

T.C.
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ



YAPAY ZEKA OPTİMİZASYON ALGORİTMASI TEMELLİ
ALDATMA VE SÖYLENTİ TESPİT YÖNTEMİ

Harun BİNGÖL

Doktora Tezi

YAZILIM MÜHENDİSLİĞİ ANABİLİM DALI

Yazılım Mühendisliği Bilim Dalı

HAZİRAN 2021

T.C.
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

Yazılım Mühendisliği Anabilim Dalı

Doktora Tezi

YAPAY ZEKA OPTİMİZASYON ALGORİTMASI TEMELLİ ALDATMA VE SÖYLENTİ TESPİT YÖNTEMİ

Tez Yazarı
Harun BİNGÖL

Danışman
Prof. Dr. Bilal ALATAŞ

HAZİRAN 2021
ELAZIĞ

T.C.
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

Yazılım Mühendisliği Anabilim Dalı

Doktora Tezi

Başlığı:	Yapay Zeka Optimizasyon Algoritması Temelli Aldatma ve Söylenti Tespit Yöntemi
Yazarı:	Harun BİNGÖL
İlk Teslim Tarihi:	14.04.2021
Savunma Tarihi:	02.06.2021

TEZ ONAYI

Fırat Üniversitesi Fen Bilimleri Enstitüsü tez yazım kurallarına göre hazırlanan bu tez aşağıda imzaları bulunan jüri üyeleri tarafından değerlendirilmiş ve akademik dinleyicilere açık yapılan savunma sonucunda OYBİRLİĞİ ile kabul edilmiştir.

Danışman:	Prof. Dr. Bilal ALATAŞ Fırat Üniversitesi, Mühendislik Fakültesi	<i>İmza</i> Onayladım
Başkan:	Dr. Öğr. Üyesi Taner UÇKAN Van Yüzüncü Yıl Üniversitesi, Mühendislik Fakültesi	Onayladım
Üye:	Doç. Dr. Fatih ÖZKAYNAK Fırat Üniversitesi, Teknoloji Fakültesi	Onayladım
Üye:	Doç. Dr. M. Tuncay GENÇOĞLU Fırat Üniversitesi, T.B. MYO.	Onayladım
Üye:	Dr. Öğr. Üyesi Muammer TÜRKOĞLU Samsun Üniversitesi, Mühendislik Fakültesi	Onayladım

Bu tez, Enstitü Yönetim Kurulunun/...../20..... tarihli toplantısında tescillenmiştir.

İmza
Doç. Dr. Kürşat Esat ALYAMAÇ
Enstitü Müdürü

BEYAN

Fırat Üniversitesi Fen Bilimleri Enstitüsü tez yazım kurallarına uygun olarak hazırladığım “Yapay Zeka Optimizasyon Algoritması Temelli Aldatma ve Söylenti Tespit Yöntemi ” Başlıklı Doktora Tezimin içindeki bütün bilgilerin doğru olduğunu, bilgilerin üretilmesi ve sunulmasında bilimsel etik kurallarına uygun davrandığımı, kullandığım bütün kaynakları atıf yaparak belirttiğimi, maddi ve manevi desteği olan tüm kurum/kuruluş ve kişileri belirttiğimi, burada sunduğum veri ve bilgileri unvan almak amacıyla daha önce hiçbir şekilde kullanmadığımı beyan ederim.

02.06.2021

Harun BİNGÖL



ÖNSÖZ

Sosyal medya, bilişim teknolojilerindeki ilerleme ile birlikte neredeyse dünyada yaşayan tüm toplulukların hayatına girmiş bulunmaktadır. Özellikle cep telefonlarının internete bağlı küçük birer bilgisayar gibi işlev görmesi, insanların tüm işlemlerini internetin olduğu her yerden çok rahat bir şekilde halledebilme kolaylığını da beraberinde getirmiştir. Bu kolaylık aynı zamanda sosyal ağlara da daha kolay ve konforlu erişmeyi sağlamıştır.

Sosyal ağlara erişimin bu denli hızlı ve rahat olması, havanın durumu, trafiğin tıkanıklığı ya da herhangi bir marketteki indirim yada kampanya gibi bilgilere erişim bakımından düşünüldüğünde insanların hayatlarını kolaylaştırdığı yadsınamaz bir gerçektir. Ancak bu olumlu taraflarının yanında olumsuz etkileri de görmezden gelinemez. Bunların başında ise sosyal medya aracılığı ile güvenilirliği tartışmalı olan bilgiye erişen kişilerin aldatılması olayıdır. Aldatma; genel olarak belirli bir amaç için kişilerin belli bir konu üzerine kitlelerin düşüncelerini yönlendirme ve toplumsal algı oluşturmayı amaçlar. Dolayısıyla bu durum hem bireysel hem de toplumsal düzeyde zarar vermektedir. Bir diğer sosyal medya problemi ise söylenti'dir. Söylenti ise olmayan bir olayın sanki olmuş gibi ağızdan ağıza yayılarak aktarılması olayıdır. Bu olayların tespiti sosyal ağları kullanan kişilerin sayısındaki devasa artış hızı ile bilgisayar bilimleri alanında gün geçtikçe daha da önemli bir hale gelmektedir.

Bu tezde sosyal ağlardaki aldatma ve söylenti problemi, literatürde ilgili ve ilişkili çalışmalar olmasına rağmen ilk kez bir optimizasyon problemi olarak değerlendirilmiştir. Bu tez çalışmasında ayrıca üç yeni kaotik metasezgisel optimizasyon algoritması önerilmiş ve bahsi geçen sosyal ağ problemlerine uygulanmıştır.

Doktora çalışmalarım süresince, değerli görüş ve katkılarıyla beni yönlendiren, takıldığım tüm konularda içtenlikle yardım eden, bilimsel çalışma disiplini bana kazandıran ve çok kıymetli tecrübelerinden yararlandığım tez danışmanım Sayın Prof. Dr. Bilal ALATAŞ'a teşekkürü borç bilirim.

Ayrıca her zaman varlıklarıyla bana destek olan sevgili eşim Tuba'ya, kızım Asel'e, oğlum Göktuğ'a, anneme, babama ve sevgili dostlarıma sonsuz minnet ve teşekkürlerimi sunarım.

Harun BİNGÖL

ELAZIĞ, 2021

İÇİNDEKİLER

	Sayfa
ÖNSÖZ.....	iv
İÇİNDEKİLER	v
ÖZET	vii
ABSTRACT	viii
ŞEKİLLER LİSTESİ	ix
TABLolar LİSTESİ	xi
SİMGELER	xiv
KISALTMALAR	xv
1. GİRİŞ	1
1.1. Tezin Amacı	4
1.2. Tezin Organizasyonu	5
2. VERİ MADENCİLİĞİ	7
2.1. Veri Madenciliği Aşamaları	7
2.1.1. Seçim.....	7
2.1.2. Ön İşlem ve Birleştirme	8
2.1.3. Dönüşüm	8
2.1.4. Modelleme.....	8
2.1.5. Yorumlama.....	8
2.2. Veri Madenciliği Modelleri	8
2.2.1. Tanımlayıcı Modeller	9
2.2.2. Tahmin Edici Modeller	9
3. OPTİMİZASYON	10
3.1. Genel Optimizasyon	10
3.2. Metasezgisel Yöntemler	11
4. ALDATMA VE SÖYLENTİ TESPİTİ	16
4.1. Aldatma Tespiti	16
4.1.1. Aldatıcı İçerik Nedir?.....	16
4.1.2. Aldatmanın Toplum Üzerindeki Etkileri	17
4.1.3. Aldatma Tespiti ile İlgili Literatür Taraması.....	17
4.2. Söylenti Tespiti.....	22
4.2.1. Söylenti Nedir?.....	22
4.2.2. Söylentinin Toplum Üzerindeki Etkileri	22
4.2.3. Söylenti ile İlgili Literatür Taraması	23
5. ALGORİTMALAR VE VERİ KÜMELERİ	26
5.1. Optik İlham Optimizasyon	26
5.2. Gri Kurt Optimizasyon	32
5.3. Denetimli Makine Öğrenme Algoritmaları.....	39
5.4. Veri Ön İşleme Aşamaları	41
5.5. Aldatma ve Söylenti Tespiti Veri Kümeleri	42
5.5.1. Aldatma Tespiti Veri Setleri.....	42
5.5.2. Söylenti Tespiti Veri Setleri	44

5.6. Performans Değerlendirme Kriterleri	45
5.7. Friedman Testi.....	46
5.8. Kutu Grafiği.....	47
6. ÖNERİLEN YÖNTEMLER	48
6.1. Kaos Tabanlı Optik İlham Optimizasyon Algoritmaları (KTOIO).....	48
6.1.1. Kaotik Haritalar	48
6.1.2. Optik İlham Optimizasyon Algoritmasına Kaotik Harita Entegrasyonu.....	50
6.2. Metasezgisel Optimizasyon Algoritmalarının Aldatma ve Söylenti Tespit Problemleri için Modellenmesi.....	66
7. DENEYSEL BULGULAR	68
7.1. Deceptive Opinion Spam Corpus v1.4 Veri Setinden Elde Edilen Sonuçlar	68
7.2. Real-Life Trial Veri Setinden Elde Edilen Sonuçlar	73
7.3. S1 Veri Setinden Elde Edilen Sonuçlar	77
7.4. Germanwings Crash Veri Setinden Elde Edilen Sonuçlar.....	82
8. SONUÇLAR.....	87
KAYNAKLAR	89
ÖZGEÇMİŞ	

ÖZET

Yapay Zeka Optimizasyon Algoritması Temelli Aldatma ve Söylenti Tespit Yöntemi

Harun BİNGÖL

Doktora Tezi

FIRAT ÜNİVERSİTESİ
Fen Bilimleri Enstitüsü

Yazılım Mühendisliği Anabilim Dalı

Haziran 2021, Sayfa: xv + 96

Günümüzde gelişen teknoloji ile birlikte bilgiye olan erişim yöntemleri de değişmiştir. Geleneksel bilgiye erişim araçlarından olan TV, radyo, gazete, dergi gibi araçlar yerini sosyal ağlara bırakmıştır. Bilgiye erişim yönteminin değişmesindeki temel faktörler arasında sosyal ağların ucuz olması, geleneksel yöntemlere göre daha hızlı erişilmesi ve internetin olduğu her yerden rahat bir şekilde ulaşılması başlıca nedenleridir. Sosyal ağların veya sosyal medyanın bu kadar çok avantajına rağmen geleneksel yöntemlere göre dezavantajlı olduğu konu ise bilginin doğruluğudur. Sosyal medyada bilgi o kadar çabuk yayılır ki ya yayılırken insanlar bilgiyi istemsizce değiştirmekte (söylenti olayı) ya da bilgi kasıtlı bir şekilde belirli bir veya bir takım amaçlar için yanlış servis edilmektedir (aldatma olayı). Bilginin bu şekilde dağıtılmasının ticari, politik, ekonomik gibi birçok nedeni olabilir. Hem söylenti hem de aldatma amaçlı olan yanlış içerikli bilgi hem bilgiyi edinen kişiyi hem de toplumu olumsuz yönde etkiler. Sosyal ağlardaki aldatma ve söylenti içerikli bilginin tespit edilmesi son derece önemlidir.

Bu tezde sosyal ağ problemlerinden olan Söylenti Tespit Problemi ve Aldatma Tespit Problemi bir optimizasyon problemi olarak ele alınmıştır. Optimizasyon çalışması sırasında üçü ilk defa bu tezde önerilen yeni kaotik optimizasyon algoritmaları olmak üzere beş adet metasezgisel optimizasyon algoritması kullanılmış ve literatürde kullanılan yedi adet denetimli yapay zekâ algoritması ile aldatma tespiti ve söylenti tespiti sonuçları karşılaştırılmıştır. Problemlerin çözümü noktasında önermiş olduğumuz yöntem çok yeni olmasına rağmen, umut verici sonuçlar elde edilmiştir.

Anahtar Kelimeler: Aldatma tespiti, Söylenti tespiti, Optimizasyon, Kaos, Sosyal ağ analizi

ABSTRACT

Deception and Rumor Detection Method based on Artificial Intelligence Optimization Algorithm

Harun BİNGÖL

Ph.D. Thesis

FIRAT UNIVERSITY
Graduate School of Natural and Applied Sciences
Department of Software Engineering

June 2021, Pages: xv + 96

Today, with the developing technology, access methods to information have also changed. Social networks fill the places of traditional information access tools such as TV, radio, newspaper, magazine. Among the main factors in changing the method of access to information, social networks are cheap, faster access than traditional methods, and easy access from anywhere where the internet is available. It is the accuracy of the information that the social networks or social media are disadvantageous compared to traditional methods. In social media, information spreads so quickly that when it spreads, people unintentionally change the information (rumor event) or information is deliberately served incorrectly for a particular or purpose (deception event). There may be many reasons for the distribution of information in this way, such as commercial, political, and economic. Information with fake content, both rumor and deception purposes, negatively affects both the person who acquires the information and the society. It is extremely important to detect deceptive and rumored information in social networks.

In this thesis, Rumor Detection problem and Deception Detection problem which is one of the social network problems are considered as an optimization problem. During the optimization, five heuristic optimization algorithms have been used, three of them have been proposed for the first time in this thesis, which are the new chaotic optimization algorithms and seven supervised artificial intelligence algorithms used in the literature and deception detection and rumor detection results have been compared. Although the method we proposed at the point of solution of the problems is very new, promising results have been obtained.

Keywords: Deception detection, Rumor detection, Optimization, Chaos, Social network analysis

ŞEKİLLER LİSTESİ

	Sayfa
Şekil 1.1. Sosyal medya platformları.....	2
Şekil 1.2. Tüm dünya geneli telefon, internet ve SM kullanıcı sayıları.....	3
Şekil 1.3. Tüm dünyada yıllara göre SM kullanıcı sayısı artış oranları.....	3
Şekil 2.1. Veri madenciliği aşamaları.....	7
Şekil 3.1. Matematiksel modeller	10
Şekil 3.2. Metasezgisel yöntemler.....	14
Şekil 3.3. Popülasyon tabanlı metasezgisel algoritmaların genel akış diyagramı.....	15
Şekil 3.4. Popülasyon tabanlı metasezgisel algoritmaların sözde kodu.....	15
Şekil 5.1.İç bükey ayna ışık ve yansıma.....	27
Şekil 5.2.İçbükey aynanın yansıma karakteristiği	28
Şekil 5.3.Dışbükey aynanın yansıma karakteristiği.....	29
Şekil 5.4.İçbükey ayna küresel sapma.....	29
Şekil 5.5.İçbükey ayna yanal sapma.....	29
Şekil 5.6.Işığın dışbükey ayna üzerine düşmesi durumunda yansıma kanunlarına göre gösterimi	31
Şekil 5.7.OIO algoritmasının akış diyagramı	32
Şekil 5.8.Gri kurt hiyerarşik yapısı (baskınlık tepeden zemine doğru azalmaktadır)	33
Şekil 5.9. (a) Avı izleme, yaklaşma ve kovalama. (b-d) Avı takip etme, etrafını sarma ve taciz etme. (e) Uygun anı bulma ve ava saldırma.....	34
Şekil 5.10.(a) İki boyutlu konum vektörleri (b) Güncellenmiş üç boyutlu konum vektörleri	35
Şekil 5.11.GKO'da pozisyon güncelleme	36
Şekil 5.12.Ava saldırı ile av arayışının karşılaştırılması	37
Şekil 5.13.GKO algoritmasının sözde kodu	39
Şekil 5.14. Ham veri üzerinde gerçekleştirilen ön işlem aşamaları.....	41
Şekil 5.15.Deceptive Opinion Spam Corpus v1.4 veri setinden bir kesit	43
Şekil 5.16.Real-Life Trial Data veri setinden bir kesit	43
Şekil 5.17. S1 veri setinden bir kesit	44
Şekil 5.18.Germanwings Crash veri setinden bir kesit.....	45
Şekil 6.1.Germe/Sıkıştırma yay tasarımı.....	57
Şekil 6.2.Germe/Sıkıştırma yay tasarımında Gauss harita kullanımı	59
Şekil 6.3.Germe/Sıkıştırma yay tasarımında Çember harita kullanımı	59
Şekil 6.4.Germe/Sıkıştırma yay tasarımında Lojistik harita kullanımı.....	60

Şekil 6.5. Germe/Sıkıştırma yay tasarımı Sinüzoidal harita kullanımı	60
Şekil 6.6. Germe/Sıkıştırma yay tasarımı Çadır harita kullanımı	61
Şekil 6.7. Kaynaklı giriş tasarımı	62
Şekil 6.8. Kaynaklı giriş tasarımı Gauss harita kullanımı.....	64
Şekil 6.9. Kaynaklı giriş tasarımı Çember harita kullanımı.....	64
Şekil 6.10. Kaynaklı giriş tasarımı Lojistik harita kullanımı	65
Şekil 6.11. Kaynaklı giriş tasarımı Sinüzoidal harita kullanımı	65
Şekil 6.12. Kaynaklı giriş tasarımı Çadır harita kullanımı.....	66
Şekil 6.13. Aldatma ve söylenti tespiti için önerilen sistemin yapısı	67
Şekil 7.1. Deceptive opinion spam corpus v1.4 veri setinde Gauss harita kullanılarak OIO, GKO, KTOIO1, KTOIO2 ve KTOIO3 algoritmalarının kutu grafiği.....	71
Şekil 7.2. Denetimli makine öğrenme algoritmaları ve gauss harita kullanılarak elde edilen metasezgisel optimizasyon algoritmalarının karşılaştırılması.....	73
Şekil 7.3. Real-Life Trial veri setinde Sinüzoidal harita kullanılarak OIO, GKO, KTOIO1, KTOIO2 ve KTOIO3 algoritmalarının kutu grafiği.....	75
Şekil 7.4. Denetimli makine öğrenme algoritmaları ve sinüzoidal harita kullanılarak elde edilen metasezgisel optimizasyon algoritmalarının karşılaştırılması.....	77
Şekil 7.5. S1 veri setinde Gauss harita kullanılarak OIO, GKO, KTOIO1, KTOIO2 ve KTOIO3 algoritmalarının kutu grafiği.....	80
Şekil 7.6. Denetimli makine öğrenme algoritmaları ve gauss harita kullanılarak elde edilen metasezgisel optimizasyon algoritmalarının karşılaştırılması.....	82
Şekil 7.7. Germanwings veri setinde Lojistik harita kullanılarak OIO, GKO, KTOIO1, KTOIO2 ve KTOIO3 algoritmalarının kutu grafiği.....	84
Şekil 7.8. Denetimli makine öğrenme algoritmaları ve lojistik harita kullanılarak elde edilen metasezgisel optimizasyon algoritmalarının karşılaştırılması.....	86

TABLÖLAR LİSTESİ

	Sayfa
Tablo 1.1. SM platformları bazında aktif kullanıcı sayıları	4
Tablo 5.1. Görüntünün tipi, boyu, konumu ve yönü	27
Tablo 5.2. Formüller ve değişkenlerin anlamları	30
Tablo 5.3. Aldatma ve söylenti tespiti için karmaşıklık matrisi	45
Tablo 5.4. Performans değerlendirme ölçütleri	46
Tablo 6.1. Kalite test fonksiyonlarına ait matematiksel ifade, tanımlı aralık ve global minimum	51
Tablo 6.2. Standart OIO'nun kısıtsız kalite test fonksiyonları üzerindeki sonuçları	51
Tablo 6.3. Rastrigin fonksiyonunda Çember, Gauss, Lojistik haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar	52
Tablo 6.4. Rastrigin fonksiyonunda Çadır, Sinüzoidal haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar	52
Tablo 6.5. Sphere fonksiyonunda Çember, Gauss, Lojistik haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar	52
Tablo 6.6. Sphere fonksiyonunda Çadır, Sinüzoidal haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar	53
Tablo 6.7. Rosenbrock fonksiyonunda Çember, Gauss, Lojistik haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar	53
Tablo 6.8. Rosenbrock fonksiyonunda Çadır, Sinüzoidal haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar	53
Tablo 6.9. Griewank fonksiyonunda Çember, Gauss, Lojistik haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar	54
Tablo 6.10. Griewank fonksiyonunda Çadır, Sinüzoidal haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar	54
Tablo 6.11. Ackley fonksiyonunda Çember, Gauss, Lojistik haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar	54
Tablo 6.12. Ackley fonksiyonunda Çadır, Sinüzoidal haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar	54
Tablo 6.13. Bohachevsky-1 fonksiyonunda Çember, Gauss, Lojistik haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar	55
Tablo 6.14. Bohachevsky-1 fonksiyonunda Çadır, Sinüzoidal haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar	55
Tablo 6.15. KTOIO algoritmalarında en iyi değerlerin toplamı ($Min + O$)	55
Tablo 6.16. KTOIO algoritmalarında en iyi değerlerin toplamı (Min)	56
Tablo 6.17. Germe/Sıkıştırma yay tasarımında Çember, Gauss, Lojistik haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar	58

Tablo 6.18. Germe/Sıkıştırma yay tasarımında Sinüzoidal ve Çadır haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar	58
Tablo 6.19. Kaynaklı giriş tasarımında Çember, Gauss, Lojistik haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar	63
Tablo 6.20. Kaynaklı giriş tasarımında Sinüzoidal ve Çadır haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar	63
Tablo 7.1. Metasezgisel optimizasyon algoritmalarına ait parametreler ve değerleri	68
Tablo 7.2. Deceptive opinion spam corpus v1.4 veri setinde Gauss harita kullanılarak elde edilen sonuçlar	69
Tablo 7.3. Deceptive opinion spam corpus v1.4 veri setinde Çember harita kullanılarak elde edilen sonuçlar	69
Tablo 7.4. Deceptive opinion spam corpus v1.4 veri setinde Lojistik harita kullanılarak elde edilen sonuçlar	69
Tablo 7.5. Deceptive opinion spam corpus v1.4 veri setinde Sinüzoidal harita kullanılarak elde edilen sonuçlar.....	70
Tablo 7.6. Deceptive opinion spam corpus v1.4 veri setinde Çadır harita kullanılarak elde edilen sonuçlar 70	
Tablo 7.7. Deceptive opinion spam corpus v1.4 veri setinde Gauss harita kullanılarak uygunluk değerlerine göre Friedman testi ile elde edilen metasezgisel algoritmaların ortalama sıralamaları.....	71
Tablo 7.8. Deceptive opinion spam corpus v1.4 veri seti için Friedman testi sonuçları	71
Tablo 7.9. Deceptive opinion spam corpus v1.4 veri setinde Gauss haritası kullanılarak elde edilen KTOIO algoritmaları, makine öğrenme algoritmaları, GKO ve OIO'nun karşılaştırılması.....	72
Tablo 7.10. Real-Life Trial veri setinde Gauss harita kullanılarak elde edilen sonuçlar	74
Tablo 7.11. Real-Life Trial veri setinde Çember harita kullanılarak elde edilen sonuçlar.....	74
Tablo 7.12. Real-Life Trial veri setinde Lojistik harita kullanılarak elde edilen sonuçlar	74
Tablo 7.13. Real-Life Trial veri setinde Sinüzoidal harita kullanılarak elde edilen sonuçlar	74
Tablo 7.14. Real-Life Trial veri setinde Çadır harita kullanılarak elde edilen sonuçlar	75
Tablo 7.15. Real-Life Trial veri setinde Sinüzoidal harita kullanılarak uygunluk değerlerine göre Friedman testi ile elde edilen metasezgisel algoritmaların ortalama sıralamaları	76
Tablo 7.16. Real-Life Trial veri seti için Friedman testi sonuçları	76
Tablo 7.17. Real-Life Trial veri setinde Sinüzoidal harita kullanılarak elde edilen KTOIO algoritmaları, makine öğrenme algoritmaları, GKO ve OIO'nun karşılaştırılması	77
Tablo 7.18. S1 veri setinde Gauss harita kullanılarak elde edilen sonuçlar	78
Tablo 7.19. S1 veri setinde Çember harita kullanılarak elde edilen sonuçlar	78
Tablo 7.20. S1 veri setinde Lojistik harita kullanılarak elde edilen sonuçlar	78
Tablo 7.21. S1 veri setinde Sinüzoidal harita kullanılarak elde edilen sonuçlar	79
Tablo 7.22. S1 veri setinde Çadır harita kullanılarak elde edilen sonuçlar	79

Tablo 7.23. S1 veri setinde Gauss harita kullanılarak uygunluk değerlerine göre Friedman testi ile elde edilen metasezgisel algoritmaların ortalama sıralamaları	80
Tablo 7.24. S1 veri seti için Friedman testi sonuçları	80
Tablo 7.25. S1 veri setinde Gauss harita kullanılarak elde edilen KTOIO algoritmaları, makine öğrenme algoritmaları, GKO ve OIO'nun karşılaştırılması.....	81
Tablo 7.26. Germanwings Crash veri setinde Gauss harita kullanılarak elde edilen sonuçlar	83
Tablo 7.27. Germanwings Crash veri setinde Çember harita kullanılarak elde edilen sonuçlar	83
Tablo 7.28. Germanwings Crash veri setinde Lojistik harita kullanılarak elde edilen sonuçlar	83
Tablo 7.29. Germanwings Crash veri setinde Sinüzoidal harita kullanılarak elde edilen sonuçlar.....	83
Tablo 7.30. Germanwings Crash veri setinde Çadır harita kullanılarak elde edilen sonuçlar.....	84
Tablo 7.31. Germanwings Crash veri setinde Lojistik harita kullanılarak uygunluk değerlerine göre Friedman testi ile elde edilen metasezgisel algoritmaların ortalama sıralamaları	85
Tablo 7.32. Germanwings Crash veri seti için Friedman testi sonuçları.....	85
Tablo 7.33. Germanwings Crash veri setinde Lojistik harita kullanılarak elde edilen KTOIO algoritmaları, makine öğrenme algoritmaları, GKO ve OIO'nun karşılaştırılması	86

SİMGELER

Simgeler

alt_i	: x_i 'nin alt sınırı
f	: Odak uzaklığı
F	: Fonksiyon
$f(x)$	: Amaç fonksiyonu
$g_k(x)$	: Eşitsizlik kısıtı
$h_j(x)$	: Eşitlik kısıtı
HI	: Görüntü yüksekliği
HO	: Nesne yüksekliği
m	: Büyütme oranı
Mak	: En büyük
Min	: En küçük
O	: Ortalama
p	: Görüntü konumu
q	: Nesne konumu
r	: Kavis yarıçapı
SS	: Standart sapma
t	: İterasyon sayısı
$üst_i$	: x_i 'nin üst sınırı

KISALTMALAR

Kısaltmalar

ALP	: Artificial Light Point
ARPANET	: Advanced Research Project Agency Network
BDI	: Belief, Desire, Intention
DT	: Decision Tree
GKO	: Gri Kurt Optimizasyon
IP	: İnternet Protokolü
KNN	: K-Nearest Neighbor
KTOIO	: Kaos Tabanlı Optik İlham Optimizasyon
LR	: Logistic Regression
NB	: Naive Bayesian
NFE	: Number of Function Evaluation
NN	: Neural Networks
NSF	: National Science Foundation
OIO	: Optik İlham Optimizasyon
RF	: Random Forest
SA	: Sosyal Ağ
SM	: Sosyal Medya
SVM	: Support Vector Machine
TF	: Term Frequency
VSM	: Vector Space Model

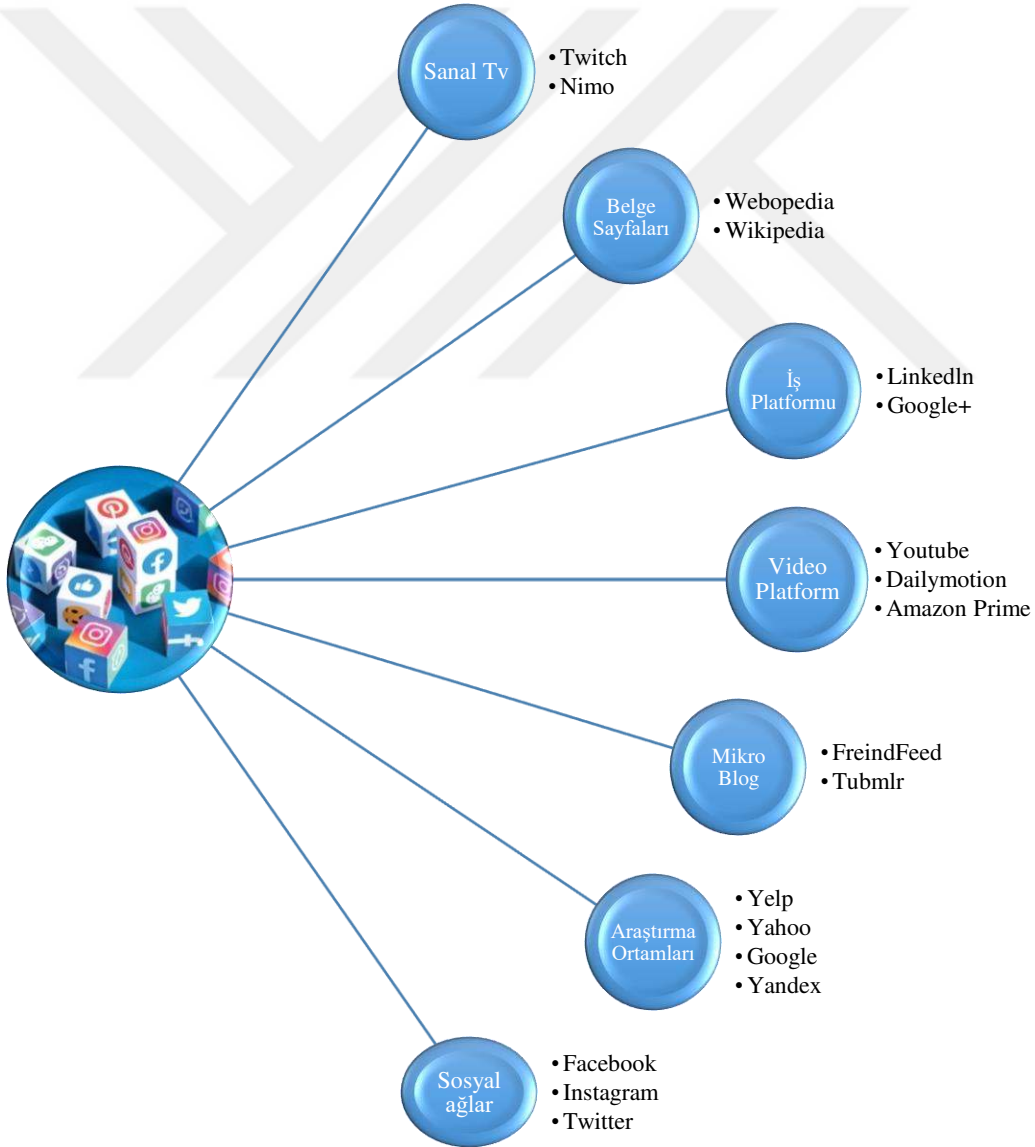
1. GİRİŞ

Toplama, çıkarma gibi temel aritmetik işlemlerin bir makine yardımı ile gerçekleştirme fikri çok eski zamanlara dayanmaktadır. Bilgisayar kelimesi Latince hesaplamak anlamına gelen “computere” kelimesinin dilimize çevrilmesiyle ortaya çıkmıştır [1]. Hesaplama işlemlerini otomatikleştiren bir bilgisayar makinesini geliştirmek amacıyla sürekli çalışan bilim insanları 1950’li yıllara kadar 1. kuşak olarak tabir edilen ve elektron tüplerini kullanan bilgisayarları geliştirdiler.

1948 yılında transistörler içerisinde kullanılacak yarı iletken maddelerin (silisyum, germanyum) bulunması bilgisayarların gelişimi açısından yeni bir çağın açılmasına olanak tanıdı. 1955 yılında elektron tüpleri yerine transistörler elektronik devrelerde kullanılmaya başlandı. Böylece daha az enerji tüketen, daha küçük boyutlu, daha hızlı 2. kuşak olarak tabir edilen bilgisayarlar geliştirilmeye başlandı. 1958 yılında transistörlerin bilgisayarlarda kullanılmasıyla internet teknolojisi de hızla gelişmeye başladı [1]. 1960 yılların başında Amerika Birleşik Devletleri tarafından fonlanarak geliştirilen, İnternet Protokolü (IP) kullanan ilk ağ ARPANET (Advanced Research Project Agency Network) geliştirildi [2]. 1970’lerin başında paket anahtarlama IP kullanan birçok ağ geliştirildi. 1980’li yılların başında NSF (National Science Foundation) tarafından birçok üniversite bünyesinde bilgisayar laboratuvarları kurulması sağlandı ve NSFNET projesi kapsamında bu üniversitelerdeki laboratuvarlar birbirlerine bağlandı. Ticari internet servis sağlayıcıları NSFNET projesinin akabinde 1980’li yılların sonuna doğru ortaya çıkmaya başladı [3]. Zaman içerisinde ticari internet servis sağlayıcıların artması internetin gelişmesine olanak sağladı. Böylelikle TCP/IP protokolünü kullanan tüm cihazlar internet üzerinden birbirleriyle haberleşebildiler. Günümüzde insanlar istedikleri yerden ister kablolu ister kablosuz bağlantı kullanarak cep telefonları, tablet ya da dizüstü bilgisayarları ile internete erişebilmekte ve bilgi alış-verişi yapabilmektedir. İnternetin dünya genelinde 1990’lı yıllardan itibaren hızla gelişmesi sonucu toplumlarda ekonomik, sosyal ve politik değişiklikler meydana gelmiştir. Bireylerin dini inançlar, özgürlükler ve tabular ile ilgili düşünceleri de zamanla değişime uğramıştır. Bu değişimin temel nedeni internet ile beraber hayatımıza giren Sosyal Ağ (SA) [4] ya da Sosyal Medya (SM) [5] kavramlarıdır. SA en basit tabiri ile bireylerin kendilerine ait sayfalarında ya da bir başkasının kişisel sayfası üzerinde fotoğraf, video ya da metin paylaşımı yapabildiği, sanal arkadaşlıklar ile insanların iletişim kurmalarına izin verilen dijital platformların genel adıdır. Birbirini tanıyan ya da bir birini hiç tanımayan insanların SA ortamları vasıtasıyla etkili, kolay ve ucuz bir şekilde haberleşebilmesi daha önceki nesiller ile karşılaştırıldığında inanılmaz bir yaşam değişikliğidir. SM ise SA ile karşılaştırıldığında daha geniş bir kavramdır. Şöyle ki SA için örnek vermek gerekirse Facebook, Instagram, Twitter, Myspace vb. verilebilir. SM ise SA’yı bünyesinde barındırmanın yanı sıra iş platformları olan (Google+, LinkedIn vb.),

mikro blog sayfaları (FreindFeed, Tubmlr vb.), belge sayfaları (Webopedia, Wikipedia vb.), araştırma ortamları (Yelp, Yahoo, Google, Yandex vb.), video platformları (Youtube, Dailymotion, Amazon Prime Video), sanal televizyon ortamları (Twitch, Nimo vb.) platformların tamamını kapsayan genel tanımın adıdır. SM bireylerin öğrenme ve öğretme süreçlerinin yeniden şekillenmesi yönünde inanılmaz imkânlar sunmaktadır [6]. Sosyal içerik üreten ağ yapıları Şekil 1.1’de gösterilmektedir. Tüm dünyada Ocak 2021 itibariyle teknoloji ve SM kullanıcı sayıları Şekil 1.2’de, yıllara göre SM kullanıcı sayısındaki artış oranları Şekil 1.3’te ve SM platformlar bazında aktif kullanıcı sayıları Tablo1.1’de gösterilmektedir [7].

Bu kadar hızlı gelişmekte olan bir yapıyı bilim dünyası da yakından takip etmektedir. Bilim dünyası bir yandan hızla gelişen SM platformlarının toplum üzerindeki kazanımlarını ve faydalarını araştırırken, diğer yandan toplum üzerindeki olumsuz etkilerini de incelemektedirler.



Şekil 1.1. Sosyal medya platformları



Toplam Dünya
Nüfusu
7,83 Milyar



Telefon Kullanıcısı
5,22 Milyar

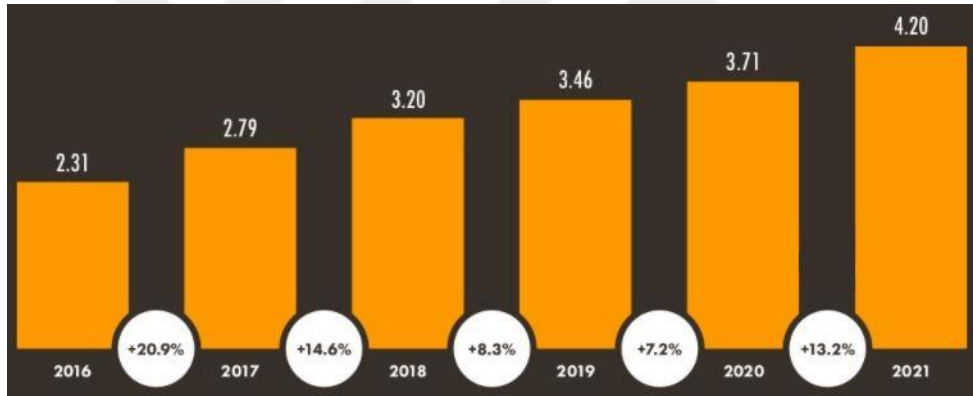


İnternet Kullanıcısı
4,66 Milyar



Sosyal Medya Kullanıcısı
4,20 Milyar

Şekil 1.2. Tüm dünya geneli telefon, internet ve SM kullanıcı sayıları



Şekil 1.3. Tüm dünyada yıllara göre SM kullanıcı sayısı artış oranları

SM'nin çok hızlı, kolay ve ucuz hizmet sağlama imkânları sadece bireylerin değil şirketlerin de dikkatini çekmiş durumdadır. Böylelikle SM ürünlerin tanıtılıp satışının yapıldığı bir platform durumuna dönüşmüştür. Milyonlarca kullanıcıya ulaşma ve onlara ürünlerini tanıtmak fikri sadece şirketlerin ilgilendiği bir konu olarak kalmayıp, kitlelere ulaşarak fikirlerini değiştirmek isteyen siyasiler, politikacılar, sivil toplum kuruluşları, dernekler, eğitim kurumları vb. birçok farklı kurum ve kuruluşun ortak noktası olmuştur. SM'nin toplum üzerinde oluşturmuş olduğu bazı sorunlar küçük çocukların teknoloji bağımlısı olması, ahlak yozlaşması, tüketim toplumu olması, toplumun hassas değerlerine dokunarak onların yanıltılması sıralanabilir. Ancak araştırmacıların dikkatle incelediği konuların başında SM platformlarında doğrulanmamış bilgilerin, doğru olan bilgilerden bile daha hızlı bir şekilde yayılması ve toplumun bu şekilde yanıltılmasıdır. Aldatma problemi SM platformlarında kasıtlı bir şekilde SM'yi bir propaganda aracı olarak, kitlelerin bir konu, olay,

durum hakkında fikrini deęiřtirmek iin kullanılması durumudur. Syleri problemi ise SM platformlarında bazı konu ve olayların aktarılırken abartılması ve aslının deęiřtirilmesi ile bilginin kirlenmesi ve toplumun yanıltılması durumudur.

Tablo 1.1. SM platformları bazında aktif kullanıcı sayıları

Sosyal medya platformu	Aktif kullanıcı sayısı (milyon)
Facebook	2740
Youtube	2291
Whatsapp	2000
Fb Messenger	1300
Instagram	1221
Weixin / wechat	1213
Tiktok	689
Qq	617
Douyin	600
Sina weibo	511
Telegram	500
Snapchat	498
Kuaishou	481
Pinterest	442
Reddit	430
Twitter	353
Quora	300

1.1. Tezin Amacı

Bu tez alışmasında elektronik bilimindeki gelişmeler sonucunda ve bilgisayar teknolojilerindeki ilerlemeler sayesinde, internet teknolojilerinin gelişimi ile hayatımıza giren ve ok önemli bir yer tutan SA ve SM kavramları, bunların birbirleri ile olan ilişkileri ve toplum üzerindeki olumsuz sonuçlarından olan aldatma ve sylenti problemleri detaylı bir şekilde incelenmiştir. Bu problemleri özmek amacıyla literatürde yer alan bilimsel yayınlar incelenerek derinlemesine bir analiz gerçekleştirilmiştir. Bu tezde ok yeni ve popüler olan, SM sorunlarını inceleyen bilim adamları açısından ok kıymetli bulunan, aldatma ve sylenti tespit problemi ilk defa bir optimizasyon problemi olarak ele alınmıştır. Bu tezde Optik İlham Optimizasyon (OIO) ve Gri Kurt Optimizasyon (GKO) algoritmaları Aldatma ve Sylenti Tespit problemine uyarlanmış ve drt farklı veri seti üzerinde algoritmaların başarımları sınanmıştır. Yine bu tez kapsamında OIO algoritmasının en iyi özümü yakınsama başarımını artırmak iin, OIO algoritması ierisinde yer

alan rassal deęiřkenler yerine, beř adet farklı kaotik harita kullanılarak üç adet yeni Kaos Tabanlı Optik İlham Optimizasyon (KTOIO) algoritmaları önerilmiştir. Bu tezde önerilen KTOIO algoritmalarının başarımları, dört farklı veri seti üzerinde sınanmıştır.

Bu tezin literatüre olan katkıları řunlardır:

- Sosyal medya ve sosyal aę kavramları hakkında temel bilgiler sunmaktır.
- Sosyal medya problemlerinden olan aldatma ve söylenti tespit problemleri ile ilgili literatür çalışması sunmak, bu problemlerin çözümü noktasında farklı bakış açıları geliştirebilmeyi sağlamaktır.
- Veri madencilięi ve optimizasyon algoritmaları hakkında temel bilgiler vermektir.
- Metasezgisel optimizasyon algoritmaları ilk defa bu tez kapsamında aldatma ve söylenti tespit probleminin çözümü için kullanılmıştır.
- Bu tez kapsamında üç adet yeni kaotik haritalı metasezgisel optimizasyon algoritması önerilmiştir.
- Bu tezde geliştirilen yöntemler çok yeni olmalarına karşın umut verici sonuçlar elde edilmiştir.

1.2. Tezin Organizasyonu

Bu tez çalışması sekiz temel bölümden oluşmaktadır. Birinci bölümde bilgisayar teknolojilerindeki ilerlemeler sayesinde, internet teknolojilerinin gelişimi ve akabinde hayatımıza giren ve çok önemli bir yer tutan SA ve SM kavramları hakkında temel bilgiler verilmiş ve tezin amacı açıklanmıştır.

İkinci bölümde, veri madencilięi hakkında temel bilgiler verilmiştir.

Üçüncü bölümde, optimizasyon ve metasezgisel optimizasyon algoritmaları hakkında temel bilgiler verilmiştir.

Dördüncü bölümde, aldatma ve söylenti tespit problemi, bu problemlerden kaynaklanan sorunlar, literatürde yer alan bakış açıları ve çözüm yöntemleri incelenerek aldatma ve söylenti tespit problemleri hakkında detaylı bir analiz yapılmıştır.

Beşinci bölümde, aldatma ve söylenti tespiti için kullanılan optimizasyon algoritmaları, denetimli makine öğrenme yöntemleri, veri ön işlem adımları, sınıflandırma amacıyla kullanılan aldatma ve söylenti tespiti veri kümeleri ve elde edilen deneysel sonuçları karşılařtırmak için kullanılan istatistiksel yöntemler hakkında detaylı bilgiler verilmiştir.

Altıncı bölümde, aldatma ve söylenti tespiti için önerilen kaos tabanlı optimizasyon algoritmaları hakkında detaylı bilgiler verilmiş ve bu algoritmaların başarımları kalite test fonksiyonları ve gerçek dünya mühendislik problemleri üzerinde sınanmıştır. Ayrıca yine bu bölümde kaos tabanlı optimizasyon algoritmasını geliřtirmek için kullanılan kaotik haritalar

hakkında temel bilgiler verilmiştir. Yine bu bölümde metasezgisel optimizasyon algoritmalarının aldatma ve söylenti tespit problemlerini çözmek amacıyla nasıl modellendiği hakkında bilgiler verilmiştir.

Yedinci bölümde, üç tanesi bu tez kapsamında önerilen kaos tabanlı metasezgisel optimizasyon algoritması olmak üzere beş farklı optimizasyon algoritması ile yedi farklı denetimli makine öğrenme yönteminin aldatma ve söylenti tespiti için kullanılan veri setlerinde elde ettiği test sonuçları karşılaştırmalı bir şekilde verilmiştir. Elde edilen sonuçların istatistiksel olarak değerlendirilmesi için friedman ve kutu grafiği yöntemleri kullanılmış ve test sonuçları verilmiştir. Ayrıca kullanılan metasezgisel optimizasyon algoritmaları ve diğer yöntemlerin, farklı veri setleri üzerinde test edilmesiyle elde edilen sonuçların dört farklı performans ölçütüne göre değerlendirildiği grafiksel gösterimler bu bölümde verilmektedir.

Sekizinci bölümde, tezde yapılan çalışmalar ve elde edilen sonuçlar değerlendirilmiştir. İleriye dönük çalışmalar için öneriler sunulmuştur.

2. VERİ MADENCİLİĞİ

Veri madenciliği büyük ve karmaşık veri kümelerindeki yapıların ve modellerin keşfidir [8]. Bilgisayar teknolojilerindeki hızlı gelişmeler sonucu, veri üreten cihaz sayısı ve insanların teknolojiye artan merakları veri miktarını da arttırmaktadır. Veri madenciliğinin temel amacı, büyük miktardaki veriden anlamlı olanlarının seçilerek, geleceğe yönelik tahmin yapmada kullanışlı bağlantı ve kuralların ortaya çıkartılmasıdır [9].

2.1. Veri Madenciliği Aşamaları

Verinin ham halinden bilgiye dönüşüncye kadar geçen aşamalarının tümünü ifade eder. Genel olarak 5 aşamada ham veri bilgiye dönüşmektedir. Bu aşamalar Şekil 2.1’de gösterilmektedir.



Şekil 2.1. Veri madenciliği aşamaları

- ✓ Verinin seçimi aşaması
- ✓ Ön işlem ve birleştirme aşaması
- ✓ Veri dönüştürme aşaması
- ✓ Modelleme aşaması
- ✓ Yorumlama aşaması

2.1.1. Seçim

Veri madenciliği işlemlerinde en fazla zaman alan aşamadır. Problem ile ilişkilendirme yapılmayan veya iyi analiz edilmeyen veri ile madencilik işleminin başarıya ulaşması mümkün değildir.

2.1.2. Ön İşlem ve Birleştirme

Ön işleme aşaması düzgün bir model oluşturmaya doğrudan etki etmektedir. Bu aşamada veri diğer aşamalarda kullanılabilecek elverişli bir hale getirilir. Veri girişi veya verinin toplanması esnasında oluşan sistem dışı hatalara gürültü denir. Bu aşamada veriden gürültülü olanlar çıkartılır [10].

Veri madenciliğinde kullanılacak olan veri genellikle tek bir kaynaktan elde edilmez. Farklı veri tabanlarından elde edilen veri, veri ambarı olarak adlandırılan büyük depolama alanları altında birleştirilmektedir. Burada dikkat edilmesi gereken konu verinin indirgenmesidir. Herhangi bir değişken başka tablolardan türetilmiş ise fazlalık olabilir. Birleştirme aşamasında bu fazlalıklar silinir.

2.1.3. Dönüşüm

Dönüşüm aşaması ile veri ambarındaki veriler, veri madenciliği için uygun olan formlara dönüştürülmektedir. Bu aşamada aykırı veriler veriden kaldırılır. Veri normalleştirilmesi bu aşamada gerçekleştirilir. *Min-Maks*, *Z skor*, *Ondalık ölçekleme* gibi bazı veri normalleştirme teknikleri mevcuttur. En çok kullanılan veri normalleştirme tekniği *Min-Maks* tekniğidir. Bu teknikte orijinal veriler yeni veri aralığına doğrusal dönüşüm ile dönüştürülürler. Bu veri aralığı genellikle 0-1 aralığıdır [11].

2.1.4. Modelleme

Verinin diğer tüm veri madenciliği aşamalarından geçip, hazır hale getirildikten sonra veri madenciliği algoritmalarının biri ya da bir kaçının uygulanıp sonuç alındığı aşama modelleme aşamasıdır.

2.1.5. Yorumlama

Veri madenciliği aşamalarının sonuncusu olan yorumlama aşamasında, önceki aşamalarda üretilmiş olan modele, veri madenciliği algoritmaları uygulandıktan sonra alınan sonuçlar bu aşamada değerlendirilir. Çalışmanın doğru sonuçlanıp sonuçlanmadığı bu aşamada irdelenmektedir.

2.2. Veri Madenciliği Modelleri

Tanımlayıcı modeller ve tahmin edici modeller olarak ikiye ayrılmaktadır. Tanımlayıcı modeller kümeleme ve birliktelik kurallarıdır. Tahmin edici modeller ise sınıflandırma ve regresyondur.

2.2.1. Tanımlayıcı Modeller

Mevcut bir veri kümesi üzerindeki örüntülerin tanınarak karar verilmesi sağlayan modellerdir. Kümeleme ve birliktelik kuralları olarak bilinen iki türü mevcuttur. Kümeleme de amaç dağınık halde duran verileri özelliklerine göre birleştirerek işlenebilir hale getirmektir. Sınıflandırma işlemine benzemektedir fakat kümelemede önceden sınıf sayısı belirlenmemektedir. Birliktelik kurallarındaki ise büyük veri tabanlarında birbirleriyle ilişkili nesnelerin ve aralarındaki bağlantının büyüklüğünü tespit etmek amacıyla kullanılan yöntemdir.

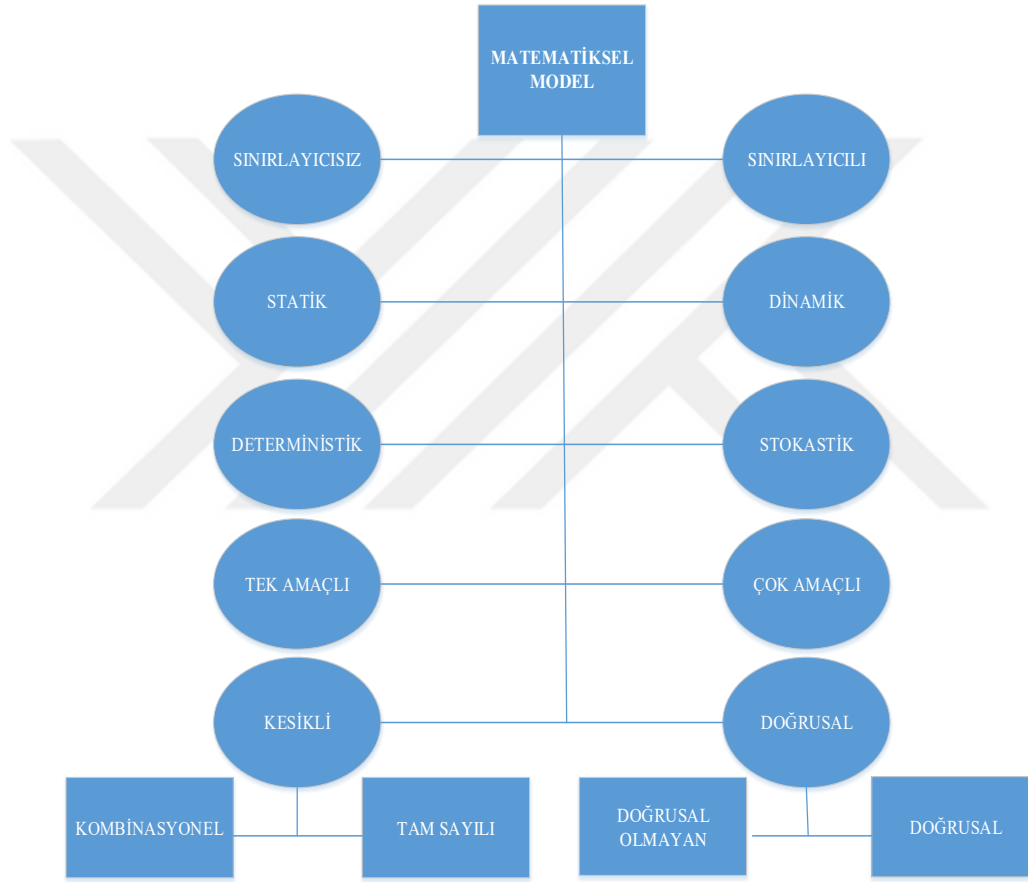
2.2.2. Tahmin Edici Modeller

Sonuçları bilinen verileri kullanarak oluşturulan bir model ile sonuçları bilinmeyen bir grup veri için sonuçların tahmin edilmesidir. Sınıflandırma ve regresyon olarak bilinen iki türü mevcuttur. Sınıflandırma, bir nesnenin niteliklerini inceleyerek gelecekteki eğilimlerini açıklamak için önceden tanımlanan sınıflara atama işlemidir. Regresyon da ise çok sayıda veri öğelerinin gerçek değerli tahmin değişkenlerine eşlenildiği bir öğrenme yöntemidir [12].

3. OPTİMİZASYON

3.1. Genel Optimizasyon

Bir problemin en iyi çözümünü elde etme işlemlerinin tümüne optimizasyon denilmektedir. Yani bir diğer deyişle optimizasyon, karar değişkenlerinin mümkün olan tüm olasılıklar arasında verilen bütün kısıtlayıcıları sağlayarak, uygunluk fonksiyonunu en iyi duruma (minimum ya da maksimum) getirme işlemidir. Şekil 3.1’de optimizasyon problemlerine ilişkin matematiksel modeller gösterilmektedir.



Şekil 3.1. Matematiksel modeller

Eğer karar değişkenleri üzerinde herhangi bir kısıtlama mevcut değilse sınırlayıcısız, herhangi bir kısıtlama varsa model sınırlayıcılı olmaktadır. Problem tek bir dönem için çözülecekse statik model, birden çok dönem dikkate alınarak çözülecekse dinamik model olmaktadır. Modelin algoritmada çalıştırılması esnasında, kesin parametreler kullanılıyorsa deterministik, olasılık durumu mevcutsa stokastiktir. Modelin birden fazla amacı mevcutsa çok amaçlı, eğer tek bir amacı mevcutsa tek amaçlı modeldir. Eğer tüm karar değişkenleri pozitif değerler alıyorsa sürekli

optimizasyon, eğer tüm karar değişkenleri tamsayı değerler alması gerekiyorsa kesikli optimizasyon modelidir.

Optimizasyon metotlarının en iyi sonucu verebilmesi için uygunluk fonksiyonu ve modelin gerçeğe uygun doğru bir şekilde belirlenmesi gerekmektedir. Matematiksel modellemeye bundan ötürü ihtiyaç vardır. Fakat gerçek dünya problemleri genellikle kısıtlı, karmaşık ve zor bir yapıya sahiptir. Bu tür problemlerin matematiksel modelini oluşturmak ya çok zordur ya da imkansızdır. Matematiksel model oluşturulmuş olsa dahi problemi çözmek için geçen zaman çok uzun olduğundan dolayı kullanılması zordur. Bu tür büyük ölçekli doğrusal olmayan problemler için klasik optimizasyon metotları yetersiz kalmaktadır.

Gerçek dünya problemlerini çözebilmek için eşitlik ve eşitsizlik kısıtları tanımlanmalıdır. Ancak bu şekilde uygunluk fonksiyonu en iyiye yakın optimal çözümü bulabilir. Bu şartlar altında Denklem 3.1 ve Denklem 3.2’de eşitsizlik kısıtlarına ait denklem verilmektedir. Denklem 3.3 ise eşitlik kısıtlarına ait denklem verilmektedir.

Eşitsizlik kısıtları:

$$g_k(x) \leq 0, \quad k = 1, 2, 3, \dots, q \quad (3.1)$$

$$low_i \leq x_i \leq up_i \quad i = 1, 2, 3, \dots, D \quad (3.2)$$

Eşitlik kısıtları:

$$h_j(x) = 0, \quad j = 1, 2, 3, \dots, m \quad (3.3)$$

$g_k(x)$ eşitsizlik kısıtı, $h_j(x)$ eşitlik kısıtı ve $f(x)$ uygunluk fonksiyonudur. $x = (x_1, x_2, \dots, x_D)$ D boyutlu bir vektördür (karar değişkeni). low_i ve up_i sırasıyla x_i ’nin alt ve üst sınırlarıdır. Kısıtlı gerçek dünya mühendislik problemlerinde amaç uygunluk fonksiyonunu eşitlik ya da eşitsizlik kısıtlarına uyarak minimum ya da maksimum yapmaktır. Gerçek dünya problemleri genellikle büyük ölçekli ve doğrusal olmayan problem türleri arasında yer almaktadır. Bu tür problemler için metasezgisel yöntemler kullanılmaktadır.

3.2. Metasezgisel Yöntemler

Metasezgisel yöntemler arama ve optimizasyon problemlerinin çözümünde kullanılan yöntemlerdir. Doğadan esinlenilerek geliştirilmiş sezgilere dayalı, basit, kolay uygulanabilir operatörleri bulunmaktadır. Problem çözüm uzayı sonsuz ya da bütün çözümlerin değerlendirilemeyeceği kadar büyük olduğu durumlarda, mevcut olan bu problemi makul bir zaman içerisinde çözebilecek ve en iyi çözüm olmasa bile en iyiye yakın optimal bir çözümü bulabilmek amacıyla metasezgisel yöntemler kullanılır [13, 14].

Yani çoğu algoritma belirli tipteki amaç fonksiyonu ya da sınırlayıcıların olduğu modelleri çözmektedir. Ancak çoğu yönetim ve organizasyon bilimi, teknoloji, çevre bilimi, ilaç sanayi, tıp, medikal, kimya gibi bir çok farklı alandaki optimizasyon problemleri eşzamanlı olarak

formülasyonlarında farklı tipteki karar değişkenleri, kısıtlayıcılar ve amaç fonksiyonu gerektirir. Bu nedenle sezgisel optimizasyon algoritmaları önerilmiştir. Bu algoritmalar son yıllarda oldukça popüler metotlar haline gelmiştir. Çünkü bu yöntemlerin hesaplama gücü oldukça iyidir ve dönüşümleri kolaydır. Şöyle ki tek amaç fonksiyonlu bir problem için yazılmış olan bir metasezgisel program, kolaylıkla çok amaçlı bir probleme ya da farklı bir probleme uyarlanabilmektedir [15].

Metasezgisel algoritmalara ihtiyaç duyulmasının nedenleri şunlardır:

- ✓ Optimizasyon problemi kesin çözümünü bulma işleminin tanımlanamadığı bir yapıya sahip olabilir.
- ✓ Metasezgisel algoritmalar, öğrenme amaçlı ve kesin çözümünü bulma işleminin bir parçası olarak kullanılabilir.
- ✓ Anlaşılabilirlik açısından metasezgisel algoritmalar karar verici açısından çok daha basit olabilir.
- ✓ Matematik formülleriyle yapılan tanımlamalarda genellikle gerçek dünya problemlerinin en zor tarafları (hangi amaçlar ve hangi sınırlamalar kullanılmalı, hangi alternatifler test edilmeli, problem verisi nasıl toplanmalı) ihmal edilir. Model parametrelerini belirleme aşamasında kullanılan verinin hatalı olması, metasezgisel yaklaşımın üretebileceği alt en uygun çözümden daha büyük hatalara sebep olabilir [16].

Metasezgisel algoritmalar değerlendirilirken şu kriterler göz önünde bulundurulmalıdır:

- ✓ *Çözüm Kalitesi ve Hesaplama Zamanı*: Herhangi bir algoritmanın etkinliğinin değerlendirilmesi için önemli bir kriterlerdir. Bir algoritma, ayarlanabilir parametreler kümesine sahip olmalı ve bu parametreler kullanıcıya önemlilik açısından hesaplama maliyeti ile çözüm kalitesi arasında bir vurgulamanın yapılabilmesine imkân sağlamalıdır. Başka bir ifade ile çözüm kalitesi ile hesaplama zamanı arasındaki ilişki kontrol edilebilmelidir.
- ✓ *Esneklik*: Algoritmalar modelde, sınırlamalarda ve amaç fonksiyonlarında yapılacak değişiklikleri kolayca karşılayabilmelidir.
- ✓ *Basitlik ve Analiz Edilebilirlik*: Karmaşık algoritmalar, esneklik ve çözüm kalitesi açısından basit algoritmalarından daha zor analiz edilebilmektedir. Algoritma kolayca analiz edilebilir olmalıdır.
- ✓ *Kod Basitliği ve Gerçeklenebilirlik*: Algoritmanın prensipleri basit olmalı ve genel olarak uygulanabilir olmalıdır. Böylece, algoritma uzman deneyimi olmadan ve problemin nasıl çözüleceği bilinmeden yeni problemlere kolaylıkla uygulanabilir.

- ✓ *Dinçlik*: Algoritma, başlangıç çözümüne bağlı olmaksızın her zaman kaliteli ve kabul edilebilir çözümler sunabilme yeteneğine sahip olmalıdır.
- ✓ *Etkileşimli Hesaplama ve Teknoloji Değişimleri*: Algoritma içinde insan-makine etkileşimini kullanma fikri çoğu sistemde yaygın olarak gerçekleştirilmektedir. Bilindiği gibi iyi bir kullanıcı ara yüzü herhangi bir bilgisayar sistemini veya algoritmayı daha çekici yapmaktadır. Bunun en önemli avantajı çözümlerin grafiksel olarak sergilenebilmesidir [16].

Optimizasyon algoritmaları, tek bir çözümünden başlayarak bunu operatörlerle iletiyorsa tek noktalı yöntemler olarak bilinir. Isıl işlem, tabu arama gibi yerel arama tabanlı sezgisel algoritmalar bu kategoriye girer. Çok noktadan yani bir popülasyon üzerinden çözüme başlayarak farklı noktalarla optimizasyon yapıyorsa çok noktalı ya da popülasyon tabanlı yöntemler denir. Genetik Algoritma ve Parçacık Sürüsü Optimizasyonu bu yöntemlere örnek verilebilir [13].

Metasezgisel algoritmaların birçoğu tek bir komşuluk yapısında çalışır ve tek komşuluk yapılı olarak sınıflandırılabilir. Bazıları ise arama işlemini sistematik bir şekilde değiştirerek birden fazla yerel arama yöntemiyle diğer çözüm alanlarına ulaşmaya çalışır ve değişken komşuluk yapılı şeklinde sınıflandırılabilir [13].

Metasezgisel yöntemlerin bir kısmı problemi gösterirken amaç fonksiyonunu sabit tutar ve bu yöntemlere sabit amaç fonksiyonlu denir. Bazıları ise, yerel minimumdan kurtulmak için amaç fonksiyonlarını değiştirir ve değişken amaç fonksiyonlu olarak adlandırılır.

Metasezgisel yöntemler çalışırken daha önceki durumlar ya da en iyi durumlar hatırlanıyorsa hafızalı, hatırlanmıyorsa hafızasız şeklinde sınıflandırılır. Örneğin, tabu arama ve Parçacık Sürüsü Optimizasyonu hafızalı, Genetik Algoritma hafızasızdır [13].

Metasezgisel yöntemlerin çoğu doğadan esinlenerek geliştirilmiştir. Metasezgisel yöntemler doğadan esinlendikleri alana göre fizik tabanlı (Optik İlham Optimizasyon (OIO) [17], Merkezi Kuvvet Optimizasyon Algoritması [18], Uzay Yerçekimi Optimizasyonu [19], Parçacık Çarpışma Algoritması [20]), sosyal tabanlı (İnsan Grup Formasyon Algoritması [21], Toplum ve Medeniyet Algoritması [22], Kültürel Algoritma [23]), müzik tabanlı (Harmoni Arama Algoritması [24], Müzikal Kompozisyon Algoritması [25], Melodi Arama Algoritması [26]), sürü tabanlı (Parçacık Sürüsü Optimizasyonu [27], Kedi Sürüsü Algoritması [28], Yapay Arı Koloni Algoritması [29]), kimya tabanlı (Kimyasal Tepkime Optimizasyonu [30], Yapay Kimyasal Tepkime Optimizasyon Algoritması [31], Yapay Atom Algoritması [32]), biyoloji tabanlı (Genetik Algoritma [33], Diferansiyel Gelişim Algoritması [34], Yapay Bağışık Sistemler Algoritması [35]), spor tabanlı (Lig Şampiyonası Algoritması [36], Altın Top Algoritması [37], Dünya Kupası Optimizasyon Algoritması [38]), matematik tabanlı (Sinüs Kosinüs Algoritması [39]) olmak üzere 8 farklı gruba ayrılabilir. Bunların dışında birden çok metasezgisel algoritmanın beraber kullanılmasıyla elde

The diagram illustrates the classification of Metaseizmic Methods (Metasezgisel Yöntemler) based on their foundational basis (Tabanlı) and their characteristics (Özellikler).

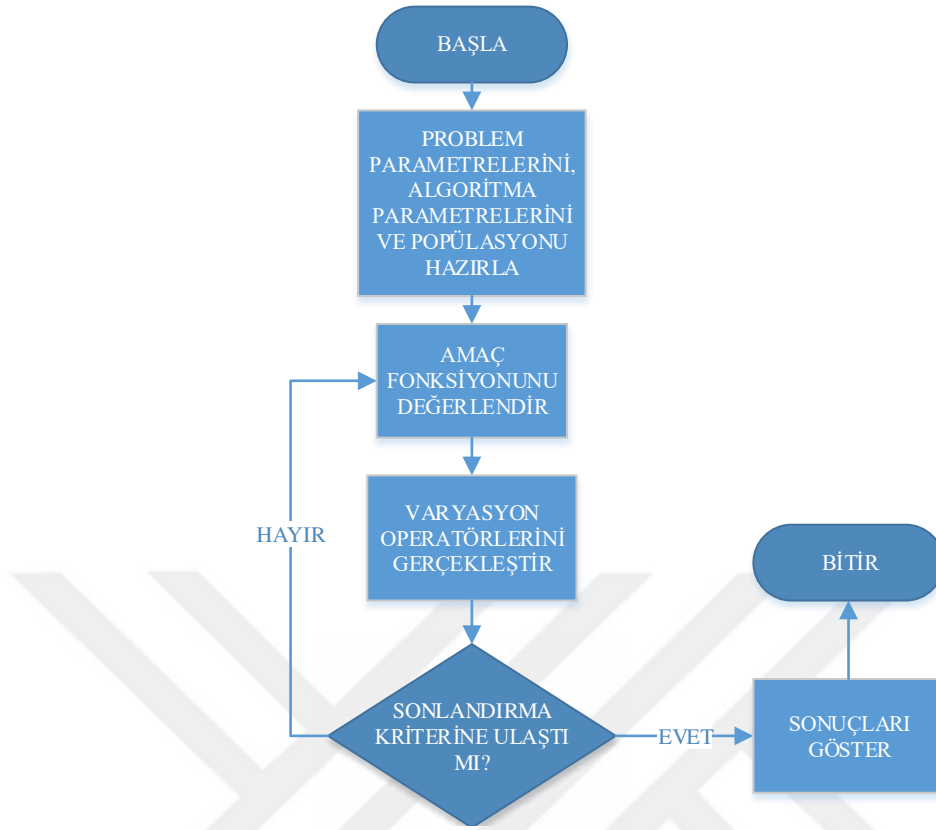
Metasezgisel Yöntemler are categorized into nine foundational bases (Tabanlı):

- Fizik Tabanlı
- Kimya Tabanlı
- Sürü Tabanlı
- Biyoloji Tabanlı
- Müzik Tabanlı
- Spor Tabanlı
- Matematik Tabanlı
- Melez Tabanlı
- Sosyal Tabanlı

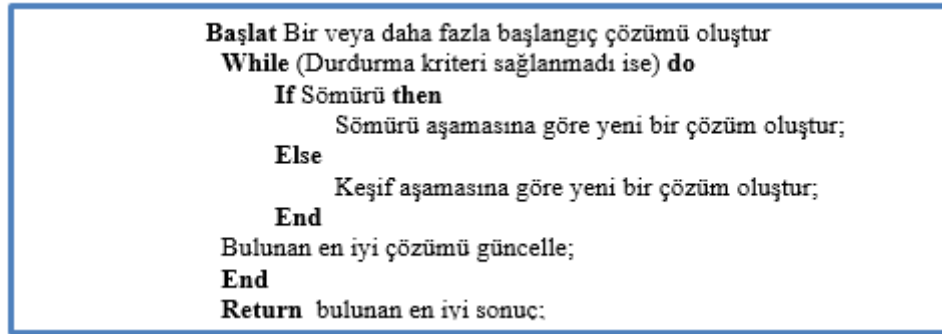
These methods are further characterized by four pairs of opposing features (Özellikler):

- Tek Noktalı** (Single Point) vs. **Çok Noktalı** (Multiple Points)
- Sabit Amaç Fonksiyonu** (Fixed Objective Function) vs. **Değişken Amaç Fonksiyonu** (Variable Objective Function)
- Tek Komşuluk Yapısı** (Single Neighborhood Structure) vs. **Değişken Komşuluk Yapısı** (Variable Neighborhood Structure)
- Hafızalı** (Memory-based) vs. **Hafızasız** (Memoryless)

14



Şekil 3.3. Popülasyon tabanlı metasezgisel algoritmaların genel akış diyagramı



Şekil 3.4. Popülasyon tabanlı metasezgisel algoritmaların sözde kodu

4. ALDATMA VE SÖYLENTİ TESPİTİ

İnternet teknolojilerinin hızla yayılması, bilgiye erişim bakımından düşünüldüğünde insanların hayatlarını kolaylaştırdığı yadsınamaz bir gerçektir. SM kullanımı internet teknolojileri içerisinde en yaygın kullanılan araçlardır. SM kullanmayan ya da en az bir tane SM hesabı olmayan birey günümüzde yok denecek kadar azdır. SM sunmuş olduğu imkanlar, insanoğlunun hayatını çok kolay bir hale getirmektedir. Bu imkanlar için bazı örnekler verecek olursak, ev hanımları artık herhangi bir yemeğin tarifini bir deftere not etmek zorunda kalmıyor, bir çok farklı yemek yapma üzerine kurulmuş olan web siteleri, bloglar ve video siteleri sayesinde kolayca bu tariflere ulaşıyorlar. Öğrenciler öğrenmek istedikleri konulara ait ders videolarından kolaylıkla mekandan bağımsız bir şekilde erişmekte ve öğrenmektedirler. Siyasetçiler ve politikacılar SM araçlarını etkili bir şekilde kullanmakta ve seçmenlerine sunmuş olduğu hizmetlerini anlatmaktadırlar. SM'yi aslında bir reklam ve tanıtım aracı olarak kullanmaktadırlar. Tüm dünya ülkeleri, insanlığın tamamını olumsuz bir şekilde etkileyen, birçok ülkede yıkıcı sonuçları olan Covid-19 hastalığını hakkındaki gelişmeleri ve bu hastalıkla ilgili yeni çıkan haberleri, Twitter başta olmak üzere birçok SM platformu kullanılarak takip etmektedirler. Bu örnekler daha çok artırılabilir. Ancak bu olumlu taraflarının yanında olumsuz etkileri de görmezden gelinemez. Bunlardan ilki SM aracılığı ile güvenilirliği tartışmalı, aldatıcı içerik barındıran yanlış bilgiye erişen kişilerin aldatılması olayıdır. Ötekisi ise yine SM aracılığı ile mevcut bir bilginin söylenti şeklinde, ağızdan ağıza dolaşırken, doğru olan bilginin değişmesi ve bilginin kirlenmesi durumudur. Aldatma ve söylenti tespiti SM problemi olarak insanların aldatılması, kandırılması, yanlış bilgilere inandırılması durumlarını ortaya çıkarır. Bu durum sosyal etkileri olan toplumsal bir sorundur. Çok tehlikeli bir durumdur. Çünkü yanlış öğrenilen bilgiler kolay kolay değiştirilemezler. İnsanlar doğru diye inandıkları bilgilerin, aslında yanlış olduğunu öğrendiklerinde dahi etkisinden kolay kolay kurtulamazlar.

4.1. Aldatma Tespiti

4.1.1. Aldatıcı İçerik Nedir?

Aldatma, genel olarak belirli bir amaç için kişilerin belli bir konu üzerine kitlelerin düşüncelerini yönlendirme ve toplumsal algı oluşturmayı amaçlar. Başka bir ifadeyle;

Aldatıcı içerik, sosyal medyada insanların herhangi bir konu hakkında, gerçek olmayan bilgiler içeren, bu sayede insanların hayatlarına doğrudan etki edecek, belki de onların daha sonra pişman olacakları kararları almasına sebep olacak sahte metin, fotoğraf ve videoların tamamına denir.

4.1.2. Aldatmanın Toplum Üzerindeki Etkileri

Aldatmanın toplum üzerindeki olumsuz etkilerini incelerken çeşitli örnekler üzerinden giderek topluma vermiş olduğu zararlara da dikkat çekilmeye çalışılmıştır. Bu örnekler şöyledir:

Yakın tarihe kadar insanlar uzaktan alışveriş yapmazlar, bir ürünü görmeden satın alma işlemini gerçekleştirmezlerdi. Fakat hem teknolojideki gelişmelerin etkisiyle hem de Covid-19 salgın hastalığının insanları evlerine kapatmasıyla, artık internet üzerinden alışveriş yapmak hem çok kolaylaştı hem de çok yaygın hale geldi. Mallarını uzaktan satmak isteyen tüccarlar çeşitli web siteleri üzerinde sanal marketler açmaya başladılar. Bu mal ve ürünleri buradan satın alan kişiler hem ürün hakkında hem de satıcı hakkında yorum yaparak, bu satın alma işlemi hakkındaki tecrübelerini diğer tüketicilerin okuması için paylaşmaktadırlar. İşte tam da bu noktada aldatma olayı meydana gelmektedir. Yani aldatıcı fikir sahteciliği [40] insanların güven duygularını istismar ederek bir mal ve ürünü kandırarak pazarlamaktır. Bu işlem esnasında firma personeli sanki daha önce bu mal ve ürünü satın almış ve memnun kalmış gibi yorumlar yapmakta, bu mal ve ürünü satın almayı planlayan kişileri kandırmaya çalışmaktadır. Bu aldatıcı yorumları okuyarak ya da bu aldatıcı görsellere itibar ederek satın alma işlemi gerçekleştiren insanların güven duyguları kökünden zedelenmekte, hem uzaktan alışveriş hakkındaki fikirleri olumsuz etkilenmekte hem de artık insanlara güvenmemeye başlamaktadır. Toplum genelinde güvensizlik oluşmakta ve insanlar böyle alışverişlerden kaynaklanan zararlar sonucunda mutsuzluk yaşamaktadırlar.

4.1.3. Aldatma Tespiti ile İlgili Literatür Taraması

Günden güne popülaritesi artan aldatma tespit problemini çözmek için literatürde çeşitli yaklaşımlar vardır. Bilim dünyası bu problemi çözmek için ciddi çaba harcamaktadır.

Ott vd. [40] yapmış oldukları çalışmada tüketici yorumlarını içeren web sitelerinin, aldatıcı içeriklerin hedefi haline geldiğini ifade etmişlerdir. Bu çalışmada aldatıcı içeriği tespit etmek için üç adet yaklaşım geliştirmişler ve bu yaklaşımların performanslarını karşılaştırmışlardır. Gerçekleştirilen deneyler sonucunda standart *n-gram* tabanlı metin kategorizasyon yönteminin en başarılı yöntem olarak bulunduğunu ifade etmişlerdir.

Feng vd. [41] çalışmalarında aldatma tespiti üzerine yapılan önceki çalışmaların çoğunun sözdizimsel kalıplara dayandığını bildirmişlerdir. Sözdizimsel stilometrinin araştırıldığı bu çalışmayla literatüre daha önce olmayan bir bakış açısı kazandırıldığı ifade edilmiştir. DeneySEL sonuçlara göre %14 hata azaltma ile %91,2 doğruluk oranının elde edildiği belirtilmiştir. Bu çalışmada büyük veri kullanılmamıştır.

Song vd. [42] yapmış oldukları çalışmada popüler aldatma tespit yöntemlerinin avantaj ve dezavantajlarını incelemişlerdir. Bu çalışmada aldatıcı davranışların zamanla değişebileceği ve bu

nedenle mevcut yöntemlerin, aldatıcı davranışları etkili bir şekilde tespit etmelerini zorlaştırdığı belirtilmiştir.

Shojaee vd. [43] çalışmalarında aldatmayı tespit etmek için denetimli makine öğrenme sınıflandırıcılarını kullanmışlardır. Mevcut bir otel incelemesi üzerinde yapılan deneyler neticesinde aldatıcı görüşlerin tespitinde ümit verici sonuçların elde edildiği belirtilmiştir.

Merritts doktora tezinde BDI ajanları kullanılarak aldatma olayının otomatik ve özerk bir şekilde gerçekleştirilmesini amaçlamıştır. Çalışma sonucunda geliştirilen prototip ile verilerin %85 doğruluk oranı ile "aldatıcı" veya "aldatıcı olmayan" olarak sınıflandırılabilceği belirtilmektedir [44]. Bu çalışmada umut verici doğruluk oranlarına ulaşılmış olsa da diğer performans değerlendirme metrikleri cinsinden de başarı oranı incelenmelidir.

Azaria vd. [45] yapmış oldukları çalışmada DIG adlı yenilikçi bir makine öğrenme tabanlı ajan geliştirmiştir. Bu çalışmada DIG'in performansı dört oyunculu bir korsan oyunuyla ölçülmüştür. DIG'in bir grup oyuncunun sıradan üyeler olarak katıldığı bir ortamda, aldatıcı oyuncuyu bulmaya yardımcı olduğunu belirtmişlerdir.

Feng ve Hirst [46], otel incelemeleri için bir veri kümesi olan op_spam_v1.3 [40] 'yi kullanarak ürünlerin orijinal olup olmadığını ayırt etmeye çalışmıştır. Deneyler sırasında profil uyumluluğu kullanılmıştır. Bu şekilde, sınıflandırma performansının önemli ölçüde iyileştirildiğini belirtmişlerdir.

Ott vd. [47] çalışmalarında duygu ve aldatma arasındaki olası etkileşimleri değerlendirmiştir. Bu ilişkinin daha fazla araştırılmasının teşvik edilmesi gerektiğini belirtmişlerdir. İlk sonuçları makine öğrenimi algoritmalarını kullanarak elde etmişlerdir. SVM sınıflandırıcı kullanılarak yapılan bu işlem sonucunda yaklaşık % 86 doğruluk oranına ulaşıldığını belirtmişlerdir. Diğer birçok çalışmada olduğu gibi başarı açısından doğruluk tek başına yeterli bir parametre değildir ancak problemin çözümü noktasında fikir sağlama açısından önemli bir parametredir.

Tsikerdekis ve Zeadally [48], yaptıkları çalışmada, birçok sosyal medya türüne uygulanabilecek kimlik aldatmacasında sözselsel olmayan davranışa dayalı bir aldatma tespit yöntemi sunmuşlardır. Deneyler sırasında Wikipedia ortamı kullanılmıştır. Önerilen yöntemin, diğer yöntemlere kıyasla daha yüksek bir doğruluk oranına ulaştığı belirtilmektedir. Bilimsel olarak hassasiyet, kesinlik ve F-ölçütü gibi parametrelerde de yüksek oranlara ulaştığı noktada önermiş oldukları yöntemin geçerliliği kabul edilebilir. Makalede bu parametrelerin yüksek başarımlara sahip olup olmadıkları noktasında her hangi bir bilgi verilmemektedir.

Li vd. [49] genel olarak yapmış oldukları çalışmada aldatıcı fikirlerin doğası hakkında daha derin bir anlayış elde etmeye çalışmışlardır. Üç farklı alandan (otel, restoran ve ilaç) toplanarak oluşturulan yeni bir veri kümesini kullanarak, çevrimiçi aldatıcı fikirleri belirlemeye yönelik genel bir yaklaşımı araştırmışlardır. Aldatıcı içeriklere ait karakteristiklerin çoğunun iyi eğitilmiş aldatıcı içerik yazarları tarafından aşılabileceğini ve bunun daha zor bir konu olan aldatıcı fikirlere sebep

olacağı hakkında uyarıda bulunmaktadırlar. Önermiş oldukları modelleri çeşitli psikolojik aldatma teorilerini keşfetmek amacıyla da kullanılabileceğini belirtmişlerdir.

Huayi Li vd. [50] önceki aldatma çalışmalarında genellikle İngilizce olarak belirtilen metinlerin kullanıldığını ve İngilizce için bir aldatma tespiti önerildiğini bildirmiştir. Ayrıca yazarlar literatüre daha önce olmayan bir yenilik ekleyerek kendi çalışmalarında Çince metinlerdeki aldatıcı içeriği tespit etmeye çalışmışlardır. Bu işlemleri gerçekleştirmek için Çin inceleme sitesi Dianping'den alınan veri seti kullanılmıştır.

Alowibdi vd. [51] çevrimiçi sosyal ağlarda aldatıcı profilleri belirlemek için yeni bir yaklaşım önermiştir. Twitter'ı çevrimiçi bir sosyal ağ ortamı olarak kullanarak, kullanıcının cinsiyeti ve konumu hakkında aldatıcı bilgileri tanımlamak için özel olarak bir dizi analiz yöntemi belirtmişlerdir. Yapılan testlerde, aldatıcı içerikleri kabul edilebilir bir doğrulukla belirlemek için Bayes sınıflandırıcı ve K-ortalamlar kümeleme algoritmalarını kullanılmışlardır.

Conroy vd. [52] yapmış oldukları çalışmada aldatıcı içeriği tespit etmek için insanların davranış verilerini, dilbilimi, makine öğrenimi ve ağ analizi yaklaşımlarını birleştirerek yenilikçi bir karma yöntem önermişlerdir. Bu yöntemde öncelikle dil kalıplarını aldatma ile ilişkilendirerek analiz edildiği dil yaklaşımları incelenmekte, daha sonra ağ sorgularına ve iletilere ait yapısal bilgiler incelenmekte ve toplu aldatma olaylarının önüne geçecek bir yöntem önerilmektedir. SVM ve NB sınıflandırıcılarını deneyler esnasında kullanmışlardır.

Rubin vd. [53] kullanıcılar için olası aldatıcı haber türlerini tanımlayan ve kullanıcılara filtreleme konusunda yardımcı olmayı amaçlayan bir sistem önermişlerdir. Çalışmada mizah, büyük aldatmacalar ve hiciv olarak üç tür yalan haber türüne detaylıca yer verilmiştir. Tahmine dayalı modelleme için artılar ve eksiler analiz edilmektedir.

Pisarevskaya vd. [54] çalışmalarında rastgele orman ve SVM sınıflandırıcıları kullanarak Rusça metinlerde aldatma tespiti yapmışlardır. Deneyler sırasında, SVM ve rastgele orman sınıflandırıcılarının doğruluğunu sırasıyla % 57 ve % 42 olduğunu belirtmişlerdir. Literatürdeki diğer yayınlanan makaleler göz önüne alındığında metriklerdeki başarı oranları genel olarak düşük kalmaktadır.

Kumari ve Srivastava [55], metin madenciliğinde yaygın olarak kullanılan ikili sınıflandırıcıları tartışmıştır. Ayrıca ikili sınıflandırma metodu kullanılarak çoklu kimlik aldatma tespiti ve metin kategorizasyonu üzerine yapılan bilimsel çalışmaların özetini vermişlerdir.

Rubin, sosyal medyadaki kullanıcıların, çevrimiçi içeriklerin kasıtlı bir şekilde manipülasyona uğradığının pek farkında olmadığını ifade etmektedir. Makalede doğrulanmamış bilgilerin, kullanıcıların en yüksek güven düzeyine ulaştığı belirtilmektedir. İçerik doğrulama sistemleri ve melez teknolojilerin bu sorunun çözümünde önemli bir rol oynayacağı ifade edilmektedir [56].

Molla vd. [57] yapmış oldukları çalışmada Samsung ürünleri ve hizmetleri hakkında olumsuz tweetler toplayarak oluşturdıkları veri setinden, her yazarın yazma stilini ve kişisel profilini analiz ederek, aldatma tespitinde kimliğin belirlenmesine yardımcı olabilecek bir yaklaşım sağlamaktadır. Testler sırasında aldatma tespit problemi için en iyi algoritmanın, maksimum entropi algoritması olduğu belirtilmektedir. Deneylerde kullanılan veri sayısı artırılarak ya da başka algoritmalar deneyerek daha iyi sonuçlar alınabilirdi.

Rastogi ve Mehrotra [58], yapmış oldukları çalışmada, aldatma tespiti sürecinde metinler incelenirken kullanılan veri seti sayısının oldukça sınırlı olduğunu belirtmektedirler. Bu nedenle, mevcut yaklaşımların çoğunun doğruluk yüzdesi düşük kalmaktadır. Ayrıca literatürde aldatmanın tespiti için yapılan çoğu çalışmada metinsel ve davranışsal özelliklerin ayrı ayrı kullanıldığı ve sadece birkaç çalışmada bu özelliklerin birlikte kullanıldığı belirtilmektedir.

Rosso ve Cagnina [59], yapmış oldukları çalışmada, psikolojik, dilbilimsel ve hesaplama süreçlerinin sürekli olarak aldatmayı tespit etmede zorluklar ortaya çıkardığını ifade etmektedir. Bu çalışmada veri sayısı yeterli değildir ve büyük veri kullanılmamıştır.

Kumar vd. [60] çalışmalarında, aldatma tespiti sorununun, çevrimiçi ağlarda işletmeler için de ilginç bir konu haline geldiğini belirtmektedir. Çevrimiçi ağlarda aldatıcı yorumları azaltmanın ve bu ağların güvenilirliği konusunda tüketicinin güvenini artırmanın önemli olduğu vurgulanmaktadır. Yelp, TripAdvisor ve Amazon'dan elde ettikleri verileri kullanarak aldatma tespiti problemini çözmeye çalışılmıştır. Sınıflandırıcı olarak makine öğrenme algoritmaları kullanılmakta ve bu algoritmaların performansları incelenmektedir.

Rangel vd. [61] yapmış oldukları çalışmada arapça bir metnin aldatıcı olup olmadığını bulmaya çalışmıştır. Ayrıca yine bu çalışmada Twitter kullanıcılarının yaşını, cinsiyetini ve ana dilini belirlemeye çalışmışlardır. Bu amaçla 13 katılımcıya ait veriler değerlendirilmektedir. Geleneksel yaklaşımların derin öğrenme yaklaşımlarından daha iyi performans gösterdiğini belirtmişlerdir. En yüksek başarı oranı SVM ile elde edildiği vurgulanmaktadır.

Nayel vd. [62] yapmış oldukları çalışmada, doğrusal sınıflandırıcı, SVM ve çok katmanlı algılayıcı sınıflandırıcı algoritmalarını kullanarak Arapça metinlerde aldatıcı içeriği sınıflandırmaya çalışmıştır. SVM algoritmasının diğer makine öğrenme algoritmalarına göre daha başarılı olduğunu belirtmektedirler.

Sen vd. [63] çalışmalarında konu düzeyinde aldatmayı sınıflandırmak için görsel, dilbilimsel ve akustik verileri birlikte kullanarak bir sınıflandırıcı geliştirdiklerini vurgulamaktadırlar. Geliştirdikleri sınıflandırıcı ile % 83.05 doğruluk oranına ulaştıkları belirtilmektedir.

Delgado vd. [64] çalışmalarında aldatıcı içeriği tespit etmek için Decision Tree (DT), Logistic Regression (LR), Naive Bayesian (NB), SVM, Neural Networks (NN), Random Forest (RF) yöntemleri gibi makine öğrenme tekniklerini kullanmışlardır. Araştırmacılar, sınıflandırma işlemi için kelime çantası (Bag of Words) ve konuşma parçası (Part of Speech) yöntemlerini

kullanmışlardır. Aldatıcı içeriğin tespiti için veri kaynağı olarak haberler ve e-postalar kullanılmıştır.

Krishnaveni ve Radha [65], yaptıkları çalışmada metin sınıflandırma algoritmalarından daha iyi sonuçlar elde etmek için kümeleme algoritmalarıyla birlikte kullanılabileceğini belirtmektedir. Deneylerde NB, SVM ve DT sınıflandırma algoritmaları ile K-ortalamlar, One-Pass ve DBScan kümeleme algoritmaları birlikte kullanılmıştır. K-ortalamlar ve SVM algoritmalarının birlikte kullanıldığı durum en başarılı durumun olduğu belirtilmektedir.

Kesarwani vd. [66] yapmış oldukları çalışmada Facebook'tan elde edilen verileri aldatmayı tespit etmek için kullanmışlardır. Deneylerde LR, RF ve SVM gibi makine öğrenme algoritmaları kullanılmıştır. LR sınıflandırma algoritması % 98,25 doğruluk oranı ile en yüksek performansı gösterdiği belirtilmektedir.

Sternglanz vd. [67] yapmış oldukları çalışmada, kolluk kuvvetleri tarafından kullanılan aldatmayı tespit etme yöntemlerini incelemektedir. Ayrıca bu çalışma kapsamında meta-analitik çalışmalar hakkında bilgi verilmektedir.

Rill-Garcia vd. [68] video verileriyle aldatma olayını tespit etmeye çalışmıştır. Elde ettikleri sonuçları çok modlu füzyon yöntemi kullanarak iyileştirmeye çalışmışlardır. İspanyolca veri kümesi deneyler sırasında kullanılmıştır.

Jaume Masip [69], çalışmasında aldatma tespitinin son zamanlarda önemli gelişmeler yaşayan canlı, dinamik bir psikoloji alanı olduğunu belirtmiştir. Ayrıca insanların aldatmayı nasıl saptadıkları üzerine araştırma ve çağdaş teoriler sunmaktadır. Bunun yanında aldatmacayı tespit etmek için stratejik görüşme konusunda son gelişmeleri vurgulamıştır. Sistematiik sözlü yalan tespit yaklaşımlarına ilişkin son meta-analizlerin bütünleştirici bulguları bu çalışma kapsamında sunulmuştur. Aldatmanın psikofizyolojik tespiti ile ilgili bazı hususlara bu çalışma içerisinde yer verilmiştir.

Litvinova vd. [70] yapmış oldukları çalışmada daha önce literatürde olmayan bir yenilik olarak Rusça olarak hazırlanan gerçek dünya veri setini çalışmalarında kullanmışlardır. Seçilen parametrelere göre önerilen sınıflandırıcı ile % 68,3 doğruluğa ulaşmışlardır.

Ding vd. [71] yapmış oldukları çalışmada aldatmayı gerçek hayattaki videolardan tespit etmenin çok zor olduğunu belirtmektedirler. İnsan vücuduna ait öğelerin aldatmadaki rolü üzerinde genel bilgiler verilmekte ve aldatıcı içeriği tespit etmek için insan vücudunu ve yüzünü birlikte değerlendirmeyi önermektedirler.

Kleinberg vd. [72] yapmış oldukları çalışmada sözlü güvenilirlik değerlendirmesi ve bilimsel yanıltıcı sözlü aldatma araştırmalarının, yasal ortamlarda hayati öneme sahip olduğunu vurgulamaktadırlar. Bu nedenle, sözlü aldatma tespit yöntemlerine ait raporların en yüksek standartları karşılaması kamuoyunun, bilim camiasının ve kanun uygulayıcılarının yararına olması

gerektiğini belirtmektedirler. Aldatıcı içeriği tespit etmek için çapraz doğrulama tekniğinin kullanılması gerektiğini öne sürmüşlerdir.

Van der Walt vd. [73] çalışmalarında profil resmi özelliklerini kullanarak aldatma olayını tespit ettiklerini belirtmektedirler. Makine öğrenme yöntemlerini kullanarak bir sınıflandırma işlemi gerçekleştirdiği belirtilmektedir.

Krishnamurthy vd. [74], çalışmalarında aldatıcı içeriğin tespiti için çok modlu bir sinir modeli geliştirdiğini belirtmektedir. Deneysel işlemlerin sonucunda % 96.14 doğruluk oranı ile bilinen tüm yöntemlerden daha iyi bir performans elde ettiklerini vurgulamaktadırlar.

4.2. Söylenti Tespiti

4.2.1. Söylenti Nedir?

Söylenti, belli bir grup için önem taşıyan, henüz gerçekleşmemiş olan bir olayın sanki olmuş gibi kontrolsüz bir şekilde kişiden kişiye aktarılan ve resmi olmayan bir iletişim türüdür. Başka bir ifadeyle;

Belli bir olayın olduğu yolunda bir topluluk ya da toplumsal kümede dolaşan, sürekli değişime uğrayan doğrulanmamış bilgidir. Söylenti genel olarak üç evreden oluşmaktadır. Bunlar;

- ✓ Ortaya çıkış
- ✓ Yayılma
- ✓ Sona erme

Ortaya çıkış evresi genellikle kötü niyetli kişilerin ortaya attıkları kirli bilgilerin, buna inanmak isteyen kişi ya da kişiler tarafından erişilmesi evresidir. Yayılma evresinde birden fazla insana gereksinim vardır. Şöyle ki bir kişi anlatma ihtiyacını giderirken, karşısındaki de konu hakkındaki merakını gidermektedir. Bu aktarımlar esnasında kişiler kendinden ifadeler de katacağı için söylenti sürekli bir değişim halinde olacaktır. Söylenti anlatıldıkça kısılacaktır ve daha az detaylar içerecektir. Genel itibari ile söylenti anlatan kişinin değerlerine, görüşlerine ve ön yargılarına göre şekillenmektedir. Sona erme evresi ise, resmi kurumların ifadeleri doğrultusunda söylentinin yanlış olduğunun anlaşılması ya da mevcut durumun değişime uğrayarak gerginlik yaratan nedenlerin ortadan kalktığı durumlarda söylentinin de sona erdiği evredir [75].

4.2.2. Söylentinin Toplum Üzerindeki Etkileri

Söylentiler toplumun tutum ve davranışlarını ciddi şekilde etkileyen bir kitle olayıdır. Söylenti gücünü ona inanmak isteyen insanlardan almaktadır. Toplumsal kaygı ve endişenin kaynakları arasında söylentiler mevcuttur. Şöyle ki Irak'ın elinde tüm dünyayı yok edecek kimyasal ve biyolojik silahların olduğu söylentisi, başta Irak'ın düşmanı olan ülkelerdeki toplumlar ve

Irak'ın komşuları olan ülkelerin toplumları olmak üzere tüm dünyadaki toplumlar üzerinde psikolojik, sosyolojik bir baskı oluşturmaktadır. Başka örnek olarak Elazığ'da yakın zamanda şiddeti 7 ve üzerinde deprem olacak söylentisiyle, bölge halkının sokaklarda yaşamaya başlaması, aslında gerçeği değil söylentinin yaymış olduğu korkuyu topluluk halinde yaşamalarına sebep olmaktadır. Bir başka örnekte ise Türkiye'nin güney sınırında oluşturulmak istenen terör koridorunu yok etmek, bölgeye barış ve huzur getirmek için başlattığı Barış Pınarı Harekatı da söylentilerin hedefinde yer aldı. Şöyle ki terör örgütlerine yakın olan SM hesapları tarafından Yemen ve Gazze de çekilmiş olan katliam fotoğraflarının sanki Barış Pınarı, Zeytin Dalı ve Fırat Kalkanı harekâtlarında çekilmiş gibi kötü niyetli olarak servis edilmeye çalışıldı. Bağımsız ve art niyetli olmayan medya kuruluşları tarafından bu söylentiler araştırılarak, toplum gerçek olan bilgiler ışığında aydınlatılmıştır. Bir başka söylenti tüm dünyayı ilgilendiren Covid-19 pandemisi ile ilgilidir ki ülkemizde bu hastalıkla ilgili alanında uzman kişilerden oluşan bir bilim kurulu olmasaydı, toplum art niyetli SM hesapları tarafından yayılan çeşitli söylentilerle ciddi bir şekilde kandırılabilirdi. Bu söylentiler karşısında bilim kurulu anında reaksiyon gösterip ortaya çıkan söylentilerin gerçeği yansıtmadığını ve doğrusunun ne olduğunu izah ederek toplumu rahatlatmaktadır. Bir üst heyet ya da idare tarafından, toplumun genelini ilgilendiren bir konu, olay hakkında söylentinin aksi iddia edilmezse ve gerçek bilgiler insanlarla paylaşılmazsa toplumsal kargaşa çıkması kaçınılmazdır. Bu nedenlerden ötürü genellikle SM platformları kullanılarak ortaya çıkan bu söylentilerin en erken zamanda tespit edilmesi, beraberinde getirebilecekleri birçok sorunun önüne geçilmesini sağlayacaktır.

4.2.3. Söylenti ile İlgili Literatür Taraması

Geçmiş zamanlarda henüz internet teknolojileri günümüzdeki etkinliğine sahip değilken, insanların kendilerini dinleyen diğer insanlara anlatmasıyla söylenti yayılmaktaydı. Artık günümüzde SM platformları kullanılarak oluşturulan söylentiler hem daha çok kişiyi etkisi altında bırakmakta hem de daha hızlı bir şekilde yayılmaktadır. Teknolojik imkanlardan ötürü günümüzde söylentilerin çoğunluğunun SM platformlarında olmasından dolayı, bilim dünyası da daha çok SM platformları üzerinde ki söylentinin tespiti üzerinde yoğunlaşmış durumdadır. Toplum üzerinde çok ciddi etkileri olan ve muhataplarını yakından ilgilendiren bu sosyal problemi çözmek için bilim dünyası ciddi çaba harcamaktadır. Her geçen gün popülaritesi daha da artan söylenti tespit problemini çözmek için literatürde çeşitli yaklaşımlar mevcuttur.

Yang vd. [76] yapmış oldukları çalışmada Çin'in en büyük mikro blog sitesi olan Sina Weibo üzerinde karışık haldeki bir dizi doğru ve yanlış bilgi içeren veriden söylentiye otomatik olarak tespit eden bir yöntem geliştirdiler. Yine bu çalışmada söylentiye tespit etmek için retweet sayısı, konum ve tweetlerin içerik türlerinin incelendiği belirtilmektedir.

Gupta vd. [77], çalışmalarında Boston etkinliği sırasında Twitter'da altı binden fazla kötü niyetli hesap tespit ettiğini belirtmektedir. Bu hesapların daha sonra Twitter tarafından askıya alındığı bildirilmektedir. Yine bu çalışmada askıya alınmış olan bu profillerin kendi aralarındaki etkileşim ağı içerisinde kapalı topluluk yapısı ve yıldız oluşumuna sahip oldukları vurgulanmaktadır.

Aggrawal [78], söylenti tespiti üzerine kapsamlı bir çalışma sunmuştur. Söylenti tespiti için deneylerde kosinüs benzerlik yöntemi kullanmıştır. %94 doğruluk oranı elde etmiştir. Ayrıca yine bu çalışmada kural tabanlı, NB 'den daha başarılı bir sözcük sınıflandırıcı geliştirdiğini belirtmektedir.

Jin vd. [79] söylentiyi tespit etmek için, söylenti içeren tweetleri, doğrulanmış söylenti makaleleriyle eşleştirerek güvenilir ve yorumlanabilir bir yaklaşım önerildiği belirtilmiştir. Söylenti tespitinde 5 adet algoritma kullanılarak, Twitter'dan toplamış oldukları 8 milyon tweet üzerinde deneyler yapılmış ve % 94,7 doğruluk oranı elde edildiği vurgulanmaktadır.

Hamidian ve Diab [80], yapmış oldukları çalışmada halka açık standart veri seti kullanarak denetimli bir söylenti sınıflandırma yöntemi önermektedir. Her tweet için Tweet Latent Vector (TLV) özelliği kullanarak, boyutu 100 olan vektör temsilcisinin oluşturulduğu belirtilmektedir. Söylenti tespit işlemlerinde hassasiyetin %97,2'ye kadar çıkarıldığı belirtilmiştir.

Takahashi ve Igata [81], yapmış oldukları çalışmada tweetler kullanarak söylentiyi tespit eden bir sistem geliştirdiler. Kabul edilebilir doğrulukta bir algoritma önerdiler. Önermiş oldukları algoritma söylenti örneklerinden bir ipucu anahtar kelimesi çıkarmaktadır. Bu anahtar kelimenin başka söylentileri bulmak için de çalıştığını ifade etmektedirler.

Shirai vd. [82], yapmış oldukları çalışmada söylentilerin yayılımını gösteren bir model geliştirildiği belirtilmektedir. Bu söylenti yayılma modeli yardımıyla Twitter'dan sağlanan Japonca tweetler üzerinde söylentilerin daha önce keşfedilmemiş bazı özelliklerini bulduklarını ifade etmektedirler. Ancak çalışmalarında söylentinin nasıl tespit edileceğine değinilmemektedir.

Bai vd. [83], yapmış oldukları çalışmada, birçok söylenti algılama yönteminin yerel yapısal özelliklere odaklandığını, kaynak tweet ile yanıtları arasındaki global yapısal özelliklerin etkin bir şekilde kullanılmadığını açıklamışlardır. Global yapısal özelliklerden ve içerik bilgilerinden tam olarak faydalanmak amacıyla her bir konuşma için Kaynak-Cevap ilişkisi Grafiğini (SR-graph) önermişlerdir. SR-graph, her bir düğümün bir tweet'i ifade ettiği, düğüm özelliğinin ağırlıklı kelime vektörlerini, kenarların ise tweet'ler arasındaki etkileşimi gösterdiği bir yapıyı önermektedirler. Önermiş oldukları modelde Word2Vec modeline dayalı olarak kelime vektörlerinin eğitildiğini ve etkili sonuçlar alındığını vurgulamışlardır.

Zhang vd. [84], yapmış oldukları çalışmada söylentileri tespit etmek için mevcut sinir ağı mimarilerini uygulamak yerine, daha düşük bir seviyede grafik sinir ağlarının kullanmışlardır. Doğru söylentilere ve yanlış söylentilere verilen yanıtların farklı yapısal özellikler gösterdiğini

belirtmektedirler. Bu tür etkileşimleri etkili bir şekilde yakalamanın, bir derin öğrenme ağının daha iyi bir şekilde eğitilmesinde ve performansının arttırmasında gerekli olduğunu vurgulamışlardır. Halka açık olan Twitter veri kümeleri üzerinde yapılan deneylerde, önerilen ağın son teknoloji grafik evrişimli ağlara eşit veya hatta daha iyi bir performansa sahip olduğunu ve bunların yanında hesaplama karmaşıklığını da önemli ölçüde azalttığını belirtmektedirler.

Providel ve Mendoza [85], söylenti tespit probleminin çözümünde LSTM (Long Short Term Memory) ve GRU (Gated Recurrent Unit) mimarileri de dahil olmak üzere yapmış oldukları çalışmada on farklı derin öğrenme mimarisini kullanmışlardır. Bu mimarilerin performanslarının karşılaştırıldığı kapsamlı bir çalışma ortaya koymuşlardır. Deneysel sonuçlara bakıldığında bazı mimarilerin, bazı özel veri setleri için daha uygun olduğu görülmektedir. Ayrıca daha iyi sonuçlar için farklı mimarilerin bir arada kullanılmasının avantaj sağlayacağı belirtilmektedir.

Huang vd. [86], yapmış oldukları çalışmada söylentiyi tespit etmek için kullanıcı davranışını etkili bir şekilde yakalayan grafik evrişimli ağlara dayalı bir model önermişlerdir. Bu çalışmada kullanıcı kodlayıcı, yayılma ağacı kodlayıcı ve iki modülün çıkışını birleştiren bir entegratörden oluşan toplam üç modüle sahip olan bir model önermişlerdir.

5. ALGORİTMALAR VE VERİ KÜMELERİ

Bu bölümde aldatma ve söylenti tespitinde kullanılan metasezgisel optimizasyon algoritmaları, ve deneyler esnasında kullanılan gerçek dünya veri kümelerine ait bilgiler verilmektedir.

5.1. Optik İlham Optimizasyon

Optik, ışığın özelliklerini, davranışını, madde ile etkileşimini inceleyen; ışığı kullanan, tespit eden optik aygıt yapımında yararlanılan fiziğin alt dalıdır. Optiğin pratik uygulamaları arasında aynalar, objektifler, teleskoplar, mikroskoplar, teknoloji vb. dahil olmak üzere bir çok çeşidi vardır. Kavisli veya küresel ayna, kavisli içbükey veya dışbükey yansıtıcı yüzeye sahiptir. Çoğu kavisli ayna bir küre parçası gibi şekillendirilir. Kavisli bir ayna tarafından yansıtılan ışığın davranışı yansıma yasaları olarak bilinen düzlem aynalarda olduğu gibi aynı kanunlara tabidir [87].

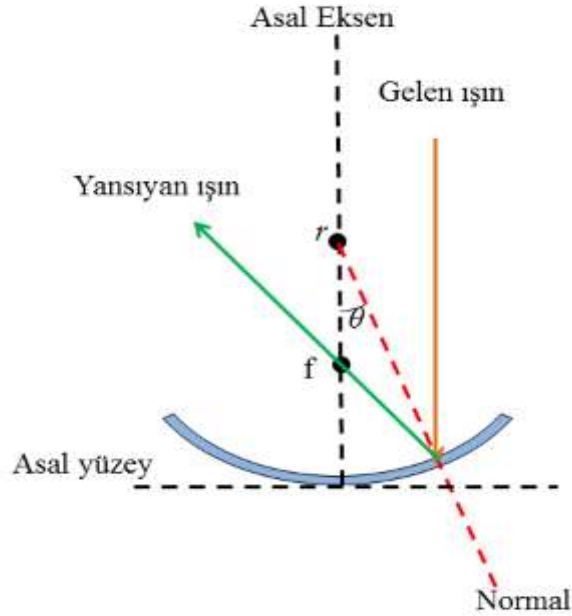
Birinci yasa: Gelen ışın, yansıyan ışın, bütün konumlar aynı düzlemedir.

İkinci yasa: Yansıyan ışın ve normal arasındaki açı gelen ışın ve normal arasındaki açıya eşittir.

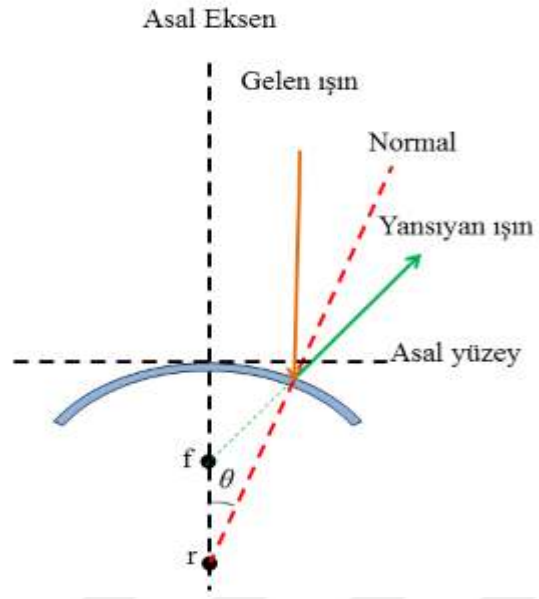
İçbükey ayna çukur şeklinde yansıtan yüzeye sahip aynadır. Odak noktasına ışığı yansıtırlar ve odaklamak için kullanılırlar. Bu aynalar nesne ve ayna arasındaki mesafeye bağlı olarak farklı resim türlerini göstermektedir. Asal eksene paralel gelen ışınları odak noktasına göndermektedir. Bu yüzden bu aynalara yakınsak ayna denir. Dışbükey ayna ise tümsek şeklinde yansıtan yüzeye sahip aynadır. Bu aynalar ışığı dışa doğru yansıtırlar ve görüntü her zaman sanaldır. Dışbükey aynada görüntü her zaman sanal, düz ve nesneden daha küçüktür. Eğer ışık ışınları aynanın aynı tarafında ise, mesafeler pozitifdir. Eğer ışık ışınları aynanın arkasında ise, mesafeler negatiftir. r (ya da f) ve q , dışbükey bir ayna için negatiftir. Nesne; tepe ve odak noktası arasında olduğu zaman q (Görüntü Pozisyonu) içbükey aynada negatiftir. Büyütme (m), küresel aynanın farklı bir özelliğidir. Görüntünün nesneye göre ne kadar büyük veya küçük olduğunu belirler. Nesne yüksekliği (HO), görüntü yüksekliği (HI) olarak Şekil 5.1’de gösterilmektedir.

Tablo 5.1’de nesne konumuna göre görüntünün tipi, boyu, konumu ve yönü hakkında bilgiler verilmiştir.

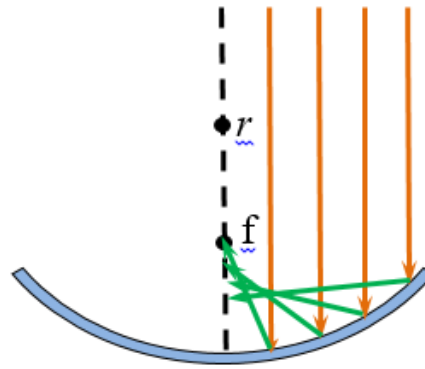
tümsek olabilmektedir. Yansıyan yapay ışığın konumuna göre bir optimum değer elde edilmektedir [88, 89]. Küresel aynalarda yansıyan ışınlarda sapmalar meydana gelebilir. Bu sapmaları düzeltmek amacı ile bazı formüller algoritmaya eklenerek sapmalar düzeltilmektedir [17]. Ana eksen, genellikle görüntüyü oluşturmak için paralel gelen ışınları kullanmaktadır. Bu durumda ışının ana eksenine açısı $\sin(\theta) = \theta$ kabul edilmektedir. Nesneden aynaya gelen ışınlar neredeyse ana eksene paraleldir. Küresel bir aynanın ana eksenine paralel olan ışınların, odak noktasına doğru tamamını yansıtamaması, *küresel sapma* (spherical aberration) olarak adlandırılır. Odak yanal sapmaları kavis yarıçapı r ile içbükey aynanın ana eksenine ışık ışınının mesafesi HO açısı kullanılarak hesaplanabilmektedir. Buna *yanal sapma* denilmektedir. Küresel sapma Şekil 5.4'te, yanal sapmalar ise Şekil 5.5'de gösterilmektedir. Bu kusurlar aynanın dış kenarlarına çarpan ışınlarda daha net görülmektedir.



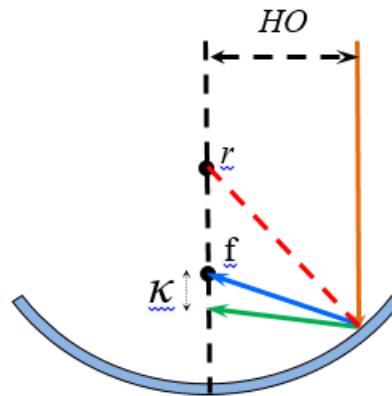
Şekil 5.2.İçbükey aynanın yansıma karakteristiği



Şekil 5.3.Dışbükey aynanın yansıma karakteristiği



Şekil 5.4.İçbükey ayna küresel sapma



Şekil 5.5.İçbükey ayna yanal sapma

Yapay bir görüntü oluşumu temsili olarak Şekil 5.6'da gösterilmektedir. Şekil 5.6'da görüldüğü üzere yapay ışık noktası ışınları aynaya çarpar ve yapay aynayla yapay bir görüntü oluşturmak için geri yansımaktadır. Yani bir dizi yapay ışık noktasının, görüntüyü yansıtan yapay dalgalı bir aynanın önünde konumlandığı varsayılmaktadır. OIO denklemlerinde kullanılan değişkenler ve anlamları Tablo 5.2'de verilmektedir. Üçgen ilişkisi ve yansıma kanunlarını kullanarak, nesne ve görüntü mesafeleri arasında nicel bir ilişki geliştirmek mümkündür. Odak uzaklığı f , r kavis yarıçapı olsun ($r = 2f$), p nesne konumu ve q ise görüntü konumudur. Yarıçap ve mesafeler sırası ile Denklem 5.1 ve Denklem 5.2'deki gibi hesaplanmaktadır [88].

$$r_{i_k}^t = m_{i_k}^t - f(\vec{F}_{i_k}^t) \quad (5.1)$$

$$p_{j,i_k}^t = s_{j,i_k}^t - f(\vec{F}_{i_k}^t) \quad (5.2)$$

Bu değerler hesaplandıktan sonar, q_{j,i_k}^t hesaplamak için Denklem 5.3 kullanılmaktadır.

$$\frac{2}{r_{i_k}^t} = \frac{1}{p_{j,i_k}^t} + \frac{1}{q_{j,i_k}^t} \Rightarrow q_{j,i_k}^t = \frac{r_{i_k}^t p_{j,i_k}^t}{2p_{j,i_k}^t - r_{i_k}^t} \quad (5.3)$$

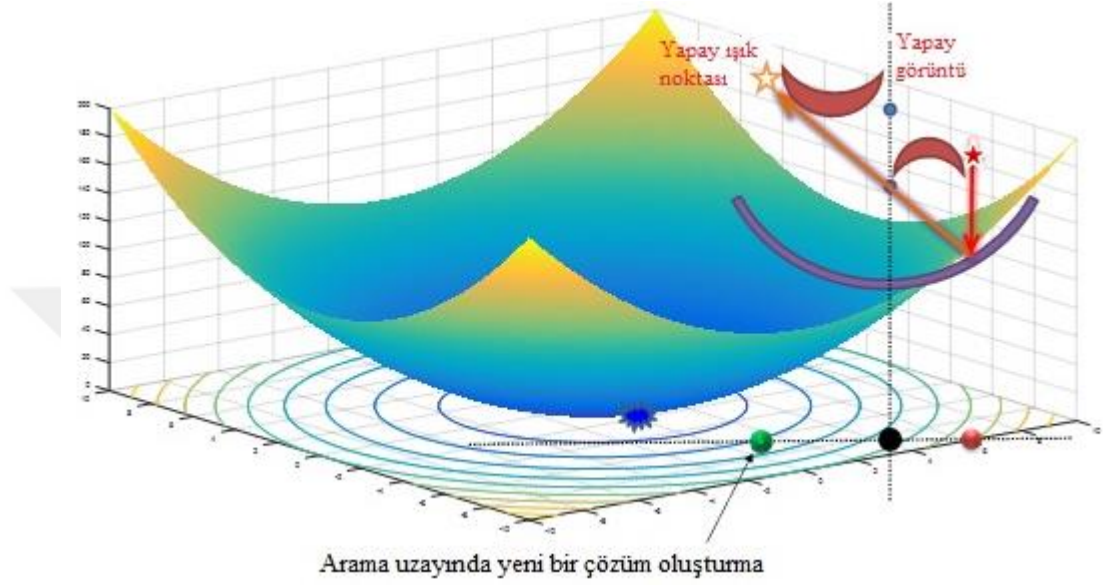
Tablo 5.2. Formüller ve değişkenlerin anlamları

Formüller ve Değişkenler	Anlamları
$\vec{O}_j^t = [O_{j1}^t \ O_{j2}^t \ \dots \ O_{jn}^t]_{1 \times n}$	t iterasyonunda n boyutlu arama uzayında yapay ışık noktası j 'nin pozisyonu [88].
$\vec{F}_{i_k}^t = [f_{i_k1}^t \ f_{i_k2}^t \ \dots \ f_{i_kn}^t]_{1 \times n}$	Ana eksen pozisyonu
$\vec{I}_j^t = [I_{j1}^t \ I_{j2}^t \ \dots \ I_{jn}^t]_{1 \times n}$	t iterasyonunda n boyutlu arama uzayında j yapay ışık noktasına ait görüntünün konumunu belirtir.
S_{j,i_k}^t	t iterasyonunda fonksiyon / amaç eksenini (amaç uzayı) üzerinde j yapay ışık noktasının (görüntüsü yapay ayna tarafından oluşturulan) konumunu belirtir [88].
p_{j,i_k}^t	t iterasyonunda fonksiyon / amaç eksenini üzerinde j yapay ışık noktasının konumu ve fonksiyon / amaç eksenini üzerindeki yapay ayna tepesinin konumu arasındaki mesafedir [88].
q_{j,i_k}^t	t iterasyonunda fonksiyon / amaç eksenini üzerinde j yapay ışık noktasının görüntüsünün konumu ve fonksiyon / amaç eksenini üzerindeki yapay ayna tepesinin konumu arasındaki mesafedir [88].
r_{j,i_k}^t	Yapay ayna eğrilik yarıçapı
$m_{i_k}^t$	Fonksiyon / amaç eksenindeki eğrilik merkezinin konumudur [88].
HO_{j,i_k}^t	t iterasyonunda yapay ana eksenden, j yapay ışık noktasının yüksekliğidir [88].
HI_{j,i_k}^t	t iterasyonunda yapay ana eksenden, j yapay ışık noktasının görüntüsünün yüksekliğidir [88].
K_{j,i_k}^t	t iterasyonunda j yapay ışık noktasının görüntüsünü yansıtan yapay ayna ile ilgili yanal sapmaların değeridir [88].

Denklem 5.4 kullanılarak yapay ışık noktası j 'nin görüntü yüksekliği Denklem 5.5 yardımıyla hesaplanmaktadır.

$$HO_{j,i_k}^t = \|\vec{o}_j^t - \vec{F}_{i_k}^t\| \quad (5.4)$$

$$HI_{j,i_k}^t = -HO_{j,i_k}^t \frac{q_{j,i_k}^t}{p_{j,i_k}^t} \quad (5.5)$$



Şekil 5.6. Işığın dışbükey ayna üzerine düşmesi durumunda yansıma kanunlarına göre gösterimi

Yapay ışık noktasının j 'nin t iterasyonundaki görüntü pozisyonu, Denklem 5.6 kullanılarak hesaplanmaktadır [88].

$$\begin{aligned} \vec{I}_{j,i_k}^t &= \vec{F}_{i_k}^t + HI_{j,i_k}^t \frac{(\vec{o}_j^t - \vec{F}_{i_k}^t)}{\|\vec{o}_j^t - \vec{F}_{i_k}^t\|} \\ &= \vec{F}_{i_k}^t - HO_{j,i_k}^t \frac{q_{j,i_k}^t}{p_{j,i_k}^t} \frac{(\vec{o}_j^t - \vec{F}_{i_k}^t)}{\|\vec{o}_j^t - \vec{F}_{i_k}^t\|} \\ &\Rightarrow \vec{I}_j^t = \vec{F}_{i_k}^t - \frac{q_{j,i_k}^t}{p_{j,i_k}^t} (\vec{o}_j^t - \vec{F}_{i_k}^t) \\ &= \vec{F}_{i_k}^t - \frac{r_{i_k}^t}{2p_{j,i_k}^t - r_{i_k}^t} (\vec{o}_j^t - \vec{F}_{i_k}^t) \end{aligned} \quad (5.6)$$

Denklem 5.7 ve Denklem 5.8 kullanılarak yeni bir sonuç görüntü pozisyonu (aday çözüm) oluşturulmaktadır [88].

$$\vec{I}_j^t = \sum_{k=1}^K w_k^t \vec{I}_{j,i_k}^t \quad (5.7)$$

$$\sum_{k=1}^K w_k^t = 1, 0 < w_k^t \leq 1 \quad (5.8)$$

t iterasyonunda j yapay ışık noktasının görüntüsünü yansıtan yapay bir ayna ile ilgili yanal sapmanın değerini hesaplamak için Denklem 5.9 kullanılmaktadır [87].

$$\vec{K}_{j,i_k}^t = \frac{(r_{i_k}^t)^2}{2\sqrt{(r_{i_k}^t)^2 - (HO_{j,i_k}^t)^2}} - \frac{|r_{i_k}^t|}{2} \quad (5.9)$$

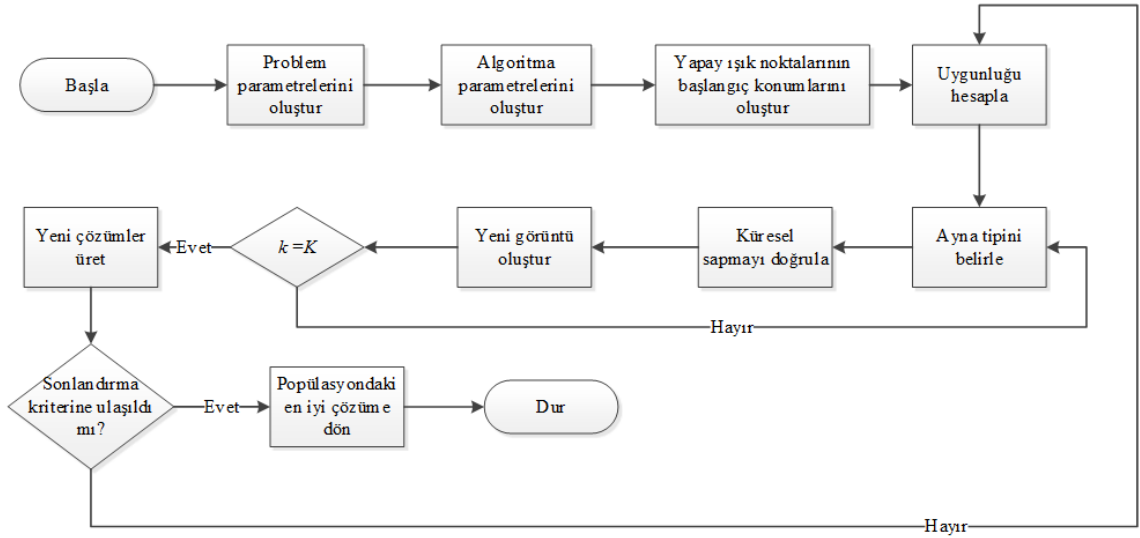
Kesik geometrik olasılık dağılımı, değişikliklerin sayısını simüle etmek için kullanılmaktadır. Q_0 genel optimizasyon sırasında gerçekleştirilen en az değişiklik sayısıdır. Hesaplamalar esnasında Q_0 Denklem 5.10 'da 1 olarak kabul edilmektedir. r rastgele bir sayı olup değeri 0 ile 1 arasındadır [88].

$$Q_i^t = \left\lceil \frac{\ln(1-(1-(1-P_c)^{n-Q_0+1})r)}{\ln(1-P_c)} \right\rceil + Q_0 - 1 \quad (5.10)$$

$$Q_i^t \in \{Q_0 | Q_0 + 1, \dots, n\}$$

$$P_c < 1, P_c \neq 0$$

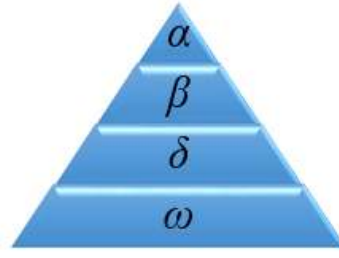
P_c kesik geometrik dağılımın başarımlı kolaylığını ifade eden bir katsayıdır. P_c için tavsiye edilen değer 0.7 dir [88]. Şekil 5.7'de OIO algoritmasının akış diyagramı gösterilmektedir.



Şekil 5.7.OIO algoritmasının akış diyagramı

5.2. Gri Kurt Optimizasyon

Canis lupus (Gri kurt) köpekçiller ailesinin üyesidir. Gri kurtlar tepe yırtıcılar olarak bilinmektedir, bu da besin zincirinin en tepesinde oldukları manasına gelmektedir. Gri kurtlar genellikle topluluk halinde yaşamayı tercih ederler. Topluluk sayıları yaklaşık 5 ila 12 arasında değişmektedir. En çok dikkat çeken yönleri ise Şekil 5.8'de gösterildiği gibi çok katı sosyal olarak baskın bir hiyerarşiye sahip olmalarıdır [90].



Şekil 5.8.Gri kurt hiyerarşik yapısı (baskınlık tepeden zemine doğru azalmaktadır)

Grubun liderleri, alfa adı verilen bir erkek ve bir dişi kurttur. Alfa, çoğu zaman uyuma mekanı, uyanma zamanı, avlanma vb. ile ilgili kararları almaktan sorumludur. Alfa'nın almış olduğu tüm kararlar gruba kabul ettirilir ve grup alfalara itaat eder. Fakat bununla birlikte, bir alfanın gruptaki başka kurtların arkasından gittiği bir tür demokratik davranış da grup içerisinde gözlemlenmiştir. Toplantı esnasında, tüm topluluk üyeleri kuyruklarını aşağıda tutarak alfalara biat ettiklerini göstermektedirler. Alfa kurda bu özelliklerinden ötürü dominant kurt da denmektedir [91]. Alfa kurtların sürü dışından bir bireyle çiftleşmesine izin verilmez. Tuhaf olan şudur ki, alfa sürünün en güçlü üyesi olmak zorunda değildir, fakat topluluğun yönetimi ve idaresi bakımından en iyisi olmalıdır. Bu gerçeklik göstermektedir ki, bir sürünün yönetim ve idaresinin sahip olduğu güçten çok daha önemlidir [90].

Gri kurt hiyerarşik yapısında ikinci önemli seviye beta kurtlardır. Beta kurtlar karar alma aşamalarında veya diğer topluluk faaliyetlerinde alfa kurtlara yardımcı olan ikinci seviye kurtlardır. Beta kurt erkek veya dişi olabilir. Alfa kurtlardan birinin sakat kalması, ölmesi ya da sürüyü önderlik edemeyecek kadar aşırı yaşlanması durumunda en güçlü olan alfa adaydır. Beta kurt alfaya saygı göstermeli ve bunun yanında diğer düşük seviye kurtları da idame edebilmelidir. Temel görevi alfa kurtlara yardım ederler ve sürünün disiplininden sorumludurlar. Beta kurtlar, alfa kurtların tüm emirlerini topluluk içerisinde destekler ve alfaya geri bildirimde bulunurlar [90].

En düşük seviyedeki gri kurt topluluğu omegalardır. Omegalar, daima sürüde baskın olan diğer bütün kurtlara itaat etmek zorundadır. Alfa ve beta kurtlar yemeklerini bitirdikten sonra yemek yemelerine izin verilir. Omega kurtlar, topluluk içinde çok önemli bir üye değilmiş gibi görünse de, topluluğun omega kaybı yaşaması durumunda tüm topluluğun iç kargaşaya sürüklendiği ve problemler yaşadığı görülmüştür. Bunun nedeni ise omegaların sürü içerisinde meydana gelen kriz anlarında tüm sürüyü yatıştırmasıdır [90].

Alfa, beta veya omega grupları içerisinde yer almayan gri kurtlara, delta adı verilmektedir. Deltalar, alfa ve beta kurtlara itaat etmek zorundadırlar. Fakat delta kurtlar, omega kurtlar üzerinde hâkimiyet kurabilmektedirler. Nöbetçiler, avcılar, izciler, bakıcılar ve yaşlılar bu kategoriye girmektedirler. İziciler, bölge sınırlarını sürekli kontrol etmek ve herhangi bir olumsuzluk ve tehlike anında topluluğu uyarmaktan sorumludur. Nöbetçiler, topluluğu korumakla ve topluluğun

güvenliğini sağlamakla sorumludurlar. Yaşlılar, önceden alfa veya beta grubu içerisinde yer almış olan tecrübeli kurtlardır. Avcıların temel görevi alfa ve beta kurtlara avlanırken veya topluluk için yiyecek besin ararken yardım etmektir. Bakıcılar ise topluluktaki hasta, zayıf ve yaralı olan kurtların bakımından sorumludur [90].

Sosyal hiyerarşilerine ek olarak, grup avı gri kurtlara ait diğer bir ilgi çekici sosyal davranıştır. Gri kurtların avlanma sürecinin ana evreleri şöyledir [92];

- ✓ Avı izlemek, yaklaşmak ve kovalamak
- ✓ Av hareketsiz kalıncaya kadar takip etmek, etrafını sarıp avı rahatsız ve taciz etmek
- ✓ En uygun anda ava saldırmak

Tüm bu evreleri içeren gri kurtlara ait avlanma davranışları Şekil 5.9’da gösterilmektedir.



Şekil 5.9. (a) Avı izleme, yaklaşma ve kovalama. (b-d) Avı takip etme, etrafını sarma ve taciz etme. (e) Uygun anı bulma ve ava saldırma

Metasezgisel bir arama ve optimizasyon algoritması olan GKO’nun sosyal hiyerarşisinden esinlenilerek geliştirilen matematiksel modelinde en iyi çözüm alfa (α) olarak isimlendirilmiştir. İkinci ve üçüncü sıradaki uygun çözümler sırasıyla beta (β) ve delta (δ) olarak isimlendirilmiştir. Geri kalan aday çözümlerin omega (ω) olduğu kabul edilmektedir. GKO’da avlanma (optimizasyon) α , β ve δ tarafından yönlendirilir. ω kurtlar bu üç kurdu izlemektedir [90].

Gri kurtların sosyal davranışları arasında olan avın çevresini sarma hareketini matematiksel modellemek amacıyla Denklem 5.11, 5.12, 5.13 ve 5.14 kullanılmaktadır.

$$\vec{D} = |\vec{C} \cdot \overline{X_p(t)} - \vec{X}(t)| \quad (5.11)$$

$$\vec{X}(t+1) = \overline{X_p(t)} - \vec{A} \cdot \vec{D} \quad (5.12)$$

Denklemlerde yineleme işlemlerini t göstermektedir. A ve C katsayı vektörlerini ifade etmektedir. X gri kurdun mevcut konum vektörünü temsil ederken, X_p ise avın konum vektörünü temsil etmektedir.

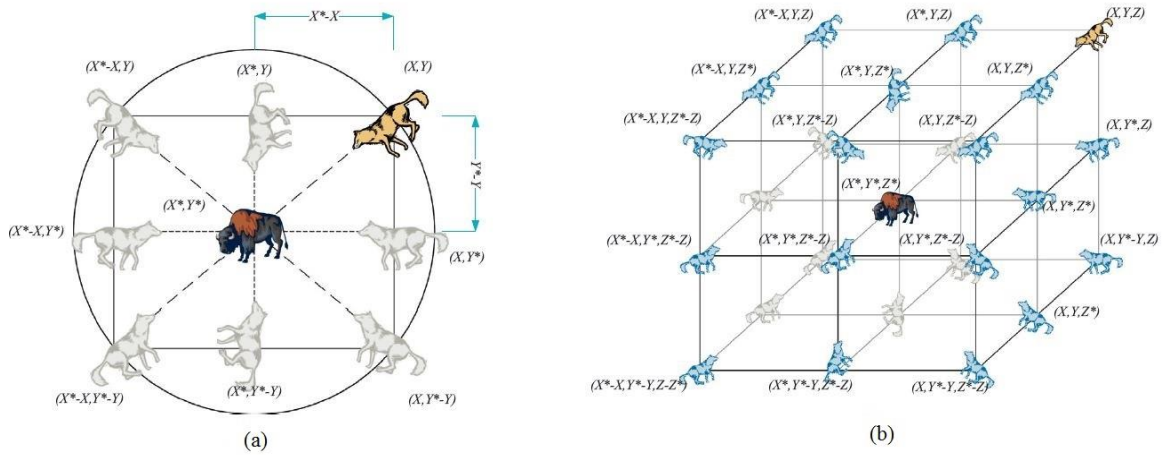
Denklem 5.13 ve Denklem 5.14'te sırasıyla A ve C vektörlerinin nasıl hesaplandığı gösterilmiştir.

$$\vec{A} = 2\vec{a} \cdot \vec{r}_1 - \vec{a} \quad (5.13)$$

$$\vec{C} = 2 \cdot \vec{r}_2 \quad (5.14)$$

Burada a 'nın bileşenleri, yinelemeler boyunca doğrusal olarak 2'den 0'a düşürülmüştür. r_1 ve r_2 ise $[0, 1]$ arasında yer alan rastgele vektörlerdir.

Denklem 5.11 ve 5.12'nin etkilerini görmek için iki boyutlu bir konum vektörü ve olası komsuların bazıları Şekil 5.10(a)'da gösterilmektedir. Görüleceği üzere (X, Y) pozisyonundaki bir birey (X^*, Y^*) pozisyonundaki avına göre kendi konumunu güncelleyebilmektedir. A ve C vektörlerinin değerleri ayarlanarak, sahip oldukları mevcut konuma göre en iyi ajanın etrafındaki farklı konumlara ulaşılabilir. Örneğin, (X^*-X, Y^*) konumuna, $A = (1, 0)$ ve $C = (1, 1)$ vektörlerinin ayarlarıyla ulaşılabilir. Şekil 5.10(b)'de gri kurdun üç boyutlu uzayda güncellenmiş konumları belirtilmektedir. Gri kurtlar r_1 ve r_2 rastgele vektörleri yardımıyla Şekil 5.10'da gösterilen noktalar arasındaki herhangi bir pozisyona erişebilmektedir. Denklem 5.11 ve 5.12 eşitlikler kullanılarak gri kurda ait herhangi bir andaki pozisyon, avın çevresinde bulunan uygun bir pozisyonla güncellenebilir. Aynı kavram, n boyutta bir arama uzayına genişletilebilir. Gri kurtlar şimdiye kadar elde edilen en iyi çözüm etrafında, hiper-küpler içinde hareket edecektir [90].



Şekil 5.10.(a) İki boyutlu konum vektörleri (b) Güncellenmiş üç boyutlu konum vektörleri

Gri kurtlar, avın pozisyonunu tanıma ve etrafını sarma kabiliyetine sahiptir. Avlanma, genellikle alfa kurtlar tarafından koordine edilmektedir. Beta ve delta kurtlarda, zaman zaman

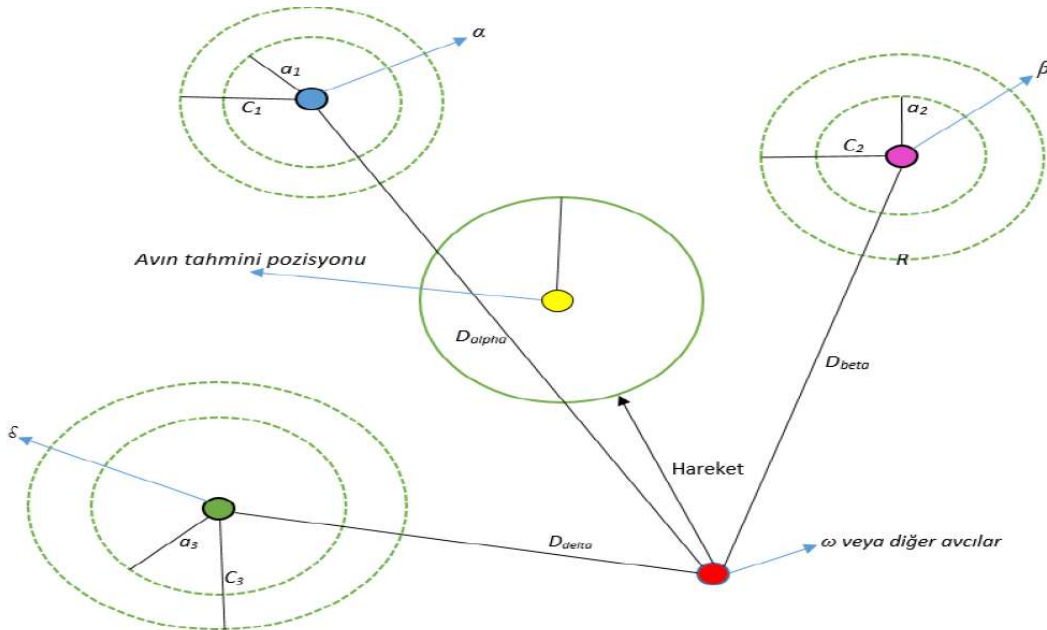
avlanmaya olayına katılabilmektedir. Fakat soyut bir arama uzayında, avın optimumum pozisyonu hakkında hiçbir bilginiz yoktur. Gri kurtların avlanma davranışını modellemek için, alfa, beta ve deltanın avın potansiyel pozisyonu hakkında daha iyi bilgiye sahip olduğunu varsayalım. Bu sebeple elde edilen en iyi ilk üç çözümü kaydedilmekte ve diğer (omegalar dâhil) arama ajanlarının pozisyonlarını, en iyi arama ajanlarının pozisyonuna göre güncellenecektir. Bu nedenle Denklem 5.15, 5.16 ve 5.17 verilmektedir [90].

$$\begin{aligned}\vec{D}_\alpha &= |\vec{C}_1 \cdot \vec{X}_\alpha - \vec{X}| \\ \vec{D}_\beta &= |\vec{C}_2 \cdot \vec{X}_\beta - \vec{X}| \\ \vec{D}_\delta &= |\vec{C}_3 \cdot \vec{X}_\delta - \vec{X}|\end{aligned}\quad (5.15)$$

$$\begin{aligned}\vec{X}_1 &= \vec{X}_\alpha - \vec{A}_1 \cdot (\vec{D}_\alpha) \\ \vec{X}_2 &= \vec{X}_\beta - \vec{A}_2 \cdot (\vec{D}_\beta) \\ \vec{X}_3 &= \vec{X}_\delta - \vec{A}_3 \cdot (\vec{D}_\delta)\end{aligned}\quad (5.16)$$

$$\vec{X}(t+1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3}\quad (5.17)$$

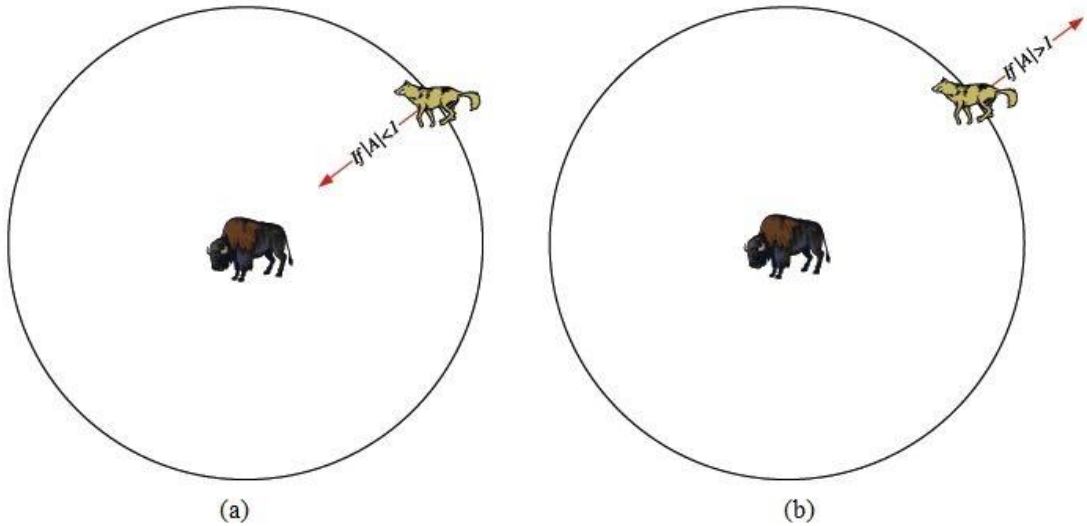
Şekil 5.11’de, bir arama ajanının pozisyonu iki boyutlu arama alanındaki alfa, beta ve deltaya göre nasıl güncellediği verilmektedir. Nihai pozisyonun, arama alanındaki alfa, beta ve delta pozisyonları tarafından tanımlanan bir daire içinde rastgele bir yerde olacağı gözlemlenebilir. Yani, alfa, beta ve delta, avın pozisyonunu tahmin eder ve diğer gri kurtlar da pozisyonlarını avın etrafında rastgele günceller [90].



Şekil 5.11.GKO’da pozisyon güncelleme

Av hareketsiz kaldığı anda daha öncede bahsedildiği üzere, gri kurtlar ava saldırarak avlanma olayını tamamlar. Ava yaklaşmayı matematiksel olarak modellemek için a 'nın değeri azaltılmıştır. A 'nın değişim aralığının da a ile birlikte azaldığına dikkat edilmelidir. Bir diğer ifadeyle, A ; iterasyon sürecinde 2'den 0'a düşen $[-2a, 2a]$ aralığında rastgele bir değerdir. A 'nın rastgele değerleri $[-1, 1]$ olduğunda, arama ajanının bir sonraki pozisyonu, mevcut pozisyonu ile avın pozisyonu arasında herhangi bir noktada olabilir. Şekil 5.12 (a)'da, $|A| < 1$ durumunun gri kurtları ava saldırmaya teşvik ettiği gösterilmektedir [90].

GKO algoritması, şimdiye kadar önerilen operatörlerle, arama ajanlarının alfa, beta ve deltanın pozisyonuna göre konumlarını güncellemelerine ve ava saldırımlarına imkan tanır. GKO algoritması bu operatörlerle yerel çözüme takılma eğilimindedir. Önerilen avı sarma mekanizmasının belli ölçüde keşif özelliği gösterdiği doğrudur, fakat GKO'nun keşfi vurgulamak için daha fazla operatöre ihtiyacı olduğu da bir gerçektir [90].



Şekil 5.12. Ava saldırı ile av arayışının karşılaştırılması

Gri kurtlar çoğunlukla alfa, beta ve deltanın pozisyonuna göre arama yaparlar. Av aramak için birbirlerinden uzaklaşıp, avlarına saldırmak için bir birlerine yaklaşırlar. Birbirinden uzaklaşmayı matematiksel olarak modellemek için, arama ajanını avdan uzaklaştırmaya zorlamak amacıyla A ile birlikte 1'den büyük veya -1'den küçük rassal değerler kullanılmıştır. Bu durum, keşfi öne çıkaracak ve GKO algoritmasının global olarak arama yapmasına olanak sağlayacaktır. $|A| > 1$ durumunun, gri kurtları olası başka daha iyi avlar bulmak için avdan ayrılmaya zorladığı Şekil 5.12 (b)'de gösterilmektedir [90].

GKO'nun keşfi destekleyen diğer bir bileşeni de C dir. Denklem 5.14'ten de görüldüğü gibi, C vektörü $[0, 2]$ aralığında rassal değerler içermektedir. Bu bileşen, stokastik şekilde avın Denklem 5.11'deki mesafeyi belirleme üzerindeki etkisini öne çıkarmak ($C > 1$) veya bunun önemini azaltmak ($C < 1$) amacıyla av için rassal değerler sağlar. Bu durum, optimizasyon boyunca GKO'

nun, keşif ve yerel optimumdan kaçınma lehinde daha rassal bir davranış göstermesine yardımcı olur. Burada, C 'nin A 'dan farklı olarak, doğrusal olarak azaltılmadığını belirtmekte fayda vardır. Keşfi vurgulamak amacıyla, kasıtlı olarak C 'nin sadece ilk iterasyonları sırasında değil, son iterasyonları sırasında da daima rassal değerler vermesi istenmiştir. Bu bileşen, özellikle son iterasyonlarda, algoritmanın yerel optimuma takılmasının önüne geçmesinde çok fayda sağlamaktadır [90].

C vektörü ayrıca doğadaki avlara yaklaşmanın önündeki engellerin etkisi olarak da düşünülebilir. Genel olarak ifade edilirse, doğadaki engeller gri kurtların avlanma yollarında ortaya çıkmaktadır. Aslında bu engeller gri kurtların avlarına hızlı ve kolay bir şekilde yaklaşmalarını engeller. Bu da, aslında C vektörünün yaptığı iştir. Gri kurdun pozisyonuna göre, a rassal bir değer verebilir ve a ulaşmayı kurtlar için zorlaştırır. Ya da a avı uzaklaştırır veya tam tersini yapar [90].

GKO algoritmasında arama süreci rassal olarak gri kurt popülasyonu oluşturmakla başlar. İterasyonlar boyunca, a , b ve d gri kurtları avın muhtemel pozisyonunu tahmin eder. Her aday çözüm, avla olan mesafesini günceller. Keşif ve sömürmeyi vurgulamak için a parametresi 2'den 0'a düşürülmüştür. Aday çözümler, $|A| > 1$ iken avdan uzaklaşma ve $|A| < 1$ iken a doğru yaklaşmaktadır. Son olarak, GKO algoritması bir sonlandırma şartını yerine getirilmesiyle sona erdirilir [90].

GKO algoritmasının gerçek dünya problemlerini teorik olarak nasıl çözdüğünü daha iyi anlamak için bazı noktalara dikkat etmek gerekmektedir. Bunlar:

- ✓ Önerilen sosyal hiyerarşi, GKO algoritmasının iterasyon aşamaları boyunca elde edilen en iyi çözümleri saklamasına yardımcı olur.
- ✓ Önerilen çevreleme mekanizması, çözümlerin çevresinde hiper-küre olarak daha üst boyutlara genişletilebilecek daire şeklinde bir alanı tanımlar.
- ✓ A ve C rassal parametreleri, farklı rassal yarıçapları olan hiper-kürelere sahip aday çözümlere yardımcı olur.
- ✓ Avlanma yöntemi, aday çözümlerin avın muhtemel pozisyonunu tespit etmesine olanak sağlar.
- ✓ Keşif ve sömürü işlemleri a ve A 'nın uyarlanabilir değerleri sayesinde güvence altındadır.
- ✓ a ve A parametrelerinin uyarlanabilir oluşu, GKO algoritmasının keşif ve sömürme olayları arasında sorunsuz geçiş yapmasına imkan sağlamaktadır.
- ✓ A 'nın azalmasıyla birlikte, iterasyonların yarısı keşfe ($|A| \geq 1$) ve öteki yarısı ise sömürmeye ($|A| < 1$) aittir.
- ✓ GKO'nun, (a ve C olmak üzere) ayarlanacak sadece iki ana parametresi vardır [90].

GKO algoritmasının sözde kodu Şekil 5.13'te gösterilmektedir.

```

Gri kurtların başlangıç popülasyonunun ayarlanması  $X_i (i = 1, 2, \dots, n)$ 
 $a$ ,  $A$  ve  $C$  parametrelerine ilk değerlerin verilmesi
Her bir arama ajanının uygunluk değerinin hesaplanması
 $X_a$  = en iyi arama ajanı
 $X_\beta$  = en iyi ikinci arama ajanı
 $X_\delta$  = en iyi üçüncü arama ajanı
while ( $t < \text{Maksimum yineleme sayısı}$ )
for her bir arama ajanı
    Mevcut arama ajanının konumunu Denklem 5.17'ye göre güncelle
end for
     $a$ ,  $A$  ve  $C$  parametrelerini güncelle
    Tüm arama ajanlarının uygunluğunu hesapla
     $X_a$ ,  $X_\beta$  ve  $X_\delta$  'yi güncelle
     $t = t + 1$ 
end while
return  $X_a$ 

```

Şekil 5.13.GKO algoritmasının sözde kodu

5.3. Denetimli Makine Öğrenme Algoritmaları

Denetimli makine öğrenme yöntemleri, sisteme eğitim veri seti ve test veri setinin yüklenmesi, veri setinde her bir veri için gerekli etiketlemenin yapılması ve bu sayede girdi veri seti ile çıktı veri seti arasında ilişki kurulması mantığına dayanır. Temel amaç sonuçları bilinen veri setinden yapılan sınıflandırma dikkate alınarak, sonuçları bilinmeyen veri setine dair etkili tahminler yapabilmektir [93]. Bu bölümde, tez çalışması esnasında kullanılan denetimli makine öğrenme algoritmaları hakkında temel bilgiler verilmiştir.

Naive Bayes: Naive Bayes (NB) algoritması belirli bir veri setindeki değerlerin sıklığını ve kombinasyonunu sayarak olasılık kümesini hesaplayan seçici bir sınıflandırıcıdır. NB sınıflandırıcısı bayes teoremi ve toplam olasılık teoremine dayanmaktadır. Sınıfların etiketleri, belirli bir örneğin olasılığı ile tahmin edilir. Gerekli olan sınıf etiketini tahmin etmek için sadece eğitim verisine ihtiyaç vardır [94].

Random Forest: Random Forest (RF), her sınıflandırıcının giriş vektöründen bağımsız olarak örneklenen rassal bir vektör kullanılarak üretildiği ve her ağacın bir girdi vektörünü sınıflandırmak için en popüler sınıfa bir birim oy kullandığı ağaç sınıflandırıcılarının birleşiminden oluşmaktadır [95]. RF algoritması, sınıflandırmada çok iyi performans sergilediği için toplu öğrenme yöntemlerinde sıklıkla tercih edilmektedir. RF sınıflandırıcısı, hem hızlı hem de yüksek doğruluk sağladığından ötürü toplu öğrenmede iyi olarak bilinen Hızlandırma [96] ve Torbalama

[97] metotlarına göre avantaj sağlamaktadır. RF, ağaç tipi sınıflandırıcılar topluluğuna üyedir. Torbalama metodunun gelişmiş bir şekli olarak bilinmektedir. RF algoritması, özellikle hızlandırma metoduna göre, eğitim aşamasında çok daha hızlıdır. Yeterliliği ve sağlamış olduğu doğruluk oranları nedeniyle çok kullanışlı bir sınıflandırıcıdır.

Decision Tree: Decision Tree (DT), çeşitli alanlarda tercihlerin, risklerin hedeflerin modellenmesinde ve birbirine bağlı, içinde belirsizlik içeren verilerle sonuçlar çıkarmakta ve karar almakta kullanılan yaygın ve başarılı bir yöntemdir. DT üç temel yapıdan oluşmaktadır. Bunlar iç karar düğümleri, dallar ve yapraklardır. İç karar düğümlerinde sınıflandırılmış veri elemanları vardır. Dallar karar seviyelerini belirtmektedir. Karar düğümünden, dalların sınıflandırmalarına göre diğer düğümlere ulaşılmaktadır. Yapraklar ise ulaşılan sonuçları temsil etmektedir [98].

Logistic Regression: Logistic Regression (LR), bir bağımlı ikili değişken ile bir veya daha fazla nominal, sıra, aralık veya oran düzeyinde bağımsız değişken arasındaki ilişkiyi açıklamak için kullanılır. Mevcut bulunan tüm regresyon analizleri gibi LR de öngörücü bir analizdir. İstatistikte, lojistik model olarak da bilinmektedir. Geçti-kaldı, kazan-kaybet, canlı-ölü veya sağlıklı-hasta gibi var olan belirli bir sınıf veya olayın olasılığını modellemek için kullanılmaktadır. Bu, bir fotoğrafın yılan, fare, kuş vb. içerip içermediğinin belirlenmesi gibi çeşitli olay sınıflarını modellemek için de genişletilebilir [99].

Support Vector Machine: Support Vector Machine (SVM), eğitim verilerindeki herhangi bir noktadan en uzak olan iki sınıf arasında bir karar sınırı bulan vektör uzayı tabanlı bir makine öğrenme metodudur. Bir SVM, verileri en uygun olarak iki kategoriye ayırarak n-boyutlu bir hiper düzlem oluşturan yöntemdir. SVM modelleri yapay sinir ağlarıyla yakından ilişkili olup, sigmoid kernel fonksiyonu kullanan SVM, iki katmanlı, ileri beslemeli bir yapay sinir ağına sahiptir [100].

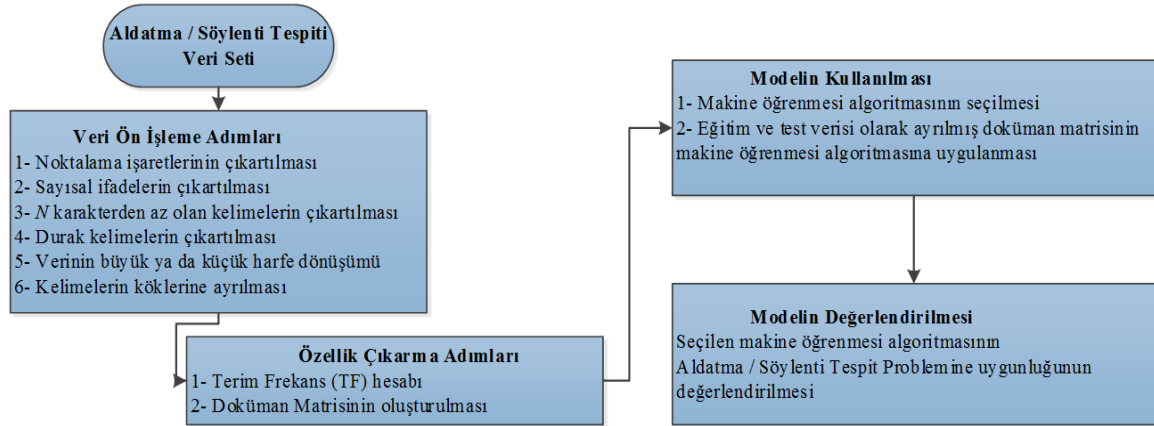
Neural Networks: Neural Network (NN), en basit bir şekilde insan beyninin çalışma şeklini taklit eden bir yapıya sahiptir. NN metodolojisi veriden öğrenebilme, genelleme yapabilme, sınırsız sayıda değişkenle çalışabilme vb. birçok önemli özelliğe sahiptir. NN, girdi ve çıktı değişkenleri arasındaki herhangi bir ön bilgiye gereksinim duymadan doğrusal ve doğrusal olmayan modellemeyi sağlayabilmektedir. NN'nin bir öngörü aracı olarak kullanılmasına ilişkin literatürde yapılmış çok sayıda yayın olmasına rağmen, NN'nin performansını etkileyen anahtar faktörlerin neler olduğu konusunda kesin bir yargı mevcut değildir. Belirtilen bu nedenlere ilave olarak, eğitim algoritması, veri setinin düzenlenmesi, kestirim dönemi uzunluğu faktörlerinin de NN'nin başarımı üzerinde etkili olduğu belirtilmektedir [101].

K-Nearest Neighbor: K-Nearest Neighbor (KNN) algoritması örnek veri noktasının bulunduğu sınıfın ve en yakın komşunun, k değerine göre belirlendiği bir sınıflandırma metodudur. KNN algoritmasında öğrenme işlemi eğitim veri kümesinde bulunan verilere dayalı olarak gerçekleştirilmektedir. Yeni karşılaşılan bir örnek, eğitim veri kümesinde yer alan örnekler ile arasındaki benzerliğe göre sınıflandırılmaktadır [102]. KNN algoritmasında, eğitim veri kümesinde

yer alan örnekler n -boyutlu sayısal nitelikler ile gösterilmektedir. Her bir örnek n -boyutlu uzayda bir noktayı temsil edecek biçimde tüm eğitim örnekleri n -boyutlu bir örnek uzayında tutulmaktadır. Bilinmeyen bir örnek ile karşı karşıya kalındığında, eğitim veri kümesinden ilgili örneğe en yakın k tane örnek belirlenerek yeni örneğin sınıf etiketi, k en yakın komşusunun sınıf etiketlerinin çoğunluk oylamasına göre atanır [103].

5.4. Veri Ön İşleme Aşamaları

Yapay zekâya dayalı öğrenen sistemlerde kullanılan verilerin temsil biçimi, öğrenme sisteminin üreteceği sonuçların doğruluğunu çok büyük ölçüde etkilemektedir. Verinin tipi ister görüntü ister metin ya da ses verisi olsun, uygulanacak yönteme uygun olan bir gösterime dönüştürülmesi gerekmektedir. SM uygulamalarında paylaşılan metinsel veriler büyük oranda yapılandırılmamış durumdadır. Bu sebeple SM platformlarından elde edilen ham veriler ön işlem adı verilen bir dizi aşamadan geçirilerek yapılandırılmaktadır. Bu tez kapsamında kullanılan veri setleri üzerinde yapılan ön işlem aşamaları Şekil 5.14'te gösterilmektedir.



Şekil 5.14. Ham veri üzerinde gerçekleştirilen ön işlem aşamaları

Aldatma / Söylenti veri setlerinden okunan ham veri bünyesinde bulunan noktalama işaretleri, satır sonu karakterleri metinsel veriden çıkartılır. Ardından sayısal ifade içeren tüm terimler veriden silinir. Daha sonra önceden belirlenmiş olan N adet, karakterden az olan terimler veriden silinir. Edat, bağlaç ve zarflar durak kelimelerdir, bu kelimeler veriden çıkartılır. Ardından tüm veri ya büyük harfe ya da küçük harfe dönüştürülür. Bu adım sayesinde veri tek tipe dönüştürülmüştür. Bazı kelimeler sonlarına ek almaktadırlar. Kelimelerin almış oldukları bu ekler kelimelerden çıkartılarak kök haline gelmesi sağlanmıştır. Bu sayede metin içinde geçen kök kelimelerin geçme sıklığı ölçülecektir [104].

Yüksek boyut metin madenciliği uygulamalarında karşılaşılabilecek olan en büyük sorunlardan biridir. Bu sebeple önermekte olunan sistemin yüksek başarımlı göstermesi için boyut indirgeme

amaçlı gereksiz özellikler veriden silinmesi gerekmektedir. Veri seti içerisindeki her doküman için terimler ağırlıklandırılır ve her bir doküman terimlerin ağırlık vektörüne dönüştürülür. Bu gösterim Vector Space Model (VSM) olarak adlandırılır [105]. Metin dokümanları VSM sayesinde bir özellik vektörüne dönüşmektedir. VSM’de her bir kelime, kelimenin doküman içerisindeki ağırlığını gösteren bir değer ile ifade edilmektedir. Bu tez kapsamında gerçekleştirilen çalışmalarda terimleri ağırlıklandırmak için Term Frequency (TF) metodu kullanılmıştır. TF bir kelimenin doküman içerisinde geçme sıklığını göstermektedir. TF’nin matematiksel ifadesi Denklem 5.18’de gösterilmektedir.

$$TF_{ij} = \frac{n_{ij}}{|d_i|} \quad (5.18)$$

d_i , i . veri kümesi içerisinde geçen tüm kelime köklerinin sayısıdır. n_{ij} ise j . kelime kökünün toplanan i . veri kümesi içerisinde kaç defa geçtiğinin sayısını belirtir [106]. Her bir dokümandaki TF değerleri hesaplandıktan sonra, terim ağırlıklarına göre Denklem 5.19’te gösterilen $m \times n$ boyutundaki Document Term Matrix (DTM) oluşturulur. Oluşturulan DTM’de her bir satır dokümanları, her bir sütun ise terimleri göstermektedir. Hücreler dokümandaki terimlerin ağırlıklarını belirtir.

$$D_1 \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{m1} & \cdots & A_{mn} \end{bmatrix} \quad (5.19)$$

5.5. Aldatma ve Söylenti Tespiti Veri Kümeleri

Bu tez çalışmasında OIO, GKO ve üç adet yeni KBOIO algoritmalarının aldatma ve söylenti tespiti problemleri üzerindeki başarımlarını gözlemlemek için halka açık dört farklı veri seti kullanılmıştır.

5.5.1. Aldatma Tespiti Veri Setleri

Deceptive Opinion Spam Corpus v1.4: Bu veri seti Amerika Birleşik Devletlerinin Chicago ilinde bulunan 20 otel hakkında doğru ve aldatıcı otel incelemelerinden oluşmaktadır. Bu veri seti 800’ü gerçek ve 800’ü aldatıcı yorum olmak üzere toplam 1600 adet yorumdan oluşmaktadır. Veri setini oluşturan bu yorumlar TripAdvisor, Mechanical Turk, Expedia, Hotels.com, Orbitz, Priceline ve Yelp gibi SM platformlarından sağlanmıştır [40, 47]. Şekil 5.15’te Deceptive Opinion Spam Corpus v1.4 veri setinden bir kesit gösterilmektedir.

S Text	S State
We really enjoyed our stay at the Palmer House Hilton. We were in Chica...	Deceptive
The Palmer House Hilton was recommended to me by a friend who had vis...	Deceptive
My stay at The Palmer House Hilton was wonderful. I was attracted to thi...	Deceptive
Amazing! I was swept away when I walked into the hotel it was gorgeous...	Deceptive
A hotel made for royalty. Decked out to the top with luxuries you could o...	Deceptive
I stayed here with my family when we were in Chicago on a family vacatio...	Deceptive
The Palmer House Hilton Hotel has to be the nicest hotel I have stayed at ...	Deceptive
My husband and I stayed at the Palmer House Hilton in Chicago for our a...	Deceptive
If anyone is traveling to Chicago, I would strongly recommend the Palmer...	Deceptive
My husband and I stayed at The Palmer House Hilton last month to celebr...	Deceptive
My husband and I snagged a great deal on a weekend trip to Chicago, an...	Deceptive
The Palmer House Hilton is located in the heart of the theatre, financial a...	Deceptive
The Palmer House Hilton in Chicago is by far the best experience I have e...	Deceptive
The Palmer House Hilton hotel has it all for the business and leisure travel...	Deceptive
Very beautiful hotel. The historic features of this hotel make it absolutely ...	Deceptive
We stay at Hilton for 4 nights last march. It was a pleasant stay. We got ...	Truthful
Thirty years ago, we had a tiny room and indifferent service. This time, t...	Truthful
We loved the hotel. When I see other posts about it being shabby I can't ...	Truthful
I took the wife and kids to Chicago for one last fling before school started...	Truthful
in the windy city ,this is a very good place.amazing rooms and service.ev...	Truthful
I only stayed out with my boyfriend for one night, however enjoyed my s...	Truthful
We booked our stay through Priceline.com and got a great deal. We were...	Truthful
Great place, great room, great location. Even though there was a big me...	Truthful
My family and I have just had a two week holiday in Chicago, and we sta...	Truthful
This is a stunning hotel in an excellent location in the greatest of US cities...	Truthful

Şekil 5.15.Deceptive Opinion Spam Corpus v1.4 veri setinden bir kesit

Real-Life Trial Data: Bu veri setinin toplanmasındaki asıl amaç mahkeme duruşmalarında hem sözlü hem de sözlü olmayan davranışları, aldatmayla ilgili olarak analiz etmek içindir. Bu veri setinde 61 adet aldatıcı, 60 adet de gerçek olan toplam 121 adet videodan oluşmaktadır [107]. Bu video içeriklerinin metne dönüştürülmüş durumları deneyler esnasında aldatıcı içerik tespitinde kullanılmıştır. Şekil 5.16’da Real-Life Trial Data veri setinden bir kesit gösterilmektedir.

S Text	S State
100 percent. Absolutely. 100 percent. 100 percent. Totall...	Deceptive
um person in the congregation was involved in activities lik...	Deceptive
The objective is as Jesus said, if you have a hundred shee...	Deceptive
So the objective is to help that individual. Whether it's to e...	Deceptive
The elder would mate with the individual based on person'...	Deceptive
My statement is this: I've never been alone with Candace ...	Deceptive
What do you mean think? I just told you. Never. I've neve...	Deceptive
Never have I been alone with Candace Conte.	Deceptive
Not at all. Oh, that's her word agains mine isn't. Who do y...	Deceptive
Since then you become the biggest (inaudible) in all of this?	Deceptive
Well, I would like to think so, but that's not for me to juge...	Deceptive
Reforming the criminal justice system matters to me becau...	Truthful
they need to believe that the law treats everyone um ...	Truthful
All of us, who have represented people for years in the sy...	Truthful
You know, this person is improperly convicted, you need t...	Truthful
But, its no one else's job to do it, other than the innocence...	Truthful
But yes, I was there. Yep, I stayed. Uh ... Yep, prob -- ye...	Truthful
I do. I was, uh... in the office, and I got a call ... uh, may...	Truthful
I don't remember, somewhere between eight and nine, fro...	Truthful
someone was just on the side of the house, and he was s...	Truthful
someone was, like, sleeping on the side of the house, and ...	Truthful
oh and then he had to go really quickly to talk to the police...	Truthful
he had put Ian in the car as he does every day around th...	Truthful
as he walked cause he sort of exits the garage a little bit t...	Truthful
So he walked over to that spot, it's near our air conditio...	Truthful

Şekil 5.16.Real-Life Trial Data veri setinden bir kesit

5.5.2. Söylenti Tespiti Veri Setleri

S1: Bu veri seti 60 adet söylenti içeren, 51 adet söylenti içermeyen toplam 111 adet olay ile ilgili tweetler barındırmaktadır [108]. Deneyler esnasında toplam 18571 adet tweet kullanılmıştır. Şekil 5.17’de S1 veri setinden bir kesit gösterilmektedir.

S	Text	S	State
	We are following an effort to get 1 Million people in Kenya...		Rumor
	The most common test for HIV is the antibody test called E...		Rumor
	LISTEN yall niggas stay away from ketchup in restaurants n...		Rumor
	he had a heart bypass and had 2 have a blood transfusion...		Rumor
	Getting my blood work checked Yeah I aint scared Anyone...		Rumor
	Funny wen HIV amp Thalassemia tests r still voluntary RT ...		Rumor
	is HIV testing today 427 from 5p7p Its Free Anonymous a...		Rumor
	Im going to sneak HIV tainted blood into your blood bank		Rumor
	good ol Gareth Lewis on the scandal of contaminated bloo...		Rumor
	We have HIV positive people putting their blood in a resta...		Rumor
	Next iPhone app maybe Mobile phone turned into a portab...		Rumor
	Got a call back from the nurse The source patients blood is...		Rumor
	Air France says at least 1 Canadian on board plane that fl...		Nonrumor
	NEWS Photos French Plane Lost in Ocean Storm An Air Fra...		Nonrumor
	The hunt for clues to the fate of Air France jet lost in Atla...		Nonrumor
	Air France Jet Missing After Hitting Storm Off Brazil Voice ...		Nonrumor
	Air France jet missing with 228 people over Atlantic after r...		Nonrumor
	Air France flight from Brazil to Paris Tropical storm caused ...		Nonrumor
	Brazilian rescuers continue the search for an Air France fli...		Nonrumor
	Air France Jet Missing After Hitting Storm Off Brazil Voice ...		Nonrumor
	Air France Jet Missing After Hitting Storm Off Brazil Voice ...		Nonrumor
	NEWS Photos French Plane Lost in Ocean Storm An Air Fra...		Nonrumor
	Brazilian aircraft searching for an Air France jet which wen...		Nonrumor
	Breaking Air France plane from Brazil to Paris is missing Los...		Nonrumor
	French plane lost in Atlantic storm An Air France airliner ca...		Nonrumor
	RT weatha CNN Breaking News Brazilian air force says deb...		Nonrumor

Şekil 5.17. S1 veri setinden bir kesit

Germanwings Crash: Veri seti Arkaitz Zubiaga vd. [109] tarafından 2016 yılında yayınlanmıştır. Bu veri seti 228 adet söylenti içeren, 231 adet söylenti içermeyen toplam 459 Twitter verilerini içeren klasörden oluşmaktadır. Bu klasörler ‘kaynak-tweet’ ve kaynak tweet’e yanıt veren bir dizi ‘tepki-tweet’ten oluşmaktadır. Şekil 5.18’de Germanwings Crash veri setinden bir kesit gösterilmektedir.

S Text	S State
German media reporting Andreas Lubitz had a serious ...	Rumor
Andreas Lubitz was deemed unsuitable for flight duties...	Rumor
NEWS Germanwings co-pilot Andreas Lubitz had seriou...	Rumor
BREAKING German media report Germanwings co-pilot ...	Rumor
Germanwings Co Pilot New Muslim Convert Islamic Stat...	Rumor
Just had an emergency meeting at Ansar, Lawyers ab...	Rumor
Report: Germanwings Co-Pilot Treated for Depression....	Rumor
Germanwings co-pilot was Muslim convert. European g...	Rumor
What is it about Islam that it attracts all the psychopat...	Rumor
GERMAN NEWS: Co-Pilot of Germanwings Airbus Was ...	Rumor
Germanwings CoPilot was a recent Muslim convert as w...	Rumor
GERMAN NEWS REPORT: Co-Pilot of Germanwings Airb...	Rumor
If the Germanwings copilot Lubitz was Muslim,it wud ha...	Rumor
BREAKING: German Media Site Says Germanwings Co-...	Rumor
Wake Up America Co-pilot was a Muslim Convert. But n...	Rumor
Route map now being shared by. 4U9525	Nonrumor
Pray for 4U9525. oh God! Now I am scared!	Nonrumor
Airbus A320 4U9525 crash: Flight tracking data here. g...	Nonrumor
Flightradar24 has this as the plane's last position. 6800...	Nonrumor
IMAGE: Flightradar24 has this as the plane's last posi...	Nonrumor
Last position Germanwings A320 via flightradar24. 800...	Nonrumor
Germanwings A320 D-AIPX is one of the oldest A320 w...	Nonrumor
Germanwings A320 D-AIPX is one of the oldest A320 a...	Nonrumor
Last few positions of the germanwings flight plotted on...	Nonrumor
JUST IN: Airbus plane operated by Lufthansa's German...	Nonrumor
BREAKING Last few positions of the germanwings flight...	Nonrumor

Şekil 5.18.Germanwings Crash veri setinden bir kesit

5.6. Performans Değerlendirme Kriterleri

Tez kapsamında aldatma ve söylenti problemlerin çözümünde kullanılan metasezgisel optimizasyon algoritmaları ve denetimli makine öğrenme yöntemlerinin, performanslarını değerlendirmek ve etkinliklerini ölçmek için dört farklı performans değerlendirme kriterinden faydalanılmıştır. Aldatma ve söylenti tespiti problemleri için Tablo 5.3'te gösterilen karmaşıklık matrisi oluşturulmaktadır.

Tablo 5.3. Aldatma ve söylenti tespiti için karmaşıklık matrisi

		Gerçek Sınıf	
		Pozitif	Negatif
Tahmin Edilen Sınıf	Pozitif	Doğru Pozitif (DP)	Yanlış Pozitif (YP)
	Negatif	Yanlış Negatif (YN)	Doğru Negatif (DN)

DP, YP, DN ve YN aldatma ve söylenti problemi için aşağıdaki gibi ifade edilmiştir.

- ✓ DP: Tahmin edilen aldatma ve söylenti gerçekte de bir aldatma ve söylenti ise, tahmin DP'dir.

- ✓ YP: Tahmin edilen aldatma ve söylenti aslında gerçek bir olay ise, tahmin YP'dir.
- ✓ DN: Tahmin edilen gerçek olay aslında da gerçek bir olay ise, tahmin DN'dir.
- ✓ YN: Tahmin edilen gerçek olay aslında bir aldatma ve söylenti ise, tahmin YN'dir.

Hem önerilen yöntemlerin hem de diğer metasezgisel optimizasyon algoritmalarının ve denetimli makine öğrenme yöntemlerinin performanslarını karşılaştırmak için Tablo 5.4'te verilen değerlendirme kriterleri kullanılmıştır.

Tablo 5.4. Performans değerlendirme ölçütleri

Performans Değerlendirme Kriterleri	Formüller
Doğruluk	$\frac{DP + DN}{DP + DN + YP + YN}$
Kesinlik	$\frac{DP}{DP + YP}$
Hassasiyet	$\frac{DP}{YN + DP}$
F-ölçütü	$\frac{2 * \left(\frac{DP}{DP + YP}\right) * \left(\frac{DP}{DP + YN}\right)}{\left(\frac{DP}{DP + YP}\right) + \left(\frac{DP}{DP + YN}\right)}$

Aldatma ve söylenti tespiti problemi için doğruluk değeri, tüm veri kümesi içerisinde doğru tahmin edilen durumların oranını belirtir. Kesinlik, aldatma ve söylenti olarak tahmin edilen durumlar içerisindeki, doğru tahmin edilen aldatma ve söylenti oranını gösterir. Hassasiyet, doğru tahmin edilen aldatma ve söylentinin, tüm aldatma ve söylenti durumlarına oranını ifade eder. F-ölçütü ise, hassasiyet ve kesinlik değerlerinin harmonik ortalamasıdır.

5.7. Friedman Testi

Bu tez kapsamında OIO, GKO ve önerilen üç yeni Kaos Tabanlı OIO algoritmalarının başarımlarını test etmek için parametrik olmayan Friedman testi kullanılmıştır. Bu test birbiri ile ilişkili ikiden fazla değişkene ait dağılımları karşılaştırılarak, dağılımlar arasında farkın anlamlı olup olmadığını ölçmek için kullanılır. Yani, örneklemin tekrarlanan ölçümleri arasında fark olup olmadığı test edilir.

Friedman testi için tekrarlanan ölçümler, rastgele örnekleme ve her bir gruptaki örneklemelerin kendi içerisinde bağımsız olduğu kabulü yapılmıştır. Friedman testi için X_R^2 analizi Denklem 5.20’de gösterilmektedir.

$$X_R^2 = \frac{12}{nk(k+1)} [\sum_{j=1}^k R_j^2] - [3n(k+1)] \quad (5.20)$$

Denklem 5.20’de n satır sayısını, k grup sayısını, R_j her bir gruba ilişkin sıra sayıları toplamalarını, X_R^2 ise istatistiği seçilen yanılma düzeyinde $k-1$ serbestlik dereceli ki-kare dağılımını ifade etmektedir. Friedman testinde gerçekleştirilen adımlar şöyledir:

- ✓ H_0 ve H_1 hipotezleri belirlenir. Tez çalışmasında aldatma ve söylenti tespit problemlerine uygulanan OIO, GKO ve üç yeni kaos tabanlı OIO algoritmaları için hipotezler şöyledir:
 H_0 : Deneylerde kullanılan beş metasezgisel optimizasyon algoritmasından alınan sonuçlar arasındaki fark istatistiksel olarak anlamlı değildir.
 H_1 : Deneylerde kullanılan beş metasezgisel optimizasyon algoritmasından alınan sonuçlar arasındaki fark istatistiksel olarak anlamlıdır.
- ✓ Alpha (α) değeri belirlenir. $\alpha = 0,05$ olarak belirlenmiştir.
- ✓ Serbestlik derecesi (df) hesaplanır. $df = k - 1$, $k = 5$ (beş adet metasezgisel optimizasyon algoritması) olduğundan ötürü, çalışmamızda $df = 4$ olarak alınmıştır.
- ✓ Karar kuralı belirlenir. α ve df değerleri baz alınarak ki-kare dağılım tablosunda X_R^2 değeri 9,488 olduğu tespit edilmiştir.
- ✓ Deneysel istatistik işlemleri gerçekleştirilir
- ✓ Bulunan sonuçlar değerlendirilir. Eğer X_R^2 değeri 9,49’dan büyükse fark anlamlıdır H_0 reddedilir H_1 kabul edilir.

5.8. Kutu Grafiği

Metasezgisel optimizasyon algoritmalarından elde edilen uygunluk fonksiyonu değerlerinin kutu grafiğinde gösterimi için öncelikle veri kümesi için aşağıdaki değerler hesaplanmalıdır.

$X_{\min}$ = Uygunluk fonksiyonunun en küçük değerini ifade eder.

$Q1$ = Birinci çeyrek (%25)

X_{med} = Ortanca değer (medyan)

$Q3$ = Üçüncü çeyrek (%75)

X_{maks} =Uygunluk fonksiyonunun en büyük değerini ifade eder.

Bu değerleri hesaplamak için içerisinde metasezgisel optimizasyon algoritmalarından elde edilen uygunluk değerlerini içeren tablo küçükten büyüğe sıralanır ve beşli özetleme tablosu çıkarılmakta ve bu veriler ışığında kutu grafiği çizdirilmektedir.

6. ÖNERİLEN YÖNTEMLER

6.1. Kaos Tabanlı Optik İlham Optimizasyon Algoritmaları (KTOIO)

OIO algoritmasının yakınsama performansı temelde bazı parametreler için rastgele sayılar kullanan stokastik yapısına bağlıdır. Ancak, diğer metasezgisel algoritmalarda olduğu gibi, OIO içinde yüksek başarıyı garanti eden, özel bir sayı üreticine dayalı, herhangi bir analitik sonuç yoktur [110].

Son zamanlarda, rastgele sayıların yerini kaotik sayılar aldı ve daha iyi sonuçlar elde etmeyi sağladı. Güvenli iletişim, doğal fenomen modelleme [111], doğrusal olmayan devreler [112], imge işleme [113] verilebilir. Ayrıca evrimsel algoritmalarının performanslarını arttırmak için de kullanılmıştır [114, 115]. Kaotik sayı dizilerinin kullanılması teorik olarak bunların tahmin edilemezliği, yayılmış spektrumlu karakteristiği ve ergodik özelliklerinden dolayı artmıştır [110].

Tezin bu kısmında OIO'nun bazı parametrelerinin belirlenmesinde rasgele tabanlı bir seçim söz konusu olduğunda farklı kaotik sistemler rasgele sayı dizilerinin yerine kullanılmış ve üç farklı OIO algoritması önerilmiştir. Böylece OIO'nun global yakınsama özelliğinin artırılması ve lokal çözümde takılıp kalması engellenmeye çalışılmıştır. Örneğin w ve r değerleri yakınsamayı etkileyen faktörlerdir. Ancak bu parametreler faz uzayında algoritmanın ergodik özelliğini garanti edemezler çünkü bunlar OIO'da rasgeledir.

6.1.1. Kaotik Haritalar

Kompleks fenomenleri modellemede, örneklemede, sayısal analizde, karar vermede ve özellikle bu tez konusunu da içeren meta sezgisel optimizasyon algoritmalarında uzun periyotlu rasgele sayı dizileri çok önemli bir yer tutmaktadır. Üretilen sayılar için fazla depolama alanı kullanılmamalı ve istenen bir doğruluğa ulaşmak için fazla zamana gereksinim duyulmamalıdır. Bu şekilde üretilen sayılar bir uygulama için yeterince “rasgele” olurken başka bir uygulama için yeterince rasgele olmayabilir [110].

Kaos periyodik olmayan, yakınsamayan ve sınırlı olan, doğrusal olmayan dinamik sistemlerde bulunan deterministik, rasgele benzeri bir süreçtir. Ayrıca başlangıç şartları ve parametrelerine oldukça bağlıdır [116]. Kaosun yapısı görünürde rasgele ve tahmin edilemezdir. Ayrıca kendi içerisinde bir düzene sahiptir. Hatta çoğu kez düzen içinde düzensizlik ya da düzensizlik içinde düzen olarak da tanımlanmaktadır. Matematiksel olarak, basit deterministik dinamik bir sistemin rasgeleliğidir ve kaotik sistem rasgelelik kaynağı olarak düşünülebilir [110].

Kaotik bir harita ayrık zamanlı dinamik bir sistemdir. Böyle bir sistem Denklem 6.1'de gösterilmektedir.

$$x_{k+1} = f(x_k), \quad 0 < x_k < 1, \quad k = 0, 1, 2, \dots \quad (6.1)$$

Kaotik sayı dizilerinin üretilmelerinin ve depolanmalarının kolay ve hızlı olduğu ispatlanmıştır. Sadece birkaç fonksiyon (kaotik harita) ve birkaç parametre (başlangıç şartı) çok uzun diziler için bile yeterlidir. Ayrıca, çok fazla sayıda farklı sayı dizisi basitçe başlangıç şartı değiştirilerek çok kolay bir şekilde üretilir. Bu sayı dizilerinin bir özelliği de deterministik olmaları ve tekrar üretilmeleridir [110].

Kaos Tabanlı Optik İlham Optimizasyon (KTOIO) algoritmasında ergodik, düzensiz ve rassallık özelliklerine sahip kaotik haritalar kullanılarak diğer OIO yöntemlerine oranla daha kolayca lokal çözümden kaçabilmeyi sağlamak amaçlanmıştır. Bu şekilde global yakınsamanın artırılması hedeflenmiştir. Rasgele sayılar özel bir kaotik harita bir adım ilerletilerek üretilmektedir. Yani, OIO’da ilk iterasyondan itibaren rasgele sayı üretimine ihtiyaç duyulduğunda seçilen kaotik harita seçilen bir başlangıç noktasından başlanarak birer adım ilerletilir. OIO parametreleri için kaotik sayı üreten haritalar olarak Gauss, Çember, Lojistik, Sinüzoidal ve Çadır haritaları kullanılmıştır.

Gauss Harita: Literatürde test amaçlı kullanılan Gauss haritası [117] Denklem 6.2 ve 6.3’teki gibi temsil edilmektedir. Bu kaotik harita 0 ile 1 arasında sayı üretmektedir.

$$X_{n+1} = \begin{cases} 0 & \text{if } X_n = 0 \\ \left(\frac{1}{X_n}\right) \bmod(1) & \text{if } 0 < X_n \leq 1 \end{cases} \quad (6.2)$$

$$\left(\frac{1}{X_n}\right) \bmod(1) = \frac{1}{X_n} - \left\lfloor \frac{1}{X_n} \right\rfloor \quad (6.3)$$

Çember Harita: Çember haritaya ait matematiksel ifade Denklem 6.4’te gösterilmektedir. $a = 0.5$ ve $b = 0.2$ olduğu durumlarda çember harita 0 ile 1 arasında değer üretmektedir [118].

$$x_{n+1} = x_n + b - (a/2\pi) \sin(2\pi x_n) \bmod(1) \quad (6.4)$$

Lojistik Harita: En kolay ve çok kullanılan kaotik haritalardandır [119]. Bu harita, kaotik davranış gösteren biyolojik popülasyonların doğrusal olmayan dinamiklerinde ortaya çıkmıştır. Bu harita 0 ile 1 arasında sayı üretmektedir. Ancak başlangıç X_0 değeri $\{0.0, 0.25, 0.5, 0.75, 1.0\}$ değerlerini içermemelidir. Ayrıca deneyler esnasında $a = 4$ alınmıştır. Lojistik haritaya ait matematiksel ifade Denklem 6.5’te gösterilmektedir.

$$x_{n+1} = ax_n(1 - x_n) \quad (6.5)$$

Sinüzoidal Harita: Kullanılan bir diğer rasgele sayı üretici, sinüzoidal yineleyici [120] olarak adlandırılmaktadır. Kaotik haritaya ait matematiksel ifade Denklem 6.6'daki gibi gösterilmektedir. $a=2.3$ ve $X_0=0.7$ seçilerek oluşturulan sadeleştirilmiş ifadeye Denklem 6.7'de gösterilmektedir. Sinüzoidal harita 0 ile 1 arasında sayı üretmektedir.

$$x_{n+1} = ax_n^2 \sin(\pi x_n) \quad (6.6)$$

$$x_{n+1} = \sin(\pi x_n) \quad (6.7)$$

Çadır Harita: Lojistik haritaya çok benzemektedir [110]. Çadır harita 0 ile 1 aralığında sayılar üretmektedir. Kaotik haritaya ait olan matematiksel ifade Denklem 6.8'de gösterilmektedir.

$$x_{n+1} = \begin{cases} x_n/0.7 & , x_n < 0.7 \\ 10/3 x_n (1 - x_n) & , \text{otherwise} \end{cases} \quad (6.8)$$

6.1.2. Optik İlham Optimizasyon Algoritmasına Kaotik Harita Entegrasyonu

Rastgele sayılar, seçilen kaotik haritanın bir adım ileriye taşınmasıyla oluşturulur. Yani, standart OIO' da ilk yinelemeden rastgele bir sayıya ihtiyaç duyulduğunda, kaotik harita seçilen başlangıç değerinden başlayarak adım adım ilerletilir [120]. Bu tezde önerilen kaotik OIO algoritmaları şu şekilde açıklanmıştır:

KTOIO1: Denklem 5.8'deki w parametresi, seçilen kaotik harita ile yer değiştirir. Bu parametre 0 ile 1 arasındadır [120].

KTOIO2: Denklem 5.10'deki r parametresi, seçilen kaotik harita ile yer değiştirir. Bu parametre 0 ile 1 arasındadır [120].

KTOIO3: Denklem 5.8'deki w parametresi ve Denklem 5.10'daki r parametresi seçilen kaotik haritalar ile yer değiştirir. Bu parametreler 0 ile 1 arasındadır [120].

Parametreleri ayarlamak için standart OIO'ya farklı kaotik haritalar entegre edilmiştir. Bu işlem, rastgele bir sayı için kaotik bir harita kullanılarak gerçekleştirilmiştir. Üç yeni kaotik OIO (KTOIO) yöntemi ve bu algoritmalar için beş harita, kalite test fonksiyonlarında ve gerçek dünya mühendislik problemlerinde kullanılmıştır. Deneyler esnasında P_c katsayısı 0,7'dir; ALP olarak 10 ve NFE olarak 500 belirlenmiştir [120].

Kalite Test Fonksiyonları Kullanarak KTOIO Algoritmalarının Performans Analizi: Bu makalede, KTOIO'ların performansı farklı özelliklere sahip kıyaslama problemlerinde test edilmiştir. Tablo 6.1'de kalite test fonksiyonlarına ait matematiksel ifade, tanımlı aralık ve global minimum değerleri gösterilmektedir. Kalite test fonksiyonlarının tümünde problem boyutu 2 olarak belirlenmiştir [120].

Tablo 6.1. Kalite test fonksiyonlarına ait matematiksel ifade, tanımlı aralık ve global minimum

Fonksiyon adı	Matematiksel ifade	Tanımlı aralık	Global minimum
Griewank	$f(x) = 1 + \frac{1}{4000} \sum_{i=1}^n x_i^2 - \prod_{i=1}^n \cos \frac{x_i}{\sqrt{i}}$	[-300, 300]	0
Ackley	$f(x) = -20 \exp \left(-0.2 \sqrt{0.5 \left(\sum_{i=1}^n x_i^2 \right)} \right) - \exp \left(0.5 \left(\sum_{i=1}^n \cos(2\pi x_i) \right) \right) + e + 20$	[-32, 32]	0
Rastrigin	$f(x) = An + \sum_{i=1}^n \left((x_i^2 - A \cos(2\pi x_i)) \right)$	[-5.12, 5.12]	0
Sphere	$f(x) = \sum_{i=1}^n (x_i^2)$	[-100, 100]	0
Bohachevsky 1	$f(x) = x_1^2 + 2x_2^2 - 0.3 \cos(3\pi x_1) - 0.4 \cos(4\pi x_2) + 0.7$	[-100, 100]	0
Rosenbrock	$f(x) = \sum_{i=1}^{n/2} (100(x_{2i-1}^2 - x_{2i})^2 + (x_{2i-1} - 1)^2)$	[-10, 10]	0

OIO ve KTOIO 30 defa bağımsız bir şekilde çalıştırılmış ve amaç fonksiyonu ($f(x)$) kaydedilmiştir. Kaydedilen sonuçlardan istatistiksel analiz yapılmıştır. Kısıtsız kalite test fonksiyonlarından Rosenbrock [121], Ackley [122], Sphere [123], Griewank [124], Bohachevsky-1 [125] ve Rastrigin [121] deneyler esnasında kullanılmıştır. Seçilen kısıtsız kalite test fonksiyonları için standart OIO'dan elde edilen sonuçlar Tablo 6.2'de gösterilmektedir.

Tablo 6.2. Standart OIO'nun kısıtsız kalite test fonksiyonları üzerindeki sonuçları

F	Rastrigin	Sphere	Ackley	Griewank	Rosenbrock	Bohachevsky1
O	1.66E-01	2.32E-06	1.27E-03	2.64E-02	1.65E+00	2.40E+01
Min	9.83E-11	1.10E-09	1.37E-05	1.79E-04	3.64E-04	3.09E-03
Mak	1.08E+00	4.58E-05	8.38E-03	9.72E-02	1.25E+01	9.90E+01
SS	3.61E-01	8.56E-06	2.04E-03	2.23E-02	0.31E+01	2.39E+01

Rastrigin fonksiyonunda; Çember, Gauss, Lojistik haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar Tablo 6.3'te gösterilirken, Çadır ve Sinüzoidal haritalar

kullanarak elde edilen sonuçlar ise Tablo 6.4'te gösterilmektedir. En iyi değerler koyu olarak tablolarda belirtilmektedir. F1 Rastrigin fonksiyonunu ifade etmektedir. En iyi değer (Min), Ortalama (O), En kötü değer (Mak) ve Standart Sapma (SS) tablolarda belirtilmektedir.

Tablo 6.3 ve Tablo 6.4 dikkatle incelendiğinde, KTOIO3 algoritması Rastrigin fonksiyonu için diğer KTOIO algoritmaları arasından en iyisi olduğu görülmektedir. KTOIO1 ve KTOIO2 algoritmaları bu fonksiyon için eşit başarımlar göstermektedir.

Tablo 6.3.Rastrigin fonksiyonunda Çember, Gauss, Lojistik haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar

	Gauss Harita $f(x)$			Çember Harita $f(x)$			Lojistik Harita $f(x)$		
F1	KTOIO1	KTOIO2	KTOIO3	KTOIO1	KTOIO2	KTOIO3	KTOIO1	KTOIO2	KTOIO3
O	8.35E-02	1.35E-01	9.34E-02	9.81E-02	1.81E-01	4.20E-02	7.43E-02	4.70E-02	8.16E-02
Min	1.64E-07	1.93E-08	7.07E-10	1.14E-07	1.13E-09	1.42E-08	6.02E-08	4.50E-09	2.39E-09
Mak	1.01E+00	9.95E-01	1.01E+00	9.95E-01	9.95E-01	9.95E-01	9.95E-01	8.50E-01	1.06E+00
SS	0.253791169	0.316792123	0.260098923	0.292631076	0.345819120	0.182518357	0.238907774	0.164943999	0.263878122

Tablo 6.4.Rastrigin fonksiyonunda Çadır, Sinüzoidal haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar

	Sinüzoidal Harita $f(x)$			Çadır Harita $f(x)$		
F1	KTOIO1	KTOIO2	KTOIO3	KTOIO1	KTOIO2	KTOIO3
O	1.19E-01	1.70E-01	1.08E-01	1.53E-01	1.70E-01	3.03E-02
Min	1.35E-08	1.13E-08	9.55E-09	9.02E-09	2.82E-07	2.36E-08
Mak	9.95E-01	9.95E-01	1.04E+00	1.27E+00	9.95E-01	6.65E-01
SS	0.275720907	0.338202979	0.306358557	0.369237512	0.375641969	0.123181363

Sphere fonksiyonunda; Çember, Gauss, Lojistik haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar Tablo 6.5'te gösterilirken, Çadır ve Sinüzoidal haritalar kullanarak elde edilen sonuçlar ise Tablo 6.6'da gösterilmektedir. En iyi değerler koyu olarak tablolarda belirtilmektedir. F2 Sphere fonksiyonunu ifade etmektedir. Tablo 6.5 ve Tablo 6.6 dikkatle incelendiğinde, KTOIO1 algoritması Sphere fonksiyonu için diğer KTOIO algoritmaları arasından en iyisi olduğu görülmektedir.

Tablo 6.5.Sphere fonksiyonunda Çember, Gauss, Lojistik haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar

	Gauss Harita $f(x)$			Çember Harita $f(x)$			Lojistik Harita $f(x)$		
F2	KTOIO1	KTOIO2	KTOIO3	KTOIO1	KTOIO2	KTOIO3	KTOIO1	KTOIO2	KTOIO3
O	1.15E-08	1.61E-02	2.06E-07	6.37E-09	6.21E-07	8.18E-07	9.85E-09	6.30E-07	8.42E-07
Min	7.37E-12	9.71E-05	1.30E-10	2.10E-12	7.26E-12	9.76E-11	8.76E-13	5.81E-12	1.32E-11
Mak	1.04E-07	5.81E-02	1.61E-06	8.02E-08	8.65E-06	1.28E-05	6.89E-08	5.31E-06	1.11E-05
SS	2.86E-08	0.011640302	3.55E-07	1.68E-08	1.74E-06	2.39E-06	1.84E-08	1.25E-06	2.32E-06

Tablo 6.6.Sphere fonksiyonunda Çadır, Sinüzoidal haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar

	Sinüzoidal Harita $f(x)$			Çadır Harita $f(x)$		
F2	KTOIO1	KTOIO2	KTOIO3	KTOIO1	KTOIO2	KTOIO3
O	9.50E-09	7.12E-07	5.23E-07	7.06E-09	8.07E-06	2.96E-07
Min	7.59E-12	8.04E-11	1.49E-09	3.33E-13	4.17E-11	3.20E-12
Mak	1.86E-07	5.26E-06	4.10E-06	8.16E-08	2.31E-04	3.94E-06
SS	3.36E-08	1.20E-06	1.10E-06	1.70E-08	4.21E-06	7.67E-07

Rosenbrock fonksiyonunda; Çember, Gauss, Lojistik haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar Tablo 6.7’de gösterilirken, Çadır ve Sinüzoidal haritalar kullanarak elde edilen sonuçlar ise Tablo 6.8’de gösterilmektedir. En iyi değerler koyu olarak tablolarda belirtilmektedir. F3 Rosenbrock fonksiyonunu ifade etmektedir. Tablo 6.7 ve Tablo 6.8 dikkatle incelendiğinde, KTOIO1 algoritması Rosenbrock fonksiyonu için diğer KTOIO algoritmaları arasından en iyisi olduğu görülmektedir.

Tablo 6.7.Rosenbrock fonksiyonunda Çember, Gauss, Lojistik haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar

	Gauss Harita $f(x)$			Çember Harita $f(x)$			Lojistik Harita $f(x)$		
F3	KTOIO1	KTOIO2	KTOIO3	KTOIO1	KTOIO2	KTOIO3	KTOIO1	KTOIO2	KTOIO3
O	4.19E-02	7.79E-02	1.38E-01	5.39E-02	1.05E-01	1.40E-01	5.92E-04	1.31E-01	1.31E-01
Min	2.42E-05	3.98E-03	3.41E-04	1.23E-04	3.60E-04	2.84E-04	8.23E-11	6.28E-04	7.28E-03
Mak	2.74E-01	2.81E-01	2.03E+00	2.31E-01	2.83E-01	1.40E+00	7.40E-03	2.91E-01	7.93E-01
SS	0.054285034	0.055069742	0.361631534	0.067870182	0.08012676	0.248295431	0.001763595	0.090743032	0.140575382

Tablo 6.8.Rosenbrock fonksiyonunda Çadır, Sinüzoidal haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar

	Sinüzoidal Harita $f(x)$			Çadır Harita $f(x)$		
F3	KTOIO1	KTOIO2	KTOIO3	KTOIO1	KTOIO2	KTOIO3
O	6.10E-02	8.17E-02	1.60E-01	3.79E-02	1.06E-01	9.16E-02
Min	4.43E-05	5.86E-04	2.48E-02	9.28E-05	1.94E-03	6.86E-04
Mak	4.93E-01	3.00E-01	1.55E+00	1.60E-01	4.54E-01	2.95E-01
SS	0.095894129	0.084982288	0.270532139	0.038819138	0.088817621	0.084821669

Griewank fonksiyonunda; Çember, Gauss, Lojistik haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar Tablo 6.9’de gösterilirken, Çadır ve Sinüzoidal haritalar kullanarak elde edilen sonuçlar ise Tablo 6.10’da gösterilmektedir. En iyi değerler koyu olarak tablolarda belirtilmektedir. F4 Griewank fonksiyonunu ifade etmektedir. Tablo 6.9 ve Tablo 6.10 dikkatle incelendiğinde, KTOIO1 algoritması Griewank fonksiyonu için diğer KTOIO algoritmaları arasından en iyisi olduğu görülmektedir. KTOIO3 algoritması KTOIO2 algoritmasından daha iyi sonuçlar verdiği görülmektedir.

Tablo 6.9.Griewank fonksiyonunda Çember, Gauss, Lojistik haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar

	Gauss Harita $f(x)$			Çember Harita $f(x)$			Lojistik Harita $f(x)$		
F4	KTOIO1	KTOIO2	KTOIO3	KTOIO1	KTOIO2	KTOIO3	KTOIO1	KTOIO2	KTOIO3
O	8.13E-04	2.58E-02	2.51E-02	3.01E-06	2.95E-02	2.80E-02	4.29E-04	2.54E-02	1.78E-02
Min	1.44E-11	1.62E-05	1.24E-03	4.53E-10	8.26E-04	4.57E-03	9.22E-11	1.81E-03	1.97E-04
Mak	7.42E-03	7.64E-02	7.10E-02	4.81E-05	1.10E-01	7.53E-02	7.40E-03	8.42E-02	5.23E-02
SS	0.00226459	0.018746693	0.01999317	9.75E-07	0.024459537	0.018637783	0.001516381	0.020801332	0.012025223

Tablo 6.10.Griewank fonksiyonunda Çadır, Sinüzoidal haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar

	Sinüzoidal Harita $f(x)$			Çadır Harita $f(x)$		
F4	KTOIO1	KTOIO2	KTOIO3	KTOIO1	KTOIO2	KTOIO3
O	5.40E-02	2.58E-02	2.37E-02	4.43E-05	3.33E-02	2.06E-02
Min	3.40E-05	6.73E-04	7.31E-05	5.30E-11	9.61E-04	2.79E-04
Mak	2.56E-01	6.68E-02	7.78E-02	8.84E-04	1.08E-01	6.66E-02
SS	0.0615282	0.018156422	0.019582672	1.72E-04	0.027867036	0.016268026

Ackley fonksiyonunda; Çember, Gauss, Lojistik haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar Tablo 6.11’de gösterilirken, Çadır ve Sinüzoidal haritalar kullanarak elde edilen sonuçlar ise Tablo 6.12’de gösterilmektedir. En iyi değerler koyu olarak tablolarda belirtilmektedir. F5 Ackley fonksiyonunu ifade etmektedir. Tablo 6.11 ve Tablo 6.12 dikkatle incelendiğinde, KTOIO1 algoritması Ackley fonksiyonu için diğer KTOIO algoritmaları arasından en iyisi olduğu görülmektedir.

Tablo 6.11.Ackley fonksiyonunda Çember, Gauss, Lojistik haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar

	Gauss Harita $f(x)$			Çember Harita $f(x)$			Lojistik Harita $f(x)$		
F5	KTOIO1	KTOIO2	KTOIO3	KTOIO1	KTOIO2	KTOIO3	KTOIO1	KTOIO2	KTOIO3
O	1.89E-04	1.37E-03	1.01E-03	6.37E-05	1.79E-03	2.31E-01	1.42E-04	2.30E-01	7.86E-04
Min	7.70E-06	2.74E-05	2.27E-05	2.10E-08	4.98E-05	1.81E-05	8.21E-07	1.22E-05	5.88E-05
Mak	1.31E-03	7.20E-03	6.37E-03	8.02E-04	1.84E-02	6.88E+00	1.15E-03	6.88E+00	3.57E-03
SS	0.00025057	0.001738479	0.00133851	1.68E-04	0.003415649	1.256264646	0.000220689	1.256436985	0.000819099

Tablo 6.12.Ackley fonksiyonunda Çadır, Sinüzoidal haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar

	Sinüzoidal Harita $f(x)$			Çadır Harita $f(x)$		
F5	KTOIO1	KTOIO2	KTOIO3	KTOIO1	KTOIO2	KTOIO3
O	7.81E-05	2.30E-01	8.81E-04	1.09E-04	1.06E-03	1.35E-03
Min	1.64E-06	3.66E-05	2.30E-05	2.48E-06	5.32E-06	6.95E-06
Mak	3.15E-04	6.88E+00	7.64E-03	6.42E-04	1.19E-02	1.19E-02
SS	8.98E-05	1.2564606	0.00148744	0.151035064	0.00219232	0.002621254

Bohachevsky-1 fonksiyonunda; Çember, Gauss, Lojistik haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar Tablo 6.13’de gösterilirken, Çadır ve Sinüzoidal haritalar kullanarak elde edilen sonuçlar ise Tablo 6.14’de gösterilmektedir. En iyi değerler koyu olarak tablolarda belirtilmektedir. F6 Bohachevsky-1 fonksiyonunu ifade etmektedir. Tablo 6.13 ve Tablo 6.14 dikkatle incelendiğinde, KTOIO2 algoritması Bohachevsky-1 fonksiyonu için diğer KTOIO algoritmaları arasından en iyisi olduğu görülmektedir.

Tablo 6.13.Bohachevsky-1 fonksiyonunda Çember, Gauss, Lojistik haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar

Gauss Harita $f(x)$				Çember Harita $f(x)$			Lojistik Harita $f(x)$		
F6	KTOIO1	KTOIO2	KTOIO3	KTOIO1	KTOIO2	KTOIO3	KTOIO1	KTOIO2	KTOIO3
O	1.09E-04	1.38E-05	1.12E-05	9.49E-05	8.70E-06	9.10E-05	1.38E-02	5.11E-05	7.42E-05
Min	3.46E-06	1.36E-09	1.89E-09	5.48E-08	3.65E-10	3.00E-09	4.86E-11	3.00E-10	5.35E-08
Mak	6.33E-04	4.13E-04	8.87E-05	3.07E-04	1.14E-04	1.21E-03	4.13E-01	9.07E-04	1.03E-03
SS	1.41E-04	7.53E-05	2.02E-05	8.85E-05	0.022673548	0.000283724	0.075386886	1.70E-04	0.000202602

Tablo 6.14.Bohachevsky-1 fonksiyonunda Çadır, Sinüzoidal haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar

Sinüzoidal Harita $f(x)$				Çadır Harita $f(x)$		
F6	KTOIO1	KTOIO2	KTOIO3	KTOIO1	KTOIO2	KTOIO3
O	1.89E-04	1.84E-03	6.34E-05	1.57E-02	3.78E-04	8.10E-04
Min	3.37E-11	1.85E-08	2.89E-10	4.26E-11	4.27E-08	1.63E-09
Mak	5.57E-03	5.38E-02	1.65E-03	4.70E-01	1.04E-02	2.28E-02
SS	0.00101562	0.009819881	0.000300929	0.085784643	0.001893205	0.004155575

Testlerde, KTOIO1 algoritmasının genel olarak, keşif ve sömürü dengesi açısından küresel arama yeteneğini geliştirerek çözüm kalitesini artırdığı gözlemlenmiştir. Bu durum Tablo 6.3, 6.4, 6.5, 6.6, 6.7, 6.8, 6.9, 6.10, 6.11, 6.12, 6.13 ve 6.14’ten de görülebilmektedir.

Tablo 6.15’te KTOIO algoritmalarından elde edilen *Min* ve *O* değerlerinin toplam sayısı gösterilmektedir. Rastgele sayılar yerine *w* için farklı kaotik haritalar kullanan KTOIO1 algoritması, Tablo 6.15’te gösterildiği gibi 43 kez en iyi *Min* ve *O* değerlerine ulaştığı için hızlı yakınsama özelliğini artırarak yerel minimumlara takılmamaktadır. KTOIO1 algoritmasından elde edilen en iyi *Min* değerlerinin toplam sayısı 23’tür buda Tablo 6.16’da gösterilmektedir.

Tablo 6.15.KTOIO algoritmalarında en iyi değerlerin toplamı (*Min* + *O*)

	KTOIO1	KTOIO2	KTOIO3	OIO
<i>Min+O</i>	43	5	7	5

Tablo 6.16.KTOIO algoritmalarında en iyi değerlerin toplamı (*Min*)

	KTOIO1	KTOIO2	KTOIO3	OIO
<i>Min</i>	23	2	0	5

Gerçek Dünya Mühendislik Tasarım Problemleri Kullanarak KTOIO Algoritmalarının Performans Analizi:

Gerçek dünyada var olan sistemler genellikle kısıtlı problemlerdir. Bu özellikteki problemler çözülmenden önce eşitlik ve eşitsizlik kısıtlamaları tanımlanmalıdır. Genel olarak Denklem 6.9 ve 6.10 eşitsizlik, Denklem 6.11 ise eşitlik kısıtlarını ifade etmektedir.

$$g_k(x) \leq 0, \quad k = 1, 2, 3, \dots, q \quad (6.9)$$

$$alt_i \leq x_i \leq üst_i \quad i = 1, 2, 3, \dots, D \quad (6.10)$$

$$h_j(x) = 0, \quad j = 1, 2, 3, \dots, m \quad (6.11)$$

$g_k(x)$ eşitsizlik kısıtı, $h_j(x)$ eşitlik kısıtı ve $f(x)$ ise amaç fonksiyonudur. $x = (x_1, x_2, \dots, x_D)$ boyutu D olan bir vektördür. alt_i ve $üst_i$ sırasıyla x_i 'nin alt ve üst sınırlarını göstermektedir. q eşitsizlik kısıtlarının sayısı ve m eşitlik kısıtlarının sayısını ifade eder. Deneylerimiz esnasında kısıtlama fonksiyonları için ceza fonksiyonu yaklaşımlarını takip ettik. Statik cezalar, ölüm cezası, dinamik cezalar, uyarlamalı cezalar, tavlama cezaları, ayrılmış cezalar, ortak evrim cezaları gibi kısıtlama işlemine yönelik birçok ceza yöntemi mevcuttur. Bu çalışmada, mühendislik problemlerinde kısıtlı fonksiyonları ele almak için statik ceza yöntemi [125] kullanılmıştır. Bu ceza yöntemi Denklem 6.12'de olduğu gibi eklemeli form kullanarak kısıtlamasız hale getirilmektedir. $p(x)$ ceza terimini ifade etmektedir. Genel olarak uygunluk fonksiyonu $eval(x)$ 'dir.

$$eval(x) = \begin{cases} f(x), & \text{if } x \in F \\ f(x) + p(x), & \text{otherwise} \end{cases} \quad (6.12)$$

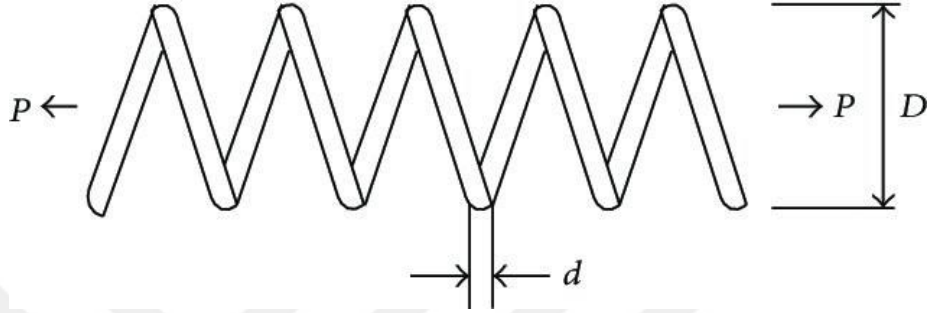
İhlal olmazsa $p(x)$ sıfır değerini alır, aksi durumda ise pozitif bir değeri almaktadır. Bu yaklaşımdaki arama ajanları genel olarak Denklem 6.13'teki gibi ifade edilmektedir.

$$eval(x) = f(x) + \sum_{k=1}^q 10^{15} \max[0, g_k(x)]^2 \quad (6.13)$$

- **Germe / Sıkıştırma Yay Tasarım Problemi**

Bu problem ilk defa Belegundu ve Arora tarafından tanımlandı. Bu problemin temel amacı, Şekil 6.1'de gösterildiği gibi dalgalanma frekansı, kesme gerilimi ve minimum sapma üzerindeki

kısıtlamalara maruz kalan bir gerilim / sıkıştırma yayının ağırlığını en aza indirmektir [126]. Problemin dört doğrusal olmayan eşitsizlik kısıtlaması vardır. x_1 , x_2 , ve x_3 olmak üzere üç sürekli değişkeni mevcuttur. Bu değişkenler, x_1 , tel çapı (w), x_2 ortalama bobin çapı (d) ve x_3 bobin numarası veya uzunluğu (L) şeklindedir. Problemin amaç fonksiyonu Denklem 6.14'te verilmektedir. Problemin kısıtları Denklem 6.15, 6.16, 6.17 ve 6.18'de verilmektedir. Problemin değişken aralığı Denklem 6.19'da verilmektedir.



Şekil 6.1. Germe/Sıkıştırma yay tasarımı

$$\min f(x_1, x_2, x_3) = (x_3 + 2)x_1^2 x_2 \quad (6.14)$$

$$g_1(x) = 1 - \frac{x_2^3 x_3}{71785 x_1^4} \leq 0 \quad (6.15)$$

$$g_2(x) = \frac{x_2(4x_2 - x_1)}{12566 x_1^3 (x_2 - x_1)} + \frac{1}{5108 x_1^2} - 1 \leq 0 \quad (6.16)$$

$$g_3(x) = 1 - \frac{140.45 x_1}{x_2^2 x_3} \leq 0 \quad (6.17)$$

$$g_4(x) = \frac{2(x_1 + x_2)}{3} - 1 \leq 0 \quad (6.18)$$

$$0.05 \leq x_1 \leq 2$$

$$0.25 \leq x_2 \leq 1.3$$

$$2 \leq x_3 \leq 15 \quad (6.19)$$

Deneyler esnasında popülasyon sayısı 40 alınmıştır. NFE sayısı 10000 ve bağımsız çalıştırma sayısı 50 olarak belirlenmiştir. Metasezgisel optimizasyon algoritmalarının performansı eşit şartlar altında ölçülmüştür. Gauss, Çember, Lojistik, Sinüzoidal ve Çadır haritaları özelinde KTOIO algoritmalarının performansları Tablo 6.17 ve 6.18'de gösterilmektedir. Tablo 6.17 ve

6.18’de Germe/Sıkıştırma yay tasarımı problemi F7 olarak belirtilmektedir. Deneylerde en iyi sonuçlar koyu olarak gösterilmektedir.

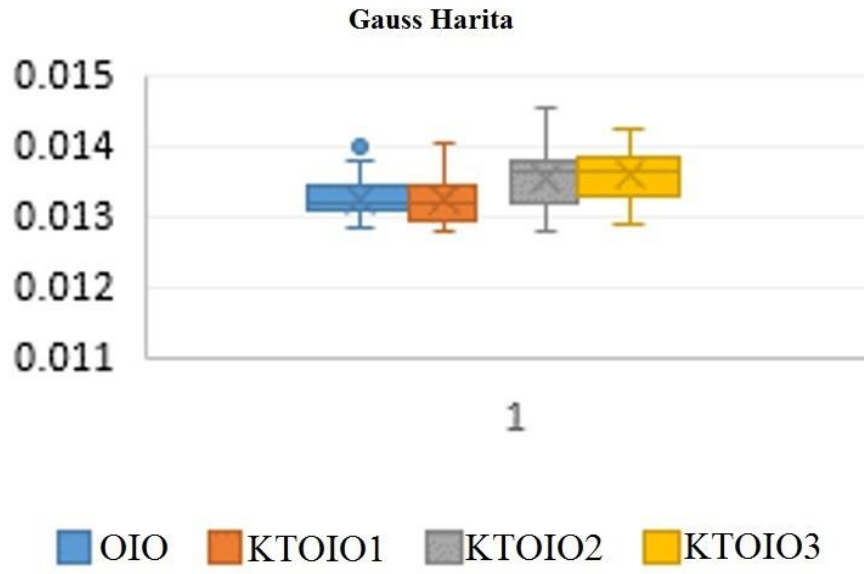
Tablo 6.17.Germe/Sıkıştırma yay tasarımında Çember, Gauss, Lojistik haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar

	Gauss Harita $f(x)$			Çember Harita $f(x)$			Lojistik Harita $f(x)$		
F7	KTOIO1	KTOIO2	KTOIO3	KTOIO1	KTOIO2	KTOIO3	KTOIO1	KTOIO2	KTOIO3
O	0.01321917	0.013547973	0.013566723	0.013228423	0.013513908	0.01352288	0.013218232	0.01353181	0.013494537
Min	0.01278969	0.012767321	0.012867007	0.012777869	0.012908009	0.012820197	0.012745497	0.012754534	0.012785203
Mak	0.01404274	0.014528654	0.01421676	0.01398381	0.014421749	0.014682316	0.014455816	0.01436604	0.014476903
SS	0.00030115	0.000389115	0.000352549	0.000278988	0.000332554	0.000390505	0.000342244	0.000382148	0.000410123

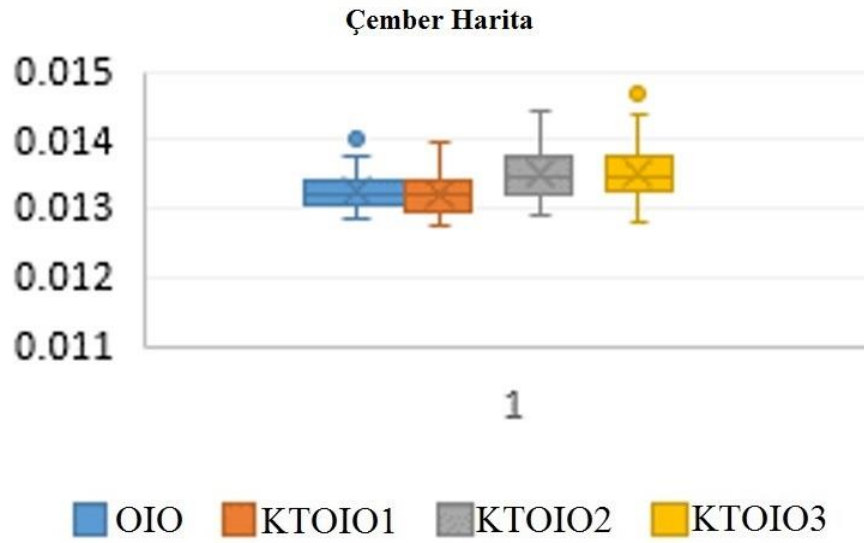
Tablo 6.18.Germe/Sıkıştırma yay tasarımında Sinüzoidal ve Çadır haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar

	Sinüzoidal Harita $f(x)$			Çadır Harita $f(x)$		
F7	KTOIO1	KTOIO2	KTOIO3	KTOIO1	KTOIO2	KTOIO3
O	0.01320715	0.013497544	0.013422002	0.013207259	0.013597303	0.013494952
Min	0.01280897	0.012797887	0.012912896	0.012680878	0.012805006	0.012778746
Mak	0.01385032	0.014588186	0.014122497	0.014120127	0.014788163	0.01437789
SS	0.00024642	0.000404327	0.000286361	0.000280725	0.000441274	0.000333813

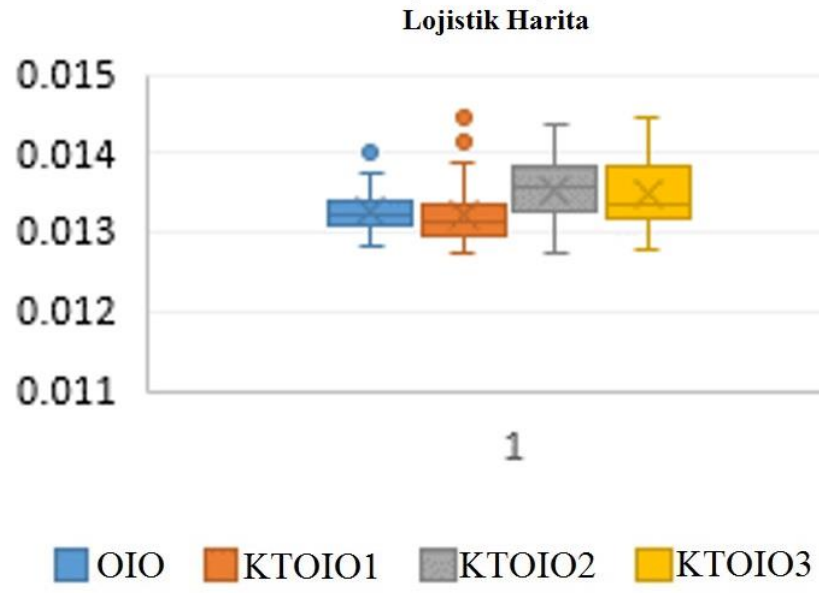
Tablo 6.17 ve 6.18 incelendiğinde KTOIO1 algoritmasının diğer KTOIO algoritmalarından daha iyi sonuç verdiği gözlemlenmektedir. KTOIO2 algoritması da KTOIO3 algoritmasından daha sonuçlar vermiştir. Farklı kaotik haritalara göre germe / sıkıştırma yay tasarım problemi için uygulanan metasezgisel algoritmaların kutu grafikleri gauss harita için Şekil 6.2’de, çember harita için Şekil 6.3’te, lojistik harita için Şekil 6.4’te, sinüzoidal harita için Şekil 6.5’te ve çadır harita için Şekil 6.6’da gösterilmektedir.



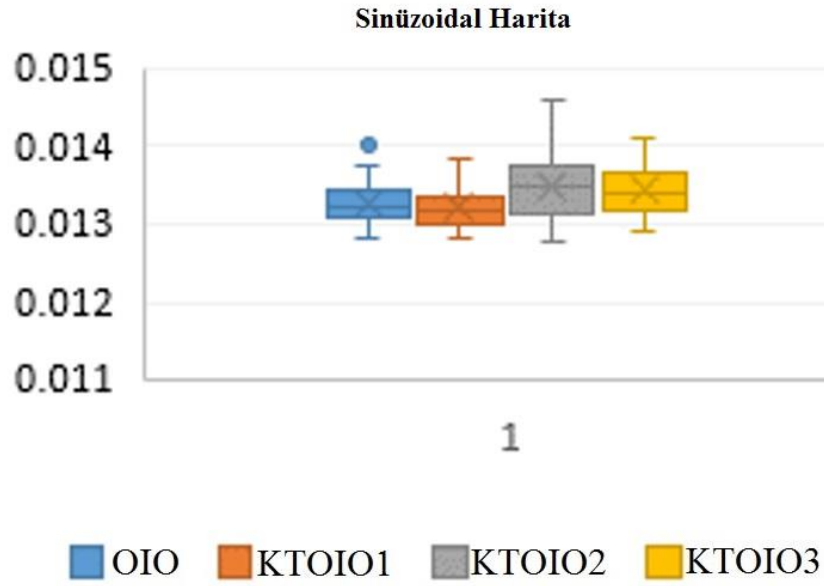
Şekil 6.2.Germe/Sıkıştırma yay tasarımında Gauss harita kullanımı



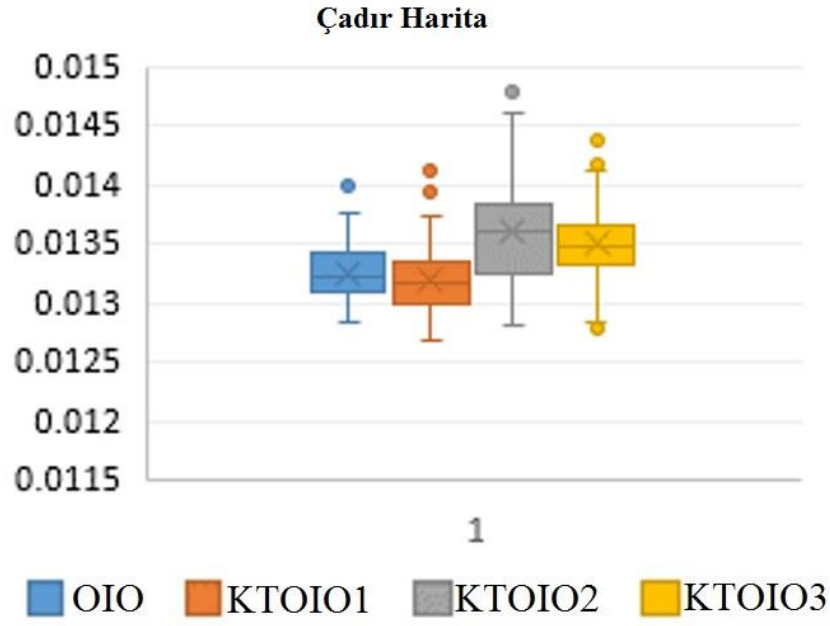
Şekil 6.3.Germe/Sıkıştırma yay tasarımında Çember harita kullanımı



Şekil 6.4.Germe/Sıkıştırma yay tasarımında Lojistik harita kullanımı



Şekil 6.5.Germe/Sıkıştırma yay tasarımında Sinüzoidal harita kullanımı



Şekil 6.6. Germe/Sıkıştırma yay tasarımında Çadır harita kullanımı

- **Kaynaklı Kiriş Tasarım Problemi**

Metasezgisel optimizasyon algoritmalarının kısıtlı problemlerdeki performansını ölçmek için kullanılan en iyi bilinen mühendislik tasarım problemlerinden biri kaynaklı kiriş tasarım problemidir [127]. Şekil 6.7'de gösterildiği gibi, problemdeki ana amaç, kayma gerilmesi (τ), hacimsel yük (P_c), uç sapma (δ), eğilme gerilmesi (σ) üzerindeki kısıtlamalara maruz kalan kaynaklı kirişin minimum üretim maliyetini bulmaktır. x_1 , x_2 , x_3 ve x_4 olmak üzere dört sürekli değişkeni mevcuttur. Bu değişkenler $x_1(h)$, $x_2(L)$, $x_3(t)$ ve $x_4(b)$ 'dir. Problemin amaç fonksiyonu Denklem 6.20'de verilmektedir. Problemin kısıtları Denklem 6.21, 6.22, 6.23, 6.24, 6.25, 6.26 ve 6.27'de verilmektedir. Problemin değişken aralığı Denklem 6.28'de verilmiştir.

Aşağıda listelenen problem için bazı sabitlere ihtiyaç vardır:

$$G = 12.106 \quad \sigma_{max} = 30000 \quad C_1 = 0.10471$$

$$E = 30.106 \quad \delta_{max} = 0.25 \quad C_2 = 0.04811$$

$$L = 14 \quad P = 6000 \quad C_3 = 1$$

$$\tau_{max} = 13600$$

$$V_w = x_1^2 x_2$$

$$V_s = x_3 x_4 (L + x_2)$$

$$\min f(x_1, x_2, x_3, x_4) = (C_1 + C_3)V_w + C_2 V_s$$

$$\min f(x_1, x_2, x_3, x_4) = 1.10471x_1^2 x_2 + 0.04811x_3 x_4 (14 + x_2) \quad (6.20)$$

$$g_1(x) = \tau_{max} - \tau(x_1, x_2, x_3, x_4) \geq 0 \quad (6.21)$$

$$g_2(x) = \sigma_{max} - \sigma(x_1, x_2, x_3, x_4) \geq 0 \quad (6.22)$$

$$g_3(x) = x_4 - x_1 \geq 0 \quad (6.23)$$

$$g_4(x) = C_1 x_1^2 + C_2 x_3 x_4 (L + x_2) - 5 \leq 0 \quad (6.24)$$

$$g_5(x) = x_1 - 0.125 \geq 0 \quad (6.25)$$

$$g_6(x) = \delta_{max} - \delta(x_1, x_2, x_3, x_4) \geq 0 \quad (6.26)$$

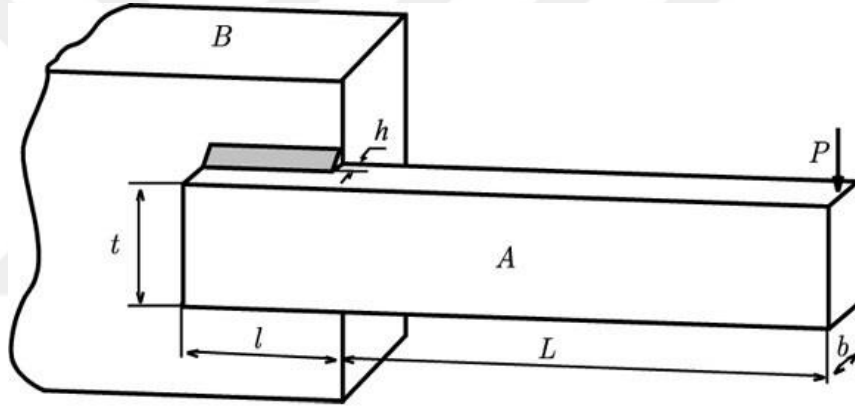
$$g_7(x) = P_c(x_1, x_2, x_3, x_4) - P \geq 0 \quad (6.27)$$

$$0.1 \leq x_1 \leq 2$$

$$0.1 \leq x_2 \leq 10$$

$$0.1 \leq x_3 \leq 10$$

$$0.1 \leq x_4 \leq 2 \quad (6.28)$$



Şekil 6.7. Kaynaklı kiriş tasarımı

Mühendislik hesaplamaları için gerekli olan bazı ifadeler Denklem 6.29, 6.30, 6.31, 6.32, 6.33, 6.34, 6.35, 6.36 ve 6.37’de gösterilmektedir.

$$\tau(x_1, x_2, x_3, x_4) = \sqrt{\tau_d^2 + 2\tau_d\tau_{dd}\frac{x_2}{2R} + \tau_{dd}^2} \quad (6.29)$$

$$\tau_d = \frac{P}{x_1 x_2 \sqrt{2}} \quad (6.30)$$

$$\tau_{dd} = \frac{MR}{J} \quad (6.31)$$

$$M = P(L + \frac{x_2}{2}) \quad (6.32)$$

$$R = \sqrt{\frac{x_2^2}{4} + (\frac{x_1 + x_3}{2})^2} \quad (6.33)$$

$$J = 2 \left(x_1 x_2 \sqrt{2} \left(\frac{x_2^2}{12} + (\frac{x_1 + x_3}{2})^2 \right) \right) \quad (6.34)$$

$$\sigma(x_1, x_2, x_3, x_4) = \frac{6PL}{x_3^2 x_4} \quad (6.35)$$

$$\delta(x_1, x_2, x_3, x_4) = \frac{4PL}{Ex_3^2 x_4} \quad (6.36)$$

$$P_c(x_1, x_2, x_3, x_4) = \frac{4.013E \sqrt{\frac{x_4^6 x_3^2}{36}}}{L^2} \left(1 - \frac{x_3}{2L} \sqrt{\frac{E}{4G}} \right) \quad (6.37)$$

Deneyler esnasında popülasyon sayısı 40 alınmıştır. NFE sayısı 10000 ve bağımsız çalıştırma sayısı 50 olarak belirlenmiştir. Metasezgisel optimizasyon algoritmalarının performansı eşit şartlar altında ölçülmüştür. Gauss, Çember, Lojistik, Sinüzoidal ve Çadır haritaları özelinde KTOIO algoritmalarının performansları Tablo 6.19 ve 6.20’de gösterilmektedir. Tablo 6.19 ve 6.20’de Kaynaklı kiriş tasarımı problemi F8 olarak belirtilmektedir. Deneylerde en iyi sonuçlar koyu olarak gösterilmektedir.

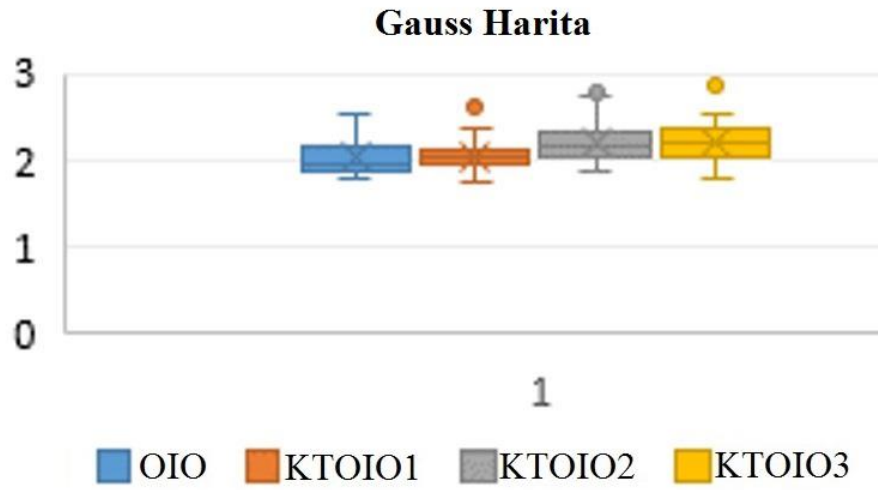
Tablo 6.19. Kaynaklı kiriş tasarımında Çember, Gauss, Lojistik haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar

Gauss Harita $f(x)$				Çember Harita $f(x)$			Lojistik Harita $f(x)$		
F8	KTOIO1	KTOIO2	KTOIO3	KTOIO1	KTOIO2	KTOIO3	KTOIO1	KTOIO2	KTOIO3
O	2.04009544	2.185342954	2.194990536	2.061999907	2.207299222	2.160653893	1.995520476	2.140300769	2.121185423
Min	1.75225438	1.869505004	1.799800714	1.748388077	1.820443997	1.843813054	1.749577289	1.786500273	1.787069796
Mak	2.61030889	2.792714931	2.880958261	2.771005611	2.636019301	3.014095094	2.511117901	2.824353719	2.602094705
SS	0.14924367	0.201874774	0.213128815	0.197275489	0.212690747	0.212647097	0.158190253	0.216360045	0.193013647

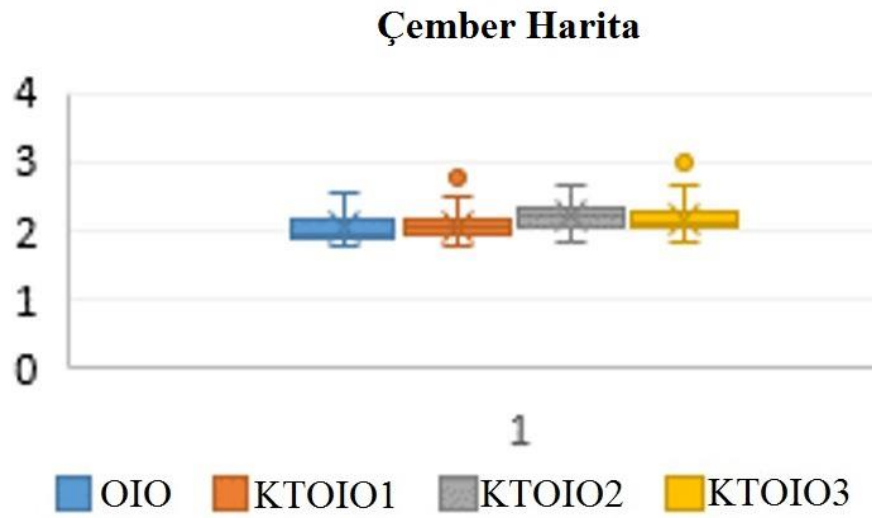
Tablo 6.20. Kaynaklı kiriş tasarımında Sinüzoidal ve Çadır haritalar kullanarak üç yeni KTOIO algoritmalarından elde edilen sonuçlar

Sinüzoidal Harita $f(x)$				Çadır Harita $f(x)$		
F8	KTOIO1	KTOIO2	KTOIO3	KTOIO1	KTOIO2	KTOIO3
O	2.03759730	2.145196912	2.173669177	2.004342651	2.168019424	2.159044184
Min	1.76168174	1.803866425	1.813135149	1.745261344	1.771983225	1.803324467
Mak	2.48293067	2.824341615	2.871448733	2.419534805	2.806242883	2.693395972
SS	0.16750174	0.198554877	0.211045784	0.152085507	0.210855489	0.242222048

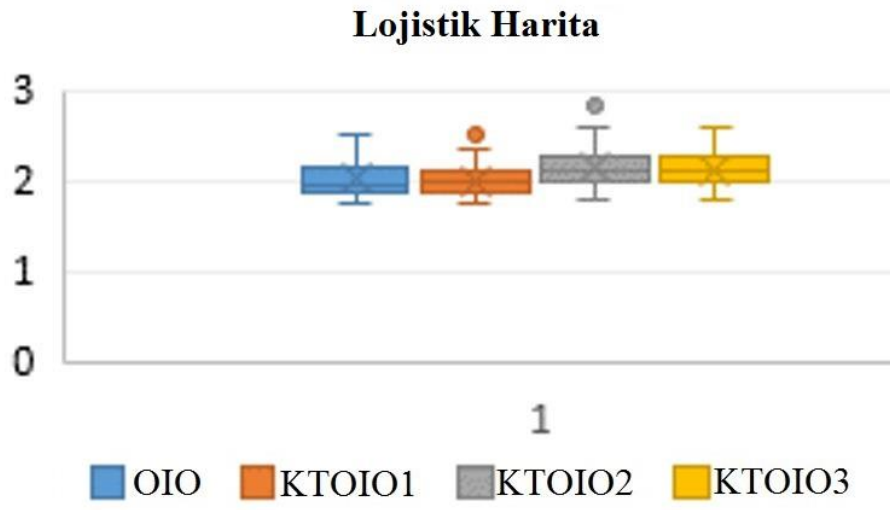
Tablo 6.19 ve 6.20 incelendiğinde KTOIO1 algoritmasının diğer KTOIO algoritmalarından daha iyi sonuç verdiği gözlemlenmektedir. KTOIO2 ve KTOIO3 algoritmaları denk sonuçlar vermiştir. Farklı kaotik haritalara göre kaynaklı kiriş tasarım problemi için uygulanan metasezgisel algoritmaların kutu grafikleri gauss harita için Şekil 6.8’de, çember harita için Şekil 6.9’da, lojistik harita için Şekil 6.10’da, sinüzoidal harita için Şekil 6.11’de ve çadır harita için Şekil 6.12’de gösterilmektedir.



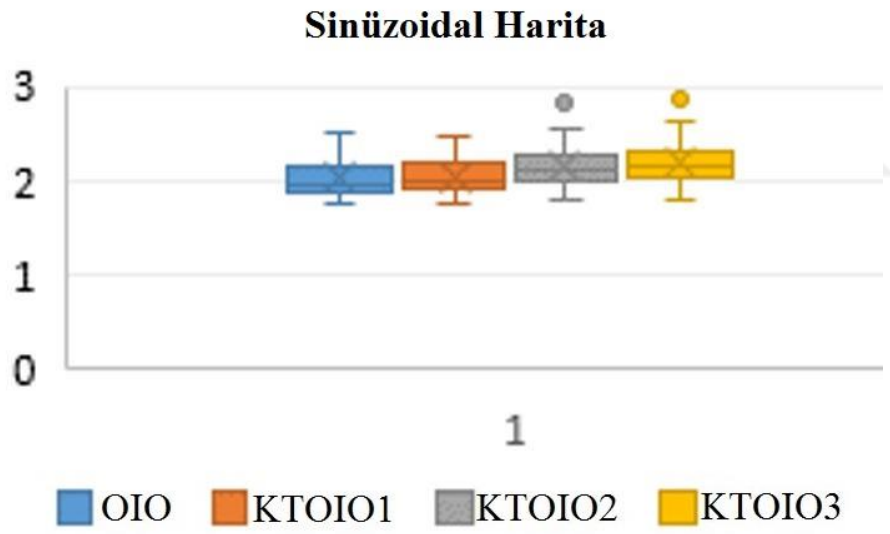
Şekil 6.8.Kaynaklı giriş tasarımında Gauss harita kullanımı



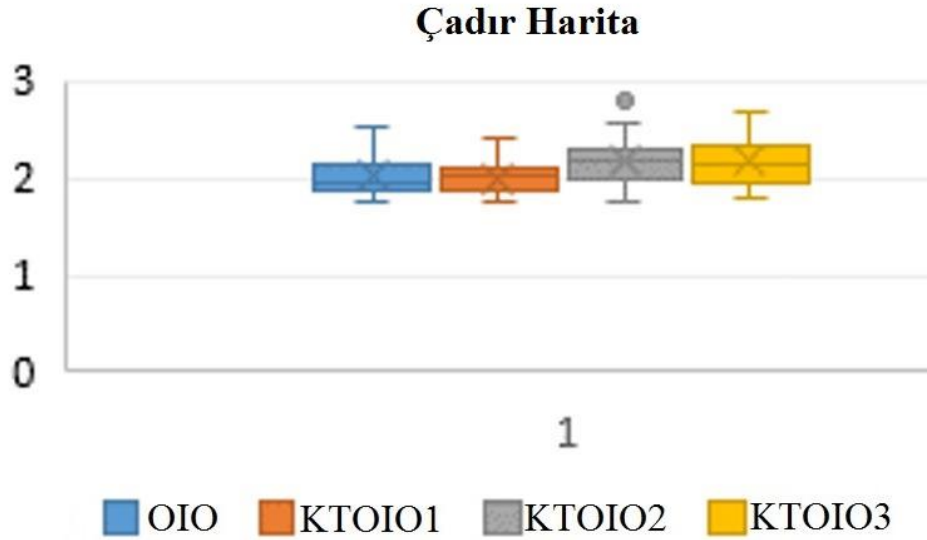
Şekil 6.9.Kaynaklı giriş tasarımında Çember harita kullanımı



Şekil 6.10.Kaynaklı giriş tasarımında Lojistik harita kullanımı



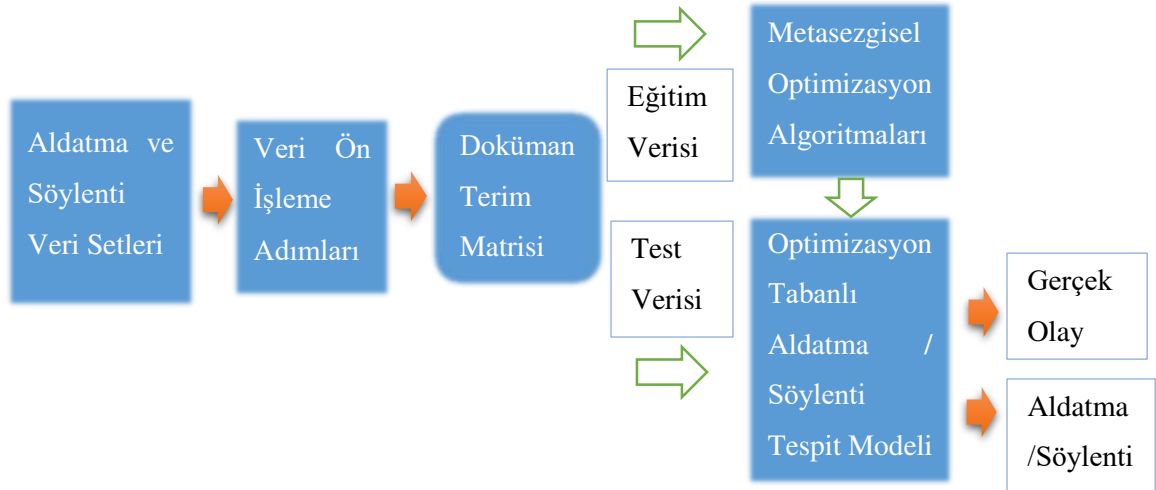
Şekil 6.11.Kaynaklı giriş tasarımında Sinüzoidal harita kullanımı



Şekil 6.12.Kaynaklı giriş tasarımında Çadır harita kullanımı

6.2. Metasezgisel Optimizasyon Algoritmalarının Aldatma ve Söylenti Tespit Problemleri için Modellenmesi

Veri setleri ön işlem adımlarından geçtikten sonra elde edilen doküman terim matrisindeki özellik sayısı, metasezgisel optimizasyon algoritmalarında uygun çözümü bulmak için kullanılacak olan arama ajanlarının boyutunun belirlenmesinde kullanılan temel kriterdir. Bu iki değer aynı olmalıdır. Yani bir diğer deyişle problemin boyutu (N) ise, bu N değeri her veri kümesi için elde edilen doküman matrisindeki terim sayısı ile ilişkilendirilmiştir. OIO, GKO ve KTOIO metasezgisel algoritmaları için başlangıç popülasyonları 0 ile 1 arasında rastgele değerlerden oluşturulmuştur. Tüm metasezgisel optimizasyon algoritmaları 20 aday çözüm, 100 iterasyon ve 30 bağımsız çalıştırma ile eşit şartlar altında test edilmiştir. Tez çalışmasında metasezgisel optimizasyon algoritmalarını kullanarak aldatma ve söylenti tespiti için izlenen adımlar Şekil 6.13'te gösterilmektedir.



Şekil 6.13. Aldatma ve söylenti tespiti için önerilen sistemin yapısı

Tez çalışmasında metinler arasındaki benzerliği ölçmek için Jaccard benzerliği kullanılmıştır [128]. $X_i = X_{i1}, X_{i2}, \dots, X_{iN}$ popülasyondaki i . birey ile $D_j = D_{j1}, D_{j2}, \dots, D_{jN}$ doküman matrisindeki j . haber satırına karşılık gelen vektör değerleri olarak alınırsa, dokümanlar arasındaki Jaccard benzerliği Denklem 6.38 ile hesaplanır [128].

$$jaccard\ değeri = \left| \frac{X_i \cap D_j}{X_i \cup D_j} \right| \quad (6.38)$$

Metasezgisel optimizasyon algoritmaları tarafından oluşturulan arama ajanlarının uygunluğunu hesaplamak için kullanılan uygunluk fonksiyonu Denklem 6.39 da gösterilmektedir. En iyi uygunluğa sahip olan arama ajanı en iyi çözüm olarak kabul edilmektedir. Denklem 6.39’da dört farklı performans değerlendirme kriteri aynı anda ele alınarak bir uygunluk fonksiyonu hesaplanmıştır.

$$Uygunluk\ fonksiyonu = a_1 * (Doğruluk) + a_2 * (Kesinlik) + a_3 * (Hassasiyet) + a_4 * (F\text{--}Ölçütü) \quad (6.39)$$

7. DENEYSEL BULGULAR

Bu tez çalışmasında, SM analiz problemleri olan “aldatma tespiti” ve “söylenti tespiti” problemlerinin sınıflandırılmasında ilk defa bir çözüm yöntemi olarak OIO, GKO ve ilk defa bu tezde önerilen üç adet yeni kaotik OIO metasezgisel optimizasyon algoritmaları kullanılmıştır. Bu metasezgisel optimizasyon algoritmaları farklı konuları baz alan gerçek dünya verilerinden elde edilen halka açık dört farklı veri seti üzerinde uygulanmıştır. Öncelikle, bu metasezgisel optimizasyon algoritmalarından elde edilen sonuçlar ile bu veri setlerinin KNIME yazılımı içerisinde yer alan 7 farklı denetimli makine öğrenme algoritmasından elde edilen sonuçlar sınıflandırma ölçütleri kullanılarak karşılaştırılmıştır. Metasezgisel optimizasyon algoritmalarının 30 defa bağımsız bir şekilde çalıştırılması ile elde edilen en iyi uygunluk değerleri, bu değerler için yapılan Friedman test sonuçları ve kutu grafikleri verilmiştir. Daha sonra, aldatma ve söylenti tespit problemleri için kullanılan beş metasezgisel optimizasyon algoritması ve yedi ayrı denetimli makine öğrenme algoritmasından elde edilen sonuçların dört farklı performans değerlendirme kriterine göre çizilmiş şekilsel gösterimleri verilmiştir. Performans karşılaştırılması yapılırken denetimli makine öğrenme algoritmaları için standart parametre değerleri kullanılmıştır. OIO, GKO ve KTOIO metasezgisel optimizasyon algoritmaları için problem ve algoritma başlangıç parametrelerinin değerleri Tablo 7.1’de gösterilmektedir.

Tablo 7.1. Metasezgisel optimizasyon algoritmalarına ait parametreler ve değerleri

OIO	$-1 < Pc < 1$ $Pc \sim 0$	Boyut: Döküman matrisinde bulunan özellik sayısı
GKO	Katsayı vektörü ($\vec{A}$): $[-2a, 2a]$	Popülasyon sayısı: 20
	Katsayı vektörü ($\vec{C}$): Rassal bir sayı	İterasyon sayısı: 100
KTOIO	w : Rassal bir sayı (0 ile 1 arasında)	Deney sayısı: 30
	r : Rassal bir sayı (0 ile 1 arasında)	Alt sınır (alt_i): 0 Üst sınır ($üst_i$): 1

7.1. Deceptive Opinion Spam Corpus v1.4 Veri Setinden Elde Edilen Sonuçlar

Deceptive opinion spam corpus v1.4 veri seti üzerinde, ön işlem adımları gerçekleştirilmiş ve sonrasında doküman terim matrisi elde edilmiştir. Gerçekleştirilen ön işlem adımları sonrasında 39 adet özellik bulunmuştur. Bu sebepten dolayı metasezgisel optimizasyon algoritmaları Deceptive opinion spam corpus v1.4 veri seti için çalıştırılırken problem boyutu 39 olarak ayarlanmıştır. Metasezgisel optimizasyon algoritmaları Deceptive opinion spam corpus v1.4 veri

seti için beş farklı kaotik harita ile 30 kez çalıştırılmış, elde edilen sonuçlar dört farklı performans değerlendirme kriterleri açısından karşılaştırılmıştır. Gauss harita, Çember harita, Lojistik harita, Sinüzoidal harita ve Çadır kullanılarak elde edilen KTOIO algoritmalarının, OIO ve GKO algoritmalarının deneysel sonuçları sırasıyla Tablo 7.2, 7.3, 7.4, 7.5 ve 7.6'da gösterilmektedir. Ayrıca en iyi değerler koyu olarak belirtilmektedir. Tablo 7.2'de görüldüğü üzere KTOIO1 algoritması tüm performans değerlendirme kriterleri açısından deneylerde kullanılan diğer tüm metasezgisel optimizasyon algoritmalarından daha yüksek başarımlar göstermiştir. OIO algoritması ise KTOIO1 algoritmasından sonra en yüksek ikinci doğruluk oranını göstermiştir.

Tablo 7.2. Deceptive opinion spam corpus v1.4 veri setinde Gauss harita kullanılarak elde edilen sonuçlar

	Doğruluk	Hassasiyet	Kesinlik	F_ölçütü
OIO	0.743	0.702	0.790	0.763
GKO	0.719	0.756	0.801	0.772
KTOIO1	0.804	0.822	0.918	0.851
KTOIO2	0.735	0.748	0.869	0.782
KTOIO3	0.741	0.762	0.859	0.779

Tablo 7.3.Deceptive opinion spam corpus v1.4 veri setinde Çember harita kullanılarak elde edilen sonuçlar

	Doğruluk	Hassasiyet	Kesinlik	F_ölçütü
OIO	0.743	0.702	0.790	0.763
GKO	0.719	0.756	0.801	0.772
KTOIO1	0.758	0.759	0.854	0.788
KTOIO2	0.684	0.731	0.759	0.718
KTOIO3	0.706	0.772	0.816	0.764

Tablo 7.4.Deceptive opinion spam corpus v1.4 veri setinde Lojistik harita kullanılarak elde edilen sonuçlar

	Doğruluk	Hassasiyet	Kesinlik	F_ölçütü
OIO	0.743	0.702	0.790	0.763
GKO	0.719	0.756	0.801	0.772
KTOIO1	0.720	0.779	0.782	0.764
KTOIO2	0.718	0.769	0.751	0.759
KTOIO3	0.706	0.768	0.785	0.750

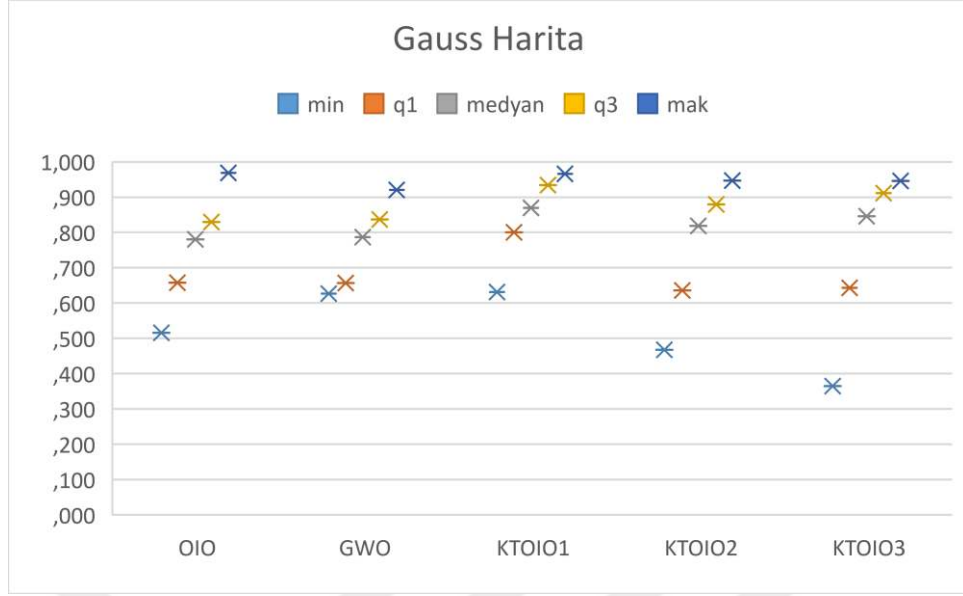
Tablo 7.5.Deceptive opinion spam corpus v1.4 veri setinde Sinüzoidal harita kullanılarak elde edilen sonuçlar

	Doğruluk	Hassasiyet	Kesinlik	F_ölçütü
OIO	0.743	0.702	0.790	0.763
GKO	0.719	0.756	0.801	0.772
KTOIO1	0.779	0.777	0.887	0.815
KTOIO2	0.718	0.730	0.795	0.747
KTOIO3	0.695	0.700	0.773	0.715

Tablo 7.6.Deceptive opinion spam corpus v1.4 veri setinde Çadır harita kullanılarak elde edilen sonuçlar

	Doğruluk	Hassasiyet	Kesinlik	F_ölçütü
OIO	0.743	0.702	0.790	0.763
GKO	0.719	0.756	0.801	0.772
KTOIO1	0.732	0.764	0.775	0.827
KTOIO2	0.705	0.754	0.804	0.754
KTOIO3	0.759	0.781	0.881	0.810

Tablo 7.2, 7.3, 7.4, 7.5 ve 7.6 da en yüksek performans değerlendirme ölçüt değerleri Gauss harita kullanılarak üretilen KTOIO algoritmalarından elde edildiği görülmektedir. Bu sebepten ötürü kutu grafikleri ve Friedman testleri gauss harita kullanılarak üretilen KTOIO algoritmalarına, GKO ve OIO algoritmalarına uygulanacaktır. Gauss harita kullanılarak elde edilen KTOIO1, KTOIO2 ve KTOIO3 algoritmaları, OIO ve GKO algoritmaları 30 kez bağımsız şekilde çalıştırılmasıyla elde edilen uygunluk değerlerine göre çizilen kutu grafiği Şekil 7.1’de verilmiştir. Grafik, uygunluk değerlerinin dağılımını ifade etmektedir. Grafikteki değerler dikkatle incelendiğinde birinci çeyrek ($q1$), ortanca değer (*medyan*) ve üçüncü çeyrek ($q3$) değerleri en yüksek metasezgisel algoritma KTOIO1 algoritmasıdır.



Şekil 7.1. Deceptive opinion spam corpus v1.4 veri setinde Gauss harita kullanılarak OIO, GKO, KTOIO1, KTOIO2 ve KTOIO3 algoritmalarının kutu grafiği

Ayrıca bu değerler Friedman testi ile değerlendirilerek metasezgisel optimizasyon algoritmalarının üretmiş oldukları sonuçlar arasında anlamlı bir fark olup olmadığı incelenmiştir. Metasezgisel optimizasyon algoritmalarının uygunluk fonksiyon değerlerine göre Friedman testi ile elde edilen algoritma ortalama sıralamaları Tablo 7.7’te gösterilmektedir.

Tablo 7.7.Deceptive opinion spam corpus v1.4 veri setinde Gauss harita kullanılarak uygunluk değerlerine göre Friedman testi ile elde edilen metasezgisel algoritmaların ortalama sıralamaları

Metasezgisel Algoritmalar	Ortalama Sıralama
OIO	2,40
GKO	2,53
KTOIO1	3,83
KTOIO2	2,97
KTOIO3	3,27

Tablo 7.8. Deceptive opinion spam corpus v1.4 veri seti için Friedman testi sonuçları

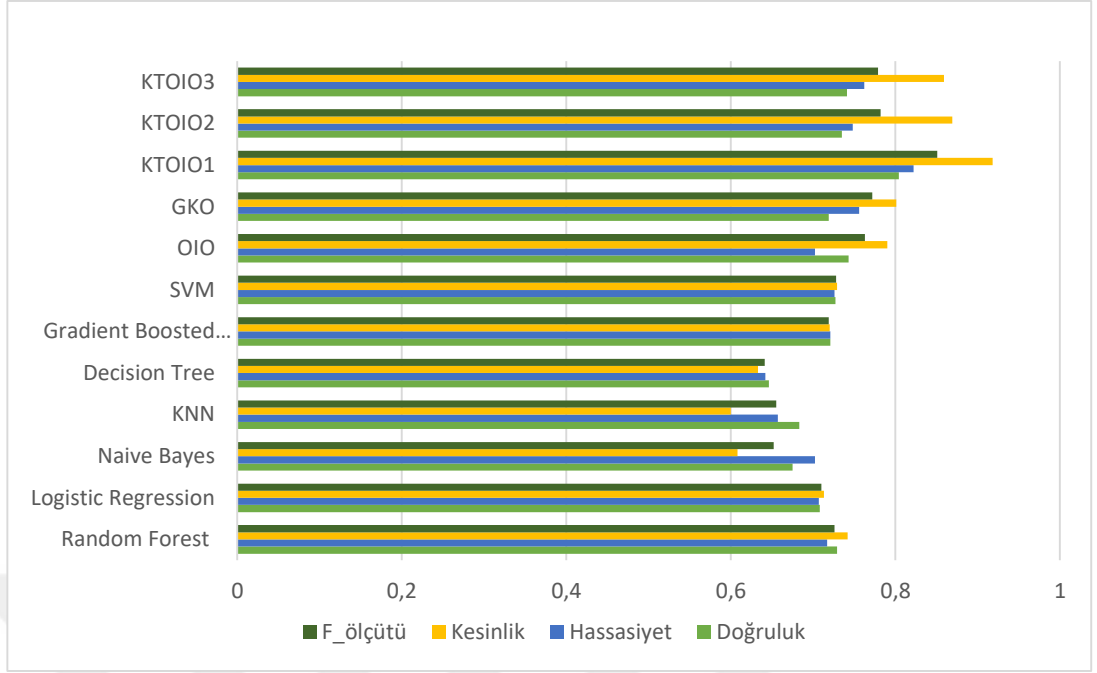
n	30
Ki-kare değeri	16,133
df	4
p -değeri	0,003

OIO, GKO, KTOIO1, KTOIO2 ve KTOIO3 metasezgisel algoritmalarının Friedman test istatistik sonuçları Tablo 7.8’te gösterilmektedir. Tablo incelendiğinde X_R^2 değeri 16,133’tür. Ki-kare tablosundan $\alpha=0,05$ ve $df=4$ değerleri için gösterilen (X_{tablo}^2) değeri 9,488 olarak görülmektedir. $16,133 > 9,488$ olduğundan dolayı H_0 hipotezi reddedilir ve H_1 hipotezi kabul edilir. Başka bir deyişle Friedman testine göre Deceptive opinion spam corpus v1.4 veri setinde gauss harita kullanılarak OIO, GKO, KTOIO1, KTOIO2 ve KTOIO3 metasezgisel algoritmalarından alınan sonuçlar arasındaki fark istatistiksel olarak anlamlıdır. Ayrıca, 7 ayrı denetimli makine öğrenme algoritması ve 5 metasezgisel optimizasyon algoritmasının 4 farklı performans değerlendirme kriterine göre elde edilen sonuçları Tablo 7.9’da ve bu tabloya ait görsel ifade Şekil 7.2’de gösterilmektedir.

Tablo 7.9 incelendiğinde KTOIO1 optimizasyon algoritmasının Deceptive opinion spam corpus v1.4 veri setinde deneylerde kullanılan tüm denetimli makine öğrenme algoritmalarından daha başarılı olduğu gözlemlenmiştir. Ayrıca denetimli makine öğrenme algoritmaları arasında RF metodu % 72,9 doğruluk oranı ile en yüksek başarıyı göstermiştir.

Tablo 7.9. Deceptive opinion spam corpus v1.4 veri setinde Gauss haritası kullanılarak elde edilen KTOIO algoritmaları, makine öğrenme algoritmaları, GKO ve OIO’nun karşılaştırılması

Algoritmalar	Doğruluk	Hassasiyet	Kesinlik	F-Ölçütü
Random Forest	0.729	0.717	0.742	0.726
Logistic Regression	0.708	0.707	0.713	0.710
Naive Bayes	0.675	0.702	0.608	0.652
KNN	0.683	0.657	0.600	0.655
Decision Tree	0.646	0.642	0.633	0.641
Gradient Boosted Tree	0.721	0.721	0.720	0.719
SVM	0.727	0.726	0.729	0.728
OIO	0.743	0.702	0.790	0.763
GKO	0.719	0.756	0.801	0.772
KTOIO1	0.804	0.822	0.918	0.851
KTOIO2	0.735	0.748	0.869	0.782
KTOIO3	0.741	0.762	0.859	0.779



Şekil 7.2. Denetimli makine öğrenme algoritmaları ve gauss harita kullanılarak elde edilen metasezgisel optimizasyon algoritmalarının karşılaştırılması

7.2. Real-Life Trial Veri Setinden Elde Edilen Sonuçlar

Real-Life Trial veri seti üzerinde, ön işlem adımları gerçekleştirilmiş ve sonrasında doküman terim matrisi elde edilmiştir. Gerçekleştirilen ön işlem adımları sonrasında 21 adet özellik bulunmuştur. Bu sebepten dolayı metasezgisel optimizasyon algoritmaları Real-Life veri seti için çalıştırılırken problem boyutu 21 olarak ayarlanmıştır. Metasezgisel optimizasyon algoritmaları Real-Life veri seti için beş farklı kaotik harita ile 30 kez çalıştırılmış, elde edilen sonuçlar dört farklı performans değerlendirme kriterleri açısından karşılaştırılmıştır. Gauss harita, Çember harita, Lojistik harita, Sinüzoidal harita ve Çadır kullanılarak elde edilen KTOIO algoritmalarının, OIO ve GKO algoritmalarının deneysel sonuçları sırasıyla Tablo 7.10, 7.11, 7.12, 7.13 ve 7.14'te gösterilmektedir. Ayrıca en iyi değerler koyu olarak belirtilmektedir. Performans değerlendirme kriterlerinin hepsinde en iyi sonuçları veren metasezgisel bir algoritmaya deneyler esnasında rastlanmadığı Tablo 7.10, 7.11, 7.12, 7.13 ve 7.14'te görülmektedir. Bu nedenden ötürü, Friedman istatistiksel test ve kutu grafik incelemeleri, Tablo 7.13'deki görüldüğü üzere en yüksek doğruluk değerinin elde edildiği Sinüzoidal harita kullanılarak elde edilen KTOIO algoritmaları, GKO ve OIO algoritmalarının uygunluk fonksiyon değerlerine göre gerçekleştirilecektir.

Tablo 7.10. Real-Life Trial veri setinde Gauss harita kullanılarak elde edilen sonuçlar

	Doğruluk	Hassasiyet	Kesinlik	F_ölçütü
OIO	0.694	0.718	0.791	0.740
GKO	0.663	0.703	0.749	0.740
KTOIO1	0.747	0.741	0.874	0.786
KTOIO2	0.703	0.721	0.814	0.750
KTOIO3	0.734	0.765	0.818	0.780

Tablo 7.11.Real-Life Trial veri setinde Çember harita kullanılarak elde edilen sonuçlar

	Doğruluk	Hassasiyet	Kesinlik	F_ölçütü
OIO	0.694	0.718	0.791	0.740
GKO	0.663	0.703	0.749	0.740
KTOIO1	0.695	0.758	0.764	0.749
KTOIO2	0.667	0.748	0.726	0.724
KTOIO3	0.680	0.746	0.770	0.743

Tablo 7.12.Real-Life Trial veri setinde Lojistik harita kullanılarak elde edilen sonuçlar

	Doğruluk	Hassasiyet	Kesinlik	F_ölçütü
OIO	0.694	0.718	0.791	0.740
GKO	0.663	0.703	0.749	0.740
KTOIO1	0.692	0.740	0.721	0.757
KTOIO2	0.708	0.760	0.782	0.768
KTOIO3	0.681	0.718	0.698	0.718

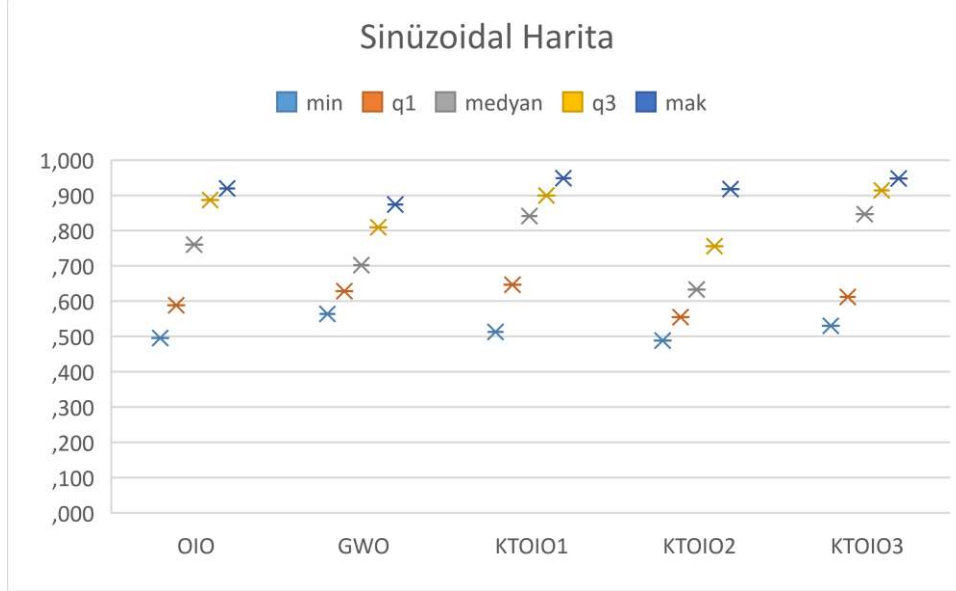
Tablo 7.13.Real-Life Trial veri setinde Sinüzoidal harita kullanılarak elde edilen sonuçlar

	Doğruluk	Hassasiyet	Kesinlik	F_ölçütü
OIO	0.694	0.718	0.791	0.740
GKO	0.663	0.703	0.749	0.740
KTOIO1	0.760	0.740	0.798	0.793
KTOIO2	0.639	0,661	0,661	0,673
KTOIO3	0.729	0.721	0.800	0.760

Tablo 7.14.Real-Life Trial veri setinde Çadır harita kullanılarak elde edilen sonuçlar

	Doğruluk	Hassasiyet	Kesinlik	F_ölçütü
OIO	0.694	0.718	0.791	0.740
GKO	0.663	0.703	0.749	0.740
KTOIO1	0.681	0.680	0.726	0.720
KTOIO2	0.634	0.688	0.663	0.680
KTOIO3	0.688	0.777	0.723	0.752

Sinüzoidal harita kullanılarak elde edilen KTOIO1, KTOIO2 ve KTOIO3 algoritmaları, OIO ve GKO algoritmaları 30 kez bağımsız şekilde çalıştırılmasıyla elde edilen uygunluk değerlerine göre çizilen kutu grafiği Şekil 7.3’de verilmiştir. Grafik, uygunluk değerlerinin dağılımını ifade etmektedir. Grafikteki değerler dikkatle incelendiğinde birinci çeyrek (*q1*) değeri en yüksek metasezgisel algoritma KTOIO1 algoritmasıdır. Ortanca değer (*medyan*) ve üçüncü çeyrek (*q3*) KTOIO1 ve KTOIO3 algoritmalarının birbirlerine çok yakın olmakla birlikte diğer algoritmalarından yukarda gözükmemektedir.



Şekil 7.3.Real-Life Trial veri setinde Sinüzoidal harita kullanılarak OIO, GKO, KTOIO1, KTOIO2 ve KTOIO3 algoritmalarının kutu grafiği

Ayrıca bu değerler Friedman testi ile değerlendirilerek metasezgisel optimizasyon algoritmalarının üretmiş oldukları sonuçlar arasında anlamlı bir fark olup olmadığı incelenmiştir.

Metasezgisel optimizasyon algoritmalarının uygunluk fonksiyon değerlerine göre Friedman testi ile elde edilen algoritma ortalama sıralamaları Tablo 7.15’te gösterilmektedir.

Tablo 7.15.Real-Life Trial veri setinde Sinüzoidal harita kullanılarak uygunluk değerlerine göre Friedman testi ile elde edilen metasezgisel algoritmaların ortalama sıralamaları

Metasezgisel Algoritmalar	Ortalama Sıralama
OIO	3,10
GKO	2,73
KTOIO1	3,57
KTOIO2	2,20
KTOIO3	3,40

Tablo 7.16.Real-Life Trial veri seti için Friedman testi sonuçları

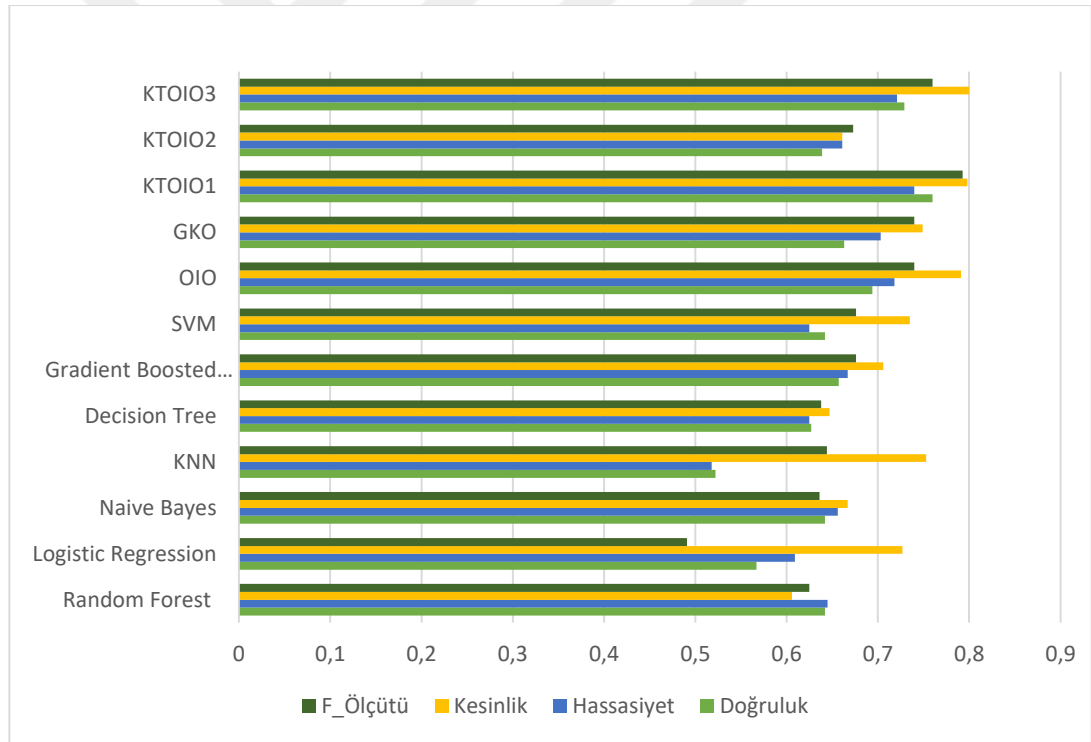
n	30
Ki-kare değeri	14,427
df	4
p -değeri	0,006

OIO, GKO, KTOIO1, KTOIO2 ve KTOIO3 metasezgisel algoritmalarının Friedman test istatistik sonuçları Tablo 7.16’da gösterilmektedir. Tablo incelendiğinde X_R^2 değeri 14,427’dir. Ki-kare tablosundan $\alpha=0,05$ ve $df=4$ değerleri için gösterilen (X_{tablo}^2) değeri 9,488 olarak görülmektedir. $14,427 > 9,488$ olduğundan dolayı H_0 hipotezi reddedilir ve H_1 hipotezi kabul edilir. Başka bir deyişle Friedman testine göre Real-Life Trial veri setinde Sinüzoidal harita kullanılarak OIO, GKO, KTOIO1, KTOIO2 ve KTOIO3 metasezgisel algoritmalarından alınan sonuçlar arasındaki fark istatistiksel olarak anlamlıdır. Ayrıca, 7 ayrı denetimli makine öğrenme algoritması ve 5 metasezgisel optimizasyon algoritmasının 4 farklı performans değerlendirme kriterine göre elde edilen sonuçları Tablo 7.17’de ve bu tabloya ait görsel ifade Şekil 7.4’te gösterilmektedir.

Tablo 7.17 incelendiğinde KTOIO1 optimizasyon algoritmasının Real-Life Trial veri setinde deneylerde kullanılan tüm denetimli makine öğrenme algoritmalarından daha başarılı olduğu gözlemlenmiştir. Ayrıca denetimli makine öğrenme algoritmaları arasında Gradient Boosted Tree metodu % 65,7 doğruluk oranı ile en yüksek başarıyı göstermiştir.

Tablo 7.17.Real-Life Trial veri setinde Sinüzoidal harita kullanılarak elde edilen KTOIO algoritmaları, makine öğrenme algoritmaları, GKO ve OIO'nun karşılaştırılması

Algoritmalar	Doğruluk	Hassasiyet	Kesinlik	F-Ölçütü
Random Forest	0.642	0.645	0.606	0.625
Logistic Regression	0.567	0.609	0.727	0.491
Naive Bayes	0.642	0.656	0.667	0.636
KNN	0.522	0.518	0.753	0.644
Decision Tree	0.627	0.625	0.647	0.638
Gradient Boosted Tree	0.657	0.667	0.706	0.676
SVM	0.642	0.625	0.735	0.676
OIO	0.694	0.718	0.791	0.740
GKO	0.663	0.703	0.749	0.740
KTOIO1	0.760	0.740	0.798	0.793
KTOIO2	0.639	0.661	0.661	0.673
KTOIO3	0.729	0.721	0.800	0.760



Şekil 7.4.Denetimli makine öğrenme algoritmaları ve sinüzoidal harita kullanılarak elde edilen metasezgisel optimizasyon algoritmalarının karşılaştırılması

7.3. S1 Veri Setinden Elde Edilen Sonuçlar

S1 veri seti üzerinde, ön işlem adımları gerçekleştirilmiş ve sonrasında doküman terim matrisi elde edilmiştir. Gerçekleştirilen ön işlem adımları sonrasında 24 adet özellik bulunmuştur. Bu sebepten dolayı metasezgisel optimizasyon algoritmaları S1 veri seti için çalıştırılırken problem

boyutu 24 olarak ayarlanmıştır. Metasezgisel optimizasyon algoritmaları S1 veri seti için beş farklı kaotik harita ile 30 kez çalıştırılmış, elde edilen sonuçlar dört farklı performans değerlendirme kriterleri açısından karşılaştırılmıştır. Gauss harita, Çember harita, Lojistik harita, Sinüzoidal harita ve Çadır kullanılarak elde edilen KTOIO algoritmalarının, OIO ve GKO algoritmalarının deneysel sonuçları sırasıyla Tablo 7.18, 7.19, 7.20, 7.21 ve 7.22’de gösterilmektedir. Ayrıca en iyi değerler koyu olarak belirtilmektedir. Performans değerlendirme kriterlerinin tümünde en iyi sonuçları veren metasezgisel bir algoritmaya deneyler esnasında rastlanmadığı Tablo 7.18, 7.19, 7.20, 7.21 ve 7.22’de görülmektedir. Bu nedenden ötürü, Friedman istatistiksel test ve kutu grafik incelemeleri, Tablo 7.18’deki görüldüğü üzere en yüksek doğruluk değerinin elde edildiği Gauss harita kullanılarak elde edilen KTOIO algoritmaları, GKO ve OIO algoritmalarının uygunluk fonksiyon değerlerine göre gerçekleştirilecektir.

Tablo 7.18.S1 veri setinde Gauss harita kullanılarak elde edilen sonuçlar

	Doğruluk	Hassasiyet	Kesinlik	F_ölçütü
OIO	0.820	0.814	0.877	0.862
GKO	0.878	0.872	0.897	0.890
KTOIO1	0.921	0.908	0.960	0.924
KTOIO2	0.749	0.744	0.752	0.783
KTOIO3	0.781	0.782	0.800	0.815

Tablo 7.19.S1 veri setinde Çember harita kullanılarak elde edilen sonuçlar

	Doğruluk	Hassasiyet	Kesinlik	F_ölçütü
OIO	0.820	0.814	0.877	0.862
GKO	0.878	0.872	0.897	0.890
KTOIO1	0.841	0.838	0.882	0.852
KTOIO2	0.733	0.775	0.767	0.770
KTOIO3	0.800	0.810	0.869	0.838

Tablo 7.20.S1 veri setinde Lojistik harita kullanılarak elde edilen sonuçlar

	Doğruluk	Hassasiyet	Kesinlik	F_ölçütü
OIO	0.820	0.814	0.877	0.862
GKO	0.878	0.872	0.897	0.890
KTOIO1	0.857	0.867	0.902	0.886
KTOIO2	0.735	0.768	0.801	0.788
KTOIO3	0.745	0.771	0.808	0.772

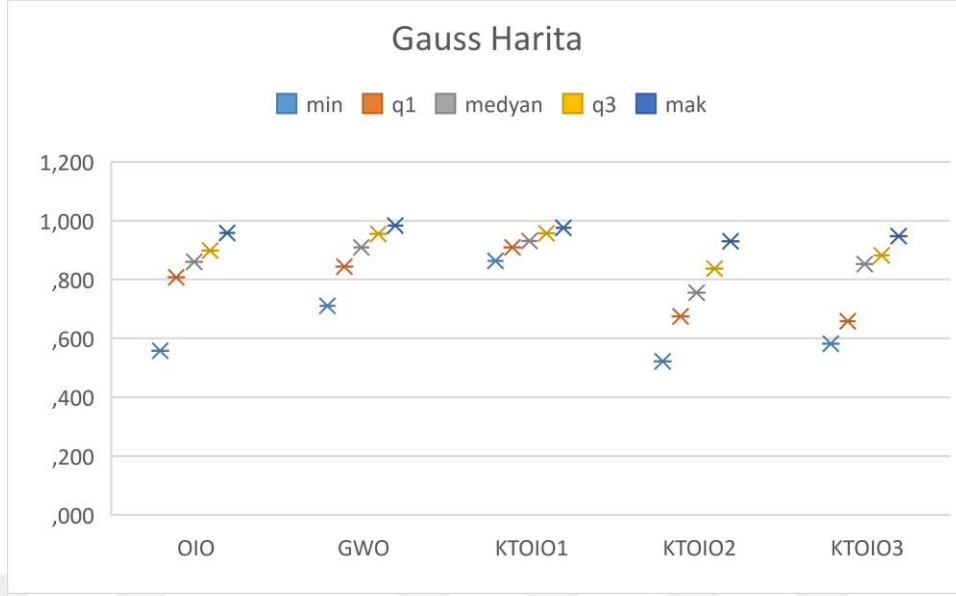
Tablo 7.21.S1 veri setinde Sinüzoidal harita kullanılarak elde edilen sonuçlar

	Doğruluk	Hassasiyet	Kesinlik	F_ölçütü
OIO	0.820	0.814	0.877	0.862
GKO	0.878	0.872	0.897	0.890
KTOIO1	0.882	0.877	0.886	0.899
KTOIO2	0.744	0.736	0.795	0.790
KTOIO3	0.773	0.777	0.833	0.813

Tablo 7.22.S1 veri setinde Çadır harita kullanılarak elde edilen sonuçlar

	Doğruluk	Hassasiyet	Kesinlik	F_ölçütü
OIO	0.820	0.814	0.877	0.862
GKO	0.878	0.872	0.897	0.890
KTOIO1	0.767	0.756	0.805	0.796
KTOIO2	0.721	0.724	0.719	0.732
KTOIO3	0.845	0.867	0.916	0.874

Gauss harita kullanılarak elde edilen KTOIO1, KTOIO2 ve KTOIO3 algoritmaları, OIO ve GKO algoritmaları 30 kez bağımsız şekilde çalıştırılmasıyla elde edilen uygunluk değerlerine göre çizilen kutu grafiği Şekil 7.5'te verilmiştir. Grafik, uygunluk değerlerinin dağılımını ifade etmektedir. Grafikteki değerler dikkatle incelendiğinde birinci çeyrek ($q1$) ve ortanca (*medyan*) değer en yüksek metasezgisel algoritma KTOIO1 algoritmasıdır. Üçüncü çeyrek ($q3$) KTOIO1 ve GWO algoritmalarının birbirlerine çok yakın olmakla birlikte diğer algoritmalar yukarıda görülmektedir. Ayrıca en yüksek *min* değeri de KTOIO1 algoritmasında bulunmaktadır.



Şekil 7.5. S1 veri setinde Gauss harita kullanılarak OIO, GKO, KTOIO1, KTOIO2 ve KTOIO3 algoritmalarının kutu grafiği

Ayrıca bu değerler Friedman testi ile değerlendirilerek metasezgisel optimizasyon algoritmalarının üretmiş oldukları sonuçlar arasında anlamlı bir fark olup olmadığı incelenmiştir. Metasezgisel optimizasyon algoritmalarının uygunluk fonksiyon değerlerine göre Friedman testi ile elde edilen algoritma ortalama sıralamaları Tablo 7.23'te gösterilmektedir.

Tablo 7.23. S1 veri setinde Gauss harita kullanılarak uygunluk değerlerine göre Friedman testi ile elde edilen metasezgisel algoritmaların ortalama sıralamaları

Metasezgisel Algoritmalar	Ortalama Sıralama
OIO	2,63
GKO	3,70
KTOIO1	4,50
KTOIO2	1,83
KTOIO3	2,33

Tablo 7.24. S1 veri seti için Friedman testi sonuçları

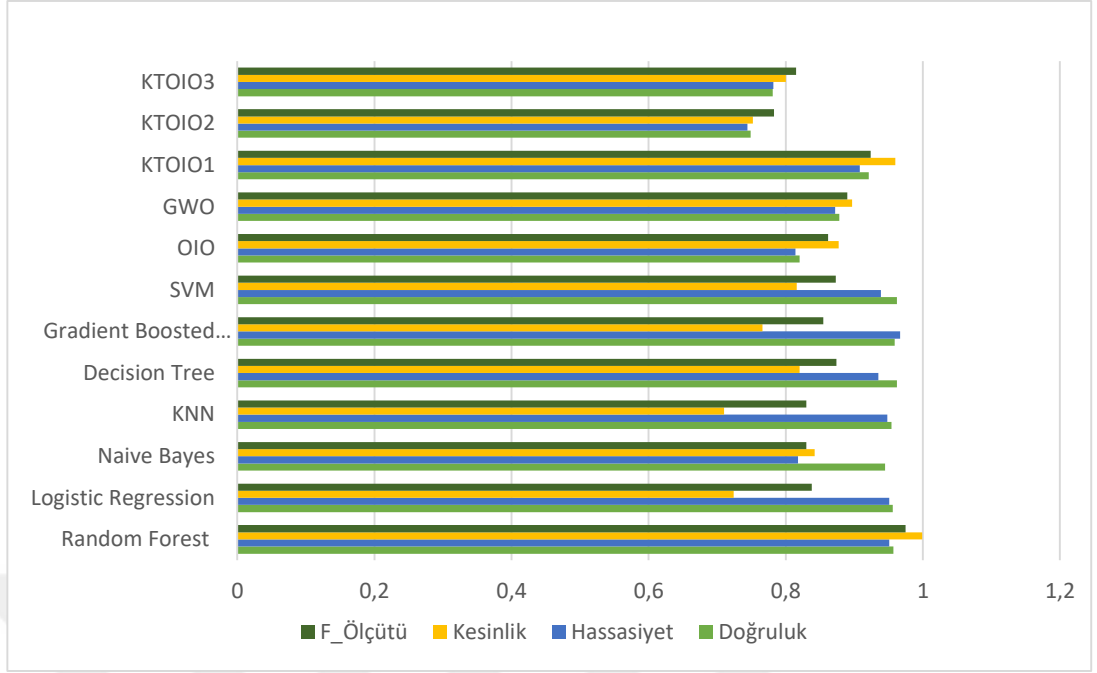
n	30
Ki-kare değeri	56,160
df	4
p -değeri	0,000

OIO, GKO, KTOIO1, KTOIO2 ve KTOIO3 metasezgisel algoritmalarının Friedman test istatistik sonuçları Tablo 7.24'te gösterilmektedir. Tablo incelendiğinde X^2_R değeri 56,160'dır. Ki-kare tablosundan $\alpha=0,05$ ve $df=4$ değerleri için gösterilen (X^2_{tablo}) değeri 9,488 olarak görülmektedir. $56,160 > 9,488$ olduğundan dolayı H_0 hipotezi reddedilir ve H_1 hipotezi kabul edilir. Başka bir deyişle Friedman testine göre S1 veri setinde Gauss harita kullanılarak OIO, GKO, KTOIO1, KTOIO2 ve KTOIO3 metasezgisel algoritmalarından alınan sonuçlar arasındaki fark istatistiksel olarak anlamlıdır. Ayrıca, 7 ayrı denetimli makine öğrenme algoritması ve 5 metasezgisel optimizasyon algoritmasının 4 farklı performans değerlendirme kriterine göre elde edilen sonuçları Tablo 7.25'de ve bu tabloya ait görsel ifade Şekil 7.6'da gösterilmektedir.

Tablo 7.25 incelendiğinde metasezgisel optimizasyon algoritmalarının S1 veri setinde denetimli makine öğrenme algoritmaları kadar başarılı olamadığı gözlemlenmiştir. Ayrıca denetimli makine öğrenme algoritmaları arasında DT ve SVM metotları % 96,2 doğruluk oranı ile en yüksek başarıyı göstermişlerdir.

Tablo 7.25. S1 veri setinde Gauss harita kullanılarak elde edilen KTOIO algoritmaları, makine öğrenme algoritmaları, GKO ve OIO'nun karşılaştırılması

Algoritmalar	Doğruluk	Hassasiyet	Kesinlik	F-Ölçütü
Random Forest	0.957	0.951	0.999	0.975
Logistic Regression	0.956	0.951	0.724	0.838
Naive Bayes	0.945	0.818	0.842	0.830
KNN	0.954	0.948	0.710	0.830
Decision Tree	0.962	0.935	0.820	0.874
Gradient Boosted Tree	0.959	0.967	0.766	0.855
SVM	0.962	0.939	0.816	0.873
OIO	0.820	0.814	0.877	0.862
GKO	0.878	0.872	0.897	0.890
KTOIO1	0.921	0.908	0.960	0.924
KTOIO2	0.749	0.744	0.752	0.783
KTOIO3	0.781	0.782	0.800	0.815



Şekil 7.6.Denetimli makine öğrenme algoritmaları ve gauss harita kullanılarak elde edilen metasezgisel optimizasyon algoritmalarının karşılaştırılması

7.4. Germanwings Crash Veri Setinden Elde Edilen Sonuçlar

Germanwings Crash veri seti üzerinde, ön işlem adımları gerçekleştirilmiş ve sonrasında doküman terim matrisi elde edilmiştir. Gerçekleştirilen ön işlem adımları sonrasında 40 adet özellik bulunmuştur. Bu sebepten dolayı metasezgisel optimizasyon algoritmaları Germanwings Crash veri seti için çalıştırılırken problem boyutu 40 olarak ayarlanmıştır. Metasezgisel optimizasyon algoritmaları Germanwings Crash veri seti için beş farklı kaotik harita ile 30 kez çalıştırılmış, elde edilen sonuçlar dört farklı performans değerlendirme kriterleri açısından karşılaştırılmıştır. Gauss harita, Çember harita, Lojistik harita, Sinüzoidal harita ve Çadır kullanılarak elde edilen KTOIO algoritmalarının, OIO ve GKO algoritmalarının deneysel sonuçları sırasıyla Tablo 7.26, 7.27, 7.28, 7.29 ve 7.30'da gösterilmektedir. Ayrıca en iyi değerler koyu olarak belirtilmektedir. Performans değerlendirme kriterlerinin tümünde en iyi sonuçları veren metasezgisel bir algoritmaya deneyler esnasında rastlanmadığı Tablo 7.26, 7.27, 7.28, 7.29 ve 7.30'da görülmektedir. Bu nedenden ötürü, Friedman istatistiksel test ve kutu grafik incelemeleri, Tablo 7.28'deki görüldüğü üzere en yüksek doğruluk değerinin elde edildiği Lojistik harita kullanılarak elde edilen KTOIO algoritmaları, GKO ve OIO algoritmalarının uygunluk fonksiyon değerlerine göre gerçekleştirilecektir.

Tablo 7.26. Germanwings Crash veri setinde Gauss harita kullanılarak elde edilen sonuçlar

	Doğruluk	Hassasiyet	Kesinlik	F_ölçütü
OIO	0.830	0.809	0.846	0.843
GKO	0.775	0.785	0.833	0.815
KTOIO1	0.877	0.847	0.891	0.895
KTOIO2	0.808	0.756	0.739	0.817
KTOIO3	0.763	0.751	0.839	0.794

Tablo 7.27. Germanwings Crash veri setinde Çember harita kullanılarak elde edilen sonuçlar

	Doğruluk	Hassasiyet	Kesinlik	F_ölçütü
OIO	0.830	0.809	0.846	0.843
GKO	0.775	0.785	0.833	0.815
KTOIO1	0.785	0.793	0.850	0.815
KTOIO2	0.707	0.745	0.746	0.701
KTOIO3	0.775	0.748	0.808	0.805

Tablo 7.28. Germanwings Crash veri setinde Lojistik harita kullanılarak elde edilen sonuçlar

	Doğruluk	Hassasiyet	Kesinlik	F_ölçütü
OIO	0.830	0.809	0.846	0.843
GKO	0.775	0.785	0.833	0.815
KTOIO1	0.923	0.911	0.950	0.928
KTOIO2	0.821	0.815	0.844	0.839
KTOIO3	0.859	0.845	0.934	0.888

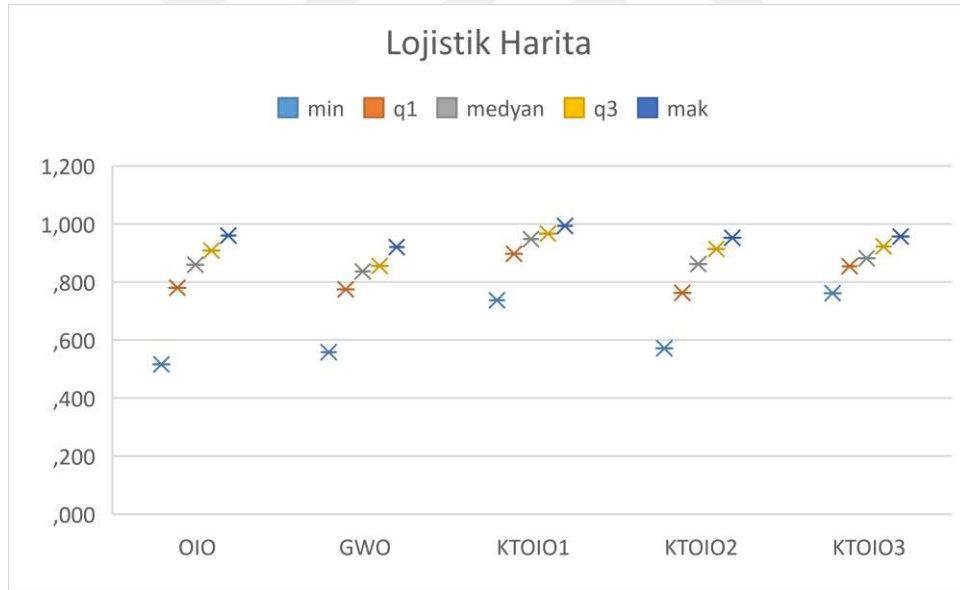
Tablo 7.29. Germanwings Crash veri setinde Sinüzoidal harita kullanılarak elde edilen sonuçlar

	Doğruluk	Hassasiyet	Kesinlik	F_ölçütü
OIO	0.830	0.809	0.846	0.843
GKO	0.775	0.785	0.833	0.815
KTOIO1	0.845	0.817	0.872	0.855
KTOIO2	0.727	0.717	0.737	0.734
KTOIO3	0.862	0.794	0.877	0.837

Tablo 7.30. Germanwings Crash veri setinde Çadır harita kullanılarak elde edilen sonuçlar

	Doğruluk	Hassasiyet	Kesinlik	F_ölçütü
OIO	0.830	0.809	0.846	0.843
GKO	0.775	0.785	0.833	0.815
KTOIO1	0.858	0.828	0.900	0.878
KTOIO2	0.754	0.775	0.847	0.792
KTOIO3	0.798	0.792	0.829	0.837

Lojistik harita kullanılarak elde edilen KTOIO1, KTOIO2 ve KTOIO3 algoritmaları, OIO ve GKO algoritmaları 30 kez bağımsız şekilde çalıştırılmasıyla elde edilen uygunluk değerlerine göre çizilen kutu grafiği Şekil 7.7’de verilmiştir. Grafik, uygunluk değerlerinin dağılımını ifade etmektedir. Grafikteki değerler dikkatle incelendiğinde birinci çeyrek (*q1*), ortanca (*medyan*) değer, üçüncü çeyrek (*q3*) ve *mak* en yüksek metasezgisel algoritma KTOIO1 algoritmasıdır. KTOIO1 ve KTOIO3 algoritmalarının birbirlerine çok yakın olmakla birlikte diğer algoritmalarından yukarda gözükmemektedir. Ayrıca en yüksek *min* değeri de KTOIO3 algoritmasında bulunmaktadır.

**Şekil 7.7.** Germanwings veri setinde Lojistik harita kullanılarak OIO, GKO, KTOIO1, KTOIO2 ve KTOIO3 algoritmalarının kutu grafiği

Ayrıca bu değerler Friedman testi ile değerlendirilerek metasezgisel optimizasyon algoritmalarının ürettiği sonuçlar arasında anlamlı bir fark olup olmadığı incelenmiştir. Metasezgisel optimizasyon algoritmalarının uygunluk fonksiyon değerlerine göre Friedman testi ile elde edilen algoritma ortalama sıralamaları Tablo 7.31’de gösterilmektedir.

Tablo 7.31. Germanwings Crash veri setinde Lojistik harita kullanılarak uygunluk değerlerine göre Friedman testi ile elde edilen metasezgisel algoritmaların ortalama sıralamaları

Metasezgisel Algoritmalar	Ortalama Sıralama
OIO	2,70
GKO	2,03
KTOIO1	4,37
KTOIO2	2,60
KTOIO3	3,30

Tablo 7.32. Germanwings Crash veri seti için Friedman testi sonuçları

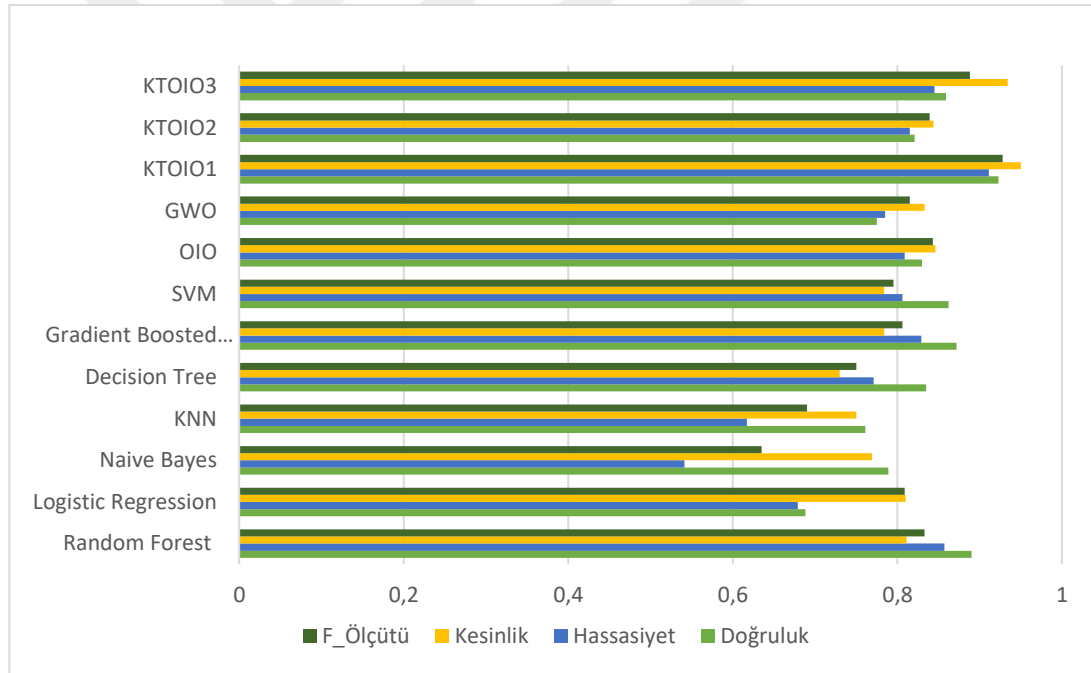
n	30
Ki-kare değeri	37,707
df	4
p -değeri	0,001

OIO, GKO, KTOIO1, KTOIO2 ve KTOIO3 metasezgisel algoritmalarının Friedman test istatistik sonuçları Tablo 7.32’de gösterilmektedir. Tablo incelendiğinde X_R^2 değeri 37,707’dir. Ki-kare tablosundan $\alpha=0,05$ ve $df=4$ değerleri için gösterilen (X_{tablo}^2) değeri 9,488 olarak görülmektedir. $37,707 > 9,488$ olduğundan dolayı H_0 hipotezi reddedilir ve H_1 hipotezi kabul edilir. Başka bir deyişle Friedman testine göre Germanwings veri setinde Lojistik harita kullanılarak OIO, GKO, KTOIO1, KTOIO2 ve KTOIO3 metasezgisel algoritmalarından alınan sonuçlar arasındaki fark istatistiksel olarak anlamlıdır. Ayrıca, 7 ayrı denetimli makine öğrenme algoritması ve 5 metasezgisel optimizasyon algoritmasının 4 farklı performans değerlendirme kriterine göre elde edilen sonuçları Tablo 7.33’de ve bu tabloya ait görsel ifade Şekil 7.8’de gösterilmektedir.

Tablo 7.33 incelendiğinde KTOIO1 optimizasyon algoritmasının Germanwings veri setinde deneylerde kullanılan tüm denetimli makine öğrenme algoritmalarından daha başarılı olduğu gözlemlenmiştir. Ayrıca denetimli makine öğrenme algoritmaları arasında RF metodu % 89 doğruluk oranı ile en yüksek başarıyı göstermiştir.

Tablo 7.33. Germanwings Crash veri setinde Lojistik harita kullanılarak elde edilen KTOIO algoritmaları, makine öğrenme algoritmaları, GKO ve OIO'nun karşılaştırılması

Algoritmalar	Doğruluk	Hassasiyet	Kesinlik	F-Ölçütü
Random Forest	0.890	0.857	0.811	0.833
Logistic Regression	0.688	0.679	0.810	0.809
Naive Bayes	0.789	0.541	0.769	0.635
KNN	0.761	0.617	0.750	0.690
Decision Tree	0.835	0.771	0.730	0.750
Gradient Boosted Tree	0.872	0.829	0.784	0.806
SVM	0.862	0.806	0.784	0.795
OIO	0.830	0.809	0.846	0.843
GKO	0.775	0.785	0.833	0.815
KTOIO1	0.923	0.911	0.950	0.928
KTOIO2	0.821	0.815	0.844	0.839
KTOIO3	0.859	0.845	0.934	0.888



Şekil 7.8. Denetimli makine öğrenme algoritmaları ve lojistik harita kullanılarak elde edilen metasezgisel optimizasyon algoritmalarının karşılaştırılması

8. SONUÇLAR

İnternet teknolojilerinin hızlı gelişimi ve yaygınlaşması, bu teknolojiyi kullanan insanların hayatlarında bir takım değişikliklere sebep olmuştur. Bu değişikliklerin başında bilgiye erişim yöntemleri gelmektedir. İnsanların alışlagelmiş bilgiye erişim yöntemleri olan televizyon, basılı medya, radyo gibi araçlar yerlerini sosyal medya yada sosyal ağ araçlarına bırakmışlardır. Sosyal medya araçlarının bilgiye erişim maliyetlerinin düşük olması, hemen hemen her yerden bu teknolojilere erişilebilir olması bunların hem yaygınlaşmasını sağlamış hemde cazip hale getirmiştir. Sosyal medya araçlarını sadece bilgiye erişim platformu olarak düşünmek doğru değildir bu araçlar aynı zamanda insanların eğlenceli vakit geçirdikleri ve sosyalleştiği sanal ortamlardır.

Bu avantajlarının yanında sosyal medya ve sosyal ağların insanlara sunduğu bilginin kalitesi geleneksel yöntemlerle kıyaslandığında oldukça düşük kalmaktadır. İçerdikleri bilginin doğru olup olmadığının belirlenmesi amacıyla genellikle bir kontrol mekanizmasının olmadığı bu ortamların insanları belirli bir doğrultuda yönlendirme, provoke etme, yanıltma, fikrini değiştirme ve aldatma gibi amaçlar içerdiği yadsınamaz bir gerçektir. Sosyal medyada sürekli olarak gerçek olmayan bilgilerle karşılaşan insanlar bir süre sonra sosyal medyada yer alan tüm haberlere kuşku ile yaklaşmaktadır. Bu durum insanların internette yer alan içeriklere olan güvenini sarsmaktadır. Böylelikle sosyal medyadaki aldatma ve söylenti barındıran içeriklerin insanlar üzerindeki olumsuz etkilerinden korunmak için bu içeriklerin tespit edilmesi önemli bir problem haline gelmiştir.

Aldatma ve Söylenti Tespiti problemleri ilk kez bu tez çalışmasında bir optimizasyon problemi olarak alınmış ve çözülmeye çalışılmıştır. Metasezgisel optimizasyon algoritmaları olan OIO ve GKO Aldatma ve Söylenti Tespiti problemleri için uyarlanarak başarımları sınanmıştır. Bunların yanında daha iyi sonuçlar alabilmek adına algoritmalar içerisinde yer alan rasgele sayılar yerine kaotik haritalar kullanılarak üretilmiş olan kaotik sayılar kullanılarak elde edilen üç yeni KTOIO1, KTOIO2 ve KTOIO3 algoritmaları ilk kez bu tez kapsamında önerilerek Aldatma ve Söylenti Tespiti problemleri için modellenmiştir. Deneyler esnasında beş farklı metasezgisel optimizasyon algoritmalarının başarımlarını sınamak amacıyla Deceptive Opinion Spam Corpus v1.4, Real-Life Trial, S1 ve Germanwings Crash olmak üzere gerçek dünya verilerini içeren dört farklı veri seti kullanılmıştır. En başta veri setlerindeki metinler ön işlem aşamalarından geçirilerek doküman terim matrisi elde edilmiştir. Sonra OIO, GKO, KTOIO1, KTOIO2 ve KTOIO3 metasezgisel optimizasyon algoritmalarının Aldatma ve Söylenti Tespiti problemlerine uyarlanabilmesi için bir uygunluk fonksiyonu belirlenerek veri setleri üzerinde aldatma ve söylenti barındıran içerik tespiti gerçekleştirilmiştir. Doğruluk, kesinlik, hassasiyet ve F-ölçütü olmak üzere dört adet performans değerlendirme kriteri için en iyi değeri elde etmeyi amaçlayan bir uygunluk fonksiyonu belirlenmiş ve bu uygunluk fonksiyonu ilk kez bu çalışmada Aldatma ve Söylenti

Tespiti problemleri için kullanılmıştır. Deneyler esnasında kullanılan uygunluk fonksiyonu yapısı gereği başka amaçlar ilave edilerek yada bazı amaçlar çıkarılarak kolaylıkla değiştirilebilecek esnek bir yapıya sahiptir. Kullanılan metasezgisel optimizasyon algoritmaları her veri seti için 30 defa bağımsız bir şekilde çalıştırılmıştır. Akabinde uygunluk fonksiyonu sonuc değerleri istatistiksel olarak yorumlamak amacıyla Friedman testi kullanılmıştır. Bu test marifetiyle metasezgisel optimizasyon algoritmalarından alınan sonuçlar arasında anlamlı bir fark olup olmadığı incelenmiştir. Metasezgisel optimizasyon algoritmalarına ilave olarak yedi farklı denetimli makine öğrenme metodu da Aldatma ve Söylenti Tespiti problemleri için çalıştırılmış ve elde edilen bütün değerler doğruluk, kesinlik, hassasiyet ve F-ölçütü performans değerlendirme kriterleri bakımından karşılaştırılmıştır.

Bu tez kapsamında ilk kez denenen Aldatma ve Söylenti Tespiti problemlerini, metasezgisel optimizasyon algoritmaları marifetiyle çözme yaklaşımı üretmiş oldukları başarılı sonuçlar bakımından oldukça umut vadetmektedir. Deneyler esnasında genel itibariyle gauss kaotik haritası ile üretilen KTOIO algoritmalarının diğer kaotik haritalar ile üretilmiş olan KTOIO algoritmalarından daha başarılı olduğu gözlemlenmiştir. Bunun yanında KTOIO1 algoritmasının diğer metasezgisel optimizasyon algoritmaları ve makine öğrenme metotlarından daha yüksek doğruluk oranlarına ulaştığı görülmüştür.

Bir problemde yüksek başarı gösteren bir metasezgisel optimizasyon algoritması doğası gereği tüm problemlerde en iyi sonucu veremeyeceğinden ötürü gelecekte Aldatma ve Söylenti Tespiti problemlerinin çözümünde daha yüksek başarıyı sağlamak adına farklı metasezgisel optimizasyon algoritmaları kullanılabilir. Deneylerde kullanılan algoritmaların adaptif ve hibrid versiyonları geliştirilerek başarımları arttırılabilir.

KAYNAKLAR

- [1] Ekiz, H., Vatansever, F., Zengin, A., Demir, Z. (2000). Hesaplamanın Tarihi ve Bilgisayarların Gelişimi. *Sakarya University Journal of Science*, 4(1-2), 73-82.
- [2] Hauben, M. (2007). *History of ARPANET*. Site de l'Instituto Superior de Engenharia do Porto, 17.
- [3] Mills, D. L., Braun, H. W. (1987). *The NSFNET backbone network*. In Proceedings of the ACM workshop on Frontiers in computer communications technology (pp. 191-196).
- [4] Scott, J. (1988). *Social network analysis*. Sociology, 22(1), 109-127.
- [5] Perrin, A. (2015). *Social media usage*. Pew research center, 125, 52-68.
- [6] Özmen, F., Aküzüm, C., Sünkür, M. (2012). Sosyal ağ sitelerinin eğitsel ortamlardaki işlevselliği. *Education Sciences*, 7(2), 496-506.
- [7] <https://datareportal.com/reports/digital-2021-global-overview-report>, Erişim: 01.02.2021
- [8] Hand, D. J., Adams, N. M. 2014. *Data mining*. Wiley StatsRef: Statistics Reference Online, 1-7.
- [9] Akpınar, H. (2000). Veri tabanlarında bilgi keşfi ve veri madenciliği. *İstanbul Üniversitesi İşletme Fakültesi Dergisi*, 29(1), 1-22.
- [10] Oğuzlar, A. (2003). Veri ön işleme. *Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, (21).
- [11] Roiger, R. J. 2017. *Data mining: a tutorial-based primer*.
- [12] Krishnaiah, V., Narsimha, G., Chandra, D. N. S. (2013). Diagnosis of lung cancer prediction system using data mining classification techniques. *International Journal of Computer Science and Information Technologies*, 4(1), 39-45.
- [13] Alataş, B. (2007). *Kaotik haritalı parçacık sürü optimizasyonu algoritmaları geliştirme*, Doktora Tezi, Fırat Üniversitesi, Fen Bilimleri Enstitüsü.
- [14] Cura, T. 2008. *Modern sezgisel teknikler ve uygulamaları*. Papatya Yayıncılık Eğitim.
- [15] Bingol, H., Alatas, B. (2015). Metasezgisel optimizasyon tekniklerine spor tabanlı yeni bir yaklaşım: lig şampiyonası algoritması. *Fırat Üniversitesi Fen Bilimleri Dergisi*, 27(1), pp. 1-11.
- [16] Karaboğa, D. 2011. *Yapay zekâ optimizasyon algoritmaları*. Nobel Yayın Dağıtım, Ankara
- [17] Kashan, A. H. (2015). A new metaheuristic for optimization: optics inspired optimization (OIO). *Computers & Operations Research*, 55, pp. 99-125.
- [18] Formato, R. A. (2007). Central force optimization. *Prog Electromagn Res*, 77, pp. 425-491.
- [19] Hsiao, Y. T., Chuang, C. L., Jiang, J. A., & Chien, C. C. (2005). A novel optimization algorithm: space gravitational optimization. In *2005 IEEE international conference on systems, man and cybernetics* (Vol. 3, pp. 2323-2328). IEEE.
- [20] Sacco, W. F., Oliveira, C. R. E. A. (2005). A new stochastic optimization algorithm based on a particle collision metaheuristic. *Proceedings of 6th WCSMO*.

- [21] Thammano, A., Moolwong, J. (2010). A new computational intelligence technique based on human group formation. *Expert Systems with Applications*, 37(2), pp. 1628-1634.
- [22] Ray, T., Liew, K. M. (2003). Society and civilization: an optimization algorithm based on the simulation of social behavior. *IEEE Transactions on Evolutionary Computation*, 7(4), pp. 386-396.
- [23] Coello, C. A. C., Becterra, R. L. (2003). *Evolutionary multiobjective optimization using a cultural algorithm. In Proceedings of the 2003 IEEE Swarm Intelligence Symposium. SIS'03 (Cat. No. 03EX706), pp. 6-13. IEEE.*
- [24] Geem, Z. W., Kim, J. H., Loganathan, G. V. (2001). A new heuristic optimization algorithm: harmony search. *simulation*, 76(2), pp. 60-68.
- [25] Mora-Gutiérrez, R.A., Ramírez-Rodríguez, J., Rincón-García, E.A. (2014). An optimization algorithm inspired by musical composition, *Artificial Intelligence Review*, 41(3), pp. 301-315.
- [26] Ashrafi, S.M., Dariane, A.B. (2011). *A novel and effective algorithm for numerical optimization: melody search (MS). 11th IEEE international conference on hybrid intelligent systems, Malaysia.*
- [27] Kennedy. J., Eberhart, R.C. (1995). *Particle swarm optimization, IEEE international conference on neural networks, Australia.*
- [28] Chu, S. C., Tsai, P. W., Pan, J. S. (2006). *Cat swarm optimization. In Pacific Rim international conference on artificial intelligence, pp. 854-858, Springer, Berlin, Heidelberg.*
- [29] Karaboga, D., Basturk, B. (2007). A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm. *Journal of Global Optimization*, 39(3), pp. 459-471.
- [30] Lam. A.Y., Li, V.O. (2010). Chemical-reaction-inspired metaheuristic for optimization, *IEEE Transactions on Evolutionary Computation*, 14(3), pp. 381-399.
- [31] Alatas, B. (2011). ACROA: artificial chemical reaction optimization algorithm for global optimization, *Expert Systems with Applications*, 38(10):13170–13180. doi: 10.1016/j.eswa.2011.04.126
- [32] Karci, A. (2012). *A new meta-heuristic algorithm based on chemical process: atom algorithm. In 1st International Eurasian Conference on Mathematical Sciences and Applications, pp. 3-7.*
- [33] Holland J.H., Goldberg, D.E. (1989). *Genetic algorithms in search, optimization and machine learning*, AddisonWesley Longman Publishing, Boston.
- [34] Storn, R., Price, K. (1997). Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces. *Journal of global optimization*, 11(4), pp. 341-359.
- [35] Dasgupta, D. (1993). An overview of artificial immune systems and their applications. *Artificial immune systems and their applications*, pp. 3-21.

- [36] Kashan, A.H. (2011). An efficient algorithm for constrained global optimization and application to mechanical engineering design: league championship algorithm (LCA), *Computer-Aided Design*, 43(12), pp. 1769-1792.
- [37] Osaba, E., Diaz, F., Onieva, E. (2014). Golden ball: a novel meta-heuristic to solve combinatorial optimization problems based on soccer concepts. *Applied Intelligence*, 41(1), pp. 145-166.
- [38] Razmjooy, N., Khalilpour, M., Ramezani, M. (2016). A new meta-heuristic optimization algorithm inspired by FIFA world cup competitions: theory and its application in PID designing for AVR system. *Journal of Control, Automation and Electrical Systems*, 27(4), pp. 419-440.
- [39] Mirjalili, S. (2016). SCA: a sine cosine algorithm for solving optimization problems, *Knowledge-Based Systems*, 96, pp. 120-133.
- [40] Ott, M., Choi, Y., Cardie, C., Hancock, J. T. (2011). Finding deceptive opinion spam by any stretch of the imagination. In Proceedings of the 49th annual meeting of the association for computational linguistics: Human language technologies-volume 1, pp. 309-319. *Association for Computational Linguistics*.
- [41] Feng, S., Banerjee, R., Choi, Y. (2012). Syntactic stylometry for deception detection. In Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers-Volume 2, pp. 171-175. *Association for Computational Linguistics*.
- [42] Song, L., Zhang, W., Liao, S. S., Kwok, R. C. W. (2012). A Critical Analysis of the State-of-The-Art on Automated Detection of Deceptive Behavior in Social Media. In *PACIS*, p. 168.
- [43] Shojaei, S., Murad, M. A. A., Azman, A. B., Sharef, N. M., Nadali, S. (2013). *Detecting deceptive reviews using lexical and syntactic features*. In *2013 13th International Conference on Intelligent Systems Design and Applications*, pp. 53-58. IEEE.
- [44] Merritts, R. A. (2013). *Online Deception Detection Using BDI Agents*.
- [45] Azaria, A., Richardson, A., Kraus, S. (2013). *Autonomous Agent for Deception Detection*, *Proc. of HAIDM (In AAMAS)*.
- [46] Feng, V. W., Hirst, G. (2013). *Detecting deceptive opinions with profile compatibility*. *Proceedings of the Sixth International Joint Conference on Natural Language Processing* (pp. 338-346).
- [47] Ott, M., Cardie, C., Hancock, J. T. (2013). *Negative deceptive opinion spam*. *Proceedings of the 2013 conference of the north American chapter of the association for computational linguistics: human language technologies*, pp. 497-501.
- [48] Tsikerdekis, M., Zeadally, S. (2014). Multiple account identity deception detection in social media using nonverbal behavior. *IEEE Transactions on Information Forensics and Security*, 9(8), pp. 1311-1321.
- [49] Li, J., Ott, M., Cardie, C., Hovy, E. (2014). Towards a general rule for identifying deceptive opinion spam. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, Vol. 1, pp. 1566-1576.

- [50] Li H., Liu, B., Mukherjee, A., Shao, J. (2014). Spotting fake reviews using positive-unlabeled learning. *Computación y Sistemas*, 18(3), pp. 467-475.
- [51] Alowibi, J. S., Buy, U. A., Philip, S. Y., Ghani, S., Mokbel, M. (2015). Deception detection in Twitter. *Social network analysis and mining*, 5(1), 32.
- [52] Conroy, N. J., Rubin, V. L., Chen, Y. (2015). Automatic deception detection: Methods for finding fake news. In *Proceedings of the 78th ASIS&T Annual Meeting: Information Science with Impact: Research in and for the Community* (p. 82). American Society for Information Science.
- [53] Rubin, V. L., Chen, Y., Conroy, N. J. (2015). Deception detection for news: three types of fakes. In *Proceedings of the 78th ASIS&T Annual Meeting: Information Science with Impact: Research in and for the Community* (p. 83). American Society for Information Science.
- [54] Pisarevskaya, D., Litvinova, T., Litvinova, O. (2017). Deception detection for the russian language: Lexical and syntactic parameters. In *Proceedings of the 1st Workshop on Natural Language Processing and Information Retrieval/RANLP*.
- [55] Kumari, R., Srivastava, S. K. (2017). Machine learning: A review on binary classification. *International Journal of Computer Applications*, 160(7).
- [56] Rubin, V. L. (2017). *Deception detection and rumor debunking for social media*. In The SAGE Handbook of Social Media Research Methods (p. 342). SAGE.
- [57] Molla, A., Biadgie, Y., Sohn, K. A. (2017). *Detecting Negative Deceptive Opinion from Tweets*. In *International Conference on Mobile and Wireless Technology*, pp. 329-339, Singapore.
- [58] Rastogi, A., Mehrotra, M., (2017). Opinion spam detection in online reviews, *Journal of Information & Knowledge Management*, vol. 16, no. 04, pp. 38.
- [59] Rosso, P., Cagnina, L. C., (2017). Deception Detection and Opinion Spam, *A practical guide to sentiment analysis*, pp. 155-171, Springer.
- [60] Kumar, N., Venugopal, D., Qiu, L., Kumar, S. (2018). Detecting review manipulation on online platforms with hierarchical supervised learning. *Journal of Management Information Systems*, 35(1), pp. 350-380.
- [61] Rangel, F., Rosso, P., Charfi, A., Zaghouani, W., Ghanem, B., Snchez-Junquera, J. (2019). Overview of the track on author profiling and deception detection in Arabic. In *Working Notes of the Forum for Information Retrieval Evaluation*. CEUR Workshop Proceedings. In: CEUR-WS.org, Kolkata, India.
- [62] Nayel, H. A. (2019). NAYEL@APDA: Machine Learning Approach for Author Profiling and Deception Detection in Arabic Texts. In *Working Notes of the Forum for Information Retrieval Evaluation*. CEUR Workshop Proceedings, Kolkata, India, pp. 12-15.
- [63] Sen, U. M., Perez-Rosas, V., Yanikoglu, B., Abouelenien, M., Burzo, M., & Mihalcea, R. (2020). Multimodal Deception Detection using Real-Life Trial Data. *IEEE Transactions on Affective Computing*.

- [64] Ceballos Delgado, A. A., Glisson, W., Shashidhar, N., McDonald, J., Grispos, G., Benton, R. (2021). *Deception Detection Using Machine Learning*. In *Proceedings of the 54th Hawaii International Conference on System Sciences* (p. 7122). Hawaii.
- [65] Krishnaveni, N., Radha, V. (2021). Performance Evaluation of Clustering-Based Classification Algorithms for Detection of Online Spam Reviews. In *Data Intelligence and Cognitive Informatics*, pp. 255-266. Springer, Singapore.
- [66] Kesarwani, A., Chauhan, S. S., Nair, A. R., Verma, G. (2021). Supervised Machine Learning Algorithms for Fake News Detection. In *Advances in Communication and Computational Technology* (pp. 767-778). Springer, Singapore.
- [67] Sternglanz, R. W., Morris, W. L., Morrow, M., Braverman, J. (2019). *A review of meta-analyses about deception detection*. The Palgrave handbook of deceptive communication, pp. 303-326
- [68] Rill-García, R., Jair Escalante, H., Villasenor-Pineda, L., Reyes-Meza, V. (2019). *High-level features for multimodal deception detection in videos*. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*.
- [69] Masip, J. (2017). Deception detection: State of the art and future prospects. *Psicothema*, 29(2), pp. 149-159.
- [70] Litvinova, O., Seredin, P., Litvinova, T., Lyell, J. (2017). *Deception detection in russian texts*. In *Proceedings of the Student Research Workshop at the 15th Conference of the European Chapter of the Association for Computational Linguistics*, pp. 43-52.
- [71] Ding, M., Zhao, A., Lu, Z., Xiang, T., Wen, J. R. (2019). *Face-focused cross-stream network for deception detection in videos*. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 7802-7811.
- [72] Kleinberg, B., Arntz, A., Verschuere, B. (2019). Being accurate about accuracy in verbal deception detection. *PloS one*, 14(8), e0220228.
- [73] Van der Walt, E., Eloff, J. H., Grobler, J. (2018). Cyber-security: Identity deception detection on social media platforms. *Computers & Security*, 78, 76-89.
- [74] Krishnamurthy, G., Majumder, N., Poria, S., Cambria, E. (2018). A deep learning approach for multimodal deception detection. *arXiv preprint arXiv:1803.00344*.
- [75] Solmaz, B. (2003). *Söylentilerin kurumsal iletişim açısından değerlendirilmesi ve bir uygulama örneği*, Doktora Tezi, Anadolu Üniversitesi, Sosyal Bilimler Enstitüsü.
- [76] Yang, F., Liu, Y., Yu, X., Yang, M. (2012). Automatic detection of rumor on sina weibo. In *Proceedings of the ACM SIGKDD workshop on mining data semantics* (pp. 1-7).
- [77] Gupta, A., Lamba, H., Kumaraguru, P. (2013). \$1.00 per rt# bostonmarathon# prayforboston: Analyzing fake content on twitter. In *2013 APWG eCrime researchers summit*, pp. 1-12. IEEE.
- [78] Aggrawal, N. (2018). Detection of offensive Tweets: A comparative study. *Computer Reviews Journal*, 1(1), pp. 75-89.

- [79] Jin, Z., Cao, J., Guo, H., Zhang, Y., Wang, Y., Luo, J. (2017). *Detection and analysis of 2016 us presidential election related rumors on twitter. In International conference on social computing, behavioral-cultural modeling and prediction and behavior representation in modeling and simulation (pp. 14-24).* Springer, Cham.
- [80] Hamidian, S., Diab, M. (2016). Rumor identification and belief investigation on twitter. *In Proceedings of the 7th Workshop on computational approaches to subjectivity, sentiment and social media analysis*, pp. 3-8.
- [81] Takahashi, T., Igata, N. (2012). *Rumor detection on twitter. In The 6th International Conference on Soft Computing and Intelligent Systems, and The 13th International Symposium on Advanced Intelligence Systems*, pp. 452-457. IEEE.
- [82] Shirai, T., Sakaki, T., Toriumi, F., Shinoda, K., Kazama, K., Noda, I., Kurihara, S. (2012). *Estimation of false rumor diffusion model and estimation of prevention model of false rumor diffusion on twitter. In Proc. the 26th Annual Conference of the Japanese Society for Artificial Intelligence*, pp. 1-4.
- [83] Bai, N., Meng, F., Rui, X., Wang, Z. (2021). Rumour Detection based on Graph Convolutional Neural Net. *IEEE Access*.
- [84] Zhang, L., Li, J., Zhou, B., Jia, Y. (2021). Rumor Detection Based on SAGNN: Simplified Aggregation Graph Neural Networks. *Machine Learning and Knowledge Extraction*, 3(1), 84-94.
- [85] Providel, E., Mendoza, M. (2020). Using Deep Learning to Detect Rumors in Twitter. *In International Conference on Human-Computer Interaction (pp. 321-334).* Springer, Cham.
- [86] Huang, Q., Zhou, C., Wu, J., Wang, M., Wang, B. (2019). *Deep structure learning for rumor detection on twitter. In 2019 International Joint Conference on Neural Networks (IJCNN) (pp. 1-8).* IEEE.
- [87] Alatas, B., Bingol, H. (2020). Comparative assessment of light-based intelligent search and optimization algorithms. *Light & Engineering*, 28(6).
- [88] Kashan, A. H. (2015). An effective algorithm for constrained optimization based on optics inspired optimization (OIO). *Computer-Aided Design*, 63, 52-71.
- [89] Alatas, B., Bingol, H. (2019). A physics based novel approach for travelling tournament problem: optics inspired optimization. *Information Technology and Control*, 48(3), 373-388.
- [90] Mirjalili, S., Mirjalili, S. M., Lewis, A. (2014). Grey wolf optimizer. *Advances in engineering software*, 69, 46-61.
- [91] Mech, L. D. (1999). Alpha status, dominance, and division of labor in wolf packs. *Canadian Journal of Zoology*, 77(8), 1196-1203.
- [92] Muro, C., Escobedo, R., Spector, L., Coppinger, R. P. (2011). Wolf-pack (*Canis lupus*) hunting strategies emerge from simple rules in computational simulations. *Behavioural processes*, 88(3), 192-197.
- [93] Aydın, S., Özkul, A. E. (2015). Veri madenciliği ve Anadolu Üniversitesi açık öğretim sisteminde bir uygulama. *Journal of Research in Education and Teaching*, 4(3), 36-44.

- [94] Shree, S. B., Sheshadri, H. S. (2014). *An initial investigation in the diagnosis of Alzheimer's disease using various classification techniques*, 2014 IEEE International Conference on Computational Intelligence and Computing Research, Tamilnadu, India.
- [95] Pal, M. (2005). Random forest classifier for remote sensing classification. *International journal of remote sensing*, 26(1), pp. 217-222.
- [96] Freund, Y., Schapire, R. E. (1996). Experiments with a new boosting algorithm. *In icml* (Vol. 96), pp. 148-156.
- [97] Breiman, L. (1996). Bagging predictors. *Machine learning*, 24(2), pp. 123-140.
- [98] Özkan, E. (2017). *Bulanık karar ağaçları ile hisse senedi fiyat tahmini*, Yüksek Lisans Tezi, Ankara Üniversitesi, Fen Bilimleri Enstitüsü.
- [99] Dreiseitl, S., Ohno-Machado, L. (2002). Logistic regression and artificial neural network classification models: a methodology review. *Journal of biomedical informatics*, 35(5-6), 352-359.
- [100] Haykin, S., Network, N. (2004). A comprehensive foundation. *Neural networks*, 2, 41.
- [101] Ataseven, B. (2013). Yapay sinir ağları ile öngörü modellemesi. *Öneri Dergisi*, 10(39), pp. 101-115.
- [102] Mitchell, T., McGraw-Hill, M. L. (1997). *Machine Learning*. Edition.
- [103] Han, J., Kamber, M., Pei, J. (2011). Data mining concepts and techniques third edition. *The Morgan Kaufmann Series in Data Management Systems*, 5(4), pp. 83-124.
- [104] Cheng, C. H., Chen, H. H. (2019). Sentimental text mining based on an additional features method for text classification. *PloS one*, 14(6), e0217591.
- [105] Salton, G., Wong A., Yang, C.S. (1975). A vector space model for automatic indexing, *Communications of the ACM*, 18(11), 613–620. doi: 10.1145/361219.361220
- [106] Çalış, K., Gazdağı, O., Yıldız, O. (2013). Reklam İçerikli Epostaların Metin Madenciliği Yöntemleri ile Otomatik Tespiti. *Bilişim Teknolojileri Dergisi*, 6(1), pp. 1-7.
- [107] Pérez-Rosas, V., Abouelenien, M., Mihalcea, R., Burzo, M. (2015). *Deception detection using real-life trial data*. *In Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*, pp. 59-66.
- [108] Kwon, S., Cha, M., Jung, K. (2017). Rumor detection over varying time windows. *PloS one*, 12(1), e0168344.
- [109] Zubiaga, A., Liakata, M., Procter, R. (2016). Learning reporting dynamics during breaking news for rumour detection in social media. *arXiv preprint arXiv:1610.07363*.
- [110] Alatas, B., Akin, E., Ozer, A. B. (2009). Chaos embedded particle swarm optimization algorithms. *Chaos, Solitons & Fractals*, 40(4), pp. 1715-1734.
- [111] Brannstrom, A. (2004). *Modelling Animal Populations*, Umea University, PhD Thesis.
- [112] Arena, P., Caponetto, R., Fortuna, L., Rizzo, A., La Rosa, M. (2000). Self-organization in nonrecurrent complex systems. *International Journal of Bifurcation and Chaos*, 10(05), pp. 1115-1125.

- [113] Behnia, S., Ahadpour, S., Ayubi, P. (2014). Design and implementation of coupled chaotic maps in watermarking. *Applied Soft Computing*, 21, pp. 481-490.
- [114] Alatas, B. (2010). Chaotic bee colony algorithms for global numerical optimization. *Expert Systems with Applications*, 37(8), pp. 5682-5687.
- [115] Altay, E. V., Alatas, B. (2020). Bird swarm algorithms with chaotic mapping. *Artificial Intelligence Review*, 53, pp. 1373-1414.
- [116] Schuster, H. G., Just, W. (2006). *Deterministic chaos: an introduction*. John Wiley & Sons.
- [117] Peitgen, H. O., Jürgens, H., Saupe, D. (2006). *Chaos and fractals: new frontiers of science*. Springer Science & Business Media.
- [118] Zheng, W. M. (1994). Kneading plane of the circle map. *Chaos, Solitons & Fractals*, 4(7), pp. 1221-1233.
- [119] May, R. M. (2004). Simple mathematical models with very complicated dynamics. *The Theory of Chaotic Attractors*, pp. 85-93.
- [120] Bingol, H., Alatas, B. (2020). Chaos based optics inspired optimization algorithms as global solution search approach. *Chaos, Solitons & Fractals*, 141, 110434.
- [121] Digalakis, J. G., Margaritis, K. G. (2002). An experimental study of benchmarking functions for genetic algorithms. *International Journal of Computer Mathematics*, 79(4), pp. 403-416.
- [122] Dieterich, J. M., Hartke, B. (2012). Empirical review of standard benchmark functions using evolutionary global optimization. *arXiv preprint arXiv:1207.4318*.
- [123] Zhu, G., Kwong, S. (2010). Gbest-guided artificial bee colony algorithm for numerical function optimization. *Applied mathematics and computation*, 217(7), pp. 3166-3173.
- [124] Shah, H., Herawan, T., Naseem, R., Ghazali, R. (2014). *Hybrid guided artificial bee colony algorithm for numerical function optimization*. In *International Conference in Swarm Intelligence*, pp. 197-206. Springer, Cham.
- [125] Homaifar, A., Qi, C. X., Lai, S. H. (1994). Constrained optimization via genetic algorithms. *Simulation*, 62(4), pp. 242-253.
- [126] Belegundu, A. D., Arora, J. S. (1985). A study of mathematical programming methods for structural optimization. Part I: Theory. *International Journal for Numerical Methods in Engineering*, 21(9), 1583-1599.
- [127] Ragsdell, K. M., Phillips, D. T. (1976). *Optimal design of a class of welded structures using geometric programming*. pp. 1021-1025.
- [128] Wang, S., Qi, H., Kong, L., Nu, C. (2013). *Combination of VSM and Jaccard coefficient for external plagiarism detection*. In *2013 International Conference on Machine Learning and Cybernetics, Vol. 4*, pp. 1880-1885. IEEE.

ÖZGEÇMİŞ

Harun BİNGÖL

ARAŞTIRMA DENEYİMİ

- ✓ C#
- ✓ Matlab
- ✓ Java

İŞ DENEYİMİ

AKADEMİK FAALİYETLER

Makaleler:

1. Bilal ALATAS, Harun BİNGÖL, (2020). Comparative Assessment of Light-Based Intelligent Search and Optimization Algorithms, *Light & Engineering*, 28(6).
2. Harun BİNGÖL, Bilal ALATAS, (2020). Chaos Based Optics Inspired Optimization Algorithms as Global Solution Search Approach, *Chaos, Solution & Fractals*, 141, 110434.

3. Bilal ALATAS, Harun BINGOL, (2019). A Physics-Based Novel Approach for Travelling Tournament Problem: Optics Inspired Optimization, *Information Technology and Control*, 48 (3), 373-388.
4. Harun BINGOL, Bilal ALATAS, (2016). Chaotic League Championship Algorithms, *Arabian Journal for Science and Engineering*, DOI 10.1007/s13369-016-2200-9 ,(41), 5123–5147, 2016.
5. Harun BINGOL, Bilal ALATAS, (2017). Investigating Performance of Intelligent Sine Cosine Algorithm within Unconstrained Unimodal and Multimodal Search And Optimization Problems”, *International Journal of Information Processing Sciences and Applications*, 1(1), 11-22.
6. Harun BINGOL, Bilal ALATAS, (2015). Metasezgisel Optimizasyon Tekniklerine Spor Tabanlı Yeni Bir Yaklaşım: Lig Şampiyonası Algoritması, *Fırat Üniv. Fen Bilimleri Dergisi*, 27(1), 1-11.
7. Harun BINGOL, Mehmet Mehdi KARAKOC, (2015). İnternet Bankacılığı ve Telekomünikasyon Alanında Siber Güvenliğe Genel Bakış, *Bingöl Üniv. Türk Doğa ve Fen Dergisi*, 4(1), 23-35.
8. Harun BINGOL, Muhammed YILDIRIM, (2021). Global Optimizasyon için Sürü Tabanlı Bir Yaklaşım: Salp Sürü Algoritması, *Fırat Üniv. Fen Bilimleri Dergisi*.
9. Harun BINGOL, Bilal ALATAS, (2021). Classification of Brain Tumor Images using Deep Learning Methods, *Fırat Univ. Turkish Journal of Science and Technology*.

Bildiriler:

1. Harun BINGOL, Bilal ALATAS, (2019). *Rumor Detection in Social Media Using Machine Learning Methods*, UBMKY, IEEE, Ankara.
2. Ulas Baran BALOGLU, Bilal ALATAS, Harun BINGOL, (2019). *Assessment of Supervised Learning Algorithms on Irony Detection within Online Social Media*, UBMKY, IEEE, Ankara.