

YSA’larda Kesir Dereceli Türev Kullanımının Öğrenmeye Etkisi

The Effect of Using Fractional-Order Derivatives on Learning in Artificial Neural Networks

Meral Karakurt^{1*} 

^{1*}Bilgisayar Teknolojileri Bölümü, Osmaniye Korkut Ata Üniversitesi Düziçi MYO, Osmaniye, Türkiye

(meralkarakurt@osmaniye.edu.tr)

Received:Oct.22,2025

Accepted:Nov.30,2025

Published:Dec.01,2025

Özetçe: YSA’ların tasarımı aşamasında kullanılan optimizasyon yöntemleri, katman ve nöron sayıları, aktivasyon ve hata fonksiyonları gibi hiperparametreler, modellerin başarımını belirlemektedir. Gradyan tabanlı optimizasyon yöntemleri, modellerin öğrenme sürecinde yaygın olarak kullanılmaktadır. SGD yönteminde ezberleme, öğrenememe, genelleme yapamama, yakınsayamama gibi önemli problemlerle karşılaşmaktadır. Bu problemleri çözmek ve kesir dereceli türevlerle sistemleri daha iyi modellemek amacıyla KarcıFANN optimizasyon yöntemi kullanılmaktadır.

Bu çalışmada, MNIST ve Dry Bean veri setlerinin sınıflandırılması amacıyla KarcıFANN ve SGD öğrenme yöntemlerinin kullanıldığı modeller tasarlanmıştır. Bu modeller, çeşitli katman sayısı konfigürasyonları ile farklı aktivasyon ve hata fonksiyonları kullanılarak eğitilmiş ve performansları karşılaştırmalı olarak değerlendirilmiştir. Elde edilen bulgular, kesir dereceli türev kullanımıyla modellerin performansının arttığını ve KarcıFANN yönteminin SGD’ye alternatif bir yaklaşım olarak değerlendirilebileceğini göstermektedir.

Anahtar Kelimeler: KarcıFANN, SGD, kesir dereceli türev, sınıflandırma.

Abstract: During the design phase of artificial neural networks (ANNs), optimization techniques, the number of layers and neurons, as well as activation and loss functions—collectively referred to as hyperparameters—play a decisive role in model performance. Gradient-based optimization methods are widely employed throughout the learning process of these models. However, Stochastic Gradient Descent (SGD) often encounters critical issues such as overfitting, underfitting, poor generalization, and failure to converge. To address these challenges and to achieve more effective modeling of systems through fractional-order derivatives, the KarcıFANN optimization method is employed.

In this study, models employing the KarcıFANN and SGD learning algorithms were designed for the classification of the MNIST and Dry Bean datasets. These models were trained using various configurations of layer numbers, as well as different activation and loss functions, and their performances were comparatively evaluated. The results demonstrate that incorporating fractional-order derivatives leads to improved model performance, highlighting the potential of the KarcıFANN method as an effective alternative to the conventional SGD algorithm.

Keywords: KarcıFANN, SGD, fractional order derivative, classification.

1. Giriş

Bilgisayar teknolojilerinin gelişmesi, paylaşılan veri miktarlarının ve bu verilerin işlenmesi ihtiyacının artmasına neden olmaktadır. Canlıların, verileri öğrenerek bilgiye dönüştürme sürecinin matematiksel formüllerle ifade edilmesi amacıyla yapay sinir ağları (YSA) kullanılmaktadır. YSA’lar; eğitim, sağlık, ekonomi, tarım, uzay araştırmaları gibi hayatın neredeyse tüm alanlarında tahmin,

sınıflandırma, karar verme, nesne tanıma, kümeleme ve daha birçok problemin çözümünde başarılı bir şekilde kullanılmaktadır.

YSA'ların tasarımında katman sayısı, nöron sayısı, optimizasyon yöntemi, aktivasyon fonksiyonu, hata fonksiyonu ve daha birçok hiperparametre kullanılmaktadır. Optimizasyon yöntemi, hata değerini en aza indirmek ve doğruluk, hassasiyet ile duyarlılık gibi metriklerin değerlerini en yüksek seviyeye çıkarmak amacıyla kullanılmaktadır (Karakurt ve diğ., 2025). Aktivasyon fonksiyonu, bir nöronun çıktı değerini belirlemektedir. Hata fonksiyonu ise bir modelin ürettiği çıktı değeri ile modelden beklenen çıktı değeri (etiket bilgisi) arasındaki farkın ölçülmesinde kullanılmaktadır.

Literatürde çeşitli optimizasyon yöntemleri ile tasarlanan modellerin kullanıldığı birçok çalışma yapılmıştır. Örneğin; Dogo vd. (2018), bir evrişimli sinir ağı (ConvNet) modelinde Stokastik Gradyan İniş (SGD), Vanilya SGD (vSGD), momentumlu SGD (SGDm), momentumlu ve Nesterovlu SGD (SGDm+n), Uyarlamalı Delta (AdaDelta), Uyarlamalı Moment Tahmini (Adam), Nesterov hızlandırılmalı Uyarlamalı Moment Tahmini (Nadam), Uyarlamalı Gradyan (AdaGrad), Ortalama Karekök Yayılımı (RMSProp) ve Sonsuzluk normuna dayalı Uyarlamalı moment tahmini Uzantısı (Adamax) yöntemlerinin performanslarını karşılaştırmışlardır. Yaptıkları deneyler sonunda, Nadam yönteminin en başarılı ve AdaDelta yönteminin de en başarısız yöntem olduğunu göstermişlerdir (Dogo vd., 2018). Diğer bir çalışmada, Tian vd. (2023), görüntü işleme ve doğal dil işleme problemlerinde SGD, Adam, AdamW, Adagrad, Adadelat ve RMSprop yöntemlerinin kullanımının avantaj ve dezavantajlarını açıklamışlardır (Tian vd., 2023).

Chen vd. (2024), Grünwald–Letnikov (G–L) kesir dereceli türevine dayalı AdaGL adlı uyarlamalı (adaptive) öğrenme oranlı bir optimizasyon yöntemi önermişlerdir. AdaGL yöntemi ile daha kararlı ve hızlı yakınsama sağlamak amacıyla uzun dönemli ve kısa dönemli gradyan bilgilerini birlikte kullanarak parametre güncellemelerini dinamik biçimde düzenlemişlerdir. Evrişimsel Sinir Ağı (CNN), Graf Sinir Ağı (GNN), Üretici Çekişmeli Ağ (GAN) ve Uzun Kısa Süreli Bellek (LSTM) modelleri üzerinde yaptıkları deneyler, AdaGL'nin doğruluk ve genelleme açısından üstün performans gösterdiğini ortaya koymuştur (Chen vd., 2024).

Cheng vd. (2025), optimizasyon sürecinde model karmaşıklığının dikkate alınmamasının yakınsama kararsızlığına yol açtığını belirtmiş ve bu soruna çözüm olarak AMC (Adam with Model Complexity) adlı yeni bir uyarlamalı optimizasyon yöntemi önermişlerdir. AMC yönteminde, modelin karmaşıklık düzeyi ölçülmekte ve bu değere göre öğrenme oranı dinamik bir şekilde azaltılmakta ya da artırılmaktadır. İlgili çalışmada, MNIST, Fashion-MNIST veri setlerini sınıflandırdıkları dört farklı VGG modelinde AMC yönteminin kullanımının Adam ve AdamW gibi yöntemlere göre daha kararlı ve daha hızlı yakınsama sağladığını göstermişlerdir (Cheng vd., 2025).

Herrera-Alcántara (2022), gradyan tabanlı optimizasyon yöntemlerine Caputo kesir dereceli türevini ekleyerek yeni optimizasyon yöntemleri (FSGD, FAdagrad, FAdadelat, FRMSProp, FAdam, FSGDP ve FAdamP) önermiştir. MNIST ve insan aktivitesi tanıma (human action recognition, HAR) veri setleri üzerinde yaptığı deneylerle kesir dereceli optimizasyon yöntemlerinin klasik yöntemlere kıyasla daha kararlı yakınsama ve daha iyi genelleme performansına sahip olduğunu ispatlamıştır (Herrera-Alcántara, 2022).

Yang vd. (2022), çok görevli öğrenme (Multi-Task Learning, MTL) modellerinde aynı parametreleri kullanan farklı görevlerin arasında oluşan öğrenme dengesizliğinin optimizasyon sürecini olumsuz etkilediğini belirterek AdaTask adında yeni bir uyarlamalı öğrenme oranı yöntemi önermişlerdir. AdaTask, her parametre için görev-bazlı kümülatif gradyan büyüklüklerini kullanarak her görev için öğrenme oranını ayrı ayrı ayarlamaktadır. İlgili çalışmada, görev bazlı uyarlamalı öğrenme oranı yönteminin etkinliğini hem sentetik bir veri seti hem de CityScapes, TikTok ve WeChat gibi gerçek veri setleri üzerinde deneysel olarak doğrulamışlardır (Yang vd., 2022).

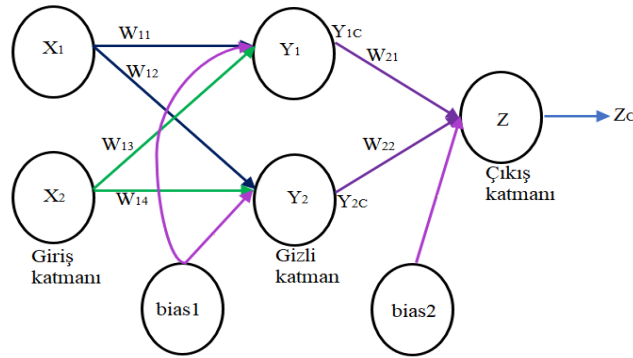
Jeong vd. (2024), klasik optimizasyon yöntemlerinin yüksek hesaplama maliyeti ve zayıf genelleme performansını iyileştirmek amacıyla AdaLo (Adaptive Learning Rate Optimizer with Loss) ismini verdikleri yeni bir uyarlamalı öğrenme oranı yöntemi önermişlerdir. Bu yöntem, öğrenme oranını doğrudan hata fonksiyonu değerinden türeterek dinamik biçimde ayarlamakta ve böylece kararlı bir optimizasyon süreci sağlamaktadır. CIFAR-10 ve CIFAR-100 veri setlerinin sınıflandırıldığı deneysel

çalışmada, AdaLo yöntemi ile Adam, AdaBelief ve diffGrad yöntemleri karşılaştırılmış ve AdaLo'nun daha yüksek doğruluk değeri ürettiği gösterilmiştir (Jeong vd., 2024).

Saygılı vd. (baskıda), KarcıFANN yöntemi ile SGD yöntemini karşılaştırmak amacıyla farklı veri setleri üzerinde yaptıkları çalışmanın sonunda KarcıFANN'ın ezberleme, öğrenememe ve sapma gibi problemleri çözebileceğini ve genelleme performansının SGD'ye kıyasla daha yüksek olduğunu göstermişlerdir. Karakurt vd. (2024), aktivasyon fonksiyonlarının KarcıFANN yöntemiyle kullanımının performansa etkilerini gözlemlemek üzere sigmoid, hiperbolik tanjant (tanh) ve hiperbolik tanjant sigmoid (tansig) aktivasyon fonksiyonlarını kullanmışlardır. Gina_Prior2 veri setini sınıflandırdıkları KarcıFANN modelinde Sigmoid fonksiyonunun diğerlerine kıyasla üstün performans gösterdiğini kanıtlamışlardır. Benzer bir çalışmada, Saygılı vd. (2024), KarcıFANN yöntemiyle birlikte farklı hata fonksiyonlarını kullandıkları modeller tasarlayarak başarımlarını karşılaştırmışlardır. Ortalama mutlak hata (MAE), ortalama karesel hata (MSE) ve ortalama karesel hatanın karekökü (RMSE) fonksiyonları ile yaptıkları deneylerde, MSE fonksiyonunun en yüksek başarıma sahip olduğunu göstermişlerdir (Saygılı ve diğ., 2024).

2. Yapay Sinir Ağı (YSA)

Biyolojik beynin çalışma mantığının bilgisayar ortamında taklit edilmesi ve öğrenme, hatırlama, nesne tanıma gibi özelliklerin bilgisayarlara kazandırılması amacıyla YSA'lar kullanılmaktadır. Özellikle karmaşık problemlerin çözümünde çok katmanlı YSA'lar başarılı olmaktadır. Şekil 1 ile gösterilen çok katmanlı bir YSA, temelde bir girdi katmanı, en az bir gizli (ara) katman ve bir çıktı katmanından oluşmaktadır. Girdi katmanı, veri setine ait sayısal değerleri içermektedir. Gizli katman, verilerin işlenerek öğrenmenin gerçekleştiği katmandır. Çıktı katmanı ise gizli katmanda üretilen çıktıları dışarıya sunan nöronlardan oluşmaktadır. Girdi ve çıktı katman nöronları veri setlerine bağlı olarak değişirken gizli katman sayısı ve bu katmanlardaki nöron sayıları problemde problemde göre değişmektedir.



Şekil 1. Üç katmanlı YSA modeli.

YSA'ların eğitim sürecinde girdi katmanından çıktı katmanına doğru tek yönde yapılan işlemler bütününe ileri besleme; ileri besleme sonucu elde edilen hata değerinin çıktı katmanından geriye doğru yayılmasına ise geri besleme denir (Rumelhart ve diğ., 1986).

İleri besleme yönteminde $X = X_1, X_2$ girdi değerleri, $W = W_{11}, W_{12}, W_{13}, W_{14}$ girdilere ait ağırlık değerleri ve f aktivasyon fonksiyonu olmak üzere modelin ürettiği Z_c çıktı değeri (1) denklemi ile hesaplanmaktadır.

$$Y_1 = X_1 \cdot W_{11} + X_2 \cdot W_{13} + bias1, Y_{1c} = f(Y_1), Y_2 = X_1 \cdot W_{12} + X_2 \cdot W_{14} + bias1, Y_{2c} = f(Y_2)$$

$$Z = Y_{1c} \cdot W_{21} + Y_{2c} \cdot W_{22} + bias2, Z_c = f(Z) \quad (1)$$

YSA'ların ürettiği çıktı değerinin beklenen çıktı değerinden uzaklaşma miktarı hata fonksiyonu ile ölçülür. Ortalama Karesel Hata (Mean Squared Error, MSE), YSA başarımını değerlendirmek amacıyla yaygın olarak kullanılmaktadır. (2) denklemi ile gösterilen MSE fonksiyonu, ağın ürettiği değerler ile üretmesi beklenen değerler arasındaki farkların karelerinin ortalamasıdır (Jadon ve diğ, 2024).

$$MSE = \frac{1}{n} \sum_{i=0}^n (R_i - Z_{ci})^2 \quad (2)$$

(2) ve (3) numaralı denklemlerde; Z_C modelin ürettiği tahmin değeri, R modelden üretmesi beklenen değeri, n örnek sayısını ve m sınıf sayısını göstermektedir.

Sınıflandırma problemlerinde yaygın olarak kullanılan diğer bir hata fonksiyonu, Çapraz Entropi (Cross-Entropy, CE)'dir. (3) denklemi ile gösterilen bu fonksiyon, modelin çıktısında elde edilen olasılık dağılımının, modelden üretmesi beklenen gerçek dağılımdan ne kadar sapma gösterdiğini hesaplamaktadır. (Koşan ve diğ, 2019; Seyyarer ve diğ, 2020).

$$CE = -\frac{1}{n} \sum_{i=1}^n \sum_{k=1}^m R_{i,k} \log(Z_{ci,k}) \quad (3)$$

F1 skoru, doğruluk (accuracy), duyarlılık (recall) ve kesinlik (precision) metrikleri modelin genelleme performansını değerlendirmede kullanılan diğer metriklerdir. Karmaşıklık matrisi (confusion matrix) kullanılarak hesaplanan bu metriklerin değerleri 1'e yaklaştıkça başarımlar artmaktadır. (4) denklemi ile sunulan Doğruluk (accuracy) değeri, modelin doğru tahmin ettiği değerlerin tüm tahminlere oranıdır (Christen ve diğ, 2023).

$$Doğruluk = \frac{DP+DN}{DP+DN+YP+YN} \quad (4)$$

Kesinlik ya da hassasiyet (precision) değeri, modelin doğru tahmin ettiği pozitif değerlerin, modelin pozitif olarak tahmin ettiği tüm değerlere oranı olarak tanımlanmakta ve $Kesinlik = \frac{DP}{DP+YP}$ denklemi ile hesaplanmaktadır (Lewis, 1991; Powers, 2020). Duyarlılık ya da hatırlama (recall, sensitivity) değeri, modelin doğru tahmin ettiği pozitif örneklerin gerçek pozitif örneklerin tümüne oranı olmakla birlikte $Duyarlılık = \frac{DP}{DP+YN}$ denklemi ile hesaplanmaktadır (Lewis, 1991; Powers, 2020). (5) denklemi ile sunulan F1 skoru (F-measure) değeri ise duyarlılık ve kesinlik değerlerinin harmonik ortalamasıdır (Christen ve diğ, 2023).

$$F1 = 2 \cdot \frac{Kesinlik \cdot Duyarlılık}{Kesinlik + Duyarlılık} \quad (5)$$

Aktivasyon fonksiyonu, bir modelde kullanılan nöronların öğrenmeye ne kadar etki edeceğini belirlemektedir (Nwankpa ve diğ, 2018). Doğrusal olmayan karmaşık yapıdaki problemlerin çözümünde doğrusal olmayan aktivasyon fonksiyonları kullanılmaktadır (Dubey ve diğ, 2022; Karakurt ve diğ, 2024). Sigmoid fonksiyonu, nöronların çıktılarını 0 ile 1 arasında sınırlayan ve doğrusal olmayan bir aktivasyon fonksiyonudur. Z , bir nöronun çıktısı olmak üzere, $f(Z) = \frac{1}{1+e^{-Z}}$ denklemi ile hesaplanan Sigmoid fonksiyonu, türevlenebilir ve sürekli bir fonksiyondur. Bu makalede, $f(Z) = Z_C$ olarak kullanılmaktadır.

ReLU (düzeltilmiş doğrusal birim - rectified linear unit) fonksiyonu, yaygın kullanıma sahip diğer bir aktivasyon fonksiyonudur. Bu fonksiyon, negatif girdi değerleri için sıfır değerini üretirken pozitif girdileri doğrudan çıktıya göndermekte ve x , bir nöronun çıktısını göstermek üzere $f(x) = \max(0, x)$ denklemi ile hesaplanmaktadır (Seyyarer ve diğ, 2020).

2.1. SGD (Stokastik Gradyan İniş – Stochastic Gradient Descent)

SGD, YSA’larda hata fonksiyonunun minimum değerini bulmak amacıyla kullanılan sezgisel bir optimizasyon yöntemidir (LeCun ve diğ, 1998). Bu yöntemde, hata değerleri geri yayılımımla tekrarlı olarak hesaplanmakta ve bu hatalar sonraki iterasyonlarda ağırlıkların güncellenmesinde kullanılmaktadır.

SGD ile tasarlanan modellerin eğitim sürecinde, veri setinin bir kısmı ya da rastgele seçilen bir örneği için hesaplamalar yapılarak ağırlıklar güncellenmektedir (Karakurt ve diğ, 2024; Saygılı ve diğ, 2024). SGD yönteminin uygulanmasında gradyan (kısmi türev) hesaplama yöntemi kullanılmaktadır. Şekil 1 ile sunulan modelde, W_{11} ağırlığının toplam hataya (H) etkisi, (6) denkleminde gösterilen Newton türevinin zincir kuralı ile hesaplanmaktadır.

$$\frac{\partial H}{\partial W_{11}} = \frac{\partial H}{\partial z_C} \cdot \frac{\partial z_C}{\partial z} \cdot \frac{\partial z}{\partial y_{1C}} \cdot \frac{\partial y_{1C}}{\partial y_1} \cdot \frac{\partial y_1}{\partial W_{11}} \quad (6)$$

α , öğrenme katsayısı olmak üzere W_{11} ağırlığının herhangi bir $t+1$ adımında güncellenmesi işlemi, $w_{11t+1} = w_{11t} - \alpha \cdot \frac{\partial H}{\partial w_{11t}}$ denklemi ile ifade edilmektedir (LeCun ve diğ, 1998).

2.2. KarcıFANN (Karcı Kesir Dereceli Yapay Sinir Ağı - Karcı Fractional Artificial Neural Network)

KarcıFANN yöntemi, Karcı (2013a, 2013b, 2015a, 2015b, 2015c, 2019, 2021) tarafından tanımlanan Karcı Kesir Dereceli Türevin YSA’ların eğitim sürecine eklendiği bir optimizasyon yöntemidir. Bu yöntemle YSA’larda karşılaşılan öğrenememe, ezberleme ve sapma gibi önemli problemlerin çözülmesi amaçlanmaktadır.

KarcıFANN yöntemiyle tasarlanan modelin eğitim sürecinde, her iterasyon için kesir dereceli türev hesaplamaları yapılarak ağırlıklar güncellenmektedir. Böylece dinamik bir yapı oluşturularak modelin kendi içinde öğrenmesi sağlanmakta ve öğrenme hızı artırılmaktadır.

Karcı kesir dereceli türevi (D_α^K) tanımından gelen $\left(\frac{H}{W_{11}}\right)^{\alpha-1}$ çarpanının Newton türevine eklenmesiyle KarcıFANN yöntemi hesaplanmaktadır. Buradaki alfa (α) değeri, kesir derecesini temsil etmektedir. Şekil 1 ile gösterilen w_{11} ağırlığının toplam hataya (H) etkisinin KarcıFANN yöntemiyle hesaplanması işlemi (7) denklemi ile formüle edilmektedir.

$$D_\alpha^K \frac{H}{W_{11}} = \left(\frac{H}{W_{11}}\right)^{\alpha-1} \cdot \frac{\partial H}{\partial z_C} \cdot \frac{\partial z_C}{\partial z} \cdot \frac{\partial z}{\partial y_{1C}} \cdot \frac{\partial y_{1C}}{\partial y_1} \cdot \frac{\partial y_1}{\partial W_{11}} \quad (7)$$

Burada kullanılan $\left(\frac{H}{W_{11}}\right)^{\alpha-1}$ çarpanı, her iterasyonda elde edilen hata ve ağırlık değerlerine göre değişmekte ve bu sayede, SGD yönteminde sabit bir sayı olarak kullanılan öğrenme katsayısı problemi çözülmektedir. (Karakurt ve diğ, 2022; Karakurt ve diğ, 2024; Saygılı ve diğ, 2024; Karakurt ve diğ, 2025).

Herhangi bir $t + 1$ adımında w_{11} ağırlık değerinin güncellenmesi işlemi, $w_{11t+1} = w_{11t} - D_\alpha^K \frac{H}{W_{11t}}$ denklemi ile hesaplanmaktadır.

3.Farklı YSA Modellerinde KarcıFANN Yönteminin Kullanımı

MNIST veri seti, el yazısı ile yazılan 70000 rakamın farklı gri tonlardaki görüntülerini içermektedir (LeCun ve diğ, 1998). 0-9 rakamlarından oluşan 10 sınıflı veri setindeki tüm görüntüler 28×28 boyutunda olup her görüntü 784 ($28 \times 28 = 784$) elemanlı tek boyutlu sayısal vektörlere dönüştürülerek

modellerin girdi katmanlarına sunulmuştur. Ayrıca, veri setinin 55000 görüntüsü eğitim, 5000 görüntüsü doğrulama (validation) ve 10000 görüntüsü de test aşaması için kullanılmıştır. Dry Bean veri seti ise 7 sınıfa (barbunya, şeker, dermosan, bombay, cali, horoz ve sıra) ait toplam 13611 fasulye görüntüsünden oluşmaktadır. Bu görüntülerden 8611 tanesi eğitim, 2500 tanesi doğrulama ve 2500 tanesi de test aşamasında kullanılmıştır. Her görüntüye özgü 16 özellik bilgisi sayısallaştırılmış ve sayısal bilgiler modellerin girdi katmanlarında kullanılmıştır.

Uygulama aşamasında, açık kaynak kodlu Anaconda tümleşik geliştirme ortamının (IDE) Jupyter notebook uygulamasında Python dili kullanılmıştır. MNIST veri seti, PyTorch'un torchvision modülü üzerinden yüklenmiş; Dry Bean veri seti ise UCI Makine Öğrenmesi Deposu'ndan (UCI Machine Learning Repository) Excel formatında indirilerek kullanılmıştır. Çeşitli hiperparametreler kullanılarak KarcıFANN ve SGD modelleri tasarlanmış; bu modeller hem kendi içlerinde hem de birbirleriyle karşılaştırılmıştır. İlk aşamada, üç katmanlı ve dört katmanlı modeller oluşturulmuştur. Üç katmanlı modeller, birer tane girdi katmanı, gizli katman ve çıktı katmanından oluşurken bu modellerin gizli katman sayısı ikiye çıkarılarak dört katmanlı modeller elde edilmiştir.

Tüm modellerin girdi ve çıktı katmanlarının nöron sayıları veri setlerine göre değişmektedir. MNIST veri setinin sınıflandırıldığı modellerin girdi katmanında 784 nöron ve çıktı katmanında 10 nöron; Dry Bean veri setinin sınıflandırıldığı modellerin girdi katmanında 16 nöron ve çıktı katmanında 7 nöron kullanılmıştır. Tüm modellerin gizli katmanlarındaki nöron sayıları 100 olarak belirlenmiştir. Eğitim tur sayısı (iterasyon sayısı) ise her bir model için 50 seçilmiştir. KarcıFANN yönteminde alfa'ya ve SGD yönteminde öğrenme katsayısına rastgele bir şekilde belirlenen 0.005, 0.05, 0.55, 0.73, 0.85, 1.1 ve 1.25 değerleri kullanılmıştır. Doğruluk, MSE, duyarlılık, F1 skoru ve kesinlik metrikleri kullanılarak modellerin performansları değerlendirilmiş; elde edilen sonuçlar ve eğitim süreleri tablolarla gösterilmiştir. Bu tablolarda, KarcıFANN'ın SGD'ye göre daha yüksek performans gösterdiği değerler, koyu tonla vurgulanmıştır. Ayrıca, tüm modellerde alfa ve öğrenme katsayısının (alfa/öğrenme katsayısı) 0.55 değeri için elde edilen grafiksel sonuçlar sunulmuştur.

MNIST veri setinin sınıflandırılması amacıyla tasarlanan ilk KarcıFANN ve SGD modellerinde, Sigmoid ve ReLU aktivasyon fonksiyonları ile MSE hata fonksiyonu kullanılmıştır. Bu modellerin eğitilmesiyle elde edilen doğruluk, MSE, kesinlik, F1 skoru ve duyarlılık değerleri ile eğitim süreleri Tablo 1; diğer parametreleri sabit tutularak MSE yerine CE fonksiyonunun kullanıldığı modellerin başarımları da Tablo 2 ile raporlanmaktadır.

MSE fonksiyonu ile tasarlanan KarcıFANN modelinde, Sigmoid fonksiyonunun kullanımının ReLU fonksiyonuna kıyasla daha başarılı olduğu; MSE fonksiyonu ile tasarlanan SGD yönteminde de Sigmoid fonksiyonunun 0.005 ve 0.05 dışındaki öğrenme katsayısı değerlerinde ReLU fonksiyonuna göre daha düşük MSE ve daha yüksek doğruluk, kesinlik, duyarlılık ve F1 skoru ürettiği, Tablo 1 ile gösterilmektedir. Ayrıca, Sigmoid fonksiyonunu içeren KarcıFANN ve SGD modelleri karşılaştırıldığında, 1.1 ve 1.25 dışındaki alfa/öğrenme katsayısı değerleri için KarcıFANN'ın genel performansının SGD'ye kıyasla daha yüksek olduğu anlaşılmaktadır. ReLU fonksiyonu ile tasarlanan modellerde ise 0.55, 0.73 ve 0.85 alfa/öğrenme katsayısı değerleri için KarcıFANN; 0.005 ve 0.05 değerleri için SGD daha üstün performans sergilemektedir.

CE hata fonksiyonu ile tasarlanan KarcıFANN ve SGD yöntemlerinde, Sigmoid fonksiyonunun kullanımının ReLU fonksiyonuna göre daha başarılı sonuçlar ürettiği, Tablo 2 ile gösterilmiştir. Öte yandan, Sigmoid fonksiyonunun kullanıldığı KarcıFANN ile SGD modellerinin genel performanslarının yakın olduğu görülmektedir. ReLU fonksiyonunun kullanıldığı KarcıFANN modeli, 0.55 alfa değerinde SGD'den daha başarılı olmakta, ancak 0.73, 0.85, 1.1 ve 1.25 alfa değerleri için öğrenememektedir. SGD modeli ise 0.005 ve 0.05 öğrenme katsayıları için KarcıFANN'a kıyasla daha üstün performans göstermekte, diğer değerler için ise öğrenme gerçekleştirememektedir.

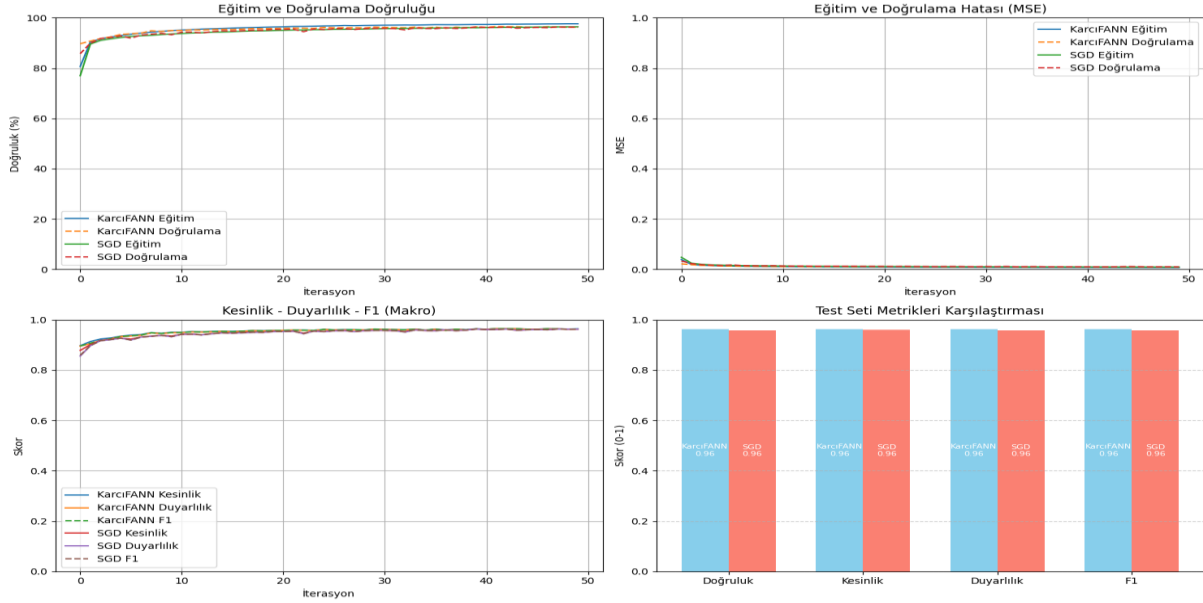
Tablo 1. MNIST veri setinin sınıflandırılması amacıyla 100 nöronlu bir gizli katman ve MSE fonksiyonuyla tasarlanan KarcıFANN ve SGD modellerinin kullanılması sonucu elde edilen doğruluk, MSE, kesinlik, duyarlılık ve F1 skoru değerleri.

	Aktivasyon Fonksiyonu	Alfa/ Öğrenme Katsayısı	Eğitim		Doğrulama					Test					
			Doğruluk	MSE	Doğruluk	MSE	Kesinlik	Duyarlılık	F1	Doğruluk	MSE	Kesinlik	Duyarlılık	F1	Süre
KarcıFANN	Sigmoid	0.005	96.88	0.0103	94.76	0.0163	0.948	0.947	0.947	95.04	0.0153	0.869	0.864	0.865	707.44
		0.05	97.07	0.0095	94.88	0.0157	0.952	0.949	0.949	94.61	0.0157	0.950	0.946	0.946	631.58
		0.55	97.61	0.0069	96.32	0.0096	0.964	0.963	0.963	96.32	0.0096	0.963	0.963	0.963	657.65
		0.73	97.74	0.0067	96.84	0.0082	0.969	0.968	0.968	96.73	0.0083	0.967	0.967	0.967	656.94
		0.85	96.50	0.0084	95.42	0.0100	0.954	0.953	0.953	95.65	0.0098	0.957	0.956	0.956	618.79
		1.1	92.57	0.0155	91.08	0.0185	0.911	0.911	0.910	91.91	0.0166	0.918	0.918	0.918	636.92
		1.25	90.97	0.0178	90.34	0.0198	0.905	0.902	0.903	90.18	0.0199	0.904	0.901	0.901	678.97
	ReLU	0.005	80.31	0.0355	74.90	0.0412	0.756	0.740	0.717	76.09	0.0399	0.766	0.754	0.732	683.27
		0.05	88.80	0.0304	81.34	0.0455	0.851	0.812	0.805	82.08	0.0450	0.855	0.818	0.810	634.63
		0.55	96.20	0.0115	95.56	0.0122	0.955	0.955	0.955	95.57	0.0122	0.956	0.955	0.955	617.48
		0.73	93.50	0.0153	92.60	0.0190	0.927	0.926	0.926	92.68	0.0189	0.929	0.926	0.927	622.07
		0.85	89.54	0.0237	87.00	0.0283	0.894	0.868	0.874	87.93	0.0274	0.901	0.878	0.883	615.46
		1.1	19.68	0.0824	19.16	0.0839	0.131	0.195	0.121	18.38	0.0833	0.137	0.200	0.124	615.73
		1.25	9.83	-	10.32	-	0.010	0.100	0.019	9.80	-	0.010	0.100	0.018	618.32
SGD	Sigmoid	0.005	85.57	0.0398	84.94	0.0403	0.849	0.845	0.845	86.68	0.0390	0.869	0.864	0.865	612.87
		0.05	93.36	0.0158	93.48	0.0161	0.935	0.935	0.934	93.61	0.0155	0.935	0.936	0.935	665.03
		0.55	96.42	0.0091	96.28	0.0095	0.963	0.963	0.963	95.93	0.0097	0.959	0.959	0.959	627.31
		0.73	95.71	0.0096	94.76	0.0113	0.950	0.947	0.948	94.98	0.0110	0.950	0.949	0.950	631.70
		0.85	94.91	0.0105	93.54	0.0120	0.936	0.935	0.935	94.13	0.0115	0.942	0.941	0.941	598.10
		1.1	93.87	0.0122	92.18	0.0148	0.924	0.921	0.922	92.90	0.0136	0.931	0.928	0.929	621.72
		1.25	93.97	0.0115	93.44	0.0130	0.935	0.934	0.934	93.74	0.0125	0.938	0.937	0.937	604.07
	ReLU	0.005	93.21	0.0188	92.84	0.0193	0.928	0.927	0.928	93.24	0.0185	0.932	0.932	0.932	667.86
		0.05	97.11	0.0099	96.32	0.0110	0.963	0.963	0.963	96.48	0.0106	0.965	0.964	0.965	598.54
		0.55	94.75	0.0125	93.96	0.0139	0.941	0.939	0.940	94.16	0.0135	0.943	0.941	0.941	599.05
		0.73	91.80	0.0199	91.98	0.0222	0.922	0.920	0.920	92.66	0.0215	0.928	0.926	0.927	597.90
		0.85	55.06	0.0516	55.28	0.0528	0.432	0.549	0.465	54.92	0.0522	0.433	0.551	0.467	597.56
		1.1	19.91	0.0814	17.74	0.0835	0.107	0.192	0.113	18.38	0.0831	0.105	0.196	0.114	625.35
		1.25	18.97	0.0826	18.46	0.0832	0.097	0.187	0.107	19.10	0.0829	0.101	0.191	0.110	604.65

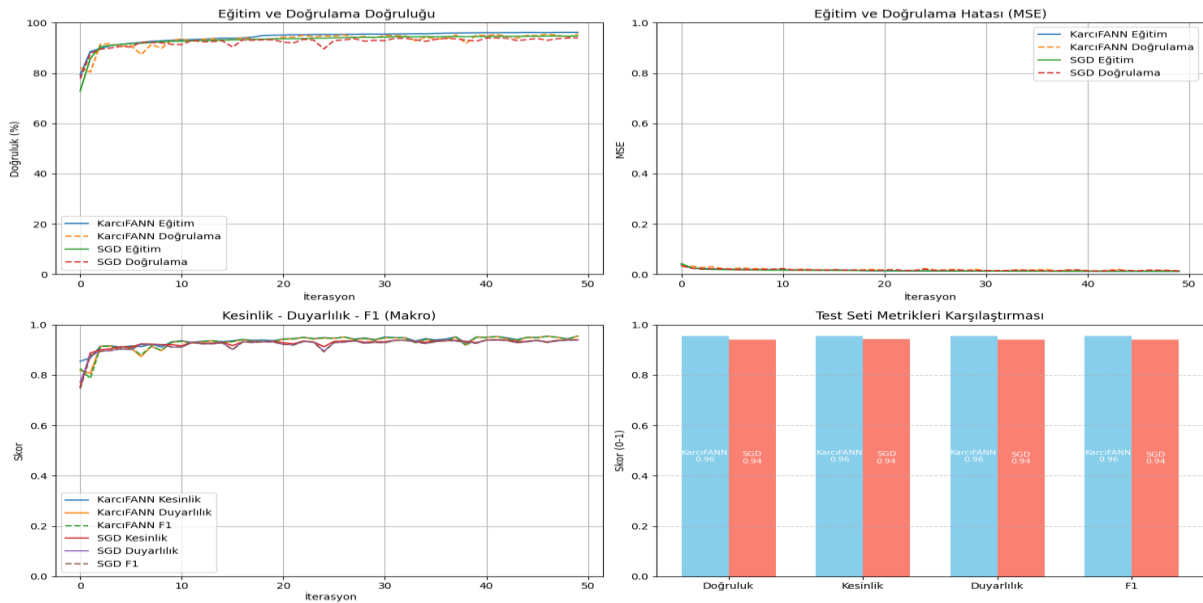
Tablo 2. MNIST veri setinin sınıflandırılması amacıyla 100 nöronlu bir gizli katman ve CrossEntropy (CE) fonksiyonuyla tasarlanan KarcıFANN ve SGD modellerinin kullanılması sonucu elde edilen doğruluk, CE, kesinlik, duyarlılık ve F1 skoru değerleri.

	Aktivasyon Fonksiyonu	Alfa/ Öğrenme Katsayısı	Eğitim		Doğrulama					Test					
			Doğruluk	CE	Doğruluk	CE	Kesinlik	Duyarlılık	F1	Doğruluk	CE	Kesinlik	Duyarlılık	F1	Süre
KarcıFANN	Sigmoid	0.005	96.56	0.1286	96.04	0.1733	0.961	0.960	0.960	96.37	0.1494	0.963	0.963	0.963	628.11
		0.05	96.86	0.1152	95.72	0.1793	0.957	0.956	0.956	95.97	0.0610	0.970	0.970	0.970	619.62
		0.55	99.56	0.0156	97.20	0.1827	0.972	0.972	0.972	96.93	0.1811	0.969	0.969	0.969	621.51
		0.73	100.00	0.0002	97.62	0.1230	0.976	0.976	0.976	97.68	0.1200	0.977	0.976	0.977	648.50
		0.85	100.00	0.0005	97.96	0.0959	0.979	0.979	0.979	97.84	0.1000	0.978	0.978	0.978	621.92
		1.1	98.10	0.0597	95.06	0.1857	0.951	0.950	0.951	95.65	0.1604	0.957	0.956	0.956	621.51
		1.25	92.32	0.2652	91.14	0.3292	0.912	0.912	0.911	91.40	0.3116	0.914	0.913	0.913	642.35
	ReLU	0.005	95.51	0.1629	94.92	0.1920	0.949	0.948	0.948	95.50	0.1623	0.955	0.954	0.955	627.72
		0.05	95.89	0.1515	95.68	0.1795	0.957	0.957	0.956	95.92	0.1587	0.959	0.959	0.959	627.17
		0.55	96.47	0.2549	94.62	0.6927	0.946	0.946	0.946	95.25	0.5290	0.952	0.952	0.952	627.36
		0.73	10.91	2.3023	9.98	2.3017	0.010	0.100	0.018	9.80	2.3032	0.010	0.100	0.021	616.60
		0.85	10.87	2.3029	11.20	2.3023	0.011	0.100	0.020	11.35	2.3032	0.011	0.100	0.020	613.91
		1.1	10.45	2.3048	9.56	2.3058	0.010	0.100	0.017	10.09	2.3053	0.010	0.100	0.018	632.48
		1.25	10.47	2.3050	11.48	2.3058	0.011	0.100	0.021	11.35	2.3072	0.011	0.100	0.019	632.35
SGD	Sigmoid	0.005	92.85	0.2500	92.64	0.2602	0.925	0.926	0.925	93.03	0.2439	0.929	0.929	0.929	601.95
		0.05	98.45	0.0610	97.00	0.0976	0.970	0.970	0.970	97.27	0.0885	0.973	0.972	0.972	601.69
		0.55	100.00	0.0027	97.50	0.0922	0.975	0.975	0.975	97.90	0.0767	0.979	0.979	0.979	602.46
		0.73	100.00	0.0019	97.70	0.0932	0.977	0.977	0.977	97.91	0.0473	0.979	0.979	0.979	602.76
		0.85	100.00	0.0023	97.56	0.0958	0.975	0.975	0.975	97.47	0.0933	0.975	0.975	0.975	604.58
		1.1	98.15	0.0580	95.32	0.1961	0.954	0.953	0.953	95.88	0.1577	0.959	0.958	0.958	605.20
		1.25	98.44	0.0486	96.50	0.1454	0.964	0.965	0.965	95.99	0.1400	0.959	0.960	0.959	613.10
	ReLU	0.005	96.91	0.1123	95.90	0.1362	0.959	0.958	0.958	96.32	0.1249	0.963	0.963	0.963	609.39
		0.05	99.92	0.0070	97.90	0.0774	0.979	0.979	0.979	97.87	0.0782	0.978	0.978	0.978	608.83
		0.55	10.55	2.3034	11.84	2.3026	0.012	0.100	0.021	11.35	2.3032	0.011	0.100	0.020	608.65
		0.73	10.48	2.3040	11.54	2.3032	0.012	0.100	0.021	11.35	2.3040	0.011	0.100	0.020	605.73
		0.85	10.64	2.3047	11.20	2.3013	0.011	0.100	0.020	11.35	2.3027	0.011	0.100	0.020	621.55
		1.1	10.36	2.3053	10.16	2.3070	0.010	0.100	0.018	10.10	2.3062	0.010	0.100	0.018	614.25
		1.25	10.37	2.3062	9.24	2.3109	0.009	0.100	0.017	10.32	2.3065	0.010	0.100	0.019	616.51

MNIST veri setinin sınıflandırılması amacıyla, MSE ve Sigmoid fonksiyonları ile tasarlanan üç katmanlı KarcıFANN ve SGD modellerinde, alfa/öğrenme katsayısının 0.55 değeri için elde edilen grafikler Şekil 2; MSE ve ReLU fonksiyonunun kullanımıyla elde edilen sonuçlar ise Şekil 3 ile sunulmaktadır. Burada ve devam eden şekillerde, eğitim ve doğrulama aşamalarının doğruluk ile hata grafikleri, doğrulama aşamasının duyarlılık, kesinlik ile F1 skoru grafiği ve test aşamasının duyarlılık, kesinlik ile F1 skoru bar grafiği gösterilmektedir.

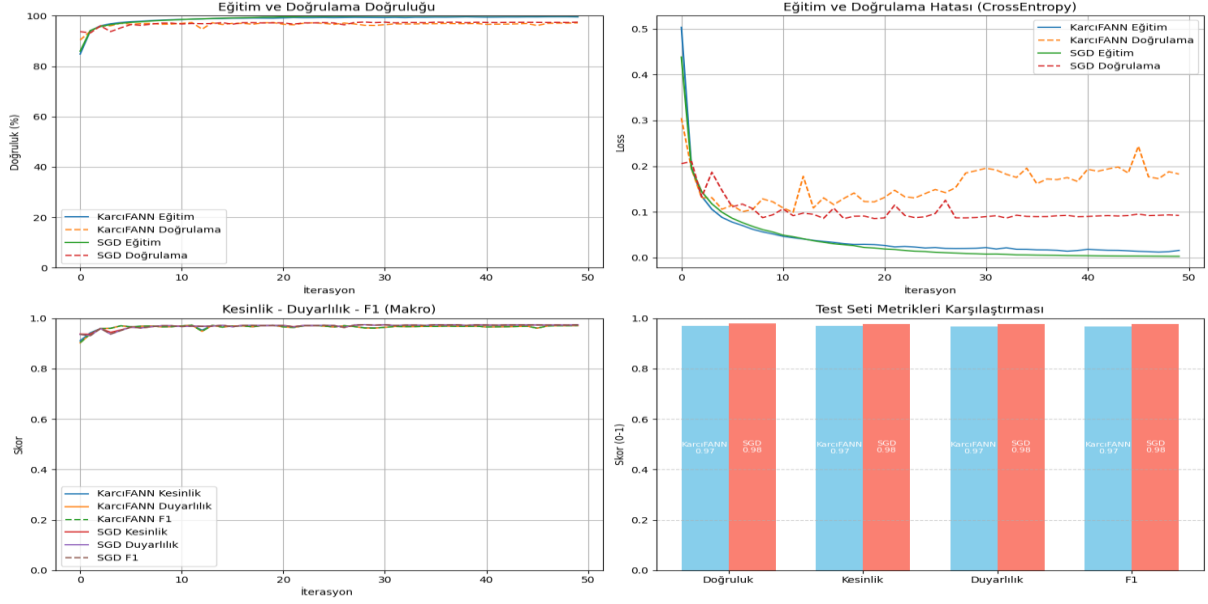


Şekil 2. MNIST veri setinin sınıflandırılması amacıyla, Sigmoid aktivasyon fonksiyonu ve MSE hata fonksiyonunun kullanıldığı; 100 nöronlu bir gizli katmana sahip KarcıFANN ve SGD modellerinde, alfa/öğrenme katsayısı = 0.55 için elde edilen grafikler.

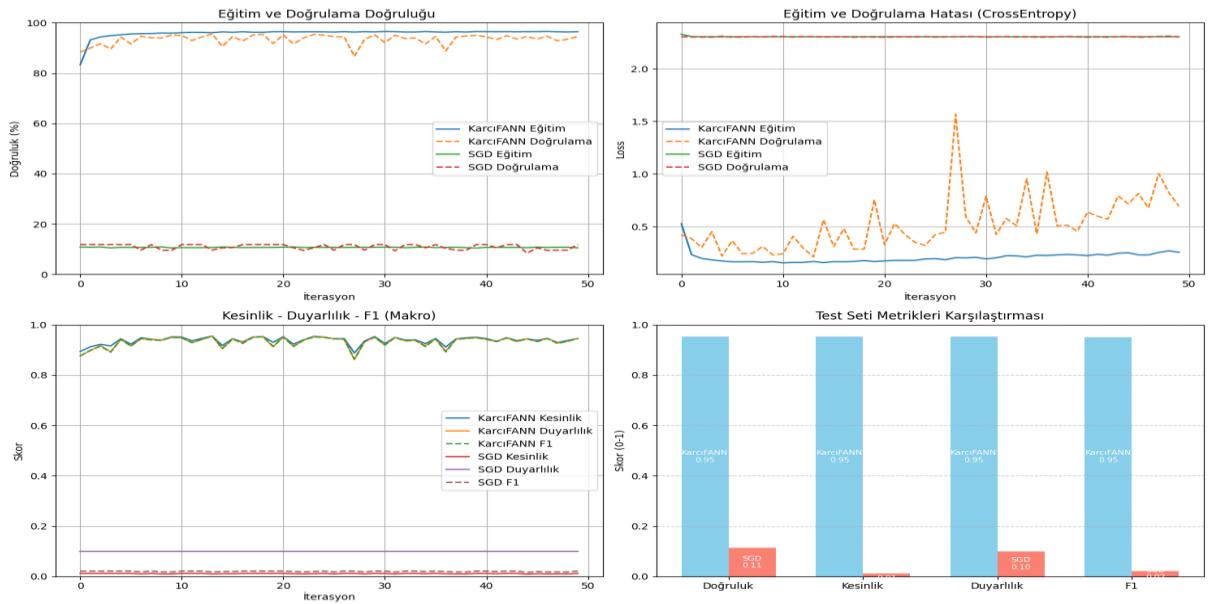


Şekil 3. MNIST veri setinin sınıflandırılması amacıyla, ReLU aktivasyon fonksiyonu ve MSE hata fonksiyonunun kullanıldığı; 100 nöronlu bir gizli katmana sahip KarcıFANN ve SGD modellerinde, alfa/öğrenme katsayısı = 0.55 için elde edilen grafikler.

CE hata fonksiyonu ve Sigmoid aktivasyon fonksiyonu ile tasarlanan modellerin alfa/öğrenme katsayısının 0.55 değeri için ürettiği değerlerin grafiksel gösterimi Şekil 4; Sigmoid yerine ReLU fonksiyonunun kullanımıyla elde edilen grafiksel sonuçlar ise Şekil 5 ile sunulmaktadır.



Şekil 4. MNIST veri setinin sınıflandırılması amacıyla, Sigmoid aktivasyon fonksiyonu ve CE hata fonksiyonunun kullanıldığı; 100 nöronlu bir gizli katmana sahip KarciFANN ve SGD modellerinde, alfa/öğrenme katsayısı = 0.55 için elde edilen grafikler.



Şekil 5. MNIST veri setinin sınıflandırılması amacıyla, ReLU aktivasyon fonksiyonu ve CE hata fonksiyonunun kullanıldığı; 100 nöronlu bir gizli katmana sahip KarciFANN ve SGD modellerinde, alfa/öğrenme katsayısı = 0.55 için elde edilen grafikler.

Katman sayısını artırmanın performansa etkisini analiz etmek amacıyla üç katmanlı KarcıFANN ve SGD modellerine birer tane gizli katman (her gizli katmanda 100 nöron) eklenerek dört katmanlı modeller tasarlanmıştır. Dört katmanlı modellerin eğitilmesi sonucunda elde edilen performans değerleri ve eğitim süreleri, Tablo 3'te yer almaktadır. Diğer parametreleri sabit tutulmak şartıyla, MSE yerine CE hata fonksiyonunun kullanıldığı modellerin performansları ise Tablo 4'te detaylandırılmıştır.

KarcıFANN ve Sigmoid aktivasyon fonksiyonu ile tasarlanan modelde, 0.005 ve 0.05 alfa değerleri için tüm aşamalarda ReLU fonksiyonundan daha başarılı sonuçlar elde edildiği Tablo 3'te görülmektedir. Öte yandan, alfa'nın 0.55, 0.73, 0.85 ve 1.1 değerlerinde ReLU fonksiyonunun genel performansının daha yüksek olduğu gözlemlenmiştir. SGD yönteminin kullanıldığı modellerde, öğrenme katsayısının 1.25 değerinde Sigmoid ve ReLU fonksiyonları ile öğrenme gerçekleştirilememiştir. Diğer öğrenme katsayılarında ise ReLU fonksiyonunun genel performansı, Sigmoid fonksiyonundan daha yüksektir.

Sigmoid ve MSE fonksiyonları ile tasarlanan KarcıFANN ve SGD modelleri karşılaştırıldığında, alfa/öğrenme katsayısının 0.005, 0.55, 0.73 ve 0.85 değerleri için KarcıFANN'ın SGD'ye göre tüm veri setlerinde daha düşük MSE ve daha yüksek doğruluk, kesinlik, duyarlılık ve F1 skoru ürettiği anlaşılmaktadır. ReLU fonksiyonu ile tasarlanan modellerde de 0.55, 0.73 ve 0.85 alfa/öğrenme katsayısı değerleri için KarcıFANN yöntemi daha üstün performans sergilemektedir.

CE fonksiyonunun kullanıldığı ve Tablo 4 ile sonuçları verilen iki gizli katmanlı KarcıFANN modellerinde, Sigmoid fonksiyonunun ReLU fonksiyonuna göre daha başarılı olduğu gözlemlenmektedir. SGD modellerinde ise 0.005 ve 0.05 öğrenme katsayıları dışında Sigmoid fonksiyonu daha yüksek performans göstermekte ve ReLU fonksiyonu 0.55, 0.73, 0.85, 1.1 ve 1.25 değerleri için öğrenme gerçekleştirememektedir.

Sigmoid ve CE fonksiyonlarının kullanıldığı KarcıFANN ve SGD modelleri karşılaştırıldığında, alfa/öğrenme katsayısının 0.005 değeri için KarcıFANN modelinin eğitim ve doğrulama aşamalarında daha başarılı olduğu; 0.73, 0.85 ve 1.1 değerleri için de her iki modelin yakın değerler ürettiği görülmektedir. ReLU fonksiyonu ile tasarlanan KarcıFANN modeli, alfa/öğrenme katsayısının 0.55 değeri için SGD modelinden daha iyi sonuçlar üretmekte ve 0.73, 0.85, 1.1 ve 1.25 değerlerinde ise öğrenme gerçekleştirememektedir. Buna karşılık SGD modeli, 0.005 ve 0.05 öğrenme katsayılarında KarcıFANN'dan daha başarılı olurken diğer öğrenme katsayılarında öğrenememektedir.

MSE fonksiyonunun kullanıldığı üç katmanlı modellerin performanslarını gösteren Tablo 1 ile dört katmanlı modellerin performanslarını gösteren Tablo 3 karşılaştırıldığında, 0.55, 0.73 ve 0.85 alfa/öğrenme katsayıları için dört katmanlı modellerin performansı artırdığı görülmektedir. Diğer taraftan, MSE yerine CE fonksiyonu ile tasarlanan modellerin sonuçlarını gösteren Tablo 2 ile dört katmanlı modellerin sonuçlarını gösteren Tablo 4 karşılaştırıldığında, 0.73 ve 0.85 dışındaki alfa değerleri için Sigmoid fonksiyonunun dört katmanlı modelde daha iyi performans sergilediği, ReLU fonksiyonunun kullanıldığı üç ve dört katmanlı modellerin de yakın sonuçlar ürettiği görülmektedir.

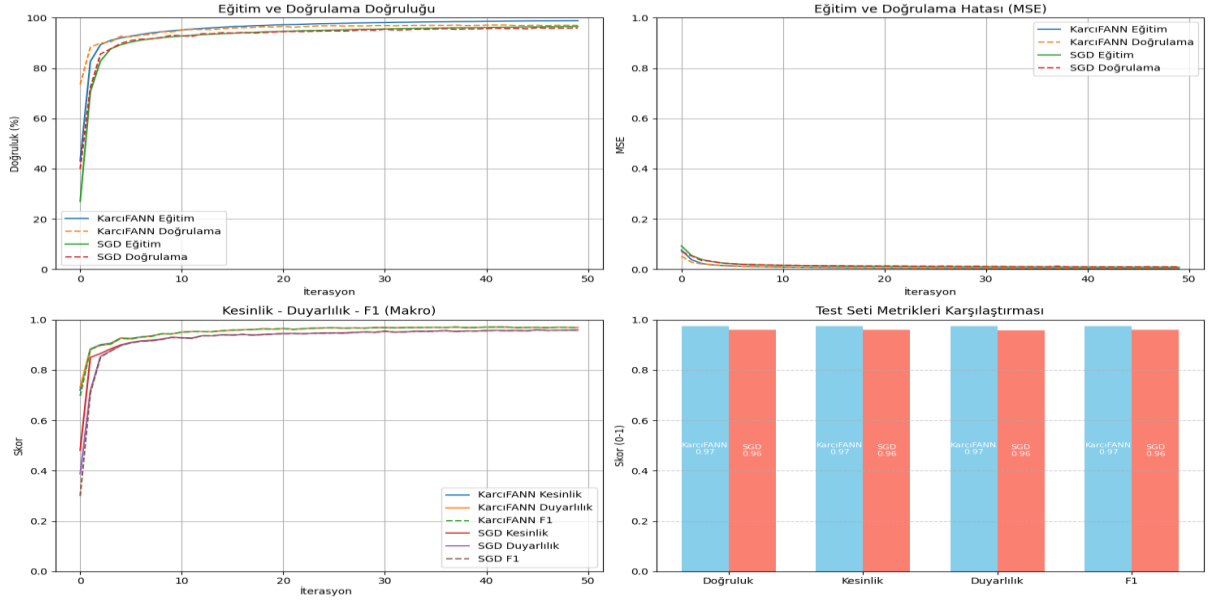
Tablo 3. MNIST veri setinin sınıflandırılması amacıyla 100'er nöronlu iki gizli katman ve MSE fonksiyonuyla tasarlanan KarcıFANN ve SGD modellerinin kullanılması sonucu elde edilen doğruluk, MSE, kesinlik, duyarlılık ve F1 skoru değerleri.

	Aktivasyon Fonksiyonu	Alfa/ Öğrenme Katsayısı	Eğitim		Doğrulama					Test					
			Doğruluk	MSE	Doğruluk	MSE	Kesinlik	Duyarlılık	F1	Doğruluk	MSE	Kesinlik	Duyarlılık	F1	Süre
KarcıFANN	Sigmoid	0.005	94.76	0.0338	96.36	0.0139	0.964	0.963	0.963	96.09	0.0138	0.961	0.960	0.961	691.45
		0.05	94.79	0.0587	86.24	0.0857	0.792	0.870	0.825	87.08	0.0843	0.801	0.873	0.832	654.21
		0.55	98.84	0.0023	96.98	0.0050	0.970	0.970	0.970	97.45	0.0044	0.974	0.974	0.974	667.11
		0.73	98.66	0.0030	97.14	0.0050	0.971	0.971	0.971	97.43	0.0046	0.974	0.974	0.974	665.52
		0.85	98.20	0.0040	97.02	0.0056	0.970	0.970	0.970	97.25	0.0051	0.972	0.972	0.972	667.31
		1.1	10.42	0.0901	9.90	0.0904	0.010	0.100	0.018	9.74	0.0904	0.010	0.100	0.018	671.70
		1.25	10.37	0.0901	10.74	0.0906	0.011	0.100	0.019	9.82	0.0905	0.010	0.100	0.018	659.19
	ReLU	0.005	88.10	0.0291	87.18	0.0200	0.831	0.866	0.840	86.93	0.0203	0.830	0.867	0.841	836.57
		0.05	93.90	0.0580	85.58	0.3573	0.776	0.870	0.818	86.04	0.3561	0.779	0.874	0.822	655.00
		0.55	99.46	0.0022	97.24	0.0056	0.972	0.972	0.972	97.09	0.0058	0.971	0.970	0.971	640.68
		0.73	99.60	0.0015	97.42	0.0060	0.975	0.974	0.974	97.83	0.0057	0.978	0.978	0.978	644.76
		0.85	99.41	0.0018	97.38	0.0050	0.974	0.973	0.973	97.55	0.0049	0.976	0.975	0.975	638.00
		1.1	98.43	0.0036	96.08	0.0088	0.962	0.0961	0.961	96.29	0.0084	0.963	0.963	0.963	640.63
		1.25	9.92	-	9.32	-	0.009	0.100	0.017	9.80	-	0.010	0.100	0.018	628.92
SGD	Sigmoid	0.005	68.68	0.0609	70.04	0.0602	0.766	0.689	0.668	70.36	0.0603	0.774	0.695	0.671	639.15
		0.05	90.87	0.0210	91.38	0.0204	0.913	0.913	0.913	91.54	0.0202	0.915	0.914	0.914	631.73
		0.55	96.53	0.0087	95.92	0.0101	0.959	0.959	0.959	95.97	0.0099	0.960	0.959	0.959	638.71
		0.73	94.69	0.0101	93.50	0.0114	0.935	0.935	0.934	93.91	0.0107	0.939	0.938	0.938	635.68
		0.85	95.95	0.0085	94.92	0.0101	0.949	0.949	0.949	95.48	0.0093	0.955	0.954	0.954	639.01
		1.1	40.81	0.0610	38.96	0.0613	0.232	0.388	0.268	38.99	0.0616	0.235	0.388	0.270	640.43
		1.25	10.45	0.0902	9.88	0.0902	0.010	0.100	0.018	9.58	0.0903	0.010	0.100	0.017	638.29
	ReLU	0.005	92.19	0.0191	92.28	0.0191	0.923	0.923	0.922	92.63	0.0184	0.926	0.926	0.925	607.24
		0.05	97.62	0.0068	96.80	0.0078	0.968	0.968	0.968	96.98	0.0076	0.970	0.970	0.970	658.00
		0.55	99.44	0.0019	97.74	0.0046	0.977	0.977	0.977	97.79	0.0044	0.978	0.978	0.978	618.19
		0.73	99.45	0.0019	97.00	0.0055	0.971	0.970	0.970	97.00	0.0057	0.970	0.970	0.970	627.78
		0.85	99.33	0.0020	97.24	0.0056	0.972	0.972	0.972	97.29	0.0053	0.973	0.973	0.973	613.82
		1.1	98.94	0.0026	96.70	0.0062	0.967	0.967	0.967	97.08	0.0053	0.971	0.971	0.971	615.66
		1.25	9.92	-	9.32	-	0.009	0.100	0.017	9.80	-	0.010	0.100	0.018	600.86

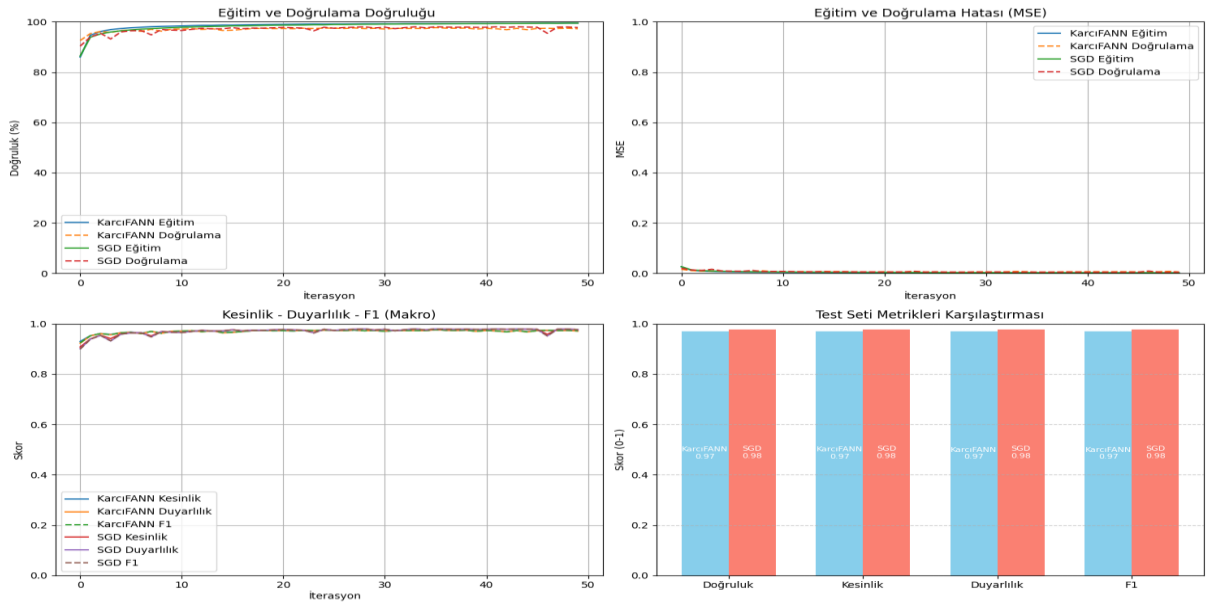
Tablo 4. MNIST veri setinin sınıflandırılması amacıyla 100'er nöronlu iki gizli katman ve CE fonksiyonuyla tasarlanan KarcıFANN ve SGD modellerinin kullanılması sonucu elde edilen doğruluk, CE, kesinlik, duyarlılık ve F1 skoru değerleri.

	Aktivasyon Fonksiyonu	Alfa/ Öğrenme Katsayısı	Eğitim		Doğrulama					Test					
			Doğruluk	CE	Doğruluk	CE	Kesinlik	Duyarlılık	F1	Doğruluk	CE	Kesinlik	Duyarlılık	F1	Süre
KarcıFANN	Sigmoid	0.005	96.61	0.1186	95.98	0.1627	0.960	0.959	0.959	95.93	0.1484	0.960	0.959	0.959	656.44
		0.05	96.91	0.1130	96.32	0.1355	0.963	0.962	0.963	96.75	0.1278	0.967	0.967	0.967	631.93
		0.55	99.64	0.0129	97.22	0.1851	0.972	0.972	0.972	94.34	0.1845	0.973	0.973	0.973	632.89
		0.73	100.00	0.0001	97.68	0.1315	0.976	0.977	0.977	97.75	0.1388	0.977	0.977	0.977	632.89
		0.85	100.00	0.0003	97.86	0.1046	0.978	0.979	0.978	97.92	0.1123	0.979	0.979	0.979	663.23
		1.1	99.79	0.0083	97.18	0.1165	0.972	0.971	0.972	97.13	0.1194	0.971	0.971	0.971	633.35
		1.25	98.43	0.0499	96.16	0.1515	0.961	0.961	0.961	96.06	0.1400	0.960	0.960	0.960	638.35
	ReLU	0.005	95.38	0.1700	94.56	0.2083	0.946	0.945	0.945	94.93	0.1920	0.950	0.948	0.949	633.80
		0.05	95.60	0.1700	93.78	0.2482	0.940	0.937	0.937	94.95	0.1924	0.951	0.949	0.949	631.82
		0.55	94.34	0.2917	92.18	0.5439	0.925	0.921	0.921	92.77	0.4263	0.932	0.927	0.927	638.43
		0.73	20.29	1.8106	19.44	1.8322	0.041	0.189	0.067	19.26	1.8347	0.040	0.191	0.066	632.86
		0.85	10.89	2.3030	10.88	2.3025	0.011	0.100	0.020	11.35	2.3019	0.011	0.100	0.020	638.99
		1.1	10.85	2.3046	11.36	2.3024	0.011	0.100	0.020	11.35	2.3026	0.011	0.100	0.020	652.31
		1.25	10.93	2.3034	10.22	2.3053	0.010	0.100	0.019	11.35	2.3037	0.011	0.100	0.020	647.73
SGD	Sigmoid	0.005	90.97	0.3227	90.38	0.3280	0.902	0.902	0.902	95.93	0.1484	0.960	0.959	0.959	602.75
		0.05	98.71	0.0490	97.60	0.0837	0.976	0.976	0.976	97.41	0.0834	0.974	0.974	0.974	605.33
		0.55	100.00	0.0008	97.60	0.1014	0.976	0.976	0.976	97.88	0.0946	0.979	0.979	0.979	607.20
		0.73	100.00	0.0008	97.74	0.0935	0.977	0.977	0.977	97.85	0.0974	0.978	0.978	0.978	607.05
		0.85	100.00	0.0007	97.58	0.0977	0.976	0.976	0.976	97.95	0.0924	0.979	0.979	0.979	607.77
		1.1	99.19	0.0251	97.02	0.1128	0.970	0.970	0.970	97.13	0.1129	0.971	0.971	0.971	606.69
		1.25	99.18	0.0257	96.88	0.1216	0.969	0.968	0.969	96.89	0.1183	0.969	0.969	0.969	605.75
	ReLU	0.005	97.68	0.0816	96.50	0.1085	0.965	0.964	0.965	96.81	0.1072	0.968	0.968	0.968	622.16
		0.05	100.00	0.0011	97.92	0.1165	0.979	0.979	0.979	97.94	0.1924	0.979	0.979	0.979	604.90
		0.55	10.72	2.3031	9.58	2.3056	0.010	0.100	0.017	9.82	2.3045	0.010	0.100	0.018	611.25
		0.73	10.62	2.3041	10.72	2.3028	0.011	0.100	0.019	11.35	2.3039	0.011	0.100	0.020	606.59
		0.85	10.85	2.3039	10.88	2.3054	0.011	0.100	0.020	11.35	2.3046	0.011	0.100	0.020	629.02
		1.1	10.75	2.3054	10.60	2.3025	0.011	0.100	0.019	10.28	2.3039	0.010	0.100	0.019	604.65
		1.25	10.62	2.3058	9.54	2.3119	0.010	0.100	0.017	9.58	2.3091	0.010	0.100	0.017	615.68

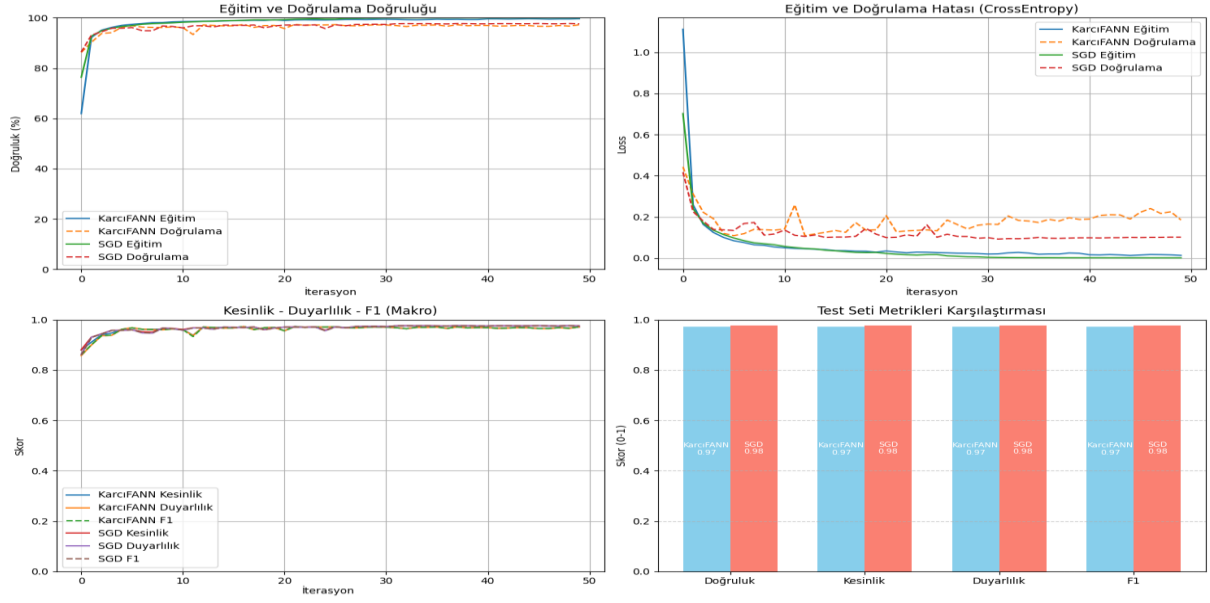
MSE ve Sigmoid fonksiyonları ile tasarlanan dört katmanlı (gizli katmanlarda 100'er nöron) KarcıFANN ve SGD modellerinin eğitimi sonunda, alfa/öğrenme katsayısının 0.55 değeri için elde edilen grafikler Şekil 6'da, Sigmoid yerine ReLU fonksiyonu ile tasarlanan modellerle elde edilen grafikler Şekil 7'de gösterilmektedir. MSE yerine CE fonksiyonunun kullanıldığı modellerde elde edilen grafikler, sırasıyla Şekil 8 ve Şekil 9 ile sunulmaktadır.



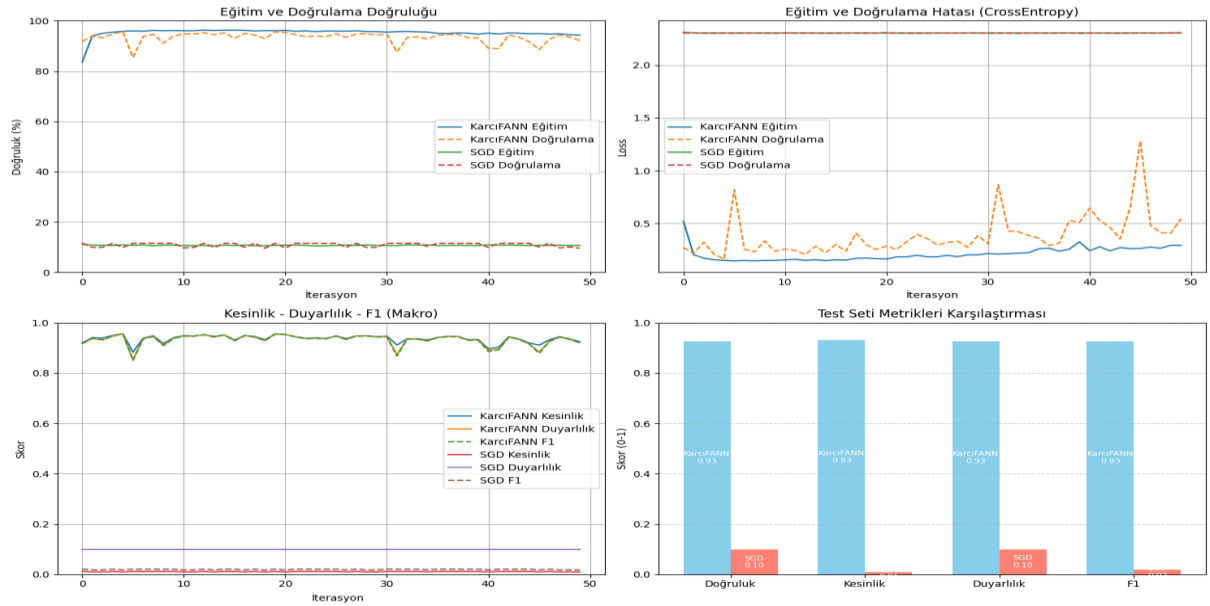
Şekil 6. MNIST veri setinin sınıflandırılması amacıyla, Sigmoid aktivasyon fonksiyonu ve MSE hata fonksiyonunun kullanıldığı; 100'er nöronlu iki gizli katmana sahip KarcıFANN ve SGD modellerinde, alfa/öğrenme katsayısı = 0.55 için elde edilen grafikler.



Şekil 7. MNIST veri setinin sınıflandırılması amacıyla, ReLU aktivasyon fonksiyonu ve MSE hata fonksiyonunun kullanıldığı; 100'er nöronlu iki gizli katmana sahip KarcıFANN ve SGD modellerinde, alfa/öğrenme katsayısı = 0.55 için elde edilen grafikler.



Şekil 8. MNIST veri setinin sınıflandırılması amacıyla, Sigmoid aktivasyon fonksiyonu ve CE hata fonksiyonunun kullanıldığı; 100'er nöronlu iki gizli katmana sahip KarciFANN ve SGD modellerinde, alfa/öğrenme katsayısı = 0.55 için elde edilen grafikler.



Şekil 9. MNIST veri setinin sınıflandırılması amacıyla, ReLU aktivasyon fonksiyonu ve CE hata fonksiyonunun kullanıldığı; 100'er nöronlu iki gizli katmana sahip KarciFANN ve SGD modellerinde, alfa/öğrenme katsayısı = 0.55 için elde edilen grafikler.

Çalışmanın bu bölümünde, MNIST veri setinin sınıflandırılmasında kullanılan üç katmanlı modellerin girdi ve çıktı katmanlarındaki nöron sayıları değiştirilerek elde edilen yeni modellerle Dry Bean veri seti sınıflandırılmıştır. Sigmoid ve ReLU aktivasyon fonksiyonları ile MSE hata fonksiyonunun kullanıldığı KarciFANN ve SGD modellerinde elde edilen sonuçlar Tablo 5, modellerin diğer parametreleri sabit tutularak MSE yerine CE hata fonksiyonunun kullanıldığı modellerde elde edilen sonuçlar da Tablo 6 ile gösterilmektedir.

Tablo 5'te, 0.05, 0.55, 0.73 ve 0.85 alfa değerleri ile tüm öğrenme katsayıları için ReLU fonksiyonunun kullanıldığı KarcıFANN ve SGD modellerinin, Sigmoid fonksiyonuyla tasarlanan modellerden daha üstün performans sergilediği görülmektedir. Ayrıca, Sigmoid fonksiyonunun kullanıldığı KarcıFANN modeli, alfa/öğrenme katsayısının 1.1 ve 1.25 dışındaki tüm değerlerinde SGD modelinden daha başarılı sonuçlar üretmektedir. ReLU fonksiyonu ile tasarlanan SGD modeli ise 0.005, 0.05, 0.55 ve 1.1 alfa/öğrenme katsayıları için KarcıFANN modelinden daha düşük MSE ve daha yüksek doğruluk, kesinlik, duyarlılık ve F1 skoru üretmektedir.

CE fonksiyonu kullanılan KarcıFANN modelinde, 1.25 dışındaki tüm alfa değerlerinde Sigmoid fonksiyonunun kullanılmasının ReLU fonksiyona kıyasla daha iyi sonuçlar ürettiği, Tablo 6'da gösterilmektedir. Ayrıca, Sigmoid fonksiyonuyla tasarlanan KarcıFANN ve SGD modelleri karşılaştırıldığında, alfa/öğrenme katsayısının 0.005 ve 0.05 değerleri için KarcıFANN'ın tüm aşamalarda daha yüksek performans sergilediği; diğer değerlerde ise SGD modeline yakın sonuçlar ürettiği gözlemlenmektedir. Öte yandan, Sigmoid yerine ReLU fonksiyonu kullanılan KarcıFANN modeli, alfa/öğrenme katsayısının 1.25 dışındaki tüm değerleri için SGD modeline kıyasla daha başarılı sonuçlar üretmektedir.

Tablo 5 ile gösterilen alfa/öğrenme katsayısının 0.55 değeri için MSE ve Sigmoid fonksiyonlarının kullanıldığı üç katmanlı KarcıFANN ve SGD modellerinde elde edilen grafiksel sonuçlar Şekil 10'da, Sigmoid yerine ReLU fonksiyonunun kullanılmasıyla elde edilen grafikler de Şekil 11'de verilmiştir. Tablo 6 ile sonuçları gösterilen ve MSE yerine CE fonksiyonunun kullanıldığı modellerde ise alfa/öğrenme katsayısının 0.55 değeri için elde edilen grafiksel sonuçlar sırasıyla Şekil 12 ve Şekil 13 ile sunulmuştur.

Dry Bean veri setinin sınıflandırıldığı üç katmanlı modellere birer gizli katman eklenerek dört katmanlı modeller oluşturulmuştur. Tablo 5 ve Tablo 6 ile sonuçları sunulan üç katmanlı modellerin katman sayısının artırılmasıyla elde edilen sonuçlar sırasıyla Tablo 7'de ve Tablo 8'de yer almaktadır. MSE ve Sigmoid fonksiyonları ile tasarlanan dört katmanlı modellerin genel performansının ReLU fonksiyonunun kullanıldığı modellere kıyasla daha yüksek olduğu Tablo 7 ile gösterilmektedir. Bununla birlikte MSE yönteminin KarcıFANN ile kullanılması, birçok alfa/öğrenme katsayısı değeri için SGD ile kullanılmasından daha başarılı sonuçlar üretmektedir. Tablo 8 ile performans sonuçları gösterilen ve CE yönteminin kullanıldığı KarcıFANN ve SGD modellerinde, 1.25 dışındaki alfa/öğrenme katsayısının değerleri için ReLU fonksiyonunun kullanımı, Sigmoid fonksiyonuna kıyasla daha başarılı sonuçlar üretilmesini sağlamaktadır. Ayrıca, KarcıFANN'ın genel performansının SGD'den daha yüksek olduğu görülmektedir.

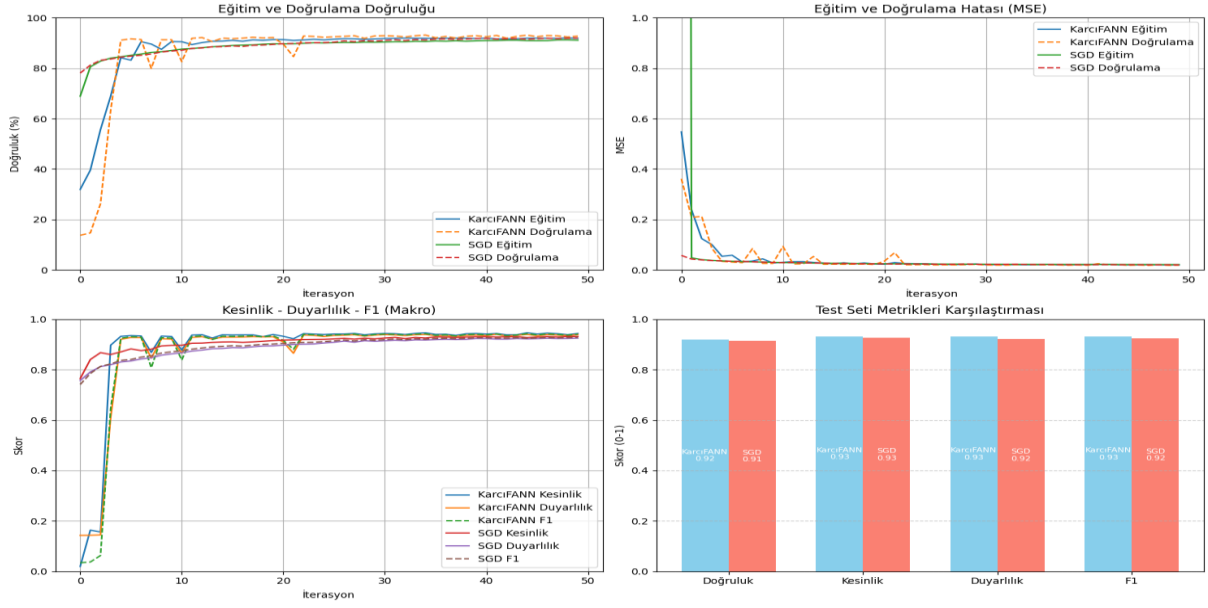
Tablo 5 ve Tablo 7 ile gösterilen üç katmanlı ve dört katmanlı modeller karşılaştırıldığında, MSE fonksiyonunu içeren KarcıFANN modellerinde katman sayısının artırılmasının tüm kategorilerde performansı artırdığı, MSE fonksiyonunu içeren SGD modellerinde ise ReLU fonksiyonu ile oluşturulan dört katmanlı modelin başarımının üç katmanlı modele kıyasla daha yüksek olduğu görülmektedir. Tablo 6 ve Tablo 8 karşılaştırıldığında ise CE ve ReLU fonksiyonlarını içeren dört katmanlı KarcıFANN modelinin 1.25 dışındaki tüm alfa değerlerinde performansı artırdığı anlaşılmaktadır.

Tablo 5. Dry Bean veri setinin sınıflandırılması amacıyla 100 nöronlu bir gizli katman ve MSE fonksiyonuyla tasarlanan KarcıFANN ve SGD modellerinin kullanılması sonucu elde edilen doğruluk, MSE, kesinlik, duyarlılık ve F1 skoru değerleri.

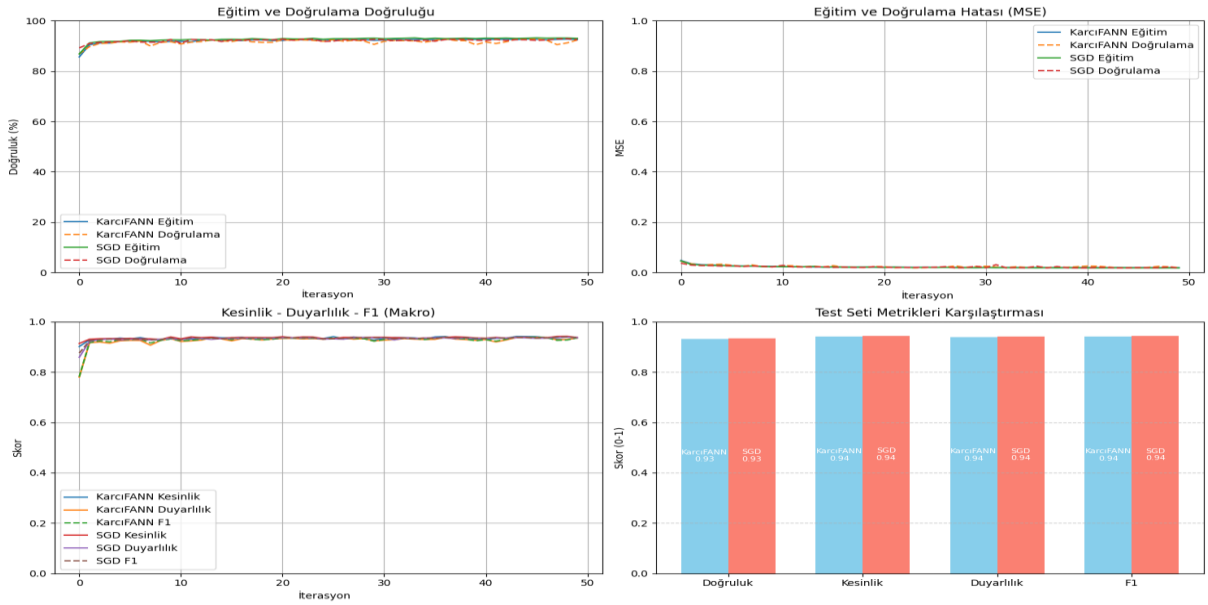
	Aktivasyon Fonksiyonu	Alfa/ Öğrenme Katsayısı	Eğitim		Doğrulama					Test					
			Doğruluk	MSE	Doğruluk	MSE	Kesinlik	Duyarlılık	F1	Doğruluk	MSE	Kesinlik	Duyarlılık	F1	Süre
KarcıFANN	Sigmoid	0.005	88.67	0.0390	90.88	0.0418	0.936	0.921	0.924	89.20	0.0411	0.923	0.906	0.908	12.39
		0.05	89.75	0.0338	91.60	0.0288	0.929	0.932	0.930	91.88	0.0281	0.934	0.934	0.933	13.86
		0.55	91.85	0.0195	92.72	0.0188	0.943	0.939	0.941	91.84	0.0195	0.933	0.930	0.931	12.32
		0.73	92.38	0.0205	91.60	0.0205	0.933	0.927	0.930	92.80	0.0187	0.940	0.936	0.938	12.11
		0.85	92.21	0.0198	92.40	0.0191	0.938	0.931	0.934	92.24	0.0189	0.936	0.931	0.933	12.02
		1.1	9.39	-	10.76	-	0.015	0.143	0.028	9.76	-	0.014	0.143	0.025	12.32
		1.25	9.74	-	9.40	-	0.013	0.143	0.025	9.92	-	0.014	0.143	0.026	12.04
	ReLU	0.005	87.55	0.0410	80.08	0.0511	0.706	0.793	0.742	80.36	0.0502	0.698	0.798	0.739	12.63
		0.05	90.29	0.0304	90.24	0.0345	0.921	0.915	0.916	89.48	0.0354	0.911	0.909	0.906	13.32
		0.55	92.54	0.0190	92.36	0.0192	0.936	0.935	0.936	93.24	0.0189	0.942	0.940	0.941	11.44
		0.73	92.63	0.0189	92.88	0.0197	0.943	0.939	0.941	93.40	0.0185	0.946	0.942	0.944	11.22
		0.85	93.02	0.0176	93.84	0.0162	0.951	0.946	0.948	91.96	0.0198	0.940	0.934	0.936	11.28
		1.1	93.32	0.0079	92.28	0.0215	0.935	0.931	0.933	92.36	0.0192	0.940	0.939	0.940	11.60
		1.25	10.02	-	9.04	-	0.013	0.143	0.024	9.32	-	0.013	0.143	0.024	13.90
SGD	Sigmoid	0.005	73.81	0.0736	74.44	0.0742	0.812	0.780	0.750	73.12	0.0741	0.828	0.769	0.739	11.40
		0.05	88.15	0.0483	86.00	0.0490	0.897	0.882	0.880	88.48	0.0466	0.918	0.895	0.899	11.53
		0.55	91.10	0.0212	91.96	0.0201	0.932	0.925	0.928	91.44	0.0209	0.926	0.922	0.924	11.05
		0.73	10.05	-	9.28	-	0.013	0.143	0.024	9.00	-	0.013	0.143	0.024	11.14
		0.85	9.87	-	9.36	-	0.013	0.143	0.024	9.52	-	0.014	0.143	0.025	10.67
		1.1	9.39	-	10.76	-	0.015	0.143	0.028	9.76	-	0.014	0.143	0.025	11.54
		1.25	9.74	-	9.40	-	0.013	0.143	0.025	9.92	-	0.014	0.143	0.026	10.75
	ReLU	0.005	90.81	0.0388	90.48	0.0390	0.923	0.910	0.916	91.08	0.0384	0.926	0.917	0.921	10.74
		0.05	92.27	0.0245	92.16	0.0247	0.939	0.932	0.935	92.76	0.246	0.942	0.937	0.939	11.71
		0.55	93.00	0.0180	92.48	0.0177	0.938	0.937	0.937	93.44	0.0173	0.945	0.942	0.943	10.74
		0.73	92.70	0.0189	92.88	0.0212	0.940	0.938	0.939	93.52	0.0196	0.944	0.942	0.943	10.90
		0.85	92.90	0.0180	93.20	0.0224	0.946	0.947	0.945	90.64	0.0256	0.930	0.928	0.927	10.44
		1.1	93.29	0.0185	91.72	0.0210	0.932	0.930	0.930	91.96	0.0198	0.940	0.939	0.938	10.64
		1.25	92.97	0.0179	92.52	0.0191	0.940	0.935	0.937	92.96	0.0200	0.942	0.937	0.939	12.17

Tablo 6. Dry Bean veri setinin sınıflandırılması amacıyla 100 nöronlu bir gizli katman ve CE fonksiyonuyla tasarlanan KarcıFANN ve SGD modellerinin kullanılması sonucu elde edilen doğruluk, CE, kesinlik, duyarlılık ve F1 skoru değerleri.

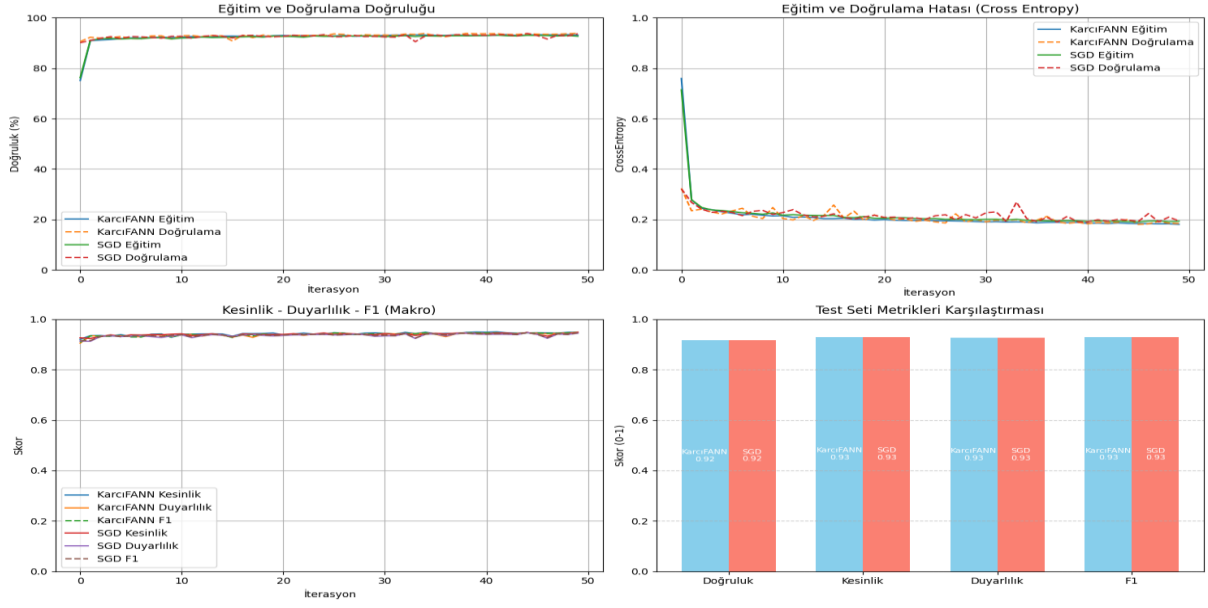
	Aktivasyon Fonksiyonu	Alfa/ Öğrenme Katsayısı	Eğitim		Doğrulama					Test					
			Doğruluk	CE	Doğruluk	CE	Kesinlik	Duyarlılık	F1	Doğruluk	CE	Kesinlik	Duyarlılık	F1	Süre
KarcıFANN	Sigmoid	0.005	92.67	0.1987	93.48	0.1723	0.945	0.944	0.944	92.72	0.2113	0.941	0.939	0.940	11.57
		0.05	93.01	0.1908	91.84	0.2289	0.934	0.934	0.933	92.64	0.1852	0.938	0.940	0.938	10.93
		0.55	93.11	0.1808	93.84	0.1843	0.948	0.948	0.948	91.76	0.2682	0.930	0.927	0.928	12.23
		0.73	93.26	0.1863	92.52	0.2186	0.937	0.932	0.934	92.56	0.2159	0.940	0.937	0.937	11.23
		0.85	93.32	0.1848	92.84	0.1907	0.943	0.939	0.941	92.80	0.2011	0.942	0.938	0.940	10.87
		1.1	92.99	0.1908	93.72	0.1847	0.950	0.946	0.947	92.80	0.1913	0.943	0.937	0.939	11.44
		1.25	92.82	0.1938	93.12	0.1875	0.939	0.939	0.938	92.60	0.1886	0.937	0.936	0.936	11.18
	ReLU	0.005	92.72	0.1932	92.72	0.1957	0.941	0.941	0.941	92.76	0.2297	0.938	0.939	0.938	11.61
		0.05	93.19	0.1869	93.08	0.2126	0.943	0.940	0.941	92.76	0.2175	0.940	0.941	0.940	11.32
		0.55	93.57	0.1786	93.40	0.1901	0.945	0.944	0.944	93.24	0.1744	0.944	0.942	0.943	19.07
		0.73	93.44	0.1726	92.28	0.1965	0.940	0.940	0.939	92.56	0.2081	0.939	0.937	0.937	11.91
		0.85	93.65	0.1709	92.56	0.2127	0.940	0.937	0.938	94.16	0.1713	0.952	0.951	0.952	12.50
		1.1	93.47	0.1830	93.68	0.1798	0.949	0.945	0.947	91.84	0.2271	0.935	0.927	0.930	11.40
		1.25	9.93	-	9.68	-	0.014	0.143	0.025	9.00	-	0.013	0.143	0.024	11.07
SGD	Sigmoid	0.005	90.59	0.4462	90.64	0.4443	0.922	0.912	0.915	89.92	0.4524	0.917	0.906	0.910	10.54
		0.05	92.27	0.2242	92.08	0.2660	0.934	0.933	0.933	92.52	0.2075	0.935	0.935	0.935	10.63
		0.55	92.68	0.1930	93.68	0.1942	0.948	0.945	0.946	91.80	0.2569	0.930	0.927	0.928	10.34
		0.73	93.00	0.1895	92.44	0.2165	0.936	0.937	0.936	91.92	0.2140	0.934	0.936	0.934	10.37
		0.85	93.01	0.1862	92.88	0.1945	0.944	0.938	0.940	92.52	0.2107	0.940	0.933	0.936	9.58
		1.1	92.96	0.1909	93.60	0.1692	0.946	0.943	0.945	93.00	0.1819	0.945	0.940	0.943	10.00
		1.25	92.72	0.1898	93.24	0.1847	0.944	0.941	0.941	93.32	0.1842	0.945	0.941	0.942	11.24
	ReLU	0.005	92.07	0.2465	92.24	0.2300	0.937	0.933	0.935	92.28	0.2368	0.935	0.931	0.933	10.16
		0.05	93.03	0.1849	92.68	0.1996	0.938	0.937	0.937	92.88	0.2045	0.940	0.941	0.940	10.83
		0.55	93.45	0.1743	92.24	0.2069	0.935	0.937	0.936	92.80	0.1846	0.942	0.943	0.941	16.87
		0.73	93.38	0.1730	88.28	0.3458	0.902	0.884	0.880	89.24	0.3262	0.910	0.886	0.887	10.00
		0.85	93.44	0.1700	92.16	0.2169	0.939	0.931	0.934	93.80	0.1783	0.950	0.946	0.948	11.21
		1.1	93.38	0.1811	92.88	0.1884	0.944	0.944	0.942	91.68	0.2214	0.934	0.930	0.931	10.26
		1.25	93.11	0.1849	93.08	0.2322	0.943	0.939	0.940	93.52	0.1918	0.945	0.942	0.943	10.36



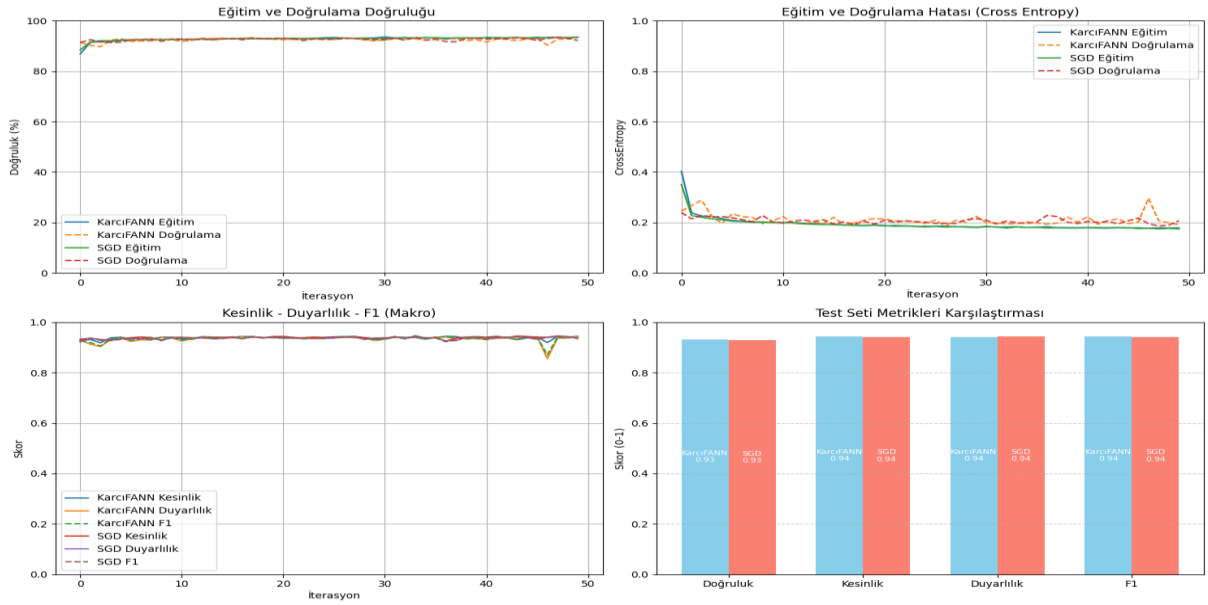
Şekil 10. Dry Bean veri setinin sınıflandırılması amacıyla, Sigmoid aktivasyon fonksiyonu ve MSE hata fonksiyonunun kullanıldığı; 100 nöronlu bir gizli katmana sahip KarciFANN ve SGD modellerinde, alfa/öğrenme katsayısı = 0.55 için elde edilen grafikler.



Şekil 11. Dry Bean veri setinin sınıflandırılması amacıyla, ReLU aktivasyon fonksiyonu ve MSE hata fonksiyonunun kullanıldığı; 100 nöronlu bir gizli katmana sahip KarciFANN ve SGD modellerinde, alfa/öğrenme katsayısı = 0.55 için elde edilen grafikler.



Şekil 12. Dry Bean veri setinin sınıflandırılması amacıyla, Sigmoid aktivasyon fonksiyonu ve CE hata fonksiyonunun kullanıldığı; 100 nöronlu bir gizli katmana sahip KarıcıFANN ve SGD modellerinde, alfa/öğrenme katsayısı = 0.55 için elde edilen grafikler.



Şekil 13. Dry Bean veri setinin sınıflandırılması amacıyla, ReLU aktivasyon fonksiyonu ve CE hata fonksiyonunun kullanıldığı; 100 nöronlu bir gizli katmana sahip KarıcıFANN ve SGD modellerinde, alfa/öğrenme katsayısı = 0.55 için elde edilen grafikler.

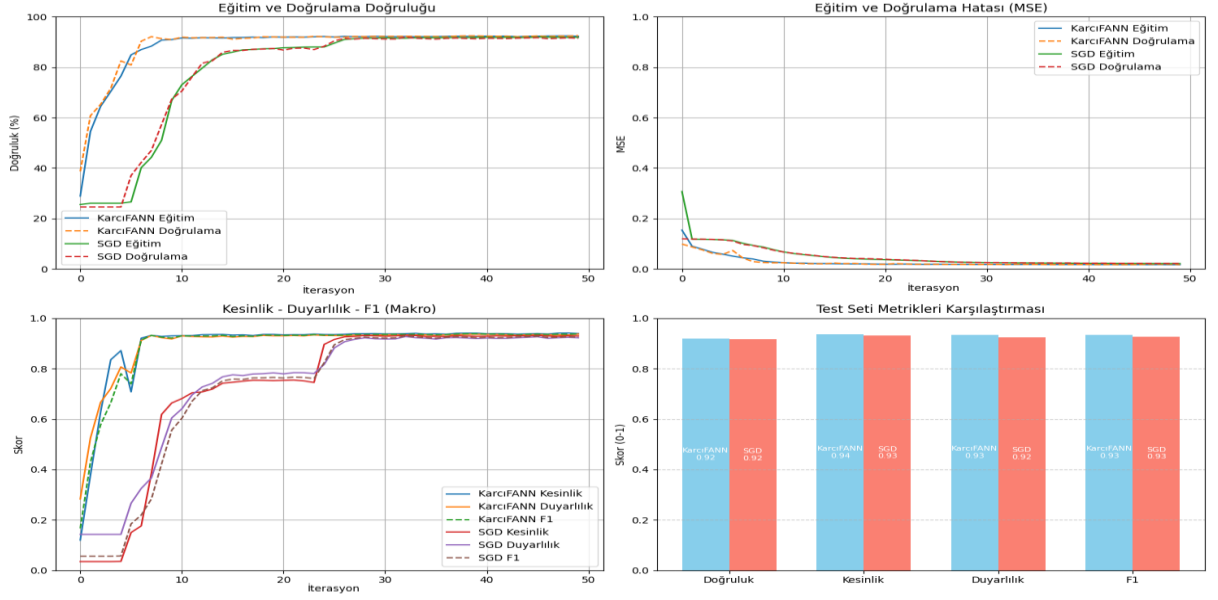
Tablo 7. Dry Bean veri setinin sınıflandırılması amacıyla 100'er nöronlu iki gizli katman ve MSE fonksiyonuyla tasarlanan KarcıFANN ve SGD modellerinin kullanılması sonucu elde edilen doğruluk, MSE, kesinlik, duyarlılık ve F1 skoru değerleri.

	Aktivasyon Fonksiyonu	Alfa/ Öğrenme Katsayısı	Eğitim		Doğrulama					Test					
			Doğruluk	MSE	Doğruluk	MSE	Kesinlik	Duyarlılık	F1	Doğruluk	MSE	Kesinlik	Duyarlılık	F1	Süre
KarcıFANN	Sigmoid	0.005	92.27	0.212	91.88	0.0191	0.933	0.928	0.930	92.32	0.0162	0.938	0.932	0.934	16.13
		0.05	92.74	0.0187	93.28	0.0199	0.947	0.941	0.943	91.56	0.0232	0.935	0.927	0.931	13.93
		0.55	92.27	0.0168	92.52	0.0168	0.939	0.937	0.938	92.00	0.0178	0.936	0.934	0.935	18.45
		0.73	92.52	0.0174	92.44	0.0178	0.939	0.934	0.937	91.68	0.0185	0.929	0.929	0.929	15.25
		0.85	92.44	0.0185	91.44	0.0206	0.928	0.926	0.927	91.84	0.0197	0.933	0.928	0.930	14.01
		1.1	59.71	0.0695	58.16	0.0699	0.493	0.544	0.457	57.48	0.0684	0.500	0.543	0.449	16.73
		1.25	38.11	0.1031	42.76	0.1029	0.215	0.339	0.251	43.44	0.1039	0.218	0.339	0.254	14.67
	ReLU	0.005	92.59	0.0386	92.88	0.0213	0.944	0.941	0.942	92.12	0.0241	0.937	0.931	0.933	14.60
		0.05	93.03	0.0188	93.52	0.0207	0.949	0.945	0.946	91.44	0.0264	0.931	0.926	0.928	16.14
		0.55	93.65	0.0145	93.40	0.0149	0.947	0.945	0.946	93.20	0.0150	0.944	0.940	0.942	15.40
		0.73	93.91	0.0138	92.32	0.0164	0.938	0.936	0.937	93.16	0.0158	0.945	0.943	0.944	14.33
		0.85	93.68	0.0142	93.20	0.0146	0.945	0.945	0.945	93.20	0.0145	0.944	0.943	0.943	12.35
		1.1	93.54	0.0144	93.52	0.0142	0.948	0.942	0.945	93.04	0.0153	0.942	0.941	0.941	14.87
		1.25	93.44	0.0144	93.00	0.0164	0.944	0.939	0.941	92.88	0.0170	0.939	0.938	0.939	14.83
SGD	Sigmoid	0.005	40.26	0.1003	42.20	0.1000	0.166	0.294	0.203	39.68	0.1003	0.165	0.298	0.199	12.34
		0.05	73.81	0.0742	72.28	0.0742	0.703	0.709	0.680	72.08	0.0752	0.701	0.709	0.680	12.75
		0.55	91.86	0.0212	91.48	0.0215	0.930	0.923	0.925	91.72	0.0212	0.933	0.925	0.928	14.40
		0.73	91.87	0.0225	92.00	0.0217	0.935	0.927	0.931	91.24	0.0231	0.926	0.920	0.921	15.34
		0.85	55.75	0.0755	54.20	0.0764	0.415	0.427	0.316	55.76	0.0747	0.432	0.430	0.327	12.33
		1.1	26.35	0.1183	25.08	0.1198	0.036	0.143	0.057	26.00	0.1192	0.037	0.143	0.059	12.14
		1.25	25.47	0.1184	25.32	0.1188	0.036	0.143	0.058	25.96	0.1190	0.037	0.143	0.059	12.66
	ReLU	0.005	91.14	0.0354	91.48	0.0360	0.931	0.924	0.927	90.68	0.0364	0.922	0.916	0.918	12.53
		0.05	92.82	0.0196	92.84	0.0190	0.940	0.937	0.939	91.44	0.0226	0.929	0.926	0.927	14.15
		0.55	93.47	0.0145	92.92	0.0150	0.943	0.942	0.942	92.64	0.0154	0.938	0.936	0.937	13.59
		0.73	93.71	0.0141	92.20	0.0173	0.936	0.936	0.935	92.76	0.0167	0.942	0.941	0.941	11.76
		0.85	93.60	0.0145	93.12	0.0146	0.945	0.944	0.944	93.16	0.0142	0.945	0.940	0.942	12.35
		1.1	93.40	0.0144	93.24	0.0140	0.946	0.940	0.943	93.24	0.0146	0.944	0.942	0.943	12.58
		1.25	93.64	0.0141	93.36	0.0155	0.951	0.942	0.946	92.84	0.0166	0.942	0.936	0.938	12.98

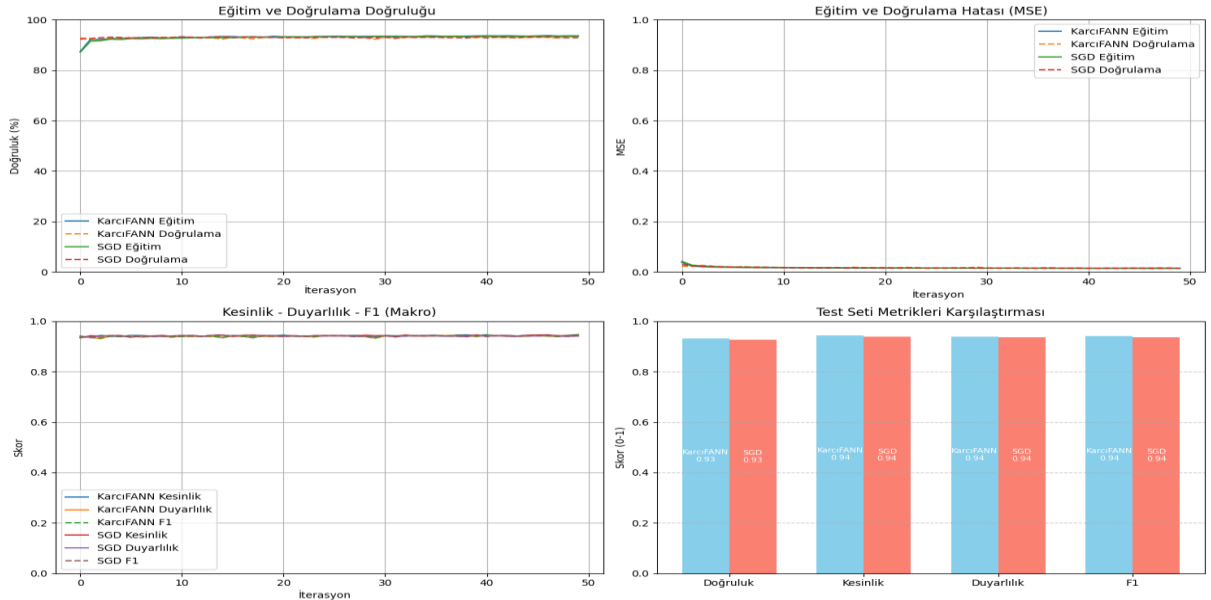
Tablo 8. Dry Bean veri setinin sınıflandırılması amacıyla 100'er nöronlu iki gizli katman ve CE fonksiyonuyla tasarlanan KarcıFANN ve SGD modellerinin kullanılması sonucu elde edilen doğruluk, CE, kesinlik, duyarlılık ve F1 skoru değerleri.

	Aktivasyon Fonksiyonu	Alfa/ Öğrenme Katsayısı	Eğitim		Doğrulama					Test					
			Doğruluk	CE	Doğruluk	CE	Kesinlik	Duyarlılık	F1	Doğruluk	CE	Kesinlik	Duyarlılık	F1	Süre
KarcıFANN	Sigmoid	0.005	92.81	0.1961	92.64	0.2029	0.943	0.933	0.937	92.52	0.1976	0.944	0.933	0.937	15.36
		0.05	93.06	0.1872	92.32	0.2193	0.939	0.931	0.934	92.64	0.2184	0.935	0.935	0.935	16.46
		0.55	93.03	0.1823	92.40	0.1963	0.939	0.939	0.939	93.76	0.1792	0.949	0.947	0.948	13.75
		0.73	93.18	0.1805	93.12	0.1906	0.943	0.941	0.942	92.00	0.2179	0.933	0.930	0.931	15.45
		0.85	93.09	0.1888	92.68	0.1973	0.943	0.937	0.939	92.64	0.1985	0.941	0.935	0.937	14.20
		1.1	93.07	0.1893	93.40	0.1795	0.947	0.945	0.946	92.84	0.1835	0.938	0.937	0.938	13.83
		1.25	92.89	0.1906	92.76	0.2020	0.941	0.933	0.936	93.04	0.2085	0.944	0.937	0.940	15.32
	ReLU	0.005	93.53	0.1706	92.24	0.2418	0.938	0.931	0.934	92.28	0.2287	0.942	0.933	0.937	17.21
		0.05	93.60	0.1760	93.36	0.1979	0.943	0.943	0.943	93.40	0.1713	0.946	0.944	0.945	14.37
		0.55	94.30	0.1519	92.92	0.2033	0.943	0.940	0.941	93.12	0.2133	0.947	0.943	0.945	14.35
		0.73	94.17	0.1507	92.28	0.2344	0.936	0.930	0.932	92.80	0.2173	0.942	0.935	0.938	13.59
		0.85	94.12	0.1530	91.96	0.2451	0.933	0.931	0.932	92.64	0.2464	0.942	0.934	0.938	14.38
		1.1	93.96	0.1508	92.16	0.2256	0.938	0.931	0.934	91.92	0.2297	0.935	0.929	0.931	16.68
		1.25	10.09	-	9.04	-	0.013	0.143	0.024	9.08	-	0.013	0.143	0.024	14.01
SGD	Sigmoid	0.005	71.54	0.9656	72.72	0.9575	0.651	0.596	0.581	71.60	0.9659	0.636	0.588	0.570	13.57
		0.05	91.89	0.2427	91.48	0.2397	0.929	0.926	0.927	91.28	0.2535	0.925	0.925	0.925	14.52
		0.55	92.79	0.1925	91.68	0.2209	0.937	0.930	0.932	92.88	0.2181	0.947	0.935	0.939	12.08
		0.73	92.85	0.1889	90.72	0.2463	0.930	0.933	0.928	90.08	0.2793	0.922	0.926	0.920	13.80
		0.85	93.28	0.1888	92.44	0.2112	0.943	0.930	0.936	92.36	0.2060	0.941	0.933	0.936	13.63
		1.1	93.09	0.1873	91.40	0.2174	0.934	0.935	0.932	91.32	0.2144	0.931	0.932	0.928	11.51
		1.25	92.60	0.1893	90.56	0.2961	0.936	0.915	0.921	90.24	0.2690	0.931	0.911	0.917	12.92
	ReLU	0.005	92.39	0.2153	92.44	0.2269	0.937	0.937	0.937	91.88	0.2269	0.934	0.932	0.933	13.46
		0.05	93.22	0.1814	92.96	0.1962	0.940	0.936	0.938	93.36	0.1778	0.946	0.942	0.944	13.25
		0.55	94.12	0.1505	93.20	0.2092	0.942	0.942	0.942	93.04	0.2160	0.943	0.943	0.913	12.43
		0.73	94.12	0.1534	92.68	0.2051	0.941	0.941	0.941	92.76	0.1976	0.940	0.936	0.938	12.40
		0.85	94.09	0.1461	91.96	0.2514	0.929	0.935	0.931	92.36	0.2318	0.936	0.935	0.935	11.97
		1.1	93.88	0.1672	92.32	0.2452	0.935	0.933	0.933	92.16	0.2325	0.935	0.933	0.934	13.72
		1.25	68.73	0.9105	76.48	0.7749	0.690	0.636	0.647	78.12	0.7344	0.700	0.654	0.663	11.65

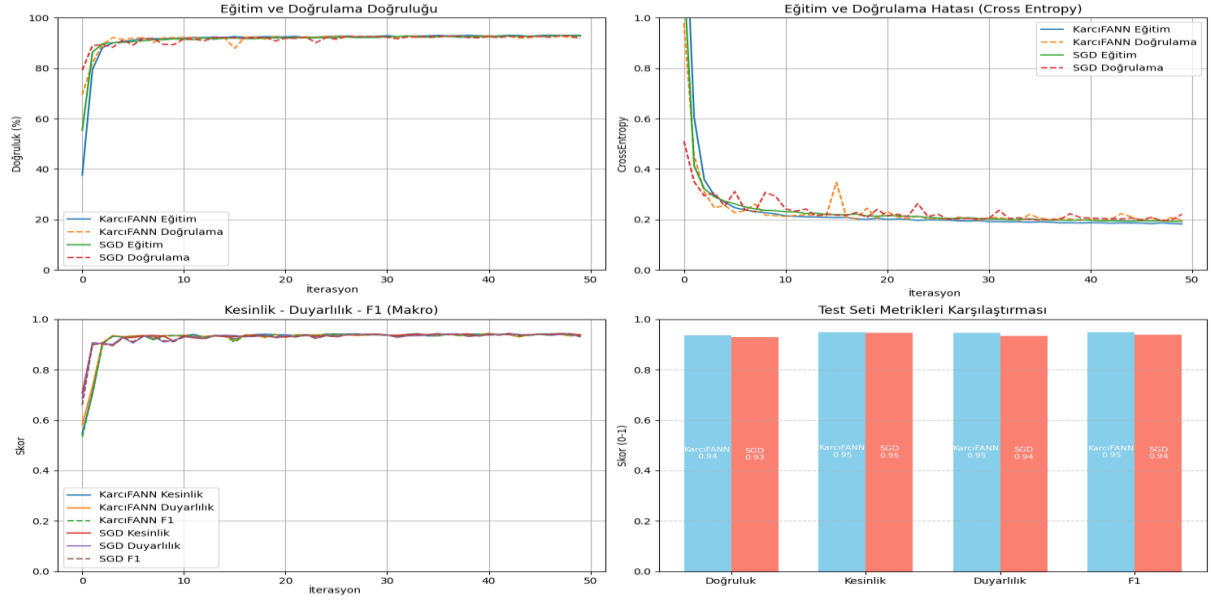
MSE fonksiyonu ile tasarlanan dört katmanlı KarcıFANN ve SGD modellerinde, alfa/öğrenme katsayısının 0.55 değeri için Sigmoid fonksiyonunun kullanılmasıyla elde edilen grafikler Şekil 14'te, Sigmoid yerine ReLU fonksiyonunun kullanılmasıyla elde edilen grafikler de Şekil 15'te gösterilmektedir. Ayrıca, MSE yerine CE fonksiyonunun kullanıldığı modellerde elde edilen grafikler sırasıyla Şekil 16 ve Şekil 17 ile sunulmaktadır.



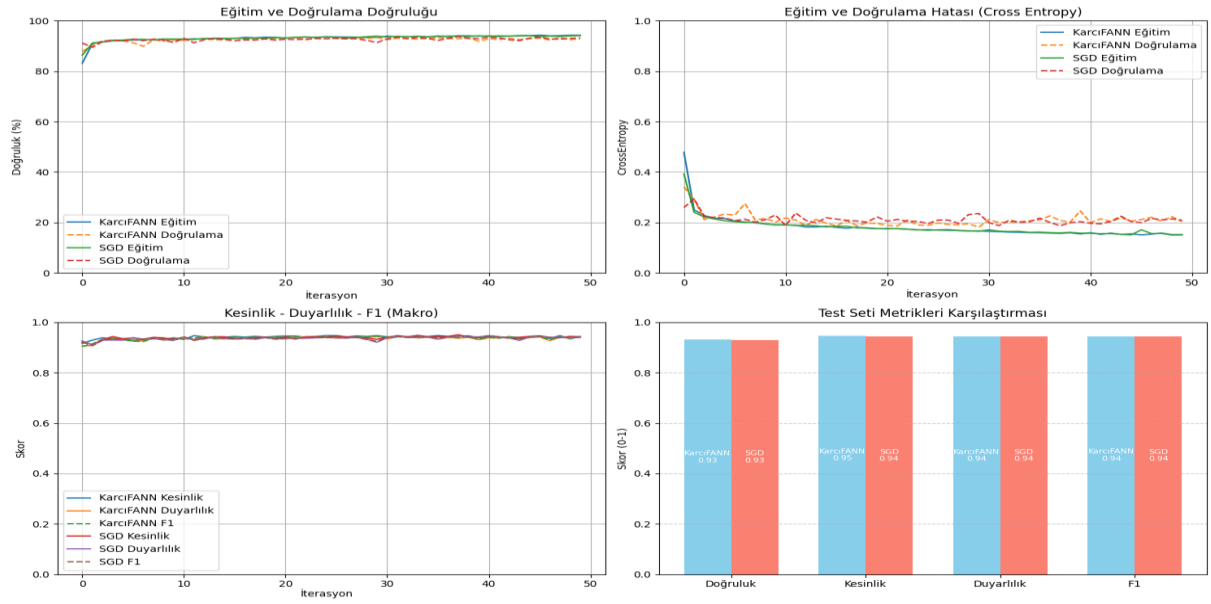
Şekil 14. Dry Bean veri setinin sınıflandırılması amacıyla, Sigmoid aktivasyon fonksiyonu ve MSE hata fonksiyonunun kullanıldığı; 100'er nöronlu iki gizli katmana sahip KarcıFANN ve SGD modellerinde, alfa/öğrenme katsayısı = 0.55 için elde edilen grafikler.



Şekil 15. Dry Bean veri setinin sınıflandırılması amacıyla, ReLU aktivasyon fonksiyonu ve MSE hata fonksiyonunun kullanıldığı; 100'er nöronlu iki gizli katmana sahip KarcıFANN ve SGD modellerinde, alfa/öğrenme katsayısı = 0.55 için elde edilen grafikler.



Şekil 16. Dry Bean veri setinin sınıflandırılması amacıyla, Sigmoid aktivasyon fonksiyonu ve CE hata fonksiyonunun kullanıldığı; 100'er nöronlu iki gizli katmana sahip KarciFANN ve SGD modellerinde, alfa/öğrenme katsayısı = 0.55 için elde edilen grafikler.



Şekil 17. Dry Bean veri setinin sınıflandırılması amacıyla, ReLU aktivasyon fonksiyonu ve CE hata fonksiyonunun kullanıldığı; 100'er nöronlu iki gizli katmana sahip KarciFANN ve SGD modellerinde, alfa/öğrenme katsayısı = 0.55 için elde edilen grafikler.

4. Sonuç

Bu çalışmada, KarciFANN ve SGD yöntemlerinin kullanıldığı farklı modeller tasarlanarak MNIST ve Dry Bean veri setlerini sınıflandırılmıştır. Yapılan deneylerde, modellerin kendi içlerinde ve karşılıklı olarak performansları değerlendirilmiştir. Farklı katman sayıları, hata ve aktivasyon fonksiyonları kullanarak tasarlanan tüm modellerin performansları, kesinlik, F1 skoru, doğruluk, duyarlılık, MSE ve CE metrikleri ile değerlendirilmiştir.

Elde edilen sonuçlar, henüz tam olarak optimize edilmemiş olmasına rağmen KarcıFANN yönteminin SGD'ye kıyasla benzer ya da üstün performans gösterebildiğini ortaya koymuştur. Bu durum, kesir dereceli türev kullanımının model performansını artırdığını ve KarcıFANN'ın SGD'ye güçlü bir alternatif oluşturabileceğini göstermesi bakımından dikkat çekicidir. Bununla birlikte, farklı model yapılarında KarcıFANN yönteminin daha başarılı sonuçlar üretebileceği ve daha karmaşık problemlere de uyarlanabileceği öngörülmektedir.

Kaynaklar

- Chen, S., Zhang, C. & Mu, H. An Adaptive Learning Rate Deep Learning Optimizer Using Long and Short-Term Gradients Based on G–L Fractional-Order Derivative. *Neural Process Lett* 56, 106 (2024). <https://doi.org/10.1007/s11063-024-11571-7>
- Cheng W, Pu R, Wang B. AMC: Adaptive Learning Rate Adjustment Based on Model Complexity. *Mathematics*. 2025; 13(4):650. <https://doi.org/10.3390/math13040650>
- Christen, P., Hand, D. J., & Kirielle, N. (2023). A review of the F-measure: its history, properties, criticism, and alternatives. *ACM Computing Surveys*, 56(3), 1-24.
- Dogo, E. M., Afolabi, O. J., Nwulu, N. I., Twala, B., & Aigbavboa, C. O. (2018). A comparative analysis of gradient descent-based optimization algorithms on convolutional neural networks. In 2018 international conference on computational techniques, electronics and mechanical systems (CTEMS,) pp. 92-99. IEEE.
- Dubey, S. R., Singh, S. K., & Chaudhuri, B. B. (2022). Activation functions in deep learning: A comprehensive survey and benchmark. *Neurocomputing*, 503, 92-108. doi: 10.1016/j.neucom.2022.06.111.
- Herrera-Alcántara, O. (2022). Fractional Derivative Gradient-Based Optimizers for Neural Networks and Human Activity Recognition. *Applied Sciences*, 12(18), 9264. <https://doi.org/10.3390/app12189264>
- Jadon, A., Patil, A., & Jadon, S. (2024). A comprehensive survey of regression-based loss functions for time series forecasting. In *International Conference on Data Management, Analytics & Innovation* (pp. 117-147). Singapore: Springer Nature Singapore.
- Jeong, J. J., & Koo, G. (2024). AdaLo: Adaptive learning rate optimizer with loss for classification. *SSRN*. <https://doi.org/10.2139/ssrn.4737256>
- Karakurt, M. & İşeri, İ. (2022). Patoloji Görüntülerinin Derin Öğrenme Yöntemleri İle Sınıflandırılması. *Avrupa Bilim ve Teknoloji Dergisi*, (33), 192-206.
- Karakurt, M., Saygılı, H. ve Karcı, A. (2024). Comparison of Activation Functions in the KarcıFANN Method. 8th International Artificial Intelligence and Data Processing Symposium (IDAP2024), IEEE, doi: 10.1109/IDAP64064.2024.10711149.
- Karakurt, M., Saygili, 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.
- Karcı, A. (2013a). “A New Approach for Fractional Order Derivative and Its Applications”, *Universal Journal of Engineering Sciences*, Vol:1, pp: 110-117.
- Karcı, A. (2013b). Generalized fractional order derivatives, its properties and applications. *arXiv preprint arXiv:1306.5672*.
- Karcı, A. (2015a). Kesir Dereceli Türevin Yeni Yaklaşımının Özellikleri. *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*, 30(3), 487-501.
- Karcı, A. (2015b). “Properties of Fractional Order Derivatives for Groups of Relations/Functions”, *Universal Journal of Engineering Sciences*, vol:3, pp:39-45.

- Karci, A. (2015c). Chain rule for fractional order derivative. *Science Innovation*, Vol:3, pp:63-67.
- Karci, A. (2019). Properties of Karcı's Fractional Order Derivative. *Universal Journal of Engineering Science*, Vol:7, pp:32-38.
- Karci, A. (2021). Fractional Order Integration: A New Perspective based on Karcı's Fractional Order Derivative. *Computer Science*, 6(2), 102-105.
- Koşan, M. A., Coşkun, A., & Karacan, H. (2019). Yapay zeka yöntemlerinde entropi. *Journal of Information Systems and Management Research*, 1(1), 15-22.
- LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278-2324.
- Lewis, D. D. (1991). Evaluating text categorization. In *Speech and Natural Language: Proceedings of a Workshop Held at Pacific Grove*, 312-138. California, <https://aclanthology.org/H91-1061.pdf>.
- McCulloch, W. S. and Pitts, W. 1943. A Logical Calculus of the Ideas Immanent in Nervous Activity. *The Bulletin of Mathematical Biophysics*, 5:4, 115-133.
- Nwankpa, C., Ijomah, W., Gachagan, A., & Marshall, S. (2018). Activation functions: Comparison of trends in practice and research for deep learning. *arXiv preprint arXiv:1811.03378*.
- Powers, D. M. (2020). Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. *arXiv preprint arXiv:2010.16061*.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *nature*, 323(6088), 533-536.
- 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.
- Saygılı, H., Karakurt, M. & Karcı, A. (baskıda). Karcı kesir dereceli yapay sinir ağı (KarcıFANN): öğrenme oranı, aşırı uyum ve yetersiz uyum sorunlarını çözme. *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*.
- Seyyarer, E., Ayata, F., Uçkan, T., & Karci, A. (2020). Derin öğrenmede kullanılan optimizasyon algoritmalarının uygulanması ve kıyaslanması. *Anatolian Journal of Computer Sciences*, 5(2), 90-98.
- Tian, Y., Zhang, Y., & Zhang, H. (2023). Recent advances in stochastic gradient descent in deep learning. *Mathematics*, 11(3), 682.
- Yang, E., Pan, J., Wang, X., Yu, H., Shen, L., Chen, X., Xiao, L., Jiang, J., & Guo, G. (2022). AdaTask: A Task-aware Adaptive Learning Rate Approach to Multi-task Learning. *AAAI Conference on Artificial Intelligence*.