





Özgün Araştırma/Research Article

Bilgisayar Uyarlamalı Test (BUT) Uygulamalarında Regresyon Ağacı Yaklaşımı: Regresyon Karar Ağaçları ile Psikometrik Model Kullanan Standart BUT Algoritmasının Yapay Bir Veri Üzerinde Değerlendirilmesi  Emrah Gökay ÖZGÜR ¹
 Beyza DOĞANAY ERDOĞAN ¹

Regression Tree Approach in Computer Adaptive Testing (BUT) Applications: Evaluation of Standard CAT Algorithm Using a Psychometric Model with Regression Decision Trees on Artificial Data

Özet

Amaç: Bu çalışma ile sağlık alanında kullanılan ölçeklerden yararlanarak, bireylerin özürsüzlük değerlendirimi gibi, incelen özellik düzeylerini belirlemede kullanılan psikometrik model temelli bilgisayar uyarlamalı test uygulamalarına regresyon ağacı yönteminin alternatif bir yaklaşım olarak tanıtılması, yapay bir veri üzerinde uygulanarak performansının incelenmesi amaçlanmıştır. Gereç ve **Yöntem:** 100 kişinin 100 maddeye verdiği iki sonuçlu (örn. evet/hayır) yanıtlar olarak türetilen yapay veri seti üzerinde Bilgisayar Uyarlamalı Test ve Regresyon Ağacı yöntemlerinin tahmin performansları incelenmiştir. İnceleme, tahmin değerleri, madde sayıları ve hata değerleri bakımından gerçekleştirilmiştir. Uygulamalarda R v.3.6.3 programı kullanılmıştır. **Bulgular:** Bilgisayar Uyarlamalı Test yönteminde hata değeri 2,45 çıkarken regresyon ağacı yönteminde 4,04 çıkmıştır. Bilgisayar Uyarlamalı Test yaklaşımı ortalama 42 (minimum:41 maksimum:43) madde ile tahmin yaparken regresyon ağacı 23 madde ile tahmin yapmıştır. **Sonuçlar:** Karşılaştırılan değerler bakımından Bilgisayar Uyarlamalı Test yaklaşımı ve Regresyon Ağacı yöntemleri benzer sonuçlar vermiştir. Kullanılmış olduğumuz veri seti için Regresyon Ağaçları yönteminin Bilgisayar Uyarlamalı Test yaklaşımına bir alternatif olarak kullanılabileceği yanında, daha geliştirilmiş ağaç algoritmalarının incelenmesi gerekliliği görülmüştür.

Anahtar Kelimeler: bilgisayar uyarlamalı test, regresyon ağacı, ölçek

Alındığı tarih/Received Date:
23.05.2020

Kabul tarihi/Accepted Date:
01.06.2020

Sorumlu yazar:

Emrah Gökay ÖZGÜR

e-mail:

emrahgokayozgur@gmail.com

¹ Ankara Üniversitesi, Tıp Fakültesi, Biyoistatistik Anabilim Dalı, Türkiye.

Abstract

Objective: By using the scales used in the field of health with this study, it is to examine whether or not the regression tree method can be an alternative approach to the computer adaptive test applications based on psychometric model used to determine the level of examined features. **Methods:** The predictive performances of Computer Adaptive Test and Regression Tree methods were evaluated on artificial data set derived from 100 people's dichotomous responses (e.g. yes/no) to 100 items. Evaluations were made in terms of estimation values, item numbers and error values. R v.3.6.3 program is used in applications. **Results:** While the error value was 2.45 in the Computer Adaptive Test method, it was 4.04 in the regression tree method. While the Computer Adaptive test approach predicted with an average of 42 items (minimum:41-maximum:43), the regression tree estimated with 23 items. **Conclusion:** In terms of the compared values, computer adaptive test approach and regression tree methods gave similar results. It has been seen that, as well as the regression trees method can be used as an alternative to the computer adaptive test approach for the data set we use, advanced tree algorithms should be examined further..

Key words: computerized adaptive testing, regression tree, scale

GİRİŞ

Herhangi bir sağlık sisteminin merkezi “hasta” olarak kabul edilebilir. Son yıllarda hasta-merkezli sağlık sistemleri giderek artan önem kazanmaktadır. Gelecekte, hastaya uygulanan klinik bir müdahalenin hasta tarafından bildirilen sonuçlarının klinik, fizyolojik ya da hastaya bakan kişinin raporu gibi diğer sonuçlara göre, daha önemli olacağı öngörülmektedir. Fiziksel fonksiyon, semptomlar, global sağlık değerlendirmesi, psikolojik iyi hal, sosyal iyi hal, bilişsel fonksiyon, sağlığa bağlı yaşam kalitesi gibi özelliklere ilişkin ölçümler hasta tarafından bildirilen sonuçlar kullanılarak yapılabilir (Deshpande ve ark., 2011).

Özellikle kanser gibi hastalıklarda, yaşam kalitesinin hastalığın ilerlemesine bağlı olduğu durumlar olabilir. Kanser ilerlemesine bağlı olarak hastalarda sıklıkla birden fazla semptom görülmesi, hastalığın getirdiği ekonomik yük, ev yaşamındaki olumsuz etkiler ve moral düzeyinin olumsuz etkilenmesi hastaların yaşam kalitesini etkileyen faktörlerdir. Bu gibi durumlarda hasta tarafından bildirilen sonuçlar kanser hastalarında yaşam kalitesinin belirlenmesinde yardımcıdır.

Bu örneklerdeki gibi, sağlık alanında doğrudan ölçüm yapılamayan örtük değişkenlerin ölçülmesinde ölçekler kullanılır. Ölçekler uygun maddelerden oluşan, güvenilir, geçerli ve değişime duyarlı ölçme araçlarıdır. Ölçeklerden elde edilen sonuçlar bireylerin incelenen özellik (θ) düzeyleri hakkında değer yargılarına varmada kullanılır.

Bu çalışmada sağlık alanında sıklıkla kullanılan ölçeklerden yararlanarak bireylerin incelenen özellik bakımından (θ) düzeylerini belirlemede kullanılan psikometrik model temelli Bilgisayar Uyarlamalı Test uygulamalarına Regresyon Ağacı yöntemi alternatif bir yaklaşım olup olamayacağı yapay bir veri üzerinde uygulanarak incelenecektir.

BİREYLER ve YÖNTEM

Geleneksel veya sabit formlu testte bir kağıt - kalem formu veya bir bilgisayar kullanılarak aynı öğeler kişiler arasında uygulanır (Rezaie ve ark., 2015). Uyarlamalı bir test, farklı kişilere aynı madde kavramının dışına çıkarak her kişiye göre zorluğu ayarlar ve kişilerin performanslarını değerlendirmek için farklı zorlukta maddeler uygular. Bu bilgisayar uyarlamalı testi (BUT) diğer geleneksel yaklaşımlardan ayıran en önemli özelliktir.

Uyarlamalı bir testin temellerinden biri, test deneyiminin dinamik ve test katılımcısının performansına duyarlı olmasıdır (Rezaie ve ark.). Uyarlanabilir test fikri ilk olarak Binet tarafından 1905 yılında zeka testi ile kullanılmıştır (Öztuna, 2008). Binet bu testte birbirini izleyen her soruyu daha önceki sorulardaki performanslarına göre seçerek uyguladı ve IQ testinin hem etkili hem de yanıt verenin becerisine uygun olmasını sağladı. Burada amaç gruptan ziyade herkes için performansına göre ayrı sorular seçerek kişilerin bilgi düzeyini (θ) tahmin edebilmektir (Rudner, 1998).

Kalibre edilmiş soru bankası kullanılarak testin uygulanacağı her kişi için belirli bir kurala bağlı olarak seçilen başlangıç sorusu ile test başlar ve kişinin incelenen özellik düzeyi (θ) tahmini yapılır (Öztuna, 2008). BUT yönteminde kişinin ilk maddeye verdiği cevap, sonraki maddelerin sıralamasının belirlenmesinde önemli rol oynar (Rudner, 1998). Eğer ilk maddeye doğru cevap verildiyse, bir sonraki madde daha zor, yanlış cevap verildiyse bir sonraki madde daha kolay olacaktır. Bu yaklaşımın arkasındaki mantık, çok kolay ya da çok zor maddelerden bireyin (θ) düzeyi hakkında bilgi edinilemeyeceği, dolayısıyla maddelerin bireyin (θ) düzeyi hakkında en fazla bilgiyi sağlayacak şekilde seçildiğidir (Rezaie ve ark., 2015). Sonrasında belirlenen madde seçim kuralına göre yeni soru seçimi yapılarak yine θ kestirimi yapılır. Soru seçimi ve θ kestirimi basamakları belirlenen durdurma kriteri sağlanana kadar iteratif olarak devam eder (Öztuna, 2008). Durdurma kuralı karşılandığında test biter ve en son yapılan θ kestirimi kişinin incelenen özellik düzeyi kestirimi olarak alınır. Bütün BUT algoritmaları bu temel formata dayalıdır. Ancak soru seçme kuralları gibi bazı detaylar değişkenlik gösterir (Rudner, 1998).

İncelenen özellik düzeyi kestirimi ve soru seçiminde, klasik istatistik ve Bayesci istatistik yaklaşımları kullanılabilir. İncelenen özellik düzeyi için en sık kullanılan kestirim yöntemleri, klasik istatistik çerçevesinde geliştirilen En Çok Olabilirlik (Maximum Likelihood, ML), Ağırlıklandırılmış En Çok Olabilirlik (Weighted Maximum Likelihood, WML) ve Bayesci istatistik çerçevesinde geliştirilen Sonsal Beklenti (Expectation a Posteriori, EAP) ile Sonsal Maksimum (Maximum a Posteriori, MAP) kestirim yöntemleridir. Soru seçme yöntemleri o anki θ kestiriminde bilgi kriterini maksimum

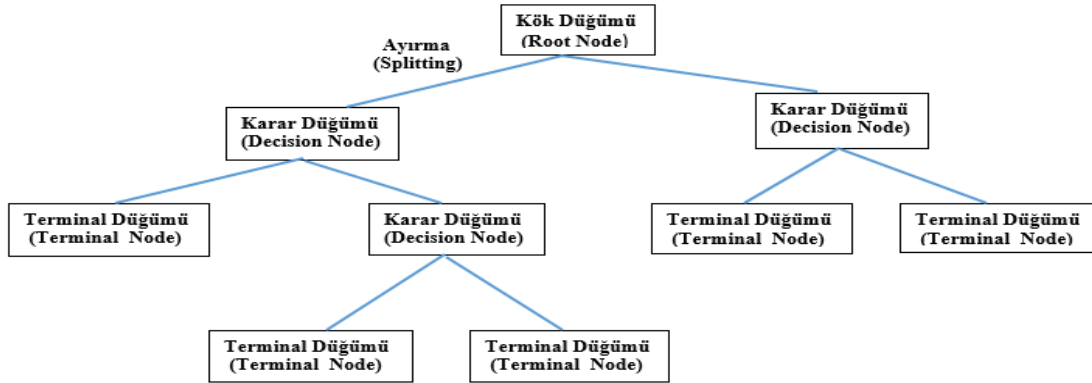
yapan sorunun seçilmesine dayalıdır. En çok kullanılan klasik istatistik soru seçme kriteri Fisher'in En Çok Bilgi Kriteri (Maximum Fisher's Information Criterion, MFI) ve Bayesci istatistik kriteri ise Sonsal Dağılım Ağırlıklı Bilgi Kriteri (Maximum Posterior Weighted Information, MPWI) yöntemleridir. Bu yöntemlere ek olarak, özellikle eğitim alanında, aynı soruya maruziyet, test kapsamındaki her içerikten soru gelmesi zorunluluğu gibi pratikte göz önünde bulundurulması gereken durumlar da vardır. Ancak sağlık alanında soru seçiminde bu tip pratikte göz önüne alınacak durumlar bulunmamaktadır.

Bu çalışma kapsamında BUT algoritması için incelenen özellik kestirimi WML ile soru seçimi ise MFI yöntemi ile gerçekleştirildi. Algoritmada durdurma kriteri olarak incelenen özellik düzeyi kesitiriminin standart hatası 0.33 olarak belirlendi. BUT algoritması R v.3.6.3 programında catR kütüphanesindeki randomCAT fonksiyonu ile gerçekleştirildi.

Regresyon Ağacı (Regression Tree)

Karar ağaçları, hem sınıflandırma hem de regresyon görevlerini gerçekleştirebilen çok yönlü makine öğrenme algoritmasıdır (Michel ve ark.,

Şekil. Karar Ağacı Algoritması



Regresyon ağaçlarında aşırı uyum sorunu ortaya çıkabilmektedir. Ayrırma işlemleri ve dallanmalar ne kadar fazla olursa o kadar aşırı uyum meydana gelecektir. Bu durumda karmaşıklık parametresi sayesinde hata belirlenen bir değere gelindiğinde dallanma durur ve böylece aşırı uyum sorunu ortadan kalmış olur(Michel ve ark., 2018).

Bu çalışmada R v.3.6.3 programında rpart kütüphanesinin predict, summary ve plot fonksiyonları kullanılmıştır. Regresyon ağacının

2018). Bunlar karmaşık veri setleri için çözüm oluşturan çok güçlü algoritmalarıdır Ağaç tabanlı öğrenme algoritmaları en iyi ve en çok kullanılan denetimli öğrenme yöntemlerinden biri olarak kabul edilir(James ve ark., 2014). Ağaç temelli yöntemler, kestirim amaçlı kullanılan modelleri yüksek doğruluk, kararlılık ve yorumlama kolaylığı ile güçlendirir.

Bağımlı değişkenin türüne göre kullanılacak karar ağacı seçilmektedir (Breimann ve ark. 1984). Bağımlı değişken kategorik olduğunda kullanılacak karar ağacı modeli sınıflandırma ağaçları, bağımlı değişken sürekli olduğunda kullanılacak karar ağacı modeli regresyon ağacıdır (Breimann ve ark. 1984). Bu yöntemde heterojen veri setleri belirlenmiş hedef değişkene (bağımlı değişken) göre homojen alt gruplara ayrılır. Alt gruplara ayırma işlemi belirli bir takım karar verme kuralları ile yapılmaktadır. Dallanma kök düğümden terminal düğümlere doğru saflık ölçümleri yapılarak gerçekleştirilir. Entropi, Ki-kare, Varyans Azaltma Kriteri gibi değerler ile oluşan düğümlerdeki homojenlik yapısı incelenerek bir dallanma olup olmayacağı kontrol edilir (Dangeti, 2017). Terminal düğümden ortaya çıkan değerler bağımlı değişkenin tahmini değerleridir.

diyagami için rpart.plot kütüphanesi ve hata değerlerini hesaplamak için caret kütüphanesinden defaultSummary fonksiyonu kullanılmıştır. Regresyon ağacı uygulamasında minsplit: 5, cp:0.001 ve max.depth: 30 olarak alınmıştır.

Yapay Veri Türetilmesi

Türetilen veri seti için madde sayısı 100, kişi sayısı 100 olarak belirlenmiştir. Kişilere ait parametreleri içeren θ vektörü, minimum -4 logit, maksimum 4 logit olacak şekilde ($\theta_{maksimum}$ -

$\theta_{\text{minimum}}/(\text{kişi sayısı}-1)$ birimlik artışlarla oluşturulmuştur. Madde zorlukları, minimum -4 logit, maksimum 4 logit; madde sayısı da 100 olacak şekilde $(\beta_{\text{maksimum}}-\beta_{\text{minimum}})/(\text{madde sayısı}-1)$ birimlik artışlarla oluşturulmuştur. Belirlenen θ ve β parametreleri ile her maddeye ait $P(y_j=1 | \theta_i)$ olasılığı iki sonuçlu Rasch modeline göre hesaplanarak Uniform(0,1) dağılımdan çekilen rastgele u değişkeni bu olasılık değeri ile karşılaştırılmıştır.

Eğer rastgele u değeri karşılaştırılan olasılıktan büyük ise o maddeye yanıt 1, küçük ise 0 olarak belirlenmiştir. Bu işlemden sonra her bireyin toplam skoru hesaplanarak aynı veri seti içine yazdırılmıştır. Regresyon ağacı yönteminde çıktı değişkeni 100 maddeden alınan toplam skor, girdi değişkenleri ise 100 maddeye verilen yanıtlar olarak alınmıştır. Klasik BUT algoritmasında ise psikometrik model olarak iki sonuçlu Rasch modeli seçilmiştir.

Çalışmada kullanılan veriler R V3.6.3 programında genresp fonksiyonu ile türetilmiştir.

BUT algoritması ile bulunan Rasch skorlarının toplam skora dönüştürülmesi için 1.1’de ki formül kullanılmıştır.

$$y = m + (s * \text{Bireyin Yetenek Düzeyi}) \quad 1.1$$

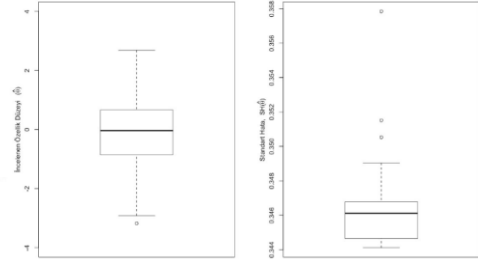
$$s = (\text{İstenen Aralık}) / (\text{Kullanılan Aralık}) \quad 1.2$$

$$m = (\text{İstenen minimum değer}) - (\text{Kullanılan minimum değer} * s) \quad 1.3$$

BULGULAR

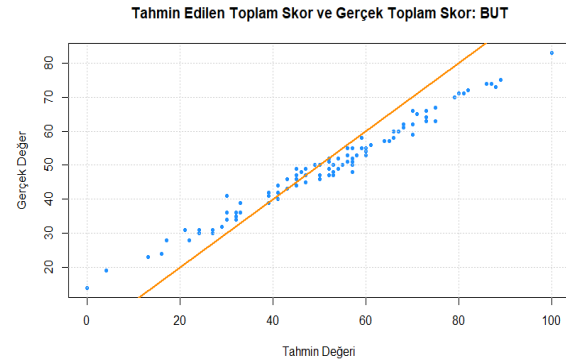
Bilgisayar Uyarlamalı test sonuçlarına göre bireylerin Rasch skorları ortalama 42 (min:41-maks:43) madde ile tahmin edilmiştir. RMSE değeri 2,45, MAE değeri ise 2,03 olarak hesaplanmıştır (Tablo 1). Grafik 1’de görüldüğü gibi BUT sonucunda bireylerin incelenen özellik değeri tahminin Rasch skoru -4 ve 4 arasında 0 civarında iken, standart hatası ise 0,346 seviyesindedir. Formül 1.1 ile Rasch skorları toplam skora dönüştürülmüştür. 49,77 olan bireylerin gerçek ortalama skoru BUT ile 52,07 olarak tahmin edilmiştir.

Grafik 1. BUT Yöntemi İle Bireylerin İncelenen Özellik Değeri Tahmini ve Standart Hatası



BUT ağacı sonucunda ortaya çıkan tahmin değerlerinin gerçek değerler ile ilişkisi Grafik 2’de verilmiştir.

Grafik 2. BUT için Tahmin Edilen Değer ve Gerçek Değer İlişkisi



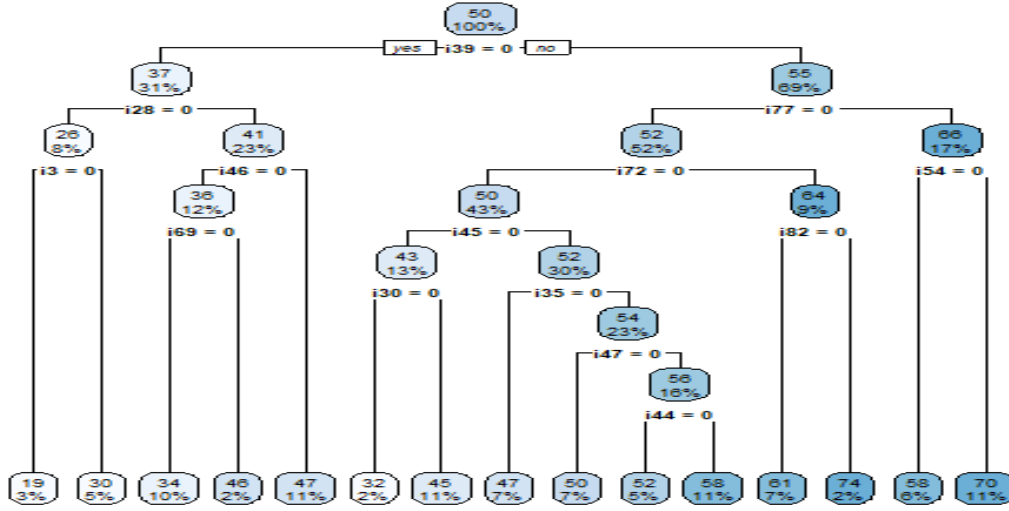
Oluşturulan regresyon ağacına göre bireylerin toplam skorları 23 madde ile tahmin edilmiştir.

RMSE değeri 4,04, R2 değeri 0,91 ve MAE değeri 3,28 olarak hesaplanmıştır (Tablo 1). 49,77 olan bireylerin gerçek ortalama skoru regresyon ağacı yöntemi ile 49,77 olarak tahmin edilmiştir. Maddelerin önem düzeyleri incelendiğinde sırasıyla 39, 77, 72, 30 ve 28. maddeler en önemli olanlarıdır. İlk ayrımı gerçekleştiren 39. sorudur. Regresyon ağacının sol tarafa doğru olan dallanmalarını incelediğimizde sırasıyla 39, 28 ve 3. sorulara sıfır yanıtı veren bireylerin toplam skoru 19 olarak tahmin edilmiştir. Regresyon ağacının sağ tarafa doğru dallanmalarını incelediğimizde sırasıyla 39, 77 ve 54. Sorulara bir yanıtı veren bireylerin toplam skoru 70 olarak tahmin edilmiştir (Şekil 2). Regresyon ağacına göre bireylerin %3’ ünün toplam skoru 19, %5’ inin 30, %10’ unun 34, %2’ sinin 46, %11’ inin 47, %2’ sinin 32, %11’ inin 45, %7’ sinin 47, %7’ sinin 50, %5’ inin 52, %11’ inin 58, %7’ sinin 61, %2’ sinin 74, %6’ sının 58 ve %11’ inin 70 olarak tahmin edilmiştir (Şekil 2).

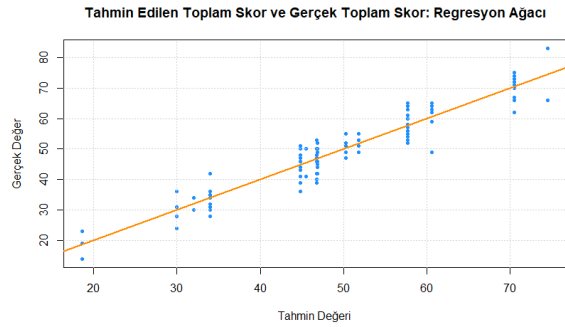
Regresyon ağacı sonucunda ortaya çıkan tahmin değerlerinin gerçek değerler ile ilişkisi

Grafik 3’ de verilmiştir. Bu grafiğe göre gerçek değerlerin tahmin değerleri ile yüksek derecede ilişkili olduğu görülmektedir.

Şekil 2. Regresyon Ağacı



Grafik3. Regresyon Ağacı için Tahmin Edilen Değer ve Gerçek Değer İlişkisi



Tablo 1. Yöntemlere Ait Tahmin Performansı Değerleri

	BUT	Regresyon Ağacı
RMSE	2,45	4,04
R ²	-	0,91
MAE	2,03	3,28
Madde Sayısı	42	23
Tahmini Değer	52,07	49,77

RMSE: Root Mean Squared Error, MAE: Mean Absolute Error

TARTIŞMA

Michel ve ark. 2016 yılında yaptığı çalışmada regresyon ağacı ve bilgisayar uyarlamalı test yaklaşımının tahmin performanslarını karşılaştırmıştır. Bu çalışmada

özellikle skorların tahmin değerleri arasında bir uyum tespit edilmiştir. Regresyon ağacı ile bireylerin toplam skorunu 50,09, bilgisayar uyarlamalı test ile ise 50.00 olarak tahmin etmiştir. Ayrıca bu iki yöntemin hata değerlerine bakıldığında ise her iki yöntem için de en iyi olduğu modellerin hata değerleri regresyon ağacı için 0,16, bilgisayar uyarlamalı test için ise 0,22 olarak bulunmuştur. Sonuç olarak Michel ve ark. regresyon ağacının bilgisayar uyarlamalı test yaklaşımına göre alternatif olarak kullanabileceğini göstermiştir (Michel ve ark., 2018). Burada sunulan sonuçlar henüz yürütülen çalışmanın ön sonuçları olmakla birlikte, ağaç temelli ileri yöntemlerin (bagging, boosting, random forest algoritmaları gibi) daha kapsamlı bir benzetim çalışması ile incelenmektedir. Ne var ki burada sunduğumuz sonuçlarda, Rasch modelden türetilen yapay veri için sağlık alanında regresyon ağacı yönteminin BUT yöntemine alternatif bir yöntem olabileceği gösterilmiştir.

Squires ve ark. 2019 yılında yaptığı çalışmada diyabetik lomber omurga cerrahi hastalarında fiziksel fonksiyonların değerlendirilmesi için BUT yöntemini kullanmışlardır (Squires ve ark., 2019).

Boudreaux ve ark. 2019 yılında yaptığı çalışmada acil serviste intihar ilişkisi taramasını BUT yöntemi ile yapmışlardır (Boudreaux ve ark., 2019).

Edmond ve ark. 2019 yılında yaptığı çalışmada kronik bel ağrılı hastaların yön tercihi, bilişsel davranışsal girişimleri ve sonuçlarını BUT yöntemi analiz etmişlerdir (Edmond ve ark., 2019).

Chester ve ark. 2019 yılında yaptığı çalışmada öz yeterlilik ve kalıcı omuz ağrısı riskini regresyon ağacı yöntemi ile analiz etmişlerdir (Chester ve ark., 2019).

SONUÇLAR VE ÖNERİLER

Sağlık alanında sıklıkla kullanılan ölçeklerden yararlanarak bireylerin incelen özellik bakımından (θ) düzeylerini belirlemede kullanılan psikometrik model temelli bilgisayar uyarlamalı test uygulamalarına regresyon ağacı yönteminin alternatif bir yaklaşım olup olamayacağı yapay bir veri üzerinde incelenerek sonuçları ortaya konulmuştur.

Bireylerin toplam skor ortalaması 49,77' dir. BUT yönteminde ise bu değer 52,07 olarak tahmin edilirken regresyon ağacında 49,77 olarak tahmin edilmiştir. Regresyon ağacı yöntemi ortalama skor bakımından mükemmel bir tahmin performansı ortaya koymuştur. İki yöntemin hata değerlerine baktığımızda ise BUT yönteminde

RMSE değeri 2,45 iken regresyon ağacı yönteminde ise 4,04' tür. Bağımlı değişkeni açıklama bakımından regresyon ağacı %91 ile yüksek bir açıklayıcılığa sahiptir. BUT yöntemi ortalama 42 (min:41-maks:43) madde ile tahminlerini gerçekleştirirken regresyon ağacı yöntemi 23 madde ile tahminlerini yapmıştır. Sonuç olarak değerlendirmeye aldığımız parametreler açısından regresyon ağacı yönteminin BUT yöntemine benzer sonuçlar verdiği görülmektedir.

Makine öğrenmesi algoritmalarının sağlık alanındaki veriler üzerinde uygulanmaya başlanması ile, farklı uygulamalara adaptasyonu ve performans değerlendirimi çalışmaları da hız kazanmıştır. Ancak çalışmamızın kapsamı incelendiğinde özellikle sağlık alanında bu tarz çalışmaların sıkça olmadığı gözlemlenmiştir. Bu açıdan çalışmamız bu konuda çalışmak isteyen araştırmacılara önemli kazanımlar sağlayacağını düşünmekteyiz. Bu çalışmanın daha iletilmesi açısından ilerleyen zamanlarda daha büyük veri setlerinde ve sağlık alanında gerçek veriler ile kullanmış olduğumuz yöntemlere ek yöntemlerle karşılaştırılmaların yapılması planlanmaktadır.

KAYNAKLAR

- Breiman L., Friedman J. H., Olshen R. A., Stone C. J. 1984. Classification and Regression Tree. 1; 3-7.
- Boudreaux E.D., De Beurs D.P., Nguyen T.H., Haskins B.L., Larkin C., Barton B.2019. Applying Computer Adaptive Testing Methods to Suicide Risk Screening in the Emergency Department. *Suicide Life Threat Behav.* 49(4):917-927.
- Chester R., Khondoker M., Shepstone L., Lewis J.S., Jerosch-Herold C.2019. *Br J Sports Med.* 53(13):825-834
- Dangeti P. 2017. *Statistical for Machine Learning.*1; p:126-127
- Deshpande G.C., Rao S.C., Keil A.D., Patole S.K. 2011. Evidence-based Guidelines for Use of Probiotics in Preterm Neonates. *BMC Med.* 2;9:92.
- Edmond S.L., Werneke M.W., Young M., Grigsby D., McGill T, McClenahan B. 2019. Directional Preference, Cognitive Behavioural Interventions, and Outcomes Among Patients With Chronic Low Back Pain. *Physiother Res Int.* 24(3): e1773.
- Healy B.C., Zurawski J., Gonzalez C.T., Chitnis T., Weiner H.L., Glanz B.I. 2019. Assessment of Computer Adaptive Testing Version of the Neuro-QQL for People With Multiple Sclerosis. *Mult. Scler.* 25(13):1791-1799.
- James G., Eitten D., Hastie T., Tibhirani R. 2014. *An Introduction to Statistical Learning with Applications in R.* 4. p:320-322
- Michel P., Baumstarck K., Loundou A., Ghattas B., Auquier P., Boyer L.2018. Computerized adaptive testing with decision regression trees: an alternative to item response theory for quality of life measurements in multiple sclerosis. *Patient Prefer Adher.*12;1043:1053.
- Rezaie M., Golshan M.,2015. Computer Adaptive Test(CAT): Advantages and Limitations. *IJEI.* 2(5),128-137.
- Rudner, L.M. (1998). *An On-Line, Interactive, Computer Adaptive Testing Tutorial.* [<http://edres.org/scripts/cat>](Accessed 2020 March 20).
- Squires M.D., Brodke D.S., Neese A., Zhang Y., Spiker W.R., Lawrence B., Spina N. 2019. Physical Function Computer Adaptive Test Outcomes in Diabetic Lumbar Spine Surgical Patients. *Spine J.* 19(6):1048-1056.