.73E-06	465	2	1.73E-06	465	2
F8	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	465	2	1.73E-06	0	1	2.84E-05	29	1	1.73E-06	0	1
F9	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	0	1	4.65E-01	268	0	1.73E-06	465	2	1.73E-06	0	1
F10	4.73E-06	10	1	4.29E-06	456	2	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	465	2	1.73E-06	0	1
F11	1.15E-04	45	1	2.43E-02	123	1	1.92E-06	1	1	1.65E-01	300	0	3.59E-04	406	1	1.73E-06	0	1
F12	1.73E-06	0	1	3.59E-04	406	1	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	0	1
F13	2.61E-04	55	1	1.65E-01	300	0	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	0	1
F14	1.00E+00	0	0	2.44E-04	0	1	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	0	1
F15	3.06E-04	57	1	4.45E-05	34	1	2.60E-06	4	1	1.48E-02	351	1	9.92E-01	232	0	1.96E-03	82	1
F16	7.00E-07	0	1	1.00E+00	0	0	1.73E-06	0	1	1.73E-06	0	1	2.56E-06	0	1	1.73E-06	0	1
F17	1.14E-04	0	1	1.00E+00	0	0	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	0	1
F18	1.60E-06	0	1	2.15E-02	5.5	1	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	0	1
F19	1.00E+00	0	0	1.00E+00	0	0	1.73E-06	465	2	1.00E+00	0	0	1.00E+00	0	0	1.00E+00	0	0
F20	6.34E-06	13	1	1.01E-02	16.5	1	3.88E-04	60	1	8.47E-06	16	1	3.06E-04	57	1	6.89E-05	39	1
F21	3.38E-03	90	1	5.20E-01	63	0	1.78E-01	167	0	2.99E-01	182	0	4.91E-01	199	0	4.95E-02	137	1
F22	8.94E-04	71	1	2.00E-02	16	1	1.17E-02	110	2	1.32E-02	112	2	1.48E-03	78	1	2.58E-03	86	1
F23	5.71E-04	65	1	2.50E-01	33	0	5.71E-02	140	0	4.49E-02	135	2	4.68E-03	95	1	1.32E-02	112	1
F24	8.94E-04	71	1	1.73E-06	465	2	1.73E-06	0	1	1.73E-06	465	2	1.73E-06	465	2	1.73E-06	0	1
F25	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	465	2	1.73E-06	465	2	1.73E-06	0	1
F26	1.73E-06	465	2	2.60E-06	4	1	3.11E-05	435	2	1.73E-06	465	2	1.73E-06	0	1	1.73E-06	465	2
F27	1.73E-06	0	1	1.57E-02	115	1	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	0	1
F28	1.73E-06	0	1	1.73E-06	465	2	1.73E-06	0	1	1.73E-06	465	2	1.73E-06	465	2	1.73E-06	0	1
F29	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	465	2	1.73E-06	465	2	1.73E-06	0	1
F30	9.75E-01	234	0	2.06E-01	294	0	1.73E-06	0	1	1.73E-06	0	1	2.60E-06	4	1	1.73E-06	0	1
Kazanılan/ Eşitlik/ Kaybedilen	21/4/5			15/8/7			21/3/6			14/4/12			17/3/10			25/1/4		

Çizelge 4.5'te tez kapsamında geliştirilen CAAAv3 yöntemi ile kıyaslanan MFO, CS, SCA, JAYA ve orijinal AAA yöntemi arasında yapılan Wilcoxon işaretli sıralar testi sonuçları gösterilmektedir.

Çizelge 4.5. Geliştirilen CAAAv3 yöntemi ile kıyaslanan yöntemlerin (SSA, PSO, FFA, GWO, WOA, MVO) Wilcoxon işaretli sıralar testi sonuçları

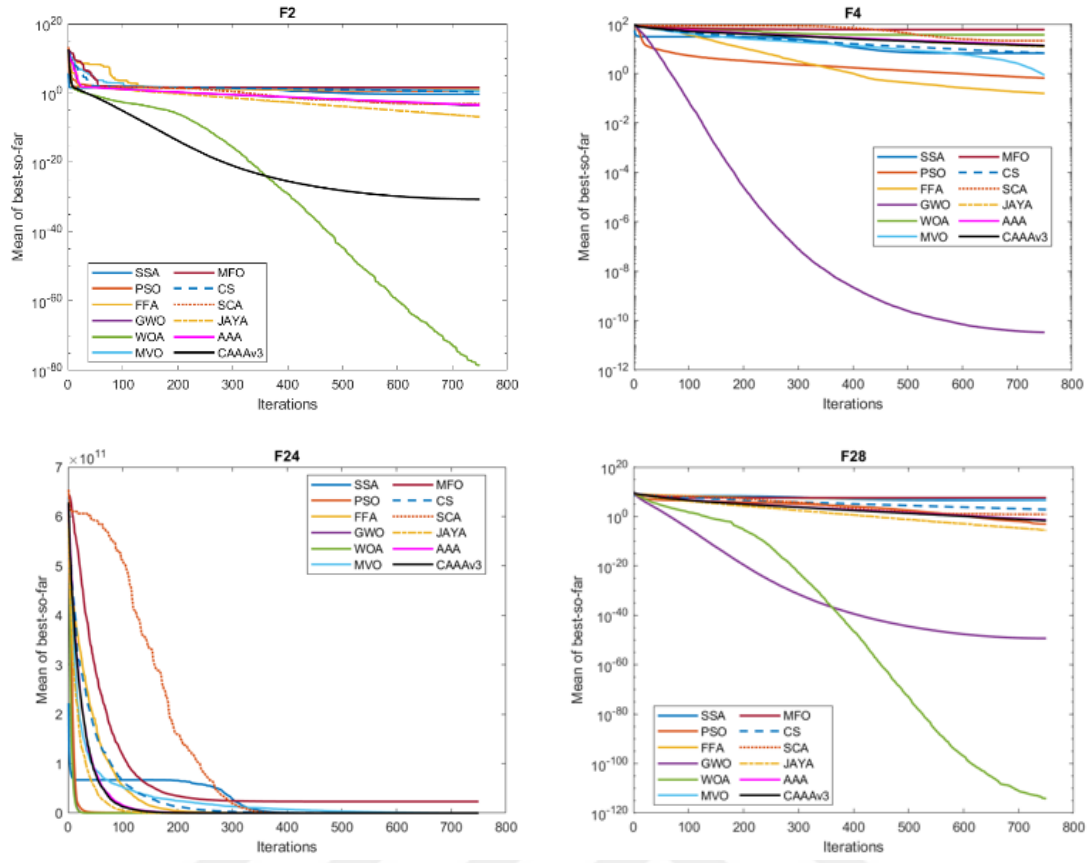
Fonksiyon Numarası	CAAAv3 - MFO			CAAAv3 -CS			CAAAv3 - SCA			CAAAv3 - JAYA			CAAAv3 - AAA		
	p-Value	T	W	p-Value	T	W	p-Value	T	W	p-Value	T	W	p-Value	T	W
F1	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	465	2	4.07E-05	33	1
F2	1.73E-06	0	1	1.73E-06	0	1	3.49E-01	278	0	1.73E-06	465	2	1.06E-04	44	1
F3	2.35E-06	3	1	1.73E-06	465	2	5.58E-01	204	0	1.92E-06	1	1	1.04E-02	108	1
F4	1.73E-06	0	1	1.73E-06	465	2	4.90E-04	63	1	2.99E-01	283	0	7.66E-01	218	0
F5	3.06E-04	57	1	5.45E-02	139	0	2.58E-03	86	1	9.37E-02	314	0	7.19E-02	145	0
F6	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	0	1	3.41E-05	31	1
F7	3.41E-05	31	1	1.59E-03	386	2	6.27E-02	323	0	5.79E-05	428	2	9.92E-01	232	0
F8	1.73E-06	0	1	1.92E-06	1	1	1.73E-06	0	1	1.73E-06	0	1	7.50E-01	217	0
F9	1.73E-06	0	1	1.73E-06	0	1	3.61E-03	91	1	1.73E-06	0	1	2.60E-05	28	1
F10	1.73E-06	0	1	1.73E-06	0	1	2.35E-06	3	1	3.11E-05	435	1	1.73E-06	0	1
F11	9.32E-06	17	1	1.73E-06	0	1	1.73E-06	0	1	6.14E-01	257	0	6.16E-04	66	1
F12	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	0	1	1.24E-05	20	1
F13	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	0	1	1.75E-02	117	1
F14	7.81E-03	0	1	1.00E+00	0	0	1.73E-06	0	1	1.73E-06	0	1	1.00E+00	0	0
F15	1.73E-06	0	1	1.73E-06	465	2	7.51E-05	40	1	2.22E-04	53	1	9.43E-01	229	0
F16	1.00E+00	0	0	1.00E+00	0	0	1.73E-06	0	1	1.73E-06	0	1	1.00E+00	0	0
F17	1.00E+00	0	0	1.00E+00	0	0	1.73E-06	0	1	1.00E+00	0	0	1.00E+00	0	0
F18	1.29E-02	15	1	1.29E-02	15	1	1.73E-06	0	1	1.73E-06	0	1	1.00E+00	12	0
F19	1.00E+00	0	0	1.73E-06	465	2	1.00E+00	0	0	1.73E-06	0	1	1.00E+00	0	0
F20	2.25E-05	0	1	5.71E-02	140	0	1.73E-06	0	1	1.73E-06	0	1	1.00E+00	18	0
F21	3.10E-03	36.5	1	6.73E-01	212	0	5.22E-06	11	1	1.92E-06	1	1	4.25E-01	50	0
F22	3.06E-02	18.5	1	1.48E-02	114	2	3.88E-06	8	1	1.73E-06	0	1	5.63E-01	10.5	0
F23	1.08E-01	14.5	0	5.71E-02	140	0	5.75E-06	12	1	2.35E-06	3	1	6.25E-02	15	0
F24	1.73E-06	0	1	1.73E-06	0	1	2.35E-06	3	1	1.73E-06	465	2	1.06E-04	44	1
F25	1.80E-05	24	1	1.73E-06	0	1	1.73E-06	0	1	8.22E-03	104	1	1.36E-04	47	1
F26	1.73E-06	0	1	1.73E-06	0	1	1.92E-06	464	2	4.29E-06	9	1	3.00E-02	127	1
F27	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	0	1	2.37E-05	27	1
F28	1.73E-06	0	1	1.73E-06	0	1	2.35E-06	3	1	1.73E-06	465	2	1.97E-05	25	1
F29	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	465	2	1.73E-06	465	2	6.34E-06	13	1
F30	6.34E-06	452	2	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	0	1	3.06E-04	57	1
Kazanılan/ Eşitlik/ Kaybedilen	25/1/4			17/7/6			24/4/2			20/4/6			16/14/0		

Wilcoxon işaretli sıralar testi sayesinde CAAAv3 yaklaşımının, literatürdeki popüler yöntemler ile bire bir kıyaslaması yapılarak detaylı bir performans analizi ortaya konulmuştur. Çizelge 4.4 ve 4.5’ deki sonuçlar analiz edildiğinde, CAAAv3 yaklaşımı kıyaslanan bütün yöntemler karşısında daha yüksek kazanılan (win) değerine sahiptir. Genel anlamda geliştirilen algoritma otuz fonksiyonun %50 den fazlasında diğer yöntemlerden istatistiksel anlamda daha başarılı sonuçlar elde etmektedir. Friedman sıralama değerine göre en iyi yöntem CAAAv3’tür. İkinci ve üçüncü sırada yer alan AAA

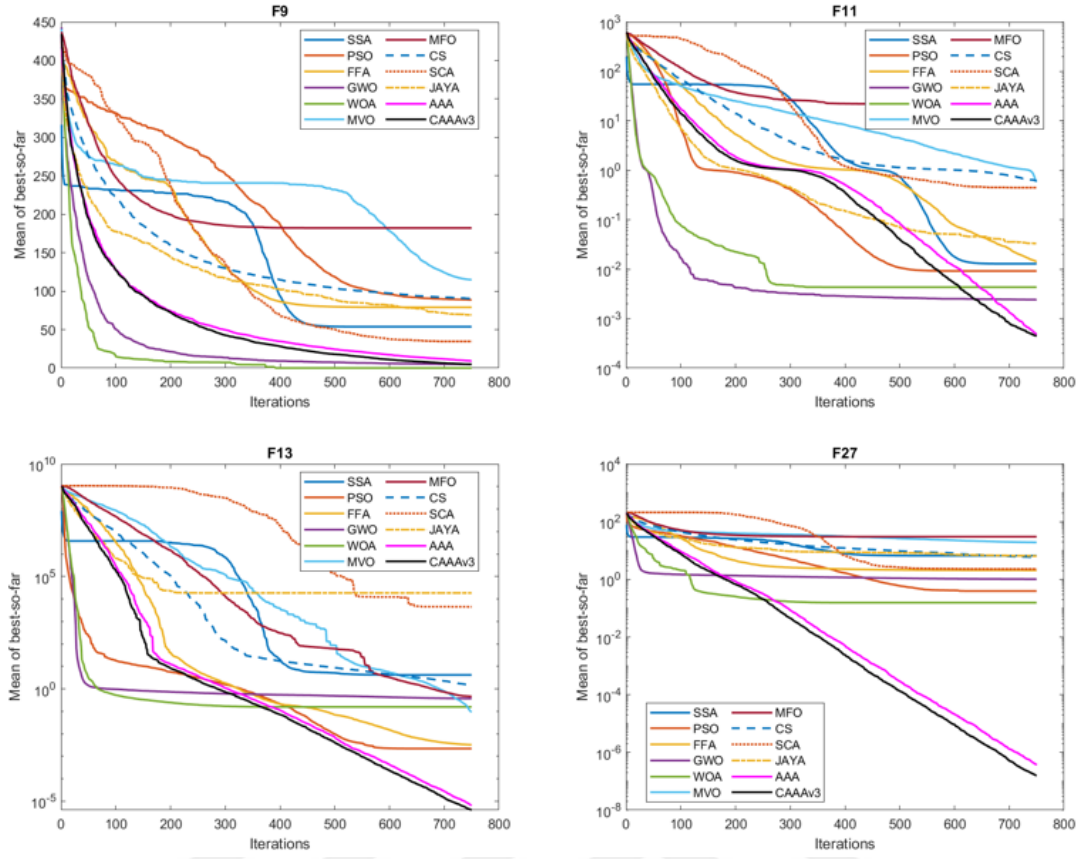
ve GWO yöntemleri ile yapılan kıyaslama sonuçlarına bakıldığında CAAAv3 yöntemi on altı fonksiyonda AAA yönteminden daha iyi sonuç elde ederken on dört fonksiyonda istatistiksel olarak aynı sonuçları elde etmiştir. GWO yöntemiyle kıyaslandığında ise on dört fonksiyonda CAAAv3 yöntemi, on iki fonksiyonda GWO yaklaşımı daha iyi sonuç elde ederken dört fonksiyonda istatistiksel açıdan aynı sonucu elde etmiştir. Yapılan istatistiksel analiz çalışmasında AAA ve GWO istatistiksel olarak rekabetçi bir yaklaşım gösterse de CAAAv3 yaklaşımının kıyaslanan yöntemlerden daha üstün sonuçlar elde ettiği görülmektedir. Detaylı analiz sonucunda CAAAv3 yönteminin çok modlu ve sabit boyutlu çok modlu fonksiyonlarda daha performanslı olduğu aşikardır. GWO yöntemi ise tek modlu fonksiyonlarda başarılı bir performans ortaya koymaktadır.

Optimizasyon yöntemleri optimum çözümü araştırırken kendilerine has arama davranışları göstermektedirler. Bu davranışlar problemin çeşidine ve yöntemin çözüm uzayının araştırılması (keşfi) ve mevcut çözümün geliştirilmesi (sömürü) stratejilerine göre şekil almaktadır. Bir optimizasyon yönteminin çözüm arama davranışının ortaya konulmasındaki en önemli araç yakınsama (convergence) eğrileridir. Yakınsama eğrileri iterasyonlar boyunca yöntemin elde ettiği en iyi çözüm temel alınarak oluşturulmaktadır. Yapılan deneysel çalışma sonucunda on iki yöntemin otuz benchmark fonksiyonunda elde ettikleri yakınsama eğrileri oluşturulmuştur. Bu eğrilerden farklı karakteristikteki on iki fonksiyonun yakınsama eğrileri Şekil 4.1-4.3 de gösterilmektedir.

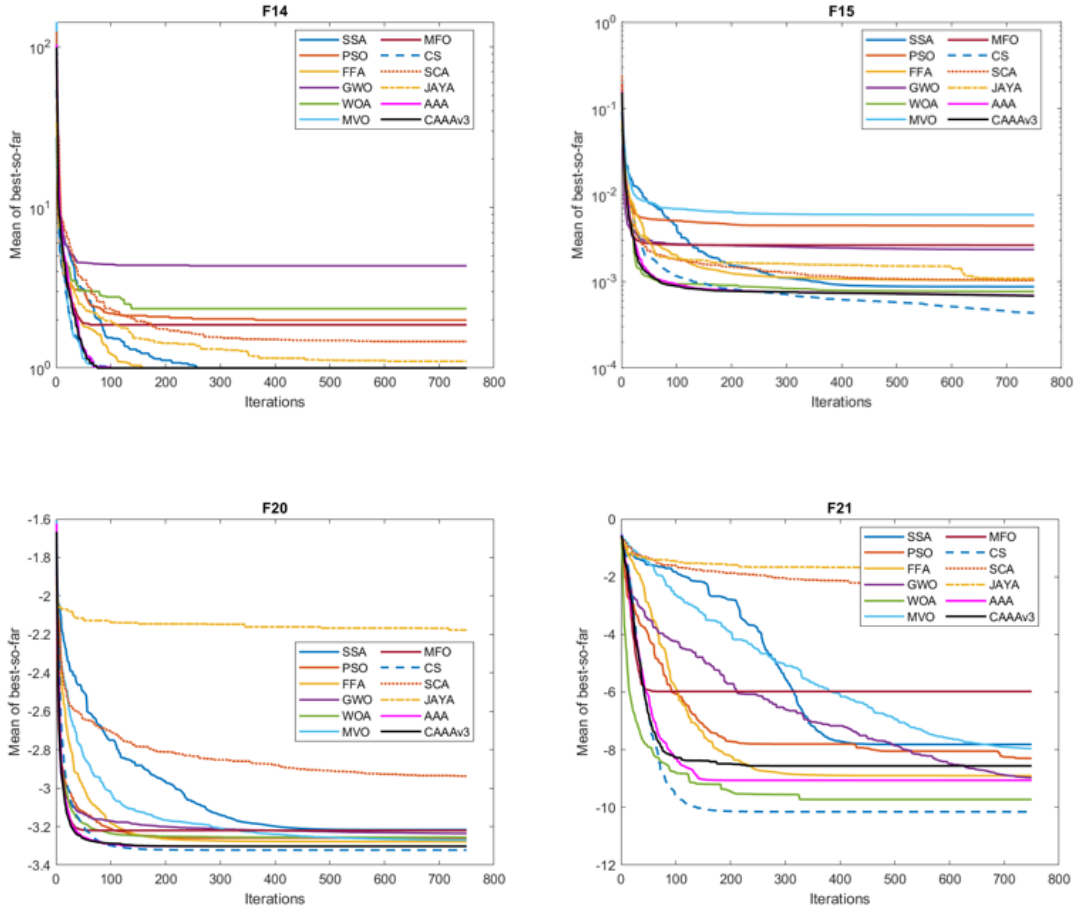
Şekil 4.1’de gösterilen dört yakınsama eğrisi tek modlu fonksiyonlara (F2, F4, F24 ve F28), Şekil 4.2’de gösterilen dört yakınsama eğrisi çok modlu fonksiyonlara (F9, F11, F13 ve F27) ve Şekil 4.3’te gösterilen dört tane yakınsama eğrisi de sabit boyutlu çok modlu fonksiyonlara (F14, F15, F20 ve F21) aittir.



Şekil 4.1. Tek modlu fonksiyonlardan F2, F4, F24 ve F28'e ait yakınsama eğrileri



Şekil 4.2. Çok modlu fonksiyonlardan F9, F11, F13 ve F27'e ait yakınsama eğrileri



Şekil 4.3. Sabit boyutlu çok modlu fonksiyonlardan F14, F15, F20 ve F21'e ait yakınsama eğrileri

Yakınsama eğrileri analiz edildiğine tek modlu fonksiyonlarda GWO ve WOA algoritmasının üstün performansı göze çarpmaktadır. Bunun temel nedeni tek bir global minimum noktasına sahip problemlerde GWO ve WOA yaklaşımlarının etkili sömürü kabiliyeti sayesinde hızlı yakınsama karakteristiklerine sahip olmalarıdır.

Çok modlu fonksiyonlar birden fazla yerel minimum noktasına sahip olduğu için çözülmesi tek modlu fonksiyonlardan daha zordur. Çok modlu fonksiyonların yakınsama grafiklerine bakıldığında CAAAv3 ve AAA yöntemleri etkili keşif ve sömürü dengesi sayesinde başarılı performanslar ortaya çıkardığı gözlemlenmektedir. CAAAv3 yaklaşımı kaotik yapısından dolayı AAA yönteminde daha başarılı bir yakınsama performansı göstermektedir. Sabit boyutlu çok modlu fonksiyonlarda ise CS yöntemi yinelemelerin erken döneminde hızlı bir şekilde etkili yakınsama göstermektedir. CAAAv3 yöntemi de bu problemlerde rekabetçi bir yakınsama performansı ortaya koymaktadır.

4.2. Gerçek Dünya Optimizasyon Problemlerinde Kaotik Yapay Alg Algoritmasının Performans Analizi

Bu bölümde tez kapsamında geliştirilen CAAAv3 yaklaşımı iyi bilinen üç mühendislik tasarımı ve ESA GTOP'daki sekiz yörünge tasarım problemi olmak üzere toplam on bir farklı gerçek dünya optimizasyon problemi üzerinde performans analizi yapılmıştır. Karşılaştırma sürecinde on farklı karakteristik yapıya sahip metasezgisel optimizasyon algoritması kullanılmıştır. Bu algoritmalar PSO (Kennedy ve Eberhart, 1995), GWO (Mirjalili ve ark., 2014a), MVO (Aljarah ve ark., 2020b), Sinüs Kosinüs Algoritması (SCA) (Mirjalili, 2016), FFA (Yang, 2010a), CS (Yang ve Deb, 2009), Jaya Algoritması (JAYA) (Rao, 2016), Güve Alev Optimizasyonu (MFO) (Mirjalili, 2015), SSA (Mirjalili ve ark., 2017) ve WOA (Mirjalili ve Lewis, 2016)'dır. Karşılaştırma sürecinde orijinal AAA ile de kıyaslanmıştır. Gerçek dünya problemleriyle gerçekleştirilen deneysel süreçte de algoritmaların parametre değerleri Çizelge 4.2'de gösterilen değerlerdir. Bu parametreler ilgili algoritma yazarlarının kendi yayınlarında tavsiye ettiği değerlerdir (Faris ve ark., 2016).

Bu problemleri çözmek için tasarlanan deneysel çalışmada kullanılan parametreler, benchmark fonksiyonlarda kullanılan parametrelerle aynıdır. Benzer şekilde otuz bağımsız çalıştırma gerçekleştirilerek ortalaması alınmıştır. Maksimum uygunluk değerlendirme sayısı olarak 50,000 (FEs), popülasyondaki birey sayısı olarak 40 tercih edilmiştir (Gandomi ve ark., 2013a; Kaur ve Arora, 2018; Kohli ve Arora, 2018; Sayed ve ark., 2018). Elde edilen sonuçlar literatürdeki on popüler algoritma ve orijinal AAA ile karşılaştırılmıştır.

Bu problemlerdeki değişkenlerin amaç fonksiyonunun kısıt sınırlarını ihlal etmesi durumunda adil bir şekilde değerlendirebilmek için bir ceza (penalty) fonksiyonuna ihtiyaç vardır. Literatürde tavlama, birlikte evrim (co-evolutionary), statik, uyarlanabilir, dinamik ve ölüm cezası gibi çeşitli ceza fonksiyonları bulunmaktadır. Bu tez çalışmasında da literatürdeki çalışmalara benzer şekilde en çok tercih edilen ölüm cezası (death penalty) yöntemi kullanılmıştır (Coello, 2002).

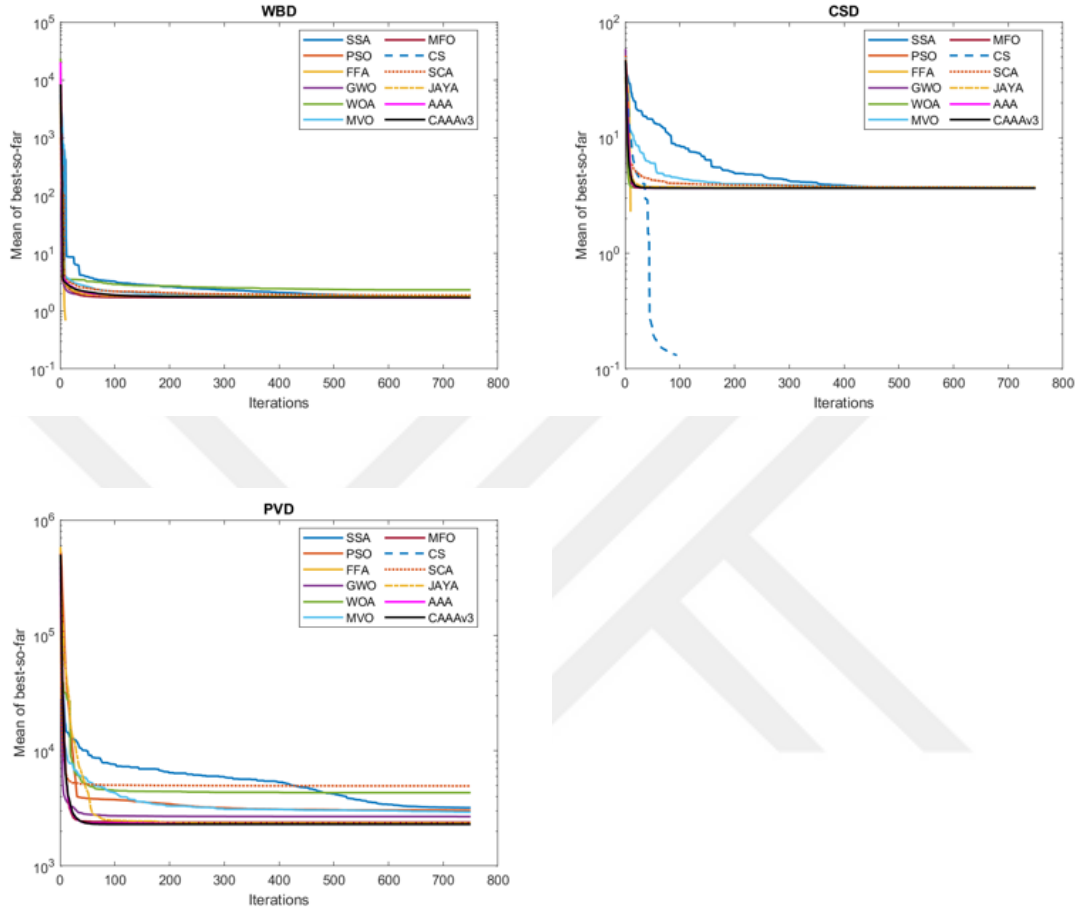
Çizelge 4.6' de bütün yöntemlerin her bir problemde elde ettiği otuz bağımsız çalışma sonuçlarının ortalaması ve standart sapması sunulmuştur. Ayrıca tablonun en alt bölümünde yöntemlerin elde ettiği en iyi sonuç sayısı (Best Count) ve Friedman sıralama testi sonuçlarının bulunduğu bir özet bilgi kısmı sunulmaktadır.

Çizelge 4.6. CAAAv3 ile literatürdeki algoritmaların gerçek dünya optimizasyon problemlerindeki sonuçları

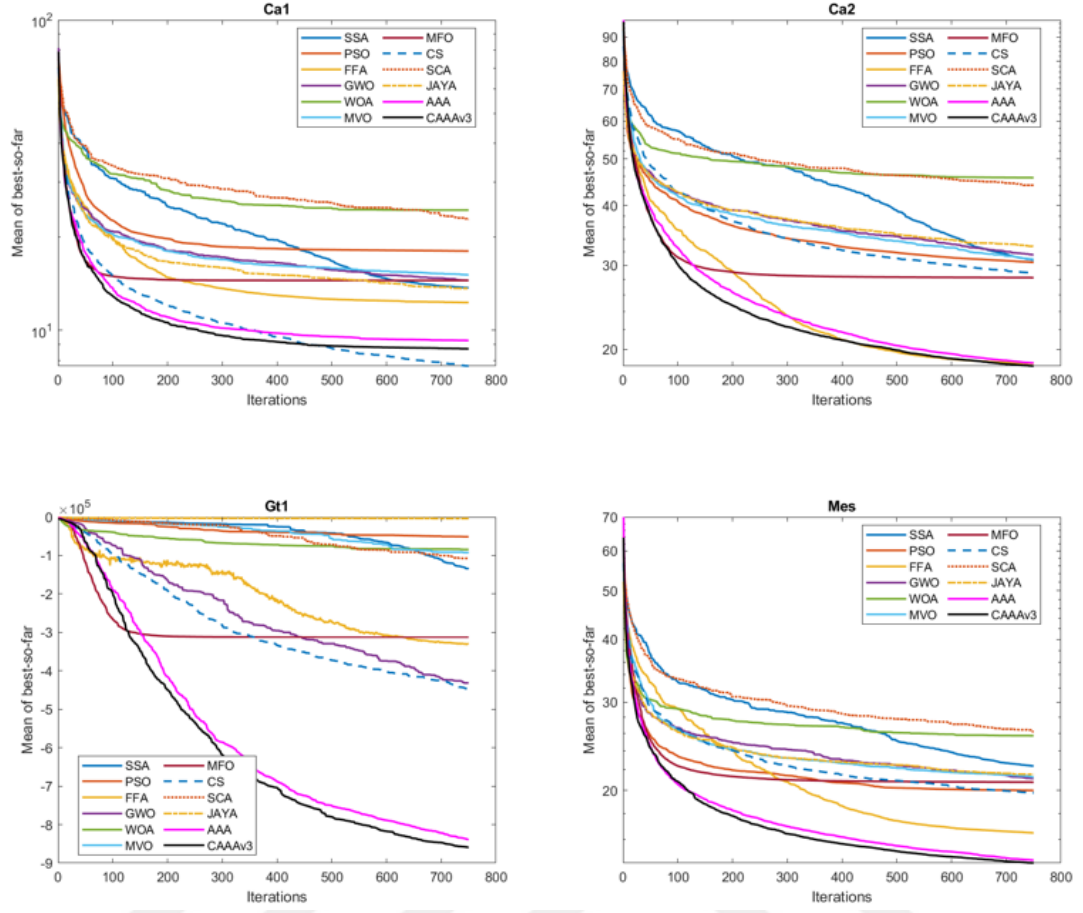
No	Problem		SSA	PSO	FFA	GWO	WOA	MVO	MFO	CS	SCA	JAYA	AAA	CAAAv3
P1	WBD	mean	1.77E+00	1.72E+00	-3.36E+05	1.74E+00	2.33E+00	1.77E+00	1.72E+00	-6.97E+01	1.88E+00	1.80E+00	1.70E+00	1.72E+00
		std	9.88E-02	4.24E-02	2.64E+05	1.53E-02	4.57E-01	3.16E-02	8.34E-02	2.90E+01	3.04E-02	2.12E-01	3.01E-02	4.18E-02
P2	CSD	mean	3.68E+00	3.66E+00	-9.58E+06	3.67E+00	3.69E+00	3.71E+00	3.67E+00	-1.04E+01	3.72E+00	3.70E+00	3.66E+00	3.66E+00
		std	2.08E-02	1.42E-02	2.51E+07	2.63E-03	2.93E-02	3.28E-02	1.98E-02	1.86E+01	4.74E-02	1.25E-01	2.06E-15	1.57E-15
P3	PVD	mean	3.20E+03	3.03E+03	-2.33E+11	2.68E+03	4.32E+03	2.94E+03	2.39E+03	-1.62E+07	4.95E+03	2.35E+03	2.30E+03	2.30E+03
		std	6.28E+02	1.42E+03	3.59E+10	1.14E+03	1.81E+03	6.12E+02	3.35E+02	3.72E+06	1.72E+03	2.42E+02	8.40E-13	8.53E-13
P4	Ca1	mean	1.37E+01	1.80E+01	1.23E+01	1.45E+01	2.44E+01	1.51E+01	1.45E+01	7.68E+00	2.28E+01	1.37E+01	9.29E+00	8.73E+00
		std	3.84E+00	7.09E+00	3.76E+00	2.04E+00	7.98E+00	4.47E+00	5.70E+00	1.72E+00	6.77E+00	3.60E+00	3.24E+00	3.17E+00
P5	Ca2	mean	3.07E+01	3.04E+01	1.87E+01	3.15E+01	4.57E+01	3.09E+01	2.83E+01	2.89E+01	4.41E+01	3.28E+01	1.87E+01	1.85E+01
		std	4.53E+00	4.60E+00	3.45E+00	3.44E+00	7.26E+00	4.83E+00	5.53E+00	2.15E+00	4.87E+00	5.20E+00	3.43E+00	3.38E+00
P6	Gt1	mean	-1.35E+05	-5.13E+04	-3.30E+05	-4.32E+05	-8.44E+04	-9.34E+04	-3.13E+05	-4.46E+05	-1.08E+05	-3.77E+03	-8.38E+05	-8.59E+05
		std	1.64E+05	1.12E+05	2.97E+05	2.40E+05	1.04E+05	9.66E+04	3.12E+05	1.58E+05	8.26E+04	6.16E+03	2.09E+05	1.88E+05
P7	Mes	mean	2.24E+01	2.00E+01	1.65E+01	2.11E+01	2.57E+01	2.13E+01	2.08E+01	1.98E+01	2.62E+01	2.15E+01	1.45E+01	1.43E+01
		std	3.37E+00	3.04E+00	4.26E+00	2.10E+00	3.25E+00	3.21E+00	4.83E+00	1.29E+00	2.01E+00	2.89E+00	1.33E+00	1.50E+00
P8	Mef	mean	3.24E+01	2.65E+01	2.25E+01	3.63E+01	4.38E+01	3.79E+01	2.66E+01	2.76E+01	5.54E+01	3.21E+01	1.74E+01	1.70E+01
		std	7.93E+00	5.19E+00	5.28E+00	4.84E+00	9.25E+00	6.25E+00	6.37E+00	2.93E+00	6.23E+00	9.14E+00	3.52E+00	2.91E+00
P9	Sag	mean	2.32E+02	2.67E+02	1.68E+02	2.39E+02	2.38E+02	2.41E+02	2.38E+02	2.30E+02	2.30E+02	2.35E+02	1.89E+02	1.99E+02
		std	9.67E+00	3.01E+01	7.23E+01	2.68E+01	1.98E+01	2.47E+01	1.92E+01	2.00E+01	2.14E+00	1.58E+01	4.74E+01	4.78E+01
P10	Tan	mean	-3.58E+00	-1.47E+01	-2.80E+01	-1.16E+01	-6.65E-01	-2.59E+00	-2.10E+01	-5.02E+00	-1.74E-01	-3.27E+00	-1.62E+02	-1.80E+02
		std	7.64E+00	3.51E+01	6.32E+01	1.96E+01	1.71E+00	2.40E+00	4.66E+01	3.97E+00	1.69E-01	4.62E+00	1.37E+02	1.23E+02
P11	Ros	mean	1.85E+01	1.51E+01	1.04E+01	1.85E+01	2.66E+01	2.04E+01	1.45E+01	1.55E+01	2.95E+01	1.66E+01	7.74E+00	6.69E+00
		std	3.99E+00	3.02E+00	3.14E+00	3.11E+00	4.10E+00	3.13E+00	4.46E+00	1.85E+00	2.47E+00	4.77E+00	2.22E+00	1.84E+00
En iyi sonuç sayısı			0	0	4	0	0	0	0	1	0	0	0	6
Friedman sıralama değerleri			8.00	7.00	2.45	7.27	10.73	9.27	5.91	3.91	10.64	8.27	2.50	2.05
Genel sıralama			8	6	2	7	12	10	5	4	11	9	3	1
P-Değeri			8.23E-15											

Çizelge 4.6'daki sonuçlar incelendiğinde tez kapsamında geliştirilen CAAAv3 yöntemi on bir gerçek dünya probleminin altısında bir diğer tabirle problem setinin yaklaşık %55'inde en iyi sonucu elde etmiştir. FFA yöntemi ise dört problemde en iyi değeri elde ederek rekabetçi bir yaklaşım göstermiştir. Friedman sıralama sonuçları incelendiğinde 2.05 skoru ile CAAAv3 yöntemi gerçek dünya optimizasyon problemlerinde en başarılı performansı göstererek birinci sırada yer almıştır. Friedman sıralama sonucunda elde edilen p-Value değerinin 0.05 değerinin altında olması sıralanan algoritmalarda istatistiksel açıdan anlamlı bir farklılığın olduğunu göstermektedir. Algoritmaların problemlerdeki performansı detaylı olarak analiz edildiğinde FFA yöntemi üç klasik mühendislik tasarım probleminin tamamında en iyi değeri elde etmektedir. Yörünge tasarım problemlerinde ise CAAAv3 yaklaşımının aşikâr bir üstünlüğü gözükmemektedir. Yörünge tasarım problemleri hem değişken sayısının fazla olması hem de karmaşıklığının yüksek olması nedeniyle mühendislik tasarım problemlerinden daha zor çözülmektedir.

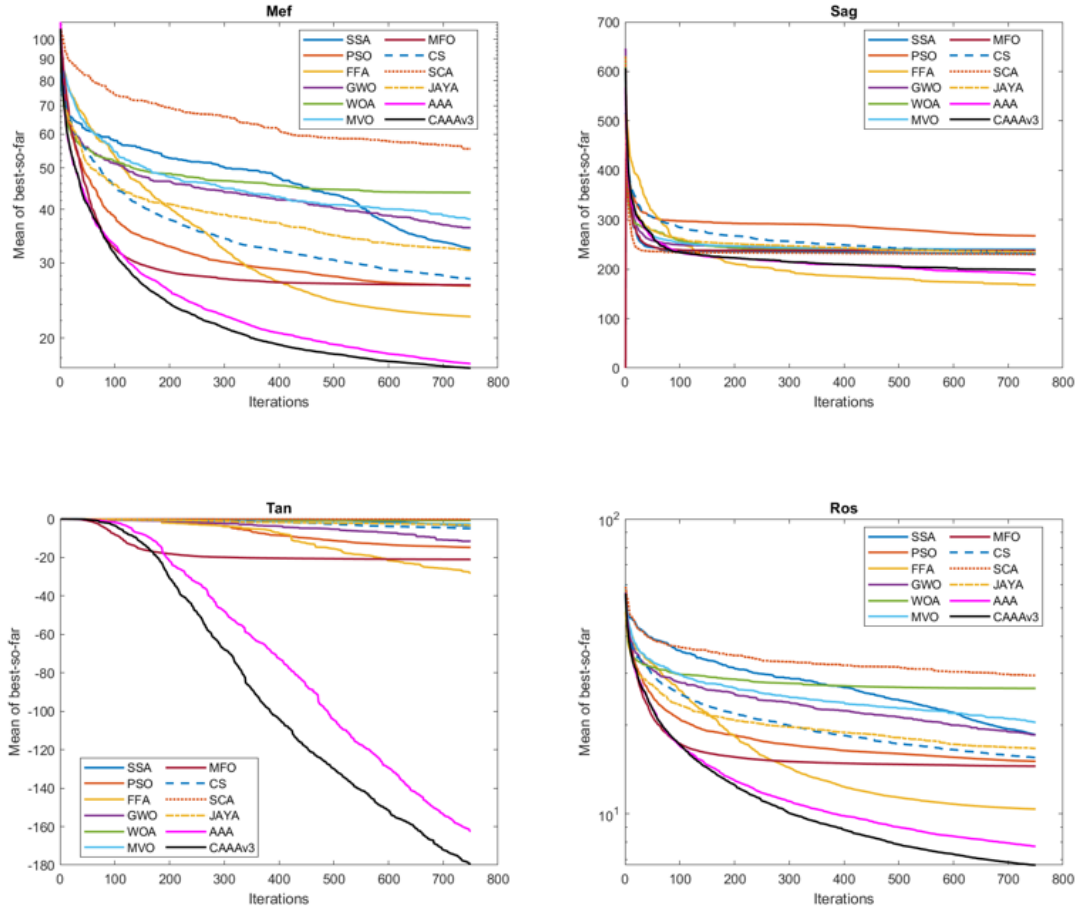
Geliştirilen CAAAv3 ve kıyaslanan diğer algoritmaların gerçek dünya problemleri üzerindeki optimum çözüme yaklaşma davranışlarını analiz edebilmek için yakınsama eğrileri Şekil 4.4-4.6 de gösterilmektedir.



Şekil 4.4. Algoritmaların mühendislik tasarım problemlerine ait yakınsama eğrileri



Şekil 4.5. Algoritmaların yörünge tasarım problemlerinden Cassini 1, Cassini 2, Gt1 ve Messenger'a ait yakınsama eğrileri



Şekil 4.6. Algoritmaların yörünge tasarım problemlerinden Messenger (full), Sagas, Tandem ve Rosetta'ya ait yakınsama eğrileri

Algoritmaların yakınsama eğrileri incelendiğinde mühendislik tasarım problemlerinde FFA yöntemi yinelemelerin erken döneminde hızlı bir yakınsama ile optimum çözüme ulaştığı görülmektedir. Yörünge tasarım problemlerinde ise genel olarak CAAAv3 ve AAA yaklaşımının etkin bir keşif ve sömürü süreci dengesinin kurulması sayesinde ideal çözüme diğer yöntemlerden daha başarılı bir şekilde yakınsamaktadırlar. Kaotik yapısından dolayı CAAAv3 yönteminin AAA yönteminden daha başarılı bir performans ortaya çıkardığı görülmektedir.

4.3. Kümeleme Analizinde Kaotik Yapay Alg Algoritmasının Performans Değerlendirmesi

Bu bölümde, öncelikle deneysel çalışmada kullanılan kümeleme veri setleri ve değerlendirme metrikleri sunulacaktır. Daha sonra bu veri setleri üzerinde popüler algoritmaların performansı ilgili metriklerle analiz edilerek kıyaslanacaktır. Tez

kapsamında geliştirilen kaotik AAA yaklaşımının kümeleme analizinde kullanımı için önerilen yaklaşım Bölüm 3.4'te detaylandırılmıştır.

Geliştirilen yöntemin değerlendirilmesinde farklı özelliklere ve farklı zorluk seviyelerine sahip yaygın olarak bilinen yirmi beş veri kümesi kullanılmıştır. Bu veri kümeleri Aggregation, Aniso, Appendicitis, Balance, Banknote, Blood, Diagnosis II, Ecoli, Flame, Heart, Iris, Iris2D, Ionosphere, Jain, Liver, Moons, Mouse, Pathbased, Smiley, Sonar, Vary Density, Vertebral 2, Vertebral 3, WDBC ve Wine'dır. Veri kümeleri, UCI makine öğrenmesi deposu² (Dhote ve ark., 2013), (Dua ve Karra Taniskidou, 2017), Doğu Finlandiya Üniversitesi Bilgisayar Bölümü (EUF)³, ELKI⁴, KEEL⁵ ve sci-kit learn⁶ kütüphanelerinden alınmıştır. Bu veri kümeleri ve özellikleri Çizelge 4.7'de gösterilmektedir.

Önerilen kaotik AAA tabanlı kümeleme yaklaşımını kıyaslamak için on farklı karakteristik yapıya sahip metasezgisel tabanlı kümeleme algoritması kullanılmıştır. Bu kümeleme yöntemleri GA (John, 1992), PSO (Kennedy ve Eberhart, 1995), SSA (Mirjalili ve ark., 2017), FFA (Yang, 2010a), GWO (Mirjalili ve ark., 2014a), WOA (Mirjalili ve Lewis, 2016), MVO (Aljarah ve ark., 2020b), MFO (Mirjalili, 2015), BAT (Yang, 2010b) ve CS (Yang ve Deb, 2009) algoritmalarına dayanmaktadır.

² <https://archive.ics.uci.edu/ml/> .

³ <http://cs.uef.fi/sipu/datasets/> .

⁴ <https://elki-project.github.io/datasets/> .

⁵ <https://sci2s.ugr.es/keel/datasets.php> .

⁶ <http://scikit-learn.org/stable/datasets/index.html> .

Çizelge 4.7. Kümeleme Sürecinde Kullanılan Veri Kümelerinin Özellikleri

Veri seti	Örnek Sayısı	Küme Sayısı (k)	Öznitelik Sayısı	Kaynağı
Aggregation	788	7	2	UEF
Aniso	1500	3	2	scikit-learn
Appendicitis	106	2	7	KEEL
Balance	625	3	4	UCI
Banknote	1372	2	4	UCI
Blood	748	2	4	UCI
Diagnosis II	120	2	6	UCI
Ecoli	327	5	7	UCI
Flame	240	2	2	UEF
Heart	270	2	13	UCI
Iris	150	3	4	UCI
Iris 2D	150	3	2	UCI
Ionosphere	351	2	34	UCI
Jain	373	2	2	UEF
Liver	345	2	7	UCI
Moons	1500	2	2	scikit-learn
Mouse	490	3	2	ELKI
Pathbased	300	3	2	UEF
Smiley	500	4	2	KEEL
Sonar	208	2	60	UCI
Vary Density	150	3	2	ELKI
Vertebral 2	310	2	6	UCI
Vertebral 3	310	3	6	UCI
WDBC	569	2	30	UCI
Wine	178	3	13	UCI

Algoritmaların adil bir şekilde kıyaslanabilmesi için aynı platform üzerinde deneysel çalışmalar gerçekleştirilmiştir. Tüm deneysel çalışmalar EvoCluster kütüphanesinden faydalanılarak yapılmıştır (Qaddoura ve ark., 2020). EvoCluster, Python yazılım dilinde geliştirilen popüler ve en yeni metasezgisel tabanlı kümeleme yaklaşımlarını barındıran açık kaynaklı bir yapıdır (framework). Algoritmaların başlangıç parametreleri olarak EvoCluster'ın varsayılan parametre değerleri kullanılmıştır (Qaddoura ve ark., 2021). Çizelge 4.8’de gösterilen bu parametreler aynı zamanda birçok çalışmada kullanılan kabul görmüş değerlerdir (Zhang ve ark., 2010; Karaboga ve Ozturk,

2011; Senthilnath ve ark., 2011; Kumar ve ark., 2017; Nasiri ve Khiyabani, 2018; Rahnama ve Gharehchopogh, 2020).

Tüm veri kümeleri [0, 1] aralığına ölçeklenecek şekilde minimum maksimum (min-max) normalizasyonu ile normalleştirilmiştir. Deneysel sürecin tamamında rastgele başlangıç parametreleriyle bağımsız otuz çalışma yapılarak ortalaması alınmıştır. Bir çok çalışmada olduğu gibi popülasyondaki birey sayısı 40 olarak seçilmiştir (Kaur ve Kumar, 2021). Algoritmalar 30.000 FEs çalıştırılmıştır.

Çizelge 4.8. Deneysel çalışmada kıyaslanan algoritmaların parametre değerleri

Algoritmalar	Parametreleri	Değerleri
AAA	Δ Share force	2
	e Energy loss	0.3
	A_p Adaptation	0.2
PSO	$c1, c2$	2, 2
	$W_{inertia}$ weight	2
	W_{max}, W_{min}	0.9, 0.4
GA	Crossover rate	0.7
	Mutation rate	0.2
BAT	A (Loudness)	0.5
	r (Pulse rate)	0.5
	Q	[0, 2]
FFA	alpha (Randomness)	0.5
	β_{min} (minimum value of beta)	0.2
	Gama (absorption coefficient)	1
GWO	alpha	2
WOA	alpha	2
MVO	wep_{max}	1
	wep_{min}	0.2
MFO	b	1
CS	Pa (Discovery rate of alien eggs)	0.25

Algoritmaların kümeleme performansını değerlendirmek için uygunluk fonksiyonu olarak küme içi mesafelerin toplamı metriğini ifade eden SSE değeri kullanılmıştır. Bu metrik aynı küme içerisindeki veri noktaları ile kümenin merkezi arasındaki mesafelerin karelerinin toplamından oluşmaktadır. SSE metriği Bölüm 3.3'te

detaylandırılmış ve Denklem 3.28'de $dist(x_i, I_j)$ olarak tanımlanmıştır. Bu metriğin değerinin düşük olması, kümeleme kalitesinin yüksek olduğu anlamına gelmektedir.

Deneyisel sonuçlar SSE metriğinin yanı sıra yerel minimumlardan kaçma kabiliyeti, yakınsama hızı ve bir sonraki bölümde detaylandırılan değerlendirme metrikleri bakımından da analiz edilmiştir.

4.3.1. Kümele Analizindeki Değerlendirme Metrikleri

Kümeleme analizinde sonuçların değerlendirilmesinde ve yorumlanmasında sıklıkla kullanılan çeşitli metrikler vardır. Bu değerlendirme metrikleri SSE, Saflık (Purity) ve V- ölçümü (V-Measure, VM)'dür. Literatürde iyi bilinen ve bu çalışmada öncelikli olarak kullanılan değerlendirme metriği SSE'dir (Turkoglu ve ark., 2022a). SSE'ye ek olarak, daha ayrıntılı bir karşılaştırma için Saflık ve V-ölçümü (VM) değerlendirme metrikleri de kullanılmıştır. Metriklerde veri noktaları N ile, noktaların gerçek sınıfları T ile ve bu noktaların tahmin edilen kümeleri P harfi ile temsil edilmektedir. Bu metrikler tezin devamında açıklanmıştır.

Saflık: Saflık, kümelerin ne ölçüde tek bir sınıf içerdiğinin ölçüsüdür. Her küme içerisinde en fazla aynı sınıfa sahip olan veri noktalarının toplam veri sayısına oranıdır. Denklem 4.1 ile hesaplanmaktadır. Saflık metriğinin derecesi arttıkça doğru kümeleme oranı yani başarısı artmaktadır (Aljarah ve Ludwig, 2013; Qaddoura ve ark., 2020).

$$Purity = \frac{1}{N} \sum_{j=1}^k \max_i (|T_i \cap P_j|) \quad (4.1)$$

Denklemlerde P_j , j kümesinin tüm noktalarını, k küme sayısını ve T_i , i kümesindeki veri noktalarını temsil etmektedir.

V-measure: VM, homojenlik ve tamlık kriterlerinin ne kadar başarılı bir şekilde karşılandığını hesaplayan, ölçen entropi tabanlı bir ölçüm metriğidir. VM, Homojenlik Skoru (HS) ve Tamlık Skoru (CS) metriklerinin harmonik ortalamasıdır. Bu metrik danışmanlı öğrenme problemlerini değerlendirilirken kullanılan F-measure metriğine benzemektedir. F-measure metriği de kesinlik ve duyarlılık metriklerinin harmonik ortalamasıdır (Rosenberg ve Hirschberg, 2007).

VM metriği Denklem 4.2’de gösterilmektedir. VM metriğinin yüksek olması daha etkin kümeleme yapıldığını gösterir.

$$V - measure = 2 \cdot \frac{HS \cdot CS}{HS + CS} \quad (4.2)$$

Homojenlik Skoru: Kümeleme analizinde homojenlik skoru, her bir kümenin aynı sınıf etiketine ait veri noktalarına sahip olma oranını göstermektedir. Homojenlik, Saflık metriğine benzemektedir ancak homojenlik hesaplanırken karar ağaçlarında olduğu gibi entropi değeri kullanılır.

$$Homojenlik\ Skoru = 1 - \frac{H(T/P)}{H(T)} \quad (4.3)$$

Denklemden $H(T)$ sınıf entropisini, $H(T/P)$ koşullu sınıf entropisini temsil etmektedir. $H(T)$ ve $H(T/P)$ Denklem 4.4 ve 4.5'deki gibi hesaplanır (Rosenberg ve Hirschberg, 2007):

$$H(T) = - \sum_{t=1}^{|T|} \frac{n_t}{N} \cdot \log\left(\frac{n_t}{N}\right) \quad (4.4)$$

$$H(T/P) = - \sum_{p=1}^{|P|} \sum_{t=1}^{|T|} \frac{n_{pt}}{N} \cdot \log\left(\frac{n_{pt}}{n_p}\right) \quad (4.5)$$

Denklemlerde n_t ve n_p , sırasıyla t sınıfındaki veri noktalarının sayısını ve tahmin edilen p kümesindeki veri noktalarının sayısını göstermektedir. n_{pt} ise tahmin edilen p kümesine kümelenen gerçek t sınıfındaki veri noktalarının sayısını temsil etmektedir.

Tamlık Skoru: Kümeleme analizinde tamlık skoru aynı sınıfa ait tüm veri noktalarının aynı kümede ne kadar kümelendiğinin ölçüsüdür. Tamlık skoru bir anlamda kümelemenin mükemmelliğe yakınlığını ifade eder. Denklem 4.6 ile hesaplanır.

$$Tamlık Skoru = 1 - \frac{H(P/T)}{H(P)} \quad (4.6)$$

Denklemden $H(P)$ küme entropisini, $H(P/T)$ koşullu küme entropisini temsil etmektedir. $H(P)$ ve $H(P/T)$ Denklem 4.7 ve 4.8'deki gibi hesaplanır (Rosenberg ve Hirschberg, 2007).

$$H(P) = - \sum_{p=1}^{|P|} \frac{n_p}{N} \cdot \log\left(\frac{n_p}{N}\right) \quad (4.7)$$

$$H(P/T) = - \sum_{t=1}^{|T|} \sum_{p=1}^{|P|} \frac{n_{pt}}{N} \cdot \log\left(\frac{n_{pt}}{n_t}\right) \quad (4.8)$$

4.3.2. Performans Değerlendirmesi

Bu bölümde geliştirilen yöntemin performansı değerlendirilmektedir. Farklı karmaşıklık seviyelerindeki (özellikler ve örnekler bakımından) Çizelge 4.7'deki veri setleri üzerinde karşılaştırılan tüm algoritmaların test sonuçları Çizelge 4.9 – 4.11'de sunulmaktadır. Çizelge 4.9, Çizelge 4.10 ve Çizelge 4.11 sırasıyla SSE, Saflık ve VM sonuçlarını göstermektedirler. Bu çizelgelerde, tüm yöntemler için otuz bağımsız çalışmanın ortalama değerleri ve standart sapmaları verilmiştir. Çizelgelerin son satırları, en iyi sonuç sayısı (best count), ortalama sıralama ve ortalama sıralamanın değerlerini göstermektedir.

Çizelge 4.9'daki sonuçlar, her bir veri seti için azaltılması gereken küme içi mesafelerin karelerinin toplamını yani SSE metriğini temsil etmektedir.

Çizelge 4.9. Kümeleme algoritmalarının SSE metriği bakımından otuz bağımsız çalışma sonuçları

Veri seti		CS	SSA	PSO	GA	BAT	FFA	GWO	WOA	MVO	MFO	CAAAv3
aggregation	mean	10.917	11.287	11.531	11.535	11.345	11.218	11.098	12.382	11.044	11.725	10.922
	±std	1.03	0.53	0.64	0.33	0.51	0.46	0.33	1.11	0.34	0.34	0.04
aniso	mean	108.429	591.670	592.072	599.476	71.640	71.640	591.816	597.249	591.670	591.804	591.670
	±std	7.58	0.00	1.22	4.40	0.00	0.00	0.00	8.65	0.00	0.73	0.00
appendicitis	mean	17.430	17.431	22.424	17.744	17.438	18.178	18.044	18.596	17.430	18.947	17.430
	±std	0.00	0.00	4.44	0.19	0.01	2.29	0.87	2.15	0.00	2.59	0.00
balance	mean	138.901	138.897	139.031	139.775	139.205	138.900	138.896	140.037	138.898	139.031	138.899
	±std	0.00	0.01	0.33	0.53	0.19	0.01	0.01	0.87	0.01	0.33	0.01
banknote	mean	1188.41	1730.94	1748.66	1747.16	785.570	785.575	1731.00	1751.65	1730.94	1730.94	1730.94
	±std	156.8	0.00	22.0	8.51	0.00	0.00	0.04	30.7	0.00	0.00	0.00
blood	mean	36.800	38.038	46.121	37.177	36.800	38.033	37.124	37.650	36.800	36.938	36.800
	±std	0.00	6.78	10.2	0.20	0.00	6.75	1.02	3.25	0.00	0.52	0.00
Diagnosis II	mean	105.126	106.192	107.354	106.980	108.229	106.754	106.035	107.776	105.758	106.712	105.124
	±std	0.04	1.75	1.87	1.66	2.42	1.46	1.03	2.23	0.87	1.54	0.02
ecoli	mean	19.765	18.854	34.336	28.650	17.897	23.434	25.955	26.685	19.530	28.147	17.160
	±std	0.39	1.20	7.89	1.80	0.73	3.50	4.85	6.57	1.69	7.12	0.22
flame	mean	7.970	7.970	7.970	7.974	7.970	7.970	7.971	7.971	7.970	7.970	7.970
	±std	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
heart	mean	245.686	258.823	300.785	253.683	251.003	255.254	255.645	278.660	249.017	272.691	245.474
	±std	0.17	17.9	23.4	2.48	11.4	18.6	14.9	26.9	8.27	17.1	0.01
iris	mean	2.830	2.930	4.890	3.237	3.031	3.176	3.132	3.118	2.830	3.341	2.830
	±std	0.00	0.55	1.36	0.18	0.76	0.94	0.69	0.86	1.69	0.65	0.00
iris2D	mean	1.410	1.410	1.458	1.438	1.410	1.410	1.442	1.432	1.410	1.464	1.410
	±std	0.39	0.00	0.14	0.14	0.00	0.00	0.12	0.08	0.00	0.17	0.00
ionosphere	mean	2585.22	2735.21	3344.60	2740.93	2560.55	2908.32	2497.89	2595.64	2480.60	2787.60	2446.78
	±std	195	297	254	68	177	362	40.7	106	114	337	0.15
jain	mean	18.200	18.200	18.200	18.208	18.201	18.200	18.200	18.202	18.200	18.200	18.200
	±std	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
liver	mean	22.920	23.641	31.080	23.583	22.920	24.743	24.200	28.871	22.920	24.484	22.920
	±std	0.00	2.24	4.56	0.33	0.00	3.40	2.56	7.18	0.00	2.71	0.00
moons	mean	232.524	246.470	246.470	247.147	231.660	231.660	246.476	246.485	246.470	246.470	246.470
	±std	0.00	0.00	0.00	0.56	0.00	0.00	0.00	0.03	0.00	0.00	0.00
mouse	mean	10.350	10.350	10.412	10.418	10.350	10.350	10.350	10.356	10.350	10.350	10.350
	±std	0.00	0.00	0.05	0.01	0.00	0.00	0.00	0.43	0.00	0.07	0.00
pathbased	mean	8.560	8.560	8.5780	8.582	8.560	8.560	8.560	8.648	8.560	8.596	8.560
	±std	0.00	0.00	0.05	0.01	0.00	0.00	0.00	0.43	0.00	0.07	0.00
smiley	mean	11.7106	11.887	11.965	12.065	12.060	11.855	11.769	12.142	11.711	11.945	11.710
	±std	0.00	0.29	0.27	0.30	0.35	0.29	0.23	0.35	0.00	0.32	0.00
sonar	mean	505.954	502.782	873.480	660.440	431.301	508.651	588.701	572.399	505.286	719.584	442.356
	±std	6.45	8.29	64.3	37.2	19.9	0.91	29.3	64.0	4.00	80.2	6.03
VaryDensity	mean	2.480	2.859	2.480	2.492	2.953	2.480	2.502	2.555	2.480	2.480	2.480
	±std	0.00	0.58	0.00	0.01	0.63	0.00	0.10	0.31	0.00	0.00	0.00
Vertebral2	mean	19.210	24.258	29.583	19.956	23.320	23.867	22.625	25.799	21.308	23.108	19.210
	±std	0.00	6.23	4.95	0.38	5.91	6.22	5.63	6.53	4.77	5.60	0.00
Vertebral3	mean	15.224	17.674	24.453	18.223	17.255	19.184	17.816	21.917	17.060	20.324	14.973
	±std	0.34	1.24	4.90	0.83	1.15	4.29	2.43	5.51	1.44	4.18	0.00
wdbc	mean	214.528	263.086	562.221	319.670	160.632	272.005	314.702	300.668	252.347	403.135	157.759
	±std	13.8	12.6	99.5	31.6	20.8	0.50	30.8	50.0	32.2	83.3	0.60
wine	mean	28.457	33.980	57.212	43.001	28.694	41.637	41.619	46.482	31.823	51.999	27.352
	±std	0.63	6.49	8.78	4.14	2.80	8.98	9.32	14.2	3.53	12.5	0.03
En iyi sonuç sayısı		12	5	3	0	10	8	4	0	10	4	19
Friedman sıralama değerleri		3.16	5.42	9.26	8.44	4.74	5.7	6.16	9.2	3.6	7.62	2.7
Genel sıralama		2	5	11	9	4	6	7	10	3	8	1

Tablo 4.9'daki SSE sonuçlarına göre tez kapsamında geliştirilen kaotik AAA tabanlı kümele yaklaşımı, yirmi beş veri kümesinden on dokuz tanesinde bir diğer tabirle veri kümelerinin %76'sında en iyi değerleri elde etmiştir. Ayrıca CS tabanlı kümele yaklaşımı, on iki veri kümesinde en iyi değeri elde ederek dikkat çekmektedir. Ancak GA ve WOA tabanlı yaklaşımlar hiçbir veri setinde en iyi değere ulaşamamışlardır. Sonuçlar ortalama sıralama açısından değerlendirildiğinde SSE metriği bakımından kaotik AAA tabanlı kümeleme yaklaşımı kıyaslanan diğer algoritmalarından daha iyi performans elde etmiştir.

Çizelge 4.10'da algoritmaların veri kümeleri üzerinde saflık metriğine göre otuz bağımsız çalışmada elde ettiği sonuçların ortalama ve standart sapma değerleri gösterilmektedir.

Çizelge 4.10'daki sonuçlara göre kaotik AAA tabanlı kümele yaklaşımı, yirmi beş veri kümesinden on beş tanesinde bir diğer tabirle veri kümelerinin %60'ında en iyi saflık değerini elde etmiştir. Saflık veri noktalarının sınıf etiketleri bilgisinden yola çıkarak aynı küme içerisinde doğru etiketle kümelenmiş örnek sayısı olarak düşünülebilir. Bu önerilen yöntem tarafından gerçekleştirilen kümeleme işleminde her kümedeki doğru kümelенmiş örnek sayısının kıyaslanan diğer yöntemlerden daha fazla olduğu anlamına gelmektedir. Ortalama sıralamaya göre de saflık metriğindeki en başarılı yöntem tez kapsamında geliştirilen kaotik AAA tabanlı kümeleme yaklaşımıdır.

Çizelge 4.11'da algoritmaların veri kümeleri üzerinde VM metriğine göre otuz bağımsız çalışmada elde ettiği sonuçların ortalama ve standart sapma değerleri gösterilmektedir.

Çizelge 4.10. Kümeleme Algoritmalarının veri kümeleri üzerinde SSE metriği bakımından otuz bağımsız çalışma ortalamaları ve standart sapma değerleri

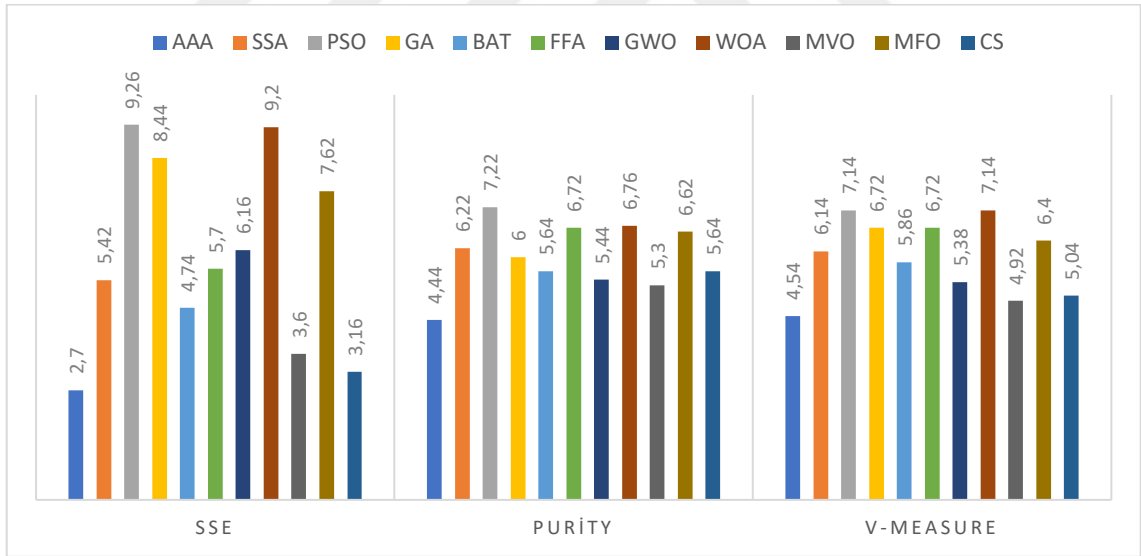
Veri Seti		CS	SSA	PSO	GA	BAT	FFA	GWO	WOA	MVO	MFO	CAAAv3
aggregation	mean	0.895	0.896	0.905	0.902	0.898	0.898	0.905	0.895	0.898	0.905	0.894
	±std	0.00	0.01	0.00	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.00
aniso	mean	0.669	0.800	0.794	0.783	0.770	0.770	0.798	0.755	0.800	0.798	0.800
	±std	0.03	0.00	0.01	0.01	0.00	0.00	0.01	0.03	0.00	0.01	0.00
appendicitis	mean	0.800	0.804	0.844	0.817	0.806	0.806	0.844	0.839	0.800	0.837	0.800
	±std	0.00	0.00	0.04	0.01	0.01	0.01	0.03	0.02	0.00	0.03	0.00
balance	mean	0.657	0.655	0.655	0.650	0.659	0.656	0.653	0.667	0.659	0.653	0.659
	±std	0.01	0.01	0.01	0.02	0.02	0.01	0.01	0.03	0.01	0.01	0.01
banknote	mean	0.572	0.790	0.734	0.785	0.560	0.560	0.790	0.743	0.790	0.790	0.790
	±std	0.02	0.00	0.09	0.01	0.00	0.00	0.00	0.08	0.00	0.00	0.00
blood	mean	0.760	0.760	0.760	0.760	0.760	0.760	0.760	0.760	0.760	0.760	0.760
	±std	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Diagnosis II	mean	0.820	0.685	0.646	0.660	0.684	0.669	0.651	0.670	0.677	0.697	0.820
	±std	0.00	0.12	0.09	0.11	0.00	0.09	0.1	0.1	0.00	0.11	0.00
ecoli	mean	0.786	0.790	0.692	0.770	0.802	0.764	0.742	0.744	0.790	0.739	0.808
	±std	0.02	0.02	0.06	0.03	0.02	0.04	0.04	0.05	0.02	0.05	0.01
flame	mean	0.640	0.640	0.640	0.640	0.640	0.640	0.640	0.640	0.640	0.640	0.640
	±std	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
heart	mean	0.706	0.625	0.631	0.649	0.659	0.648	0.656	0.639	0.653	0.650	0.710
	±std	0.02	0.06	0.06	0.05	0.06	0.05	0.05	0.07	0.05	0.06	0.00
iris	mean	0.960	0.950	0.821	0.944	0.941	0.927	0.950	0.941	0.960	0.924	0.960
	±std	0.00	0.05	0.1	0.02	0.07	0.08	0.05	0.07	0.00	0.06	0.00
iris2D	mean	0.960	0.960	0.960	0.961	0.960	0.960	0.960	0.960	0.960	0.956	0.960
	±std	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00
ionosphere	mean	0.687	0.670	0.655	0.691	0.678	0.662	0.701	0.702	0.696	0.672	0.700
	±std	0.01	0.02	0.01	0.00	0.03	0.03	0.00	0.00	0.01	0.02	0.00
jain	mean	0.870	0.870	0.870	0.870	0.870	0.870	0.870	0.870	0.870	0.870	0.870
	±std	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
liver	mean	0.580	0.580	0.580	0.580	0.580	0.581	0.580	0.581	0.580	0.580	0.580
	±std	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
moons	mean	0.853	0.850	0.850	0.850	0.860	0.860	0.850	0.850	0.850	0.850	0.85
	±std	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
mouse	mean	0.750	0.750	0.749	0.747	0.750	0.750	0.750	0.744	0.750	0.748	0.750
	±std	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00
pathbased	mean	0.840	0.840	0.841	0.832	0.840	0.840	0.840	0.836	0.840	0.840	0.840
	±std	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
smiley	mean	0.770	0.770	0.770	0.770	0.770	0.770	0.770	0.770	0.770	0.770	0.770
	±std	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
sonar	mean	0.530	0.531	0.532	0.535	0.531	0.531	0.535	0.531	0.530	0.533	0.532
	±std	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
VaryDensity	mean	0.950	0.866	0.950	0.951	0.847	0.950	0.952	0.942	0.950	0.950	0.950
	±std	0.00	0.13	0.00	0.00	0.13	0.00	0.00	0.05	0.00	0.00	0.00
Vertebral2	mean	0.680	0.680	0.680	0.680	0.680	0.680	0.680	0.680	0.680	0.680	0.680
	±std	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Vertebral3	mean	0.640	0.632	0.553	0.615	0.632	0.614	0.623	0.571	0.634	0.597	0.641
	±std	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
wdbc	mean	0.774	0.645	0.630	0.813	0.910	0.630	0.630	0.651	0.671	0.632	0.919
	±std	0.04	0.02	0.00	0.05	0.05	0.00	0.00	0.06	0.08	0.00	0.00
wine	mean	0.916	0.705	0.450	0.748	0.858	0.563	0.579	0.579	0.727	0.520	0.922
	±std	0.01	0.15	0.06	0.1	0.13	0.11	0.13	0.11	0.14	0.09	0.00
En iyi sonuç sayısı		8	8	8	7	7	8	11	8	9	7	15
Friedman sıralama değerleri		5.64	6.22	7.22	6	5.64	6.72	5.44	6.76	5.3	6.62	4.44
Genel sıralama		5	7	11	6	4	9	3	10	2	8	1

Çizelge 4.11. Kümeleme algoritmalarının VM (V-measure) metriği bakımından çalışma sonuçları

Veri Seti		CS	SSA	PSO	GA	BAT	FFA	GWO	WOA	MVO	MFO	CAAAv3
aggregation	mean	0.833	0.829	0.842	0.836	0.831	0.830	0.841	0.830	0.833	0.841	0.833
	±std	0.00	0.02	0.00	0.01	0.02	0.02	0.01	0.03	0.01	0.01	0.00
aniso	mean	0.586	0.530	0.527	0.524	0.590	0.590	0.529	0.499	0.530	0.529	0.530
	±std	0.06	0.00	0.00	0.01	0.00	0.00	0.00	0.03	0.00	0.00	0.00
appendicitis	mean	0.190	0.194	0.195	0.209	0.198	0.182	0.255	0.238	0.190	0.231	0.190
	±std	0.00	0.00	0.14	0.03	0.01	0.05	0.04	0.06	0.00	0.08	0.00
balance	mean	0.118	0.119	0.113	0.109	0.119	0.117	0.114	0.136	0.118	0.114	0.120
	±std	0.01	0.01	0.01	0.02	0.02	0.01	0.01	0.05	0.01	0.01	0.01
banknote	mean	0.015	0.300	0.232	0.293	0.000	0.000	0.300	0.240	0.300	0.300	0.300
	±std	0.02	0.0	0.13	0.02	0.0	0.0	0.0	0.11	0.0	0.0	0.0
blood	mean	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	±std	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Diagnosis II	mean	0.440	0.213	0.147	0.160	0.198	0.183	0.163	0.173	0.207	0.232	0.440
	±std	0.0	0.24	0.17	0.23	0.0	0.18	0.17	0.19	0.0	0.19	0.0
ecoli	mean	0.603	0.617	0.539	0.584	0.615	0.621	0.597	0.568	0.638	0.590	0.615
	±std	0.02	0.02	0.08	0.04	0.03	0.04	0.06	0.07	0.03	0.07	0.01
flame	mean	0.000	0.002	0.001	0.004	0.006	0.001	0.001	0.005	0.000	0.003	0.000
	±std	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
heart	mean	0.136	0.063	0.080	0.087	0.094	0.087	0.101	0.080	0.091	0.095	0.140
	±std	0.02	0.05	0.06	0.04	0.05	0.05	0.04	0.06	0.05	0.05	0.0
iris	mean	0.860	0.853	0.755	0.840	0.846	0.836	0.853	0.849	0.860	0.826	0.860
	±std	0.0	0.03	0.06	0.02	0.05	0.06	0.03	0.04	0.0	0.05	0.0
iris2D	mean	0.860	0.860	0.860	0.864	0.860	0.860	0.860	0.860	0.860	0.856	0.860
	±std	0.0	0.0	0.0	0.01	0.0	0.0	0.0	0.0	0.0	0.02	0.0
ionosphere	mean	0.106	0.071	0.049	0.114	0.080	0.048	0.126	0.124	0.115	0.082	0.127
	±std	0.03	0.05	0.04	0.0	0.04	0.05	0.0	0.0	0.02	0.05	0.0
jain	mean	0.510	0.510	0.510	0.505	0.508	0.510	0.510	0.506	0.510	0.510	0.510
	±std	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
liver	mean	0.000	0.000	0.005	0.000	0.000	0.001	0.001	0.004	0.000	0.001	0.000
	±std	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
moons	mean	0.399	0.400	0.400	0.397	0.410	0.410	0.399	0.394	0.400	0.400	0.400
	±std	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
mouse	mean	0.550	0.550	0.549	0.549	0.550	0.550	0.550	0.546	0.550	0.548	0.550
	±std	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.02	0.0	0.0	0.0
pathbased	mean	0.620	0.620	0.621	0.607	0.620	0.620	0.620	0.614	0.620	0.620	0.620
	±std	0.0	0.0	0.0	0.01	0.0	0.0	0.0	0.0	0.0	0.0	0.0
smiley	mean	0.240	0.235	0.243	0.228	0.236	0.240	0.239	0.238	0.240	0.239	0.240
	±std	0.0	0.01	0.01	0.01	0.01	0.01	0.0	0.01	0.0	0.01	0.0
sonar	mean	0.026	0.011	0.010	0.013	0.002	0.010	0.012	0.018	0.013	0.011	0.005
	±std	0.0	0.0	0.0	0.01	0.0	0.0	0.0	0.01	0.0	0.0	0.0
VaryDensity	mean	0.860	0.797	0.860	0.860	0.783	0.860	0.864	0.857	0.860	0.860	0.860
	±std	0.0	0.09	0.0	0.0	0.1	0.0	0.02	0.03	0.0	0.0	0.0
Vertebral2	mean	0.150	0.090	0.043	0.123	0.100	0.095	0.113	0.073	0.126	0.106	0.150
	±std	0.0	0.07	0.05	0.01	0.07	0.07	0.06	0.06	0.05	0.06	0.0
Vertebral3	mean	0.212	0.218	0.145	0.212	0.219	0.197	0.208	0.161	0.218	0.189	0.212
	±std	0.01	0.0	0.06	0.03	0.0	0.06	0.02	0.07	0.0	0.05	0.0
wdbc	mean	0.288	0.039	0.008	0.375	0.570	0.009	0.007	0.045	0.091	0.015	0.587
	±std	0.08	0.05	0.0	0.11	0.1	0.0	0.0	0.12	0.15	0.01	0.0
wine	mean	0.753	0.542	0.114	0.574	0.704	0.350	0.383	0.382	0.582	0.260	0.767
	±std	0.02	0.16	0.15	0.08	0.11	0.21	0.23	0.21	0.13	0.2	0.01
En iyi sonuç sayısı		6	3	6	1	5	4	5	1	5	2	10
Friedman sıralama değerleri		5.04	6.14	7.14	6.72	5.86	6.72	5.38	7.14	4.92	6.4	4.54
Genel sıralama		3	6	10	8	5	9	4	11	2	7	1

Çizelge 4.11'daki sonuçlara göre kaotik AAA tabanlı kümele yaklaşımı VM metriği açısından değerlendirildiğinde, AAA en başarılı algoritma olarak görünmektedir. On veri kümesinde en iyi değeri elde ederek öne çıkmaktadır. Friedman sıralama sonuçları bu başarıyı doğrulamaktadır.

Gerçekleştirilen deneysel çalışmalarda SSE, saflık ve VM metriklerine göre tez kapsamında geliştirilen kaotik AAA tabanlı kümeleme yaklaşımı kıyaslan diğer yöntemlere göre daha başarılıdır. En iyi sonuç ve ortalama sıralama değerleri, geliştirilen yöntemin başarısını açıkça göstermektedir. Şekil 4.8'de tüm algoritmalar için SSE, saflık ve VM metriklerinin ortalama sıralama sonuçları gösteren çubuk diyagramı sunulmaktadır. Kaotik AAA tabanlı kümeleme yaklaşımının üç metrikte de başarılı sonuçlar elde ettiği görülmektedir. Bununla birlikte, MVO ve CS tabanlı kümeleme algoritmaları da benzer bir yaklaşım benimsemiştir. Ayrıca her üç metrik sonucunun standart sapmaları incelendiğinde kaotik AAA tabanlı yönteminin standart sapma değerleri genellikle diğer yöntemlerin değerlerinden daha küçüktür. Bu sonuçlar geliştirilen kaotik AAA tabanlı kümeleme yaklaşımının kararlı olduğunu göstermektedir.



Şekil 4.8. Kümeleme analizinde kullanılan yaklaşımların deneysel çalışmanın metrik sonuçlarına göre oluşturulan çubuk diyagramı

Deneysel çalışmalarda kullanılan veri kümeleri farklı karakteristiklere (küme sayısı ve öz nitelik sayısı) sahiptir. Küme sayısı ve öz nitelik sayısı arttıkça problemin boyutu büyür ve daha karmaşık hale gelmesine neden olur. Aggregation, Ecoli ve Smiley veri kümelerinin küme sayısı diğer veri kümelerine nispeten daha küçüktür ve sırasıyla

7, 5 ve 4'tür. Kaotik AAA tabanlı kümeleme yaklaşımı, Aggregation veri kümesinde ikinci en iyi SSE değerini elde ederken, Ecoli ve Smiley veri kümelerinde en iyi SSE değerine ulaşmıştır. Öte yandan Heart, Ionosphere, Sonar, WDBC ve Wine veri kümeleri diğer veri kümelerine göre daha fazla özneliğe sahiptir ve öznelik sayıları sırasıyla 13, 34, 60, 30 ve 13'tür. Kaotik AAA tabanlı yaklaşım, Sonar hariç diğer dört veri kümesinde (Heart, Ionosphere, WDBC ve Wine) en iyi SSE değerini elde ederken, Sonar veri kümesinde en iyi ikinci SSE değerini elde etmiştir.

Deneysel çalışmanın sonuçları kümeleme probleminin boyutu açısından değerlendirildiğinde, tezde geliştirilen kaotik AAA tabanlı yaklaşımın kıyaslanan güçlü algoritmalarından daha performanslı olduğu, dikkate değer bir başarıya ulaştığı gözükmemektedir.

Çizelge 4.9'da gösterilen yirmi beş veri kümesi üzerinde kaotik AAA tabanlı yaklaşımla kıyaslanan on farklı algoritmanın otuz bağımsız çalışma sonucunda oluşan SSE değerleri gösterilmektedir. Bazı veri kümelerinde, geliştirilen yöntem ile kıyaslanan algoritmaların elde ettiği sonuçlar birbirine oldukça yakındır. Bu noktada, sonuçlar arasındaki farkın anlamlı olup olmadığı problemi ortaya çıkmaktadır. Bu problemi çözmek, iki algoritma arasındaki anlamlı farkı bulmak için, parametrik olmayan istatistiksel bir test olan Wilcoxon işaretli sıralar testi literatürde yaygın olarak kullanılmaktadır. Bu test, iki algoritmanın otuz çalıştırma sonucunu kullanarak P-değerini (P-value) hesaplamaktadır. Eğer P-değeri anlamlılık düzeyinden büyükse H_0 (Boş hipotez) kabul edilir. Anlamlılık düzeyi literatürde 0,05 olarak kabul görmektedir. Eğer P-değeri anlamlılık düzeyinden küçükse H_0 hipotezi reddedilir ve H_1 hipotezi (alternatif hipotez) kabul edilir. H_0 hipotezi "iki yöntemin sonuçları arasında istatistiksel olarak anlamlı bir fark yoktur" anlamını taşır. H_1 hipotezi, H_0 hipotezinin tersidir ve anlamlı bir fark olduğu gösterir (Demsar, 2006). Testlerde anlam seviyesi birçok çalışmada olduğu gibi $\alpha=0,05$ olarak kullanılmıştır. P-değeri (p-value) bu α değeri ile kıyaslanır. R^+ değeri ilk algoritmanın ikinciden daha iyi olduğu sonuçlardaki sıralama değerlerinin toplamlarını ifade eder ve R^- ise ikinci algoritmanın daha iyi olduğu sonuçlardaki sıralama değerlerinin toplamlarını ifade eder. T ise bu toplamların küçüğünü temsil eder ($T=\min(R^+,R^-)$) (Derrac ve ark., 2011).

Kaotik AAA tabanlı kümeleme yaklaşımı ile kıyaslanan diğer algoritmaların SSE değerlendirme metriği sonuçları, Wilcoxon işaretli sıra testi kullanılarak karşılaştırılmıştır. Anlamlılık düzeyi test değeri 0,05 olarak belirlenmiştir. Wilcoxon

işaretli sıra testinden elde edilen sonuçlar, Çizelge 4.12'de (SSA, PSO, GA, BAT ve FFA için) ve Çizelge 4.13'te (GWO, WOA, MVO, MFO ve CS) sunulmaktadır.

Çizelge 4.12- 4.13'ün son satırında Wilcoxon işaretli sıralar testi sonucunda karşılaştırılan iki yöntem için eşitlik, kazanılan ve kaybedilen değerlerin toplamı ayrı ayrı sunulmuştur. Çizelge 4.12'de tez kapsamında geliştirilen kaotik AAA tabanlı kümeleme yöntemi ile kıyaslanan SSA, PSO, GA, BAT ve FFA tabanlı kümeleme yöntemleri arasında yapılan Wilcoxon işaretli sıralar testi sonuçları gösterilmektedir.

Çizelge 4.12. Geliştirilen kaotik AAA tabanlı kümeleme yöntemi ile kıyaslanan SSA, PSO, GA, BAT ve FFA tabanlı kümeleme yöntemleri arasında yapılan Wilcoxon işaretli sıralar testi sonuçları

Veri Seti	CAAAv3 - SSA			CAAAv3- PSO			CAAAv3- GA			CAAAv3- BAT			CAAAv3- FFA		
	p-Value	T	W	p-Value	T	W	p-Value	T	W	p-Value	T	W	p-Value	T	W
aggregation	0.00014	17	1	3.20E-05	8	1	1.73E-06	0	1	2.64E-05	3	1	9.10E-05	21	1
aniso	1	0	0	0.25	0	0	1.73E-06	0	1	4.32E-08	465	2	4.32E-08	465	2
appendicitis	1	0	0	1.73E-06	0	1	1.73E-06	0	1	0.0625	0	0	0.003906	0	1
balance	5.51E-01	71	0	0.008195	29	1	1.73E-06	0	1	4.07E-06	1	1	5.27E-01	97	0
banknote	1	0	0	0.000255	0	1	1.73E-06	0	1	6.80E-08	465	2	4.77E-07	465	2
Blood	1	0	0	3.47E-06	0	1	1.73E-06	0	1	1.00E+00	0	0	1.00E+00	0	0
Diagnosis II	7.46E-05	0	1	5.27E-06	0	1	1.73E-06	0	1	8.08E-06	0	1	7.60E-06	0	1
ecoli	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	0	1	2.60E-05	28	1	1.73E-06	0	1
flame	1.00E+00	0	0	1.00E+00	0	0	1.22E-04	0	1	1.00E+00	0	0	1.00E+00	0	0
heart	3.72E-05	3	1	1.73E-06	0	1	1.73E-06	0	1	2.75E-05	12	1	0.000135	1	1
iris	1.00E+00	0	0	2.53E-06	0	1	1.73E-06	0	1	5.00E-01	0	0	1.25E-01	0	0
iris2D	1	0	0	2.50E-01	0	0	1.46E-06	0	1	1	0	0	1	0	0
ionosphere	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	0	1	8.73E-03	105	1	0.000114	45	1
jain	1	0	0	1	0	0	2.19E-06	0	1	0.0625	0	0	1	0	0
liver	0.125	0	0	1.73E-06	0	1	1.73E-06	0	1	1	0	0	0.007812	0	1
moons	1	0	0	1.00E+00	0	0	1.73E-06	0	1	4.32E-08	465	2	4.32E-08	465	2
mouse	1	0	0	1	0	0	1.69E-06	0	1	1	0	0	1	0	0
pathbased	1	0	0	0.25	0	0	4.48E-06	0	1	1	0	0	1	0	0
smiley	0.00048	0	1	9.72E-05	0	1	1.72E-06	0	1	5.94E-06	0	1	0.001953	0	1
sonar	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	0	1	7.16E-04	397	2	1.73E-06	0	1
vary-density	0.00390	0	1	1	0	0	1.05E-04	0	1	0.000976	0	1	1	0	0
vertebral2	1.22E-04	0	1	3.78E-06	0	1	1.73E-06	0	1	0.001953	0	1	9.77E-04	0	1
vertebral3	9.28E-06	1	1	1.73E-06	0	1	1.73E-06	0	1	2.70E-06	5	1	4.10E-06	4.5	1
wdbc	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	0	1	3.11E-05	435	1	1.73E-06	0	1
wine	3.40E-05	21	1	1.73E-06	0	1	1.73E-06	0	1	0.27844	268	0	2.56E-06	0	1
Kazanılan/ Eşitlik/ Kaybedilen	12/13/0			17/8/0			25/0/0			11/10/4			13/9/3		

Çizelge 4.13'te tezde geliştirilen kaotik AAA tabanlı kümeleme yöntemi ile kıyaslanan GWO, WOA, MVO, MFO ve CS tabanlı kümeleme yöntemleri arasında yapılan Wilcoxon işaretli sıralar testi sonuçları gösterilmektedir.

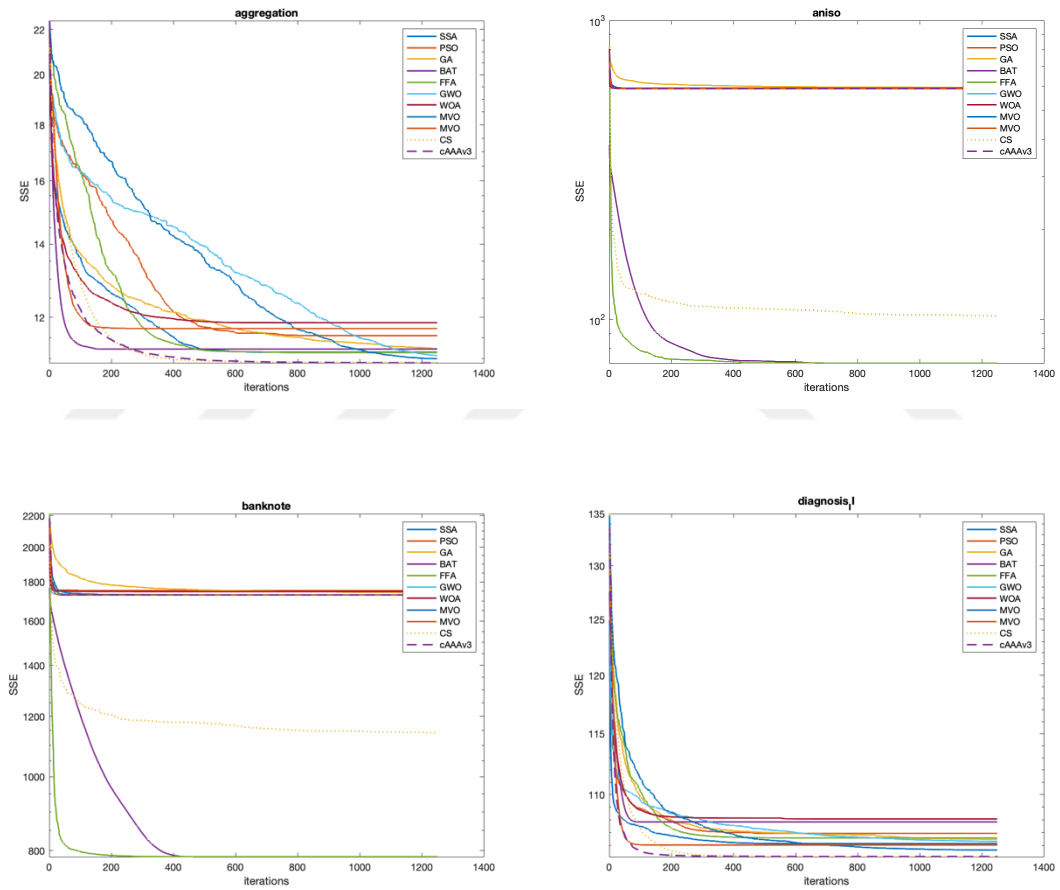
Çizelge 4.13. Geliştirilen kaotik AAA tabanlı kümeleme yöntemi ile kıyaslanan GWO, WOA, MVO, MFO ve CS tabanlı kümeleme yöntemleri arasında yapılan Wilcoxon işaretli sıralar testi sonuçları

Veri Seti	CAAAv3-GWO			CAAAv3-WOA			CAAAv3-MVO			CAAAv3-MFO			CAAAv3-CS		
	p-Value	T	W	p-Value	T	W	p-Value	T	W	p-Value	T	W	p-Value	T	W
aggregation	0.00510	65	1	1.73E-06	0	1	0.048895	66	1	0.00011	15	1	0.418367	144	0
aniso	3.17E-05	0	1	4.33E-05	0	1	1	0	0	1	0	0	1.73E-06	465	2
appendicitis	1.82E-05	0	1	1.73E-06	0	1	1	0	0	3.99E-05	0	1	1.00E+00	0	0
balance	5.47E-01	79	0	1.73E-06	0	1	9.44E-01	61	0	0.01876	55	1	0.15768	77	0
banknote	3.65E-06	0	1	8.73E-05	0	1	1	0	0	1	0	0	1.73E-06	465	2
Blood	2.50E-01	0	0	6.10E-05	0	1	1	0	0	0.5	0	0	1.00E+00	0	0
Diagnosis II	5.82E-06	4	1	1.73E-06	0	1	4.18E-05	1	1	7.82E-06	0	1	3.93E-04	20	1
ecoli	1.73E-06	0	1	1.73E-06	0	1	2.60E-06	4	1	1.73E-06	0	1	1.73E-06	0	1
flame	1.00E+00	0	0	5.00E-01	0	0	1.00E+00	0	0	1.00E+00	0	0	1.00E+00	0	0
heart	1.73E-06	0	1	1.73E-06	0	1	1.47E-05	1	1	1.73E-06	0	1	2.02E-06	1.5	1
iris	8.12E-05	0	1	1.50E-06	0	1	1.00E+00	0	0	1.57E-04	0	1	1.00E+00	0	0
iris2D	2.50E-01	0	0	1.34E-04	0	1	1	0	0	0.5	0	0	1.00E+00	0	0
ionosphere	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	0	1
jain	1.00E+00	0	0	1.56E-02	0	1	1	0	0	1	0	0	1	0	0
liver	0.015625	0	1	1.73E-06	0	1	1	0	0	0.00012	0	1	1	0	0
moons	0.00012	0	1	3.13E-02	0	1	1	0	0	1	0	0	1.73E-06	465	2
mouse	1	0	0	8.02E-05	0	1	1	0	0	1	0	0	1	0	0
pathbased	1	0	0	2.68E-05	0	1	1	0	0	0.03125	0	1	1	0	0
smiley	0.125	0	0	2.41E-06	0	1	0.0625	0	0	0.00012	0	1	0.5	0	0
sonar	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	0	1
vary-density	0.5	0	0	2.50E-01	0	0	1	0	0	1	0	0	1	0	0
vertebral2	1.21E-05	0	1	1.73E-06	0	1	0.0625	0	0	1.85E-04	0	1	1.00E+00	0	0
vertebral3	1.73E-06	0	1	1.73E-06	0	1	1.76E-05	6.5	1	1.73E-06	0	1	1.72E-06	0	1
wdbc	1.73E-06	0	1	1.73E-06	0	1	1.92E-06	1	1	1.73E-06	0	1	1.73E-06	0	1
wine	1.92E-06	1	1	1.73E-06	0	1	4.40E-04	55	1	1.73E-06	0	1	1.73E-06	0	1
Kazanılan/ Eşitlik/ Kaybedilen	16/9/0			23/2/0			9/16/0			16/9/0			8/14/3		

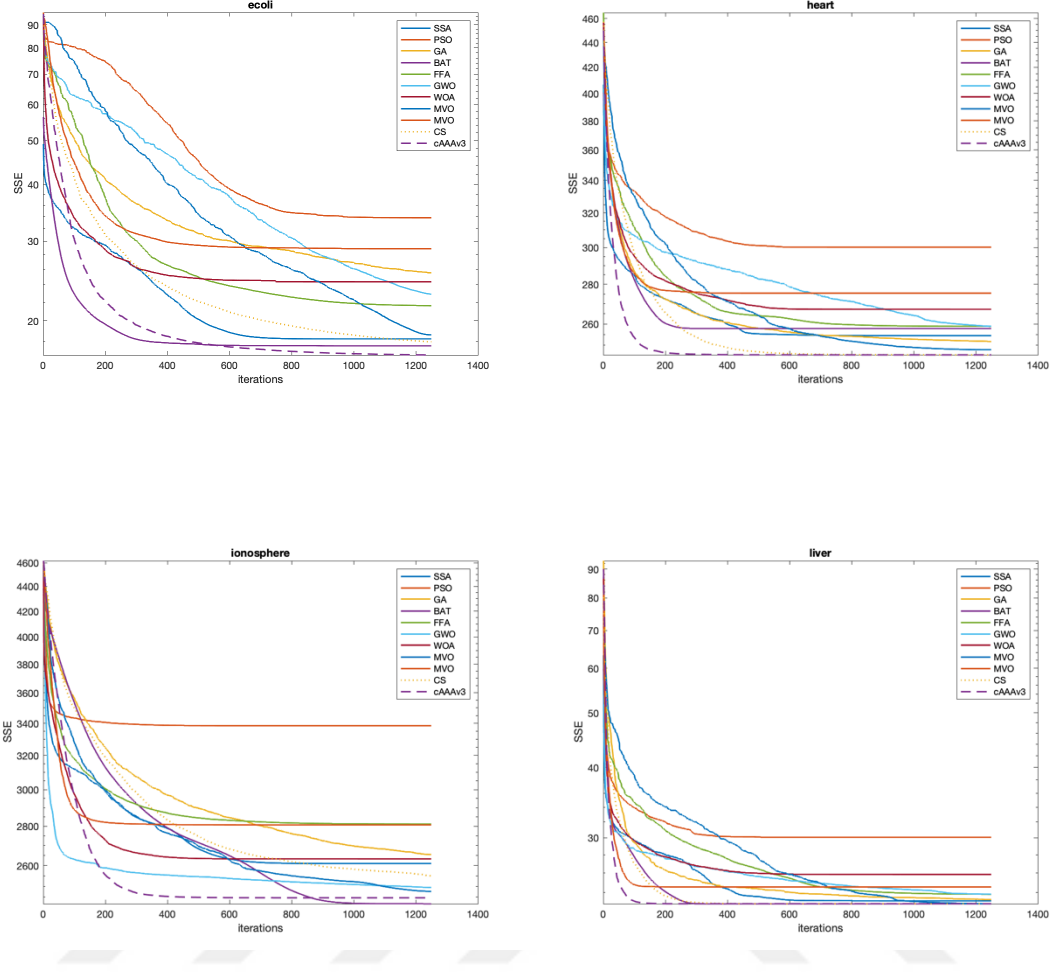
Wilcoxon işaretli sıralar testi sonuçları analiz edildiğinde SSA, PSO, GA, GWO, WOA, MVO ve MFO tabanlı kümeleme algoritmaları, bazı veri kümelerinde geliştirilen kaotik AAA tabanlı kümeleme yöntemine karşı iddialı sonuçlar elde ettiği, ancak herhangi bir veri kümesinde geliştirdiğimiz yöntemden daha iyi sonuçlar elde etmediği görülmektedir. BAT, FFA ve CA tabanlı kümeleme yaklaşımları sırasıyla dört (Aniso, Banknote, Moons, ve Sonar), üç (Aniso, Banknote ve Moons) ve üç (Aniso, Banknote ve

Moons) veri kümesinde kaotik AAA tabanlı yaklaşımdan daha iyi sonuçlar elde etmiştir. Öte yandan, herhangi bir yöntemin Wilcoxon işaretli sıralar testinde kazananların toplam sayısı, tezde geliştirilen kaotik AAA tabanlı kümeleme yönteminden daha anlamlı değildir. Dolayısıyla tezde geliştirilen yöntem diğer yöntemlere göre istatistiksel olarak daha başarılıdır.

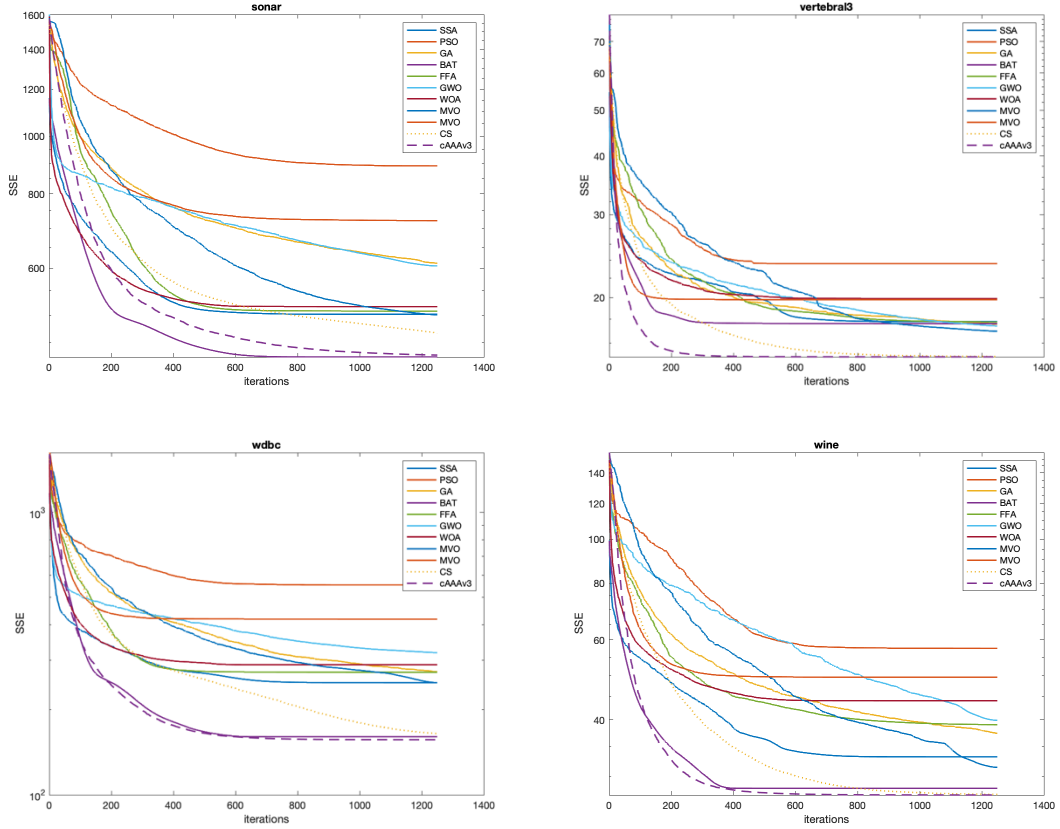
Yakınsama eğrileri, en iyi çözüm arayışı bitene kadar her adımda algoritmanın elde ettiği en iyi çözüm değerlerinden oluşur ve algoritmalarının problem çözme davranışını gösterir. Deneysel çalışmada elde edilen farklı özelliklere sahip on iki veri kümesi üzerinde algoritmaların göstermiş oldukları yakınsama eğrileri Şekil 4.9- 4.11 gösterilmektedir.



Şekil 4.9. Algoritmaların kümeleme veri setlerinden Aggregation, Aniso, Banknote ve Diagnositise ait yakınsama eğrileri



Şekil 4.10. Algoritmaların kümeleme veri setlerinden Ecoli, Heart, Ionosphere ve Liver'e ait yakınsama eğrileri



Şekil 4.11. Algoritmaların kümeleme veri setlerinden Sonar, Vertebral 3, WDBC ve Wine'ye ait yakınsama eğrileri

Şekil 4.9-4.11'deki yakınsama eğrileri incelendiğinde Diagnosis, Heart, Liver, Vertebral 3 ve Wine veri kümelerinde, kaotik AAA tabanlı kümeleme yaklaşımının yinelemelerin hemen başlangıcında iyi bir şekilde optimum çözüme yakınsandığı ve yinelemelerin sonlarına kadar gelişimini sürdürdüğü görülmektedir. Diagnosis, Heart, Vertebral 3 ve Wine veri kümelerinde geliştirilen kaotik AAA tabanlı kümeleme yaklaşımının hem yakınsama eğrisi bakımından hem de istatistiksel olarak diğer yöntemlerden daha başarılıdır. Liver veri kümesinde SSA, BAT ve CS tabanlı kümeleme algoritmaları istatistiksel olarak kaotik AAA ile aynı sonuçları elde etmiş ve kıyaslanan diğer algoritmaların hiç birisi kaotik AAA tabanlı yöntemden daha iyi sonuçlar elde edememiştir.

Kaotik AAA tabanlı kümeleme yaklaşımının sonucu ile CS tabanlı kümeleme algoritması arasındaki fark istatistiksel olarak anlamlı değildir. Bununla birlikte, CS tabanlı yöntem bu veri kümeleri arasında yalnızca Aggregation veri kümesinde en iyi değeri elde etmiştir. Ecoli, Ionosphere ve WDBC veri setlerinde, AAA yöntemi

yinelemelerin başlangıcında hızlı bir yakınsama sağlamasa da yinelemelerin ilerlemesiyle iyileşme göstererek başarılı sonuçlar elde etmiştir.

Aniso ve Banknote veri kümelerinde BAT, FFA ve CS tabanlı algoritmalar yinelemeler boyunca mevcut en iyi çözümü geliştirmeye devam ederken, diğer algoritmalar yerel minimumdan kaçınamamış ve çözümü geliştirmişlerdir. Sonar veri kümesinde, kaotik AAA tabanlı kümeleme yöntemi de sürekli iyileştirme göstermiştir, ancak BAT tabanlı kümeleme yöntemi daha iyi bir geliştirme süreci göstermiştir.

Yakınsama eğrilerinde genel olarak değerlendirildiğinde geliştirilen kaotik AAA tabanlı kümeleme yaklaşımı, kıyaslanan popüler algoritmaların birçoğundan iyi bir yakınsama davranışı sergilemiş, güçlü yöntemlere nispeten de oldukça rekabetçi bir karakteristik göstermiştir.

4.4. Öznitelik Seçme Probleminde İkili Kaotik Yapay Alg Algoritmasının Performansı

Bu bölümde öncelikle deneysel çalışmada kullanılan veri setleri ve değerlendirme kriterleri sunulacaktır. Sonrasında, geliştirilen sekiz farklı ikili varyasyonun analizi yapılacak ve en iyi varyasyon tespit edilecektir. Daha sonra, seçilen en iyi varyasyon ile literatürdeki popüler algoritmaların veri setlerindeki performansları değerlendirme kriterleri üzerinden analiz edilerek kıyaslanacaktır. Tez kapsamında geliştirilen ikili kaotik AAA yaklaşımının, öznitelik seçimi problemlerine nasıl uygulanacağı Bölüm 3.5'te detaylandırılmıştır.

Geliştirilen ikili algoritmanın değerlendirilmesinde farklı özelliklere ve farklı zorluk seviyelerine sahip yaygın olarak bilinen ve birçok çalışmada kullanılan yirmi beş veri seti tercih edilmiştir (Emine ve Ülker, 2020). Bu veri kümeleri Breast cancer Wisconsin diagnostic, Climate model simulation crashes, Dermatology, Cnae-9, Ecoli, Electrical grid, Heart, Hepatitis, Hillvalley, Image segmentation, Ionosphere, Letter recognition, Ozone level detection eight, Liver, Optical recognition handwritten digits, Pen-based recognition handwritten digits, Planning relax, Qsar Biodegradation, Sonar, Spambase, Statlog project Landsat satellite, Thyroid Disease, Waveform, WDBC ve Weight lifting exercises'dır. Çizelge 4.14' de özellikleri gösterilen bu veri setleri UCI Machine Learning Repository⁷ (Dhote ve ark., 2013)'den alınmıştır.

⁷ <https://archive.ics.uci.edu/ml/>.

Bu veri setlerinin tamamı [0,1] aralığına minimum-maksimum normalleştirilmesi kullanılarak ölçeklenmiş ve %66'sı eğitim ve %34'ü test olmak üzere iki gruba ayrılmıştır. Algoritmaların adil bir şekilde kıyaslanabilmesi için deneysel çalışmaların tamamı aynı platform üzerinde EvolopyFS (Khurma ve ark., 2020) kütüphanesi kullanılarak gerçekleştirilmiştir. EvolopyFS, Python'da geliştirilmiş, iyi bilinen ikili metasezgisel optimizasyon algoritmalarını içeren açık kaynak kodlu bir kütüphanedir. Tüm algoritmalar, 10.000 uygunluk hesaplaması (FEs) ile çalıştırılmış ve otuz bağımsız çalışma yapılarak değerlendirilmiştir. Ayrıca, algoritmaların başlangıç popülasyonları rastgele oluşturulmuş ve popülasyondaki birey sayısı 40 olarak belirlenmiştir. Bu çalışmadaki parametrelerin bu değerleri literatürdeki birçok çalışma tarafından kullanılmaktadır (Nakamura ve ark., 2012; Rodrigues ve ark., 2014; Emary ve ark., 2016; Mafarja ve Mirjalili, 2018; Khurma ve ark., 2020).

Çizelge 4.14. Öznitelik seçim probleminde kullanılan veri setleri ve özellikleri

No	Veri seti ismi	Boyutu	Öznitelik	Sınıf	Alanı
1	Breast cancer Wisconsin diagnostic	569	30	2	Biyoloji
2	Climate model simulation crashes	540	18	2	Biyoloji
3	Cnae - 9	1080	856	9	Ekonomi
4	Dermatology	358	34	6	Biyoloji
5	Ecoli	1372	4	2	Biyoloji
6	Electrical grid	10000	12	2	Elektromanyetik
7	Heart	120	6	2	Biyoloji
8	Hepatitis	80	19	2	Biyoloji
9	Hillvalley	606	100	2	Coğrafya
10	Image segmentation	210	19	7	Yapay Zeka
11	Ionosphere	351	34	2	Coğrafya
12	Letter recognition	2000	16	26	Yapay Zeka
13	Liver	345	7	2	Biyoloji
14	Optical recognition handwritten digits	3823	64	10	Yapay Zeka
15	Ozone level detection eight	1847	72	2	Coğrafya
16	Pen based recognition handwritten digits	7494	16	10	Yapay Zeka
17	Planning relax	182	12	2	Biyoloji
18	Qsar Biodegradation	1055	41	2	Kimya
19	Sonar	208	60	2	Coğrafya
20	Spambase	4601	57	2	Yapay Zeka
21	Statlog Project landsat satellite	4435	36	6	Coğrafya
22	Thyroid Disease	3772	21	3	Biyoloji
23	Waveform	5000	40	3	Fizik
24	WDBC	569	30	2	Biyoloji
25	Weight lifting exercises	4024	54	5	Biyoloji

Bölümün en sonunda deneysel çalışma sonuçlarının istatistiksel testleri (Wilcoxon işaretli sıralar testi ve Friedman testi) yapılmış, yerel minimumlardan kaçınma kabiliyeti, kutu grafikleri ve yakınsama eğrileri kullanılarak, değerlendirme kriterleri (uygunluk değeri, sınıflandırma doğruluğu ve ortalama öznitelik sayısı) üzerinden analiz edilmiştir.

4.4.1. Değerlendirme Kriterleri

Öznitelik seçme işleminin ne kadar başarılı bir şekilde yapıldığını değerlendirmek için farklı metrikler vardır. Literatürde çalışmaların neredeyse hepsinde kullanılan üç metrik vardır (Arora ve Anand, 2019b; Emine ve Ülker, 2020; Pan ve ark., 2021; Awadallah ve ark., 2022; Hu ve ark., 2022; Turkoglu ve ark., 2022b). Bu metrikler sınıflandırma doğruluğu, seçilen öznitelik sayısı ve uygunluk fonksiyonudur. Bu metrikler sınıflandırma doğruluğu, seçilen öznitelik sayısı ve uygunluk değeridir.

Sınıflandırma Doğruluğu: öznitelik seçimi işleminin ardından oluşan veri setinin, bir sınıflandırma algoritması ile sınıflandırıldığında elde edilen test başarısını gösterir. Sınıflandırma doğruluğu metriğinin matematiksel ifadesi Denklem 4.9'da gösterilmektedir.

$$\uparrow \text{Ortalama Performans} = \frac{1}{M} \sum_{j=1}^M \frac{1}{N} \sum_{i=1}^N \text{match}(C_i, L_i) \quad (4.9)$$

Denklemlerde M kaç bağımsız çalışma yapıldığını, N test kümesindeki veri sayısını, C_i tahmin edilen sınıf etiketini, L_i gerçek sınıf etiketini göstermektedir.

Seçilen Öznitelik Sayısı: Öznitelik seçim işlemi sonrasında, seçilen özniteliklerin sayısıdır. Kaç bağımsız çalışma yapıldıysa ortalaması alınır. Denklem 4.10'daki matematiksel ifade kullanılarak hesaplanır.

$$\downarrow \text{Mean} = \frac{1}{M} \sum_{i=1}^M \text{size}(g^*) \quad (4.10)$$

Denklemden M kaç bağımsız çalışma yapıldığını, $\text{size}(g^*)$ test verilerinde seçilen öznitelik sayısını gösterir.

Uygunluk Değeri: Sınıflandırma doğruluğu ve seçilen öznitelik sayısını dengeleyerek en ideal öznitelikleri tespit etmek için geliştirilmiş fonksiyondur. Denklem 4.11’de matematiksel ifadesi gösterilmektedir.

$$\downarrow \text{Uygunluk Değeri} = 0.99 * (1 - \text{Sınıflandırma Başarısı}) + 0.01 * \frac{|\text{Seçilen Öznitelik Sayısı}|}{|\text{Toplam Öznitelik Sayısı}|} \quad (4.11)$$

4.4.2. Performans Değerlendirmesi

Bu bölümde, ikili AAA varyasyonları ve literatürdeki popüler algoritmalar, veri setleri üzerinde çalıştırılarak, deneysel çalışma gerçekleştirilmiştir. Deneysel çalışmadan elde edilen sonuçlar detaylı olarak analiz edilmiştir. İlk olarak, sekiz farklı transfer fonksiyonunu kullanan AAA varyasyonlarının performansı analiz edilmiştir. Gerçekleştirilen analizde en başarılı “ikili AAA” versiyonuna karar verilmiştir. Daha sonra en başarılı ikili AAA varyasyonu, yirmi beş veri seti üzerinde, literatürdeki popüler algoritmalarla karşılaştırılmıştır. Farklı zorluk seviyelerine (özellik ve boyut) sahip veri setleri üzerinde karşılaştırılan tüm algoritmaların test sonuçları çizelgelerde sunulmuştur. Tüm yöntemler için, otuz bağımsız çalışma gerçekleştirilerek ortalama değerleri ve standart sapmaları bu çizelgelerde verilmiştir. Çizelgelerin son satırlarında en iyi sonuç sayısı, ortalama sıralama değerleri ve genel sıralama gösterilmiştir.

4.4.3. İkili Versiyonlarının Performans Analizi

Deneysel çalışma analizinin bu ilk bölümünde, ideal öznitelik alt kümesini belirlerken diğer bir tabirle öznitelik seçme problemini çözerken, sekiz farklı transfer fonksiyonunu esas alan ikili AAA versiyonlarının performansları değerlendirilecektir. Bu varyasyonlardan dört tanesi S şekilli transfer fonksiyonlarından geliştirilen bAAAs1, bAAAs2, bAAAs3, bAAAs4 ve dört tanesi V şekilli transfer fonksiyonlarından geliştirilen bAAAv1, bAAAv2, bAAAv3 ve bAAAv4 yaklaşımlardır. Analiz sürecinde, Çizelge 4.14’te gösterilen yirmi beş veri seti kullanılarak sekiz ikili AAA versiyonundan elde edilen sonuçlar analiz edilmiştir. İkili AAA versiyonlarının karşılaştırılmasında,

birbirinden bağımsız otuz çalışma sonucunda elde edilen uygunluk değerlerinin ortalama ve standart sapma değerleri kullanılmıştır.

Çizelge 4.15. bAAA versiyonlarının veri setleri üzerindeki çalışma sonuçları

Veri Seti		bAAAs1	bAAAs2	bAAAs3	bAAAs4	bAAAv1	bAAAv2	bAAAv3	bAAAv4
Breast cancer Wisconsin diagnostic	mean	0.0983	0.1035	0.1069	0.1075	0.1079	0.1079	0.1079	0.1079
	±std	0.0042	0.0067	0.0039	0.0041	0.0014	0.0014	0.0014	0.0014
Climate model simulation crashes	mean	0.0646	0.0668	0.0622	0.0668	0.0665	0.0691	0.0644	0.0643
	±std	0.0109	0.0113	0.0122	0.0136	0.0139	0.0125	0.0150	0.0119
Cnae - 9	mean	0.1571	0.1735	0.1921	0.1984	0.2665	0.2696	0.2708	0.2697
	±std	0.0194	0.0187	0.0233	0.0252	0.0341	0.0291	0.0423	0.0365
Dermatology	mean	0.0405	0.0409	0.0476	0.0441	0.0533	0.0701	0.0638	0.0678
	±std	0.0165	0.0181	0.0187	0.0213	0.0398	0.0451	0.0463	0.0447
Ecoli	mean	0.1547	0.1547	0.1547	0.1547	0.1547	0.1547	0.1547	0.1547
	±std	0.0028	0.0028	0.0028	0.0028	0.0028	0.0028	0.0028	0.0028
Electrical grid	mean	0.2203	0.2203	0.2203	0.2203	0.2203	0.2203	0.2203	0.2202
	±std	0.0113	0.0113	0.0113	0.0113	0.0113	0.0113	0.0113	0.0004
Heart	mean	0.2286	0.2250	0.2247	0.2230	0.2588	0.2533	0.2582	0.2527
	±std	0.0150	0.0130	0.0139	0.0105	0.0105	0.0168	0.0109	0.0169
Hepatitis	mean	0.1749	0.2047	0.1948	0.1760	0.2580	0.2533	0.2675	0.258
	±std	0.0550	0.0573	0.0576	0.0551	0.0713	0.0718	0.0693	0.0713
Hillvalley	mean	0.4830	0.485	0.4823	0.4819	0.5027	0.4950	0.4958	0.4929
	±std	0.0115	0.0124	0.0138	0.0154	0.0232	0.0213	0.0237	0.0163
Image segmentation	mean	0.1194	0.1230	0.1235	0.1231	0.1245	0.1236	0.1253	0.1232
	±std	0.0091	0.0068	0.0064	0.0065	0.0051	0.0057	0.0041	0.0060
Ionosphere	mean	0.1402	0.1423	0.1323	0.1385	0.0947	0.0982	0.0971	0.0955
	±std	0.0181	0.0215	0.0239	0.0212	0.0093	0.0139	0.0124	0.0100
Letter recognition	mean	0.0529	0.0530	0.0554	0.0539	0.0529	0.0533	0.0564	0.0641
	±std	0.0035	0.0005	0.0019	0.0014	0.0035	0.0010	0.0028	0.0056
Liver	mean	0.3520	0.3520	0.3520	0.3520	0.3520	0.3520	0.3520	0.3520
	±std	0.0169	0.0169	0.0169	0.0169	0.0169	0.0169	0.0169	0.0169
Optical recognition handwritten digits	mean	0.0301	0.0301	0.0302	0.0307	0.0394	0.0403	0.0382	0.041
	±std	0.0029	0.0032	0.0033	0.0032	0.0053	0.0053	0.0061	0.0067
Ozone level detection eight	mean	0.0783	0.0774	0.0777	0.0789	0.0749	0.0755	0.0740	0.0764
	±std	0.0034	0.0041	0.0048	0.0059	0.0052	0.0053	0.0054	0.0047
Pen based recognition handwritten digits	mean	0.0198	0.0197	0.0189	0.0187	0.0188	0.0184	0.0191	0.0211
	±std	0.0008	0.0009	0.0016	0.0018	0.0018	0.0020	0.0021	0.0033
Planning relax	mean	0.3530	0.3531	0.3553	0.3530	0.3534	0.3535	0.3491	0.3539
	±std	0.0206	0.0121	0.0031	0.0064	0.0145	0.0093	0.0084	0.0160
Qsar Biodegradation	mean	0.1623	0.1634	0.1636	0.1586	0.1697	0.1727	0.1737	0.173
	±std	0.0081	0.0124	0.0146	0.0118	0.0149	0.0180	0.0160	0.0134
Sonar	mean	0.2720	0.2663	0.2770	0.2588	0.2648	0.2579	0.2630	0.2750
	±std	0.0340	0.0369	0.0458	0.0463	0.0375	0.0377	0.0387	0.0331
Spambase	mean	0.0845	0.0863	0.0898	0.0894	0.0973	0.1012	0.1004	0.1136
	±std	0.0062	0.0058	0.0069	0.0049	0.0069	0.0179	0.0078	0.0376
Statlog Project landsat satellite	mean	0.1099	0.1097	0.1106	0.1110	0.1154	0.1153	0.1147	0.1160
	±std	0.0043	0.0049	0.0064	0.0066	0.0054	0.0043	0.0048	0.0050
Thyroid Disease	mean	0.0269	0.0242	0.0245	0.0243	0.0196	0.0195	0.0199	0.0199
	±std	0.0049	0.0028	0.0048	0.0030	0.0012	0.0012	0.0011	0.0012
Waveform	mean	0.1971	0.1956	0.1980	0.1971	0.1982	0.2008	0.2003	0.2096
	±std	0.0071	0.0069	0.0073	0.0111	0.0087	0.0112	0.0113	0.0117
WDBC	mean	0.1002	0.1045	0.1080	0.1075	0.1079	0.1079	0.1079	0.1079
	±std	0.0057	0.0065	0.0025	0.0035	0.0014	0.0014	0.0014	0.0014
Weight lifting exercises	mean	0.0170	0.0157	0.0156	0.0149	0.0002	0.0002	0.0002	0.0002
	±std	0.0036	0.0023	0.0037	0.0032	0.0000	0.0000	0.0000	0.0000
En iyi sonuç sayısı		11	5	3	5	5	6	5	4
Sıralama değerleri		3.48	3.72	4.52	3.70	7.78	5.06	5.16	5.58
Genel sıralama		1	3	4	2	5	6	7	8

Çizelge 4.15'te, geliştirilen ikili kaotik AAA versiyonlarının uygunluk değerlerinin ortalamaları ve standart sapma değerleri gösterilmektedir. En iyi sonuçlar kalın yazı tipiyle vurgulanmıştır. Ortalama değerler göz önüne alındığında, S şekilli bAAA varyasyonları (bAAAs1, bAAAs2, bAAAs3, bAAAs4) V şekilli varyasyonlardan (bAAAv1, bAAAv2, bAAAv3, bAAAv4) daha iyi sonuçlar ettiği görülmektedir. Bu varyasyonlar arasından en başarılısı bAAAs1 varyasyonudur. Test için kullanılan yirmi beş veri setinin on bir tanesinde diğer varyasyonlardan daha iyi sonuçlar elde etmiştir. (Breast cancer Wisconsin diagnostic, Cnae-9, Dermatology, Ecoli, Hepatitis, Image segmentation, Letter recognition, Liver, Optical recognition handwritten digits, Spam base, WDBC). bAAAs1 varyasyonundan sonra en başarılı varyasyon bAAAs4 varyasyonudur ve altı veri setinde diğer yaklaşımlardan daha iyi performans göstermiştir. Ardından sırasıyla bAAAs2, bAAAs3, bAAAv1, bAAAv2, bAAAv3 ve bAAAv4 varyasyonları başarı göstermektedir. Uygunluk (fitness) fonksiyonu sonuçlarının standart sapma değerlerine göre, bAAAs1 varyasyonu dokuz veri setinde en düşük değeri yani yüksek başarı göstermektedir. Ardından, bAAAv4 ve bAAAv3 varyasyonları 21 veri setinden, sırasıyla beş ve altı tanesinde en iyi değere ulaşmışlardır. Standart sapma sonuçları bAAAs1 versiyonunun kararlı olduğunu doğrulamıştır.

Geliştirilen varyasyonlardan en başarılısı olan bAAAs1'in test başarısı ve öznitelik sayısı Çizelge 4.16'da gösterilmektedir. Bu sonuçlar otuz bağımsız çalışmanın ortalama, en iyi, en kötü ve standart sapma değerleridir.

bAAAs1 yaklaşımı on bir veri setinde en iyi uygunluk değerini elde ederek öznitelik seçme problemini diğer varyasyonlardan daha iyi çözebildiğini göstermiştir. Bununla birlikte Çizelge 4.16'daki değerlere göre test doğruluğu açısından da en başarılı yöntem bAAAs1'dir. Genel bir değerlendirme yapmak gerekirse deneysel çalışmanın varyasyon analizinde, bAAAs1 yaklaşımı en iyi performansı göstermiştir.

Çizelge 4.16. bAAAs1 versiyonunun otuz bağımsız çalışmadaki test başarısı ve indirilen öznitelik sayısı açısından ortalama, en iyi, en kötü ve standart sapma değerleri

bAAAs1 Veri Seti	Test Başarısı				İndirilen Öznitelik Sayısı			
	ortalama	en iyi	en kötü	std	ortalama	en iyi	en kötü	std
Breast cancer Wisconsin diagnostic	0.9036	0.9072	0.8918	0.0051	13.74	9.00	16.00	1.48
Climate model simulation crashes	0.9433	0.9620	0.9185	0.0123	11.08	8.00	14.00	1.35
Cnae - 9	0.8593	0.8886	0.7989	0.0162	691.52	671.00	713.00	11.32
Dermatology	0.9713	0.9836	0.9016	0.0134	21.65	19.00	25.00	1.79
Ecoli	0.8522	0.8522	0.8522	0.0000	5.00	5.00	5.00	0.00
Electrical grid	0.7839	0.7839	0.7839	0.0000	7.00	7.00	7.00	0.00
Heart	0.7667	0.7826	0.7283	0.0192	5.71	4.00	7.00	0.63
Hepatitis	0.8190	0.9286	0.7143	0.0471	7.82	5.00	9.00	1.1
Hillvalley	0.5529	0.5811	0.4928	0.0130	67.46	49.00	81.00	7.58
Image segmentation	0.8824	0.9167	0.8750	0.0070	9.29	7.00	10.00	1.1
Ionosphere	0.8664	0.9000	0.8167	0.0224	17.42	9.00	26.00	3.63
Lettter recognition	0.9540	0.9540	0.9540	0.0000	11.00	11.00	11.00	0.00
Liver	0.6525	0.6525	0.6525	0.0000	4.00	4.00	4.00	0.00
Optical recognition handwritten digits	0.9782	0.9831	0.9700	0.0034	50.12	41.00	54.00	2.66
Ozone level detection eight	0.9274	0.9331	0.9188	0.0038	44.23	27.00	58.00	8.06
Pen based recognition handwritten digitis	0.9883	0.9902	0.9878	0.0006	11.81	11.00	12.00	0.40
Planning relax	0.6468	0.7097	0.5806	0.0084	4.92	4.00	5.00	0.3
Qsar Biodegradation	0.8435	0.8607	0.8301	0.0102	29.1	25.00	34.00	1.77
Sonar	0.7352	0.8028	0.6761	0.0301	40.03	36.00	47.00	5.72
Spambase	0.9244	0.9310	0.9073	0.0051	43.91	38.00	46.00	2.43
Statlog Project landsat satellite	0.8995	0.9032	0.8859	0.0032	27.62	23.00	31.00	1.95
Thyroid Disease	0.9793	0.9844	0.9618	0.0031	9.00	5.00	10.00	1.3
Waveform	0.8094	0.8242	0.7948	0.0069	29.61	27.00	34.00	1.97
WDBC	0.9046	0.9072	0.8918	0.0048	14.02	9.00	16.00	1.44
Weight lifting exercises	0.9902	0.9942	0.9825	0.0027	37.92	33.00	41.00	2.23

4.4.4. Geliştirilen Varyasyonun Literatürdeki Algoritmalar ile Performans

Kıyaslaması

Bir önceki bölümde en başarılı ikili varyasyonun hem uygunluk fonksiyonu hem sınıflandırma doğruluğu açısından bAAAs1 olduğu ispat edilmiştir. Bu bölümde bAAAs1 varyasyonunun literatürdeki popüler yöntemlerle performansı kıyaslanmıştır. Karşılaştırma sürecinde on farklı karakteristik yapıya sahip metasezgisel optimizasyon algoritması kullanılmıştır. Hem sınıflandırma doğruluğu hem de uygunluk değerlendirmesi bakımından kıyaslanacak algoritmalar; PSO (Kennedy ve Eberhart, 1995), BAT (Yang, 2010b), GWO (Mirjalili ve ark., 2014a), MVO (Aljarah ve ark., 2020b), FFA (Yang, 2010a), CS (Yang ve Deb, 2009), MFO (Mirjalili, 2015) ve WOA

(Mirjalili ve Lewis, 2016) algoritmalarının s1 transfer fonksiyonunu kullanan ikili yaklaşımlardır.

Adil bir karşılaştırma için, bAAAs1 ve diğer yöntemler Çizelge 4.17'deki parametre seti kullanılarak aynı koşullar altında gerçekleştirilmiştir. Bu parametreler ilgili algoritmaların kendilerine has çalışma parametreleridir ve yazarlarının kendi yayınlarında tavsiye ettiği değerlerdir (Mirjalili ve Lewis, 2013; Mirjalili ve ark., 2014b; Crawford ve ark., 2015; Hussien ve ark., 2019; Hans ve Kaur, 2020; Abdel-Basset ve ark., 2021b; Nadimi-Shahraki ve ark., 2021).

Çizelge 4.17. Deneysel çalışmada karşılaştırılan algoritmaların parametre değerleri

Algoritmalar	Parametreler	Değerleri
AAA	Δ Share force	2
	e Energy loss	0.3
	A_p Adaptation	0.2
BAT	A (Loudness)	0.5
	r (Pulse rate)	0.5
	Q	[0, 2]
CS	Pa (Discovery rate of alien eggs)	0.25
PSO	c1,c2	2, 2
	$W_{inertia}$ weight	2
	W_{max} , W_{min}	0.9, 0.4
GWO	alpha	2
WOA	alpha	2
FFA	alpha (Randomness)	0.5
	β_{min} (minimum value of beta)	0.2
	Gama (absorption coefficient)	1
MVO	wep_{max}	1
	wep_{min}	0.2
MFO	b	1

Çizelge 4.18'de sınıflandırma doğruluğu açısından bAAAs1 ve diğer ikili BAT, CS, FFA, GWO, MFO, MVO, WOA ve PSO algoritmalarının sonuçlarını özetlemektedir. Çizelgede tüm yöntemler için otuz bağımsız çalışmanın ortalama değerleri ve standart

sapmaları verilmiştir. Çizelgelerin son satırları, en iyi sonuç sayısı, ortalama sıralama değerleri ve genel sıralama göstermektedir.

Çizelge 4.18. İkili algoritmaların veri setleri üzerinde sınıflandırma doğruluğu bakımından otuz bağımsız çalışma ortalamaları ve standart sapma değerleri

Veri Seti		bWOA	bBAT	bCS	bFFA	bGWO	bMFO	bMVO	bPSO	bAAs1
Breast cancer Wisconsin diagnostic	mean	0.8933	0.9074	0.8933	0.8931	0.9041	0.9038	0.8959	0.8981	0.9036
	±std	0.0034	0.0178	0.0025	0.0020	0.0082	0.005	0.0062	0.0164	0.0051
Climate model simulation crashes	mean	0.9422	0.9261	0.9415	0.9424	0.9357	0.9404	0.9382	0.9384	0.9433
	±std	0.0127	0.0101	0.0169	0.0131	0.0151	0.0135	0.0152	0.0115	0.0123
Cnae - 9	mean	0.8517	0.5490	0.7891	0.8038	0.7512	0.8526	0.8351	0.7965	0.8593
	±std	0.0258	0.0722	0.0292	0.0224	0.0353	0.0181	0.0295	0.0384	0.0162
Dermatology	mean	0.9492	0.7213	0.95195	0.9544	0.9587	0.9664	0.9617	0.9161	0.9713
	±std	0.0251	0.0654	0.02141	0.0224	0.0185	0.0161	0.0163	0.07054	0.0134
Ecoli	mean	0.8522	0.5843	0.8522	0.8522	0.8516	0.8522	0.8522	0.8217	0.8522
	±std	0.0047	0.1133	0.0000	0.0029	0.0796	0.0025	0.0012	0.0569	0.0062
Electrical grid	mean	0.7839	0.6127	0.7839	0.7841	0.7757	0.7839	0.7839	0.7539	0.7839
	±std	0.0089	0.0361	0.0005	0.0002	0.0096	0.0032	0.0033	0.0232	0.0001
Heart	mean	0.7678	0.6500	0.7732	0.7678	0.7446	0.7728	0.7761	0.6290	0.7667
	±std	0.0215	0.0552	0.0191	0.0223	0.0492	0.0140	0.0152	0.0771	0.0192
Hepatitis	mean	0.8060	0.8524	0.7726	0.8048	0.7786	0.8226	0.7964	0.8298	0.8190
	±std	0.0664	0.0245	0.0622	0.0681	0.0535	0.0492	0.0813	0.0614	0.0471
Hillvalley	mean	0.5174	0.5113	0.5185	0.5209	0.5177	0.5214	0.5229	0.5227	0.5529
	±std	0.0147	0.0164	0.0172	0.0130	0.0125	0.0121	0.0119	0.0148	0.0130
Image segmentation	mean	0.8815	0.6269	0.8810	0.8815	0.8755	0.8819	0.8787	0.8046	0.8824
	±std	0.0708	0.1223	0.0073	0.0072	0.0338	0.0078	0.0069	0.0705	0.0070
Ionosphere	mean	0.8742	0.8531	0.8658	0.8642	0.8667	0.8550	0.8681	0.8569	0.8664
	±std	0.0225	0.0242	0.0202	0.0288	0.0249	0.0214	0.0233	0.0235	0.0224
Letter recognition	mean	0.9538	0.3367	0.9497	0.9516	0.9242	0.9540	0.9540	0.9302	0.9540
	±std	0.0007	0.2281	0.0039	0.0025	0.0141	0.0004	0.0037	0.0162	0.002
Liver	mean	0.6525	0.5610	0.6525	0.6525	0.5949	0.6525	0.6525	0.6367	0.6525
	±std	0.0003	0.0572	0.0002	0.003	0.0662	0.0008	0.0036	0.0148	0.0035
Optical recognition handwritten digits	mean	0.9779	0.8682	0.9741	0.9739	0.9669	0.9776	0.9752	0.9735	0.9782
	±std	0.0021	0.0794	0.0036	0.0042	0.0053	0.0028	0.0031	0.0063	0.0034
Ozone level detection eight	mean	0.9260	0.9290	0.9267	0.9256	0.9277	0.9286	0.9257	0.9276	0.9274
	±std	0.0048	0.0037	0.0036	0.0052	0.0034	0.0031	0.0043	0.0034	0.0038
Pen based recognition handwritten digitis	mean	0.9891	0.7738	0.9889	0.9888	0.9870	0.9881	0.9884	0.9859	0.9883
	±std	0.0019	0.1536	0.0017	0.0015	0.0024	0.0003	0.0017	0.0067	0.0006
Planning relax	mean	0.6462	0.6505	0.6457	0.6511	0.6419	0.6435	0.6430	0.6366	0.6468
	±std	0.0047	0.0327	0.0023	0.0136	0.0285	0.0083	0.0088	0.0345	0.0084
Qsar Biodegradation	mean	0.8407	0.7877	0.8391	0.8419	0.8350	0.8436	0.8430	0.8259	0.8435
	±std	0.0163	0.0259	0.0123	0.0162	0.0186	0.0112	0.0116	0.0185	0.0102
Sonar	mean	0.7357	0.7192	0.7399	0.7202	0.7268	0.7362	0.7305	0.7188	0.7352
	±std	0.0354	0.0376	0.0382	0.0438	0.0394	0.0341	0.0385	0.0314	0.0301
Spambase	mean	0.9211	0.7720	0.9119	0.9118	0.8944	0.9216	0.9173	0.8493	0.9244
	±std	0.0074	0.0216	0.0071	0.0113	0.0217	0.0051	0.0725	0.0587	0.0051
Statlog Project landsat satellite	mean	0.8965	0.8868	0.8940	0.8935	0.8939	0.8972	0.8947	0.8986	0.8995
	±std	0.0064	0.0103	0.0050	0.0069	0.0079	0.0044	0.0053	0.0048	0.0032
Thyroid Disease	mean	0.9798	0.9350	0.9807	0.9798	0.9736	0.9789	0.9810	0.9500	0.9793
	±std	0.0024	0.0043	0.0020	0.0037	0.0048	0.0025	0.0017	0.0142	0.0031
Waveform	mean	0.8089	0.6242	0.8046	0.8077	0.7932	0.8010	0.8084	0.7953	0.8094
	±std	0.0077	0.0744	0.009	0.0087	0.0124	0.0065	0.0081	0.0151	0.0069
WDBC	mean	0.8945	0.9096	0.8935	0.8992	0.9081	0.9027	0.8962	0.8998	0.9046
	±std	0.0043	0.0182	0.002	0.0042	0.0086	0.0066	0.0062	0.0171	0.0048
Weight lifting exercises	mean	0.9894	0.9536	0.9890	0.9885	0.9803	0.9900	0.9897	0.9745	0.9902
	±std	0.0043	0.0142	0.004	0.0046	0.0062	0.0029	0.0024	0.0082	0.0027
En iyi sonuç sayısı		4	4	3	4	0	5	6	0	12
Friedman sıralama değerleri		4.16	7.36	5.06	5.02	6.44	3.34	4.16	6.66	2.8
Genel sıralama		4	9	6	5	7	2	3	8	1

Çizelge 4.18'deki sonuçlar incelendiğinde yirmi beş veri setinden on ikisinde (Climate model simulation crashes, Cnae-9, Dermatology, Ecoli, Hillvalley, Image segmentation, Letter recognition, Liver, Optical recognition handwritten digits, Spambase, Statlog project Landsat satellite ve Weight lifting exercises) bAAAs1 versiyonun, kıyaslanan diğer yöntemlerden daha iyi sonuçlar elde ettiği görülmektedir. Diğer bir tabirle geliştirilen ikili kaotik bAAAs1 yöntemi, veri setlerinin %48'inde en iyi değeri elde ederek, kıyaslanan yöntemlerden daha iyi performans göstermiştir.

MVO yöntemi ise bAAAs1 yöntemine karşı kısmen rekabetçi bir yaklaşım sergileyerek, yirmi beş veri setinin yedisinde (Ecoli, Heart, Hillvalley, Letter recognition, Liver ve Thyroid Disease) en iyi performansı sergilemiştir. MFO yaklaşımı beş veri setinde, MVO yaklaşımı altı veri setinde en iyi değeri elde etmiş olmasına rağmen MVO yaklaşımı ortalama sıralamada daha başarılıdır. Ortalama sıralamayı sırasıyla WOA ve CS yaklaşımları takip etmektedir.

Çizelge 4.19'da ikili algoritmaların veri setleri üzerinde uygunluk fonksiyonu değerlerine göre otuz bağımsız çalışmada elde ettiği sonuçların ortalama ve standart sapma değerleri gösterilmektedir.

Çizelge 4.19'daki sonuçlar değerlendirildiğinde, bAAAs1 yaklaşımı sekiz veri setinde en iyi sonuçları elde ederek, kıyaslanan ikili algoritmalarından daha iyi performans sergilemiştir. MFO ve MVO yöntemi de tezde geliştirilen yönteme uygunluk değerleri açısından rekabetçi bir yaklaşım göstermiştir. Ortalama sıralama sırasıyla WOA, CS ve FFA algoritmalarını takip etmektedir.

Tezde geliştirilen ikili kaotik bAAAs1 ayrıca Cnae-9, Dermatology, Ecoli, Letter recognition, Liver, Optical recognition handwritten digits, Spambase ve Statlog Project Landsat Satellite veri setleri için en iyi uygunluk elde etmiştir. GWO ve PSO hiçbir veri setinde en iyi değere ulaşamamıştır. Bununla birlikte BAT yaklaşımı dört veri setinde iyi değeri elde etse de ortalama sıralamada son sıradadır.

Çizelge 4.20'de ikili algoritmaların veri setleri üzerinde indirilen öznitelik sayısı değerlerine göre otuz bağımsız çalışmada elde ettiği sonuçların ortalama ve standart sapma değerleri gösterilmektedir.

Çizelge 4.19. İkili algoritmaların veri setleri üzerinde uygunluk fonksiyonun otuz bağımsız çalışma ortalamaları ve standart sapma değerleri

Veri Seti		bWOA	bbAT	bCS	bFFA	bGWO	bMFO	bMVO	bPSO	bAAAs1
Breast cancer Wisconsin diagnostic	mean	0.1084	0.0965	0.1083	0.1091	0.1000	0.0995	0.1058	0.1068	0.1002
	±std	0.0033	0.0177	0.0025	0.0022	0.0079	0.0057	0.0058	0.0167	0.0055
Climate model simulation crashes	mean	0.0619	0.0779	0.0621	0.0621	0.0691	0.0647	0.0659	0.0682	0.0626
	±std	0.0123	0.0105	0.0159	0.0134	0.0149	0.0138	0.0148	0.0115	0.0121
Cnae - 9	mean	0.1549	0.4504	0.2138	0.1999	0.2514	0.1534	0.1694	0.2088	0.1474
	±std	0.0255	0.0711	0.0289	0.0227	0.0384	0.0187	0.0294	0.0386	0.0160
Dermatology	mean	0.0552	0.2803	0.0523	0.0505	0.0461	0.0394	0.0434	0.0901	0.0349
	±std	0.0249	0.0636	0.0211	0.0226	0.0180	0.0165	0.0162	0.0694	0.0137
Ecoli	mean	0.1547	0.4149	0.1547	0.1547	0.1558	0.1547	0.1547	0.1849	0.1547
	±std	0.0028	0.1119	0.0028	0.0028	0.0073	0.0028	0.0028	0.0498	0.0028
Electrical grid	mean	0.2203	0.3867	0.2203	0.2201	0.2295	0.2203	0.2203	0.2509	0.2203
	±std	0.0113	0.0354	0.0113	0.0007	0.0092	0.0113	0.0113	0.0223	0.0113
Heart	mean	0.2344	0.3517	0.2293	0.2344	0.2573	0.2297	0.2265	0.3744	0.2357
	±std	0.0205	0.0550	0.0186	0.0212	0.0485	0.0141	0.0146	0.0763	0.0185
Hepatitis	mean	0.1945	0.1518	0.2272	0.1962	0.2229	0.1793	0.2038	0.1748	0.1835
	±std	0.0654	0.0234	0.0618	0.0681	0.0528	0.0487	0.0805	0.0613	0.0471
Hillvalley	mean	0.4822	0.4882	0.4812	0.4793	0.4824	0.4801	0.4777	0.4795	0.4793
	±std	0.0145	0.0167	0.0172	0.0128	0.0122	0.0125	0.0115	0.0143	0.0132
Image segmentation	mean	0.1214	0.3733	0.1217	0.1217	0.1286	0.1211	0.1236	0.1997	0.1215
	±std	0.0069	0.1207	0.0069	0.0069	0.0328	0.0069	0.0061	0.0701	0.0076
Ionosphere	mean	0.1275	0.1506	0.1359	0.1382	0.1362	0.1488	0.1343	0.1480	0.1375
	±std	0.0227	0.0246	0.0200	0.0288	0.0246	0.0212	0.0237	0.0232	0.0221
Letter recognition	mean	0.0531	0.6594	0.0572	0.0555	0.0827	0.0529	0.0529	0.0770	0.0529
	±std	0.0008	0.2254	0.0036	0.0025	0.0139	0.0035	0.0035	0.0163	0.0035
Liver	mean	0.3520	0.4396	0.3520	0.3520	0.4080	0.3520	0.3520	0.3686	0.3520
	±std	0.0169	0.0566	0.0019	0.0169	0.0650	0.0169	0.0169	0.0146	0.0169
Optical recognition handwritten digits	mean	0.0299	0.1342	0.0316	0.0321	0.0385	0.0296	0.0311	0.0337	0.0295
	±std	0.0024	0.0779	0.0030	0.0040	0.0046	0.0028	0.0031	0.0063	0.0029
Ozone level detection eight	mean	0.0781	0.0744	0.0770	0.0786	0.0765	0.0771	0.0789	0.0787	0.0781
	±std	0.0052	0.0037	0.0036	0.0054	0.0033	0.0034	0.0041	0.0032	0.0038
Pen based recognition handwritten digitis	mean	0.0185	0.2268	0.0187	0.0189	0.0206	0.0197	0.0192	0.0218	0.0194
	±std	0.0016	0.1514	0.0019	0.0016	0.0026	0.0011	0.0013	0.0065	0.0012
Planning relax	mean	0.3547	0.3512	0.3553	0.3497	0.3588	0.3574	0.3579	0.3668	0.3541
	±std	0.0043	0.0314	0.0031	0.0132	0.0285	0.0086	0.0079	0.0340	0.0089
Qsar Biodegradation	mean	0.1634	0.2146	0.1639	0.1618	0.1685	0.1618	0.1613	0.1795	0.1621
	±std	0.0156	0.0253	0.0128	0.0158	0.0180	0.0115	0.0116	0.0183	0.0105
Sonar	mean	0.2661	0.2825	0.2618	0.2819	0.2751	0.2344	0.2719	0.2854	0.2689
	±std	0.0357	0.0367	0.0376	0.0434	0.0390	0.0344	0.0385	0.0318	0.0300
Spambase	mean	0.0855	0.2300	0.0926	0.0932	0.1097	0.0848	0.0880	0.1564	0.0827
	±std	0.0069	0.0208	0.0070	0.0113	0.0212	0.0057	0.0069	0.0575	0.0053
Statlog Project Landsat Satellite	mean	0.1091	0.1165	0.1105	0.1109	0.1103	0.1090	0.1103	0.1077	0.1073
	±std	0.0058	0.0096	0.0050	0.0068	0.0075	0.0045	0.0053	0.0042	0.0031
Thyroid Disease	mean	0.0229	0.0698	0.0216	0.0233	0.0306	0.0251	0.0213	0.0551	0.0250
	±std	0.0027	0.0039	0.0022	0.0032	0.0050	0.0030	0.0015	0.0144	0.0037
Waveform	mean	0.1970	0.3757	0.1994	0.1966	0.2104	0.1956	0.1961	0.2102	0.1963
	±std	0.0074	0.0731	0.0093	0.0087	0.0121	0.0065	0.0083	0.0150	0.0063
WDBC	mean	0.1073	0.0950	0.1080	0.1081	0.0960	0.1005	0.1054	0.1051	0.0992
	±std	0.0041	0.0187	0.0025	0.0040	0.0083	0.0060	0.0058	0.0174	0.0046
Weight lifting exercises	mean	0.0160	0.0504	0.0155	0.0166	0.0245	0.0162	0.0157	0.0322	0.0168
	±std	0.0043	0.0135	0.0046	0.0041	0.0063	0.0027	0.0028	0.0083	0.0027
En iyi sonuç sayısı		5	4	4	4	0	5	7	0	8
Friedman sıralama değerleri		3.96	7.36	4.60	4.82	6.46	3.50	3.94	6.96	3.40
Genel sıralama		4	9	5	6	7	2	3	8	1

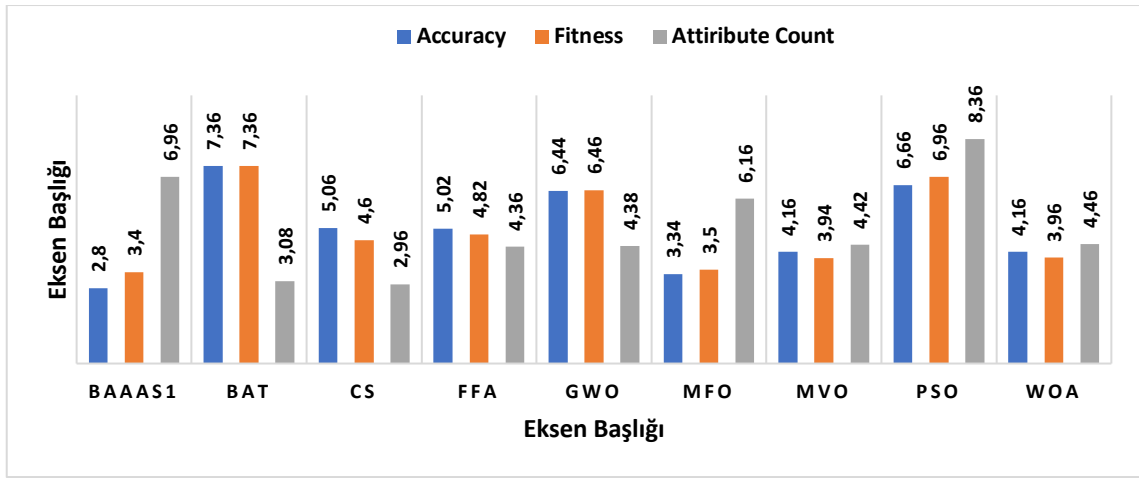
Çizelge 4.20. İkili algoritmaların veri setleri üzerinde seçilen öznitelik sayısı açısından otuz bağımsız çalışma ortalamaları ve standart sapma değerleri

Veri Seti		bWOA	bbAT	bCS	bFFA	bGWO	bMFO	bMVO	bPSO	bAAAs1
Breast cancer Wisconsin diagnostic	mean	7.9	14.1	7.8	9.4	14.8	12.3	8.0	17.2	13.7
	±std	1.3	3.1	1.0	1.1	2.6	0.9	1.2	2.5	1.4
Climate model simulation crashes	mean	7.9	8.1	7.0	8.6	9.3	9.6	8.1	12.2	11.0
	±std	1.6	2.9	1.1	1.5	1.8	1.2	1.6	2.3	1.3
Cnae - 9	mean	689.6	336.9	432.7	483.6	428.9	637.4	525.8	627.6	691.5
	±std	25.17	52.3	16.6	16.5	13.1	13.4	13.7	14.8	11.3
Dermatology	mean	16.0	14.6	15.6	17.5	17.5	20.3	18.3	23.3	21.6
	±std	2.3	4.2	2.2	2.2	3.0	2.2	2.8	2.3	1.7
Ecoli	mean	5.0	2.1	5.0	5.0	5.3	5.0	5.0	5.0	5.0
	±std	0.0	0.7	0.0	0.0	0.6	0.0	0.0	1.0	0.0
Electrical grid	mean	7.0	3.6	7.0	7.0	8.1	7.0	7.0	7.9	7.0
	±std	0.0	1.5	0.0	0.0	11.3	0.0	0.0	1.5	0.0
Heart	mean	5.4	6.2	5.7	5.4	5.3	5.7	5.7	8.5	5.7
	±std	0.9	1.9	0.7	0.8	1.3	0.7	0.6	1.5	0.6
Hepatitis	mean	4.3	10.1	3.7	5.2	6.6	6.7	4.1	11.3	7.8
	±std	0.7	2.8	0.9	1.0	1.4	0.8	0.8	2.2	1.1
Hillvalley	mean	44.1	43.6	45.2	49.8	48.4	62.0	52.4	69.0	67.4
	±std	4.3	7.4	4.7	5.0	4.5	5.9	4.8	4.7	7.5
İmage segmentation	mean	7.4	7.1	7.1	7.8	9.5	7.6	6.4	11.2	9.2
	±std	0.7	2.1	0.8	0.8	2.2	0.7	0.5	1.8	1.1
Ionosphere	mean	9.7	17.0	10.1	12.3	13.7	17.3	12.0	20.9	17.4
	±std	1.9	4.0	2.9	2.2	2.6	2.6	2.0	2.4	3.6
Lettter recognition	mean	11.1	4.1	11.2	11.4	11.6	11.0	11.0	12.0	11.0
	±std	0.2	2.0	0.9	1.0	1.3	0.0	0.0	1.3	0.0
Liver	mean	4.0	2.5	4.0	4.0	3.5	4.0	4.0	4.5	4.0
	±std	0.0	1.1	0.0	0.0	0.5	0.0	0.0	0.8	0.0
Optical recognition handwritten digits	mean	51.0	23.7	37.3	39.2	35.9	47.0	41.4	47.3	50.1
	±std	3.1	5.1	3.3	3.0	3.3	2.3	3.2	3.5	2.6
Ozone level detection eight	mean	34.1	29.0	31.4	35.1	34.7	45.2	37.7	50.0	44.2
	±std	6.3	6.1	4.3	3.6	4.9	4.0	3.6	3.1	8.0
Pen based recognition handwritten digitis	mean	11.6	4.3	11.6	11.7	11.5	11.8	11.7	11.8	11.8
	±std	0.5	1.6	0.4	0.4	1.2	0.4	0.4	1.4	0.4
Planning relax	mean	4.9	5.7	5.0	4.7	4.7	5.0	4.9	7.7	4.9
	±std	0.2	1.6	0.1	0.4	1.2	0.1	0.2	1.7	0.3
Qsar Biodegradation	mean	22.7	17.8	18.4	20.8	20.6	27.9	23.3	28.4	29.1
	±std	5.0	3.0	2.3	3.2	2.6	2.0	2.7	3.2	1.7
Sonar	mean	26.3	26.8	25.7	28.9	27.4	38.4	30.4	41.5	40.0
	±std	4.5	4.2	3.4	3.2	3.4	2.9	3.3	4.0	5.7
Spambase	mean	41.1	23.7	30.0	32.9	28.7	40.4	34.3	40.1	43.9
	±std	4.7	5.3	2.8	3.4	3.7	2.3	2.9	3.5	2.4
Statlog Project landsat satellite	mean	23.4	15.4	19.6	19.3	18.5	25.6	21.0	25.6	27.6
	±std	3.7	4.3	2.4	3.5	2.7	2.3	2.3	2.2	1.9
Thyroid Disease	mean	5.8	10.9	5.1	6.6	8.8	8.2	5.0	11.3	9.0
	±std	1.0	2.2	0.9	1.4	1.8	1.2	0.9	2.2	1.3
Waveform	mean	30.4	14.4	23.5	24.5	21.9	29.1	25.0	29.4	29.6
	±std	2.9	4.1	2.9	2.4	3.4	2.0	2.2	3.4	1.9
WDBC	mean	8.3	16.0	7.5	9.6	14.5	12.3	7.8	17.1	14.0
	±std	1.2	3.3	1.3	1.3	2.6	0.9	1.2	2.3	1.4
Weight lifting exercises	mean	29.4	23.7	24.4	27.7	26.6	33.4	29.0	37.1	37.9
	±std	6.3	4.8	3.3	3.5	2.8	3.8	3.4	2.7	2.2
En iyi sonuç sayısı		1	15	5	1	2	0	2	0	0
Friedman sıralama değerleri		4.46	3.08	2.96	4.36	4.38	6.16	4.42	8.36	6.96
Genel sıralama		3	2	1	4	5	7	6	9	8

Çizelge 4.20’de, seçilen öznitelik değerlerinin sonuçları göz önüne alındığında, bAAAs1 yöntemi diğer yöntemlerle rekabet halinde olmasına rağmen en iyi sıralamayı

elde edememiştir. Bunun temel nedeni, geliştirilen yöntemin sınıflandırma başarısını en üst düzeye çıkaracak şekilde öznitelikleri seçmeyi hedeflemesidir.

Çizelge 4.18–20'deki sonuçlara göre tezde geliştirilen bAAAs1 yaklaşımı öznitelik seçimi probleminde diğer bir tabirle ideal öznitelik alt kümesini belirlenmesinde etkin bir performans sergilemiştir. bAAAs1'den sonra en performanslı ikinci algoritma MFO'dur. Yöntemlerin genel performansını değerlendirmek için test doğruluğu, uygunluk fonksiyonu ve seçilen öznitelik sayısını üzerinden ortalama sıralama sonuçları gösteren çubuk diyagram Şekil 4.12'de sunulmaktadır.



Şekil 4.12. Öznitelik seçimi probleminde kıyaslanan yaklaşımların deneysel çalışma sonuçlarına göre oluşturulan çubuk diyagramı

Şekil 4.12'deki çubuk diyagramı incelendiğinde, bAAAs1 yöntemi nitelik sayısı açısından iyi performans göstermemiş olsa da test doğruluğu bakımında oldukça etkili bir performans sergilemiştir. Seçilen öznitelik sayısı ile test doğruluğunu başarısını birlikte gösteren uygunluk değeri, yöntemlerin gerçek performansını sunmaktadır. bAAAs1 yöntemi uygunluk fonksiyonu değeri bakımında değerlendirildiğinde ortalama sıralamada en başarılı yöntemdir.

BAT ve CS algoritmaları öznitelik sayısı açısından iyi başarı değeri göstermelerine rağmen, test doğruluğundaki eksikliklerinden dolayı etkili bir çözüm sağlayamamışlardır. Performans için önemli olan, uygunluk fonksiyonunda en iyi değere ulaşmak için test doğruluğu başarısını ve seçilen öznitelik sayısını dengelemektir.

Geliştirilen bAAAs1 yönteminin başarılı sonuçları, kaotik AAA yaklaşımının ikili uzayda keşif ve sömürü sürecini etkin bir şekilde dengeleyebildiğini göstermektedir. Bu sayede diğer yöntemlere göre daha başarılı sonuçlar elde etmiştir.

Çizelge 4.19’da gösterilen yirmi beş veri seti üzerinde kıyaslanan algoritmaların otuz bağımsız çalışma sonucunda oluşan uygunluk değerleri sonuçları birbirine oldukça yakındır. Bu algoritmaların arasındaki farkın anlamlı olup olmadığını tespit etmek için tezin tüm deneysel çalışmalarında yapıldığı gibi bu bölümde de Wilcoxon işaretli sıralar testi yapılmıştır. Geliştirilen kaotik ikili AAA yaklaşımı ile kıyaslanan diğer algoritmaların sonuçlarının Wilcoxon işaretli sıralar testi sonuçları test başarısı ve uygunluk değerlendirmesi açısından sırasıyla Çizelge 4.21-4.22 ve Çizelge 4.23-4.24 de verilmiştir. İlgili çizelgelerde yapılan Wilcoxon testine ait p-Value, T ve W değerleri sunulmaktadır. Çizelgelerin son satırında Wilcoxon işaretli sıralar testi sonucunda karşılaştırılan iki yöntem için eşitlik, kazanılan ve kaybedilen değerlerin toplamı ayrı ayrı sunulmuştur. Çizelge 4.21’de tez kapsamında geliştirilen bAAAs1 yöntemi ile BAT, CS, FFA ve GWO yöntemleri arasında test doğruluğu üzerinden yapılan Wilcoxon işaretli sıralar testi sonuçları gösterilmektedir.

Çizelge 4.21. Geliştirilen bAAAs1 yöntemi ile kıyaslanan diğer yöntemlerin test doğruluğu açısından Wilcoxon işaretli sıralar testi sonuçları (BAT, CS, FFA ve GWO için)

Datasets	bAAAs1 - BAT			bAAAs1 - CS			bAAAs1 - FFA			bAAAs1 -GWO		
	p-Value	T	S	p-Value	T	S	p-Value	T	S	p-Value	T	S
breast-cancer-wisconsin-diagnostic	9.44E-02	151.5	0	5.70E-06	398	1	5.22E-06	351	1	8.57E-01	121	0
climate-model-simulation-crashes	3.36E-05	295	1	5.03E-01	265	0	7.10E-01	163	0	4.97E-02	308	1
cnae	1.73E-06	465	1	1.73E-06	465	1	1.72E-06	465	1	1.73E-06	465	1
dermatology	1.72E-06	465	1	2.92E-04	256.5	1	2.50E-03	274.5	1	2.07E-03	257.5	1
ecoli	1.54E-06	465	1	1.00E+00	0	0	1.00E+00	0	0	1.00E+00	11	0
electrical-grid	1.73E-06	465	1	1.00E+00	0	0	2.50E-01	0	0	1.09E-04	420.5	1
heart	1.71E-06	465	1	2.88E-01	69	0	9.08E-01	123	0	2.34E-02	322	1
hepatitis	1.85E-03	22.5	2	2.39E-03	294.5	1	6.25E-01	141.5	0	3.20E-03	271	1
hillvalley	1.06E-02	276	1	2.02E-01	242	0	7.80E-01	186.5	0	1.17E-01	254	0
image-segmentation	1.72E-06	465	1	3.71E-01	93.5	0	6.55E-01	85	0	3.63E-01	141.5	0
ionosphere	7.50E-03	300	1	9.31E-01	221.5	0	7.45E-01	232.5	0	9.62E-01	191	0
letter-recognition	1.73E-06	465	1	1.72E-05	300	1	5.29E-05	231	1	1.73E-06	465	1
liver	3.04E-06	459	1	1.00E+00	0	0	1.00E+00	0	0	3.50E-04	147	1
optical-recognition-handwritten-digits	1.73E-06	465	1	1.01E-04	373	1	5.11E-04	355.5	1	1.91E-06	464	1
ozone-level-detection-eight	7.42E-02	115	0	6.55E-01	207.5	0	2.06E-01	241.5	0	6.47E-01	145.5	0
pen-based-recognition-handwritten-digits	1.73E-06	465	1	1.02E-01	68.5	0	1.77E-01	69	0	1.47E-02	271.5	1
planning-relax	3.50E-01	139	0	7.66E-01	17.5	0	3.05E-01	20	0	4.26E-01	191.5	0
qsar-biodegradation	1.72E-06	465	1	1.45E-01	249.5	0	9.59E-01	230	0	2.58E-02	320.5	1
sonar	8.06E-02	279.5	0	7.18E-01	174	0	2.03E-01	225.5	0	3.78E-01	225.5	0
spambase	1.73E-06	465	1	6.31E-06	452	1	3.54E-05	433.5	1	1.73E-06	465	1
statlog-project-landsat-satellite	8.83E-06	423	1	1.22E-04	326.5	1	5.10E-04	355.5	1	1.71E-03	362.5	1
thyroid-disease-ann-thyroid	1.72E-06	465	1	1.09E-01	103	0	7.13E-01	200.5	0	8.85E-05	423	1
waveform	1.73E-06	465	1	6.35E-02	284.5	0	4.97E-01	265.5	0	1.80E-05	441	1
wdbc	3.99E-02	133	2	4.29E-06	351	1	6.88E-06	325	1	8.80E-02	82.5	0
weight-lifting-exercises	1.73E-06	465	1	3.25E-01	263	0	1.18E-01	254	0	4.78E-06	429	1
Kazanılan/ Eşitlik/ Kaybedilen	19/4/2			9/16/0			8/17/0			16/9/0		

Çizelge 4.22’de tez kapsamında geliştirilen bAAAs1 yöntemi ile MFO, MVO, PSO ve WOA yöntemleri arasında test doğruluğu üzerinden yapılan Wilcoxon işaretli sıralar testi sonuçları gösterilmektedir.

Çizelge 4.22. Geliştirilen bAAAs1 yöntemi ile kıyaslanan diğer yöntemlerin test doğruluğu açısından Wilcoxon işaretli sıralar testi sonuçları (MFO, MVO, PSO ve WOA için)

Datasets	bAAAs1 - MFO			bAAAs1 - MVO			bAAAs1 - PSO			bAAAs1 - WOA		
	p-Value	T	W	p-Value	T	W	p-Value	T	W	p-Value	T	W
breast-cancer-wisconsin-diagnostic	1.00E+00	61	0	9.90E-04	281.5	1	1.42E-01	186	0	8.02E-06	322	1
climate-model-simulation-crashes	4.26E-01	222	0	1.69E-01	198	0	1.36E-01	234	0	7.63E-01	201.5	0
cnae	6.60E-02	265.5	0	2.79E-04	385.5	1	1.73E-06	465	1	5.16E-01	247.5	0
dermatology	2.64E-01	189	0	1.91E-02	214.5	1	2.35E-04	299	1	2.98E-04	361.5	1
ecoli	1.00E+00	0	0	1.00E+00	0	0	6.36E-05	210	1	1.00E+00	0	0
electrical-grid	1.00E+00	0	0	1.00E+00	0	0	2.88E-06	460	1	1.00E+00	0	0
heart	1.43E-01	46	0	2.67E-02	30.5	2	1.89E-06	464	1	9.75E-01	137	0
hepatitis	7.40E-01	162.5	0	2.51E-01	205	0	3.48E-01	162	0	4.59E-01	204.5	0
hillvalley	3.62E-01	226.5	0	8.55E-01	211	0	7.49E-01	217	0	1.27E-01	270	0
image-segmentation	7.81E-01	27	0	3.89E-02	104	1	1.51E-04	392.5	1	8.17E-01	67.5	0
ionosphere	1.49E-02	235	1	8.08E-01	153.5	0	6.47E-02	248	0	1.18E-01	134.5	0
letter-recognition	1.00E+00	0	0	1.00E+00	0	0	1.73E-06	465	1	5.00E-01	3	0
liver	1.00E+00	0	0	1.00E+00	0	0	1.50E-05	297	1	1.00E+00	0	0
optical-recognition-handwritten-digits	4.90E-01	266	0	2.70E-03	356	1	3.25E-04	383.5	1	7.45E-01	232.5	0
ozone-level-detection-eight	1.32E-01	126.5	0	9.18E-02	259	0	5.25E-01	150.5	0	3.25E-01	263	0
pen-based-recognition-handwritten-digits	4.67E-01	42	0	8.49E-01	64.5	0	1.63E-01	247	0	5.42E-03	27.5	2
planning-relax	2.66E-01	22	0	2.03E-01	22.5	0	2.45E-01	190.5	0	1.00E+00	12	0
qsar-biodegradation	8.01E-01	178.5	0	9.64E-01	201	0	8.04E-04	372.5	1	6.65E-01	253.5	0
sonar	9.52E-01	191.5	0	8.10E-01	199	0	3.11E-02	260	1	8.19E-01	213	0
spambase	5.43E-02	326	0	1.24E-03	389.5	1	2.60E-06	461	1	6.56E-02	322	0
statlog-project-landsat-satellite	5.57E-02	287	0	6.36E-04	398.5	1	4.28E-01	271	0	3.76E-02	333.5	1
thyroid-disease-ann-thyroid	4.80E-01	234	0	1.82E-02	67.5	2	1.73E-06	465	1	5.07E-01	149.5	0
waveform	7.13E-01	161	0	5.66E-01	244	0	1.01E-04	421.5	1	7.81E-01	246	0
wdbc	2.25E-01	73	0	2.66E-04	294	1	1.76E-01	298	0	5.72E-06	374.5	1
weight-lifting-exercises	9.82E-01	204	0	2.39E-01	254.5	0	2.56E-06	435	1	8.80E-01	224.5	0
Kazanılan/ Eşitlik/ Kaybedilen	1/24/0			8/15/2			15/10/0			4/20/1		

Çizelge 4.23’te tez kapsamında geliştirilen bAAAs1 yöntemi ile BAT, CS, FFA ve GWO yöntemleri arasında uygunluk fonksiyonu üzerinden yapılan Wilcoxon işaretli sıralar testi sonuçları gösterilmektedir.

Çizelge 4.23. Geliştirilen bAAAs1 yöntemi ile kıyaslanan diğer yöntemlerin uygunluk fonksiyonu açısından Wilcoxon işaretli sıralar testi sonuçları (BAT, CS, FFA ve GWO)

Datasets	bAAAs1 - BAT			bAAAs1 - CS			bAAAs1 - FFA			bAAAs1 - GWO		
	p-Value	T	W	p-Value	T	W	p-Value	T	W	p-Value	T	W
breast-cancer-wisconsin-diagnostic	1.36E-01	305	0	1.47E-05	22	1	4.63E-06	10	1	8.37E-01	227	0
climate-model-simulation-crashes	3.45E-05	26	1	8.61E-01	224	0	8.71E-01	210	0	7.52E-02	146	0
cnae	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	0	1	1.73E-06	0	1
dermatology	1.73E-06	0	1	2.50E-03	85.5	1	4.16E-03	85	1	7.57E-03	94	1
ecoli	1.66E-06	0	1	1.00E+00	0	0	1.00E+00	0	0	4.34E-01	46	0
electrical-grid	1.73E-06	0	1	1.00E+00	0	0	2.55E-01	6	0	3.11E-05	30	1
heart	1.73E-06	0	1	2.92E-01	121	0	7.37E-01	149	0	3.22E-02	118.5	1
hepatitis	1.17E-02	334	2	8.71E-03	105	1	9.57E-01	215	0	9.44E-03	97.5	1
hillvalley	4.95E-02	137	1	3.71E-01	189	0	8.13E-01	244	0	2.80E-01	180	0
image-segmentation	1.73E-06	0	1	4.38E-01	206	0	2.83E-01	267	0	5.93E-01	206.5	0
ionosphere	8.46E-03	104.5	1	8.77E-01	240	0	9.92E-01	233	0	8.77E-01	240	0
letter-recognition	1.73E-06	0	1	1.72E-05	0	1	5.29E-05	0	1	1.73E-06	0	1
liver	3.08E-06	6	1	1.00E+00	0	0	1.00E+00	0	0	4.97E-04	6	1
optical-recognition-handwritten-digits	1.73E-06	0	1	3.68E-02	131	1	1.48E-02	114	1	2.60E-06	4	1
ozone-level-detection-eight	8.94E-04	394	2	1.41E-01	304	0	6.00E-01	207	0	1.03E-01	293	0
pen-based-recognition-handwritten-digits	1.73E-06	0	1	1.06E-01	162	0	1.77E-01	141	0	4.38E-02	114.5	1
planning-relax	7.50E-01	248	0	7.19E-01	11.5	0	1.21E-01	51	0	3.14E-01	171	0
qsar-biodegradation	1.73E-06	0	1	5.30E-01	202	0	5.44E-01	262	0	9.78E-02	152	0
sonar	1.31E-01	159	0	4.91E-01	266	0	3.60E-01	188	0	6.58E-01	211	0
spambase	1.73E-06	0	1	3.41E-05	31	1	8.91E-05	42	1	2.13E-06	2	1
statlog-project-landsat-satellite	2.37E-05	27	1	4.68E-03	95	1	6.03E-03	99	1	1.06E-01	154	0
thyroid-disease-ann-thyroid	1.73E-06	0	1	2.08E-04	389	2	6.26E-02	323	0	2.05E-04	52	1
waveform	1.73E-06	0	1	1.99E-01	170	0	9.10E-01	227	0	8.19E-05	41	1
wdbc	4.27E-02	331	2	4.64E-06	10	1	7.05E-06	10	1	3.77E-02	333.5	2
weight-lifting-exercises	1.73E-06	0	1	1.47E-01	303	0	6.44E-01	255	0	1.02E-05	18	1
Kazanılan/ Eşitlik/ Kaybedilen	19/3/3			9/15/1			8/17/0			13/11/1		

Çizelge 4.24'te tez kapsamında geliştirilen bAAAs1 yöntemi ile MFO, MVO, PSO ve WOA yöntemleri arasında uygunluk fonksiyonu üzerinden yapılan Wilcoxon işaretli sıralar testi sonuçları gösterilmektedir.

Wilcoxon işaretli sıralar testi ile bAAAs1 yaklaşımının, literatürdeki popüler yöntemler ile bire bir kıyaslaması yapılarak detaylı bir performans analizi ortaya konulmuştur. Çizelge 4.21 ile Çizelge 4.24 arasındaki sonuçlar analiz edildiğinde, bAAAs1 yaklaşımı kıyaslanan bütün yöntemlerden daha yüksek kazanan (win) değerine sahiptir. BAT, CS, FFA, GWO, MFO, MVO, PSO ve WOA algoritmaları bazı veri kümelerinde AAA'ya karşı uygunluk değeri açısından rekabetçi sonuçlar etmiştir ancak toplam kazanan değerinde AAA'yı geçememiştir.

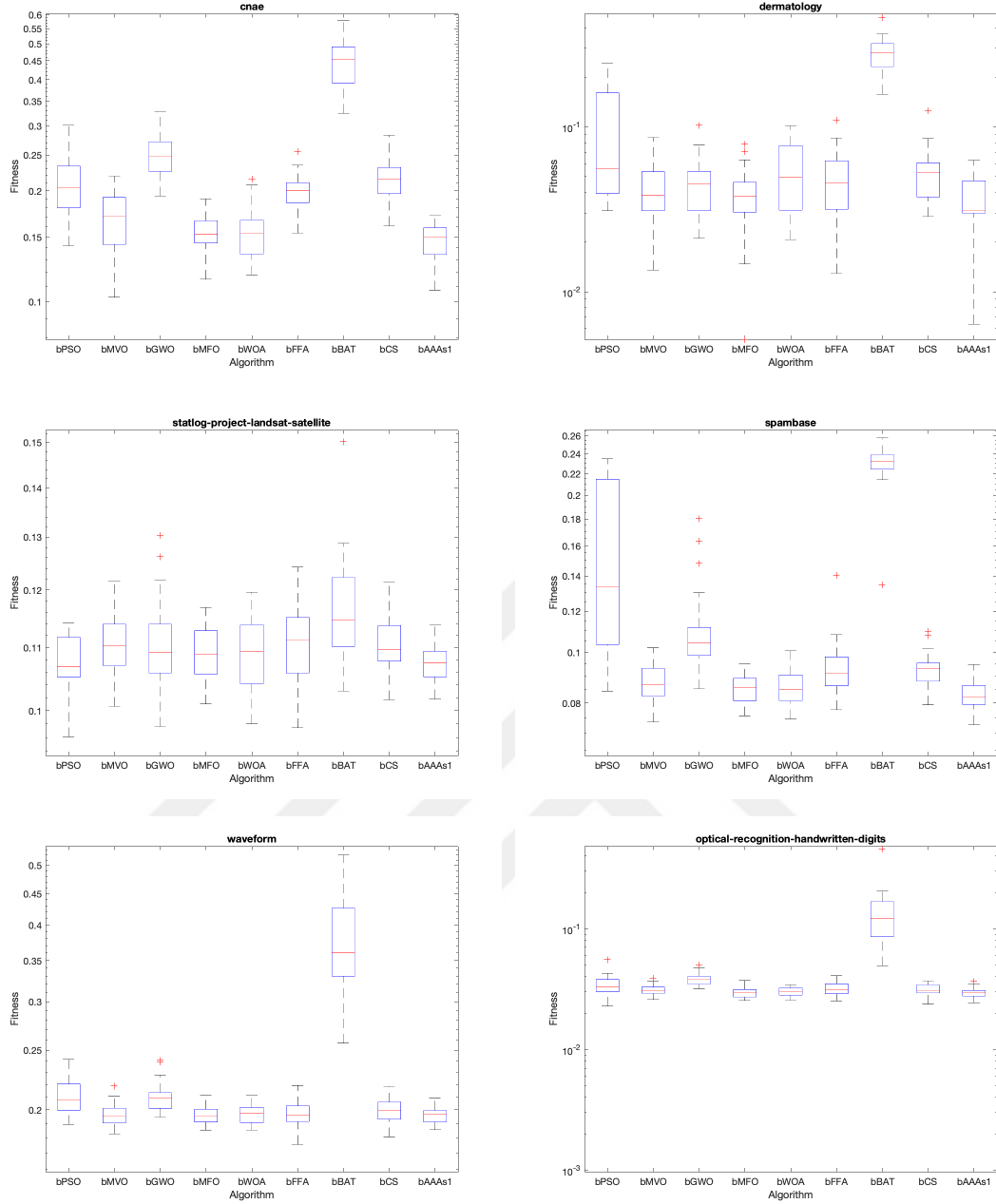
Test doğruluğu açısından (Çizelge 4.21-4.22) BAT, MVO ve WOA algoritmaları sırasıyla iki (hepatitis ve wdbc), iki (heart ve waveform) ve bir (pen-based recognition handwritten digits) veri setinde AAA'dan daha iyi sonuçlar elde etmiştir.

Çizelge 4.24. Geliştirilen bAAAs1 yöntemi ile kıyaslanan diğer yöntemlerin uygunluk fonksiyonu açısından Wilcoxon işaretli sıralar testi sonuçları (MFO, MVO, PSO ve WOA)

Datasets	bAAAs1 - MFO			bAAAs1 - MVO			bAAAs1 - PSO			bAAAs1 - WOA		
	p-Value	T	W	p-Value	T	W	p-Value	T	W	p-Value	T	W
breast-cancer-wisconsin-diagnostic	1.87E-01	296.5	0	3.13E-03	89	1	5.29E-02	118	0	2.09E-05	26	1
climate-model-simulation-crashes	6.42E-01	196	0	3.05E-01	158	0	1.22E-01	146	0	8.45E-01	242	0
cnae	9.37E-02	151	0	5.28E-04	64	1	1.73E-06	0	1	5.04E-01	200	0
dermatology	3.71E-01	189	0	5.06E-02	137.5	0	1.91E-04	45	1	6.63E-04	67	1
ecoli	1.00E+00	0	0	1.00E+00	0	0	1.78E-05	20	1	1.00E+00	0	0
electrical-grid	1.00E+00	0	0	1.00E+00	0	0	2.88E-06	5	1	1.00E+00	0	0
heart	1.34E-01	145	0	1.96E-02	125.5	2	1.92E-06	1	1	9.76E-01	139	0
hepatitis	3.87E-01	274.5	0	4.91E-01	199	0	7.19E-01	250	0	8.77E-01	240	0
hillvalley	7.50E-01	217	0	7.50E-01	248	0	7.50E-01	217	0	3.71E-01	189	0
image-segmentation	7.26E-02	263.5	0	7.21E-01	201	0	1.67E-04	49.5	1	2.01E-01	276.5	0
ionosphere	2.18E-02	121	1	4.11E-01	272.5	0	2.85E-02	126	1	5.71E-02	325	0
letter-recognition	1.00E+00	0	0	1.00E+00	0	0	1.73E-06	0	1	5.00E-01	0	0
liver	1.00E+00	0	0	1.00E+00	0	0	2.55E-05	10	1	1.00E+00	0	0
optical-recognition-handwritten-digits	8.77E-01	225	0	1.31E-01	159	0	1.11E-03	74	1	6.27E-01	195	0
ozone-level-detection-eight	1.59E-01	301	0	4.11E-01	192.5	0	9.02E-01	226.5	0	9.43E-01	229	0
pen-based-recognition-handwritten-digits	4.67E-01	24	0	8.49E-01	71.5	0	1.44E-01	150	0	5.78E-03	162	2
planning-relax	2.19E-01	6	0	1.88E-01	6	0	3.07E-02	127.5	1	8.76E-01	12.5	0
qsar-biodegradation	7.66E-01	247	0	6.58E-01	254	0	7.70E-04	69	1	9.02E-01	226.5	0
sonar	8.22E-01	243.5	0	8.45E-01	242	0	2.01E-02	119.5	1	8.77E-01	240	0
spambase	1.25E-01	158	0	4.99E-03	96	1	2.60E-06	4	1	1.25E-01	158	0
statlog-project-landsat-satellite	1.25E-01	158	0	9.84E-03	107	1	6.96E-01	213.5	0	1.47E-01	162	0
thyroid-disease-ann-thyroid	7.54E-01	203	0	2.59E-05	437	2	1.73E-06	0	1	1.32E-02	353	2
waveform	6.49E-01	223	0	9.59E-01	235	0	9.71E-05	43	1	7.58E-01	217.5	0
wdbc	8.73E-01	210	0	1.70E-03	80	1	9.37E-02	151	0	7.36E-06	15	1
weight-lifting-exercises	4.53E-01	269	0	8.59E-02	316	0	1.73E-06	0	1	1.31E-01	306	0
Kazanılan/ Eşitlik/ Kaybedilen	1/24/0			5/18/2			17/8/0			3/20/2		

Kutu grafikleri (boxplot), çeyrekler ve ortalamaları görüntüleyerek sayısal verilerin görsel dağılımını göstermek için kullanılan grafiklerdir. Kutu grafikleri, bir veri kümesi üzerinde algoritmalarının otuz bağımsız çalışma sonucunda oluşan beş özelliğini (minimum değer, ilk çeyrek, medyan, üçüncü çeyrek ve maksimum değer) gösterir. Minimum ve maksimum değer, aykırı değerler hariç en düşük ve en yüksek değeri gösterir. İlk çeyrek ve üçüncü çeyrek, %25 ve %75’lik dilimdeki puanların bir parçasıdır. Medyan verilerin orta noktasını gösterir. Ayrıca, bu değerlerin dışında kalan ve ortalamadan önemli sapmalar gösteren aykırı değerler işaretleyicilerle gösterilmektedir.

Algoritmaların veri setleri üzerinde performanslarının daha detaylı analizi için Şekil 4.13 'de bazı veri setlerine ait kutu grafikleri sunulmuştur. Grafikler otuz bağımsız çalışmadan elde edilen sonuçların dağılımını göstermektedir.

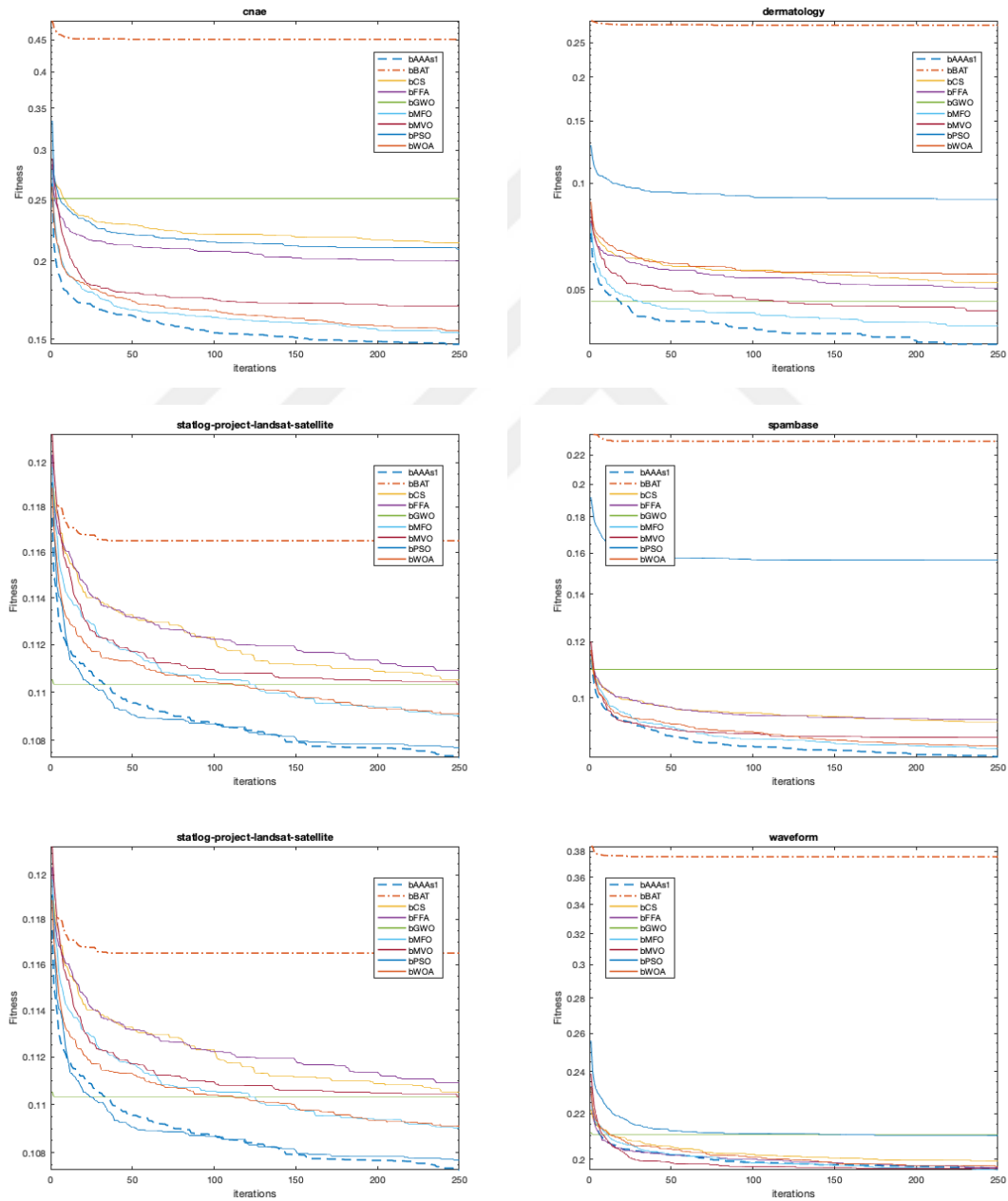


Şekil 4.13. Algoritmaların otuz bağımsız çalışma sonucundaki değerlerine göre oluşturulmuş kutu grafikleri

Şekil 4.13'teki kutu grafiklerine göre, BAT ve PSO tabanlı öznelilik seçimi algoritmaları birçok problemde çok çeşitli sonuçlar üretmiştir. Dolayısıyla standart sapma değeri yüksek ve parametrelere yüksek bağımlı olarak yorumlanabilir. Tezde geliştirilen ikili kaotik AAA yaklaşımının çözüm aralığı daha az sapma ile istikrarlı bir seyir izlemektedir. Daha dar bir aralıkta çözüm üreten ve aykırı değeri neredeyse bulunmayan bu algoritmanın ve ikili MFO algoritmalarının daha optimum sonuçlara ulaştığı

görülmektedir. Kutu grafikleri değerlendirildiğinde tezde geliştirilen yöntemin, kıyaslanan diğer yöntemlere göre daha kararlı bir davranış sergilediği görülmektedir.

Öznitelik seçiminde kullanılan metasezgisel yaklaşımlar, kendilerine has stratejilerinden dolayı ideal çözümü araştırırken farklı karakteristikler gösterirler. Algoritmaların bu karakteristiklerini gözlemlemek için yaygın olarak kullanılan teknik yakınsama grafikleridir. Deneysel çalışmada elde edilen farklı özelliklere sahip altı veri kümesi üzerindeki yakınsama eğrileri Şekil 4.14'te gösterilmektedir.



Şekil 4.14. Algoritmaların çeşitli veri setleri üzerindeki yakınsama eğrileri

Cnae-9, Dermatology, Statlog project landsat satellite, Spambase, Optical recognition handwritten digits ve Waveform veri setlerinde ikili kaotik AAA yaklaşımı, yinelemelerin başlangıcında hızlı şekilde yakınsama göstermiş ve yinelemelerin sonuna kadar gelişimini sürdürmüştür. Bu veri setlerinde, geliştirilen yaklaşım hem istatistiksel olarak hem de yakınsama eğrisi olarak kıyaslanan diğer yöntemlerden daha başarılıdır. Waveform veri setinde MVO, MFO ve FFA tabanlı yaklaşımlarda rekabetçi sonuçlar elde etmiştir. Geliştirilen yöntem ve FFA ve CS tabanlı algoritmalar, yinelemeler boyunca mevcut en iyi çözümü geliştirmeye devam etmiştir. Diğer algoritmalar yerel minimumlarda sıkışmış ve mevcut çözümü global optimumuma taşıyamamışlardır.



5. SONUÇLAR VE ÖNERİLER

5.1 Sonuçlar

Son yıllarda, gerçek dünya problemlerinin giderek artan karmaşıklığı ve zorluğu, daha güvenilir optimizasyon tekniklerine, özellikle metasezgisel optimizasyon algoritmalarına daha fazla ihtiyaç duyulmasına neden olmuştur. Metasezgisel teknikler sezgisel stratejiler içermesi ve yerel optimumlardan kaçınma kabiliyetlerinin yüksek olması nedeniyle geleneksel optimizasyon algoritmalarının yerini almaktadır. Optimizasyon, bir fonksiyonun veya bir problemin amaç fonksiyonunu minimize veya maksimize ederek optimal karar değişkenlerinin tespit edilme sürecidir. Genel olarak gerçek dünya optimizasyon problemleri lineer olmayan kısıtlara, karmaşık ve uzun hesaplama sürelerine, üstel olarak artan geniş arama uzaylarına sahiptirler. Bu nedenlerle kaba kuvvet yaklaşımı gibi geleneksel tekniklerle çözümleri pratikte mümkün değildir. Metasezgisel yaklaşımlar en optimum çözümü sunmayı garanti etmemekle birlikte makul süre zarfında kabul edilebilir çözümler sunmayı vaat etmektedir.

Metasezgisel optimizasyon algoritmalarının iki önemli arama stratejisi vardır. Bunlardan birisi keşif/çeşitlendirme diğeri sömürü/yoğunlaştırmadır. Keşif süreci, arama uzayını küresel olarak keşfetme yeteneğidir. Bu yetenek, yerel optimumdan kaçınma ve yerel optimuma sıkışınca kurtulabilme kabiliyetidir. Sömürü süreci ise çözümün uygunluğunu yerel olarak iyileştirmek için mevcut çözümün yakınındaki muhtemel çözümleri keşfetme yeteneğidir. Bir algoritmanın performansının mükemmel olması bu iki strateji arasındaki dengeye bağlıdır (Sadollah ve ark., 2018; Faramarzi ve ark., 2020b). Tüm popülasyon tabanlı algoritmalar bu iki stratejiyi geliştirmek için farklı operatörler ve mekanizmalar kullanırlar.

Metasezgisel algoritmaların popüler bir kategorisi canlı sürülerinin karakteristik özelliklerinden, yaşam davranışlarından, hareket süreçlerinden ve avlanma stratejilerinden ilham alan sürü zekâsı yaklaşımlarıdır. Bu grubun temel özelliği, optimizasyon süreci boyunca sürüdeki tüm canlıların arasında bilgi paylaşılması yapılmasıdır. Bu grubun geleneksel yöntemlerine kuşlar ve balıklardan esinlenen PSO, arılardan ilham alan ABC, kurt sürüsünü taklit eden GWO, balinaları taklit eden WOA gibi algoritmalar gösterilebilir.

Bu algoritmalarından birisi de AAA'dır. AAA mikro algilerin karakteristik özelliklerinden ve yaşam davranışlarından esinlenen doğadan ilhamlı metasezgisel bir optimizasyon algoritmasıdır. AAA'nın da birçok metasezgisel algoritmaya benzer şekilde yerel minimumlara sıkışma ve erken yakınsama problemi mevcuttur. Bu problemlerin üstesinden gelmek için sömürü ve keşif kabiliyetini çeşitli stratejilerle güçlendirmek gereklidir. Bu stratejilerden birisi kaos teorisinden ilham alan kaotik haritalardır.

Bu tez çalışmasında AAA'yı kaotik haritalarla donatarak keşif ve sömürü dengesinin güçlendirilmesi ve optimizasyon problemlerindeki performansının artırılması amaçlanmıştır. On iyi bilinen kaotik harita AAA'nın adaptasyon sürecinde kullanılmıştır. Geliştirilen kaotik AAA algoritmasının performansı on bir tek modlu, dokuz çok modlu ve on sabit boyutlu çok modlu toplam otuz benchmark fonksiyonu kullanılarak değerlendirilmiştir. İlk olarak farklı kaotik haritaları baz alan AAA varyasyonlarının performans analizi yapılmıştır. En başarılı varyasyonun gauss kaotik haritasıyla donatılan AAA varyasyonu CAAAv3 olarak gözlemlenmiştir. Daha sonra bu model PSO, GWO, MVO, SCA, FFA, CS, JAYA, MFO, SSA, WOA ve orijinal AAA olmak üzere on bir popüler sürü zekâsı tekniği ile bahsedilen otuz benchmark fonksiyonu üzerinde kıyaslanarak performansı doğrulanmıştır.

Benchmark fonksiyonlarının yanı sıra basınçlı tank tasarımı, kaynaklı kiriş tasarımı ve germe sıkıştırma yay tasarımı olmak üzere üç klasik mühendislik optimizasyon problemi ve Avrupa Uzay Ajansı'ndan alınan sekiz yörünge tasarım problemi üzerinde de test edilmiştir. Çalışma sonuçları hem benchmark setinde hem tasarım optimizasyon problemlerinde geliştirilen gauss tabanlı kaotik AAA'nın kıyaslanan on yöntemden ve orijinal AAA'dan daha performanslı olduğu göstermiştir. Wilcoxon işaretli sıralar testi ve Friedman istatistiksel testleri ile kaotik AAA'nın performansını doğrulamıştır.

Bu tez kapsamında çözüm geliştirilen bir diğer problem kümeleme problemidir. Kümeleme veri kümesi içerisindeki örneklerin sınıf etiketlerine bakılmaksızın yalnızca özelliklerine dayanarak gruplayan bir danışmansız öğrenme tekniğidir. Kümeleme yapılırken, aynı küme içerisindeki örneklerin birbirine benzerliğinin en fazla, farklı kümelerdeki örneklerle benzerliğinin en az olması prensibi esas alınır. Literatürdeki k-means, k-medoids gibi basit ve kolay uygulanabilen klasik kümeleme tekniklerinin karmaşık veri kümelerinde istenen performansı vermemesi, başlangıç parametrelerine bağımlılığının yüksek olması, erken yakınsaması ve yerel optimumlarda sıkışması

dolayısıyla araştırmacılar son yıllarda metasezgisel kümeleme yaklaşımlarına yönelmişlerdir. Bu sebeplerden motivasyonla bu tez kapsamında geliştirilen kaotik AAA yaklaşımı kümeleme analizinde optimum küme merkezlerinin tespit edilmesinde kullanılmak üzere organize edilmiştir. Yirmi beş farklı zorluk seviyesine sahip veri seti üzerinde önerilen yöntemin performansı PSO, GWO, MVO, GA, FFA, CS, BAT, MFO, SSA ve WOA ile kıyaslanarak değerlendirilmiştir. Çalışma sonuçları farklı zorluk seviyelerindeki kümeleme problemlerinin çoğunluğunda SSE, VM gibi birçok metrik açısından geliştirilen yöntemin kıyaslanan on popüler yöntemden daha başarılı olduğunu göstermiştir. Wilcoxon işaretli sıralar testi ve Friedman istatistiksel testleri yapılarak kaotik AAA yaklaşımının kümeleme analizindeki performansı doğrulanmıştır.

Tez kapsamında çözülen başka bir problem de öznitelik seçimidir. Öznitelik bir veri seti içerisinde bulunan her özellik, nitelik diğer bir tabirle verideki her bir kolon, sütundur. Öznitelik seçimi ise veri seti içerisindeki en yararlı öznitelikleri tespit etme işlemidir. Veri boyutunun artmasıyla gereksiz öznitelikler çoğalmakta, modellerin eğitim süresini artmakta ve performansını düşürmektedir. Dolayısıyla öznitelik seçimi yapay zekâ ve veri madenciliği için vazgeçilmez kritik bir süreç haline gelmiştir. Veri setindeki öznitelik sayısının doğrusal artışı, öznitelik alt kümesi sayısının üstel artışına neden olduğu için filtre tabanlı klasik öznitelik seçme algoritmalarının makul süre zarfında çözüm bulamamasına yol açmıştır. Dolayısıyla metasezgisel tabanlı öznitelik seçme yöntemlerine duyulan ihtiyaç artmaktadır. Tez kapsamında kaotik AAA tabanlı yeni bir öznitelik seçme yaklaşımı geliştirilmiştir. Kaotik AAA tabanlı yaklaşım sekiz farklı transfer fonksiyonu aracılığıyla ikili uzaya taşınarak çeşitli versiyonları geliştirilmiştir. Yirmi beş farklı büyüklükteki veri seti üzerinde önerilen yöntem ile öznitelik seçim işlemi yapılmıştır. İlk olarak sekiz farklı ikili versiyonun analizi yapılmış en performanslı modelin s1 transfer fonksiyonu ile geliştirilen versiyon olduğu görülmüştür. Daha sonra bu model veri setlerinde s1 fonksiyonunu ile transfer edilen BAT, CS, FFA, GWO, MFO, MVO, PSO ve WOA algoritmaları ile kıyaslanmıştır. Çalışma sonuçları kaotik bAAAs1 algoritmasının kıyaslanan sekiz popüler algoritmadan daha performanslı olduğu göstermiştir. Wilcoxon işaretli sıralar testi ve Friedman istatistiksel testleri ile kaotik bAAAs1 yaklaşımının güvenilirliği ve performansı kanıtlanmıştır.

5.2 Öneriler

Kaotik operatörler kullanılarak geliştirilen AAA yaklaşımı, global optimizasyon problemlerini, gerçek dünya problemlerini, öznitelik seçimi problemini ve kümeleme problemini başarı ile çözebildiği bu tez çalışması ile ispatlanmıştır. Literatürde AAA'nın ayırık versiyonu ile ilgili çalışmalar henüz bulunmamaktadır. Gelecek çalışmalarda kaotik AAA'nın ayırık optimizasyon problemlerine yönelik versiyonu geliştirilerek gezgin satıcı problemi çözülebilir.

Bu tez çalışmasında, AAA'ya kaos teorisini esas alan bir geliştirme yapılmıştır. Literatürde birçok algoritma geliştirme stratejisi bulunmaktadır. AAA'ya levy uçuşu, kuantum davranışı ve zeki arama gibi süreçler dahil edilerek optimizasyon problemlerindeki performansı artırılabilir. Yeni geliştirilecek modeller ile çok amaçlı optimizasyon problemleri de çözülebilir.

AAA gerek literatürdeki çalışmalarda gerek bu tez kapsamında mühendislik tasarım problemleri ve yörünge optimizasyon problemleri gibi geniş bir yelpazedeki gerçek dünya optimizasyon problemlerini başarıyla çözdüğü gösterilmiştir. Bu çalışmalarda gösterdiği performans yapılacak yeni çalışmalara ilham kaynağı olmaktadır. Literatürde henüz AAA ile çözülmemiş yapısal tasarım problemleri gibi gerçek dünya problemleri bulunmaktadır. AAA yeni problemlere uyarlanabilir ve çeşitli veri setlerinde performans çalışmaları yapılabilir.

Makine öğrenmesinin danışmanlı öğrenme, danışmansız öğrenme ve pekiştirmeli öğrenme olmak üzere üç temel öğrenme stratejisi vardır. Bu tez kapsamında danışmanlı öğrenme ve danışmansız öğrenme süreçlerinde AAA başarıyla uygulamıştır. Lakin önemli bir makine öğrenmesi süreci olan pekiştirmeli öğrenme alanında henüz çalışma yapılmamıştır. Derin Pekiştirmeli öğrenme sürecindeki ağırlık eğitimi ve parametrelerin optimizasyonu sürecinde AAA destekli yaklaşımlar kullanılabilir.

KAYNAKLAR

- Abdel-Basset, M., Ding, W. ve El-Shahat, D., 2021a, A hybrid Harris Hawks optimization algorithm with simulated annealing for feature selection, *Artificial Intelligence Review*, 54 (1), 593-637.
- Abdel-Basset, M., Sallam, K. M., Mohamed, R., Elgendi, I., Munasinghe, K. ve Elkomy, O. M., 2021b, An Improved Binary Grey-Wolf Optimizer With Simulated Annealing for Feature Selection, *IEEE Access*, 9, 139792-139822.
- Abdulwahab, H. A., Noraziah, A., Alsewari, A. A. ve Salih, S. Q., 2019, An enhanced version of black hole algorithm via levy flight for optimization and data clustering problems, *IEEE Access*, 7, 142085-142096.
- Abualigah, L., Diabat, A., Mirjalili, S., Abd Elaziz, M. ve Gandomi, A. H., 2021, The arithmetic optimization algorithm, *Computer Methods in Applied Mechanics and Engineering*, 376, 113609.
- Abualigah, L. M., Khader, A. T., Hanandeh, E. S. ve Gandomi, A. H., 2017, A novel hybridization strategy for krill herd algorithm applied to clustering techniques, *Applied Soft Computing*, 60, 423-435.
- Agbaje, M. B., Ezugwu, A. E. ve Els, R., 2019, Automatic data clustering using hybrid firefly particle swarm optimization algorithm, *IEEE Access*, 7, 184963-184984.
- Akhtar, T., Din, N. ve Uddin, J., 2019, Substitution box design based on chaotic maps and cuckoo search algorithm, *2019 International Conference on Advanced Communication Technologies and Networking (Commnet)*, 1-7.
- Alatas, B., Akin, E. ve Ozer, A. B., 2009, Chaos embedded particle swarm optimization algorithms, *Chaos, Solitons & Fractals*, 40 (4), 1715-1734.
- Aljarah, I. ve Ludwig, S. A., 2013, A new clustering approach based on glowworm swarm optimization, *2013 IEEE Congress on Evolutionary Computation*, 2642-2649.
- Aljarah, I., Mafarja, M., Heidari, A. A., Faris, H. ve Mirjalili, S., 2020a, Clustering analysis using a novel locality-informed grey wolf-inspired clustering approach, *Knowledge and Information Systems*, 62 (2), 507-539.
- Aljarah, I., Mafarja, M., Heidari, A. A., Faris, H. ve Mirjalili, S., 2020b, Multi-verse optimizer: theory, literature review, and application in data clustering, In: *Nature-Inspired Optimizers*, Eds: Springer, p. 123-141.
- Amaldi, E. ve Kann, V., 1998, On the approximability of minimizing nonzero variables or unsatisfied relations in linear systems, *Theoretical Computer Science*, 209 (1-2), 237-260.
- Arora, S. ve Anand, P., 2019a, Chaotic grasshopper optimization algorithm for global optimization, *Neural Computing & Applications*, 31 (8), 4385-4405.
- Arora, S. ve Anand, P., 2019b, Binary butterfly optimization approaches for feature selection, *Expert Systems with Applications*, 116, 147-160.
- Arora, S. ve Anand, P., 2019c, Chaotic grasshopper optimization algorithm for global optimization, *Neural Computing and Applications*, 31 (8), 4385-4405.
- Atalı, G., Pehlivan, I., GÜREVİN, B. ve ŞEKER, H. İ., 2021, Chaos in metaheuristic based artificial intelligence algorithms: a short review, *Turkish Journal of Electrical Engineering & Computer Sciences*, 29 (3), 1354-1367.
- Awadallah, M. A., Hammouri, A. I., Al-Betar, M. A., Braik, M. S. ve Abd Elaziz, M., 2022, Binary Horse herd optimization algorithm with crossover operators for feature selection, *Computers in biology and medicine*, 141, 105152.

- Azhagusundari, B. ve Thanamani, A. S., 2013, Feature selection based on information gain, *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, 2 (2), 18-21.
- Baalamurugan, K. ve Bhanu, S. V., 2019, An efficient clustering scheme for cloud computing problems using metaheuristic algorithms, *Cluster Computing*, 22 (5), 12917-12927.
- Beşkirli, M., Koç, İ., Hakkı, H. ve Kodaz, H., 2018, A new optimization algorithm for solving wind turbine placement problem: Binary artificial algae algorithm, *Renewable energy*, 121, 301-308.
- Billah, M. ve Waheed, S., 2020, Minimum redundancy maximum relevance (mRMR) based feature selection from endoscopic images for automatic gastrointestinal polyp detection, *Multimedia Tools and Applications*, 79 (33), 23633-23643.
- Blocho, M., 2020, Heuristics, metaheuristics, and hyperheuristics for rich vehicle routing problems, *Smart Delivery Systems: Solving Complex Vehicle Routing Problems*, 101-156.
- Cheng, M.-Y. ve Huang, K.-Y., 2010, Genetic algorithm-based chaos clustering approach for nonlinear optimization, *Journal of Marine Science and Technology*, 18 (3), 15.
- Coello, C. A. C., 2002, Theoretical and numerical constraint-handling techniques used with evolutionary algorithms: a survey of the state of the art, *Computer Methods in Applied Mechanics and Engineering*, 191 (11-12), 1245-1287.
- Crawford, B., Soto, R., Riquelme-Leiva, M., Peña, C., Torres-Rojas, C., Johnson, F. ve Paredes, F., 2015, Modified binary firefly algorithms with different transfer functions for solving set covering problems, In: *Software Engineering in Intelligent Systems*, Eds: Springer, p. 307-315.
- Çelik, Y. ve Kutucu, H., 2018, Solving the Tension/Compression Spring Design Problem by an Improved Firefly Algorithm, *IDDM*, 1 (2255), 1-7.
- Demir, F. B., Tuncer, T. ve Kocamaz, A. F., 2020, A chaotic optimization method based on logistic-sine map for numerical function optimization, *Neural Computing and Applications*, 32 (17), 14227-14239.
- Demsar, J., 2006, Statistical comparisons of classifiers over multiple data sets, *Journal of Machine Learning Research*, 7, 1-30.
- Derrac, J., García, S., Molina, D. ve Herrera, F., 2011, A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms, *Swarm and evolutionary computation*, 1 (1), 3-18.
- Dhote, C. A., Thakare, A. D. ve Chaudhari, S. M., 2013, Data Clustering Using Particle Swarm Optimization and Bee Algorithm, *2013 Fourth International Conference on Computing, Communications and Networking Technologies (Icccnt)*.
- dos Santos Coelho, L. ve Mariani, V. C., 2009, A novel chaotic particle swarm optimization approach using Hénon map and implicit filtering local search for economic load dispatch, *Chaos, Solitons & Fractals*, 39 (2), 510-518.
- dos Santos Coelho, L. ve Pessôa, M. W., 2011, A tuning strategy for multivariable PI and PID controllers using differential evolution combined with chaotic Zaslavskii map, *Expert Systems with Applications*, 38 (11), 13694-13701.
- dos Santos Coelho, L. ve Mariani, V. C., 2012, Firefly algorithm approach based on chaotic Tinkerbell map applied to multivariable PID controller tuning, *Computers & Mathematics with Applications*, 64 (8), 2371-2382.
- Dua, D. ve Karra Taniskidou, E., 2017, UCI Machine Learning Repository [<http://archive.ics.uci.edu/ml>]. Irvine, CA: University of California, *School of Information and Computer Science*.

- Elaskar, S. ve Del Río, E., 2017, New advances on chaotic intermittency and its applications, Springer, p.
- Emary, E., Zawbaa, H. M. ve Hassanien, A. E., 2016, Binary grey wolf optimization approaches for feature selection, *Neurocomputing*, 172, 371-381.
- Emine, B. ve Ülker, E., 2020, An efficient binary social spider algorithm for feature selection problem, *Expert Systems with Applications*, 146, 113185.
- Faramarzi, A., Heidarinejad, M., Mirjalili, S. ve Gandomi, A. H., 2020a, Marine Predators Algorithm: A nature-inspired metaheuristic, *Expert Systems with Applications*, 152, 113377.
- Faramarzi, A., Heidarinejad, M., Stephens, B. ve Mirjalili, S., 2020b, Equilibrium optimizer: A novel optimization algorithm, *Knowledge-based systems*, 191, 105190.
- Faris, H., Aljarah, I., Mirjalili, S., Castillo, P. A. ve Guervós, J. J. M., 2016, EvoloPy: An Open-source Nature-inspired Optimization Framework in Python, *IJCCI (ECTA)*, 171-177.
- Gandomi, A. H., Yang, X.-S., Talatahari, S. ve Alavi, A. H., 2013a, Firefly algorithm with chaos, *Communications in Nonlinear Science and Numerical Simulation*, 18 (1), 89-98.
- Gandomi, A. H., Yun, G. J., Yang, X.-S. ve Talatahari, S., 2013b, Chaos-enhanced accelerated particle swarm optimization, *Communications in Nonlinear Science and Numerical Simulation*, 18 (2), 327-340.
- Gündüz, M., Kiran, M. S. ve Özceylan, E., 2015, A hierarchic approach based on swarm intelligence to solve the traveling salesman problem, *Turkish Journal of Electrical Engineering & Computer Sciences*, 23 (1), 103-117.
- Haklı, H. ve Uğuz, H., 2014, A novel particle swarm optimization algorithm with Levy flight, *Applied Soft Computing*, 23, 333-345.
- Han, X. ve Chang, X., 2012, A chaotic digital secure communication based on a modified gravitational search algorithm filter, *Information Sciences*, 208, 14-27.
- Hans, R. ve Kaur, H., 2020, Binary Multi-Verse Optimization (BMVO) Approaches for Feature Selection, *International Journal of Interactive Multimedia & Artificial Intelligence*, 6 (1).
- Hassan, S., Kumar, K., Raj, C. D. ve Sridhar, K., 2014, Design and optimisation of pressure vessel using metaheuristic approach, *Applied Mechanics and Materials*, 401-406.
- Ho, Y.-C. ve Pepyne, D. L., 2002, Simple explanation of the no-free-lunch theorem and its implications, *Journal of optimization theory and applications*, 115 (3), 549-570.
- Houssein, E. H., Saad, M. R., Hashim, F. A., Shaban, H. ve Hassaballah, M., 2020, Lévy flight distribution: A new metaheuristic algorithm for solving engineering optimization problems, *Engineering Applications of Artificial Intelligence*, 94, 103731.
- Hu, G., Du, B., Wang, X. ve Wei, G., 2022, An enhanced black widow optimization algorithm for feature selection, *Knowledge-based systems*, 235, 107638.
- Hussain, K., Salleh, M. N. M., Cheng, S. ve Shi, Y. H., 2019, Metaheuristic research: a comprehensive survey, *Artificial Intelligence Review*, 52 (4), 2191-2233.
- Hussien, A. G., Hassanien, A. E., Houssein, E. H., Bhattacharyya, S. ve Amin, M., 2019, S-shaped binary whale optimization algorithm for feature selection, In: Recent trends in signal and image processing, Eds: Springer, p. 79-87.
- Izzo, D., 2007, 1st ACT global trajectory optimisation competition: Problem description and summary of the results, *Acta Astronautica*, 61 (9), 731-734.

- John, H., 1992, Holland. genetic algorithms, *Scientific american*, 267 (1), 44-50.
- Jordehi, A. R., 2015, Chaotic bat swarm optimisation (CBSO), *Applied Soft Computing*, 26, 523-530.
- Jothiprakash, V. ve Arunkumar, R., 2013, Optimization of hydropower reservoir using evolutionary algorithms coupled with chaos, *Water Resources Management*, 27 (7), 1963-1979.
- Kamil, A. T., Saleh, H. M. ve Abd-Alla, I. H., 2021, A Multi-Swarm Structure for Particle Swarm Optimization: Solving the Welded Beam Design Problem, *Journal of Physics: Conference Series*, 012012.
- Karaboga, D. ve Basturk, B., 2007, A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm, *Journal of global optimization*, 39 (3), 459-471.
- Karaboga, D. ve Ozturk, C., 2011, A novel clustering approach: Artificial Bee Colony (ABC) algorithm, *Applied Soft Computing*, 11 (1), 652-657.
- Karegowda, A. G., Manjunath, A. ve Jayaram, M., 2010, Comparative study of attribute selection using gain ratio and correlation based feature selection, *International Journal of Information Technology and Knowledge Management*, 2 (2), 271-277.
- Kaur, A., Pal, S. K. ve Singh, A. P., 2020, Hybridization of Chaos and Flower Pollination Algorithm over K-Means for data clustering, *Applied Soft Computing*, 97, 105523.
- Kaur, A. ve Kumar, Y., 2021, A new metaheuristic algorithm based on water wave optimization for data clustering, *Evolutionary Intelligence*, 1-25.
- Kaur, G. ve Arora, S., 2018, Chaotic whale optimization algorithm, *Journal of Computational Design and Engineering*, 5 (3), 275-284.
- Kechagiopoulos, P. N. ve Beligiannis, G. N., 2014, Solving the urban transit routing problem using a particle swarm optimization based algorithm, *Applied Soft Computing*, 21, 654-676.
- Kennedy, J. ve Eberhart, R., 1995, Particle swarm optimization, *Proceedings of ICNN'95-international conference on neural networks*, 1942-1948.
- Kennedy, J. ve Eberhart, R. C., 1997, A discrete binary version of the particle swarm algorithm, *1997 IEEE International conference on systems, man, and cybernetics. Computational cybernetics and simulation*, 4104-4108.
- Khurma, R. A., Aljarah, I., Sharieh, A. ve Mirjalili, S., 2020, Evolopy-fs: An open-source nature-inspired optimization framework in python for feature selection, In: *Evolutionary Machine Learning Techniques*, Eds: Springer, p. 131-173.
- Kıran, M. S., Özceylan, E., Gündüz, M. ve Paksoy, T., 2012, A novel hybrid approach based on particle swarm optimization and ant colony algorithm to forecast energy demand of Turkey, *Energy conversion and management*, 53 (1), 75-83.
- Kiran, M. S., 2015, TSA: Tree-seed algorithm for continuous optimization, *Expert Systems with Applications*, 42 (19), 6686-6698.
- Kiran, M. S., Hakli, H., Gunduz, M. ve Uguz, H., 2015, Artificial bee colony algorithm with variable search strategy for continuous optimization, *Information Sciences*, 300, 140-157.
- Kiran, M. S. ve Hakli, H., 2021, A tree-seed algorithm based on intelligent search mechanisms for continuous optimization, *Applied Soft Computing*, 98, 106938.
- Kohli, M. ve Arora, S., 2018, Chaotic grey wolf optimization algorithm for constrained optimization problems, *Journal of Computational Design and Engineering*, 5 (4), 458-472.
- Kumar, M. ve Dhillon, J., 2018, Hybrid artificial algae algorithm for economic load dispatch, *Applied Soft Computing*, 71, 89-109.

- Kumar, R., Sharma, D. ve Sadu, A., 2011, A hybrid multi-agent based particle swarm optimization algorithm for economic power dispatch, *International journal of electrical power & energy systems*, 33 (1), 115-123.
- Kumar, V., Chhabra, J. K. ve Kumar, D., 2017, Grey wolf algorithm-based clustering technique, *Journal of Intelligent Systems*, 26 (1), 153-168.
- Kuwil, F. H., Shaar, F., Topcu, A. E. ve Murtagh, F., 2019, A new data clustering algorithm based on critical distance methodology, *Expert Systems with Applications*, 129, 296-310.
- Lange, K., 2013, Optimization, Springer Science & Business Media, p.
- Li, Y., Li, T. ve Liu, H., 2017, Recent advances in feature selection and its applications, *Knowledge and Information Systems*, 53 (3), 551-577.
- Liu, B., Wang, L., Jin, Y.-H., Tang, F. ve Huang, D.-X., 2005, Improved particle swarm optimization combined with chaos, *Chaos, Solitons & Fractals*, 25 (5), 1261-1271.
- Liu, H. ve Setiono, R., 1995, Chi2: Feature selection and discretization of numeric attributes, *Proceedings of 7th IEEE International Conference on Tools with Artificial Intelligence*, 388-391.
- Mafarja, M., Eleyan, D., Abdullah, S. ve Mirjalili, S., 2017, S-shaped vs. V-shaped transfer functions for ant lion optimization algorithm in feature selection problem, *Proceedings of the international conference on future networks and distributed systems*, 1-7.
- Mafarja, M., Aljarah, I., Heidari, A. A., Faris, H., Fournier-Viger, P., Li, X. ve Mirjalili, S., 2018, Binary dragonfly optimization for feature selection using time-varying transfer functions, *Knowledge-based systems*, 161, 185-204.
- Mafarja, M. ve Mirjalili, S., 2018, Whale optimization approaches for wrapper feature selection, *Applied Soft Computing*, 62, 441-453.
- Mafarja, M., Aljarah, I., Faris, H., Hammouri, A. I., Ala'M, A.-Z. ve Mirjalili, S., 2019, Binary grasshopper optimisation algorithm approaches for feature selection problems, *Expert Systems with Applications*, 117, 267-286.
- Mafarja, M. M. ve Mirjalili, S., 2017, Hybrid whale optimization algorithm with simulated annealing for feature selection, *Neurocomputing*, 260, 302-312.
- Mageshkumar, C., Karthik, S. ve Arunachalam, V., 2019, Hybrid metaheuristic algorithm for improving the efficiency of data clustering, *Cluster Computing*, 22 (1), 435-442.
- May, R. M., 2004, Simple mathematical models with very complicated dynamics, *The Theory of Chaotic Attractors*, 85-93.
- Mirjalili, S. ve Hashim, S. Z. M., 2012, BMOA: binary magnetic optimization algorithm, *International Journal of Machine Learning and Computing*, 2 (3), 204.
- Mirjalili, S. ve Lewis, A., 2013, S-shaped versus V-shaped transfer functions for binary particle swarm optimization, *Swarm and evolutionary computation*, 9, 1-14.
- Mirjalili, S., Mirjalili, S. M. ve Lewis, A., 2014a, Grey wolf optimizer, *Advances in Engineering Software*, 69, 46-61.
- Mirjalili, S., Mirjalili, S. M. ve Yang, X.-S., 2014b, Binary bat algorithm, *Neural Computing and Applications*, 25 (3), 663-681.
- Mirjalili, S., 2015, Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm, *Knowledge-Based Systems*, 89, 228-249.
- Mirjalili, S., 2016, SCA: a sine cosine algorithm for solving optimization problems, *Knowledge-based systems*, 96, 120-133.
- Mirjalili, S. ve Lewis, A., 2016, The whale optimization algorithm, *Advances in Engineering Software*, 95, 51-67.

- Mirjalili, S., Mirjalili, S. M. ve Hatamlou, A., 2016, Multi-verse optimizer: a nature-inspired algorithm for global optimization, *Neural Computing and Applications*, 27 (2), 495-513.
- Mirjalili, S. ve Gandomi, A. H., 2017, Chaotic gravitational constants for the gravitational search algorithm, *Applied Soft Computing*, 53, 407-419.
- Mirjalili, S., Gandomi, A. H., Mirjalili, S. Z., Saremi, S., Faris, H. ve Mirjalili, S. M., 2017, Salp Swarm Algorithm: A bio-inspired optimizer for engineering design problems, *Advances in engineering software*, 114, 163-191.
- Mitić, M., Vuković, N., Petrović, M. ve Miljković, Z., 2015, Chaotic fruit fly optimization algorithm, *Knowledge-based systems*, 89, 446-458.
- Mohammadzadeh, H. ve Gharehchopogh, F. S., 2021, A novel hybrid whale optimization algorithm with flower pollination algorithm for feature selection: Case study Email spam detection, *Computational Intelligence*, 37 (1), 176-209.
- Nadimi-Shahraki, M. H., Banaie-Dezfouli, M., Zamani, H., Taghian, S. ve Mirjalili, S., 2021, B-MFO: A Binary Moth-Flame Optimization for Feature Selection from Medical Datasets, *Computers*, 10 (11), 136.
- Nakamura, R. Y., Pereira, L. A., Costa, K. A., Rodrigues, D., Papa, J. P. ve Yang, X.-S., 2012, BBA: a binary bat algorithm for feature selection, *2012 25th SIBGRAPI conference on graphics, Patterns and Images*, 291-297.
- Nasiri, J. ve Khiyabani, F. M., 2018, A whale optimization algorithm (WOA) approach for clustering, *Cogent Mathematics & Statistics*, 5 (1), 1483565.
- Neggaz, N., Houssein, E. H. ve Hussain, K., 2020, An efficient henry gas solubility optimization for feature selection, *Expert Systems with Applications*, 152, 113364.
- Nouri-Moghaddam, B., Ghazanfari, M. ve Fathian, M., 2021, A novel multi-objective forest optimization algorithm for wrapper feature selection, *Expert Systems with Applications*, 175, 114737.
- Ouannas, A., Khennaoui, A.-A., Odibat, Z., Pham, V.-T. ve Grassi, G., 2019, On the dynamics, control and synchronization of fractional-order Ikeda map, *Chaos, Solitons & Fractals*, 123, 108-115.
- Ozsari, S., Uguz, H. ve Hakli, H., 2021, Implementation of meta-heuristic optimization algorithms for interview problem in land consolidation: A case study in Konya/Turkey, *Land Use Policy*, 108, 105511.
- Pan, J.-S., Tian, A.-Q., Chu, S.-C. ve Li, J.-B., 2021, Improved binary pigeon-inspired optimization and its application for feature selection, *Applied Intelligence*, 51 (12), 8661-8679.
- Panagant, N., Pholdee, N., Bureerat, S., Kaen, K., Yıldız, A. R. ve Sait, S. M., 2020, Seagull optimization algorithm for solving real-world design optimization problems, *Materials Testing*, 62 (6), 640-644.
- Peng, H., Long, F. ve Ding, C., 2005, Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy, *IEEE transactions on pattern analysis and machine intelligence*, 27 (8), 1226-1238.
- Pernkopf, F., 2005, Bayesian network classifiers versus selective k-NN classifier, *Pattern recognition*, 38 (1), 1-10.
- Qaddoura, R., Faris, H., Aljarah, I. ve Castillo, P. A., 2020, EvoCluster: An Open-Source Nature-Inspired Optimization Clustering Framework in Python, *International Conference on the Applications of Evolutionary Computation (Part of EvoStar)*, 20-36.
- Qaddoura, R., Faris, H., Aljarah, I. ve Castillo, P. A., 2021, EvoCluster: An Open-Source Nature-Inspired Optimization Clustering Framework, *SN Computer Science*, 2 (3), 1-12.

- Rahnema, N. ve Gharehchopogh, F. S., 2020, An improved artificial bee colony algorithm based on whale optimization algorithm for data clustering, *Multimedia Tools and Applications*, 79 (43), 32169-32194.
- Rao, R., 2016, Jaya: A simple and new optimization algorithm for solving constrained and unconstrained optimization problems, *International Journal of Industrial Engineering Computations*, 7 (1), 19-34.
- Rashedi, E., Nezamabadi-Pour, H. ve Saryazdi, S., 2009, GSA: a gravitational search algorithm, *Information Sciences*, 179 (13), 2232-2248.
- Rashedi, E., Nezamabadi-Pour, H. ve Saryazdi, S., 2010, BGSA: binary gravitational search algorithm, *Natural Computing*, 9 (3), 727-745.
- Rather, S. A. ve Bala, P. S., 2020, Swarm-based chaotic gravitational search algorithm for solving mechanical engineering design problems, *World Journal of Engineering*.
- Rodrigues, D., Pereira, L. A., Nakamura, R. Y., Costa, K. A., Yang, X.-S., Souza, A. N. ve Papa, J. P., 2014, A wrapper approach for feature selection based on bat algorithm and optimum-path forest, *Expert Systems with Applications*, 41 (5), 2250-2258.
- Rosenberg, A. ve Hirschberg, J., 2007, V-measure: A conditional entropy-based external cluster evaluation measure, *Proceedings of the 2007 joint conference on empirical methods in natural language processing and computational natural language learning (EMNLP-CoNLL)*, 410-420.
- Sadollah, A., Sayyaadi, H., Lee, H. M. ve Kim, J. H., 2018, Mine blast harmony search: a new hybrid optimization method for improving exploration and exploitation capabilities, *Applied Soft Computing*, 68, 548-564.
- Sala, R. ve Müller, R., 2020, Benchmarking for metaheuristic black-box optimization: perspectives and open challenges, *2020 IEEE Congress on Evolutionary Computation (CEC)*, 1-8.
- Saremi, S., Mirjalili, S. ve Lewis, A., 2014, Biogeography-based optimisation with chaos, *Neural Computing and Applications*, 25 (5), 1077-1097.
- Sayed, G. I., Khoriba, G. ve Haggag, M. H., 2018, A novel chaotic salp swarm algorithm for global optimization and feature selection, *Applied Intelligence*, 48 (10), 3462-3481.
- Sayed, G. I., Hassanien, A. E. ve Azar, A. T., 2019, Feature selection via a novel chaotic crow search algorithm, *Neural Computing and Applications*, 31 (1), 171-188.
- Schlueter, M., Neshat, M., Wahib, M., Munetomo, M. ve Wagner, M., 2021, GTOPX space mission benchmarks, *Software*, 14.
- Senthilnath, J., Omkar, S. ve Mani, V., 2011, Clustering using firefly algorithm: performance study, *Swarm and evolutionary computation*, 1 (3), 164-171.
- Senthilnath, J., Kulkarni, S., Suresh, S., Yang, X. ve Benediktsson, J., 2021, FPA clust: evaluation of the flower pollination algorithm for data clustering, *Evolutionary Intelligence*, 14 (3), 1189-1199.
- Sharma, M. ve Chhabra, J. K., 2021, An efficient hybrid PSO polygamous crossover based clustering algorithm, *Evolutionary Intelligence*, 14 (3), 1213-1231.
- Singh, H., Kumar, Y. ve Kumar, S., 2019, A new meta-heuristic algorithm based on chemical reactions for partitional clustering problems, *Evolutionary Intelligence*, 12 (2), 241-252.
- Singh, T., 2020, A chaotic sequence-guided Harris hawks optimizer for data clustering, *Neural Computing and Applications*, 32 (23), 17789-17803.
- Snaselova, P. ve Zboril, F., 2015, Genetic algorithm using theory of chaos, *Procedia computer science*, 51, 316-325.

- Song, F., Guo, Z. ve Mei, D., 2010, Feature selection using principal component analysis, *2010 international conference on system science, engineering design and manufacturing informatization*, 27-30.
- Stork, J., Eiben, A. E. ve Bartz-Beielstein, T., 2020, A new taxonomy of global optimization algorithms, *Natural Computing*.
- Storn, R. ve Price, K., 1997, Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces, *Journal of global optimization*, 11 (4), 341-359.
- Sun, J., Xu, W. ve Feng, B., 2004, A global search strategy of quantum-behaved particle swarm optimization, *IEEE Conference on Cybernetics and Intelligent Systems*, 2004., 111-116.
- Talatahari, S., Azar, B. F., Sheikholeslami, R. ve Gandomi, A., 2012, Imperialist competitive algorithm combined with chaos for global optimization, *Communications in Nonlinear Science and Numerical Simulation*, 17 (3), 1312-1319.
- Tang, R., Fong, S. ve Dey, N., 2018, Metaheuristics and chaos theory, *Chaos theory*, 182-196.
- Tarkhaneh, O. ve Moser, I., 2019, An improved differential evolution algorithm using Archimedean spiral and neighborhood search based mutation approach for cluster analysis, *Future Generation Computer Systems*, 101, 921-939.
- Tsai, C.-W., Chang, W.-Y., Wang, Y.-C. ve Chen, H., 2019, A high-performance parallel coral reef optimization for data clustering, *Soft Computing*, 23 (19), 9327-9340.
- Turabieh, H., Mafarja, M. ve Li, X., 2019, Iterated feature selection algorithms with layered recurrent neural network for software fault prediction, *Expert Systems with Applications*, 122, 27-42.
- Turkoglu, B. ve Kaya, E., 2020, Training multi-layer perceptron with artificial algae algorithm, *Engineering Science and Technology, an International Journal*.
- Turkoglu, B., Uymaz, S. A. ve Kaya, E., 2022a, Clustering analysis through artificial algae algorithm, *International Journal of Machine Learning and Cybernetics*, 13 (4), 1179-1196.
- Turkoglu, B., Uymaz, S. A. ve Kaya, E., 2022b, Binary Artificial Algae Algorithm for feature selection, *Applied Soft Computing*, 120, 108630.
- Uymaz, S. A., Tezel, G. ve Yel, E., 2015a, Artificial algae algorithm (AAA) for nonlinear global optimization, *Applied Soft Computing*, 31, 153-171.
- Uymaz, S. A., Tezel, G. ve Yel, E., 2015b, Artificial algae algorithm with multi-light source for numerical optimization and applications, *Biosystems*, 138, 25-38.
- Wang, G.-G., Guo, L., Gandomi, A. H., Hao, G.-S. ve Wang, H., 2014, Chaotic krill herd algorithm, *Information Sciences*, 274, 17-34.
- Wolpert, D. H. ve Macready, W. G., 1997, No free lunch theorems for optimization, *IEEE transactions on evolutionary computation*, 1 (1), 67-82.
- Wu, G., Qiu, D., Yu, Y., Pedrycz, W., Ma, M. ve Li, H., 2014, Superior solution guided particle swarm optimization combined with local search techniques, *Expert Systems with Applications*, 41 (16), 7536-7548.
- Yang, X.-S., 2009, Firefly algorithms for multimodal optimization, *International symposium on stochastic algorithms*, 169-178.
- Yang, X.-S. ve Deb, S., 2009, Cuckoo search via Lévy flights, *2009 World congress on nature & biologically inspired computing (NaBIC)*, 210-214.
- Yang, X.-S., 2010a, Firefly algorithm, stochastic test functions and design optimisation, *International journal of bio-inspired computation*, 2 (2), 78-84.

- Yang, X.-S., 2010b, A new metaheuristic bat-inspired algorithm, In: Nature inspired cooperative strategies for optimization (NICSO 2010), Eds: Springer, p. 65-74.
- Yildiz, A. R., 2009, An effective hybrid immune-hill climbing optimization approach for solving design and manufacturing optimization problems in industry, *Journal of Materials Processing Technology*, 209 (6), 2773-2780.
- Zhang, C., Ouyang, D. ve Ning, J., 2010, An artificial bee colony approach for clustering, *Expert Systems with Applications*, 37 (7), 4761-4767.
- Zhao, J. ve Gao, Z.-M., 2021, Liebovitch map enabled Sine Cosine algorithm, *2021 International Conference on Machine Learning and Intelligent Systems Engineering (MLISE)*, 309-312.
- Zhao, W., Wang, L. ve Mirjalili, S., 2022, Artificial hummingbird algorithm: A new bio-inspired optimizer with its engineering applications, *Computer Methods in Applied Mechanics and Engineering*, 388, 114194.
- Zhenyu, G., Bo, C., Min, Y. ve Binggang, C., 2006, Self-adaptive chaos differential evolution, *International Conference on Natural Computation*, 972-975.
- Zhou, Y., Wu, H., Luo, Q. ve Abdel-Baset, M., 2019, Automatic data clustering using nature-inspired symbiotic organism search algorithm, *Knowledge-based systems*, 163, 546-557.
- Zhu, L., Wang, J., Wang, H., Guo, S.-S., Guo, M. ve Xie, W., 2020, Data clustering method based on improved bat algorithm with six convergence factors and local search operators, *IEEE Access*, 8, 80536-80560.
- Zimmerman, D. W. ve Zumbo, B. D., 1993, Relative power of the Wilcoxon test, the Friedman test, and repeated-measures ANOVA on ranks, *The Journal of Experimental Education*, 62 (1), 75-86.