

## KarcıFANN Yönteminde Ağırlık Sönümlemenin Etkisi

### The Effect of Weight Decay in the KarcıFANN Method

Meral Karakurt<sup>1\*</sup> 

<sup>1\*</sup>Bilgisayar Teknolojileri Bölümü, Osmaniye Korkut Ata Üniversitesi Düziçi MYO, Osmaniye, Türkiye

(meralkarakurt@osmaniye.edu.tr)

Received:Oct.22,2025

Accepted:Nov.18,2025

Published:Dec.01, 2025

**Özetçe:** Yapay sinir ağlarının (YSA) modellenmesinde, optimizasyon yöntemi, aktivasyon ve hata fonksiyonu gibi hiperparametrelerle birlikte çeşitli düzenleme (düzenleştirme) yöntemleri kullanılmaktadır. Bu yöntemlerden biri olan ağırlık sönümleme işlemi, modellerin eğitimi aşamasında ağırlık vektörlerinin çok fazla büyümesiyle ortaya çıkan gradyan patlaması ve ezberleme gibi önemli problemlerin çözülmesi ve modellerin genelleme performanslarının artırılması amacıyla uygulanmaktadır.

Bu çalışmada, ağırlık sönümleme hiperparametresinin, KarcıFANN yönteminin performansına etkileri analiz edilmektedir. Bu amaçla, KarcıFANN yöntemi ile tasarlanan çeşitli modellere ağırlık sönümleme işlemi uygulanmıştır. MNIST ve Dry Bean veri setlerinin sınıflandırılması sonucu elde edilen bulgular, ağırlık sönümleme hiperparametresinin, modellerin başarımını ve genelleme kabiliyetini önemli ölçüde iyileştirdiğini göstermiştir.

**Anahtar Kelimeler:** KarcıFANN, sınıflandırma, düzenleme yöntemleri, ağırlık sönümleme.

**Abstract:** In the modeling of artificial neural networks (ANNs), various regularization methods are employed in conjunction with hyperparameters such as optimization methods, activation functions, and loss functions. Among these methods, weight decay is applied to address critical issues such as gradient explosion and overfitting, which arise from excessive growth of weight vectors during training, and to enhance the generalization performance of models.

In this study, the effects of the weight decay hyperparameter on the performance of the KarcıFANN method are analyzed. For this purpose, weight decay was applied to various models designed using the KarcıFANN method. The empirical results derived from the classification on the MNIST and Dry Bean datasets demonstrate that weight decay hyperparameter markedly improves the predictive performance and the generalization capacity of the models.

**Keywords:** KarcıFANN, classification, regularization methods, weight decay.

## 1. Giriş

Günümüzde artan veri miktarlarının işlenmesi ihtiyacı, YSA'lara duyulan ilgiyi artırmaktadır. YSA'lar; örüntü tanıma, tahmin, kümeleme ve sınıflandırma problemlerine başarılı çözümler getirmeleri sayesinde neredeyse tüm bilim dallarında kullanılmaktadır.

Optimizasyon yöntemi, hata (kayıp) fonksiyonu, aktivasyon fonksiyonu ve katman sayısı gibi birçok hiperparametre kullanılarak YSA modelleri tasarlanmaktadır. Aktivasyon fonksiyonu, bir yapay sinir hücresinin (nöronun) aktiflik durumunu ve öğrenmeye ne oranda katkıda bulunacağını belirlemek amacıyla kullanılmaktadır. Optimizasyon yöntemi, modelin başarımını artırmak amacıyla hata değerini en aza indirme (veya kazancı en yüksek seviyeye çıkarma) işlemleridir (Karakurt vd., 2025). Hata fonksiyonu ise bir modelin eğitilmesiyle elde edilen çıktı değerinin, modelden beklenen çıktı değerinden (etiket bilgisi) ne kadar farklı olduğunu belirlemektedir.

YSA'ların tasarımı aşamasında, problemin yapısı ve kullanılan veri seti göz önünde bulundurularak optimum çözümlere ulaşılmaktadır. Burada belli bir standardın olmayışı, çeşitli hiperparametrelerin değiştirilerek en uygun modelin elde edildiği bir optimizasyon süreci gerektirmektedir. YSA'ların başarımı, bu hiperparametrelerle beraber eğitim sürecinde kullanılan düzenleme yöntemlerine de bağlıdır. YSA modellerinde ağırlık vektörlerinin kontrolsüz biçimde büyümesi (olması gerekenden çok büyük olması), modellerin ezberleme eğilimini artırarak genelleme kabiliyetini azaltmaktadır (Outmezguine ve Levi, 2024).

Ağırlık sönümleme hiperparametresi, YSA modellerinde ezberlemeyi önlemek amacıyla yaygın olarak kullanılan düzenleme yöntemlerinden biridir (Zhang vd., 2018). Bu yöntem, eğitim sürecinde ağırlık vektörlerinin büyüklüğünü sınırlayarak modelin gereksiz karmaşıklık artışı engellemenekte ve böylece genelleme kabiliyetini artırmaktadır (Loshchilov ve Hutter, 2019; Kosson vd., 2023; Galanti, 2025).

Literatürde KarıcıFANN yöntemi ile yapılan ilk çalışmalar, Karakurt vd. (2022), Karakurt vd. (2024, 2025) ve Saygılı vd. (2024, baskıda) tarafından sunulmuştur. Ağırlık sönümleme ile birlikte çeşitli hiperparametrelerin kullanıldığı modellerle de birçok çalışma yapılmıştır. Bu çalışmalardan birinde; Xie vd. (2020), uyarlamalı optimizasyon algoritmalarıyla eğitilen modellerde, ağırlık sönümleme hiperparametresinin sabit bir değer olmasının gradyanların büyümesine neden olabileceğini vurgulayarak *Zaman-planlı Ağırlık Sönümleme (Stable Weight Decay, SWD)* adlı bir yöntem geliştirmişlerdir. Bu yaklaşım, sönümleme katsayısını gradyanlara göre dinamik olarak ayarlamakta; büyük gradyanların sebep olacağı yakınsayamama ve genelleme performansının düşmesi gibi problemleri çözme amaçlamaktadır (Xie vd., 2020).

Loshchilov ve Hutter (2017), uyarlamalı (adaptive) optimizasyon algoritmalarında ağırlık sönümleme işleminin uygulanmasında *Decoupled Weight Decay (AdamW)* adını verdikleri yeni bir yöntem önermişlerdir. Bu yöntem ile ağırlık güncelleme iki adımda yapılmaktadır. İlk adımda, klasik uyarlamalı yöntemlerdeki hesaplamalar yapılarak yeni ağırlık değeri elde edilmekte, ikinci adımda ise bu yeni ağırlık değerine ağırlık sönümleme işlemi uygulanarak güncelleme işlemi tamamlanmaktadır. İlgili çalışmada, AdamW yönteminin klasik L2 düzenleme yöntemine kıyasla daha iyi genelleme sağladığı gösterilmektedir (Loshchilov ve Hutter, 2017). Diğer bir çalışmada, Li vd. (2020), SGD yöntemini genelleştirebilmek amacıyla "*içsel öğrenme katsayısı (intrinsic learning rate)*" adını verdikleri yeni bir hiperparametre önermişlerdir. İlgili çalışmada, öğrenme katsayısı ile ağırlık sönümleme katsayısının çarpımını kullanarak küçük öğrenme katsayılarıyla da genelleştirmenin mümkün olduğunu savunmuşlardır (Li vd., 2020).

Kosson vd. (2023), ağırlık sönümlemenin güncelleme davranışlarını nasıl etkilediğini inceledikleri çalışmada, "*döner denge (rotational equilibrium)*" adını verdikleri kavramı açıklamışlardır. Her ağırlık vektörü için büyüklük ve yön (açısal) güncellemelerinin büyüklüklerinin zamanla bir dengeye geldiğini ve bir ağırlık vektörünün yön değiştirmesinin, modelin yeni yönlere doğru öğrenme yaptığına işaret ettiğini vurgulamışlardır. Yaptıkları deneylerle ağırlık sönümleme işleminin, nöron ağırlıklarının büyüklük ve yön güncellemelerini kararlı bir dengeye yaklaştırdığını ve optimizasyonun kararlılığını artırdığı göstermişlerdir (Kosson vd., 2023).

Zhang vd. (2018), ağırlık sönümlemenin model eğitimindeki düzenleme etkisini inceledikleri çalışmada, ağırlık sönümlemenin genelleme başarımını üç farklı mekanizma aracılığıyla artırdığını belirtmişlerdir: Etkili öğrenme oranını yükseltme, Jacobian normunu düzenleme ve ikinci mertebe optimizasyonda sönüm katsayısını azaltma. SGD, ADAM ve K-FAC optimizasyon yöntemlerini kullanarak CIFAR-10 ve CIFAR-100 veri setlerini sınıflandırdıkları çalışmada elde ettikleri bulgular ile ağırlık sönümlemenin optimizasyonda önemli bir rol oynadığını ortaya koymuşlardır (Zhang vd., 2018).

Outmezguine ve Levi (2024), klasik L2 düzenleme yaklaşımını genelleştirerek *p-norm Weight Decay (pWD)* adında yeni bir yöntem önermişlerdir. Bu yöntemde, ağırlıkların bir kısmı sıfıra veya çok küçük değerlere çekilerek modellerin seyrekleştirilmesi (sparsification) sağlanmaktadır. CIFAR-10 veri setinin sınıflandırılmasında ResNet18 ve Tiny Shakespeare veri setinin sınıflandırılmasında NanoGPT modellerini kullanarak yaptıkları deneyler sonucunda pWD'nin yüksek seyrekleştirme oranlarında bile yüksek doğruluk performansını koruduğunu göstermişlerdir (Outmezguine ve Levi, 2024).

Galanti vd. (2025), SGD yönteminde ağırlık sönümleme işlemini uygularken öğrenme katsayısı, bir iterasyonda işlenen parça boyutu (batch size) ve ağırlık sönümleme katsayısının başarıma etkilerini incelemişlerdir. İlgili çalışmada, parça boyutunun küçültülmesi, öğrenme katsayısı ve ağırlık sönümleme katsayısının da büyütülmesi ile ağırlık vektörünün düşük oran (low-rank) eğiliminin artmasını sağladığını göstermişlerdir (Galanti vd., 2025).

Bu çalışmada, ağırlık sönümleme işleminin KarcıFANN yöntemine etkilerinin incelenmesi amacıyla farklı modeller tasarlanarak deneyler gerçekleştirilmiştir. Uygulama kısmında, açık kaynak kodlu Anaconda tümleşik geliştirme ortamının (IDE) Jupyter notebook uygulaması üzerinde Python dili kullanılmıştır.

## 2. Yapay Sinir Ağının (YSA) Modellenmesi

Canlılara özgü bilişsel davranışlar olan öğrenme, hatırlama, sınıflandırma, nesne tanıma ve karar verme gibi özelliklerin makinelerle yaptırılması ihtiyacının sonucu olarak YSA'lar kullanılmaktadır. YSA, beyindeki nöronlar ve nöronların arasındaki bağlantıların matematiksel olarak modellenmesidir. Bir YSA, veri setlerinin sunulduğu bir girdi katmanı, verilerin işlendiği en az bir gizli katman ve işlenmiş verilerin sunulduğu bir çıktı katmanından oluşmaktadır. Bu katmanlardaki nöronlar, veri setine ve probleme göre değişmektedir.

YSA'lar, öğrenme şekillerine göre girdi ve çıktı değerlerinin bilindiği denetimli öğrenme, sadece girdi değerlerinin bilindiği denetimsiz öğrenme, denetimli ve denetimsiz öğrenmenin birleştirildiği hibrit öğrenme ile bir ödül veya ceza sistemine göre öğrenmenin gerçekleştiği takviyeli öğrenme olarak dört gruba ayrılır.

Bir YSA'nın girdi katmanından çıktıya doğru yapılan işlem dizisine ileri yayılım denir. İleri yayılım sonucunda ağın çıktısının beklenen çıktı değerinden farkı hata değerini vermektedir. Hata hiperparametresinin, çıktı katmanından geriye doğru iletilerek düşürülmesi işlemlerine geri yayılım denir (Rumelhart ve diğ., 1986).

Denetimli öğrenme gerçekleştirilen bu çalışmada, hata değerinin hesaplanmasında Ortalama Karesel Hata (Mean Squared Error, MSE) ve Çapraz Entropi (Cross-Entropy, CE) yöntemleri kullanılmaktadır. MSE, modelin çıktısının etiket bilgisinden (modelin üretmesi gereken bilgi) farkının karesinin ortalamasıdır (Jadon ve diğ., 2024). CE ise modelin çıktısının oluşturduğu olasılık dağılımının, etiket bilgilerini içeren gerçek dağılımdan sapma miktarı olarak hesaplanan bir hata fonksiyonudur (Koşan ve diğ., 2019; Seyyarer ve diğ., 2020).  $n$  örnek sayısı ve  $m$  sınıf sayısı,  $Z$  modelin çıktısı ve  $R$  etiket bilgisi olmak üzere MSE değeri (1) denklemi, CE değeri de (2) denklemi ile hesaplanmaktadır.

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

$$CE = -\frac{1}{n} \sum_{i=1}^n \sum_{k=1}^m R_{i,k} \log(Z_{i,k}) \quad (2)$$

Bu çalışma kapsamında tasarlanan modellerin genelleme kabiliyetini ölçmek amacıyla doğruluk, F1 skoru (F-measure), duyarlılık (hatırlama) ve kesinlik (hassasiyet) metrikleri kullanılmaktadır. Bu metriklerin hesaplanmasında, karmaşıklık matrisi temel alınmaktadır.  $Doğruluk = \frac{DP+DN}{DP+DN+YP+YN}$  denklemi ile hesaplanan doğruluk değeri, modelin doğru tahminlerinin tüm tahminlere bölünmesidir (Christen ve diğ., 2023).

$Kesinlik = \frac{DP}{DP+YP}$  denklemi ile hesaplanan kesinlik değeri, modelin tahmin ettiği pozitif değerler arasındaki doğru pozitiflerin oranıdır (Lewis, 1991; Powers, 2020).  $Duyarlılık = \frac{DP}{DP+YN}$  denklemi ile gösterilen duyarlılık değeri, gerçek pozitif değerler arasında modelin doğru tahmin ettiği pozitiflerin oranıdır (Lewis, 1991; Powers, 2020).  $F1 = 2 \cdot \frac{Kesinlik \cdot Duyarlılık}{Kesinlik + Duyarlılık}$  denklemi ile gösterilen F1 skoru değeri de duyarlılık ve kesinlik metriklerinin harmonik ortalamasıdır (Christen ve diğ., 2023).

Modellerde kullanılan nöronların öğrenmeye ne kadar etki edeceğini belirlemek üzere Sigmoid ve ReLU (Düzeltilmiş Doğrusal Birim - Rectified Linear Unit) aktivasyon fonksiyonları kullanılmaktadır. Sigmoid, her nöronun çıktısını 0 ile 1 arasında sınırlayan bir aktivasyon fonksiyonudur. ReLU

fonksiyonu ise negatif girdiler için sıfır, pozitif girdiler için aynı değeri çıktı olarak sunmaktadır (Seyyarer vd., 2020).  $x$ , bir nöronun çıktısını göstermek üzere Sigmoid fonksiyonu,  $sigmoid(x) = \frac{1}{1+e^{-x}}$  denklemi ve ReLU fonksiyonu,  $relu(x) = \max(0, x)$  denklemi ile hesaplanmaktadır.

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

KarcıFANN optimizasyon yönteminin matematiğinde Karcı kesir dereceli türevi kullanılmaktadır (Karcı, 2013a, 2013b, 2015a, 2015b, 2015c, 2019, 2021). Kesir dereceli türevlerin global modeller oluşturabilmesi fikrine dayanan KarcıFANN yönteminde amaç, öğrenememe, gradyan kaybolması, gradyan patlaması ve ezberleme gibi önemli problemlerin çözülmesidir.

SGD (Stokastik Gradyan İniş, Stochastic Gradient Descent) yönteminde ağırlık güncellemeleri yapılırken herhangi bir  $w$  ağırlığının toplam hataya ( $H$ ) etkisinin hesaplanmasında kullanılan Newton türevine  $(\frac{\partial H}{\partial w})$ , Karcı kesir dereceli türev  $(D_\alpha^K)$  çarpanı  $(\frac{H}{w})^{\alpha-1}$  eklenerek (3) denkleminde gösterildiği gibi hesaplanmaktadır. Böylece, modelin kendi dinamikleriyle öğrenmesi sağlanmaktadır.

$$D_\alpha^K \frac{H}{w} = \left(\frac{H}{w}\right)^{\alpha-1} \cdot \frac{\partial H}{\partial w} \quad (3)$$

Burada kullanılan alfa ( $\alpha$ ) değeri, kesir derecesini ifade etmektedir. Bir  $t$  adımında (3) denklemi ile elde edilen değer ile bir sonraki adımda  $(t + 1)$  kullanılacak olan ağırlık değerinin ( $w$ ) güncellenmesi işlemi,  $w_{t+1} = w_t - D_\alpha^K \frac{H}{w_t}$  denklemi ile gösterilmektedir.

## 2.2. Ağırlık Sönümlleme (Weight Decay)

Ağırlık cezalandırma, ağırlık küçültme veya ağırlık bozunması tanımlarıyla da bilinen ağırlık sönümlleme işlemi, eğitim aşamasında ağırlıkların bir kısmının işlem dışı bırakıldığı bir düzenleme yöntemidir (Xie vd., 2020). Bu işlem, ağırlık güncelleme adımında, hata fonksiyonuna eklenen bir ceza terimiyle gerçekleştirilmekte ve modelin ağırlıklarını sınırlandırmaktadır. Böylece, öğrenme süreci daha kararlı hale getirilerek ezberleme ve gradyan patlaması gibi riskler azaltılmaktadır (Van Laarhoven, 2017; Zhang ve diğ., 2018).

KarcıFANN yönteminde bir  $t + 1$  anındaki ağırlık sönümlleme işlemi, ağırlık sönümlleme katsayısı  $\lambda$  olmak üzere ağırlık güncelleme bağıntısına (4) denkleminde gösterildiği gibi eklenmektedir.

$$w_{t+1} = w_t - D_\alpha^K \cdot \left(\frac{H}{w_t} + \lambda \cdot w_t\right) \quad (4)$$

## 3. Farklı KarcıFANN Modellerinde Ağırlık Sönümlleme İşleminin Uygulanması

MNIST veri seti, toplam 10 sınıfa ayrılan 70000 tane el yazısı rakamın gri tonlu görüntülerinden oluşmaktadır (LeCun ve diğ., 1998). Boyutları  $28 \times 28$  olan tüm görüntüler, KarcıFANN modellerinde işlenmek üzere tek boyutlu ( $28 \times 28 = 784$ ) sayısal vektörlere dönüştürülmüştür. Veri setindeki 55000 örnek eğitim, 10000 örnek test ve 5000 örnek doğrulama (validation) adımlarında kullanılmıştır.

Dry Bean veri seti, barbunya, şeker, dermosan, bombay, cali, horoz ve sira olarak 7 sınıfa ayrılan 13611 fasulye görüntüsünü içermektedir. Bu görüntülerin de 8611 tanesi eğitim, 2500'er tanesi doğrulama ve test için ayrılmıştır. Tüm örneklerle ait 16 sayısallaştırılmış özellik bilgisi modellere sunulmuştur.

Bu çalışmada, MNIST ve Dry Bean veri setlerini sınıflandırmak amacıyla gizli katmanlarında 100 nöronun kullanıldığı üç katmanlı (ağırlık sönümlleme işleminin uygulanmadığı) KarcıFANN modellerinin sonuçları ile bu modellere ağırlık sönümlleme işleminin uygulanmasıyla elde edilen

sonuçlar karşılaştırılmaktadır. Ağırlık sönümlleme işleminde, sönümlleme katsayısına  $\lambda = 1e - 4$  rastgele değeri verilmiş ve tüm modellerde bu değer kullanılmıştır.

Tüm modellerin girdi katmanlarındaki nöron sayıları, veri setlerindeki görüntülerin boyutu ya da özellik bilgisi sayısı; çıktı katmanlarındaki nöron sayıları ise sınıf sayısı olarak belirlenmektedir. Bu bağlamda, girdi katmanlarında MNIST veri seti için 784 nöron ve Dry Bean veri seti için 16 nöron; çıktı katmanlarında ise MNIST veri seti için 10 nöron ve Dry Bean veri seti için 7 nöron kullanılmıştır. Her model için iterasyon (adım, eğitim tur) sayısına 50 değeri ve alfa kesir derecesine rastgele olarak seçilen 0.005, 0.05, 0.55, 0.73, 0.85, 1.1 ve 1.25 değerleri verilmiştir. Modellerin performanslarının ölçülmesinde kullanılan doğruluk, MSE, CE, kesinlik, duyarlılık ve F1 skoru değerleri ve eğitim süreleri tablolar ile kayıt altına alınmıştır. Ayrıca, alfa'nın 0.73 değeri için elde edilen sonuçların grafikleri sunulmuştur.

MNIST veri setinin sınıflandırılmasında Sigmoid ve ReLU aktivasyon fonksiyonları ile MSE ve CE hata fonksiyonları kullanılarak tasarlanan ağırlık sönümllemesiz KarcıFANN modellerinde elde edilen doğruluk, MSE, CE, F1 skoru, kesinlik ve duyarlılık değerleri Tablo 1'de kaydedilmiştir. İlgili tabloda, Sigmoid ve MSE fonksiyonlarının kullanıldığı KarcıFANN modelinin genel performansının, ReLU ve MSE fonksiyonlarının kullanıldığı modele kıyasla daha başarılı olduğu görülmektedir. Benzer şekilde, CE fonksiyonu ile tasarlanan modellerde de Sigmoid fonksiyonu ile daha yüksek performans elde edilmiştir.

Hata fonksiyonları açısından bakıldığında, Sigmoid fonksiyonunun CE fonksiyonu ile birlikte kullanıldığı modelin tüm aşamalarındaki başarımının, MSE ile kullanıldığı modele kıyasla daha yüksek olduğu görülmektedir. Öte yandan, ReLU fonksiyonunun CE fonksiyonuyla kullanıldığı model, alfa'nın 0.005, 0.05 ve 0.55 değerlerinde MSE fonksiyonuyla kullanıldığı modelden genel olarak daha iyi sonuçlar üretmektedir. Ayrıca, MSE ve ReLU fonksiyonlarının kullanıldığı modellerde, 1.1 ve 1.25 alfa değerleri, CE ve ReLU fonksiyonlarının kullanıldığı modellerde ise 0.73, 0.85, 1.1 ve 1.25 alfa değerleri için öğrenmenin gerçekleşmediği gözlemlenmektedir.

MNIST veri setinin sınıflandırılması amacıyla Sigmoid ve ReLU aktivasyon fonksiyonları ile MSE ve CE hata (kayıp) fonksiyonlarının kullanıldığı KarcıFANN modellerinde, alfa'nın 0.73 değeri için elde edilen grafiksel sonuçlar Şekil 1 ve Şekil 2 ile sunulmaktadır. Tüm şekillerde, eğitim ve doğrulama süreçlerinin doğruluk ve hata grafikleri, doğrulama sürecinin kesinlik, duyarlılık ve F1 skoru grafiği ile test sürecinin kesinlik, duyarlılık ve F1 skoru bar grafiği yer almaktadır.

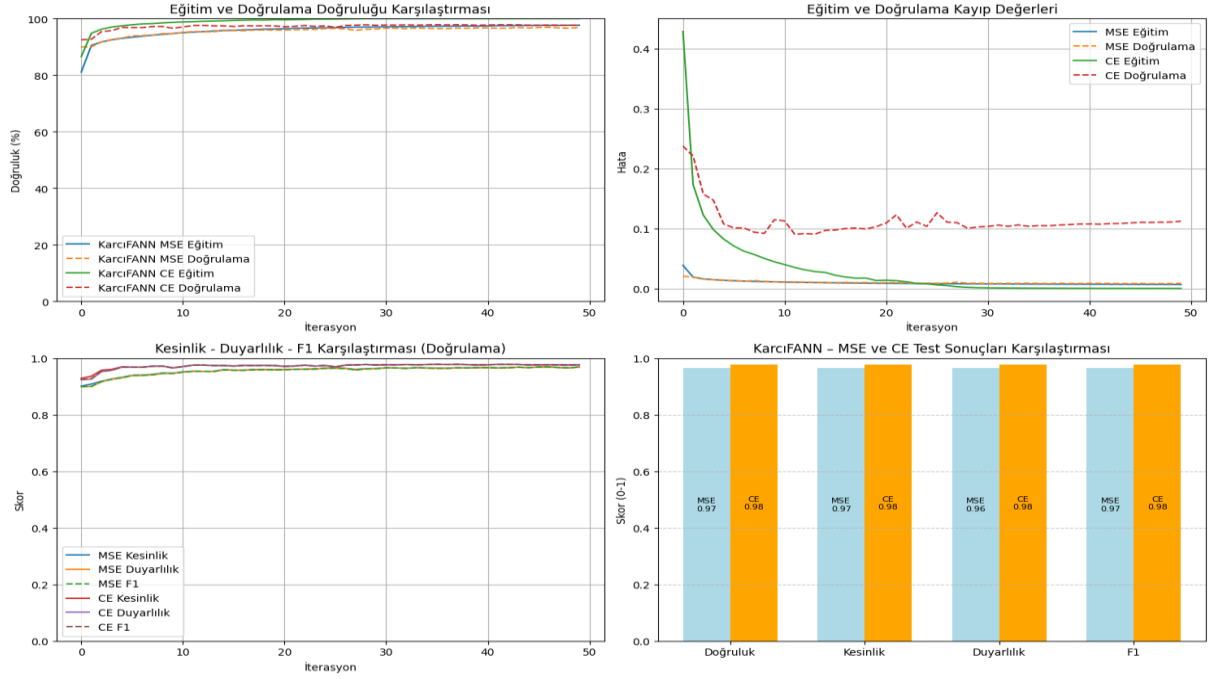
Tablo 1 ile sonuçları gösterilen modellere ağırlık sönümlleme işlemi uygulanarak elde edilen doğruluk, hata, kesinlik, duyarlılık ve F1 skoru değerleri Tablo 2'de sunulmaktadır. Tablo 2'de ve Tablo 4'te ağırlık sönümlleme işleminin uygulandığı KarcıFANN modellerinin, ağırlık sönümllemesiz modellere kıyasla daha başarılı olduğu değerler kırmızı tonla vurgulanmaktadır.

MSE ve CE fonksiyonlarıyla tasarlanan modellerde Sigmoid fonksiyonunun kullanımının, ReLU fonksiyonunun kullanımına kıyasla daha yüksek performans sağladığı Tablo 2 ile gösterilmektedir. Sigmoid ve CE fonksiyonlarıyla tasarlanan model, Sigmoid ve MSE fonksiyonlarıyla tasarlanan modelden daha başarılı olmaktadır. Alfa'nın 0.005, 0.05 ve 0.55 değerleri için ReLU fonksiyonu ile CE fonksiyonunun kullanıldığı modelin, MSE ile birlikte kullanıldığı modelden daha iyi sonuçlar ürettiği görülmektedir. Ayrıca, 1.1 ve 1.25 alfa değerleri için ReLU fonksiyonunun kullanıldığı tüm modellerde öğrenme gerçekleşmemektedir.

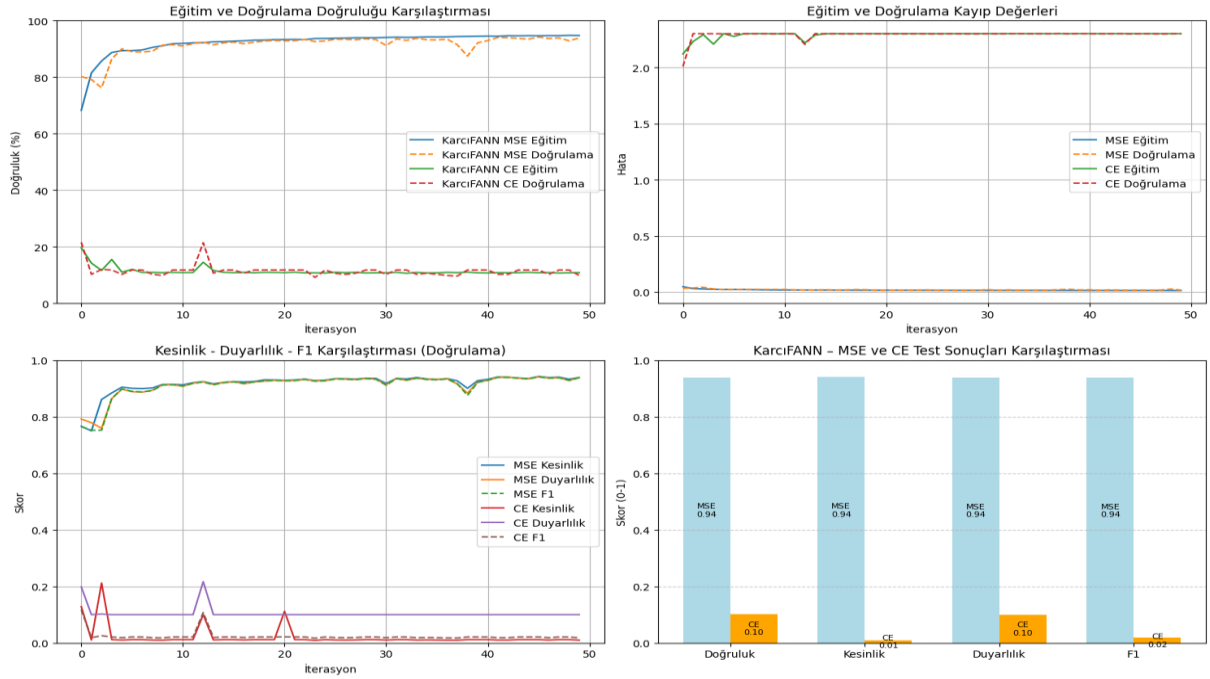
Ağırlık sönümllemesiz KarcıFANN modellerin sonuçlarının sunulduğu Tablo 1 ile ağırlık sönümlleme işleminin uygulandığı modellerin sonuçlarının sunulduğu Tablo 2 incelendiğinde, alfa'nın 1.1 ve 1.25 değerleri için Sigmoid ve MSE fonksiyonlarını, 0.005 ve 0.05 değerleri için ReLU ve MSE fonksiyonlarını, 0.005 değeri için de ReLU ve CE fonksiyonlarını içeren modellere ağırlık sönümllemenin eklenmesiyle tüm aşamalarda performans artışı sağlandığı görülmektedir. Ayrıca, bazı alfa değerleri için farklı metriklerde başarımın arttığı gözlemlenmektedir.

**Tablo 1.** MNIST veri setinin sınıflandırılması amacıyla kullanılan KarcıFANN modelleriyle elde edilen doğruluk, MSE, CE, kesinlik, duyarlılık ve F1 skoru değerleri.

| Aktivasyon Fonksiyonu | Alfa  | Eğitim   |        | Doğrulama |        |          |            |       | Test     |        |          |            |       |        |
|-----------------------|-------|----------|--------|-----------|--------|----------|------------|-------|----------|--------|----------|------------|-------|--------|
|                       |       | Doğruluk | MSE    | Doğruluk  | MSE    | Kesinlik | Duyarlılık | F1    | Doğruluk | MSE    | Kesinlik | Duyarlılık | F1    | Süre   |
| 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  | 97.66    | 0.0068 | 96.92     | 0.0087 | 0.969    | 0.968      | 0.969 | 96.53    | 0.0083 | 0.967    | 0.967      | 0.967 | 647.15 |
|                       | 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  | 94.73    | 0.0138 | 93.86     | 0.0170 | 0.927    | 0.926      | 0.938 | 93.93    | 0.0189 | 0.929    | 0.926      | 0.939 | 641.34 |
|                       | 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 |
|                       | Alfa  | Doğruluk | CE     | Doğruluk  | CE     | Kesinlik | Duyarlılık | F1    | Doğruluk | CE     | Kesinlik | Duyarlılık | F1    | Süre   |
| 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 | 97.68     | 0.1121 | 0.976    | 0.976      | 0.977 | 97.79    | 0.1200 | 0.977    | 0.976      | 0.978 | 651.23 |
|                       | 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  | 10.84    | 2.3023 | 9.84      | 2.3023 | 0.010    | 0.100      | 0.018 | 10.31    | 2.3032 | 0.010    | 0.100      | 0.019 | 626.82 |
|                       | 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 |



**Şekil 1.** MNIST veri setinin sınıflandırılması amacıyla Sigmoid aktivasyon fonksiyonu ile MSE ve CE hata fonksiyonlarının kullanıldığı KarıcıFANN modellerinde,  $\alpha = 0.73$  için elde edilen grafikler.



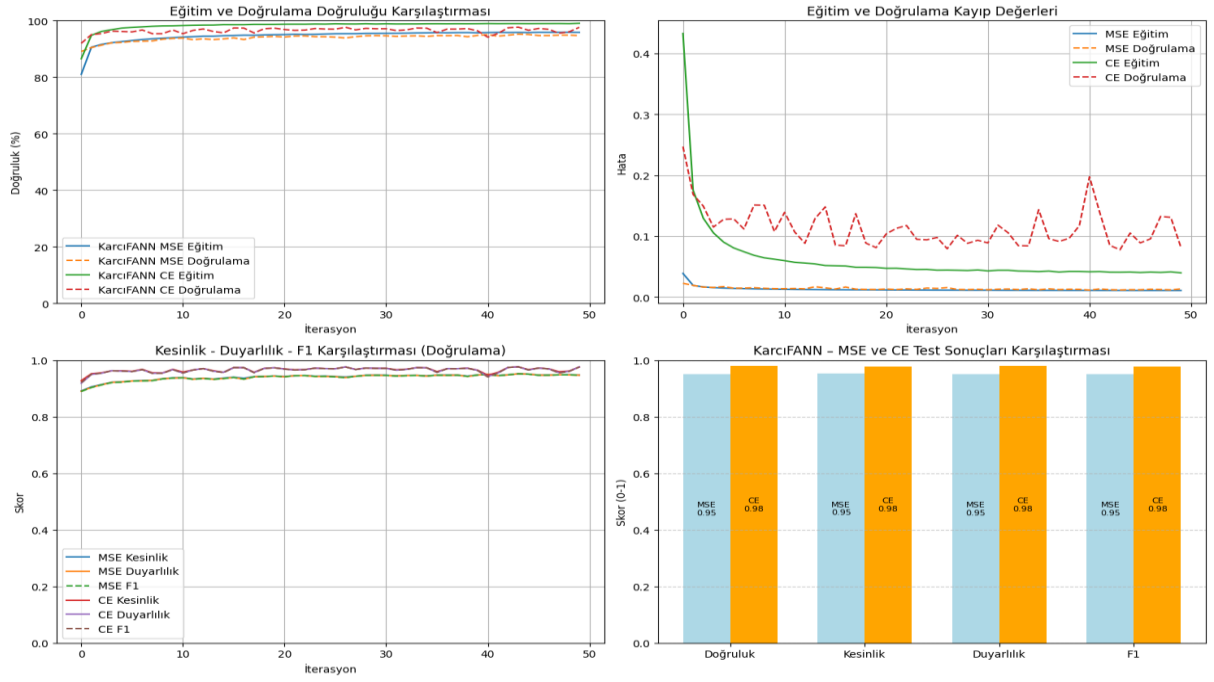
**Şekil 2.** MNIST veri setinin sınıflandırılması amacıyla ReLU aktivasyon fonksiyonu ile MSE ve CE hata fonksiyonlarının kullanıldığı KarıcıFANN modellerinde,  $\alpha = 0.73$  için elde edilen grafikler.

**Tablo 2.** MNIST veri setinin sınıflandırılması amacıyla kullanılan ve ağırlık sönümlemenin uygulandığı KarcıFANN modelleriyle elde edilen doğruluk, MSE, CE, kesinlik, duyarlılık ve F1 skoru değerleri.

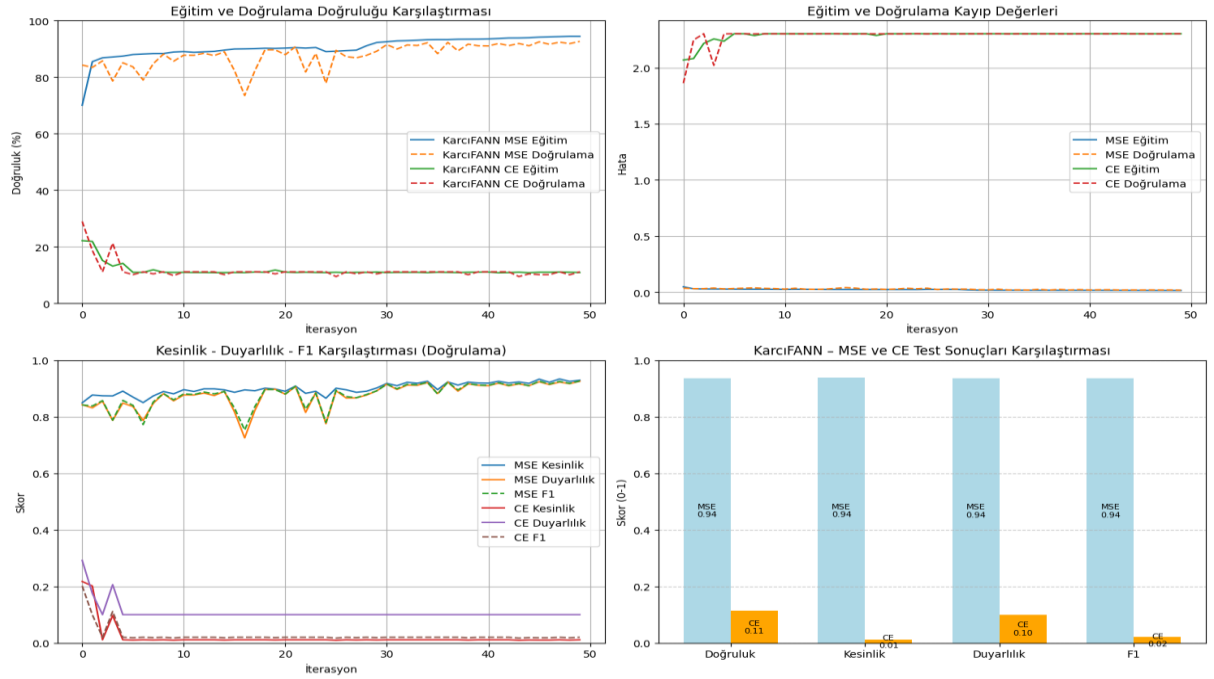
| Aktivasyon Fonksiyonu | Alfa  | Eğitim   |        | Doğrulama |        |          |            |       | Test     |        |          |            |       |        |
|-----------------------|-------|----------|--------|-----------|--------|----------|------------|-------|----------|--------|----------|------------|-------|--------|
|                       |       | Doğruluk | MSE    | Doğruluk  | MSE    | Kesinlik | Duyarlılık | F1    | Doğruluk | MSE    | Kesinlik | Duyarlılık | F1    | Süre   |
| Sigmoid               | 0.005 | 93.82    | 0.0153 | 84.86     | 0.0352 | 0.873    | 0.849      | 0.845 | 86.93    | 0.0343 | 0.889    | 0.868      | 0.865 | 686.41 |
|                       | 0.05  | 94.29    | 0.0146 | 93.14     | 0.0167 | 0.935    | 0.931      | 0.931 | 93.60    | 0.0159 | 0.938    | 0.935      | 0.935 | 674.88 |
|                       | 0.55  | 95.47    | 0.0116 | 94.98     | 0.0121 | 0.949    | 0.949      | 0.949 | 95.10    | 0.0122 | 0.951    | 0.951      | 0.950 | 664.20 |
|                       | 0.73  | 95.83    | 0.0109 | 95.83     | 0.0135 | 0.952    | 0.951      | 0.947 | 95.13    | 0.0122 | 0.955    | 0.954      | 0.951 | 626.20 |
|                       | 0.85  | 95.11    | 0.0119 | 93.80     | 0.0155 | 0.940    | 0.937      | 0.938 | 94.02    | 0.0153 | 0.943    | 0.939      | 0.940 | 666.99 |
|                       | 1.1   | 94.65    | 0.0124 | 94.12     | 0.0128 | 0.942    | 0.941      | 0.941 | 94.40    | 0.0126 | 0.945    | 0.944      | 0.944 | 670.08 |
|                       | 1.25  | 94.07    | 0.0131 | 94.14     | 0.0130 | 0.943    | 0.941      | 0.941 | 93.66    | 0.0131 | 0.938    | 0.936      | 0.936 | 676.04 |
| ReLU                  | 0.005 | 88.31    | 0.0308 | 85.28     | 0.0355 | 0.869    | 0.855      | 0.851 | 85.11    | 0.0352 | 0.866    | 0.855      | 0.851 | 670.99 |
|                       | 0.05  | 91.36    | 0.0265 | 88.36     | 0.0295 | 0.898    | 0.886      | 0.882 | 88.64    | 0.0294 | 0.900    | 0.888      | 0.884 | 670.45 |
|                       | 0.55  | 96.11    | 0.0124 | 94.94     | 0.0160 | 0.950    | 0.949      | 0.949 | 94.80    | 0.0156 | 0.950    | 0.946      | 0.947 | 669.84 |
|                       | 0.73  | 94.45    | 0.0149 | 92.64     | 0.0171 | 0.941    | 0.940      | 0.927 | 93.71    | 0.0149 | 0.946    | 0.945      | 0.945 | 633.39 |
|                       | 0.85  | 87.49    | 0.0267 | 82.20     | 0.0408 | 0.868    | 0.823      | 0.831 | 83.04    | 0.0405 | 0.872    | 0.831      | 0.839 | 666.56 |
|                       | 1.1   | 29.05    | 0.0742 | 30.00     | 0.0745 | 0.195    | 0.285      | 0.200 | 30.25    | 0.0742 | 0.195    | 0.290      | 0.202 | 648.11 |
|                       | 1.25  | 9.86     | -      | 9.96      | -      | 0.010    | 0.100      | 0.018 | 9.80     | -      | 0.010    | 0.100      | 0.018 | 700.23 |
|                       | Alfa  | Doğruluk | CE     | Doğruluk  | CE     | Kesinlik | Duyarlılık | F1    | Doğruluk | CE     | Kesinlik | Duyarlılık | F1    | Süre   |
| Sigmoid               | 0.005 | 97.58    | 0.0861 | 94.04     | 0.1959 | 0.952    | 0.940      | 0.943 | 94.35    | 0.1816 | 0.953    | 0.942      | 0.945 | 643.62 |
|                       | 0.05  | 97.76    | 0.0793 | 96.62     | 0.1104 | 0.966    | 0.966      | 0.966 | 96.96    | 0.0951 | 0.970    | 0.969      | 0.969 | 631.53 |
|                       | 0.55  | 98.87    | 0.0459 | 97.44     | 0.0806 | 0.974    | 0.974      | 0.974 | 97.79    | 0.0712 | 0.978    | 0.978      | 0.978 | 639.63 |
|                       | 0.73  | 99.05    | 0.0398 | 97.62     | 0.0814 | 0.978    | 0.978      | 0.976 | 97.89    | 0.0649 | 0.979    | 0.979      | 0.979 | 626.69 |
|                       | 0.85  | 98.92    | 0.0414 | 96.90     | 0.0933 | 0.969    | 0.969      | 0.969 | 97.29    | 0.0890 | 0.973    | 0.973      | 0.973 | 640.21 |
|                       | 1.1   | 97.03    | 0.0964 | 96.16     | 0.1296 | 0.962    | 0.961      | 0.961 | 96.53    | 0.1159 | 0.965    | 0.965      | 0.965 | 642.56 |
|                       | 1.25  | 90.97    | 0.3090 | 92.22     | 0.2755 | 0.924    | 0.920      | 0.921 | 92.33    | 0.2586 | 0.925    | 0.922      | 0.922 | 653.72 |
| ReLU                  | 0.005 | 96.03    | 0.1368 | 95.64     | 0.1445 | 0.956    | 0.956      | 0.956 | 95.82    | 0.1367 | 0.958    | 0.958      | 0.958 | 645.23 |
|                       | 0.05  | 96.21    | 0.1273 | 95.18     | 0.1758 | 0.953    | 0.952      | 0.952 | 95.79    | 0.1447 | 0.959    | 0.958      | 0.958 | 647.15 |
|                       | 0.55  | 96.43    | 0.1220 | 95.10     | 0.2080 | 0.952    | 0.951      | 0.950 | 95.02    | 0.1922 | 0.952    | 0.951      | 0.950 | 644.76 |
|                       | 0.73  | 10.89    | 2.3023 | 11.14     | 2.3023 | 0.011    | 0.100      | 0.020 | 11.35    | 2.3022 | 0.011    | 0.100      | 0.020 | 629.08 |
|                       | 0.85  | 10.73    | 2.3029 | 11.06     | 2.3030 | 0.011    | 0.100      | 0.020 | 11.35    | 2.3030 | 0.011    | 0.100      | 0.020 | 642.27 |
|                       | 1.1   | 10.54    | 2.3051 | 10.02     | 2.3114 | 0.010    | 0.100      | 0.018 | 10.09    | 2.3095 | 0.010    | 0.100      | 0.018 | 643.11 |
|                       | 1.25  | 10.58    | 2.5893 | 10.72     | 2.3074 | 0.011    | 0.100      | 0.019 | 10.28    | 2.3089 | 0.010    | 0.100      | 0.019 | 743.64 |



Grafiksel sonuçları Şekil 1 ve Şekil 2 ile gösterilen KarcıFANN modellerine ağırlık sönümleme hiperparametresinin eklenmesiyle elde edilen grafikler sırasıyla Şekil 3 ve Şekil 4 ile gösterilmektedir.



**Şekil 3.** MNIST veri setinin sınıflandırılması amacıyla Sigmoid aktivasyon fonksiyonu, MSE ve CE hata fonksiyonları ile ağırlık sönümlemenin kullanıldığı KarcıFANN modellerinde,  $\alpha = 0.73$  için elde edilen grafikler.



**Şekil 4.** MNIST veri setinin sınıflandırılması amacıyla ReLU aktivasyon fonksiyonu, MSE ve CE hata fonksiyonları ile ağırlık sönümlemenin kullanıldığı KarcıFANN modellerinde,  $\alpha = 0.73$  için elde edilen grafikler.

KaracıFANN modellerinde ağırlık sönümlleme işleminin analizinde ikinci veri seti olarak Dry Bean kullanılmıştır. Üç katmandan oluşan ağırlık sönümllemesiz modellerin Sigmoid ve ReLU aktivasyon fonksiyonları, MSE ve CE hata fonksiyonları, doğruluk, duyarlılık, kesinlik ve F1 skoru değerleri ile modellerin eğitim süreleri Tablo 3'te kaydedilmektedir. İlgili tabloda, MSE fonksiyonunun kullanıldığı modellerde Sigmoid ve ReLU fonksiyonlarının 1.1 ve 1.25 dışındaki alfa değerlerinde yakın sonuçlar ürettikleri görülmektedir. Ayrıca, Sigmoid fonksiyonunu içeren model, 1.1 ve 1.25 alfa değerlerinde, ReLU fonksiyonunu içeren model ise 1.25 alfa değerinde öğrenememektedir. CE fonksiyonunun kullanıldığı modellerde ReLU kullanımı, Sigmoid kullanımına kıyasla daha yüksek performans sağlamaktadır. Ayrıca, Sigmoid fonksiyonunun kullanıldığı model, alfa'nın 0.85, 1.1 ve 1.25 değerleri için öğrenme gerçekleştirilememektedir.

Sigmoid ve MSE fonksiyonları ile tasarlanan modelin genel performansı, 0.73, 1.1 ve 1.25 dışındaki alfa değerlerinde Sigmoid ve CE fonksiyonları ile tasarlanan modele kıyasla daha yüksektir. Diğer taraftan, ReLU ve CE fonksiyonlarının kullanıldığı model, ReLU ve MSE fonksiyonlarının kullanıldığı modelden genel olarak daha başarılı sonuçlar üretmektedir.

Tablo 3 ile gösterilen 0.73 alfa değeri için Sigmoid ve ReLU aktivasyon fonksiyonları ile MSE ve CE hata fonksiyonlarının kullanıldığı modellerin eğitilmesiyle elde edilen grafiksel sonuçlar Şekil 3 ve Şekil 4 ile sunulmaktadır.

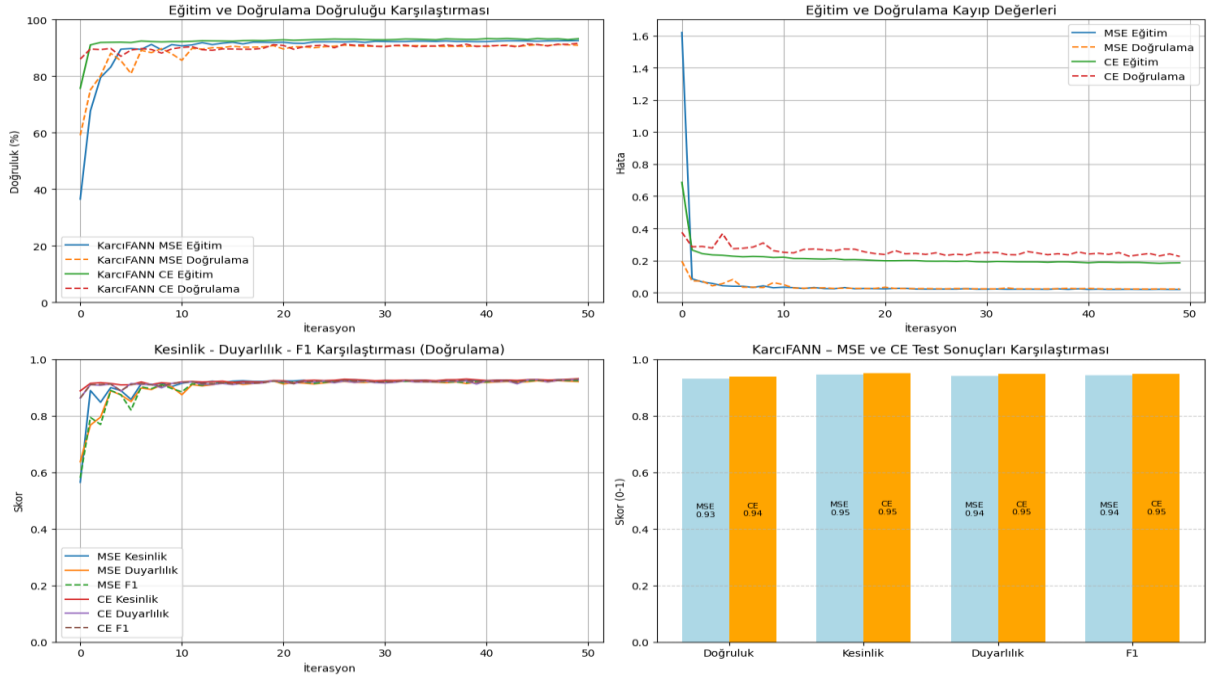
Dry Bean veri setinin sınıflandırıldığı modellere ağırlık sönümlleme işlemi uygulanarak elde edilen sonuçlar Tablo 4 ile raporlanmaktadır. Bu tablo, MSE ve CE fonksiyonlarını içeren modellerde ReLU fonksiyonu kullanıldığında, Sigmoid fonksiyonuna kıyasla genel olarak daha iyi sonuçlar üretildiğini göstermektedir. Sigmoid ve CE fonksiyonlarıyla tasarlanan model, tüm aşamalarındaki hata değerleri dışında Sigmoid ve MSE fonksiyonlarıyla tasarlanan modelden genel olarak daha yüksek performans sergilemektedir. Benzer şekilde, ReLU fonksiyonunun CE fonksiyonu ile birlikte kullanıldığı model, MSE fonksiyonu ile kullanıldığı modele kıyasla genel olarak daha yüksek doğruluk, kesinlik, duyarlılık ve F1 skoru değerleri üretmektedir.

Tablo 3 ile gösterilen ağırlık sönümllemesiz KaracıFANN modellerinin sonuçları ile bu modellere ağırlık sönümlleme işlemi uygulanarak elde edilen sonuçların sunulduğu Tablo 4 incelendiğinde, birçok alfa değeri için ağırlık sönümlleme işleminin farklı metrikler için performansı artırdığı görülmektedir.

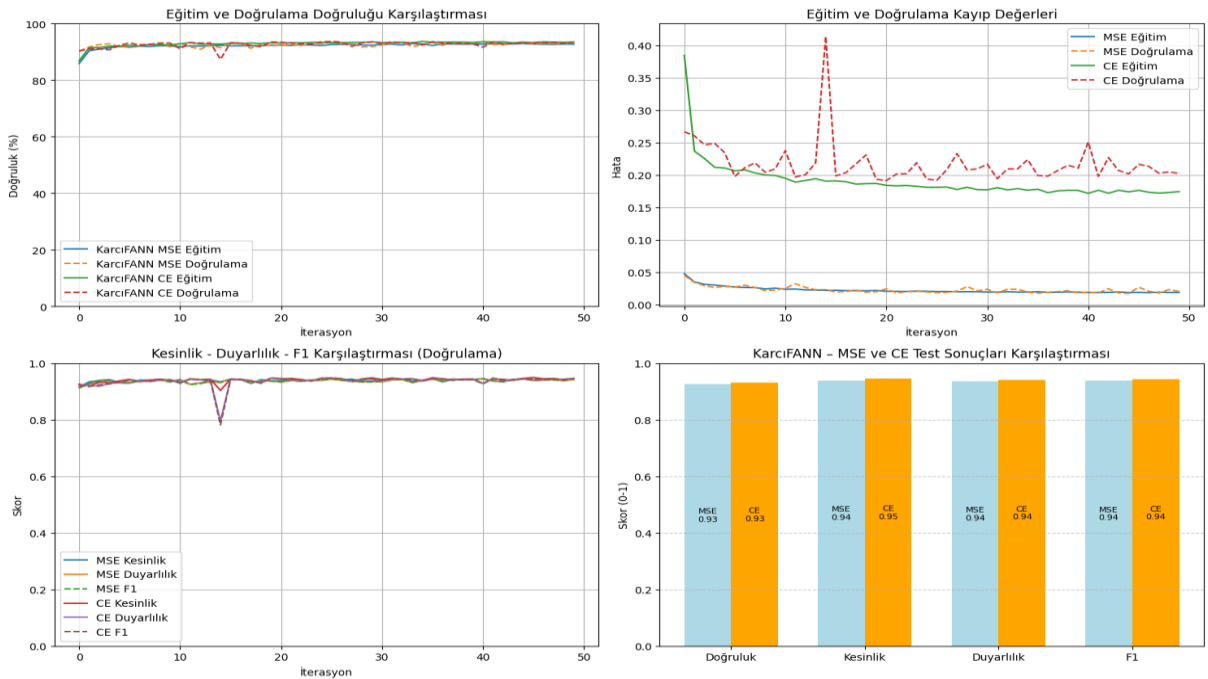
Dry Bean veri setinin sınıflandırılması amacıyla kullanılan ve grafiksel sonuçları Şekil 5 ve Şekil 6 ile gösterilen KaracıFANN modellerine ağırlık sönümlleme parametresinin eklenmesiyle elde edilen grafikler sırasıyla Şekil 7 ve Şekil 8 ile sunulmaktadır.

**Tablo 3.** Dry Bean veri setinin sınıflandırılması amacıyla kullanılan KarcıFANN modellerinin doğruluk, MSE, CE, kesinlik, duyarlılık ve F1 skoru değerleri.

| Aktivasyon Fonksiyonu | Alfa  | Eğitim   |        | Doğrulama |        |          |            |       | Test     |        |          |            |       |       |
|-----------------------|-------|----------|--------|-----------|--------|----------|------------|-------|----------|--------|----------|------------|-------|-------|
|                       |       | Doğruluk | MSE    | Doğruluk  | MSE    | Kesinlik | Duyarlılık | F1    | Doğruluk | MSE    | Kesinlik | Duyarlılık | F1    | Süre  |
| 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  | 92.56    | 0.0193 | 90.92     | 0.0223 | 0.927    | 0.921      | 0.924 | 93.20    | 0.0185 | 0.946    | 0.940      | 0.943 | 12.17 |
|                       | 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  | 92.73    | 0.0186 | 93.36     | 0.0200 | 0.945    | 0.943      | 0.944 | 92.52    | 0.0225 | 0.940    | 0.936      | 0.938 | 12.56 |
|                       | 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 |
| Aktivasyon Fonksiyonu | Alfa  | Doğruluk | CE     | Doğruluk  | CE     | Kesinlik | Duyarlılık | F1    | Doğruluk | CE     | Kesinlik | Duyarlılık | F1    | Süre  |
|                       |       | Doğruluk | CE     | Doğruluk  | CE     | Kesinlik | Duyarlılık | F1    | Doğruluk | CE     | Kesinlik | Duyarlılık | F1    | Süre  |
| 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  | 93.22    | 0.1853 | 91.60     | 0.2248 | 0.931    | 0.928      | 0.929 | 93.84    | 0.1757 | 0.951    | 0.948      | 0.949 | 11.43 |
|                       | 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    | 0.0246 | 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  | 93.46    | 0.1742 | 93.52     | 0.2020 | 0.947    | 0.946      | 0.946 | 93.08    | 0.0239 | 0.946    | 0.942      | 0.944 | 11.37 |
|                       | 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 |



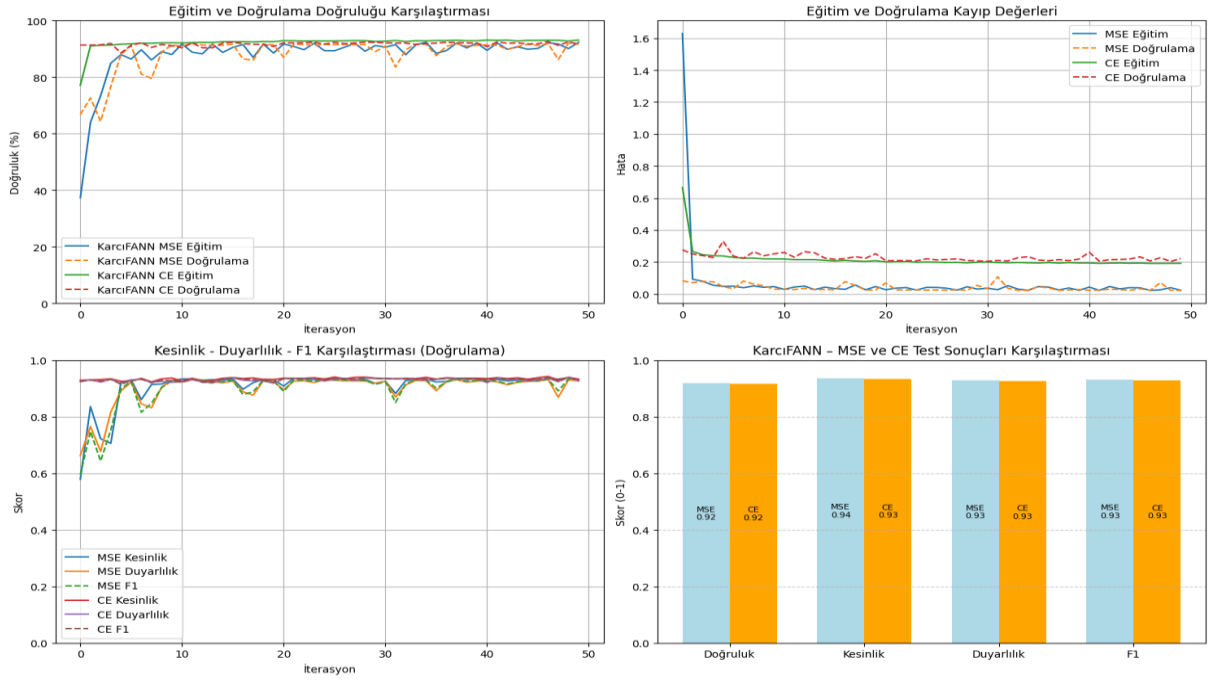
**Şekil 5.** Dry Bean veri setinin sınıflandırılması amacıyla Sigmoid aktivasyon fonksiyonu ile MSE ve CE hata fonksiyonlarının kullanıldığı KarcıFANN modellerinde,  $\alpha = 0.73$  için elde edilen grafikler.



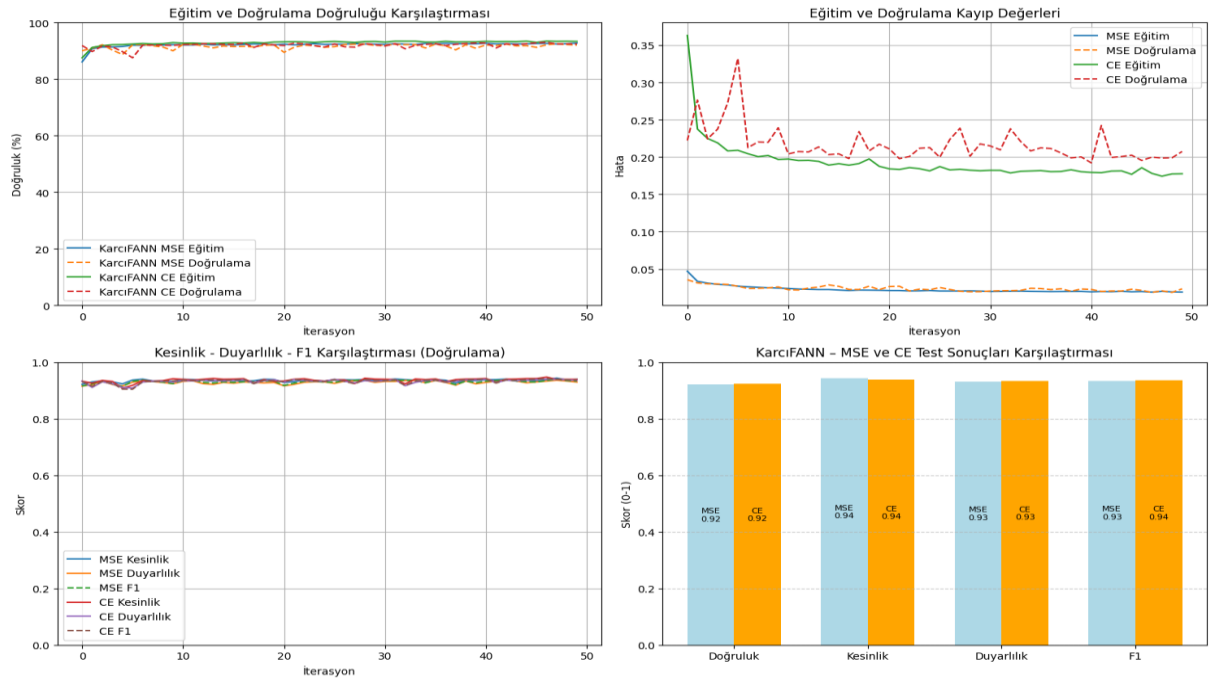
**Şekil 6.** Dry Bean veri setinin sınıflandırılması amacıyla ReLU aktivasyon fonksiyonu ile MSE ve CE hata fonksiyonlarının kullanıldığı KarcıFANN modellerinde,  $\alpha = 0.73$  için elde edilen grafikler.

**Tablo 4.** Dry Bean veri setinin sınıflandırılması amacıyla kullanılan ve ağırlık sönümlemenin uygulandığı KarcıFANN modellerinin doğruluk, MSE, CE, kesinlik, duyarlılık ve F1 skoru değerleri.

| Aktivasyon Fonksiyonu | Alfa  | Eğitim   |        | Doğrulama |        |          |            |       | Test     |        |          |            |       |       |
|-----------------------|-------|----------|--------|-----------|--------|----------|------------|-------|----------|--------|----------|------------|-------|-------|
|                       |       | Doğruluk | MSE    | Doğruluk  | MSE    | Kesinlik | Duyarlılık | F1    | Doğruluk | MSE    | Kesinlik | Duyarlılık | F1    | Süre  |
| Sigmoid               | 0.005 | 87.85    | 0.0416 | 89.00     | 0.0352 | 0.922    | 0.906      | 0.907 | 88.64    | 0.0355 | 0.916    | 0.896      | 0.900 | 14.02 |
|                       | 0.05  | 88.82    | 0.0360 | 88.48     | 0.0339 | 0.915    | 0.893      | 0.898 | 89.28    | 0.0341 | 0.923    | 0.897      | 0.906 | 13.18 |
|                       | 0.55  | 88.90    | 0.0416 | 92.52     | 0.0219 | 0.939    | 0.934      | 0.936 | 91.40    | 0.0229 | 0.929    | 0.924      | 0.926 | 12.98 |
|                       | 0.73  | 91.46    | 0.0320 | 90.64     | 0.0440 | 0.925    | 0.921      | 0.920 | 90.92    | 0.0434 | 0.929    | 0.923      | 0.922 | 14.36 |
|                       | 0.85  | 91.52    | 0.0274 | 93.12     | 0.0209 | 0.946    | 0.941      | 0.943 | 91.88    | 0.0236 | 0.937    | 0.935      | 0.936 | 15.45 |
|                       | 1.1   | 9.81     | -      | 9.04      | -      | 0.013    | 0.143      | 0.024 | 10.04    | -      | 0.014    | 0.143      | 0.026 | 15.33 |
|                       | 1.25  | 10.06    | -      | 8.80      | -      | 0.013    | 0.143      | 0.023 | 9.08     | -      | 0.013    | 0.143      | 0.024 | 12.73 |
| ReLU                  | 0.005 | 90.79    | 0.0311 | 91.60     | 0.0289 | 0.929    | 0.927      | 0.928 | 92.12    | 0.0278 | 0.931    | 0.931      | 0.931 | 13.89 |
|                       | 0.05  | 90.93    | 0.0311 | 90.44     | 0.0286 | 0.926    | 0.917      | 0.918 | 90.36    | 0.0280 | 0.923    | 0.922      | 0.919 | 13.25 |
|                       | 0.55  | 92.41    | 0.0210 | 91.60     | 0.0265 | 0.934    | 0.924      | 0.928 | 91.92    | 0.0261 | 0.934    | 0.924      | 0.929 | 13.42 |
|                       | 0.73  | 92.64    | 0.0203 | 91.84     | 0.0200 | 0.932    | 0.931      | 0.931 | 92.08    | 0.0203 | 0.939    | 0.936      | 0.937 | 12.71 |
|                       | 0.85  | 92.52    | 0.0195 | 93.16     | 0.0190 | 0.945    | 0.941      | 0.943 | 93.16    | 0.0191 | 0.945    | 0.940      | 0.942 | 14.61 |
|                       | 1.1   | 93.28    | 0.0179 | 91.64     | 0.0233 | 0.934    | 0.930      | 0.931 | 91.76    | 0.0224 | 0.937    | 0.935      | 0.935 | 14.02 |
|                       | 1.25  | 93.00    | 0.0178 | 93.12     | 0.0170 | 0.943    | 0.941      | 0.941 | 92.28    | 0.0178 | 0.941    | 0.935      | 0.937 | 12.87 |
|                       | Alfa  | Doğruluk | CE     | Doğruluk  | CE     | Kesinlik | Duyarlılık | F1    | Doğruluk | CE     | Kesinlik | Duyarlılık | F1    | Süre  |
| Sigmoid               | 0.005 | 92.59    | 0.2010 | 90.68     | 0.2609 | 0.923    | 0.929      | 0.925 | 90.76    | 0.2336 | 0.919    | 0.928      | 0.922 | 13.73 |
|                       | 0.05  | 92.80    | 0.2000 | 92.12     | 0.2116 | 0.934    | 0.938      | 0.936 | 92.44    | 0.2105 | 0.933    | 0.935      | 0.934 | 14.64 |
|                       | 0.55  | 92.55    | 0.1931 | 92.72     | 0.2024 | 0.939    | 0.936      | 0.937 | 92.16    | 0.2086 | 0.936    | 0.929      | 0.931 | 12.76 |
|                       | 0.73  | 92.80    | 0.1958 | 92.48     | 0.2142 | 0.943    | 0.938      | 0.939 | 90.00    | 0.2902 | 0.924    | 0.923      | 0.921 | 11.23 |
|                       | 0.85  | 92.96    | 0.1971 | 92.56     | 0.1986 | 0.940    | 0.939      | 0.938 | 90.52    | 0.2595 | 0.923    | 0.920      | 0.920 | 12.23 |
|                       | 1.1   | 93.08    | 0.1916 | 92.40     | 0.2152 | 0.941    | 0.932      | 0.936 | 92.00    | 0.2319 | 0.936    | 0.930      | 0.931 | 12.95 |
|                       | 1.25  | 92.79    | 0.1954 | 93.40     | 0.1805 | 0.942    | 0.943      | 0.943 | 91.84    | 0.2208 | 0.928    | 0.931      | 0.929 | 12.20 |
| ReLU                  | 0.005 | 92.71    | 0.1952 | 92.80     | 0.2086 | 0.936    | 0.940      | 0.938 | 91.68    | 0.2237 | 0.929    | 0.931      | 0.929 | 12.02 |
|                       | 0.05  | 92.94    | 0.1935 | 92.56     | 0.2029 | 0.940    | 0.937      | 0.937 | 93.36    | 0.1916 | 0.947    | 0.942      | 0.944 | 12.08 |
|                       | 0.55  | 93.15    | 0.1763 | 92.84     | 0.1936 | 0.941    | 0.942      | 0.941 | 93.28    | 0.1944 | 0.946    | 0.944      | 0.945 | 12.14 |
|                       | 0.73  | 92.94    | 0.1876 | 93.64     | 0.1702 | 0.947    | 0.945      | 0.945 | 92.92    | 0.1934 | 0.944    | 0.938      | 0.940 | 11.95 |
|                       | 0.85  | 93.35    | 0.1773 | 92.40     | 0.2106 | 0.938    | 0.937      | 0.936 | 92.12    | 0.2182 | 0.936    | 0.934      | 0.934 | 12.30 |
|                       | 1.1   | 93.49    | 0.1798 | 91.96     | 0.2120 | 0.935    | 0.928      | 0.930 | 93.12    | 0.2030 | 0.946    | 0.941      | 0.943 | 12.52 |
|                       | 1.25  | 92.92    | 0.1785 | 93.68     | 0.1805 | 0.951    | 0.944      | 0.947 | 92.64    | 0.2020 | 0.943    | 0.940      | 0.941 | 12.06 |



**Şekil 7.** Dry Bean veri setinin sınıflandırılması amacıyla Sigmoid aktivasyon fonksiyonu, MSE ve CE hata fonksiyonları ile ağırlık sönümlemenin kullanıldığı KarıcıFANN modellerinde,  $\alpha = 0.73$  için elde edilen grafikler.



**Şekil 8.** Dry Bean veri setinin sınıflandırılması amacıyla ReLU aktivasyon fonksiyonu, MSE ve CE hata fonksiyonları ile ağırlık sönümlemenin kullanıldığı KarıcıFANN modellerinde,  $\alpha = 0.73$  için elde edilen grafikler.

#### 4. Sonuç

Düzenleme yöntemleri, YSA'ların optimizasyonu sürecinde önemli bir yer tutmaktadır. Bu yöntemlerden biri olan ağırlık sönümleme işleminde, ağırlıkları belli oranlarda azaltılarak (cezalandırılarak) aşırı büyümelerin sebep olduğu problemlerin çözülmesi sağlanmaktadır.

Bu çalışmada, ağırlık sönümleme hiperparametresinin, KarcıFANN yönteminin başarımına etkileri analiz edilmiştir. MNIST ve Dry Bean veri setlerini sınıflandırmak amacıyla önce ağırlık sönümlemesiz KarcıFANN modelleri tasarlanmış, daha sonra bu modellere ağırlık sönümleme işlemi uygulanarak elde edilen sonuçlar karşılaştırılmıştır. Sigmoid ve ReLU aktivasyon fonksiyonlarıyla birlikte, MSE ve CE hata fonksiyonları da kullanılarak model çeşitliliği artırılmıştır. Her modelin eğitim performansını değerlendirmek için MSE ve CE hata fonksiyonları ile F1 skoru, kesinlik, doğruluk ve duyarlılık metrikleri kullanılmıştır.

Elde edilen bulgular, birçok alfa değerinde ağırlık sönümleme hiperparametresinin kullanılmasının, KarcıFANN modellerinin başarımını anlamlı düzeyde iyileştirdiğini ortaya koymaktadır. Bu sonuç, henüz optimize edilmemiş KarcıFANN yönteminin, uygun düzenleme yöntemleriyle geliştirilmeye açık olduğunu göstermektedir. Gelecekteki çalışmalarda, daha çeşitli KarcıFANN mimarilerinde, ağırlık sönümleme ve diğer düzenleme tekniklerinin uygulanmasıyla daha yüksek başarımlar elde edilebileceği öngörülmektedir.

#### Kaynaklar

- 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.
- Galanti, T., Siegel, Z. S., Gupte, A., & Poggio, T. A. (2025, March). SGD with weight decay secretly minimizes the ranks of your neural networks. *In The Second Conference on Parsimony and Learning (CPAL 2025 Proceedings Track)*.
- 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.
- Karakurt, M., Hark, H., Erdoğan, M. C., & Karcı, A. (2022). Karcı Sinir Ağlarının Uygulanması ve Performans Analizi. *Computer Science*, 7(2), 68-80.
- 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., 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.
- 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.
- Karcı, A. (2015c). Chain rule for fractional order derivative. *Science Innovation*, Vol:3, pp:63-67.
- Karcı, 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 Karci'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.
- Kosson, P., Messmer, P., & Jaggi, M. (2023). Rotational equilibrium: How weight decay balances learning across neural networks. *arXiv preprint arXiv:2305.17212*.
- 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>.
- Li, Z., Lyu, K., & Arora, S. (2020). Reconciling modern deep learning with traditional optimization analyses: The intrinsic learning rate. *Advances in Neural Information Processing Systems*, 33, 14544-14555.
- Loshchilov, I., & Hutter, F. (2017). Decoupled weight decay regularization. *International Conference on Learning Representations 2019*. *arXiv preprint arXiv:1711.05101*.
- Outmezguine, N. J., & Levi, N. (2024). Decoupled weight decay for any p norm. *arXiv preprint arXiv:2404.10824*.
- 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. & Karci, A. (2024). Comparison of Loss Functions in the KarciFANN Method. In *2024 8th International Artificial Intelligence and Data Processing Symposium (IDAP)*, (pp. 1-5). IEEE.
- Saygılı, H., Karakurt, M. & Karci, A. (baskıda). Karci kesir dereceli yapay sinir ağı (KarciFANN): 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.
- Xie, Z., Sato, I., & Sugiyama, M. (2020). Understanding and scheduling weight decay. Retrieved from [https://openreview.net/forum?id=J7V\\_4aauV6B](https://openreview.net/forum?id=J7V_4aauV6B)
- Van Laarhoven, T. (2017). L2 regularization versus batch and weight normalization. *arXiv preprint arXiv:1706.05350*. <https://doi.org/10.48550/arXiv.1706.05350>
- Zhang, G., Wang, C., Xu, B., & Grosse, R. (2018). Three Mechanisms of Weight Decay Regularization. *arXiv preprint arXiv:1810.12281*.