EĞİTİM FONKSİYONU VE NÖRON SAYISININ YAPAY SİNİR AĞI PERFORMANSINA ETKİSİ: OPERASYON BAZLI TEZGÂH ATAMA PROBLEMİ ÖRNEĞİ

Burcu ADIGÜZEL[1], Yunus KAYIR[2]

[1-2]Gazi Üniversitesi, Teknoloji Fakültesi, Ankara / Türkiye

Öz: Günümüzde Yapay Sinir Ağları pek çok farklı disiplinde, geniş uygulama alanına sahiptir. Ancak bu kara kutu halen analitik olarak açıklanamayan bileşenler içermektedir. Bunlardan biri veri setinin yapısına ve uygulayıcının tecrübesine bağlı olan "eğitim fonksiyonu"dur. Bir diğeri ise ağ yapısındaki tüm bileşenlerle ilişkili olan gizli katman "nöron sayısı"dır. Bu çalışmada, operasyon bazlı tezgâh atama probleminin çözümü için oluşturulan bir yapay sinir ağı, farklı eğitim fonksiyonları ve gizli katman nöron sayılarıyla eğitilmiş ve bu iki faktörün ağ performansı üzerindeki etkisi değerlendirilmiştir. Değerlendirmeler sonucunda, tek gizli katmanlı bir ağın eğitimde, gizli katman nöron sayısını arttırmanın öğrenmeyi olumlu yönde etkilediği ancak eşik değerin aşıldığı durumlarda, eğitim fonksiyonuna bağlı olarak, negatif etki yaratabileceği belirlenmiştir. Seçilen eğitim fonksiyonu türüne göre değişmekle birlikte, tekrarlı denemelerin ağ performansını artırabildiği, eğitim fonksiyonu seçiminde veri seti yapısının ve ağın kullanım amacının dikkate alınması gerektiği görülmüştür.

Anahtar Kelimeler: Yapay Sinir Ağları, Eğitim Fonksiyonu, Nöron Sayısı

GİRİŞ

Çağımız teknolojisinin geldiği noktada, matematiksel formülasyonu kurulamayan ve çözülmesi mümkün olmayan problemler, "yapay zekâ" olarak bilinen sezgisel yöntemler yolu ile bilgisayarlar tarafından çözülebilmektedir. Günümüzde karşılaşılan problemlerin karmaşıklığı ve değişkenliği karşısında yapay zekâ; öğrenme, analiz etme, sınıflandırma, optimizasyon, problem çözme ve karar verme becerileri ile hayatımızın her alanında uygulanabilecek çözümler sağlamaktadır.

Yapay zekâ biliminin bir alt dalı olarak ele alınan Yapay Sinir Ağları (YSA), "makine öğrenmesi" olarak bilinen ve bilgisayarların olayları öğrenmesini sağlayan yaklaşımlardan biridir (Öztemel, 2012). Örneklerden öğrenmeye dayanan bu yaklaşım biyolojik sinir ağı yapısını taklit etmektedir. Nasıl ki insan, tecrübelerine dayanarak yeni kararlar alabiliyorsa, bir yapay sinir ağı da öğrendiklerinden yola çıkarak daha önce karşılaşmadığı durumlar karşısında karar verebilmektedir. Bunun için sinir ağındaki her bir sinir hücresi, kendisine sunulan bilgiyi bir diğer sinir hücresine iletir ve bu iletim ağ öğrenene kadar devam eder. Çalışma prensibi basitçe bu şekilde açıklanabilmesine rağmen YSA halen "kara kutu" olarak adlandırılmaktadır. Çünkü bir ağın sonuca ulaşmasında etkili olan tüm bileşenler (girdiler, ağırlıklar, toplama fonksiyonu, aktivasyon fonksiyonu, gizli katman sayısı, nöron sayısı, öğrenme algoritması) matematiksel olarak açıklanabilir olsa da bir yapay sinir ağının tüm bunları nasıl gerektiği şekilde bir araya getirerek bir çıktı elde ettiği bilinmemektedir.

Yapay Sinir Ağları, MathWorks firması tarafından 1992'de ilk kez MATLAB 4.0 yazılımına dahil bir modül olarak (Neural Network Toolbox) sunulduğundan beri daha pratik şekilde uygulanabilir hale gelmiş; doğrusal olmayan ve büyük veriler içeren karmaşık problemlerin çözümü için akademide ve endüstride geniş kullanım alanı bulmuştur (Anifowose vd., 2017). MATLAB Neural Network modülü, bir yapay sinir ağı oluşturmak için matematiksel formülleri literatürde tanımlanmış tüm gerekli bileşenleri kullanıcıya sunmaktadır. Ağ oluşturulduktan sonra, yapılandırılması ve eğitilmesi gerekmektedir. Yapılandırmadan kasıt, örnek verilerin çözmek istediğimiz probleme uygun şekilde ayarlanmasıdır. Eğitim ise ağ içinde değiştirilebilir

olan ağırlık ve bias değerleriyle oynayarak ağın öğrenmesini sağlamaktır (Beale vd., 2021). Bu noktada karşımıza YSA'nın en temel sorusu çıkmaktadır: Bir yapay sinir ağı hangi algoritma ile eğitilmelidir? Bu sorunun cevabı veri setinin yapısı, ağdaki ağırlıklar, nöron sayıları, ağın kullanım amacı gibi pek çok etkene bağlı olduğundan hangi eğitim algoritmasının en iyi sonucu vereceğini belirlemek çok zordur. Bu çalışmada kullanılan 5 eğitim fonksiyonu dahil olmak üzere, MATLAB YSA modülünde yer alan eğitim fonksiyonları, kullandıkları algoritmalar ve öncelikli kullanım alanları Tablo 1'de sunulmuştur. [32]

Eğitim Fonksiyonu	Algoritması	Öncelikli Kullanım Amacı
trainlm	Levenberg-Marquardt	Regresyon
trainbfg	BFGS Quasi-Newton	Regresyon
trainbr	Bayesian Regularization	Örüntü Tanıma
trainrp	Resilient Backpropagation	Örüntü Tanıma
trainscg	Scaled Conjugate Gradient	Örüntü Tanıma
traincgb	Conjugate Gradient with Powell/Beale Restarts	Regresyon/Örüntü Tanıma
traincgf	Fletcher-Powell Conjugate Gradient	Regresyon/Örüntü Tanıma
traincgp	Polak-Ribiére Conjugate Gradient	Regresyon/Örüntü Tanıma
trainoss	One Step Secant	Regresyon/Örüntü Tanıma
traingdx	Variable Learning Rate Gradient Descent	Regresyon/Örüntü Tanıma
traingdm	Gradient Descent with Momentum	Regresyon/Örüntü Tanıma
traingd	Gradient Descent	Regresyon/Örüntü Tanıma

Tablo 1. MATLAB Eğitim fonksiyonları, algoritmaları ve öncelikli kullanım alanları

Bir yapay sinir ağı oluşturulurken kullanılacak eğitim fonksiyonu ile birlikte akla gelen bir diğer soru da gizli katman ve nöron sayılarının ne olması gerektiğidir. İleri beslemeli (feed forward) yapay sinir ağlarında tek gizli katmanın yeterli olduğu kanıtlanmış olmasına rağmen, teoride tek gizli katmanda yer alan nöronlarla çözülebilecek bir problem için pratikte çok daha fazla katmanda yer alan nöronlara gerek duyulabilmektedir (Bebis vd.,1994).

Gizli katman sayısında olduğu gibi, katmanlardaki nöron sayısının ne olması gerektiği konusunda da literatürde farklı yaklaşımlar mevcuttur (Xu vd., 2008). En iyi gizli nöron sayısına ulaşmak için ağ mimarisindeki birçok etkeni (örnek veri yapısı, girdi ve çıktı sayısı, eğitim algoritması, aktivasyon fonksiyonu gibi) bir arada değerlendirecek çok karmaşık bir yaklaşım gerekmektedir (Sarle, 2002).

Genelde incelenen çalışmalarda, gizli nöron sayısının, genel kabul görmüş 3 temel yaklaşıma dayanarak belirlendiği görülmüştür:

1. Gizli nöron sayısı, girdi katmanı ve çıktı katmanı arasında bir değer olmalıdır.
2. Gizli nöron sayısı girdi katmanın 2/3'ü değerinde olmalıdır.
3. Gizli nöron sayısı girdi katmanının 2 katından az olmalıdır (Panchal vd., 2011).

[32] https://www.mathworks.com/help/deeplearning/ug/choose-a-multilayer-neural-network-training-function.html

AMAÇ

Bu çalışmanın amacı, tek gizli katmanlı ileri beslemeli bir yapay sinir ağında "eğitim fonksiyonu" ve "nöron sayısı"nın, ağ performansı üzerindeki etkilerini ortaya koymaktır.

KAPSAM

Bu çalışma "Birleşik Yapay Sinir Ağları ile Dinamik Sanal İmalat Hücresi Oluşturma Problemi" başlıklı yüksek lisans tezinden üretilmiştir.

Bir talaşlı imalat atölyesinde, değişen talebe göre, parçaları parça ailelerine ve makineleri yeni gruplara atayarak dinamik şekilde çalışan sistemler "Dinamik Sanal Hücresel İmalat Sistemi" olarak bilinmektedir. Birleşik yapay sinir ağları ile sanal imalat hücrelerini oluşturmak için öncelikle ağ kurulmalı ve eğitilmelidir. Bu çalışma her bir imalat hücresi için bilinmesi gereken parça, tezgâh ve operasyon bilgilerinin derlenmesi, ardından yapay sinir ağının oluşturulması ve eğitilmesini kapsamaktadır.

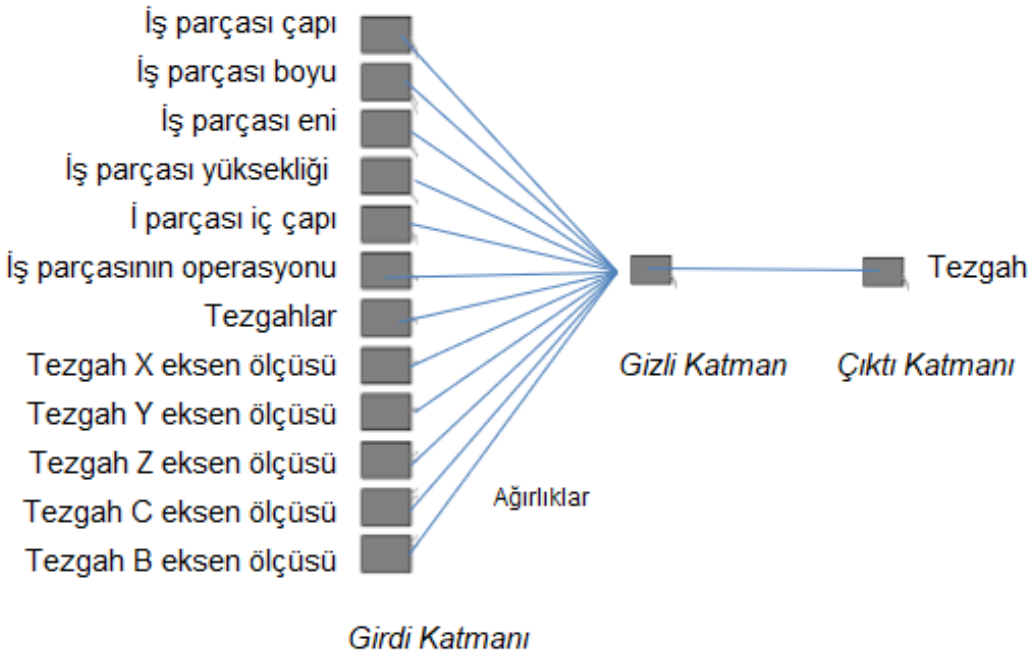
YÖNTEM

Çalışmada bir talaşlı imalat atölyesinde toplanan veriler, operasyon bazlı tezgâh atama problemine uygun şekilde derlenerek 13 sütun, 1639 satırdan oluşan veri seti hazırlanmıştır. Sütunlarda yer alan değerler sırasıyla parçanın çap, boy, en, yükseklik ve iç çap ölçülerini, parçanın talaşlı imalat operasyonunu, atölyede mevcut tezgâhları ve eksen ölçülerini ve son olarak ilgili operasyonun o tezgâhta yapılıp yapılamayacağı bilgisini içermektedir. Veri setinde yer alan örneklerden bir kısmı Tablo 2'de sunulmuştur. Tabloda görüldüğü üzere toplanan veriler, YSA'ya uygun şekilde kodlanmış, çoğaltılmış ve sıralanmıştır.

41014	42014	43000	44000	45000	10003	511	511440	511200	511130	511000	511000	11
41014	42014	43000	44000	45000	10003	510	510700	510420	510380	510200	510200	0
41014	42014	43000	44000	45000	10006	502	502610	502455	502510	502000	502000	2
41014	42014	43000	44000	45000	10006	501	501610	501305	501520	501000	501000	1
41010	42020	43000	44000	45000	10003	511	511440	511200	511130	511000	511000	11

Tablo 2. Veri setinde yer alan örnekler

Veri seti MATLAB 2021a programı Neural Network Toolbox'a, Şekil 1'de gösterildiği gibi, 12 girdi ve 1 çıktı (tezgâh) şeklinde sunulmuş ve oluşturulan tek katmanlı ağın eğitiminde 5 farklı eğitim fonksiyonu kullanılmıştır. Her bir eğitim, literatürde genel kabul görmüş yaklaşımlarla elde edilen 3 farklı gizli katman nöron sayısı ile yapılmıştır. Farklı koşullarda eğitilen ağın öğrenme performansları Ortalama Kare Hata (Mean Square Error) yöntemiyle karşılaştırılmıştır.



Şekil 1. Kurulan ağın yapısı

Oluşturulan yapay sinir ağından beklenti, verilen örneklerden yola çıkarak, iş parçasının hangi tezgâhta işlenebileceğini tahmin etmesidir. Bu doğrultuda ağın eğitiminde Levenberg-Marquardt (trainlm), BFGS Quasi-Newton (trainbfg), Resilient Backpropagation (trainrp), Scaled Conjugate Gradient (trainscg) ve Fletcher-Powell Conjugate Gradient (traincgf) algoritmaları kullanılmıştır.

Gizli katmandaki nöron sayıları giriş bölümünde açıklanan 3 temel yaklaşıma dayanarak 9, 12 ve 23 olarak belirlenmiştir. Yapılan deneme sayısı ve eşik değerin üstündeki gizli katman nöron sayısı uygulayıcının inisiyatifine bağlı olarak artırılmış ve farklı zamanlarda toplam 172 deneme yapılarak eğitim ve validasyon için korelasyon katsayı değerleri (R) kayıt altına alınmıştır.

Korelasyon katsayısı değeri -1 ile +1 arasında değişmekte ve ağın tahmin ettiği çıktı ile gerçek çıktıların ne kadar iyi eşleştiğini ifade etmektedir. Değerin +1'e yakın olması ağ performansının yüksek olduğunu göstermektedir.

BULGULAR

Her bir eğitim fonksiyonu için 4 farklı nöron sayısı ile 7 tekrarlı deneme yapılmıştır. Tablo 3'te trainlm eğitim fonksiyonu için elde edilen sonuçlar verilmiştir. Tekrarlı denemelerle performansın arttığı tek eğitim fonksiyonu olan trainlm ile yapılan denemelerde, çalışmanın en iyi performansı olan R=0,91 değerine, 23 nöron sayısında ve 7. tekrarda ulaşıldığı görülmektedir.

Gizli Katman Nöron Sayısı	Eğitim R degeri	Validasyon R degeri	Gizli Katman Nöron Sayısı	Eğitim R degeri	Validasyon R degeri
9	0,86	0,84	23	0,91	0,80
	0,82	0,82		0,87	0,87
	0,81	0,80		0,89	0,87
	0,89	0,83		0,91	0,82
	0,82	0,55		0,89	0,87
	0,85	0,89		0,90	0,89
	0,71	0,62		0,92	0,91
12	0,87	0,88	35	0,88	0,77
	0,90	0,89		0,88	0,88
	0,78	0,71		0,90	0,87
	0,88	0,84		0,84	0,82
	0,43	0,38		0,86	0,81
	0,79	0,85		0,89	0,72
	0,87	0,78		0,91	0,87

Tablo 3. Trainlm eğitim fonksiyonu ile elde edilen sonuçlar

Tablo 4'te tüm denemeler sonunda, her bir eğitim fonksiyonu için ilgili nöron sayısında elde edilen en iyi sonuçlar görülmektedir. Eğitim fonksiyonlarından trainlm hariç tümünde en yüksek ağ performansı, en yüksek nöron sayısı olan 35'te elde edilmiştir. Trainlm eğitim fonksiyonu, 12 girdi için eşik değer olan 24 nöron sayısının aşıldığı denemelerde R=0,91 değerini yakalayamamıştır. En iyi performansın trainlm eğitim fonksiyonunda elde edildiği, ardından sırasıyla trainbfg, trainscg, trainrp ve traincgf eğitim fonksiyonlarının geldiği görülmektedir.

Eğitim Fonksiyonu	Gizli Katman Nöron Sayısı	Eğitim R değeri	Validasyon R değeri
trainlm	9	0,85	0,89
trainlm	12	0,90	0,89
trainlm	***23***	0,92	***0,91***
trainlm	35	0,88	0,88
trainbfg	9	0,79	0,80
trainbfg	12	0,82	0,81
trainbfg	23	0,87	0,85
trainbfg	***35***	0,83	***0,89***
trainrp	9	0,65	0,66
trainrp	12	0,70	0,68
trainrp	23	0,59	0,68
trainrp	***35***	0,63	***0,79***
trainscg	9	0,67	0,66
trainscg	12	0,74	0,77
trainscg	23	0,77	0,78
trainscg	***35***	0,79	***0,81***
traincgf	9	0,53	0,55
traincgf	12	0,67	0,60
traincgf	23	0,6	0,69
traincgf	***35***	0,68	***0,75***

Tablo 4. Elde edilen sonuçlar

SONUÇ

Yapay sinir ağlarının dezavantajlarından olan, uygun ağ yapısının ve parametrelerinin belirlenmesinde deneme yanılma yolu kullanılması ve ağın gelişiminin kullanıcı tecrübesine bağlı olması bu çalışmanın kısıtlarını belirlemektedir. Bu çalışma kapsamında ele alınan problem YSA'da regresyon tipi bir problem olduğundan en iyi sonuçları trainlm ve trainbfg fonksiyonları vermiştir. Bu da oluşturulan veri setinin ve seçilen eğitim fonksiyonunun ağın kullanım amacına uygun olması gerektiğini göstermektedir. Çalışmada gizli katman nöron sayısı arttıkça genel olarak ağ performansı artmış olmasına rağmen, çok fazla sayıda nöronun kullanılması sebebiyle ağın aşırı öğrenme (overfitting) durumuna geçeceği unutulmamalıdır. Ağın her eğitimde farklı ağırlık değerleri atadığı göz önüne alınarak tekrarlı denemelerle performansın artması beklenebilir, ancak bu durum eğitim süresini uzatacaktır.

KAYNAKÇA

Öztemel, E., (2012), Yapay Sinir Ağları, 3.Baskı, Papatya Yayıncılık, İstanbul.

Anifowose, F., Labadin, J., & Abdulraheem, A. (2017). Towards an improved ensemble learning model of artificial neural networks: lessons learned on using randomized numbers of hidden neurons. In Artificial Intelligence: Concepts, Methodologies, Tools, and Applications. *IGI Global*.

Beale, M. H., Hagan, M. T., & Demuth, H. B. (2021). *Deep Learning Toolbox User's Guide R2021a* (2021). The MathWorks.

Sharma, B., & Venugopalan, K. J. I. J. C. E. (2014). Comparison of neural network training functions for hematoma classification in brain CT images. *IOSR Journal of Computer Engineering*, 16(1), 31-35.

Bebis, G., & Georgiopoulos, M. (1994). Feed-forward neural networks. *IEEE Potentials*, *13*(4), 27-31.

Xu, S., & Chen, L. (2008). A novel approach for determining the optimal number of hidden layer neurons for FNN's and its application in data mining.

Sarle, W. S., (2002), Neural Network FAQ. ftp.sas.com/pub/neural/FAQ.html.

Panchal, G., Ganatra, A., Kosta, Y. P., & Panchal, D. (2011). Behaviour analysis of multilayer perceptrons with multiple hidden neurons and hidden layers. *International Journal of Computer Theory and Engineering*, *3*(2), 332-337.

İNTERNET KAYNAKLARI

https://www.mathworks.com/help/deeplearning/ug/choose-a-multilayer-neural-network-training-function.html (E.T. 01.01.2021)