



**KUANTUM TÜREVİN YAPAY SİNİR AĞI ÖĞRENME
ALGORİTMALARINA ENTEGRASYONU VE
MONOTON KAPSAMA PROBLEMİNİN ÇÖZÜMÜ
İÇİN YENİ ALGORİTMALAR**

İrfan IŞIK

Doktora Tezi

Matematik Anabilim Dalı

Danışman: Doç. Dr. İbrahim KARAHAN

Eş Danışman: Doç. Dr. Okan ERKAYMAZ

2025

Her hakkı saklıdır.



FEN BİLİMLERİ ENSTİTÜSÜ
DOKTORA TEZİ

KUANTUM TÜREVİN YAPAY SİNİR AĞI ÖĞRENME ALGORİTMALARINA
ENTEGRASYONU VE MONOTON KAPSAMA PROBLEMİNİN ÇÖZÜMÜ
İÇİN YENİ ALGORİTMALAR

İrfan IŞIK

Tez Danışmanı: Doç. Dr. İbrahim KARAHAN

Eş Danışman: Doç. Dr. Okan ERKAYMAZ

Anabilim Dalı: Matematik

Erzurum

2025

Her hakkı saklıdır

T.C.
ERZURUM TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
TEZ ONAY FORMU

**KUANTUM TÜREVİN YAPAY SİNİR AĞI ÖĞRENME ALGORİTMALARINA
ENTEGRASYONU VE MONOTON KAPSAMA PROBLEMİNİN ÇÖZÜMÜ
İÇİN YENİ ALGORİTMALAR**

Doç. Dr. İbrahim KARAHAN danışmanlığında, İrfan IŞIK tarafından hazırlanan bu çalışma 15 / 09 / 2025 tarihinde aşağıdaki jüri tarafından Matematik Ana Bilim Dalı'nda Doktora tezi olarak **Oy birliği ile (5/5)** kabul edilmiştir.

Başkan	: Prof. Dr. Murat ÖZDEMİR	İmza	:
Üye	: Prof. Dr. Ceren Sultan ELMALI	İmza	:
Üye	: Prof. Dr. Murat ÇAĞLAR	İmza	:
Üye	: Doç. Dr. İbrahim KARAHAN	İmza	:
Üye	: Doç. Dr. Okan ERKAYMAZ	İmza	:
Üye	: Dr. Öğr. Üyesi Sait ALP	İmza	:

Yukarıdaki sonucu onaylıyorum

Doç. Dr. Bünyamin ÖZGERİŞ
Enstitü Müdürü

Bu tez çalışması tarafından nolu proje ile desteklenmiştir.

ÜRETKEN YAPAY ZEKÂ VE ETİK KURALLARA UYGUNLUK BEYANI

Erzurum Teknik Üniversitesi Fen Bilimleri Enstitüsü tez yazım kurallarına uygun olarak hazırladığım bu tez içindeki tüm bilgilerin doğru ve tam olduğunu, bilgilerin üretilmesi aşamasında bilimsel etik ilkelere uygun davrandığımı ve tezin hazırlanması sürecinde üretken yapay zekâ programlarından, içeriğin ve sonuçların özgünlük sorumluluğunun tarafıma ait olması ve çalışmanın bilimsel niteliğini etkilemeyecek düzeyde, yalnızca yardımcı bir araç olarak sınırlı düzeyde destek alındığımı beyan ederim.

15 / 09 / 2025

İmzası

İrfan IŞIK

ÖZET

DOKTORA TEZİ

KUANTUM TÜREVİN YAPAY SİNİR AĞI ÖĞRENME ALGORİTMALARINA ENTEGRASYONU VE MONOTON KAPSAMA PROBLEMİNİN ÇÖZÜMÜ İÇİN YENİ ALGORİTMALAR

İrfan IŞIK

Erzurum Teknik Üniversitesi

Fen Bilimleri Enstitüsü

Matematik Ana Bilim Dalı

Danışman: Doç. Dr. İbrahim KARAHAN

Bu tezde öncelikle, öğrenme sürecini daha esnek ve uyumlu hâle getiren q -türev kavramı yapay sinir ağlarıyla bir araya getirilmiş ve literatürde yalnızca aktivasyon fonksiyonları düzeyinde ele alınan q -türevin, bu çalışmada geri yayılım temelli algoritmalarda (geri yayılım, stokastik geri yayılım, esnek geri yayılım) türev yerine kullanılmasıyla yeni yaklaşımlar geliştirilmiştir. Geliştirilen bu algoritmalar sınıflandırma problemleri üzerinde Digits, MNIST ve Iris veri setleri kullanılarak klasik versiyonları ile karşılaştırılmıştır. Bu karşılaştırma neticesinde q -türev temelli algoritmalar ile doğruluk, kesinlik, duyarlılık ve $F1$ skorunda klasik yaklaşımlara göre daha başarılı sonuçlar elde edilmiştir. Diğer taraftan reel Hilbert uzayında tanımlı monoton kapsama problemlerinin çözümüne yönelik yeni iki algoritma geliştirilmiş; mevcut ileri-geri ayrıştırma algoritmalarına çift eylemsizlik ve viskozite terimleri eklenerek daha genel bir yapı tasarlanmış ve bu algoritmaların yakınsaklık durumları ifade edilerek ispatlanmıştır. Geliştirilen her iki yaklaşım da California Housing, Diabetes, Iris, Digits, MNIST, Anacalt, Delta Ail, Machine CPU, BCWD ve Fashion MNIST veri kümeleri üzerinde regresyon ve sınıflandırma problemlerine uygulanmıştır. Elde edilen sonuçlar önerilen yöntemlerin kararlı yakınsama sağladığını, yüksek doğruluk elde ettiğini ve mevcut yöntemlerle rekabet edebilecek düzeyde performans sergilediğini göstermiştir.

2025, 131 sayfa

Anahtar Kelimeler: q -türev, Yapay sinir ağları, Monoton kapsama problemleri, Regresyon, Veri sınıflandırma.

ABSTRACT

Ph.D

INTEGRATION OF QUANTUM DERIVATIVE INTO ARTIFICIAL NEURAL NETWORK LEARNING ALGORITHMS AND NEW ALGORITHMS FOR SOLVING THE MONOTONE INCLUSION PROBLEM

İrfan IŞIK

Erzurum Technical University
Graduate School of Natural and Applied Sciences
Department of Maths

Supervisor: Assoc. Prof. Dr. İbrahim KARAHAN

In this thesis, the concept of the q -derivative, which renders the learning process more flexible and adaptable, is integrated with artificial neural networks. While the q -derivative has previously been considered in the literature only at the level of activation functions, this study introduces novel approaches by employing it in place of the ordinary derivative within backpropagation-based algorithms (backpropagation, stochastic backpropagation, and resilient backpropagation). The proposed algorithms were evaluated on classification problems using the Digits, MNIST, and Iris datasets, and were compared against their classical counterparts. The results demonstrate that the q -derivative-based algorithms achieved superior performance in terms of accuracy, precision, recall, and $F1$ -score compared to traditional methods. On the other hand, two new algorithms were developed to address monotone inclusion problems defined in real Hilbert spaces. By incorporating double inertial and viscosity terms into the existing forward–backward splitting algorithms, a more general framework was constructed, and the convergence properties of these algorithms were formally established and proven. Both approaches were applied to regression and classification tasks using the California Housing, Diabetes, Iris, Digits, MNIST, Anacalt, Delta Ail, Machine CPU, BCWD, and Fashion MNIST datasets. The findings indicate that the proposed methods ensure stable convergence, attain high accuracy, and exhibit performance competitive with existing techniques.

2025, 131 page

Keywords: q -derivative, Neural networks, Monotone inclusion problems, Regression, Data classification.

TEŞEKKÜR

Bu tezin ortaya çıkmasında ve zorlu doktora eğitimim süresince desteklerini benden esirgemeyen, hayatıma dokunan tüm değerli insanlara en içten teşekkürlerimi sunmayı bir borç bilirim.

Bu uzun ve meşakkatli yolculuğun her aşamasında, engin bilgisi, yol göstericiliği ve yapıcı eleştirileriyle araştırmama yön veren, akademik gelişimimde en büyük pay sahibi olan değerli danışmanım Sayın Doç. Dr. İbrahim KARAHAN hocama sonsuz teşekkür ederim. Kendisinin sabrı, teşviki ve bilimsel vizyonu, bu çalışmanın tamamlanmasındaki en önemli itici güç olmuştur.

Araştırmam boyunca karşılaştığım her sorunda bana destek olan, sabrı ve yol göstericiliğiyle güven veren değerli eş danışmanım Doç. Dr. Okan ERKAYMAZ hocama içten teşekkürlerimi iletmek isterim.

Doktora eğitimim boyunca derslerini alma fırsatı bulduğum, bilgi ve tecrübelerinden faydalandığım Erzurum Teknik Üniversitesi ve Atatürk Üniversitesi Fen Fakültesi Matematik Bölümünün saygıdeğer hocalarına ve tez izleme komitemde yer alarak değerli geri bildirimleriyle çalışmama katkıda bulunan hocalarıma en içten şükranlarımı sunarım.

Eğitim hayatım boyunca mensubu olmaktan gurur duyduğum Erzurum Teknik Üniversitesi ve Atatürk Üniversitesi'ne, bana sundukları bilimsel ortam ve olanaklar için teşekkür ederim.

Bu süreçte bana her türlü esnekliği ve desteği sağlayan, çalıştığım kurum olan Milli Savunma Üniversitesi yönetimine ve değerli mesai arkadaşlarıma, gösterdikleri anlayış için minnettarım.

Bu yolculuğun en büyük kahramanları ise şüphesiz ailemdir. Her anımda yanımda olan, sevgisi, sabrı ve fedakârlığıyla bana güç veren, en büyük destekçim sevgili eşim Dilara'ya ne kadar teşekkür etsem azdır. Varlığı, bu zorlu yolda en büyük motivasyon

kaynađım olmuştur. Hayatıma neşe katan, varlığıyla bana her şeyi unutturan canım ođlum İlteriş Baybars’a, ona ayırmam gereken zamanlarda gösterdiği anlayış için minnettarım.

Dođduğum günden bu yana maddi ve manevi desteklerini hiçbir zaman esirgemeyen, bana çalışma azmi, sabrı ve hayat boyu yol gösterecek değeri aşılayan canım annem Elif ve babam Ali’ye en derin şükranlarımı sunarım. Varlıkları, sevgileri, özverileri ve dualarıyla her adımda yanımda oldular. Onların desteđi ve inancı olmasaydı bu başarı mümkün olmazdı.

Sevgi dolu kalpleriyle bana güç veren, şefkatleriyle ruhumu saran, en hassas anlarımda yanımda olup beni anlayan, ince ruhları ve moral destekleriyle her zaman arkamda duran kıymetli kardeşlerim Semanur ve Furkan’a da kalpten teşekkür ederim. Sizlerin varlığı, bana yalnız olmadığımı hissettirdi ve en zorlu süreçlerde dahi umutla yoluma devam etmemi sağladı. İyi ki varsınız, iyi ki benim kardeşlersiniz.

Ve son olarak, bu uzun süreçte bir telefon kadar uzağımda olan; kahkahalarımızı ve hüzünlerimizi paylaştığımız, varlıklarıyla hayatımı zenginleştiren can dostlarıma... Desteđiniz ve dostluđunuz için her birinize ayrı ayrı teşekkür ederim.

Bu tez, adını andığım ve anamadığım tüm bu değeri insanların destek ve katkılarıyla tamamlanmıştır. Hepsine ayrı ayrı sonsuz teşekkürler.

Bu tez, sevgili eşim ve canım ođluma ithaf edilmiştir...

İrfan IŞIK
Eylül 2025

İÇİNDEKİLER

	<u>Sayfa</u>
ÖZET	i
ABSTRACT	ii
TEŞEKKÜR	iii
İÇİNDEKİLER	iv
SİMGELER VE KISALTMALAR DİZİNİ	v
ŞEKİLLER DİZİNİ	vi
ÇİZELGELER DİZİNİ	viii
1. GİRİŞ	1
2. KAYNAK ÖZETLERİ	7
3. KURAMSAL TEMELLER	22
3.1. Temel Tanım ve Kavramlar	22
3.2. Monoton Kapsama Problemi	30
3.3. Quantum Türev (q -Türev)	34
3.4. Yapay Sinir Ağlarının Temelleri	41
3.5. Yapay Sinir Ağları Algoritmaları	50
3.5.1. Geri yayılım	50
3.5.2. Stokastik geri yayılım	53
3.5.3. Esnek geri yayılım	56
3.5.4. Ekstrem öğrenme makineleri	58
4. İKİ EYLEMSİZLİK TERİMLİ İLERİ-GERİ AYRIŞTIRMA ALGORİTMALARI	61
4.1. Algoritmaların Teorisi ve İspatları	61
4.2. Algoritmaların Bilgisayar Programı İle Benzetimi	84
5. q-TÜREV VE YAPAY SİNİR AĞLARI UYGULAMALARI	112
5.1. q -Geri Yayılım Algoritması	112
5.2. q -Stokastik Geri Yayılım Algoritması	118
5.3. q -Esnek Geri Yayılım Algoritması	120
6. SONUÇ ve ÖNERİLER	124
KAYNAKLAR	126

SİMGELER VE KISALTMALAR DİZİNİ

<u>Simgeler</u>	<u>Açıklama</u>
J_λ^A	A operatörünün γ indeksli resolvent operatörü
Id	Birim operatör
∂f	f fonksiyonunun alt diferansiyeli
$d_q f(x)$	f fonksiyonunun q -diferansiyeli
$D_q f(x)$	f fonksiyonunun q -türevi
$prox_f(\rho)$	f fonksiyonunun ρ noktasındaki proksimal operatörü
$R(T)$	Görüntü kümesi
∇f	Gradyan operatör
$\mathcal{H}$	Hilbert uzay
$\langle \cdot, \cdot \rangle$	İç çarpım
$\ \cdot \ $	Norm
$\mathbb{R}$	Reel uzay
$D(T)$	Tanım kümesi
$Gph(T)$	T dönüşümünün grafiği
$F(T)$	T dönüşümünün sabit noktalarının kümesi
V	Vektör uzayı
$P_C \rho$	ρ 'nun C üzerine metrik izdüşümü

Kısaltmalar

ELM	Ekstrem Öğrenme Makineleri
Rprop	Esnek Geri Yayılım
BP	Gerilim Yayılım
FBS	İleri-Geri Ayrıştırma
SBP	Stokastik Geri Yayılım
SGD	Stokastik Gradyan İniş
YSA	Yapay Sinir Ağı

ŞEKİLLER DİZİNİ

Şekil 3.1. Temsili biyolojik nöron	42
Şekil 3.2. İki gizli katmanlı bir yapay sinir ağı modeli (Uğuz 2019)	44
Şekil 3.3. Koordinat düzleminde XOR (Soruç 2020)	47
Şekil 3.4. Ağırlıkların güncellenmesi (Uğuz 2019)	52
Şekil 3.5. Varyasyonel oto kodlayıcı mimarisi (Anonim 2025)	55
Şekil 4.1. 10 sinüs örneği için lineer aktivasyonlu regresyon.....	86
Şekil 4.2. 100 sinüs örneği için lineer aktivasyonlu regresyon	87
Şekil 4.3. 10 sinüs örneği için sigmoid aktivasyonlu regresyon.....	88
Şekil 4.4. 100 sinüs örneği için sigmoid aktivasyonlu regresyon.....	89
Şekil 4.5. California Housing veri seti: 100 rastgele örnekte regresyon sonuçları.....	90
Şekil 4.6. California Housing veri seti: Algoritmalarla göre hata dağılımı	91
Şekil 4.7. Diabetes veri seti: 100 rastgele örnekte regresyon sonuçları.....	92
Şekil 4.8. Diabetes veri seti: Algoritmalarla göre hata dağılımı	92
Şekil 4.9. Iris veri seti: Algoritma 4.1 için karmaşıklık matrisi.....	94
Şekil 4.10. Digits veri seti: DIPFBSM algoritması için karmaşıklık matrisi.....	95
Şekil 4.11. Digits veri seti: FISTA algoritması için karmaşıklık matrisi.....	96
Şekil 4.12. Digits veri seti: Algoritma 4.1 için karmaşıklık matrisi	96
Şekil 4.13. MNIST veri seti: DIPFBSM algoritması için karmaşıklık matrisi.....	97
Şekil 4.14. MNIST veri seti: FISTA algoritması için karmaşıklık matrisi.....	98
Şekil 4.15. MNIST veri seti: Algoritma 4.1 için karmaşıklık matrisi.....	99
Şekil 4.16. Anacalt veri seti: 100 rastgele örnekte regresyon sonuçları	100
Şekil 4.17. Anacalt veri seti: Algoritmalarla göre hata dağılımı	101
Şekil 4.18. Delta Ail veri seti: 100 rastgele örnekte regresyon sonuçları.....	102
Şekil 4.19. Delta Ail veri seti: Algoritmalarla göre hata dağılımı	102
Şekil 4.20. Machine CPU veri seti: 100 rastgele örnekte regresyon sonuçları.....	103
Şekil 4.21. Machine CPU veri seti: Algoritmalarla göre hata dağılımı	104
Şekil 4.22. California Housing veri seti: 100 rastgele örnekte regresyon sonuçları.....	105
Şekil 4.23. California Housing veri seti: Algoritmalarla göre hata dağılımı	105
Şekil 4.24. BCWD veri seti: Algoritma 4.1 için karmaşıklık matrisi	107
Şekil 4.25. BCWD veri seti: IVFBS algoritması için karmaşıklık matrisi	107
Şekil 4.26. BCWD veri seti: DIVFBS algoritması için karmaşıklık matrisi	108

Şekil 4.27. Fashion MNIST veri seti: Algoritma 4.1 için karmaşıklık matrisi.....	109
Şekil 4.28. Fashion MNIST veri seti: IVFBS algoritması için karmaşıklık matrisi	110
Şekil 4.29. Fashion MNIST veri seti: DIVFBS algoritması için karmaşıklık matrisi ..	110
Şekil 5.1. Digits veri seti: Geri yayılım algoritması için karmaşıklık matrisi	114
Şekil 5.2. Digits veri seti q -geri yayılım algoritması için karmaşıklık matrisi.....	114
Şekil 5.3. Iris veri seti: Geri yayılım algoritması için karmaşıklık matrisi.....	116
Şekil 5.4. Iris veri seti: q -geri yayılım algoritması için karmaşıklık matrisi	116
Şekil 5.5. Iris veri seti: q -aktivasyon fonksiyonlu q -geri yayılım algoritması için karmaşıklık matrisi	117
Şekil 5.6. MNIST veri seti: Stokastik geri yayılım algoritması için karmaşıklık matrisi	118
Şekil 5.7. MNIST veri seti: q -stokastik geri yayılım algoritması için karmaşıklık matrisi	119
Şekil 5.8. MNIST veri seti: q -aktivasyon fonksiyonlu q -stokastik geri yayılım algoritması için karmaşıklık matrisi	119
Şekil 5.9. MNIST veri seti: Esnek geri yayılım algoritması için karmaşıklık matrisi..	121
Şekil 5.10. MNIST veri seti: q -esnek geri yayılım algoritması için karmaşıklık matrisi	121
Şekil 5.11. MNIST veri seti: q -aktivasyon fonksiyonlu q -esnek geri yayılım algoritması için karmaşıklık matrisi.....	122

ÇİZELGELER DİZİNİ

Çizelge 3.1. Örnek 3.2.2 de verilen iterasyonun adımları	32
Çizelge 3.2. XOR doğruluk çizelgesi	46
Çizelge 4.1. Şekil 4.1'deki modele göre algoritmaların performans metrikleri	87
Çizelge 4.2. Şekil 4.2'deki modele göre algoritmaların performans metrikleri	88
Çizelge 4.3. Şekil 4.3'teki modele göre algoritmaların performans metrikleri	89
Çizelge 4.4. Şekil 4.4'teki modele göre algoritmaların performans metrikleri	90
Çizelge 4.5. California Housing veri seti: 100 rastgele örnekte regresyon performansı	91
Çizelge 4.6. Diabetes veri seti: 100 rastgele örnekte regresyon performansı	93
Çizelge 4.7. Iris veri seti: Algoritma 4.1 sınıflandırma metrikleri.....	94
Çizelge 4.8. Digits veri seti: Algoritmaların sınıflandırma metrikleri.....	97
Çizelge 4.9. MNIST veri seti: Algoritmaların sınıflandırma metrikleri	99
Çizelge 4.10. Anacalt veri seti: 100 rastgele örnekte regresyon performansı	101
Çizelge 4.11. Delta Ail veri seti: 100 rastgele örnekte regresyon performansı	103
Çizelge 4.12. Machine CPU veri seti: 100 rastgele örnekte regresyon performansı	104
Çizelge 4.13. California Housing veri seti: 100 rastgele örnekte regresyon performansı	106
Çizelge 4.14. BCWD veri seti: Algoritmaların sınıflandırma metrikleri.....	108
Çizelge 4.15. Fashion MNIST veri seti: Algoritmaların sınıflandırma metrikleri	111
Çizelge 5.1. XOR çözümü: $\tanh$ aktivasyon fonksiyonu ile geleneksel geri yayılım ve q -geri yayılım algoritması ile elde edilen sonuçlar	113
Çizelge 5.2. Digits veri seti: Algoritmaların sınıflandırma metrikleri.....	115
Çizelge 5.3. Iris veri seti: Geri yayılım algoritmalarının sınıflandırma metrikleri.....	117
Çizelge 5.4. Stokastik geri yayılım algoritmalarının MNIST veri seti ile sınıflandırma metrikleri	120
Çizelge 5.5. Esnek geri yayılım algoritmalarının MNIST veri seti ile sınıflandırma metrikleri	122

1. GİRİŞ

Yapay sinir ağları günümüz bilim insanları tarafından ilgi odağı haline gelmiş bir araştırma alanıdır. Bilgisayarların öğrenmesini konu eden bu alan bunu da insan vücudunda bulunan biyolojik nöronlardan esinlenerek yapay nöronlar aracılığıyla yapmaktadır. Günümüzde güncel teknolojik gelişmeler takip edildiğinde ilk yıllarda yalnızca veri transferi ve karmaşık hesaplamalar için kullanılan bilgisayarların, zamanla büyük veri setlerini çok kısa sürede özetleyebilen ve mevcut bilgileri kullanıp gerçekleşen olaylar hakkında ise çıkarımlar yapabilen yetkinlikler kazandığı görülmektedir (Öztemel 2003). Yapay sinir ağları dış dünyadan verileri alıp öğrenmekte ve sonrasında ise karşılaştığı probleme nasıl çözüm üretebileceğine karar verebilen aslında zeki sistem diyebileceğimiz bir yapıdır. Bu sayede öğrenmeyi, olaylar arasında ilişki kurabilmeyi, verileri sınıflandırabilmeyi, olaylar hakkında genleme yapabilmeyi ve verilerin özelliklerini belirlemeyi başarıyla sağlamaktadır.

1940 yılından itibaren McCulloch ve Pitts (1943) yaptıkları çalışmalar ile hem yapay sinir ağlarının temellerini hemde mühendislik alanına uygulanabilir olduklarını göstermişlerdir. Hebb (1949) ise yapay sinir ağı elemanlarının değerlerini değiştiren bir öğrenme kuralı tanımlayarak günümüzde var olan çokça öğrenme kuralının temelini de inşa etmiştir. Bilgisayarlarında yaygınlaşması ve geliştirilmesiyle beraber hızlanan çalışmalar ile hem uygulamada hem de teorik alt yapıda ciddi derecede mesafe kat edilmiştir.

Bu teorik çerçevenin güncel ve etkili uygulama alanlarından biri aslında yapay zekânın temelini oluşturan yapay sinir ağlarının eğitilmesi sürecidir. Esasen, bir yapay sinir ağının öğrenme süreci temelde şu soruya cevap arar; “Ağa ait ağırlık değerleri belirli bir hata fonksiyonunu minimize edecek şekilde nasıl optimize edilir?” Özellikle derin ağ mimarilerinde karşılaşılan bu büyük ölçekli minimizasyon problemleri matematiksel olarak varyasyonel eşitsizlik problemleri, sabit nokta problemleri ve monoton kapsama problemleri ile analiz edilebilen bir yapı haline gelmiştir. Tarihsel olarak yapay sinir ağlarının gelişimindeki en önemli dönüm noktalarından biri tek katmanlı mimarilerin doğrusal olarak ayıramayan "XOR" problemini çözmede yetersiz kalmasıyla yaşanmıştır. Bu problem sayesinde doğrusal olmayan yapıları da öğrenebilen çok

katmanlı mimarilerin ve bu mimarileri etkin bir şekilde eğitecek daha gelişmiş optimizasyon algoritmalarının geliştirilmesini de tetiklemiştir. Böylece yapay zekânın günümüzdeki seviyesine hızla ulaşmasının önünü açmıştır.

Yapay sinir ağlarının eğitiminde en temel yaklaşımlardan biri olan geri yayılım algoritmasının (Backpropagation-BP) sinir ağları bağlamındaki ilk formülasyonu Werbos (1982) tarafından sunulmuştur. Algoritmanın yaygınlaşmasını ve sinir ağları alanında devrim niteliğinde bir gelişme yaşanmasını sağlayan öncü çalışma 1986'da ortaya konmuştur (Rumelhart et al. 1986). BP tüm veri setini kullanarak kararlı ancak yavaş güncellemeler yaparken stokastik geri yayılım (Stochastic Backpropagation-SBP), her bir veri örneğiyle hem modelin her eğitim örneğini gördüğünde hemen ağırlıklarını güncelleyerek daha hızlı uyum sağlamasını, öğrenme sürecinde daha erken sonuçlar üretebilmesini ve hemde yerel minimumlardan kaçınma eğilimi gösterir. SBP temelinin stokastik gradyan inişi (Stochastic Gradient Descent-SGD) optimizasyon yönteminden alır. SGD'nin matematiksel temelleri ise 1951 yılında atılmıştır (Herbert et al. 1951). Bu yöntemin BP ile birleştirilerek sinir ağlarına uygulanması ise 1990'ların başında, özellikle Yann LeCun'un öncülüğünde, büyük veri setleriyle çalışmayı mümkün kılan bir yaklaşım olarak ortaya çıkmıştır (LeCun et al. 1989). Bu yöntemlerden farklı olarak esnek geri yayılım (Resilient Backpropagation-Rprop) ise gradyanın büyüklüğü yerine yalnızca yönünü dikkate alarak öğrenme oranını adaptif bir şekilde ayarlayan sağlam bir mekanizma sunar (Riedmiller and Braun 1993). Bu, onu standart geri yayılımdan çok daha esnek hale getirir.

Literatüre kazandırılan bu modeller optimizasyon problemlerine algoritmik düzeyde yenilikçi çözümler sunarken devam eden bilimsel çalışmalar aynı zamanda bu problemlerin analizinde kullanılan matematiksel teoremin de gelişmesini sağlamıştır. Özellikle optimizasyon problemlerinin temelinin oluşturan fonksiyonların her zaman pürüzsüz (smooth) veya kolayca türevlenebilir olmadığı durumlar klasik analiz yöntemlerinin sınırlarını zorlamaktadır. Bu tür durumlarla başa çıkabilmek için klasik kalkülsten farklı olarak daha genel ve esnek teorilere ihtiyaç duyulmuştur. Bu ihtiyaca cevap veren yaklaşımlardan biri olarak klasik kalkülüsün parametrik bir genellemesi şeklinde değerlendirilen q -kalkülüs kavramıdır. İlk olarak Jackson (1908) tarafından tanımlanan q -türev kavramı özellikle sayısal analiz ve özel fonksiyonlar teorisinde de

geniş bir uygulama alanı bulmuştur. 2016 yılında adi türev gerektirmeyen iki küresel optimizasyon yöntemi sunulmuştur. Bunların ilki en dik iniş (Steepest Descent) yönteminin bir q -analoğu olan $q - G$ yöntemi, ikincisi ise eşlenik gradyan yönteminin (Conjugate Gradient) bir q -analoğu olan $q - CG$ yöntemidir (Gouvêa et al. 2016). 2020 yılına gelindiğinde ise q -nöron olarak adlandırılan yeni bir yapay nöron türü önerilmiştir. Bu q -nöronlar aktivasyon fonksiyonu stokastik bir q parametresi için q türeve dayanan bir nöron olup sinir ağı mimarilerinde nasıl kullanılacağı gösterilmiştir (Nielsen and Sun 2020). 2021 yılında ise kısıtsız optimizasyon problemlerini çözmek için q -gradyan temelli olan bir q -Dai-Yuan (q -DY) eşlenik gradyan algoritması (Conjugate Gradient Algorithm) önerilmiştir. Algoritmanın kuvvetli yakınsaması standart Wolfe koşulları altında kanıtlanmıştır. Ayrıca sayısal sonuçlar ile önerilen algoritmanın verimliliği gösterilmiştir (Mishra et al. 2021). Disiplinler arası bir uygulamaya da sahip olan bu q yaklaşımı 2025 yılında yapılan bir çalışmada yeni enerjiyi daha iyi tüketmek ve yeni enerjiden tam olarak yararlanmak için, rüzgar enerjisi belirsizliğini dikkate alan entegre enerji sistemi için aralıklı bir optimal programlama yöntemi önerilmiş ve bu yöntemin optimizasyon doğruluğunu daha da artırmak için kuantum türev kullanmıştır (Sun et al. 2025).

Bu tezde q -türev temelli modelleri kullanmamızda ki temel motivasyon kaynağımız kuantum türevin disiplinler arası uygulamada başarılı sonuçlar veren esnek ve uyarlanabilir olmasıdır. Buradan hareketle literatürde Nielsen and Sun (2020) tarafından sadece aktivasyon fonksiyonlarında kullanılan kuantum türev kavramı ilk kez bu tez ile birlikte adi türevde ki gibi zincir kuralı olmamasına rağmen öğrenme algoritmalarının doğrudan ağırlık güncellemesinde başarı ile kullanılmıştır. Geleneksel geri yayılım, stokastik geri yayılım ve esnek geri yayılım algoritmalarının q -türev ile q -aktivasyon fonksiyonu içeren versiyonları tanımlanarak sınıflandırma problemleri üzerinde Digits, Iris ve MNIST gibi büyük örnek içeren veri setleri üzerinde uygulanmıştır. Elde edilen sonuçlar ışığında geleneksel yöntemlere göre hem q -türev içeren modellerin hemde q -türev ile q -aktivasyon fonksiyonu içeren modellerin yüksek başarı gösterdiği ortaya konulmuştur.

Öte yandan farklı teorik yapıların kesişiminde bulunan monoton kapsama problemleri ise bu yönüyle birçok problemin çözümünde merkezi bir rol üstlenmektedir.

Monoton kapsama problemleri ile sabit nokta problemleri ve konveks minimizasyon problemleri arasında ilişki kurulabilmektedir (Rockafellar 1970b; Bauschke and Combettes 2017). Bu ilişki sayesinde çözümü zor olan veya çözülemeyen problemler birbirlerine rahatlıkla dönüştürülüp çözüm kümeleri bulunabilmektedir. Monoton kapsama problemi, ilk kez 20. yüzyılın ortalarında Minty (1962) ve sonrasında Rockafellar (1970a) tarafından kuramsal düzeyde ele alınmıştır. Ardından Lions and Mercier (1979) tarafından geliştirilen ayrıştırma teknikleri aracılığıyla problemin çözümüne yönelik algoritmalar önerilmiştir. Bu problem $\mathcal{H}$ bir reel Hilbert uzayı, $A: \mathcal{H} \rightarrow \mathcal{H}$ tek değerli bir operatör ve $B: \mathcal{H} \rightarrow 2^{\mathcal{H}}$ küme değerli ve maksimal monoton bir operatör olmak üzere, $0 \in (A + B)z$ koşulunu sağlayan bir $z \in \mathcal{H}$ noktasının bulunması problemidir.

Konveks ve türevlenebilir olan fonksiyonların gradyanlarının monoton karakterde olması pek çok konveks minimizasyon probleminin monoton kapsama problemi şeklinde ifade edilmesini sağlamıştır. Böylece aslında birer konveks optimizasyon problemi olan regresyon ve sınıflandırma gibi istatistiksel öğrenme alanlarında da monoton kapsama problemlerinin izleri gözlemlenmektedir. Dolayısıyla, gradyan temelli algoritmalar ve ileri-geri ayrıştırma gibi yöntemler, bu tür problemlere uygulanabilmektedir.

Literatürde var olan bu problemler arasındaki teorik geçiş ise bizim monoton kapsama probleminin çözümüne yönelik yaptığımız çalışmalara ilham olmuştur. Bu tezde çift eylemsizlik terimi ile viskozite terimi içeren ve monoton kapsama probleminin çözümüne yakınsak olan iki farklı algoritma önerilmiştir. Bu algoritmaların teorileri ile yakınsaklık durumları ifade ve ispat edilmiştir. Ayrıca regresyon ve sınıflandırma problemleri üzerinde benzetimleri yapılmıştır. Bu benzetimler yapılırken Huang et al. (2006) tarafından tanımlanan ekstrem öğrenme makineleri modifiye edilerek kullanılmıştır.

Gelişmiş optimizasyon algoritmalarının büyük çoğunluğu temelde ağırlık öğrenilebilir parametreleri olan ağırlık ve eşik değerlerini (bias) iteratif olarak güncelleyen geri yayılım mekanizmasına dayanmaktadır. Süreç, ağırlık bir tahmin üretmesi, bu tahminin gerçek değerden sapmasının bir hata fonksiyonu ile ölçülmesi ve ardından hesaplanan hatanın ağırlık boyunca geriye doğru yayılmasıyla işler. Bu geri yayılım adımı

hatayı azaltacak şekilde her bir parametrenin güncellenmesini ya belirlenen hata değerine ulaşana kadar ya da belirlenen iterasyon sayısına ulaşana kadar sağlar. Eğitim veri seti üzerinde defalarca tekrarlanan bu döngüsel süreç, ağın hatayı minimize etmesini sağlasa da önemli dezavantajları da beraberinde getirmektedir. Örneğin yavaş yakınsama eğilimi, yüksek hesaplama maliyeti ve lokal minimumlarda sıkışma problemi bu iteratif yöntemlerin temel zorlukları arasında yer almaktadır. Bu zorluklar algoritmanın öğrenmesini etkileyen sebepler olup böylece yapay sinir ağı potansiyel olarak ulaşabileceği en iyi performansa ulaşamamaktadır.

Bu zorluklara radikal bir çözüm önerisi olarak sunulan ekstrem öğrenme makineleri (Extreme Learning Machines-ELM), geleneksel iteratif eğitim paradigmasını temelden değiştirmektedir. Tek gizli katmanlı yapıya sahip olan bu ağ mimarisinde gizli katmanın girdi ağırlıkları ve eşik değerleri eğitilmemektedir. Bunun yerine sürecin başında rastgele atanarak sabitlenmektedir. Böylece ağıın öğrenme süreci aslında gizli katman ile çıktı katmanı arasındaki ağırlıkların tek bir adımda analitik olarak hesaplandığı bir lineer denklem sisteminin çözümüne indirgenmektedir. Çıktı katmanının ağırlıkları ise Moore-Penrose genelleştirilmiş matrisinin tersi kullanılarak hesaplanmaktadır. İteratif ayarlamalara ve geri yayılıma ihtiyaç duymayan ELM bu sayede olağanüstü bir eğitim hızı sunarken lokal minimum problemini de tamamen ortadan kaldırır. Ancak bu yöntemde iki farklı problemle karşılaşılabilir. Birincisi H matrisinin Moore-Penrose genelleştirilmiş tersini bulmak her zaman mümkün olmayabilir. İkincisi ekstrem öğrenme makinesinin çalışabilmesi için gizli düğüm sayısının örnek sayısından küçük ya da eşit olması gerekmektedir. Ancak birçok gerçek yaşam probleminde bu sayı daha büyük olmaktadır. Bu da modelin aşırı öğrenmesine (overfitting) sebep olmaktadır ve istenilen bir durum değildir. Bu tezde monoton kapsama problemlerinin çözümüne yönelik yapılan çalışmalarda belirtilen problemlerle karşılaşmamak için ELM yapısındaki ağırlık ve eşik değerlerinin rastgele seçilmesi kısmı modellere entegre edilmiştir. Bu sayede çıktı katmanı ağırlıklarını ELM ile değil önerilen algoritmalar kullanılarak iteratif yaklaşımlarla hesaplanılmıştır. Yapılan benzetimler aşamasında açık erişim kaynaklı California Housing, Diabetes, İris, Digits, MNIST, Anacalt, Delta Ail, Machine CPU, BCWD ve Fashion MNIST veri kümeleri kullanılmıştır. Yine literatürde var olan ve benzetimleri sonucunda başarılı olan diğer algoritmalar ile yeni önerilen iki algoritma kıyaslanmış ve onlardan daha başarılı sonuçlar elde edilmiştir.

Özetle bu çalışma kapsamında q -türev literatürde ilk kez öğrenme algoritmasının aktivasyon fonksiyonu aşamasına değil, doğrudan ağırlıkların optimizasyon aşamasına entegre edilmiştir. Literatürde var olan BP, SBP, Rprop algoritmalarının q versiyonları tanımlanmış ve sınıflandırma problemlerine uygulanmıştır. Ayrıca monoton kapsama problemlerinin çözümü için yeni algoritmalar önerilmiştir. Önerilen yöntemler hem çift eylemsizlikli hızlandırma hemde viskozite yaklaşımlarını içermektedir. Geliştirilen algoritmaların teorik geçerliliği ve pratik performansı sayısal deneyler aracılığıyla detaylı bir şekilde incelenmiştir. Ayrıca ifade edilen yöntemin etkinliği makine öğrenmesindeki regresyon ve sınıflandırma problemleri üzerinde test edilmiştir. Yapılan tüm benzetim sonuçlarında başarılı sonuçlar elde edilmiştir.



2. KAYNAK ÖZETLERİ

Bu bölümde, monoton kapsama probleminin çözümüne yönelik literatürde yer alan, iki eylemsizlik terimi içeren çalışmalar ile kuantum türev kavramını ele alan araştırmalardan bahsedilecektir. Bu noktada öncelikle iki eylemsizlik terimine dayalı yaklaşımlar verilmiştir.

Wang et al. (2022) çalışmalarında, ilgili dönüşümlerin ters kuvvetli monoton dönüşüm özelliği varsayılmadan, yeni algoritmalar tarafından üretilen yinelemeli dizilerin, sonsuz boyutlu reel Hilbert uzaylarında ayrıştırılmış monoton kapsama probleminin çözümüne kuvvetli biçimde yakınsadığı gösterilmiştir. Yeni algoritmaların adım büyüklükleri, operatör normunun önceden bilinmesine gerek kalmadan, her yinelemede basit bir hesaplama ile güncellenmektedir. İlgili adım büyüklüklerinin değer aralığını genişletmek amacıyla bazı parametreler gevşetilmiştir. Yakınsama hızını artırmak için yeni algoritmalara çift eylemsizlik ekstrapolasyon adımları dâhil edilmiştir. Uygulama olarak, ayrıştırılmış varyasyonel kapsama problemi, monoton kapsama problemi ve ayrıştırılmış varyasyonel eşitsizlik problemi incelenmiştir. Yeni algoritmaların hesaplama verimliliğini göstermek amacıyla bazı sayısal deneyler gerçekleştirilmiştir. Algoritmaları ise şu şekildedir.

Başlatma Kriteri: $x_0, x_1 \in H_1$, $0 < \kappa_1 < 2$, $0 < \kappa_2 < \kappa_1' < 2$, $\varrho \in (0,1)$, $\lambda_1, \mu_1 > 0$, $\alpha > 0$, $b_0 > 0$, $b_0 = x_0$ ve x_1 seçilir ve $n := 1$ olarak ayarlanır.

Adım 1. $0 \leq \alpha_n \leq \bar{\alpha}_n$ olacak şekilde α_n seçilir. Burada

$$\bar{\alpha}_n = \begin{cases} \min \left\{ \alpha, \frac{\varrho \|x_n - x_{n-1}\|}{\|x_n - x_{n-1}\|} \right\}, & x_n \neq x_{n-1} \\ \alpha, & \text{aksi halde} \end{cases}$$

dır.

Adım 2.

$$\omega_n = x_n + \alpha_n(x_n - x_{n-1}),$$

$$y_n = J_{\lambda_n F}(T\omega_n - \lambda_n FT\omega_n),$$

$$\lambda_{n+1} = \begin{cases} \min \left\{ (\varrho + s_n) \frac{\|T\omega_n - y_n\|^2}{\|FT\omega_n - Fy_n, T\omega_n - y_n\|}, \lambda_n + p_n \right\}, & |FT\omega_n - Fy_n, T\omega_n - y_n| \neq 0 \\ \lambda_n + p_n, & \text{aksi halde} \end{cases}$$

ve

$$r_n = T\omega_n - y_n - \lambda_n(FT\omega_n - Fy_n)$$

olmak üzere

$$z_n = T\omega_n - \beta_n r_n,$$

ifadesi ve

$$\beta_n = \begin{cases} (\kappa_1 + s_n) \frac{\|T\omega_n - z_n\|}{\|r_n\|}, & r_n \neq 0 \\ 0, & \text{aksi halde} \end{cases}$$

hesaplanır.

Adım 3. Ardından

$$b_n = \omega_n + \eta_n T^*(z_n - T\omega_n),$$

$$\eta_n = \begin{cases} \frac{\varrho \|T\omega_n - z_n\|^2}{\|T^*(T\omega_n - z_n)\|^2}, & \text{eğer } \|T^*(T\omega_n - z_n)\| \neq 0 \\ 0, & \text{aksi halde} \end{cases}$$

bulunur.

Adım 4.

$$\bar{\rho}_n = \begin{cases} \min\left\{\rho, \frac{\varrho \|b_n - b_{n-1}\|}{\|b_n - b_{n-1}\|}\right\}, & b_n \neq b_{n-1} \\ \rho, & \text{aksi halde} \end{cases}$$

olmak üzere $0 \leq \rho_n \leq \bar{\rho}_n$ olacak şekilde ρ_n seçilir.

Adım 5.

$$\xi_n = b_n + \rho_n(b_n - b_{n-1}),$$

$$u_n = J_{\mu_n A}(\xi_n - \mu_n A\xi_n),$$

μ_{n+1}

$$= \begin{cases} \min\left\{(\varrho + q_n) \frac{\|\xi_n - u_n\|^2}{\|Au_n - A\xi_n, u_n - \xi_n\|}, \mu_n + q_n\right\}, & \text{eğer } |Au_n - A\xi_n, u_n - \xi_n| \neq 0 \\ \mu_n + q_n, & \text{aksi halde} \end{cases}$$

ve

$$v_n = \xi_n - u_n - \mu_n(A\xi_n - Au_n)$$

olmak üzere

$$t_n = \xi_n - v_n v_n,$$

ifadesi ve

$$v_n = \begin{cases} (\kappa_2 + \sigma_n) \frac{\|t_n - u_n\|}{\|v_n\|^2}, & v_n \neq 0 \\ 0, & \text{aksi halde} \end{cases}$$

hesaplanır.

Adım 6. Daha sonra ise

$$x_{n+1} = (1 - \theta_n - \phi_n - \delta_n)b_n + \theta_n r_n + \phi_n s_n.$$

bulunur. $n := n + 1$ olarak ayarlanır ve Adım 1'e geri dönülür.

Yao et al. (2022) çalışmalarında psödo-monoton ve Lipschitz sürekli amaç fonksiyonu için reel Hilbert uzaylarında varyasyonel eşitsizliklerin çözümü hedeflenmiştir. Çift eylemsizlikli ekstrapolasyon ve uyarlanabilir adım büyüklüğü içeren yeni bir alt diferansiyel ve ekstra gradyan yöntemi algoritması geliştirilmiştir. Bu algoritmanın, mevcut yaklaşımlara göre yakınsama hızını artırdığı ve sayısal deneylerle etkinliği ortaya konmuştur. Ayrıca, yöntemin daha esnek parametrelerle uygulanabildiği ve literatürdeki özel durumları kapsadığı belirtilmiştir. Algoritmaları ise aşağıda verilmiştir.

Adım 1. $\mu \in (0,1)$, $\lambda_1 > 0$, $x_0, x_1 \in H$ noktaları olarak seçilir. $n := 1$ olarak ayarlanır.

Adım 2.

$$\begin{aligned} z_n &= x_n + \delta(x_n - x_{n-1}), \\ w_n &= x_n + \theta_n(x_n - x_{n-1}), \\ y_n &= P_C(w_n - \lambda_n A w_n), \end{aligned}$$

eşitlikleri yardımıyla z_n, w_n ve y_n hesaplanır. Eğer $w_n = y_n = x_n$ ise algoritma durdurulur. Aksi halde Adım 3'e gidilir.

Adım 3.

$$\lambda_{n+1} = \begin{cases} \min \left\{ \frac{\mu \|w_n - y_n\|}{\|Aw_n - Ay_n\|}, \lambda_n \right\}, & \text{eğer } Aw_n \neq Ay_n \\ \lambda_n, & \text{aksi halde} \end{cases}$$

ve

$$T_n := \{w \in H: \langle w_n - \lambda_n Aw_n - y_n, w - y_n \rangle \leq 0\}$$

olmak üzere

$$x_{n+1} = (1 - \alpha_n)z_n + \alpha_n P_{T_n}(w_n - \lambda_n Ay_n)$$

hesaplanır.

Adım 4. $n \leftarrow n + 1$ olarak ayarlanır ve 2. adıma geri dönülür.

Viet et al. (2022) tarafından yapılan çalışmada reel Hilbert uzaylarında Lipschitz sürekli olmayan psödo-monoton dönüşümler için varyasyonel eşitsizlik problemleri incelenmiştir. Bunların çözümü için çift eylemsizlik terimli iz düşüm türünde bir yöntem önerilmiştir. Önerilen algoritmada, çift eylemsizlikli adımlar ile iz düşüm ve daraltanlık tekniği birleştirilmiştir. Her iterasyonda yalnızca bir iz düşüm işlemi uygulanmakta olup, algoritmanın çalışması için Lipschitz süreklilik koşulu gerekmemektedir. Zayıf yakınsama için yeterli koşul, psödo-monotonluk ve düzgün süreklilik altında verilmiştir. Güçlü psödo-monotonluk ve Lipschitz sürekliliği varsayımları altında ise R-lineer yakınsama hızı elde edilmiştir. Son olarak, algoritmanın hesaplama etkinliğini göstermek amacıyla sayısal sonuçlara yer verilmiştir. Elde edilen temel bulguların, ilgili literatürdeki bazı çalışmaları genişlettiği ve iyileştirdiği belirtilmiştir. Algoritmaları ise aşağıda sunulmuştur.

Başlatma kriterleri: $\rho, l \in (0,1), \mu \in (0,1), \gamma \in (0,2), \theta \in [0,1], \alpha, \beta \in (0,1), s_0, s_1 \in H$ rastgele seçilir.

Adım 1.

$$\begin{cases} u_n = s_n + \alpha(s_n - s_{n-1}) \\ w_n = s_n + \theta(s_n - s_{n-1}) \end{cases}$$

ve

$$y_n = P_C(w_n - \tau_n F w_n),$$

hesaplanır. Burada τ_n ise

$$\tau \parallel F y_n - F w_n \parallel \leq \mu \parallel y_n - w_n \parallel$$

eşitsizliğini sağlayan $\{\rho, \rho l, \rho l^2, \dots\}$ kümesindeki en büyük τ olarak seçilir. Eğer $w_n = y_n$ veya $F w_n = 0$ ise durdurulur ve bu durumda w_n ele alınan problemin bir çözümüdür. Aksi takdirde:

Adım 2.

$$\eta_n := \begin{cases} \frac{\langle w_n - y_n, d_n \rangle}{\parallel d_n \parallel^2} & \text{eğer } d_n \neq 0 \\ 0 & \text{eğer } d_n = 0 \end{cases}$$

olmak üzere

$$z_n = w_n - \gamma \eta_n d_n,$$

ve

$$d_n := w_n - y_n - \tau_n(Fw_n - Fy_n)$$

hesaplanır.

Adım 3. Son olarak

$$s_{n+1} = (1 - \beta)u_n + \beta z_n$$

hesaplanır. $n := n + 1$ alınıp tekrar Adım 1'e dönülür.

Suantai et al. (2023) yapmış oldukları bu çalışmada monoton kapsama problemleri için ileri-geri-ileri algoritmalarının yeni bir türünü sunmuşlardır. Lipschitz sabitinin önceden seçilmesi gerekliliğini ortadan kaldırmak amacıyla kendiliğinden uyarlanabilen bir tekniği temel almışlardır. Algoritmanın yakınsama performansını artırmak için çift eylemsizlik ekstrapolasyon metodunu kullanılmışlardır. Önerilen algoritmanın zayıf yakınsaklığı ise hafif varsayımlar altında kanıtlanmıştır. Ayrıca, görüntü bulanıklığının giderilmesi ve sinyal kurtarma gibi uygulamalarda sayısal testler gerçekleştirilmiştir. Elde edilen sonuçların, algoritmanın literatürdeki bazı bilinen algoritmalarından daha iyi performans gösterdiğini ortaya koyduğu belirtilmiştir. Bahse konu bu algoritmaları ise aşağıda verilmiştir.

Başlatma kriteri: $\vartheta_1 > 0$, $\mu \in (0,1)$, $\{\lambda_n\}$, $\{\zeta_n\}$ pozitif reel diziler olmak üzere $v_0, v_1 \in \mathcal{H}$ ve $\{\gamma_n\}$ sıfırdan büyük bir dizi olup $\sum_{n=1}^{\infty} \gamma_n < \infty$ olacak şekilde seçilir.

Adım 1. v_n, v_{n-1} verildiğinde

$$\begin{aligned} u_n &= v_n + \lambda_n(v_n - v_{n-1}), \\ z_n &= u_n + \zeta_n(u_n - v_{n-1}), \\ y_n &= (I + \vartheta_n \tilde{B})^{-1}(I - \vartheta_n \tilde{A})z_n \end{aligned}$$

hesaplanır. Eğer $y_n = z_n$ ise algoritma durdurulur ve $y_n \in \Gamma$ elde edilir. Aksi halde Adım 2'ye geçilir.

Adım 2.

$$v_{n+1} = y_n - \vartheta_n(\tilde{A}y_n - \tilde{A}z_n).$$

hesaplanır.

Adım 3.

$$\vartheta_{n+1} = \begin{cases} \min \left\{ \frac{\mu \|z_n - y_n\|}{\|\tilde{A}z_n - \tilde{A}y_n\|}, \vartheta_n + \gamma_n \right\}, & \text{eğer } \|\tilde{A}z_n - \tilde{A}y_n\| \neq 0; \\ \vartheta_n + \gamma_n, & \text{aksi halde.} \end{cases}$$

dir. Daha sonra ise Adım 1'e geri dönülür.

Thong et al. (2023) çalışmalarında, Hilbert uzaydaki psödo-monoton varyasyonel eşitsizlik problemlerini çözmek üzere, çift eylemsizlikli ekstrapolasyon adımları ve kendinden uyarlamalı adım boyutuna sahip tek projeksiyonlu bir algoritma önermişlerdir. Tseng'in (2000) ekstra gradyan yönteminin tek projeksiyonlu versiyonu temel alınarak geliştirilen bu yöntem, iki farklı eylemsizlik parametresi ve gevşeme katsayısı ile iyileştirilmiştir. Bahse konu algoritmaları ise şu şekilde verilmiştir. $\mu \in (0,1)$, $\alpha, \beta \in [0,1]$, $\theta \in (0,1)$ ve $\tau_1 > 0$ verilsin. $z_0, z_1 \in H$ rastgele seçilsin. $\{\alpha_n\}$ pozitif reel sayılardan oluşan ve

$$\sum_{n=1}^{\infty} \alpha_n < +\infty$$

koşulunu sağlayan bir dizi olsun. Bu durumda

$$w_n = z_n + \alpha(z_n - z_{n-1}),$$

$$u_n = z_n + \beta(z_n - z_{n-1}),$$

$$\begin{aligned} y_n &= P_C(u_n - \tau_n F(u_n)), \\ z_{n+1} &= (1 - \theta)w_n + \theta(y_n - \tau_n(F(y_n) - F(u_n))) \end{aligned}$$

dır. Adım boyutu güncellemesi

$$\tau_{n+1} = \begin{cases} \min \left\{ \frac{\mu \|u_n - y_n\|}{\|F(u_n) - F(y_n)\|}, \tau_n + \alpha_n \right\}, & \text{eğer } F(u_n) \neq F(y_n), \\ \tau_n + \alpha_n, & \text{aksi halde.} \end{cases}$$

ile yapılır. Bu algortmada çift eylemsizlik adımı ile hızlandırma sağlanmakta; adım boyutu her iterasyonda uyarlamalı olarak güncellenmekte ve gerektiğinde artırılabilir. Yöntem yalnızca tek projeksiyon gerektirdiğinden hesaplama maliyeti düşüktür.

Wang et al. (2023) çalışmalarında, reel Hilbert uzaylarındaki monoton kapsama problemlerini çözmek üzere çift eylemsizlik ekstrapolasyon adımları ve kendinden uyarlamalı adım boyutları içeren değiştirilmiş bir Tseng ayrıştırma yöntemi sunulmaktadır. Makul ve standart varsayımlar altında, önerilen algoritmanın sırasıyla zayıf yakınsaması, asimptotik olmayan $O\left(\frac{1}{\sqrt{n}}\right)$ yakınsama hızı, güçlü yakınsaması ve doğrusal yakınsama oranı kurulmaktadır. Son olarak, algoritmanın başarımını ve kuramsal sonuçlarını göstermek üzere çeşitli sayısal deneyler sunulmaktadır. Bahse konu algoritmaları ise şu şekilde verilmiştir. $\mu \in (0,1)$ ve $\lambda_1 > 0$ verilsin. Başlangıç noktaları $x_0, x_1 \in H$ seçilsin. Sırasıyla $\{\alpha_n\}, \{\beta_n\} \subset [0,1]$, $\{\theta_n\} \subset (0,1]$, $\{p_n\} \subset (0, \infty)$ ve $\{\mu_n\} \subset [0,1)$ olsun. Ayrıca $\sum_{n=1}^{\infty} p_n < \infty$ ve $\mu_n \rightarrow 0$ olduğu varsayılın. Bu durumda

$$\begin{aligned} w_n &= x_n + \alpha_n(x_n - x_{n-1}), \\ z_n &= x_n + \beta_n(x_n - x_{n-1}), \\ y_n &= (I + \lambda_n B)^{-1}(I - \lambda_n A) w_n, \\ x_{n+1} &= (1 - \theta_n) z_n + \theta_n(y_n - \lambda_n(Ay_n - Aw_n)) \end{aligned}$$

dir. Adım boyutu güncellemesi

$$\lambda_{n+1} = \begin{cases} \min \left\{ \frac{(\mu_n + \mu) \|w_n - y_n\|}{\|Aw_n - Ay_n\|}, \lambda_n + p_n \right\}, & \text{eğer } Aw_n \neq Ay_n, \\ \lambda_n + p_n, & \text{aksi halde} \end{cases}$$

ile yapılır. Eğer $w_n = y_n$ olursa algoritma durdurulur ve y_n bir çözüm kabul edilir; değilse $n \leftarrow n + 1$ yapılarak aynı adımlar tekrarlanır. Böylece çift eylemsizlik sayesinde hızlandırma amaçlanırken, Lipschitz sabiti bilinmeden adım boyutu da her iterasyonda uyarlamalı biçimde güncellenir.

Jolaoso et al. (2023) çalışmalarında, reel Hilbert uzaylarında maksimal monoton ve ters kuvvetli monoton operatörlerin toplamının bir sıfır noktasını bulmak amacıyla, iki eylemsizlik parametrelili (biri negatif olmayan ve diğeri pozitif olmayan) bir ekstrapolasyon adımına sahip ileri-geri ayrıştırma (FBS) algoritması önerilmektedir. Önerilen algoritmanın ilginç özelliklerinden biri, eylemsizlik parametreleri üzerinde iterasyonlarla birlikte çevrimiçi bir kurala ihtiyaç duyulmamasıdır. Önerilen algoritmanın zayıf yakınsaklık sonucu, bazı standart varsayımlar altında ortaya konulmuştur. Algoritmaları ise $\theta \in [0,1)$, $\delta \leq 0$, $\lambda \in (0,2\alpha)$ ve $x^{-1}, x^0, x^1 \in H$ keyfi başlangıç noktaları olsun. $k = 1$ ile başlamak üzere önerilen iterasyon

$$\begin{cases} w^k = x^k + \theta(x^k - x^{k-1}) + \delta(x^{k-1} - x^{k-2}) \\ x^{k+1} = J_\lambda^B(w^k - \lambda Aw^k) \end{cases}$$

dir. Burada döngü tamamlandığında $k \leftarrow k + 1$ olarak ayarlanır ve iterasyon bu şekilde tekrar yeni bir döngüde çalıştırılır.

Kankam and Cholanjiak (2023) tarafından yapılan çalışmada, kısıtsız minimizasyon problemleri için çift eylemsizlikli ileri-geri algoritmaları ve kısıtlı problemler için izdüşümlü çift eylemsizlikli ileri-geri algoritmaları önerilmektedir. Uygun varsayımlar altında bu yöntemler için yakınsama teoremleri kanıtlanmıştır. Çeşitli sayısal deneyler yürütülmekte ve elde edilen sonuçlarla, önerilen algoritmaların

literatürdeki bilinen yöntemlere kıyasla daha verimli olduğu ortaya konulmaktadır. Algoritmaları ise aşağıda sunulmuştur.

Başlatma kriteri: α_n , θ_n ve η_n reel pozitif diziler olmak üzere $x^0, x^1 \in H$ keyfi olarak seçilir.

Adım 1. $n = 1$ den başlanarak

$$w^n = x^n + \theta_n(x^n - x^{n-1})$$

hesaplanır.

Adım 2.

$$v^n = \text{prox}_{\alpha_n g}(w^n - \alpha_n \nabla f(w^n)),$$

ve

$$y^n = v^n + \eta_n(v^n - x^{n-1})$$

eşitliği hesaplanır. Sonraki yineleme ise

$$x^{n+1} = \text{prox}_{\alpha_n g}(y^n - \alpha_n \nabla f(y^n)).$$

şeklinde hesaplanır. Daha sonra $n \leftarrow n + 1$ olarak güncellenir ve Adım 1'e dönülür.

Wang et al. (2024) çalışmalarında, reel Hilbert uzaylarında yarı-monotonluk koşulunu sağlayan varyasyonel eşitsizliklerin çözümü için, çift eylemsizlikli ekstrapolasyon adımı ve kendinden uyarlamalı adım büyüklükleri içeren bir yöntemi önermiştir. Bazı koşullar altında zayıf ve kuvvetli yakınsama sonuçlarının elde edildiği,

yöntemin özel bir durumunda ise doğrusal yakınsama sağlandığı belirtilmiştir. Yapılan sayısal deneylerle yöntemin literatürdeki diğer yöntemlerle rekabetçi olduğu gösterilmiştir. Algoritmaları ise aşağıda sunulmuştur.

Adım 1. $\gamma \in (0,2)$, $\mu \in (0,1)$ ve $\lambda_1 > 0$ olmak üzere $x_0, x_1 \in H$ başlangıç noktaları seçilir.

Adım 2. $n := 1$ den başlanarak

$$\begin{cases} z_n = x_n + \beta(x_n - x_{n-1}), \\ w_n = x_n + \theta_n(x_n - x_{n-1}), \\ y_n = P_C(w_n - \lambda_n A w_n), \end{cases}$$

eşitlikleri hesaplanır. Eğer $w_n = y_n = x_n$ ise işlem durdurulur. Eşitlik sağlanmaz ise Adım 3'e geçilir. Burada

$$\lambda_{n+1} = \begin{cases} \min \left\{ \mu \frac{\|w_n - y_n\|}{\|Aw_n - Ay_n\|}, \lambda_n \right\}, & Aw_n \neq Ay_n, \\ \lambda_n, & \text{diğer durumda} \end{cases}$$

dır.

Adım 3.

$$d_n = w_n - y_n - \lambda_n(Aw_n - Ay_n), \quad \forall n \geq 1$$

hesaplanır.

Adım 4.

$$\eta_n = \begin{cases} \frac{\langle w_n - y_n, d_n \rangle}{\|d_n\|^2}, & d_n \neq 0, \\ 0, & d_n = 0. \end{cases}$$

olmak üzere

$$x_{n+1} = (1 - \alpha)z_n + \alpha(w_n - \gamma\eta_n d_n), \quad \forall n \geq 1,$$

dır.

Adım 5. $n \leftarrow n + 1$ olarak ayarlanır ve 2. adıma dönülür.

Li et al. (2025) tarafından yapılan çalışmada, Hilbert uzaylarında tanımlı quasi-monoton varyasyonel eşitsizlikler için yeni bir algoritma önerilmiştir. Bu algoritma, çift eylemsizlik (double inertial) terimi, altgradyan ekstrapolasyon yaklaşımı ve projeksiyon kontraksiyon yöntemini birleştirerek geliştirilmiştir. Çalışmada algoritmanın uygun koşullar altında zayıf yakınsadığı, özel parametre seçimlerinde ise kuvvetli yakınsadığı gösterilmektedir.

Adım 1: $\sigma > 0, l, \mu \in (0,1), \gamma \in (0,2)$ olmak üzere başlangıç noktaları $x^0, x^1 \in H$ olsun. $n = 1$ 'den başlayarak

$$\begin{aligned} z^n &= x^n + \delta_n (x^n - x^{n-1}), \\ w^n &= x^n + \theta_n (x^n - x^{n-1}), \\ y^n &= P_C(w^n - \lambda_n A w^n), \\ u^n &= P_{T_n}(w^n - \gamma \lambda_n \tau_n A y^n) \end{aligned}$$

eşitlikleri hesaplanır. Buradan

$$\lambda_n = \sigma l^{m_n}, \quad m_n = \min\{m \geq 0 : \sigma l^m \|Aw^n - Ay^n\| \leq \mu \|w^n - y^n\|\},$$

$$T_n := \{w \in H : \langle w^n - \lambda_n A w^n - y^n, w - y^n \rangle \leq 0\},$$

$$d(w^n, y^n) = w^n - y^n - \lambda_n (A w^n - A y^n),$$

$$\tau_n = \begin{cases} \frac{\langle w^n - y^n, d(w^n, y^n) \rangle}{\|d(w^n, y^n)\|^2}, & d(w^n, y^n) \neq 0, \\ 0, & d(w^n, y^n) = 0 \end{cases}$$

dır.

Adım 2: Eğer $w^n = y^n$ veya $Aw^n = 0$ ise durdurulur.

Adım 3: Aksi halde yeni iterasyon

$$x^{n+1} = (1 - \alpha_n)z^n + \alpha_n u^n$$

olarak güncellenir.

Adım 4: $n \leftarrow n + 1$ yapılarak Adım 1'e dönülür.

İki eylemsizlik terimine dayalı çalışmaların incelenmesinin ardından, kuantum türevini esas alan araştırmalar değerlendirilecektir.

Gouvêa et al. (2016) tarafından yapılan çalışmada, q -türev temelli q -gradient vektörü kullanılarak, klasik türev gerektirmeyen iki global optimizasyon yöntemi geliştirilmiştir. q -G ve q -CG olarak adlandırılan bu yöntemler, $q=1$ için klasik versiyonlarına indirgenmektedir. Arama süreci, başlangıçta global iken, adım adım lokal aramaya dönüşecek şekilde tasarlanmıştır. Bazı yinelemelere Gauss pertürbasyonları eklenerek yöntemlerin olasılıksal olarak global minimuma yakınsaması garanti edilmiştir. q -G ve q -CG yöntemleri, klasik versiyonlarının sonlu fark içeren iç nokta yöntemi gibi diğer yöntemlerle, 27 farklı test problemi üzerinde karşılaştırılmıştır. Genel olarak, q -G ve q -CG yöntemlerinin özellikle çok modellenli problemlerde oldukça rekabetçi ve umut vadeden sonuçlar verdiği gösterilmiştir.

Nielsen and Sun (2020) çalışmalarında, klasik aktivasyon fonksiyonlarına alternatif olarak q -aktivasyon adı verilen yeni bir yaklaşım geliştirilmiştir. Yöntem, Jackson'ın ayrık türevine dayalı olarak tanımlanmış ve aktivasyon fonksiyonlarının stokastik bir biçimde çeşitlendirilmesini sağlamıştır. Böylece derin sinir ağlarının eğitimine, fonksiyonun eğriselliğine duyarlı bir gürültü doğal biçimde eklenmiştir.

Önerilen yöntemde, q parametresi rastgele seçilmiş olup 1'e yaklaştığında sistem normal aktivasyonlara yaklaşmaktadır. Eğrisel aktivasyon fonksiyonlarında (örneğin ELU) q -aktivasyonun etkisi belirgin olurken, doğrusal veya parçalı doğrusal fonksiyonlarda (örneğin ReLU) bu etkinin sınırlı kaldığı belirtilmiştir. Ayrıca, yöntem ek parametre gerektirmemekte ve mevcut ağlara düşük maliyetle uyarlanabilmektedir. Önerilen q -aktivasyonlar, MNIST ve CIFAR-10 veri kümelerinde çok katmanlı algılayıcı, evrimsel ağ ve ResNet gibi farklı mimariler üzerinde sınanmıştır. Elde edilen sonuçlarda, özellikle ELU'nun q -aktivasyonu (q ELU) kullanıldığında test hatalarının azaldığı ve genelleme performansının arttığı görülmüştür. Bağlantı koparma (dropout) veya gradyan gürültüsü gibi yöntemlere kıyasla, q -aktivasyonların daha tutarlı ve ayar bağımlılığı düşük sonuçlar verdiği rapor edilmiştir. Sonuç olarak, q -aktivasyon

yaklaşımı, derin öğrenme modellerine basit biçimde entegre edilebilen ve genelleme yeteneğini artıran etkili bir yöntem olarak sunulmuştur.



3. KURAMSAL TEMELLER

3.1. Temel Tanım ve Kavramlar

Bu bölümde, yapılan çalışmaların temelini oluşturan kuramsal ifadeler verilmiştir. Verilen tanımlar ise reel bir Hilbert uzay çerçevesinde ele alınmıştır.

Tanım 3.1.1: X, K cismi üzerinde bir vektör uzayı olmak üzere her $\rho, \varphi, \omega \in X$ ve $\lambda \in K$ için,

- i. $\langle \rho, \rho \rangle \geq 0$ ve $\langle \rho, \rho \rangle = 0 \Leftrightarrow \rho = 0$,
- ii. $\langle \rho, \varphi \rangle = \overline{\langle \varphi, \rho \rangle}$,
- iii. $\langle \lambda \rho, \varphi \rangle = \lambda \langle \rho, \varphi \rangle$,
- iv. $\langle \rho + \varphi, \omega \rangle = \langle \rho, \omega \rangle + \langle \varphi, \omega \rangle$

şartını sağlayan $\langle \cdot, \cdot \rangle: X \times X \rightarrow K$ fonksiyonuna X üzerinde bir iç çarpım, bu fonksiyon ile donatılmış olan $(X, \langle \cdot, \cdot \rangle)$ yapısına ise bir iç çarpım uzayı denir. Tanımda (ii) ile verilen özellik, temel alınan cismin reel sayılar yani $K = \mathbb{R}$ olması durumunda $\langle \rho, \varphi \rangle = \langle \varphi, \rho \rangle$ simetri koşuluna indirgenir (Bayraktar 2006).

Tanım 3.1.2: $(X, \langle \cdot, \cdot \rangle)$ bir iç çarpım uzayı ve $\rho \in X$ olmak üzere $\|\rho\| = \sqrt{\langle \rho, \rho \rangle}$ fonksiyonu bir norm fonksiyonudur. Bu norma iç çarpım normu denir (Bayraktar 2006).

Tanım 3.1.3: $(X, \langle \cdot, \cdot \rangle)$ iç çarpım uzayı iç çarpım normuna göre tam ise bu uzaya Hilbert uzay denir (Bayraktar 2006).

Tanım 3.1.4: $\mathcal{H}$ bir Hilbert uzay, $f: \mathcal{H} \rightarrow [-\infty, +\infty]$ bir dönüşüm ve $D(f) \subset \mathcal{H}$, f 'nin tanım kümesi olmak üzere $f(\rho) < +\infty$ olacak şekilde en az bir $\rho \in D(f)$ varsa f fonksiyonuna has (proper) fonksiyon denir (Bauschke and Combettes 2017).

Tanım 3.1.5: C, V vektör uzayının bir alt kümesi olsun. Eğer C kümesi her $\rho_1, \rho_2 \in C$ ve her $\theta \in [0,1]$ için

$$\theta\rho_1 + (1 - \theta)\rho_2 \in C$$

şartını sağlarsa C ye V de bir konveks küme denir (Boyd and Vandenberghe 2004).

Tanım 3.1.6: $f: A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ bir fonksiyon olsun. Eğer A kümesi konveks bir küme ve her $\rho, \varphi \in A$ ve her $\theta \in [0,1]$ için

$$f(\theta\rho + (1 - \theta)\varphi) \leq \theta f(\rho) + (1 - \theta)f(\varphi)$$

eşitsizliği sağlanıyorsa f fonksiyonuna konveks fonksiyon denir (Boyd and Vandenberghe 2004).

Tanım 3.1.7: $f: \mathbb{R}^n \rightarrow \mathbb{R}$ bir fonksiyon olsun. Eğer her $\bar{\rho} \in \mathbb{R}^n$ için

$$\liminf_{\rho \rightarrow \bar{\rho}} f(\rho) \geq f(\bar{\rho})$$

eşitsizliği sağlanıyorsa f fonksiyonu $\mathbb{R}^n$ de alttan yarı sürekli denir (Rockafellar and Wets 1998).

Tanım 3.1.8: $\mathcal{H}$ bir Hilbert uzay, $f: \mathcal{H} \rightarrow (-\infty, +\infty]$ aralığından tanımlı has, alttan yarı sürekli konveks bir fonksiyon olsun. ve $\rho \in \mathcal{H}$ olsun. Bu durumda f fonksiyonunun $\rho \in \mathcal{H}$ noktasındaki proksimal (yakınlık) operatörü $prox_f \rho$

$$prox_f \rho = \operatorname{argmin}_{\varphi \in \mathcal{H}} \left\{ f(\varphi) + \frac{1}{2} \|\varphi - \rho\|^2 \right\}$$

şeklinde tanımlanır (Bauschke and Combettes 2017).

Örnek 3.1.9: a , Hilbert uzayının bir elemanı ve β ise reel bir sabit olmak üzere; sürekli ve doğrusal $f(\varphi) = \langle a, \varphi \rangle - \beta$ fonksiyonu tanımlansın. Proksimal operatörün tanımını gereği, herhangi bir ρ noktası için

$$\text{prox}_f \rho = \underset{\varphi \in \mathcal{H}}{\operatorname{argmin}} \left\{ f(\varphi) + \frac{1}{2} \|\varphi - \rho\|^2 \right\}$$

şeklinde ifade edilmektedir. Bu durumda, aşağıdaki fonksiyonun minimumunu sağlayan

$$\delta(\varphi) = \langle a, \varphi \rangle - \beta + \frac{1}{2} \|\varphi - \rho\|^2$$

φ değeri aranır. Minimumu elde etmek için, $\delta(\varphi)$ fonksiyonunun φ 'ya göre türevi alınmakta ve sıfıra eşitlenmektedir. Buradan

$$\nabla_{\varphi} \delta(\varphi) = a + (\varphi - \rho)$$

ve

$$a + (\varphi - \rho) = 0 \Rightarrow \varphi = \rho - a$$

sonucu elde edilir. Dolayısıyla, $f(\varphi) = \langle a, \varphi \rangle - \beta$ için, ρ noktasındaki proksimal operatör

$$\text{prox}_f(\rho) = \rho - a$$

şeklinde bulunmuş olur. Elde edilen bu sonuç, Hilbert uzayında doğrusal bir fonksiyon için proksimal operatörün, giriş değerinden a vektörünün çıkarılmasıyla uygulanan bir öteleme işlemi olduğunu göstermektedir. Sonuç olarak, Hilbert uzayında yapılan her ötelemenin bir proksimal operatör uygulaması olduğu ifade edilebilir (Moreau 1965).

Tanım 3.1.10: $\mathcal{H}$ bir Hilbert uzay olsun. Her $\rho \in \mathcal{H}$ için

$$T\rho = \begin{cases} \left(1 - \frac{1}{\|\rho\|}\right)\rho, & \text{eğer } \|\rho\| > 1 \\ 0, & \text{eğer } \|\rho\| \leq 1 \end{cases}$$

şeklinde tanımlı $T: \mathcal{H} \rightarrow \mathcal{H}$ dönüşümüne yumuşak eşikleme (soft thresholding) operatörü denir. Bu operatör, normu 1'den büyük olan vektörlerin büyüklüğünü 1 birim azaltarak orijine doğru çekerken yönlerini korur. Normu 1'e eşit veya daha küçük olan vektörleri ise doğrudan sıfır vektörüne eşler.

Tanım 3.1.11: $C \neq \emptyset$, $\mathcal{H}$ Hilbert uzayının kapalı ve konveks bir alt kümesi ve $\rho \in \mathcal{H}$ olsun. $\|\rho - w\| = \inf \{\|\rho - \varphi\|: \varphi \in C\}$ olacak şekilde bir $w \in C$ noktası varsa bu w noktasına ρ 'nun C üzerine metrik izdüşümü denir ve $P_C\rho$ ile gösterilir. $P_C: \mathcal{H} \rightarrow C$ operatörüne de C üzerinde metrik izdüşüm operatörü denir (Cegielski 2012).

Tanım 3.1.12: $\mathcal{H}$ bir Hilbert uzay ve $f: \mathcal{H} \rightarrow (-\infty, +\infty]$ bir has fonksiyon olsun. f fonksiyonunun alt diferansiyeli (subdifferential), ∂f ile gösterilen ve her bir $\rho \in \mathcal{H}$ noktasına aşağıdaki koşulu sağlayan $u \in \mathcal{H}$ vektörlerinin kümesini atayan bir operatördür:

$$\partial f(\rho) := \{u \in \mathcal{H} \mid \forall \varphi \in \mathcal{H}, \quad f(\varphi) \geq f(\rho) + \langle \varphi - \rho, u \rangle\}.$$

Eğer bir $\rho \in \mathcal{H}$ noktasında $\partial f(\rho)$ kümesi boş değilse f fonksiyonu bu noktada alt diferansiyellenebilir kabul edilir. $\partial f(\rho)$ kümesini oluşturan her bir elemana ise f fonksiyonunun ρ noktasındaki bir alt gradyanı (subgradient) denir (Bauschke and Combettes 2017).

Örneğin $f(\rho) = |\rho|$ mutlak değer fonksiyonunun alt diferansiyeli

$$\partial f(\rho) = \begin{cases} \{-1\}, & \rho < 0 \text{ ise} \\ [-1,1], & \rho = 0 \text{ ise} \\ \{1\}, & \rho > 0 \text{ ise} \end{cases}$$

şeklindedir.

Tanım 3.1.13: $\mathcal{H}$ bir Hilbert uzay olmak üzere, $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$ küme değerli bir operatör ve $\gamma > 0$ pozitif bir skaler olsun. A operatörünün γ indeksli resolvent operatörü J_{λ}^A ile gösterilir

$$J_{\lambda}^A := (Id + \gamma A)^{-1}$$

şeklinde tanımlanır. Burada Id , $\mathcal{H}$ üzerindeki birim operatördür. Bu tanıma göre, bir $\rho \in \mathcal{H}$ için $\varphi = J_{\lambda}^A(\rho)$ eşitliği ile aşağıdaki ifade

$$\rho \in \varphi + \gamma A(\varphi)$$

birbirine denktir (Bauschke and Combettes 2017).

Tanım 3.1.14: $\mathcal{H}$ bir Hilbert uzay ve X , $\mathcal{H}$ 'in boş olmayan bir alt kümesi olsun. $T: X \rightarrow \mathcal{H}$ bir operatör olmak üzere her $\rho, \varphi \in X$ için

$$\|T(\rho) - T(\varphi)\| \leq K\|\rho - \varphi\|$$

eşitsizliğini sağlayan bir $K \geq 0$ varsa T ye Lipschitz sürekli denir. Bu eşitsizliği sağlayan en küçük K değerine, T operatörünün Lipschitz sabiti denir. Eğer $K < 1$ ise operatöre daralma (contraction), $K = 1$ ise genişlemeyen (nonexpansive) denir (Cegielski 2012). Diğer yandan T dönüşümünün en az bir sabit noktası olduğu kabul edilsin. $f \in F(T)$ için

$$\|T(\rho) - f\| \leq \|\rho - f\|$$

eşitsizliği sağlanırsa T dönüşümüne yarı genişlemeyen (quasi-nonexpansive) dönüşüm denir (Aleyner and Reich 2009).

$\mathcal{H}$ bir Hilbert uzay ve $C \subseteq \mathcal{H}$ kapalı ve konveks bir küme ise, P_C ile gösterilen metrik izdüşüm operatörü genişlemeyendir (Cegielski 2012).

Tanım 3.1.15: $T, \mathcal{H}$ Hilbert uzayında $D(T)$ tanım ve $R(T)$ görüntü kümelerine sahip lineer olmayan bir dönüşüm olsun. Bu durumda,

- i. Her $\rho, \varphi \in D(T)$ için, $\langle T\rho - T\varphi, \rho - \varphi \rangle \geq 0$ eşitsizliği sağlanıyorsa T 'ye monoton dönüşüm denir (Zeidler 2013).
- ii. Her $\rho, \varphi \in D(T)$ için, $\langle T\rho - T\varphi, \rho - \varphi \rangle \geq \eta \|\rho - \varphi\|^2$ olacak şekilde $\eta > 0$ sabiti varsa T 'ye η -kuvvetli monoton dönüşüm denir (Zeidler 2013).
- iii. Her $\rho, \varphi \in D(T)$ için, $\langle T\rho - T\varphi, \rho - \varphi \rangle \geq \alpha \|T\rho - T\varphi\|^2$ olacak şekilde $\alpha > 0$ sabiti varsa T 'ye α -ters kuvvetli monoton dönüşüm veya α -cocoercive dönüşüm denir. Tanımdan yola çıkarak, her α -cocoercive dönüşüm $\frac{1}{\alpha}$ -Lipschitz sürekli ve monoton dönüşüm olduğu söylenebilir (Cegielski 2012).
- iv. $S: \mathcal{H} \rightarrow \mathcal{H}$ genişlemeyen bir dönüşüm ve $\kappa \in (0,1)$ olsun. Bu durumda, $T = (1 - \kappa)I + \kappa S$ eşitliği ile tanımlanan T 'ye κ ortalamalı operatör denir (Cegielski 2012).
- v. Bir T monoton dönüşümün $Gph(T)$ grafiği başka bir monoton dönüşümün grafiği tarafından içerilmiyorsa T 'ye maksimal monoton dönüşüm denir (Rockafellar and Wets 1998).

Tanım 3.1.16: $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$ küme değerli bir operatör olsun. A operatörünün grafiği

$$Gph(A) := \{(\rho, z) \in \mathcal{H} \times 2^{\mathcal{H}} : z \in A\rho\}$$

şeklinde tanımlanır. Her $z \in A\rho$ ve $w \in A\varphi$ için

$$\langle z - w, \rho - \varphi \rangle \geq 0$$

eşitsizliği sağlanıyorsa A küme değerli operatörü monoton operatör olarak adlandırılır (Bauschke and Combettes 2017).

Önerme 3.1.17: Her $\rho, \varphi \in \mathcal{H}$ ve $\alpha \in \mathbb{R}$ için aşağıdakiler geçerlidir:

- i. $\|\rho + \varphi\|^2 \leq \|\rho\|^2 + 2\langle \varphi, \rho + \varphi \rangle,$
- ii. $\|\rho + \varphi\|^2 = \|\rho\|^2 + 2\langle \rho, \varphi \rangle + \|\varphi\|^2,$
- iii. $\|\alpha\rho + (1 - \alpha)\varphi\|^2 = \alpha\|\rho\|^2 + (1 - \alpha)\|\varphi\|^2 - \alpha(1 - \alpha)\|\rho - \varphi\|^2.$

Lemma 3.1.18: $\mathcal{H}$ reel bir Hilbert uzay olsun. Her $\rho, \varphi \in \mathcal{H}$ ve $\theta \in [0,1]$ için aşağıdaki ifadeler sağlanır:

- i. $\|\varphi - \rho\|^2 = \|\varphi\|^2 - \|\rho\|^2 - 2\langle \varphi - \rho, \rho \rangle,$
- ii. $\|\varphi - \rho\|^2 \leq \|\varphi\|^2 + 2\langle \rho, \varphi + \rho \rangle,$
- iii. $\|\theta\varphi + (1 - \theta)\rho\|^2 = \theta\|\varphi\|^2 + (1 - \theta)\|\rho\|^2 - \theta(1 - \theta)\|\varphi - \rho\|^2$

(Bauschke and Combettes 2017).

Lemma 3.1.19: $\mathcal{H}$ reel bir Hilbert uzay olmak üzere $A: \mathcal{H} \rightarrow \mathcal{H}$ α -ters kuvvetli monoton bir operatör ve $B: \mathcal{H} \rightarrow 2^{\mathcal{H}}$ maksimal monoton bir operatör olsun. $T_\lambda = J_\lambda^B(Id - \lambda A)$ olmak üzere

- i. $\lambda > 0$ için $F(T_\lambda) = (A + B)^{-1}(0)$ eşitliği geçerlidir.
- ii. $0 < \tau \leq \lambda$ ve $z \in \mathcal{H}$ için $\|z - T_\tau z\| \leq 2\|z - T_\lambda z\|$

dır (López et al. 2012).

Lemma 3.1.20: $\{a_n\}$ sıfırdan büyük reel sayılardan oluşan bir dizi olsun ve aşağıdaki bağıntıyı sağlasın:

$$a_{n+1} = (1 - \alpha_n)a_n + \alpha_n b_n + c_n, \quad \forall n \in \mathbb{N}.$$

Burada $\{a_n\}, \{b_n\}$ ve $\{c_n\}$ reel sayı dizileri olup aşağıdaki koşulları sağlamaktadır:

- i. $\{a_n\} \subset [0,1]$ ve $\sum_{n=1}^{\infty} \alpha_n = \infty$,
- ii. $\limsup_{n \rightarrow \infty} b_n < \infty$,
- iii. $c_n \geq 0$ ve $\sum_{n=1}^{\infty} c_n < \infty$.

Bu durumda $a_n \rightarrow 0$ dır (Marino and Xu 2004).

Lemma 3.1.21: $\mathcal{H}$ reel bir Hilbert uzay, $T: \mathcal{H} \rightarrow \mathcal{H}$ genişlemeyen bir dönüşüm ve $F(T) \neq \emptyset$ olsun. Bu durumda $Id - T$ dönüşümü sıfırda yarı kapalıdır (demiclosed), yani $\mathcal{H}$ içinde herhangi bir $\{\rho_n\}$ dizisi için, $\rho_n \rightarrow \rho \in \mathcal{H}$ ve $\|\rho_n - T\rho_n\| \rightarrow 0$ ise $\rho \in F(T)$ dir (Puangpee and Suantai 2020).

Lemma 3.1.22: (S_n) azalmayan ve reel sayılardan oluşan bir dizi olsun. Bu durumda (S_n) dizisinin, $S_{n_j} < S_{n_{j+1}}$ eşitsizliğini sağlayan en az bir $(S_{n_j})_{j \geq 0}$ alt dizisi mevcuttur. Ayrıca, $(\tau(n))_{n \geq n_0}$ tamsayılar dizisi ise

$$\tau(n) = \max\{k \leq n \mid S_k < S_{k+1}\}$$

şeklinde tanımlansın. Bu durumda $(\tau(n))_{n \geq n_0}$ monoton artan bir dizi olup,

$$\lim_{n \rightarrow \infty} \tau(n) = \infty$$

dur. Ayrıca her $n \geq n_0$ için

$$S_{\tau(n)} \leq S_{\tau(n)+1}$$

ve

$$S_n \leq S_{\tau(n)+1}$$

eşitsizlikleri geçerlidir (Maingé, 2008).

3.2. Monoton Kapsama Problemi

Monoton kapsama problemi, konveks minimizasyon problemi ve sabit nokta problemi birbirleri ile matematikte birbirleri ile yakın ilişkeli olan problemlerdir. Bu problemler arasındaki bağlantılar aracılığı ile farklı problem türleri birbirine indirgenmektedir. Bunlar içerisinde monoton kapsama problemi 20. yüzyılın ikinci yarısında fonksiyonel analizin gelişimiyle birlikte Minty (1962) ve Browder'in (1967) çalışmalarıyla monoton operatörlerin sıfır noktalarını bulma problemi olarak formüle edilmiş ve matematiksel altyapısı oluşturulmuştur.

Tanım 3.2.1: $\mathcal{H}$ bir reel Hilbert uzay olsun. A ve B , $\mathcal{H}$ üzerinde tanımlı monoton operatörler olmak üzere, iki operatörün toplamı için monoton kapsama problemi,

$$0 \in (A + B)z \tag{3.1}$$

koşulunu sağlayan bir $z \in \mathcal{H}$ noktasının bulunması problemidir. Bu problemin çözüm kümesi $(A + B)^{-1}(0)$ ile gösterilir (Minty 1962).

Monoton kapsama problemi; konveks minimizasyon problemi, varyasyonel eşitsizlik problemi ve denge problemlerini içerir. Aynı zamanda makine öğrenmesi, görüntü işleme, sinyal işleme ve lineer ters problemler matematiksel olarak kapsama problemi şeklinde formülize edilebilir (Kitkuan et al. 2019). Monoton kapsama probleminin çözümü için genel olarak A 'nın ters kuvvetli monoton operatör ve B 'nin ise

maksimal monoton operatör olduğu kabul edilir. Nümerik olarak çözüme ulaşmak için en temel yöntem proksimal nokta algoritmasıdır. Bu algoritma, sabit nokta iterasyonuna dayanmaktadır.

Örnek 3.2.2: $\mathcal{H} = \mathbb{R}$ ve $c > 0$ olmak üzere $f(\rho) = \frac{1}{2}c\rho^2$ konveks fonksiyonunun minimizasyon problemini ele alalım. Bu problem, $0 \in \partial f(\rho)$ monoton kapsama problemini çözmeye denktir. Buradan hareketle $A(\rho) = \partial f(\rho) = \{c\rho\}$ olur. Bu operatör, tek değerli ve maksimal monotondur. Problemin analitik çözümü açıkça $\rho^* = 0$ 'dır. Resolvent hesabı için $\lambda > 0$ için J_λ^A resolventi

$$\varphi = (I + \lambda A)(\rho)$$

$$= \rho + \lambda(c\rho)$$

$$= (1 + \lambda c)\rho$$

$$\rho = \frac{\varphi}{1 + \lambda c}$$

dir. Dolayısıyla, $J_\lambda^A(\varphi) = \frac{\varphi}{1 + \lambda c}$ olarak bulunur. Picard iterasyonuna göre algoritma, $\rho_{k+1} = J_\lambda^A(\rho_k)$ şeklindedir. Yani

$$\rho_{k+1} = \frac{\rho_k}{1 + \lambda c}$$

dir. Şimdi sayısal uygulama için parametrelerimiz olan $c = 1$, $\lambda = 1$ ve başlangıç noktası $\rho_0 = 10$ olarak seçilsin. Bu durumda iterasyon $\rho_{k+1} = \rho_k/2$ olur. Bu iterasyonun bazı adımları Çizelge 3.1. de verilmiştir.

Çizelge 3.1. Örnek 3.2.2 de verilen iterasyonun adımları

k	ρ_k
1	5
2	2.5
3	1.25
4	0.625
5	0.3125
10	0.009765625
20	9.5367431640625E-06
30	9.313225746154785E-09
40	9.094947017729282E-12
50	8.881784197001252E-15
100	7.888609052210118E-30

Görüldüğü üzere, üretilen dizi $\{\rho_k\}$ geometrik olarak problemin bilinen çözümü olan $\rho^* = 0$ noktasına yakınsamaktadır.

Konveks minimizasyon problemi ise 1940'larda matematiksel optimizasyon alanının temel dallarından biri olan doğrusal programlamanın gelişimi ile başlayıp, 1950-1960'lı yıllarda konveks analizin kurulması ile şekillenmiştir. Bu alanın teorik çerçevesi, Fenchel (1949), Moreau (1962) ve Rockafellar (1970b) gibi matematikçilerin gerçekleştirdiği çalışmalar ile oluşturulmuştur.

$\mathcal{H}$ bir reel Hilbert uzay olsun. $f: \mathcal{H} \rightarrow \mathbb{R}$ sürekli türevlenebilir bir konveks fonksiyon ve ∇f gradyan operatörü L-Lipschitz sürekli olsun. Ayrıca, $g: \mathcal{H} \rightarrow (-\infty, +\infty]$

alttan yarı-süreklı ve konveks bir fonksiyon olsun. Bu durumda konveks minimizasyon problemi

$$\min_{\rho \in \mathcal{H}} \{F(\rho) := f(\rho) + g(\rho)\}$$

şeklinde verilen iki fonksiyonun toplamının minimumunu bulma problemidir. Bu problemin bir $\rho^* \in \mathcal{H}$ çözümünü bulmak, $F(\rho^*) \leq F(\rho)$ koşulunun her $\rho \in \mathcal{H}$ için sağlanması anlamına gelmektedir.

Fermat kuralından, ρ^* noktasının konveks minimizasyon probleminin bir çözümü olması için gerek ve yeter şart F 'nin alt diferansiyelinin sıfırı içermesidir. Yani $0 \in \partial F(\rho^*)$ olmasıdır. f fonksiyonunun türevlenebilir olması ve alt diferansiyel kavramından

$$\partial F(\rho^*) = \nabla f(\rho^*) + \partial g(\rho^*)$$

olup, minimizasyon problemimiz

$$0 \in \nabla f(\rho^*) + \partial g(\rho^*)$$

şeklinde ifade edilen bir monoton kapsama problemine dönüşür. Bu ifade aslında (3.1) ile verilen problemde $A = \nabla f$ ve $B = \partial g$ olarak alınmış özel bir halidir.

Konveks minimizasyon ve sabit nokta problemleri arasındaki geçiş ise proximal operatörler kullanılarak yapılır. Sabit nokta problemleri Brouwer'ın (1911) geliştirdiği sabit nokta teoremi ile başlatılmış ardından Banach'ın (1922) daralma dönüşümü prensibi ile güçlendirilmiş ve Browder (1965) ile Kirk'ün (1965) çalışmaları aracılığıyla monoton operatör teorisiyle bağlantısı kurulmuştur. Sabit nokta problemleri ve monoton kapsama problemleri arasındaki bağlantı ise resolvent operatörlerinin sabit noktalarının monoton operatörün sıfır noktalarına karşılık gelmesi prensibi ile açıklanmaktadır.

Boş kümeden farklı bir X kümesi ve bu küme üzerinde tanımlı bir $T: X \rightarrow X$ dönüşümü verildiğinde, sabit nokta problemi,

$$T(\rho) = \rho$$

eşitliğini sağlayan bir $\rho \in X$ elemanı bulmayı amaçlar. Bu koşulu sağlayan tüm noktaların oluşturduğu küme, T 'nin sabit nokta kümesi olarak adlandırılır ve

$$F(T) = \{\rho \in X \mid T(\rho) = \rho\}$$

şeklinde gösterilir (Agarwal et al. 2007).

Birçok durumda, $0 \in (A + B)z$ gibi bir monoton kapsama problemi, doğrudan çözülmesi zor olduğu için, eşdeğer bir sabit nokta problemine dönüştürülür. $A: \mathcal{H} \rightarrow \mathcal{H}$ tek değerli bir operatör ve $B: \mathcal{H} \rightarrow 2^{\mathcal{H}}$ küme değerli ve maksimal monoton bir operatör olmak üzere $(A + B)^{-1}(0)$ çözüm kümesini bulmak,

$$T = J_{\lambda}^B(Id - \lambda A)$$

şeklinde tanımlanan bir T operatörünün sabit noktalarını bulma problemine denktir. Burada J_{λ}^B , B operatörünün resolvent operatörü, Id birim operatör ve $\lambda > 0$ ise adım büyüklüğüdür.

Sonuç olarak yukarıda verilen geçişler sayesinde monoton kapsama problemleriyle konveks minimizasyon problemleri ve sabit nokta problemleri arasındaki bağlantı kurulmuş olur. Böylece herhangi bir problem türü başka bir problem türüne dönüştürülebilir. Bu dönüşüm ise optimizasyon teorisinin gelişiminde önemli katkılar sağlamaktadır.

3.3. Quantum Türev (q -Türev)

q -türev kavramı, adi türevin bir genelleştirilmesi olarak matematikte önemli bir yer tutmaktadır. Jackson (1908) tarafından yapılan çalışma öncesinde ilk olarak Thomae (1870) ve Heine (1878) tarafından q -hipergeometrik serilerin ve bazı q -fonksiyonlarının özellikleri incelenmiştir. Yine ilerleyen dönemlerde q -fark operatörlerinin temel cebirsel

yapıları ve q -integralin, klasik integral ile ilişkisi araştırılmıştır. Ancak q -integral ile kurulan ilişkiler ise yeterince netleştirilememiştir.

Jackson (1908) tarafından ise q -türev ve q -fark denklemleri teorisi ortaya konulmuştur. q -türev operatörünün tanımı, temel özellikleri, adi türev ile ilişkisi ve limit geçişleri ayrıntılı biçimde ortaya konulmuştur. Bu bağlamda q -parametresinin $q \rightarrow 1$ için limitinin alınması durumunda fonksiyonun adi türevi elde edilmektedir. Yine q -diferansiyel denklemlerinin çözümünde kullanılabilecek yeni ifadeler tanıtılmıştır. Ayrıca q -integral kavramının q -hipergeometrik serilerle olan ilişkisi ifade edilmiştir.

Bu bölümde çalışmalarımızda kullandığımız kuantum türev ve özellikleri verilecektir. Aşağıda ilk olarak q -diferansiyel kavramı verilmiştir.

Tanım 3.3.1: Herhangi bir $f(x)$ fonksiyonunu ele alalım.

$$d_q f(x) = f(qx) - f(x) \quad (3.2)$$

şeklinde tanımlanan $d_q f(x)$ ifadesine f in q -diferansiyeli denir. İki fonksiyonun çarpımının q -diferansiyeli ise

$$\begin{aligned} d_q(f(x)g(x)) &= f(qx)g(qx) - f(x)g(x) \\ &= f(qx)g(qx) - f(qx)g(x) + f(qx)g(x) - f(x)g(x) \\ &= f(qx)d_q g(x) + g(x)d_q f(x) \end{aligned} \quad (3.3)$$

şeklinde tanımlanır (Jackson,1908).

Şimdi q -diferansiyel kavramını kullanarak q -türev kavramını tanımlayalım.

Tanım 3.3.2: $f(x)$ fonksiyonunun q -türevi $q \neq 1$ olmak üzere

$$D_q f(x) = \begin{cases} \frac{f(qx) - f(x)}{(q-1)x}, & x \neq 0, \\ f'(0), & x = 0 \text{ ise} \end{cases}$$

şeklinde tanımlanır (Jackson,1908).

Benzer şekilde, $f(x)$, n değişkenli türevlenebilir bir fonksiyonu olsun. O hâlde, f fonksiyonunun x_i değişkenine göre birinci mertebeden kısmi q -türevi

$$\frac{\partial_{qi} f(x_i)}{\partial_{qi} x_i} = \begin{cases} \frac{f(x_1, \dots, q_i x_i, \dots, x_n) - f(x_1, \dots, x_i, \dots, x_n)}{q_i x_i - x_i}, & x_i \neq 0 \text{ ve } q_i \neq 1, \\ \frac{\partial f(x)}{\partial x_i}, & \text{aksi halde} \end{cases}$$

şeklinde tanımlanır. Buna göre, q -gradyanı, f fonksiyonunun n tane birinci mertebeden kısmi q -türevinden oluşan vektör olup

$$\nabla_q f(x)^T = \left[\frac{\partial_{q1} f(x_1)}{\partial_{q1} x_1} \dots \frac{\partial_{qi} f(x_i)}{\partial_{qi} x_i} \dots \frac{\partial_{qn} f(x_n)}{\partial_{qn} x_n} \right]$$

şeklinde tanımlanır. Burada q parametresi, $q = (q_1, \dots, q_i, \dots, q_n) \in \mathbb{R}^n$ şeklinde bir vektördür. Her $i = 1, \dots, n$ için $q_i \rightarrow 1$ limiti alındığında klasik gradyan elde edilir (Soterroni et al. 2011)

Eğer $f(x)$ türevlenebilir ise q değerinin aşağıdaki gibi seçilmesi neticesinde q -türev adi türeve indirgenmiş olur:

$$\lim_{q \rightarrow 1} D_q f(x) = \frac{df(x)}{dx}.$$

Tıpkı klasik türevde olduğu gibi, bir fonksiyonun q türevi bir doğrusal operatördür. Başka bir deyişle, herhangi sabit a ve b için

$$D_q(af(x) + bg(x)) = aD_qf(x) + bD_qg(x)$$

ifadesi sağlanır.

Örnek 3.3.3: $f(x) = x$ olsun.

$$D_qf(x) = \frac{f(x) - f(qx)}{x - qx} = \frac{x - qx}{x - qx} = 1$$

bulunur.

Örnek 3.3.4: $f(x) = x^2$ olsun.

$$D_qf(x) = \frac{f(x) - f(qx)}{x - qx} = \frac{x^2 - (qx)^2}{x - qx} = \frac{(x - qx)(x + qx)}{x - qx} = (1 + q)x$$

bulunur. $q \rightarrow 1$ için sonucun $2x$ olduğu görülür.

Örnek 3.3.5: $f(x) = x^3$ olsun.

$$\begin{aligned} D_qf(x) &= \frac{f(x) - f(qx)}{x - qx} \\ &= \frac{x^3 - (qx)^3}{x - qx} \\ &= \frac{(x - qx)(x^2 + x^2q + x^2q^2)}{x - qx} \\ &= (1 + q + q^2)x^2 \end{aligned}$$

bulunur. $q \rightarrow 1$ için sonucun $3x^2$ olduğu görülür.

Örnek 3.3.6: n pozitif bir tamsayı olmak üzere $f(x) = x^n$ fonksiyonunun q -türevi

$$D_q x^n = \frac{(qx)^n - x^n}{(q-1)x} = \frac{q^n - 1}{q-1} x^{n-1} \quad (3.4)$$

şeklinde bulunur. Yukarıdaki ifadede $\frac{q^n - 1}{q-1}$ terimi sıklıkla geçtiğinden, bu ifade ortak çarpan parantezine alınırsa aşağıdaki gibi

$$[n] = \frac{q^n - 1}{q - 1} = q^{n-1} + \dots + 1.$$

yazılır. Bu gösterime n 'nin q -analogu adı verilir. Böylece (3.4) eşitliği

$$D_q x^n = [n]x^{n-1}$$

halini alır. Bu ifade, x^n in klasik türevine benzemektedir. $q \rightarrow 1$ olması durumunda, $[n] = q^{n-1} + \dots + 1 \rightarrow n$ olur. Bu durumda, $[n]$ ifadesi q -kalkülüste, n 'nin klasik analizdeki rolünü üstlenir.

$f(x)$ ve $g(x)$ fonksiyonlarının çarpımının ve bölümünün q -türevleri (3.3) ifadesi de kullanılarak sırasıyla

$$D_q(f(x)g(x)) = \frac{d_q(f(x)g(x))}{(q-1)x} = \frac{f(qx)d_q g(x) + g(x)d_q f(x)}{(q-1)x} \quad (3.5)$$

yazılır. Şimdi her bir terimi q -türev cinsinden yeniden yazalım:

$$D_q g(x) = \frac{d_q g(x)}{(q-1)x} \Rightarrow d_q g(x) = (q-1)x D_q g(x)$$

$$D_q f(x) = \frac{d_q f(x)}{(q-1)x} \Rightarrow d_q f(x) = (q-1)x D_q f(x)$$

Bu ifadeleri yukarıdaki (3.5) eşitliğine yerleştirilirse

$$D_q(f(x)g(x)) = \frac{f(qx)[(q-1)x D_q g(x)] + g(x)[(q-1)x D_q f(x)]}{(q-1)x}$$

elde edilir. Ortak olan $(q-1)x$ ifadeleri sadeleştirilirse

$$D_q(f(x)g(x)) = f(qx)D_q g(x) + g(x)D_q f(x) \quad (3.6)$$

bulunur. Başlangıçta

$$d_q(f(x)g(x)) = g(qx)d_q f(x) + f(x)d_q g(x)$$

olup bunu q -türev ile yazarsak

$$D_q(f(x)g(x)) = \frac{g(qx)d_q f(x) + f(x)d_q g(x)}{(q-1)x}$$

ifadesine ulaşılır. Aynı şekilde, $d_q f(x)$ ve $d_q g(x)$ 'yi yukarıda gösterildiği gibi yerine konulursa:

$$\begin{aligned} D_q(f(x)g(x)) &= \frac{g(qx)[(q-1)x D_q f(x)] + f(x)[(q-1)x D_q g(x)]}{(q-1)x} \\ &= g(qx)D_q f(x) + f(x)D_q g(x) \end{aligned} \quad (3.7)$$

olur. Ayrıca bu ifade (3.6) ile eşdeğerdir. Şimdi (3.6) ifadesini kullanarak

$$g(x) \cdot \frac{f(x)}{g(x)} = f(x)$$

ifadesinin q -türevi

$$g(qx)D_q\left(\frac{f(x)}{g(x)}\right) + \frac{f(x)}{g(x)}D_qg(x) = D_qf(x)$$

dir. Dolayısıyla:

$$D_q\left(\frac{f(x)}{g(x)}\right) = \frac{g(x)D_qf(x) - f(x)D_qg(x)}{g(x)g(qx)} \quad (3.8)$$

olur. Öte yandan benzer şekilde (3.7) ifadesi kullanılırsa

$$g(x)D_q\left(\frac{f(x)}{g(x)}\right) + \frac{f(qx)}{g(qx)}D_qg(x) = D_qf(x)$$

elde edilir. Buradan da

$$D_q\left(\frac{f(x)}{g(x)}\right) = \frac{g(qx)D_qf(x) - f(qx)D_qg(x)}{g(x)g(qx)} \quad (3.9)$$

olur. Yukarıdaki (3.8) ve (3.9) eşitliklerinin her ikisi de geçerlidir. Karşılaşılan problem durumuna göre birinin yerine diğeri kullanılabilir.

Çarpımın ve bölümün q -türev ifadelerinden sonra, q -türevde zincir kuralının olup olmadığı sorusu merak edilebilir. Bilinmesi gerekir ki q -türevler için genel bir zincir kuralı yoktur. Ancak istisna olarak α, β sabit sayılar ve $z(x) = \alpha x^\beta$ olmak üzere $f(z(x))$ fonksiyonun zincir kuralı mevcuttur.

3.4. Yapay Sinir Ağlarının Temelleri

Yapay sinir ağı (YSA) bilgisayarların öğrenme süreçlerini konu edinen bir araştırma alanıdır. Günümüz dünyasında bilgisayar sistemlerinde, mobil cihazlarda, otonom ve askeri sistemlerde ve hatta en yakınımızdaki ev aletlerinde dahi kullanılarak günlük yaşantımızın ayrılmaz bir parçası haline gelmiştir. Bu sistemler iş dünyasından kamu hizmetlerine, sağlık organizasyonlarından ekonomik uygulamalara kadar sayamayacağımız geniş bir çerçevede kullanılmaktadır (Öztemel 2003).

Yapay zekâ çalışmaları ile matematiksel formülasyonu kurulamayan veya çözümü mümkün olmayan problemler sezgisel yöntemlerle bilgisayarlar tarafından çözülebilmektedir (Öztemel 2003). İlk olarak 1950'li yıllarda ortaya atılan yapay zekâ kavramı, geçen zaman içinde hayatın vazgeçilmez bir parçası haline gelmiş ve yeni teknolojilerin doğmasına öncülük etmiştir.

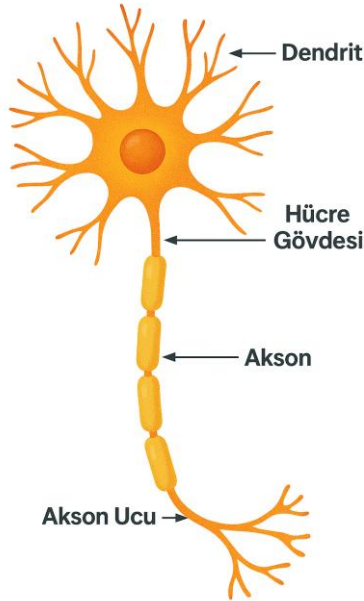
Yapay zekâyâ dair tartışmaların Türkiye'deki yansımaları da erken bir tarihte görülmüştür. Örneğin, Arf (1959) eğitim öğretim yılı açılış töreninde sunduğu "Makine Düşünebilir mi ve Nasıl Düşünebilir?" başlıklı konferansında, düşünen makinelerin ve elektronik beyinlerin önemine dikkat çekmiştir. Arf, makinelerin belirli girdilere verdiği yanıtlar ile insanların çevresel uyaranlara gösterdiği tepkileri arasında paralellikler kurmuştur. Bununla birlikte, insan beyninin en ayırt edici özelliğinin yeni ve öngörülemeyen problemlere uyum sağlama yeteneği olduğunu ifade etmiştir. Dönemin makinelerinin ise sadece önceden programlanmış görevleri yerine getirebildiğini vurgulanmıştır. Elektronik makinelerin hesaplama hızına rağmen insan beyninin yetenekleri olan öğrenme, estetik algı veya inisiyatif kullanma gibi karmaşık zihinsel görevleri çözmeye yaklaşmadığı ifade edilmiştir.

Makine öğrenmesinin temelini oluşturan "öğrenme" kavramı, Simon (1990) tarafından "zaman içinde yeni bilgilerin keşfedilmesi yoluyla davranışların iyileştirilmesi süreci" olarak tanımlanmıştır. Makine öğrenmesi ise öğrenme sürecinin aslında bilgisayarlar tarafından gerçekleştirilmesidir. Burada temel amaç bilgisayarların insanlar gibi zamanla tecrübe kazanarak bir konu hakkındaki gelecekteki benzer olaylar hakkında kararlar verebilmesi ve karşılaşılan problemlere çözümler üretebilmesidir. Bilgisayarların

bu yeteneği kazanması ancak konu hakkında yeterli veriye sahip olmasına bağlıdır. Yapay sinir ağları söz konusu olduğunda bu veriler genellikle örnekler aracılığıyla sağlanır.

YSA öğrenme, ilişkilendirme, sınıflandırma, genelleme ve optimizasyon gibi alanlarda başarılı uygulamalara sahiptir. Bu süreçte kullanılan bilgisayar sistemleri, önce bir örneği analiz ederek belli başlı bilgileri alır, ardından ikinci bir örneği inceleyerek mevcut bilgisini iyileştirir. Bu süreç konuyla alakalı tüm örneklerin tekrar tekrar gözden geçirilmesiyle süreç yinelenir. Sonuç olarak sistemin artık olaylar hakkında genel çıkarımlar yapabilmesi beklenir. Bu yaklaşım aslında tecrübelerden öğrenme olarak görülebilir.

YSA'nın işleyişini anlamak için biyolojik sinir ağlarının yapısını anlamak gerekmektedir. Biyolojik sinir ağları, canlıların sinir sistemini oluşturan biyolojik nöronlardan meydana gelir. Her bir nöron, dendritler aracılığıyla aldığı elektriksel sinyalleri hücre gövdesinde (soma) işler ve akson vasıtasıyla diğer nöronlara iletir. Nöronlar arasındaki kimyasal iletişimin gerçekleştiği bu bağlantı noktalarına ise sinaps adı verilir.



Şekil 3.1. Temsili biyolojik nöron

Yapay sinir ağı (YSA) insan beynindeki biyolojik sinir ağlarının işleyişinden esinlenilerek geliştirilmiş matematiksel modeller olarak tanımlanmaktadır (McCulloch ve Pitts, 1943). Bu modellerin temel yapı taşı ise haliyle "yapay nöron" olarak adlandırılır. Bu nöronlar, karmaşık veri işleme görevlerini yerine getirmek üzere katmanlar halinde organize edilmektedir.

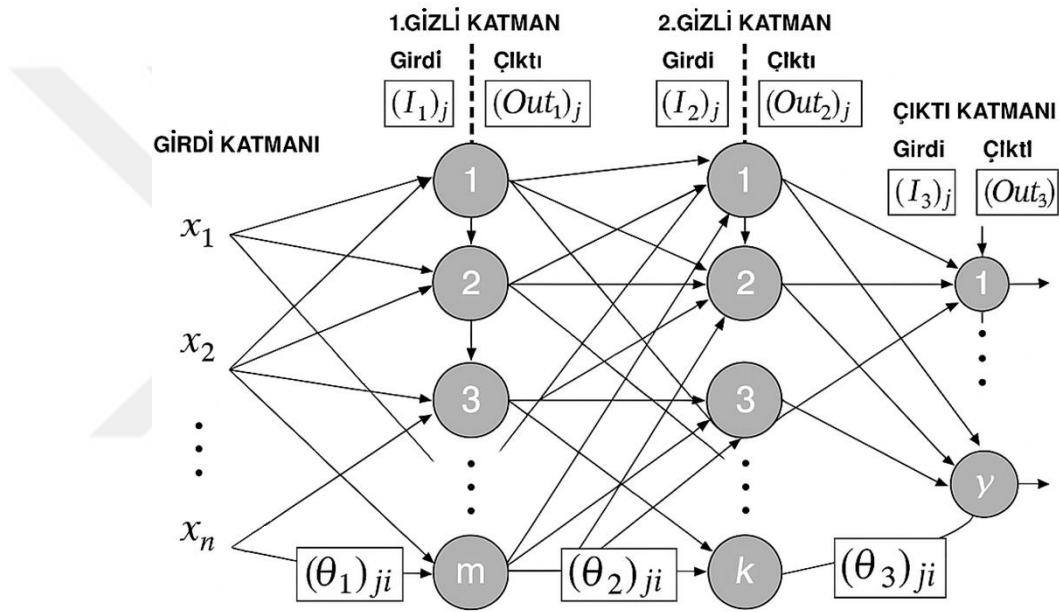
Günümüzde YSA pek çok probleme çözüm sunma potansiyeline sahip sistemler olarak kabul edilmektedir. Teknik olarak bir YSA'nın temel görevi, kendisine sunulan bir girdi setine karşılık gelebilecek uygun bir çıktı seti belirlemektir. YSA'nın karakteristik özellikleri, uygulanan ağ modeline göre farklılık gösterebilmektedir. Ancak literatürde var olan tüm modeller için geçerli sayılabilecek temel nitelikler mevcuttur. Bu nitelikler şu şekilde sıralanabilir:

- Bilgi, ağ bağlantılarının sayısal değerleri ile temsil edilir ve bu bağlantılarda saklanır.
- Öğrenme, bilinen programlama yöntemlerinden farklı olarak, sunulan örnekler üzerinden gerçekleştirilir.
- Güvenilir bir şekilde çalışabilmesi için sistemin öncelikle eğitilmesi ve performansının test edilmesi zorunludur.
- Daha önce görülmemiş örnekler hakkında genelleme yaparak bilgi üretebilme yeteneğine sahiptir.
- Örüntü ilişkilendirme, sınıflandırma ve tamamlama gibi algılamaya yönelik görevlerde kullanılabilir.
- Kendi kendini organize etme ve öğrenebilme gibi otonom yeteneklere sahiptir.
- Eksik, belirsiz veya tam olmayan bilgilerle çalışabilme ve hata toleransı gösterebilme kabiliyetleri vardır.

- İşlemlerini yalnızca nümerik veriler üzerinden yürütebilirler.

Yukarıda sıralanan özellikler ışığında YSA'nın, özellikle eksik, gürültülü ve belirsiz verileri işleme yeteneği ile aslında güçlü bir problem çözme tekniği olduğu ifade edilebilir.

Uğuz (2019) eseri baz alınarak aşağıda bir YSA'nın yapısı ve matematiksel modeli verilmiştir.



Şekil 3.2. İki gizli katmanlı bir yapay sinir ağı modeli (Uğuz 2019)

Şekil 3.2. de verilen modelin bir girdi katmanı, iki gizli katman ve bir çıktı katmanından oluştuğu görülmektedir. Parametrelerin hangi katmana ait olduğunun belirtilmesinde, alt indisler (parantez içinde) kullanılmaktadır. Bir yapay nörona ait giriş değeri " I " harfi ile, çıkış değeri ise " Out " ifadesi ile belirtilmektedir. Örneğin, $(I_2)_3$ ifadesi, ikinci gizli katmandaki üçüncü yapay nöronun girişini temsil ederken, $(Out_2)_1$ ifadesi ikinci gizli katmandaki birinci yapay nöronun çıktısını ifade etmektedir. Katmanlar arasındaki bağlantı ağırlıkları ise θ sembolü ile gösterilmektedir. Genel olarak $(\theta_2)_{ji}$ biçiminde yazılan bu ifade, ikinci gizli katmandaki j 'inci yapay nöron ile birinci gizli katmandaki i 'inci yapay nöron arasındaki ağırlığı belirtmektedir.

Şekil 3.2. de verilen modele göre ileri yayılım sürecinde hesaplanması gereken terimler

$$(I_1)_j = \sum_{i=1}^n (\theta_1)_{ji} \cdot x_i, \quad (Out_1)_j = f((I_1)_j),$$

$$(I_2)_j = \sum_{i=1}^m (\theta_2)_{ji} \cdot (Out_1)_i, \quad (Out_2)_j = f((I_2)_j),$$

$$(I_3)_j = \sum_{i=1}^k (\theta_3)_{ji} \cdot (Out_2)_i, \quad (Out_3)_j = f((I_3)_j),$$

eşitlikleri yardımıyla bulunur. Buna göre ağıın hatası

$$E(t) = \frac{1}{2} \sum_{j=1}^v \left(y_j^{(t)} - (Out_3)_j^{(t)} \right)^2,$$

formülü ile hesaplanır. Burada t numaralı eğitim örneği için $y_j^{(t)}$ ifadesi, hedeflenen gerçek çıkış değerini; $(Out_3)_j^{(t)}$ ise ağ tarafından tahmin edilen değeri göstermektedir.

YSA yapılarına göre ve öğrenme algoritmalarının türüne göre iki ana başlıkta sınıflandırılabilir. Bu başlıklar da kendi arasında yapılarına göre iki alt başlıkta incelenir; ileri beslemeli ve geri beslemelidir. Öğrenme algoritmalarına göre ise üç alt başlıkta incelenir; denetimli öğrenme, denetimsiz öğrenme ve pekiştirmeli (takviyeli) öğrenmedir.

Bizim çalışmalarımız denetimli öğrenme grubuna girmekte olup bu tip öğrenmede YSA kullanılmaya başlanmadan önce eğitilmektedir. Eğitim anında hem girdi değerleri hem de o girdi değerlerine karşılık istenen çıktı değerleri sisteme verilir. İstenilen çıktı değeri ile YSA'nın ürettiği çıktı karşılaştırılarak hata hesaplaması yapılır ve buna göre ağırlıklar güncellenir. Yine denetimli öğrenme; regresyon problemlerini (hava durumu tahmini, popülasyon büyümesi tahmini, bir ürüne ait ticari pazar tahmini vb.) ve

sınıflandırma problemlerini (tıpta tanı ve teşhis koyma, ticari olarak müşteri tutma, kriminal olarakta kimlik dolandırıcılığı tespiti, vb.) çözmek için kullanılır.

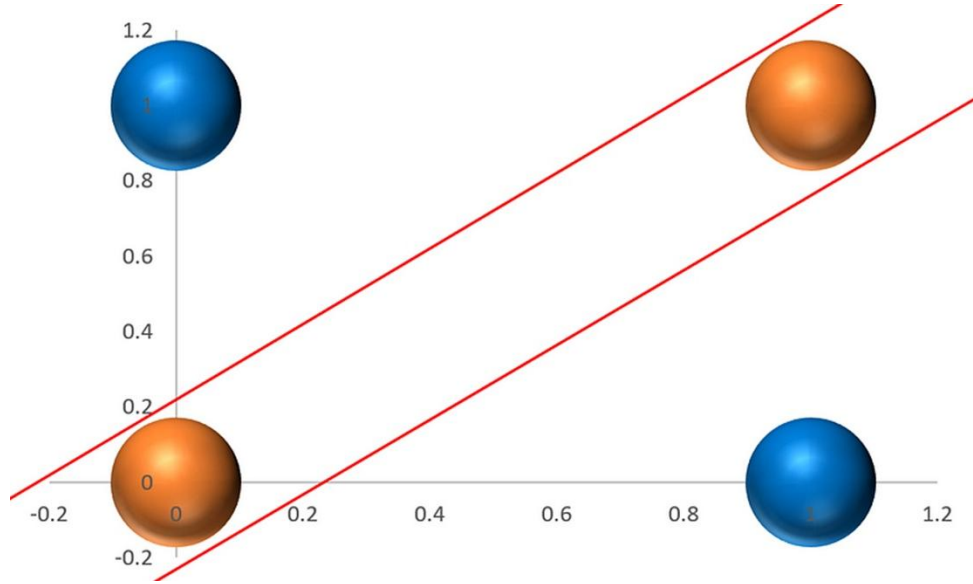
Makine öğrenmesi, veri analizi ve modelleme için istatistik, optimizasyon ve hesaplamalı yöntemlerin birleşiminden oluşan disiplinler arası bir alandır (Hastie et al. 2009). Makine öğrenmesi problemleri olarak özellikle XOR, regresyon ve sınıflandırma problemleri belli başlı problemlerdir. Regresyon problemleri, çıktı değerlerinin tahminine odaklanırken; sınıflandırma problemleri ise gözlemlerin önceden tanımlanmış kategorilere atanmasını amaçlar. Bu iki temel problem tipi, makine öğrenmesinin birçok uygulamasında karşılık bulmaktadır.

XOR problemi ise YSA ve makine öğrenmesi tarihinde bir dönüm noktası olarak kabul edilir. XOR problemi, yapay zekâ ve özellikle YSA alanında, ilk dönem modellerinin sınırlılıklarını ortaya koyan ve daha gelişmiş mimarilerin (derin öğrenmenin temelleri) doğmasına neden olan temel bir problemdir. Bu açıdan basit görünmesine rağmen tarihsel ve teorik önemi büyüktür. XOR aslında bir mantık operatörüdür. İki girdili bir XOR fonksiyonu, girdilerden yalnızca biri doğru (1) olduğunda doğru (1) çıktısı verir. Her iki girdi de aynıysa (ikisi de 0 veya ikisi de 1 ise) çıktı yanlıştır (0).

Çizelge 3.2. XOR doğruluk çizelgesi

Girdi A	Girdi B	A XOR B (Çıkış)
0	0	0
0	1	1
1	0	1
1	1	0

Doğruluk çizelgesinden yararlanılarak bu veriler koordinat sisteminde aşağıdaki gibi gösterilmiştir:



Şekil 3.3. Koordinat düzleminde XOR (Soruç 2020)

Problemin özü, XOR fonksiyonunun doğrusal olarak ayrılabilir olmasından kaynaklanır. Eğer bu dört veri noktasını Şekil 3.3.'teki gibi $((0,0), (0,1), (1,0), (1,1))$ iki boyutlu bir kartezyen düzleme yerleştirirsek, '1' çıktısı veren noktaları $((0,1)$ ve $(1,0))$ '0' çıktısı veren noktalardan $((0,0)$ ve $(1,1))$ tek bir düz çizgi ile ayırmanın imkânsız olduğunu görürüz.

1958 yılında sonunda Frank Rosenblatt tarafından geliştirilen perceptron (Algılayıcı), YSA'nın ilk ve en basit modellerinden biriydi (Rosenblatt 1958). Bir perceptron, girdileri alır, bunları belirli ağırlıklarla çarpar, bir eşik değeri (bias) değeri ekler ve bir basamak aktivasyon fonksiyonundan (step function) geçirerek bir çıktı üretir. Bir perceptronun yaptığı şey, aslında bir karar sınırı oluşturmaktır. İki boyutlu uzayda bu karar sınırı bir doğrudur. Bu nedenle, tek katmanlı bir perceptron yalnızca doğrusal olarak ayrılabilir problemleri çözebilir.

Hızla gelişen çalışmalar neticesinde Marvin and Seymour (1969) "Perceptrons" adlı meşhur kitaplarını yayınladılar. Bu kitapta, perceptron modelinin temel sınırlılıklarını ispatlamışlardır. XOR problemini, bir perceptronun çözemeyeceği problemlerin en bilinen örneği olarak sundular. Bu kitap, o dönemde yapay zekâ alanında büyük bir hayal kırıklığı yarattı. Araştırmacılar, perceptron gibi basit modellerin yeteneklerinin ne kadar kısıtlı olduğunu gördüler. Bu durum, YSA araştırmalarına olan

ilginin ve fonların önemli ölçüde azalmasına yol açtı ve bu dönem "Yapay Zekâ Kışı" (AI Winter) olarak anılmaya başlandı.

XOR probleminin çözümü, perceptron modelini daha karmaşık hale getirmekle bulundu. Tek bir katman yerine, girdi ve çıktı katmanları arasına bir veya daha fazla gizli katman (hidden layer) eklenmesiyle oluşturulan çok katmanlı algılayıcı (Multi-Layer Perceptron-MLP), bu sorunu aşmayı başardı. Özetle, XOR problemi, tek katmanlı doğrusal modellerin yetersizliğini matematiksel olarak ortaya koymuş ve yapay zekânın ilerlemesi için gizli katmanlara sahip, hiyerarşik ve doğrusal olmayan modellere (yani derin sinir ağlarına) ihtiyaç olduğunu kanıtlamış tarihsel bir mihenk taşıdır.

YSA'nın öğrenme kapasitesini ve problem çözme yeteneğini belirleyen temel unsurlardan biri de aktivasyon fonksiyonlarıdır. Aktivasyon fonksiyonları, her bir yapay nöronun çıktısını belirleyen matematiksel işlemlerdir. Nöronun aldığı toplam girdiye göre bir çıktı üretmesini sağlayarak, ağı doğrusal olmayan problemlere karşı daha esnek ve güçlü hâle getirir. Başka bir ifadeyle, sinir ağına doğrusal olmayan özellikler ekleyerek, sadece doğrusal olarak ayrılabilir problemlerle sınırlı kalmadan daha karmaşık örüntüleri öğrenmesini mümkün kılar. Aktivasyon fonksiyonlarının seçiminde, ağın çözmekte olduğu problem türü belirleyici rol oynamaktadır.

Yaygın olarak kullanılan aktivasyon fonksiyonlarının ilki sigmoid aktivasyon fonksiyonudur. Formülasyonu

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

şeklinde olup girdiyi 0 ile 1 arasında sıkıştırır. Sigmoid aktivasyon fonksiyonu genellikle sınıflandırma problemlerinde kullanılmaktadır. Bir diğer aktivasyon fonksiyonu ise tanjant hiperbolik ($\tanh$) aktivasyon fonksiyonudur. Formülasyonu

$$\tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$$

şeklinde olup çıktıyı -1 ile 1 arasında dönüştürür. Sigmoid benzer şekilde çalışır; ancak sıfır merkezli olması sebebiyle daha kararlı bir öğrenme süreci sağlar. Bir diğer aktivasyon fonksiyonu ise ReLU (Rectified Linear Unit) aktivasyon fonksiyonudur. Formülasyonu

$$f(z) = \max(0, z)$$

şeklinde olup negatif girdileri sıfıra eşitler, pozitifleri olduğu gibi bırakır. Hesaplama açısından verimli olması ve gradyan sönmesi problemini azaltması nedeniyle derin ağlarda sıklıkla kullanılabilmektedir (Nwankpa 2018). Bir diğer aktivasyon fonksiyonu ise Leaky ReLU aktivasyon fonksiyonudur. Formülasyonu

$$f(z) = \max(\alpha z, z)$$

şeklinde olup ReLU'nun negatif değerlerde sabit sıfır üretilmesi sorununu gidermek için geliştirilmiştir. Negatif bölgede küçük bir eğim bırakılır. Böylece bilgi kaybı azaltılır (Xu et al. 2020). Bir diğer aktivasyon fonksiyonu ise softmax aktivasyon fonksiyonudur. Formülasyonu

$$\sigma(z_i) = \frac{e^{z_i}}{\sum_{j=1}^n e^{z_j}}$$

şeklinde olup büyük sınıflı sınıflandırma problemlerinde kullanılabilmektedir. Girdileri olasılık dağılımına dönüştürerek her sınıfa ait olasılık değerini verir. Son olarak lineer aktivasyon fonksiyonu

$$f(z) = z$$

biçiminde formülize edilir. Lineer aktivasyon fonksiyonu, sinir ağına nonlinear (doğrusal olmayan) davranış veya karmaşıklık kazandırmaz. Genellikle regresyon problemlerinde, kullanılmaktadır. 2021 yılında ise Nielsen and Sun (2020) aktivasyon fonksiyonlarının q -hallerini tanımlamıştır. Bunlardan bazıları şöyle verilmiştir. q -tanh aktivasyon fonksiyonu

$$g_q^{\tanh}(x) = \frac{\tanh(x) - \tanh(qx)}{1 - q}$$

şeklinde, q -ReLU aktivasyon fonksiyonu

$$g_q^{\text{ReLU}}(x) = \frac{\text{ReLU}(x) - \text{ReLU}(qx)}{1 - q}$$

şeklinde tanımlanmış, $F(x)$ aktivasyon fonksiyonu olmak üzere genel olarak diğer tüm aktivasyon fonksiyonlarının q halinin ise

$$g_q^F(x) = \frac{F(x) - F(qx)}{1 - q}$$

şeklinde tanımlanabileceği ifade edilmiştir (Nielsen and Sun 2020).

3.5. Yapay Sinir Ağları Algoritmaları

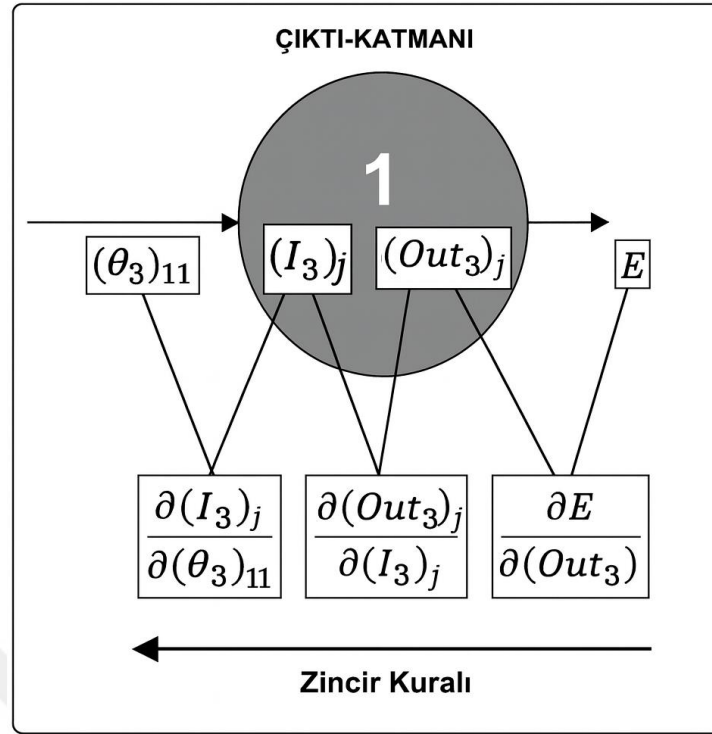
3.5.1. Geri yayılım

Geri yayılım algoritması (Werbos 1982) çok katmanlı algılayıcıların (MLP) eğitilmesinde yaygın olarak kullanılmakla beraber birçok problem türünün çözümünde de yüksek başarı sağlamıştır. Literatürde güncel olarak da sıkça kullanılmakta olan geri yayılım algoritması bir çok farklı disiplinde karşılaşılan problemlerin çözümünde kullanılmaktadır. Örneğin (2025) kullanıcı desteği olmadan ultrason görüntülerinden elde edilen meme kanseri görüntülerinde tümörün büyümesini tahmin etmek amacıyla uyguladıkları sistematik öğrenmede geri yayılım algoritmasını kullanmıştır. Bu algoritma, literatürde sıklıkla "genelleştirilmiş delta kuralı" olarak da adlandırılmaktadır. Temel olarak geri yayılım algoritması iki aşamadan oluşmaktadır. İlk aşamada, "ileri yayılım" (forward propagation) gerçekleştirilmektedir. Bu süreçte, girdi verisi ağın giriş katmanından başlayarak sırasıyla gizli katmanlar üzerinden geçip çıkış katmanına doğru aktarılmaktadır. Bu aktarım sırasında, her bir girdi için başlangıçta rastgele seçilmiş ağırlıklar kullanılmaktadır. İkinci aşamada ise, çıkış katmanında hesaplanan çıkış değerleri ile gerçek değerler arasındaki fark belirlenmektedir. Bu fark (hata) daha sonra

ağ boyunca geriye doğru yayılmaktadır. Hatanın geriye yayılması sayesinde katmanlar arasındaki ağırlıklar ve bias terimleri güncellenmektedir. İşte bu sürece "geri yayılım" (backward propagation) adı verilmektedir. Verilen bu iki aşamanın döngüsel olarak tekrarı ise ya belirli bir iterasyon (epoch) sayısına ulaşıncaya kadar ya da hata kabul edilen bir sınırın altına düşünceye kadar tekrar edilir.

Şekil 3.2. de verilen YSA modelinde hata, çıktı katmanından geriye doğru ilk olarak ikinci gizli katmana aktarılır. Daha sonra ise birinci gizli katmana ve son olarakta giriş katmanına aktarılmaktadır. Bu ileri ve geri devam eden süreç sayesinde tüm ağ ağırlıkları güncellenmektedir. Aslında burada asıl amaç ağırlık değerlerinin hata fonksiyonunu minimize edecek biçimde güncellenmesidir. Bu bilgiler ışığında ikinci gizli katmandaki ağırlıkların güncellenmesi için çıktı katmanındaki hata değerleri kullanılmaktadır. Birinci gizli katmandaki ağırlıkların güncellenmesi için ise ikinci gizli katmandaki hata değerleri kullanılmaktadır. Söz konusu süreç, gizli katman sayısından bağımsız olarak aynı prensiple sürdürülmektedir. Çıktı katmanında geri yayılım ise şu şekilde gerçekleşmektedir; bir önceki aşamada belirlenen hata değerine bağlı olarak, katmanlar arasındaki ağırlıkların ve bias değerlerinin güncellenmesi gerekmektedir. Geri yayılım algoritmasında bu güncellemeler gradyan iniş yöntemi kullanılarak yapılmaktadır.

Örneğin Şekil 3.4. te, çıktı katmanındaki birinci yapay nörona ait $(\theta_3)_{11}$ ağırlık değeri gösterilmektedir. $(\theta_3)_{11}$ ağırlığında yapılacak güncellenmenin, hata üzerindeki etkisi $\frac{\partial E}{\partial (\theta_3)_{11}}$ ifadesi ile belirtilir.



Şekil 3.4. Ağırlıkların güncellenmesi (Uğuz 2019)

Bu ifadenin hesaplanması

$$\frac{\partial E}{\partial (\theta_3)_{ji}} = \frac{\partial E}{\partial (Out_3)_j} \cdot \frac{\partial (Out_3)_j}{\partial (I_3)_j} \cdot \frac{\partial (I_3)_j}{\partial (\theta_3)_{ji}}$$

şeklinde verilen zincir kuralı kullanılarak yapılabilmektedir. Böylece elde edilen hata değeri ile öğrenme oranının çarpımı, eski ağırlık değerine eklenerek

$$(\theta_3)_{ji} = (\theta_3)_{ji} + \gamma \cdot \delta_{3j} \cdot (Out_2)_i$$

formülü ile güncellenmiş ağırlık değeri elde edilir. Burada δ_{3j}

$$\delta_{3j} = -(y_j - (Out_3)_j) f'((I_3)_j)$$

şeklinde bulunup yerel eğim olarak ifade edilmektedir. İkinci gizli katmandaki ağırlıkların güncellenmesinde

$$\frac{\partial E}{\partial(\theta_2)_{ji}} = \frac{\partial E}{\partial(Out_2)_j} \cdot \frac{\partial(Out_2)_j}{\partial(I_2)_j} \cdot \frac{\partial(I_2)_j}{\partial(\theta_2)_{ji}}$$

ile verilen bağıntı kullanılmaktadır. Yukarıda elde edilen hata değeri ile öğrenme oranının çarpımı, önceki ağırlık değerine eklenerek

$$(\theta_2)_{ji} = (\theta_2)_{ji} + \gamma \cdot \delta_{2j} \cdot (Out_1)_i$$

ifadesindeki güncellenmiş ağırlık değeri elde edilir. Benzer şekilde birinci gizli katmandaki ağırlıkların güncellemesi de aynı şekilde yapılmaktadır.

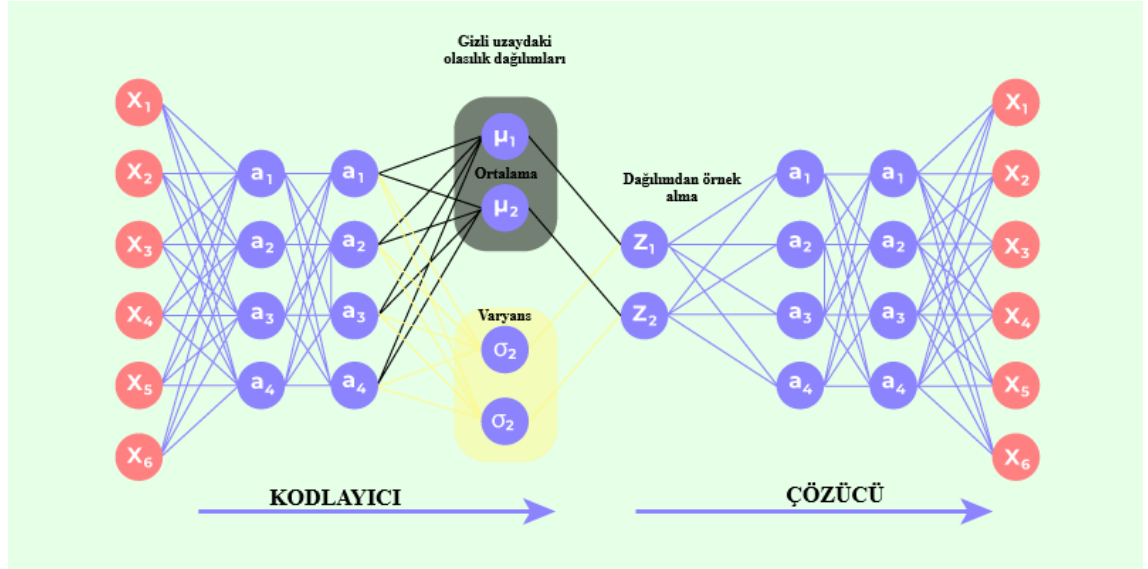
3.5.2. Stokastik geri yayılım

Stokastik geri yayılım, sinir ağlarının eğitiminde ağırlıkların güncellenmesinde kullanılan ve geleneksel geri yayılımdan ayrılan bir yöntemdir. 1989 yılında LeCun ve arkadaşlarının el yazısı posta kodu tanıma üzerine yaptıkları çalışmada, stokastik geri yayılım algoritmasını tanımlamışlardır. Bu yöntemde, eğitim verileri ardışık olarak ve sabit bir sırayla ağa sunulmakta; her bir örnek için ağın çıktısı hesaplanmakta ve hata fonksiyonu elde edilen bu tekil örnek üzerinden değerlendirilerek, ağ ağırlıkları her bir örneğin sunumundan hemen sonra güncellenmektedir. Buna karşın, geleneksel geri yayılımda ağırlık güncellemeleri tüm eğitim seti üzerindeki hata fonksiyonunun ortalamasına göre bir defada yapılmaktadır. Stokastik geri yayılımın özellikle büyük ve tekrarlı veri kümelerinde çok daha hızlı yakınsadığı ve daha kararlı çözümler bulduğu ve lokal minimumlara sıkışmanın daha az olduğu ifade edilmiştir (LeCun et al. 1989).

Bu tezde stokastik geri yayılım algoritması ile varyasyonel oto kodlayıcı (Variational Auto Encoder-VAE) birlikte kullanılmıştır. Güncel olarak literatürde birçok çalışmada bu şekliyle kullanılmaya devam edilmektedir. Örneğin Wang et al. (2025) çalışmalarında üç gerçek dünya veri kümesi üzerinde yapılan deneysel çalışmalar sonucunda önerilen modelin görüntüleri yeniden yapılandırmasında ince ayrıntıları koruduğu ve kalite ölçütlerinde başarılı sonuçlar verdiğini ifade etmişlerdir.

VAE, gözlemlenen verilerden olasılıksal ve sürekli bir gizli uzay (latent space) öğrenilmesini sağlayan üretken modelleme araçları olarak tanımlanmaktadır (Doersch 2016). Klasik oto kodlayıcılardan farklı olarak VAE’lerde kodlayıcı (encoder) ve çözücü (decoder) yapay sinir ağları kullanılır. Bu ağlar veriyi gizli değişkenler üzerinden parametrik bir olasılık dağılımı ile modelleyebilmektedir. Böylece modelin hem veriyi yeniden üretmede başarılı olması, hem de öğrenilen dağılımdan örneklenerek gerçekçi ve tutarlı yeni örnekler üretebilmesi hedeflenir.

Kingma ve Welling (2013), VAE mimarisinin temelinde ‘yeniden parametrelendirme triki’ olarak bilinen yaklaşımın yer aldığını ifade etmiştir. Bu yöntem sayesinde karmaşık olasılıksal çıkarım süreçleri gradyan temelli optimizasyon ile daha verimli bir şekilde çözülebilmektedir. Böylece yaklaşık posterior dağılımlarının (yani veri gözlemlendikten sonra bir parametre ya da değişkenin olasılık dağılımı) eğitimi sırasında üretilen örnekler doğrudan türevlenebilir olur ve standart öğrenme algoritmalarıyla kullanılabilir. Bu mimari ile VAE’ler özellikle yüksek boyutlu ve karmaşık veri kümelerinde başarılı sonuçlar vermektedir. Sadece veri sıkıştırma ve boyut indirgeme için değil, aynı zamanda yeni veri örnekleri üretimi, eksik veri tamamlama ve temsili öğrenme gibi alanlarda da geniş uygulama olanakları sunmaktadır (Kingma and Welling 2013; Doersch 2016). VAE’lerin bu özellikleri onları derin öğrenme modelleri arasında önemli bir yere taşımış ve makine öğrenmesinde hızla yaygınlaşmalarını sağlamıştır. VAE üç ana bileşenden oluşur. İlk olarak kodlayıcı giriş verisinden iki vektör çıkarır. Bu vektörler ortalama (μ) ve standart sapma (σ) dır. Bu sayede her özellik için belirli bir aralık tanımlanır. İkinci olarak gizli uzay aşamasında; kodlayıcının ürettiği (μ) ve (σ) değerlerine göre, bu aralıkta rastgele bir nokta seçilir. Bu rastgelelik sayesinde model yeni örnekler üretebilir hâle gelmektedir. Üçüncüsü olarak çözücü gizli uzayda seçilen rastgele bir noktayı alarak giriş verisine benzer ama aynı olmayan bir çıktı üretir.



Şekil 3.5. Varyasyonel oto kodlayıcı mimarisi (Anonim 2025)

Matematiksel süreç ise şöyle işlemektedir. VAE, gözlemler (x) ile gizli değişkenler (z) arasındaki olasılık ilişkisini modellemek için

$$p(x) = \int p(x|z) p(z) dz$$

formunu kullanır. Burada $p(z)$ genellikle standart normal öncül (prior) dağılım olarak seçilir, $p(x|z)$ ise çözücünün tanımladığı koşullu dağılımdır. Ancak yukarıda verilen integral çoğu zaman kapalı formda hesaplanamamaktadır. Dolayısıyla gerçek posterior

$$p(z|x) = \frac{p(x|z) p(z)}{p(x)}$$

doğrudan elde edilemez. Bunun yerine, her bir veri örneği için ayrı çıkarım yapmak yerine, tüm veri noktaları üzerinde ortak parametrelerle çalışan bir yaklaşımcı posterior modeli $q_\phi(z|x)$ tanımlanır. Bu yöntem, literatürde amortize varyasyonel çıkarım olarak bilinir ve esasen tek seferde öğrenilen parametrelerle tüm örnekler için yaklaşık çıkarım yapma anlamına gelir.

Teorik olarak amaç, $KL(q_\phi(z|x) \parallel p(z|x))$ değerini minimize etmektir; fakat $p(z|x)$ bilinmediği için bu ifade doğrudan optimize edilemez. Bunun yerine, kanıt alt sınırı (Evidence Lower Bound – ELBO) maksimize edilir. ELBO

$$\text{ELBO} = \mathbb{E}_{q_{\phi}(z|x)}[\log p_{\theta}(x|z)] - KL(q_{\phi}(z|x) \parallel p(z))$$

şeklindedir. Burada ilk terim yeniden yapılandırma bir modelin ürettiği olasılık değerinin logaritmasının alınmış hâli log-olasılığını (log-likelihood) temsil eder ve modelin veriyi ne kadar iyi üretebildiğini ölçer; ikinci terim ise yaklaşımcı posteriorun bir olasılık modelinde gizli değişkenler olan z hakkında, veriyi görmeden önce sahip olduğumuz ön bilgiyi tanımlayan dağılım olan öncül dağılıma yakın olmasını sağlar.

Toplam kayıp, negatif log-olasılık (uygulamada L2, Binary Cross-Entropy gibi ölçülerle eşdeğer) ile KL (Kullback–Leibler Divergence) sapma teriminin toplamıdır. Yani

$$\mathcal{L}(\theta, \phi; x) = -\mathbb{E}_{q_{\phi}(z|x)}[\log p_{\theta}(x|z)] + KL(q_{\phi}(z|x) \parallel p(z))$$

dir. Buradaki KL terimi veri dağılımı ile değil, $q(z|x)$ ile $p(z)$ arasındaki farkı ölçer ve yalnızca olasılık modellerinde karmaşık ortak dağılımları basitleştirmek için kullanılan bir bağımsızlık varsayımı yöntemi olan mean-field bağımsızlık varsayımı altında

$$KL(q(z|x) \parallel p(z)) = \sum_j KL(q_j(z_j|x) \parallel p(z_j))$$

biçiminde ayrışabilir. Bu şekilde VAE, hem veri uzayında yeniden yapılandırma başarısını hem de gizli uzayda düzenliliği aynı anda optimize eder. Dolayısıyla hem veri oluşturabilir hem de anlamlı bir latent uzay sağlayarak veri hakkında sezgisel kontrol sunar.

3.5.3. Esnek geri yayılım

Esnek geri yayılım algoritması (Resilient Backpropagation-Rprop) , çok katmanlı yapay sinir ağlarının eğitimi için geliştirilmiştir (Riedmiller and Braun 1993). Güncel olarak da literatürde disiplinler arası birçok çalışmada bu algoritma kullanılmaktadır. Örneğin Shobug et al. (2025) çalışmalarında güneş ve rüzgâr enerjisi tabanlı üretimin yoğun olduğu izole mikro şebekelerde şebekenin frekans tepkisini ve genel kararlılığını

artırmak için bu algoritmadan faydalanmıştır. Çünkü bu algoritma doğrudan ağırlık güncellemelerine uyum sağlamakta ve öğrenme sürecini hızlandırmaktadır.

Bilindiği üzere klasik gradyan iniş temelli algoritmalarda, öğrenme oranının (learning rate) seçimi büyük önem taşımaktadır. Bu oran çok küçük seçildiğinde öğrenme süreci yavaşlarken, çok büyük seçildiğinde ise ağda osilasyonlar veya kararsızlıklar olabilmektedir. Yine klasik algoritmalarda kullanılan momentum parametresi de genellikle problemde problemde değişmektedir. Öğrenme sürecinde istenilen başarıya ulaşmak için momentum parametresinde dikkatli bir seçim yapılması gerekmektedir.

Esnek geri yayılım algoritması ise, bu geleneksel problemlerin üstesinden gelmek amacıyla geliştirilmiştir. Her ağırlık için güncelleme adımının büyüklüğünü, ayrı ayrı ve yerel olarak ayarlamaktadır. Temel olarak bu algoritmada ağırlık güncellemesi için gradyanın sadece işaretine bakılır. Türevin mutlak değeri ise dikkate alınmaz. Böylelikle öğrenme hızı türevin büyüklüğündeki ani değişimlerden etkilenmemektedir. Ağırlık güncellemesi ise türevin yönüne göre yapılmaktadır. Her ağırlık için ayrı ayrı izlenen güncelleme adımı (Δ_{ij}), hata fonksiyonunun gradyanındaki işaret değişimine bağlı olarak şu şekilde uyarlanır: eğer bir ağırlığın gradyanı ardışık iki adımda aynı işareti koruyorsa (yani hata fonksiyonu hâlâ aynı yönde azalmakta), ilgili ağırlık güncelleme adımı belirli bir oranda artırılır. Eğer gradyan işareti değişirse (yani minimum atlanmış ve yön tersine dönmüşse), adım boyutu belirli bir çarpanla azaltılır. Ayrıca, bir önceki adım geri alınır ve bu adımda güncelleme miktarı değiştirilmez. Gradyan sıfırsa, ağırlık güncellemesi yapılmaz.

Bu uyarlama sayesinde, ağırlık güncellemeleri, türevin büyüklüğünden bağımsız olarak yalnızca yönüne (pozitif ya da negatif) bağlı olacak şekilde belirlenir. Böylece, özellikle derin katmanlardaki küçük gradyanların veya yüzeysel bölgelerdeki gradyan azalmalarının öğrenmeyi yavaşlatması önlenmiş olur.

Esnek geri yayılım algoritmasının dikkat çekici özelliklerinden biri de parametre seçiminin çok daha az kritik olmasıdır. Başlangıçtaki güncelleme değeri (Δ_0), çoğu problemde geniş bir aralıkta seçilebilmektedir ve algoritmanın yakınsama süresinde anlamlı bir fark oluşturmamaktadır. Deneysel çalışmalarda, Rprop algoritmasının hem

geleneksel geri yayılım algoritmasına hem de diğer adaptif yöntemlere göre çok daha hızlı ve kararlı bir şekilde yakınsadığı rapor edilmiştir (Riedmiller and Braun 1993). Algoritmanın matematiksel olarak bir θ_{ij} ağırlığını güncellemesi gradyanın işaretine göre

$$\Delta_{ij}(t) = \begin{cases} \eta^+ \cdot \Delta_{ij}(t-1), & \text{eğer } \frac{\partial E(t-1)}{\partial \theta_{ij}} \cdot \frac{\partial E(t)}{\partial \theta_{ij}} > 0 \\ \eta^- \cdot \Delta_{ij}(t-1), & \text{eğer } \frac{\partial E(t-1)}{\partial \theta_{ij}} \cdot \frac{\partial E(t)}{\partial \theta_{ij}} < 0 \\ \Delta_{ij}(t-1), & \text{aksi halde} \end{cases}$$

şeklinde yapılır. Burada; $\Delta_{ij}(t)$, t adımındaki ağırlık değişim miktarı, $\eta^+ > 1$ artış katsayısı ve $\eta^- < 1$ ise azalış katsayısıdır. Güncellenmiş ağırlık değeri ise

$$\theta_{ij}(t+1) = \begin{cases} \theta_{ij}(t) - \Delta_{ij}(t), & \text{eğer } \frac{\partial E(t)}{\partial \theta_{ij}} > 0 \\ \theta_{ij}(t) + \Delta_{ij}(t), & \text{eğer } \frac{\partial E(t)}{\partial \theta_{ij}} < 0 \\ \theta_{ij}(t), & \text{aksi halde} \end{cases}$$

şeklinde hesaplanır.

3.5.4. Ekstrem öğrenme makineleri

Ekstrem öğrenme makineleri (Extreme Learning Machines-ELM) tek katmanlı klasik ileri beslemeli yapay sinir ağlarındaki eğitim sürecinin temel zorluklarına doğrudan çözüm sunan yeni nesil bir öğrenme algoritması olarak tanımlanmaktadır (Huang et al. 2006). Geleneksel ağ eğitiminde, tüm ağ parametrelerinin (ağırlıklar ve biaslar) iteratif ve genellikle yavaş ilerleyen gradyan temelli yöntemlerle ayarlanması gerekirken, ELM yaklaşımında yalnızca çıktı katmanına ait ağırlıklar analitik olarak belirlenmekte; girişten gizli katmana ait ağırlıklar ile bias değerleri ise rastgele atanmakta ve sabit tutulmaktadır. Bu sayede eğitim süresi son derece hızlanmakta, klasik yöntemlerde karşılaşılan yerel minimumlara saplanma veya öğrenme hızının yetersizliği gibi sorunlar büyük ölçüde ortadan kalkmaktadır.

ELM modelinin matematiksel altyapısı ise tek gizli katmanlı ileri beslemeli yapay sinir ağı (Single Hidden Layer Feedforward Neural Network - SLFN) yapısına dayanır. N eğitim örneği sayısı, n giriş boyutu, L gizli katmandaki nöron sayısı, m çıktı boyutu eğitim veri kümesi ise

$$\{(x_i, y_i) \mid i = 1, 2, \dots, N\}, \quad x_i \in \mathbb{R}^n, \quad y_i \in \mathbb{R}^m$$

olsun. $f(\cdot)$ aktivasyon fonksiyonu (sigmoid, tanh, ReLU, vb.), θ_j girişten j 'inci gizli nörona giden ağırlık vektörü, b_j ise j 'inci gizli nöronun bias değeri olmak üzere gizli katmandaki j 'inci nöronun çıktısı

$$h_j(x_i) = f(\theta_j \cdot x_i + b_j), \quad j = 1, 2, \dots, L$$

şeklinde ifade edilir. ELM de θ_j ve b_j rastgele atanır ve eğitim boyunca değiştirilmez. Gizli katman çıktı matrisi ise

$$H = \begin{bmatrix} f(\theta_1 \cdot x_1 + b_1) & \dots & f(\theta_L \cdot x_1 + b_L) \\ \vdots & \ddots & \vdots \\ f(\theta_1 \cdot x_N + b_1) & \dots & f(\theta_L \cdot x_N + b_L) \end{bmatrix}_{N \times L}$$

şeklinde verilir. Çıktı katmanı ağırlıkları θ bulmak için $H\theta = Y$ den faydalanılır. Buradan

$$\theta = H^\dagger Y$$

olarak elde edilir. Burada $H^\dagger$, H 'nin Moore–Penrose genelleştirilmiş tersidir.

Yapılan deneysel çalışmalar sonucunda, ELM algoritmasının hem doğruluk hem de genelleme performansında, klasik ileri beslemeli ağlara ve destek vektör makinelerine (SVM) karşı rekabetçi, çoğu durumda ise üstün sonuçlar verdiği gösterilmiştir. Özellikle yüksek boyutlu ve büyük ölçekli veri setlerinde, hızlı eğitim ve test yanıtı sayesinde gerçek zamanlı uygulamalar için de oldukça uygundur. Ayrıca, ELM nin basitliği ve kolay uygulanabilirliği, literatürde çok sayıda farklı uygulama alanında tercih edilmesine olanak sağlamaktadır (Huang et al. 2006).

Avantajlarının yanısıra bu yöntemde iki farklı problemle karşılaşılabilir. Birincisi, H^+ matrisini bulmak her zaman mümkün olmayabilir. İkincisi, ekstrem öğrenme makinesinin çalışabilmesi için gizli düğüm sayısı olan M nin örnek sayısı olan N den küçük ya da eşit olması gerekmektedir ki birçok gerçek yaşam problemlerinde bu sayı daha büyük olabilmektedir. Bu da modelin aşırı öğrenmesine (overfitting) sebep olmaktadır ki istenen bir durum değildir. Bu nedenle θ parametresini bulmak için ELM yöntemi farklı şekillerde modifiye edilmiştir.



4. İKİ EYLEMSİZLİK TERİMLİ İLERİ-GERİ AYRIŞTIRMA ALGORİTMALARI

4.1. Algoritmaların Teorisi ve İspatları

Bu bölümde bir reel Hilbert uzayında tanımlı maksimal monoton bir operatör ile ters kuvvetli monoton bir operatörün toplamının sıfır noktasını belirlemeyi amaçlayan yani monoton kapsama problemini çözmeyi amaçlayan iki eylemsizlik terim içeren ileri-geri ayrıştırma algoritmaları tanımlanmış ve önerilen yöntemler tarafından üretilen dizilerin yakınsaklık teoremleri ifade ve ispat edilmiştir. Bu kapsamda ilk algoritma aşağıda verilmiştir.

Algoritma 4.1: $\delta \leq 0$, $\lambda \in (0, 2\alpha)$, $\vartheta \in [0, 1)$ ve bir $\{\eta_k\} \subset (0, 1)$ dizisi alınsın. Başlangıç noktaları olarak keyfi $p_{-1}, p_0, p_1 \in \mathcal{H}$ seçilsin ve $k = 1$ olarak başlatılsın. $J_\lambda^B = (I + \lambda B)^{-1} B$ operatörünün resolventi olmak üzere

$$\begin{cases} w_k = \vartheta(p_k - p_{k-1}) + \delta(p_{k-1} - p_{k-2}) + p_k \\ p_{k+1} = (1 - \eta_k)w_k + \eta_k \left(J_\lambda^B(w_k - \lambda A w_k) \right) \end{cases}$$

eşitlikleri yardımıyla p_{k+1} hesaplınsın. Ardından $k = k + 1$ alınarak bu işleme devam edilsin.

Tanımlanan algoritmanın yakınsaklık teoremleri için aşağıdaki şartlar kabul edilecektir.

Kabul 4.2:

- i. $A: \mathcal{H} \rightarrow \mathcal{H}$ α -ters kuvvetli monoton ve $B: \mathcal{H} \rightarrow 2^{\mathcal{H}}$ maksimal monoton bir operatör olsun.
- ii. (3.1) probleminin çözüm kümesi boş kümeden farklı olsun.

Ayrıca, eylemsizlik parametreleri olan ϑ , δ ve $\{\eta_k\}$ dizisinin aşağıda belirtilen koşulları sağlaması gerekmektedir.

Kabul 4.3:

- i. Her $k \geq 1$ için $0 < \varepsilon_1 < \eta_k < \varepsilon_2 < 1$ eşitsizlikleri sağlansın.
- ii. ϑ parametresi $\kappa := \frac{2\alpha}{4\alpha-\lambda}$ olmak üzere

$$0 \leq \vartheta < \min \left\{ \frac{1}{3}, \frac{\varepsilon_1(1-\kappa)}{\varepsilon_1(1-\kappa) + 2\kappa} \right\}$$

eşitsizliğini sağlansın.

- iii. $\delta \leq 0$ skaleri için

$$\max \left\{ -\frac{\varepsilon_1(1-\kappa)(1-\vartheta) - 2\kappa\vartheta}{\varepsilon_1(1-\kappa)}, \frac{\vartheta(1+\vartheta) - \varepsilon_1 \left(\frac{1-\kappa}{\kappa} \right) (1-\vartheta)^2}{(1+\vartheta) \left(1 + \varepsilon_1 \left(\frac{1-\kappa}{\kappa} \right) \right)} \right\} < \delta,$$

ve

$$\begin{aligned} \kappa\vartheta(1+\vartheta) - \varepsilon_1(1-\vartheta)^2(1-\kappa) &< \kappa\delta(1+2\vartheta-\delta) + 2\varepsilon_1(1-\kappa)(1+\vartheta)\delta \\ &+ \varepsilon_1(1-\kappa)\delta^2 \end{aligned}$$

eşitsizlikleri sağlansın.

Lemma 4.4: Kabul 4.2 ve Kabul 4.3 altında, Algoritma 4.1 tarafından üretilen $\{p_k\}$ dizisi sınırlıdır.

İspat: $J_\lambda^B(I - \lambda A)$ operatörü $\frac{2\alpha}{4\alpha-\lambda}$ ortalamalıdır. Algoritma 4.1, $\mathcal{T} := J_\lambda^B(I - \lambda A)$ olmak üzere

$$w_k = \vartheta(p_k - p_{k-1}) + \delta(p_{k-1} - p_{k-2}) + p_k \quad (4.1)$$

$$p_{k+1} = (1 - \eta_k)w_k + \eta_k \mathcal{T}w_k \quad (4.2)$$

şeklinde ifade edilebilir. $p^* \in F(\mathcal{T}) = (A + B)^{-1}(0)$ olduğunu varsayalım. Buna göre:

$$\begin{aligned} w_k - p^* &= p_k + \vartheta(p_k - p_{k-1}) + \delta(p_{k-1} - p_{k-2}) - p^* \\ &= (1 + \vartheta)(p_k - p^*) - (\vartheta - \delta)(p_{k-1} - p^*) - \delta(p_{k-2} - p^*) \end{aligned}$$

dır. Buna bağlı olarak

$$\begin{aligned} \|w_k - p^*\|^2 &= \|(1 + \vartheta)(p_k - p^*) - (\vartheta - \delta)(p_{k-1} - p^*) \\ &\quad - \delta(p_{k-2} - p^*)\|^2 \\ &= (1 + \vartheta)^2 \|p_k - p^*\|^2 - (\vartheta - \delta)^2 \|p_{k-1} - p^*\|^2 \\ &\quad - \delta^2 \|p_{k-2} - p^*\|^2 + (1 + \vartheta)(\vartheta - \delta) \|p_k - p_{k-1}\|^2 \\ &\quad + \delta(1 - \vartheta) \|p_k - p_{k-2}\|^2 - \delta(\vartheta - \delta) \|p_{k-1} - p_{k-2}\|^2 \end{aligned} \quad (4.3)$$

olur. Ayrıca

$$\begin{aligned} 2\vartheta \langle p_{k+1} - p_k, p_k - p_{k-1} \rangle &= 2\langle \vartheta(p_{k+1} - p_k), p_k - p_{k-1} \rangle \\ &\leq 2|\vartheta| \|p_{k+1} - p_k\| \|p_k - p_{k-1}\| \\ &= 2\vartheta \|p_{k+1} - p_k\| \|p_k - p_{k-1}\| \end{aligned}$$

olduğu göz önünde bulundurulursa buradan

$$-2\vartheta \langle p_{k+1} - p_k, p_k - p_{k-1} \rangle \geq -2\vartheta \|p_{k+1} - p_k\| \|p_k - p_{k-1}\| \quad (4.4)$$

ifadesi elde edilir. Bunun yanı sıra

$$\begin{aligned} 2\delta\langle p_{k+1} - p_k, p_{k-1} - p_{k-2} \rangle &= 2\langle \delta(p_{k+1} - p_k), p_{k-1} - p_{k-2} \rangle \\ &\leq 2|\delta| \|p_{k+1} - p_k\| \|p_{k-1} - p_{k-2}\| \end{aligned}$$

olur. Böylelikle

$$-2\delta\langle p_{k+1} - p_k, p_{k-1} - p_{k-2} \rangle \geq 2|\delta| \|p_{k+1} - p_k\| \|p_{k-1} - p_{k-2}\| \quad (4.5)$$

sonucu elde edilir. Benzer şekilde

$$\begin{aligned} 2\delta\vartheta\langle p_{k-1} - p_k, p_{k-1} - p_{k-2} \rangle &= 2\langle \delta\vartheta(p_{k-1} - p_k), p_{k-1} - p_{k-2} \rangle \\ &\leq 2|\delta|\vartheta \|p_{k-1} - p_k\| \|p_{k-1} - p_{k-2}\| \\ &= 2|\delta|\vartheta \|p_k - p_{k-1}\| \|p_{k-1} - p_{k-2}\| \end{aligned}$$

dir. Böylece

$$\begin{aligned} 2\delta\vartheta\langle p_k - p_{k-1}, p_{k-1} - p_{k-2} \rangle &= -2\delta\vartheta\langle p_{k-1} - p_k, p_{k-1} - p_{k-2} \rangle \\ &\geq -2|\delta|\vartheta \|p_k - p_{k-1}\| \|p_{k-1} - p_{k-2}\| \end{aligned} \quad (4.6)$$

olur. (4.4), (4.5) ve (4.6) numaralı eşitsizliklerinden

$$\begin{aligned} \|p_{k+1} - w_k\|^2 &= \|p_{k+1} - (p_k + \vartheta(p_k - p_{k-1}) + \delta(p_{k-1} - p_{k-2}))\|^2 \\ &= \|p_{k+1} - p_k\|^2 - 2\vartheta\langle p_{k+1} - p_k, p_k - p_{k-1} \rangle \\ &\quad - 2\delta\langle p_{k+1} - p_k, p_{k-1} - p_{k-2} \rangle + \vartheta^2 \|p_k - p_{k-1}\|^2 \\ &\quad - 2\delta\vartheta\langle p_k - p_{k-1}, p_{k-1} - p_{k-2} \rangle + \delta^2 \|p_{k-1} - p_{k-2}\|^2 \\ &\geq \|p_{k+1} - p_k\|^2 - 2\vartheta \|p_{k+1} - p_k\| \|p_k - p_{k-1}\| \end{aligned}$$

$$\begin{aligned}
 & -2|\delta|\|p_{k+1} - p_k\|\|p_{k-1} - p_{k-2}\| + \vartheta^2\|p_k - p_{k-1}\|^2 \\
 & -2|\delta|\vartheta\|p_k - p_{k-1}\|\|p_{k-1} - p_{k-2}\| \\
 & + \delta^2\|p_{k-1} - p_{k-2}\|^2 \\
 & \geq \|p_{k+1} - p_k\|^2 - \vartheta\|p_{k+1} - p_k\|^2 - \vartheta\|p_k - p_{k-1}\|^2 \\
 & - |\delta|\|p_{k+1} - p_k\|^2 - |\delta|\|p_{k-1} - p_{k-2}\|^2 \\
 & - \vartheta^2\|p_k - p_{k-1}\|^2 - |\delta|\vartheta\|p_k - p_{k-1}\|^2 \\
 & + \delta^2\|p_{k-1} - p_{k-2}\|^2 - |\delta|\vartheta\|p_k - p_{k-1}\|^2 \\
 & + \delta^2\|p_{k-1} - p_{k-2}\|^2 \\
 & \geq (1 - |\delta| - \vartheta)\|p_{k+1} - p_k\|^2 \\
 & - (\vartheta^2 - \vartheta - |\delta|\vartheta)\|p_k - p_{k-1}\|^2 \\
 & + (\delta^2 - |\delta| - |\delta|\vartheta)\|p_{k-1} - p_{k-2}\|^2
 \end{aligned} \tag{4.7}$$

elde edilir. $\mathcal{T}$ operatörünün κ -ortalamalı ve yarı genişlemeyen dönüşüm olduğu dikkate alındığında

$$\begin{aligned}
 \|p_{k+1} - p^*\|^2 &= \|(1 - \eta_k)w_k + \eta_k\mathcal{T}w_k - p^*\|^2 \\
 &= (1 - \eta_k)\|w_k - p^*\|^2 + \eta_k\|\mathcal{T}w_k - p^*\|^2 \\
 &\quad - \eta_k(1 - \eta_k)\|w_k - \mathcal{T}w_k\|^2 \\
 &\leq (1 - \eta_k)\|w_k - p^*\|^2 \\
 &\quad + \eta_k\left(\|w_k - p^*\|^2 - \left(\frac{1-\kappa}{\kappa}\right)\|w_k - \mathcal{T}w_k\|^2\right)
 \end{aligned} \tag{4.8}$$

eşitsizliği elde edilir. (4.3) ve (4.7) eşitsizlikleri (4.8) ifadesine uygulandığında

$$\begin{aligned}
 \|p_{k+1} - p^*\|^2 &\leq (1 - \eta_k) \|w_k - p^*\|^2 \\
 &\quad + \eta_k \left(\|w_k - p^*\|^2 - \left(\frac{1-\kappa}{\kappa} \right) \|w_k - \mathcal{T}w_k\|^2 \right) \\
 &\leq \|w_k - p^*\|^2 - \varepsilon_1 \left(\frac{1-\kappa}{\kappa} \right) \|w_k - \mathcal{T}w_k\|^2 \\
 &\leq (1 + \vartheta) \|p_k - p^*\|^2 - (\vartheta - \delta) \|p_{k-1} - p^*\|^2 - \delta \|p_{k-2} - p^*\|^2 \\
 &\quad + (1 + \vartheta)(\vartheta - \delta) \|p_k - p_{k-1}\|^2 + \delta(1 + \vartheta) \|p_k - p_{k-2}\|^2 \\
 &\quad - \delta(\vartheta - \delta) \|p_{k-1} - p_{k-2}\|^2 + \delta(1 + \vartheta) \|p_k - p_{k-2}\|^2 \\
 &\quad - \varepsilon_1 \left(\frac{1-\kappa}{\kappa} \right) ((1 - |\delta| - \vartheta) \|p_{k+1} - p_k\|^2 \\
 &\quad - (\vartheta^2 - \vartheta - |\delta|\vartheta) \|p_k - p_{k-1}\|^2 \\
 &\quad + (\delta^2 - |\delta| - |\delta|\vartheta) \|p_{k-1} - p_{k-2}\|^2) \\
 &= (1 + \vartheta) \|p_k - p^*\|^2 - (\vartheta - \delta) \|p_{k-1} - p^*\|^2 \\
 &\quad - \delta \|p_{k-2} - p^*\|^2 \\
 &\quad + \left((1 + \vartheta)(\vartheta - \delta) - \varepsilon_1 \left(\frac{1-\kappa}{\kappa} \right) (\vartheta^2 - \vartheta - |\delta|\vartheta) \right) \|p_k - p_{k-1}\|^2 \\
 &\quad - \left(\varepsilon_1 \left(\frac{1-\kappa}{\kappa} \right) (1 - |\delta| - \vartheta) \right) \|p_{k+1} - p_k\|^2 \\
 &\quad - \left(\delta(\vartheta - \delta) + \varepsilon_1 \left(\frac{1-\kappa}{\kappa} \right) (\delta^2 - |\delta| - |\delta|\vartheta) \right) \|p_{k-1} - p_{k-2}\|^2 \\
 &\quad + \delta(1 + \vartheta) \|p_k - p_{k-2}\|^2 \\
 &\leq (1 + \vartheta) \|p_k - p^*\|^2 - (\vartheta - \delta) \|p_{k-1} - p^*\|^2 - \delta \|p_{k-2} - p^*\|^2 \\
 &\quad + \left((1 + \vartheta)(\vartheta - \delta) - \varepsilon_1 \left(\frac{1-\kappa}{\kappa} \right) (\vartheta^2 - \vartheta - |\delta|\vartheta) \right) \|p_k - p_{k-1}\|^2 \\
 &\quad - \left(\varepsilon_1 \left(\frac{1-\kappa}{\kappa} \right) (1 - |\delta| - \vartheta) \right) \|p_{k+1} - p_k\|^2
 \end{aligned}$$

$$-\left(\delta(\vartheta - \delta) + \varepsilon_1 \left(\frac{1-\kappa}{\kappa}\right) (\delta^2 - |\delta| - |\delta|\vartheta)\right) \|p_{k-1} - p_{k-2}\|^2$$

bulunur. Dolayısıyla

$$\begin{aligned} & \|p_{k+1} - p^*\|^2 - \vartheta \|p_k - p^*\|^2 - \delta \|p_{k-1} - p^*\|^2 + \left(\varepsilon_1 \left(\frac{1-\kappa}{\kappa}\right) (1 - |\delta| - \vartheta)\right) \|p_{k+1} - p_k\|^2 \\ & \leq \|p_k - p^*\|^2 - \vartheta \|p_{k-1} - p^*\|^2 - \delta \|p_{k-2} - p^*\|^2 \\ & \quad + \left(\varepsilon_1 \left(\frac{1-\kappa}{\kappa}\right) (1 - |\delta| - \vartheta)\right) \|p_k - p_{k-1}\|^2 \\ & \quad + \left((1 + \vartheta)(\vartheta - \delta) - \varepsilon_1 \left(\frac{1-\kappa}{\kappa}\right) (\vartheta^2 - 2\vartheta - |\delta|\vartheta - |\delta| + 1)\right) \|p_k - p_{k-1}\|^2 \\ & \quad - \left(\delta(\vartheta - \delta) + \varepsilon_1 \left(\frac{1-\kappa}{\kappa}\right) (\delta^2 - |\delta| - |\delta|\vartheta)\right) \|p_{k-1} - p_{k-2}\|^2 \end{aligned} \quad (4.9)$$

eşitsizliğini elde edilir. Her $k \geq 1$ için

$$\begin{aligned} \Gamma_k &:= \|p_k - p^*\|^2 - \vartheta \|p_{k-1} - p^*\|^2 - \delta \|p_{k-2} - p^*\|^2 \\ &\quad + \left(\varepsilon_1 \left(\frac{1-\kappa}{\kappa}\right) (1 - |\delta| - \vartheta)\right) \|p_k - p_{k-1}\|^2 \end{aligned}$$

olarak tanımlansın. Öncelikle $\Gamma_k \geq 0$ olduğunu gösterelim. Bunun için

$$\|p_{k-1} - p^*\|^2 \leq 2\|p_k - p_{k-1}\|^2 + 2\|p_k - p^*\|^2$$

olduğunu göz önünde bulundurursak

$$\Gamma_k = \|p_k - p^*\|^2 - \vartheta \|p_{k-1} - p^*\|^2 - \delta \|p_{k-2} - p^*\|^2$$

$$\begin{aligned}
 & + \left(\varepsilon_1 \left(\frac{1-\kappa}{\kappa} \right) (1 - |\delta| - \vartheta) \right) \|p_k - p_{k-1}\|^2 \\
 & \geq \|p_k - p^*\|^2 - 2\vartheta \|p_k - p_{k-1}\|^2 - 2\vartheta \|p_k - p^*\|^2 - \delta \|p_{k-2} - p^*\|^2 \\
 & + \left(\varepsilon_1 \left(\frac{1-\kappa}{\kappa} \right) (1 - |\delta| - \vartheta) \right) \|p_k - p_{k-1}\|^2 \\
 & = (1 - 2\vartheta) \|p_k - p^*\|^2 \\
 & + \left(\varepsilon_1 \left(\frac{1-\kappa}{\kappa} \right) (1 - |\delta| - \vartheta) - 2\vartheta \right) \|p_k - p_{k-1}\|^2 - \delta \|p_{k-2} - p^*\|^2 \quad (4.10)
 \end{aligned}$$

ifadesi elde edilir. Kabul 4.3'te (i)-(iii) ten

$$|\delta| < \frac{\varepsilon_1(1-\kappa)(1-\vartheta) - 2\kappa\vartheta}{\varepsilon_1(1-\kappa)}$$

olduğu görülür. Böylece (4.10) eşitsizliğinden her $k \geq 0$ için $\Gamma_k \geq 0$ eşitsizliği elde edilir.

Ayrıca Kabul 4.3'te (i)-(iii) ifadesine dayanarak $\vartheta < \frac{1}{3}$ ve

$$- \frac{\varepsilon_1(1-\kappa)(1-\vartheta) - 2\kappa\vartheta}{\varepsilon_1(1-\kappa)} < \delta$$

olduğundan (4.9) eşitsizliği kullanılarak

$$\begin{aligned}
 \Gamma_{k+1} - \Gamma_k & \leq \left((1 + \vartheta)(\vartheta - \delta) - \varepsilon_1 \left(\frac{1-\kappa}{\kappa} \right) \right. \\
 & \quad \cdot (\vartheta^2 - 2\vartheta - |\delta|\vartheta - |\delta| + 1) \left. \right) \|p_k - p_{k-1}\|^2 \\
 & \quad - \left(\delta(\vartheta - \delta) + \varepsilon_1 \left(\frac{1-\kappa}{\kappa} \right) (\delta^2 - |\delta| - |\delta|\vartheta) \right) \|p_{k-1} - p_{k-2}\|^2 \\
 & \quad - \left((1 + \vartheta)(\vartheta - \delta) - \varepsilon_1 \left(\frac{1-\kappa}{\kappa} \right) (\vartheta^2 - 2\vartheta - |\delta|\vartheta - |\delta| + 1) \right)
 \end{aligned}$$

$$\begin{aligned}
 & \cdot (\|p_{k-1} - p_{k-2}\|^2 - \|p_k - p_{k-1}\|^2) \\
 & + \left((1 + \vartheta)(\vartheta - \delta) - \varepsilon_1 \left(\frac{1-\kappa}{\kappa} \right) (\vartheta^2 - 2\vartheta - |\delta|\vartheta - |\delta| + 1) \right. \\
 & \left. - \delta(\vartheta - \delta) - \varepsilon_1 \left(\frac{1-\kappa}{\kappa} \right) (\delta^2 - |\delta| - |\delta|\vartheta) \right) \|p_{k-1} - p_{k-2}\|^2 \\
 & \leq c_1 (\|p_{k-1} - p_{k-2}\|^2 - \|p_k - p_{k-1}\|^2) - c_2 \|p_{k-1} - p_{k-2}\|^2 \quad (4.11)
 \end{aligned}$$

ifadesi bulunur. Burada c_1 ve c_2 terimleri ise

$$\begin{aligned}
 c_1 &= - \left((1 + \vartheta)(\vartheta - \delta) - \varepsilon_1 \left(\frac{1-\kappa}{\kappa} \right) (\vartheta^2 - 2\vartheta - |\delta|\vartheta - |\delta| + 1) \right) \\
 c_2 &= - \left((1 + \vartheta)(\vartheta - \delta) - \varepsilon_1 \left(\frac{1-\kappa}{\kappa} \right) (\vartheta^2 - 2\vartheta - |\delta|\vartheta - |\delta| + 1) - \delta(\vartheta - \delta) \right. \\
 & \quad \left. - \varepsilon_1 \left(\frac{1-\kappa}{\kappa} \right) (\delta^2 - |\delta| - |\delta|\vartheta) \right)
 \end{aligned}$$

şeklindedir. $|\delta| = -\delta$ olduğu göz önüne alınarak

$$- \left((1 + \vartheta)(\vartheta - \delta) - \varepsilon_1 \left(\frac{1-\kappa}{\kappa} \right) (\vartheta^2 - 2\vartheta - |\delta|\vartheta - |\delta| + 1) \right) > 0,$$

bulunmuş olur. Bu son eşitsizlik ise

$$\delta > \frac{\vartheta(1 + \vartheta) - \varepsilon_1 \left(\frac{1-\kappa}{\kappa} \right) (1 - \vartheta)^2}{(1 + \vartheta) \left(1 + \varepsilon_1 \left(\frac{1-\kappa}{\kappa} \right) \right)} \quad (4.12)$$

eşitsizliğine denktir. Buna göre, Kabul 4.3 te (iii) ifadesinden $c_1 > 0$ olduğu elde edilir.

Şimdi, $c_2 > 0$ olduğunu gösterelim. Buradan hareketle

$$(1 + \vartheta)(\delta - \vartheta) + \varepsilon_1 \left(\frac{1 - \kappa}{\kappa} \right) (\vartheta^2 - 2\vartheta - |\delta|\vartheta - |\delta| + 1) + \delta(\vartheta - \delta) \\ + \varepsilon_1 \left(\frac{1 - \kappa}{\kappa} \right) (\delta^2 - |\delta| - |\delta|\vartheta) > 0 \quad (4.13)$$

olup bu ise

$$\delta(\vartheta - \delta) + \delta(1 + \vartheta) \left[1 + \varepsilon_1 \left(\frac{1 - \kappa}{\kappa} \right) \right] + \varepsilon_1 \left(\frac{1 - \kappa}{\kappa} \right) (\delta^2 + \delta(1 + \vartheta)) \\ < \vartheta(1 + \vartheta) - \varepsilon_1 \left(\frac{1 - \kappa}{\kappa} \right) (\vartheta - 1)^2 \quad (4.14)$$

olmasını gerektirir. Dolayısıyla, Kabul 4.3'te (iii) ifadesinden yine açıkça görülür ki $c_2 > 0$ 'dır. Öte yandan, (4.11)'den

$$\Gamma_{k+1} + c_1 \|p_k - p_{k-1}\|^2 \leq \Gamma_k + c_1 \|p_{k-1} - p_{k-2}\|^2 - c_2 \|p_{k-1} - p_{k-2}\|^2 \quad (4.15)$$

sonucu elde edilir. $\bar{\Gamma}_k := \Gamma_k + c_1 \|p_{k-1} - p_{k-2}\|^2$ olmak üzere (4.15)'ten $\bar{\Gamma}_{k+1} \leq \bar{\Gamma}_k$ olduğu sonucuna ulaşılır. Bu da $\{\bar{\Gamma}_k\}$ dizisinin azalan olduğunu ve dolayısıyla $\lim_{k \rightarrow \infty} \bar{\Gamma}_k$ limitinin var olduğunu gösterir. Sonuç olarak, (4.15) eşitsizliğinden

$$\lim_{k \rightarrow \infty} \Gamma_k + c_1 \|p_{k-1} - p_{k-2}\|^2 = 0$$

olduğu elde edilir. Dolayısıyla:

$$\lim_{k \rightarrow \infty} \|p_{k-1} - p_{k-2}\| = 0 \quad (4.16)$$

dır. (4.16) ifadesi ve $\{\bar{\Gamma}_k\}$ dizisinin limitinin var olduğu dikkate alındığında

$$\begin{aligned} \lim_{k \rightarrow \infty} \Gamma_k &= \lim_{k \rightarrow \infty} (\|p_k - p^*\|^2 - \vartheta \|p_{k-1} - p^*\|^2 - \delta \|p_{k-2} - p^*\|^2 \\ &\quad + \varepsilon_1 \left(\frac{1-\kappa}{\kappa} \right) (1 - |\delta| - \vartheta) \|p_k - p_{k-1}\|^2) \end{aligned} \quad (4.17)$$

limiti vardır. Bunun yanı sıra

$$\begin{aligned} \|p_{k+1} - w_k\| &= \|p_{k+1} - p_k - \vartheta(p_k - p_{k-1}) - \delta(p_{k-1} - p_{k-2})\| \\ &\leq \|p_{k+1} - p_k\| + \vartheta \|p_k - p_{k-1}\| + |\delta| \|p_{k-1} - p_{k-2}\| \rightarrow 0, \quad k \rightarrow \infty \end{aligned}$$

dır. Dolayısıyla

$$\lim_{k \rightarrow \infty} \|w_k - \mathcal{T}w_k\| = 0$$

limiti elde edilir. Ayrıca

$$\begin{aligned} \|w_k - p_k\| &= \|p_k + \vartheta(p_k - p_{k-1}) + \delta(p_{k-1} - p_{k-2}) - p_k\| \\ &= \|\vartheta(p_k - p_{k-1}) + \delta(p_{k-1} - p_{k-2})\| \\ &\leq \vartheta \|p_k - p_{k-1}\| + |\delta| \|p_{k-1} - p_{k-2}\| \rightarrow 0, \quad k \rightarrow \infty. \end{aligned}$$

dır.

$\lim_{k \rightarrow \infty} \Gamma_k$ limitinin var olması ve $\lim_{k \rightarrow \infty} \|p_k - p_{k-1}\| = 0$ olduğu göz önünde bulundurulduğunda, (4.10) ifadesinden $\{p_k\}$ dizisinin sınırlı olduğu elde edilir.

Aşağıda Algoritma 4.1 tarafından üretilen dizinin zayıf yakınsaklığı ispatlanmıştır.

Teorem 4.5: Kabul 4.2'deki şartlar altında Algoritma 4.1 tarafından üretilen $\{p_k\}$ dizisi (3.1) probleminin bir çözümüne zayıf yakınsaktır.

İspat: (4.16) eşitsizliği (4.17) içinde kullanıldığında, $k \rightarrow \infty$ için

$$\lim_{k \rightarrow \infty} [\|p_k - p^*\|^2 - \vartheta \|p_{k-1} - p^*\|^2 - \delta \|p_{k-2} - p^*\|^2]$$

limiti mevcuttur. Lemma 4.4'e göre, $\{p_k\}$ dizisinin sınırlı olduğu bilinmektedir. İlk olarak, $\{p_k\}$ dizisinin herhangi bir zayıf yığılma noktasının $F(\mathcal{T})$ kümesine ait olduğunu gösterelim. $\{p_{k_n}\} \subset \{p_k\}$ ve $p_{k_n} \rightharpoonup v^* \in \mathcal{H}$ olduğunu varsayalım. $k \rightarrow \infty$ için $\|w_k - p_k\| \rightarrow 0$ olduğundan, $w_{k_n} \rightarrow v^* \in \mathcal{H}$ elde edilir. Sonuç olarak $k \rightarrow \infty$ için $\|w_k - \mathcal{T}w_k\| \rightarrow 0$ olduğundan ve $\mathcal{T}$ operatörünün yarı kapalı olmasından dolayı $v^* \in F(\mathcal{T}) = (A + B)^{-1}(0)$ olduğu sonucuna varılır.

Şimdi $p_k \rightharpoonup p^* \in F(\mathcal{T})$ olduğunu gösterelim. $\{p_{k_n}\} \subset \{p_k\}$ ve $\{p_{k_j}\} \subset \{p_k\}$ olmak üzere $n \rightarrow \infty$ için $p_{k_n} \rightharpoonup v^*$ ve $j \rightarrow \infty$ için $p_{k_j} \rightharpoonup p^*$ olduğunu kabul edelim. $v^* = p^*$ olduğunu göstermek istiyoruz. Bunun için

$$2\langle p_k, p^* - v^* \rangle = \|p_k - v^*\|^2 - \|p_k - p^*\|^2 - \|v^*\|^2 + \|p^*\|^2, \quad (4.18)$$

$$\begin{aligned} 2\langle -\vartheta p_{k-1}, p^* - v^* \rangle &= -\vartheta \|p_{k-1} - v^*\|^2 + \vartheta \|p_{k-1} - p^*\|^2 \\ &\quad + \vartheta \|v^*\|^2 - \vartheta \|p^*\|^2 \end{aligned} \quad (4.19)$$

ve

$$\begin{aligned} 2\langle -\delta p_{k-2}, p^* - v^* \rangle &= -\delta \|p_{k-2} - v^*\|^2 + \delta \|p_{k-2} - p^*\|^2 \\ &\quad + \delta \|v^*\|^2 - \delta \|p^*\|^2 \end{aligned} \quad (4.20)$$

dir. (4.18), (4.19) ve (4.20) ifadelerinin toplamı

$$\begin{aligned}
 2\langle p_k - \vartheta p_{k-1} - \delta p_{k-2}, p^* - v^* \rangle &= (\|p_k - v^*\|^2 - \vartheta \|p_{k-1} - v^*\|^2 \\
 &\quad - \delta \|p_{k-2} - v^*\|^2) \\
 &\quad - (\|p_k - p^*\|^2 - \vartheta \|p_{k-1} - p^*\|^2 \\
 &\quad - \delta \|p_{k-2} - p^*\|^2) \\
 &\quad + (1 - \vartheta - \delta)(\|p^*\|^2 - \|v^*\|^2)
 \end{aligned}$$

eşitliğini verir. (4.17) ifadesine göre

$$\lim_{k \rightarrow \infty} (\|p_k - p^*\|^2 - \vartheta \|p_{k-1} - p^*\|^2 - \delta \|p_{k-2} - p^*\|^2),$$

ve

$$\lim_{k \rightarrow \infty} (\|p_k - v^*\|^2 - \vartheta \|p_{k-1} - v^*\|^2 - \delta \|p_{k-2} - v^*\|^2)$$

ifadeleri yazılır. Buradan

$$\lim_{k \rightarrow \infty} \langle p_k - \vartheta p_{k-1} - \delta p_{k-2}, p^* - v^* \rangle$$

limiti mevcuttur. Dolayısıyla

$$\begin{aligned}
 \langle v^* - \vartheta v^* - \delta v^*, p^* - v^* \rangle &= \lim_{n \rightarrow \infty} \langle p_{k_n} - \vartheta p_{k_n-1} - \delta p_{k_n-2}, p^* - v^* \rangle \\
 &= \lim_{k \rightarrow \infty} \langle p_k - \vartheta p_{k-1} - \delta p_{k-2}, p^* - v^* \rangle \\
 &= \lim_{j \rightarrow \infty} \langle p_{k_j} - \vartheta p_{k_j-1} - \delta p_{k_j-2}, p^* - v^* \rangle \\
 &= \langle p^* - \vartheta p^* - \delta p^*, p^* - v^* \rangle
 \end{aligned}$$

eşitliği elde edilir. Böylece

$$(1 - \vartheta - \delta)\|p^* - v^*\|^2 = 0.$$

olduğu sonucuna varılır. $\delta \leq 0 < 1 - \vartheta$ olduğu için $p^* = v^*$ elde edilir. Sonuç olarak, $\{p_k\}$ dizisi $F(\mathcal{T}) = (A + B)^{-1}(0)$ kümesindeki bir noktaya zayıf yakınsar.

Şimdi ise yine reel Hilbert uzayında monoton kapsama problemlerini çözmek amacıyla hem çift eylemsizlik terimli hemde viskozite temelli ileri-geri ayrıştırma algoritması (Double Inertial Viscosity Forward-Backward Splitting Algorithm-DIVFBS) verilecektir. Bu algoritmanın yapısı aşağıda ayrıntılı olarak sunulmaktadır.

Algoritma 4.6: $h, \mathcal{H}$ üzerinde tanımlı c -daraltan bir dönüşüm olsun. $\{\alpha_k\}, [0,1]$ kümesinde bir dizi, $0 < \lambda_k \leq 2\alpha, \tau \in [0,1]$ olmak üzere $\{\theta_k\} \subset [0, \tau]$ ve $\{\delta_k\} \leq 0$ olsun. Bu durumda DIVFBS algoritması

$$\begin{cases} \omega_k = x_k + \theta_k(x_k - x_{k-1}) + \delta_k(x_{k-1} - x_{k-2}), \\ x_{k+1} = \alpha_k h(x_k) + (1 - \alpha_k)J_{\lambda_k}^B(\omega_k - \lambda_k A\omega_k), \end{cases}$$

şeklinde tanımlanmaktadır.

Aşağıda bu algoritmanın yakınsaklık teoremi ifade ve ispat edilmiştir.

Teorem 4.7: $\mathcal{H}$ reel Hilbert uzayı üzerinde $A: \mathcal{H} \rightarrow \mathcal{H}$ operatörü α -ters kuvvetli monoton operatör ve $B: \mathcal{H} \rightarrow 2^{\mathcal{H}}$ operatörü ise $\Omega := (A + B)^{-1}(0) \neq \emptyset$ şartını sağlayan maksimal monoton bir operatör olsun. DIVFBS algoritması ile üretilen $\{x_k\}$ dizisi,

- i. $\sum_{k=1}^{\infty} \theta_k \|x_k - x_{k-1}\| < \infty, \sum_{k=1}^{\infty} |\delta_k| \|x_{k-1} - x_{k-2}\| < \infty,$
- ii. $\lim_{k \rightarrow \infty} \alpha_k = 0, \sum_{k=1}^{\infty} \alpha_k = \infty,$
- iii. $0 < \liminf_{k \rightarrow \infty} \lambda_k \leq \limsup_{k \rightarrow \infty} \lambda_k < 2\alpha$

kabulleri altında $x^\circ = P_\Omega(h(x^\circ))$ eşitliğini sağlayan bir $x^\circ \in \Omega$ noktasına kuvvetli yakınsar.

İspat: $Id, \mathcal{H}$ reel Hilbert uzayı üzerinde birim dönüşüm olmak üzere $T_k = J_{\lambda_k}^B(Id - \lambda_k A)$ dönüşümü genişlemeyen bir dönüşüm olduğu bilinmektedir. Öncelikle $\{x_k\}$ dizisinin sınırlı olduğunu gösterelim. $x^\circ \in \Omega = F(T_k) = (A + B)^{-1}(0)$ alınırsa

$$\|w_k - x^\circ\| \leq \|x_k - x^\circ\| + \theta_k \|x_k - x_{k-1}\| + |\delta_k| \|x_{k-1} - x_{k-2}\| \quad (4.21)$$

eşitsizliği elde edilir. Burada h fonksiyonunun c -daraltan bir dönüşüm olması ve T_k dönüşümünün genişlemeyen bir dönüşüm olması nedeniyle, (4.21) eşitsizliğinden

$$\begin{aligned} \|x_{k+1} - x^\circ\| &= \|\alpha_k h(x_k) + (1 - \alpha_k) T_k w_k - x^\circ\| \\ &\leq \alpha_k \|h(x_k) - x^\circ\| \\ &\quad + (1 - \alpha_k) \|T_k(x_k + \theta_k(x_k - x_{k-1}) + \delta_k(x_{k-1} - x_{k-2})) - x^\circ\| \\ &\leq \alpha_k \|h(x_k) - x^\circ\| + (1 - \alpha_k) \|x_k - x^\circ\| \\ &\quad + (1 - \alpha_k) \|\theta_k(x_k - x_{k-1})\| + (1 - \alpha_k) \|\delta_k(x_{k-1} - x_{k-2})\| \\ &\leq \alpha_k \|h(x_k) - h(x^\circ)\| + \alpha_k \|h(x^\circ) - x^\circ\| + (1 - \alpha_k) \|x_k - x^\circ\| \\ &\quad + (1 - \alpha_k) \|\theta_k(x_k - x_{k-1})\| + (1 - \alpha_k) \|\delta_k(x_{k-1} - x_{k-2})\| \\ &\leq \alpha_k c \|x_k - x^\circ\| + \alpha_k \|h(x^\circ) - x^\circ\| + (1 - \alpha_k) \|x_k - x^\circ\| \\ &\quad + (1 - \alpha_k) \|\theta_k(x_k - x_{k-1})\| + (1 - \alpha_k) \|\delta_k(x_{k-1} - x_{k-2})\| \\ &= (1 - \alpha_k(1 - c)) \|x_k - x^\circ\| + \alpha_k \|h(x^\circ) - x^\circ\| \\ &\quad + (1 - \alpha_k) \theta_k \|x_k - x_{k-1}\| + (1 - \alpha_k) |\delta_k| \|x_{k-1} - x_{k-2}\| \end{aligned}$$

sonucu elde edilir. Teoremdeki (i) şartından

$$\theta_k \|x_k - x_{k-1}\| \leq M_1$$

ve

$$|\delta_k| \|x_{k-1} - x_{k-2}\| \leq M_2$$

eşitsizliğini sağlayan $M_1, M_2 \in \mathbb{R}$ mevcuttur. Böylece

$$\begin{aligned} \|x_{k+1} - x^o\| &\leq \max \left\{ \|x_k - x^o\|, \frac{\|h(x^o) - x^o\| + M_1 + M_2}{1 - c} \right\} \\ &\quad \vdots \\ &\leq \max \left\{ \|x_1 - x^o\|, \frac{\|h(x^o) - x^o\| + M_1 + M_2}{1 - c} \right\} \end{aligned}$$

yazılır. Dolayısıyla, $\{x_k\}$ ve $\{w_k\}$ dizileri sınırlıdır. Şimdi $\{x_k\}$ dizisinin $x^o = P_\Omega h(x^o)$ noktasına kuvvetli yakınsadığını gösterelim. x^o noktası $P_\Omega h(x)$ dönüşümünün bir sabit noktası olduğu için ve T_k genişlemeyen bir dönüşüm olduğundan, Lemma 3.1.18-(ii)'den

$$\begin{aligned} \|T_k w_k - x^o\|^2 &= \|w_k - x^o\|^2 \\ &= \|x_k + \theta_k(x_k - x_{k-1}) + \delta_k(x_{k-1} - x_{k-2}) - x^o\|^2 \\ &\leq \|x_k - x^o\|^2 + 2\theta_k \langle x_k - x_{k-1}, w_k - x^o \rangle \\ &\quad + 2\delta_k \langle x_{k-1} - x_{k-2}, w_k - x^o \rangle \end{aligned} \tag{4.22}$$

eşitsizliği elde edilir. Böylece, (4.22)'den ve Cauchy-Schwarz eşitsizliğinden

$$\|x_{k+1} - x^o\|^2 = \langle \alpha_k h(x_k) + (1 - \alpha_k)T_k w_k - x^o, x_{k+1} - x^o \rangle$$

$$\begin{aligned}
 &= \langle \alpha_k(h(x_k) - x^o) + (1 - \alpha_k)(T_k w_k - x^o), x_{k+1} - x^o \rangle \\
 &= \alpha_k \langle h(x_k) - h(x^o), x_{k+1} - x^o \rangle + \alpha_k \langle h(x_k) - x^o, x_{k+1} - x^o \rangle \\
 &\quad + (1 - \alpha_k) \langle T_k w_k - x^o, x_{k+1} - x^o \rangle \\
 &\leq \frac{\alpha_k}{2} (\|h(x_k) - h(x^o)\|^2 + \|x_{k+1} - x^o\|^2) \\
 &\quad + \alpha_k \langle h(x^o) - x^o, x_{k+1} - x^o \rangle \\
 &\quad + \frac{(1-\alpha_k)}{2} (\|T_k w_k - x^o\|^2 + \|x_{k+1} - x^o\|^2) \\
 &\leq \frac{\alpha_k}{2} (\|h(x_k) - h(x^o)\|^2 + \|x_{k+1} - x^o\|^2) \\
 &\quad + \alpha_k \langle h(x^o) - x^o, x_{k+1} - x^o \rangle \\
 &\quad + \frac{(1-\alpha_k)}{2} (\|x_k - x^o\|^2 + 2\theta_k \langle x_k - x_{k-1}, w_k - x^o \rangle \\
 &\quad + 2\delta_k \langle x_{k-1} - x_{k-2}, w_k - x^o \rangle + \|x_{k+1} - x^o\|^2) \\
 &\leq \frac{(\alpha_k c^2)}{2} \|x_k - x^o\|^2 + \frac{\alpha_k}{2} \|x_{k+1} - x^o\|^2 \\
 &\quad + \alpha_k \langle h(x^o) - x^o, x_{k+1} - x^o \rangle + \frac{(1-\alpha_k)}{2} \|x_k - x^o\|^2 \\
 &\quad + (1 - \alpha_k) \theta_k \langle x_k - x_{k-1}, w_k - x^o \rangle \\
 &\quad + (1 - \alpha_k) \delta_k \langle x_{k-1} - x_{k-2}, w_k - x^o \rangle + \frac{(1-\alpha_k)}{2} \|x_{k+1} - x^o\|^2 \\
 &\leq \frac{(1 - \alpha_k(1 - c^2))}{2} \|x_k - x^o\|^2 + \frac{1}{2} \|x_{k+1} - x^o\|^2 \\
 &\quad + \alpha_k \langle h(x^o) - x^o, x_{k+1} - x^o \rangle + (1 - \alpha_k) \theta_k \langle x_k - x_{k-1}, w_k - x^o \rangle \\
 &\quad + (1 - \alpha_k) \delta_k \langle x_{k-1} - x_{k-2}, w_k - x^o \rangle
 \end{aligned}$$

sonucu elde edilir. Böylece

$$\begin{aligned}
 \|x_{k+1} - x^o\|^2 &\leq (1 - \alpha_k(1 - c^2))\|x_k - x^o\|^2 \\
 &\quad + 2(1 - \alpha_k)\theta_k \langle x_k - x_{k-1}, w_k - x^o \rangle \\
 &\quad + 2(1 - \alpha_k)\delta_k \langle x_{k-1} - x_{k-2}, w_k - x^o \rangle \\
 &\quad + \alpha_k(1 - c^2) \left(\frac{2}{(1 - c^2)} \langle h(x^o) - x^o, x_{k+1} - x^o \rangle \right)
 \end{aligned} \tag{4.23}$$

ifadesi yazılır. Şimdi aşağıdaki iki durumu inceleyelim.

Durum 1: $n_0 \in \mathbb{N}$ olmak üzere $\{\|x_k - x^o\|\}$ dizisinin monoton olarak azaldığını varsayalım. $\{x_k\}$ dizisi sınırlı bir dizi olduğundan $\{\|x_k - x^o\|\}$ dizisi yakınsak bir dizidir. Dolayısıyla Lemma 3.1.18-(iii)'ten

$$\begin{aligned}
 \|x_{k+1} - x^o\|^2 &= \|\alpha_k h(x_k) + (1 - \alpha_k)T_k w_k - x^o\|^2 \\
 &= \alpha_k \|h(x_k) - x^o\|^2 + (1 - \alpha_k) \|T_k w_k - x^o\|^2 \\
 &\quad - (1 - \alpha_k) \alpha_k \|T_k w_k - h(x_k)\|^2 \\
 &\leq \alpha_k \|h(x_k) - x^o\|^2 + (1 - \alpha_k) \|w_k - x^o\|^2
 \end{aligned}$$

eşitsizliği elde edilir. Buradan hareketle

$$\|w_k - x^o\|^2 \geq \|x_{k+1} - x^o\|^2 - \alpha_k (\|h(x_k) - x^o\|^2 - \|w_k - x^o\|^2)$$

olarak yazılır. Öte yandan teoremdaki (i) koşulundan

$$\begin{aligned}
 \|w_k - x_k\| &= \|x_k + \theta_k(x_k - x_{k-1}) + \delta_k(x_{k-1} - x_{k-2}) - x_k\| \\
 &\leq \theta_k \|x_k - x_{k-1}\| + |\delta_k| \|x_{k-1} - x_{k-2}\|
 \end{aligned} \tag{4.24}$$

bulunur. Böylece

$$\lim_{k \rightarrow \infty} \|w_k - x_k\| = 0 \quad (4.25)$$

dır. Ayrıca, x_{k+1} 'in tanımından

$$\begin{aligned} \|x_{k+1} - w_k\| &= \|\alpha_k h(x_k) + (1 - \alpha_k)T_k w_k - w_k\| \\ &\leq \alpha_k \|h(x_k) - w_k\| + (1 - \alpha_k) \|T_k w_k - w_k\| \\ &\leq \alpha_k c \|x_k - x^o\| + \alpha_k \|h(x_k) - w_k\| \\ &\quad + (1 - \alpha_k) \|T_k w_k - w_k\| \end{aligned} \quad (4.26)$$

elde edilir. Ayrıca

$$\|T_k w_k - x^o\| \geq \|x_{k+1} - x^o\| - \alpha_k \|h(x_k) - T_k w_k\|,$$

eşitsizliği kullanılırsa

$$\begin{aligned} 0 &\leq \|w_k - x^o\| - \|T_k w_k - T_k x^o\| \\ &= \|w_k - x^o\| - \|T_k w_k - x^o\| \\ &\leq \|w_k - x^o\| + \alpha_k \|h(x_k) - T_k w_k\| - \|x_{k+1} - x^o\| \end{aligned}$$

ifadesi elde edilir. $\{\|x_k - x^o\|\}$ dizisi yakınsak olduğundan teoremdaki (ii) şartından ve $\{w_k\}$ dizisinin sınırlı olmasından

$$\|w_k - x^o\| - \|T_k w_k - T_k x^o\| \rightarrow 0, \quad k \rightarrow \infty.$$

yazılabilir. T_k dönüşümünün genişlemeyen dönüşüm olmasından dolayı

$$\lim_{k \rightarrow \infty} \|T_k w_k - w_k\| = 0 \quad (4.27)$$

eşitliği elde edilir. (4.26), (4.27) ve teoremdaki (i) şartından

$$\lim_{k \rightarrow \infty} \|x_{k+1} - w_k\| = 0$$

olduğu sonucuna ulaşılır. (4.25) ifadesinden ise

$$\|x_{k+1} - x_k\| \leq \|x_{k+1} - w_k\| + \|w_k - x_k\| \rightarrow 0 \text{ as } k \rightarrow \infty \quad (4.28)$$

olur. Ayrıca

$$\begin{aligned} \|x_k - T_k x_k\| &= \|x_k - x_{k+1} + x_{k+1} + T_k w_k - T_k w_k - T_k x_k\| \\ &\leq \|x_k - x_{k+1}\| + \|x_{k+1} - T_k w_k\| + \|T_k w_k - T_k x_k\| \\ &\leq \|x_k - x_{k+1}\| + \|x_{k+1} - T_k w_k\| + \|w_k - x_k\| \end{aligned} \quad (4.29)$$

dır. x_{k+1} 'in tanımı ve (4.29) ifadesinden

$$\begin{aligned} \|x_{k+1} - T_k w_k\| &= \|\alpha_k h(x_k) + (1 - \alpha_k)T_k w_k - T_k w_k\| \\ &= \alpha_k \|h(x_k) - T_k w_k\| \\ &\leq \alpha_k \|h(x_k) - h(x^o)\| + \alpha_k \|h(x^o) - T_k w_k\| \\ &\leq \alpha_k c \|x_k - x^o\| + \alpha_k \|h(x^o) - T_k w_k\| \rightarrow 0 \text{ as } k \rightarrow \infty \end{aligned} \quad (4.30)$$

sonucuna ulaşılır. (4.25), (4.28) ve (4.30) ifadelerini (4.29) ifadesinde yerine koyarsak

$$\lim_{k \rightarrow \infty} \|x_k - T_k x_k\| = 0 \quad (4.31)$$

sonucu elde edilir. $\{x_k\}$ dizisi sınırlı olduğundan, bu dizinin $\bar{x} \in \mathcal{H}$ elemanına zayıf olarak yakınsayan bir $\{x_{k_i}\}$ alt dizisi mevcuttur. Şimdi $\bar{x} \in \Omega$ olduğu gösterilecektir. Her $i \in \mathbb{N}$ için

$$\begin{aligned} \|(I - T_k)\bar{x}\|^2 &\leq \langle (I - T_k)\bar{x}, \bar{x} - x_{k_i} \rangle + \langle (I - T_k)\bar{x}, T_k x_{k_i} - T_k \bar{x} \rangle \\ &\quad + \langle (I - T_k)\bar{x}, x_{k_i} - T_k x_{k_i} \rangle \end{aligned} \quad (4.32)$$

eşitsizliği geçerlidir. (4.31)'den

$$\lim_{k \rightarrow \infty} \|x_{k_i} - T_k x_{k_i}\| = 0 \quad (4.33)$$

eşitliği elde edilir. (4.33) ifadesi ve $\{x_{k_i}\} \rightarrow \bar{x}$ olduğundan

$$\lim_{k \rightarrow \infty} \|(I - T_k)\bar{x}\| = 0 \quad (4.34)$$

dır. Böylece $\bar{x} = T_k \bar{x}$ ve $\bar{x} \in \Omega$ elde edilir. Şimdi, $\{x_k\}$ dizisinin $x^o = P_\Omega(x^o)$ noktasına kuvvetli yakınsadığı gösterilecektir. (4.28) ifadesinden $k \rightarrow \infty$ iken $\|x_{k+1} - x_k\| \rightarrow 0$ olduğu bilinmektedir. Bu durumda, $\{x_{k+1}\}$ dizisinin bir alt dizisi olan $\{x_{k_{i+1}}\}$ alt dizisi için

$$\limsup_{k \rightarrow \infty} \frac{2}{1 - c^2} \langle h(x^o) - x^o, x_{k+1} - x^o \rangle = \lim_{i \rightarrow \infty} \frac{2}{1 - c^2} \langle h(x^o) - x^o, x_{k_{i+1}} - x^o \rangle \quad (4.35)$$

eşitliği yazılabilir. $k \rightarrow \infty$ iken $\|x_{k+1} - x_k\| \rightarrow 0$ ve $x_{k_i} \rightarrow \bar{x}$ olduğundan, $\{x_{k_{i+1}}\}$ dizisinin de $\bar{x}$ noktasına yakınsadığı sonucu çıkarılır. Bu durumda, (4.35) ifadesinden

$$\limsup_{k \rightarrow \infty} \frac{2}{1-c^2} \langle h(x^o) - x^o, x_{k+1} - x^o \rangle = \frac{2}{1-c^2} \langle h(x^o) - x^o, \bar{x} - x^o \rangle$$

elde edilir. (4.23) ifadesi, $\{w_k\}$ ve $\{x_k\}$ dizilerinin sınırlılığı ile $\lim_{k \rightarrow \infty} \alpha_k = 0$ olduğu göz önüne alınarak

$$\sum_{k=1}^{\infty} 2(1-\alpha_k) \theta_k \langle x_k - x_{k-1}, w_k - x^o \rangle < \infty$$

ve

$$\sum_{k=1}^{\infty} 2(1-\alpha_k) \delta_k \langle x_{k-1} - x_{k-2}, w_k - x^o \rangle < \infty$$

sonuçları elde edilir. Lemma 3.1.20 ve (4.23) ifadesi kullanılarak $k \rightarrow \infty$ iken $x_k \rightarrow x^o$ olur. $k \rightarrow \infty$ iken $\|w_k - x_k\| \rightarrow 0$ olduğundan, $w_k \rightarrow x^o$ olduğu görülür.

Durum 2: $\{\|x_k - x^o\|\}$ dizisinin monoton azalan bir dizi olmadığı varsayalım. $\forall k \in \mathbb{N}$ için $\gamma_k = \|x_k - x^o\|$ olarak tanımlansın ve tüm $k \geq k_0$ değerleri için $v: \mathbb{N} \rightarrow \mathbb{N}$ dönüşümü

$$v(k) = \max\{\ell \in \mathbb{N}: \ell \leq k, \gamma_\ell \leq \gamma_{\ell+1}\}$$

şekilde verilsin. Açıkça görülmektedir ki, $\{v(k)\}$ azalmayan bir dizidir. Ayrıca

$$\lim_{k \rightarrow \infty} v(k) = \infty$$

ve her $k \geq k_0$ için

$$0 \leq \gamma_{v(k)} \leq \gamma_{v(k)+1}$$

dır. Benzer şekilde, Durum 1 de yapılan çıkarımlar doğrultusunda sırayla aşağıdaki sonuçlar elde edilir:

- $\lim_{k \rightarrow \infty} \|x_{v(k)+1} - x_{v(k)}\| = 0,$
- $\limsup_{k \rightarrow \infty} \frac{2}{1-k^2} \langle h(x^o) - x^o, x_{v(k)+1} - x^o \rangle \leq 0,$
- $\sum_{k=1}^{\infty} 2(1 - \alpha_{v(k)}) \vartheta_{v(k)} \langle x_{v(k)} - x_{v(k)-1}, w_{v(k)} - x^o \rangle < \infty,$
- $\sum_{k=1}^{\infty} 2(1 - \alpha_{v(k)}) \delta_{v(k)} \langle x_{v(k)-1} - x_{v(k)-2}, w_{v(k)} - x^o \rangle < \infty.$

$\{x_{v(k)}\}$ dizisi sınırlı olduğundan, bu dizinin $\bar{x} \in \Omega$ elemanına zayıf olarak yakınsayan bir alt dizisi mevcuttur. Böylece, (4.23) ifadesi kullanılarak

$$\begin{aligned} \|x_{v(k)+1} - x^o\|^2 &\leq \left(1 - \alpha_{v(k)}(1 - c^2)\right) \|x_{v(k)} - x^o\|^2 \\ &\quad + \alpha_{v(k)}(1 - c^2) \left(\frac{2}{1-c^2} \langle h(x^o) - x^o, x_{v(k)+1} - x^o \rangle \right) \\ &\quad + 2(1 - \alpha_{v(k)}) \vartheta_{v(k)} \langle x_{v(k)} - x_{v(k)-1}, w_{v(k)} - x^o \rangle \\ &\quad + 2(1 - \alpha_{v(k)}) \delta_{v(k)} \langle x_{v(k)-1} - x_{v(k)-2}, w_{v(k)} - x^o \rangle \end{aligned} \quad (4.36)$$

eşitsizliği elde edilir. Her $v(k) \in \mathbb{N}$ için $\alpha_{v(k)}(1 - c^2) > 0$ olduğundan,

$$\limsup_{k \rightarrow \infty} \|x_{v(k)} - x^o\|^2 \leq 0$$

dır. Böylece

$$\lim_{k \rightarrow \infty} \|x_{v(k)} - x^o\|^2 = 0$$

ve bunun doğal bir sonucu olarak

$$\lim_{k \rightarrow \infty} \|x_{v(k)} - x^o\| = 0 \quad (4.37)$$

ifadesi elde edilir. $\lim_{k \rightarrow \infty} \|x_{v(k)+1} - x_{v(k)}\| = 0$ limiti ve (4.37) ifadesi kullanılarak $k \rightarrow \infty$ için,

$$\|x_{v(k)+1} - x_{v(k)}\| \leq \|x_{v(k)+1} - x^o\| + \|x^o - x_{v(k)}\| \rightarrow 0 \quad (4.38)$$

sonucuna varılır. Ek olarak $k \geq k_0$ için, $k \neq v(k)$ ise $\gamma_{v(k)} \leq \gamma_{v(k)+1}$ olduğu kolayca görülür, çünkü $v(k) + 1 \leq j \leq k$ için $\gamma_j \geq \gamma_{j+1}$ olur. Sonuç olarak her $k \geq k_0$ için,

$$0 \leq \gamma_k \leq \max\{\gamma_{v(k)}, \gamma_{v(k)+1}\} = \gamma_{v(k)+1}$$

dir. (4.38) kullanılarak $\lim_{k \rightarrow \infty} \gamma_k$ ve $\{x_k\}$ dizisinin x^o noktasına kuvvetli yakınsadığı görülür.

4.2. Algoritmaların Bilgisayar Programı İle Benzetimi

Bu bölümde tanımlanan algoritmaların regresyon ve sınıflandırma problemlerine uygulamaları verilmiştir. Regresyon ve veri sınıflandırma problemlerine yönelik benzetim aşamasında, Keel, UC Irvine Machine Learning Repository (Asuncion and Newman 2007) ile Scikit-learn kütüphanesi (Pedregosa et al. 2011) gibi açık erişimli veri kaynaklarında sunulan ve gerçek dünya senaryolarından elde edilen karşılaştırmalı veri kümeleri kullanılmıştır. Bu süreçte yararlanılan başlıca veri kümeleri şunlardır: Analysing Categorical Data (Anacalt) (Simonoff 2003), Delta Ailerons (Alcalá-Fdez et al. 2009), MachineCPU (Feldmesser 1987), California Housing (Pace and Barry 1997),

Breast Cancer Wisconsin (BCW) (Wolberg et al. 1993), Fashion-MNIST (Xiao et al. 2017), MNIST (LeCun et al. 1998), Digits (Pedregosa et al. 2011), Diabetes (Kahn 1994) ve Iris (Fisher 1936).

Bu algoritmaların literatürde verilen bazı algoritmalar ile performanslarını kıyaslamak amacıyla ortalama kare hatası (Mean Squared Error, MSE) (Sammur 2010b), ortalama mutlak hata (Mean Absolute Error, MAE) (Sammur 2010a), R-kare (R-squared, R^2) (Wright 1921), ortalama kare hatanın karekökü (Root Mean Square Error, RMSE) (Nevitt and Hancock 2000), işlem süresi (CPU time) ve çapraz doğrulama (Cross Validation, CV) (Stone 1974) gibi ölçümler gerçekleştirilmiştir. Karşılaştırma yapılan algoritmalar ise FISTA (Beck and Teboulle 2009), VIFBA (Verma et al. 2018), DIPFBSM (Jolaoso et al. 2023), Algoritma 5 (Inthakon et al. 2020) ve IVFBS (Kitkuan et al. 2020) olarak belirlenmiştir.

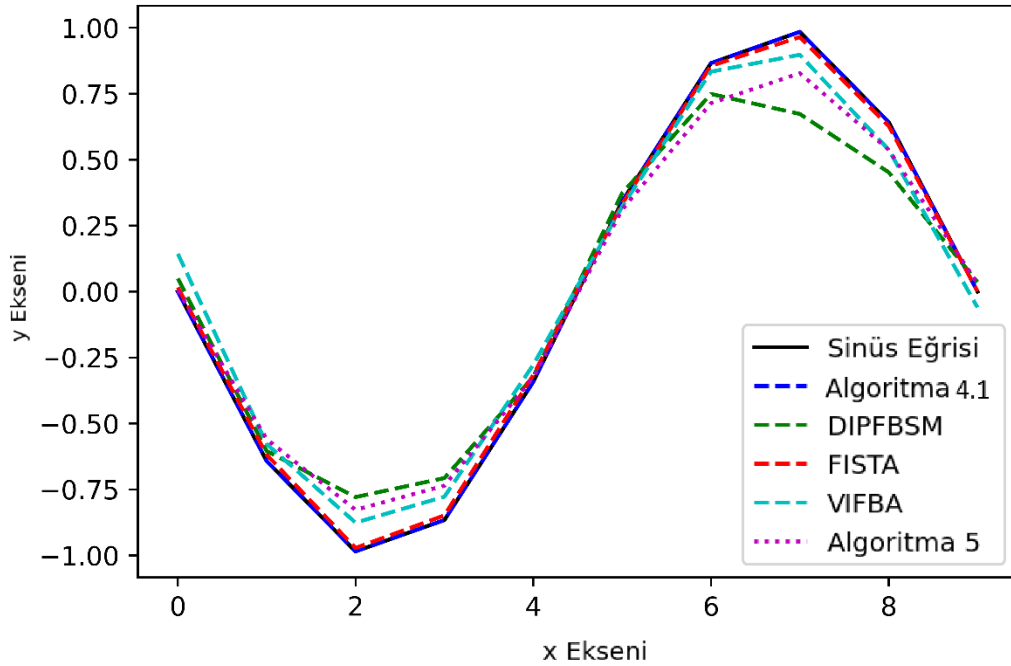
Bu başlık altında ele alınacak problemler

$$\min_{\theta} \|H\theta - Y\|_2^2 + \lambda \|\theta\|_1$$

formülasyonu ile konveks minimizasyon problemi şeklinde ifade edilebilen regresyon ve sınıflandırma problemleridir. Burada $\|\cdot\|_1$, l_1 normu, $\|\rho\|_1 = \sum_{i=1}^n |\rho_i|$ biçiminde tanımlanmakta olup, $\lambda > 0$ değeri düzenleme (regularization) parametresi olarak kullanılmaktadır. Bu problem, literatürde LASSO (Least Absolute Shrinkage and Selection Operator) olarak adlandırılmaktadır (Tibshirani 1996). Uygulamalarda λ düzenleme parametresi, modelin karmaşıklığının kontrol edilmesi ve aşırı öğrenmenin (overfitting) önlenmesi açısından kritik rol oynamaktadır.

Bu tez kapsamında yürütülen tüm sayısal uygulamalar ve analizler, karmaşık ve yoğun veri içermeyen veri setlerinde 13. Nesil Intel Core i9-13900H (2.60 GHz) işlemci, 16 GB RAM ve 64 bit işletim sistemi bulunan bir bilgisayarda; karmaşık ve yoğun veri içeren veri setlerinde ise Google Colab platformundan erişilebilen T4-GPU çalışma ortamında gerçekleştirilmiştir.

Bu kapsamda Algoritma 4.1 ile yapılan regresyon ve sınıflandırma çalışmalarında, gizli katman nöron sayısı 100 (sadece MNIST veri seti için 300), düzenleme parametresi olan $\lambda = 0.01$ ve iterasyon sayısı 1000 olarak alınmıştır. Algoritma 4.1 için $\vartheta = \frac{1}{4}$ ve $\delta = -\frac{1}{10}$ olarak seçilmiştir. İlk olarak Algoritma 4.1'in performansını görmek adına rastgele seçilen 10 sinüs değeri için lineer aktivasyon fonksiyonu kullanılarak elde edilen sonuçlar aşağıda verilmiştir.



Şekil 4.1. 10 sinüs örneği için lineer aktivasyonlu regresyon

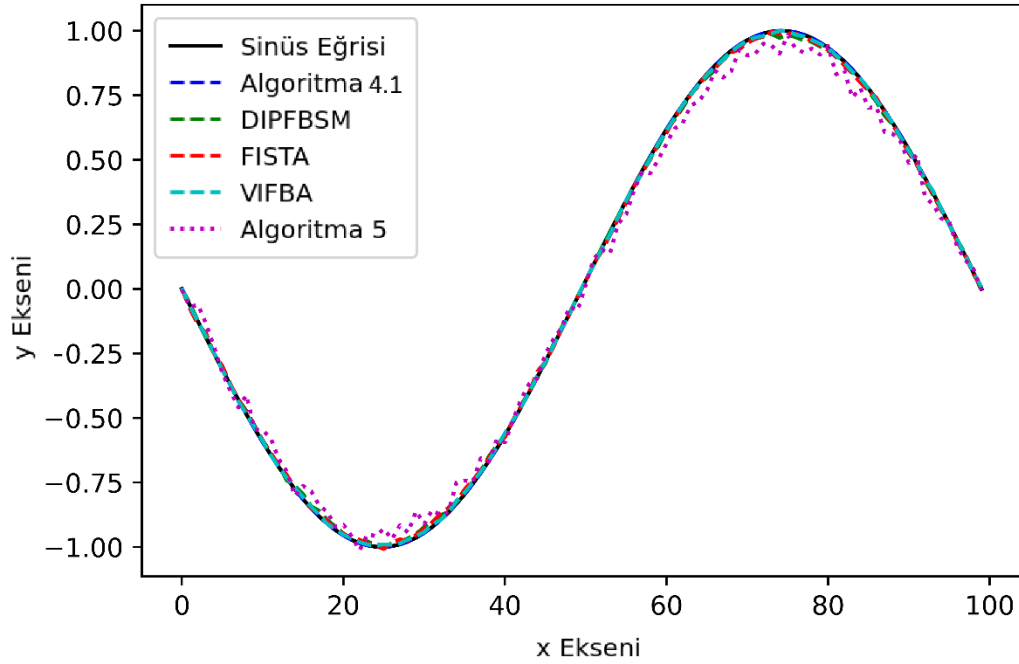
Şekil 4.1'de, Algoritma 4.1'in sinüs dalgasına en iyi yakınsama davranışını sergilediği görülmektedir. Hata test sonuçlarına göre başarı sıralaması ve süre açısından diğer algoritmalarla karşılaştırmanın verildiği Çizelge 4.1'de analiz sonuçları sunulmuştur. Çizelge 4.1'den görülebileceği üzere, Algoritma 4.1 tüm diğer algoritmalarından daha hızlı ve daha başarılı çalışmaktadır.

4. İKİ EYLEMSİZLİK TERİMLİ İLERİ-GERİ AYRIŞTIRMA ALGORİTMALARI

Çizelge 4.1. Şekil 4.1’deki modele göre algoritmaların performans metrikleri

Algoritmalar	İşlem Süresi	MSE	RMSE	R^2 Skoru	MAE
Algoritma 4.1	0,0061	5,37E-10	2,32E-5	0,9999	2,08E-5
FISTA	0,0175	2,78E-4	1,67E-2	0,9993	1,52E-2
VIFBA	0,0146	7,23E-3	8,50E-2	0,9839	7,75E-2
Algoritma 5	0,0314	1,10E-2	1,05E-1	0,9756	8,70E-2
DIPFBSM	0,0157	2,21E-2	1,49E-1	0,9509	1,16E-1

Bu sonuçlardan sonra ise yine lineer aktivasyon fonksiyonu kullanılarak 1000 iterasyon ve rastgele seçilen 100 sinüs değeri için benzetimler gerçekleştirilmiştir. Elde edilen sonuçlar Şekil 4.2’de sunulmuştur. Benzetimlerin istatistiksel sonuçları ise Çizelge 4.2’de verilmiştir.



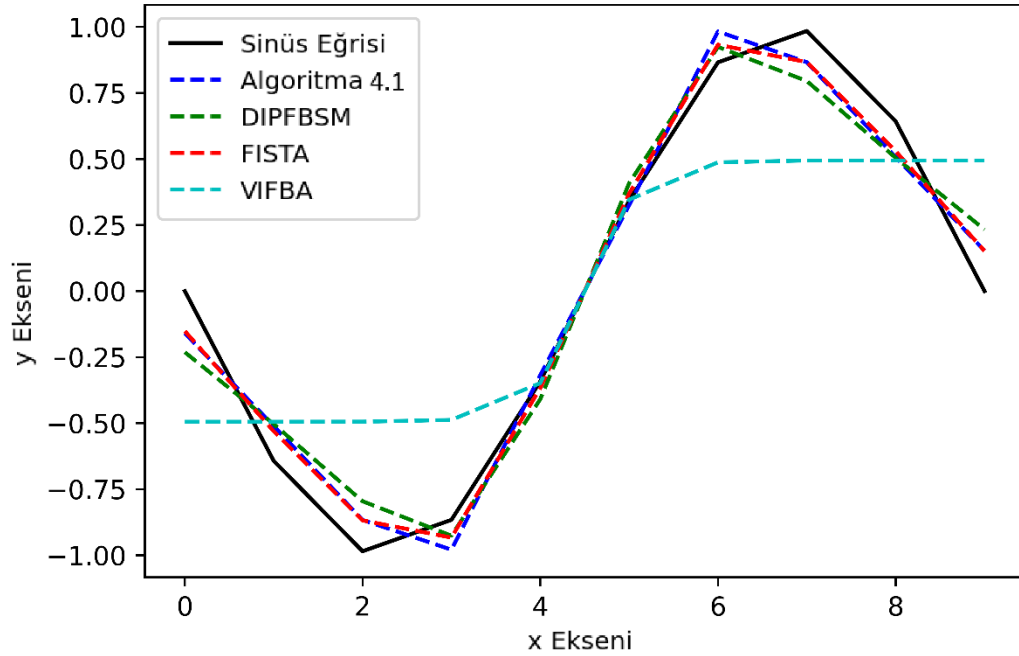
Şekil 4.2. 100 sinüs örneği için lineer aktivasyonlu regresyon

4. İKİ EYLEMSİZLİK TERİMLİ İLERİ-GERİ AYRIŞTIRMA ALGORİTMALARI

Çizelge 4.2. Şekil 4.2’deki modele göre algoritmaların performans metrikleri

Algoritmalar	İşlem Süresi	MSE	RMSE	R^2 Skoru	MAE
Algoritma 4.1	0,0114526	2,36E-10	1,54E-5	0,9999	1,31E-5
VIFBA	0,0312581	2,50E-5	5,00E-3	0,9999	4,17E-3
FISTA	0,0138959	9,83E-5	9,92E-3	0,9998	8,38E-3
Algoritma 5	0,0157158	9,92E-5	9,96E-3	0,9998	7,87E-3
DIPFBSM	0,0785269	2,30E-3	4,79E-2	0,9954	4,07E-2

Şekil 4.2 ve Çizelge 4.2’den de görülebileceği üzere, Algoritma 4.1 diğer algoritmalarla kıyasla hem daha yüksek başarı göstermiş hemde daha hızlı çalışmıştır. Bu başarılı sonuçların ardından, sinüs eğrisi için sigmoid aktivasyon fonksiyonu kullanılarak regresyon analizi gerçekleştirilmiştir. Bu karşılaştırma sonucunda, rastgele seçilen 10 sinüs değeri için 5000 yinleme ile elde edilen Şekil 4.3 aşağıda sunulmuştur.



Şekil 4.3. 10 sinüs örneği için sigmoid aktivasyonlu regresyon

Aşağıda verilen Çizelge 4.3’teki performans metriklerine Algoritma 4.1, FISTA dışında tüm diğer algoritmalarla kıyasla daha başarılı çalışmaktadır. Zaman değerlerine

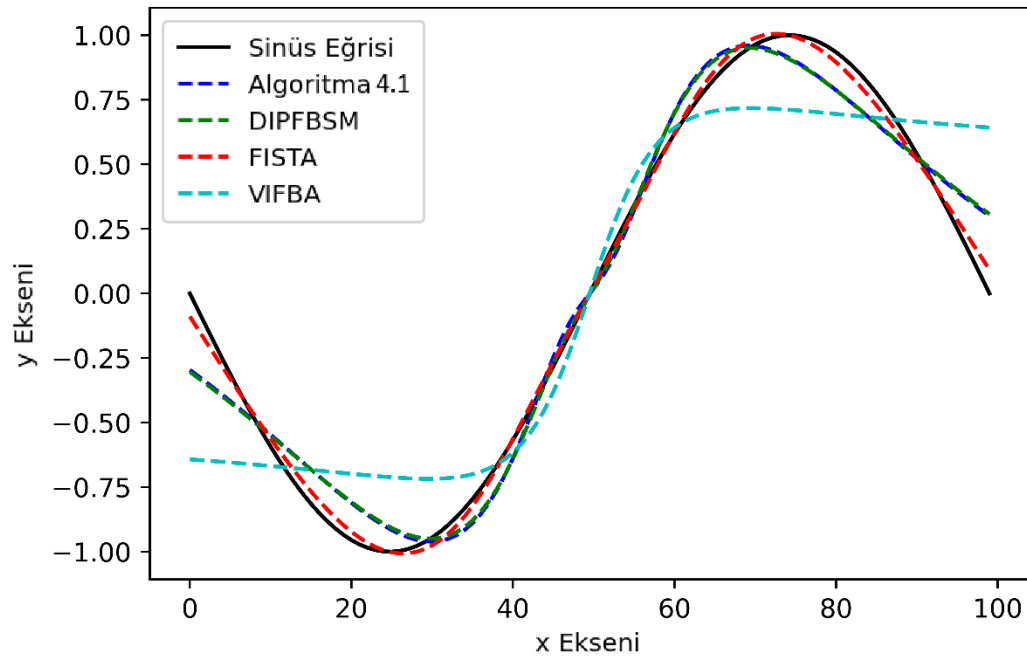
4. İKİ EYLEMSİZLİK TERİMLİ İLERİ-GERİ AYRIŞTIRMA ALGORİTMALARI

bakıldığında ise, farklar ihmal edilebilir düzeyde olsada Algoritma 4.1 en geç tepki süresine sahiptir.

Çizelge 4.3. Şekil 4.3'teki modele göre algoritmaların performans metrikleri

Algoritmalar	İşlem Süresi	MSE	RMSE	R^2 Skoru	MAE
FISTA	0,04846715	1,06E-2	1,03E-1	0,9763	9,34E-2
Algoritma 4.1	0,06319165	1,51E-2	1,23E-1	0,9664	1,11E-1
DIPFBSM	0,06293702	2,24E-2	1,50E-1	0,9503	1,35E-1
VIFBA	0,06139302	1,30E-1	3,61E-1	0,7110	3,03E-1

Son olarak sinüs eğrisi için elde edilen sonuçların ardından, sigmoid aktivasyon fonksiyonu kullanılarak 100 rastgele sinüs değeri için 5000 iterasyon sonucunda regresyon analizi gerçekleştirilmiştir. Elde edilen sonuçlar diğer algoritmalar ile karşılaştırılmış ve ilgili şekiller ile çizelgeler aşağıda sunulmuştur.



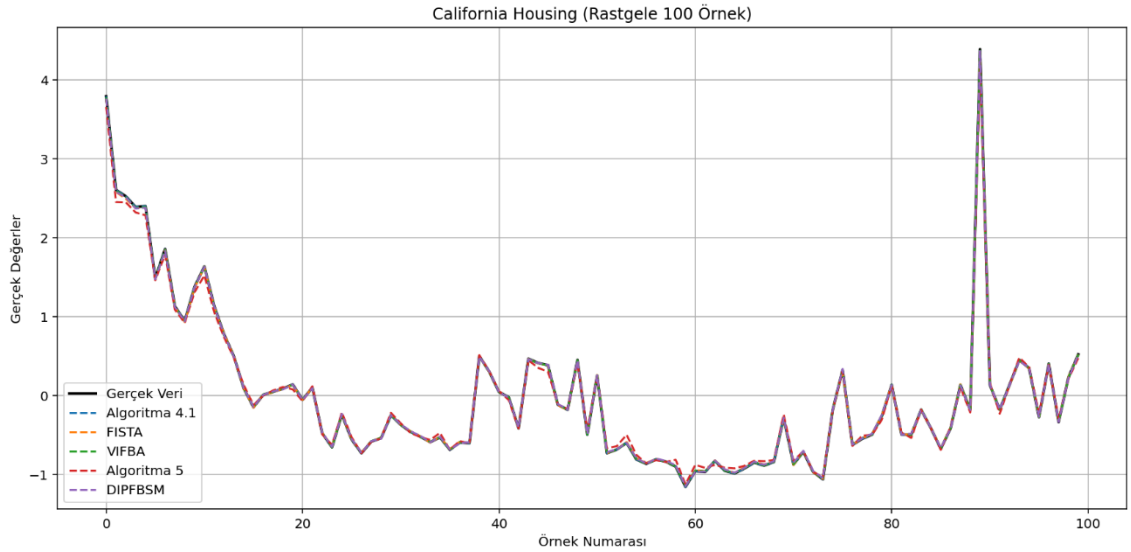
Şekil 4.4. 100 sinüs örneği için sigmoid aktivasyonlu regresyon

4. İKİ EYLEMSİZLİK TERİMLİ İLERİ-GERİ AYRIŞTIRMA ALGORİTMALARI

Çizelge 4.4. Şekil 4.4'teki modele göre algoritmaların performans metrikleri

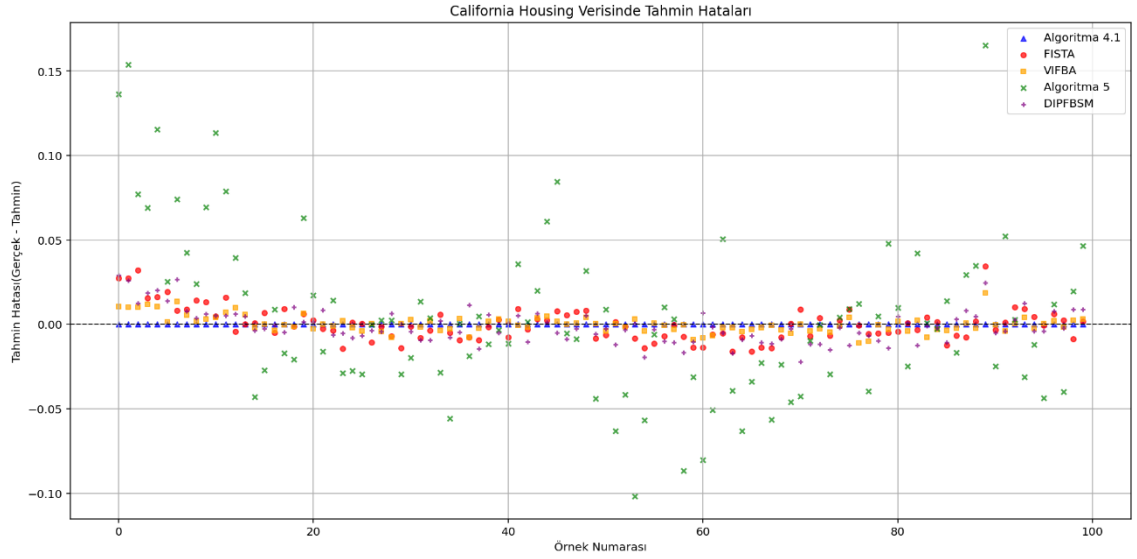
Algoritmalar	İşlem Süresi	MSE	RMSE	R^2 Skoru	MAE
FISTA	0,3938	1,04E-3	3,22E-2	0,9979	2,67E-2
Algoritma 4.1	0,4539	1,22E-2	1,11E-1	0,9753	8,99E-2
DIPFBSM	0,3770	1,28E-2	1,13E-1	0,9741	9,12E-2
VIFBA	0,3823	5,68E-2	2,38E-1	0,8852	1,89E-1

Ayrıca, algoritmamızın daha karmaşık regresyon görevlerini yerine getirmesi amacıyla, Scikit-learn kütüphanesi üzerinden elde edilen California Housing ve Diabetes benchmark veri setleri üzerinde de benzetimler gerçekleştirilmiştir. Basitten karmaşığa doğru yapılandırılan bu çalışmada, böylelikle algoritmamızın yüksek veri hacmine sahip veri setlerinde nasıl performans gösterdiği de gözlemlenmiştir. Rastgele seçilen 100 örnekten elde edilen grafikler ve çizelgeler aşağıda sunulmuştur.



Şekil 4.5. California Housing veri seti: 100 rastgele örnekte regresyon sonuçları

4. İKİ EYLEMSİZLİK TERİMLİ İLERİ-GERİ AYRIŞTIRMA ALGORİTMALARI



Şekil 4.6. California Housing veri seti: Algoritmalarla göre hata dağılımı

Elde edilen sonuçlarda, Algoritma 4.1'in diğer algoritmalarından daha başarılı sonuçlar verdiği görülmektedir. Özellikle, rastgele seçilen 100 örnek üzerinde yapılan tahmin hatası analizinde, Algoritma 4.1 tarafından üretilen hataların sıfır etrafında dar bir şekilde dağıldığı görülmektedir. Bu durumun tahmin doğruluğundaki başarıyı işaret ettiği tespit edilmiştir. Ayrıca tahmin edilen değerler ile gerçek hedef değerlerin karşılaştırıldığı Şekil 4.5'te Algoritma 4.1'in gerçek eğriye en yakın uyumu sağladığı gözlemlenmektedir.

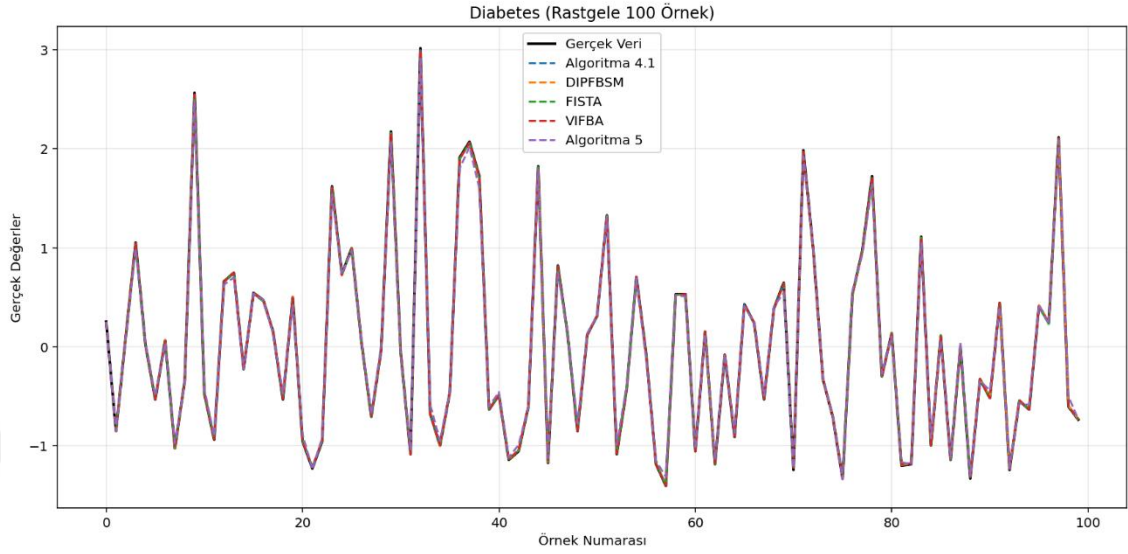
Çizelge 4.5. California Housing veri seti: 100 rastgele örnekte regresyon performansı

Algoritmalar	İşlem Süresi	MSE	RMSE	R^2 Skoru	MAE
Algoritma 4.1	0,0120	3,06E-10	1,75E-5	0,9999999997	1,24E-5
VIFBA	0,0090	2,50E-5	5,00E-3	0,999975	3,74E-3
FISTA	0,0090	9,88E-5	9,94E-3	0,999901235	7,06E-3
DIPFBSM	0,0087	9,92E-5	9,96E-3	0,999900829	7,66E-3
Algoritma 5	0,0434	2,23E-3	4,81E-2	0,997678434	3,60E-2

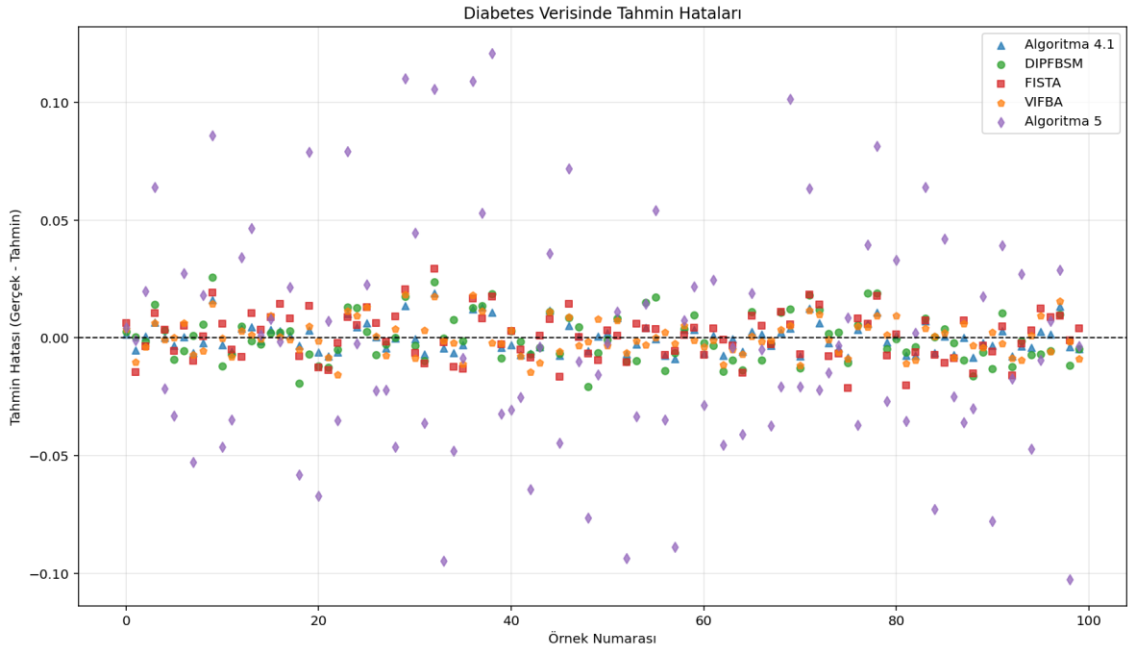
Çizelge 4.5'te de gösterildiği üzere Algoritma 4.1 karşılaştırılan algoritmalar arasında en başarılı sonuçları veren algoritmadır. $3,06 \times 10^{-10}$ gibi son derece düşük bir MSE değerine ve yüksek $R^2 = 0,9999999997$ değerine sahiptir. Ayrıca, RMSE ve MAE değerlerinin diğer algoritmalarla göre anlamlı derecede daha küçük olduğu gözlemlenmiştir. FISTA ve VIFBA algoritmalarına kıyasla hesaplama süresi biraz daha

4. İKİ EYLEMSİZLİK TERİMLİ İLERİ-GERİ AYRIŞTIRMA ALGORİTMALARI

yüksek olsada başarıdaki önemli artış nedeniyle bu farkın kabul edilebilir olduğu sonucuna varılmıştır.



Şekil 4.7. Diabetes veri seti: 100 rastgele örnekte regresyon sonuçları



Şekil 4.8. Diabetes veri seti: Algoritmalarla göre hata dağılımı

Şekiller incelendiğinde Algoritma 4.1'in Diabetes veri seti üzerinde oldukça başarılı sonuçlar verdiği görülmektedir. Hata dağılımı büyük ölçüde sıfırın çevresinde toplanmış bu da tahminlerin gerçek değerlere çok yakın olduğunu göstermiştir. Tahmin çıktılarının çoğu durumda gerçek değerlerle neredeyse bire bir örtüştüğü söylenebilir.

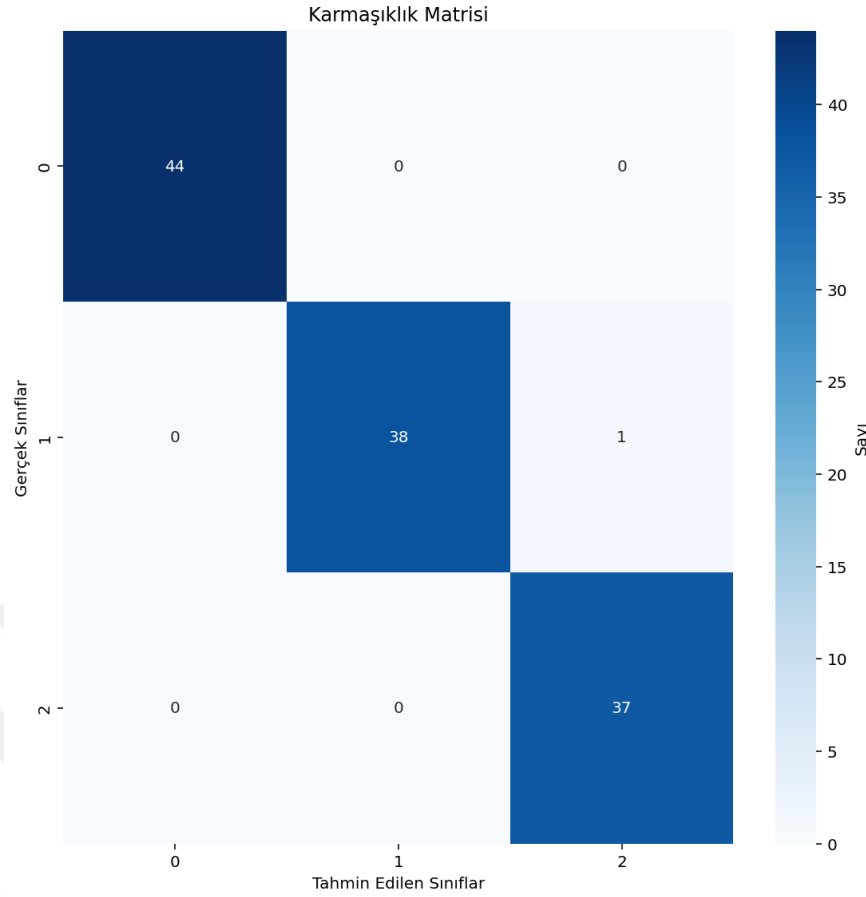
Diğer algoritmalar ile kıyaslandığında ise Algoritma 4.1'in hem dayanıklı hem de hassas olduğu açıkça ortaya çıkmaktadır.

Çizelge 4.6. Diabetes veri seti: 100 rastgele örnekte regresyon performansı

Algoritmalar	İşlem Süresi	MSE	RMSE	R^2 Skoru	MAE
Algoritma 4.1	0.028443	0.000040	0.006312	0.999960	0.005091
VIFBA	0.000185	0.000056	0.007500	0.999944	0.005976
FISTA	0.003174	0.000099	0.009949	0.999901	0.008301
DIPFBSM	0.000237	0.000100	0.010000	0.999900	0.008177
Algoritma 5	0.000279	0.002331	0.048278	0.997669	0.038269

Çizelge 4.6'da gösterildiği üzere, Algoritma 4.1'in diğer algoritmalarından başarılı olduğu görülmektedir. Diğer algoritmalarla göre düşük olan MSE, MAE ve RMSE değerine sahip olduğu açıktır. Yine diğer algoritmalarından yüksek R^2 değerine sahip olması hedef değişkendeki varyansın model tarafından iyi bir şekilde açıklandığını göstermektedir. Algoritma 4.1'in biraz daha fazla hesaplama süresi gerektirmesine rağmen, tahmin hassasiyetindeki önemli artış gösterdiği göz önüne alındığında bu artışın ihmal edilebilir olduğu sonucuna varılmaktadır.

Bu bölümde Algoritma 4.1 için elde edilen veri sınıflandırma sonuçları sunulacaktır. Tüm simülasyonlarda, her bir veri setinin %80'i eğitim, %20'si ise test için kullanılmıştır. Öncelikle Iris veri setini kullanarak elde ettiğimiz performans sonuçları sunulacaktır.



Şekil 4.9. Iris veri seti: Algoritma 4.1 için karmaşıklık matrisi

Çizelge 4.7. Iris veri seti: Algoritma 4.1 sınıflandırma metrikleri

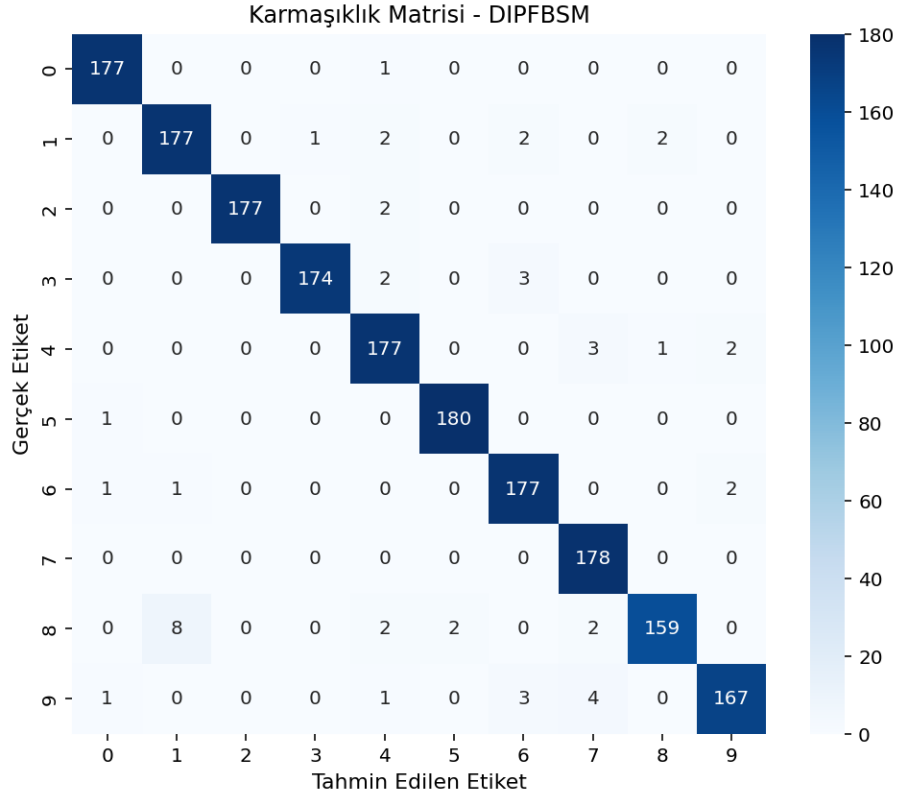
Sınıf	Kesinlik (Precision)	Duyarlılık (Recall)	F1 Skoru
0 (Setosa)	1,00	1,00	1,00
1 (Versicolor)	1,00	0,97	0,99
2 (Virginica)	0,97	1,00	0,99
Doğruluk (Accuracy)		0,99	

Çizelge ve karmaşıklık matrisi birlikte incelendiğinde, modelin sınıflandırma performansının yüksek olduğu görülmektedir. Sonuç olarak, modelin sınıflar arası ayrımı gerçekleştirdiği ve özellikle Setosa sınıfında mükemmel bir performans sergilediği gözlemlenmektedir. Versicolor ve Virginica sınıflarında ise küçük sınıflandırma hatalarına rağmen genel performansın yüksek düzeyde olduğu anlaşılmaktadır.

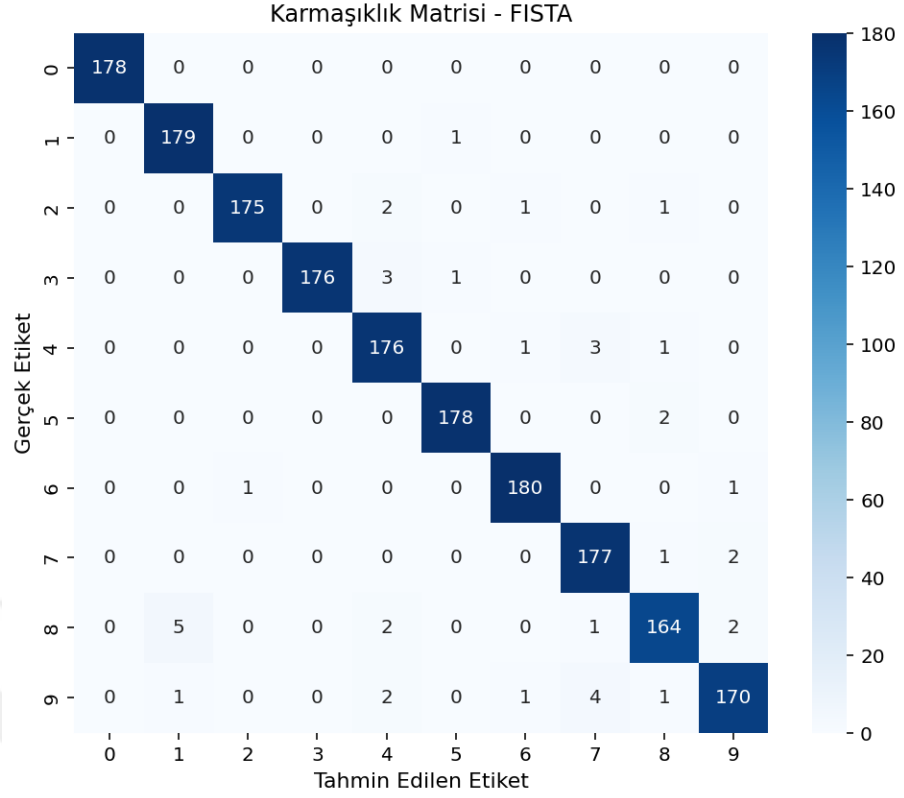
Sınıflandırma problemi için daha yüksek veri hacmine ve sınıf yoğunluğuna sahip olan Digits (Katman yapısı: 100 gizli nöron – 10 çıkış nöronu, durdurma kriteri: 1000

4. İKİ EYLEMSİZLİK TERİMLİ İLERİ-GERİ AYRIŞTIRMA ALGORİTMALARI

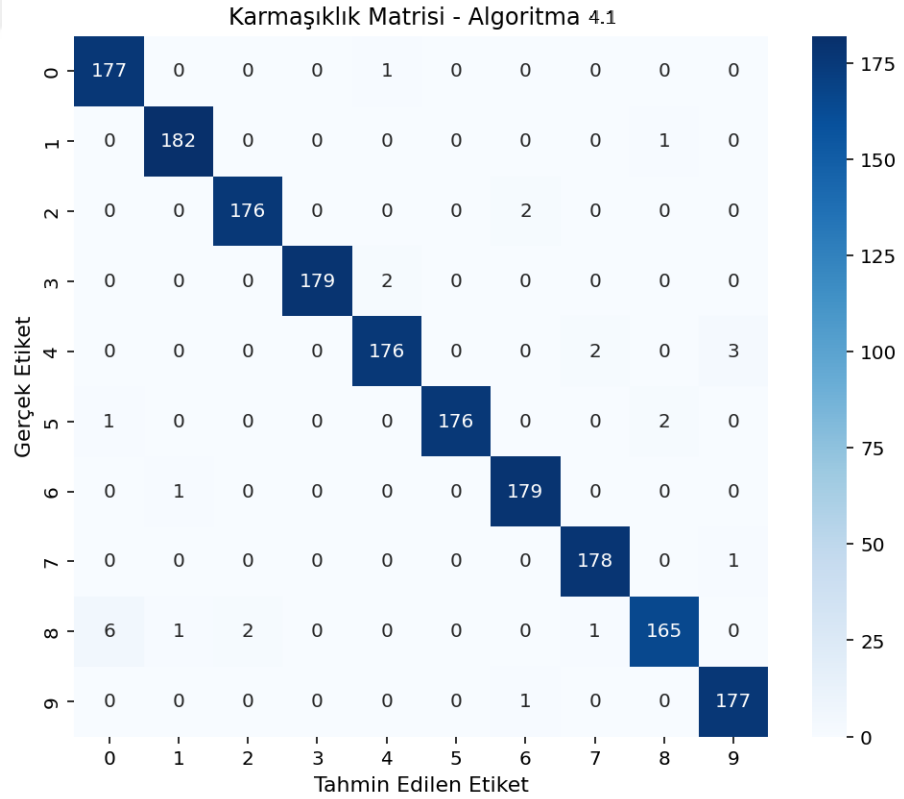
iterasyon) ve MNIST (Katman yapısı: 300 gizli nöron – 10 çıkış nöronu, durdurma kriteri: 2000 iterasyon Doğrulama: 3-fold cross-validation) veri setleri de kullanılmıştır. Bu veri setleri sırasıyla Scikit-learn ve TensorFlow/Keras kütüphaneleri aracılığıyla elde edilmiştir. Aşağıda öncelikle Digits veri seti için DIPFBSM, FISTA ve Algoritma 4.1'in karmaşıklık matrisleri ve performans metrikleri verilmiştir.



Şekil 4.10. Digits veri seti: DIPFBSM algoritması için karmaşıklık matrisi



Şekil 4.11. Digits veri seti: FISTA algoritması için karmaşıklık matrisi



Şekil 4.12. Digits veri seti: Algoritma 4.1 için karmaşıklık matrisi

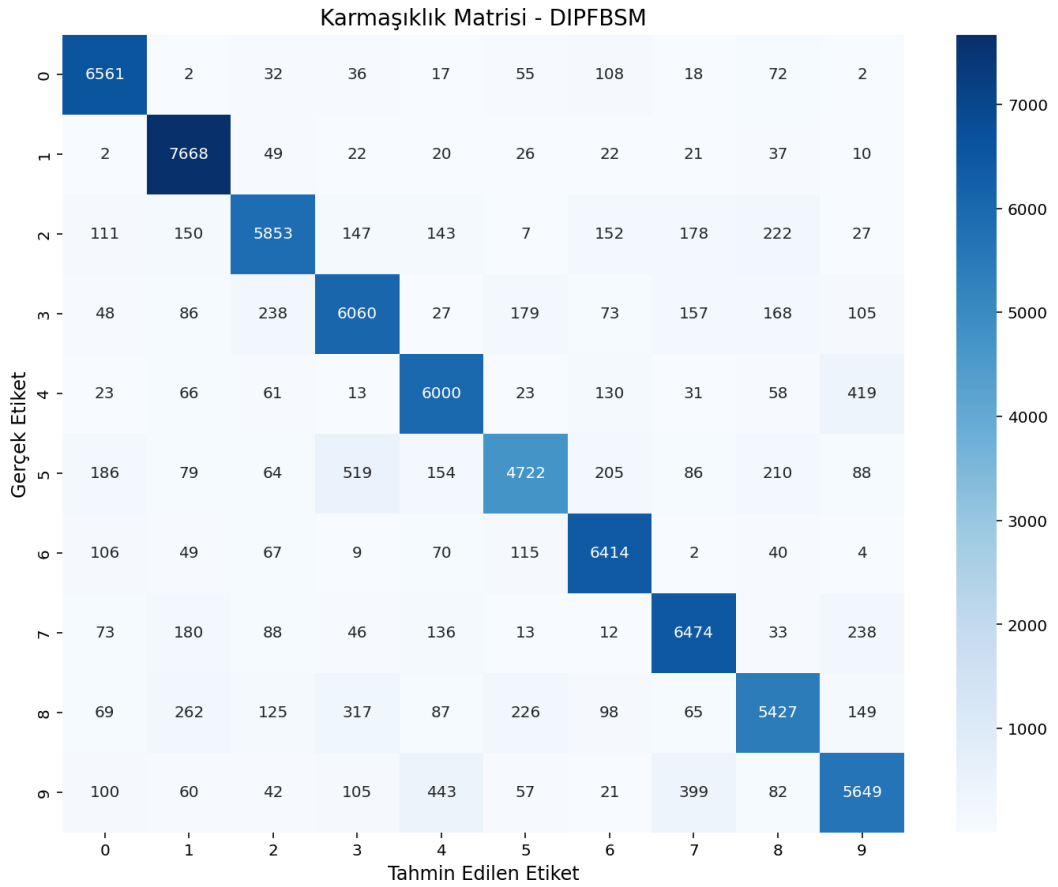
4. İKİ EYLEMSİZLİK TERİMLİ İLERİ-GERİ AYRIŞTIRMA ALGORİTMALARI

Çizelge 4.8. Digits veri seti: Algoritmaların sınıflandırma metrikleri

Algoritma	Doğruluk	Kesinlik	Duyarlılık	F1 Skoru
Algoritma 4.1	0,9833	0,98	0,98	0,98
FISTA	0,9755	0,98	0,98	0,98
DIPFBSM	0,9700	0,97	0,97	0,97

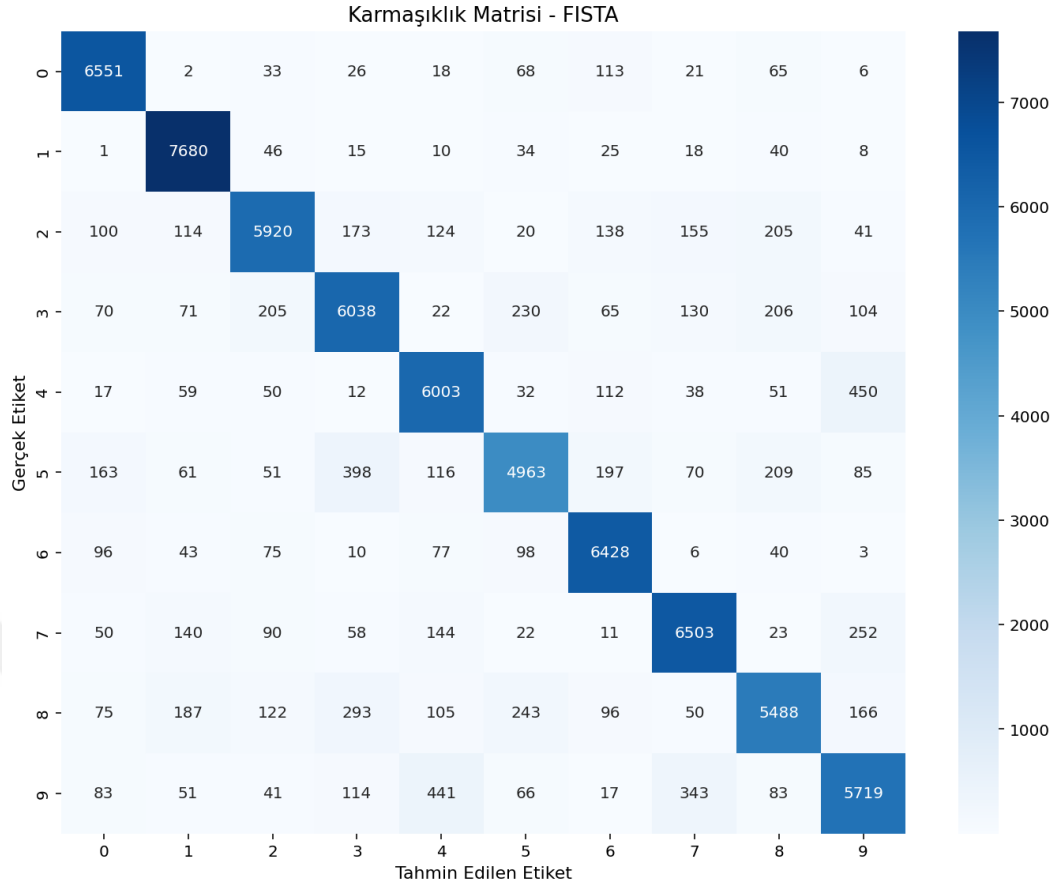
Çizelge 4.8'de sunulduğu üzere Algoritma 4.1'in diğer yöntemlere kıyasla daha yüksek kesinlik, duyarlılık ve *F1* skoru ile diğer algoritmalarından başarılı sınıflandırma doğruluğuna (%98,33) ulaştığı görülmektedir. Bu durum algoritmanın çok sınıflı sınıflandırma görevlerindeki etkinliğini ve sağlamlığını açıkça göstermektedir.

Şimdi ise MNIST veri seti için aynı algoritmaların karmaşıklık matrisleri ve performans metrikleri verilmiştir.



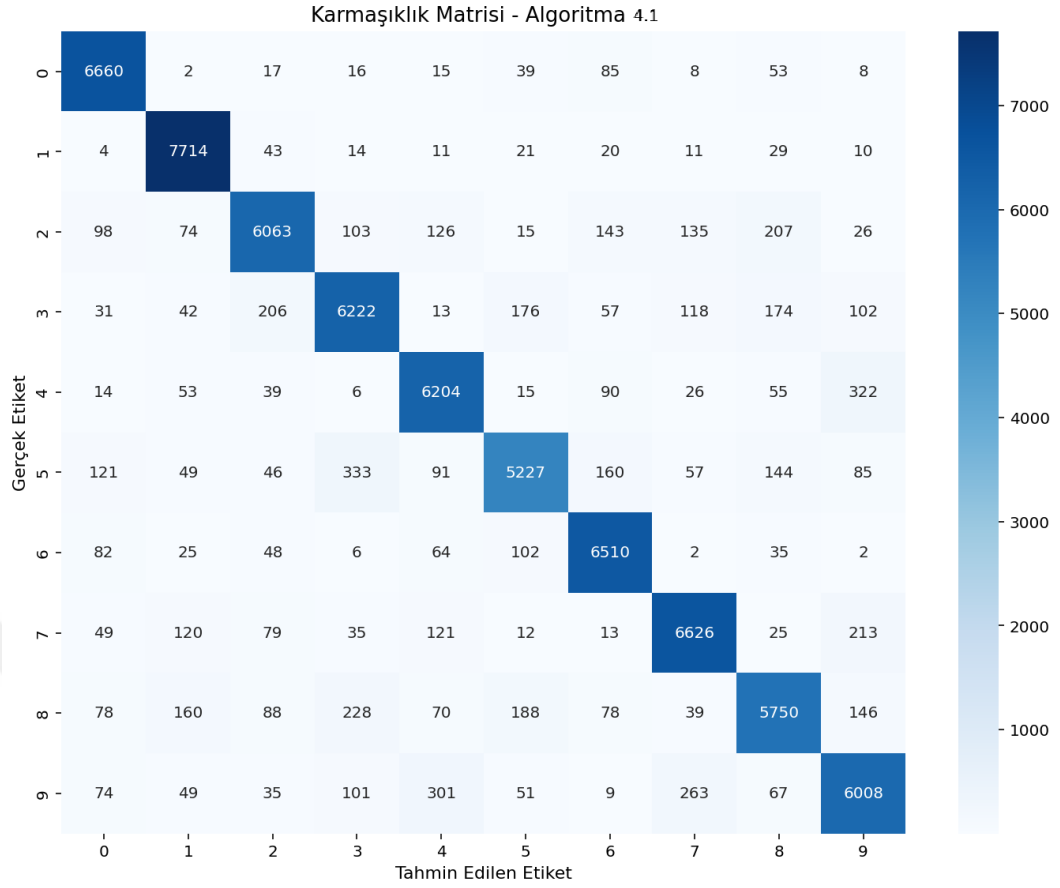
Şekil 4.13. MNIST veri seti: DIPFBSM algoritması için karmaşıklık matrisi

4. İKİ EYLEMSİZLİK TERİMLİ İLERİ-GERİ AYRIŞTIRMA ALGORİTMALARI



Şekil 4.14. MNIST veri seti: FISTA algoritması için karmaşıklık matrisi

4. İKİ EYLEMSİZLİK TERİMLİ İLERİ-GERİ AYRIŞTIRMA ALGORİTMALARI



Şekil 4.15. MNIST veri seti: Algoritma 4.1 için karmaşıklık matrisi

Karmaşıklık matrisinde köşegen boyunca çoğu rakam sınıfı genelinde diğer algoritmalarla göre dikkate değer ölçüde daha az yanlış sınıflandırma yapıldığı ortaya konulmuştur. MNIST veri setinin yüksek sınıf içi değişkenlik, rakam benzerlikleri ve gerçek dünya el yazısı olmasından dolayı karmaşık bir veri seti olduğu bilinmektedir. Bu zorluklara rağmen diğer algoritmalarla göre Algoritma 4.1'in başarılı sınıf ayrımı yaptığı görülmektedir. Özellikle 5, 8 ve 9 gibi ayrılması zor sınıflarda dahi algoritmaların her ikisinden de daha iyi performans gösterdiği tespit edilmiştir.

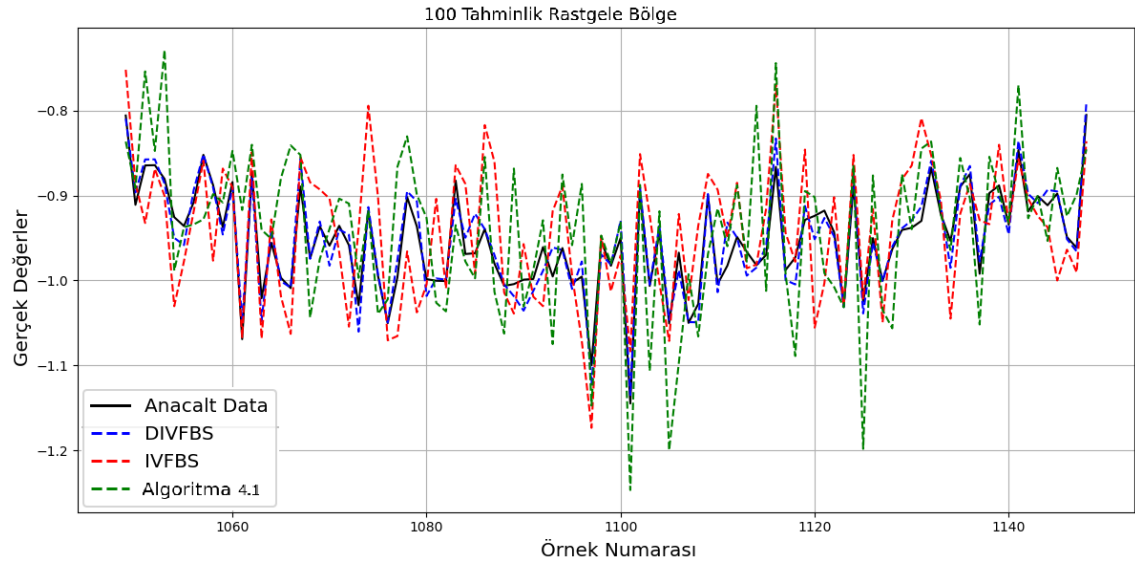
Çizelge 4.9. MNIST veri seti: Algoritmaların sınıflandırma metrikleri

Algoritma	Doğruluk	Kesinlik	Duyarlılık	F1 Skoru
Algoritma 4.1	0,8998	0,90	0,90	0,90
FISTA	0,8756	0,87	0,87	0,87
DIPFBSM	0,8690	0,87	0,87	0,87

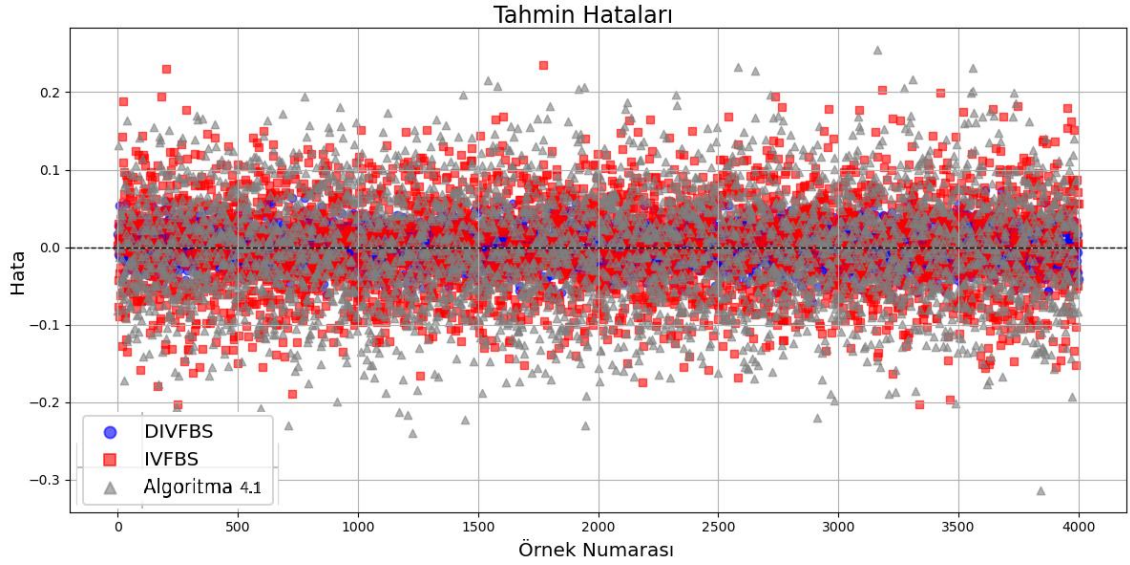
4. İKİ EYLEMSİZLİK TERİMLİ İLERİ-GERİ AYRIŞTIRMA ALGORİTMALARI

Yine Çizelge 4.9'da gösterildiği üzere Algoritma 4.1'in tüm yöntemler arasında en yüksek sınıflandırma doğruluğuna (%89,98) ulaştığı ve en iyi kesinlik, duyarlılık ve $F1$ skoru değerlerini elde ettiği görülmektedir.

Bu aşamada ise DIVFBS ismini verdiğimiz algoritmamızın regresyon ve sınıflandırma problemlerine yönelik benzetimleri sunulmuştur. Tanımladığımız bu algoritma yine tanımladığımız bir diğer algoritma olan Algoritma 4.1 ve IVFBS algoritmaları ile karşılaştırılacaktır. Öncelikle Anacalt veri seti için düzenlileştirme parametresi $\lambda = 0.01$, iterasyon sayısı 1000, gizli nöron sayısı 400, $\{\theta_k\} = \{0.99\}$, $\{\delta_k\} = \{-1/200\}$, $\{\alpha_k\} = \{0.1\}$ ve $h(x) = 0.89x$ olarak seçilmiştir. Bu seçimlere karşılık gelen sonuçlar aşağıda sunulmuştur.



Şekil 4.16. Anacalt veri seti: 100 rastgele örnekte regresyon sonuçları



Şekil 4.17. Anacalt veri seti: Algoritmalarla tahmin hatalarının dağılımı

Şekil 4.16'dan görüldüğü gibi DIVFBS algoritması tarafından üretilen tahminlerin gerçek veri setini yakından takip ettiği tespit edilmiştir. Yine Şekil 4.17'den görülebileceği üzere DIVFBS algoritmasına ait tahmin hatalarının sıfır etrafında yoğunlaştığı, diğer algoritmalarla tahmin hatalarının ise geniş bir alana yayıldığı görülmektedir.

Çizelge 4.10. Anacalt veri seti: 100 rastgele örnekte regresyon performansı

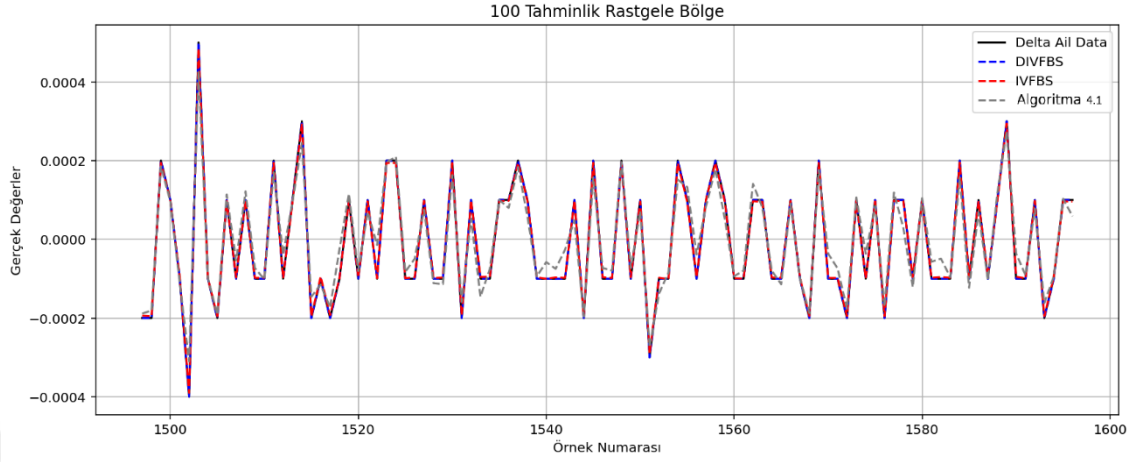
Algoritmalar	İşlem Süresi	MSE	RMSE	R ² Skoru	MAE
DIVFBS	0,3225	0,020130	0,141879	0,933318	0,079130
Algoritma 4.1	0,3410	0,022424	0,149747	0,925717	0,077531
IVFBS	0,3319	0,025342	0,159193	0,916050	0,103190

En düşük MSE, RMSE ve MAE değeri ile en yüksek R² değeri DIVFBS algoritması ile elde edilmiştir. Böylece seçilen algoritmalar arasında DIVFBS başarılı bir performans sergilemiştir. Son olarak işlem süresi bakımından en hızlı ikinci algoritma olduğu saptanmıştır.

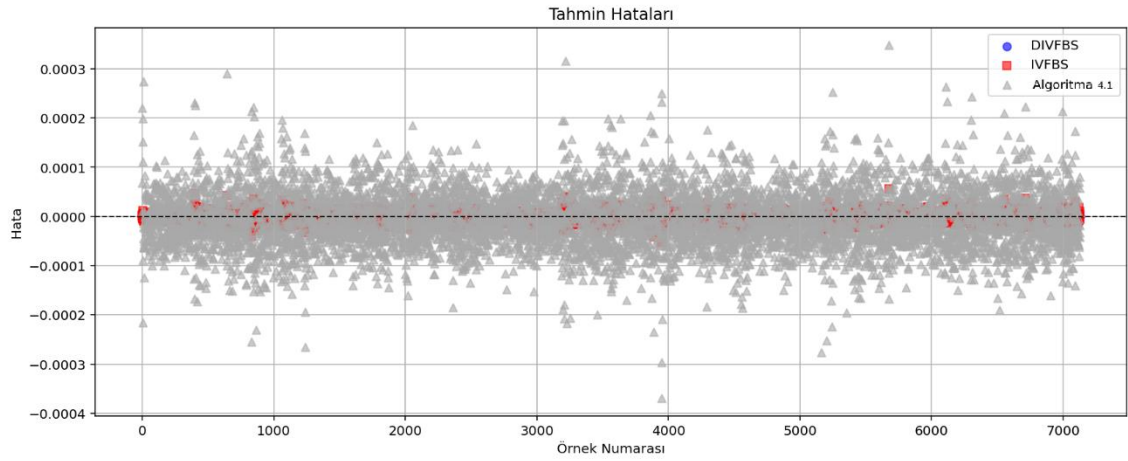
Şimdi ise Delta Ail veri seti için düzenleme parametresi $\lambda = 1e - 6$, iterasyon sayısı 1000, gizli nöron sayısı 400 olarak belirlenmiştir. Başlangıç parametreleri $\theta_1 = 0.99$ ve $\delta_1 = -1/200$ olarak tanımlanırken, diğer parametreler

4. İKİ EYLEMSİZLİK TERİMLİ İLERİ-GERİ AYRIŞTIRMA ALGORİTMALARI

$\alpha_k = 0.01$ ve $h = 0.89$ olarak seçilmiştir. Bu seçimlerine karşılık gelen sonuçlar aşağıda sunulmuştur.



Şekil 4.18. Delta Ail veri seti: 100 rastgele örnekte regresyon sonuçları



Şekil 4.19. Delta Ail veri seti: Algoritmalarla göre hata dağılımı

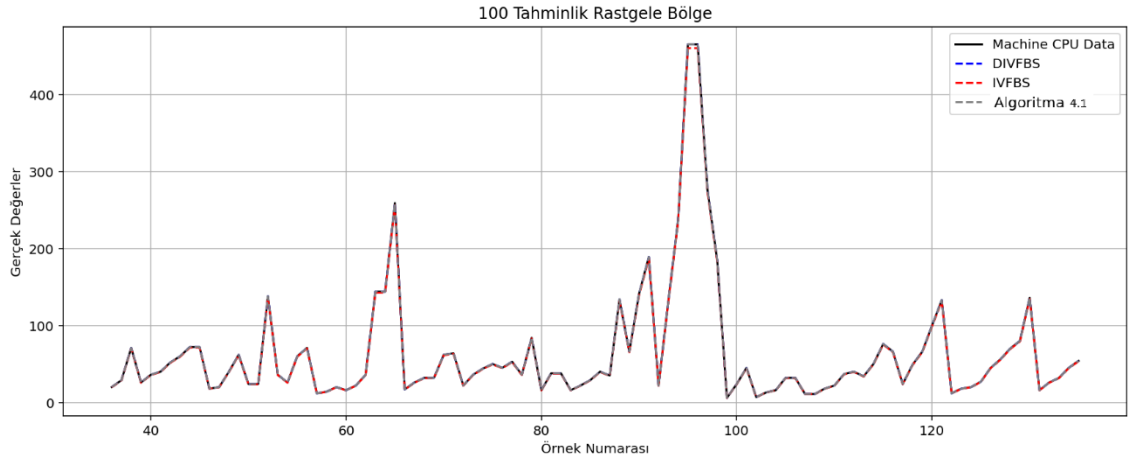
Şekil 4.18’de DIVFBS algoritması tarafından üretilen tahminlerin, rastgele seçilen 100 örneklilik bölgede Delta Ail veri setini yakından takip ettiği görülmektedir. Şekil 4.19’da ise DIVFBS algoritmasına ait hataların sıfır etrafında olduğu görülmektedir. Diğer algoritmalar ise daha dağınık bir tutum sergilemiştir.

Çizelge 4.11. Delta Ail veri seti: 100 rastgele örnekte regresyon performansı

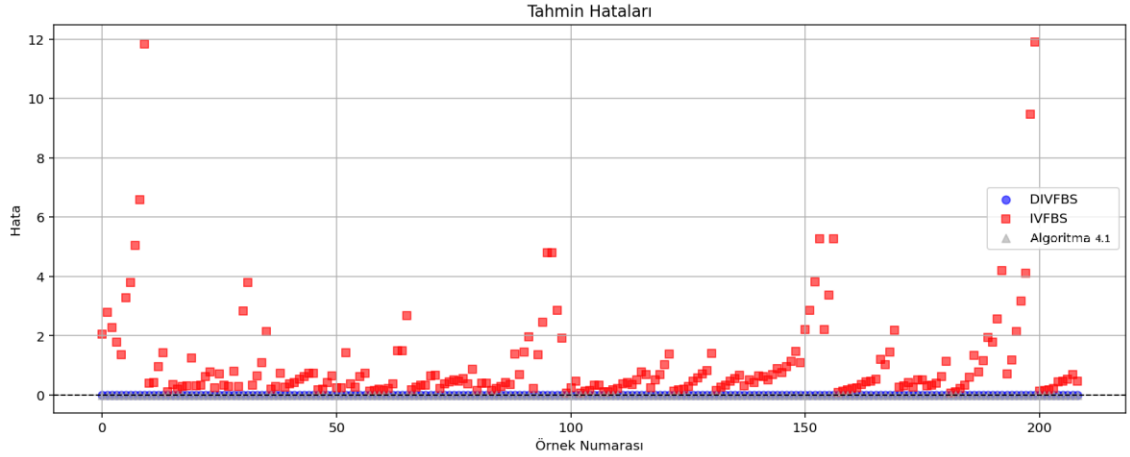
Algoritmalar	İşlem Süresi	MSE	RMSE	R ² Skoru	MAE
DIVFBS	2.362689	6.09E-14	2.47E-07	0.999999	1.89E-07
IVFBS	3.576176	6.25E-11	7.91E-06	0.999318	6.12E-06
Algoritma 4.1	0.762797	3.18E-09	5.64E-05	0.965276	4.31E-05

Çizelge 4.11’de DIVFBS algoritması diğer algoritmalarla göre düşük MSE, RMSE ve MAE değerlerine ve yüksek R^2 skoruna ulaşmıştır. Böylece bu veri seti ile en başarılı sonuçları DIVFBS algoritması elde etmektedir.

Şimdi ise Machine CPU veri seti için düzenleme parametresi $\lambda = 1e - 6$, iterasyon sayısı 1000, gizli nöron sayısı 400 olarak belirlenmiştir. Başlangıç parametreleri $\theta_1 = 0.99$ ve $\delta_1 = -1/2000$ olarak tanımlanırken, diğer parametreler $\alpha_k = 0.001$ ve $h = 0.89$ olarak seçilmiştir. Bu seçimlerine karşılık gelen sonuçlar aşağıda sunulmuştur.

**Şekil 4.20.** Machine CPU veri seti: 100 rastgele örnekte regresyon sonuçları

4. İKİ EYLEMSİZLİK TERİMLİ İLERİ-GERİ AYRIŞTIRMA ALGORİTMALARI



Şekil 4.21. Machine CPU veri seti: Algoritmalarla ilgili hata dağılımı

Şekil 4.20’de görüleceği üzere DIVFBS algoritmasının Machine CPU veri setine ait rastgele noktaları minimum sapmayla yakından takip ettiği görülmektedir. Şekil 4.21’de verilen hata dağılım grafiğinde ise DIVFBS algoritmasına ait hataların sıfır ekseninde toplandığı, diğer algoritmaların ise aykırı değerler sergilediği görülmektedir. Bu durum ise diğer algoritmaların karmaşık değişimler altında daha az kararlı olduklarını göstermektedir.

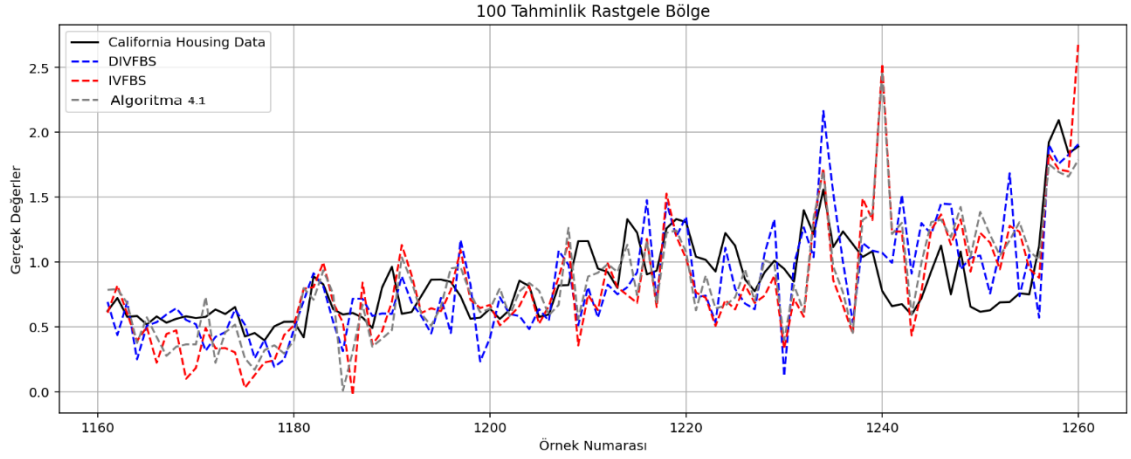
Çizelge 4.12. Machine CPU veri seti: 100 rastgele örnekte regresyon performansı

Algoritmalar	İşlem Süresi	MSE	RMSE	R ² Skoru	MAE
DIVFBS	0.292945	0.000000	0.000468	1.000000	0.000257
IVFBS	0.276610	3.947349	1.986794	0.999847	1.092450
Algoritma 4.1	0.066224	0.000000	0.000593	1.000000	0.000414

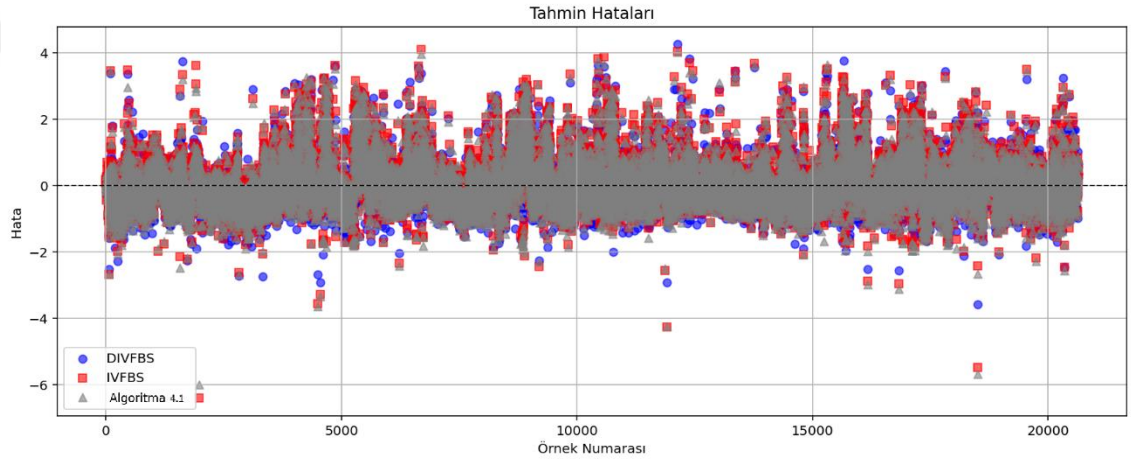
Çizelge 4.12’deki performans metrikleri incelendiğinde DIVFBS algoritmasının diğer algoritmalarla göre başarılı olduğu görülmektedir.

Şimdi ise California Housing veri seti için düzenleme parametresi $\lambda = 1e - 6$, iterasyon sayısı 1000, gizli nöron sayısı 1000 olarak belirlenmiştir. Başlangıç parametreleri $\theta_1 = 0.99$ ve $\delta_1 = -1/20000$ olarak tanımlanırken, diğer parametreler $\alpha_k = 0.0001$ ve $h = 0.99$ olarak seçilmiştir. Bu seçimlerine karşılık gelen sonuçlar aşağıda sunulmuştur.

4. İKİ EYLEMSİZLİK TERİMLİ İLERİ-GERİ AYRIŞTIRMA ALGORİTMALARI



Şekil 4.22. California Housing veri seti: 100 rastgele örnekte regresyon sonuçları



Şekil 4.23. California Housing veri seti: Algoritmalarla göre hata dağılımı

California Housing veri seti karmaşık yapısı, yüksek örnek sayısı ve hedef değerlerdeki büyük varyans nedeniyle regresyon görevleri için zor bir veri setidir. Bu veri seti ile yapılan regresyon benzetiminde DIVFBS algoritmasının diğer yöntemlere kıyasla daha kararlı ve tutarlı tahmin hataları sergilediği görülmüştür. Şekil 4.23'te DIVFBS algoritmasına ait hataların sıkı bir şekilde kümелendiği gözlemlenmiştir. Buna karşılık, IVFBS algoritması ve Algoritma 4.1 yöntemlerinin daha geniş ve dağınık hata dağılımları sergilediği saptanmıştır.

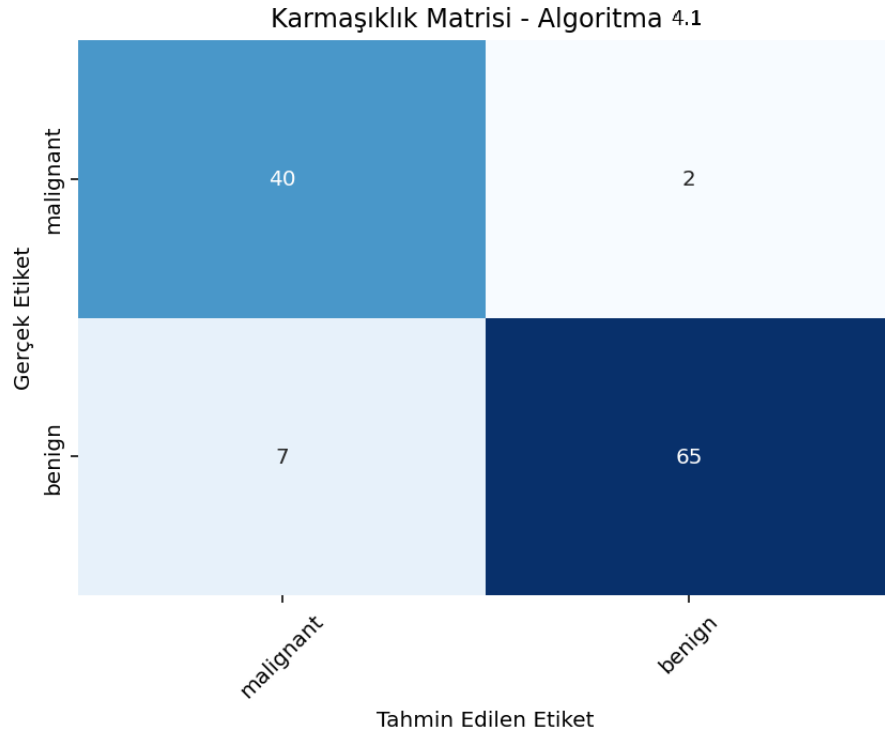
Ayrıca, rastgele seçilen 100 tahminlik bir bölgede, DIVFBS algoritması tarafından gerçek değerlere en yakın tahminlerin üretildiği belirlenmiştir. Özellikle hızlı dalgalanmaların olduğu bölgelerde, DIVFBS'nin gerçek eğilimi daha iyi takip ettiği, diğer yöntemlerin ise daha fazla sapma gösterdiği görülmüştür.

Çizelge 4.13. California Housing veri seti: 100 rastgele örnekte regresyon performansı

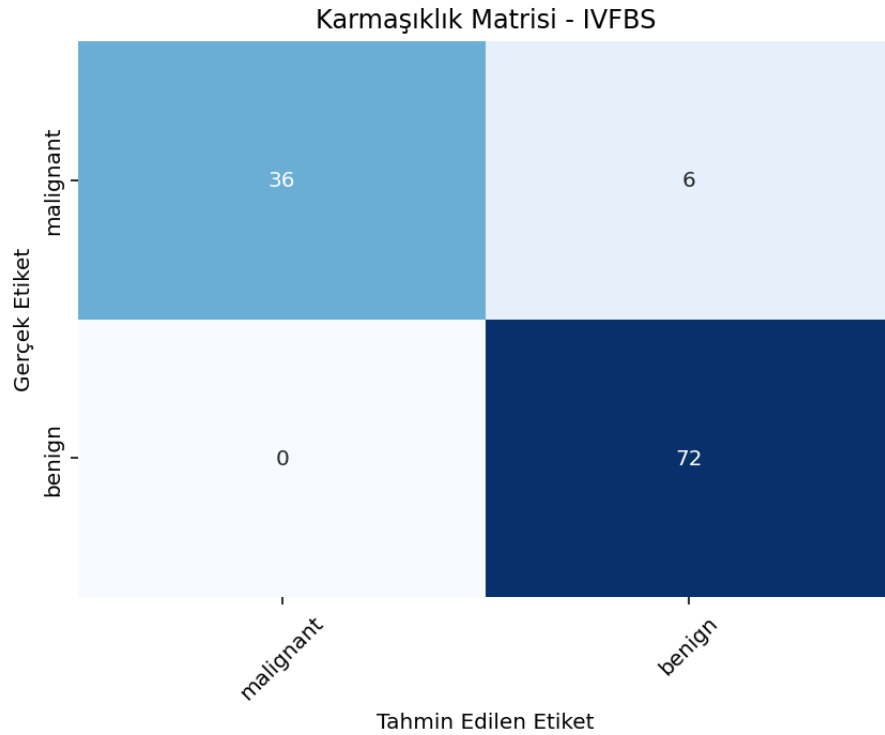
Algoritmalar	İşlem Süresi	MSE	RMSE	R ² Skoru	MAE
DIVFBS	0.939424	0.462944	0.680400	0.652327	0.501876
IVFBS	72.091820	0.502667	0.708990	0.622495	0.514297
Algoritma 4.1	74.283803	0.498119	0.705776	0.625910	0.521041

Tüm algoritmalar arasında en düşük MSE, RMSE ve MAE değeri ile birlikte en yüksek R^2 skoruna sahip olan DIVFBS algoritması bu metrikler ile diğer algoritmalara göre başarılı olmuştur. Ayrıca işlem süresi olarakta DIVFBS algoritması diğer algoritmalara göre önemli ölçüde daha kısa sürede süreci tamamlanmıştır. IVFBS algoritması ve Algoritma 4.1 ise daha yüksek tahmin hatası ve uzun hesaplama süresine sahip olmuştur.

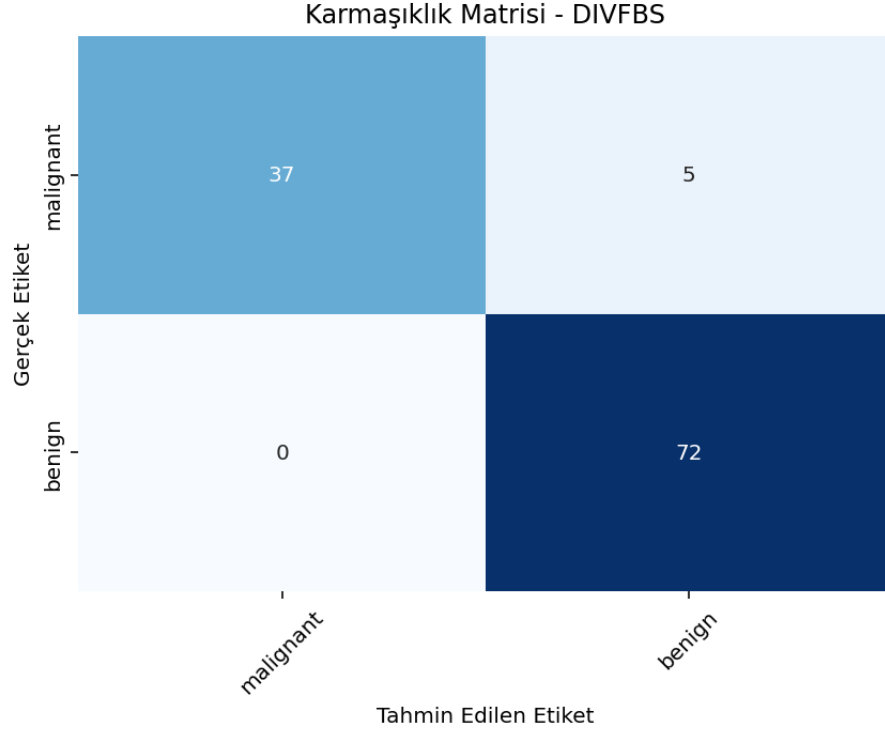
Veri sınıflandırma sürecinde algoritmaların performansını değerlendirmek amacıyla karmaşıklık matrisi kullanılmıştır. Bu metrikleri açıkça sunmak için, her bir algoritmanın performans sonuçları çizelgelerde özetlenmiştir. Tüm simülasyonlarda, her bir veri setinin %80'i eğitim, %20'si ise test için kullanılmıştır. BCWD veri seti veri seti için düzenleme parametresi $\lambda = 0.0005$, gizli nöron sayısı 1000 olarak belirlenmiştir. Başlangıç parametreleri $\theta_1 = 0.00001$ ve $\delta_1 = -1/10000$ olarak tanımlanırken, diğer parametreler $\alpha_k = 0.1$ ve $F = 0.99$ olarak seçilmiştir. Bu seçimlerine karşılık gelen sonuçlar aşağıda sunulmuştur.



Şekil 4.24. BCWD veri seti: Algoritma 4.1 için karmaşıklık matrisi



Şekil 4.25. BCWD veri seti: IVFBS algoritması için karmaşıklık matrisi



Şekil 4.26. BCWD veri seti: DIVFBS algoritması için karmaşıklık matrisi

Karmaşıklık matrislerinde görüldüğü üzere tüm algoritmalar içerisinde daha yüksek sınıflandırma başarımı DIVFBS algoritması ile elde edilmiştir.

Çizelge 4.14. BCWD veri seti: Algoritmaların sınıflandırma metrikleri

Algoritma	Doğruluk	Kesinlik	Duyarlılık	F1 Skoru
DIVFBS	0.96	0.97	0.94	0.95
IVFBS	0.95	0.96	0.93	0.94
Algoritma 4.1	0.92	0.91	0.93	0.92

BCWD veri seti üzerinde yapılan benzetim sonucunda, algoritmalar arasında performans açısından belirgin farklılıklar gözlemlenmektedir. DIVFBS algoritması doğruluk, kesinlik, duyarlılık ve F1 skoru bakımından daha yüksek bir performans sergileyerek öne çıkmıştır. IVFBS algoritması ise DIVFBS algoritmasına oldukça yakın sonuçlar elde ederken özellikle kesinlik açısından güçlü bir performans ortaya koymuştur. Ama genel başarı ölçütü olan F1 skorunda DIVFBS algoritmasından bir miktar geride kalmıştır. Algoritma 4.1 ise diğer tüm algoritmalara göre daha düşük bir performans göstermiştir.

4. İKİ EYLEMSİZLİK TERİMLİ İLERİ-GERİ AYRIŞTIRMA ALGORİTMALARI

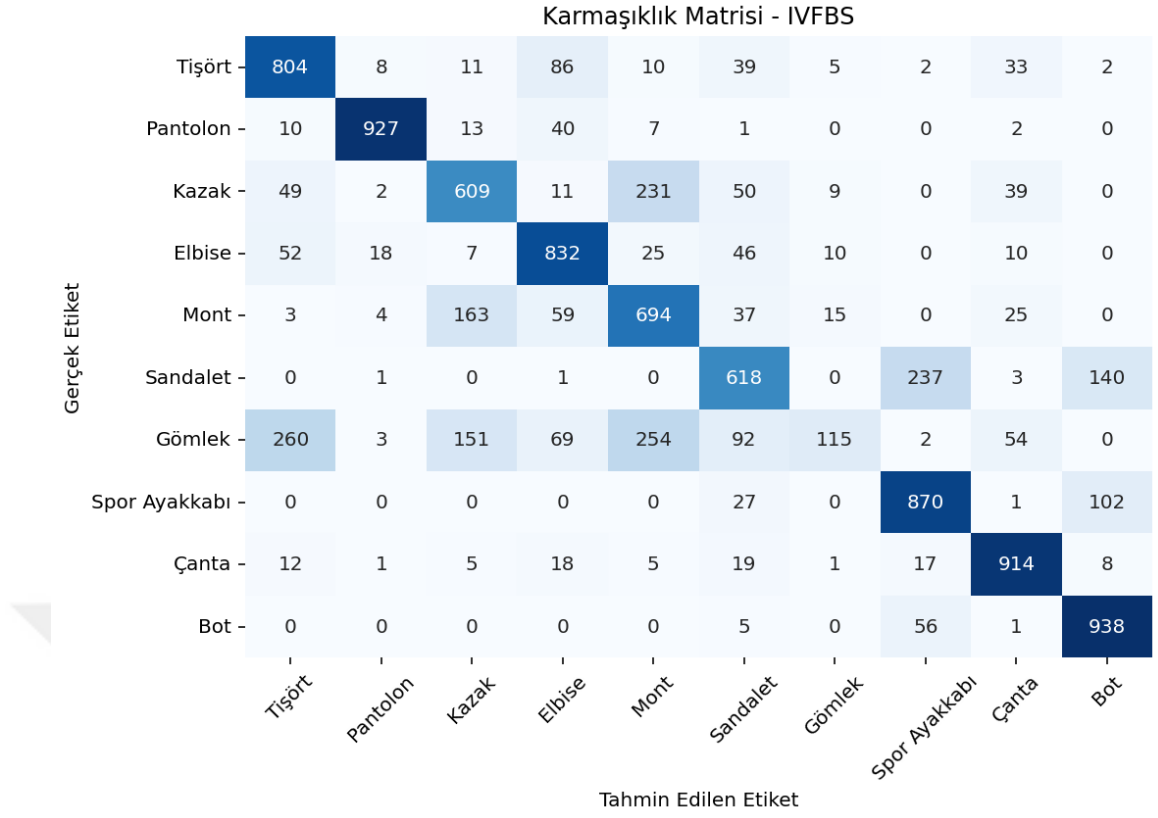
Fashion MNIST veri seti için düzenleme parametresi $\lambda = 0.0005$, gizli nöron sayısı 2500, iterasyon sayısı 2000 olarak belirlenmiştir. Başlangıç parametreleri $\theta_1 = 0.5$ ve $\delta_1 = -1/100$ olarak tanımlanırken, diğer parametreler $\alpha_k = 0.1$ ve $F = 0.99$ olarak seçilmiştir. Bu seçimlerine karşılık gelen sonuçlar aşağıda sunulmuştur.

Karmaşıklık Matrisi - Algoritma 4.1

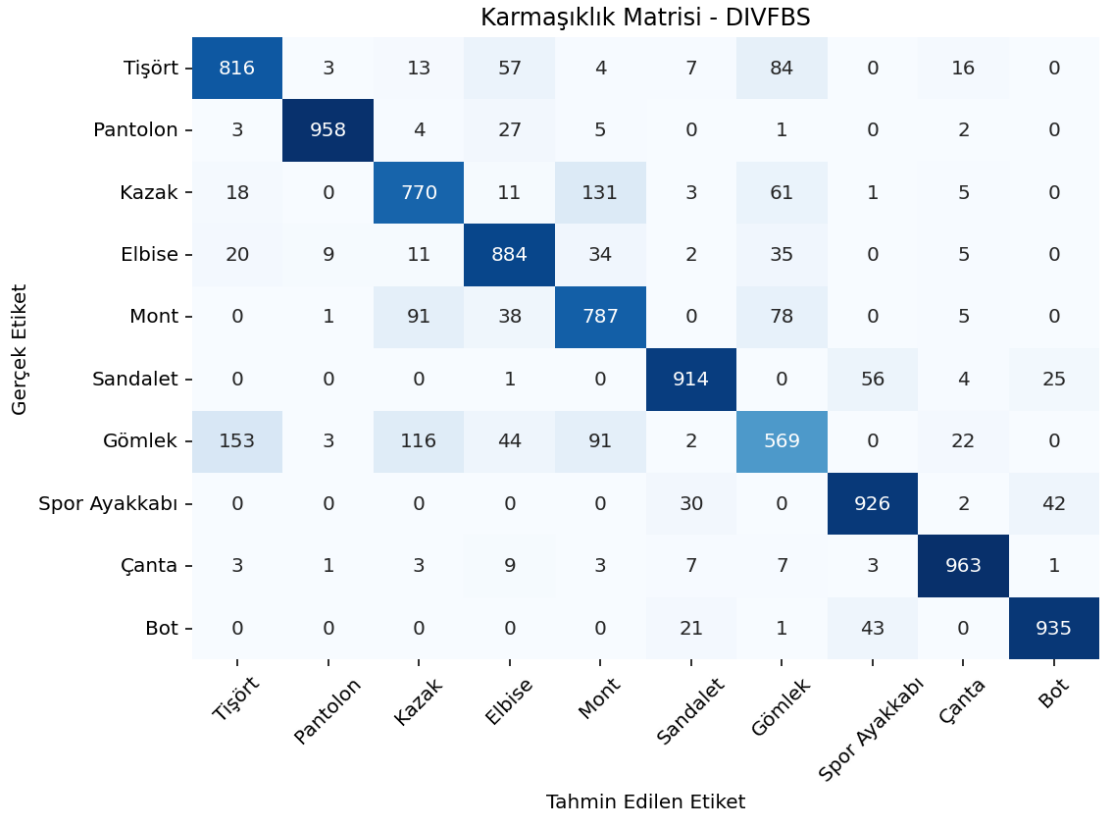
Gerçek Etiket	Tişört	Pantolon	Kazak	Elbise	Mont	Sandalet	Gömlek	Spor Ayakkabı	Çanta	Bot
Tişört	789	7	9	98	10	58	2	1	25	1
Pantolon	15	930	14	32	7	0	0	0	2	0
Kazak	48	0	585	12	243	71	5	0	36	0
Elbise	53	20	7	810	35	59	7	0	9	0
Mont	1	11	171	48	684	53	6	0	26	0
Sandalet	0	0	0	3	0	581	0	257	2	157
Gömlek	259	3	149	72	261	128	64	0	64	0
Spor Ayakkabı	0	0	0	0	0	19	0	877	0	104
Çanta	9	1	3	19	4	25	1	23	904	11
Bot	0	0	0	0	0	4	0	51	1	944
	Tişört	Pantolon	Kazak	Elbise	Mont	Sandalet	Gömlek	Spor Ayakkabı	Çanta	Bot
	Tahmin Edilen Etiket									

Şekil 4.27. Fashion MNIST veri seti: Algoritma 4.1 için karmaşıklık matrisi

4. İKİ EYLEMSİZLİK TERİMLİ İLERİ-GERİ AYRIŞTIRMA ALGORİTMALARI



Şekil 4.28. Fashion MNIST veri seti: IVFBS algoritması için karmaşıklık matrisi



Şekil 4.29. Fashion MNIST veri seti: DIVFBS algoritması için karmaşıklık matrisi

4. İKİ EYLEMSİZLİK TERİMLİ İLERİ-GERİ AYRIŞTIRMA ALGORİTMALARI

Fashion MNIST veri setinde bulunan sınıflar içerisinde gömlek, mont ve tişört gibi sınıflar görsel benzerliklere sahiptir. Dolayısıyla sınıflandırılması zor bir problem olarak kabul edilmektedir. Buna rağmen, DIVFBS algoritmasının diğer algoritmalara göre daha iyi sonuçlar verdiği görülmektedir.

Çizelge 4.15. Fashion MNIST veri seti: Algoritmaların sınıflandırma metrikleri

Algoritma	Doğruluk	Kesinlik	Duyarlılık	F1 Skoru
DIVFBS	0.85	0.85	0.85	0.85
IVFBS	0.73	0.74	0.73	0.71
Algoritma 4.1	0.72	0.72	0.72	0.69

Çizelge 4.15'e göre DIVFBS algoritması doğruluk, kesinlik, duyarlılık ve F1 skoru açısından diğer algoritmalara göre daha başarılı yöntem olmuştur. Genel olarak tablodan DIVFBS algoritmasının zorlayıcı Fashion MNIST veri setinde dahi diğer yöntemlere kıyasla daha etkili ve güvenilir bir sınıflandırma performansı sunduğu görülmektedir.

5. q-TÜREV VE YAPAY SİNİR AĞLARI UYGULAMALARI

q -türev kavramı matematik açısından adi türev yaklaşımına alternatif bir yapı sunmaktadır. Parametrik olarak esnek olması sayesinde farklı problem alanlarında uygulanabilir. Yapay sinir ağları bağlamında inceleyecek olursak q -türev kavramının bugüne kadar yalnızca aktivasyon fonksiyonlarının q -versiyonlarının geliştirilmesi aşamasında kullanıldığı bilinmektedir. Bu çalışmada ise bahse konu uygulama genişletilerek q -türev temelli geleneksel geri yayılım algoritması geliştirilmeye çalışılmıştır. Bilindiği kadarıyla literatürde ilk kez gerçekleştirilen bu uygulama, yapay sinir ağlarının öğrenme sürecine farklı bir bakış açısı kazandırmıştır. Ayrıca eğitim algoritmalarının performansını artırmaya yönelik alternatif bir yöntem ortaya koymaktadır.

5.1. q -Geri Yayılım Algoritması

Bu bölümde ilk olarak q -türevin klasik geri yayılım algoritmasına entegre edilmesi q -geri yayılım algoritması olarak adlandırdığımız model kullanılarak klasik yapay sinir ağları için önemli olan XOR probleminin başarıyla çözülebildiği gösterilmiştir. q -geri yayılım algoritmasında ağırlık güncellemeleri

$$(\theta_n)_{ji} = (\theta_n)_{ji} - \gamma \cdot \frac{\partial_q E}{\partial_q (\theta_n)_{ji}}$$

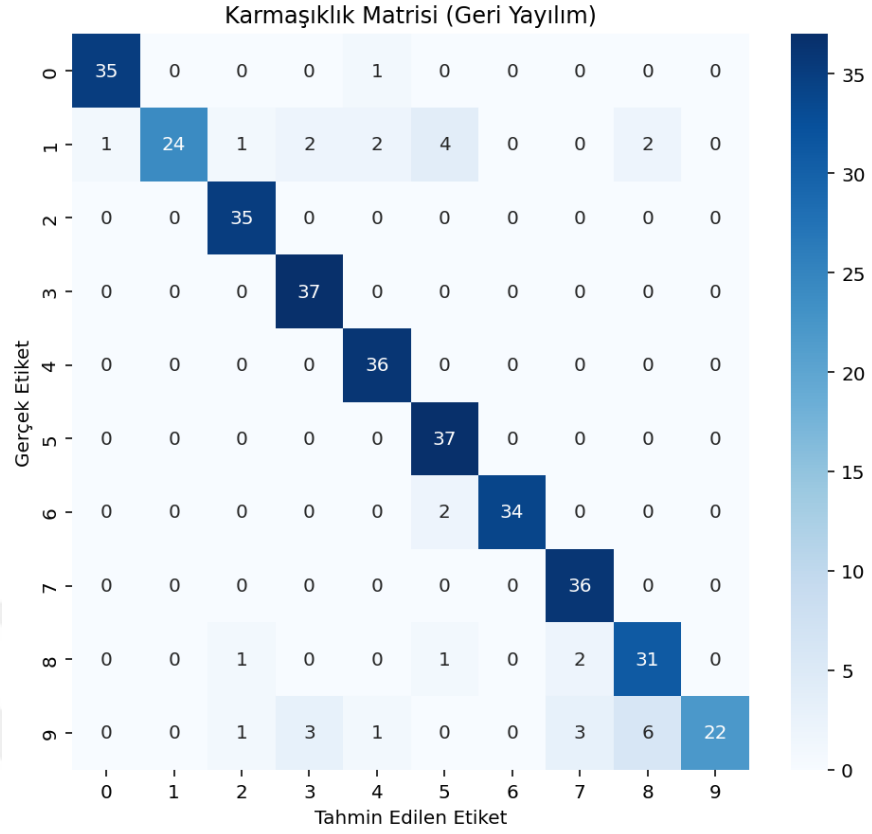
eşitliği yardımıyla yapılmaktadır. $\tanh$ aktivasyon fonksiyonu ile geleneksel geri yayılım yöntemi kullanılarak gerçekleştirilen benzetim sonuçları ve q -geri yayılım algoritması kullanılarak gerçekleştirilen benzetim sonuçları aşağıda verilmiştir. Kullanılan yapay sinir ağı mimarisi; 2 giriş, 2 gizli katman ve 1 çıkış katmanı nöronlarından oluşan çok katmanlı bir mimari şeklinde tasarlanmıştır. Böylece hem önerilen algoritmanın hem de q -geri yayılımın XOR probleminin çözümündeki etkisi değerlendirilmiştir.

Çizelge 5.1. XOR çözümü: *tanh* aktivasyon fonksiyonu ile geleneksel geri yayılım ve *q*-geri yayılım algoritması ile elde edilen sonuçlar

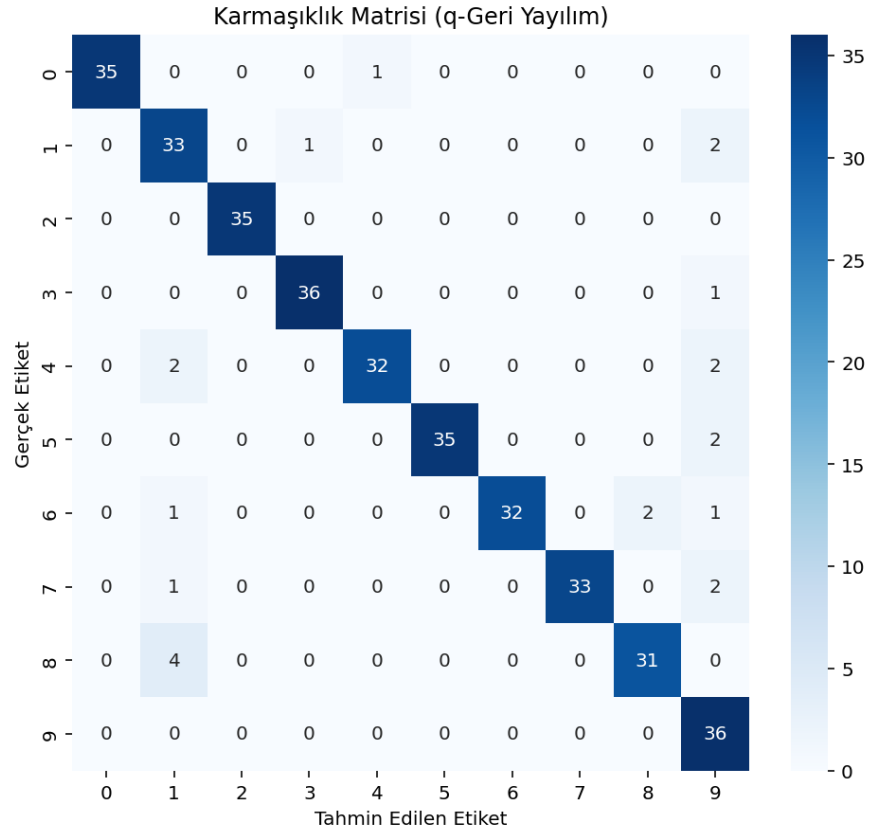
İterasyon sayısı	Geri yayılım için toplam hata	<i>q</i> - geri yayılım için toplam hata
500	0,386509	0,418860933
1000	0,334832	0,353893229
1500	0,258804	0,305025711
2000	0,127468	0,248005742
2500	0,047618	0,125516658
3000	0,021223	0,035014914
3500	0,011946	0,010772076
4000	0,007847	0,004593886
4500	0,005678	0,002475000
5000	0,004378	0,001543891
Toplam çalışma süresi:	0,0978 saniye	0.1275 saniye
Modelin son adımdaki doğruluğu:	100.00%	100.00%

Çizelge 5.1 incelendiğinde geri yayılım algoritması ile çalışan modelin hata değerlerinin iterasyonlar boyunca düzenli olarak azaldığı ve modelin 5000’inci iterasyonda %100 doğruluk oranına ulaştığı görülmektedir. Benzer şekilde *q*-geri yayılım modelinde ise hata değerleri geri yayılım algoritmasına göre daha hızlı bir biçimde azalmıştır. Yine bu model için de 5000’inci iterasyonda %100 doğruluk oranına ulaşılmıştır. Bu bulgular, *q*-türev yaklaşımının öğrenme sürecini hızlandırarak toplam hatayı daha kısa sürede küçülttüğünü göstermektedir.

Geri yayılım ile *q*-geri yayılım algoritmalarının başarılarını karşılaştırmak için sınıflandırma veri seti olan Digits kullanılmıştır. Elde edilen sonuçlar aşağıdaki şekil ve çizelgelerde sunulmuştur.



Şekil 5.1. Digits veri seti: Geri yayılım algoritması için karmaşıklık matrisi



Şekil 5.2. Digits veri seti: q-geri yayılım algoritması için karmaşıklık matrisi

Şekil 5.1’de modelin genel olarak yüksek doğruluk oranına sahip olduğu görülmektedir. Özellikle 0, 2, 3, 4, 5, 6 ve 7 numaralı sınıflar için doğru sınıflandırmaların baskın olduğu ve doğal olarak bu sınıflarda hataların minimum düzeyde kaldığı ifade edilebilir. Ancak bazı sınıflarda sınıflandırma hatalarının yoğunlaştığı tespit edilmiştir. Özellikle 1 numaralı sınıfın daha çok 5, 3, 2 ve 8 numaralı sınıflarla karıştırıldığı gözlenmiştir. Bu durum aslında modelin 1 numaralı sınıf ile benzer görsel özellikler taşıyan sınıfları ayırt etmede güçlük yaşadığını ifade etmektedir. Benzer şekilde 9 numaralı sınıf da çeşitli sınıflarla (özellikle 3, 4, 7 ve 8) karıştırılmıştır. Bu tür karışıklıklar ilgili sınıfların görsel yakınlığından ve modelin öğrenme kapasitesinin bu ayrımları netleştirmede yetersiz kalmasından kaynaklanmaktadır.

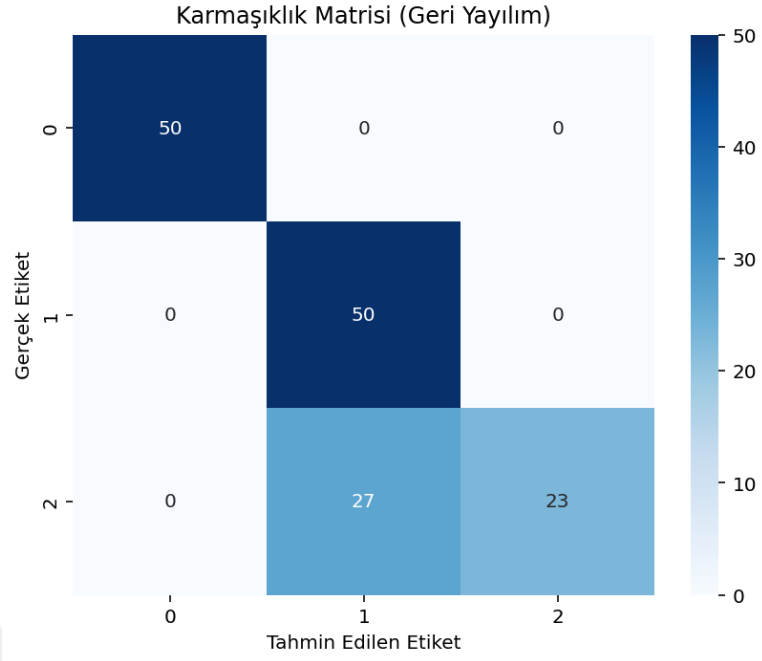
Şekil 5.2’de q -geri yayılım algoritmasının sınıflar arasında daha homojen ve daha başarılı sınıflandırmalar gerçekleştirmiştir. Modelin özellikle 0, 2, 3, 4, 5, 6, 7 numaralı sınıflarda hiç hata yapmadığı veya yalnızca bir örnek düzeyinde hata yaptığı görülmektedir. Bu sonuçlar incelendiğinde q -geri yayılım modelinin geleneksel geri yayılım modeline göre daha başarılı olduğu görülmektedir.

Çizelge 5.2. Digits veri seti: Algoritmaların sınıflandırma metrikleri

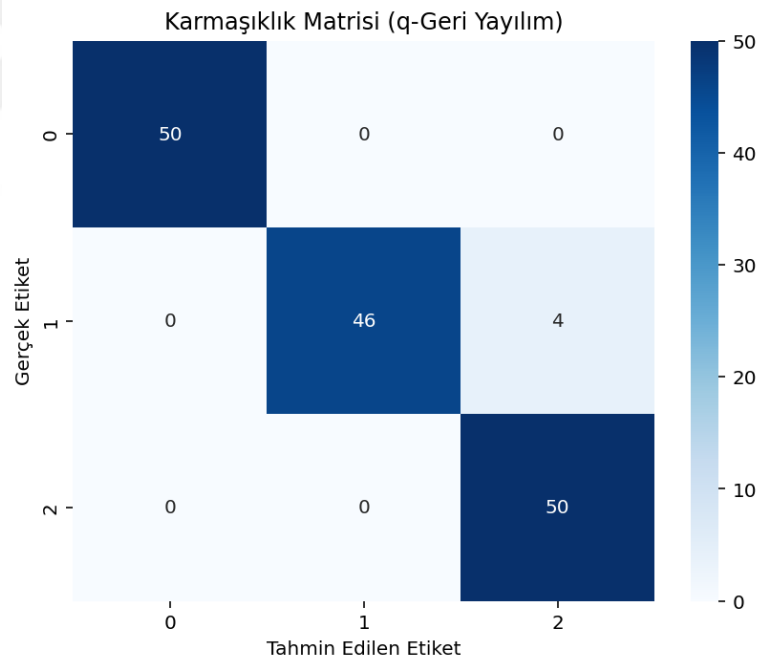
Algoritma	Doğruluk	Kesinlik	Duyarlılık	F1 Skoru
q -Geri yayılım	0.94	0.95	0.94	0.94
Geri Yayılım	0.91	0.92	0.91	0.90

Çizelge 5.2’de q -geri yayılım algoritmasının tüm performans metriklerinde standart geri yayılım algoritmasından daha başarılı olduğu görülmektedir. Bu sonuçlar yöntemin yalnızca genel doğruluk oranını değil aynı zamanda modelin doğru sınıfları yakalama ve hatalı sınıflandırmaları azaltma becerisini de geliştirdiğini göstermektedir.

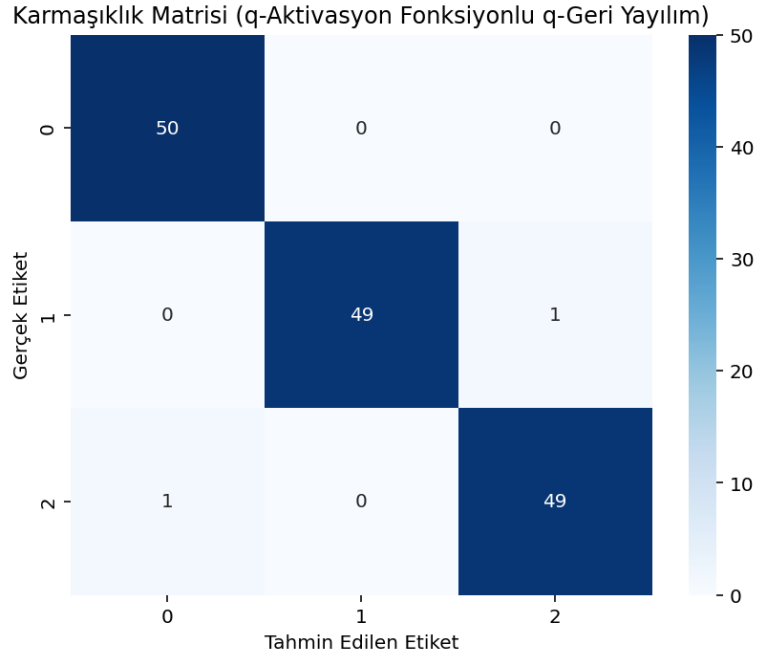
Çalışmanın devamında geri yayılım, q -geri yayılım ve q -aktivasyon fonksiyonlu q -geri yayılım algoritmalarının performanslarını karşılaştırmak adına Iris veri seti kullanılmıştır. Elde edilen sonuçlar aşağıdaki şekil ve çizelgelerde sunulmuştur.



Şekil 5.3. Iris veri seti: Geri yayılım algoritması için karmaşıklık matrisi



Şekil 5.4. Iris veri seti: q -geri yayılım algoritması için karmaşıklık matrisi



Şekil 5.5. Iris veri seti: q -aktivasyon fonksiyonlu q -geri yayılım algoritması için karmaşıklık matrisi

Iris veri seti ile yukarıda üç farklı geri yayılım modeli karşılaştırılmıştır. İlk modelde standart $\tanh$ aktivasyon fonksiyonu, ikinci modelde q -türev, üçüncü modelde ise $q - \tanh$ ile q -türev kullanılmıştır. Elde edilen karmaşıklık matrisleri incelendiğinde geri yayılım algoritmasının bazı sınıflarda ayırma zorlandığı, q -türev kullanılan modelde bu durumun kısmen iyileştiği, $q - \tanh$ ve q -türevin birlikte kullanıldığı modelde ise her sınıf için sınıflandırma başarımının daha yüksek olduğu görülmektedir.

Çizelge 5.3. Iris veri seti: Geri yayılım algoritmalarının sınıflandırma metrikleri

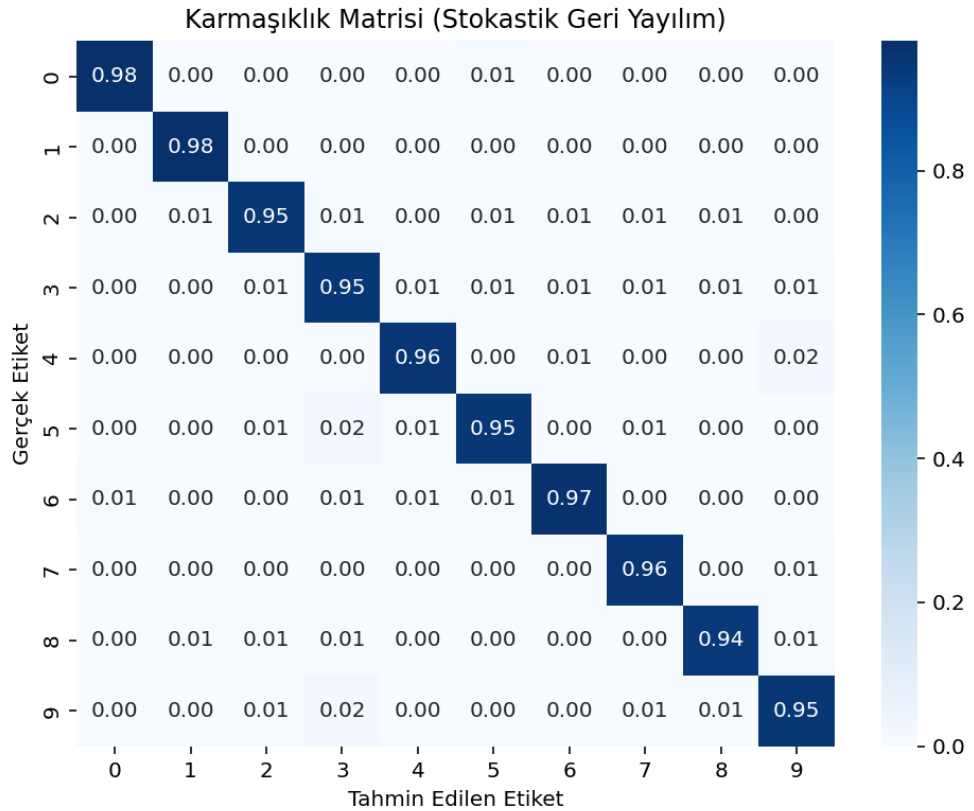
Algoritma	Doğruluk	Kesinlik	Duyarlılık	F1 Skoru
q -Aktivasyon Fonksiyonlu q -Geri Yayılım	0.99	0.99	0.99	0.99
q -Geri Yayılım	0.97	0.98	0.97	0.97
Geri Yayılım	0.82	0.88	0.82	0.81

Çizelge 5.3'te q -türev barındıran yaklaşımların standart geri yayılım algoritmasına göre başarılı sonuçlar verdiği görülmektedir. Özellikle de q -aktivasyon fonksiyonlu q -geri yayılım modeli diğer tüm modellerden daha başarılı olmuştur. Bu

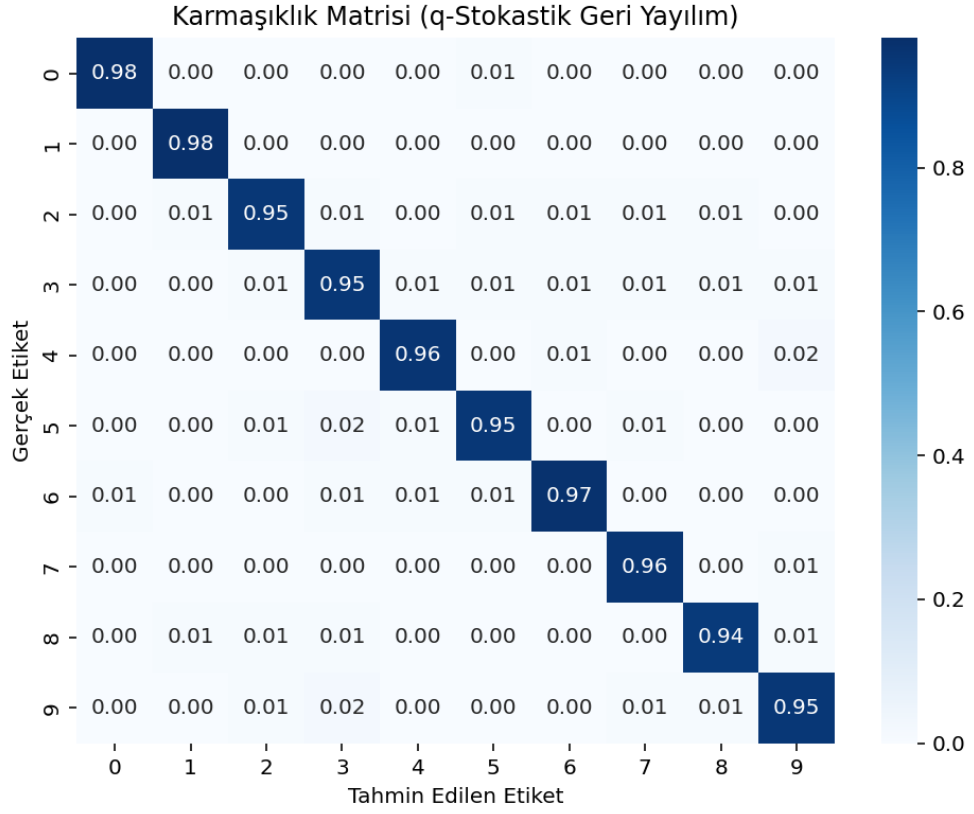
sonuçlar ışığında q -türev temelli yöntemlerin sınıflandırma başarısını anlamlı ölçüde artırdığı da görülmektedir.

5.2. q -Stokastik Geri Yayılım Algoritması

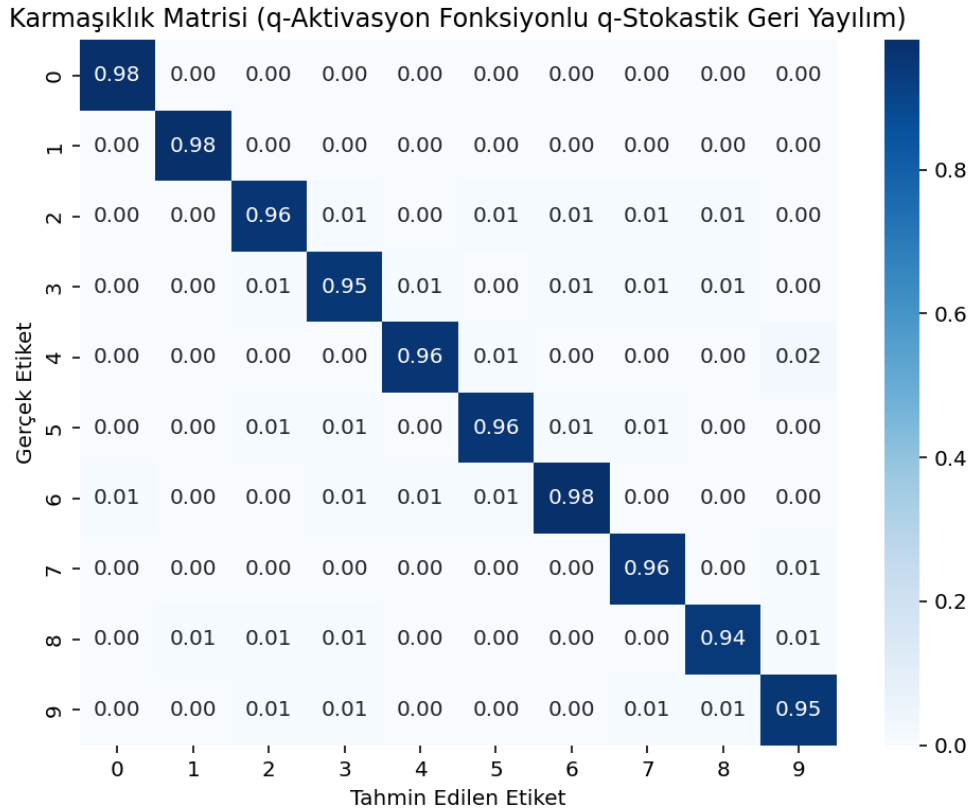
Bu bölümde geleneksel stokastik geri yayılım algoritmasında alınan adi türevler yerine q -türev alınarak q -stokastik geri yayılım algoritması oluşturulmuştur. Ayrıca q -aktivasyon fonksiyonları yeni tanımlanan bu q -stokastik geri yayılım algoritmasına entegre edilerek yeni bir yöntem sunulmuştur. Sınıflandırma problemini çözmek için ise stokastik geri yayılım, q -stokastik geri yayılım ve q -aktivasyon fonksiyonlu q -stokastik geri yayılım algoritmaları kullanılmış ve performanslarını karşılaştırmak adına MNIST veri seti kullanılmıştır. Elde edilen sonuçlar aşağıdaki şekil ve çizelgelerde sunulmuştur.



Şekil 5.6. MNIST veri seti: Stokastik geri yayılım algoritması için karmaşıklık matrisi



Şekil 5.7. MNIST veri seti: q -stokastik geri yayılım algoritması için karmaşıklık matrisi



Şekil 5.8. MNIST veri seti: q -aktivasyon fonksiyonlu q -stokastik geri yayılım algoritması için karmaşıklık matrisi

Çizelge 5.4. Stokastik geri yayılım algoritmalarının MNIST veri seti ile sınıflandırma metrikleri

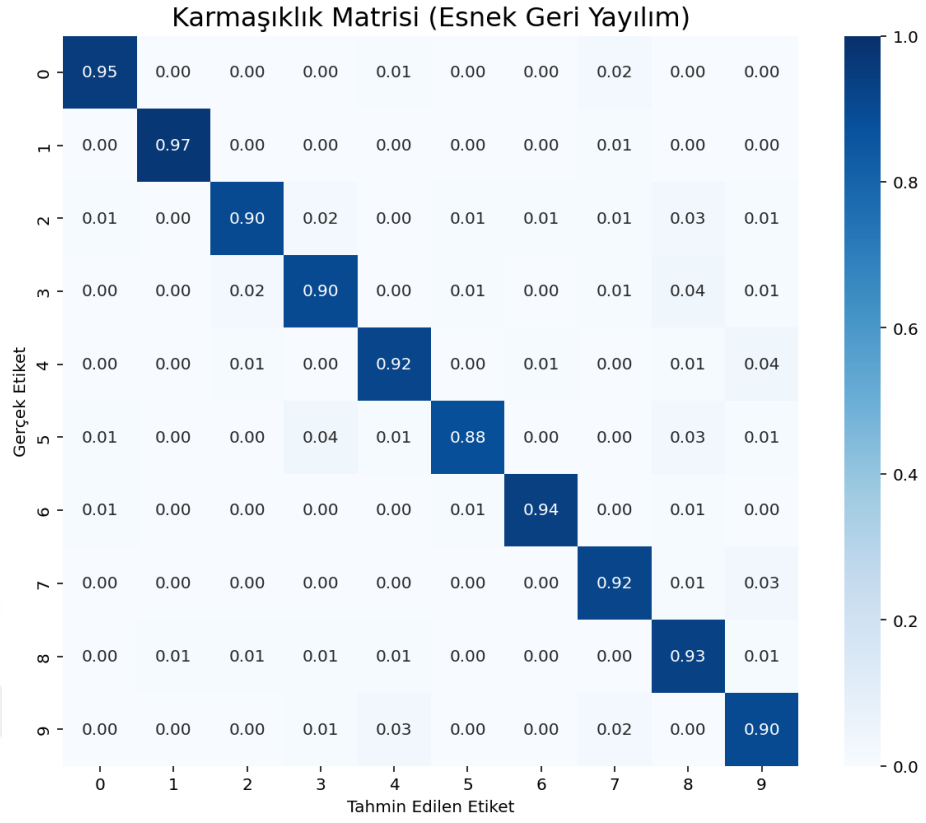
Algoritma	Doğruluk	Kesinlik	Duyarlılık	F1 Skoru
q -Aktivasyon Fonksiyonlu q -Stokastik Geri Yayılım	0.9625	0.96	0.96	0.96
q - Stokastik Geri Yayılım	0.9580	0.96	0.96	0.96
Stokastik Geri Yayılım	0.9582	0.96	0.96	0.96

Şekil 5.6’da verilen modelin genel olarak başarıyla sınıflandırma yaptığı görülmektedir. Ancak özellikle 3, 5 ve 9 numaralı sınıflarda köşegen dışındaki alanlarda hatalar artmıştır. Bu durum söz konusu rakamların benzer yapısal özelliklerinden kaynaklı olarak model tarafından daha zor ayırt edildiğini göstermektedir. Şekil 5.7’de verilen modelin ise Şekil 5.6’da verilen modelde olduğu gibi ana köşegen üzerindeki başarı oranlarının benzer olduğu gözlemlenmektedir. Şekil 5.8’de ise ana köşegen üzerindeki değerlerin daha başarılı ve köşegen dışındaki hataların daha düşük olduğu görülmektedir. Bu durum, q -aktivasyon fonksiyonlu q -stokastik geri yayılım modelinin resim verileri üzerindeki renk tonlamalarını daha etkili ayırt edebildiğini göstermektedir.

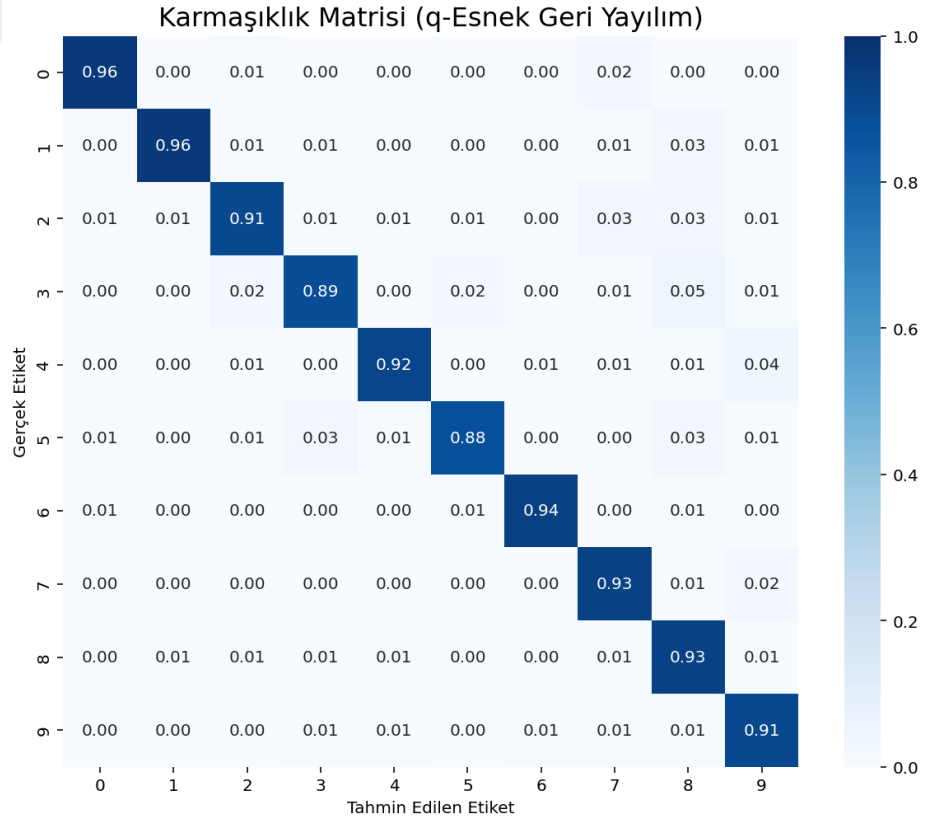
Son olarak Çizelge 5.4 incelendiğinde tüm algoritmaların yüksek doğruluk, kesinlik, duyarlılık ve $F1$ skoru değerlerine ulaştığı görülmektedir. Ancak q -aktivasyon fonksiyonlu q -stokastik geri yayılım algoritması diğer yöntemlere göre küçükte olsa üstünlük sağlamıştır.

5.3. q -Esnek Geri Yayılım Algoritması

Bu bölümde geleneksel esnek geri yayılım algoritmasında alınan adi türevler yerine q -türev alınarak q -esnek geri yayılım algoritması oluşturulmuştur. Ayrıca q -aktivasyon fonksiyonları yeni tanımlanan bu q -esnek geri yayılım algoritmasına entegre edilerek yeni bir yöntem sunulmuştur. Sınıflandırma problemini çözmek için ise esnek geri yayılım, q -esnek geri yayılım ve q -aktivasyon fonksiyonlu q -esnek geri yayılım algoritmaları kullanılmış ve performanslarını karşılaştırmak adına MNIST veri seti kullanılmıştır. Elde edilen sonuçlar aşağıdaki şekil ve çizelgelerde sunulmuştur.

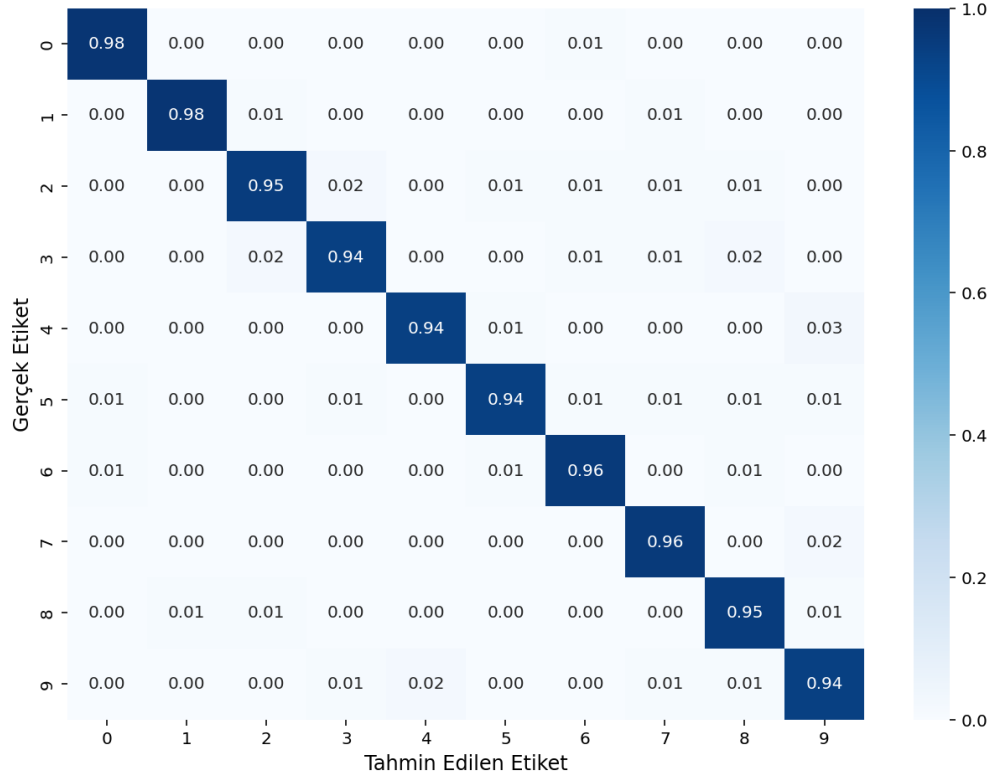


Şekil 5.9. MNIST veri seti: Esnek geri yayılım algoritması için karmaşıklık matrisi



Şekil 5.10. MNIST veri seti: q-esnek geri yayılım algoritması için karmaşıklık matrisi

Karmaşıklık Matrisi (q-Aktivasyon Fonksiyonlu q-Esnek Geri Yayılım)

Şekil 5.11. MNIST veri seti: q -aktivasyon fonksiyonlu q -esnek geri yayılım algoritması için karmaşıklık matrisi

Çizelge 5.5. Esnek geri yayılım algoritmalarının MNIST veri seti ile sınıflandırma metrikleri

Algoritma	Doğruluk	Kesinlik	Duyarlılık	F1 Skoru
q -Aktivasyon Fonksiyonlu q -Esnek Geri Yayılım	0.9528	0.9529	0.9528	0.9528
q -Esnek Geri Yayılım	0.9249	0.9263	0.9249	0.9252
Esnek Geri Yayılım	0.9218	0.9230	0.9218	0.9220

Şekil 5.9’da esnek geri yayılım algoritmasına ait karmaşıklık matrisi verilmiştir. Sonuçlar genel olarak yüksek doğruluk değerlerine işaret etse de bazı sınıflarda hataların yoğunlaştığı görülmektedir. Özellikle 3, 7 ve 8 rakamlarının sınıflandırılmasında yanlış eşleştirmeler daha belirgin olmuştur. Bu rakamların yapısal benzerliklerinden kaynaklı olup modelin bu rakamları ayırt etme gücü ise sınırlı kalmıştır.

Şekil 5.10’da q -esnek geri yayılım algoritmasının sonuçları sunulmuştur. Burada standart yönteme kıyasla bazı sınıflarda iyileşme gözlemlenmiştir. Genel olarak model

daha kararlı sonuçlar üretmiş ve doğruluk açısından da standart esnek geri yayılım modelini geçmiştir.

Şekil 5.11’de q -aktivasyon fonksiyonunun q -esnek geri yayılımla birlikte kullanıldığı durumda ise en yüksek başarı elde edilmiştir. Doğru sınıflandırmaların oranı artmış özellikle bazı sınıflardaki hata oranları azalmıştır.

Çizelge 5.5’te verilen sayısal metrikler de bu gözlemleri desteklemektedir. Tüm yöntemler yüksek başarı göstermesine rağmen q -aktivasyon fonksiyonlu q -esnek geri yayılım yöntemi en iyi performansı sunmuştur. Bu bulgu q -parametresinin yalnızca geri yayılım sürecinde değil aynı zamanda aktivasyon fonksiyonuna entegre edildiğinde de sınıflandırma doğruluğunu artırdığını ortaya koymaktadır.

6. SONUÇ ve ÖNERİLER

Bu tezde yapay sinir ağı algoritmaları olan geri yayılım, stokastik geri yayılım ve esnek geri yayılım algoritmalarının q -türevli versiyonları tanımlanmıştır. Literatürde q -türev yapay sinir ağlarında sadece aktivasyon fonksiyonları aşamasında kullanılırken bu tezde ilk kez q -türev zincir kuralını barındırmamasına rağmen doğrudan ağırlıkların optimizasyon aşamasına entegre edilmiştir. Tanımlanan bu yeni algoritmaların sınıflandırma problemlerini içeren veri setleri üzerinde benzetimleri gerçekleştirilmiştir. Elde edilen sonuçlar kapsamında q -türev temelli algoritmaların geleneksel yöntemlere kıyasla daha başarılı sonuçlar verdiği tespit edilmiştir. Öte yandan reel Hilbert uzaylarında monoton kapsama problemlerinin çözümüne yönelik literatürde var olan ileri-geri ayrıştırma algoritmalarına çift eylemsizlik terimi ile viskozite terimi eklenerek yeni iki algoritma tanımlanmıştır. Bu algoritmaların teorik alt yapıları ile yakınsaklıkları ifade ve ispat edilmiştir. Yine bu iki algoritma regresyon ve sınıflandırma problemleri üzerinde iyi bilinen veri setleri kullanılarak test edilmiştir. Oluşturulan mimaride tarafımızca modifiye edilen ekstrem öğrenme makinesi kullanılmış ve bu mimaride kullanılan H matrisi önerilen algoritmalar yardımıyla iteratif olarak oluşturulmuştur. Literatürde var olan ve iyi sonuçlar veren başkaca algoritmalar ile yapılan performans kıyaslamalarında önerilen iki algoritmanın da daha başarılı olduğu görülmüştür.

Ele alınan problemlerin doğası gereği bu tez kapsamında önerilen yeni algoritmaların disiplinler arası pek çok uygulamada kullanılabileceği düşünülmektedir. Bu yönüyle bu tez hem teorik hem de uygulama açısından gelecekte yapay sinir ağı öğrenme algoritmalarına önemli bir zemin oluşturmaktadır.

Tez kapsamında yapılan deneyler çok katmanlı algılayıcılar (MLP) üzerinde gerçekleştirilmiştir. Gelecek çalışmalarda, q -türev ve q -aktivasyon içeren yaklaşımlarının evrimsel sinir ağları (CNN) üzerinde, tekrarlayan yapılar (RNN, LSTM) üzerinde ve günümüzde öne çıkan dönüşüm tabanlı modeller (transformer) üzerinde test edilmesi, önerilen yöntemin derin öğrenme alanındaki başarımını daha görünür kılacaktır. Yine bunun yanı sıra, q -türevli yöntemlerin bazı durumlarda ek hesaplama maliyeti getirdiği göz önüne alındığında, gelecekte GPU, TPU ve benzeri paralel işlemci tabanlı donanımlarda verimliliklerinin test edilmesi gerekmektedir.

Dolayısıyla donanım dostu iyileştirmelerin geliştirilmesi de önemli bir araştırma alanı olarak görülmektedir.

Son olarak, bu çalışmada önerilen yöntemlerin yalnızca yapay sinir ağları ile sınırlı kalmayıp mühendislik, ekonomi ve fizik gibi farklı disiplinlerde var olan güncel problemlere de uygulanması elde edilen başarının daha geniş alanlara taşınmasını mümkün kılacaktır. Böylece, geliştirilen yapıların esnekliği sayesinde disiplinler arası problemlerin çözümünde de kullanılabilecektir.



KAYNAKLAR

- Alcalá-Fdez, J., Sánchez, L., García, S., del Jesus, M. J., Ventura, S., Garrell, J. M., Otero, J., Romero, C., Bacardit, J., Rivas, V. M., Fernández, J. C. ve Herrera, F. 2009. KEEL: a software tool to assess evolutionary algorithms for data mining problems. *Soft Computing*, 13(3), 307-318.
- Aleyner, A. ve Reich, S. 2009. Random products of quasi-nonexpansive mappings in Hilbert space. *Journal of Convex Analysis*, 16, 633-640.
- Anandan, H. 2025. Tumor growth stage prediction on ultrasound breast cancer images with backpropagation network. *Grn Int J Apl Med Sci*, 3(3), 118–121.
- Anonim, 2025. Web Sitesi: <https://www.geeksforgeeks.org/machine-learning/variational-autoencoders/>, Erişim Tarihi: 20.08.2025.
- Arf, C. 1959. Makine Düşünebilir mi ve Nasıl Düşünebilir?, In: Atatürk Üniversitesi Halk Konferansları 1958-1960. Gözütok, T. T. (ed), Atatürk Üniversitesi Yayınları, 91-103, Erzurum.
- Asuncion, A. ve Newman, D. 2007. Web Sitesi: <https://archive.ics.uci.edu>, Erişim Tarihi: 18.08.2025.
- Banach, S. 1922. Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales. *Fundamenta mathematicae*, 3(1), 133-181.
- Bauschke, H. H. ve Combettes, P. L. 2017. Correction to: Convex Analysis and Monotone Operator Theory in Hilbert Spaces, In: Convex Analysis and Monotone Operator Theory in Hilbert Spaces. Bauschke, H. H. ve Combettes, P. L. (eds), Springer International Publishing, C1-C4, Cham.
- Bayraktar, M. 2006. Fonksiyonel analiz. Gazi Kitabevi, 456, Ankara.
- Beck, A. ve Teboulle, M. 2009. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM journal on imaging sciences*, 2(1), 183-202.
- Boyd, S. P. ve Vandenberghe, L. 2004. Convex optimization. Cambridge University Press, 714, Cambridge.
- Brouwer, L. E. J. 1911. Über abbildung von mannigfaltigkeiten. *Mathematische annalen*, 71(1), 97-115.
- Browder, F. E. 1965. Fixed-point theorems for noncompact mappings in Hilbert space. *Proceedings of the National Academy of Sciences*, 53(6), 1272-1276.
- Browder, F. E. 1967. Existence and perturbation theorems for nonlinear maximal monotone operators in Banach spaces. *Bulletin of the American Mathematical Society*, 73(3), 322-327.

- Cegielski, A. 2012. Iterative methods for fixed point problems in Hilbert spaces. Springer, 298, Heidelberg.
- Doersch, C. 2016. Tutorial on variational autoencoders. Web Sitesi: <https://arxiv.org/abs/1606.05908>, Erişim Tarihi: 18.08.2025.
- Feldmesser, J. 1987. Computer Hardware [Dataset]. Web Sitesi: <https://doi.org/10.24432/C5830D>, Erişim Tarihi: 18.08.2025.
- Fenchel, W. 1949. On conjugate convex functions. Canadian Journal of Mathematics, 1(1), 73-77.
- Fisher, R. 1936. Iris [Dataset]. Web Sitesi: <https://doi.org/10.24432/C56C76>, Erişim Tarihi: 18.08.2025.
- Gouvêa, É. J., Regis, R. G., Soterroni, A. C., Scarabello, M. C. ve Ramos, F. M. 2016. Global optimization using q -gradients. European Journal of Operational Research, 251(3), 727-738.
- Hastie, T., Tibshirani, R. ve Friedman, J. H. 2009. The elements of statistical learning: data mining, inference, and prediction. Springer, 758, New York.
- Hebb, D. O. 1949. The first stage of perception: growth of the assembly. The Organization of Behavior, 60, 78–104.
- Heine, E. 1878. Theorie der Kugelfunctionen und der verwandten Functionen. Walter de Gruyter GmbH & Co KG, 342, Berlin.
- Huang, G. B., Zhu, Q. Y. ve Siew, C. K. 2006. Extreme learning machine: Theory and applications. Neurocomputing, 70(1), 489-501.
- Inthakon, W., Suantai, S., Sarnmeta, P. ve Chumpungam, D. 2020. A New Machine Learning Algorithm Based on Optimization Method for Regression and Classification Problems. Mathematics, 8(6), 1007.
- Jackson, F. H. 1908. On q -functions and a certain difference operator. Earth and Environmental Science Transactions of the Royal Society of Edinburgh, 46(2), 253-281.
- Jolaoso, L. O., Shehu, Y., Yao, J. C. ve Xu, R. 2023. Double Inertial Parameters Forward-Backward Splitting Method: Applications to Compressed Sensing, Image Processing, and SCAD Penalty Problems. Journal of Nonlinear & Variational Analysis, 7(4), 627-646.
- Kahn, M. 1994. Web Sitesi: <https://doi.org/10.24432/C5T59G>, Erişim Tarihi: 18.08.2025.
- Kankam, K. ve Cholanjiak, P. 2023. Double inertial proximal gradient algorithms for convex optimization problems and applications. Acta Mathematica Scientia, 43(3), 1462-1476.

- Kingma, D. P. ve Welling, M. 2013. Web Sitesi: <https://arxiv.org/abs/1312.6114>, Erişim Tarihi: 18.08.2025.
- Kirk, W. A. 1965. A fixed point theorem for mappings which do not increase distances. *The American mathematical monthly*, 72(9), 1004-1006.
- Kitkuan, D., Kumam, P., Padcharoen, A., Kumam, W. ve Thounthong, P. 2019. Algorithms for zeros of two accretive operators for solving convex minimization problems and its application to image restoration problems. *Journal of Computational and Applied Mathematics*, 354, 471-495.
- Kitkuan, D., Kumam, P., Martínez-Moreno, J. ve Sitthithakerngkiet, K. 2020. Inertial viscosity forward-backward splitting algorithm for monotone inclusions and its application to image restoration problems. *International Journal of Computer Mathematics*, 97(1-2), 482-497.
- LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W. ve Jackel, L. D. 1989. Backpropagation applied to handwritten zip code recognition. *Neural computation*, 1(4), 541-551.
- LeCun, Y., Bottou, L., Bengio, Y. ve Haffner, P. 1998. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278-2324.
- Li, H., Wang, X. ve Wang, F. 2025. Projection and contraction method with double inertial steps for quasi-monotone variational inequalities. *Optimization*, 74(7), 1643-1674.
- Lions, P. L. ve Mercier, B. 1979. Splitting algorithms for the sum of two nonlinear operators. *SIAM Journal on Numerical Analysis*, 16(6), 964-979.
- López, G., Martín-Márquez, V., Wang, F. ve Xu, H. K. 2012. Forward-backward splitting methods for accretive operators in Banach spaces. *Abstract and Applied Analysis*, 2012(1), 109236.
- Maingé, P. E. 2008. Strong convergence of projected subgradient methods for nonsmooth and nonstrictly convex minimization. *Set-valued analysis*, 16(7), 899-912.
- Marino, G. ve Xu, H. K. 2004. Convergence of generalized proximal point algorithms. *Commun. Pure Appl. Anal*, 3(4), 791-808.
- Marvin, M. ve Seymour, A. P. 1969. *Perceptrons*. The MIT Press, 258, Cambridge.
- Minty, G. J. 1962. Monotone (nonlinear) operators in Hilbert space. *Duke Mathematical Journal*, 29(3), 341-346.
- Mishra, S. K., Samei, M. E., Chakraborty, S. K. ve Ram, B. 2021. On q-variant of Dai-Yuan conjugate gradient algorithm for unconstrained optimization problems. *Nonlinear Dynamics*, 104(3), 2471-2496.

- Moreau, J. J. 1962. Fonctions convexes duales et points proximaux dans un espace hilbertien. Comptes rendus hebdomadaires des séances de l'Académie des sciences, 255(23), 2897-2899.
- Moreau, J. J. 1965. Proximité et dualité dans un espace hilbertien. Bulletin de la Société mathématique de France, 93, 273-299.
- Nevitt, J. ve Hancock, G. R. 2000. Improving the root mean square error of approximation for nonnormal conditions in structural equation modeling. The Journal of experimental education, 68(3), 251-268.
- Nielsen, F. ve Sun, K. 2020. Q-neurons: Neuron activations based on stochastic Jackson's derivative operators. IEEE Transactions on Neural Networks and Learning Systems, 32(6), 2782-2789.
- Nwankpa, C. 2018. Web Sitesi: <https://arxiv.org/abs/1811.03378>, Erişim Tarihi: 18.08.2025.
- Öztemel, E. 2003. Yapay sinir ağları. Papatya Yayıncılık, 230, İstanbul.
- Pace, R. K. ve Barry, R. 1997. Sparse spatial autoregressions. Statistics & Probability Letters, 33(3), 291-297.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. ve Duchesnay, É. 2011. Scikit-learn: Machine learning in Python. The Journal of Machine Learning Research, 12, 2825-2830.
- Puangpee, J. ve Suantai, S. 2020. A new accelerated viscosity iterative method for an infinite family of nonexpansive mappings with applications to image restoration problems. Mathematics, 8(4), 615.
- Riedmiller, M. ve Braun, H. 1993. A direct adaptive method for faster backpropagation learning: the RPROP algorithm, IEEE International Conference on Neural Networks, 28 Mart-1 Nisan, 586-591, San Francisco, USA.
- Rockafellar, R. T. 1970a. On the maximal monotonicity of subdifferential mappings. Pacific Journal of Mathematics, 33(1), 209-216.
- Rockafellar, R. T. 1970b. Convex Analysis. Princeton University Press, 451, Princeton.
- Rockafellar, R. T. ve Wets, R. J. 1998. Variational analysis. Springer, 735, Berlin, Heidelberg.
- Rosenblatt, F. 1958. The perceptron: a probabilistic model for information storage and organization in the brain. Psychological review, 65(6), 386-408.
- Rumelhart, D. E., Hinton, G. E. ve Williams, R. J. 1986. Learning representations by back-propagating errors. Nature, 323(6088), 533-536.

- Sammut, C. ve Webb, G. I. 2010a. Mean absolute error, In: Encyclopedia of machine learning. Sammut, C. ve Webb, G. I. (eds), Springer, 652, New York.
- Sammut, C. ve Webb, G. I. 2010b. Mean squared error, In: Encyclopedia of machine learning. Sammut, C. ve Webb, G. I. (eds), Springer, 653, New York.
- Shobug, M. A., Hossain, M. A., Yang, F. ve Lu, J. 2025. Dynamic control of isolated network microgrids: A resilient backpropagation neural network based virtual inertia control approach. IEEE Access.
- Simon, H. 1990. Reason in human affairs. Stanford University Press, 128, Stanford.
- Simonoff, J. S. 2003. Analyzing categorical data. Springer, 498, New York.
- Soruç, U., 2020. Web Sitesi: <https://medium.com/@ugurcan.soruc/xor-problemi-nedir-ve-cozumü-197f9b110053>, Erişim Tarihi: 18.08.2025.
- Soterroni, A. C., Galski, R. L. ve Ramos, F. M. 2011. The q-Gradient Vector for Unconstrained Continuous Optimization Problems, In: Operations Research Proceedings 2010. Hu, B., Morasch, K., Pickl, S., Siegle, M. (eds), Springer, 365-370, Berlin.
- Stone, M. 1974. Cross-validatory choice and assessment of statistical predictions. Journal of the royal statistical society: Series B (Methodological), 36(2), 111-133.
- Sun, T., Wu, Y., Qian, X. ve Wang, B. 2025. Interval optimal scheduling method of integrated energy system with ground source heat pump based on quantum derivative algorithm. *Energy Reports*, 13: 4151–4161.
- Suantai, S., Inkrong, P. ve Chalamjiak, P. 2023. Forward–backward–forward algorithms involving two inertial terms for monotone inclusions. *Computational and Applied Mathematics*, 42(6), 255.
- Thomae, J. 1870. Les séries Heinéennes supérieures, ou les séries de la forme. *Annali di Matematica Pura ed Applicata* (1867-1897), 4(1), 105-138.
- Thong, D. V., Dung, V. T., Anh, P. K. ve Thang, H. V. 2023. A single projection algorithm with double inertial extrapolation steps for solving pseudomonotone variational inequalities in Hilbert space. *Journal of Computational and Applied Mathematics*, 426, 115099.
- Tibshirani, R. 1996. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 58(1), 267-288.
- Tseng, P. 2000. A modified forward-backward splitting method for maximal monotone mappings. *SIAM Journal on Control and Optimization*, 38(2), 431-446.
- Uğuz, S. 2019. Makine öğrenmesi teorik yönleri ve Python uygulamaları ile bir yapay zekâ ekolü. Nobel Yayıncılık, 312, Ankara.

- Verma, M., Sahu, D. R. ve Shukla, K. K. 2018. VAGA: a novel viscosity-based accelerated gradient algorithm: Convergence analysis and applications. *Applied Intelligence*, 48(9), 2613-2627.
- Viet, T. D., Tien, D. V., Huong, H. P. T. ve Thanh, T. H. T. 2022. Web Sitesi: <https://doi.org/10.21203/rs.3.rs-2198022/v1>, Erişim Tarihi: 18.08.2025.
- Wang, H., Wang, L., Wang, Z., Ma, L. ve Luo, Y. 2025. SSC-VAE: Structured sparse coding based variational autoencoder for detail preserved image reconstruction. *Proceedings of the AAAI Conference on Artificial Intelligence*, 39(7), 7665–7673.
- Wang, K., Wang, Y., Iyiola, O. S. ve Shehu, Y. 2024. Double inertial projection method for variational inequalities with quasi-monotonicity. *Optimization*, 73(3), 707-739.
- Wang, Z. B., Long, X., Lei, Z. Y. ve Chen, Z. Y. 2022. New self-adaptive methods with double inertial steps for solving splitting monotone variational inclusion problems with applications. *Communications in Nonlinear Science and Numerical Simulation*, 114, 106656.
- Wang, Z. B., Lei, Z. Y., Long, X. ve Chen, Z. Y. 2023. A modified Tseng splitting method with double inertial steps for solving monotone inclusion problems. *Journal of Scientific Computing*, 96(3), 92.
- Werbos, P. J. 1982. Applications of advances in nonlinear sensitivity analysis, In: *System Modeling and Optimization*. Drenick, R. F. ve Kozin, F. (eds), Springer Berlin Heidelberg, 762-770, Berlin, Heidelberg.
- Wolberg, W., Mangasarian, O., Street, N. ve Street, W. 1993. Web Sitesi: <https://doi.org/10.24432/C5DW2B>, Erişim Tarihi: 18.08.2025.
- Wright, S. 1921. Correlation and causation. *Journal of agricultural research*, 20(7), 557-585.
- Xiao, H., Rasul, K. ve Vollgraf, R. 2017. Web Sitesi: <https://arxiv.org/abs/1708.07747>, Erişim Tarihi: 18.08.2025.
- Xu, J., Li, Z., Du, B., Zhang, M. ve Liu, J. 2020. Reluplex made more practical: Leaky ReLU, 2020 IEEE Symposium on Computers and Communications (ISCC), 7-10 Temmuz, 1-7, Rennes, Fransa.
- Yao, Y., Iyiola, O. S. ve Shehu, Y. 2022. Subgradient extragradient method with double inertial steps for variational inequalities. *Journal of Scientific Computing*, 90(2), 71.
- Zeidler, E. 2013. *Nonlinear functional analysis and its applications: II/B: nonlinear monotone operators*. Springer Science & Business Media, 741, New York.