

Yükseköğretimde Öğrenci Başarılarının Sınıflandırılmasında Yapay Sinir Ağları ve Lojistik Regresyon Yöntemlerinin Kullanılması

The Usage of Artificial Neural Network and Logistic Regression Methods in the Classification of Student Achievement in Higher Education

Gülçin ÇIRAK*
Ömay ÇOKLUK**

Özet: Bu araştırmada, araştırmacı tarafından geliştirilen “Üniversite Öğrencilerinin Akademik Başarılarını Etkileyen Değişkenler Anketi” kullanılarak elde edilen bilgilerle farklı fakültelere ve programlara yeni kayıt yaptıran öğrencilerin gelecekteki başarılarının tahmin edilmesine olanak sağlayacak sınıflandırma modellerinin elde edilmesinde lojistik regresyon analizi ve yapay sinir ağları yöntemlerinin performanslarının karşılaştırılması amaçlanmıştır. Araştırmada öğrencilerin genel akademik başarı not ortalaması bağımlı değişken olarak alınmıştır. Sürekli bir değişken olan genel akademik başarı not ortalamaları iki kategorili bir süreksiz değişken haline getirilmiştir. İlişkisel tarama modelinde olan araştırmanın çalışma grubunu 2011-2012 Eğitim-Öğretim Yılı Bahar Dönemi’nde Ankara Üniversitesi Eğitim Bilimleri Fakültesi ile Dil ve Tarih-Coğrafya Fakültesi’nin bazı programlarında lisans öğrenimi gören toplam 419 üçüncü sınıf öğrencisi oluşturmaktadır. Araştırma sonucunda yapay sinir ağlarının lojistik regresyon analizine göre daha yüksek doğru sınıflandırma olasılığı sunduğu belirlenmiştir.

Anahtar sözcükler: Öğrenci Başarısı, Sınıflandırma, Tahmin, Yapay Sinir Ağları, Lojistik Regresyon

Abstract: In this study, according to the results of the survey conducted by the researcher entitled "Variables Which Affect the Success of University Students" artificial neural networks and logistic regression methods' performances were compared to indicate the estimated future success of the students who are registered to different faculties and programs. The overall academic grades of the students are taken as the dependent variable. The general academic achievement grade average is a permanent variable, it turns out to be a discontinuous variable. The group which was studied with the research which is a relationally browsing model consists of 419 students who were at their 3rd year in the 2011-2012 education and teaching year. This group were students of Ankara University in the departments of educational sciences and Language and History-Geography. According to this study, neural network analysis has a higher right classification probability when compared to logistic regression analysis.

Keywords: Student Achievement, Classification, Prediction, Artificial Neural Network, Logistic Regression

Başarı, istenen sonuca ulaşma, güdülen amaca erişme, isteneni elde etme olarak tanımlanabilir. Eğitim açısından düşündüğümüzde başarı; program hedefleriyle tutarlı davranışlar bütünüdür (Demirtaş, & Güneş, 2002). Bir başka deyişle, bir öğrenci belli bir programdaki hedef davranış-

* PhD., Uzman, Ankara Üniversitesi, EBF., Ölçme ve Değerlendirme ABD., Ankara, cirakgulcin@gmail.com

** Doç. Dr., Ankara Üniversitesi, EBF., Ölçme ve Değerlendirme ABD., Ankara, cokluk@education.ankara.edu.tr
Bu çalışma Gülçin Çırak tarafından Ankara Üniversitesi Eğitim Bilimleri Enstitüsü'nde Doç. Dr. Ömay Çokluk danışmanlığında hazırlanan ve Haziran-2012'de tamamlanan Yüksek Lisans Tezi'nin bir bölümüdür.

ları sergilemesi durumunda başarılı sayılabilir. Başarı, okul ortamında belirli bir ders ya da akademik programdan bireyin ne derece yararlandığının bir ölçüsü ya da göstergesidir. Okuldaki başarı ise bir akademik programdaki derslerden öğrencinin aldığı notların ya da puanların ortalaması olarak düşünülebilir (Özgüven, 1974). Eğitimde başarı denildiğinde, okulda dersler aracılığıyla geliştirilen ve öğretmenlerce takdir edilen notlarla, test puanlarıyla ya da her ikisi ile belirlenen beceriler veya kazanılan bilgilerin ifadesi olan “*akademik başarı*” kastedilmektedir (Carter, & Good, 1973).

Akademik başarı diye nitelendirilebilecek öğrenme ürünü, genel anlamı ile bireyin içinde tekrarlar ve enerji sarfı sonunda belli bir konu veya konularda sergilediği, anlamlandırdığı davranış değişikliğidir (Özoğlu, & Koç, 1995). Akademik başarı genellikle, öğrencinin psikomotor ve duyuşsal gelişiminin dışında kalan, bütün program alanlarındaki bilişsel davranış değişmelerini ifade eder (Ahmann, & Glock, 1967). Geleneksel kullanımda “*akademik başarı*” terimi, öğrencinin okula ilişkin son durumunun bazı terimlerle ve tekniklerle ifade edilmesidir. Bu genellikle tek bir ders için o derse ilişkin geçme notu, birden fazla ders için o derslerde alınan notların ortalaması, tüm dersler için ise ağırlıklı olarak hesaplanan ve genel akademik başarı not ortalaması olarak adlandırılan ortalamadır (Gülleroğlu, 2005).

Okulların insan davranışlarında istenilen değişimler meydana getirmede diğer kurumlardan üstün yönleri vardır; ancak farklı okullarda öğrenim gören öğrencilerin başarısı kaçınılmaz olarak farklı olabilmektedir (Marzano, Pickering, & Pollock, 2001). Bu farklılıkların yalnızca %10’u öğrencilerin gittiği okulların kalitesinden kaynaklanmaktadır. Başka bir deyişle, en iyi okula gitmekle, en kötü okula gitmek arasındaki fark, öğrenci başarısında yalnızca %10’luk bir değişime neden olmaktadır. Geriye kalan %90’lık kısmı etkileyen değişkenlerin öğrencinin doğal yetenek ve eğilimleri, sosyo-ekonomik durumu, ev ortamı vb. olduğu ifade edilmektedir (Akt. Sakacı, 2008). Araştırmalar sınıf ortamında eşit olduğu varsayılan öğrencilerin, aslında bilgiyi edinme yollarının farklı olduğunu ortaya koymaktadır. Bunun en iyi kanıtı, sınıfta aynı süreçlerden geçen öğrencilerin farklı başarı düzeylerine sahip olmasıdır. Sınıfta başarıyı ve daha da önemlisi öğrenmeyi etkileyen çeşitli değişkenler vardır (Saracaloğlu, & Yenice, 2009). Bu bağlamda, üniversite öğrencilerinin başarısızlığa uğramasının önlenmesi de, başarıyı etkileyen değişkenlerin belirlenmesine bağlıdır. Başarıyı etkileyen değişkenler bilinirse, başarısızlığa yol açan etkenler kontrol altına alınabilir ve öğrencilerin niteliklerine göre doğru bir sınıflandırma yapılması sağlanabilir.

Sınıflama amaçlı bazı yöntemler ve regresyon analizleri, önemli veri sınıflarını ortaya koyan veya gelecekteki eğilimlerini tahmin etmeyi sağlayan modeller kurabilen veri analiz yöntemleridir. Sınıflama, kategorik değerleri tahmin ederken, regresyon süreklilik gösteren değerlerin tahmin edilmesinde kullanılır (Kamber, & Morgan, 2000). Sınıflama, veri sınıfı ve kavramlarını tanımlama ve ayırt etmeyi sağlayan bir model kümesini bulma sürecidir. Sınıflama modelleri, sınıflar önceden incelenen veriler aracılığıyla oluşturulduğundan “*denetimli öğrenmeler*” olarak da ifade edilir (Aydın, 2002).

Bir nesnenin belirli sınıflar içerisinde hangisine ait olduğunu belirleyecek bir sınıflayıcı oluşturmak amacıyla farklı analiz teknikleri (lojistik regresyon analizi, diskriminant analizi, yapay sinir ağları vb.) kullanılmaktadır. Bu çalışmada yapay sinir ağları ve lojistik regresyon analizi yöntemleri ile sınıflandırma yapılmış olduğundan aşağıda kısaca bu yöntemlerin tanıtımına yer verilmiştir.

Yapay sinir ağları, insan beyninin özellikleri arasında yer alan; öğrenme yolu ile yeni bilgiler türetebilme, yeni bilgiler oluşturabilme ve keşfedebilme gibi yetenekleri herhangi bir yardım almadan otomatik olarak gerçekleştirmek amacıyla geliştirilen bilgisayar sistemleridir.

Yapay sinir ağları yöntemi, verilen örneklerin kümelendirilmesi ve belirli sınıflara ayrıştırılması ile daha sonra gelen örneklerin hangi sınıflara gireceğine karar verilmesini hedefler (Öztemel, 2003). Yapay sinir ağları tekniğinde n adet girdi değeri, kendi ağırlık değeri ile çarpılarak toplanır ve çıktığı elde etmek için aktivasyon fonksiyonu ile işlem yapılır (Gürsoy, 2009). Yapay sinir ağları ile eğitim alanında sınırlı sayıda uygulama yapılmış olmasına rağmen bu yöntemin başarı ile kullanıldığı birçok alan (ulaştırma ve havacılık, biomedikal ve ilaç sanayi, finans, borsa ve kredi kartı uygulamaları gibi) vardır. Başarılı uygulamalar incelendiğinde, yapay sinir ağlarının özellikle doğrusal olmayan, çok boyutlu, eksik, kusurlu ve hata olasılığı yüksek verilerin olduğu durumlar ile problemin çözümü için bir matematiksel modelin bulunmadığı durumlarda kullanıldığı belirtilebilir.

Lojistik regresyon analizi sınıflandırma ve atama işlemlerini yapmak için kullanılan çok değişkenli istatistik yöntemlerinden biridir. Bu yöntem, bağımsız değişkenlerin dağılımına ilişkin araştırmacılarca karşılanması gereken normal dağılım, doğrusallık ve varyans-kovaryans matrislerinin eşitliği gibi sayıtların karşılanmasını gerektirmez. Lojistik regresyon analizi, bireylerin hangi grubun üyesi olduğunu kestirmede kullanılabilir bir regresyon denklemi oluşturmayı amaçlamaktadır (Çokluk *et alii.*, 2010). Lojistik regresyon en az değişkeni kullanarak en iyi uyuma sahip olacak şekilde bağımlı değişken ile bağımsız değişkenler kümesi arasındaki ilişkiyi tanımlayabilen ve genel olarak kabul edilebilir bir model kurmayı amaçlar (Aktaş, & Erkuş, 2009). Lojistik regresyon analizinde amaç, kategorik bağımlı değişkenin değerini tahmin etmek olduğundan, aslında burada yapılmaya çalışılan iki ya da daha fazla gruba ilişkin “üyelik” tahminidir. Buna göre analizin amaçlarından birinin sınıflandırma, diğerinin ise bağımlı ve bağımsız değişkenler arasındaki ilişkileri araştırma olduğu ifade edilebilir (Mertler, & Vannatta, 2005; Akt. Çokluk *et alii.*, 2010).

Öğrenci başarısızlıklarının tahmininde yapay sinir ağlarının kullanımı ile ilgili çalışmalar ilk kez, 1994 yılında Gorr, Nagin ve Szczyplu’ın öğrencilerin ağırlıklı not ortalamalarının tahmininde çoklu doğrusal regresyon ve aşamalı doğrusal regresyon analizleri ile yapay sinir ağları analizlerini karşılaştırması ile gerçekleştirilmiştir. Araştırma bulgularına göre yapay sinir ağları analizi ile yapılan tahminlerin daha iyi olduğu sonucuna varılmıştır. Subbanarasimha, Arinzeb, & Anandarajanb (2000), MBA öğrencilerinin akademik performanslarını tahmin etmede yapay sinir ağları ve regresyon tekniklerini karşılaştırmak için iki farklı veri kümesini kullanmışlardır. Araştırma sonucunda yapay sinir ağları analizinin doğru tahmin etme yüzdesi daha yüksek bulunmuştur. Güneri, & Apaydın (2004), öğrencileri başarı durumlarına göre sınıflandırmada yapay sinir ağları ve lojistik regresyon yöntemlerinden yararlanmışlardır. Başarı sınıflandırmasında lojistik regresyon analizi ve sinir ağları yaklaşımı sonucunda elde edilen genel doğru sınıflandırma oranı %95 olarak bulunmuştur. Tosun (2007), öğrenci başarısı ile ilgili yaptığı çalışmada yapay sinir ağları ve karar ağaçları teknikleri karşılaştırmıştır. Karar ağaçları ile öğrenci başarılarına göre sınıflandırma işlemi sonrasında %86 oranında başarı elde edilmiştir. Aynı verilerle yapılan yapay sinir ağları uygulaması sonrasında başarı oranı yaklaşık %92 olarak bulunmuştur. Ibrahim ve Rusli (2007) öğrenci başarılarının tahmininde yapay sinir ağları, karar ağaçları ve doğrusal regresyon yöntemlerini karşılaştırmışlardır. Genel akademik başarının tahmininde yapay sinir ağları analizinin daha iyi sonuçlar verdiği görülmüştür. Oladokun, Adebajo ve Charles-Owaba (2008), yaptıkları çalışmada öğrenci başarısı üzerinde etkili olan değişkenleri belirlemek ve öğrenci performansını tahmin etmede yapay sinir ağları analizini test etmeyi amaçlamışlardır. Modele eklenecek yeni bir öğrencinin gelecekteki başarısı tahmin edilmek istendiğinde, yapay sinir ağları kullanıldığı zaman tahminin doğru olma olasılığı %74 olarak belirlenmiştir.

Üniversitede okuyan öğrencilerin başarısızlığa uğramasının önlenmesi, başarıyı etkile-

yen faktörlerin belirlenmesine bağlıdır. Başarıyı etkileyen faktörler bilinirse, başarısızlığa yol açan etkenler kontrol altına alınabilir ve öğrencilerin niteliklerine göre daha doğru bir sınıflandırma yapılması sağlanabilir. Yapılan alanyazın incelemeleri sonucunda, özellikle eğitim alanında öğrenci başarılarının sınıflandırılmasında yapay sinir ağlarının performansını geleneksel istatistiksel yöntemlerle karşılaştıran sınırlı sayıda çalışmaya rastlanmıştır. Bu nedenle bu çalışmada öğrencilerin akademik başarılarına göre sınıflandırılmasında lojistik regresyon analizi ve yapay sinir ağları yaklaşımlarının sınıflama performanslarının karşılaştırılması problemi üzerinde durulmuştur.

Amaç

Bu araştırmanın genel amacı öğrenci başarısı üzerinde etkili olan değişkenlerin belirlenmesi ve farklı fakültele ve programlara yeni kayıt yaptıran öğrencilerin gelecekteki başarılarının tahmin edilmesine olanak sağlayacak sınıflandırma modellerinin elde edilmesinde lojistik regresyon analizi ve yapay sinir ağları yöntemlerinin performanslarının karşılaştırılmasıdır.

Bu genel amaç doğrultusunda aşağıdaki sorulara yanıt aranmıştır:

1. Öğrencilerin ortaöğretim mezuniyet ortalaması, mezun oldukları lise, üniversiteye giriş puanı, öğrenim gördükleri bölümü tercih sırası, üniversiteye yerleştirilmede esas alınan puan türü, vize/final gibi sınavlara hazırlanma zamanı, sınavlara hazırlanırken tercih ettiği ortam, düzenli ders çalışma alışkanlığına sahip olma durumu, sahip olunan olanaklar (çalışma odası, çalışma masası, internet bağlantısı, bilgisayar, kitaplık), kardeş sayısı, ailenin ortalama toplam aylık geliri, anne eğitim durumu, baba eğitim durumu, anne çalışma durumu ve baba çalışma durumu değişkenlerine göre:
 - a. Lojistik regresyon analizi ile sonucunda elde edilecek olan model, öğrencileri başarı durumlarına göre hangi doğruluk oranında sınıflandırmaktadır?
 - b. Yapay sinir ağları analizi öğrencileri başarı durumlarına göre hangi doğruluk oranında sınıflandırmaktadır? Eğitim, geçerlik ve test süreçleri ile genel doğru sınıflandırma oranları nelerdir?
2. Lojistik regresyon ve yapay sinir ağları yaklaşımının öğrencileri başarı durumlarına göre genel doğru sınıflandırma oranlarının karşılaştırılmasına ilişkin sonuçlar nasıldır?

Yöntem

Çalışma Grubu

Araştırmanın çalışma grubunu 2011-2012 Eğitim-Öğretim Yılı Bahar Dönemi'nde Ankara Üniversitesi Eğitim Bilimleri Fakültesi ile Dil ve Tarih-Coğrafya Fakültesi'nin bazı programlarında lisans öğrenimi gören toplam 419 üçüncü sınıf öğrencisi oluşturmaktadır. Bu öğrencilerin 304'ü (%72.60) kız ve 115'i (%27.40) ise erkektir. Tablo 1'de çalışma grubunu oluşturan öğrencilerin fakültele ve programlara göre dağılımı sunulmaktadır.

Tablo 1. Çalışma Grubunu Oluşturan Öğrencilerin Fakülte ve Programlara Göre Dağılımı

Fakülte	Program	f	%
Eğitim Bilimleri Fakültesi	Rehberlik ve Psikolojik Danışmanlık	47	11.20
	Bilgisayar Öğretimi ve Teknolojileri Eğitimi	45	10.70
	Sınıf Öğretmenliği	29	6.90
	Sosyal Bilgiler Öğretmenliği	38	9.10
	Zihinsel Engelliler Öğretmenliği	34	8.10

	Okul Öncesi Öğretmenliği	52	12.40
	Toplam	245	58.50
Dil ve Tarih- Coğrafya Fakültesi	Japon Dili ve Edebiyatı	5	1.20
	Fransız Dili ve Edebiyatı	23	5.50
	İngiliz Dili ve Edebiyatı	28	6.70
	İtalyan Dili ve Edebiyatı	18	4.30
	Amerikan Kültürü ve Edebiyatı	14	3.30
	Türk Dili ve Edebiyatı	22	5.30
	Psikoloji	22	5.30
	Sosyoloji	15	3.60
	Felsefe	27	6.40
	Toplam	174	41.50
	Toplam	419	% 100.00

Veri Toplama Aracı

Araştırmada veri toplama aracı olarak; araştırmacı tarafından geliştirilen “*Üniversite Öğrencilerinin Akademik Başarılarını Etkileyen Değişkenler Anketi*” kullanılmıştır. Anket; öğrenim görülen fakülte, bölüm, cinsiyet, ortaöğretim mezuniyet ortalaması, mezun olunan lise, üniversiteye giriş puanı, öğrenim gördüğü bölümü tercih sırası, üniversiteye yerleştirilmede esas alınan puan türü, vize/final gibi sınavlara hazırlanma zamanı, sınavlara hazırlanırken tercih edilen ortam, düzenli ders çalışma alışkanlığına sahip olma durumu, sahip olunan olanaklar (çalışma odası, çalışma masası, internet bağlantısı, bilgisayar, kitaplık), kardeş sayısı, ailenin ortalama toplam aylık geliri, anne öğrenim durumu, baba öğrenim durumu, anne çalışma durumu, baba çalışma durumu hakkında bilgi toplamaya yönelik kapalı uçlu sorulardan oluşmaktadır.

Verilerin Analizi

İstatistiksel çözümleme sürecinde “*Lojistik Regresyon Analizi*” ve “*Yapay Sinir Ağları Analizi*” kullanılmıştır. Araştırmada öğrencilerin üç sene sonundaki not ortalamaları, lojistik regresyon analizi için bağımlı değişken, yapay sinir ağları analizi için ise çıktı değişkeni olarak belirlenmiştir. Not ortalaması (1.00-2.72) arasında olan öğrenciler için “*başarısız (0)*”, 2.73-4.00 arasında olan öğrenciler için “*başarılı (1)*” olmak üzere iki kategorili bir süreksiz değişken oluşturulmuştur. Ortaöğretim mezuniyet ortalaması, mezun olunan lise, üniversiteye giriş puanı, öğrenim gördüğü bölümü tercih sırası, üniversiteye yerleştirilmede esas alınan puan türü, üniversiteye giriş puanı, vize/final gibi sınavlara hazırlanma zamanı, sınavlara hazırlanırken tercih edilen ortam, düzenli ders çalışma alışkanlığına sahip olma durumu, sahip olunan olanaklar (çalışma odası, çalışma masası, internet bağlantısı, bilgisayar, kitaplık), kardeş sayısı, ailenin ortalama toplam aylık geliri, anne öğrenim durumu, baba öğrenim durumu, anne çalışma durumu, baba çalışma durumu lojistik regresyon analizi için bağımsız değişkenler, yapay sinir ağları analizi için girdi değişkeni olarak araştırmaya dahil edilmiştir.

Bu çalışmada lojistik regresyon analizi için adımsal yöntemlerden ileriye doğru yöntemi, İleri Beslemeli Yapay Sinir Ağları’ndan ise Çok Katmanlı Algılayıcı (MLR- Multiplayer Perceptron) kullanılmıştır. Girdi katmanındaki yapay sinir hücrelerinin aktivasyon fonksiyonu olarak “*Hiperbolik Tanjant Fonksiyonu*” ve çıktı katmanındaki yapay sinir hücrelerinin aktivasyon fonksiyonu olarak da “*Softmax Fonksiyonu*” kullanılmıştır. Ağ modelinde, sınırlar arasındaki

bağlantıların ağırlık değerleri, uygulamanın başında SPSS 20.00'de rastgele üretilmiştir.

Bulgular ve Yorum

Araştırmanın amaçları doğrultusunda verilerin analiz edilmesi sonucunda elde edilen bulgular ve yorumlar üç temel başlık altında ele alınmıştır. İlk olarak Lojistik Regresyon analizi ile elde edilen modele ilişkin bulgular yorumlanmıştır. İkinci bölümde yapay sinir ağları analizi ile elde edilen modele ilişkin bulgular yorumlanmıştır. Üçüncü olarak lojistik regresyon analizi ve yapay sinir ağları sınıflandırma performansları bakımından karşılaştırılmasına ilişkin bulgulara ve yorumlara yer verilmiştir.

Tablo 2. Lojistik Regresyon Modeli Sonucu Elde Edilen Sınıflandırma Tablosu

Gerçek / Gözlenen Durum	Kestirilen Durum		Doğru Sınıflandırma Yüzdesi
	Başarısız	Başarılı	
Başarısız	187	56	77.00
Başarılı	86	90	51.10
Toplam Doğru Sınıflandırma Yüzdesi			66.10

Tablo 2 incelendiğinde, lojistik regresyon modeli ile yordayıcı değişkenlere göre yapılan sınıflandırma sonucunda, başarısız grupta olan 243 öğrenciden 187'si doğru, 56'sı yanlış sınıflandırılmış olup, doğru sınıflandırılma oranı %77.00'dir. Başarılı olan 176 öğrenciden 90'ı doğru, 86'sı yanlış sınıflandırılmış olup, doğru sınıflandırma oranı %51.10'dur. Amaçlanan modele ilişkin toplam doğru sınıflandırma oranı ise %66.10'dur.

Yapay sinir ağı ile öğrencilerin akademik başarı durumlarına göre sınıflandırılmasına ilişkin yapılan analizle elde edilen bulgular Tablo 3'te sunulmuştur.

Tablo 3. Yapay Sinir Ağı Modeli Sonucu Elde Edilen Sınıflandırma Tablosu

	Gerçek / Gözlenen Durum	Kestirilen Durum		Doğru Sınıflandırma Yüzdesi
		Başarısız	Başarılı	
Eğitim	Başarısız	117	18	%86.70
	Başarılı	47	45	%48.90
	Toplam	%72.20	%27.80	%71.40
Test	Başarısız	35	4	%89.70
	Başarılı	9	15	%62.50
	Toplam	%69.80	%30.20	%79.40
Geçerlilik	Başarısız	51	17	%75.00
	Başarılı	30	30	%50.00
	Toplam	%63.30	%36.70	%63.30

Tablo 3 incelendiğinde, oluşturulan yapay sinir ağı modelinin eğitim setinde bulunan 135 başarısız öğrenciden 117'si doğru, 18'i yanlış sınıflandırılmış olup, doğru sınıflandırma oranı %86.70'dir. Başarılı olan 92 öğrenciden, 45'i doğru, 47'si yanlış sınıflandırılmış olup, doğru sınıflandırılma oranı %48.90'dır. Toplam doğru sınıflandırılma oranı %71.40'dır.

Test setinde 39 başarısız öğrenciden 35'i doğru, 4'ü yanlış sınıflandırılmış olup, doğru sınıflandırma oranı %89.70'tir. Başarılı olan 24 öğrenciden 15'i doğru, 9'u yanlış sınıflandırılmış olup, doğru sınıflandırılma oranı %62.50'dir. Toplam doğru sınıflandırma oranı %79.40'dır. Test setinin eğitim uygulamasının performansını ölçmede kullanıldığı hatırlanacak olursa, doğru

sınıflandırma oranının yüksek olduğu ifade edilebilir.

Geçerlilik setinde 68 başarısız öğrenciden 51'i doğru, 17'si yanlış sınıflandırılmış olup, doğru sınıflandırma oranı %75.00'dır. Başarılı olan 60 öğrenciden 30'u doğru, 30'u yanlış sınıflandırılmış olup, doğru sınıflandırılma oranı %50.00'dır. Toplam doğru sınıflandırma oranı %63.30'dur.

Lojistik regresyon analizi ve yapay sinir ağları analizi yöntemlerinin performanslarını karşılaştırmak amacıyla her iki yöntemle elde edilen sınıflandırma tablolarından yararlanılmıştır. Yapay sinir ağları analizi sonucunda, her bir set için (eğitim, test ve geçerlilik) ayrı sınıflandırma tabloları elde edilmektedir ve her bir setten elde edilen toplam doğru sınıflandırma yüzdeleri farklıdır. Bu nedenle genel doğruluk yüzdesi hesaplanırken bu üç set birleştirilmiştir. Birleştirme işlemi yapılırken, aynı hücrelerde yer alan atama değerleri toplanmış ve genel doğru sınıflandırma yüzdesi elde edilmiştir. Tablo 4'te her iki analiz ile elde edilen doğru sınıflandırma yüzdeleri karşılaştırmalı olarak sunulmuştur.

Tablo 4. Yapay Sinir Ağı (YSA) ve Lojistik Regresyon Analizi (LRA) Sınıflandırma Yüzdesi Karşılaştırması

Gerçek / Gözlenen Durum	Kestirilen Durum				Doğru Sınıflandırma Yüzdesi	
	Başarısız		Başarılı		LRA	YSA
	LRA	YSA	LRA	YSA		
Başarısız	187	204	56	39	%77.00	%83.95
Başarılı	86	86	90	90	%51.10	%51.10
Toplam Doğru Sınıflandırma Yüzdesi					%66.10	%70.16

Tablo 4'te yer alan lojistik regresyon modeli sonucunda elde edilen sınıflandırma sonucu incelendiğinde, yordayıcı değişkenlere göre yapılan sınıflandırma ile başarısız grubunda olan 243 öğrenciden 187'si doğru, 56'sı yanlış sınıflandırılmış olup, doğru sınıflandırma oranı %77.00'dır. Başarılı olan 176 öğrenciden 90'ı doğru, 86'sı yanlış sınıflandırılmış olup, doğru sınıflandırma oranı %51.10'dur. Amaçlanan modele ilişkin toplam doğru sınıflandırma oranı ise %66.10'dur. Yapay sinir ağı analizi sonucu elde edilen sınıflandırma sonucu incelendiğinde, başarısız grubunda olan 243 öğrenciden 204'ü doğru, 39'u yanlış sınıflandırılmış olup, doğru sınıflandırma oranı %83.95'tir. Başarılı olan 176 öğrenciden 90'ı doğru, 86'sı yanlış sınıflandırılmış olup doğru sınıflandırma oranı %51.10'dur. Toplam doğru sınıflandırma oranı ise %70.16'dır. Tablo 4 incelendiğinde "başarılı" öğrencilerin sınıflandırılmasında lojistik regresyon analizi ile yapay sinir ağları analizinin toplam doğru sınıflandırma yüzdelerinin aynı olduğu görülmektedir. Bu durum yapay sinir ağları analizinin lojistik regresyon analizine alternatif bir yöntem olarak kullanılabilmesi şeklinde yorumlanabilir.

Sonuç

Lojistik regresyon modeli ile öğrencileri başarı durumlarına göre sınıflandırma işlemi sonrasında %66.10 oranında bir başarı elde edilmiştir. Aynı verilerle yapılan yapay sinir ağları uygulaması sonrasında başarı oranı %70.16 olarak bulunmuştur. Modele eklenecek yeni bir öğrencinin gelecekteki başarısı tahmin edilmek istendiğinde, lojistik regresyon analizi kullanıldığı zaman tahminin doğru olma olasılığı %66.10 iken yapay sinir ağları kullanıldığında bu oran %70.16'dır. Bu sonuçlarla birlikte yapay sinir ağlarının sınıflandırma analizinde lojistik regresyon analizine göre daha başarılı sonuçlar ortaya koyduğu görülmektedir.

Lojistik regresyon analizi ile yapılan öğrenci başarılarını sınıflandırmada, oluşturulan lojistik regresyon modelinin, başarılı öğrencilerin %51.10'unu doğru sınıflandırdığı görülmektedir. Başarısız öğrencilerin %77.00'ını doğru sınıflandırmıştır. Bu durum başarısız öğrencilerin sınıflandırılmasındaki performansın, başarılı öğrencilerin sınıflandırılmasındaki performansa göre daha yüksek olduğunu göstermektedir. Buna göre, oluşturulan lojistik regresyon modelinin özellikle başarısız öğrencilerin sınıflandırılmasında daha iyi sonuçlar verdiğini ifade etmek mümkündür.

Yapay sinir ağları ile oluşturulan çok katmanlı algılayıcı model, başarılı öğrencilerin %51.10'unu doğru sınıflandırırken, başarısız öğrencilerin %83.95'ini doğru sınıflandırmaktadır. Bu durum başarısız öğrencilerin sınıflandırılmasındaki performansın, başarılı öğrencilerin sınıflandırılmasındaki performansa göre daha yüksek olduğu ve bu nedenle yapay sinir ağı ile kurulan modelin özellikle başarısız öğrencilerin sınıflandırılmasında daha iyi sonuçlar verdiği şeklinde yorumlanabilir.

Bu araştırma bir çalışma grubundan elde edilen sınırlı sayıda veri ile gerçekleştirilen bir örnek uygulama niteliğindedir. Bu çalışma kapsamında, iki farklı analiz tekniğinin birbirleri ile karşılaştırılması yapılmış ve birinin diğerine göre daha doğru sınıflandırma oranları sunduğundan söz edilmiş olsa da, sonuçların genellenmesi, bir tekniğin daha başarılı sonuçlar verdiğinin ifade edilebilmesi için, kapsamlı çalışmalara (büyük veri setleri, gerçek ve simülatif verilerle yapılan çalışmalar vb.) ihtiyaç olduğu belirtilmelidir. O nedenle de çalışmadan elde edilen sonuçların, söz konusu sınırlılıklar dikkate alınarak değerlendirilmesi yararlı olacaktır.

KAYNAKÇA

- Ahmann, J. S., & Glock M. D., (1967). *Evaluating Pupil Growth: Principles of Tests and Measurement*. (3. Baskı). Boston: Allyn and Bacon Inc.
- Aktaş, C., & Erkuş, O. (2009). "Lojistik Regresyon Analizi ile Eskişehir'in Sis Kestiriminin İncelenmesi". *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi*, 8 (16), 23-46.
- Aydın, B., (2002). *Basit ve Geri Yayılımlı Yapay Sinir Ağları, Uygulama Alanları*. Yayınlanmamış Yüksek Lisans Tezi, Çanakkale Onsekiz Mart Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı, Çanakkale.
- Carter, V., & Good, E. (1973). *Dictionary of Education*. (4. Baskı). New York: McGraw Hill Book Company.
- Çokluk, Ö., Şekercioglu, G., & Büyükoztürk Ş., (2010). *Sosyal Bilimler İçin Çok Değişkenli İstatistik*. Ankara: Pegem Akademi.
- Demirtaş H., & Güneş, H. (2002). *Eğitim Yönetimi ve Denetimi Sözlüğü*. Ankara: Anı Yayıncılık.
- Gorr, W. L., Nagin, D., & Szczypula, J. (1994). "Comparative study of Artificial Neural Network and Statistical Models for Predicting Student Grade Point Averages". *International Journal of Forecasting*, (10), 17-34.
- Gülleroğlu, D. (2005). *Üniversite Öğrencilerinin Akademik Başarılarının Yordanmasına İlişkin Karşılaştırmalı Bir Araştırma*. Yayınlanmamış Doktora Tezi. Ankara Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.
- Güneri, N., & Apaydın, A. (2004). "Öğrenci Başarılarının Sınıflandırılmasında Lojistik Regresyon Analizi ve Sinir Ağları Yaklaşımı". *Ticaret ve Turizm Eğitim Fakültesi Dergisi*, (1), 170-188.
- Gürsoy, T. Ş. (2009). *Veri Madenciliği ve Bilgi Keşfi*. (1. Basım). Ankara: Pegem A Yayıncılık.
- Ibrahim, Z., & Rusli, D. (2007). "Predicting Students' Academic Performance: Comparing Artificial Neural Network, Decision Tree and Linear Regression". *21st Annual SAS Malaysia Forum*, Shangri-La Hotel, Kuala Lumpur, September, 5.
- Kamber, J. H., & Morgan, M. (2000). *Data Mining Concept and Techniques*. San Francisco, USA: Kaufmann Publishers.
- Marzano, R. J., Pickering D. J., & Pollock J. E. (2008). *Öğrenci Başarısını Artıran Öğretim Stratejileri*. Çev. Sakacı, S. İstanbul: Sev Yayınevi.
- Oladokun, V. O., Adebajo, A. T., & Charles-Owaba, O. E. (2008). "Predicting Students' Academic Performance using Artificial Neural Network: A Case Study of an Engineering Course". *The Pacific Journal of Science and Technology*, 9 (1), 72-79.
- Özguven, İ. E. (1974). "Üniversite Öğrencilerinin Akademik Başarılarını Etkileyen Zihinsel Olmayan faktörler". *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*. Ankara: Hacettepe Üniversitesi Basımevi.
- Özoğlu, S. Ç., & Koç, N. (1995). "Çağdaş Üniversitede Öğrencinin Başarısının Ölçülmesi ve Değerlendirilmesi". *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Yayınları*, Ankara.
- Öztemel, E. (2003). *Yapay Sinir Ağları*. (1. Basım). İstanbul: Papatya Yayıncılık.
- Saracaloğlu, A. S., Yenice, N. (2009). Sınıf Öğretmeni Adaylarının Öğrenme Stilleri ile Fen Başarıları Arasındaki İlişki. *Yüzüncü Yıl Üniversitesi, Eğitim Bilimleri Fakültesi Dergisi*, 6 (1), 162-173.
- Subbanarasimha, P. N., Arinzeb, B., & Anandarajanb, M. (2000). "The Predictive Accuracy of Artificial Neural Networks and Multiple Regression in the Case of Skewed Data". *Exploration of Some Issues. Expert Systems with Applications*, 19, 117-123.
- Tosun, S. (2007). *Sınıflandırmada Yapay Sinir Ağları ve Karar Ağaçları Karşılaştırması: Öğrenci Başarıları Üzerine Bir Uygulama*. Yayınlanmamış Yüksek Lisans Tezi. İstanbul Teknik Üniversitesi Fen Bilimleri Enstitüsü, İstanbul.