

OKUMA BECERİLERİNE YÖNELİK MADDELERİ YANITLAMA HIZLARINI YORDAYAN ÖZELLİKLERİN BELİRLENMESİ¹

DETERMINATION OF THE FEATURES PREDICTING THE SPEED OF RESPONDING TO THE ITEMS RELATED TO READING SKILLS

İzzettin AYDOĞAN²

Selahattin GELBAL³

Başvuru Tarihi: 16.12.2021 Yayına Kabul Tarihi: 27.05.2022 DOI: 10.21764/maeuefd.1037681
(Araştırma Makalesi)

Özet: Bu araştırmayla öğrencilerin okuma becerilerine yönelik maddeleri yanıtlama hızlarını yordayan özelliklerin belirlenmesi amaçlanmıştır. Araştırmanın çalışma grubunu, PISA 2015 programına katılan 5232 onbeş yaş grubu öğrenci oluşturmuştur. Araştırma verileri, PISA 2015 programı verileri üzerinden sağlanmış olup genetik algoritmalar yöntemi kestirimine dayalı regresyon modeli esasıyla analiz edilmiştir. Analizler R programı kullanılarak gerçekleştirilmiştir. Genetik algoritmalar yöntemi ile okuma becerilerine yönelik maddeleri yanıtlama hızlarını en iyi derecede yordayan değişkenlerden oluşan regresyon modeli için değişken seçim işlemi yapmak istenmiştir. Ulaşılan sonuçlara göre, cinsiyet, evdeki kitap sayısı, evde konuşulan dil, okuma becerisi, eylem sayısı ve okulda okuma becerileri için ayrılan haftalık ders saati değişkenlerinin öğrencilerin okuma becerilerine yönelik maddeleri yanıtlama hızlarını istatistiksel olarak anlamlı düzeyde yordadığı saptanmıştır. Yordama düzeyi anlamlı bulunan değişkenlerdeki farklılaşmanın öğrencilerin okuma becerilerini ölçen maddeleri yanıtlama hızlarında da anlamlı düzeyde farklılaşmaya yol açtığı anlaşılmıştır. Öğrencilerin okuma becerilerine yönelik maddeleri yanıtlama hızlarını istatistiksel olarak anlamlı yordayan değişkenlerin okuma becerilerini ölçen maddeleri yanıtlama hızlarındaki değişkenliğin %8.53'ünü açıkladığı gözlenmiştir.

Abstract: The current study aimed to determine the features predicting students' response speed for items related the reading skills. The study group of the study is comprised of 5232 students aged 15 years participating in the PISA 2015 program. The data of the study were obtained over the PISA 2015 program and analyzed by using the regression model based on the genetic algorithms method estimation. Analyzes were carried out using the R program. It was intended to perform a feature selection process for the regression model, which consisted of the variables that best predicted response speed for items related the reading skills with the method of genetic algorithms. The results obtained revealed that the variables of gender, number of books at home, language spoken at home, reading skills, number of actions and weekly lesson hours for reading skills at school significantly predicted the students' response speed for items related the reading skills. The variation in the variables whose prediction power was found to be significant was understood to lead to a significant variation in students' response speed for items measuring the reading skills. It was observed that the variables that statistically significantly predicted students' response speed explained 8.53% of the variability in response speed.

Anahtar Sözcükler: *Yanıtlama hızı, yordayıcı özellikler, regresyon, genetik algoritmalar*

Keywords: *Response speed, predictor features, regression, genetic algorithms*

¹ Bu çalışma, birinci yazarın ikinci yazar danışmanlığında hazırladığı doktora tezinden üretilmiştir.

² Dr., Milli Eğitim Bakanlığı, Van İl Müdürlüğü, Van, ayd.izzet@gmail.com, ORCID: 0000-0002-5908-1285

³ Prof. Dr., Hacettepe Üniversitesi, Eğitim Bilimleri Bölümü, Eğitimde Ölçme ve Değerlendirme Bilim Dalı, Ankara, gelbal@hacettepe.edu.tr, ORCID: 0000-0001-5181-7262

Giriş

Okuma, bireylerin yazılı olan bir metni harf ve sözcük işaretlerini tanıyarak anlamlandırması olup; ses, dokunma, görme gibi fizyolojik ve anlamlandırma yönüyle zihinsel bir süreç gerektirmektedir (Gürses, 1996; Kantemir, 1995). Okuma eylemi, yazılı ya da basılı olarak karşılaşılabilen işaretlerin duyu organlarınca algılanmasıyla başlar, bu durum psikomotor bir süreç gerektirmektedir. Daha sonra sinirlerce algılanan işaretler zihinde değerlendirilerek anlamlandırılır, bu durum ise zihinsel süreç içerisinde gerçekleşir (Temizkan, 2007). Okuma becerileri, bireylerin kelime ve cümleleri zihinde anlamlı bir bütün haline getirmelerine yönelik yetenekler içerir. Bu durum zihinsel bir model oluşumuyla tarif edilmektedir. Bireylerin zihinlerinde oluşturdukları bu model, tecrübe edilmiş ya da yordanan bilgi, görüntü ya da seslerden oluşan ve sözcük ya da cümleleri birleştirerek anlamlandırılan bir sürece karşılık gelmektedir. Okunan metin içerisindeki olay ve kurguların aktarılacağı uygun bir zihinsel modelin oluşturulması anlamayı, düşünmeyi ve hatırlamayı daha kolay hale getirmektedir (Oakhill vd., 2015).

Yüksek düzeyde bir okuma becerisi cümle bilgisi, içerik, hedef vb. öğelerin sağladığı ipuçları ve hız olmak üzere büyük ölçüde iki ölçütle ilişkili olmaktadır. Okuma hızı okuduğunu anlama açısından oldukça önemli olup, düşük okuma hızı okuduğunu anlama etkinliğini olumsuz etkilemektedir. Bu durum, işleyen belleğin bilgiyi işleme ve depolama biçimiyle açıklanmaktadır. Okuma hızının düşük olmasıyla bilgi transferinin yavaş olacağı ve buna bağlı olarak işleyen belleğe aşırı yük binmesi ve okuduğunu anlamada hatalara sebep olacağı görüşü savunulmaktadır (Williams, 2003). Hızı, akıcı okumanın bir ögesi olarak kabul eden yaklaşımlar da söz konusudur (Wolf & Katzir-Cohen, 2001). Akıcı okuma; metnin hızlı, doğru ve içerdiği ifadelerin uygun şekilde okunması olarak tanımlanmaktadır (National Reading Panel, 2000). Akıcı okumanın üç bileşeninden biri olan hız, okumanın otomatikleşmesi olarak ifade edilebilir. Otomatikleşme, kelimelerin az bir bilişsel çaba ve dikkatle okunmasına karşılık gelmektedir. Otomatiklik, kelimelerin doğru ve hızlı şekilde tanınmasının yanı sıra metnin hızlı bir şekilde anlaşılıp, anlamın zihinde takip edilmesine bağlıdır. Otomatiklik okuma becerisi ile yakından ilişkilidir. Metnin yavaş ve kesik kesik okunması durumunda anlam bütünlüğü kurulamaz, cümle içi ve cümleler arasındaki anlam bağlantıları anlaşılabilir (Başaran, 2013). Akıcı okumanın okuma

becerileri ile ilişkili olduğunu vurgulayan farklı çalışmalar (Canizo vd., 2015; Veenendaal vd., 2015) söz konusudur.

Hız ile paralel olan yanıtlanma (tepki) sürelerinin, zihinsel aktivitelerin daha iyi anlaşılması amacıyla son zamanlarda yapılan araştırmalara konu olduğu gözlenmektedir (Schnipke & Scrams, 2002). Maddelerin cevaplanmasında yanıtlanma sürecinin yanıtlanma süreleriyle ilişkisinin önemi birçok araştırmada ön plana çıkmaktadır (Dodonova & Dodonov, 2013; Goldhammer vd., 2015; Greiff vd., 2016; Lee & Haberman, 2016; Michaelides vd., 2020; Wise, 2017). Alanyazın incelendiğinde, yanıtlanma sürelerinin farklı değişken durumlarla ilişkilendirildiği gözlenmektedir. Bunlardan biri, yanıtlanma sürelerinin test katılımında istekli olma ya da öğrenci motivasyonunun bir göstergesi olduğu üzerinedir. Yanıtlanma sürelerinin aşırı düşük ya da yüksek olmasının öğrencinin test performansı için bir geçerlilik kanıtı olabileceği savunulmaktadır (Rios vd., 2017; Wise, 2017; Wise & DeMars, 2005; Wise & Kong, 2005). Michaelides ve diğerlerinin (2020) yapmış oldukları çalışmada maddeleri yanıtlanma sürelerinin yüksek ve düşük olduğu öğrencilerin test performanslarındaki farklılaşma dikkat çekmektedir. Bu görüşe göre yanıtlanma süreleri hızlı tahmin davranışının bir sonucu olarak değerlendirilmektedir (Lee & Jia, 2014; Wise, 2017). Diğer taraftan yanıtlanma sürelerinin bireysel bir yetenek olarak değerlendirildiğine yönelik başka bir görüşe rastlanmaktadır (Kyllonen & Zu, 2016). Fox (2017) ve Lee ve Haberman'ın (2016) çalışmalarında ulaşılan sonuçların bu görüşü desteklediği gözlenmektedir.

Özetle, öğrencilerin maddeleri yanıtlanma süresi ve öğrenci performansları arasında önemli bir ilişkinin söz konusu olduğu anlaşılmaktadır. Dolayısıyla öğrencilerin maddeleri yanıtlanma hızları ve ilişkili özelliklerin incelenmesinin, yanıtlanma süresi ile ilgili görüşleri ve çalışmaları destekleyici bir rol üstlenecektir. Belirtilen görüşle, bu araştırmanın amacı; yordanan değişkenin okuma becerilerine yönelik maddeleri yanıtlanma hızı olduğu bir regresyon modeli oluşturularak yanıtlanma hızını yordayan özelliklerin belirlenmesi olmaktadır. Araştırmanın, öğrencilerin maddeleri yanıtlanma süreçlerinin ilişkili olduğu özellikleri açıklaması yönüyle alan yazına anlamlı bir katkı sunacağı düşünülmektedir. Ayrıca oluşturulan regresyon modelinin çözümlenmesinde kullanılan genetik algoritma yönteminin sağladığı avantaj ve regresyon modellerindeki başarılı performansı da araştırmanın diğer bir önemini ortaya koymaktadır.

Regresyon analizleri; oluşturulan regresyon modeli için tanımlanan fonksiyonun belirlenmesi, değişken seçimi ve parameter kestirimi olmak üzere üç temel amaç için kullanılmaktadır

(Paterlini & Minerva, 2010; Yang vd., 2011). Genetik algoritmalar yöntemi ise regresyon analizlerinde en iyi modelin oluşturulması ya da regresyon modeli için en iyi yordayıcı değişkenlerin seçilmesinde tercih edilen bir yöntemdir (Vasant, 2013; Žegklitz & Pošík, 2015). Bu bağlamda genetik algoritma tekniği, yordanan değişkenin öğrencilerin okuma becerilerine yönelik maddeleri yanıtlama hızı olduğu en iyi doğrusal regresyon modelinin optimize (en iyileme) edilmesi amacıyla kullanılmıştır. Daha açık bir ifadeyle, okuma becerilerine yönelik maddeleri yanıtlama hızını yordayan değişkenlerden oluşan en iyi regresyon modeli için değişken seçimi (feature selection) işleminin gerçekleşmesi genetik algoritma tekniği kullanılarak gerçekleştirilmiştir. Doğal bir seçim ve arama yöntemi olan genetik algoritmalar, Charles Darwin'in evrim teorisi ilkelerince şekillenen doğayı taklit etme ve güçlü olanın hayatta kalması esasına dayanmaktadır. Canlıların biyolojik gelişimlerini esas alarak doğal genetik ve seçim mekaniğine bağlı arama yapan olasılık temelli algoritmalarlardır. Canlılardaki gen yapısına benzer kodlama esasına dayalı kromozom düzeni ile işlemler gerçekleşir. Popülasyonu oluşturan grubun özelliklerine göre güçlü olan bireylerin hayatta kalması esas alınır. Bir takım gelişim ve değişimlere maruz kalan bireyler arasından dayanıklılık gösterenlerin bir araya gelmesiyle bir sonraki nesil oluşturulur (Michalewich, 1996). Genetik algoritmaları deterministik yapıda olan diğer temel yöntemlerden ayıran en önemli özellik olasılık temelli işlev göstermesidir. Bu özellik en yüksek uygunluk değerine sahip bireylerin seçilmesi olanağı sağlamaktadır. Problem durumlarına sunduğu ikili kodlama prensibine dayalı çözümler ise yöntemin diğer bir avantajı olarak değerlendirilmektedir (Örkcü, 2009). Regresyon modellerinde optimal model (değişken) seçimine ilişkin çalışmalarda (Leardi vd., 1992; Minghua vd., 2017; Paterlini & Minerva, 2010; Tolvi, 2004; Trejos vd., 2016) performansı, genetik algoritmalar yönteminin başarısını açık bir şekilde ortaya koymaktadır.

Araştırmanın amacı doğrultusunda, öğrencilerin okuma becerilerine yönelik maddeleri yanıtlama hızını yordayan özellikler nelerdir?, sorusuna cevap aranacaktır. Problem durumunun çözümlenmesinde kullanılan regresyon modelinin değişkenleri PISA (Programme for International Students Assessment) 2015 uygulamasından yararlanılarak oluşturulan veri seti üzerinden elde edilmiştir. Modelin yordanan değişkeni öğrencilerin okuma becerilerine yönelik maddeleri yanıtlama hızları olurken, yordayıcı değişkenleri ise; yordanan değişkenle ilişkili olduğu varsayılan cinsiyet, anne ve baba eğitim durumu, evde internet kullanımı, evde konuşulan dil, sahip olunan e-kitap okuyucu sayısı, evdeki kitap türü ve sayısı, okulda okuma becerileri için

ayrılan haftalık ders saati, okuma becerileri, sınıf düzeyi ve eylem sayısı olmaktadır. Aralarındaki ilişkinin alanyazındaki çalışmalarla (Goldhammer vd., 2015; Kaan vd., 2015; Kroehne vd., 2019; Michaelides vd., 2020; Ponce vd., 2020; Yavuz, 2019) desteklendiği üzere bu araştırmayla, öğrencilerin okuma becerilerine yönelik maddeleri yanıtlama hızları ile ilişkili olduğu varsayılan bu değişkenler arasından, öğrencilerin okuma becerilerine yönelik maddeleri yanıtlama hızlarını yordayan en iyi regresyon modeli için yordayıcı değişkenlerin seçilmesi planlanmaktadır.

Yöntem

Araştırma Türü

İki değişken arasında ilişki söz konusu olduğu durumlarda, değişkenlerden biri üzerinden diğerinin yordanması mümkün olabilmektedir. Bu amaçla tasarlanan araştırmalar, korelasyonel (ilişkisel) araştırmalar olarak ifade edilmektedir (Marczyk vd., 2010). Bu bağlamda, öğrencilerin okuma becerilerine yönelik maddeleri yanıtlama hızlarını yordayan özelliklerin belirlenmesi amacıyla tasarlanan bu araştırma bir korelasyonel araştırma olmaktadır.

Çalışma Grubu

Araştırma, PISA 2015 verilerinden yararlanılarak oluşturulan veri seti ile gerçekleştirilmiştir. Bu bağlamda araştırmanın çalışma grubunu PISA 2015 uygulamasına katılan 2699'u kız (%51.6) ve 2533'ü erkek (%48.4) olmak üzere 42 farklı ülkeden 5232 on beş yaş grubu öğrenci oluşturmuştur. PISA 2015 uygulamasında okuma becerileri ile ilgili maddeleri aynı sırada yanıtlayan öğrenci verileri dikkate alınmıştır. En çok yanıtlanan kitapçık olma özelliğinde olduğundan 36. kitapçığın bir formu tercih edilmiştir. Bilgisayar tabanlı (computer based) olarak uygulanan bu formu 9396 öğrenci yanıtlamıştır. Araştırmaya konu olan değişkenlerin bir araya getirildiği veri setindeki kayıp veriler liste bazlı veri silme tekniğiyle temizlenmiş ve bu işlem sonrasında 5232 öğrenci verisine ulaşılmıştır.

Veri Toplama Araçları

Araştırmada veri toplama aracı olarak PISA 2015 uygulamasında kullanılan, okuma becerilerine yönelik bilişsel maddeleri içeren 36. kitapçığın bir formu ve öğrenme ortamı, duyuşsal, demografik bilgilerle ilgili maddelerin yer aldığı öğrenci anketlerinden yararlanılmıştır. Bilgisayar ortamında uygulanan 36. kitapçığın bu formunda farklı yanıtlama türünde okuma

becerilerine yönelik 30 madde yer almaktadır. Bu maddeler arasından; madde konum etkisinin engellenmesi ve ortak maddelerin yanıtlanması dikkate alınarak uygun olan 19 madde kullanılmıştır. Öğrencilerin yanıt örüntüleri ve süre miktarları bu maddeler aracılığıyla, öğrencilere ait diğer veriler ise ilgili maddeleri içeren öğrenci anketleri üzerinden sağlanmıştır.

Araştırmanın yordanan değişkeni olan okuma becerilerine yönelik maddeleri yanıtlama hızı, öğrencilerin her bir madde için harcadığı sürelerin toplamına karşılık gelen, dakika cinsinden süre miktarını ifade etmektedir. Araştırmada 19 madde yer aldığından bu değişken 19 madde için ayrılan toplam süre miktarı olmaktadır. Yordayıcı değişkenlerden biri olan okuma becerileri, PISA 2015 uygulamasındaki 36. kitapçıkta yer alan 19 maddeye verilen cevaplara ait öğrenci ham puanları toplamı olarak değerlendirilmiştir. Ham puanlar hesaplanırken; doğru yanıtlar 1, yanlış ve boş yanıtlar 0 puan olarak puanlanmıştır. Diğer yordayıcı değişken olan eylem sayısı, öğrencilerin her bir madde için hangi sayıda eylem gerçekleştirdiğini göstermektedir. Eylem sayısı; tıklama, tuşlama, sürükleyip bırakma gibi bilgisayar ortamındaki ekran özelliklerine yönelik işlemler olarak tanımlanmaktadır (OECD, 2017).

Araştırmanın diğer yordayıcı değişkenleri ise cinsiyet, evde internet kullanımı, e-kitap okuyucu sayısı, evdeki kitap türü ve sayısı, evde konuşulan dil, okuma becerilerine yönelik okutulan haftalık ders saati, anne ve baba eğitim durumu ve sınıf düzeyi olarak sıralanmaktadır. Cinsiyet, veri grubunu oluşturan öğrencilerin cinsiyetlerini (kız-erkek); evde internet kullanımı, öğrencilerin evlerinde internet kullanımına sahip olma durumunu (var-yok); e-kitap okuyucu sayısı, öğrencilerin sahip oldukları elektronik kitap okuyucu sayısını; evdeki kitap türü, evlerinde bulunun kitap çeşidi sayısını; evdeki kitap sayısı, evlerinde bulunun kitap miktarını ifade etmektedir. Diğer taraftan evde konuşulan dil, öğrencilerin evlerinde konuştukları dilin test (eğitim) dilinden farklı olup olmama durumunu (test dili-diğer); okuma becerilerine yönelik haftalık ders saati, öğrencilerin okullarında bu kapsamda ayrılan ders saati miktarını; anne ve baba eğitim durumu, ebeveynlerin tamamladıkları son eğitim basamağını ve sınıf düzeyi, öğrencilerin hangi sınıf düzeyine devam ettiklerini işaret etmektedir.

Verilerin Toplanması

Araştırmada kullanılan veriler, PISA 2015 programına katılan öğrencilere ait verilerden oluşmaktadır. PISA, üç yılda bir 15 yaş grubu öğrencilere OECD (Organisation for Economic

Co-operation and Development) tarafından uygulanan dünya çapında bir öğrenci değerlendirme programıdır. Fen, matematik ve okuma becerileri olmak üzere üç yetenek grubuna yönelik ölçme işlemleri gerçekleştirilmektedir. Öğrenci yeteneklerinin yanı sıra öğrenci, öğrenme ortamları ve ailelere yönelik demografik bilgiler de toplanmaktadır. PISA programı, farklı ülkelerin eğitim sistemleriyle ilişkili bilgilerden yola çıkarak; eğitim süreçlerinin etkinliği ve etkililiği üzerine eğitimi, araştırmacı ve politikacılara önemli bir rehber görevi üstlenmektedir. PISA 2015 uygulamasına 72 ülkeden yaklaşık 540 bin öğrenci katılmıştır (OECD, 2017).

Verilerin Analizi

Oluşturulan regresyon modeli ile öğrencilerin okuma becerilerine yönelik maddeleri yanıtlama hızlarını yordayan en iyi değişkenlerin seçilmesinde kullanılan genetik algoritmalar tekniğine ilişkin analizler *R* programı aracılığıyla gerçekleştirilmiş olup, çözümlenmeler için *caret* (Kuhn vd., 2021b), *GA* (Scrucca, 2021), *mlbench* (Leisch & Dimitriadou, 2021); *modelr* (Wickham & Rstudio, 2020); *recipes* (Kuhn vd., 2021a) ve *rsample* (Silge vd., 2021) paketlerinden yararlanılmıştır. Genetik algoritmalar tekniğinde parametrik istatistik yöntemlerinin gerektirdiği değişkenler arası ilişkinin doğrusallığı, verilerin normal dağılımı ve homojenlik vs. gibi sayıtlara ihtiyaç duyulmadığından belirtilen sayıtlara yönelik her hangi bir test işlemi gerçekleştirilmemiştir (Altunkaynak, 2009; Şen & Öztopal, 2001).

Ancak genetik algoritmalar tekniğinin de örneklerinden biri olduğu makine öğrenmelerine dayalı yöntemlerde modeli oluşturan tüm değişkenlerin sayısal (numeric) değerlerle ifade edilmesi gerekmektedir. Dolayısıyla kategorik yapıdaki değişkenlerin sayısal değerlere dönüştürülmesi sağlanmalıdır. Bu amaç için tercih edilen popüler dönüşüm tekniklerinden biri sıralı kodlama (ordinal encoding) yöntemi olmaktadır. Sıralı kodlama yönteminde kategorik değişkenin sahip olduğu her bir düzey 1-2-3 gibi ardışık tamsayılarla ifade edilmektedir (Brownlee, 2020). Bu doğrultuda genetik algoritmalar esasıyla analiz edilmek üzere oluşturulan regresyon modelinde yer alan kategorik değişkenler sıralı kodlama tekniği esasıyla ardışık tamsayılarla dönüştürülmüştür.

Genetik Algoritmalar

Genetik algoritmalar, Darwin'in evrim teorisine dayanan biyolojik sürecin bilgisayar ortamlarına aktarılmasıyla makine öğrenmelerine uyarlanmasıyla ortaya çıkmıştır. Genetik algoritmalara yönelik

analizlerde amaç problem durumunu yansıtan fonksiyonlara en iyi çözümü üretmektir. Veri grubunun özellikleri dikkate alınarak güçlü verilerin hayatta kalarak zayıf verilerin ölmesi mantığına göre hareket edilir. Canlılardaki gen özelliğine benzer şekilde bilginin kromozom yapısına dönüştürülmesi esas alınır. Kromozomlar birey rolüyle popülasyonlar oluşturur. Problem durumunun çözümünü sağlayacak veriler, kromozomların hayatta kalma mücadelesiyle belirlenir (Michalewich, 1996). Çözüm süreci olarak ifade edilen bu mücadele bireylerin kopyalama, çaprazlama, mutasyon gibi genetik operatörlere maruz kalmasıyla gerçekleşir (Goldberg, 1989). Bu işlemler sonrasında her bir kromozom (birey) için uygunluk değeri hesaplanır. Uygunluk değeri, problemin çözümü için oluşturulan uygunluk (amaç) fonksiyonu üzerinden belirlenir (Haupt & Haupt, 1998). Hesaplanan uygunluk değeri ile güçlü ve zayıf bireyler belirlenerek, güçlü bireylerin üremesi; zayıf bireylerin ise öldürülmesi sağlanır. İstenen durdurma koşulu sağlanınca işlemler durdurulur ve en güçlü bireylere yani en iyi çözüme ulaşılmış olur (Michalewich, 1996). Genetik algoritma yöntemiyle gerçekleşen analiz işlem basamakları aşağıda açıklanmıştır (Örkcü, 2009):

1. Adım: Olası çözümlerin kodlandığı bir çözüm kümesi oluşturulur. Bu çözüm kümesi popülasyon; çözüm kodları ise kromozom olarak adlandırılır.
2. Adım: Her bir birey için uygunluk fonksiyonu değerleri hesaplanarak, en iyi sonuçları veren kromozom yapısına sahip bireyler bir sonraki popülasyon için seçilir.
3. Adım: Seçilen kromozomlar üreme, çaprazlama ve mutasyon işlemleri uygulanarak yeni yapıya sahip popülasyon için bireyler belirlenir.
4. Adım: Tüm bireyler için uygunluk değerleri tekrar hesaplanır.
5. Adım: Hedef nesil için istenen birey sayısı sağlanmışsa ya da popülasyon içinde bir durağanlık mevcut ise işlem durur ve o ana kadar hesaplanan en iyi uygunluk değeri optimal sonucu verir. Aksi durumda 2. Adımdan itibaren süreç yenilenir.

Bulgular

Araştırmanın çalışma grubunu oluşturan öğrenci verileri üzerinden, problem durumu olan “Öğrencilerin okuma becerilerine yönelik maddeleri yanıtlama hızlarını yordayan özellikler

nelerdir?”, sorusuna yanıt oluşturmak adına ulaşılan betimsel özelliklere ve genetik algoritmalar tekniği kullanılarak en iyi değişkenlerin seçimi amacıyla analiz edilen regresyon modelinden elde edilen bulgulara aşağıda yer verilmiştir.

Genetik algoritmalar yöntemi esasıyla analiz edilen ve yordanan değişkeninin okuma becerilerine yönelik maddeleri yanıtlama hızları; yordayıcı değişkenlerinin ise cinsiyet, anne ve baba eğitim durumu, evde internet kullanımı, evde konuşulan dil, sahip olunan e-kitap okuyucu sayısı, evdeki kitap türü ve sayısı, okulda okuma becerileri için ayrılan haftalık ders saati, okuma becerileri, sınıf düzeyi ve eylem sayısı olan regresyon modelinin değişkenlerine ilişkin betimsel özellikler Tablo 1a ve Tablo 1b ile sunulmuştur.

Tablo 1a

Yanıt Hızları İle İlişkili Değişkenlere Ait Betimsel Özellikler

Değişken		N	%	Değişken		N	%
Cinsiyet	Kız	2699	51.6	İnternet kullanımı	Var	5072	96.9
	Erkek	2533	48.4		Yok	160	3.1
Anne eğitim düzeyi	Tamamlamadı	83	1.6	Konuşulan dil	Test dili	4683	89.5
	İlkokul	229	4.4		Diğer	549	10.5
	Ortaokul	739	14.1	Sınıf düzeyi	7	13	0.2
	Lise	1313	25.1		8	164	3.1
	Yüksekokul	743	14.2		9	1718	32.9
	Lisans	981	18.8		10	2995	57.2
	Lisansüstü	1144	21.9		11	336	6.4
	Tamamlamadı	74	1.4		12	6	0.1
İlkokul	256	4.9	Kitap sayısı	0-10	520	9.9	
Ortaokul	884	16.9		11-25	747	14.3	
Lise	1240	23.7		26-100	1581	30.2	
Yüksekokul	775	14.8		101-200	1029	19.7	
Lisans	867	16.6		201-500	888	17	
Lisansüstü	1136	21.7	501-üstü	467	8.9		

Tablo 1b

Yanıt Hızları İle İlişkili Değişkenlere Ait Betimsel Özellikler

Değişken	N	Min.	Maks.	Ortalama	S. sapma
Yanıt hızları	5232	2.65	38.38	16.98	4.52
Okuma becerileri	5232	0	19	13.14	3.8
E-kitap okuyucu sayısı	5232	0	3	0.35	0.67
Kitap türü sayısı	5232	0	6	4.11	1.61
Haftalık ders saati	5232	2	10	4.43	1.41
Eylem sayısı	5232	19	50	21.97	3.86

Makine öğrenmelerine yönelik yöntemlerle yapılan analizlerde sık rastlanan problemlerin başında, yanlış hatalara neden olan aşırı uyum problemi gelmektedir. Bu sorunun çözümündeki en etkili yöntem olarak verilerin eğitim (training) ve test (testing) seti olarak ayrılması ve analiz işlemleri eğitim seti üzerinden gerçekleştirildikten sonra diğer veri setleri aracılığıyla kontrollerin sağlanması önerilir (Kushchu, 2002; Žegklitz & Pošík, 2015). Bu gerekçeyle, verilerin %70'i eğitim seti ($n=3663$) ve %30'u test seti ($n=1569$) olmak üzere (Ahmad vd., 2018; Ahmed & Elaraby, 2014) ikiye ayrılarak analiz işlemleri gerçekleştirilmiştir. Regresyon modelinin analiz edilmesinde eğitim seti verileri kullanılırken; analiz edilen modelin farklı veri grupları üzerinden performans düzeyinin belirlenmesinde test ve tüm veri setinden yararlanılmıştır. Eğitim veri seti ile gerçekleştirilen genetik algoritmalar yöntemine yönelik analizde; başlangıç topluluğu büyüklüğü 50, maksimum nesil sayısı 100, mutasyon işlemi olasılığı .1, çaprazlama işlemi olasılığı .8 ve elitizm sayısı 2 olarak belirlenen ikili kodlama esasına dayalı genetik işlemler uygulanmıştır. En iyi regresyon modelinin belirlenmesini sağlayan değişkenlerin seçimi amacıyla analiz edilen model için uyum kriteri olan Bayes Bilgi Kriteri (BIC) değerini en küçük yapan uygunluk (amaç) fonksiyonu kullanılmıştır (Minghua vd., 2017; Trejos vd., 2016). Analizde kullanılan R programı GA paketi uygunluk (amaç) fonksiyonunu maksimum yapma özelliğine sahiptir. BIC değerinin minimum olması arzulandığından uygunluk fonksiyonu içerisinde yer alan negatif BIC değerinin maksimum değerinin elde edilmesi sağlanmıştır. Bu bağlamda en iyi model için BIC değeri -21,243.7 olarak kestirilmiş olup genetik algoritmalar yöntemine ait kestirim özellikleri ve en iyi regresyon modeli için seçilen değişkenlere ilişkin bilgileri içeren R programı analiz çıktısı Şekil 1 ile sunulmuştur.

```

-- Genetic Algorithm -----

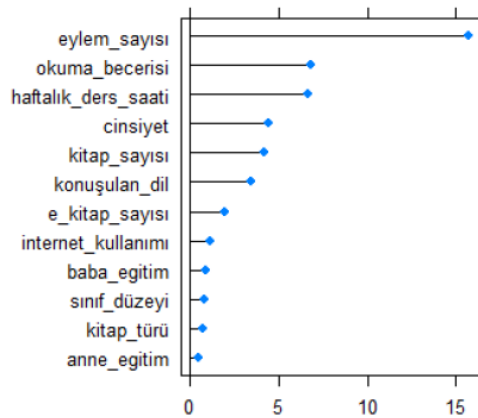
GA settings:
Type           = binary
Population size = 50
Number of generations = 100
Elitism        = 2
Crossover probability = 0.8
Mutation probability = 0.1

GA results:
Iterations      = 100
Fitness function value = -21243.7
Solution =
  cinsiyet internet_kullanımı e_kitap_sayısı kitap_sayısı konuşulan_dil okuma_becerisi eylem_sayısı
[1,]      1             0             0             1             1             1             1
  haftalık_ders_saati anne_egitim baba_egitim sınıf_düzeysi kitap_türü
[1,]      1             0             0             0             0

```

Şekil 1. Genetik algoritmalar analizi R programı çıktısı

İkili kodlama esasıyla gerçekleştirilen genetik algoritmalar yöntemine yönelik analiz çıktılarında, model için seçilen en iyi yordayıcı değişkenler 1 olarak kodlanırken; yordayıcı olarak seçilmeyen değişkenler ise 0 olarak kodlanır. Şekil 1 incelendiğinde, en iyi regresyon modeli için okuma becerilerine yönelik maddeleri yanıtlama hızını istatistiksel olarak anlamlı şekilde yordayan değişkenlerin; cinsiyet, evdeki kitap sayısı, evde konuşulan dil, okuma becerisi, eylem sayısı ve haftalık ders saati olmak üzere altı özelliğin olduğu saptanmıştır. İnternet kullanımı, e-kitap okuyucu sayısı, anne ve baba eğitim düzeyi, sınıf düzeyi ve evdeki kitap türü sayısı değişkenleri ise model için öğrencilerin yanıtlama hızlarını istatistiksel olarak anlamlı düzeyde yordamamıştır. Seçilen yordayıcı değişkenlerin regresyon modeli için 0 ile 16 değerleri arasında ölçeklenen önem düzeyleri Şekil 2 ile sunulmuştur.



Şekil 2. Yordayıcı değişkenlerin regresyon modeli için önem düzeyleri

Öğrencilerin okuma becerilerine yönelik maddeleri yanıtlama hızlarını yordayan değişkenlerin regresyon modeli için önem düzeylerini içeren Şekil 2 incelendiğinde, öğrencilerin okuma becerilerine yönelik maddeleri yanıtlama hızlarını istatistiksel olarak anlamlı düzeyde yordayan en önemli değişkenin eylem sayısı olduğu anlaşılmaktadır. Bu değişkeni önem sırasına göre; okuma becerisi, haftalık ders saati, cinsiyet, evdeki kitap sayısı ve evde konuşulan dil takip etmektedir. Önem derecesi en düşük olan altı değişken; e-kitap okuyucu sayısı, evde internet kullanımı, baba eğitim düzeyi, sınıf düzeyi, evdeki kitap türü sayısı ve anne eğitim düzeyi değişkenlerinin ise öğrencilerin okuma becerilerine yönelik maddeleri yanıtlama hızlarını istatistiksel olarak anlamlı düzeyde yordamadığı saptanmıştır. Eğitim seti üzerinden analiz edilen regresyon modelinin performansı eğitim seti, test seti ve tüm veri seti için R^2 , RMSE ve MAE metrikleri ile araştırılmış (Özdemir, 2017; Paterlini & Minerva, 2010) olup, elde edilen değerlere ilişkin bilgiler Tablo 2 ile sunulmuştur.

Tablo 2

Regresyon Modelinin Veri Setleri İçin Performans Metrikleri Değerleri

Veri seti	N	R^2	RMSE	MAE
Eğitim seti	3663	.0853	4.36	3.47
Test seti	1569	.0811	4.27	3.42
Tüm veri seti	5232	.0841	4.33	3.45

Öğrencilerin okuma becerilerine yönelik maddeleri yanıtlama hızlarını yordayan en iyi değişkenlerin seçilmesini sağlayan regresyon modelinin üç farklı veri seti için performans değerlerine yer verilen Tablo 2 incelendiğinde, eğitim seti için ulaşılan performans metrikleri değerlerinin test ve tüm veri seti için yakın değerlerde olduğu gözlenmektedir. Bu bağlamda, oluşturulan modelin farklı veri yapılarında benzer derecede performans gösterdiği anlaşılmaktadır. Diğer taraftan eğitim seti üzerinden elde edilen değerler dikkate alındığında, öğrencilerin okuma becerilerine yönelik maddeleri yanıtlama hızlarını istatistiksel olarak anlamlı düzeyde yordayan değişkenlerin, öğrencilerin okuma becerilerine yönelik maddeleri yanıtlama hızlarındaki değişkenliğin % 8.53'ünü ($R^2=.0853$) açıkladığı anlaşılmaktadır.

Sonuç, Tartışma ve Öneriler

PISA 2015 programına katılan öğrencilere ait verilerden yararlanılarak oluşturulan çalışma grubundaki öğrencilerin okuma becerilerine yönelik maddeleri yanıtlama hızlarını yordayan değişkenlerin belirlenmesi amacıyla genetik algoritmalar tekniği kullanılarak en iyi (optimal) regresyon modelinin kestirilmesi sonucunda; öğrencilerin okuma becerilerine yönelik maddeleri yanıtlama hızlarını altı değişkenin istatistiksel olarak anlamlı düzeyde yordadığı saptanmıştır.

Öğrencilerin okuma becerilerine yönelik maddeleri yanıtlama hızlarını istatistiksel olarak anlamlı düzeyde yordayan özelliklerden ikisi; öğrencilerin okuma becerileri ve eylem sayısı olmuştur. Bu bağlamda öğrencilerin okumaya yönelik becerileri ve okuma becerilerine yönelik maddeleri yanıtlamak üzere gerçekleştirdikleri eylem sayısındaki farklılaşma öğrencilerin okuma becerilerine yönelik maddeleri yanıtlama hızlarında anlamlı düzeyde farklılaşmaya yol açmaktadır. Dodonova ve Dodonov (2013); Goldhammer ve diğerleri (2015); Greiff ve diğerleri (2016); Michaelides ve diğerleri (2020) ve Yavuz (2019) çalışmalarında araştırmanın bulgularına paralel olarak, öğrenci performanslarının maddeleri yanıtlama süreleri ile ilişkisini vurgulamaktadırlar. Arslan ve diğerleri (2020) ve Ponce ve diğerleri (2020) gerçekleştirdikleri araştırmada, bilgisayar tabanlı test uygulamalarındaki ekran hareketlerine yönelik eylem sayılarının öğrencilerin maddeleri yanıtlama sürelerini anlamlı düzeyde yordadığı sonucuna ulaşmışlardır.

Öğrencilerin okuma becerilerine yönelik maddeleri yanıtlama hızlarını istatistiksel olarak anlamlı düzeyde yordayan diğer iki değişken cinsiyet ve evde konuşulan dil olmuştur. Elde edilen sonuçlara göre; öğrencilerin cinsiyetlerinin ve evde konuştukları dilin eğitim aldıkları (test dili) dilden farklı olması öğrencilerin okuma becerilerine yönelik maddeleri yanıtlama hızlarında farklılaşmaya yol açmaktadır. Arnett ve diğerleri (2017); Karia ve diğerleri (2012) ve Kroehne ve diğerlerinin (2019) çalışmalarında cinsiyetin öğrencilerin yanıtlama hızlarını yordadığı anlaşılmaktadır. Kaan ve diğerleri (2015) ve Van Gelderen ve diğerlerinin (2004) çalışmalarında, test dilinin ikinci dili olduğu öğrenciler ile birinci dili olan öğrencilerin okuma becerilerine yönelik maddeleri yanıtlama hızlarının farklılaştığı gözlenmektedir. Taguchi (2005) ise ikinci dil ile öğrencilerin okuma becerilerine yönelik maddeleri yanıtlama hızları arasında anlamlı düzeyde bir ilişki saptamamıştır.

Araştırma sonuçlarına göre öğrencilerin maddeleri yanıtlama hızlarını istatistiksel olarak anlamlı düzeyde yordayan diğer değişken haftalık ders saati olmuştur. Okullarında okuma becerilerine yönelik okutulan haftalık ders saati, öğrencilerin okuma becerilerine yönelik maddeleri yanıtlama hızlarında anlamlı düzeyde farklılaşmaya yol açmaktadır. Lai (1993) gerçekleştirmiş olduğu çalışmada yaz aylarında ek olarak yapılan dört haftalık okumaya yönelik kurs programının öğrencilerin hızlarında değişime yol açtığı sonucuna ulaşmıştır. Tanaka ve Stapleton'ın (2007) beş ay boyunca derslerin 5 ile 10 dakikasını öğrencilere okumaya yönelik etkinlikler yaptırarak gerçekleştirdikleri çalışmada, etkinlik yaptırılan öğrencilerin okumalarının hızlandığı sonucuna ulaşmışlardır. Yen (2012) çalışmasında, okuma becerilerine yönelik ekstra kurs alan öğrencilerin okuma hızlarının arttığına yönelik bir sonuç elde etmiştir.

Yordama düzeyi istatistiksel olarak anlamlı bulunan son değişken evdeki kitap sayısı olmuştur. Öğrencilerin evlerinde bulunan kitap sayısındaki farklılaşma okuma becerilerine yönelik maddeleri yanıtlama hızlarında farklılaşmaya yol açtığı anlaşılmaktadır. Anderson ve diğerleri (1988) okul dışında yürütülen kitap okuma ve okumaya yönelik etkinliklerdeki sıklığın öğrencilerin okuma hızlarına etki ettiğine yönelik bir çalışma gerçekleştirmişlerdir. Van Bergen ve diğerlerinin (2017) çalışmasında öğrencilerin evlerinde bulunan kitap sayısı ile okuma hızları arasında anlamlı düzeyde ilişki olduğu anlaşılmaktadır.

Araştırmada öğrencilerin okuma becerilerine yönelik maddeleri yanıtlama hızlarını istatistiksel olarak anlamlı düzeyde yordayan değişkenlerin; cinsiyet, evdeki kitap sayısı, evde konuşulan dil, okuma becerisi, eylem sayısı ve okutulan haftalık ders saati olduğu saptanmış olup, bu değişkenlerin öğrencilerin yanıtlama hızlarındaki değişkenliğin %8.53'ünü açıkladığı kestirilmiştir. Ulaşılan sonuçların alanyazındaki benzer araştırma sonuçlarıyla tutarlılık gösterdiği gözlenmekle birlikte cinsiyet, evdeki kitap sayısı, evde konuşulan dil, okuma becerisi, eylem sayısı ve okulda okuma becerileri için ayrılan haftalık ders saatinin farklılaşması öğrencilerin okuma becerilerine yönelik maddeleri yanıtlama hızlarında da anlamlı düzeyde farklılaşmaya yol açabileceği söylenebilmektedir.

Bu bağlamda elde edilen sonuçların alanyazındaki çalışmalarla tutarlılığından yola çıkılarak genetik algoritmalar yönteminin en iyi regresyon modellerinin belirlenmesinde ya da regresyon modelleri için en iyi değişkenlerin seçilmesinde kullanılabileceği önerilmektedir. Diğer taraftan okuma becerilerine yönelik yanıt hızlarıyla ilgili yapılacak uygulama ve araştırmalarda cinsiyet,

evde bulunan kitap sayısı, evde konuşulan dil, okuma becerisi, elektronik ortamda gerçekleştirilen eylem sayısı ve okulda okuma becerilerine ilişkin okutulan haftalık ders saati özelliklerinin dikkate alınması gerektiği önerilmektedir.

Kaynakça

- Ahmad, M., Reynolds, J., & Rezgui, Y. (2018). Predictive modelling for solar thermal energy systems: A comparison of support vector regression, random forest, extra trees and regression trees. *Journal of Cleaner Production*, 203, 810-821.
- Ahmed, A. B., & Elaraby, I. S. (2014). Data mining: A prediction for student's performance using classification method. *World Journal of Computer Application and Technology*, 2(2), 43-47.
- Altunkaynak, A. (2009). Sediment load prediction by genetic algorithms. *Advances in Engineering Software*, 40(9), 928-934.
- Anderson, R. C., Wilson, P. T., & Fielding, L. G. (1988). Growth in reading and how children spend their time outside of school. *Reading Research Quarterly*, 285-303.
- Arnett, A. B., Pennington, B. F., Peterson, R. L., Willcutt, E. G., DeFries, J. C., & Olson, R. K. (2017). Explaining the sex difference in dyslexia. *Journal of Child Psychology and Psychiatry*, 58(6), 719-727.
- Arslan, B., Jiang, Y., Keehner, M., Gong, T., Katz, I. R., & Yan, F. (2020). The effect of drag-and-drop item features on test-taker performance and response strategies. *Educational Measurement: Issues and Practice*, 39(2), 96-106.
- Başaran, M. (2013). Okuduğunu anlamanın bir göstergesi olarak akıcı okuma. *Kuram ve Uygulamada Eğitim Bilimleri*, 13(4), 2277-2290.
- Brownlee, J. (2020). *Data preparation for machine learning: data cleaning, feature selection, and data transforms in Python*. Machine Learning Mastery.
- Cañizo, M. A., Suárez-Coalla, P., & Cuetos, F. (2015). The role of reading fluency in children's text comprehension. *Frontiers in Psychology*, 6, 1810.
- Dodonova Y. A., & Dodonov Y. S. (2013). Faster on easy items, more accurate on difficult ones: Cognitive ability and performance on a task of varying difficulty. *Intelligence*, 41(1), 1-10.
- Fox, J. P., & Mariani, S. (2017). Person-fit statistics for joint models for accuracy and speed. *Journal of Educational Measurement*, 54(2), 243-262.
- Goldberg, D. E. (1989). *Genetic algorithms in search, optimization and machine learning*. Addison-Wesley.

- Goldhammer, F., Naumann, J., & Greiff, S. (2015). More is not always better: The relation between item response and item response time in Raven's matrices. *Journal of Intelligence*, 3(1), 21-40.
- Greiff, S., Niepel, C., Scherer, R., & Martin, R. (2016). Understanding students' performance in a computer-based assessment of complex problem solving: An analysis of behavioral data from computer-generated log files. *Computers in Human Behavior*, 61(2016), 36-46.
- Gürses, R. (1996). Okuma anlama üzerine. *Atatürk Kültür, Dil ve Tarih Yüksek Kurumu Bülteni*, 28(9), 98-103.
- Haupt, R. L., & Haupt, S. E. (1998). *Practical genetic algorithms*. Willey-Interscience Publication.
- Kaan, E., Ballantyne, J. C., & Wijnen, F. (2015). Effects of reading speed on second-language sentence processing. *Applied Psycholinguistics*, 36(4), 799-830.
- Kantemir, E. (1995). *Yazılı ve sözlü anlatım*. Engin Yayınevi.
- Karia, R. M., Ghuntla, T. P., Mehta, H. B., Gokhale, P. A., & Shah, C. J. (2012). Effect of gender difference on visual reaction time: A study on medical students of Bhavnagar region. *IOSR Journal of Pharmacy*, 2(3), 452-454.
- Kroehne, U., Hahnel, C., & Goldhammer, F. (2019). Invariance of the response processes between gender and modes in an assessment of reading. *Frontiers in Applied Mathematics and Statistics*, 5, 2.
- Kuhn, M., Wickham, H., & RStudio. (2021a). Preprocessing and feature engineering steps for modeling. R package version 0.1.17. <https://cran.r-project.org/web/packages/recipes/index.html>
- Kuhn, M., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T., Mayer, Z., Kenkel, B., R Core Team, Benesty, M., Lescarbeau, R., Ziem, A., Scrucca, L., Tang, Y., Candan, C., & Hunt, T. (2021b). Classification and regression training. R package version 6.0-90. <https://cran.r-project.org/web/packages/caret/index.html>
- Kushchu, I. (2002). Genetic programming and evolutionary generalization. *IEEE Transactions on Evolutionary Computation*, 6(5), 431-442.
- Kyllonen, P., & Zu, J. (2016). Use of response time for measuring cognitive ability. *Journal of Intelligence*, 4(4), 1-29.
- Lai, F. K. (1993). The effect of a summer reading course on reading and writing skills. *System*, 21(1), 87-100.
- Learidi, R., Boggia, R., & Terrile, M. (1992). Genetic algorithms as a strategy for feature selection. *Journal of Chemometrics*, 6(5), 267-281.

- Lee, Y. H., & Haberman, S. J. (2016) Investigating test-taking behaviors using timing and process data. *International Journal of Testing*, 16(3), 240-267.
- Lee, Y. H., & Jia, Y. (2014). Using response time to investigate students' test-taking behaviors in a NAEP computer-based study. *Large-scale Assessments in Education*, 2(8), 1-24.
- Leisch, F., & Dimitriadou, E. (2021). Machine learning benchmark problems. R package version 2.1-3. <https://cran.r-project.org/web/packages/mlbench/index.html>
- Marczyk, G., DeMatteo, D., & Festinger, D. (2005). *Essentials of research design and methodology* (Vol. 2). John Wiley & Sons.
- Michaelides, M. P., Ivanova, M., & Nicolaou, C. (2020). The relationship between response-time effort and accuracy in PISA science multiple choice items. *International Journal of Testing*, 20(3), 187-205.
- Michalewicz, Z. (1996). *Genetic algorithms + data structures = evolution programs* (3th ed.). Springer.
- Minghua, S., Qingxian, X., Benda, Z., & Feng, Y. (2017). Regression modelling based on improved genetic algoritm. *Tehnicki vjesnik/Technical Gazette*, 24(1).
- National Research Council (2001). *Knowing what students know: The science and design of educational assessment*. J. Pelligrino, N. Chudowsky, & R. Glaser (Eds.), National Academy Press.
- Oakhill, J., Cain, K., & Elbro, C. (2015). *Understanding and teaching reading comprehension: A handbook*. Routledge.
- OECD. (2017). *PISA 2015 technical report*. OECD Publishing.
- Örkcü, H. H. (2009). *Ayırma analizine matematiksel programlama ve yapay sinir ağları yaklaşımları* (Doktora tezi). Gazi Üniversitesi, Ankara.
- Özdemir, M. (2017). Genetik algoritma ile doğrusal regresyonda tahmin amaçlı model seçimi. *Pamukkale Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 28, 213-233.
- Paterlini, S., & Minerva, T. (2010, June). Regression model selection using genetic algorithms. In *Proceedings of the 11th WSEAS International Conference on Neural Networks and 11th WSEAS International Conference on Evolutionary Computing and 11th WSEAS International Conference on Fuzzy Systems* (pp. 19-27). World Scientific and Engineering Academy and Society (WSEAS).
- Ponce, H. R., Mayer, R. E., Sitthiworachart, J., & López, M. J. (2020). Effects on response time and accuracy of technology-enhanced cloze tests: An eye-tracking study. *Educational Technology Research and Development*, 68(5), 2033-2053.

- Rios, J. A., Guo, H., Mao, L., & Liu, O. L. (2017). Evaluating the impact of careless responding on aggregated-scores: To filter unmotivated examinees or not?. *International Journal of Testing*, 17(1), 74-104.
- Schnipke, D. L., & Scrams, D. J. (2002). Exploring issues of examinee behavior: Insights gained from response-time analyses. *Computer-Based Testing: Building the Foundation for Future Assessments*, 237-266.
- Scrucca, L. (2021). Genetic algorithms. R package version 3.2.2. <https://cran.r-project.org/web/packages/GA/index.html>
- Silge, J., Chow, F., Kuhn, M., Wickham, H., & RStudio. (2021). General resampling infrastructure. R package version 0.1.1. <https://cran.r-project.org/web/packages/rsample/index.html>
- Şen, Z., & Öztopal, A. (2001). Genetic algorithms for the classification and prediction of precipitation occurrence. *Hydrological Sciences Journal*, 46(2), 255-267.
- Taguchi, N. (2005). Comprehending implied meaning in English as a foreign language. *The Modern Language Journal*, 89(4), 543-562.
- Tanaka, H., & Stapleton, P. (2007). Increasing reading input in Japanese high school EFL classrooms: An empirical study exploring the efficacy of extensive reading. *The Reading Matrix*, 7(1).
- Temizkan, M. (2007). *İlköğretim ikinci kademe Türkçe derslerinde okuma stratejilerinin okuduğunu anlama üzerindeki etkisi* (Doktora tezi). Gazi Üniversitesi, Ankara.
- Tolvi, J. (2004). Genetic algorithms for outlier detection and variable selection in linear regression models. *Soft Computing*, 8(8), 527-533.
- Trejos, J., Villalobos-Arias, M. A., & Espinoza, J. L. (2016). Variable selection in multiple linear regression using a genetic algorithm. In *Handbook of Research on Modern Optimization Algorithms and Applications in Engineering and Economics* (pp. 133-159). IGI Global.
- Van Bergen, E., Van Zuijen, T., Bishop, D., & De Jong, P. F. (2017). Why are home literacy environment and children's reading skills associated? What parental skills reveal. *Reading Research Quarterly*, 52(2), 147-160.
- Van Gelderen, A., Schoonen, R., De Glopper, K., Hulstijn, J., Simis, A., Snellings, P., & Stevenson, M. (2004). Linguistic knowledge, processing speed, and metacognitive knowledge in first and second language reading comprehension: A componential analysis. *Journal of Educational Psychology*, 96(1), 19.
- Vasant, P. (2013). Hybrid linear search, genetic algorithms, and simulated annealing for fuzzy non-linear industrial production planning problems. In *Meta-Heuristics Optimization Algorithms in Engineering, Business, Economics, and Finance* (pp. 87-109). IGI Global.

- Veenendaal, N. J., Groen, M. A., & Verhoeven, L. (2015). What oral text reading fluency can reveal about reading comprehension. *Journal of Research in Reading*, 38(3), 213-225.
- Wickham, H., & RStudio. (2020). Modelling functions that work with the pipe. R package version 0.1.8. <https://cran.r-project.org/web/packages/modelr/index.html>
- Williams, J. P. (2003). Teaching text structure to improve reading comprehension. H. L. Swanson, K. R. Harris, & S. Graham (Eds.). In: *Handbook of Learning Disabilities* (pp. 293–305). The Guilford Press.
- Wise, S. L. (2017). Rapid-guessing behavior: Its identification, interpretation, and implications. *Educational Measurement: Issues and Practice*, 36(4), 52–61.
- Wise, S. L., & Kong, X. (2005). Response time effort: A new measure of examinee motivation in computer-based tests. *Applied Measurement in Education*, 18(2), 163–183.
- Wise, S. L., & DeMars, C. E. (2005). Low examinee effort in low stakes assessment: Problems and potential solutions. *Educational Assessment*, 10(1), 1-17.
- Wolf, M., & Katzir-Cohen, T. (2001). Reading fluency and its intervention. *Scientific Studies of Reading*, 5(3), 211-239.
- Yang, C. Y., Jeng, J. T., Chuang, C. C., & Tao, C. W. (2011, June). Constructing the linear regression models for the symbolic interval-values data using PSO algorithm. In *Proceedings 2011 International Conference on System Science and Engineering* (pp. 177-181). IEEE.
- Yavuz, H. C. (2019). The effects of log data on students' performance. *Journal of Measurement and Evaluation in Education and Psychology*, 10(4), 378-390.
- Yen, T. T. N. (2012). The effects of a speed reading course and speed transfer to other types of texts. *RELC Journal*, 43(1), 23-37.
- Žegklitz, J., & Pošík, P. (2015, July). Model selection and overfitting in genetic programming: Empirical study. In *Proceedings of the Companion Publication of the 2015 Annual Conference on Genetic and Evolutionary Computation* (pp. 1527-1528).

Extended Abstract

Introduction

One of the three components of fluent reading, speed can be expressed as the automation of reading. Automation corresponds to reading words with little cognitive effort and attention. Automation relies on the recognition of words accurately and quickly, as well as understanding

the text quickly and following the meaning in the mind. Automation is closely related to reading skill. If the text is read slowly and intermittently, semantic integrity cannot be established, and the semantic connections within sentences and between sentences cannot be understood (Başaran, 2013). It has been observed that response times, which are parallel to speed, have been the subject to recent studies aiming to better understand mental activities (Schnipke & Scrams, 2002). When the literature is reviewed, it is seen that response times are associated with different variables. One of these variables is the willingness for test participation or respondent motivation. It is argued that excessively low or high response times can be a proof of validity for the student's test performance (Rios et al., 2017; Wise, 2017; Wise & DeMars, 2005; Wise & Kong, 2005). On the other hand, there is another view arguing that response times should be considered to be related to an individual ability (Kyllonen & Zu, 2016).

It is understood that there is a significant relationship between students' response time to the items and their performance. Therefore, examining students' response times to items and related features will contribute to opinions and studies on response time. Thus, the purpose of this study is to determine the features that predict response speed by creating a regression model in which the predicted variable is the speed of responding to the items related to reading skills. To this end, an answer to the research question "What are the variables predicting students' response time to the items related to reading skills and the features of these variables?" will be sought. It is thought that the study will make a significant contribution to the literature in terms of explaining the features related to processes involved in students' responding to items. In addition, the advantage to be brought about by the genetic algorithm method used in the analysis of the created regression model and its successful performance in the regression models are other reasons making the current study important.

Method

The study group of the current research was formed by using the data from PISA 2015. There are a total of 5232 students in the study group; 2699 females (51.6%) and 2533 males (48.4%). The speed of responding to the items related to reading skills, which is the predicted variable of the study, refers to the amount of time in minutes that corresponds to the sum of the time spent by the students for each item. The predictor variables of the study are gender, use of the internet at home, the number of e-book readers at home, the number and types of books at home, the language used

at home, the number of class hours devoted to reading skills, mother's education level, father's education level, grade level, reading skills and the number of acts of reading.

The genetic algorithms method used in the analysis of the regression model does not require the assumptions of normality, linearity and homogeneity required by parametric statistical methods (Altunkaynak, 2009; Şen & Öztopal, 2001). Thus, no assumption test was applied before the analysis performed by using the genetic algorithms method. Genetic algorithms analysis was carried out using the *R* program.

Results, Discussion and Suggestions

In light of the obtained data, the variables significantly predicting the speed of responding to the item related to reading skills for the best regression model were listed as follows; gender, the number of books at home, the language used at home, reading skills, the number of acts of reading and the number of class hours devoted to reading skills. The performance of the regression model created on the training data set was investigated with R^2 , RMSE and MAE metrics for the training set, test set and the entire data set. It was observed that the performance metrics values reached for the training set are close to those of the test data and the entire data set having different data set characteristics, so the model created performs similarly in different data structures. Moreover, when the values obtained from the training set are considered, it is understood that the variables that significantly predicted the speed of students' responding to the items related to reading skills explained 8.53% ($R^2=.0853$) of the variance in the speed of response to the items related to reading skills.

It was observed that the results obtained in the current study are consistent with the results of similar studies in the literature. Thus, it can be argued that variation in gender, the number of books at home, the language used at home, reading skills, the number of acts of reading and weekly class hours allocated to reading skills might lead to variation in students' speed of responding to the items related to reading skills. In this context, it is suggested that gender, number of books at home, language spoken at home, reading skills, number of actions performed in the electronic environment, and weekly course hours taught at school should be taken into consideration in the practices and researches on response speeds for reading skills.

Etik Beyan: “Okuma Becerilerine Yönelik Maddeleri Yanıtlama Hızlarını Yordayan Özelliklerin Belirlenmesi” başlıklı çalışmanın yazım sürecinde bilimsel, etik ve alıntı kurallarına uyulmuş; toplanan veriler üzerinde herhangi bir tahrifat yapılmamıştır. Çalışmada kamuya açık durumda olan, kişisel bilgi ve etik değerlerin uluslararası standartlarda korunduğu PISA verileri (<https://www.oecd.org/pisa/data/2015database/>) kullanıldığından etik kurul izni gerekmemektedir. Karşılaşılacak tüm etik ihlallerde “Mehmet Akif Ersoy Üniversitesi Eğitim Fakültesi Dergisi Yayın Kurulunun” hiçbir sorumluluğunun olmadığı, tüm sorumluluğun Sorumlu Yazara ait olduğu ve bu çalışmanın herhangi başka bir akademik yayın ortamına değerlendirme için gönderilmemiş olduğunu taahhüt ederim.