

T.C.
İNÖNÜ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

ADAPTİF ÖĞRENME ORANI VE KARMAŞIK SAYILARIN MAKİNE
ÖĞRENMESİNE ETKİSİ

DOKTORA TEZİ

Hülya SAYGILI

Bilgisayar Mühendisliği Anabilim Dalı

Tez Danışmanı: Prof. Dr. Ali KARCI

KASIM 2025

T.C
İNÖNÜ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

ADAPTİF ÖĞRENME ORANI VE KARMAŞIK SAYILARIN MAKİNE
ÖĞRENMESİNE ETKİSİ

DOKTORA TEZİ

Hülya SAYGILI
(36213619002)

Bilgisayar Mühendisliği Anabilim Dalı

Tez Danışmanı: Prof. Dr. Ali KARCI

KASIM 2025

İNÖNÜ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜ

Adaptif Öğrenme Oranı ve Karmaşık Sayıların Makine Öğrenmesine Etkisi

Hülya SAYGILI
Doktora Tezi

TEŞEKKÜR VE ÖNSÖZ

Bu tez çalışmasının her aşamasında bilgi, tecrübe ve desteğini esirgemedi beni yönlendiren, her konuda yol gösterici olan değerli danışman hocam Sayın Prof. Dr. Ali KARCI'ya,

Çalışmalarım süresince desteğini ve teşvik edici yönlendirmelerini esirgemeyen, Tez İzleme Komitesinde yer alan Prof. Dr. M. Fatih TALU ve Doç. Dr. Kazım HANBAY hocalarıma,

Doktora sürecim boyunca birlikte fikir alışverişinde bulunduğum, ortak çalışmalar yaptığım değerli arkadaşım Meral KARAKURT'a,

Hayatım boyunca beni her zaman destekleyen biricik anneme, babama ve kardeşlerime, çalışmalarım bana manevi güç veren kızım Elif'e ve doktora yapmam için beni teşvik eden oğlum Eyüp Kerem'e

Sonsuz teşekkürlerimi sunarım...

ONUR SÖZÜ

Doktora olarak sunduğum “Adaptif Öğrenme Oranı ve Karmaşık Sayıların Makine Öğrenmesine Etkisi” başlıklı bu çalışmanın bilimsel ahlak ve geleneklere aykırı düşecek bir yardıma başvurmaksızın tarafımdan yazıldığına ve yararlandığım bütün kaynakların hem metin içinde hem de kaynakçada yöntemine uygun biçimde gösterilenlerden oluştuğunu belirtir, bunu onurumla doğrularım.

Hülya SAYGILI



İÇİNDEKİLER

TEŞEKKÜR VE ÖNSÖZ	i
ONUR SÖZÜ.....	ii
İÇİNDEKİLER	iii
ÇİZELGELER DİZİNİ	v
ŞEKİLLER DİZİNİ	vi
SEMBOLLER VE KISALTMALAR.....	viii
ÖZET.....	ix
ABSTRACT	x
1.1 Amaç ve Kapsam.....	2
1.2 Literatür Çalışması	3
1.3 Araştırma Probleminin Tanımı.....	9
1.4 Tezde Yapılan Çalışmalar ve Literatüre Katkı.....	9
1.5 Tezin Yapısı	10
2. YAPAY SİNİR AĞLARI, ÖĞRENME KATSAYISI VE PROBLEMLERİ..	12
2.1 Yapay Sinir Ağları (YSA)	12
2.2 Öğrenme Katsayısı ve Problemleri.....	13
2.2.1 Sabit öğrenme katsayısı	14
2.2.2 Öğrenme çizelgesi	14
2.2.2.1 Adım azaltması	15
2.2.2.2 Üssel azalma	15
2.2.2.3 Kosinüs azaltma.....	15
2.2.2.4 Polinomsal azalma	16
2.3 Adaptif Öğrenme Katsayısı Kullanan Yöntemler	16
2.3.1 Adagrad yöntemi	16
2.3.2 Adadelta yöntemi.....	17
2.3.3 ADAM yöntemi.....	17
2.3.4 KarcıFANN (Karcı kesir dereceli yapay sinir ağları) yöntemi	18
3. KARMAŞIK DEĞERLİ SİNİR AĞLARI (CVNN).....	21
3.1 CVNN Türleri	21
3.1.1 Split-CVNN.....	21
3.1.2 Tam karmaşık CVNN	22
3.2 CVNN’lerde Gradyan Tabanlı Öğrenme	22
3.2.1 CVNN’lerde ileri besleme	22
3.2.2 CVNN’lerde geri besleme.....	22
4. KARCIFANN YÖNTEMİNİN KARMAŞIK SAYILARLA ANALİZİ.....	24
4.1 KarcıFANN Yönteminde Karmaşık Durumların Tanımı	24
4.2 Karcı Kesir Dereceli Türev ile Karmaşık Sayılar Arasındaki İlişki.....	24
4.3 Karmaşık Değerler ile KarcıFANN Yöntemi	26
5. UYGULAMALARDA KULLANILAN VERİ SETLERİ	30
6. UYGULAMALAR VE SONUÇLAR	31
6.1 Karcı Kesir Dereceli YSA (KarcıFANN) yönteminin Reel Sayılar ile Uygulanması	31
6.2 KarcıFANN Yönteminde Kompleks Değerlerin Uygulanması	33
6.3 Karcı kesir dereceli YSA (KarcıFANN) ile öğrenme katsayısı, aşırı uyum ve yetersiz uyum sorunlarını çözme	46
6.3.1 SGD yönteminin tüm veri setlerinde yetersiz kaldığı durumlar	60
6.3.2 Önerilen yönteminin genelleme ve aşırı uyum açısından değerlendirilmesi	61

6.3.2.1 Modelin genelleme yeteneđi üzerinden deđerlendirme	61
6.3.2.2 Öğrenememe(yetersiz uyum) açısından deđerlendirme	62
6.4 KarcıFANN Yönteminin Yakınsama Kabiliyetinin İncelenmesi.....	63
6.5 KarcıFANN Yönteminin Farklı Alfa Deđerleri ile MAE, MSE ve RMSE Hata Fonksiyonlarının Performanslarının Deđerlendirilmesi	71
7. SONUÇ VE ÖNERİLER	76
KAYNAKLAR	81



ÇİZELGELER DİZİNİ

Çizelge 6. 1: Digits ve MNIST Rotation 45° veri setlerinde KarcıFANN yönteminin farklı alfa değerlerindeki doğruluk ve hata fonksiyonu değerleri.	32
Çizelge 6. 2: Digits veri setinde KarcıFANN yöntemi ile Reel, Kompleks ve Abs modellerine ait farklı alfa değerlerindeki Doğruluk ve Hata fonksiyonu değerleri.	37
Çizelge 6. 3: MNIST Rotation 45° veri setinde KarcıFANN yöntemi ile Reel, Kompleks ve Abs modellerine ait farklı alfa değerlerindeki Doğruluk ve Hata fonksiyonu değerleri.	39
Çizelge 6. 4: 2BFT-Digits veri setinde KarcıFANN yöntemi ile Reel, Kompleks ve Abs modellerine ait farklı alfa değerlerindeki Doğruluk ve Hata fonksiyonu değerleri.	41
Çizelge 6. 5: 2BFT- MNIST Rotation 45° veri setinde KarcıFANN yöntemi ile Reel, Kompleks ve Abs modellerine ait farklı alfa değerlerindeki Doğruluk ve Hata fonksiyonu değerleri.	43
Çizelge 6. 6: Kuzushiji_MNIST veri setinde KarcıFANN ve Klasik NN yöntemi için alfa değerlerindeki değişimlere göre elde edilen Doğruluk ve MSE değerleri.	49
Çizelge 6. 7: Gina_prior2 veri setinde KarcıFANN ve Klasik YSA yöntemi için alfa parametre değerlerindeki değişimlere göre elde edilen Doğruluk ve MSE değerleri.	53
Çizelge 6. 8: SignMNIST veri setinde KarcıFANN yöntemi için alfa parametre değerlerindeki değişimlere göre elde edilen doğruluk ve hata değerleri.	57
Çizelge 6. 9: Alfa ve Öğrenme katsayısının Tüm Veri Setlerinin Sınıflandırılmasında Elde Edilen Test Sonuçları	61
Çizelge 6. 10: KarcıFANN, SGD, ADAM ve Momentumlu GD yöntemleriyle farklı alfa ve öğrenme katsayı değerlerinde elde edilen MSE değerlerinin karşılaştırılması.	66
Çizelge 6. 11: KarcıFANN Yöntemi ve RMSE Hata Fonksiyonu Kullanılarak Elde Edilen Doğruluk (Acc) ve RMSE Değerleri.	72
Çizelge 6. 12: KarcıFANN Yöntemi ve MSE Hata Fonksiyonu Kullanılarak Elde Edilen Doğruluk (Acc) ve MSE Değerleri.	73
Çizelge 6. 13: KarcıFANN Yöntemi ve MAE Hata Fonksiyonu Kullanılarak Elde Edilen Doğruluk (Acc) ve MAE Değerleri.	74

ŞEKİLLER DİZİNİ

Şekil 4. 1: Karmaşık sayıların gösterimi.....	24
Şekil 6. 1: a) Digits b) MNIST Rotation 45° veri setlerinin Sınıflandırılmasında Kullanılan Ağ Mimarisi.....	31
Şekil 6. 2: Digits veri setinin 1.2 alfa değeri için KarcıFANN yönteminin a) doğruluk ve b) hata grafikleri.....	33
Şekil 6. 3: MNIST Rotation 45° veri setinin 0.9 alfa değeri için KarcıFANN yönteminin a) doğruluk ve b) hata grafikleri.	33
Şekil 6. 4: a) Digits, b) 2BFT-Digits veri setlerinin sınıflandırılmasında kullanılan model mimarisi.	34
Şekil 6. 5: a) Mnist Rotation 45, b) 2BFT-Mnist Rotation 45 veri setlerinin sınıflandırılmasında kullanılan model mimarisi.	35
Şekil 6. 6: Digits veri setinin 9/8 alfa değeri için KarcıFANN yönteminin a) doğruluk ve b) hata grafikleri	44
Şekil 6. 7: 2BFT- Digits veri setinin 9/8 alfa değeri için KarcıFANN yönteminin a) doğruluk ve b) hata grafikleri	45
Şekil 6. 8: MNIST Rotation 45° veri setinin 9/8 alfa değeri için KarcıFANN yönteminin a) Doğruluk ve b) Hata grafikleri	45
Şekil 6. 9: 2BFT- MNIST Rotation 45° veri setinin 9/8 alfa değeri için KarcıFANN yönteminin a) doğruluk ve b) hata grafikleri	46
Şekil 6. 10: Veri Setinin Sınıflandırılmasında Kullanılan Ağ Modelleri a) Gina_prior2 ve Kuzushiji-MNIST veri setleri b) Sign-MNIST veri seti.	47
Şekil 6. 11: Kuzushiji_Mnist veri setinde alfa ve öğrenme katsayılarının 1.1 değeri için KarcıFANN ve Klasik YSA yönteminin a) Doğruluk ve b) MSE grafikleri.	51
Şekil 6. 12: Kuzushiji_Mnist veri setinde alfa ve öğrenme katsayılarının 2.8 değeri için KarcıFANN ve Klasik YSA Yönteminin a) Doğruluk ve b) MSE grafikleri.	51
Şekil 6. 13: Kuzushiji_Mnist veri setinde alfa ve öğrenme katsayılarının 3.5 değeri için KarcıFANN ve Klasik YSA Yönteminin a) Doğruluk ve b) MSE grafikleri.	52
Şekil 6. 14: GinaPrior2 veri setinde alfa ve öğrenme katsayılarının 0.9 değeri için KarcıFANN ve Sgd yöntemlerinin a) Doğruluk ve b) MSE grafikleri.	55
Şekil 6. 15: GinaPrior2 veri setinde alfa ve öğrenme katsayılarının 2.9 değeri için KarcıFANN ve SGD Yöntemlerinin a) Doğruluk ve b) MSE Grafikleri.	55
Şekil 6. 16: GinaPrior2 veri setinde Alfa ve öğrenme katsayılarının 3.4 değeri için KarcıFANN ve SGD yöntemlerinin a) Doğruluk ve b) MSE Grafikleri.	56
Şekil 6. 17: SignMNIST veri setinde alfa ve öğrenme katsayılarının 0.8 değeri için KarcıFANN ve Klasik YSA yöntemlerinin a) Doğruluk ve b) MSE Grafikleri.	59
Şekil 6. 18: SignMNIST veri setinde alfa ve öğrenme katsayılarının 1.1 değeri için KarcıFANN ve Klasik YSA yöntemlerinin a) Doğruluk ve b) MSE Grafikleri.	59

Şekil 6. 19: SignMNIST veri setinde alfa ve öğrenme katsayılarının 3.5 değeri için KarcıFANN ve Klasik YSA yöntemlerinin a) Doğruluk ve b) MSE Grafikleri.	60
Şekil 6. 20: Kuzushiji-MNIST veri setinde alfa ve öğrenme katsayılarının 3.3 değeri için KarcıFANN ve Klasik YSA yöntemlerinin a) Doğruluk ve b) MSE Grafikleri.	63
Şekil 6. 21: Gina_prior2 veri setinde alfa ve öğrenme katsayılarının 3.1 değeri için KarcıFANN ve Klasik YSA yöntemlerinin a) Doğruluk ve b) MSE Grafikleri.	63
Şekil 6. 22: XOR probleminin çözümünde kullanılan YSA modeli.	64
Şekil 6. 23: Momentum GD yöntemine ait 1.4 öğrenme katsayısı için ağırlık değişimleri... ..	68
Şekil 6. 24: ADAM yöntemine ait 1.4 öğrenme katsayısı için ağırlık değişimleri.	69
Şekil 6. 25: SGD yöntemine ait 1.4 öğrenme katsayısı için ağırlık değişimleri.	70
Şekil 6. 26: KarcıFANN yöntemine ait 1.4 öğrenme katsayısı için ağırlık değişimleri.	71
Şekil 6. 27: Ginaprior2 veri setinin farklı hata fonksiyonları ile sınıflandırılması sonucu alfa 0.9 değeri için elde edilen en başarılı (a) Doğruluk ve (b) Hata grafikleri..	75

SEMBOLLER VE KISALTMALAR

YSA	: Yapay Sinir Ağları
ENN	: Elman Sinir Ağı
PSO	: Parçacık Sürü Optimizasyonu
SGD	: Stokastik Gradyan İniş
GD	: Gradyan İniş
MSE	: Ortalama karesel hata
MAE	: Ortalama mutlak hata
RMSE	: Ortalama karesel hatanın karekökü
KarcıFANN	: Karcı Kesir Dereceli Yapay Sinir Ağı
ADAM	: Uyarlanabilir Moment Tahmini

ÖZET

ADAPTİF ÖĞRENME ORANI VE KARMAŞIK SAYILARIN MAKİNE ÖĞRENMESİNE ETKİSİ

Doktora Tezi

Hülya SAYGILI

İnönü Üniversitesi
Fen Bilimleri Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı

89+X sayfa

2025

Danışman: Prof. Dr. Ali KARCI

Öğrenme katsayısı, yapay sinir ağlarının eğitim performansını doğrudan etkileyen temel hiperparametrelerden biridir. Bu parametrenin uygun seçilmemesi, ağırlık yakınsama hızını, genelleme kabiliyetini ve eğitim sürecini olumsuz yönde etkilemektedir. Literatürde bu problemin çözümü için adaptif öğrenme katsayıları, hata temelli dinamik ayarlamalar ve katman bazlı yöntemler önerilmiş olmakla birlikte kesin ve genel bir çözüm geliştirilememiştir.

Sabit bir öğrenme katsayısının kullanımından kaynaklanan sınırlamaları gidermek amacıyla KarcıFANN yöntemi kullanılmaktadır. KarcıFANN yöntemi, kesir dereceli türev yapısı nedeniyle bazı durumlarda karmaşık değerler üretmektedir. Eğitim sürecinde reel bileşenlerin işlenmesi, bilgi kaybına ve optimizasyon kararsızlıklarına yol açabilmektedir. Bu tezde, YSA’larda karşılaşılan bu problemleri çözmek amacıyla karmaşık değer kullanımının öğrenmeye etkisi incelenmektedir. Bu doğrultuda reel ve sanal bileşenleri birlikte işleyen bir yöntem önerilmektedir. Digits, MNIST Rotation 45°, 2BFT-Digits ve 2BFT-MNIST Rotation 45° veri setleri kullanılarak yapılan sınıflandırma sürecinde; reel değerlerin kullanıldığı Reel model, reel ve sanal bileşenlerin birlikte kullanıldığı Kompleks model ile karmaşık değerlerin genlik bilgisini değerlendiren Abs modeli kullanılmıştır.

Elde edilen sonuçlar, Kompleks modelin en yüksek performansa sahip olduğunu ve Reel modelin frekans dönüşümlü verilerde düşük performans sergilediğini göstermiştir. Ayrıca, Abs modelinde faz bilgisi kullanılmadığından başarımın düştüğü görülmüştür. Bu bulgular, karmaşık sayılarla hesaplamaların, KarcıFANN’ın öğrenme performansını belirgin biçimde artırdığını kanıtlamıştır.

KarcıFANN ve klasik YSA yöntemlerinin performansları, Kuzushiji-MNIST, GinaPrior2 ve Sign-MNIST veri setlerinin sınıflandırılmasında karşılaştırmalı olarak değerlendirilmiştir. Ayrıca, SGD, Momentumlu GD ve Adam optimizasyon yöntemleri ile KarcıFANN’ın yakınsama davranışları XOR problemi üzerinde incelenmiştir. Son olarak KarcıFANN’ın farklı hata fonksiyonları (MSE, RMSE ve MAE) ile kullanımının modellerin performansına etkileri incelenmiştir.

Anahtar Kelimeler: KarcıFANN, yapay sinir ağları, SGD, Momentumlu GD, ADAM

ABSTRACT

ADAPTIVE LEARNING RATE AND THE EFFECT OF COMPLEX NUMBERS ON MACHINE LEARNING

Inonu University
Graduate School of Nature and Applied Sciences
Department of Computer Engineering

89+X pages

2025

Supervisor: Prof. Dr. Ali KARCI

The learning coefficient is one of the fundamental hyperparameters that directly affects the training performance of artificial neural networks. An improper selection of this parameter can negatively impact the network's convergence rate, generalization ability, and overall training process. Although adaptive learning coefficients, error-based dynamic adjustments, and layer-based methods have been proposed in the literature to solve this problem, a precise and universal solution has not yet been developed.

To overcome the limitations arising from the use of a fixed learning coefficient, the KarcıFANN method is employed. Due to its fractional-order derivative structure, the KarcıFANN method can produce complex values in certain cases. Processing only the real components during training can lead to information loss and optimization instabilities. In this thesis, the effect of using complex values on learning in artificial neural networks (ANNs) is examined. Accordingly, a method that processes real and imaginary components together is proposed. In the classification process carried out using the Digits, MNIST Rotation 45°, 2BFT-Digits, and 2BFT-MNIST Rotation 45° datasets, three models were used: the Real model, which uses only real values, the Complex model, which uses both real and imaginary components together, and the Abs model, which evaluates only the magnitude information of the complex values.

The results showed that the Complex model achieved the highest performance, while the Real model exhibited lower performance on frequency-transformed data. Moreover, since the Abs model does not utilize phase information, its performance was found to decrease. These findings demonstrate that computations using complex numbers significantly enhance the learning performance of the KarcıFANN method.

The performances of KarcıFANN and classical ANN methods were comparatively evaluated on the Kuzushiji-MNIST, GinaPrior2, and Sign-MNIST datasets. In addition, the convergence behaviors of KarcıFANN with SGD, Momentum-based GD, and Adam optimization methods were examined using the XOR problem. Finally, the effects of using different loss functions (MSE, RMSE, and MAE) with KarcıFANN on model performance were analyzed.

Keywords: KarcıFANN, artificial neural networks, SGD, momentum-based GD, ADAM

1. GİRİŞ

Yapay zeka, insan zekasını taklit ederek bilgisayar sistemlerinin karmaşık problemleri çözme, akıl yürütme ve doğal dil işleme gibi görevleri makineler vasıtasıyla yerine getirmeyi amaçlayan bir bilim dalıdır. Günümüzde veri miktarının hızlı artışı ve bu verilerin işlenmesinin önem kazanması, yapay zekanın bir alt dalı olan makine öğrenmesi disiplinlerine olan ilgiyi önemli ölçüde artırmıştır. Makine öğrenmesi, istatistiksel modeller ve hesaplama teknikleri aracılığıyla verilerden örüntüler çıkarmayı, tahminlerde bulunmayı ve öğrenme performansını artırmayı amaçlayan yöntemler bütünüdür.

Makine öğrenmesi yöntemin alt dalı olan yapay sinir ağları (YSA) insan beynindeki sinir hücrelerinin yapısından esinlenerek geliştirilen matematiksel hesaplama yöntemidir. YSA; görüntü işleme, doğal dil işleme, tahmin, sınıflandırma gibi özellikle karmaşık ve doğrusal olmayan problemlerin çözümünde kullanılmaktadır.

Makine öğrenmesi algoritmalarının başarımı, doğru model seçimi, hiperparametrelerin ayarlaması ve eğitim sürecinde kullanılan optimizasyon tekniklerine bağlıdır. YSA'ların eğitiminde gradyan iniş tabanlı yöntemler yaygın olarak kullanılmaktadır. Gradyan iniş tabanlı algoritmalarla tasarlanan YSA modeline ait ağırlıkların güncellemesinde öğrenme katsayısı parametresinden yararlanılmaktadır. Ağın eğitim sürecinde öğrenme katsayısının belirlenmesi büyük bir problem olmaktadır. Öğrenme katsayısının çok büyük seçilmesi, hata fonksiyonunun değerinin artmasına, öğrenme sürecinin başarısız olmasına, salınımlara ve modelin genelleme yeteneğinin zayıflamasına neden olmaktadır. Öğrenme katsayısının çok küçük seçilmesi de eğitim sürecinin yavaşlamasına ve modelin yerel minimum noktalarda takılı kalarak öğrenme performansının düşmesine neden olmaktadır. Ayrıca, öğrenme katsayısının çok küçük seçilmesi büyük veri setlerinde hesaplama maliyetini artırmaktadır.

YSA'ların matematiksel yapısı, ele alınan problemin gereksinimlerine bağlı olarak reel ya da karmaşık sayılarla çalışabilecek biçimde tasarlanmaktadır. Reel değerli YSA'lar (RVNN- Reel Valued Neural Networks), giriş, çıkış ve ağ parametrelerinin tamamı reel

sayılardan oluşmaktadır. Bu durum RVNN uygulamalarının daha basit bir yapıya sahip olmasını ve hesaplama maliyetinin daha düşük olmasını sağlamaktadır (Abdalla, 2023). Fakat elektromanyetik, sinyal ve görüntü işleme gibi bir çok uygulamalar doğası gereği karmaşık sayılar ile ifade edilmektedir. Karmaşık değerli YSA'lar (CVNN- Complex Valued Neural Networks) girişler, ağ parametreleri ve çıkışlar karmaşık sayılar ile temsil edilmektedir. Karmaşık sayıların yapısı, modellerin eğitim süresinin daha uzun olmasına ve hesaplama maliyetinin artmasına neden olmaktadır. CVNN'ler görüntü işleme, robotik, biyoinformatik, radar ve kuantum hesaplamalar gibi alanlarda RVNN'lerden daha iyi performans göstermektedir (Nitta, 2003).

Bu tez çalışmasının ana konusu kesir dereceli türevin negatif tabanlı ifadelerde kullanılmasıyla ortaya çıkan karmaşık sayıların modellerin eğitim sürecine etkilerini incelemektir. Ayrıca, KarcıFANN yönteminin farklı optimizasyon yöntemleri, farklı veri setleri ve farklı hata fonksiyonlarıyla kullanımının başarıma etkileri analiz edilmiştir.

Tez kapsamındaki gerçekleştirilen tüm deneysel analizler, açık kaynak tabanlı Anaconda geliştirme platformunun Jupyter Notebook bileşeni kullanılarak ve Python dili ile programlanarak yürütülmüştür.

1.1 Amaç ve Kapsam

YSA, büyük verilerin analizi, örüntü tanıma, doğal dil işleme ve ses tanıma gibi birçok alanda başarılı bir şekilde kullanıldığından makine öğrenmesi uygulamalarında önemli bir yere sahiptir. YSA'ların eğitimi sırasında, ağırlık parametrelerini optimize etmek amacıyla gradyan tabanlı optimizasyon yöntemleri yaygın olarak kullanılmaktadır. Bu yöntemlerde ağırlıkların güncelleme işlemlerinde reel sayılar kullanılmakta ve modellerin üretebileceği sanal değerler göz ardı edilmektedir.

Bu çalışmanın temel amacı, kesir dereceli türevin negatif tabanlı ifadelerde ortaya çıkardığı karmaşık sayıların YSA'ların eğitim sürecine etkilerini incelemektir. Bu kapsamda, Reel, Karmaşık ve Abs modellerinin kullanıldığı farklı güncelleme yöntemleri karşılaştırılmıştır. Ayrıca, KarcıFANN yöntemi ile klasik YSA'larda karşılaşılan ezberleme,

yetersiz öğrenme, gradyan patlaması ve gradyan kaybolması gibi problemlerin çözülmesi amacıyla çeşitli çalışmalar yapılmıştır.

1.2 Literatür Çalışması

Bu bölümde, karmaşık değerli YSA'lar, gradyan tabanlı optimizasyon yöntemleri ile öğrenme katsayısı problemlerine çözümler ve farklı aktivasyon fonksiyonlarının modelin öğrenme performansına etkisi üzerine yapılan literatür çalışmaları ele alınmıştır.

Nitta (2003) çalışmasında, XOR problemi gibi doğrusal olarak ayrıştırılamayan problemlerin çözümü, tek bir reel değerli nöron ile çözülemezken karmaşık değerli tek bir nöron problemin çözümünün başarılı şekilde gerçekleştirdiğini göstermiştir (Nitta, 2003).

Bassey ve diğ. (2021) tarafından gerçekleştirilen çalışmada, karmaşık değerli yapay sinir ağlarının (Complex-Valued Neural Networks, CVNNs) kuramsal temelleri ele alınmıştır. Karmaşık sayıların genlik ve faz bileşenlerini içeren yapıları, sinyal işleme, radar ve iletişim gibi faz duyarlılığı gerektiren uygulama alanlarında etkin biçimde kullanılabildiğini göstermişlerdir. Öğrenme süreçlerinde gradyan hesaplamalarının Wirtinger hesaplama yöntemi aracılığıyla gerçekleştirildiği belirtilmiştir (Bassey ve diğ., 2021).

Smith (2023), CVNN'in sinyal işleme ve sinyal yorumlama alanlarındaki uygulamalarını ele aldığı çalışmasında, bu yapılar için PyTorch tabanlı modüler bir geliştirme paketi sunmuş; lineer ve konvolüsyonel katmanların yanı sıra aktivasyon ve normalizasyon işlemlerinde de karmaşık değerler kullanmıştır (Smith, 2023).

Abdalla (2023) çalışmasında, karmaşık değerli sinir ağlarının (CVNN), hem reel hem de kompleks verilerin işlenmesinde, geleneksel reel değerli sinir ağlarına (RVNN) göre daha yüksek temsil kapasitesi, daha hızlı öğrenme ve daha güçlü genelleme yeteneği sunduğunu belirtmektedir. CVNN'lerin, faz ve genlik gibi bileşenleri doğrudan modelleyebilmesi sayesinde, biyolojik sinir yapılarıyla daha uyumlu bir yapı ortaya koyduğunu ifade etmektedir (Abdalla, 2023).

Barrachina ve diğ. (2023) çalışmalarında, karmaşık değerli sinir ağlarının (CVNN) matematiksel temellerini detaylandırarak Wirtinger hesaplama yöntemi, karmaşık geri yayılım, ağırlık başlatma ve karmaşık aktivasyon fonksiyonları gibi temel bileşenleri açıklamışlardır. Yapılan çalışmada, bu yapıların uygulamaları Python tabanlı CVNN

kütüphanesi aracılığıyla sunulmuş ve kütüphanenin kullanımına yönelik örneklerle yer verilmiştir. Hilbert dönüşümü ile kompleks düzleme aktarılan reel veriler üzerinde yapılan deneylerde, CVNN modellerinin daha yüksek sınıflandırma başarımı gösterdiği gözlemlenmiştir (Barrachina ve diğ.,2023).

Trabelsi ve diğ. (2018), karmaşık evrişim katmanları, karmaşık batch-normalization algoritmaları ve karmaşık ağlar için özel ağırlık başlatma stratejileri geliştirerek ileri beslemeli karmaşık evrişimsel ağlarda ve karmaşık değerli LSTM modellerinde test etmişlerdir. Önerdikleri yöntem ile karmaşık değerli derin ağların ses ve sinyal işleme gibi alanlarda reel değerli modellere göre daha başarılı performans sergilediğini belirtmişlerdir (Trabelsi ve diğ., 2018).

Lee ve diğ. (2022), karmaşık değerli derin sinir ağları üzerine kapsamlı bir inceleme sunarak CNN(Evrişimsel Sinir Ağları), RNN (Tekrarlayan Sinir Ağları) ve karmaşık ResNet gibi farklı mimari türlerini karşılaştırmalı olarak incelemişlerdir. Çalışmada, farklı uygulama alanlarında CVNN'lerin sağladığı avantajları ayrıntılı biçimde sunmuşlardır (Lee ve diğ., 2022).

Smith ve Topin (2018), yüksek öğrenme katsayıları kullanarak YSA'ların eğitimi aşamasında süper-yakınsama yöntemini önermişlerdir. Tek bir öğrenme katsayısı döngüsü ve büyük bir maksimum öğrenme katsayısı ile ağların çok daha hızlı eğitilebileceğini göstermişlerdir. CIFAR-10/100, MNIST ve ImageNet veri kümeleri üzerinde; ResNet, Wide ResNet ve DenseNet mimarileriyle gerçekleştirdikleri deneylerde, büyük öğrenme katsayısı kullanıldığında ağırlıkların azaltılması gerektiğini ve standart yöntemlere kıyasla daha yüksek başarı elde edildiğini rapor etmişlerdir (Smith ve Topin ,2018).

Igiri ve diğ. (2015), YSA modelinin eğitim sürecinde öğrenme katsayısının etkisini incelemek amacıyla yaptıkları çalışmada 15 farklı öğrenme katsayısı kullanılmışlardır. Uygun öğrenme katsayısı kullanımının modelin performansını önemli ölçüde etkilediğini yaptıkları deneysel çalışmalar sonucunda ortaya koymuşlardır (Igiri ve diğ. ,2015) .

Smith (2017), YSA'da kullanılan öğrenme katsayısını belirli bir alt ve üst değer arasında döngüsel olarak değiştiren daha az yineleme içeren bir yöntem önermiştir. Öğrenme katsayısı için en uygun değeri ve zamanlamayı deneme-yanılma yoluyla bulmayı ortadan kaldırmayı amaçladığı çalışmada, CIFAR10/100 veri kümelerinde ResNet, Stochastic Depth ağları ve DenseNet, ImageNet veri kümesinde ise AlexNet ve GoogLeNet mimarileri

üzerinde deneysel çalışmalar yapmıştır. Döngüsel öğrenme katsayısının performansı olumlu bir şekilde etkilediğini savunmuştur (Smith, 2017).

Smith (2018), yaptığı çalışmada öğrenme sürecinde hiperparametreleri ayarlamak için, döngüsel öğrenme katsayıları kullanarak yetersiz öğrenme ve aşırı öğrenme problemleri arasında optimal bir dengeye ulaşmak için öğrenme katsayısının nasıl artırılıp azaltılması gerektiği detaylı olarak incelemiştir. Deneysel çalışmalar ile kullanılan veri kümeleri ve ağ mimarilerine özgü düzenlemelerin model başarımı açısından kritik öneme sahip olduğunu ortaya koymuştur (Smith, 2018).

Li ve diğ (2018), “küçük bir öğrenme katsayısı daha hızlı eğitim ve daha iyi test performansı sağlarken, büyük bir öğrenme katsayısı ise eğitimden sonra daha güçlü bir genelleme yeteneği kazandırır” şeklindeki gözlemi açıklamaya yönelik bir çalışma gerçekleştirmişlerdir. CIFAR-10 veri seti ile önerilen yöntem test edilmiş ve küçük bir öğrenme katsayısının model tarafından hemen hafızaya alınabildiğini, büyük öğrenme katsayısına sahip modelin ise ezberlemek yerine genelleştirmeye yönelttiğini göstermişlerdir (Li ve diğ, 2018).

You ve diğ (2019), YSA modelini eğitmek amacıyla başlangıçta büyük öğrenme katsayısı ile başlayarak zamanla öğrenme katsayısının azaltıldığı LR Decay yönteminin çalışma mantığını ve modern YSA’ların eğitiminde yetersiz kaldığı durumları ortaya koymak için deneysel çalışmalar yapmışlardır. LR Decay 'in başlangıçta büyük öğrenme katsayısı kullanarak ağır gürültülü verileri ezberlemesini baskıladığını ve öğrenme katsayısı azaldıkça karmaşık desenlerin öğrenilmesinin kolaylaştırdığını gerçek veriler üzerinde yaptıkları testler ile ortaya koymuşlardır (You ve diğ, 2019).

Iranmanesh ve Mahdavi (2009), diferansiyel uyarlamalı öğrenme katsayısı kullanarak DALRM adını verdikleri yüksek hızlı bir öğrenme yöntemi önermişlerdir. XOR probleminin çözümü için farklı ağırlık başlatma yöntemleri ile karşılaştırmışlar, yöntemin yerel minimumlardan kaçabildiğini, mutlak minimuma yakınsayabildiği ve öğrenme hızının arttığını göstermişlerdir (Iranmanesh ve Mahdavi, 2009).

Wilson ve Martinez (2001), gradyan tabanlı optimizasyon algoritmalarında öğrenme katsayısının doğruluk, yakınsama hızı ve genelleme üzerine etkisini incelemek amacıyla yapılan çalışmada öğrenme katsayısına 0.00001-100 aralığında değerler vererek, uygun öğrenme katsayısı ile çalışmanın başarıma etkisini incelemişlerdir. Düşük öğrenme katsayısı

kullanımının büyük ve karmaşık problemlerde genelleme doğruluğunu artırdığı fakat daha fazla eğitim süresi gerektirdiğini yaptıkları deneysel çalışmalarla göstermektedirler. Tüm eğitim boyunca tek bir öğrenme katsayısı kullanmak yerine başlangıçta yüksek bir öğrenme katsayısı belirleyip daha sonra kademeli olarak azaltılması gerektiğini savunmuşlardır (Wilson ve Martinez, 2001).

Takase ve diğ. (2018), uyarlanabilir öğrenme katsayısı ile eğitim hata fonksiyon değerini azaltmak amacıyla ağırlıkların güncellenmesi aşamasında Breadth-first beam search yöntemi kullanmışlardır. Hata değerini azaltmak için öğrenme katsayısını artırıp azaltan bu yaklaşım her ağırlık için aynı öğrenme katsayısını kullanmaktadır. Farklı veri kümeleri üzerinde yaptıkları deneysel çalışmalar sonucunda test hata oranlarında azalmalar meydana geldiğini, yerel minimumlara sapma riskinin azaldığını ve daha optimal çözümlere ulaşıldığını göstermişlerdir (Takase ve diğ., 2018).

Sheel ve diğ. (2007), YSA'ların öğrenme katsayısını belirlemek için sezgisel (heuristik) yöntemler kullanarak ağır diğer yöntemlere göre daha hızlı eğitildiğini ve daha düşük hata değerleri ürettiğini göstermişlerdir (Sheel ve diğ., 2007).

Üstün ve Yıldız (2009), yaptıkları çalışmada, öğrenme parametrelerinden olan öğrenme ve momentum katsayılarını genetik algoritma ile optimize ederek klasik öğrenme yöntemlerine göre daha başarılı sonuçlar ürettiğini gözlemlemişlerdir (Üstün ve Yıldız, 2009).

Wei ve diğ. (2022), öğrenme katsayısını belirlemek için Parçacık Sürü Optimizasyonu (PSO) ile genetik algoritmaların mutasyon işlevini birleştirerek bir DPLRs yöntemini önermişlerdir. CIFAR-10 ve ImageNet veri kümeleri üzerinde AlexNet, VGG16 ve ResNet18 veri modelleri ile DPLRs'nin çalışma süresi ve performans bakımından SGDR ve ADAM yöntemlerine yakın sonuçlar elde etmişlerdir (Wei ve diğ., 2022).

Zheng ve diğ. (2023), YSA'daki salınımları azaltmak ve yakınsama hızını artırmak için Gauss fonksiyonuna dayalı uyarlanabilir bir bileşik öğrenme katsayısını kullanan algoritma önermişlerdir. Bu yöntemin diğer yöntemlere kıyasla doğruluğunun arttığını ve daha hızlı yakınsadığını teorik olarak kanıtlamışlardır (Zheng ve diğ., 2023).

Karakurt ve diğ. (2022), YSA'ların optimizasyon sürecinde öğrenme katsayısının belirlenmesindeki belirsizliklerden dolayı bu katsayı yerine Karcı kesir dereceli türev kullandıkları bir yöntem önermişlerdir. XOR problemi üzerinde yapılan deneysel çalışmalar

sonucunda KarcıFANN yönteminin Klasik YSA yöntemine göre kesir derecesinin 0.8-1.8 aralığındaki değerlerde daha başarılı olduğunu göstermişlerdir (Karakurt ve diğ., 2022).

Karakurt ve diğ. (2024), KarcıFANN yöntemi ile farklı aktivasyon fonksiyonlarının modelin performansına etkisini incelemek için sigmoid, hiperbolik tanjant (tanh) ve hiperbolik tanjant sigmoid (tansig) fonksiyonlarını kullandıkları bir çalışma gerçekleştirmişlerdir. Yaptıkları deneysel çalışmalar sonucunda, KarcıFANN yönteminin sigmoid aktivasyon fonksiyonu ile daha başarılı olduğunu göstermişlerdir (Karakurt ve diğ., 2024).

Ren ve diğ. (2018), Elman sinir ağında (ENN) kullanılan geri yayılım algoritmasının düşük yakınsama hızı ve zayıf genelleme yeteneği sorunlarını çözmek amacıyla yeni bir öğrenme katsayısı yöntemi önermişlerdir. ENN algoritmasını modifiye ederek öğrenme katsayısı için yeni bir denklem geliştirmiş ve matematiğini ispatlamışlardır. Önerdikleri yöntemi ENN ve PSO algoritmasını kullanan PSO-ENN gibi yöntemlerle yakınsama hızı ve zaman tüketimi açısından karşılaştırılarak yöntemin üstünlüğünü göstermişlerdir (Ren ve diğ., 2018)

Starodub ve diğ. (2021), YSA'ların öğrenme sürecini optimize etmeyi amaçlayan bir algoritma önermişlerdir. Algoritma, geriye dönük öğrenme verilerinin analizine dayanmakta olup, ağırlık değişimlerinin dinamiklerini inceleyerek eğitim verimliliğini artırmayı amaçlamışlardır. Yapılan çalışmalar sonucunda, ağırlık öğrenme sürecinin daha verimli hale geldiğini göstermişlerdir (Starodub ve diğ., 2021).

Li ve diğ. (2009), Hopfield sinir ağı modelini, ağırlıkların zamanla değişmesini sağlayacak şekilde uyarlayarak "Ağırlık Fonksiyonlu Matrisli Ayrık Hopfield Sinir Ağı" (DHNNWFM) olarak adlandırdıkları yeni bir yöntem önermişlerdir. Kullanılan ağırlık fonksiyon matrisi negatif olmayan ve monoton artan bir yapıda tasarlayarak modelin başarılı sonuçlar ürettiğini sergilemişlerdir (Li ve diğ., 2009).

Nagel ve diğ. (2022), YSA'ların yüksek doğruluk ve daha hızlı şekilde öğrenmesini sağlamak için ağırlık ve aktivasyon fonksiyonu değerlerini sıkıştırarak nicemlemeye dayalı salınım sönümler ve yinelenmeli ağırlık dondurma adında iki yeni QAT (Nicemleme Farkındalık Eğitimi) algoritma önermişlerdir. MobileNetV2, MobileNetV3 ve EfficientNet-lite ağı modelleri üzerinde deneyler yapmışlar ve başarılı sonuçlar elde etmişlerdir (Nagel ve diğ. 2022).

Zeiler (2012), her bir parametre boyutu için ayrı ayrı öğrenme katsayısı belirleyen ve bu katsayıyı dinamik olarak güncelleyebilen yeni bir optimizasyon yöntemi olan ADADELTA algoritmasını önermiştir. ADADELTA, birinci derece gradyan bilgisine dayalı olarak SGD'ye kıyasla daha az hesaplama maliyeti gerektirmektedir. MNIST veri seti üzerinde test ettiği bu yöntemi SGD, Momentum, ADAGRAD yöntemleri ile karşılaştırmış ve daha başarılı sonuçlar ürettiğini göstermiştir (Zeiler, 2012).

Yedida ve diğ. (2021), hata fonksiyonunun Lipschitz sabitine dayalı olarak öğrenme katsayılarını dinamik olarak hesaplayan bir yöntem önermişlerdir. ResNet20 üzerinde önerdikleri yöntem ile SGD, Momentum ve ADAGRAD yöntemlerini karşılaştırarak yöntemin etkinliğini göstermişlerdir (Yedida ve diğ., 2021).

Luo ve diğ. (2019), yaptıkları çalışmada, ADAM ve AMSGRAD'ın yeni varyantları olarak önerdikleri ADABOUND ve AMSBOUND algoritmaları ile öğrenme katsayısını dinamik olarak alt ve üst sınırlarla kısıtlamışlardır. Önerdikleri yöntemleri CIFAR-10 veri kümesi üzerinde ResNet-34 modeli kullanarak hem adaptif yöntemler gibi hızlı öğrenmeyi gerçekleştirdiğini hem de SGD yöntemine benzer genelleme yapabildiğini gözlemlemişlerdir (Luo ve diğ., 2019).

Senior ve diğ. (2013), konuşma tanıma sistemlerinde kullanılan öğrenme katsayısı ayarlama yöntemlerini inceleyerek AdaGrad'ın sınırlamalarını analiz etmişlerdir. Bu sınırlamalara çözüm olarak uzun vadeli öğrenme katsayısını ayarlamak için AdaDec yöntemini önermişlerdir. AdaDec yöntemini AdaGrad yöntemi ile karşılaştırarak uygun öğrenme katsayısı ile daha yüksek doğruluk ve daha düşük hata katsayısı sağladığını göstermişlerdir (Senior ve diğ. 2013).

Terven ve diğ.(2023), derin öğrenme yöntemlerinde kullanılan hata fonksiyonların performanslarını ve faydalarını inceledikleri bir çalışma sunmuşlardır (Terven ve diğ, 2023).

Ciampiconi ve diğ. (2023) makine öğrenmesinde kullanılan hata fonksiyonlarının kullanım alanlarını ve başarılı olduğu alanları, bir çok farklı uygulama alanlarında test etmişlerdir (Ciampiconi ve diğ. 2023).

Qi ve diğ.(2020), ortalama mutlak hata (MAE) fonksiyonunun performansı inceleyerek ortalama karesel hata (MSE) fonksiyonu ile karşılaştırarak deneysel çalışmalar gerçekleştirilmiş ve bunun sonucunda MAE ile optimize edilmiş DNN tabanlı regresyonun daha iyi performansa sahip olduğunu göstermişlerdir (Qi ve diğ., 2020).

Kornblith ve diğ. (2020), görüntü sınıflandırma üzerinde farklı katmanlarda farklı hata fonksiyonu kullanmanın önemini gösteren bir çalışma sunmuşlardır (Kornblith ve diğ., 2020).

Jadon ve diğ. (2022), hata fonksiyonlarının avantaj ve dezavantajlarının farklı veri kümelerinde nasıl etkiye sahip olduğu, hangi hata fonksiyonun daha hızlı ve daha iyi yakınsama yaptığı gibi durumları incelemek amacıyla bir çalışma gerçekleştirmişlerdir (Jadon ve diğ., 2022).

1.3 Araştırma Probleminin Tanımı

Bu çalışmanın iki tane temel amacı vardır. Öğrenme oranının dış müdahaleden bağımsız hale getirmek ve ikinci amaç ise, YSA eğitim sürecinde negatif tabanlı kesir dereceli türevden kaynaklanan karmaşık sayıların modelin performansını nasıl etkilediğini belirlemektir. Sistemin ürettiği karmaşık değerlerin sanal bileşenlerini öğrenmeye dahil ederek farklı güncelleme yaklaşımları ile öğrenme sürecini incelenmektedir. Ayrıca KarciFANN yönteminin öğrenememe, ezberleme ve gradyan kaynaklı kararsızlık sorunlarına bir alternatif sunup sunmadığı değerlendirilmektedir.

1.4 Tezde Yapılan Çalışmalar ve Literatüre Katkı

Bu tez kapsamında, kesir dereceli türevin negatif tabanlı ifadelerde ortaya çıkardığı karmaşık sayıların yapay sinir ağlarının (YSA) eğitim sürecine etkileri incelenmiştir. Bu amaçla, karmaşık sayıların yalnızca reel bileşeninin dikkate alındığı, reel ve sanal bileşenlerinin birlikte değerlendirildiği ve karmaşık sayıların modülünün kullanıldığı üç farklı güncelleme yöntemi geliştirilmiş ve deneysel olarak karşılaştırılmıştır. Bu çalışmalar sonucunda, karmaşık sayıların YSA'ların doğruluk oranları, hata değerleri ve yakınsama davranışları üzerindeki etkileri kapsamlı biçimde ortaya konulmuştur. Önerilen yöntemi değerlendirmek amacıyla Digits ve MNIST Rotation 45° veri setleri ile bu veri setlerinin frekans uzayına dönüştürülmeleri ile oluşturulan 2BFT-Digits ve 2BFT-MNIST Rotation 45° kullanılarak deneysel çalışmalar gerçekleştirilmiştir. Karmaşık sayılarla yapılan bu deneyler, kesir

dereceli türev yaklaşımının negatif tabanlı ifadelerde optimizasyon sürecine ekileri oraya koymaktadır.

Tez kapsamında yapılan diğer bir çalışmada, Karcı kesir dereceli türev yaklaşımına dayanan KarcıFANN yöntemi ile modeller, sabit bir öğrenme katsayısına ihtiyaç duyulmadan optimize edilmiştir. Farklı kesir dereceli türev değerleriyle yapılan deneyler, yöntemin eğitim performansını artırdığını göstermiştir. Literatürde sabit öğrenme katsayısı kullanan yöntemlerin neden olduğu ezberleme, öğrenememe, gradyan patlaması, gradyan kaybolması gibi sorunlara çözüm olacağı düşünülmektedir. Yöntemin uygulanabilirliğini göstermek amacıyla yapılan çalışmada ise Kuzushiji_MNIST, GinaPrior2 ve SignMnist veri setlerinde farklı kesir dereceleri ile test çalışmaları yapılmıştır. KarcıFANN yönteminin yakınsama, ezberleme durumu ve öğrenme performansına etkilerini incelemek amacıyla yapılan çalışmada, XOR problemi üzerinde YSA'ların ağırlık değişimleri incelenmiş, literatürde bulunan Momentumlu GD, SGD ve ADAM optimizasyon yöntemleri ile karşılaştırılmıştır. KarcıFANN yönteminin Momentumlu GD yöntemine yakın ama gerisinde kaldığı, SGD ve ADAM yönteminden ise daha başarılı sonuçlar verdiği gösterilmiştir. Ayrıca, KarcıFANN yönteminin literatürde kullanılan RMSE, MSE ve MAE hata fonksiyonları ile kullanımı değerlendirilmiştir. Yapılan deneysel çalışmalarda, MSE hata fonksiyonu ile KarcıFANN yönteminin kullanıldığı modelin başarılı sonuçlar ürettiği gözlemlenmiştir.

1.5 Tezin Yapısı

Doktora tezi çalışması olarak hazırlanan bu çalışma yedi bölümden oluşmaktadır. İlk bölümde teze giriş yapılmış, tezin amacına, kapsamına, yenilikçi yönüne ve literatüre katkısına değinilmiştir. Ayrıca literatürde bu alanda mevcut çalışmalara yer verilmiştir.

İkinci bölümde yapay sinir ağlarına, öğrenme katsayısı kullanımına, bu kullanımın neden olduğu problemlere ve bu problemleri çözmek için geliştirilen yöntemlere değinilmiştir.

Üçüncü bölümde karmaşık değerli sinir ağları detaylı olarak incelenmiştir.

Dördüncü bölümde KarcıFANN yönteminin kesir dereceli türev ifadelerinde yer alan üs ve taban parametrelerinin reel veya karmaşık düzlemde tanımlanma koşulları sistematik

olarak incelenmiştir. Ayrıca, karmaşık değerler ile KarcıFANN uygulanma yöntemlerine değinilmiştir.

Beşinci bölümde tez çalışması kapsamında kullanılan veri setleri ve özelliklerine değinilmiştir.

Altıncı bölümde tez çalışması kapsamında geliştirilen uygulamalara ve uygulamalar sonucunda elde edilen deneysel sonuçlara yer verilmiştir.

Yedinci bölümde tez çalışması ile ilgili sonuçlara ve önerilere yer verilmiştir. Bu bölümün devamında ise kaynakça sunulmuştur.



2. YAPAY SİNİR AĞLARI, ÖĞRENME KATSAYISI VE PROBLEMLERİ

2.1 Yapay Sinir Ağları (YSA)

Makine öğrenmesi, bilgisayarların insan benzeri düşünme ve karar verme yetenekleri kazanmasını sağlayan yapay zekanın alt dalıdır. Mevcut veriden bilgi çıkarımı yapmak ve bu çıkarımdan veriler arasındaki ilişkileri bulmak, veriler arasındaki ilişkilerin tahmin etmek ve verileri sınıflandırmak gibi görevleri gerçekleştirmektedir. Makine öğrenmesi yöntemleri; etiketlenmiş veri seti kullanan denetimli öğrenme, etiketlenmemiş veriler kullanılarak denetimsiz öğrenme ve bir "ajan" adı verilen öğrencinin bir ortamda etkileşimde bulunarak öğrenmeyi gerçekleştiren pekiştirmeli öğrenme olarak üç sınıfta incelenmektedir (Coşkun ve Güleroğlu, 2021; Pirim, 2006; You ve diğ., 2023; Alam, 2023).

YSA, biyolojik sinir hücrelerinden esinlenilerek geliştirilmiş olan bilgiyi öğrenme ve işleme amacıyla yapılan matematiksel hesaplama modelidir. Bir YSA modelinin yapısı giriş, gizli ve çıkış olmak üzere üç ana katmandan oluşmaktadır. Giriş katmanı, veri girişinin yapıldığı katmandır. Gizli katman, giriş katmanından gelen verilerin ağırlıklar ile çarpılarak ve aktivasyon fonksiyonundan geçirilerek çıkış katmanına iletildiği katmandır. Ağırlıklar, girişin o nöron üzerindeki etkisini gösteren değerlerdir. Bir YSA’da birden fazla gizli katman bulunabilir ve çıkış katmanı ağın sonuçlarının üretildiği katmandır. Tek katmanlı YSA’lar, giriş ve çıkış katmanlarından oluşan ve doğrusal problemlerin çözümünde kullanılan en basit yapıdır. Daha karmaşık örüntüleri modelleyebilen çok katmanlı yapay sinir ağları; bir giriş katmanı, bir veya birden fazla gizli katman ve bir çıkış katmanından oluşan hiyerarşik bir yapıya sahiptir. İleri beslemeli ağlarda bilgi akışı yalnızca girişten çıkışa doğru ilerlerken, geri beslemeli yapılarda çıkışta elde edilen hata değeri geriye doğru iletilerek ağ parametreleri güncellenmektedir. Böylece ağırlıklar, hesaplanan hata bilgisine bağlı olarak düzenlenir ve öğrenme süreci gerçekleştirilir (Günler Pirim, 2017; Öztürk ve Şahin, 2018; Arı ve Berberler, 2017; Keskenler ve Keskenler, 2017).

Hata fonksiyonu, YSA modelinin eğitimi sonucunda elde edilen çıktı ile beklenen çıktı değer arasındaki farkı ölçmektedir. Ağırlık eğitim sırasında modelin parametrelerini optimize etmek ve eğitim sürecinin sonlandırılmasına karar vermek için kullanılmaktadır. Literatürde birçok hata fonksiyonu kullanılmaktadır. Elde edilen çıktı değeri ile gerçek değer arasındaki farkın mutlak değerinin ortalamasını hesaplayan MAE (Ortalama Mutlak Hata) fonksiyonu, karesel farkların ortalamasını hesaplayan MSE (Ortalama Karesel Hata) fonksiyonu, karesel farkların ortalamasının karekökünü hesaplayan Kök Ortalama Kare Hata (RMSE), iki olasılık dağılımı arasındaki farkı ölçen Cross Entropy gibi hata fonksiyonları yaygın bir şekilde kullanılmaktadır (Azizoğlu ve Ünsal, 2022; Koşan vd.,2019).

YSA'da nöronların davranışlarını belirleyen nöronun çıkış değerini istenilen değerler arasında sınırlamak için aktivasyon fonksiyonları kullanılmaktadır (Çelik, 2022). Sürekli ve türevi alınabilir bir fonksiyon olan Sigmoid aktivasyon fonksiyonu, giriş değeri sıfırdan küçük olması durumunda 0, diğer durumda çıkışı girişe eşit değer üreten ReLU aktivasyon fonksiyonu gibi fonksiyonlar kullanılmaktadır (Çelik, 2022; Ser ve Bati,2019). Kompleks sayılarda Denklem 2.1 ve 2.2 ile hesaplanan Complex ReLU (CReLU) aktivasyon fonksiyonu ise reel ve sanal bileşenlere ayrı ayrı uygulanmaktadır (Trabelsi, 2018).

$$z=z_r + jz_i \quad (2.1)$$

$$CReLU(z) = ReLU(z_r) + ReLU(z_i) \quad (2.2)$$

Burada z çıkış değerini, z_r reel bileşeni, z_i sanal bileşen olarak tanımlanmaktadır.

2.2 Öğrenme Katsayısı ve Problemleri

Öğrenme katsayısı, YSA modeline ait ağırlıkları güncellemek için kullanılan bir hiperparametredir. Elde edilen hata sonucunda model parametreleri güncellenirken atılacak adımın büyüklüğünü belirlemektedir. Öğrenme katsayısının uygun şekilde belirlenmemesi, modelin global minimum yerine uygun olmayan yerel minimumlara veya gradyanın sıfır olduğu noktalarda takılı kalmasına sebep olmaktadır (Goodfellow ve diğ., 2016). Yüksek bir öğrenme katsayısı, uygun olmayan bir ağırlık kümesini çok hızlı ve kontrolsüz bir şekilde güncelleyerek başarısız bir öğrenmeye ya da kararsız bir eğitim sürecine neden olmaktadır. Hata fonksiyonun değerinin minimuma doğru kararlı bir şekilde yakınsamak yerine aşırı

artmasına, parametre güncellemelerinde sapmalara ve gradyanların aşırı büyüyerek gradyan patlaması gibi problemlere neden olmaktadır (Ruder, 2016).

Düşük bir öğrenme katsayısı seçilmesi, modelin uygun olmayan yerel minimumlarda takılı kalmasına ve bu noktalardan çıkmak için gerekli momentumu elde edememesine neden olur. Parametrelerin çok küçük öğrenme katsayısı ile güncellenmesi, öğrenme sürecini yavaşlatmakta, modelin uygun çözüme ulaşmasını engellemekte, karmaşık ve büyük boyutlu veri setlerinde öğrenmeyi geciktirmektedir (Goodfellow ve diğ., 2016; Ruder, 2016). Düşük öğrenme katsayısı hata fonksiyonun minimuma ilerlemesini ve global optimuma ulaşmasını yavaşlatarak eğitim sürecini uzatmaktadır.

Uygun bir öğrenme katsayısı, modelin doğruluğunu, genel performansını, hesaplama maliyetini ve genelleme yeteneğini iyileştirirken, yanlış bir öğrenme katsayısının seçimi modelin aşırı öğrenmesine ya da öğrenememesine neden olmaktadır (Üstün ve Yıldız, 2009).

YSA'ların optimizasyon sürecindeki öğrenme katsayısı belirsizliklerini gidermek için çeşitli teknikler önerilmiştir:

2.2.1 Sabit öğrenme katsayısı

Öğrenme katsayısı parametresi için tek bir değer belirlenerek, bu değer eğitimi süreci boyunca kullanıldığı yaklaşımdır (GeeksforGeeks, 2023). Bu yöntem, basit ve anlaşılır olması, model sonuçlarını yorumlarken ayarlanabilen tek değer olması ve hızlı bir eğitim süreci sağlamaktadır. Fakat öğrenme katsayısının yüksek olması modelin yakınsamamasına ve sapmasına, düşük olması ise öğrenmenin gerçekleşmemesine neden olmaktadır. Ayrıca bu değer manuel olarak ayarlanması, ağı dışarıdan müdahale edilmesi anlamına gelmektedir.

2.2.2 Öğrenme çizelgesi

YSA modelinin yakınsamasını artırmak için öğrenme katsayısı azaltma çizelgeleri, eğitim sürecinde öğrenme katsayısını kademeli olarak düşürüldüğü yöntemdir (GeeksforGeeks, 2023). Öğrenme katsayısının azaltılması için kullanılan çizelge yöntemleri aşağıda verilmektedir:

2.2.2.1 Adım azaltması

Adım azaltması yöntemi, YSA modelinin eğitim sürecinde öğrenme katsayısının aşamalı olarak azaltıldığı yöntemdir (GeeksforGeeks, 2023). Bu yöntem ile eğitim başlangıç aşamalarında daha büyük adımlar atarak hızlı yakınsamayı sağlarken, ileriki aşamalarında ise daha başarılı yakınsama sağlanmaktadır. Adım azaltmasının matematiksel formülü Denklem 2.3 ile gösterilmektedir.

$$\alpha = \alpha_{başlangıç} * dr^{\frac{t}{s}} \quad (2.3)$$

Burada α öğrenme katsayısı, $\alpha_{başlangıç}$: başlangıç öğrenme katsayısını, dr : azalma katsayısı ($0 < dr < 1$), t : iterasyonun adım sayısını, s : kaçınıcı adımda öğrenme katsayısının azalacağını göstermektedir.

2.2.2.2 Üssel azalma

Öğrenme katsayısının her iterasyonda üssel olarak azaltıldığı yöntemdir (GeeksforGeeks, 2023). Öğrenme katsayısı değişimi Denklem 2.4 ile gösterildiği gibi hesaplanmaktadır.

$$\alpha = \alpha_{başlangıç} * e^{-dr*t} \quad (2.4)$$

Burada α öğrenme katsayısı, $\alpha_{başlangıç}$: başlangıç öğrenme katsayısını, dr : azalma katsayısı ($0 < dr < 1$), t : iterasyonun adım sayısını göstermektedir.

2.2.2.3 Kosinüs azaltma

Öğrenme katsayısının kosinüs fonksiyonu ile azaltıldığı yöntemdir (Mallick, 2024). Matematiksel formülü Denklem 2.5 ile gösterilmektedir.

$$\alpha = \alpha_{min} + \frac{1}{2}(\alpha_{max} - \alpha_{min}) \left(1 + \cos\left(\frac{T_{cur}}{T_{max}}\pi\right)\right) \quad (2.5)$$

Burada α mevcut adımındaki öğrenme katsayısı, α_{min} zamanlayıcının düşmesi istenilen minimum öğrenme katsayısı, α_{max} döngünün başlangıcındaki başlangıç veya maksimum öğrenme katsayısı, T_{cur} mevcut kosinüs döngüsünün başlangıcından bu yana

geçen iterasyon sayısı, T_{max} döngünün toplam uzunluğu veya öğrenme katsayısının potansiyel olarak sıfırlanmadan önce kosinüs azalmasının sürdüğü iterasyon sayısıdır.

2.2.2.4 Polinomsal azalma

Öğrenme katsayısı zamanla polinomsal olarak azaltılmaktadır. Böylece, daha düzgün azalma sağlanmaktadır (GeeksforGeeks, 2023). Polinomsal azaltma yönteminin matematiksel formülü Denklem 2.6 ile gösterilmektedir.

$$\alpha = \alpha_{başlangıç} * \left(1 - \frac{t}{T}\right)^p \quad (2.6)$$

Burada α öğrenme katsayısı, $\alpha_{başlangıç}$ başlangıç öğrenme katsayısını, dr azalma katsayısı ($0 < dr < 1$), t iterasyonun adım sayısını, T toplam iterasyon sayısını, p öğrenme katsayısının ne kadar hızlı değişimini belirleyen polinom derecesini göstermektedir.

2.3 Adaptif Öğrenme Katsayısı Kullanan Yöntemler

2.3.1 Adagrad yöntemi

Gradyan tabanlı bir optimizasyon yöntemi olan Adagrad, öğrenme katsayısını geçmiş parametrelere göre dinamik olarak ayarlandığı yöntemdir. Her parametre için öğrenme katsayısının gradyan karelerinin birikmiş toplamına bölünmesi ile hesaplanarak her iterasyonda öğrenme katsayısını değiştirmektedir. Bu yöntem, seyrek veriler için daha büyük güncellemeler daha sık veriler için ise daha küçük güncellemeler yapmaktadır (Ruder, 2016). Sabit girilen bir öğrenme katsayısının kullanılmaması algoritmayı diğer adaptif yöntemlere göre daha avantajlı yapmış olsa da formülünde olan paydadaki kare gradyanların birikmesi, öğrenme katsayısının çok fazla küçülmesine sebep olmaktadır. Bu durum öğrenme sürecini erken durdurmasına ya da yerel minimumlarda takılı kalarak yakınsama hızının yavaşlamasına neden olmaktadır. Ayrıca, kareli gradyanların fazla olması büyük modeller için artan bellek tüketimine ve hesaplama maliyetinin artmasına sebep olmaktadır (Alok,2024). Ağırlık güncelleme işlemi Denklem 2.7 ile gösterildiği gibi hesaplanmaktadır:

$$w_{t+1} = w_t - \frac{\alpha}{\sqrt{G_t + \epsilon}} \odot g_t \quad (2.7)$$

Burada G_t , tüm parametrelere göre geçmiş gradyanların karelerinin toplamını, g_t gradyanları, w_t ağırlık değerini ve ϵ paydanın 0 olmaması için verilen (genellikle 0.01) değeri göstermektedir. α Adagrad' a dışarıdan verilen başlangıç öğrenme katsayısı olsa da SGD yöntemindeki gibi sabit bir değer değildir; her parametre için uyarlanmış katsayının çarpanı olarak kullanılmaktadır.

2.3.2 Adadelta yöntemi

Adagrad yönteminin dezavantajlarını ortadan kaldırmak amacıyla geliştirilmiş olan optimizasyon yöntemidir. Adagrad yönteminde öğrenme katsayısının gradyanlara bağlı olarak sürekli azalması eğitim sürecini bir süre sonra durdurmaktadır. Bu sorunu ortadan kaldırmak için Zeller (2012) tarafından önerilen Adadelta yöntemi, sabit bir öğrenme katsayısı yerine önceki parametre güncellemelerinin karelerinin üstel ağırlıklı ortalamalarını kullanarak her iterasyonda değişen bir güncelleme miktarı belirlemektedir (Ruder, 2016).

Denklem 2.8 ile gradyan karelerinin üstel ortalaması, Denklem 2.9 ile parametre güncellemesi için gerekli adım büyüklüğü, Denklem 2.10 ile ağırlık güncelleme işlemi hesaplanmaktadır:

$$E[g^2]_t = p \cdot E[g^2]_{t-1} + (1 - p) \cdot g_t^2 \quad (2.8)$$

$$\Delta w_t = \frac{\sqrt{E[\Delta w^2]_{t-1} + \epsilon}}{\sqrt{E[g^2]_t + \epsilon}} \cdot g_t \quad (2.9)$$

$$w_{t+1} = w_t + \Delta w_t \quad (2.10)$$

Burada p : geçmiş bilgilere ne kadar ağırlık verileceğini belirleyen katsayı (genellikle 0.95 alınır), E gradyan karelerinin üstel ortalaması, g_t gradyanları, w_t ise ağırlık değerlerini göstermektedir.

2.3.3 ADAM yöntemi

Adam (Uyarlanabilir Moment Tahmini), Adagrad ve RMSProp algoritmalarının birleşiminden oluşan gradyan tabanlı bir optimizasyon yöntemidir. Gradyanların birinci ve ikinci momentlerini kullanarak her parametreye özel uyarlanabilir öğrenme oranları belirlemektedir. Karmaşık verilerde optimum çözüme ulaşabilmesi ve adım büyüklüğünü otomatik olarak düzenlemesi, daha hızlı ve kararlı bir yakınsama sağlamaktadır (Kingma ve

Ba, 2014). Moment değerleri Denklem 2.11 ve Denklem 2.12 ile, ağırlıkların güncellenmesi Denklem 2.13 ile hesaplanmaktadır.

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t, v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \quad (2.11)$$

$$\widehat{m}_t = \frac{m_t}{1 - \beta_1^t}, \widehat{v}_t = \frac{v_t}{1 - \beta_2^t} \quad (2.12)$$

$$w_t = w_{t-1} - \frac{\alpha}{\sqrt{\widehat{v}_t + \epsilon}} \widehat{m}_t \quad (2.13)$$

m_t ilk moment, v_t ikinci momenti, g_t gradyan, β_1 ve β_2 moment tahminleri için çarpanlar, α öğrenme katsayısı, ϵ sayısal kararlılığı artırmak için küçük bir değeri ifade etmektedir.

Öğrenme katsayısı ile ilgili yapılan çalışmaların hepsinde ağırlıkların güncellenmesine göre yapılan bir güncellemeden çok tahmin üzerine yapılan güncellemedir. Bundan dolayı öğrenme katsayısını elde edilen hatanın azaltılmasına yönelik hassas değişim olarak tanımlanmamıştır. Bu tez çalışmasında YSA eğitiminin sonucunda elde edilen hatadan yola çıkarak öğrenme katsayısı hesaplanacaktır.

2.3.4 KarcıFANN (Karcı kesir dereceli yapay sinir ağları) yöntemi

KarcıFANN yöntemi, gradyan tabanlı optimizasyon yöntemlerinde ağırlık güncelleme sürecinde öğrenme katsayısı kullanmak yerine, kesir dereceli türevden faydalanılarak geliştirilmiş yenilikçi bir yaklaşımdır (Karakurt vd.,2022;Karakurt vd., 2024; Karakurt vd., 2025; Saygılı vd.,2025). Bu yöntem, sabit bir öğrenme katsayısı kullanmaktan kaynaklanan dezavantajları ortadan kaldırmayı amaçlamaktadır. Ağırlık güncelleme süreci sistemden gelen gerçek verilere dayalı olarak, her iterasyon sonunda elde edilen hata değerine göre dinamik olarak gerçekleştirilmektedir. Her bir ağırlık değeri bir önceki hata değerinin değişimine göre güncellenmektedir.

Gradyan tabanlı yöntemlerde kullanılan sabit bir öğrenme katsayısının kullanımı global optimum yerine yerel bir minimumda takılıp kalması, eğitim verilerine aşırı uyum sağlayarak görülmemiş veriler üzerinde kötü performans sergilemesi, modelin hiç öğrenememesi ya da optimuma yakınsayamaması gibi sorunlara sebep olmaktadır.

Sabit bir öğrenme katsayısının kullanılmadığı KarcıFANN yöntemi ağ modeline;

- sisteme dışarıdan müdahaleyi en aza indirmek,
- dinamik ve hataya duyarlı güncellemeler sayesinde optimum çözüme erken süreçte ulaşmak,
- yerel minimumlara takılma riskini azaltmak,
- kararlı bir yakınsama sağlamak,
- modelin görülmemiş veriler üzerindeki genelleme yeteneğini artırmak

gibi faydalar sunmaktadır.

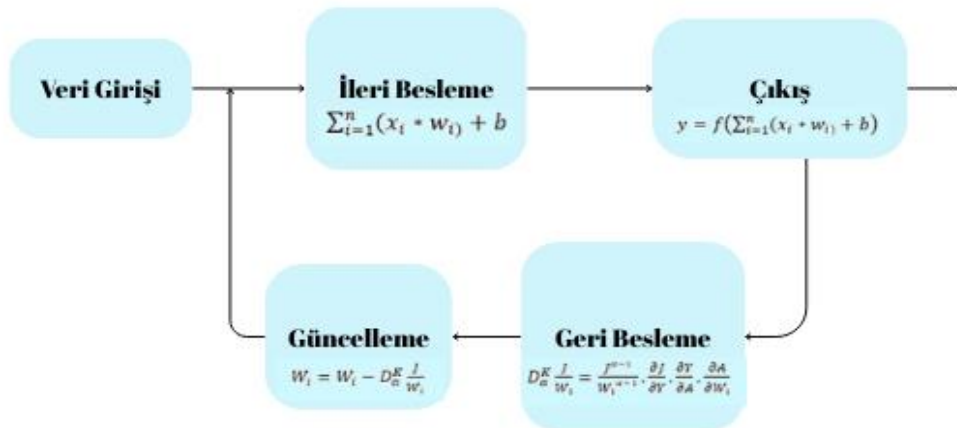
Karcı kesir dereceli türev Denklem 2.14, ağırlık güncelleme işlemi ise Denklem 2.15 ile hesaplanmaktadır.

$$D_{\alpha}^K f(x) = \left(\frac{f(x)}{x}\right)^{\alpha-1} \cdot f'(x) \quad (2.14)$$

$$w_{t+1} = w_t - D_{\alpha}^K \frac{f(x)}{w_t} \quad (2.15)$$

Burada $f(x)$, hata fonksiyonunu; $f'(x)$, hata fonksiyonunun klasik Newton türevini ve α , Karcı kesir dereceli türevin derecesini göstermektedir.

Önerilen KarcıFANN yönteminin akış modeli Şekil 2.1 ile gösterildiği gibidir:



Şekil 2. 1: KarcıFANN iş akış modeli diyagramı.

Algoritma 1 KarcıFANN

Girişler: X veri seti, Y veri setinin etiketleri,

$w_{\text{hidden}} \in (0.1,1]$, $w_{\text{output}} \in (0.1,1]$ ağırlık değerleri,

Bias parametreleri $b_{\text{hidden}} \in (0.1,1]$, $b_{\text{output}} \in (0.1,1]$,

k maksimum iterasyon sayısı,

kesir dereceli türevin derecesi alfa (α)

Çıkış: sınıflandırılmış veri

for e $\leftarrow$ 0 to k-1 do

doğruluk, hata $\leftarrow$ compute_mse_and_acc()

for (X,Y):

if $w_{\text{hidden}} \neq 0$, $w_{\text{çıkış}} \neq 0$, $b_{\text{hidden}} \neq 0$

$w_{\text{hidden}} \leftarrow w_{\text{hidden}} - ((\text{hata}[e]/w_{\text{hidden}})^{*(\alpha-1)} * (\delta(j) / \delta(w_{\text{hidden}})))$

$w_{\text{output}} \leftarrow w_{\text{output}} - ((\text{hata}[e]/w_{\text{output}})^{*(\alpha-1)} * (\delta(j) / \delta(w_{\text{output}})))$

$b_{\text{hidden}} \leftarrow b_{\text{hidden}} - ((\text{hata}[e]/b_{\text{hidden}})^{*(\alpha-1)} * (\delta(j) / \delta(b_{\text{hidden}})))$

$b_{\text{output}} \leftarrow b_{\text{output}} - ((\text{hata}[e]/b_{\text{output}})^{*(\alpha-1)} * (\delta(j) / \delta(b_{\text{output}})))$

Algoritma 1 ile KarcıFANN yönteminin temel çalışma prensibi gösterilmektedir. Geleneksel gradyan tabanlı optimizasyon algoritmalarının aksine, bu yaklaşım ağırlık güncellemeleri için sabit bir öğrenme katsayısı kullanmaz; bunun yerine, Karcı kesir dereceli türev ile mevcut hata değerlerine dayalı dinamik bir yaklaşım sunmaktadır. Her iterasyonda öğrenme katsayısı yerine geçen $(\frac{f(x)}{w})^{\alpha-1}$ teriminde hatanın değişimi ile beraber ağırlık da değişmektedir. Bu durumda, her iterasyon ve her ağırlık için Newton türevi farklı bir katsayı ile çarpılmış olur. Bu katsayı doğrudan sistemden gelen bir değer olup dışarıdan müdahaleye kapalıdır.

3. KARMAŞIK DEĞERLİ SİNİR AĞLARI (CVNN)

Reel değerli YSA'lar (RVNN), giriş, çıkış ve tüm parametreleri sadece reel sayılardan oluşan literatürde en yaygın kullanılan YSA modelidir. Karmaşık değerli YSA'lar (CVNN) ise giriş ve ağ parametrelerinin karmaşık sayılarla ifade edildiği bir YSA modelidir. CVNN'ler genlik ve faz bilgisini doğrudan işleyebilme yeteneğine sahip olduklarından elektromanyetik dalgalar, ses ve konuşma sinyalleri, Manyetik Rezonans (MR) ve tıbbi görüntüleme teknikleri, sonar ve denizaltı algılama ile optik iletişim gibi birçok RVNN'lere göre daha başarılı performans sergilemektedir (Lee, 2022).

Literatürde reel verilerin karmaşık formata dönüştürülmesi için iki yaygın yöntem bulunmaktadır. İlk yöntemde reel özellikler ikili gruplar halinde eşleştirilerek karmaşık sayılar oluşturulur (Gürüler, 2017). İkinci yöntemde ise her bir reel değer faz eşleştirme ile birim çember üzerine taşınarak faz bilgisi doğrudan modele aktarılır (Amin ve Murase, 2009).

3.1 CVNN Türleri

CVNN mimarileri iki ana gruba ayrılmaktadır: Split-CVNN ve Tam Karmaşık CVNN.

3.1.1 Split-CVNN

Split-CVNN mimarisi, başlangıçta karmaşık değerli girişin iki ayrı reel bileşene ayrılması ve bu verinin RVNN yapısı ile eğitilmesi şeklinde tasarlanmıştır. Bu ayrıştırma işlemi Denklem 3.1 ve Denklem 3.2 ile gösterildiği gibi iki farklı biçimde yapılmaktadır (Lee, 2022).

$$\{z_1, z_2\} = \{(x_1, y_1), (x_2, y_2)\} \quad (3.1)$$

$$\{z_1, z_2\} = \{(r_1, \phi_1), (r_2, \phi_2)\} \quad (3.2)$$

Burada $z_1, z_2 \in \mathbb{C}$ ve $z = x + iy$, x karmaşık girişin reel bileşeni ve y karmaşık girişin sanal bileşenlerini; r karmaşık girişin genlik ve ϕ karmaşık girişin faz bileşenlerini ifade etmektedir.

3.1.2 Tam karmaşık CVNN

Tam karmaşık CVNN mimarisi, ağırlıkların ve aktivasyon fonksiyonlarının karmaşık sayı kümesinde tanımlı olan karmaşık değerli sinyallerle doğrudan çalışan mimarilerdir. Karmaşık türevlenebilir aktivasyon fonksiyonları kullanıldığı için Split-CVNN'lere göre daha fazla hesaplama maliyetine sebep olmaktadır (Abdalla, 2023).

3.2 CVNN'lerde Gradyan Tabanlı Öğrenme

CVNN öğrenme süreci RVNN'lere benzer olarak ileri besleme ve geri yayılım adımlarından oluşmaktadır. Fakat karmaşık sayıların doğası gereği gradyanların hesaplanması için Wirtinger türevinden yararlanılmaktadır (Abdalla, 2023).

3.2.1 CVNN'lerde ileri besleme

CVNN'de ileri besleme, karmaşık değerli giriş verisinin ağ katmanları boyunca karmaşık ağırlıklar ve karmaşık aktivasyon fonksiyonları kullanılarak işlenmesi ve çıktı katmanında genlik ve faz bilgisini koruyan tahminlerin üretilmesi sürecinden oluşmaktadır (Abdalla, 2023). İleri besleme sürecinde çıkış değeri, klasik YSA'larda olduğu gibi hesaplanmaktadır. Ancak burada girişler ve ağırlıklar karmaşık sayı değerlerinden oluşmaktadır.

3.2.2 CVNN'lerde geri besleme

CVNN'lerde geri yayılım, RVNN'de olduğu gibi hata fonksiyonun karmaşık ağırlıklar üzerinden türevinin hesaplanarak ağırlıkların güncellenmesi sürecidir. Türev hesaplanmasında karmaşık sayıların getirdiği kısıtlar nedeniyle Wirtinger türevi yaygın olarak kullanılmaktadır (Abdalla, 2023). Wirtinger türevi Denklem 3.3'te belirtildiği şekilde tanımlanmaktadır.

$$\frac{\partial f}{\partial z} = \frac{1}{2} \left(\frac{\partial f}{\partial x} - i \frac{\partial f}{\partial y} \right), \quad \frac{\partial f}{\partial z^*} = \frac{1}{2} \left(\frac{\partial f}{\partial x} + i \frac{\partial f}{\partial y} \right) \quad (3.3)$$

Burada f , skaler değerli bir fonksiyon, $z = x + iy$ şeklinde karmaşık değişken ve z^* z 'nin karmaşık eşleniğini ifade etmektedir. Ağırlıkların güncellenmesi işlemi Denklem 3.4 ile gösterilmektedir.

$$w_t = w_{t-1} - \alpha \frac{\partial E}{\partial w_{t-1}} \quad (3.4)$$

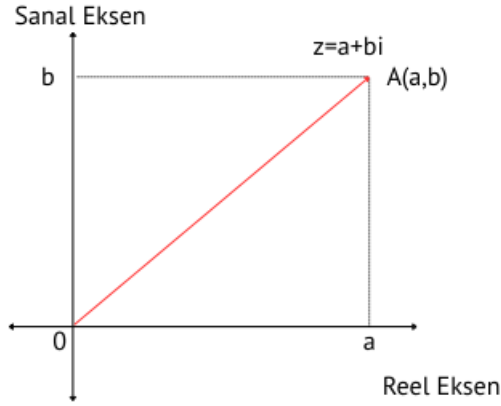
Burada w_t ağırlık değeri, $\frac{\partial E}{\partial w_{t-1}}$ hata fonksiyonunun ilgili ağırlığa göre Wirtinger türevi, α öğrenme oranını temsil etmektedir.



4. KARCIFANN YÖNTEMİNİN KARMAŞIK SAYILARLA ANALİZİ

4.1 KarcıFANN Yönteminde Karmaşık Durumların Tanımı

Karmaşık sayılar, reel sayıların yetersiz kaldığı durumlarda matematiksel modellemeyi genişletmek amacıyla kullanılmaktadır. Reel sayı kümesine sanal birim (i) eklenmesi ile karmaşık sayı kümesi ($\mathbb{C} = \{a + bi : a, b \in \mathbb{R}\}$) elde edilir. Burada, $i^2 = -1$ olmak üzere a sayının reel, b ise sanal bileşenini ifade etmektedir (Tektaş, 2024).



Şekil 4. 1: Karmaşık sayıların gösterimi.

Kesir dereceli türev tanımlarında yer alan kesir derece parametresi, bu tez kapsamında $\alpha = \frac{p}{q}$ biçiminde rasyonel sayı olarak ifade edilmektedir. Bu parametre, yalnızca türev işleminin büyüklüğünü değil aynı zamanda bu işlemin tanım kümesini de doğrudan etkilemektedir. Özellikle $(\frac{f(x)}{x})^{\alpha-1}$ biçimindeki ifadelerde, üs parametresinin işaretine ve rasyonel bileşenlerinin (p ve q) paritelerine bağlı olarak reel ya da karmaşık sayı kümeleri arasında geçiş söz konusu olabilmektedir.

4.2 Karcı Kesir Dereceli Türev ile Karmaşık Sayılar Arasındaki İlişki

E: hata, W: ağırlık ve $p, q \in \mathbb{Z}$ olmak üzere $\alpha = \frac{p}{q}$ türev derecesini göstermektedir.

Durum 1: $E, W \geq 0$ olmak üzere p ve q sayılarının çift ve tek olduğu bütün durumlar için Denklem 4.1, 4.2, 4.3 ve 4.4 elde edilir.

$$a) \quad p, q \in \mathbb{Z}^+ \text{ için } \left(\frac{E}{W}\right)^{\frac{p}{q}-1} = \sqrt[q]{\left(\frac{E}{W}\right)^{p-q}}, \quad \left(\frac{E}{W}\right)^{p-q} \in R^+ \setminus \{0\} \xrightarrow{D_\alpha^K\left(\frac{E}{W}\right)} R \quad (4.1)$$

$$b) \quad p, q \in \mathbb{Z}^- \text{ için } \left(\frac{E}{W}\right)^{\frac{p}{q}-1} = \sqrt[q]{\left(\frac{E}{W}\right)^{p-q}}, \quad \left(\frac{E}{W}\right)^{p-q} \in R^+ \setminus \{0\} \xrightarrow{D_\alpha^K\left(\frac{E}{W}\right)} R \quad (4.2)$$

$$c) \quad p \in \mathbb{Z}^+ \text{ ve } q \in \mathbb{Z}^- \text{ için } \left(\frac{E}{W}\right)^{\frac{p}{q}-1} = \sqrt[q]{\left(\frac{W}{E}\right)^{p-q}}, \quad \left(\frac{W}{E}\right)^{p-q} \in R^+ \setminus \{0\} \xrightarrow{D_\alpha^K\left(\frac{W}{E}\right)} R \quad (4.3)$$

$$d) \quad p \in \mathbb{Z}^- \text{ ve } q \in \mathbb{Z}^+ \text{ için } \left(\frac{E}{W}\right)^{\frac{p}{q}-1} = \sqrt[q]{\left(\frac{W}{E}\right)^{p-q}}, \quad \left(\frac{W}{E}\right)^{p-q} \in R^+ \setminus \{0\} \xrightarrow{D_\alpha^K\left(\frac{W}{E}\right)} R \quad (4.4)$$

Durum 2: $E, W < 0$ olmak üzere p ve q sayılarının çift ve tek olduğu bütün durumlar için Denklem 4.5, 4.6, 4.7 ve 4.8 elde edilir.

$$a) \quad p, q \in \mathbb{Z}^+ \text{ için } \left(\frac{E}{W}\right)^{\frac{p}{q}-1} = \sqrt[q]{\left(\frac{E}{W}\right)^{p-q}}, \quad \left(\frac{E}{W}\right)^{p-q} \in R^+ \setminus \{0\} \xrightarrow{D_\alpha^K\left(\frac{E}{W}\right)} R \quad (4.5)$$

$$b) \quad p, q \in \mathbb{Z}^- \text{ için } \left(\frac{E}{W}\right)^{\frac{p}{q}-1} = \sqrt[q]{\left(\frac{E}{W}\right)^{p-q}}, \quad \left(\frac{E}{W}\right)^{p-q} \in R^+ \setminus \{0\} \xrightarrow{D_\alpha^K\left(\frac{E}{W}\right)} R \quad (4.6)$$

$$c) \quad p \in \mathbb{Z}^+ \text{ ve } q \in \mathbb{Z}^- \text{ için } \left(\frac{E}{W}\right)^{\frac{p}{q}-1} = \sqrt[q]{\left(\frac{W}{E}\right)^{p-q}}, \quad \left(\frac{W}{E}\right)^{p-q} \in R^+ \setminus \{0\} \xrightarrow{D_\alpha^K\left(\frac{W}{E}\right)} R \quad (4.7)$$

$$d) \quad p \in \mathbb{Z}^- \text{ ve } q \in \mathbb{Z}^+ \text{ için } \left(\frac{E}{W}\right)^{\frac{p}{q}-1} = \sqrt[q]{\left(\frac{W}{E}\right)^{p-q}}, \quad \left(\frac{W}{E}\right)^{p-q} \in R^+ \setminus \{0\} \xrightarrow{D_\alpha^K\left(\frac{W}{E}\right)} R \quad (4.8)$$

Durum 3: $E > 0$ ve $W < 0$ için, $p, q > 0$ olmak üzere (negatif değerler yalnızca kesrin yönünü ters çevireceği için değerlendirme dışı bırakılmıştır) p ve q 'nin tek ya da çift olmasına bağlı olarak Denklem 4.9, 4.10, 4.11 ve 4.12'de verilen durumlar elde edilmektedir.

a) p, q çift olduğu durum için

$$\left(\frac{E}{W}\right)^{\frac{p}{q}-1} = \sqrt[q]{\left(\frac{E}{W}\right)^{p-q}}, \quad \left(\frac{E}{W}\right)^{p-q} \in R^+ \xrightarrow{D_\alpha^K\left(\frac{E}{W}\right)} R \quad (4.9)$$

b) p ve q tek olduğu durumlar için

$$\left(\frac{E}{W}\right)^{\frac{p}{q}-1} = \sqrt[q]{\left(\frac{E}{W}\right)^{p-q}}, \quad \left(\frac{E}{W}\right)^{p-q} \in R^+ \setminus \{0\} \xrightarrow{D_\alpha^K\left(\frac{E}{W}\right)} R \quad (4.10)$$

c) p tek ve q çift olduğu durum için

$$\left(\frac{E}{W}\right)^{\frac{p}{q}-1} = \sqrt[q]{\left(\frac{E}{W}\right)^{p-q}}, \quad \left(\frac{E}{W}\right)^{p-q} \in R^- \xrightarrow{D_{\alpha}^K\left(\frac{E}{W}\right)} C \quad (4.11)$$

d) p çift ve q tek olduğu durum için

$$\left(\frac{E}{W}\right)^{\frac{p}{q}-1} = \sqrt[q]{\left(\frac{E}{W}\right)^{p-q}}, \quad \left(\frac{E}{W}\right)^{p-q} \in R^- \xrightarrow{D_{\alpha}^K\left(\frac{E}{W}\right)} R \quad (4.12)$$

Durum 4: $E < 0$ ve $W > 0$ için, $p, q > 0$ olmak üzere (negatif değerler yalnızca kesrin yönünü ters çevireceği için değerlendirme dışı bırakılmıştır) p ve q'nun tek ya da çift olmasına bağlı olarak Denklem 4.13, 4.14, 4.15 ve 4.16'da verilen durumlar elde edilmektedir

a) p, q çift olduğu durum için

$$\left(\frac{E}{W}\right)^{\frac{p}{q}-1} = \sqrt[q]{\left(\frac{E}{W}\right)^{p-q}}, \quad \left(\frac{E}{W}\right)^{p-q} \in R^+ \xrightarrow{D_{\alpha}^K\left(\frac{E}{W}\right)} R \quad (4.13)$$

b) p ve q tek olduğu durumlar için

$$\left(\frac{E}{W}\right)^{\frac{p}{q}-1} = \sqrt[q]{\left(\frac{E}{W}\right)^{p-q}}, \quad \left(\frac{E}{W}\right)^{p-q} \in R^+ \xrightarrow{D_{\alpha}^K\left(\frac{E}{W}\right)} R \quad (4.14)$$

c) p tek ve q çift olduğu durum için

$$\left(\frac{E}{W}\right)^{\frac{p}{q}-1} = \sqrt[q]{\left(\frac{E}{W}\right)^{p-q}}, \quad \left(\frac{E}{W}\right)^{p-q} \in R^- \xrightarrow{D_{\alpha}^K\left(\frac{E}{W}\right)} C \quad (4.15)$$

d) p çift ve q tek olduğu durum için

$$\left(\frac{E}{W}\right)^{\frac{p}{q}-1} = \sqrt[q]{\left(\frac{E}{W}\right)^{p-q}}, \quad \left(\frac{E}{W}\right)^{p-q} \in R^- \xrightarrow{D_{\alpha}^K\left(\frac{E}{W}\right)} R \quad (4.16)$$

4.3 Karmaşık Değerler ile KarcıFANN Yöntemi

YSA'larda karmaşık değerli işlem yapısı, reel eksenindeki doğrusal işlemlerin karmaşık düzleme genişletilmesiyle elde edilir. Bu amaçla KarcıFANN yöntemi karmaşık uzaya genelleştirilmekte, ağırlıkların reel (w_r) ve sanal (w_i) bileşenleri birlikte optimize edilmektedir. Ağın ileri yayılım işlem akışı Denklem 4.17 ve 4.18'de verilmiştir.

$$z = x(w_r + iw_i) + b, \quad b = (b_r + ib_i) \quad (4.17)$$

$$z_r = xw_r + b_r, \quad z_i = xw_i + b_i \quad (4.18)$$

Burada x giriş, w_r, w_i ağırlık matrisinin reel ve sanal bileşenlerini, z çıkış değerini ve b bias değerinin göstermektedir. Bias için başlangıç değeri 0 olarak belirlenmiştir.

Aktivasyon fonksiyonları (f) Denklem 4.19 ile gösterildiği gibi bileşen bazlı uygulanmaktadır.

$$zz_r = f(z_r), zz_i = f(z_i) \quad (4.19)$$

Elde edilen zz çıkış değerinin reel ve sanal bileşenlerini birlikte değerlendirebilmek ve karar aşamasında bu iki bileşenin katkısını değerlendirmek amacıyla Karmaşık Karar Katmanı adı verilen bir katman tasarlanmıştır. Bu katman, her sınıf için iki ayrı ölçek parametresi (u_r ve u_i) kullanmaktadır. Karmaşık Karar Katmanı'ndan elde edilen çıktı değeri Denklem 4.20 ile ifade edilmektedir.

$$y_{bv,c} = (zz_{r,bv,c} * u_{r,c}) + (zz_{i,bv,c} * u_{i,c}) + b \quad (4.20)$$

Burada bv örnek (batch) indeksini, c sınıf indeksini göstermektedir. u_r ve u_i her sınıf için reel ve sanal bileşenin katkısını ayrı ayrı ayarlayan adaptif parametrelerdir. Başlangıçta $u_r = 1$ ve $u_i = 0$ olarak atanmıştır. Eğitim süreci boyunca bu parametreler, KarcıFANN güncelleme kuralı kullanılarak iteratif biçimde güncellenmektedir. Bu sayede sanal bileşenin etkisi, eğitim sürecinde kademeli ve denetimli olarak devreye alınmaktadır. Modelin reel bileşeni karar sürecinin temelini oluştururken sanal bileşen, verinin faz bilgisini eğitime dahil etmektedir. Bu yöntemde ara katmanlarda hesaplanan çıktı hem genlik hem de faz bilgisi taşımaktadır. Literatürde sık başvurulmuş genlik tabanlı yöntemler faz bilgisini büyük ölçüde kaybederek karar yüzeyini daraltır. Bu çalışmada, faz bilgisini kaybetmeden çıkış katmanında sınıf-başına uyarlanabilir bir Re/Im karışımı öğrenen Karmaşık Karar Katmanı önerilmektedir.

Karmaşık Karar Katmanı'nın amacı modelin karmaşık uzaydaki temsilini reel karar uzayına taşımakta ve faz duyarlılığını korumaktadır. Bu yaklaşım, özellikle reel ve sanal bileşenler arasında ilişki bulunan verilerde modelin ayırt edici gücünü artırmaktadır. Ayrıca, KarcıFANN'ın kesir dereceli türev tabanlı dinamik güncelleme kuralı, bu katmanda kullanılan parametrelerin hata büyüklüğüne ve parametre ölçeğine bağlı olarak adaptif biçimde ayarlanmasını sağlamaktadır. Böylece model, erken dönemlerde daha kararlı, ilerleyen aşamalarda ise daha duyarlı bir öğrenme davranışı sergilemektedir.

Karmaşık karar katmanından elde edilen Y çıkış değeri ve T beklenen değer olmak üzere elde edilen hata değeri E Denklem 4.21 ile hesaplanmaktadır:

$$E = \mathcal{L}(Y, T) = \sum_{bv=1}^B \ell(Y_{bv}, T_{bv}) \quad (4.21)$$

Elde edilen hata değerine göre geri yayılım işlemi yapılırken ağırlık değerleri karmaşık değer olduğu için klasik türev tanımlı değildir. Wirtinger türevinde, $L(w, w^*)$ iki bağımsız değişken kullanılmaktadır. Reel bileşen için $\frac{\partial E}{\partial w_r}$, sanal bileşen için $\frac{\partial E}{\partial w_i}$ ayrı ayrı hesaplanır ve sonra bu iki bileşen Denklem 4.22 ve 4.23 ile gösterilen Wirtinger türevi hesaplamalarında kullanılır.

$$\frac{\partial E}{\partial w} = \frac{1}{2} \left(\frac{\partial E}{\partial w_r} - i \frac{\partial E}{\partial w_i} \right) \quad (4.22)$$

$$\frac{\partial E}{\partial w^*} = \frac{1}{2} \left(\frac{\partial E}{\partial w_r} + i \frac{\partial E}{\partial w_i} \right) \quad (4.23)$$

CVNN optimizasyonunda, reel değerli hata $L(W, W^*)$ için en hızlı azalma yönü Wirtinger eşlenik gradyanla belirlenir. Δw , Δw_r ve Δw_i hesaplama yöntemleri Denklem 4.24'te, u_r ve u_i gradyanları 4.25 ve 4.26 denklemlerinde gösterildiği gibi hesaplanmaktadır.

$$\Delta w = \frac{\partial E}{\partial w^*} = \frac{1}{2} \left(\frac{\partial E}{\partial w_r} + i \frac{\partial E}{\partial w_i} \right), \Delta w_r = \Re(\Delta w), \Delta w_i = \Im(\Delta w) \quad (4.24)$$

$$\frac{\partial E}{\partial u_{r,c}} = \sum_{bv=1}^B \frac{\partial E}{\partial Y_{bv,c}} \cdot \frac{\partial Y_{bv,c}}{\partial u_{r,c}} \quad (4.25)$$

$$\frac{\partial E}{\partial u_{i,c}} = \sum_{bv=1}^B \frac{\partial E}{\partial Y_{bv,c}} \cdot \frac{\partial Y_{bv,c}}{\partial u_{i,c}} \quad (4.26)$$

Hata fonksiyonu (E) için ağırlık parametrelerinin ve Karmaşık Karar Katmanının iki ayrı ölçek parametresi için güncelleme işlemi 4.27-4.31 denklemlerinde gösterilmektedir.

$$D_\alpha^K = \left(\frac{E}{w} \right)^{\alpha-1} \frac{\partial E}{\partial w^*} \quad (4.27)$$

$$w_{r_{t+1}} = w_{r_t} - \left(\frac{E}{w_{r_t}} \right)^{\alpha-1} \Delta w_r \quad (4.28)$$

$$w_{i_{t+1}} = w_{i_t} - \left(\frac{E}{w_{i_t}} \right)^{\alpha-1} \Delta w_i \quad (4.29)$$

$$u_{r_{t+1}} = u_{r_t} - \left(\frac{E}{u_{r_t}} \right)^{\alpha-1} \frac{\partial E}{\partial u_{r_t}} \quad (4.30)$$

$$u_{i_{t+1}} = u_{i_t} - \left(\frac{E}{u_{i_t}} \right)^{\alpha-1} \frac{\partial E}{\partial u_{i_t}} \quad (4.31)$$

Kompleks modelin çıkış değerine reel ve sanal bileşenlerin etkilerini gözlemlemek için öğrenme eğilimi değerine bakılmaktadır. Çıkış değeri $z' = z'_r + jz'_i$ biçiminde ifade edildiğinde sanal bileşenin etkisi Denklem 4.32 ve reel bileşenin etkisi 4.33 ile tanımlanmaktadır.

$$\text{sanal_etki} = \frac{z'_i}{\sqrt{z'^2_r + z'^2_i}} \quad (4.32)$$

$$\text{reel_etki} = \frac{z'_r}{\sqrt{z'^2_r + z'^2_i}} \quad (4.33)$$

$\frac{\text{reel_etki}}{\text{sanal_etki}} > 1$ olması durumunda reel bileşenin daha etkili olduğu, $\frac{\text{sanal_etki}}{\text{reel_etki}} > 1$ olması durumunda sanal bileşenin daha etkili olduğu anlamına gelmektedir.

5. UYGULAMALARDA KULLANILAN VERİ SETLERİ

Tez kapsamında geliştirilen çalışmalarda kullanılan veri setleri aşağıda özetlenmiştir:

- Digits veri kümesi 8x8 piksel boyutunda ve gri-seviyeli el yazısı rakam görüntüsünden oluşmaktadır. Görüntüler gri tonlamalı ve piksel değerleri 0 (beyaz) ile 16 (siyah) arasında değişir. Veri kümesi 0–9 arasındaki rakamları temsil eden 10 sınıf içermektedir. Toplamda yaklaşık 1797 adet görüntü örneği bulunmaktadır (UCI Machine Learning Repository).
- MNIST Rotation 45°, orijinal MNIST veri kümesindeki 28×28 piksel el yazısı rakam görüntülerinin, her birinin 45° döndürülmüş versiyonlarını içermektedir. Bu veri kümesi, 60000 eğitim ve 10000 test örneği olmak üzere toplam 70000 görüntüden oluşmaktadır (Teenakunsoth, 2020).
- SignMNIST, Amerikan İşaret Dili'ne ait A'dan Z'ye (J ve Z haricindeki) harfleri temsil eden el işareti görsellerini içermektedir. Görüntüler gri tonlamalı ve 28*28 pikselden oluşmaktadır. 27455 eğitim ve 7172 test verisi olmak üzere toplam 34627 içermektedir (OpenML, 2023).
- Kuzushiji-MNIST, Japonca el yazması karakterlerin sınıflandırılması amacıyla oluşturulmuş bir görüntü veri setidir. Japonca dilinde yazı yazmak için kullanılan Hiragana hece yazısı stilinin 10 satırı temsilen 10 sınıfa ayrılmıştır. Görüntüler gri tonlamalı ve 28*28 pikselden oluşmaktadır. 60000 eğitim ve 10000 test verisi içermektedir (OpenML, 2019).
- Gina_prior2 veri seti, el yazısıyla yazılmış rakamlardan oluşan ve 10 sınıfa ayrılmış 3468 görüntü içermektedir. 28*28 pikselden oluşmaktadır (OpenML, 2014).
- XOR, iki giriş değerinin aynı olması durumunda 0, farklı olması durumunda ise 1 çıktısı üreten basit bir sınıflandırma problemidir.

Ayrıca MNIST Rotation 45° ve Digits veri setlerinin 2BFT dönüşümü uygulanmış frekans uzayı karşılıkları da (2BFT-Digits, 2BFT-MNIST Rotation 45°) çalışmalara dahil edilmiştir. Tüm çalışmalarda görüntüler, tek boyutlu satır vektörlerine dönüştürüldükten sonra oluşturulan ağ modeline giriş verisi olarak verilmektedir.

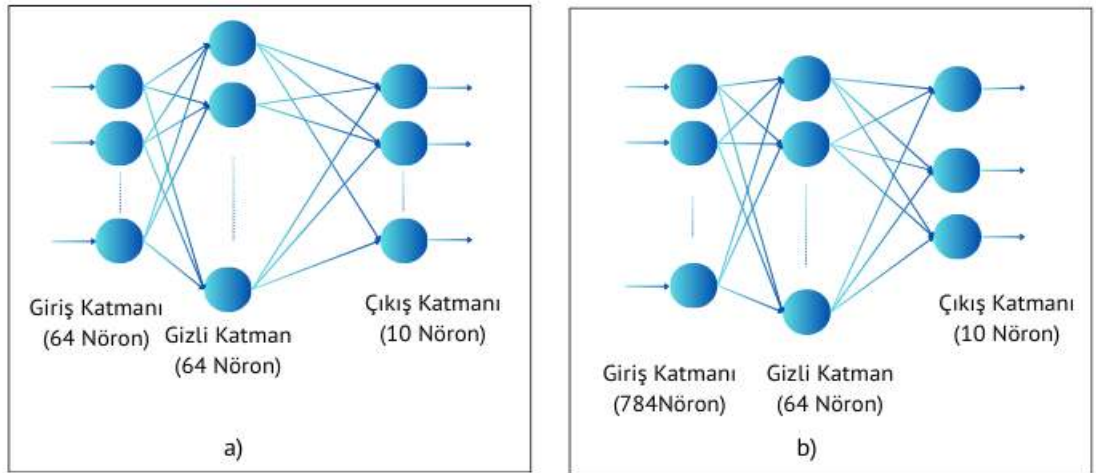
6. UYGULAMALAR VE SONUÇLAR

6.1 Karcı Kesir Dereceli YSA (KarcıFANN) yönteminin Reel Sayılar ile Uygulanması

KarcıFANN yönteminin reel sayılardaki performansını değerlendirmek amacıyla, Digits ve 45° döndürülmüş MNIST (MNIST Rotation 45°) veri kümeleri üzerinde deneysel çalışmalar gerçekleştirilmiştir.

Eğitim sürecinde her iki veri setinde iterasyon (epoch) sayısı 150 olarak belirlenmiştir. Verilerin %30'u doğrulama amacıyla ayrılmış, başlangıç ağırlıkları ise rastgele ve aynı değerler seçilmiştir.

Bu çalışmada kullanılan iki veri setinin girdi boyutu ve sınıf sayıları dikkate alınarak uygun YSA mimarileri tasarlanmıştır. Şekil 6.1'de gösterildiği gibi, Digits veri seti için giriş katmanı 64, gizli katman 64 ve çıkış katmanı 10 nörondan oluşacak şekilde tasarlanmıştır. MNIST Rotation 45° veri setinde ise giriş katmanı 784 nöron, gizli katman 64 nöron ve çıkış katmanı 10 nöron olarak yapılandırılmıştır.



Şekil 6. 1: a) Digits b) MNIST Rotation 45° veri setlerinin Sınıflandırılmasında Kullanılan Ağ Mimarisi

Çizelge 6.1'de KarcıFANN yöntemi ile kesir derecesi parametresi alfanın 5 farklı değeri kullanılarak gerçekleştirilen deneysel çalışmalara ait sonuçlar sunulmaktadır. Digits

veri seti incelendiğinde en yüksek test doğruluğu %83.89 ve en düşük test hatası 0.0674 alfanın 1.2 değerinde elde edilmiştir.

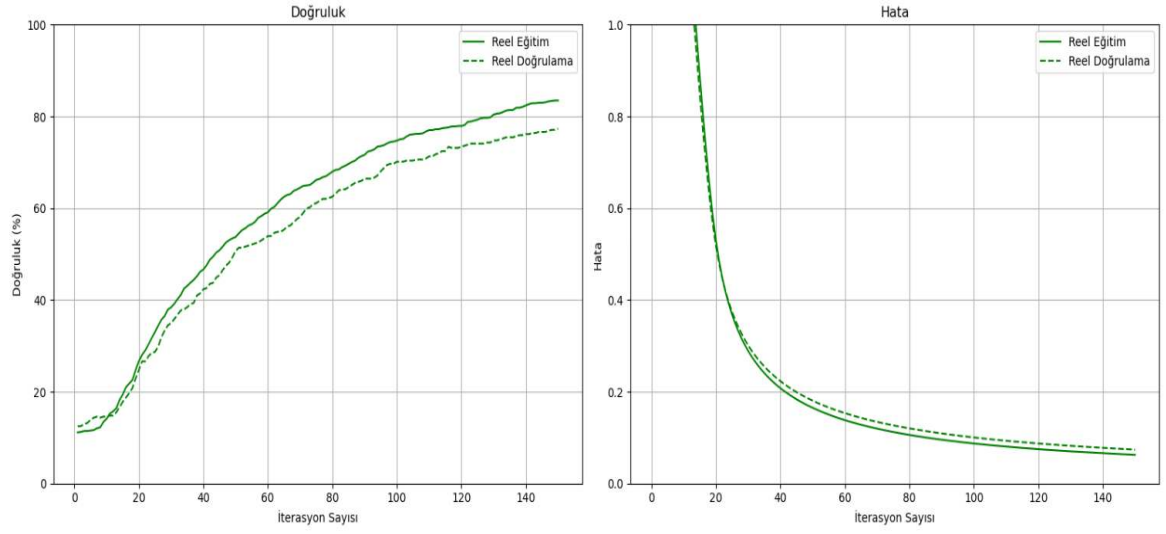
Alfa 0.5 ve 0.9 değerlerinde eğitim doğruluklarının yüksek olmasına rağmen doğrulama ve test doğruluklarının düşük kalması, modelin aşırı uyum eğiliminde olduğuna işaret etmektedir. Alfa 1.8 değerinde doğrulama doğruluğu %81.71'e yükselmiş olmasına rağmen test doğruluğunun %78.89'a düşmesi, modelin eğitim verisine fazla uyum sağladığını ve genelleme kabiliyetinin zayıfladığını ortaya koymaktadır.

MNIST Rotation 45° veri seti sonuçlarında ise alfa 0.9 ve 1.2 değerlerinin aynı performansı sergilediği görülmektedir. Bu iki durumda test doğruluğu %88.92 ve test hatası 0.0228 olarak gerçekleşmiş; eğitim, doğrulama ve test doğruluklarının birbirine yakın olması modelin dengeli bir öğrenme süreci yürüttüğünü ve kararlı bir genelleme yeteneğine sahip olduğunu göstermiştir. Alfa 0.5 değerinde doğrulama ve test doğruluklarının bir miktar düşmesi, düşük alfa değerlerinde modelin genelleme yeteneğinin sınırlı kaldığını göstermektedir.

Çizelge 6. 1: Digits ve MNIST Rotation 45° veri setlerinde KarcıFANN yönteminin farklı alfa değerlerindeki doğruluk ve hata fonksiyonu değerleri.

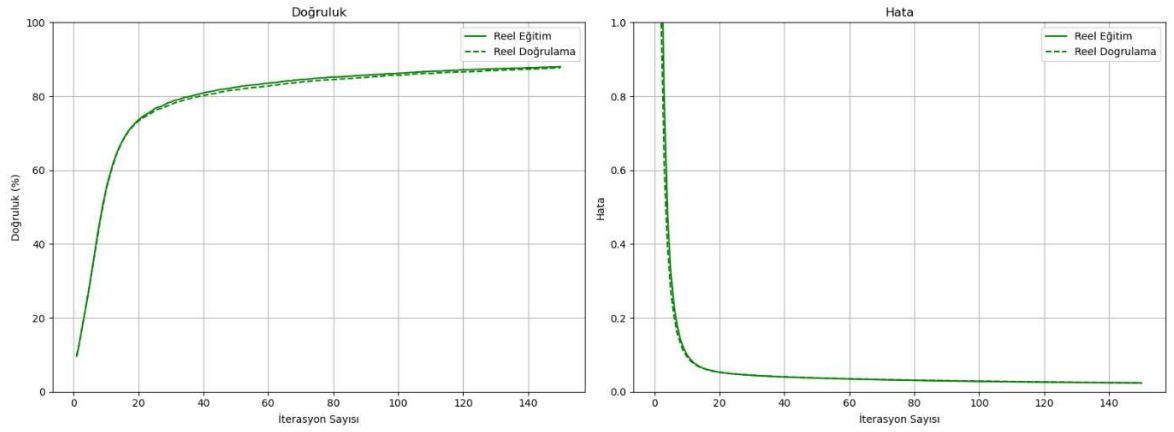
Veri Seti	Alfa	Eğitim		Doğrulama		Test	
		Doğruluk	Hata (MSE)	Doğruluk	Hata (MSE)	Doğruluk	Hata (MSE)
Digits	0.1	83.78	0.0614	77.31	0.0736	81.94	0.0710
	0.5	85.37	0.0605	79.40	0.0754	78.33	0.0724
	0.9	84.78	0.0617	77.08	0.0788	78.33	0.0793
	1.2	84.08	0.0611	76.16	0.0733	83.89	0.0674
	1.8	84.48	0.0601	81.71	0.0689	78.89	0.0740
Mnist Rotation	0.1	87.94	0.0247	87.54	0.0253	88.84	0.0238
	0.5	88.02	0.0240	87.68	0.0246	88.84	0.0229
	0.9	88.00	0.0239	87.72	0.0245	88.92	0.0228
	1.2	88.00	0.0239	87.72	0.0245	88.92	0.0228
	1.8	87.99	0.0239	87.70	0.0245	87.70	0.0229

Digits veri seti için KarcıFANN yönteminin eğitim ve doğrulama aşamalarında elde ettiği doğruluk ve hata grafikleri alfanın 1.2 değeri için Şekil 6.2 ile gösterilmektedir.



Şekil 6. 2: Digits veri setinin 1.2 alfa değeri için KarcıFANN yönteminin a) doğruluk ve b) hata grafikleri

MNIST Rotation 45° veri seti için KarcıFANN yönteminin eğitim ve doğrulama aşamalarında elde ettiği doğruluk ve hata grafikleri alfanın 0.9 değeri için Şekil 6.3 ile gösterilmektedir.



Şekil 6. 3: MNIST Rotation 45° veri setinin 0.9 alfa değeri için KarcıFANN yönteminin a) doğruluk ve b) hata grafikleri.

6.2 KarcıFANN Yönteminde Kompleks Değerlerin Uygulanması

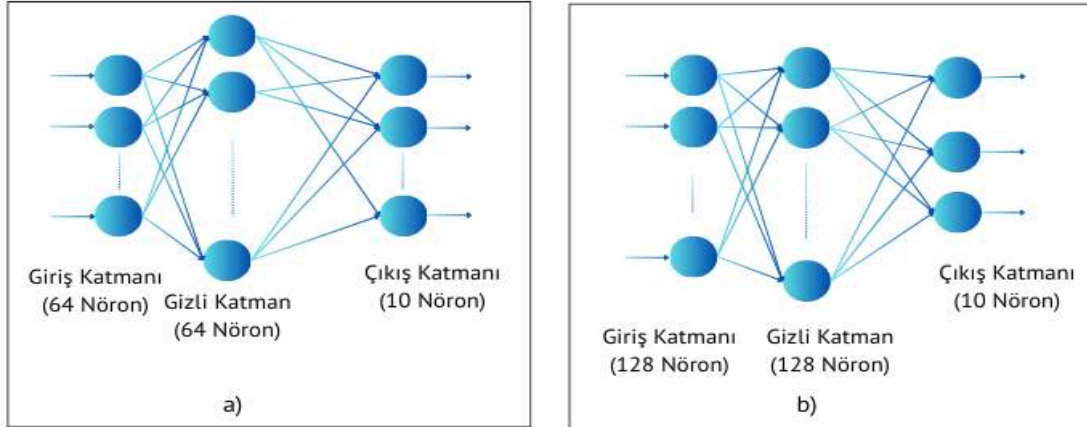
Çalışmanın bu bölümünde, KarcıFANN yöntemi ile YSA'ya ait ağırlıkların güncellenmesinde kullanılan alfa kesir derecesi parametresinin negatif tabanlı ifadelerde kompleks değerlerin ortaya çıkışını incelemek amacıyla yapılan deneysel çalışmalar sunulmaktadır. Bu amaçla geliştirilen yöntemin genelleme kabiliyeti ve farklı girdi temsilleri

üzerindeki performansları değerlendirilmek için yürütülen deneysel çalışmalarda Digits ve MNIST Rotation 45° veri setleri ve bu veri setlerinin 2BFT dönüşümü uygulanmış frekans uzayı karşılıkları ile oluşturulan 2BFT-Digits, 2BFT-MNIST Rotation 45° kullanılmıştır.

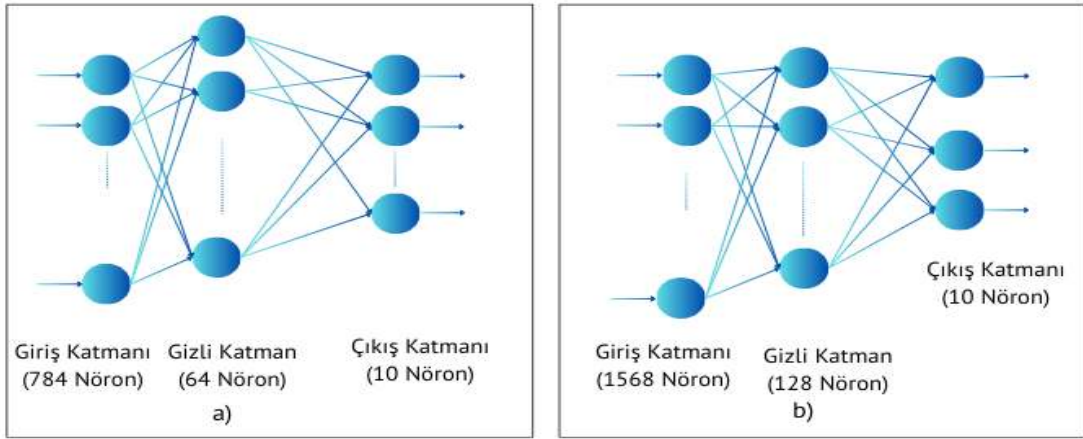
2BFT dönüşümlü veri setleri, modelin reel ve kompleks girdilerle çalışma performansını karşılaştırmak ve negatif tabanlı durumların oluşturduğu sanal bileşenin etkileri daha net gözlemlenmek amacıyla kullanılmıştır. Bu sayede KarcıFANN yönteminin hem uzaysal alanda hem de frekans alanında genelleme kabiliyeti değerlendirilmiştir.

Bu çalışmada kullanılan veri setleri için girdi boyutu ve sınıf sayıları dikkate alınarak uygun YSA mimarileri Şekil 6.4 ile gösterildiği gibi Digits ve 2BFT-Digits veri setleri için giriş katmanında sırasıyla 64 ve 128 nöron, gizli katmanda ise 128 nöron ve çıkış katmanlarında 10 nöron kullanılmıştır. MNIST Rotation 45° veri setinde giriş katmanında 784 nöron, gizli katmanda 64 nöron, 2BFT-MNIST Rotation 45° veri setinde ise frekans dönüşümü sonucu girdi boyutu iki katına çıkmış, giriş katmanı 1568 nöron, gizli katman ise 128 nöron olarak tasarlanmış ve Şekil 6.5 ile gösterilmektedir.

Eğitim ve test sürecinde alfa parametresi için rastgele belirlenen $\frac{3}{8}, \frac{5}{8}, \frac{7}{12}, \frac{9}{14}, \frac{5}{6}, \frac{3}{4}$ değerleri kullanılmıştır.



Şekil 6. 4: a) Digits, b) 2BFT-Digits veri setlerinin sınıflandırılmasında kullanılan model mimarisi.



Şekil 6. 5: a) Mnist Rotation 45, b) 2BFT-Mnist Rotation 45 veri setlerinin sınıflandırılmasında kullanılan model mimarisi.

Tüm veri setleri üzerinde gerçekleştirilen deneysel çalışmalarda, KarcıFANN yöntemi Reel, Kompleks ve Abs olmak üzere üç farklı ağ modeli ile test edilmiştir. Reel değerlerin işlendiği Reel modelde, ReLU aktivasyon fonksiyonu ve MSE hata fonksiyonu kullanılmıştır. Ağın eğitim sürecinde ağırlıkların güncellenmesi işlemi KarcıFANN yöntemi ile gerçekleştirilmiştir. Kompleks modelde, reel ve sanal bileşenler birlikte işlenmiştir. Karcı kesir dereceli türev yapısının doğası gereği, ağırlık güncelleme sürecinde ortaya çıkan karmaşık sayılar doğrudan çıkış katmanına aktarılmaktadır. Bu modelde Complex ReLU (CReLU) aktivasyon fonksiyonu ve Cross Entropy hata fonksiyonu kullanılmıştır. Kompleks katmanlar reel ve sanal bileşenleri ayrı ayrı saklayarak öğrenme sürecini yürütmektedir. Abs modelde ise, kompleks modelle aynı yapıda katmanlar kullanılmış, çıkış katmanında karmaşık çıktının modülü ($|z|$) alınarak sınıflandırma yapılmıştır. Böylece faz bilgisi göz ardı edilerek yalnızca genlik bilgisi üzerinden sınıflandırma gerçekleştirilmiştir. Bu modelde CReLU aktivasyon fonksiyonu ve MSE hata fonksiyonu kullanılmıştır.

Eğitim sürecinde iterasyon sayısı Digits, MNIST Rotation 45° ve 2BFT- MNIST Rotation 45° veri setlerinde iterasyon sayısı 150, 2BFT-Digits veri setinde ise öğrenme süreci tamamlanmadığı için iterasyon sayısı 300 olarak belirlenmiştir. Verilerin %80'i eğitim verisi, %20'si test verisi ve eğitim verilerinin %30'u doğrulama için ayrılmıştır. Başlangıçta tüm ağırlıklar için aynı reel değerler seçilmiştir. Başlangıçta, ağırlıkların reel bileşeni rastgele olarak, sanal bileşen ise 0 olarak başlatılmıştır. Eğitim süreci boyunca KarcıFANN yöntemi ile yapılan ağırlık güncelleme işleminden dolayı sanal bileşen zamanla değer kazanarak öğrenmeye katkıda bulunmuştur. Kompleks ve ABS modellerinde gizli katmanda CReLU aktivasyon fonksiyonu ve Dropout ($p = 0.15$) uygulanmıştır. Çıkış

katmanında elde edilen karmaşık değerler, faz ve genlik bileşenlerinin birlikte değerlendirilmesiyle sınıflandırmayı sağlamıştır.

Abs modeli, yapısal olarak Kompleks model ile benzer katmanlardan oluşmaktadır. Ancak, bu modelde sınıflandırma kararı verilirken karmaşık sayının sadece genlik bilgisi kullanılmaktadır. Bu genlik değeri, modelin çıktısı olarak kullanılır ve MSE hata fonksiyonu aracılığıyla gerçek etiketlerle karşılaştırılarak hata değeri hesaplanmaktadır.

Çizelge 6.2' de, Digits veri seti üzerinde KarıcıFANN yöntemi ile gerçekleştirilen deneyler sonucunda Reel, Kompleks ve Abs modellerine ait farklı alfa değerlerinde eğitim ve test doğruluk oranları, hata fonksiyonu değerleri ve kompleks model için reel ve sanal bileşenlerin etkisini gösteren sanal etki değeri sunulmuştur.

Reel model sonuçları incelendiğinde alfanın $\frac{3}{8}$ değeri için en yüksek eğitim doğruluğu %87.99, en yüksek test doğruluğu ise %88.92 olarak, en düşük eğitim hatası 0.0241 ve en düşük test hatası 0.0231 olarak elde edilmiştir.

Kompleks modelde eğitim doğruluğunun en yüksek değerlere alfanın $\frac{1}{2}, \frac{9}{14}, \frac{3}{4}, \frac{9}{8}$ değerleri için elde edildiği görülmektedir. Test veri kümesinde ise alfanın $\frac{3}{2}$ değeri için en yüksek test doğruluğu %97.78 ve en düşük hata değeri ise 0.1018 olarak kaydedilmiştir. Alfa'nın $\frac{5}{2}$ değeri için doğruluk değerinde düşüş ve hata değerinde artış gözlemlenmiştir. Sanal ve Reel bileşenlerin etkisi incelendiğinde alfanın $\frac{9}{8}$ değeri için sanal bileşenin en az etkili olduğu, $\frac{3}{2}$ değer için ise sanal etkinin en yüksek olduğu görülmektedir. Alfa= $\frac{3}{2}$ değeri için sanal bileşenin etkisi en yüksektir. Alfa değeri büyüdükçe reel baskınlığın arttığı ve sanal bilgi kaybının modelin kompleks temsil gücünü zayıflattığı gözlemlenmektedir.

Abs modelde ise $\frac{1}{2}, \frac{9}{14}, \frac{3}{4}, \frac{5}{6}, \frac{9}{8}$ değerleri için öğrenmenin başarılı bir şekilde gerçekleştiği görülmektedir. Fakat alfanın çok küçük ya da çok büyük olduğu durumlarda, modelin eğitim sürecinin başarısız olduğu gözlemlenmiştir.

Digits veri setinde yapılan sınıflandırmalar incelendiğinde Reel modelde öğrenme daha kararlı ve sınırlı bir doğruluk aralığında gerçekleşirken Kompleks modelde alfa parametresinin seçimi, modelin performansını ve sanal bileşenlerin katkısını önemli ölçüde etkilemektedir. Abs modelde ise karmaşık çıktının modülü ile sınıflandırma yapılması, kompleks model kadar başarılı bir öğrenme gerçekleştiremediğini göstermektedir.

Çizelge 6. 2: Digits veri setinde KarcıFANN yöntemi ile Reel, Kompleks ve Abs modellerine ait farklı alfa değerlerindeki Doğruluk ve Hata fonksiyonu değerleri.

Alfa	Reel				Kompleks							Abs			
	Eğitim Doğruluk %	Hata	Test Doğruluk %	Hata	Eğitim Doğruluk %	Hata	Test Doğruluk %	Hata	Reel Etki	Sanal Etki	Reel/Sanal	Eğitim Doğruluk %	Hata	Test Doğruluk %	Hata
3/8	87.99	0.0241	88.92	0.0231	99.50	0.1031	96.16	0.1833	0.742	0.525	1.413	52.01	0.0831	56.30	0.0720
1/2	83.18	0.0625	77.22	0.0751	99.90	0.0458	95.28	0.1504	0.738	0.522	1.414	91.04	0.0495	85.28	0.0455
9/14	82.29	0.0636	76.67	0.0793	99.90	0.0354	94.72	0.2623	0.753	0.506	1.488	90.35	0.0498	83.61	0.1779
3/4	81.89	0.0633	76.94	0.0734	99.90	0.0365	95.83	0.2606	0.802	0.425	1.887	92.34	0.0493	87.22	0.1506
5/6	85.27	0.0616	81.94	0.0711	99.80	0.0497	96.94	0.1493	0.817	0.416	1.964	92.44	0.0500	89.44	0.0478
9/8	83.48	0.0616	79.72	0.0730	99.90	0.0267	95.83	0.1893	0.863	0.353	2.445	91.44	0.0467	85.83	0.0450
3/2	83.78	0.0617	77.78	0.0742	99.80	0.0490	97.78	0.1038	0.702	0.567	1.238	14.13	1.4178	11.39	1.7642
5/2	82.69	0.0664	78.61	0.0855	91.94	0.3987	90.28	0.3592	0.731	0.529	1.382	8.26	248.3893	15.56	4.7795

Çizelge 6.3'te MNIST Rotation 45° veri seti üzerinde KarıcıFANN yöntemi ile test edilen Reel, Kompleks ve Abs modellerine ait farklı alfa değerlerindeki eğitim ve test doğrulukları, hata fonksiyonu değerleri ve Kompleks model için sanal etki değerleri sunulmuştur.

Reel modelde, alfa değeri değiştikçe eğitim ve test doğrulukları ile hata değerlerinin birbirine yakın ve dengeli kaldığı görülmektedir. Bu durum, Reel modelin daha kararlı bir performans sergilediği fakat sınırlı duyarlılıkta bir performans sergilediğine işaret etmektedir.

Kompleks modelde, alfanın $\frac{9}{8}$ değeri için en yüksek performans elde edilmiştir. Eğitim ve test veri kümelerinde doğruluk oranının en yüksek seviyede olduğu görülürken hata değerinin ise en düşük seviyede olduğu gözlemlenmektedir. Fakat alfanın $\frac{5}{2}$ değeri için modelin öğrenemediği çizelgeden anlaşılmaktadır. Reel/Sanal etki oranları incelendiğinde, alfa değerinin artışına paralel olarak reel bileşen belirgin biçimde etkili olduğu gözlenmiştir. Bu durum, faz etkisinin zayıflaması ve sanal bileşen akışının giderek sınırlanması ile ilişkilendirilebilir. Dolayısıyla, bu veri setinde bilgi taşınımında reel bileşenin daha etkin bir rol üstlendiği, buna karşılık karmaşık düzlemdeki faz katkısının azaldığı sonucuna ulaşılmaktadır. Genel olarak, elde edilen bulgular alfa parametresinin karmaşık uzaydaki bilgi akışını ve bileşenler arasındaki dengeyi doğrudan etkileyen belirleyici bir faktör olduğunu ortaya koymaktadır.

Abs modelde, $\frac{1}{2}$ ve $\frac{9}{8}$ aralığında sınırlı bir öğrenme gerçekleştiği doğruluk oranlarından anlaşılmaktadır. Fakat $\frac{5}{2}$ alfa değeri için elde edilen doğruluk oranları ve hata değerlerinden modelin eğitilemediği ve öğrenme sürecinin başarısız olduğu görülmektedir.

Çizelge 6. 3: MNIST Rotation 45° veri setinde KarcıFANN yöntemi ile Reel, Kompleks ve Abs modellerine ait farklı alfa değerlerindeki Doğruluk ve Hata fonksiyonu değerleri.

Alfa Değeri	Reel				Kompleks							Abs			
	Eğitim		Test		Eğitim		Test		Reel Etki	Sanal Etki	Reel/ Sanal	Eğitim		Test	
	Doğruluk %	Hata	Doğruluk %	Hata	Doğruluk %	Hata	Doğruluk %	Hata				Doğruluk %	Hata	Doğruluk %	Hata
3/8	87.99	0.0241	88.92	0.0231	98.05	0.1031	96.16	0.1833	0.881	0.346	2.546	52.01	0.0831	56.30	0.0720
1/2	88.02	0.0240	88.84	0.0229	99.42	0.0571	96.89	0.1617	0.906	0.288	3.146	70.66	0.0748	74.76	0.0617
9/14	88.00	0.0239	88.92	0.0228	99.66	0.0453	96.78	0.1694	0.898	0.292	3.075	73.43	0.0744	77.48	0.0569
3/4	88.00	0.0239	88.92	0.0228	99.67	0.0398	96.33	0.1856	0.852	0.360	2.367	73.68	0.0753	77.96	0.0582
5/6	88.00	0.0239	88.92	0.0228	99.81	0.0368	96.94	0.1794	0.844	0.388	2.175	74.02	0.0747	78.07	0.0572
9/8	88.00	0.0239	88.92	0.0228	99.96	0.0261	97.05	0.1907	0.798	0.463	1.724	71.73	0.0730	77.84	0.0569
3/2	88.00	0.0239	88.92	0.0228	99.80	0.0391	96.82	0.2032	0.886	0.360	2.461	37.51	0.1446	36.01	0.2394
5/2	86.35	0.0285	87.51	0.0275	13.56	1101.8954	10.09	2.3905	0.730	0.538	1.357	9.93	70068990883.5474	10.50	799446839.0912

Çizelge 6.4'te, 2BFT-Digits veri setinde KarcıFANN yöntemi kullanılarak gerçekleştirilen deneysel çalışmalara ait sonuçlar sunulmaktadır. Bu çizelge, Reel, Kompleks ve Abs modellerinin farklı alfa değerlerindeki eğitim ve test doğruluk oranları ile hata fonksiyonu değerlerini ve Kompleks model için sanal etki değerlerini göstermektedir.

Reel model, tüm alfa değerlerinde eğitim doğruluk oranı %73.13 değerinde sabit kaldığı, test doğruluk oranı birçok alfa değeri için %61.67 olarak sabit kaldığı ve hata değerlerinin tüm alfa değerlerinde birbirine yakın değerler olduğu görülmektedir. Bu durum, Reel modelin 2BFT-Digits veri setinde öğrenme kapasitesinin sınırlı olduğunu ve yeterli olmadığını, alfa değişiminin performans üzerinde bir etkisinin olmadığını göstermektedir.

Kompleks modelde eğitim doğruluk oranının %96.12 - %99.90, test doğruluk oranının ise %90.56 ile %95.00 arasında değiştiği görülmektedir. Alfa'nın $\frac{9}{8}$ değeri için eğitim veri kümesinde en yüksek doğruluk oranı ve en düşük hata değerleri elde edilmiştir.

Reel/sanal değerleri incelendiğinde alfa $\frac{1}{2}$ değerinde Reel/Sanal oranının 1'in altında olması, sanal bileşenin reel bileşene göre daha etkin olduğunu göstermektedir. Bu durum, faz bileşeninin öğrenme sürecine belirgin katkı sunduğunu ve modelin bu aralıkta faz duyarlılığının en yüksek seviyede olduğunu göstermektedir. Alfa'nın diğer değerlerinde reel bileşen baskın olsa da oranların 1'e yakın olması, faz bileşeninin öğrenme sürecine anlamlı düzeyde katkı sunmuştur. Bu durum, modelin karmaşık uzaydaki bilgi akışında reel ve sanal bileşenlerin birlikte aktif rol oynadığını, fakat reel bileşenin genlik tabanlı bilgi taşınımında daha belirleyici konumda olduğunu göstermektedir.

Abs modelde ise tüm alfa değerleri için doğruluk oranları ve hata değerleri incelendiğinde öğrenmenin gerçekleşmediği anlaşılmaktadır. Bu durumun sebebi, modelin sadece genlik $|z|$ bilgisinden yararlanması ve 2BFT dönüşümündeki faz bilgisini kaybetmesinden kaynaklanmaktadır.

Çizelge 6. 4: 2BFT-Digits veri setinde KarcıFANN yöntemi ile Reel, Kompleks ve Abs modellerine ait farklı alfa değerlerindeki Doğruluk ve Hata fonksiyonu değerleri.

Alfa	Reel				Kompleks							Abs			
	Eğitim Doğruluk %	Hata	Test Doğruluk %	Hata	Eğitim Doğruluk %	Hata	Test Doğruluk %	Hata	Reel Etki	Sanal Etki	Reel/ Sanal	Eğitim Doğruluk %	Hata	Test Doğruluk %	Hata
3/8	73.13	0.1839	63.06	0.1110	96.12	0.2137	91.39	0.2837	0.672	0.601	1.118	12.64	1.3535	16.11	0.7519
1/2	73.13	0.1831	61.67	0.1165	98.21	0.1287	94.44	0.2272	0.632	0.637	0.992	17.61	0.9224	18.06	0.4709
9/14	73.13	0.1831	61.67	0.1165	99.60	0.0889	94.17	0.2150	0.682	0.600	1.137	22.39	0.5747	18.06	0.2916
3/4	73.13	0.1831	61.67	0.1165	99.70	0.0721	94.72	0.2068	0.694	0.585	1.186	24.78	0.4555	19.17	0.2205
5/6	73.13	0.1831	61.67	0.1165	99.90	0.0641	95.00	0.2014	0.691	0.586	1.179	27.46	0.4000	20.00	0.2049
9/8	73.13	0.1831	61.67	0.1165	99.90	0.0614	95.00	0.2027	0.671	0.593	1.132	26.27	0.3746	18.89	0.1865
3/2	73.13	0.1831	61.67	0.1165	98.11	0.1390	92.78	0.2554	0.669	0.606	1.104	13.83	2.6284	11.11	4.0135
5/2	73.13	0.1831	61.67	0.1165	96.22	0.2303	90.56	0.3142	0.650	0.619	1.050	9.85	15.1791	9.17	5.1568

Çizelge 6.5'te, 2BFT-MNIST Rotation 45° veri seti üzerinde KarıcıFANN yöntemi ile eğitilen Reel, Kompleks ve Abs modellerinin farklı alfa değerleri için elde edilen eğitim ve test doğruluk oranları, hata fonksiyonu değerleri ile kompleks model için sanal etkinin sayısal karşılığı sunulmaktadır.

Reel model, eğitim ve test doğruluk oranları yaklaşık %50 civarında sabit kaldığı, hata değerlerinin ise 0.071 civarında sabit olup alfa değerinin değişiminden etkilenmediği görülmektedir. Bu durum, Reel modelin 2BFT-MNIST Rotation 45° veri setinde karmaşık verinin faz bilgisini işleyemediğini ve öğrenmenin gerçekleşmediğini göstermektedir.

Kompleks modelin eğitim aşamasındaki doğruluk oranları %79.96 – %95.74 aralığında, test aşamasında ise %79.04 – %92.03 arasında değişmektedir. En iyi test doğruluğu %92.03 ile alfanın $\frac{9}{8}$ değerinde elde edilmiştir. Alfa değeri $\frac{5}{2}$ için eğitim ve test doğruluklarında belirgin bir düşüş gözlemlenmiştir. Alfanın $\frac{3}{8}$, $\frac{3}{2}$ ve $\frac{5}{2}$ değerlerinde sanal bileşenin etkisinin yüksek olduğu, $\frac{1}{2}$ değeri için sanal ve reel bileşenlerin öğrenme üzerinde dengeli bir etkiye sahip olduğu ve diğer değerler için de reel bileşenin öğrenme üzerinde daha etkili olduğu görülmektedir.

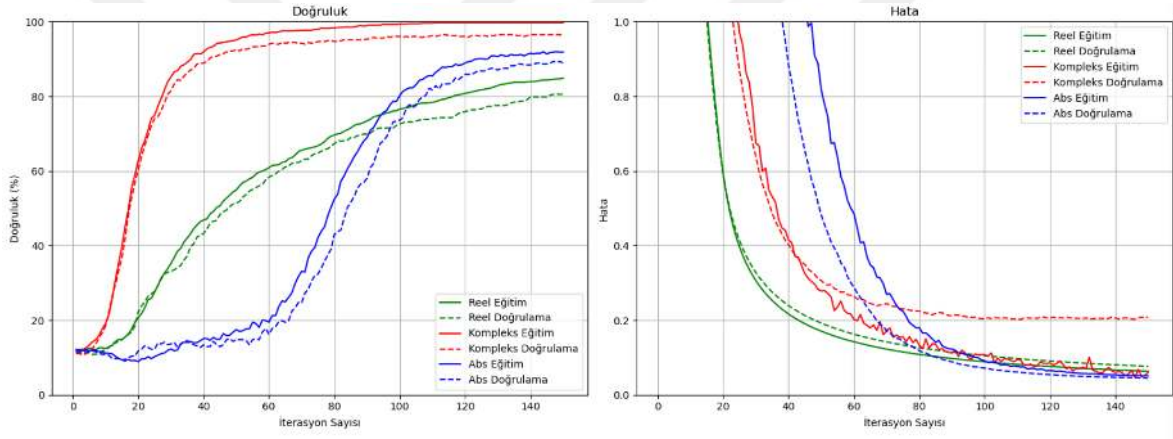
Abs modelinde ise eğitim ve test doğruluklarının %64 - %71 aralığında olduğu, öğrenmenin yeterli seviyeye ulaşmadığı görülmektedir. Özellikle $\frac{3}{2}$ ve $\frac{5}{2}$ değerlerinde modelin tamamen başarısız olduğu görülmektedir. Bu durum, sadece genlik bilgisinin kullanımının yüksek alfa değerlerinde yeterli olmadığını ve modelin eğitilemediğini göstermektedir.

Çizelge 6. 5: 2BFT- MNIST Rotation 45° veri setinde KarcıFANN yöntemi ile Reel, Kompleks ve Abs modellerine ait farklı alfa değerlerindeki Doğruluk ve Hata fonksiyonu değerleri.

Alfa Değeri	Reel				Komplex							Abs			
	Eğitim		Test		Eğitim		Test		Reel Etki	Sanal etki	Reel/ Sanal	Eğitim		Test	
	Doğruluk %	Hata	Doğruluk %	Hata	Doğruluk %	Hata	Doğruluk %	Hata				Doğruluk %	Hata	Doğruluk %	Hata
3/8	50.70	0.0716	50.28	0.0690	90.33	0.4132	88.42	0.3792	0.627	0.638	0.983	63.05	0.0695	64.01	0.0667
1/2	50.45	0.0717	50.12	0.0690	92.84	0.3177	90.36	0.3214	0.647	0.622	1.040	67.24	0.0671	68.77	0.0637
9/14	50.45	0.0717	50.12	0.0690	94.23	0.2602	91.08	0.3017	0.694	0.576	1.205	68.40	0.0654	69.81	0.0618
3/4	50.45	0.0717	50.12	0.0690	95.05	0.2305	91.77	0.2847	0.689	0.583	1.182	68.97	0.0648	70.50	0.0611
5/6	50.45	0.0717	50.12	0.0690	95.49	0.2154	91.81	0.2772	0.696	0.575	1.210	69.27	0.0644	70.84	0.0607
9/8	50.45	0.0717	50.12	0.0690	95.74	0.2066	92.03	0.2079	0.703	0.563	1.249	69.41	0.0643	70.95	0.0606
3/2	50.45	0.0717	50.12	0.0690	92.48	0.3283	90.34	0.3293	0.619	0.661	0.936	12.73	1.6152	11.87	9.4243
5/2	50.45	0.0717	50.12	0.0690	79.96	0.7862	79.04	0.6990	0.670	0.615	1.089	10.39	347.4918	11.71	7.9350

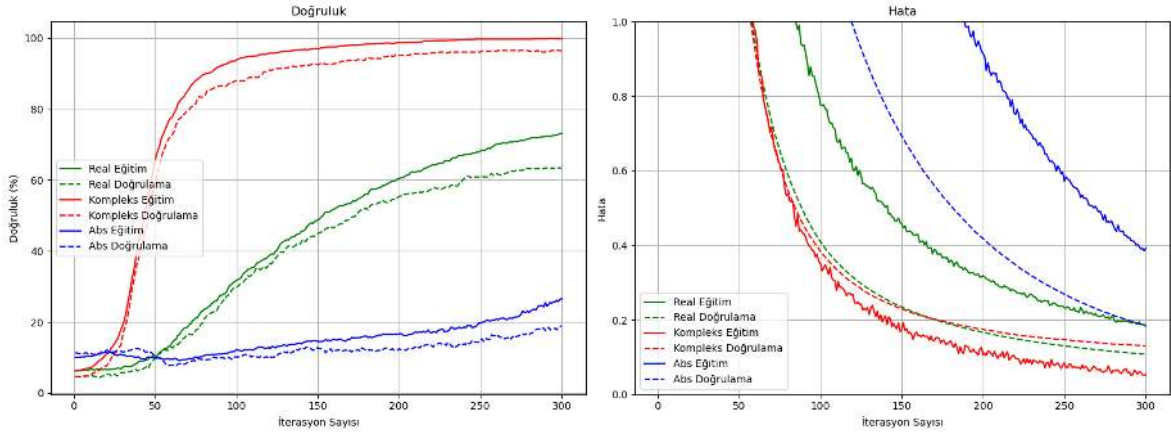
Genel olarak tüm veri setlerinde 3 model değerlendirildiğinde, Kompleks model en iyi performansı sergilerken Reel model, temel ve düşük karmaşıklıkta veri setlerinde kısmen başarılı olmuş ancak 2BFT dönüşümlü veri setlerinde yetersiz kalmıştır. Abs model ise orta düzeyde performans göstermiş, ancak faz bilgisini kullanamadığı için karmaşık verilerde yüksek doğruluk seviyelerine ulaşamamaktadır. Bu sonuçlar, karmaşık verilerin işlenmesinde reel ve sanal bileşenlerin birlikte değerlendirilmesinin öğrenme başarısı açısından önemli olduğunu ortaya koymaktadır.

Şekil 6.6'da gösterilen Digits veri setine ait alfanın $\frac{9}{8}$ değeri için doğruluk ve hata grafikleri incelendiğinde Kompleks modelin eğitim sürecinde daha kararlı ve hızlı bir öğrenme performansı sergilediği görülmektedir.



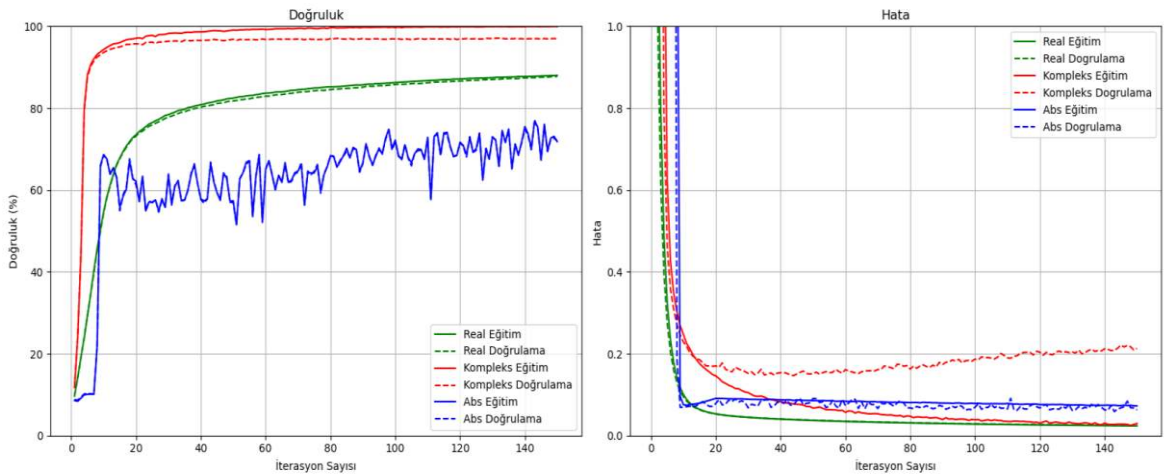
Şekil 6. 6: Digits veri setinin 9/8 alfa değeri için KarşıFANN yönteminin a) doğruluk ve b) hata grafikleri

Şekil 6.7'de 2BFT-Digits veri seti üzerinde üç modelin alfa 9/8 değeri için eğitim ve doğrulama süreçlerine ait doğruluk ve hata değerleri gösterilmektedir. Kompleks model, hızlı öğrenme, yüksek doğruluk ve düşük hata değerleri ile diğer modellere daha başarılı bir performans göstermiştir.



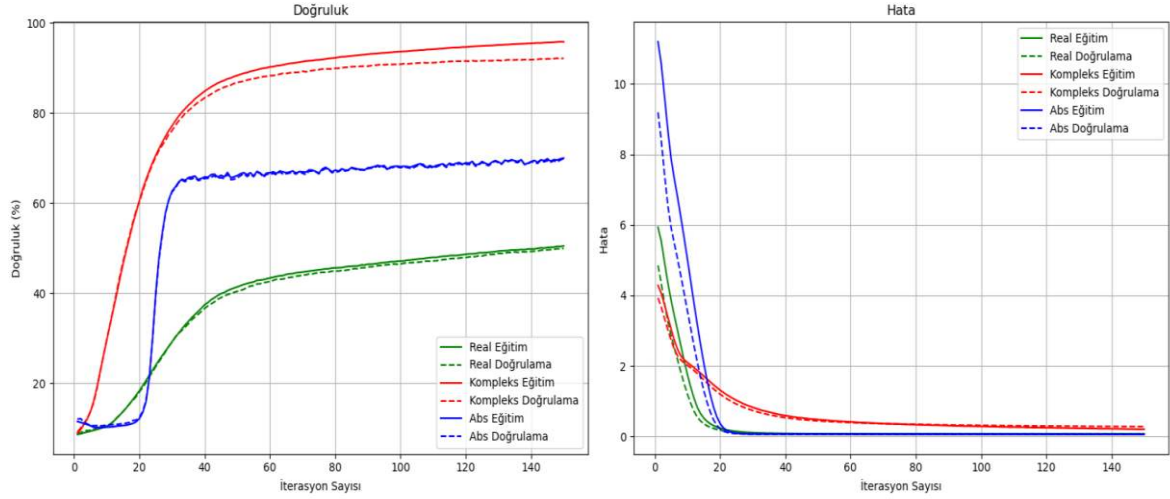
Şekil 6. 7: 2BFT- Digits veri setinin 9/8 alfa değeri için KarcıFANN yönteminin a) doğruluk ve b) hata grafikleri

Şekil 6.8’de MNIST Rotation 45° veri seti üzerinde üç modelin alfa $\frac{9}{8}$ değeri için eğitim ve doğrulama süreçlerine ait doğruluk ve hata değerleri gösterilmektedir. Karmaşık bileşenlerin birlikte işlendiği Kompleks model, öğrenme hızı ve doğruluk performansı açısından diğer modellere göre belirgin bir üstünlük göstermiştir. Fakat doğrulama hatasının eğitim hatasının üzerinde olması, aşırı uyum ihtimaline işaret etmektedir. Reel model, daha yavaş bir öğrenme sergilemiş olsa da dengeli ve genelleme açısından kabul edilebilir sonuçlar üretmiştir. Abs model ise doğruluk grafiğindeki dalgalanmadan öğrenme başarısının yetersiz olduğu anlaşılmaktadır.



Şekil 6. 8: MNIST Rotation 45° veri setinin 9/8 alfa değeri için KarcıFANN yönteminin a) Doğruluk ve b) Hata grafikleri

Şekil 6.9’da 2BFT- MNIST Rotation 45° veri seti üzerinde üç modelin alfa $\frac{9}{8}$ değeri için eğitim ve doğrulama veri kümelerine ait doğruluk ve hata değerleri gösterilmektedir. Kompleks model, en başarılı model olarak öne çıkmaktadır. Reel model, düzenli ve stabil bir öğrenme göstermesine rağmen doğruluk oranı sınırlı kalmış ve yeterli bir öğrenme gerçekleştirememiştir. Abs model, erken doygunluğa ulaşmış öğrenme yetersiz kalmıştır.



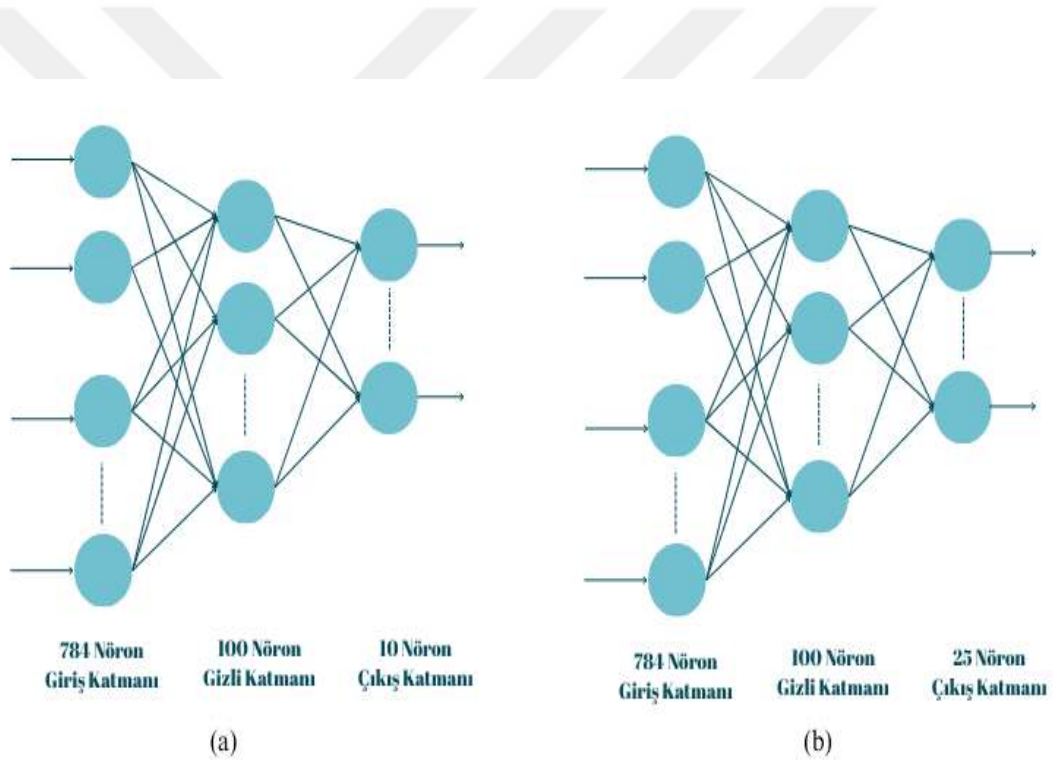
Şekil 6. 9: 2BFT- MNIST Rotation 45° veri setinin 9/8 alfa değeri için KarcıFANN yönteminin a) doğruluk ve b) hata grafikleri

6.3 Karcı kesir dereceli YSA (KarcıFANN) ile öğrenme katsayısı, aşırı uyum ve yetersiz uyum sorunlarını çözme

Bu tez kapsamında Karcı kesir dereceli türev kullanımının öğrenme katsayısı yerine kullanılması durumunda modelin başarısı test edilirken yapılan çalışmalarda SignMNIST, Kuzushiji-MNIST ve Gina_prior2 veri setleri kullanılmıştır.

YSA’larda hiperparametre seçimi öğrenme başarısını önemli ölçüde etkilemektedir. Uygun öğrenme katsayısının seçilmemesi önerilen modelin ezberleme, öğrenememe, gradyan kaybolması, gradyan patlaması ya da öğrenme sürecini geciktirme gibi birçok problemlere neden olmaktadır. Bu çalışma kapsamında ağırlık güncelleme sürecinde kesir dereceli türev kullanan KarcıFANN yöntemi ile öğrenme katsayısı kullanan Klasik YSA yöntemi ile karşılaştırılarak aşırı uyum ve yetersiz uyum problemlerine çözüm sunmayı amaçlamaktadır. Çalışmada Kuzushiji_MNIST, GinaPrior2 ve SignMnist veri setleri kullanılarak sonuçları Klasik YSA yöntemi ile karşılaştırılmıştır.

Gina_prior2 ve Kuzushiji-MNIST veri setleri için 784 nöron içeren giriş katman, 100 nöron içeren gizli katman ve 10 nöron içeren çıkış katmanı, Sign-MNIST veri seti için 784 nöron içeren giriş katman, 100 nöron içeren gizli katman ve 25 nöron içeren çıkış katmanı olmak üzere 3 katmanlı bir YSA modeli tasarlanmıştır. Tüm veri setlerinde eşit şartlar altında karşılaştırmayı yapmak için aynı rastgele başlangıç değerleri seçilerek başlatılmış, ağların eğitim tur sayısı (epoch) 1000 olarak ayarlanmış ve veri setlerinin %30'u doğrulama sürecinde kullanılmıştır. Eğitim ve test sürecinde KarcıFANN kullanılan Karcı kesir derecesi ve Klasik YSA yönteminde kullanılan öğrenme katsayısı değeri için 0.1-5.0 aralığında değerler verilmiştir. Şekil 6.10 veri setlerini sınıflandırmada kullanılan ağ modellerini göstermektedir.



Şekil 6. 10: Veri Setinin Sınıflandırılmasında Kullanılan Ağ Modelleri a) Gina_prior2 ve Kuzushiji-MNIST veri setleri b) Sign-MNIST veri seti.

Çizelge 6.6'da, Kuzushiji-MNIST veri seti üzerinde yapılan sınıflandırma sonuçları KarcıFANN ve Klasik YSA yöntemleri için karşılaştırmalı olarak sunulmuştur. Alfa ve öğrenme katsayısı parametrelerinin 0.1–0.6 aralığı ile 0.8, 2.0, 2.1, 2.2, 2.3, 2.4, 2.6, 2.7 ve 3.2 değerlerinde, Klasik YSA yöntemi tüm veri kümelerinde daha iyi performans göstermiştir. Buna karşılık, alfa ve öğrenme katsayısı 0.7 değeri için hem eğitim hem de test aşamalarında, 0.9 için ise eğitim aşamasında KarcıFANN yöntemi daha düşük hata değeri

üretmiştir. Alfa değeri 1.0 olduğunda, KarcıFANN yöntemi Klasik YSA yöntemi ile aynı şekilde çalışmaktadır. Alfa değeri 1.1 için eğitim aşamasında, 1.2, 1.7 ve 2.5 için test aşamasında KarcıFANN yöntemi daha başarılı sonuçlar elde etmiştir. Ayrıca, 1.3, 1.4, 1.5, 1.6, 1.8 ve 1.9 değerleri için test aşamasında KarcıFANN yöntemi daha yüksek doğruluk değerleri sağlamış olsa da, diğer veri kümelerinde Klasik YSA yöntemi daha düşük hata ve daha yüksek doğruluk değerleriyle başarılı performans göstermiştir. Alfa ve öğrenme katsayısı değerleri 2.8, 2.9, 3.0 ile 3.3–5.0 aralığında incelendiğinde, hem eğitim hem de test aşamalarında KarcıFANN yöntemi tüm veri kümelerinde Klasik YSA'ya kıyasla daha başarılı sonuçlar üretmiştir. Özellikle 3.1 değeri için KarcıFANN yöntemi, eğitim ve test doğruluk değerleri ile ve test hata değeri incelendiğinde Klasik YSA yöntemine kıyasla daha üstün performans sergilemiştir. KarcıFANN yönteminin üstün olduğu alfa değerlerinde, Karcı kesir dereceli türevinin Newton türevine kıyasla ağırlık güncelleme sürecinde daha etkili olduğu görülmektedir.

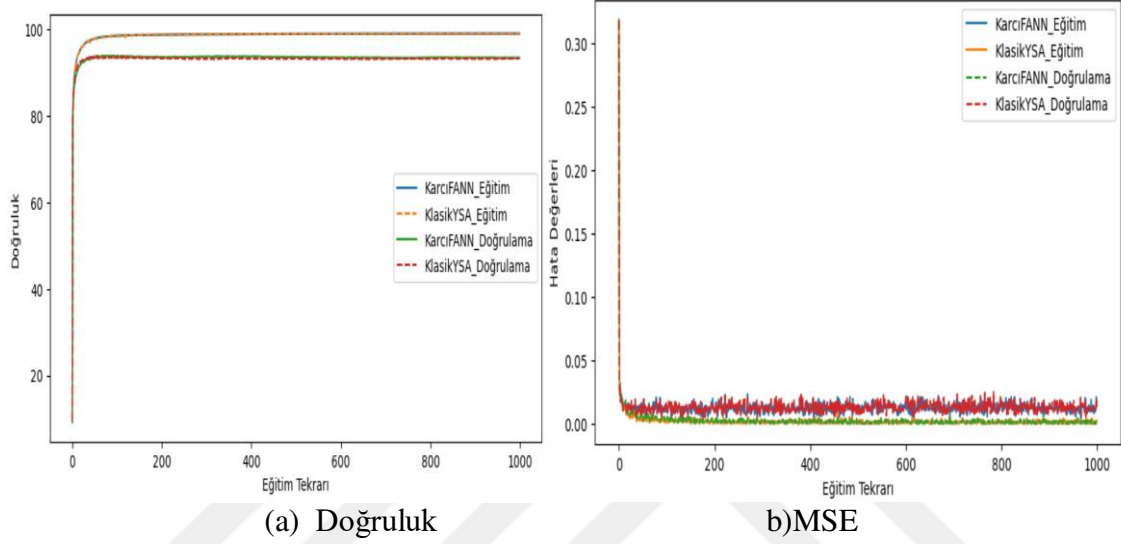
Çizelge 6. 6: Kuzushiji_MNIST veri setinde KarcıFANN ve Klasik NN yöntemi için alfa değerlerindeki değişimlere göre elde edilen Doğruluk ve MSE değerleri.

Alfa/ Öğrenme Katsayısı	KarcıFANN Eğitim		Klasik YSA Eğitim		KarcıFANN Test		Klasik YSA Test	
	Doğruluk%	MSE	Doğruluk%	MSE	Doğruluk%	MSE	Doğruluk%	MSE
0.1	10.00	0.100000	98.58	0.003504	10.00	0.100000	92.87	0.012386
0.2	10.00	0.100000	98.79	0.002167	10.00	0.100000	93.08	0.013525
0.3	10.01	0.100000	98.89	0.002120	10.00	0.100000	92.84	0.013983
0.4	90.38	0.017966	98.97	0.002096	86.65	0.022683	93.01	0.014912
0.5	88.97	0.016096	99.00	0.001090	82.54	0.022048	92.92	0.017496
0.6	88.45	0.007513	99.00	0.001055	82.60	0.028565	92.60	0.016700
0.7	98.59	0.000174	99.03	0.002046	90.44	0.009429	92.79	0.016246
0.8	98.90	0.002007	98.96	0.001030	92.31	0.016840	92.79	0.016575
0.9	98.91	0.002011	98.97	0.002034	92.66	0.015820	92.96	0.015083
1.0	98.91	0.001029	98.91	0.001029	92.89	0.013561	92.89	0.013561
1.1	98.95	0.001126	98.92	0.002021	92.82	0.016851	93.05	0.015817
1.2	98.85	0.002133	98.91	0.001026	93.21	0.013824	92.89	0.014343
1.3	98.76	0.003077	98.92	0.002023	93.29	0.014654	92.88	0.013345
1.4	98.57	0.003817	98.95	0.001019	93.17	0.017301	92.51	0.013230
1.5	98.39	0.005080	98.96	0.001019	93.01	0.017436	92.91	0.016314
1.6	98.10	0.005108	98.91	0.002037	93.04	0.015569	92.18	0.014794
1.7	97.65	0.006979	98.87	0.003017	93.20	0.016265	92.65	0.018867
1.8	97.07	0.007644	98.91	0.003976	93.02	0.017887	92.36	0.014313
1.9	96.27	0.008584	98.87	0.002013	92.80	0.018364	92.35	0.015353
2.0	95.42	0.009870	98.86	0.003016	92.30	0.018054	92.49	0.013079
2.1	94.48	0.011905	98.90	0.003012	91.87	0.017395	92.65	0.012580
2.2	93.36	0.013851	98.91	0.001014	91.01	0.017655	92.58	0.013432
2.3	92.23	0.015957	98.89	0.000017	90.33	0.018744	92.48	0.012764
2.4	91.03	0.018012	98.88	0.001019	89.24	0.020154	92.30	0.014367
2.5	89.69	0.019724	90.90	0.012023	88.03	0.021742	86.51	0.027128

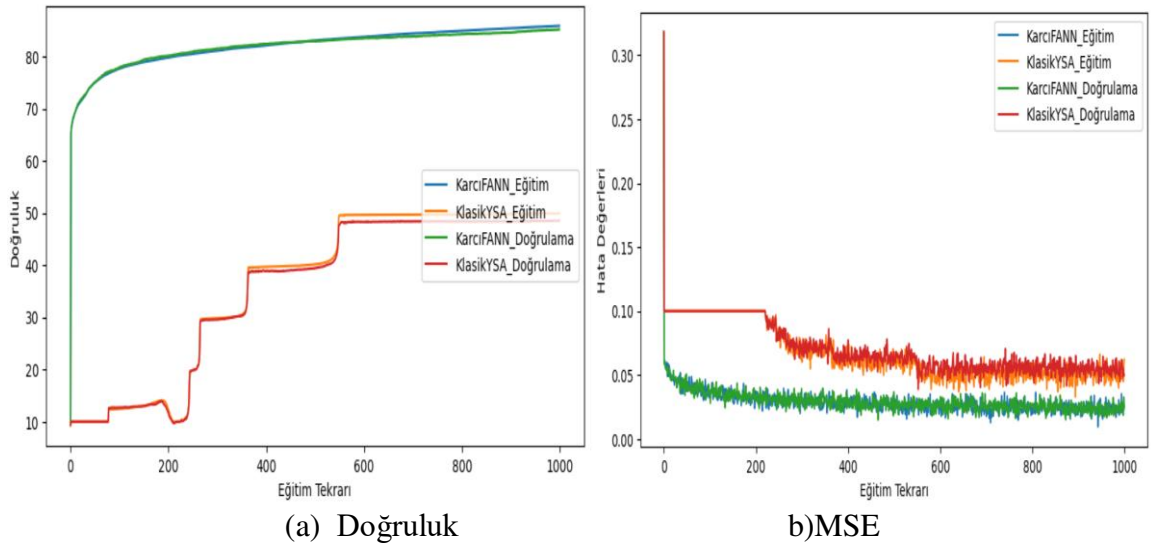
Çizelge 6. 7 (devam): Kuzushiji_MNIST veri setinde KarcıFANN ve Klasik NN yöntemi için alfa değerlerindeki değişimlere göre elde edilen Doğruluk ve MSE değerleri.

Alfa/ Öğrenme Katsayısı	KarcıFANN Eğitim		Klasik YSA Eğitim		KarcıFANN Test		Klasik YSA Test	
	Doğruluk%	MSE	Doğruluk%	MSE	Doğruluk%	MSE	Doğruluk%	MSE
2.6	88.40	0.021515	99.02	0.001012	86.88	0.023445	92.42	0.017539
2.7	87.11	0.023623	99.07	0.001017	85.76	0.025461	92.24	0.016670
2.8	85.86	0.025699	49.92	0.062004	84.74	0.027795	48.60	0.058331
2.9	87.11	0.023623	40.04	0.067003	85.76	0.025461	39.18	0.066229
3.0	83.39	0.029214	10.00	0.100000	82.38	0.031704	10,00	0.100000
3.1	82.23	0.030874	78.76	0.030430	82.38	0.031704	73.11	0.045205
3.2	72.80	0.042910	79.45	0.025037	71.81	0.045744	75.19	0.034090
3.3	82.23	0.030874	20.97	0.083002	81.44	0.033274	20.85	0.084815
3.4	79.06	0.036421	60.22	0.050006	78.50	0.037268	58.05	0.051181
3.5	74.29	0.042523	10,00	0.100000	73.10	0.045885	10,00	0.100000
3.6	74.09	0.050010	10,00	0.100000	73.56	0.049867	10,00	0.100000
3.7	75.54	0.043006	10,00	0.100000	74.77	0.043241	10,00	0.100000
3.8	74.47	0.045128	10,00	0.100000	73.71	0.044851	10,00	0.100000
3.9	72.93	0.046428	10,00	0.100000	72.38	0.046775	10,00	0.100000
4.0	71.10	0.050192	10,00	0.100000	70.66	0.050247	10,00	0.100000
4.1	70.09	0.050819	10,00	0.100000	70.05	0.051685	10,00	0.100000
4.2	59.22	0.065300	10,00	0.100000	58.99	0.067195	10,00	0.100000
4.3	67.79	0.056553	10,00	0.100000	67.80	0.056646	10,00	0.100000
4.4	66.69	0.059353	10,00	0.100000	66.60	0.059131	10,00	0.100000
4.5	66.65	0.059081	10,00	0.100000	66.65	0.058858	10,00	0.100000
4.6	63.17	0.065494	10,00	0.100000	62.81	0.064726	10,00	0.100000
4.7	60.89	0.068290	10,00	0.100000	60.88	0.067453	10,00	0.100000
4.8	28.36	0.035152	10,00	0.100000	28.57	0.035554	10,00	0.100000
4.9	56.03	0.074459	10,33	0.100000	55.87	0.073311	10,31	0.100000
5.0	53.77	0.077255	10,00	0.100000	53.81	0.075860	10,00	0.100000

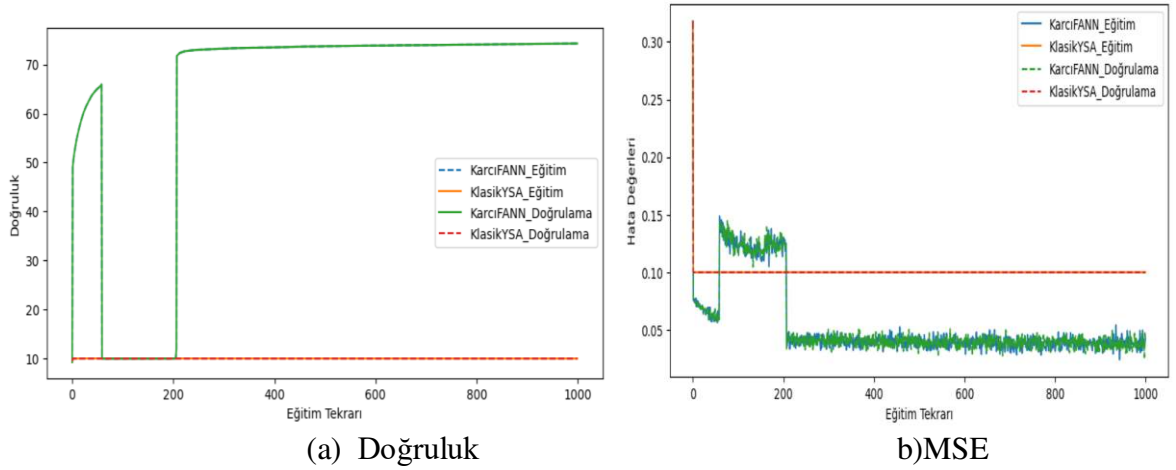
Kuzushiji-MNIST veri seti için KarcıFANN ve Klasik YSA yöntemlerinin eğitim ve doğrulama aşamalarında elde ettiği doğruluk ve hata grafikleri alfa ve öğrenme katsayısı 1.1 değeri için Şekil 6.11, 2.8 değeri için Şekil 6.12, 3 ve 3.5 değeri için Şekil 6.13 ile gösterilmektedir.



Şekil 6. 11: Kuzushiji_Mnist veri setinde alfa ve öğrenme katsayılarının 1.1 değeri için KarcıFANN ve Klasik YSA yönteminin a) Doğruluk ve b) MSE grafikleri.



Şekil 6. 12: Kuzushiji_Mnist veri setinde alfa ve öğrenme katsayılarının 2.8 değeri için KarcıFANN ve Klasik YSA Yönteminin a) Doğruluk ve b) MSE grafikleri.



Şekil 6. 13: Kuzushiji_Mnist veri setinde alfa ve öğrenme katsayılarının 3.5 değeri için KarciFANN ve Klasik YSA Yönteminin a) Doğruluk ve b) MSE grafikleri.

Çizelge 6.7’de, Gina_prior2 veri setinde KarciFANN ve Klasik YSA yöntemleri için sınıflandırma sonucunda elde edilen doğruluk ve MSE değerleri karşılaştırmalı olarak sunulmuştur. Alfa ve öğrenme katsayısı parametrelerinin 0.1–0.8 aralığında ve ayrıca 1.4 ile 1.5 değerlerinde, Klasik YSA yöntemi tüm veri kümelerinde KarciFANN yöntemine kıyasla daha başarılı performans sergilemiştir. Alfa değeri 0.9 için, KarciFANN yöntemi eğitim aşamasında Klasik YSA’ya göre daha yüksek doğruluk ve daha düşük hata değeri üretmiştir.

Test aşamasında ise her iki yöntem benzer doğruluk seviyelerine ulaşmış, ancak KarciFANN yöntemi daha düşük hata değeri üretmiştir. Alfa 1.0 için, KarciFANN yöntemi Klasik YSA yöntemi ile aynı şekilde çalışmaktadır. Alfa değerleri 1.1, 1.2 ve 1.3 olduğunda KarciFANN yöntemi, eğitim aşamasında daha başarılı olurken Klasik YSA, test aşamasında daha yüksek doğruluk ve daha düşük hata değerleri üretmiştir. Alfa ve öğrenme katsayısı 1.6–2.8 aralığında ve 3.2 ile 3.3 değerlerinde ise KarciFANN yöntemi, test aşamasında daha yüksek başarımlar sağlarken Klasik YSA yöntemi, eğitim aşamasında daha iyi performans göstermiştir. Alfa değerleri 2.9, 3.0, 3.1 ile 3.4–5.0 aralığındaki değerlerde, KarciFANN yöntemi tüm veri kümelerinde Klasik YSA’ya göre daha başarılı olmuştur. Ayrıca, çizelge incelendiğinde 3.1 ve 3.6–5.0 aralığındaki alfa değerlerinde KarciFANN yönteminin performansında bir düşüş gözlemlenmekte fakat, Klasik YSA’nın bu değerlerde neredeyse hiç başarılı sonuç üretmediği görülmektedir.

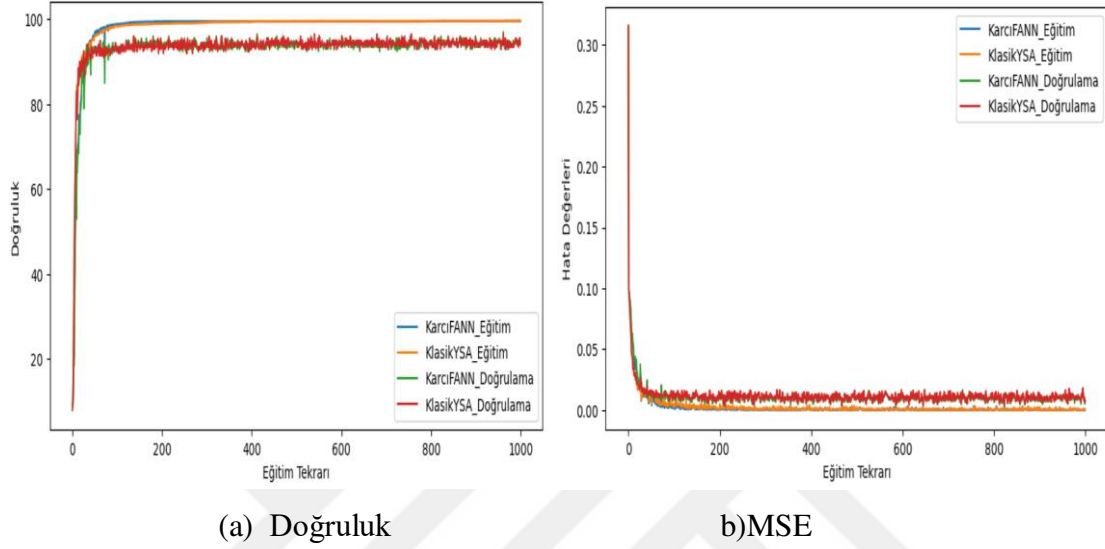
Çizelge 6. 8: Gina_prior2 veri setinde KarcıFANN ve Klasik YSA yöntemi için alfa parametre değerlerindeki değişimlere göre elde edilen Doğruluk ve MSE değerleri.

Alfa/ Öğrenme Katsayısı	KarcıFANN Eğitim		Klasik YSA Eğitim		KarcıFANN Test		Klasik YSA Test	
	Doğruluk%	MSE	Doğruluk%	MSE	Doğruluk%	MSE	Doğruluk%	MSE
0.1	10.52	0.100001	98.58	0.003504	10.60	0.100000	92.60	0.011445
0.2	9.19	0.100000	98.79	0.002167	9.20	0.100000	93.20	0.009315
0.3	10.56	0.100000	98.89	0.002120	10.60	0.100000	93.00	0.009314
0.4	9.19	0.100000	98.97	0.002096	9.20	0.100000	93.60	0.009405
0.5	17.26	0.089001	99.00	0.001090	17.80	0.089001	93.60	0.010048
0.6	10.22	0.083005	99.00	0.001055	10.20	0.090497	93.60	0.010732
0.7	69.93	0.025069	99.03	0.002046	66.80	0.039352	93.00	0.010445
0.8	68.63	0.043038	98.96	0.001030	64.40	0.036140	92.80	0.009729
0.9	99.63	0.000032	98.97	0.002034	92.60	0.009418	92.60	0.010726
1.0	99.56	0.000077	98.91	0.001029	92.80	0.010433	92.80	0.010433
1.1	99.59	0.000195	98.92	0.002021	92.60	0.011040	92.80	0.010405
1.2	99.30	0.000419	98.91	0.001026	92.40	0.010265	93.80	0.009844
1.3	99.04	0.000767	98.92	0.002023	92.80	0.010659	93.60	0.008863
1.4	98.89	0.001401	99.44	0.000066	93.20	0.011162	93.60	0.009947
1.5	98.52	0.002325	99.56	0.001019	92.80	0.011963	93.40	0.011077
1.6	98.07	0.003630	90.11	0.003126	92.80	0.012639	86.00	0.020492
1.7	97.22	0.005433	80.37	0.014038	92.40	0.013503	76.20	0.029073
1.8	96.33	0.007407	69.41	0.029041	91.60	0.014184	65.40	0.036627
1.9	94.85	0.009390	69.30	0.029082	91.60	0.014849	65.80	0.037323
2.0	93.70	0.011467	20.70	0.074388	91.20	0.015693	20.40	0.079602
2.1	92.74	0.013549	28.56	0.074082	91.00	0.016731	28.20	0.074631
2.2	91.67	0.015582	9.19	0.099985	90.40	0.018012	9.20	0.099952
2.3	91.04	0.017623	10.26	0.100000	89.20	0.019554	10.20	0.100000
2.4	89.85	0.019739	10.52	0.100000	88.60	0.021266	10.60	0.100000
2.5	89.04	0.022045	10.52	0.100000	87.00	0.023215	10.60	0.100000

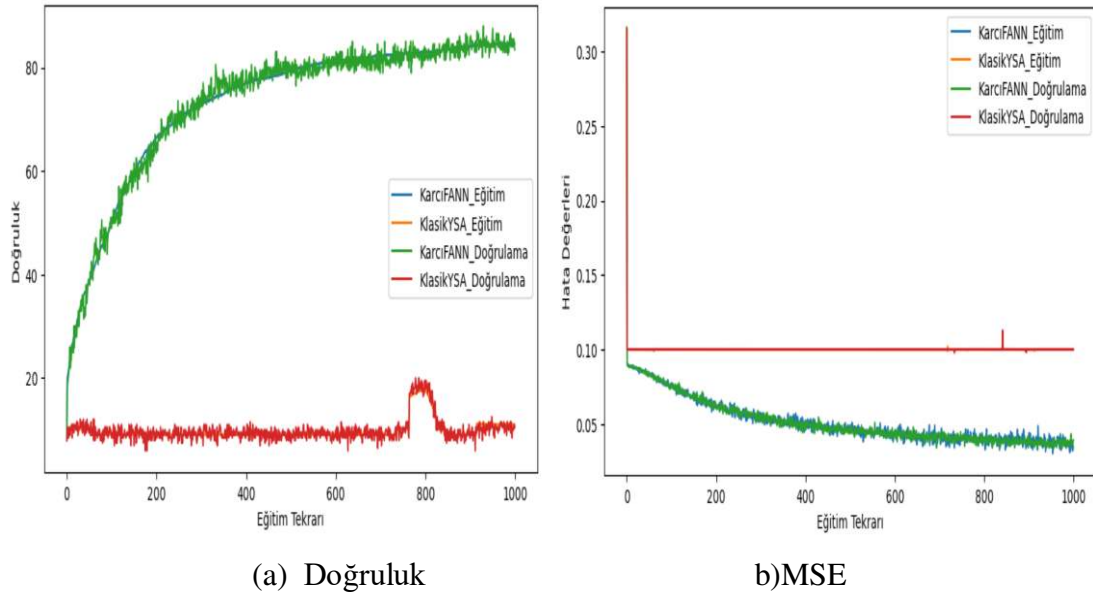
Çizelge 6. 9 (devam): Gina_prior2 veri setinde KarcıFANN ve Klasik YSA yöntemi için alfa parametre değerlerindeki değişimlere göre elde edilen Doğruluk ve MSE değerleri.

Alfa/ Öğrenme Katsayısı	KarcıFANN		Klasik YSA		KarcıFANN		Klasik YSA	
	Eğitim Doğruluk%	MSE	Eğitim Doğruluk %	MSE	Test Doğruluk%	Test Doğruluk%	Test MSE	Test Doğruluk %
2.6	88.22	0.024631	10.52	0.100000	86.40	0.025439	10.60	0.100000
2.7	87.07	0.027559	13.44	0.001012	85.40	0.027972	13.60	0.100000
2.8	86.00	0.030847	10.96	0.100000	83.80	0.030834	13.60	0.100000
2.9	84.63	0.034465	10.96	0.100000	81.80	0.034015	11.40	0.100000
3.0	82.78	0.038325	9.07	0.100000	80.60	0.037470	9.00	0.100000
3.1	80.78	0.042333	14.26	0.100000	79.00	0.041198	15.00	0.100000
3.2	77.85	0.046468	9.07	0.100000	76.60	0.045233	9.00	0.100000
3.3	74.59	0.050734	9.07	0.100000	74.60	0.049601	9.00	0.100000
3.4	71.48	0.055129	10.26	0.100000	70.80	0.054261	10.20	0.100000
3.5	68.00	0.059624	9.07	0.100000	67.40	0.059099	9.00	0.100000
3.6	62.89	0.064179	9.07	0.100000	62.80	0.063952	9.00	0.100000
3.7	57.63	0.068688	9.07	0.100000	57.80	0.068645	9.00	0.100000
3.8	51.93	0.072941	9.07	0.100000	52.40	0.072999	9.00	0.100000
3.9	46.85	0.076677	9.07	0.100000	46.80	0.076823	9.00	0.100000
4.0	42.48	0.079746	10.26	0.100000	43.60	0.079971	10.20	0.100000
4.1	39.04	0.082170	10.26	0.100000	37.80	0.082419	10.20	0.100000
4.2	36.37	0.084062	10.20	0.100000	35.20	0.084259	10.20	0.100000
4.3	33.56	0.085511	10.26	0.100000	32.60	0.085619	10.20	0.100000
4.4	30.67	0.086605	10.26	0.100000	28.80	0.086621	10.20	0.100000
4.5	28.07	0.087434	10.26	0.100000	26.20	0.087369	10.20	0.100000
4.6	25.74	0.088075	10.26	0.100000	24.40	0.087944	10.20	0.100000
4.7	23.70	0.088594	10.26	0.100000	21.60	0.088410	10.20	0.100000
4.8	21.96	0.089045	10.26	0.100000	20.60	0.088820	10.20	0.100000
4.9	20.63	0.089478	10.26	0.100000	18.40	0.089219	10.20	0.100000
5.0	19.26	0.089934	10.26	0.100000	17.40	0.089650	10.20	0.100000

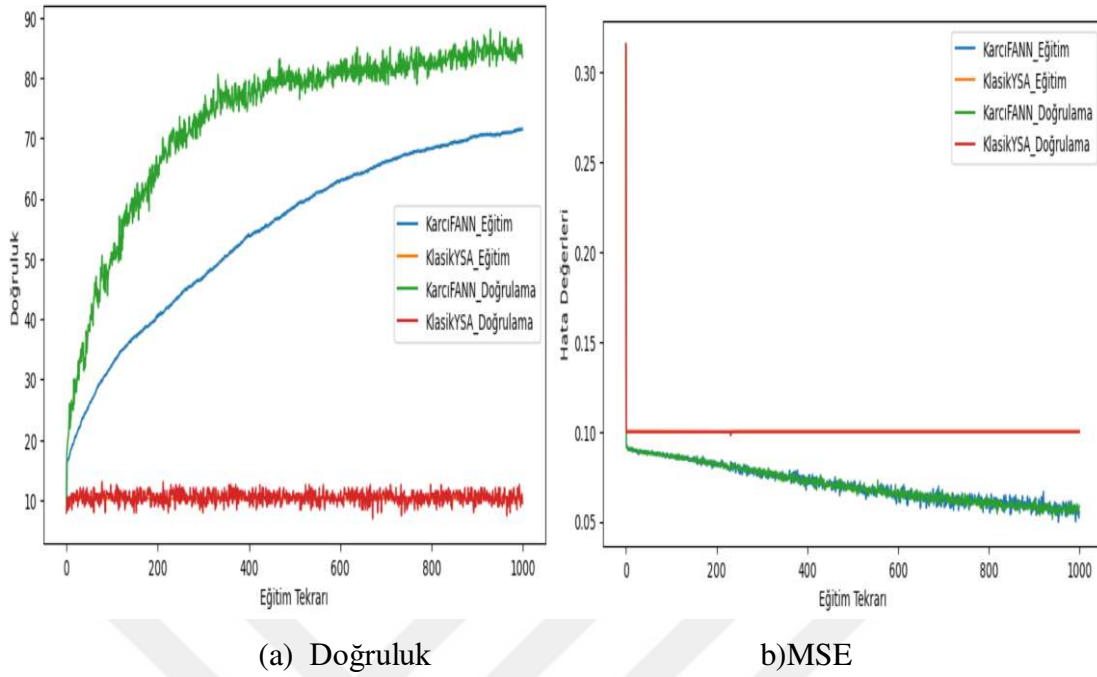
Gina_prior2 veri seti için KarcıFANN ve Klasik YSA yöntemlerinin eğitim veri kümesinde yapılan sınıflandırma işleminde elde edilen doğruluk ve hata grafikleri alfa ve öğrenme katsayısı 0.9 değeri için Şekil 6.14'te, 2.9 değeri için Şekil 6.15'te, 3.4 değeri için Şekil 6.16 ile gösterilmektedir.



Şekil 6. 14: GinaPrior2 veri setinde alfa ve öğrenme katsayılarının 0.9 değeri için KarcıFANN ve Sgd yöntemlerinin a) Doğruluk ve b) MSE grafikleri.



Şekil 6. 15: GinaPrior2 veri setinde alfa ve öğrenme katsayılarının 2.9 değeri için KarcıFANN ve SGD Yöntemlerinin a) Doğruluk ve b) MSE Grafikleri.



Şekil 6. 16:GinaPrior2 veri setinde Alfa ve öğrenme katsayılarının 3.4 değeri için KarcıFANN ve SGD yöntemlerinin a) Doğruluk ve b) MSE Grafikleri.

Çizelge 6.8’de, SignMNIST veri seti üzerinde gerçekleştirilen sınıflandırma sonucunda KarcıFANN ve SGD yöntemleriyle elde edilen doğruluk ve MSE değerleri sunulmuştur. Alfa ve öğrenme katsayısının 0.1–0.7 aralığı, 1.3 değeri ile 1.5–5.0 aralığındaki tüm değerlerde, SGD yöntemi KarcıFANN’a kıyasla daha yüksek doğruluk ve daha düşük hata değerleri üretmiştir. Alfa katsayısı 0.8, 0.9 ve 1.1 olduğunda ise KarcıFANN eğitim aşamasında daha yüksek hata üretmekte ve diğer veri kümeleri üzerinde daha iyi genelleme yeteneği göstermektedir. 1.0 değeri için KarcıFANN yöntemi, SGD ile aynı şekilde çalışmaktadır. 1.2 ve 1.4 değerlerinde ise KarcıFANN test aşamasında daha düşük hata değeri üretmiş fakat, eğitim kümesinde ise SGD daha iyi performans sergilemiştir. Bu sonuçlar, KarcıFANN yönteminin önerilen model için SignMNIST veri seti üzerinde sınıflandırma açısından yeterli performansı gösteremediğini ortaya koymaktadır. Bu durum, model mimarisinin ya da hiperparametrelerinin bu veri setine özgü yapısal özelliklere uygun olmadığını göstermektedir.

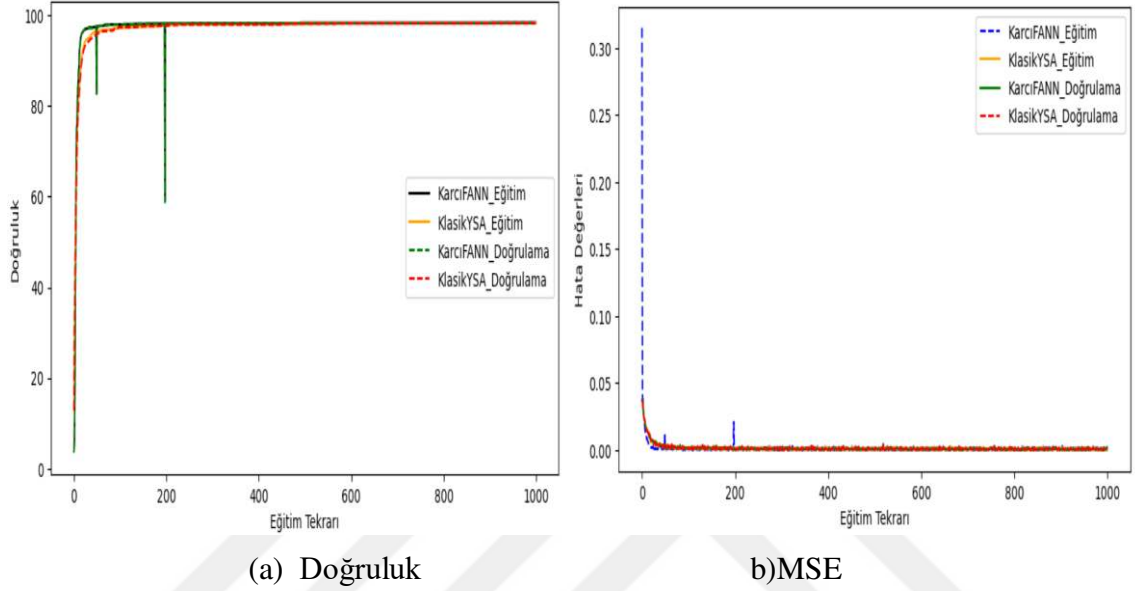
Çizelge 6. 10: SignMNIST veri setinde KarcıFANN yöntemi için alfa parametre değerlerindeki değişimlere göre elde edilen doğruluk ve hata değerleri.

Alfa/ Öğrenme Katsayısı	KarcıFANN		Klasik YSA		KarcıFANN		Klasik YSA	
	Eğitim		Eğitim		Test		Test	
	Doğruluk%	MSE	Doğruluk %	MSE	Doğruluk %	MSE	Doğruluk %	MSE
0.1	25.20	0.032254	97.60	0.000198	24.41	0.033112	97.70	0.001512
0.2	52.39	0.024669	97.85	0.000071	52.35	0.019446	97.99	0.000937
0.3	81.01	0.006402	97.96	0.000041	81.24	0.006405	98.11	0.000883
0.4	44.30	0.024800	97.82	0.000028	44.42	0.019602	98.10	0.000858
0.5	90.15	0.004800	98.07	0.000025	90.26	0.005605	98.32	0.000853
0.6	27.94	0.028401	98.13	0.000019	28.29	0.028404	98.35	0.000439
0.7	57.01	0.020402	98.13	0.000014	57.21	0.016805	98.35	0.000429
0.8	98.28	0.001601	98.17	0.000012	98.56	0.000405	98.44	0.000828
0.9	98.15	0.001603	97.76	0.000411	98.44	0.000411	98.06	0.001621
1.0	97.76	0.002006	97.65	0.000409	98.02	0.000426	97.89	0.001617
1.1	97.92	0.001616	97.83	0.000408	98.19	0.000448	98.10	0.001616
1.2	97.59	0.002028	97.89	0.000409	97.74	0.000888	98.14	0.001620
1.3	97.49	0.002059	97.90	0.000007	97.60	0.000953	98.03	0.002013
1.4	97.33	0.002114	97.74	0.000007	97.39	0.001043	97.86	0.001214
1.5	96.97	0.002218	98.17	0.000008	97.08	0.001185	98.23	0.000819
1.6	96.78	0.002404	97.35	0.000406	96.96	0.001833	97.65	0.001613
1.7	96.29	0.002662	97.83	0.000007	96.24	0.003120	98.18	0.000823
1.8	95.67	0.003404	97.36	0.000005	95.33	0.004027	97.59	0.001609
1.9	94.44	0.004552	97.41	0.000005	93.69	0.005576	97.66	0.001610
2.0	92.63	0.006171	97.68	0.000005	91.44	0.007480	98.10	0.000811
2.1	90.75	0.008020	98.11	0.000004	89.14	0.009580	98.30	0.002010
2.2	87.85	0.010189	98.16	0.000005	85.97	0.011552	98.42	0.002012
2.3	83.59	0.012310	97.66	0.000005	81.99	0.013386	97.92	0.002012
2.4	79.35	0.014648	97.59	0.000005	77.79	0.015294	97.89	0.002008
2.5	75.23	0.017426	97.49	0.000004	73.49	0.017466	97.70	0.000809

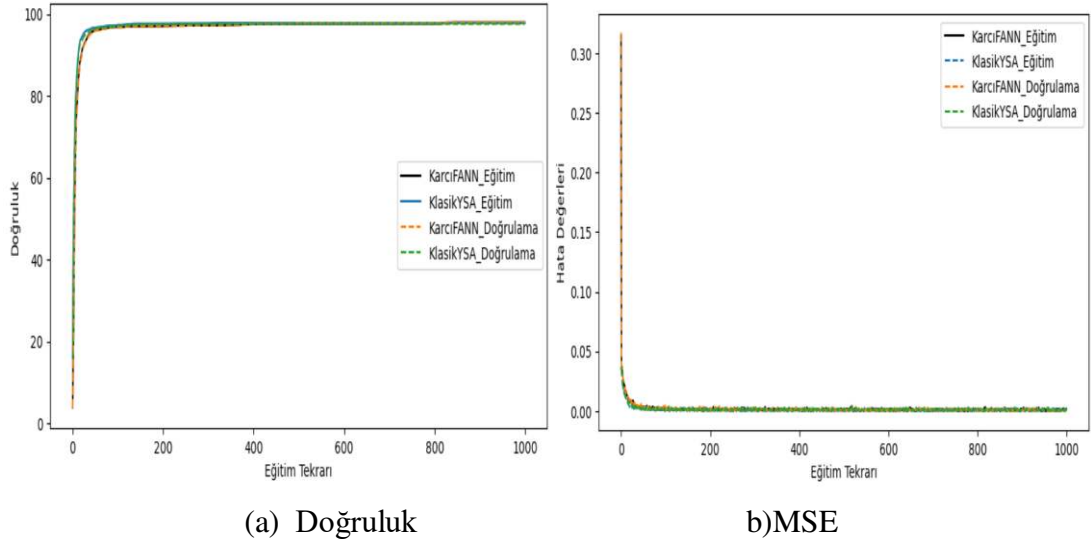
Çizelge 6. 11 (devam): SignMNIST veri setinde KarcıFANN yöntemi için alfa parametre değerlerindeki değişimlere göre elde edilen doğruluk ve hata değerleri.

Alfa/ Öğrenme Katsayısı	KarcıFANN		Klasik YSA		KarcıFANN		Klasik YSA	
	Eğitim		Eğitim		Test		Test	
	Doğruluk%	MSE	Doğruluk %	MSE	Doğruluk %	MSE	Doğruluk %	MSE
2.6	69.39	0.020484	96.83	0.000404	67.65	0.019882	97.14	0.002807
2.7	62.61	0.023620	96.64	0.000803	61.26	0.022526	97.00	0.002006
2.8	55.52	0.026863	98.69	0.000003	54.52	0.025435	98.72	0.001205
2.9	48.54	0.030095	98.10	0.000003	48.16	0.028431	98.14	0.002007
3.0	40.82	0.032952	98.32	0.000005	40.71	0.031083	98.42	0.001209
3.1	35.12	0.035057	90.04	0.001203	35.00	0.033277	90.20	0.002407
3.2	30.66	0.036358	90.07	0.001205	30.48	0.034892	90.23	0.002809
3.3	25.06	0.037100	89.74	0.001208	24.96	0.035943	89.99	0.002810
3.4	18.57	0.037555	85.89	0.004004	18.64	0.036651	86.06	0.005214
3.5	15.86	0.037854	81.62	0.005607	15.94	0.037128	81.83	0.006012
3.6	15.34	0.038071	73.72	0.007602	15.59	0.037473	73.94	0.008804
3.7	14.78	0.038245	69.67	0.009202	14.41	0.037735	69.93	0.010005
3.8	12.89	0.038390	65.84	0.011602	12.62	0.037937	65.86	0.011606
3.9	11.45	0.038517	65.71	0.011604	15.59	0.037473	65.73	0.012012
4.0	10.63	0.038636	64.86	0.013589	10.81	0.038239	65.03	0.013032
4.1	9.80	0.038758	61.41	0.012003	9.86	0.038372	61.52	0.014408
4.2	9.01	0.038893	53.05	0.016005	8.80	0.038511	53.21	0.016009
4.3	7.85	0.039051	48.92	0.018402	7.76	0.038670	49.04	0.018004
4.4	7.25	0.039241	48.82	0.018403	7.01	0.038859	48.90	0.018404
4.5	6.72	0.038893	41.06	0.021603	6.61	0.039090	41.03	0.021608
4.6	6.31	0.039755	28.64	0.029202	6.22	0.039373	28.59	0.027602
4.7	6.04	0.040098	28.52	0.029209	5.92	0.039720	28.48	0.027606
4.8	5.84	0.040511	20.46	0.032810	5.67	0.040140	20.51	0.031990
4.9	5.74	0.041007	16.46	0.034000	5.59	0.040648	16.46	0.034400
5.0	5.62	0.041598	16.06	0.034805	5.42	0.041255	16.21	0.034403

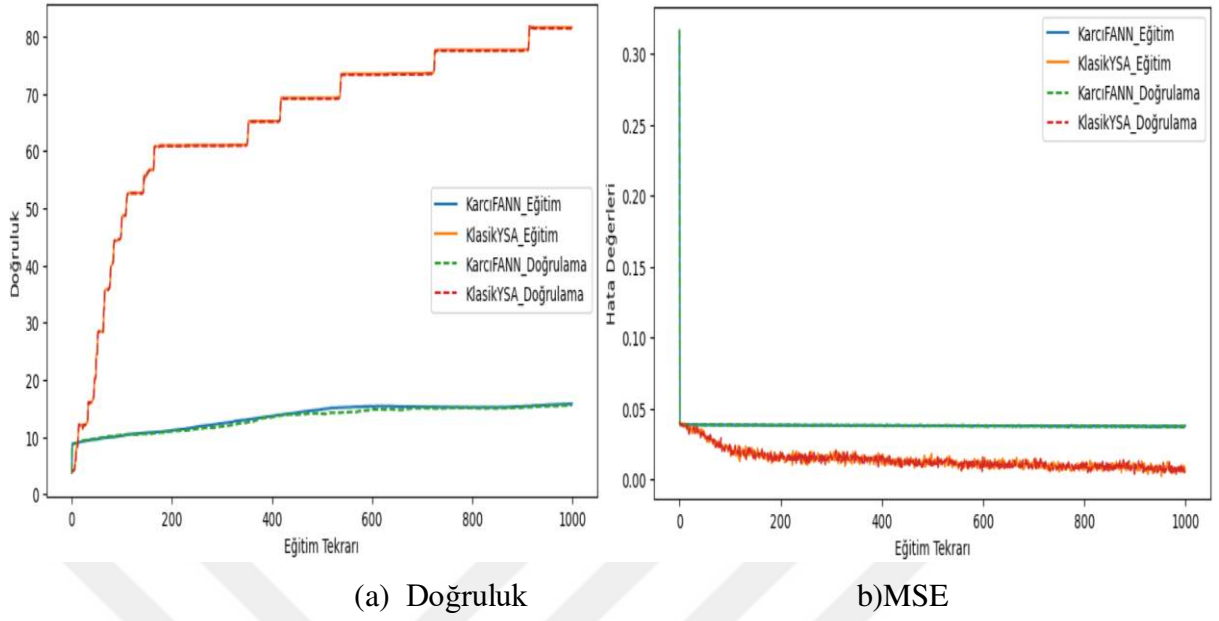
SignMNIST veri seti için KarcıFANN ve Klasik YSA yöntemlerinin eğitim veri kümesinde yapılan sınıflandırma işleminde elde edilen doğruluk ve hata grafikleri alfa ve öğrenme katsayısı 0.8 değeri için, Şekil 6.17’de, 1.1 değeri için Şekil 6.18’da, 3.5 değeri için Şekil 6.23 ile gösterilmektedir.



Şekil 6. 17: SignMNIST veri setinde alfa ve öğrenme katsayılarının 0.8 değeri için KarcıFANN ve Klasik YSA yöntemlerinin a) Doğruluk ve b) MSE Grafikleri.



Şekil 6. 18: SignMNIST veri setinde alfa ve öğrenme katsayılarının 1.1 değeri için KarcıFANN ve Klasik YSA yöntemlerinin a) Doğruluk ve b) MSE Grafikleri.



Şekil 6. 19: SignMNIST veri setinde alfa ve öğrenme katsayılarının 3.5 değeri için KarciFANN ve Klasik YSA yöntemlerinin a) Doğruluk ve b) MSE Grafikleri.

6.3.1 SGD yönteminin tüm veri setlerinde yetersiz kaldığı durumlar

Yapılan çalışmada gerçekleştirilen deneyler sonucunda, alfa ve öğrenme katsayısının 3.0–5.0 (3.1 ve 3.2 hariç) aralığında olduğu durumlarda Klasik YSA yönteminin Kuzushiji-MNIST ve Gina_prior2 veri setlerinin sınıflandırılmasında öğrenme işlemini gerçekleştiremediği gözlemlenmiştir. Bu aralıkta elde edilen test sonuçlarına ait doğruluk değerleri Çizelge 6.9’da sunulmaktadır. SignMNIST veri seti üzerinde gerçekleştirilen deneylerde Klasik YSA yöntemi, KarciFANN yöntemine göre daha yüksek performans gösterdiği görülmektedir.

Çizelge 6. 12: Alfa ve Öğrenme katsayısının Tüm Veri Setlerinin Sınıflandırılmasında Elde Edilen Test Sonuçları

Alfa/ Öğrenme Katsayısı	Kuzushiji-MNIST		Gina_prior2		SignMnist	
	KarcıFANN	Klasik YSA	KarcıFANN	Klasik YSA	KarcıFANN	Klasik YSA
3.0	82.38	10,00	80.60	9.07	40.71	98.42
3.1	82.38	73.11	79.00	14.26	35.00	90.20
3.2	71.81	75.19	76.60	9.07	30.48	90.23
3.3	81.44	20.85	74.60	9.07	24.96	89.99
3.4	78.50	58.05	70.80	10.26	18.64	86.06
3.5	73.10	10,00	67.40	9.07	15.94	81.83
3.6	73.56	10,00	62.80	9.07	15.59	73.94
3.7	74.77	10,00	57.80	9.07	14.41	69.93
3.8	73.71	10,00	52.40	9.07	12.62	65.86
3.9	72.38	10,00	46.80	9.07	15.59	65.73
4.0	70.66	10,00	43.60	10.26	10.81	65.03
4.1	70.05	10,00	37.80	10.26	9.86	61.52
4.2	58.99	10,00	35.20	10.20	8.80	53.21
4.3	67.80	10,00	32.60	10.26	7.76	49.04
4.4	66.60	10,00	28.80	10.26	7.01	48.90
4.5	66.65	10,00	26.20	10.26	6.61	41.03
4.6	62.81	10,00	24.40	10.26	6.22	28.59
4.7	60.88	10,00	21.60	10.26	5.92	28.48
4.8	28.57	10,00	20.60	10.26	5.67	20.51
4.9	55.87	10,31	18.40	10.26	5.59	16.46
5.0	53.81	10,00	17.40	10.26	5.42	16.21

6.3.2 Önerilen yönteminin genelleme ve aşırı uyum açısından değerlendirilmesi

6.3.2.1 Modelin genelleme yeteneği üzerinden değerlendirme

YSA'da önerilen modelin yeni verilere karşı gösterdiği performansını belirlemede genelleme yeteneği temel bir ölçüttür. Bu yeteneğin başarılı olarak değerlendirilebilmesi için tasarlanan bir modelin farklı veri setleri üzerinde başarılı sonuçlar üretmesi gerekmektedir.

Çalışma kapsamında gerçekleştirilen deneylerde, farklı öğrenme oranları ve farklı kesir dereceleri kullanılarak üç farklı veri seti üzerinde yapılan sınıflandırma sonuçları değerlendirilmiştir. Kuzushiji-MNIST veri seti için alfa ve öğrenme katsayısının 1.1 olarak belirlendiği durumda, KarcıFANN yöntemi eğitim kümesinde %98.95, test kümesinde ise %92.89 doğruluk sağlamıştır. Aynı parametre değeriyle Klasik YSA yöntemi, eğitim verisinde %98.92 ve test verisinde %93.05 doğruluk değeri ulaşmıştır. Gina_prior2 veri seti üzerinde alfa ve öğrenme katsayısı 0.9 olarak alındığında, KarcıFANN yöntemi eğitim

verisinde %99.63, test verisinde ise %92.60 doğruluk sağlamıştır. Klasik YSA yöntemi aynı parametrelerle eğitim kümesinde %98.97 ve test kümesinde %92.60 doğruluk değeri elde etmiştir. SignMNIST veri setinde ise alfa ve öğrenme katsayısı 0.8 olarak ayarlandığında KarcıFANN yöntemi, eğitim kümesinde %98.28 ve test kümesinde %98.56 doğruluk sağlamıştır. Klasik YSA yöntemi ise sırasıyla %98.17 ve %98.44 doğruluk değerleri ile benzer bir performans göstermiştir. Çizelge 6.6, Çizelge 6.7 ve Çizelge 6.8 incelendiğinde tüm veri setleri ve parametre değerleri için hem eğitim hem de test doğruluklarının yüksek olduğu ve öğrenme katsayısının yüksek değerlerinde KarcıFANN yönteminin iyi bir genelleme yeteneğine sahip olduğu görülmektedir.

6.3.2.2 Öğrenememe(yetersiz uyum) açısından değerlendirme

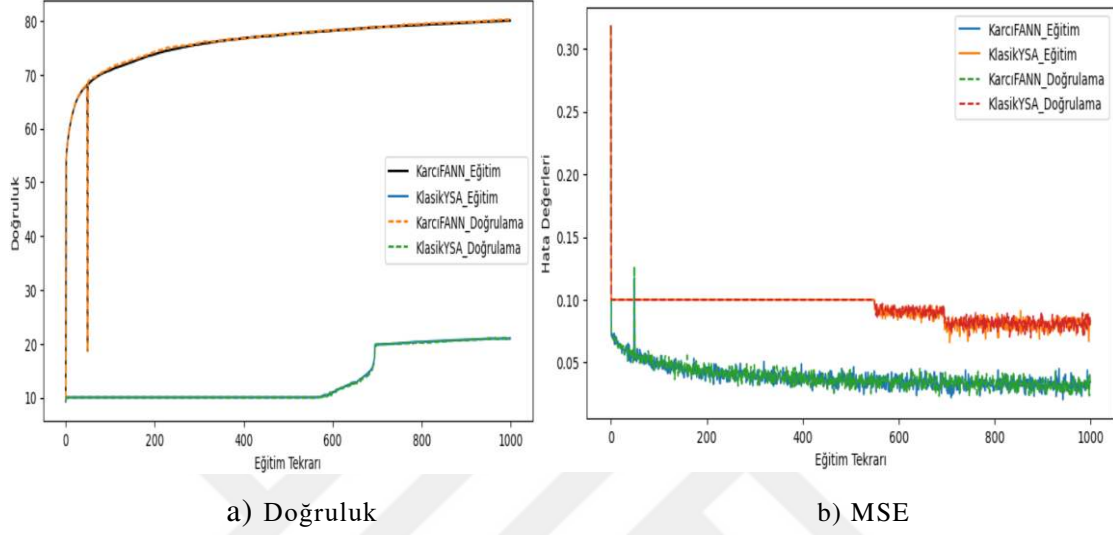
Eğitim ve test verileri üzerinde elde edilen düşük doğruluk ve yüksek hata oranları, YSA modelinin öğrenme kapasitesinin yetersiz kaldığını ortaya koymaktadır.

Kuzushiji-MNIST veri seti üzerinde KarcıFANN yöntemi kullanılarak yapılan sınıflandırma sonuçları Çizelge 6.6’da incelendiğinde, alfa katsayısının 0.1–0.3 aralığındaki değerlerinde elde edilen düşük doğruluk ve yüksek hata oranları modelin başarılı bir öğrenme gerçekleştiremediği; 0.4-5.0 arasındaki değerlerinde öğrenmenin başarıyla gerçekleştiği gözlemlenmektedir. Benzer şekilde, Klasik YSA yönteminde öğrenme katsayısının 0.1–2.7 aralığında başarılı bir öğrenme gerçekleştirdiği ancak, 2.8, 2.9, 3.0, 3.3 ve 3.5–5.0 aralığındaki değerlerinde doğruluk değerlerinin düşmesi ve hata değerlerinin artması ile modelin öğrenemediği anlaşılmaktadır. 3.3 değeri için elde edilen doğruluk ve hata grafikleri Şekil 6.20 ile verilmekte ve modelin başarısız olduğunu göstermektedir.

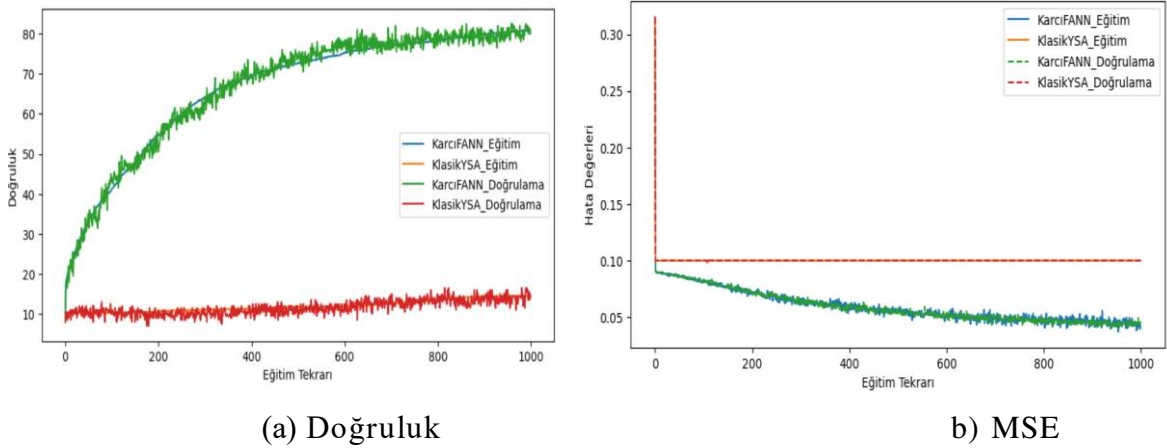
Gina_prior2 veri setinde gerçekleştirilen sınıflandırma sonuçlarının verildiği Çizelge 6.7 incelendiğinde KarcıFANN yöntemi, alfa değerinin 0.1–0.6 ile 3.9–5.0 aralığındaki değerlerde elde edilen doğruluk ve hata değerleri modelin öğrenmeyi gerçekleştiremediğini göstermektedir. 0.7–0.8 aralığındaki değerler incelendiğinde ise başarılı bir öğrenme performansı gösterdiği anlaşılmaktadır. Klasik YSA yöntemi öğrenme katsayısının 0.1-1.7 aralığında öğrenmenin başarılı bir şekilde gerçekleştirildiğini, 3.0, 3.1, 3.4 ve 3.6-5.0 aralığındaki değerlerde ise öğrenmenin yetersiz olduğu görülmektedir. 3.1 alfa ve öğrenme katsayısı için elde edilen doğruluk ve hata grafikleri Şekil 6.21’de sunulmaktadır.

SignMNIST veri seti kullanılarak gerçekleştirilen sınıflandırma sonuçlarının sunulduğu Çizelge 6.8 incelendiğinde KarcıFANN yöntemi, alfa değerinin 0.2, 0.3, 0.5, 0.7

ve 0.8-2.8 aralığında başarılı bir öğrenme işlemi gerçekleştirdiği; 0.1, 0.4, 0.6 ve 2.9-5.0 aralığındaki alfa değerlerinde ise öğrenmenin gerçekleşmediği görülmektedir. Klasik YSA yönteminde de öğrenme katsayısının 0.1–4.2 aralığındaki değerlerde başarılı bir öğrenmenin gerçekleştiği; 4.3–5.0 aralığındaki değerlerde modelin öğrenemediği görülmektedir.



Şekil 6. 20: Kuzushiji-MNIST veri setinde alfa ve öğrenme katsayılarının 3.3 değeri için KarciFANN ve Klasik YSA yöntemlerinin a) Doğruluk ve b) MSE Grafikleri.

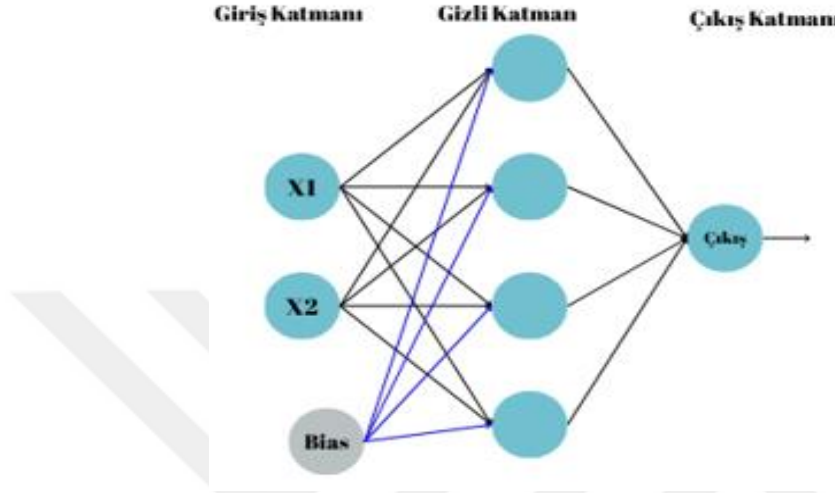


Şekil 6. 21: Gina_prior2 veri setinde alfa ve öğrenme katsayılarının 3.1 değeri için KarciFANN ve Klasik YSA yöntemlerinin a) Doğruluk ve b) MSE Grafikleri.

6.4 KarciFANN Yönteminin Yakınsama Kabiliyetinin İncelenmesi

Bu çalışmada geri yayımlı YSA'ların eğitim sürecinde kullanılan gradyan tabanlı optimizasyon yöntemlerinden olan SGD, Momentumlu GD, ADAM algoritmaları ve

KarcıFANN yönteminin ağırlıkların yakınsamasına, ezberleme durumuna ve öğrenme performansına etkileri incelenmiştir. Doğrusal olmayan problem olan XOR problemi üzerinde deneysel çalışmalar gerçekleştirilmiştir. XOR probleminin çözümünde kullanılan YSA modeli Şekil 6.22 ile göstermektedir.



Şekil 6. 22:XOR probleminin çözümünde kullanılan YSA modeli.

Modeldeki başlangıç ağırlık değerleri, tüm optimizasyon yöntemleri için aynı değerler alınarak rastgele atanmıştır. Ağların eğitimi süresince iterasyon sayısı 10.000 olarak belirlenmiştir. Eğitim sürecinde performans değerlendirmesi için hata fonksiyonu olarak MSE kullanılmıştır.

Çizelge 6.10’da alfa ve öğrenme katsayısının 0.1-3.5 arasındaki değerleri için SGD, Momentumlu GD, Adam ve SGD yöntemlerine ait MSE değerleri sunulmuştur. Momentumlu GD algoritmasının 0.9 ve 1.0 dışındaki tüm alfa ve öğrenme katsayısı değerlerinde en düşük MSE değerlerini üreterek en başarılı performansı sergilemiştir. ADAM algoritması 0.1 ve 0.2 hariç tüm değerlerde en yüksek MSE sonuçlarını üretmiş ve bu durum ADAM yönteminin XOR problemini çözmede yüksek öğrenme katsayısının kullanıldığı durumlarda yetersiz olduğunu göstermektedir. Alfa ve öğrenme katsayısı 0.1 ve 0.2 için KarcıFANN, SGD ve ADAM yöntemleriyle elde edilen eğitim aşamasındaki MSE değerleri, test aşamasındakilere göre daha düşük olarak ölçülmüştür. ADAM algoritması 0.1 ve 0.2 için KarcıFANN ve SGD’ye kıyasla daha düşük MSE sonuçları ürettiği gözlenmiştir. 0.3 değeri için KarcıFANN ve SGD yöntemleri önceki değerlere benzer MSE sonuçları verirken, ADAM algoritması eğitim ve test aşamalarında birbirine oldukça yakın MSE değerleri üretmiştir. 0.5–0.8 ve 2.1–3.5 aralığındaki değerlerde ise SGD yöntemi,

KarcıFANN ve ADAM yöntemlerine kıyasla daha düşük MSE değerleri üretmiştir. ADAM algoritması, 0.5 ile 3.5 arasındaki tüm alfa değerlerinde eğitim ve test aşamalarında neredeyse aynı MSE değerlerini üreterek başarılı bir öğrenme gerçekleştiremediği görülmektedir. 0.9 ve 1.0 olduğu durumlarda ise KarcıFANN yöntemi en düşük MSE değerlerini üretmiştir. Alfa ve öğrenme katsayısının 1.0 olduğunda, KarcıFANN yöntemi SGD yöntemi ile aynı şekilde çalışmaktadır. 1.1–2.0 aralığında olduğu durumlarda KarcıFANN yöntemi, Momentumlu GD'ye göre daha yüksek MSE değeri üretse de ADAM ve SGD'ye göre daha düşük MSE değerleri ürettiği görülmektedir.



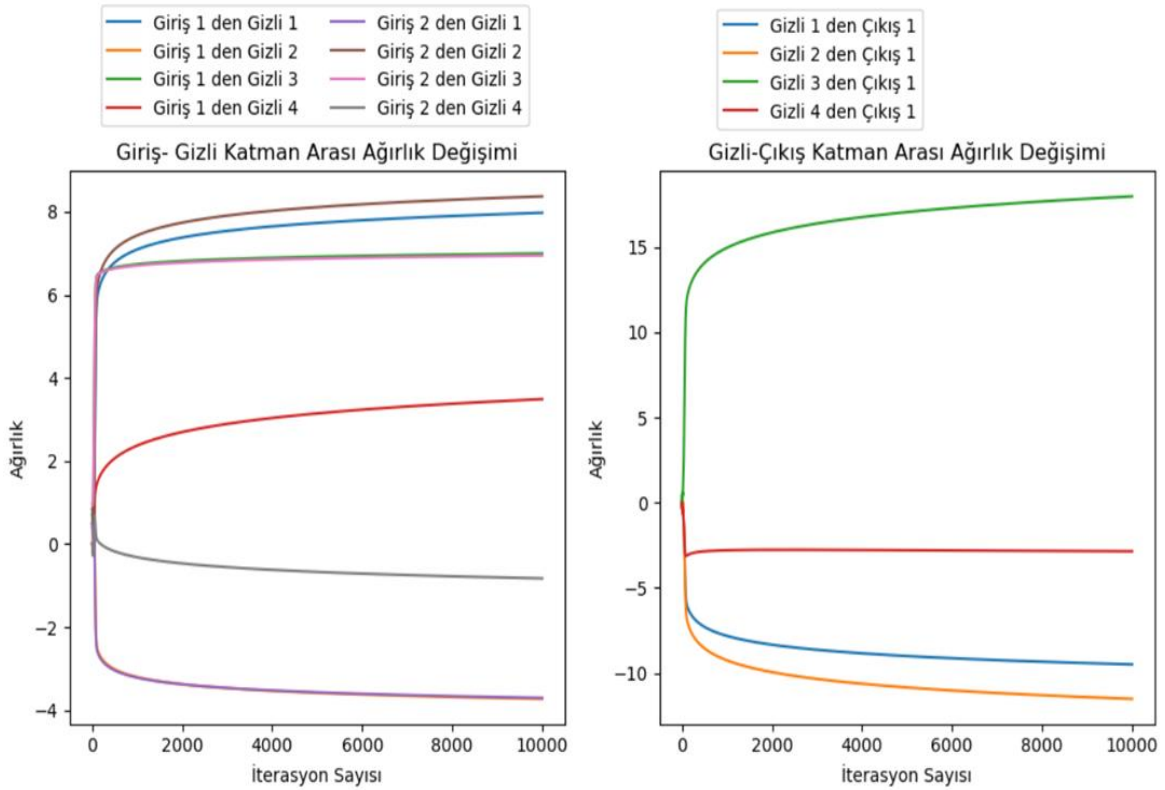
Çizelge 6. 13: KarcıFANN, SGD, ADAM ve Momentumlu GD yöntemleriyle farklı alfa ve öğrenme katsayı değerlerinde elde edilen MSE değerlerinin karşılaştırılması.

Alfa/ Öğrenme Katsayısı	KarcıFANN		SGD		ADAM		Momentumlu GD	
	Eğitim	Test	Eğitim	Test	Eğitim	Test	Eğitim	Test
0.1	0.07812500	0,500000	0.08253006	0.24845703	0.062652000	0.166669377	0.000217	0.00021726
0.2	0.07812500	0.49999998	0.08094572	0.24608712	0.062652000	0.12500013	0.000103	0.000102621
0.3	0.07812499	0.49999993	0.06762222	0.2129957	0.125027000	0.125000085	6.7e-05	6.6765e-05
0.4	0.07812470	0.49999813	0.02841464	0.07935195	0.125016000	0.062512554	4.9e-05	4.9365e-05
0.5	0.22438208	0.10426082	0.01887384	0.05040344	0.187504000	0.187500025	3.9e-05	3.9121e-05
0.6	0.06295320	0.06284979	0.01489248	0.03912254	0.187502000	0.187500018	3.2e-05	3.2386e-05
0.7	0.06282608	0.06283014	0.01261580	0.03284678	0.187502000	0.187500014	2.8e-05	2.7628e-05
0.8	0.06256070	0.06256367	0.01109465	0.02871668	0.187501000	0.187500011	2.4e-05	2.4094e-05
0.9	1.89923899e-06	8.06849866e-06	0.00998274	0.02572742	0.187501000	0.187500007	2.1e-05	2.1372e-05
1.0	4.95884825e-06	1.40047073e-05	0.00912247	0.02343168	0.187501000	0.187500005	1.9e-05	1.9215e-05
1.1	0.00016967	0.00031379	0.00843138	0.02159879	0.187501000	0.187500004	1.7468e-05	1.7e-05
1.2	0.00031259	0.00058045	0.00786209	0.02009758	0.187500000	0.187500003	1.6e-05	1.6029e-05
1.3	7.85829338e-05	0.00038912	0.00738564	0.01884862	0.187500000	0.187500003	1.5e-05	1.4828e-05
1.4	9.31395461e-05	0.00033893	0.00698383	0.01780243	0.187500000	0.187500002	1.4e-05	1.3816e-05
1.5	0.00012893	0.00050246	0.00664565	0.0169295	0.1875	0.187500002	1.3e-05	1.2956e-05
1.6	0.00203343	0.00433123	0.0063659	0.016217	0.1875	0.187500002	1.2e-05	1.222e-05
1.7	0.00073380	0.0020462	0.00614618	0.01567255	0.1875	0.187500001	1.2e-05	1.1591e-05
1.8	0.00145636	0.00398954	0.00600067	0.01534443	0.1875	0.187500001	1.1e-05	1.1054e-05
1.9	0.00304841	0.00690758	0.00598272	0.01541972	0.1875	0.187500001	1.1e-05	1.0594e-05
2.0	0.0003015	0.00258429	0.006415	0.01737656	0.1875	0.187500001	1e-05	1.0189e-05
2.1	0.03057281	0.09036979	0.00624748	0.01619678	0.1875	0.187500001	1e-05	9.774e-06
2.2	0.06158396	0.18229702	0.00559155	0.01450733	0.1875	0.187500001	9e-06	9.215e-06
2.3	0.05739571	0.17876757	0.00524042	0.01360238	0.1875	0.187500001	8e-06	8.372e-06
2.4	0.07754852	0.24460873	0.00498979	0.01295627	0.1875	0.1875	8e-06	7.6e-06
2.5	0.06711669	0.21677706	0.00479145	0.01244448	0.1875	0.1875	7e-06	7.264e-06

Çizelge 6. 14 (devam): KarcıFANN, SGD, ADAM ve Momentumlu GD yöntemleriyle farklı alfa ve öğrenme katsayı değerlerinde elde edilen MSE değerlerinin karşılaştırılması.

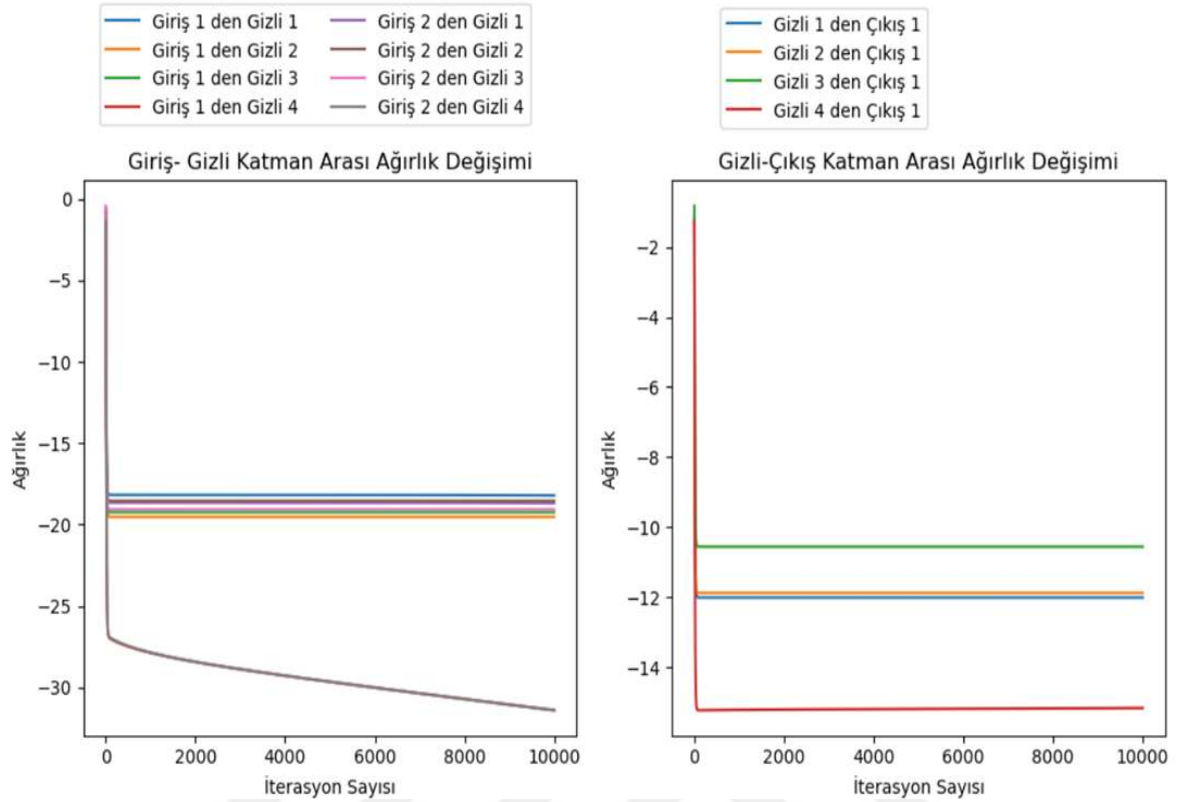
Alfa/ Öğrenme Katsayısı	KarcıFANN		SGD		ADAM		Momentumlu GD	
	Eğitim	Test	Eğitim	Test	Eğitim	Test	Eğitim	Test
2.6	0.07536766	0.23818847	0.00462662	0.01201858	0.1875	0.1875	6e-06	5.778e-06
2.7	0.07536766	0.23818847	0.00448617	0.01165502	0.1875	0.1875	5e-06	5.358e-06
2.8	0.07405632	0.23346845	0.00436522	0.01134117	0.1875	0.1875	5e-06	5.233e-06
2.9	0.08426657	0.24887302	0.00426122	0.0110704	0.1875	0.1875	5e-06	5.343e-06
3	0.07340475	0.24726232	0.00417332	0.01084049	0.1875	0.1875	7e-06	6.872e-06
3.1	0.08165853	0.2497479	0.0041026	0.010654	0.1875	0.1875	7e-06	6.724e-06
3.2	0.08171047	0.24941668	0.0040533	0.01052175	0.1875	0.1875	7e-06	6.523e-06
3.3	0.08655396	0.24944721	0.00403802	0.01047583	0.1875	0.1875	7e-06	6.327e-06
3.4	0.08627174	0.24771767	0.00410504	0.01064029	0.1875	0.1875	6e-06	6.141e-06
3.5	0.08494966	0.25033277	0.00532371	0.01368371	0.1875	0.1875	6e-06	5.964e-06
2.6	0.07536766	0.23818847	0.00462662	0.01201858	0.1875	0.1875	6e-06	5.778e-06
2.7	0.07536766	0.23818847	0.00448617	0.01165502	0.1875	0.1875	5e-06	5.358e-06
2.8	0.07405632	0.23346845	0.00436522	0.01134117	0.1875	0.1875	5e-06	5.233e-06

Ağın yakınsama durumu analiz edilerek aşırı öğrenme, yetersiz öğrenme, veri setinin modele uygunluğu ve parametrelerin uygunluğu incelenmiştir. Çizelge 6.10'da sunulan MSE değerlerine bakılarak Momentumlu GD yönteminin eğitim ve test aşamasının MSE değerleri incelendiğinde birbiriyle uyumlu değerler olduğundan modelin ezberleme yapmadığı öğrenmeyi başarılı bir şekilde gerçekleştirdiği görülmektedir. Momentumlu GD yöntemine ait ağın katmanlar arası ağırlıkların değişimini gösteren grafikler Şekil 6.23 ile gösterilmektedir. Modele ait ilgili grafiklere incelendiğinde ağırlık değerlerinin kararlı bir şekilde yakınsadığı, bazı ağırlık değerlerinin ise değişimlerinin devam etmesi daha fazla eğitime ihtiyaç duyduğu görülmektedir. Bu durum kullanılan veri setinin modele uygun olmaması ya da hiperparametrelerin değişmesi gerektiğini göstermektedir.



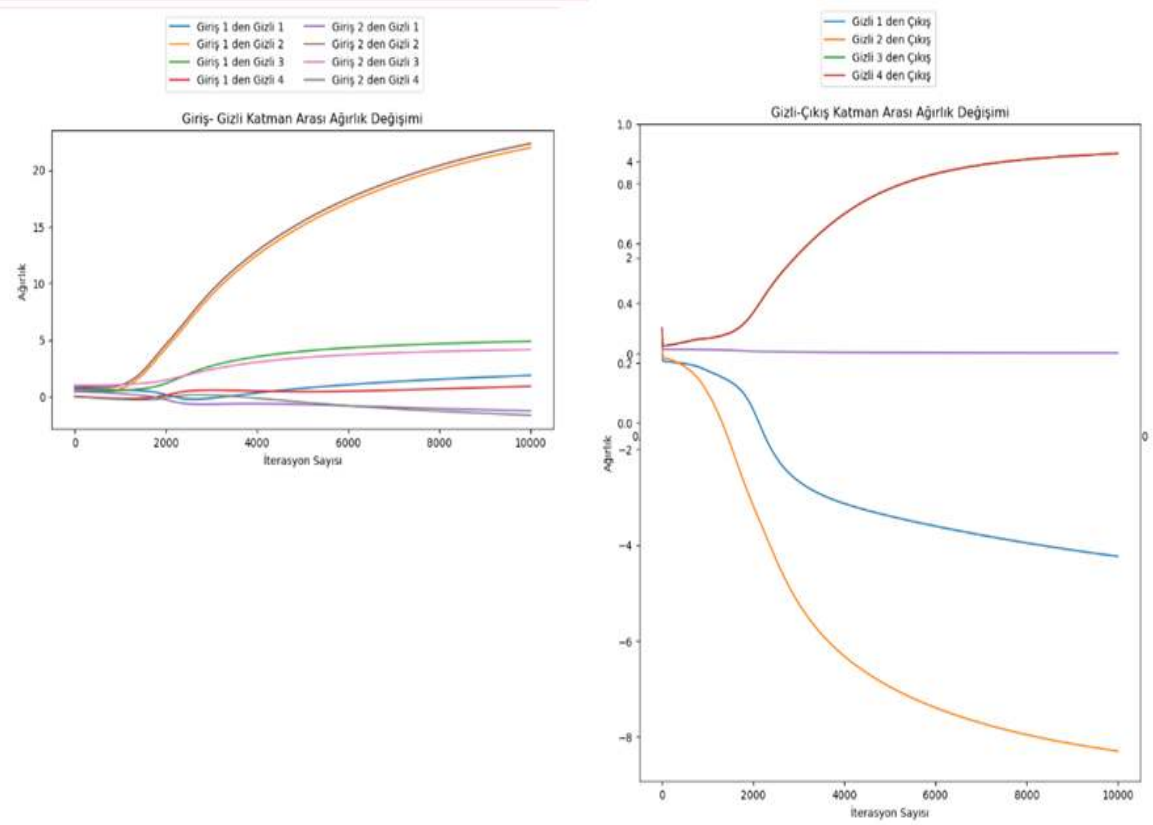
Şekil 6. 23:Momentum GD yöntemine ait 1.4 öğrenme katsayısı için ağırlık değişimleri.

ADAM yöntemi öğrenme katsayısının 0.5-3.5 aralığındaki değerlerde eğitim ve test sürecinde MSE değerlerinin birbirine yakın olduğu ve özellikle 0.5-2.1 aralığında diğer yöntemlere göre daha yüksek olduğu görülmektedir. Ağırlık değişimlerini gösteren grafik Şekil 6.24 ile gösterilmektedir. Çizelge ve grafiklere göre modelin ezberleme yapmadığı fakat yerel minimumda takıldığı anlaşılmaktadır.



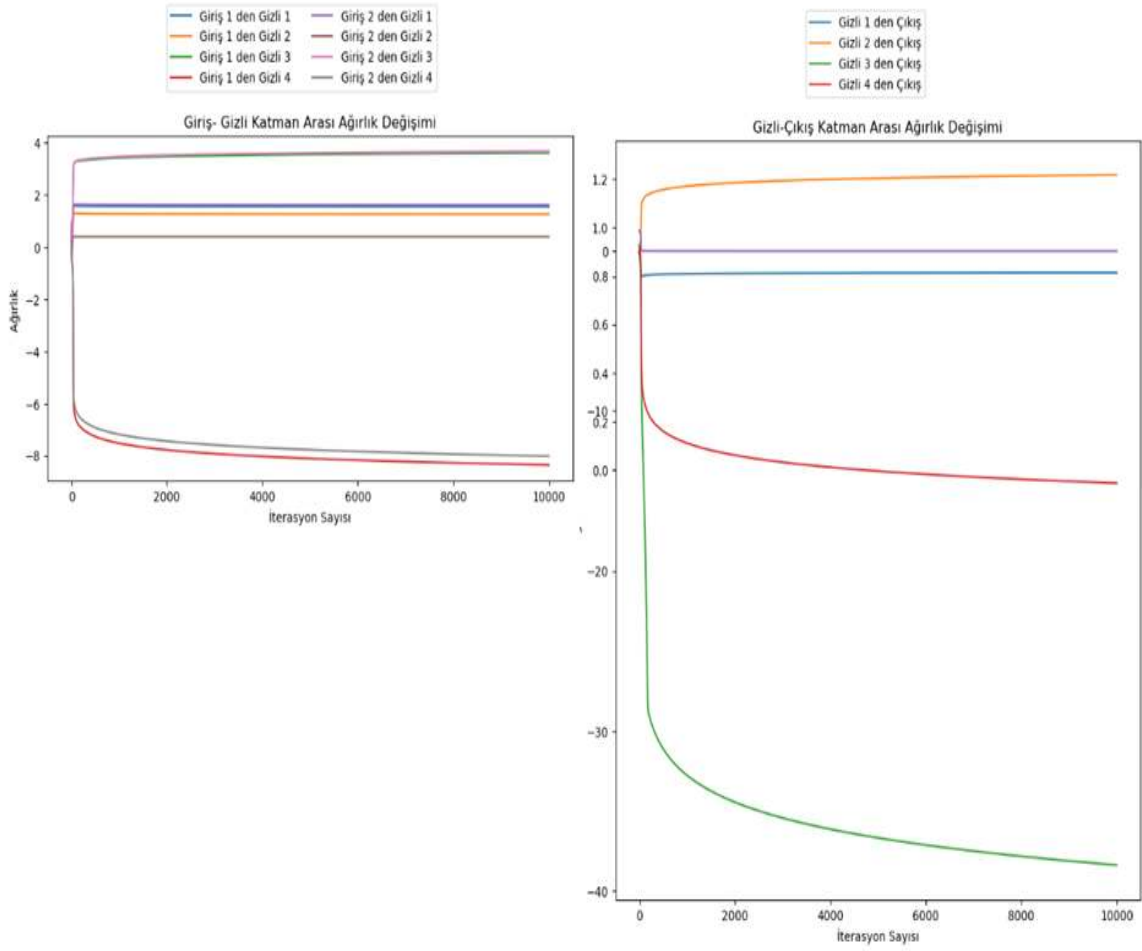
Şekil 6. 24:ADAM yöntemine ait 1.4 öğrenme katsayısı için ağırlık değişimleri.

SGD yöntemine ait Çizelge 6.10'daki MSE değerleri incelendiğinde, bazı öğrenme katsayı değerleri için eğitim MSE değerlerinin test MSE değerlerinden düşük olması modelin eğitim verilerini ezberlediğini göstermektedir. Şekil 6.25'deki ağırlık değişim grafiklerine bakıldığında ağırlıkların değişime devam etmesi, öğrenme sürecinin kararsız olduğunu, modelin yeterince öğrenemediğini ve yakınsayamadığını göstermektedir. Bu problemi çözmek için modelin daha uzun süre eğitilmesi, modele uygun bir veri seti kullanılması veya hiperparametrelerin yeniden ayarlanması gibi metotlar uygulanabilir.



Şekil 6. 25: SGD yöntemine ait 1.4 öğrenme katsayısı için ağırlık değişimleri.

Çizelge 6.10'daki KarcıFANN yöntemine ait eğitim ve test MSE değerleri incelendiğinde, bazı alfa değerlerinde öğrenmenin gerçekleşmediği fakat, alfa değerinin 0.8 ile 2.1 aralığında olduğu durumlarda yöntemin oldukça başarılı bir öğrenme performansı sergilediği görülmektedir. Şekil 6.26'da verilen ağırlık değişim grafiği incelendiğinde ağırlıkların büyük çoğunluğunun kararlı bir şekilde yakınsadığı ve bu kararlılığın eğitim süreci boyunca korunması modelin etkili bir öğrenme süreci geçirdiğini göstermektedir.



Şekil 6. 26:KarcıFANN yöntemine ait 1.4 öğrenme katsayısı için ağırlık değişimleri.

6.5 KarcıFANN Yönteminin Farklı Alfa Değerleri ile MAE, MSE ve RMSE Hata Fonksiyonlarının Performanslarının Değerlendirilmesi

Gina_prior2 veri seti kullanılarak yapılan çalışmada, KarcıFANN yönteminin farklı alfa parametresi ile MAE, MSE ve RMSE hata fonksiyonların etkisi incelenmiştir. Şekil 6.10 a ile gösterilen ve bir giriş katmanı, bir gizli katman ile bir çıkış katmanından oluşan ağ modeli kullanılmıştır. 1000 iterasyon ile sınırlandırılan eğitim süreci ve sigmoid aktivasyon fonksiyonu kullanılarak sınıflandırma işlemi gerçekleştirilmiştir. Sınıflandırma sürecini eşit şartlar altında değerlendirmek amacıyla başlangıç ağırlıkları rastgele fakat her model için aynı olacak şekilde belirlenmiştir.

Çizelge 6.11’de RMSE hata fonksiyonu ile farklı alfa değeri alınarak yapılan sınıflandırma işleminde eğitim, doğrulama ve test veri kümelerinde elde edilen hata fonksiyonu sonuçları gösterilmektedir. 0.1-6.0 aralığındaki alfa değerlerinde elde edilen

düşük doğruluk ve yüksek hata değerlerine bakılarak ağın öğrenemediği görülmektedir. 7 değeri için tüm veri kümelerinde elde edilen yüksek doğruluk ve düşük hata değerlerinden tüm alfa değerleri arasında en başarılı sonuçları verdiği görülmektedir. Alfa değerinin 8.0, 9.0 ve 10.0 değerleri için öğrenme gerçekleşmiş olsa da yeterli düzeyde başarı elde edilememiştir.

Çizelge 6. 15: KarcıFANN Yöntemi ve RMSE Hata Fonksiyonu Kullanılarak Elde Edilen Doğruluk (Acc) ve RMSE Değerleri.

Alfa	Eğitim		Doğrulama		Test	
	ACC%	RMSE	ACC%	RMSE	ACC%	RMSE
0.1	10.56	1.069579	9.50	1.069579	10.60	1.069579
0.2	11.07	1.052616	11.50	1.044988	11.00	1.052616
0.3	11.07	1.039230	11.50	1.027619	11.00	1.039230
0.4	9.19	1.048809	10.00	1.052616	9.20	1.048809
0.5	10.52	1.046900	10.50	1.054514	10.60	1.046900
0.6	11.07	1.052616	11.50	1.050714	11.00	1.052616
0.7	11.07	1.011920	11.50	1.007968	11.00	1.011920
0.8	9.44	1.035374	10.00	1.025671	9.60	1.035374
0.9	11.07	1.046900	11.50	1.046900	11.00	1.046900
1	9.44	1.029563	9.50	1.033441	9.40	1.029563
1.1	11.07	1.033441	11.50	1.029563	11.00	1.033441
1.2	11.07	1.033441	11.50	1.029563	11.00	1.033441
1.3	11.07	1.031504	11.50	1.025671	11.00	1.031504
1.4	10.52	1.041153	10.50	1.050714	10.60	1.041153
1.5	10.52	1.027618	10.50	1.027618	10.60	1.027618
1.6	10.52	0.963328	10.50	0.973653	10.60	0.963328
1.7	10.52	0.985901	10.50	0.989949	10.60	0.985901
1.8	11.07	1.001998	11.50	1.011929	11.00	1.001998
1.9	11.07	1.035374	11.50	1.025670	11.00	1.035374
2	11.07	1.033441	11.50	1.023719	11.00	1.033441
3	11.07	1.007968	11.50	1.007968	11.00	1.007968
4	-	-	-	-	-	-
5	11.07	1.091788	11.50	1.077033	11.00	1.091788
6	11.07	1.046900	11.50	1.044988	11.00	1.046900
7	76.11	0.228661	71.50	0.246963	73.20	0.228661
8	66.78	0.253319	63.50	0.271442	61.80	0.253319
9	59.37	1.069579	55.50	0.288161	53.40	0.274260
10	52.30	1.052616	51.00	0.307721	47.00	0.296160

MSE hata fonksiyonu kullanılarak yapılan sınıflandırma işleminde üretilen doğruluk ve hata değerleri Çizelge 6.12 ile gösterilmektedir. Tüm veri kümelerinde 0.1-0.6 aralığında ve 5.0-10.0 aralığındaki alfa değerler için düşük doğruluk değerleri ve yüksek hata değerleri, öğrenme işleminin gerçekleşmediğini göstermektedir. 0.7 ve 0.8 alfa değerleri için modelin

öğrenmeyi düşük bir performansta da olsa gerçekleştirdiği görülmektedir. 0.9 ve 2.0 aralığındaki tüm değerler için başarılı bir öğrenme gerçekleşmiştir. 3.0 ve üzeri değerler için modelin doğruluk değerinin giderek azalması ve hata değerinin artması modelin öğrenemediğini göstermektedir.

Çizelge 6. 16: KarcıFANN Yöntemi ve MSE Hata Fonksiyonu Kullanılarak Elde Edilen Doğruluk (Acc) ve MSE Değerleri.

Alfa	Eğitim		Doğrulama		Test	
	ACC%	MSE	ACC%	MSE	ACC%	MSE
0.1	10.52	0.100001	10.50	0.100000	10.60	0.100000
0.2	9.19	0.100000	10.00	0.100000	9.20	0.100000
0.3	10.56	0.100000	9.50	0.100000	10.60	0.100000
0.4	9.19	0.100000	10.00	0.100000	9.20	0.100000
0.5	17.26	0.089001	17.50	0.092363	17.80	0.089001
0.6	10.22	0.083005	9.50	0.090083	10.20	0.090497
0.7	69.93	0.025069	68.00	0.033414	66.80	0.039352
0.8	68.63	0.043038	65.00	0.039324	64.40	0.036140
0.9	99.63	0.000032	94.00	0.006879	92.60	0.009418
1	99.56	0.000077	94.00	0.008364	92.80	0.010433
1.1	99.59	0.000195	94.50	0.008507	92.60	0.011040
1.2	99.30	0.000419	94.00	0.009423	92.40	0.010265
1.3	99.04	0.000767	93.50	0.010233	92.80	0.010659
1.4	98.89	0.001401	93.00	0.010918	93.20	0.011162
1.5	98.52	0.002325	93.50	0.011604	92.80	0.011963
1.6	98.07	0.003630	93.00	0.012028	92.80	0.012639
1.7	97.22	0.005433	92.50	0.012600	92.40	0.013503
1.8	96.33	0.007407	93.00	0.013665	91.60	0.014184
1.9	94.85	0.009390	91.50	0.014927	91.60	0.014849
2	93.70	0.011467	91.00	0.016322	91.20	0.015693
3	82.78	0.038325	82.00	0.042518	80.60	0.037470
4	42.48	0.079746	44.50	0.080934	43.60	0.079971
5	19.26	0.089934	18.50	0.089571	17.40	0.089650
6	14.93	0.100814	14.50	0.099807	11.80	0.100364
7	13.67	0.123968	14.00	0.122395	11.40	0.123076
8	11.63	0.153478	12.50	0.151316	10.20	0.151799
9	9.81	0.185308	10.50	0.182524	8.80	0.182698
10	9.56	0.217133	10.50	0.213760	8.40	0.213642

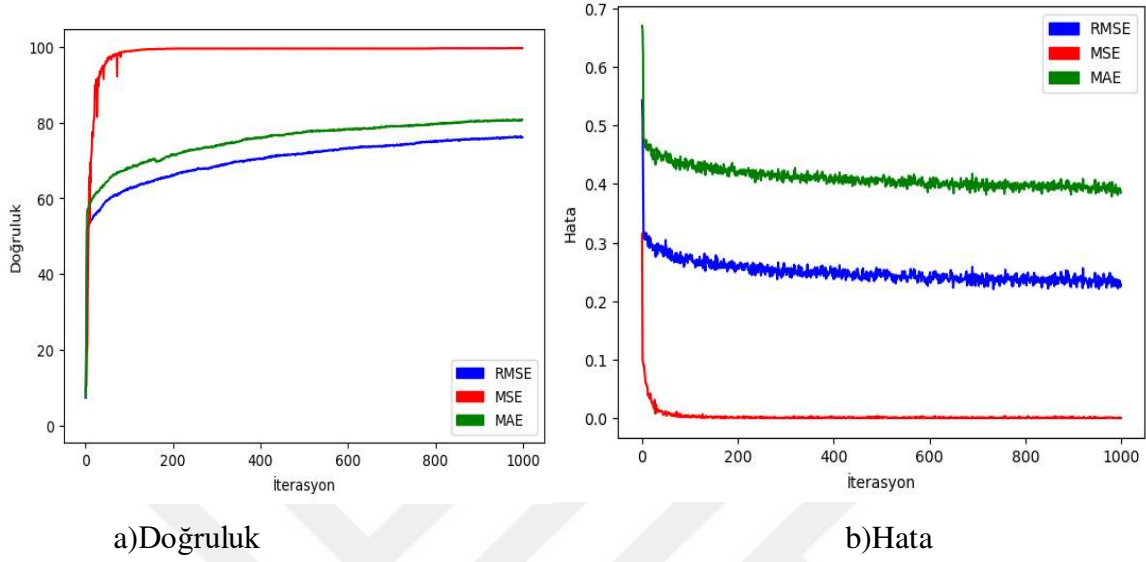
MAE hata fonksiyonu kullanılarak yapılan sınıflandırma işleminde üretilen doğruluk ve hata değerleri Çizelge 6.13 ile gösterilmektedir. MAE hata fonksiyonunun KarcıFANN yöntemi ile yapılan sınıflandırma işlemi sonucunda alfa kesir derecesinin 0.1-4 aralığında ve 6 değeri için düşük doğruluk ve yüksek hata değeri elde edilmiştir. 5 ve 7-9 aralığında yöntemin çalışmadığı görülmektedir. 10 alfa değeri için ise oldukça yüksek doğruluk değeri ile diğer

alfa kesir derecelerine göre daha düşük hata değeri ürettiği görülmektedir. Tüm alfa değerleri için bakıldığında yöntemin MAE hata fonksiyonu ile öğrenemediği anlaşılmaktadır.

Çizelge 6. 17: KarcıFANN Yöntemi ve MAE Hata Fonksiyonu Kullanılarak Elde Edilen Doğruluk (Acc) ve MAE Değerleri.

Alfa	Eğitim		Doğrulama		Test	
	ACC%	MAE	ACC%	MAE	ACC%	MAE
0.1	9.44	1.007968	9.50	1.001000	9.40	0.998999
0.2	10.52	0.980816	10.50	0.978775	10.60	0.982853
0.3	10.52	0.973653	10.50	0.970567	10.60	0.971597
0.4	9.19	0.991968	10.00	0.989949	9.20	0.990959
0.5	9.44	1.010940	9.50	1.009950	9.40	1.011929
0.6	11.07	0.992975	11.50	0.990959	11.00	0.987927
0.7	11.07	0.984886	11.50	0.987927	11.00	0.992975
0.8	9.44	1.000000	10.00	1.000000	9.60	1.001998
0.9	11.07	1.009950	11.50	1.018823	11.00	1.021763
1	9.44	0.986914	9.50	0.991968	9.40	0.989949
1.1	11.07	0.964365	11.50	0.968504	11.00	0.969536
1.2	9.07	0.986914	9.50	0.991968	9.00	0.989949
1.3	9.07	0.986914	9.50	0.991968	9.00	0.989949
1.4	9.44	0.990959	9.50	0.992975	9.40	0.985901
1.5	9.44	0.975705	9.50	0.980816	9.40	0.976729
1.6	11.07	0.985901	11.50	0.986914	11.00	0.988939
1.7	10.52	1.005982	10.50	1.000000	10.60	0.998999
1.8	10.52	1.020784	10.50	1.011929	10.60	1.007968
1.9	10.52	0.978775	10.50	0.979796	10.60	0.976729
2	10.52	0.970567	10.50	0.966437	10.60	0.970567
3	10.52	0.992975	10.50	0.981835	10.60	0.978775
4	10.52	0.991968	10.50	0.986914	10.60	0.990959
5	-	-	-	-	-	-
6	9.19	0.995992	10.00	0.989949	9.20	0.992975
7 – 9	-	-	-	-	-	-
10	80.70	0.385221	76.50	0.403658	80.00	0.389233

Şekil 6.27 incelendiğinde sunulan doğruluk ve hata grafikleri incelendiğinde, MSE hata fonksiyonunun MAE ve RMSE hata fonksiyonlarına göre daha başarılı olduğu gözlemlenmektedir. KarcıFANN yöntemi, MSE hata fonksiyonu ile kullanıldığında daha hızlı bir şekilde öğrenme eğilimi göstermektedir. Doğruluk değerinin kararlı şekilde artması ve hata değerinde meydana gelen dengeli düşüş, modelin eğitim verisine daha iyi uyum sağladığını göstermektedir. Elde edilen bu bulgular, MSE'nin hem doğruluk hem de hata metriği açısından üstün olduğunu göstermektedir.



Şekil 6. 27: Ginaprior2 veri setinin farklı hata fonksiyonları ile sınıflandırılması sonucu alfa 0.9 değeri için elde edilen en başarılı (a) Doğruluk ve (b) Hata grafikleri.

7. SONUÇ VE ÖNERİLER

YSA'ların eğitim sürecinde optimizasyon işlemi için gradyan tabanlı yöntemler kullanılmakta ve bu yöntemlerle ağırlık güncelleme işlemi gerçekleştirilmektedir.

Bu tez kapsamında kullanılan KarcıFANN yönteminin kesir dereceli türev yapısı, ağırlık güncellemeleri sırasında bazı durumlarda negatif taban ve kesirli üs kombinasyonlarının ortaya çıkmasına neden olmaktadır. Bu tür ifadeler reel sayılar kümesinde tanımsız kaldığından, elde edilen değerlerin doğru ve eksiksiz bir biçimde işlenebilmesi için karmaşık sayıların kullanılmasını gerekli kılmaktadır. Başlangıçta reel değerli modellerde bu karmaşık değerlerin sadece reel bileşenleri ile işlenmiştir. Bu yaklaşım, sanal bileşenlerin ihmal edilmesine bağlı olarak bilgi kaybına ve optimizasyon sürecinde kararsızlıklara yol açabilmektedir. Bu çalışmada, karmaşık sayıların doğrudan eğitim sürecine dahil edildiği bir ağ yapısı tasarlanmıştır. Böylelikle yöntemin matematiksel temeline tam uyum sağlanmış ve faz bilgisinin korunması sayesinde öğrenme sürecinin daha doğru, dengeli ve genelleme kabiliyeti artırılması amaçlanmıştır.

Tez kapsamında gerçekleştirilen birinci çalışmada KarcıFANN'ın reel sayılardaki ve karmaşık sayılardaki performansını değerlendirmek üzere Digits, MNIST Rotation 45°, 2BFT-Digits ve 2BFT-MNIST Rotation 45° veri kümeleri üzerinde deneyler gerçekleştirilmiştir.

İlk adımda, KarcıFANN'ın reel sayılarla kullanıldığı model ile Digits ve MNIST Rotation 45° veri setleri sınıflandırılmıştır. Digits veri setinin sınıflandırılması sonucu elde edilen en yüksek test doğruluğu %83.89 ve en düşük test hatası 0.0674 ile alfa 1.2 değerinde elde edilmiştir. Alfa 0.5 ve 0.9 değerlerinde eğitim doğruluğunun yüksek olmasına karşın doğrulama ve test doğruluklarının düşük olması modelin ezberleme eğiliminde olduğuna işaret etmektedir. Alfa 1.8 için doğrulama veri kümesinde doğruluk oranı artmış görünse de test doğruluğunun düşmesi, modelin eğitim verisine fazla uyum sağladığını göstermektedir.

MNIST Rotation 45° veri setinin sınıflandırıldığı sonuçlar incelendiğinde alfanın 0.9–1.2 değerleri için dengeli ve kararlı bir öğrenmeye işaret etmektedir. Her iki değerde de test doğruluğu %88.92, test hatası 0.0228 olup eğitim, doğrulama ve test değerlerinin birbirine yakın seyretmesi iyi bir genelleme yeteneğine sahip olduğunu göstermektedir.

İkinci adımda, KarcıFANN yönteminin kesir dereceli türev yapısının doğası gereği ortaya çıkan karmaşık değerlerin öğrenme sürecine katkısını incelemek amacıyla deneyler gerçekleştirilmiştir. Bu amaçla Digits, MNIST Rotation 45°, 2BFT-Digits ve 2BFT-MNIST Rotation 45° veri kümeleri üzerinde Reel, Kompleks ve Abs olarak üç farklı model ile karşılaştırılmıştır. Reel modelde yalnızca reel bileşenler kullanılırken Kompleks modelde reel ve sanal bileşenler birlikte işlenmiştir. Böylece, sadece reel bileşeni kullanmanın neden olduğu bilgi kaybı ortadan kaldırılarak faz bilgisinin doğrudan öğrenmeye yansıtılması sağlanmıştır. Abs modelde ise karmaşık çıktının yalnızca modülü değerlendirilmiştir.

Elde edilen sonuçlar, tüm veri setlerinde Kompleks modelin Abs ve Reel modelden daha başarılı performans sergilediğini göstermiştir. Digits veri setinde Kompleks model, alfanın $\frac{3}{2}$ değerinde %97.78 test doğruluğu üretmiştir. 2BFT-Digits veri setinde ise alfa $\frac{5}{6}$ ve $\frac{9}{8}$ değerlerinde en yüksek test doğruluğu %95.00 olarak kaydedilmiştir. MNIST Rotation 45° veri setinde alfa $\frac{9}{8}$ değerinde en iyi performans elde edilmiş ve test doğruluğu %97 seviyelerine ulaşmıştır. 2BFT-MNIST Rotation 45° veri setinde ise alfa $\frac{9}{8}$ değeriyle %92.03 test doğruluğu elde edilmiştir. Reel model, özellikle uzaysal veri setlerinde kabul edilebilir sonuçlar üretmiş ancak frekans dönüşümlü veri setlerinde faz bilgisini işleyemediği için yetersiz kalmıştır. Abs model ise yalnızca genlik bilgisine dayandığından özellikle karmaşık veri yapılarında düşük doğruluk ve yüksek hata değerleriyle sınırlı kalmıştır.

Deneylerde ayrıca “sanal etki” adı verilen bir ölçüt hesaplanarak, sanal bileşenlerin öğrenme sürecine katkısı incelenmiştir. Bu analiz, alfa değerinin modelin dengesini doğrudan etkilediğini göstermektedir. Alfanın $\frac{3}{8}$ ve $\frac{3}{2}$ değerlerinde sanal bileşenlerin daha baskın olduğu, alfanın $\frac{3}{4}$, $\frac{5}{6}$ ve $\frac{9}{8}$ aralıklarında ise reel ve sanal bileşenlerin daha dengeli bir öğrenme sağladığı gözlemlenmiştir. Çok düşük veya çok yüksek alfa değerlerinde ise Kompleks modelin öğrenme başarısının düştüğü ve hata seviyelerinin yükseldiği tespit edilmiştir.

Genel olarak, karmaşık bileşenlerin doğrudan işlendiği Kompleks model, tüm veri setlerinde en yüksek doğruluk ve en kararlı yakınsama performansını sergilemiştir. Reel model, temel veri setlerinde kısmen başarılı olurken karmaşık yapıdaki veri setlerinde yetersiz kalmıştır. Abs model ise faz bilgisini kaybettiğinden özellikle frekans tabanlı veri setlerinde düşük performans göstermiştir. Bu sonuçlar, KarcıFANN yönteminin yalnızca

reel bileşen ile sınırlı kalmaması gerektiğini ve karmaşık bileşenlerin öğrenme sürecine doğrudan katılmasının optimizasyonu güçlendirdiğini ortaya koymaktadır. Böylece, kesirli türevlerin oluşturduğu karmaşık değerlerin korunarak işlenmesi, daha yüksek genelleme kabiliyeti ve doğruluk elde edilmesine imkan tanımaktadır.

Tez kapsamında gerçekleştirilen ikinci çalışmada, öğrenme katsayısı kullanımından kaynaklanan problemlere çözümler sunmak amacıyla Karcı kesir dereceli YSA (KarcıFANN): öğrenme katsayısı, aşırı uyum ve yetersiz uyum sorunlarını çözme adı verilen çalışma sunulmuştur. Bu kapsamda Kuzushiji-MNIST, Gina_prior2, Sign-MNIST veri setlerinde sınıflandırma çalışması yapılmıştır. İki yöntemi aynı şartlarda başlatmak amacıyla aynı ağırlık değerleri, aynı ağ modeli, sigmoid aktivasyon fonksiyonu ve MSE hata fonksiyonu kullanılmıştır. 0.1-5 aralığında öğrenme katsayısı ve Karcı kesir derecesi verilerek deneysel çalışmalar gerçekleştirilmiştir. Kuzushiji-MNIST veri setinin sınıflandırıldığı KarcıFANN yönteminde en yüksek doğruluk, alfanın 1.1 değeri için %98.95 olarak ölçülmüştür. Klasik YSA yönteminde ise öğrenme katsayısının 2.7 değeri için en yüksek başarımlar %99.07 olarak elde edilmiştir. Gina_prior2 veri setinin sınıflandırıldığı KarcıFANN yönteminde en yüksek eğitim doğruluğu, 0.9 alfa değeri için %99.63 olarak kaydedilmiştir. Klasik YSA'da ise en yüksek eğitim doğruluğu, öğrenme katsayısının 0.7 değeri için %99.03 olarak ölçülmüştür. SignMnist veri setinin sınıflandırıldığı KarcıFANN yönteminde en yüksek eğitim doğruluğu alfa 0.8 değeri için %98.28 olarak elde edilmiştir. Klasik YSA yönteminde ise en yüksek eğitim doğruluğu, öğrenme katsayısının 2.2 değerinde %98.16 olarak ölçülmüştür. SignMNIST veri seti için genel olarak değerlendirme yapıldığında kullanılan ağ modelinin uygun olmadığı sonucuna varılmıştır.

Karcı kesir dereceli türev yaklaşımını kullanan KarcıFANN yöntemi, öğrenme katsayısının belirlenmesindeki problemleri azaltarak özellikle Klasik YSA'nın başarısız olduğu veya ezberlediği durumlarda daha istikrarlı ve başarılı sonuçlar üretme potansiyeline sahip olduğu yapılan deneysel çalışmalardan anlaşılmaktadır.

Tez kapsamında gerçekleştirilen üçüncü çalışmada, KarcıFANN yönteminin yakınsama kabiliyetinin analiz edilmesi adı verilen çalışmada XOR problemi üzerinde Momentumlu GD, ADAM, SGD ve KarcıFANN yöntemlerinin öğrenme, ezberleme ve ağırlık değerleri üzerindeki etkileri, MSE metriği ve ağırlık değişim grafikleri üzerinden ele alınmıştır. Hata değerleri bakımından değerlendirildiğinde Momentumlu GD yönteminin

öğrenme katsayısının 0.9 ve 1.0 değerleri dışındaki tüm değerlerde en düşük MSE değerini ürettiği, KarcıFANN yönteminin ise 0.9-2.0 aralığındaki değerleri için SGD'den ve 0.3 ile 0.6-3.5 aralığındaki değerleri için ADAM'dan daha düşük MSE değerleri ürettiği görülmektedir. ADAM yöntemi 0.1 ve 0.2 dışındaki tüm öğrenme katsayısı değerleri için en başarısız sonuçları üretmiştir. Eğitim ve test aşamalarında elde edilen MSE değerlerinin birbirine yakın değerler olması modelin ezberleme yapmadığını gösterse de bu değerlerin genel olarak yüksek değerler oluşu öğrenme başarısının düşük olduğunu ifade etmektedir. Çalışmada kullanılan 4 yöntemin ağırlık değişim grafikleri incelendiğinde Momentumlu GD ve Adam yöntemlerinin daha iyi yakınsama sağladığı görülmektedir. SGD yönteminde, ağırlık değerlerinin değişime devam etmesi tam bir yakınsama sağlamadığı yöntemin kararsız olduğu görülmektedir. KarcıFANN yönteminde ise ağırlık değişimlerinin az da olsa devam etmesi yakınsamanın henüz gerçekleşmediğini göstermektedir. Daha iyi bir yakınsama sağlamak amacıyla, modelin daha fazla eğitim iterasyonu ile çalıştırılması, kullanılan veri setinin optimize edilmesi veya model hiperparametrelerinin yeniden düzenlenmesi gibi yöntemlerin çözüm sunabileceği düşünülmektedir. Bu çalışmada elde edilen MSE değerleri ve ağırlık değişim grafiklerinin birlikte değerlendirilmesi sonucunda, öğrenme ve yakınsama açısından en başarılı yöntemin Momentumlu GD, ikinci en başarılı yöntemin ise KarcıFANN olduğu görülmektedir. ADAM yönteminin ise yerel bir minimum noktada takılı kaldığı, elde edilen MSE değerleri ve ağırlık değişim grafiklerinden anlaşılmaktadır.

Tez kapsamında geliştirilen dördüncü çalışmada, KarcıFANN yönteminin önceki çalışmalarda kullanılan MSE hata fonksiyonu yerine RMSE ve MAE hata fonksiyonları kullanılması durumunda öğrenme başarısına etkileri Gina_prior2 veri seti üzerinde incelenmiştir. Adil bir karşılaştırma gerçekleştirmek amacıyla tüm hata fonksiyonları için aynı ağ modeli, aynı başlangıç rastgele ağırlık değerleri ve aynı alfa kesir derecesi kullanılarak eşit çalışma koşulları sağlanmıştır. 0.9-1.3 aralığında MSE hata fonksiyonunun kullanıldığı modelin en başarılı sonuçları ürettiği gözlemlenmiştir. Eğitim aşamasındaki maksimum doğruluk değerleri incelendiğinde, alfa kesir derecesinin 0.9 değeri için MSE hata fonksiyonlu modelde %99.63, alfa kesir derecesinin 10 değeri için MAE hata fonksiyonlu modelde %80.70, alfa kesir derecesinin 7 değeri için RMSE hata fonksiyonlu modelde %76.11 olarak en yüksek doğruluk değerleri elde edilmiştir. Tüm değerlere bakıldığında MSE hata fonksiyonunun diğer iki yönteme göre daha başarılı performans

sergilediđi gör÷lmektedir. Model parametreleri optimize edilerek MAE ve RMSE hata fonksiyonlarının da başarılı sonuçlar üretebileceđi öngör÷lmektedir. Yapılan deneysel çalışmaların sonucunda, hata fonksiyonu seçiminin model eğitimi ve başarısı üzerinde önemli etkiye sahip olduđu anlaşılmaktadır.



KAYNAKLAR

- Abdalla, R. (2023).** Complex-valued neural networks -- Theory and analysis. arXiv. <https://arxiv.org/abs/2312.06087>
- Amin, D. F., & Murase, K. (2009).** Single-layered complex-valued neural network for real-valued classification problems. *Neurocomputing*, 72(4–6), 945–955. <https://doi.org/10.1016/j.neucom.2008.04.006>
- Alam, A. (2023).** What is Machine Learning? <https://doi.org/10.5281/zenodo.8231581>.
- Alok, Y. (2024).** *Adaptive gradient optimizer*. <https://www.analyticsvidhya.com/blog/2024/11/adaptive-gradient-optimizer/> (E.T.: 04/05/2025)
- Alpaydın, E., & Kaynak, C. (1998).** Optical Recognition of Handwritten Digits [Dataset]. UCI Machine Learning. <https://doi.org/10.24432/C50P49>
- Arı, A., & Berberler, M. E. (2017).** Yapay sinir ağları ile tahmin ve sınıflandırma problemlerinin çözümü için arayüz tasarımı. *Acta Infologica*, 1(2), 55–73.
- Azizoğlu, F., & Ünsal, E. (2022).** Hata IoT verilerinin makine öğrenmesi teknikleri ile tahmini. *European Journal of Science and Technology (ECJSE)*, 9(4), 1388–1397. <https://doi.org/10.31202/ecjse.1135485>.
- Bassey, J., Qian, L., & Li, X. (2021).** *A Survey of Complex-Valued Neural Networks*. arXiv:2101.12249.
- Barrachina, J. A., Ren, C., Vieillard, G., Morisseau, C., & Ovarlez, J.-P. (2023).** *Theory and implementation of complex-valued neural networks* (arXiv:2302.08286). arXiv. <https://doi.org/10.48550/arXiv.2302.08286>
- Ciampiconi, L., Elwood, A., Leonardi, M., Mohamed, A., & Rozza, A. (2023).** A survey and taxonomy of loss functions in machine learning.
- Coşkun, F., & Gülleroğlu, H. D. (2021).** Yapay zekanın tarih içindeki gelişimi ve eğitimde kullanılması. *Ankara University Journal of Faculty of Educational Sciences (JFES)*, 54(3), 947–966.
- Çelik, Ş. (2022).** Yapay sinir ağlarının farklı aktivasyon fonksiyonlarında uygulaması: Türkiye’de ördek popülasyonu öngörüsü. *Osmaniye Korkut Ata Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 5(2), 800-811. <https://doi.org/10.47495/okufbed.990995>
- GeeksforGeeks, (2023).** (n.d.). *Learning rate decay*. <https://www.geeksforgeeks.org/learning-rate-decay/> (E.T.: 14/05/2025)
- Goodfellow, I., Bengio, Y., & Courville, A. (2016).** *Deep learning*. MIT Press.
- Günler Pirim M. A. (2017).** Neural network based feature extraction for handwritten digit recognition (Doctoral dissertation, Atılım University). Modeling and Design

of Engineering Systems (MODES), Electrical & Electronics Engineering, Computer Engineering.

- Gürüler, H. (2017).** A novel diagnosis system for Parkinson's disease using complex-valued artificial neural network with k-means clustering feature weighting method. *Neural Comput & Applic* **28**, 1657–1666. <https://doi.org/10.1007/s00521-015-2142-2>
- Igiri, C. P., Anyama, O. U., & Silas, A. I. (2015).** Effect of learning rate on artificial neural network in machine learning.
- Iranmanesh, S., & Mahadevi, M. A. (2009).** Differential adaptive learning rate method for back propagation neural networks. *World Academy of Science, Engineering and Technology*, *50*, 285–288.
- Jadon, A., Patil, A., & Jadon, S. (2022).** A comprehensive survey of regression based loss functions for time series forecasting. *arXiv preprint arXiv:2211.02989*. <https://arxiv.org/abs/2211.02989>.
- Karakurt, M., Oymak, E. A., Hark, H., Erdoğan, M. C., & Karci, A. (2022).** Karci sinir ağlarının uygulaması ve performans analizi. *Journal of Computer Science*, *7*, 68–80.
- Karakurt, M., Saygılı, H., & Karci, A. (2024).** KarciFANN yönteminde aktivasyon fonksiyonlarının karşılaştırılması. In *8th International Artificial Intelligence and Data Processing Symposium (IDAP2024)*. IEEE.
- Karakurt, M., Saygili, H., & Karci, A. (2025).** Karci fractional artificial neural networks (KarciFANN): a new artificial neural networks model without learning rate and its problems. *Turkish Journal of Electrical Engineering and Computer Sciences* *33* (3): 248-263.
- Keskenler, M. F., & Keskenler, E. F. (2017).** Geçmişten günümüze yapay sinir ağları ve tarihçesi. *Takvim-i Vekayi*, *5*(2), 8–18.
- Kingma, D. P., & Ba, J. (2014).** Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*. <https://arxiv.org/abs/1412.6980>
- Koşan, M. A., Coşkun, A., & Karacan, H. (2019).** Yapay Zeka Yöntemlerinde Entropi. *Journal of Information Systems and Management Research*, *1*(1), 15-22.
- Kornblith, S., Lee, H., Chen, T., & Norouzi, M. (2020).** What's in a loss function for image classification? *arXiv preprint arXiv:2010.16402*. <https://arxiv.org/abs/2010.16402>
- Lee, C. Y., Hasegawa, H., & Gao, S. C. (2022).** Complex-valued neural networks: A comprehensive survey,” *IEEE/CAA J. Autom. Sinica*, vol. 9, no. 8, pp. 1406–1426, Aug. doi: 10.1109/JAS.2022.105743
- Li, J., Diao, Y., Li, M., & Yin, X. (2009).** Stability analysis of discrete Hopfield neural networks with the nonnegative definite monotone increasing weight function matrix. *Discrete Dynamics in Nature and Society*, 2009, Article 673548.
- Li, Y., Wei, C., & Ma, T. (2019).** Towards explaining the regularization effect of initial large learning rate in training neural networks. *Proceedings of the 33rd International Conference on Neural Information Processing Systems*. Curran Associates Inc., Red Hook, NY, USA, Article 1047, 11674–11685, 2019.

- Luo, L., Xiong, Y., Liu, Y., & Sun, X. (2019).** Adaptive gradient methods with dynamic bound of learning rate. *arXiv preprint* arXiv:1902.09843. <https://arxiv.org/abs/1902.09843>
- Mallick, U. (2024).** *Cosine learning rate schedulers in PyTorch.* <https://medium.com/@utkrisht14/cosine-learning-rate-schedulers-in-pytorch-486d8717d541> (E.T.: 30/05/2025).
- Nagel, M., Fournarakis, M., Bondarenko, Y., & Blankevoort, T. (2022).** Overcoming oscillations in quantization-aware training. In *International Conference on Machine Learning (ICML)*.
- Nitta, T. (2003).** The Computational Power of Complex-Valued Neuron. In O. Kaynak, E. Alpaydin, E. Oja, & L. Xu (Eds.), **Artificial Neural Networks and Neural Information Processing — ICANN/ICONIP 2003** (Lecture Notes in Computer Science, vol. 2714, pp. 118–125). Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-44989-2_118
- OpenML. (2014).** *Dataset ID: 1041.* <https://www.openml.org/search?type=data&status=active&id=1041> (E.T.: 17/05/2025)
- OpenML. (2019).** *Dataset ID: 41982.* <https://www.openml.org/search?type=data&status=active&id=41982> (E.T.: 17/05/2025)
- OpenML. (2023).** *Dataset ID: 45082.* <https://www.openml.org/search?type=data&status=active&id=45082> (E.T.: 17/05/2025)
- Özcan, A. (2024).** *Yapay sinir ağları ve bulanık mantık ile sanayi üretim endeksi tahmini* (Yüksek lisans tezi). Tekirdağ Namık Kemal Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Bilimleri ve Mühendisliği Ana Bilim Dalı.
- Öztürk, K., & Şahin, M. E. (2018).** Yapay sinir ağları ve yapay zekaya genel bir bakış. *Takvim-i Vekayi*, 6(2), 25–36.
- Pirim, H. (2006).** Yapay zeka. *Journal of Yaşar University*, 1, 81–93.
- Qi, J., Du, J., Siniscalchi, S. M., Ma, X., & Lee, C.-H. (2020).** On mean absolute error for deep neural network based vector-to-vector regression. *IEEE Signal Processing Letters*, 27, 1485–1489.
- Ren, G., Cao, Y., Wen, S., Huang, T., & Zeng, Z. (2018).** A modified Elman neural network with a new learning rate scheme. *Neurocomputing*, 286, 11–18.
- Ruder, S. (2016).** *An overview of gradient descent optimization algorithms.* *arXiv preprint* arXiv:1609.04747. <https://arxiv.org/abs/1609.04747>
- Saygili, H., Karakurt, M., & Karci, A. (2025).** KarcıFANN Yönteminin Yakınsama Kabiliyetinin Analiz Edilmesi. *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, 37(2), 897-907. <https://doi.org/10.35234/fumbd.1661969>
- Senior, A. W., Heigold, G., Ranzato, M., & Yang, K. (2013).** An empirical study of learning rates in deep neural networks for speech recognition. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing* (pp. 6724–6728). IEEE.
- Ser, G., & Bati, C. T. (2019).** Derin Sinir Ağları ile En İyi Modelin Belirlenmesi: Mantar Verileri Üzerine Keras Uygulaması. *Yuzuncu Yıl University Journal of*

- Sheel, S., Varshney, T., & Varshney, R. (2007).** Accelerated learning in MLP using adaptive learning rate with momentum coefficient. In *2007 International Conference on Industrial and Information Systems* (pp. 307–310). IEEE.
- Smith, J.W. (2023).** https://complextorch.readthedocs.io/en/stable/nn/modules/activation/complex_relu.html (E.T.: 08/09/2025)
- Smith, L. N. (2017).** Cyclical Learning Rates for Training Neural Networks," 2017 IEEE Winter Conference on Applications of Computer Vision (WACV), Santa Rosa, CA, USA, pp. 464-472, doi: 10.1109/WACV.2017.58.
- Smith, L. N. (2018)** A Disciplined Approach To Neural Network Hyper-parameters: part 1 – learning rate, Batch size, momentum, and weight decay.
- Smith, L. N. & Topin, N. (2018).** Super-Convergence: Very Fast Training of Neural Networks Using Large Learning Rates.
- Starodub, A., Eliseeva, N., & Georgiev, M. (2021).** Gradient-based algorithm for tracking the activity of neural network weights changing. *EPJ Web of Conferences*, 248, 01012.
- Takase, T., Oyama, S., & Kurihara, M. (2018).** Effective neural network training with adaptive learning rate based on training loss. *Neural Networks*, 101, 68–78. <https://doi.org/10.1016/j.neunet.2018.01.016>.
- Teenakunsoth, T. (2020).** *MNIST Rotated*. Kaggle. <https://www.kaggle.com/code/teenakunsoth/mnist-rotated> (E.T.: 17 Eylül 2025)
- Tektaş, M., & Tektaş, N. (2024).** Matematik 1 – karmaşık sayılar [PDF]. Tektasi.net. <https://tektasi.net/wp-content/uploads/2024/11/Matematik-1-Karmasik-Sayilar.pdf>, (E.T.: 24/08/2025)
- Terven, J., Cordova-Esparza, D.-M., Ramirez-Pedraza, A., & Chávez Urbiola, E. (2023).** Loss functions and metrics in deep learning: A review.
- Trabelsi, C., Bilaniuk, O., Zhang, Y., Serdyuk, D., Subramanian, S., Santos, J. F., Mehri, S., Rostamzadeh, N., Bengio, Y., & Pal, C. J. (2018).** Deep complex networks. *International Conference on Learning Representations (ICLR)*. <https://arxiv.org/abs/1705.09792>
- Üstün, O., & Yıldız, İ. (2009).** Geri-yayılmalı öğrenme algoritmasındaki öğrenme parametrelerinin genetik algoritma ile belirlenmesi. *Uluslararası Teknoloji Bilim Dergisi (UTBD)*, 1(2), 61–73.
- UCI Machine Learning Repository (1).** <https://archive.ics.uci.edu/dataset/94/spambase>, (E.T.: 17/07/2025)
- UCI Machine Learning Repository (2).** <https://archive.ics.uci.edu/dataset/80/optical+recognition+of+handwritten+digits>, (E.T.: 17/07/2025)
- UCI Machine Learning Repository (3).** <https://archive.ics.uci.edu/dataset/31/coverttype>, (E.T.: 17/07/2025)
- Wei, J., Zhang, X., Ji, Z., Wei, Z., & Li, J. (2022).** DPLRS: Distributed population learning rate schedule. *Future Generation Computer Systems*, 132, 40–50.

- Wilson, D. R., & Martinez, T. R. (2001).** The need for small learning rates on large problems. In *Proceedings of the International Joint Conference on Neural Networks (IJCNN'01)* (Vol. 1, pp. 115–119). Washington, DC, USA: IEEE. <https://doi.org/10.1109/IJCNN.2001.939002>
- Yedida, R., Saha, S., & Prashanth, T. (2021).** LipschitzLR: Using theoretically computed adaptive learning rates for fast convergence. *Applied Intelligence*, 51, 1460–1478.
- You, K., Long, M., Wang, J., & Jordan, M. I. (2019).** *How does learning rate decay help modern neural networks?* arXiv. <https://arxiv.org/abs/1908.01878>
- You, L., Zhao, X., Huang, J., Hu, H., & Liu, B. (2023).** Research on machine learning with algorithms and development. *Journal of Theory and Practice of Engineering Science*, 3(12), 7–14.
- Zeiler, M. D. (2012).** *ADADELTA: An adaptive learning rate method.* arXiv preprint arXiv:1212.5701. <https://arxiv.org/abs/1212.5701>
- Zheng, D. D., Li, X., Ren, X., & Na, J. (2023).** Intelligent control for robotic manipulator with adaptive learning rate and variable prescribed performance boundaries. *Journal of the Franklin Institute*, 360(11), 7037–7062.

ÖZGEÇMİŞ

Ad Soyad	: Hülya SAYGILI
Lisans	: Fırat Üniversitesi Bilgisayar Mühendisliği (2008)
Yüksek Lisans	: Fırat Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği ABD (2012)
Mesleki Deneyimler	: Karayolları 17. Bölge Müdürlüğü Mühendis (2009-2010) Karayolları 11. Bölge Müdürlüğü Tünel Bakım İşl. Şefi (2010-2020) Karayolları 11. Bölge Müdürlüğü Mühendis (2020-)

DOKTORA TEZİNDEN TÜRETİLEN ÇALIŞMALAR

- Karakurt, M., Oymak, E.A., **Hark, H.**, Erdoğan, M.C. & Karcı, A. (2022). Karcı Sinir Ağlarının Uygulaması ve Performans Analizi. *Journal of Computer Science*, 7(2), 68-80. DOI: 10.53070/bbd.1194017.
- Karakurt, M., **Saygılı, H.** ve Karcı, A. (2024). RMSE Yönteminde Aktivasyon Fonksiyonlarının Karşılaştırılması, *8th International Artificial Intelligence and Data Processing Symposium (IDAP2024)*, IEEE, doi: 10.1109/IDAP64064.2024.10711149. <https://ieeexplore.ieee.org/document/10711149>.
- Karakurt, M., **Saygılı, H.** & Karcı, A. (2025). Karcı fractional artificial neural networks (KarcıFANN): a new artificial neural networks model without learning rate and its problems. *Turkish Journal of Electrical Engineering and Computer Sciences*, 33(3), 248-263. doi: 10.55730/1300-0632.4125.
- **Saygılı, H.**, Karakurt, M., & Karcı, A. (2024). Comparison of Loss Functions in the KarcıFANN Method. *In 2024 8th International Artificial Intelligence and Data Processing Symposium (IDAP)*, (pp. 1-5). IEEE.
- **Saygili, H.**, Karakurt, M., & Karcı, A. (2025). KarcıFANN Yönteminin Yakınsama Kabiliyetinin Analiz Edilmesi. *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, 37(2), 897-907. <https://doi.org/10.35234/fumbd.1661969>
- **Saygılı, H.**, Karakurt, M. & Karcı, A. (2025). Karcı kesir dereceli yapay sinir ağı (KarcıFANN): öğrenme katsayısı, aşırı uyum ve yetersiz uyum sorunlarını çözme. *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*.