

ROC Analizi ve R Yazılımı ile Verilerin Sınıflama Doğruluklarının Karşılaştırılması

Fazilet TAŞDEMİR¹ 

¹Dr. Öğr. Üyesi., Recep Tayyip Erdoğan Üniversitesi, fazilet.tasdemir@erdogan.edu.tr, ORCID ID:0000-0002-0430-9094

Makale Bilgisi

ÖZET

Geliş Tarihi:
25.08.2022

Kabul Tarihi:
19.10.2022

© UEAD 2022
Tüm hakları saklıdır.

ROC analizi, bir tanı testi olarak, bir sınıflandırıcıyı temsil eden sürekli bir değişkenin ayırt edici performansını değerlendirmek için sıkça kullanılır. Verilerin sınıflandırılmasında, kümelenmesinde kullanımı artan bir diğer program ise R Studio ortamıdır. Bu çalışmada her iki istatistiksel analizden yararlanarak, daha önceden var olan kesme puanlarla aynı veri grubu üzerinde aynı özelliği belirlemedeki sınıflandırma doğruluklarının incelenmesi amaçlanmıştır. Programlama diline de uygun olan ve araştırmacı tarafından üretilen simülatif veri üzerinden araştırma yürütülmüştür. Araştırma, 1500 kişilik veri seti üzerinden ROC ve R ortamında gerçekleştirilmiştir. Ayrıca, veri grubundan çekilen 25 kişilik bir yargıcı grubu üzerinden de veriler tekrar ikinci bir analizle incelenmiştir. Yargıcıların sınıflama doğruluklarına Öklid uzaklığı ile bakılmıştır. Araştırma sonucunda AUC=0.73 olarak hesaplanmıştır ($p<.05$). Bu bulgu, ROC analizi ile sınıflama doğrulunun orta düzeyde tespit edilebildiğini göstermiştir. R ortamında yapılan analizde sadece yargıcılar arasında Y1-Y24 arasında benzerlik bulunmuştur. Bu durum, R programının yanı sıra karar vermede başka programların da ona eşlik etmesi gerektiği sonucunu ortaya çıkarmıştır. Her iki analiz sonucunda da iki uygulama bulguları arasında uyum görülmemiştir. Araştırmacılara bu iki program dışında diğer sınıflandırıcılarla da çalışmalarını önerilir.

Anahtar Sözcükler: ROC analizi, R program dili, sınıflama doğruluğu

Comparison of Classification Accuracy of Data with ROC Analysis and R Software

Article Information

ABSTRACT

Received:
25 August 2022

Accepted:
19 October 2022

© UEAD 2022
All rights reserved.

ROC analysis is frequently used as a diagnostic test to evaluate the discriminant performance of a continuous variable representing a classifier. Another program that is increasingly used in the classification and clustering of data is the R Studio environment. In this study, it was aimed to examine the classification accuracies in determining the same feature on the same data group with the previously existing cut-off scores by using both statistical analyzes. The research was carried out on the simulative data produced by the researcher, which is also suitable for the programming language. The research was carried out in the ROC and R environment on the data of 1500 people. In addition, the data were re-examined with a second analysis on a group of 25 judges who were withdrawn from the data group. The classification accuracy of the judges was checked with the Euclidean distance. As a result of the research, it was calculated as AUC=0.73 ($p<.05$). This finding showed that the classification accuracy could be determined moderately by ROC analysis. In the analysis made in the R environment, there was a similarity between Y1-Y24 only among the judges. This situation led to the conclusion that besides the R program, other programs should accompany it in decision making. As a result of both analyzes, there was no agreement between the findings of the two applications. Researchers are advised to work with other classifiers besides these two programs.

Keywords: ROC analysis, R program language, classification accuracy

DOI: 10.32960/uead.1166987

Makale Türü (Article Type): Araştırma Makalesi

Kaynakça Gösterimi: Taşdemir, F. (2022). ROC analizi ve R yazılımı ile verilerin sınıflama doğruluklarının karşılaştırılması. *Ulusal Eğitim Akademisi Dergisi (UEAD)*, 6(2), 230-241.

Citation Information: Taşdemir, F. (2022). Comparison of classification accuracy of data with ROC analysis and R software. *National Journal of Education Academy*, 6(2), 230-241.

1. GİRİŞ

Araştırmalarda elde edilen verilerin sınıflandırılma doğruluklarının tespiti önemli bir amaç olarak çoğu kez çalışmalarda yer alır. Veriyi sınıflandırma; nesnelere varolan tanımlı kategorilere yerleştirme esasına dayalı, farklı türden de uygulamaları kapsayabilecek nitelikteki geniş ölçekli bir sorundur (Tan vd., 2006). Sınıflandırma, veri dizisinin çeşitli istatistiksel teknikler ve makine öğrenmesi gibi yöntemlerden yararlanılarak önceden oluşturulmuş olan sınıflara atanması işlemidir (Akpınar, 2014). Verinin sınıflanmasına ve sınıflama doğruluklarına göre yapılmış araştırmaların (Hess, Subhiyah & Giordano, 2007; Hertz & Hertz, 1999; Koyuncu, 2015; Ömür & Selvi, 2010; Taşdelen, 2009; Vivo & Franco, 2008;) olduğu görülmektedir. Sınıflara tasnif etme başarısını ölçmeyi amaçlayan R dili ile R studio ortamında veri seti üzerinde (Kılınç vd., 2016) yapılan araştırmalar da bulunmaktadır. Karar ağacı ve ROC eğrisi sınıflandırma düzeyleri ile yapılan araştırmalara da rastlanmaktadır (Ferri, Flach, Hernández-Orallo & Senad, 2005; Provost & Fawcett, 2001). Performans ve akademik başarı durumu gibi kavramlar çoğu zaman kullanılan farklı türden ölçme araçları ile belirlenebilmektedir. Birey hakkında başarılı-başarısız, geçti-kaldı, yeterli-yetersiz gibi yargılara ulaşabilmek için sınıflandırmanın doğruluğu çok önem kazanmaktadır. Standart belirleme süreci, kişilerin başarı düzeylerini ayırıştırabilmek için son derece önemlidir. Her düzeyin birbirinden farklılaştığı noktanın belirlenmesi gerekmektedir. Bu açıdan kullanılacak olan standart belirleme yönteminin seçimi de önem taşımaktadır (Cizek, 2001).

Ölçme araçlarından elde edilen verilerin kategorize edilmesinde kullanılan veri analizlerinden biri de Alıcı İşletim Özellikleri [Receiver Operating Characteristics]- (ROC) analizidir. ROC analizi, sınıflandırıcıları tahmin performanslarına göre değerlendirmek, karşılaştırmak ve seçmek için kullanılan bir metodolojidir. ROC analizi iki kategorili, sınıflamalı durum tespitlerinin incelenmesi için geliştirilmiştir. Ölçülen özellikle ilgili ikili sınıflandırma hatasının olası tüm kombinasyonları için sınıflandırıcı eşliğinin yanlış pozitif oranı ve yanlış negatif oranı ile çalışır (Fawcett, 2005; Metz, 1978). ROC analizi, optimal davranacak ve optimal olmayan sınıflandırıcıları reddedecek bir dizi sınıflandırıcı seçmek için araçlar sağlar. Bunu yapmak için, tüm sınıflandırıcıların dışbükey gövdesi bir "eğri" vererek oluşturulur (Ferri vd., 2005).

Bir ROC grafiği, performanslarına göre sınıflandırıcıları görselleştirme, düzenleme ve seçme tekniğidir (Swets, Dawes & Monahan, 2000). ROC grafikleri, x ekseninde Yanlış Pozitif (YP) oranları ve y ekseninde Gerçek Pozitif (GP) oranları çizer. ROC eğrileri, bir puanlama modelinin çıktı aralığı boyunca bir eşik değiştirerek ve ilgili sınıflandırma performanslarını gözlemleyerek, kesinlik/hatırlama eğrilerine benzer bir şekil oluşturur. ROC eğrileri, marjinal sınıf dağılımındaki değişikliklere karşı duyarsız olmaları gibi ek avantajlara sahiptirler. Genellikle iki veya daha fazla ROC eğrisinin karşılaştırılması ya eğrinin altındaki alana bakmaktan ya da eğrilerin belirli bir kısmına odaklanmaktan ve en iyi performans gösteren algoritmayı seçmek için hangi eğrinin diğerine hakim olduğunu belirlemekten oluşur (Macskassy & Provost, 2004).

ROC eğrisinin değerlendirilmesinde kullanılan göstergelerden birisi olan ROC Eğrisinin Altındaki Alan [Area Under the ROC Curve]- (AUC), bir sınıflandırıcının bir sıralayıcı olarak ne kadar iyi

davrandığını, bir sınıflandırıcı olarak ne kadar iyi davrandığını tahmin etmek için kullanılabilir. ROC analizi ve AUC ölçümü, bilgi keşfi, veri madenciliği alanında yaygın olarak kullanılmıştır (Adams & Hand, 1999).

ROC eğrisi, bir karar eşiğine başvurmadan, puanlama sınıflandırıcısının sınıflandırma gücünün toplu bir değerlendirmesini verir. Eğrinin altındaki alan, rastgele seçilmiş bir pozitif örneğin, rastgele seçilmiş bir negatif örnekten daha yüksek bir puan alma olasılığını tahmin eder (Hand & Till, 2001). İki kategorili bir ROC eğrisi, Gerçek Pozitif Oranın (GPO) Y ekseninde ve Yanlış Pozitif Oranın (YPO) X ekseninde çizildiği iki boyutlu bir egridir. Bu oranlar aşağıdaki gibi tahmin edilmektedir (Lachiche & Flach, 2003):

$$\text{Gerçek Pozitif Oran} = \frac{\text{Doğru sınıflandırılmış pozitif örneklerin sayısı}}{\text{Toplam pozitif örnek sayısı}} \quad (1.1)$$

$$\text{Yanlış Pozitif Oran} = \frac{\text{Yanlış sınıflandırılmış negatif örneklerin sayısı}}{\text{Toplam negatif örnek sayısı}} \quad (1.2)$$

ROC analizi, iki sınıflı sınıflandırma problemlerini ortadan kaldırmak için kullanılır. ROC analizinin her biri pozitif veya negatif sınıfa ait olan, başlangıçta mevcut göstergeleri olan durumları tanımlayan veri grubunun pozitif ve geri kalanların ise negatif olarak gösterildiği uygulamalarda kullanımı yaygındır. Böylece ilgili değişkenle üzerinde çalışan bir sınıflandırıcı, daha önce tanımlanmayan veya görünmeyen örneklerin sınıf değerini tahmin edebilir. Bu analizde tahmin edilen sınıflar gerçek sınıflarla mutlaka aynı olmadığından, tahmin hatalarının sayısının kaydını tutmak için bir matris kullanılır (Majnik & Bosnić, 2013). Bu matris, Tablo 1'de gösterildiği gibi bir olasılığı gösteren kontenjans tablosu olarak adlandırılır.

		Tahmini Sınıflama	
		Pozitif (1)	Negatif (0)
Gerçek Sınıflama	Pozitif (1)	Gerçek Pozitif (GP)	Yanlış Pozitif (YP)
	Negatif (0)	Yanlış Negatif (YN)	Gerçek Negatif (GN)

Şekil 1. Sınıflandırıcı İçin İkili Sınıflandırma Problemleri Beklenmedik Kontenjans

Kaynak: Green, D.M ve Swets, J.A. (1988). *Signal Detection Theory and Psychophysics*. Reprint Edition. Los Altos, CA: Peninsula Publishing.

Yapılan bu sınıflamalarda bir tanımlama amacı yatar. Bir tanı testinin doğruluğu duyarlılık (sensitivity) ve seçicilik (specificity) başlıklarında değerlendirilebilir (Krzanowski & David, 2009).

Duyarlılık: Bireyin başarılı olduğu durumda testin sonucunun pozitif yani başarılı olabilme olasılığıdır.

$$GPO = P[D = 1 | S = 1] = P(x_i > t | S^+) = GP/GP + YN \quad (1.3)$$

Seçicilik: Bireyin başarısız olduğu durumda testin sonucunun negatif yani başarısız olabilme olasılığıdır.

$$GNO = P[D = 0 | S = 0] = P(x_i \leq t | S^-) = GN/GN + YP \quad (1.4)$$

Sınıflandırma çalışmalarında en iyi kesme noktasını saptamada iki farklı yaklaşım belirlenebilir. Bunlardan ilki; eğrinin grafiğın sol üst köşesine (0,1) koordinatlarına en yakın olduğu noktayı kesme noktası olarak belirlemektir. Diğeri ise, arařtırmacının kullandığı ölçme aracı için öngörülen bir duyarlılık, seçicilik değeri bulunmadığında, duyarlılık ve seçicilik değeri toplamının en yüksek bulunduğu noktayı en iyi kesim noktası olarak saptamaktır (Alpar, 2010).

Elde edilen verilerin sınıflandırılmasında kullanılan ortamlardan biri de günümüzde kullanımı artan R Studio ortamıdır. Bu paket programı, bir istatistiksel analiz aracı olarak veri üzerinde istenilen değışikliklerin yapılabilmesine ve uygun istatistiksel yöntemlerin kullanılabilmesine olanak sağlar. Bu ortam ayrıca büyük verilerde sınıflandırma, kümeleme gibi analizleri yapabilme imkânı da vermektedir (Ihaka & Gentleman, 1996). Bu arařtırma kapsamında, ROC analizinin yanı sıra bu programa dilinden de simülatif olarak veri giriři yapılan yargıcı verilerini sınıflandırma gücünü belirlemek amacıyla yararlanılmıştır. R programlama dili, verinin sınıflandırılmasında sıklıkla kullanım göstermektedir. Bu programlama dilinin literatürde; veri profillerinin sınıflandırılması (Iglesias vd., 2016), akademik makalelerin sınıflara tasnif edilmesi (Kılınç vd., 2016), farklı sınıflandırma algoritmalarının veri üzerinde sınanması (Çınar, 2019) amacıyla kullanıldığı da görülmektedir.

Sınıflama amacıyla yapılan analizlerin, istatistiksel yöntemlerin arařtırmalarda kullanıldığı görülmektedir. Ancak, bir teste göre belirlenen başarılı/başarısız yargısının ROC analizi ve R programı ile birlikte incelenmemiştir. Bu gerekçeyle, arařtırmanın problem, ROC analizi ve R yazılımı ile verilerin sınıflama doğruluklarının karşılaştırılması olarak belirlenmiştir.

1.2. Arařtırmanın Amacı

Bu arařtırmanın genel amacı, ölçülen özellikle ilgili önceden var olan kesme puanı ile aynı özelliğ ölçen bir teste göre belirlenen başarılı/başarısız olma sınıflandırılma doğruluklarının ROC analizi ve R programı ile incelenmesidir. Bu genel amaç doğrultusunda ařağıdaki sorulara yanıt aranmıştır:

1. ROC eğrisinin altında kalan alan nedir?
2. ROC analizinde belirlenen kesme puanı nedir?
3. R'da farklı yargıcıların lisans not ortlamasına göre sınıflama benzerlikleri var mıdır?
4. ROC ve R'da aynı bireyler için sınıflama uyumu var mıdır?

2. YÖNTEM

Bu arařtırma, bireylerin başarı durumlarının en yüksek doğru sınıflama oranları ile ayırt edebileceğ kesim noktalarının belirlenmesi amaçlandığı için tarama modelinde bir arařtırmadır. Ayrıca, kullanılan analizlerin uyumu da incelendiğ için tarama modelinde korelasyonel bir arařtırmadır. Korelasyonel

araştırmalar, iki ya da daha çok değişken arasındaki ilişkinin birlikte değişiminin varlığının veya derecesinin belirlenmeye çalışıldığı ve bu değişkenlere müdahale edilmeden incelendiği araştırmalardır (Karasar, 2018).

2.1. Araştırma Grubu

Araştırma için öğrencilerin lisans akademik not ortalaması ve Kamu Personeli Seçme Sınavı (KPSS) değişkenlerinin ele alındığı simülatif verilerden oluşan 1500 kişilik bir veri seti oluşturulmuştur. Araştırma grubunu oluşturan veri setinin simülatif tercih edilmesinin sebebi kullanılan analiz programlarının (ROC ve R programlama dili) büyük örneklemelere duyarlı olmasıdır. Böylece, R ortamında kullanılan veri setinde simülatif olarak veri girişi yapılan 25 yargıcının veri setinden random çekilen 30 öğrenciye ait düzenlenen lisans akademik not ortalamalarına göre başarılı-başarısız yargısında bulunma durumları da incelenmiştir.

Araştırmada kullanılan simülatif veri seti, R programında üretilmiştir. Veri üretilirken R programlama dili, veri üretimi sürecinde önemli esneklikler sağlar. “rnorm” komutu normal dağılım altında rastgele veriler üretimi için kullanılmıştır. Fonksiyon içerisindeki “r” (random) seçkisizliği “norm” ise verinin normal dağılım altında veri üretileceğini gösterir. Dağılımın normalliğinin sağlanmasının ardından ROC ve R için analizlere başlanmıştır.

3. BULGULAR

Bu araştırmanın genel amacı doğrultusunda birinci araştırma sorusuna yanıt aranmıştır. Bu amaçla, öncelikle, 1500 kişilik veri seti üzerinden her iki değişken “lisans ve KPSS” için 1-0 tanımlamasına gidilmiştir. (1=başarılı,0= başarısız). Buradaki “1” 50 ve üzeri geçme notunu temsil etmektedir. Lisans akademik başarı durumlarına göre KPSS başarı durumlarının sınıflama doğruluğu incelemesi için ROC analizi yapılmıştır. ROC eğrisinin altında kalan alanın incelenmesi için öncelikle başarılı-başarısız sınıflamasına ilişkin dağılıma ihtiyaç vardır. Tablo 1’de dağılıma yer verilmiştir.

Tablo 1. Başarılı ve Başarısız Olma Durumuna Göre Dağılım

Lisans akademik başarı durumu	N	%
Pozitif (başarılı)	627	41.8
Negatif (başarısız)	873	58.2
Toplam	1500	100.0

Tablo 1 incelendiğinde, veri setinde lisans akademik başarı durumunda belirlenen başarı kesme puanına göre başarılı olan N= 627 (%41.8) kişi görülmektedir. Başarısız olan kişi sayısı ise N= 873 (%58.2) olarak görülmektedir. Veri setinde uç değer ve kayıp veri yer almamaktadır.

Tablo 2. ROC Eğrisi Altında Kalan Alan (Tüm Veri)

AUC	S. Hata	%95 Güven Aralığı Alt Sınır	%95 Güven Aralığı Üst Sınır	p
.730	.013	.704	.756	.000

(p<0.05)

ROC eğrisinin altında kalan alan nedir? sorusuna ilişkin bulgular şöyledir;

ROC eğrisinin altında kalan alan ne kadar büyük ise, bu durum, testin doğru sınıflamada o kadar iyi olduğunu gösterir (Grove, 2006; Hanley & McNeil, 1982;). ROC eğrisinin altında kalan alanın büyüklüğü,

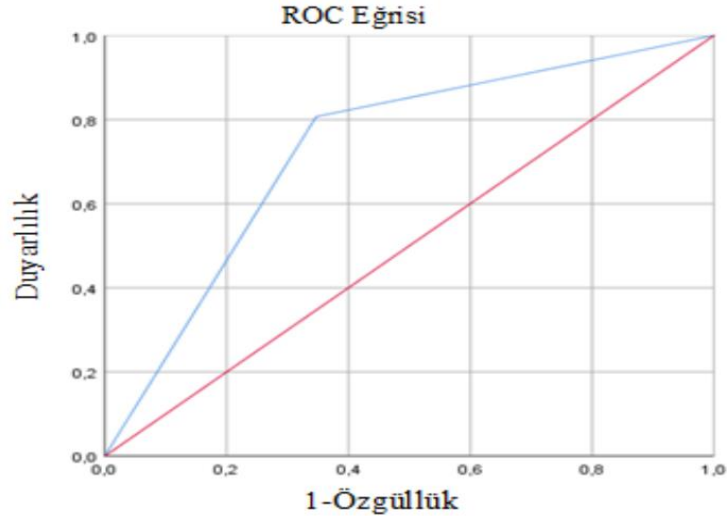
kullanılan testin sınıflama yeteneğinin istatistiksel olarak önemini açıklar. Ayırma, sınıflama gücü olmadığı durumda ROC eğrisi altındaki alanın beklenen değeri 0.50'dir.

Tablo 2 incelendiğinde, ROC analizi sonucunda işlem karakteristik eğrisi altında kalan alan (AUC)=0.73 olarak hesaplanmış ve manidar ($p<.05$) bulunmuştur. Bu değer, oluşturulan simülatif testin öğrenci başarısını belirlemede doğru ayırım yapabildiğini, yani başarılı ve başarısız öğrencileri %73 oranında (orta düzeyde) doğru sınıflandırabildiğini göstermektedir. Bu bulguyla benzer şekilde Takahashi, Uchiyama, Yanagisawa & Kamae (2006), ROC eğrisi altında kalan alanın diyabet sınıflamasında yüksek düzeyde ayırım yaptığı sonucuna ulaşmıştır.

AUC indeksi değer aralığı, $0,5 \leq AUC \leq 1$ 'dir. Eğri altındaki alanların yorumlanmasında “.90-1.00: mükemmel; .80-.90: iyi; .70-.80: orta; .60-.70: zayıf; .50-.60: başarısız” derecelendirmeleri kullanılabilir (Erden & Kanık, 2003).

ROC analizinde belirlenen kesme puanı nedir? sorusuna ilişkin bulgular şöyledir;

Araştırmada oluşturulan veri seti için kesme puanı olarak hangi değer alınmalı sorusunun yanıtını vermek için, analiz sonucunda verilen her bir duyarlılık ve 1-özgüllük değerleri incelenmiştir. Test için bir optimum nokta seçilmiştir. Bu seçim yapılırken duyarlılığın yüksek, 1-özgüllüğün düşük olduğu en iyi nokta belirlenmiştir. Teorik ve ampirik olarak AUC' nin çalışmalarda doğruluk aramalarına göre daha fazla tercih edildiğini göstermektedir (Huang & Ling, 2005). Şekil 1'deki ROC eğrisinde görüldüğü gibi, bu kesme değerinin tespitinde eğrinin önce pik ve ardından plato yaptığı nokta dikkate alınır.



Şekil 1. Başarı Durumu ROC Eğrisi (%95 CI)

Tablo 3. ROC Analizi Eşik Değer Sonuçları

Eşik değer	Duyarlılık	Seçicilik
-1.00	1.000	1.000
.50	.807	.347
2.00	.000	.000

Tablo 3 incelendiğinde, ölçme aracı için duyarlılık değerinin .807 ve seçicilik değerinin .347 olduğu .50, en uygun sınıflandırmayı yapan eşik değer olarak görülmektedir. ROC analizi, duyarlılık ve seçicilik değerlerini de dikkate alıp kesme puan belirlediği için diğer sınıflama amaçlı kullanılan yöntemlere göre daha hassastır denilebilir. Tavakol & Dennick (2012), benzer şekilde bir testin ROC analizi ile kesme puanını belirleme üzerine yaptıkları araştırmada duyarlılık ve özgülük değerlerini dikkate alan ROC'un, diğer sınıflama yöntemlerine göre daha objektif sonuçlar verdiğini ifade etmiştir.

R'da farklı yargıcıların lisans not ortlamasına göre sınıflama benzerlikleri var mıdır? sorusuna ilişkin bulgular aşağıda verilmiştir;

R ortamında 1500 kişiden random seçilen 30 kişinin başarılı-başarısız sınıflaması için 25 yargıcıdan (Y) alınan kararlar aşağıdaki fonksiyon üzerinde girilip, incelenmiştir. Yargıcıların, belirlenen kişilerin lisans akademik başarı durumuna göre KPSS başarı durumlarının (başarılı-başarısız) sınıflamasında araştırma grubunda yer alan her veri için 1-0 girişi yapılmıştır.

Daha sonra "Veri" isimli dosya, sınıflandırma sürecinde kullanılmak amacıyla ".csv" uzantılı olarak kaydedilmiştir. Bu dosya, aşağıda verilen komut dizilimi aracılığıyla, R çalışma ortamına aktarılmıştır. İşlem adımları:

İlk işlem adımında yargıcı kararlarına ilişkin vektörlerin girişleri yapılmıştır. Bu adımda, her bir yargıcının ölçme aracı puanına ilişkin kişileri başarılı olma durumuna göre sıralamaları görülmektedir.

Tablo 4. Yargıcı Sıralamalarının Vektör Gösterimi

Yargıcı	Sıralamalar
>Y1=c(27,163,59,368,324,281,575,1009,690,1110,1371,1228,1260,1422,1379,1490,19,1481,1295,90,202,440,81,517,721,1473,934,868,15,1167)#	
>Y2=c(27,440,59,1167,517,281,90,202,1295,81,1371,1228,1260,1422,934,1490,19,1481,690,575,1009,163,1110,324,721,1473,1379,868,15,368)#	
>Y3=c(1110,163,59,81,324,1422,575,1009,440,27,1371,1228,1260,281,19,1490,1379,1481,1295,90,202,690,368,517,1167,1473,934,868,15,721)#	
>Y4=c(868,163,721,368,324,281,517,1009,690,1110,1490,1228,1260,1422,202,1167,19,1481,1295,90,1379,440,81,575,59,1473,934,27,15,1371)#	
>Y5=c(517,324,59,368,163,1110,575,1260,690,281,1371,1228,1009,81,1167,1490,19,1481,1295,90,202,440,1422,27,934,1473,721,868,15,1379)#	
>Y6=c(27,281,59,368,324,163,575,1167,690,1110,1260,1228,1371,1422,1379,1490,19,440,1295,90,202,1481,81,517,721,1473,934,868,15,1009)#	
>Y7=c(81,163,59,368,721,281,575,1009,27,1110,1371,1228,19,1422,1379,1490,1260,1481,202,90,1295,440,690,517,324,1473,934,868,15,1167)#	
>Y8=c(27,59,163,1228,368,324,281,690,868,575,1110,1371,1260,1422,1379,1490,19,1167,1295,90,202,440,81,517,721,1473,934,1009,15,1481)#	
>Y9=c(281,368,59,163,324,27,575,1260,690,1110,1371,1228,1009,1422,1379,1490,202,1481,1295,90,19,440,81,517,15,1473,934,868,1167,721)#	
>Y10=c(163,27,281,59,368,324,1009,575,1167,690,1110,1490,721,1228,1371,1422,1379,1260,19,440,1295,90,202,1481,81,517,1473,934,868,15)#	
>Y11=c(868,721,163,368,324,281,1009,690,1110,1490,1228,1260,1422,202,1167,19,1481,1295,90,1379,440,81,575,59,1473,934,1371,27,15,517)#	
>Y12=c(59,81,163,27,368,721,281,575,1009,1379,1110,1371,1228,19,1422,440,1490,1260,1481,202,90,1295,15,690,517,324,1473,934,868,1167)#	
>Y13=c(368,163,27,281,59,1167,324,1009,575,1379,690,1110,1490,721,1228,1371,1422,90,1260,19,440,1295,202,934,1481,81,517,1473,868,15)#	
>Y14=c(368,59,281,163,324,27,575,1260,1379,1110,690,1371,1228,1009,1422,868,1490,202,1481,1295,90,19,440,81,517,15,1473,934,1167,721)#	
>Y15=c(27,281,59,368,1110,163,575,1167,690,324,1260,1228,1379,1422,1371,1490,19,440,1295,90,202,1481,81,517,868,1473,934,721,15,1009)#	
>Y16=c(163,575,59,368,324,281,27,1228,690,1110,1371,1009,1260,1422,1379,90,19,1481,1295,1490,202,440,81,517,721,1473,934,868,15,1167)#	

```
>Y17=c(59,163,81,368,721,281,575,1009,1110,27,1228,1371,1260,19,1422,1379,1490,934,1481,202,90,1295,440,690,517,324,1473,868,15,1167)#
>Y18=c(368,27,59,163,1228,1260,324,281,690,868,575,1110,1371,440,1422,1379,1490,19,1167,1295,90,202,81,517,721,1473,934,1009,15,1481)#
>Y19=c(281,368,163,27,690,59,1167,324,1009,575,1379,1371,1110,1490,721,1228,1473,1422,90,1260,19,440,1295,202,934,1481,81,517,868,15)#
>Y20=c(27,517,1295,440,59,1167,934,281,90,202,81,1371,1228,1260,1422,575,1490,19,1481,690,1009,163,1110,324,721,1473,1379,868,15,368)#
>Y21=c(368,59,575,281,163,324,27,690,1260,1379,1110,1371,1228,1009,1295,868,1490,202,1481,1422,90,19,440,81,517,15,1167,1473,934,721)#
>Y22=c(59,163,81,721,368,281,575,1009,1110,1228,27,1371,1260,19,1422,1379,1490,934,1481,202,517,1295,440,90,690,324,1473,868,15,1167)#
>Y23=c(59,163,575,27,368,324,281,1110,1228,690,1371,1009,1260,1422,1379,90,19,1481,1295,1490,202,440,81,517,721,868,1473,934,15,1167)#
>Y24=c(27,163,59,368,324,281,1110,1009,690,1228,1371,575,1260,1422,1379,1490,19,1481,1295,90,202,440,81,517,721,1167,934,868,15,1473)#
>Y25=c(281,27,163,59,368,324,1371,575,1009,690,1110,1379,1228,1260,1422,1295,1490,19,1481,90,202,440,81,517,721,1473,934,868,1167,15)#
```

```
> veri<-data.frame(Y1,Y2,Y3,Y4,Y5,Y6,Y7,Y8,Y9,Y10,Y11,Y12,Y13,Y14,Y15,Y16,Y17,Y18,Y19,Y20,Y21,Y22,Y23,Y24,Y25)#
```

İkinci adımda, Y1-Y25 arasındaki vektörler veri çerçevesine “data.frame” birleştirilmiştir. Sonrasında veri çerçevesi alt komutta yazdırılmıştır. Aşağıda her komut dizilimi de görülmektedir.

```
> veri#
```

>dist(veri,method="euclidian")# komutu ile satırlar arasındaki benzerliklerin Öklid uzaklığına göre bulunması sağlanmıştır.

> veri<-as.matrix(veri)# veri çerçevesini matrise dönüştürme işlemi yapılmışve alt komutta veri yazdırılmıştır.

```
> t(veri)# komutu ile matsin devriği yazdırılmıştır.
```

> dist(t(veri),method="euclidian")# Bu komutla çalışılan matrisin devriğinin satırları arasındaki benzerlik Öklid uzaklığına göre bulunmuştur.

Tablo 5. Öklid Uzaklığına Göre Benzerlik Gösterimi

Y1	Y2	Y3	Y4	Y5	Y6	Y7
Y2	2554.1120					
Y3	3055.5376	3466.5337				
Y4	2286.6005	3016.1137	3211.5850			
Y5	2410.3348	2897.5738	2447.4199	3157.1164		
Y6	1506.5789	3050.7602	3348.7207	2754.8510	2877.5927	
Y7	2568.6658	3070.9227	3480.6505	2938.9127	3224.2013	3023.7933
Y8	1255.6934	2269.7846	3243.2804	2608.5774	2550.0616	1835.6405
Y9	1559.0119	3025.8543	3378.2809	2628.6571	2929.8959	2104.1492
Y10	3319.5063	3213.6621	3820.9054	3581.0066	4125.5729	3665.3174
Y11	3672.4605	3712.1595	3652.7784	3890.6909	3764.3140	3976.5485
Y12	4905.6666	4905.6666	4905.6666	4905.6666	4905.6666	4905.6666
Y13	3490.4550	4221.5173	3347.4354	4297.2666	3846.1711	3069.1540
Y14	3412.9931	3882.1538	4001.8281	4025.9372	3932.2614	3425.2323
Y15	1884.2755	2942.7450	3261.7915	2966.2198	2906.7439	1130.9014
Y16	2122.3727	3043.6182	3680.5163	2921.6927	3152.5066	2570.5661
Y17	2932.0283	3447.8103	2861.9235	3712.0633	2913.0688	2813.5138
Y18	3193.7004	3758.4478	3193.0409	3926.4804	3414.4206	3211.2602
Y19	3416.1068	3143.2531	3501.9440	3912.9217	3661.2984	3688.2869
Y20	3776.7727	3601.3239	3732.8772	4020.3965	3796.3830	3726.9910
Y21						3733.0406

Y21	3410.3601	3769.4703	3945.1684	4030.5161	3930.4432	3472.4202	3923.9228
Y22	3064.4859	3715.4709	3410.9474	3762.0513	3239.7759	2905.7739	3480.3382
Y23	2361.9251	3025.7059	3729.7598	3040.1862	3283.6672	2797.6243	3584.9310
Y24	955.9446	2947.5583	3283.0900	2465.6812	2605.6562	1811.1615	2740.7798
Y25	2905.4053	3572.2175	3304.6007	3701.9887	3683.2703	2726.1317	3475.0827
Y8	Y9	Y10	Y11	X12	Y13	Y14	
Y9	2205.5825						
Y10	3572.0168	3047.0409					
Y11	3988.9435	3941.6974	3617.2061				
Y12	4905.6666	4905.6666	4905.6666	4905.6666			
Y13	3660.9026	3553.0660	3546.8803	4038.5406	4905.6666		
Y14	3536.7587	3110.2794	3214.3155	3249.1833	4905.6666	3125.5838	
Y15	1944.6223	2433.0512	3768.5432	4149.2995	4905.6666	3389.4250	3618.4955
Y16	2459.0771	2566.5226	3899.8426	3194.4154	4905.6666	4016.6914	3314.0652
Y17	2892.5169	3184.9060	3328.8124	3665.6320	4905.6666	3131.8196	2933.5600
Y18	3086.9415	3612.1935	3810.1835	3536.8772	4905.6666	3360.1211	3079.2853
Y19	3716.8105	3257.5948	3205.3056	3094.4983	4905.6666	4002.1199	3554.7602
Y20	3652.2092	4052.4879	3737.7576	3864.2505	4905.6666	3695.1057	3506.9756
Y21	3453.5144	3267.7362	3368.9446	3376.1674	4905.6666	2962.0395	1254.3205
Y22	3040.2286	3388.6868	3598.7817	3458.5910	4905.6666	2868.6579	2748.1008
Y23	2630.9196	2797.1922	3579.9925	3219.4978	4905.6666	4045.0377	2942.6777
Y24	1710.5631	1901.9201	3468.4025	3735.8346	4905.6666	3605.7182	3482.7814
Y25	3121.8937	2583.6807	2873.0169	3714.2420	4905.6666	2771.8730	2629.9015
Y15	Y16	Y17	Y18	Y19	Y20	Y21	
Y16	2808.6516						
Y17	2632.6181	3539.8850					
Y18	3135.5637	3321.2624	2964.3681				
Y19	3646.0743	3596.9626	3839.8081	3720.3365			
Y20	3763.1875	3874.2411	3722.2096	3166.4747	3861.8485		
Y21	3775.1013	3202.8915	3133.3340	2910.0942	3591.9201	3471.2027	
Y22	3138.5293	3650.3520	1925.1696	3021.5321	4128.7313	3545.8235	2988.2761
Y23	2897.8288	1331.1904	3197.0383	3390.2985	3688.0369	3761.0278	2908.1128
Y24	2178.6500	2391.3682	3041.3448	3330.6093	3601.8209	3960.2710	3554.5289
Y25	2825.6504	3683.9473	3031.3410	3022.4871	2984.7660	3050.9920	2888.8313
Y22	Y23	Y24					
Y23	3506.5636						
Y24	3124.2001	2499.9892					
Y25	3257.9822	3613.5838	3082.9921				

Tablo 5 incelendiğinde, Öklid uzaklığına göre Yargıcı 1 ve Yargıcı 24'ün başarılı-başarısız sınıflama kararlarına ait vektörler arasında benzerlik olduğu $d(Y1, Y2)=955.9446$ görülmektedir. Bu araştırmada da simülatif çok sayıda yargıcıya ait veri girişi yapılmasına rağmen küçük örneklemlerle çalışmalarındaki durum yinelenmiştir. Bu bulgu, sıkça simülatif çalışmalarda kullanılan R ortamında üretilen verilerin kullanımı açısından da soru işareti yaratmıştır. ROC ve R'da aynı bireyler için sınıflama doğruluğu uyumu var mıdır? sorusuna ilişkin bulgular aşağıda sunulmuştur;

ROC analizi ile R programlama dilini sınıflama doğrulukları açısından karşılaştırabilmek için tüm veri seti için yapılan ROC analizi, R ortamında sınırlandırılmış olan 30 kişilik veri üzerinde de yapılmıştır.

Tablo 6. Başarılı ve Başarısız Olma Durumuna Göre Dağılım

Lisans akademik başarı durumu	N	%
Pozitif (başarılı)	12	40
Negatif (başarısız)	18	60
Toplam	30	100

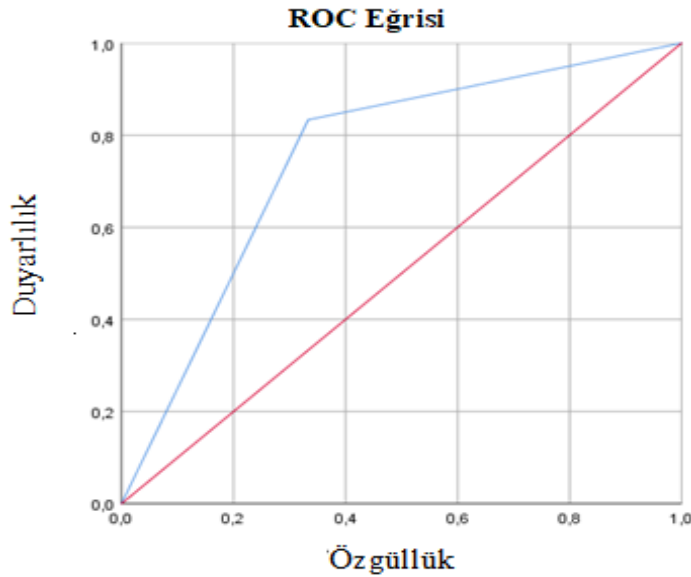
Tablo 6 incelendiğinde, 30 kişilik veri setinde lisans akademik başarı durumunda belirlenen başarı kesme puanına göre başarılı olan N= 12 (%40), başarısız olan N= 18 (%60) kişi görülmektedir.

Tablo 7. ROC Eğrisi Altında Kalan Alan

AUC	S.Hata	%95 Güven Aralığı Alt Sınır	%95 Güven Aralığı Üst Sınır	p
.750	.093	.568	.932	.022

($p < 0.05$)

Tablo 7 incelendiğinde, ROC analizi sonucunda, eğri altında kalan alan (AUC)=0.75 olarak hesaplanmış ve manidar ($p < .05$) bulunmuştur. Bu değer R ortamında da çalışılan 30 kişilik simülatif veri setine göre yapılan başarılı-başarısız sınıflamasının ROC analizinde %75 oranında orta düzeyde doğru sınıflandırıldığını ortaya koymuştur. Şekil 2’de görülen ROC eğrisinin, tüm veri setinden elde edilen Şekil 1’de görülen ROC eğrisine çok benzer olduğu görülmüştür.



Şekil 2. Karşılaştırma Yapılan Veri Setinin ROC Eğrisi (%95 Cl)

Küçük veri setinde yapılan ROC analizi sonucunda, başarılı olarak atanmış “12” ve başarısız olarak atanmış olan “18” kişi için lisans akademik başarı durumlarının, KPSS başarı sınıflamasına yönelik doğruluğu incelendiğinde, tüm veride olduğu gibi orta düzeyde bir sınıflama doğruluğunun olduğu görülmüştür. ROC analizi sınıflama, kümeleme konusunda oldukça iyi sonuçlar verebilen bir analiz olduğunu bu veri seti üzerinde de göstermiştir.

ROC ve R ortamında aynı veri grubu üzerinden sınıflama doğruluklarının karşılaştırılması yapıldığında ise, R programlama dilinde karara varmadan önce, simülatif verinin yanı sıra, uygulamaya dayalı verilerin de kullanımının yorumlamada sorun yaşanmaması için uygun olduğu ortaya çıkmıştır.

4. SONUÇ, TARTIŞMA VE ÖNERİLER

Sınıflama doğruluğunun belirlenmesi ve kullanılan istatistiklere göre karşılaştırılma amacı taşıyan bu araştırma sonucunda,

1. Varolan veriye dayalı olarak yapılan başarılı/başarısız olma sınıflandırılma doğruluklarının ROC analizi ile orta düzeyde saptanabildiği ortaya çıkmıştır.
2. Kesme puan belirlemede ROC analizinin duyarlılık, özgüllük başlıkları altında da iyi işlediği sonucuna ulaşılmıştır.
3. R programında simülatif veri setinde çok sayıda yargıcı ile yapılan sınıflama doğruluğu çalışmasında öklid uzaklığı ele alınmıştır. Bu istatistik sonucunda, yargıcıların karar verme aşamasında benzerlik gösterenlerin sayısı sadece bir çift olarak ortaya çıkmıştır. Diğer yargıcılar arasında sınıflama benzerliğine rastlanmamıştır.
4. ROC ve R'da aynı bireyler için sınıflama uyumu incelendiğinde ise uyuma rastlanmamıştır.

Bu sonuçlardan hareketle, ROC analizinin kümeleme, sınıflama ve veriyi kategorize etmek için kullanımının uygun olduğu söylenebilir. ROC ile belirlenen kesme puanlarının çalışma öncesinde belirlenen geçme puanı ile uyumlu olduğu da görülmektedir. R programlama dilinde yapılan çalışmaların örneklem büyüklüğü açısından sayısal üstünlüğü olduğu bir gerçektir. Ancak, pratikteki bilgilerin uygulamaya yansımalarının her zaman aynı olmayacağı düşünüldüğünde, tek başına kullanımından ziyade başka analizler ya da uygulamalarla desteklenmesi veri güvenilirliğini saptamak açısından önerilebilir.

5. KAYNAKÇA

- Adams, N. M., & Hand, D.J. (1999). Comparing classifiers when the misallocation costs are uncertain. *Pattern Recognition*, Vol. 32 (7) (1999) pp. 1139-1147.
- Akpınar, H. (2014), *Data: Veri madenciliği, veri analizi*. İstanbul: Papatya Yayıncılık.
- Alpar, R. (2010). Basit doğrusal regresyon çözümlemesi. spor, sağlık ve eğitim bilimlerinden örneklerle uygulamalı istatistik ve geçerlik güvenilirlik., Ankara: Detay Yayıncılık, 338-42.
- Cizek, G. J. (2001). Conjectures on the rise and fall of standard setting: an introduction to context and practice. G. Cizek (Eds.). *In setting Performance Standards: Concepts, Methods and Perspectives* (3-17). Mahwah, N. J.: Erlbaum.
- Çınar, A. (2019). Veri madenciliğinde sınıflandırma algoritmalarının performans değerlendirilmesi ve R dili ile bir uygulama. *Marmara Üniversitesi Öneri Dergisi*, 14(51), 90-111.
- Fawcett, T. (2005). An introduction to ROC analysis. *Pattern Recognition Lett.* 27, 861–874 (Special issue on ROC analysis).
- Ferri, C., Flach, P., Hernández-Orallo, J., & Senad, A. (2005). Modifying ROC curves to incorporate predicted probabilities. *In Proceedings of the Second Workshop on ROC Analysis in Machine Learning*.
- Green, D. M., & Swets, J. A. (1988). *Signal detection theory and psychophysics*. Reprint Edition. Los Altos, CA: Peninsula Publishing.
- Grove, W. M. (2006). Mathematical aspects of diagnosis. United States of America, 50-75
- Hand, D., & Till, R. (2001). A simple generalisation of the Area Under the ROC Curve for multiple class classification problems. *Machine Learning*, 45, 171–186.
- Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143, 29- 36.
- Hess, B., Subhiyah, R. G., & Giordano, C. (2007). Convergence between cluster analysis and the Angoff method for setting minimum passing scores on credentialing examinations. *Evaluation and the Health Professions*, 30(4), 362-375.
- Huang, J., & Ling, C.X. (2005). Using auc and accuracy in evaluating learning algorithms. *IEEE Transactions on knowledge and Data Engineering*, 17(3):299–310.

- Hurtz, G. M., & Hertz, N. R. (1999). How many raters should be used for establishing cutoff scores with the Angoff method? A Generalizability Theory Study. *Educational and Psychological Measurement*, 59 (6),885-897.
- Iglesias, J. A., Garcia-Cuerva, A., Ledezma, A., Sanchis, A., & Ieee. (2016, Oct 09-12). Social network analysis: evolving twitter. Paper presented at the IEEE International Conference on Systems, Man, and Cybernetics (SMC), Budapest, Hungary.
- Ihaka, R., & Gentleman, R. (1996). R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, 5(3), 299-314.
- Kanık E.A., & Erden S.(2003). Tanı testlerinin değerlendirilmesinde ROC eğrisinin kullanımı, *Mersin Üniversitesi Tıp Fakültesi Dergisi*.3:260-264.
- Karasar, N. (2018). *Bilimsel araştırma yöntemi*, Ankara, Nobel Yayınevi,33.Baskı.
- Kılınç, D., Borandağ, E., Yücalar, F., Tunalı, V., Şimşek, M., & Özçift, A. (2016). KNN algoritması ve R dili ile metin madenciliği kullanılarak bilimsel makale tasnifi. *Marmara Fen Bilimleri Dergisi*, 28(3), 89-94.
- Koyuncu, S. (2015). *Psikolojik ölçeklerde ROC analizi yöntemiyle standart belirleme* (Yüksek lisans tezi). Gazi Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.
- Krzanowski, J.W., & David, J.H. (2009). *ROC curves for continuous data*. Chapman and Hall/CRC.
- Lachiche, N., & Flach, P. (2003).Improving accuracy and cost of two-class and multi-class probabilistic classifiers using ROC curves. *In Proceedings of the Twentieth International Conference on Machine Learning*, Menlo Park, CA, USA, pp. 416–423. AAAI Press.
- Macskassy, S.A., & Provost, F. (2004). Confidence bands for ROC curves: Methods and an empirical study. *In Proceedings of the First Workshop on ROC Analysis in Artificial Intelligence*, pp. 61–70.
- Majnik, M., & Bosnić, Z. (2013). ROC analysis of classifiers in machine learning: A survey. *Intelligent data analysis*, 17(3):531–558.
- Metz, C. (1978). Basic principles of ROC analysis. *Semin. Nucl. Med.* 3 (4).
- Ömür, S., & Selvi, H. (2010). Angoff, Ebel ve Nedelsky yöntemleriyle belirlenen kesme puanlarının sınıflama tutarlılıklarının karşılaştırılması. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 1(2),109-113.
- Provost, F., & Fawcett, T. (2001). Robust classification for imprecise environments. *Machine Learning* 42, 203- 231.
- Swets, J. A., Dawes, R. M., & Monahan, J. (2000). Better decisions through science. *Scientific American*, 283, 82–87.
- Takahashi, K., Uchiyama, H., Yanagisawa, S., & Kamae, I. (2006). The logistic regression and ROC analysis of group-based screening for predicting diabetes incidence in four years applied medical statistics and decision sciences, *The Kobe Journal of Medical Sciece*, 52 (6), 171-80.
- Tan, P., Steinbach, M., & Kumar, V., (2006). *Introduction to Data Mining*. Boston: Pearson Education.
- Taşdelen, G. (2009). *Nedelsky ve Angoff standart belirleme yöntemlerinin genellenebilirlik kuramı ile karşılaştırılmasına ilişkin bir araştırma*. (Yayınlanmamış yüksek lisans tezi) Hacettepe Üniversitesi, Ankara.
- Tavakol, M., & Dennick, R. (2012). Standard setting: The application of the receiver operating characteristic method. *International Journal of Medical Education*, 3,198-200.
- Vivo, J. M., & Franco, M. (2008). How does one assess the accuracy academic success predictors? ROC analysis applied to university entrance factors. *International Journal of Mathematical Education in Science and Technology*, 39(3), 325-340.