



Van Yüzüncü Yıl Üniversitesi Eğitim Bilimleri Enstitüsü

Eğitim Bilimleri Anabilim Dalı

Ölçme ve Değerlendirme Bilim Dalı

PISA BAŞARISINI TAHMİN ETMEDE GENETİK ALGORİTMA YAKLAŞIMI

Yasemin YETKİN

Yüksek Lisans Tezi

Van, 2019

PISA BAŞARISINI TAHMİN ETMEDE GENETİK ALGORİTMA YAKLAŞIMI

Yasemin YETKİN

Danışman

Dr. Öğr. Üyesi Gürol ZIRHLIOĞLU

Van Yüzüncü Yıl Üniversitesi Eğitim Bilimleri Enstitüsü

Eğitim Bilimleri Anabilim Dalı

Ölçme ve Değerlendirme Bilim Dalı

Yüksek Lisans Tezi


Van, 2019

KABUL VE ONAY

Yasemin YETKİN tarafından hazırlanan “PISA Başarısını Tahmin Etmede Genetik Algoritma Yaklaşımı” başlıklı bu çalışma, 04/07/2019 tarihinde yapılan savunma sınavı sonucunda başarılı bulunarak jürimiz tarafından Yüksek Lisans Tezi olarak kabul edilmiştir.

[İ m z a] 
Prof. Dr. Selahattin GELBAL (Başkan)

[İ m z a] 
Dr. Öğr. Üyesi Gürol ZIRHLIOĞLU (Danışman)

[İ m z a] 
Doç. Dr. Hayati ÇAVUŞ

Yukarıdaki imzaların adı geçen öğretim üyelerine ait olduğunu onaylarım.

Doç. Dr. Fuat TANHAN
Enstitü Müdürü

BİLDİRİM

Hazırladığım tezin tamamen kendi çalışmam olduğunu ve her alıntıya kaynak gösterdiğimi taahhüt eder, tezimin kâğıt ve elektronik kopyalarının Van Yüzüncü Yıl Üniversitesi Eğitim Bilimleri Enstitüsü arşivlerinde aşağıda belirttiğim koşullarda saklanmasına izin verdiğimi onaylarım:

- Tezimin/Raporumun tamamı her yerden erişime açılabilir.
- Tezimin/Raporum sadece Van Yüzüncü Yıl Üniversitesi yerleşkesinden erişime açılabilir.
- Tezimin/Raporumun ... ay süreyle erişime açılmasını istemiyorum. Bu sürenin sonunda uzatma için başvuruda bulunmadığım takdirde, tezimin/raporumun tamamı her yerden erişime açılabilir.

04.07.2019



Yasemin YETKİN

Abim Murat'a



TEŞEKKÜR

Tez çalışmam sırasında bilgi ve deneyimini benimle paylaşarak bana destek olan danışmanım Dr. Öğr. Üyesi Gürol ZIRHLIOĞLU'na yardımları için teşekkür ederim.

Tez savunma jürimde bulunan değerli hocalarım Prof. Dr. Selahattin GELBAL'a ve Doç. Dr. Hayati ÇAVUŞ'a katkılarından dolayı teşekkür ederim.

Hayatımın her alanında beni destekleyen, tezimin yazımında bana yardımcı olan, işimi kolaylaştıran sevgisini her zaman hissettiğim abim Murat YETKİN'e sonsuz minnet ve teşekkürlerimi sunuyorum.

Hayatım boyunca maddi ve manevi desteklerini her zaman hissettiğim sevgili ailem M. Muntaz YETKİN'e, Bedriye YETKİN'e, Yasin YETKİN'e, teyzem Muazzez İSTEMKAR'a ve kuzenlerim Semiha YETKİN, Gülperi YALÇINKAYA ve Özlem KESKİN'e teşekkürlerimi sunuyorum.

Tezimin her kısmında bana destek olan, yardımlarını benden esirgemeyen Boğaziçi Üniversite'sinin bana kattığı dostlarım Bedriye DOLUZENGİN'e ve Şafak Cansu DOĞRU'ya sonsuz teşekkür ediyorum.

Bu zorlu süreçte bana hep destek olan, beni hep güldüren, yorulduğumda devam etmem için beni cesaretlendiren ve iyi ki tanımışım dediğim değerli arkadaşlarım Selda&Murat BAYAZIT çiftine, Sevda KAYA'ya, Ayşe Özlem&Yusuf ERGÜL çiftine, Okan AYDIN'a, Merve Nazmiye SEVİM'e, Elif ÖZGÜNER'e, Uğurkan SEKMEN'e ve İsmail ERYİĞİT'e sonsuz teşekkür ederim.

ÖZET

YETKİN, Yasemin. *PISA Başarısını Tahmin Etmede Genetik Algoritma Yaklaşımı*, Yüksek Lisans Tezi, Van, 2019.

Bu çalışmanın amacı genetik algoritma yönteminin öğrenci başarısını yordamada performansını değerlendirmektir. Öğrenci başarısını etkilediği düşünülen değişkenlere ait parametreler genetik algoritma yöntemini kullanarak tespit etmek amaçlanmıştır. Eğitim alanında çok az kullanılan tahminleme yöntemi olan genetik algoritmaların ne ölçüde başarılı çalıştığının belirlenmesi amaçlanmıştır. Birçok alanda başarılı bir şekilde performans gösteren genetik algoritma yönteminin eğitim alanındaki başarı öngörüsünü ortaya koymak amaçlanmıştır. Çalışmada bir bağımlı değişkenle birden fazla bağımsız değişkene bağlı olarak bağımlı değişkenin tahmin edilmesine yönelik yapılan analiz türünde En Küçük Kareler (EKK) yöntemine alternatif bir yöntem olarak genetik algoritma yöntemi kullanılmıştır. Araştırmanın betimsel bir çalışmadır. Araştırmanın veri grubunu PISA 2015 Türkiye uygulamasına katılan öğrenciler oluşturmuştur. Veri seti olarak ise PISA 2015 uygulamasına katılan Türk öğrencilerin matematik, fen ve okuma alanlarındaki başarı testi ve öğrencilere uygulanan öğrenci anket sonuçları kullanılmıştır. Uygulamaya 61 ilden 187 okul ve toplam 5895 öğrenci katılmıştır. Veri seti içerisinde boş ve cevaplanmayan örnekler silinmiş ve çalışmada 5346 öğrenciden elde edilen veriler kullanılmıştır. Araştırmada beş tane bağımsız değişken ve matematik, fen ve okuma puanları bağımlı değişkenler olarak tanımlanmış ve her bir puanın tahminlenmesi için üç ayrı model oluşturulmuştur. Genetik algoritma yöntemiyle yapılan parametre tahmini için Matlab paket programı kullanılmıştır. Genetik algoritma yöntemiyle elde edilen üç farklı regresyon modelinin sonuçları değerlendirilmiştir.

Anahtar Sözcükler

Genetik algoritmalar, PISA, parametre tahmini.

ABSTRACT

YETKİN, Yasemin. *Genetic Algorithm Approach to Estimate PISA Success*, Yüksek Lisans Tezi, Van, 2019.

The aim of this study is to evaluate the performance of the genetic algorithm method in predicting student achievement. It is aimed to determine the parameters of the variables which are thought to affect student achievement by using genetic algorithm method. Another target of the study is also to identify how genetic algorithm method about estimation is successful in the educational field compared to the other fields. It is aimed to reveal prediction of success in education with genetic algorithm method that performs successfully in many fields. In this study, genetic algorithm method was used as an alternative method to Least Squares method for the analysis of estimating dependent variable with one dependent variable and more than one independent variable. The research is a descriptive study. The data group of the research was composed of students who participated PISA 2015 Turkey application. As the data set, the achievement test of mathematics, science and reading of Turkish students participating in PISA 2015 application and student survey results applied to students were used. 187 schools from 61 provinces and a total of 5895 students participated in the application. In the data set, empty and unresponded samples were deleted and the data obtained from 5346 students were used. In the study, five independent variables and mathematics, science and reading scores were defined as dependent variables and three separate models were developed to estimate each score. Matlab package program was used for parameter estimation by genetic algorithm method. The results of three different regression models obtained by genetic algorithm method were evaluated.

Key Words

Genetic algorithm, PISA, parameter estimation.

İÇİNDEKİLER

KABUL VE ONAY	i
BİLDİRİM	ii
<i>Atıf Sayfası</i>	iii
TEŞEKKÜR	iv
ÖZET	v
ABSTRACT	vi
İÇİNDEKİLER	vii
KISALTMALAR	x
TABLolar DİZİNİ	xi
1. BÖLÜM: GİRİŞ	1
1.1. Problem	2
1.1.1. Alt Problemler	3
1.2. Çalışmanın Amacı	3
1.3. Çalışmanın Önemi	3
1.4. Varsayımlar	4
1.5. Sınırlılıklar	4
2. BÖLÜM: KURUMSAL ÇERÇEVE	5
2.1. Genetik Algoritmanın Tanımı	5
2.1.1. Genetik Algoritmaların Özellikleri.....	6
2.1.1.1. <i>Genetik Algoritmaların Avantajları</i>	6
2.2. Genetik Algoritma Tarihçesi	7
2.3. Genetik Algoritmada Temel Kavramlar	7
2.4. Genetik Algoritmada Kodlama	8
2.4.1. İkili Kodlama	8
2.4.2. Permütasyon Kodlama	9

2.4.3. Değer Kodlama.....	9
2.4.4. Ağaç Kodlama.....	10
2.5. Genetik Algoritmada Seçim Metotları.....	10
2.5.1. Rulet Çarkı.....	10
2.5.2. Sıralama Seçimi.....	11
2.5.3. Turnuva Seçimi.....	13
2.6. Genetik Operatörler.....	14
2.6.1. Üreme.....	15
2.6.2. Çaprazlama.....	15
2.6.2.1. Tek Noktalı Çaprazlama.....	16
2.6.2.2. İki noktalı çaprazlama.....	16
2.6.2.3. Çok Noktalı Çaprazlama.....	16
2.6.2.4. Tekdüze (Üniform) Çaprazlama.....	17
2.6.3. Mutasyon (Mutation).....	17
2.7. Genetik Algoritmanın Adımları ve Akış Şeması.....	18
2.7.1. Genetik Algoritmanın Sonlandırılması.....	19
2.8. Genetik Algoritmada Performansı Etkileyen Faktörler.....	19
2.9. Genetik Algoritmaların Kullanım Alanları.....	19
2.10. İlgili Araştırmalar.....	20
3. BÖLÜM: YÖNTEM.....	25
3.3. Veri Toplama Süreci ve Araçları.....	27
3.4. Uygulama.....	28
3.5. Verilerin Analizi.....	29
4.BÖLÜM: BULGULAR.....	35
4.1. Birinci Alt Probleme İlişkin Bulgular.....	35
4.2. İkinci Alt Probleme İlişkin Bulgular.....	41

4.3. Üçüncü Alt Probleme İlişkin Bulgular	47
5.BÖLÜM: SONUÇ, TARTIŞMA VE ÖNERİLER.....	53
5.1. Sonuç ve Tartışma	53
5.2. Öneriler	56
5.2.1. Araştırmacılara Yönelik Öneriler	56
5.2.2. Uygulamaya Yönelik Öneriler	56
ÖZ GEÇMİŞ	63



KISALTMALAR

Bu çalışmada kullanılan bazı simgeler ve kısaltmalar, açıklamaları ile birlikte aşağıda sunulmuştur.

MEB:	Milli Eğitim Bakanlığı
PISA:	Uluslararası Öğrenci Değerlendirme Programı
TIMSS:	Uluslararası Fen ve Matematik Çalışması
PIRL:	Okuma Becerilerinde Gelişim
EKK:	En Küçük Kareler
ço:	Çaprazlama Oranı
mo:	Mutasyon Oranı
n:	Kromozom Sayısı
its:	İterasyon Sayısı
OECD:	Ekonomik İşbirliği ve Kalkınma Örgütü
Matlab:	Matrix Laboratory
ABİDE:	Akademik Becerilerin İncelenmesi ve Değerlendirilmesi

TABLOLAR DİZİNİ

Tablo 1. <i>Cinsiyet Değişkenine Göre Kategorik Dağılımlar</i>	26
Tablo 2. <i>Sınıf Değişkenine Göre Kategorik Dağılımlar</i>	26
Tablo 3. <i>Baba Eğitim Değişkenine Göre Kategorik Dağılımlar</i>	26
Tablo 4. <i>Araştırmada Kullanılan Maddelere Ait Bilgiler</i>	28
Tablo 5. <i>Matematik Puanına Göre EKK Tahminlerinin Güven Aralıkları</i>	31
Tablo 6. <i>Başlangıç Popülasyonu İçin Oluşturulan 1. Kromozom</i>	31
Tablo 7. <i>Fen Puanına Göre EKK Tahminlerinin Güven Aralıkları</i>	32
Tablo 8. <i>Başlangıç Popülasyonu İçin Oluşturulan 1. Kromozom</i>	32
Tablo 9. <i>Okuma Puanına Göre EKK Tahminlerinin Güven Aralıkları</i>	32
Tablo 10. <i>Başlangıç Popülasyonu İçin Oluşturulan 1. Kromozom</i>	32
Tablo 11. <i>Korelasyon Matrisi</i>	36
Tablo 12. <i>Varyans Homojenliği Testi</i>	36
Tablo 13. <i>Matematik Puanına İlişkin Oluşturulan Başlangıç Popülasyonu</i>	37
Tablo 14. <i>Başlangıç Popülasyonuna Ait Uygunluk Değerleri (Matematik)</i>	39
Tablo 15. <i>Genetik Algoritma Yöntemi İle Regresyonda Model Parametre Tahminleri</i> . 41	41
Tablo 16. <i>Korelasyon Matrisi</i>	42
Tablo 17. <i>Varyans Homojenliği Testi</i>	43
Tablo 18. <i>Fen Puanına İlişkin Oluşturulan Başlangıç Popülasyonu</i>	43
Tablo 19. <i>Başlangıç Popülasyonuna Ait Uygunluk Değerleri (Fen)</i>	45
Tablo 20. <i>Genetik Algoritma Yöntemi İle Regresyonda Model Parametre Tahminleri</i> . 46	46
Tablo 21. <i>Korelasyon Matrisi</i>	48

Tablo 22. <i>Varyans Homojenliđi Testi</i>	48
Tablo 23. <i>Okuma Puanına İlişkin Oluşturulan Başlangıç Popülasyonu</i>	49
Tablo 24. <i>Başlangıç Popülasyonuna Ait Uygunluk Deđerleri (Okuma)</i>	50
Tablo 25. <i>Genetik Algoritma Yöntemi İle Regresyonda Model Parametre Tahminleri</i> .	52



ŞEKİLLER DİZİNİ

Şekil 1. İkili kodlanmış Kromozom Örnekleri.....	9
Şekil 2. Permütasyon kodlama ile kodlanmış kromozom örnekleri.....	9
Şekil 3. Değer kodlama ile kodlanmış kromozom örnekleri	9
Şekil 4. Rulet çarkı uygunluk değerleri	10
Şekil 5. Rulet çarkı kromozom dağılımı.....	11
Şekil 6. Sıralama seçimi uygunluk değerleri.....	12
Şekil 7. Sıralamadan önce yüzde değerleri	12
Şekil 8. Sıralamadan sonra yüzde değerleri	13
Şekil 9. Turnuva seçim yöntemi (Bayata, 2012).....	14
Şekil 10. Tek noktalı çaprazlama	16
Şekil 11. Genetik algoritma şeması (Çetin, 2006).....	18

1. BÖLÜM

GİRİŞ

Bireylerin deęişen dünyaya uyum saęlayabilmesi için meydana gelen deęişim ve gelişimlerden haberdar olması gerekmektedir. Eęitimin gelişen ve deęişen dünyadan en çok etkilenen mekanizmalardan bir olduęu yadsınamaz bir gerçekliktir. Geçmişten bu yana her alanda meydana gelen deęişimler ve toplumu etkileyen gelişmeler hiç kuşkusuz ülkeleri eęitim hedef ve ihtiyaçlarını deęiştirmeye yönlendirdięi açıkça görölmektedir. Özellikle ekonomi ve teknoloji gibi alanlarda meydana gelen deęişimleri takip edebilecek bireylerin yetişmesi noktasında eęitim önemli bir role sahiptir (MEB, 2016).

Bu bağlamda, gelişen dünyaya ayak uydurabilmek için ülkeler eęitim politikalarını sürekli gözden geçirmek durumundadır. Bilgi çaęı olarak nitelendirilen bu dönemde bireyler niteliklerini çaęa uygun olarak güncellemek durumundadır (Erginer, 2006). Bireylerin niteliklerini geliştirebilmeleri için ülkelerin eęitim politikalarının etkisi büyüktür. Bu yüzden ülkeler eęitim sistemlerini gözden geçirebilmesi için deęerlendirme yöntemlerine başvururlar. Bu yöntemlerden biri de öğrenci başarılarının incelenmesidir.

Ülkelerin eęitimdeki başarılarını ve bu başarıları etkileyen faktörleri doęru bir şekilde deęerlendirebilmesi gerekmektedir. Bu noktada, uluslararası düzeyde yapılan uygulamalar ülkelerin dięer ülkelere göre konumlarını görmesi açısından önemli bir fırsattır. Bu amaçla yapılan PISA (Uluslararası Öğrenci Deęerlendirme Programı), TIMSS (Uluslararası Fen ve Matematik Çalışması) ve PIRLS (Okuma Becerilerinde Gelişim) uygulamaları eęitim düzeyini belirlemede ve eksikleri gidermede ülkelere yol gösteren araştırmalardır.

Ülkeler arasındaki başarıyı yordama ve karşılaştırma yapmak için uygulanan uluslararası sınavlar öğrenci başarısını yordamada önemli bir kaynaktır. PISA süresince yapılan uygulamalar ve uygulama esnasında elde edilen anket sonuçları bir sonraki uygulama hakkında yordama yapılabilmesini saęlar. Eęitimde yapılan bu tür tahminler sayesinde ülkeler eęitimde başarıyı etkileyen faktörleri belirleyip gerekli tedbirleri alabilirler.

İstatistiksel olarak önemli bir kavram olan tahmin, bir değişkenin gelecekteki dönemler için belirli varsayımlar altında yaklaşık değerinin tahmin edilmesi olarak tanımlanabilir (Kayım, 1985). Gelecekle ilgili kestirim yapabilmek için var olan durumu ortaya çıkaran ya da açıklayan değişkenlerin belirlenmesi ve bu değişkenlerin durumu ne ölçüde açıkladığının tespit edilmesi gerekmektedir (Yıldız Aybek, 2016). Eğitimde tahmin eğitim sistemlerini düzenlemek amacıyla kullanılan önemli bir araçtır. Ülkeler eğitim alanındaki başarılarını arttırmak ve geliştirmek için yordamaya ihtiyaç duyarlar. Çünkü eğitimde öğrenci başarısı kadar başarının tespiti de önemlidir. Eğitimde öğrenci başarı seviyesini en üst düzeye çıkarmak bütün akademik bölümlerin üzerinde durduğu en önemli amaçlardan biridir. Öğrenci başarısını öngörmek amacıyla bir tahmin modeli geliştirmek için çeşitli çalışmalar yapılmaktadır. Tahmin modeli geliştirmek için birçok yöntem vardır. Regresyon analizi tahminleme modellerinde kullanılan analiz türlerinden bir tanesidir. Regresyon analizi değişkenler arasındaki ilişkiyi modellemek ve araştırmak için kullanılan en yaygın istatistiksel yöntemdir. Bu analizde oluşturulacak modelde gerekli olan regresyon katsayılarını hesaplamak için kullanılan en yaygın ve klasik yöntem ise En Küçük Kareler (EKK) yöntemidir (Coşkuntuncel, 2010). Genetik algoritmalar EKK yöntemine alternatif olarak kullanılan bir yöntemdir.

Genetik algoritma arama ve optimizasyon algoritmaları olup evrim kurallarına göre bir problem için en iyi çözümü bulmada kullanılan sezgisel bir yöntemdir (Goldberg, 1989; Gen ve Cheng, 2000; Metin, 2012; Dilaver, 2015). Genetik algoritma yöntemi optimizasyon, finans, gezgin satıcı problemleri, çizelgeleme problemleri ve tahminlemede kullanılır. Birçok alanda olduğu gibi regresyon modeli oluşturmak için gerekli olan parametreleri yordamada genetik algoritma yöntemi kullanılabilir.

1.1. Problem

Bu çalışmanın ana problemi “Türk öğrencilerin PISA başarılarının tahminlenmesinde genetik algoritma yönteminin performansına ilişkin sonuçlar nasıldır?” şeklindedir.

1.1.1. Alt Problemler

PISA 2015'e katılan Türk Öğrencilerin;

1. Matematik Becerilerine ilişkin olarak genetik algoritma ile yapılan başarı yordamasının performansı nasıldır?
2. Fen Bilimleri Yeterliğine ilişkin olarak genetik algoritma ile yapılan başarı yordamasının performansı nasıldır?
3. Okuma Yeterliğine ilişkin olarak genetik algoritma ile yapılan başarı yordamasının performansı nasıldır?

1.2. Çalışmanın Amacı

Çalışmamızın amacı farklı alanlarda tahminlemede kullanılan genetik algoritmaları eğitim alanında geleneksel yöntemlere bir alternatif olarak kullanıp, PISA 2015 uygulamasına katılan öğrencilerin başarılarına yönelik bir yordama gerçekleştirebilmektir. Öğrenci başarısını etkilediği düşünülen değişik değişkenleri genetik algoritma yöntemi kullanılarak tespit etmek ve eğitim alanında çok az kullanılan tahminleme yöntemi olan genetik algoritmaların ne ölçüde başarılı çalıştığının belirlenmesi amaçlanmıştır. Birçok alanda başarılı performans gösteren genetik algoritmaların eğitim alanındaki öngörü performansını ortaya koymak amaçlanmıştır. Tahminleme yapılırken, sayısal verilerin yanı sıra uygulamaya katılan öğrencilere yapılan çeşitli anketlerden yararlanılmıştır. Bu sayede PISA uygulamalarına katılan öğrencilerin başarılarını etkileyen değişkenlerin etkili bir tahmin metodu ile belirlenmesi ve başarıyı arttırmak için gerekli önlemlerin alınması hedeflenmiştir.

1.3. Çalışmanın Önemi

Bu araştırmanın genetik algoritmalar ile yapılacak başarı yordamasını değerlendirmek ve eğitim alanında öngörü konusunda kullanılabilir yeni bir değerlendirme yöntemi olması nedeniyle literatüre ışık tutacağı düşünülmektedir.

Yapılan literatür tarama çalışmalarında Türkiye’de daha önce genetik algoritma yöntemi kullanılarak eğitimde yordama ile ilgili bir çalışmaya rastlanmamış olması çalışmanın özgün yanını oluşturmaktadır. Bu çalışma sayesinde genetik algoritma yöntemini kullanılarak geliştirilen modellerin, eğitim alanında geçmişten beri kullanılan klasik yöntemlere bir alternatif oluşturup oluşturamayacağının belirlenmesi açısından önemlidir. Ülkemizde PISA uygulamalarına katılan öğrencilerin sosyo-ekonomik, kültürel ve çevresel değişkenlere tutumları göz önünde bulundurularak başarı öngörüsünün eğitim alanındaki kullanım alanlarıyla ilgili yapılan araştırmalara katkı sağlaması açısından önemlidir.

1.4. Varsayımlar

Araştırmada PISA 2015 projesine katılan öğrencilerin test maddelerine ve sınav öncesinde yapılan anketlere verilen cevapların samimi olduğu varsayılmıştır.

1.5. Sınırlılıklar

Bu araştırma 2015 PISA uygulamasına katılan Türk öğrencilerine ait olan başarı testi ve öğrencilere uygulanan anketlerle, başarı öngörüsünde kullanılan genetik algoritma yönteminin yorumlanmasıyla sınırlıdır.

2. BÖLÜM

KURUMSAL ÇERÇEVE

2.1.Genetik Algoritmanın Tanımı

Yapay zeka ve yapay yaşamın oluşturulmasının amaçları bilgisayar çağının başlangıcının izini sürebilir. En eski bilgisayar bilimcileri (Alan Turing, John von Neumann, Nobert Wiener ve diğerleri) bilgisayar programları büyük ölçüde çevresini kontrol etmek ve öğrenmek için zeka ile, kendi kendini kopyalayabilme becerisi ile ve uygun yetenekleri ile özümseme vizyonları tarafından harekete geçirilmiştir. Bilgisayar biliminin öncüleri elektronikte olduğu kadar biyoloji ve psikoloji ile de ilgilenmişler ve doğal sistemlere, vizyonlarına nasıl ulaşacaklarına rehberlik eden metaforlar olarak bakmışlardır. Böylece, bilgisayarlara beynin modellenmesi, insan öğrenimini taklit etmesi ve biyolojik evrimin benzerini yapma uygulandı. İlk olarak sinir ağlarında, ikinci olarak makine öğreniminde ve üçüncü olarak en belirgin örnekleri genetik algoritmalar olan “evrimsel hesaplama” olarak adlandırılan alanlarında uygulandı (Melanie, 1999).

Genetik algoritmalar arama algoritmaları olup doğal genetiğin ve doğal seçilimin tekniklerine dayanır (Goldberg, 1989). Yani, genetik algoritmaların temeli doğal seçilime dayanmaktadır. Doğal dünyadaki kısıtlamalar ve gerilmeler farklı türleri en uygun nesilleri üretmek için rekabete zorlar. Genetik algoritmalar da çeşitli çözümlerin en uygunlarını karşılaştırır ve potansiyeli çok güçlü olan çözümler daha uygun çözümler için evrimleşir (Metin ,2012).

Gen ve Cheng’e (2000) göre güçlü ve geniş skolastik arama ve optimizasyon teknikleri olarak genetik algoritmalar belki de en yaygın olarak bilinen evrimsel hesaplama yöntemlerinden biridir.

Genetik algoritmalar evrim kurallarına göre bir problem için en iyi çözümü bulmada kullanılan sezgisel bir arama yöntemidir (Ghareb, Bakar ve Hamdan, 2016; Dilaver, 2015).

Genetik algoritmalar doğadaki canlıların geçirdiği süreci örnek alır ve güçlü nesillerin hayatta kalıp, güçsüzlerin yok olması prensibine dayanır (Michalewicz, 1996; Biroğul, 2005). Bu prensibin yaklaşımıyla genetik kurallara uyum sağlayan canlıların

hayatta kalması diğerlerinin ise elenmesine neden olur.

2.1.1. Genetik Algoritmaların Özellikleri

David Tate'e göre genetik algoritmaların özellikleri aşağıdaki gibidir (1992).

- Uygun çözümlerin bir veya daha fazla "popülasyon" u vardır.
- Önceden bilinen çoklu çözümlerin özelliklerini birleştirerek, yeni uygun çözümler üreten mekanizmaya sahiptir.
- Önceden bilinen bir çözümün rastgele bir şekilde düzenini değiştirerek yeni uygun bir çözüm üreten mekanizmaya sahiptir.
- Daha iyi objektif fonksiyon değerlerine sahip olanlara öncelik vererek popülasyonlardan bireysel çözümler seçen bir mekanizmaya sahiptir.
- Popülasyonlardan çözümlerin çıkarılmasını sağlayan mekanizmaya sahiptir.

Genetik algoritmaları diğer arama yöntemlerinden ayıran en belirgin özellik çözüm arama şeklidir. Goldberg bu farklılıkları aşağıdaki şekilde açıklamaktadır (1989):

1. Genetik algoritmalar parametrenin kendisiyle değil, parametre kodlarıyla uğraşır.
2. Genetik algoritmalar tek bir noktadan değil, popülasyonun büyüklüğü kadar noktadan arama yapar.
3. Genetik algoritmalar sonuç bilgisini kullanır, yani önceden edinilmiş bilgi yerine amaç fonksiyonundan elde edilen bilgiyi kullanır.
4. Genetik algoritmalar olasılıksal dönüşüm kurallarını kullanır, belli ve kesin yöntemleri kullanmazlar.

2.1.1.1. Genetik Algoritmaların Avantajları

Genetik algoritmaların avantajları arasında;

- Parametrelerin sürekli veya ayrık olmasının fark etmemesi,
- Türevsel bilgilerin gerekmemesi,
- Çok sayıda parametre ile çalışma imkanı olması,
- Paralel kodlama imkanının bulunması,
- Lokal optimumlara takılmaması,

Sayılabılır (Svanandam ve Deepa, 2008; Kaya, 2010).

2.2. Genetik Algoritma Tarihçesi

Genetik algoritmalar ilk olarak John Holland tarafından 1960'lı yıllarda tanımlandı ve daha sonra Holland, öğrencileri ve çalışma arkadaşları tarafından Michigan Üniversitesi'nde 1960'lı ve 1970'li yıllarda geliştirilmiştir (Melanie, 1999). Mekanik öğrenme konusunda çalışan Holland, Darwin'in evrim kuramından etkilenmiş ve canlılardaki genetik süreci bilgisayar ortamında gerçekleştirmeyi denemiştir (Reeves ve Rowe, 2002; Daban, 2004). Çalışmalarının sonucu olarak 1975'te "Doğal ve Yapay Sistemlerde Adaptasyon" (Adaptation in Natural and Artificial Systems) isimli kitabı yayınlanmıştır ve bu kitap genetik algoritmaların birincil yazarı olarak kabul edilmiştir (Goldberg, 1989). Holland'ın kitabında genetik algoritmaları biyolojik evrimin bir soyutlaması olarak sunulmuş ve genetik algoritma altında adaptasyona teorik çerçeve vermiştir (Melanie, 1999).

Holland'ın öğrencisi olan David E. Goldberg'in 1985 yılında hazırlamış olduğu "Gaz boru hatlarının genetik algoritma kullanılarak denetlenmesi" konusundaki tezi sayesinde genetik algoritmaların gelişimi sağlanmıştır. Bu uygulamadan sonra Goldberg'in "Makine Öğrenmesi ve Arama Optimizasyonu için Genetik Algoritmalar" adlı kitabında genetik algoritma ile ilgili farklı uygulamaların gösterilmesiyle genetik algoritmaya yeni bir boyut kazandırmıştır (Goldberg, 1989). Bunların yanı sıra John Koza (1995) genetik algoritmaları kullanarak program tasarlamış ve bu yöntem genetik programlama demıştır. Daha sonra pek çok araştırmacı tarafından araştırılmış ve geliştirilmiştir. Mühendislik alanı başta olmak üzere birçok alanda uygulanmıştır. Genetik algoritmalar günümüzden yaklaşık 40 yıl önce kullanılmaya başlanmıştır. Pek çok çalışmada alternatif metot olmak yerine diğer sezgisel yöntemlere üstün geldiği saptanmıştır (Bayata, 2012).

2.3. Genetik Algoritmada Temel Kavramlar

Gen: Kendi başına anlamı olan ve genetik bilgi taşıyan en küçük genetik birimdir. Kromozomların belli konumunda bulunan özel işlev taşıyan her genin bir değeri vardır. Bu değerlere "allele" denir. Örneğin; göz rengini temsil eden bir gende siyah, kahverengi, mavi veya yeşil bu genin allelleri olur. Genetik algoritmalarda her problemin karar değişkeni kadar geni vardır.

Kromozom: Genlerin bir dizi halinde sıralanmasıyla ortaya çıkan genler dizisine kromozom denir. Kromozomlar toplumdaki bireyler ya da üyelere karşılık gelirler. Genetik algoritmalarda her kromozom potansiyel optimum çözümü temsil eder (Genel,2004; Biroğul, 2005). Algoritma başlangıcında genetik algoritma işlemleri sırasında kromozomların gen sırası asla değiştirilmemelidir.

Popülasyon: Kromozomlar veya bireyler topluluğudur. Popülasyon üzerinde durulan alternatif çözüm kümesidir.

Uygunluk Değeri ve Uygunluk Fonksiyonu: Popülasyondaki her bir bireyin yani kromozomun hesap değeridir. Bireyin uyumunu gösteren değerdir. Bu değerlerin belirlenebilmesi için bir fonksiyona ihtiyaç duyulur. Uygunluk fonksiyonları türev ya da analitik işlemler gerektirmeden istenilen şekilde oluşturulabilir (Dilaver, 2015).

Uygunluk değeri bireyin yaşama şansı ile orantılıdır. Uygunluk değeri yüksek olan bireyin yaşama ve çoğalma şansı daha fazla olur ve yeni nesle aktarılma oranı da daha yüksek olur.

2.4. Genetik Algoritmada Kodlama

Genetik algoritma işlemine öncelikle parametrelerin kodlanmasıyla başlanır. Kodlama parametrelerin genlere yani kromozomlara nasıl aktarılması gerektiğini belirlemektir. Holland kromozomları iki basamaklı diziler şeklinde kodlanmıştır (Davis, 1991).

Bir kromozom temsil ettiği çözüm hakkında bilgi içermelidir. En çok kullanılan kodlama ikili dizidir. Böylece kromozom şöyle görülmektedir.

Kromozom1: 1101100100110110

Kromozom2 : 1101111100000110

2.4.1. İkili Kodlama

En çok kullanılan kodlama yöntemidir. Diğer yöntemlere göre daha basittir. İkili kodlamada her kromozom bit (0 veya 1) karakter dizilerinden oluşmaktadır. İkili kodlamalar çok çeşitli kromozomlar verdiği için genetik işlemlerden sonra kodlamalarda düzeltmeler yapılması gerekir (Bayata, 2012).

KROMOZOM A	1100101110100100
KROMOZOM B	10000110101010100

Şekil 1. İkili kodlanmış Kromozom Örnekleri

2.4.2. Permütasyon Kodlama

Sıralama problemlerinde ve düzenleme problemlerinde kullanılır. Permütasyon kodlama gezgin satıcı ve çizelgeleme problemleri için kullanılan bir yöntemdir. Permütasyon kodlamada, her kromozom sırada konum belirten numara karakter dizisinden oluşur.

KROMOZOM A	1 4 7 3 9 2 9 4 5 2 2
KROMOZOM B	8 7 5 3 1 9 8 8 2 7 6

Şekil 2. Permütasyon kodlama ile kodlanmış kromozom örnekleri

2.4.3. Değer Kodlama

Gerçek sayılar gibi karmaşık değişkenlerin kullanıldığı problemlerde, ikili kodlama zor olacağı için değer kodlama kullanılır. Her kromozomun bazı değerlere eşit olduğu kodlama türüdür. Değerler problemle ilgili herhangi bir şeyi belirtebilir. Bu kodlama da gerçek sayılar, karakterler, harfler veya nesnelere kullanılabilir (Bayata, 2012).

KROMOZOM A	1,256 4,456 1,742 7,987
KROMOZOM B	FKLFDKFKDFKDFKLDJKLJ
KROMOZOM C	(AŞAĞI),(YUKARI),(SAĞ),(SAĞ)

Şekil 3. Değer kodlama ile kodlanmış kromozom örnekleri

2.4.4. Ağaç Kodlama

Genellikle programlamada ve ifadeler için kullanılmaktadır. Her kromozom nesnelere ya da nesnelere arası işlem içeren bir ağaç yapısından oluşmaktadır.

2.5. Genetik Algoritmada Seçim Metotları

Yeni nesil için seçilecek olan kromozomlar belirlenirken başlangıç popülasyonundaki kromozomlardan en iyi uygunluk değerine sahip olanlardan bazıları yeni topluma aktarılır. Bazıları ise ebeveyn olarak seçilmektedir. Ebeveyn olarak seçilen kromozomlara uygulanan genetik işlemler sonucunda yeni nesle aktarılır. Seçim metotları olarak geliştirilen birçok yöntem vardır. Bu yöntemlerden en yaygın olanları rulet çarkı, sıralama yöntemi ve turnuva seçim yöntemidir (Yeo ve Agyei, 1998).

2.5.1. Rulet Çarkı

İlk defa Holland tarafından ortaya çıkarılan bir yöntemdir. Bu yöntemde toplumdaki her bireyin uygunluk değeri hesaplanır. Hesaplanan uygunluk değerleri toplanarak popülasyonun uygunluk değeri elde edilir. Her kromozomun uygunluk değeri toplam uygunluk değerine bölünür ve bireyin seçilme olasılığı elde edilir. Seçilme olasılıkları ölçüsünde kromozomlar rulet çarkında yer alırlar. Böylece olasılık değeri yüksek olan bireylerin seçilme olasılıkları artmaktadır. Rulet çarkı üzerinden seçimler rastgele, en yüksek olasılık değerine göre veya eklemeli oranlardan yararlanılarak yapılır. Eklemeli oranlar, seçilme olasılıklarının toplanmasıyla elde edilir. Rastgele bir sayıya erişinceye kadar eklemeli oranlar elde edilir ve son eklenen sayının çözümü seçilerek işlem tamamlanır (Bayata,2012).

	UYGUNLUK DEĞERİ
KROMOZOM A	15
KROMOZOM B	58
KROMOZOM C	22
KROMOZOM D	5
TOPLAM UYGUNLUK	100

Şekil 4. Rulet çarkı uygunluk değerleri

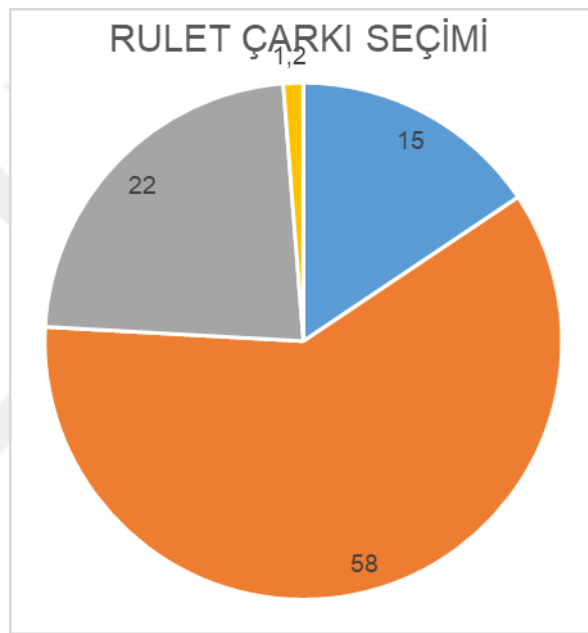
Rulet arkı algoritması:

Adım 1: Popülasyondaki tüm bireylerin uygunluk değeri hesaplanır ve toplanır (S).

Adım 2: N kez tekrarlar (0, s) aralığında rassal sayı (r) üretilir.

Adım 3: Popülasyon boyunca toplam uygunluk 0-S arasında arama yapılır. Toplam s değeri r değerinden büyük olduğunda dur ve elde edilen kromozomu yeni birey olarak al.

Bu şekilde her popülasyon için bir kez Adım 1'e gidilir (Dilaver, 2015).



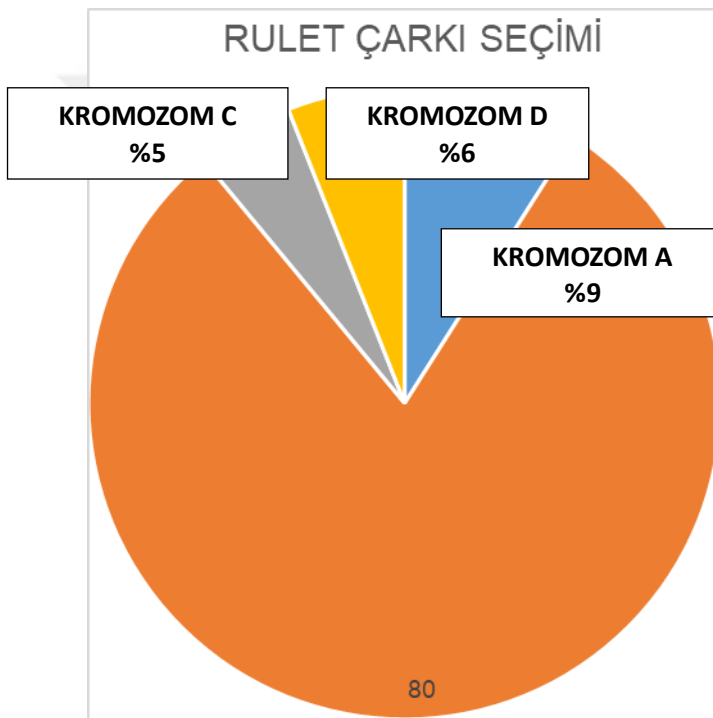
Şekil 5. Rulet arkı kromozom dağılımı

2.5.2. Sıralama Seçimi

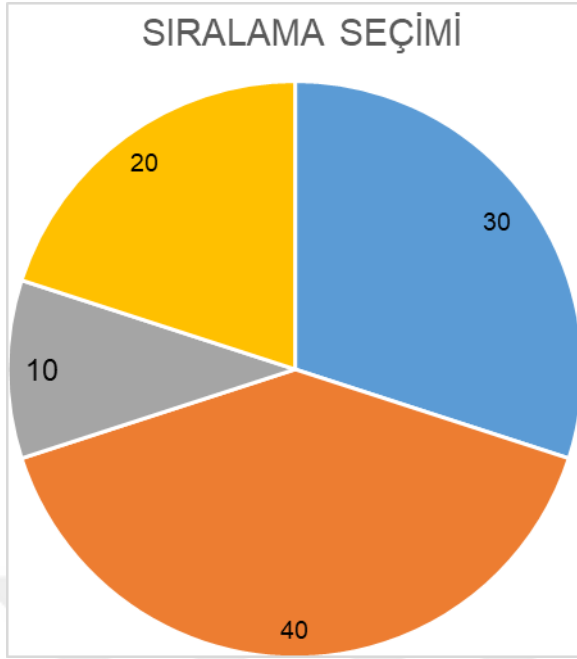
Uygunluk değeri arasındaki farklılığın çok fazla olduğu durumlarda kullanılır. Örneğin uygunluk değeri diğer tüm uygunluk değerlerinin %90'ını oluşturan bireyin seçilme olasılığı çok büyükken diğerlerinin seçilme olasılığı çok küçük olur. Sıralama seçiminde önce popülasyon sıralanır. Sıralama uygunluk değeri en kötü olandan en iyi olana doğru yapılır (Bayata, 2012).

	UYGUNLUK DEĞERİ	SIRALAMA DEĞERİ
KROMOZOM A	9	3
KROMOZOM B	80	4
KROMOZOM C	5	1
KROMOZOM D	6	2
TOPLAM UYGUNLUK	100	10

Şekil 6. Sıralama seçimi uygunluk değerleri



Şekil 7. Sıralamadan önce yüzde değerleri



Şekil 8. Sıralamadan sonra yüzde değerleri

2.5.3. Turnuva Seçimi

Turnuva yönteminde popülasyondan kromozomlar rastgele seçilir ve oluşturulan grup içerisindeki kromozomlar kıyaslanarak en yüksek uygunluk değerine sahip birey yeni popülasyona aktarılır. Böylece oluşturulan yeni popülasyon, ilk popülasyondaki kötü bireylerden arındırılmaya çalışılır. Kolaylığı ve uygunluğu açısından en yaygın seçim yöntemidir (Dilaver, 2015).

Turnuva seçimi için birçok teknik geliştirilmiştir. Bunlardan biri uygunluk değeri belirlenen ve rastgele seçilen n adet kromozoma, 1 den başlayarak sıra numarası verilir. Daha sonra sıra numarası aralığında rastgele sayılar üretilir. Rastgele sayılardan yararlanarak belirlenen iki birey kıyaslanır ve uygunluk değeri yüksek olan birey seçilir (Bayata, 2012).

	UYGUNLUK DEĞERİ	NO	RASTGELE SAYILAR
KROMOZOM A	15	1	5
KROMOZOM B	26	2	4
KROMOZOM C	5	3	1
KROMOZOM D	43	4	8
KROMOZOM E	26	5	9
KROMOZOM F	30	6	7
KROMOZOM G	66	7	6
KROMOZOM H	21	8	2
KROMOZOM I	6	9	3

TURNUVA 1
KROMOZOM A (1)- KROMOZOM E (5)

KAZANAN
KROMOZOM E

Şekil 9. Turnuva seçim yöntemi (Bayata, 2012)

Seçilimli üremenin amacı yüksek uyumluluğa sahip bireylerin yükselmesinin yanı sıra ölümcül bireylerin üremesinden kaçınmaktadır (Gen ve Cheng, 2000).

2.6. Genetik Operatörler

Genetik algoritmalarda problemlerin çözümünde kullanılan üç temel genetik işlem uygulaması bulunmaktadır. Bunlar üreme (reproduction), çaprazlama (crossover) ve mutasyon (mutation) operatörleridir.

2.6.1. Üreme

Popülasyondaki kromozomlardan üreme uygunluk kriterlerine uyan bireylerin özelliklerinin yeni popülasyona aktarılmasını sağlayan operatördür. Seçilen kromozom yeni toplum için oluşturulacak kromozomlara katkı sağlar.

Üreme işlemi genellikle üç şekilde yapılır. Birincisi, bir tanesinin tamamen yer değiştirmesidir. İlk nesil yeni oluşturulan nesil ile tamamen yer değiştirir. En önemli dezavantajı, yüksek uygunluk değerine sahip bireylerin kaybolma ihtimallerinin olmasıdır (Bayata,2012). İkincisinde belirli sayı veya oranda yavru üretilerek yeni oluşan nesilde eskileriyle beraber yer almasıdır. Örneğin eski nesildeki bireylerin %45 inin yeni nesilde yer alması gibi. Üçüncüsü sabit durum seçimi özel bir ata seçme yöntemi değildir. Bu seçimin amacı, bireylerin çoğunluğunun yeni nesle aktarılmasıdır (Bayata,2012).

Üreme sırasında ilk önce çaprazlama yapılır. Atalardan gelen genler yeni kromozom için bir araya gelirler. Daha sonra bu genler mutasyona uğrayabilir (Bayata, 2012).

Tüm canlı topluluğu bir kere oluştuğunda, her birinin uygunluk değeri ölçülerek toplam uyum bulunabilir. Eğer toplam uyumluluk istenen düzeyde değilse, uygunluk değeri kötü olan birey nesli tükenmek üzere seçilebilir. Bu olay elitist (seçkin) doğal seleksiyon işlemi olarak adlandırılır (Davis, 1991).

2.6.2. Çaprazlama

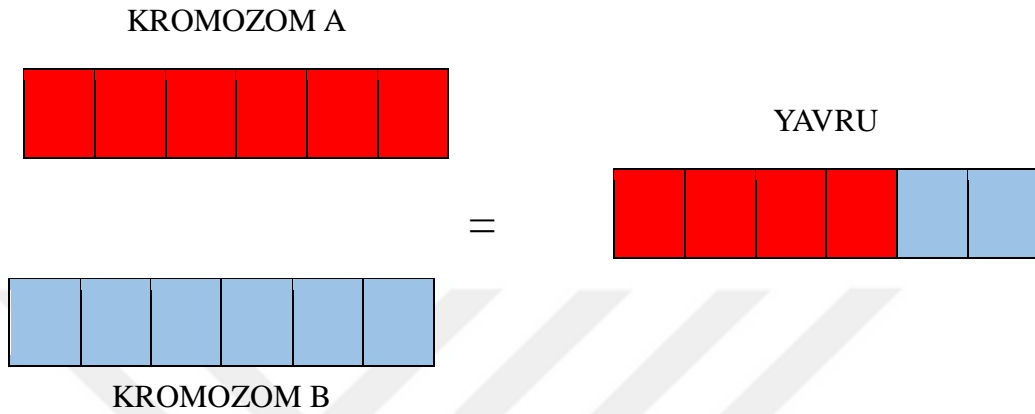
Seçilim sürecinde seçilen eski kromozomlardan gelen genlerin yer değiştirmesi ile yeni nesil için iki yeni kromozom oluşturulmasıdır. Kromozomların iyi özelliklerinin bir araya gelmesini kolaylaştırır (Taştan, 2015).

Genetik algoritmanın performansını etkileyen en önemli parametreden biri de çaprazlamadır (Dilaver, 2015). Popülasyondan rastgele seçilen iki bireyin belli bölümlerinin karşılıklı yer değiştirmesiyle iki yeni birey elde edilir.

Çaprazlama işlemleri farklı şekillerde gerçekleştirilebilir. En sık kullanılan çaprazlama yöntemleri: tek noktalı çaprazlama, iki noktalı çaprazlama, çok noktalı çaprazlama ve üniform çaprazlamadır.

2.6.2.1. Tek Noktalı Çaprazlama

En basit ve en geleneksel çaprazlama yöntemidir (Gen ve Cheng, 2000; Davis, 1991). Rastgele seçilen kromozom çiftinden çaprazlama yapılacak bölge de rastgele seçilir. Seçilen bölgeler yer değiştirir. Oluşan yeni bireyler ebeveynlerinin bazı özelliklerini alarak iki bireyin kopyasını oluşturur (Metin, 2012).



Şekil 10. Tek noktalı çaprazlama

Şekil 10'da görüldüğü gibi çaprazlamada oluşan yavru kromozomun başlangıcı bir atanın geri kalanı tamamen diğer atanın genleri taşımaktadır.

Tek noktalı çaprazlamada amaç farklı dizilerdeki yapı taşlarını birleştirmektir. Ama çaprazlamanın bazı eksiklikleri vardır. Bunlardan biri, çaprazlama bütün olası şemaları birleştiremez.

2.6.2.2. İki noktalı çaprazlama

Çaprazlama işleminin birden fazla çaprazlama noktasıyla gerçekleştiği çaprazlama türüdür (Gen ve Cheng, 2000). Bu işlemde iki kesme noktası belirlenir. Kromozomun başından ilk kesme noktasına kadar birinci bireyden, iki kesme noktası arası ikinci bireyden ve ikinci kesme noktasından sonraki kısım yine birinci bireyden alınır ve yeni birey oluşturulur.

2.6.2.3. Çok Noktalı Çaprazlama

İki noktalı çaprazlama yöntemine benzemektedir. Farkı ikiden fazla kesme noktasının olmasıdır. Parçalar çiftler arasında yer değiştirilerek oluşturulur. Çok noktalı

çaprazlamada parça sayısının karesi kadar farklı kromozom oluşturulur (Şen, 2004; akt: Bayata, 2012).

2.6.2.4. Tekdüze (Üniform) Çaprazlama

Çok amaçlı sistemlerde kullanılan yöntemdir. Bu çaprazlamada her bir genin iki atadan gelme olasılıkları eşittir. Bitler atalardan rastgele seçilerek kopyalanır.

Çaprazlama maskesi ikili sayı sisteminde rastgele oluşturulan diziyeye denir. Tekdüze çaprazlamada bit sayısı kadar çaprazlama maskesi kullanılır. Çaprazlama maskesinin ilgili geni yerine, birinci ve ikinci bireye karşılık gelen genlerin kopyalanması mantığına dayanır. Tekdüze çaprazlamada birinci birey, maskede 1 kodu görüldüğü yere, birinci bireyden karşılık gelen gen kopyalanırken; 0 kodunun olduğu yere ikinci bireydeki karşılığı yazılır (Bolat, Erol ve İmrak, 2004)

2.6.3. Mutasyon (Mutation)

Çaprazlama işlemi yaptığımızda tek noktalı çaprazlamada iki ata da aynı alele sahipse çaprazlama işlemi ne kadar yapılırsa yapılsın her zaman genin alelesi aynı kalır. Bu tür durumlarda mutasyon popülasyona çeşitlilik katmak için işleve girmektedir (Goldberg, 1989). Yani, çaprazlama sonucu elde edilemeyecek farklılıkları oluşturmak için mutasyon kullanılır. Mutasyon yoluyla kromozomdaki genler değiştirilir.

Mutasyon operatörleri; ters çevirme, yer değişikliği, ekleme, karşılıklı değişimdir.

Ters Çevirme (invertion): Rastgele seçilen alt dizideki genlerin sıraları tamamen tersine çevrilir ve yeniden yerleştirilir. Kodlanması diğerlerine göre daha zordur.

Yer Değişikliği: Mutasyonda rastgele bir alt dizi seçilir ve rastgele bir yere yerleştirilir.

Karşılıklı Değişim: Rastgele seçilen iki genin yerleri değiştirilmektedir.

Ekleme: Rastgele seçilen parça yine rastgele bir konuma yerleştirilir.

Mutasyonda aynı kromozomdaki genler yer değiştirirken çaprazlamada iki kromozom üzerindeki genler yer değiştirirler.

Çaprazlama Oranı: Çaprazlama işleminin popülasyonda bulunan kromozomların belirli bir oranına uygulanmasıdır. Çaprazlama oranı algoritmanın başında ya da her yeni toplum oluşturulmadan önce belirlenmelidir. Bu oranın yüksek olması değişime uğrayan kromozom sayısının fazla olacağını göstergesidir (Paksoy, 2007).

Mutasyon Oranı: Mutasyon oranı algoritmanın başında ya da her yeni toplum

oluşturulmadan önce belirlenmelidir. Mutasyon oranının çok yüksek seçilmemesi gerekmektedir.

2.7. Genetik Algoritmanın Adımları ve Akış Şeması

Adım1: Tüm muhtemel çözümlerin bir dizi olarak kodlandığı bir çözüm kümesi oluşturulur. Problemin türüne göre değişik kodlama türleri vardır.

Adım2: Başlangıç yığını olarak seçilecek olan popülasyon rastgele seçilir.

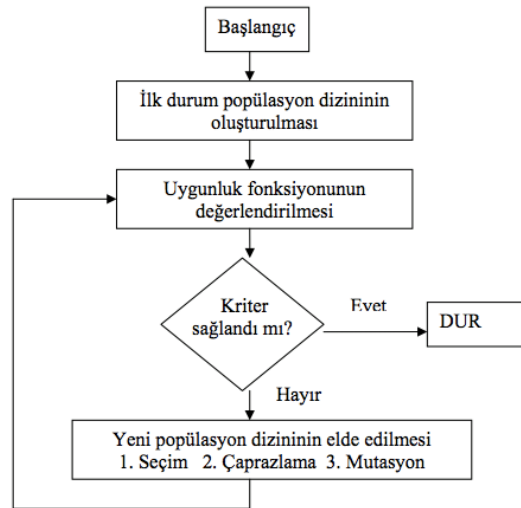
Adım3: Popülasyondaki her bir dizinin uygunluk değerine göre ne kadar iyi olduğu belirlenir. Bu değere göre iyi sonuçlar veren kromozomlar yeni popülasyona için seçilir. Uygunluk değeri popülasyonun kalitesini gösterir.

Adım4: Belli bir olasılık değerine göre bir grup dizi rastsal olarak seçilir ve çoğalma gerçekleşir.

Adım5: Yeni dizilerin uygunluk değerleri hesaplanır ve çaprazlama, mutasyon gibi genetik işlemlere tabi tutulur.

Adım6: Önceden belirlenen kuşak sayısına ulaşıncaya kadar veya durdurma kriterleri sağlanıncaya kadar yukarıdaki işlemler devam edilir. (Adım3'e gidilir)

Adım7: Kuşak sayısına ulaşıldığında veya durdurma kriteri sağlandığında, oluşan kuşaktaki en iyi uygunluk değerine sahip olan kromozom çözüm olarak seçilir.



Şekil 11. Genetik algoritma şeması (Çetin, 2006)

2.7.1. Genetik Algoritmanın Sonlandırılması

Genetik algoritmayı sonlandırmak için kullanılacak en iyi yöntemler:

- ✓ Seçilen en iyi kromozom, bilenen en iyi çözüme ulaştığında,
- ✓ İterasyonlarda tekrarlı bir şekilde, en iyi kromozom aynı olduğunda,
- ✓ Belirlenen istatistiksel değerlendirme erişildiğinde (toplum maliyet ortalaması, standart sapma),
- ✓ Belirlenen iterasyon sayısı tamamlandığında
- ✓ İterasyon sonuçları açısından bir gelişme olmadığında,
- ✓ Optimuma yakın bir değere erişildiğinde şeklinde olmaktadır (Paksoy, 2007).

2.8. Genetik Algoritmada Performansı Etkileyen Faktörler

1. Popülasyon büyüklüğü/kromozom sayısı: Kromozom sayısını arttırmak çalışma süresini arttırırken, azaltmak da kromozom çeşitliliğini yok eder.

2. Mutasyon Oranı: Kromozomlar birbirine benzemeye başladığında hala çözüm noktalarının uzağında bulunuyorsa mutasyon işlemi genetik algoritmanın sıkıştığı yerden kurtulması için tek yoldur. Ancak yüksek bir değer vermek genetik algoritmanın kararlı bir noktaya gelmesini engelleyecektir.

3. Kaç Noktalı Çaprazlama Yapılacağı: Normal olarak çaprazlama tek noktada gerçekleştirilmekte fakat yapılan çalışmalar bazı problemlerde çok noktalı çaprazlamanın yararlı olduğunu göstermiştir.

4. Çaprazlamanın sonucu elde edilen bireylerin nasıl değerlendirileceği: Elde edilen iki bireyin birden kullanılıp kullanılmayacağı bazen önemli olmaktadır.

5. Durum kodlamasının nasıl yapıldığı: Bir parametrenin doğrusal ya da logaritmik kodlanması genetik algoritmanın performansında önemli bir farka yok açar.

6. Başarı değerlendirilmesinin nasıl yapıldığı: Akıllıca yazılmamış bir değerlendirme işlevi, çalışma zamanını uzatabileceği gibi çözüme hiçbir zaman ulaşamamasına neden olabilir.

2.9. Genetik Algoritmaların Kullanım Alanları

- ✓ Optimizasyon
- ✓ Otomatik Programlama ve Bilgi Sistemleri

- ✓ Mekanik Öğrenme
- ✓ Ekonomik ve Sosyal Sistem Modelleri
- ✓ Finans
- ✓ Pazarlama
- ✓ Üretim/İşlemler
- ✓ Montaj Hattı Dengeleme Problemi
- ✓ Çizelgeleme Problemleri
- ✓ Tesis Yerleşim Problemi
- ✓ Atama Problemi
- ✓ Hücresel Üretim Problemi
- ✓ Gezgin Satıcı Problemi
- ✓ Araç Rotalama Problemi

2.10. İlgili Araştırmalar

Aydın (1998) tez çalışmasında çok değişkenli çoklu regresyon problemi olarak hedef programlama problemi modellemiştir. Değişken sayısının klasik regresyon çözümlemesindeki değişken sayısından fazla olduğu görülmüş ve genetik algoritmalar kullanılarak problemi çözmüştür. Uygulamada klasik yöntem ve genetik algoritma yöntemi kullanılarak problem çözülmüş ve sonuçlar karşılaştırılmıştır. Genetik algortmada üç farklı uygulama yapılmış ve olasılık ya da nesil sayısındaki değişikliklerin önemli bir değişiklik olmadığını göstermiştir. Araştırmanın sonucunda genetik algoritma yönteminin klasik yöntemlere göre daha iyi sonuç verdiği görülmüştür.

Daban (2004) çalışmasında yapay zeka yöntemlerinden biri olan genetik algoritma kullanarak, öğretim elemanları ve öğrencilerin ihtiyaçlarına göre daha etkin bir eğitim-öğretim sağlamak için ders programı hazırlayan ve optimize eden bir program geliştirmiştir. Ders programının uygunluğu, öğretim elemanı ve eğitim-öğretim verimliliği kriterlerine göre belirlenmektedir. Pedagojik esaslara göre, öğrencilerin algısının haftanın günleri ve gün içindeki saatlere göre değişimi göz önünde bulundurularak verimliliğin artması sağlanmıştır. Standart genetik algoritma operatörlerinden farklı olarak probleme özgü genetik algoritma operatörleri

geliştirilmiştir. Belirlenen kriterlerde başlangıca göre % 40-60 oranlarında iyileşme sağlamıştır.

Altunkaynak ve Esin'in 2004 yılında yaptıkları araştırmada doğrusal olmayan regresyonda parametre tahmini için alternatif bir yöntem olarak genetik algoritmayı önermişlerdir. Genetik algoritmada parametre tahmini için diğer yöntemler gibi yardımcı bilgiye ihtiyaç duyulmadığını bu yüzden pratik uygulamalar için kullanışlı olduğunu belirtmişlerdir. Çalışmada Gauss-Newton yöntemi ile genetik algoritma yöntemine ilişkin elde edilen sonuçlar karşılaştırılmıştır. Sonuçların birbirine yakın olduğu görülmüş fakat Gauss-Newton yönteminde parametre tahminin başarılı olması için aramanın başlangıç noktasının belirlenmesi gibi yardımcı bilgiler gerektirdiğini ve araştırma evrenin tek bir noktasını incelediğini fakat genetik algoritma hiçbir yardımcı ya da ön bilgiye ihtiyaç duymadan evrenin tamamında arama yaparak sonuca ulaştığını söylemişlerdir. Bu yüzden genetik algoritma yönteminin iyi bir alternatif olduğu sonucuna varmışlardır.

Roeva (2005) yaptığı araştırmada parametre tahmininde genetik algoritmayı kullanmıştır. Fermantasyon probleminde doğrusal olmayan bir E modelinin parametrelerini tahminini değerlendirmiştir. Sınırlamaların üstesinden gelebilmek için araştırmada farklı genetik algoritmaları global optimizasyon yöntemleri olarak kullanımı araştırılmıştır. Üç farklı, basit, düzeltilmiş ve çok popülasyonlu genetik algoritmalar incelenmiş ve karşılaştırılmıştır. Sonuçlar genetik algoritmaların belirlenen problemi çözmeye çok iyi olduğunu göstermiştir. Sapma miktarı %0,4 olarak bulunmuş ve önemsiz olduğu görülmüştür. Düzeltilmiş genetik algoritma modeli diğer iki modele göre daha hızlı sonuç vermiştir. Araştırmanın sonucunda çok popülasyonlu genetik algoritma modeli fermantasyon süreci modeli maliyet fonksiyon değerini minimum yapan değer olmasına rağmen fermantasyon işlemlerinin parametre tahmini için en uygun model olarak düzeltilmiş genetik algoritma modeli önerilmiştir.

Akyol (2006) çalışmasında doğrusal olmayan ekonometrik modellerin çözümünde genetik algoritma yönteminin etkin bir çözüm aracı olarak kullanılıp kullanılmayacağını araştırmıştır. Araştırmanın sonucunda doğrusal olmayan modeller için genetik algoritmanın etkin bir çözüm aracı olarak kullanılabileceği sonucuna varılmıştır.

Öztürkler ve Altan (2008) arařtırmalarında doğrusal olmayan ekonometrik modellerde genetik algoritma yaklaşımıyla parametre tahmini yapmışlardır. Çalışmadan EKK çözümü ile genetik algoritma çözümü gösterilmiş ve sonuç olarak genetik algoritma yöntemi karmaşık doğrusal olmayan ilişkiler içeren ekonometrik problemlerin çözümünde bir araç olarak kullanılabileceđi önerilmiştir.

Karakoca 2009 yılında yürüttüğü tez çalışmasında çok deđişkenli doğrusal olmayan modellerde genetik algoritma yöntemi ile En Küçük Kareler (EKK) yöntemiyle elde edilen parametre tahmin sonuçları karşılaştırılmıştır. Klasik yöntemlerde parametre tahmini için başlangıç noktası ve fonksiyonların türevlenebilir olması koşulu aranırken genetik algoritmada böyle bir öncelik yoktur. Bu yüzden genetik algoritma alternatif yöntem olarak önerilmiştir. Ayrıca arařtırmanın sonucunda genetik algoritma ile elde edilen parametre tahmini klasik yöntemden daha iyi sonuç verdiđi görülmüştür.

Yiđit 2011 yılında Türkiye'nin 2020 yılına kadar ihtiyaç duyduđu elektrik enerjisi tüketimi genetik algoritmalarla tahmin edilmiştir. Geliştirilen modellerin uygunluđu 1979 ile 2009 yılları arasındaki gerçek deđerlerle gösterildikten sonra farklı çalışmalarda karşılaştırılmıştır. Sonuçlar, geliştirilen modellerin Türkiye'nin gelecekteki Elektrik enerjisi ihtiyacını tahmin etmek için uygun ve geçerli olduđunu göstermektedir.

Moreno, Ovalle ve Vicari (2012) çalışmalarında grup oluşumunun işbirlikçi öğrenmede kilit süreçlerden biri olduđu düşünöldüğünde, bu yazının amacı homojen ve heterojen olmayan gruplar için genetik algoritma yaklaşımına dayanan bir yöntem önermektir. Böyle bir yöntemin ana özelliđi, arzu edilen çok sayıda öğrenci karakteristiđinin dikkate alınmasına, gruplama probleminin çok amaçlı optimizasyondan birine dönüştürölmesine izin vermesidir. Yaklaşımlarını dođrulamak için, üç özellik göz önünde bulundurularak 135 üniversite birinci sınıfıyla bir deney tasarlandı: öğrenci bilgi düzeylerinin tahmini, öğrenci iletişim becerilerinin tahmini ve öğrenci liderlik becerilerinin tahmini. Böyle bir denemenin sonuçları, algoritmik performansı ölçerek sadece hesaplama bakış açısından deđil, aynı zamanda öğrenci çıktılarını ölçerek pedagojik bakış açısından ve bunları iki geleneksel grup oluşturma stratejisiyle karşılaştırarak dođrulamayı mümkün kılmıştır: rasgele ve kendini organize etme.

Metin 2012 yılındaki araştırmasında bir sinir ağı eğitimde kullanılacak olan parametreleri genetik algoritma kullanarak belirlemiştir. Sonuç olarak genetik algoritma yapay sinir ağlarını eğiterek modellemede parametre tahmini için kullanılabileceğini göstermiştir.

Bayata (2012) çalışmasında çözümü zor problemler arasında yer alan ders çizelgeleme probleminin genetik algoritmalar kullanılarak çözümünü araştırmıştır. Problemin çözümü için genetik programlama ile C++ dili kullanılarak bir yazılım geliştirmiştir. Yazılımın uygulaması Gazi Üniversitesi Endüstri Mühendisliği bölümü ders programı verileri kullanılarak yapılmış ders programı hazırlanmış ve sunulmuştur. Algoritmanın performansını ölçmek amacıyla genetik parametreler üzerinde deneyler yapılmış ve elde edilen sonuçlar üzerinden genetik algoritma performansı değerlendirilmiştir.

Gerşil ve Palamutçuoğlu (2013) çalışmalarında eğitimde zamanlama konusunun alt bölümü sayılan ders programlarının verimliliğinin artırılması için genetik algoritma kullanılarak ders programı hazırlayan ve daha etkin eğitim-öğretim için geliştirilen yazılım programı verimlilik ve ders ağırlığı kriter olarak kullanılmıştır.

Lakshmi, Martin ve Venkatesan'ın 2013 yılında yaptıkları çalışmada nicel veri analizleri ile öğrencilerin müfredattaki performanslarındaki en etkili faktörleri tanımlamak için yapmışlardır. Bu öğrenci analizi modeli, genetik algoritmayı kullanarak en çok etkileyen faktörleri bulmak için teorik, matematiksel, pratik, bölüm ve diğer bölüm notları gibi kantitatif faktörleri dikkate almaktadır. Sonuçlar, öğrencilerin akademik kariyerlerinde edindikleri notları değerlendirdikten sonra eğitim kurumlarının öğretim kalitesini iyileştirmelerine yardımcı olması için yapılmış bir çalışmadır.

Kabra ve Bichkar (2014) çalışmalarında mühendislik öğrencilerinin iletişim eğitimindeki akademik performansını öngören karar ağacı modellerinin elde edilmesini sağlamayı amaçlamışlardır. Genetik algoritma, iyi karar ağaçlarının elde edilmesinde umut vaat eden güçlü bir arama ve optimizasyon tekniği olduğu için kullanılmıştır. Karar ağaçları, evrimsel algoritmaların yanı sıra açgözlü kullanılarak evrimleşmiştir. Sonuçlar, genetik algoritma ve J48 (WEKA'dan) kullanılarak indüklenen ağacın doğruluğu ve büyüklüğü açısından tartışılmıştır. Ayrıca, birinci sınıf mühendislik öğrencilerinin sonuçları için önemli olan özellikler de tanımlanmışlardır.

Şenol (2014) yaptığı çalışmada koroner arter hastalığının risk faktörlerinin irdelenmesinde genetik algoritma yaklaşımını kullanmış ve bu yöntemi karar ağaçlarında Random Forest (RF) sonuçlarıyla kıyaslamıştır. Genetik algoritma ile optimize edilerek sınıflama yapılmış RF modeli, optimize edilmemiş olan RF modellerine göre çok daha yüksek başarıya sahip olduğunu göstermiştir. Ağaç sayısı kaç olursa olsun kurulan modellerin genetik algoritma yöntemiyle optimize edilmesinin çalışmanın en önemli sonucu olduğunu ifade etmiştir.

2015 yılında Dilaver'in yaptığı çalışmada çizelgeleme problemlerinin bir türü olan iş atölye çizelgeleme problemlerinde, siparişlerin en kısa sürede tamamlanması için sezgisel yöntemlerden biri olan genetik algoritmayı kullanan üç farklı çalışma incelenmiştir. Bu üç çalışmanın her biri atölye çizelgeleme problemi olup, genetik algoritmanın farklı metotları ile çözümlenmiştir. Bu çalışmalarda genetik algoritma ile elde edilen sonuçların ve klasik yöntemlerle olan kıyaslamaları incelenmiştir. Çalışmaların birbirleri arasındaki ortak yönleri ve farklılıkları değerlendirilmiştir. Bu tez sonucunda, genetik algoritmanın farklı metotları, farklı genetik operatörler, farklı varsayım ve farklı koşulları ile elde edilen sonuçlar karşılaştırılmıştır.

Demir ve Akkuş'un 2015 yılında yürüttükleri çalışmada ikili lojistik regresyon modelinin parametre tahmininde iki farklı optimizasyon tekniği tanıtılmış ve karşılaştırılmıştır. Bu tekniklerden birincisi diferansiyel amaç fonksiyonu ve parametrelerle ilgili uygun başlangıç değerleri gerektiren geleneksel Newton-Raphson (NR) algoritmasıdır. Diğer teknik ise genetik algoritma yöntemidir. Araştırma sonucuna göre iki tekniğin sonuçları birbirine çok yakın bulunmuştur. Fakat NR algoritmasında gerekli ön şartlar olması gerekirken genetik algoritma için böyle bir kural yoktur ve genetik algoritma yöntemi lojistik regresyon modeli için parametre tahmininde başarılı sonuçlar verdiği görülmüştür.

Doğan 2016 yılında, çoklu regresyon analizinde çoklu bağlantı sorunu ortaya çıktığında EKK yönteminin tutarsız sonuçlar vermesine neden olduğu ifade edilmiş ve yanlış tahmin yöntemlerinden LASSO tahmin edicisinin uygun çözümü için genetik algoritma kullanılmıştır. LASSO tahmin edicisi sonuçlarının model parametre tahminlerinin güven aralıklarını daralttığı görülmüş ve standart hatalar incelendiğinde EKK yöntemine göre daha küçük değerler aldığı ifade edilmiştir.

3. BÖLÜM

YÖNTEM

Bu bölümde çalışmada kullanılacak araştırma yöntemi, araştırma grubu, verilerin toplanması ve düzenlenmesi, uygulama süreci ve verilerin analizi ile ilgili bilgiler verilmiştir.

3.1. Araştırmanın Yöntemi ve Deseni

Bu çalışmada PISA 2015 uygulamasına katılan öğrencilerin öğrenci anketinde yer alan sorulara verdiği cevaplar yardımıyla Matematik, Fen ve Okuma okuryazarlığı bakımından başarılarını tahmin etmek amaçlanmıştır. Çalışmada genetik algoritma yöntemi kullanılarak belirlenen değişkenler yardımıyla başarı puanlarına ilişkin tahminleme yapılması bakımından betimsel araştırma modelidir. Fraenkel, Wallen ve Hyun (2011)'e bir kuramın modelini test etme amacı altında yapılan çalışmalar betimsel çalışma niteliğindedir.

3.2. Evren ve Örneklem

Araştırma kapsamında, araştırmanın evrenini OECD tarafından düzenlenen PISA 2015 uygulamasına katılan 15 yaş grubu öğrencileri oluşturmaktadır. Uygulamaya 35'i OECD ülkesi olmak üzere toplam 72 ülkeden yaklaşık 540.000 öğrenci katılmıştır. PISA 2015 Türkiye uygulamasının evreni ise 15 yaş grubu öğrencilerden oluşan 1.324.089 öğrenciden oluşmakta ve uygulamaya katılabilecek öğrenci evreni ise 925.366 olarak belirlenmiştir. PISA 2015 Türkiye uygulamasına 61 ilden 187 okul ve 5895 öğrenci katılmıştır. Okullar belirlenirken tabakalı seçkisiz örnekleme yöntemi kullanılmış daha sonra bu okullardan seçilecek öğrenciler yine seçkisiz yöntemler belirlenmiştir. Okullar, istatistiki bölge birimleri sınıflamasına göre belirlenene 12 bölgeden eğitim türü, okul türü, okulların buldukları yer ve okulların idari biçimleri baz alınarak belirlenmiştir (MEB,2016).

PISA uygulaması öğrencilerin fen, matematik ve okuma becerileri temel alanlarındaki becerilerini değerlendirmek üzere yapılan bir çalışmadır. Araştırmada temel alanlarla birlikte çeşitli anketlerle öğrencilerin motivasyonları, okul ve aile

ortamları, kendileri hakkındaki düşünceleri ve öğrenme süreçlerini değerlendirdikleri değerlendirilmektedir.

Çalışma kapsamında belirlenen cinsiyet, sınıf ve baba eğitim durumu değişkenlerine ilişkin betimsel istatistik Tablo 1, Tablo 2 ve Tablo 3’ te verilmiştir.

Tablo 1. *Cinsiyet Değişkenine Göre Kategorik Dağılımlar*

Cinsiyet Değişkeni	Frekans	Yüzde (%)
Erkek	2632	49,2
Kız	2714	50,8

Tablo 2. *Sınıf Değişkenine Göre Kategorik Dağılımlar*

Sınıf Değişkeni	Frekans	Yüzde (%)
10. Sınıf Olma	3955	74
10. Sınıf Olmama	1391	26

Tablo 3. *Baba Eğitim Değişkenine Göre Kategorik Dağılımlar*

Baba Eğitim Değişkeni	Frekans	Yüzde (%)
Liseyi tamamlamış	1876	35,1
Liseyi tamamlamamış	3470	64,9

Tablo 1 incelendiğinde veri setini %50,8’ini kız öğrenciler, %49,2’sini ise erkek öğrencilerin oluşturduğu görülmektedir. Uygulamaya katılan öğrencilerin %74 ü onuncu sınıfta eğitim görüyorken, %26’lık kısım onuncu sınıf öğrencisi olmadığı Tablo 2’de verilmiştir. Onuncu sınıf olmayan öğrenciler kategorisinde yedinci, sekizinci, dokuzuncu, on birinci ve on ikinci sınıf öğrenciler bulunmaktadır. PISA uygulaması 15 yaş grubu öğrencilere yapıldığı için onuncu sınıf öğrencilerin sayısının fazla olması beklenen bir durumdur. Öğrencilerin baba eğitim durumları incelendiğinde yaklaşık olarak %70 öğrencinin babası lise eğitimini tamamlamadığı Tablo 3’te görülmektedir.

3.3. Veri Toplama Süreci ve Araçları

Bu arařtırmada üç temel alan olan matematik, fen ve okuma becerileri üzerine toplanan verilerden hem de anket sonuçlarından faydalanılacaktır. PISA uygulamasında döngüsel olarak üç temel alandan birine odaklanılmaktadır. 2015 yılında gerçekleştirilen PISA uygulamasında fen alanına odaklanılmıştır.

Arařtırmada kullanılan veriler OECD'nin resmi internet sayfası olan <http://www.oecd.org/pisa/data/2015database/> adresinden elde edilmiştir.

SPSS veri dosyası formatında yer alan öğrenci anketinden ülke numarası 792 olan 5895 öğrenciye ilişkin veriler analiz yapmak amacıyla veri kaynağı olarak düzenlenmiştir.

Arařtırmada veri toplama aracı olarak da uygulama esnasında yapılan öğrenci anketinden elde edilen değişkenler kullanılmıştır.

Öğrenci anketinde yer alan maddelerden uygun olanları seçilip düzenlenerek bağımsız değişkenler elde edildi. Bağımlı değişken olarak ise öğrencilerin matematik, fen ve okuma alanlarında aldıkları puanların ortalaması alındı.

Tablo 4. *Araştırmada Kullanılan Maddelere Ait Bilgiler*

Maddeler	Kodu
Cinsiyet	ST004Q01TA
Babanız, zorunlu eğitim ya da ortaöğretim döneminde en son hangi okulu bitirmiştir?	ST007Q01TA
Kaçıncı sınıftasınız?	ST001Q01TA
Ailem, okulumdaki faaliyetlerle ilgilenir.	ST123Q01NA
Ailem, eğitim çabalarımı ve başarılarımı destekler.	ST123Q02NA
Ailem, okulda zorluklarla karşılaştığımda bana destek olur.	ST123Q03NA
Ailem, kendime güvenmem konusunda beni cesaretlendirir.	ST123Q04NA
Okulda kendimi yabancı (ya da dışlanmış gibi) hissederim.	ST034Q01TA
Okulda diğer öğrencilerle kolaylıkla arkadaşlık kurarım.	ST034Q02TA
Kendimi okulun bir parçası olarak hissederim.	ST034Q03TA
Kendimi aykırı ve okula ait değilmiş gibi hissederim.	ST034Q04TA
Diğer öğrencilerin beni sevdiğini düşünüyorum.	ST034Q05TA
Okulda kendimi yalnız hissediyorum.	ST034Q06TA

Tablo 4 incelendiğinde genetik algoritma yöntemi ile öğrencilerin üç temel alan olan matematik, fen ve okumadaki başarılarını yordamak amacıyla kullanılacak olan bağımsız değişkenler ve bu değişkenlerin kodları görülmektedir.

3.4. Uygulama

Türkiye’de 5895 öğrenciye uygulanan PISA 2015 verileri içerisinde boş bırakılan veya cevabına ulaşamayan 549 örnek silinmiş ve 5346 sayıda örnek bulunan veri seti elde edilmiştir. Bu veri setinden yola çıkılarak Tablo 1 ‘ de verilen maddeler düzenlenerek 5 bağımsız ve 3 bağımlı değişken tanımlanmıştır. Bağımsız değişkenler Cinsiyet, Sınıf, Baba Eğitim, Okula İlgi ortalaması ve Ailenin ortalama ilgisidir. Bütün değişkenler düzenlenmiştir. Cinsiyet değişkeni; 1: Kız ve 0: Erkek, Sınıf değişkeni; 1: Onuncu sınıf olması ve 0: Onuncu sınıf olmama durumu, Baba Eğitim 1: Liseyi

tamamlamış ve 0: Liseyi tamamlamamış olarak düzenlenmiştir. Öğrencilerin okula ilgi ortalaması belirlenirken “Okulda kendimi yabancı (ya da dışlanmış gibi) hissederim.”, “Okulda diğer öğrencilerle kolaylıkla arkadaşlık kurarım.”, “Kendimi okulun bir parçası olarak hissederim.”, “Kendimi aykırı ve okula ait değilmiş gibi hissederim.”, “Diğer öğrencilerin beni sevdiğini düşünüyorum.” ve “Okulda kendimi yalnız hissediyorum.” Maddelerine verdikleri puanlar toplanıp madde sayısına bölünmüştür. Maddelere verilen cevaplar 1:Kesinlikle katılıyorum, 2:Katılıyorum, 3:Katılmıyorum ve 4:Kesinlikle katılmıyorum şeklindedir. Olumsuzluk ifadesine 4 puan verildiği için olumlu cümle içeren maddeler tersine çevrilip tekrar kodlanmıştır. Aile ilgi ortalaması “Ailem, okulumdaki faaliyetlerle ilgilenir.”, “Ailem, eğitim çabalarımı ve başarılarımı destekler.”, “Ailem, okulda zorluklarla karşılaştığımda bana destek olur.” ve “Ailem, kendime güvenmem konusunda beni cesaretlendirir.” maddeleri toplanarak hesaplanmıştır. Bu maddelere verilen cevaplar ise 1:Kesinlikle katılmıyorum, 2:Katılmıyorum, 3:Katılıyorum ve 4:Kesinlikle katılıyorum şeklindedir. Tüm maddeler olumlu cümleler içerdiği için herhangi bir tersine çevirme işlemi yapılmamıştır. Ayrıca her öğrenci için matematik, fen ve okuma alanlarının her birinde on farklı puan bulunmaktadır. Bu puanların da ortalaması alınarak her alan için tek bir puan elde edilmiştir.

5 bağımsız değişken ve 3 bağımlı değişken düzenlenerek yeniden tanımlanmıştır.

3.5. Verilerin Analizi

Araştırmanın amacı öğrencilerin PISA 2015 verilerini kullanarak başarı puanlarını tahmin edecek bir model oluşturmaktır. Öğrencilerin matematik, fen ve okuma alanından aldıkları puanlar tahmin edileceği için oluşturulacak model bir regresyon modelidir. Bağımlı ve bağımsız değişkenler arasındaki fonksiyonel ilişkiyi belirleyen istatistiksel yöntem regresyon analizidir (Erar, 2013). Regresyon modelinde bir bağımlı ve iki ya da daha fazla bağımsız (yordayıcı) değişken bulunuyorsa ve bağımsız değişkenler bağımlı değişkenin tahmin edilmesine sağlıyorsa bu analiz türü çoklu doğrusal regresyon analizidir (Büyüköztürk, 2015). Çoklu doğrusal regresyonda araştırmacının iki ana amacı bulunmaktadır. Birinci amaç bağımsız değişkenler aracılığıyla bağımlı değişkenin değerini tahmin etmek; ikinci amaç ise bağımsız

değişkenlerin hangisinin veya hangilerinin bağımlı değişken üzerinde daha fazla etkiye sahip olduğunu belirlemek ve aralarındaki ilişkiyi tanımlamaktır (Edwards, 1984; Şıklar, 2000; Orhunbilge, 2017).

Çoklu doğrusal regresyon modeli :

$$Y_i = \beta_0 + \beta_1 * X_{1i} + \beta_2 * X_{2i} + \dots + \beta_k * X_{ki} \quad i= 1,2,3,4,\dots,n$$

Eşitliği ile ifade edilmektedir. Modeli inceleyecek olursak, modeldeki Y_i tahmin etmeye çalışılan bağımlı değişkeni, $\beta_0, \beta_1, \beta_2 \dots \beta_k$ parametreleri bilinmeyen sabitler ve $X_{1i}, X_{2i}, \dots X_{ki}$ araştırmacı tarafından değeri bilinen bağımsız değişkenlerdir. Çoklu doğrusal regresyonun varsayımları ;

1. Model doğrusal kurulmuştur.
 2. Bağımlı değişken normal dağılım göstermelidir.
 3. Bağımsız değişkenler araştırmacı tarafından önceden belirlenmelidir.
 4. Bağımsız değişkenler arasındaki korelasyon sıfır veya sıfıra yakın olmalıdır.
- Yani çoklu bağlantı olmamalıdır (İşçioğlu, Kırık ve Deveci, 2019).

Çalışmamızda oluşturacağımız model;

$$Y_i = \beta_0 + \beta_1 * X_{1i} + \beta_2 * X_{2i} + \dots + \beta_k * X_{5i} \quad i= 1,2,3,4,\dots,n$$

Y_i = tahmin edilmeye çalışılan başarı puanı

X_1 = Cinsiyet değişkeni

X_2 = Okul ilgi ortalaması değişkeni

X_3 = Aile ilgi ortalaması değişkeni

X_4 = Baba eğitim değişkeni

X_5 = Sınıf değişkeni

Genetik algoritma uygulama adımları aşağıda verilmiştir.

1. Adım: Genetik algortmada kullanılacak parametrelerin belirlenmesi

Genetik algoritmada amaç uygunluk (amaç) fonksiyonunun değerini optimize etmektir (Öztürkler ve Altan, 2008). İlk olarak genetik algoritmada kullanılacak olan kromozomlara ait parametreler belirlenir. Bu parametreler kromozom sayısı (n), çaprazlama oranı (ço), mutasyon oranı (mo) ve iterasyon sayısı (its) şeklinde belirlenir.

2.Adım: Başlangıç popülasyonunun oluşturulması

Genetik algoritmada kromozomları oluşturan her bir birime gen denir. Bu çalışmada kullanacağımız genler parametre tahmini için bulmaya çalıştığımız β değerleridir. 5 bağımsız değişken ve bir sabit değişkenden oluşan modelimizde toplam altı tane gen bulunmaktadır. Bu genler $\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5$ parametreleridir. Araştırmamızda üç alt problemin her biri için bir başlangıç popülasyonu oluşturulmuştur. Başlangıç popülasyonları SPSS.15 programı kullanılarak yapılan çoklu doğrusal regresyonda EKK sonucunda elde edilen güven aralıkları içinden rastgele sayılar seçilerek oluşturulmuştur.

Tablo 5. *Matematik Puanına Göre EKK Tahminlerinin Güven Aralıkları*

Parametre tahminlerinin EKK güven aralıkları	β_0	β_1	β_2	β_3	β_4	β_5
Alt Sınır	312,69	-18,60	37,41	11,13	2,53	28,70
Üst Sınır	337,50	-11,29	45,72	16,36	9,35	36,33

Tablo 5 Matematik başarısına ait her bir bağımsız değişken için elde edilen güven aralıkları verilmiştir. Araştırma için 6 genden oluşan 1. kromozom üretilir.

Tablo 6. *Başlangıç Popülasyonu İçin Oluşturulan 1. Kromozom*

β_0	β_1	β_2	β_3	β_4	β_5
325,10	-14,94	41,56	13,75	5,94	32,51

Tablo 7. Fen Puanına Göre EKK Tahminlerinin Güven Aralıkları

Parametre tahminlerinin EKK güven aralıkları	β_0	β_1	β_2	β_3	β_4	β_5
Alt Sınır	313,86	-6,21	39,67	10,51	2,34	27,61
Üst Sınır	338,57	1,07	47,95	15,73	9,13	35,21

Tablo 7 Fen başarısına ait her bir bağımsız değişken için elde edilen güven aralıkları verilmiştir. Araştırma için 6 genden oluşan 1. kromozom üretilir.

Tablo 8. Başlangıç Popülasyonu İçin Oluşturulan 1. Kromozom

β_0	β_1	β_2	β_3	β_4	β_5
326,21	-2,57	43,81	13,12	5,73	31,41

Tablo 9. Okuma Puanına Göre EKK Tahminlerinin Güven Aralıkları

Parametre tahminlerinin EKK güven aralıkları	β_0	β_1	β_2	β_3	β_4	β_5
Alt Sınır	282,97	13,62	43,68	14,09	6,60	28,00
Üst Sınır	306,95	20,68	51,72	19,16	13,19	35,37

Tablo 9 Okuma başarısına ait her bir bağımsız değişken için elde edilen güven aralıkları verilmiştir. Araştırma için 6 genden oluşan 1. kromozom üretilir.

Tablo 10. Başlangıç Popülasyonu İçin Oluşturulan 1. Kromozom

β_0	β_1	β_2	β_3	β_4	β_5
294,96	17,15	47,70	16,62	9,90	31,69

Toplam “n” sayısı kadar kromozom oluşturulur ve bu kromozomlara başlangıç popülasyonu adı verilir.

3.Adım: Uygunluk değerinin hesaplanması

Bu adımda başlangıç popülasyonunda üretilen değerler kullanılarak her bir kromozom için amaç fonksiyonu hesaplanır. Hesaplanan bu değere uygunluk değeri denir. Böylece popülasyondaki birey sayısı kadar uygunluk değeri elde edilir.

Amaç fonksiyonumuz regresyon modelinde kullanılan EKK yönteminde tahmin edilmiş Y_i ile gözlenen Y_i arasındaki mutlak sapmaların en küçük olması istenmektedir. Amacımız hata sapma değerini minimum seviyeye indirmektir. Yani hatayı olabildiğince azaltmaktır.

$$\text{Min } \sum_{i=1}^n |Y_i - \sum_{j=1}^p X_{ij} * \beta_j|/n \quad , i=1, \dots, n; \quad j= 1, 2, \dots, p \text{ dir.}$$

n:gözlem sayısı

p: değişken sayısıdır.

4.Adım: Doğal seleksiyon işlemi

Bu aşamada genetik algoritmanın çalışma prensibi olan güçlü olan bireylerin hayatta kalmasına dayanan doğal seleksiyon işlemi yapılır. Kromozomların bir sonraki nesle aktarılmasının ölçüsü amaç fonksiyonu yardımıyla bulduğumuz uyum değerine bağlıdır. Amacımız uyum değerinin minimum seviyeye indirilmesi olduğu için uygunluk değeri yüksek olan bireyler bir sonraki nesle aktarılmaz ve elenir. Elenen kromozomlar yerine başlangıç popülasyonunda verilen alt ve üst değerler aralığında yeni bireyler eklenir.

5.Adım: Çaprazlama ve mutasyon işlemleri

Seçilen “ço” ve “mo” değerleri dikkate alınarak popülasyondaki kromozomlar üzerinde işlem yapmak ve yeni nesil için bireyler üretmek için yapılan işlemlerdir. Çaprazlamada bireylerdeki iyi özellikler birleştirilerek daha iyi çözümler üretmesi beklenir (Bolat, Erol ve İmrak 2004). Mutasyonda ise yeni çözümlerin önceki çözümlerden kopyalanmasını önlemek ve çözüme daha hızlı bir şekilde ulaşmak amacıyla kullanılır (Kurt ve Semetay, 2001).

6.Adım: Amaç fonksiyonunun minimum değeri

Yapılan işlemler belirlenen iterasyon sayısı kadar tekrar edilir ya da genetik algoritma en iyi sonuca ulaştığında çalışmayı durdurur. Amaç fonksiyonuyla elde edilen uygunluk değerlerinden en küçük değeri veren tahmin parametresi çözüm olarak kabul edilir.



4.BÖLÜM

BULGULAR

4.1. Birinci Alt Probleme İlişkin Bulgular

Birinci alt problemimiz “*Matematik Becerilerine ilişkin olarak genetik algoritma ile yapılan başarı yordamasının performansı nasıldır?*” şeklindedir.

Uygulama verisi PISA 2015 uygulamasına katılan 5346 öğrencinin verdiği cevaplardan oluşmaktadır. Amaç tahmin yapabilmek için bir model oluşturmaktır. Tahminleme için kullanılan regresyon modelinde parametreler EKK yöntemine alternatif olacak genetik algoritma yöntemiyle yapılmıştır. Araştırmada bağımlı değişken (Y) öğrencinin matematik puanı olarak belirlenmiştir. Öğrencinin başarısını etkilediği düşünülen bağımsız değişkenler ise cinsiyet (X_1), okula ilgi ortalaması (X_2), aile ilgisi ortalaması (X_3), baba eğitim durumu (X_4) ve sınıf (X_5) olarak ifade edilmiştir.

Uygulama verisinin öncelikle regresyon modeline uygunluğu test edilmeye çalışılmıştır. Çözümleme esnasında SPSS 15.0 paket programı ve Matlab programından yararlanılmıştır.

SPSS 15.0 paket programı kullanılarak veri setine EKK regresyon yöntemi uygulanmış ve veri setinin çoklu doğrusal regresyon varsayımlarını karşılayıp karşılamadığı belirlenmiştir.

1.varsayım bağımlı değişken olan matematik puanlarının normal dağılıp dağılmadığıdır. Yapılan normallik testi sonucunda matematik puanına ilişkin çarpıklık değeri 0,229 iken basıklık değeri -0,267 olduğu görülmüştür. Büyüköztürk (2015) bu durumu çarpıklık ve basıklık katsayılarının -1 ile +1 sınırları arasında kalan dağılımların normal dağılımdan önemli bir sapma göstermediği şeklinde yorumlamıştır. Analizlerde temel amacın puanların normalden aşırı sapma göstermemesidir. Bu bilgi doğrultusunda veri setinden elde edilen matematik puanlarının normal dağıldığı söylenebilir.

2.varsayım bağımsız değişkenler arasındaki korelasyon sıfır ya da sıfıra yakın olmalıdır. Yani bağımsız değişkenler arasında yüksek ilişkiden kaçınılmalıdır.

Tablo 11. Korelasyon Matrisi

Bağımsız Değişkenler	X₁	X₂	X₃	X₄	X₅
X ₁	1,00	-0,09	-0,05	0,04	0,08
X ₂	-0,09	1,00	-0,02	-0,32	-0,04
X ₃	-0,05	-0,02	1,00	-0,09	-0,06
X ₄	0,04	-0,32	-0,01	1,00	-0,09
X ₅	-0,08	-0,04	-0,06	-0,09	1,00

Tablo 11’de matris olarak gösterilen bağımsız değişkenler arasındaki korelasyonlar incelendiğinde hiçbir değişken arasında güçlü bir ilişki olmadığı anlaşılmaktadır. Bu sonuç ise bağımsız değişkenler arasında çoklu bağlantı sorununun olmadığı ortaya çıkmaktadır.

3.varsayım ise oluşturulacak regresyon modelinden varyansların homojen dağılmasıdır.

Tablo 12. Varyans Homojenliği Testi

Model	Kareler toplamı	Sd	Kareler ortalaması	F	p
Regresyon	4281935	5	856387	188,35	0,000
Artık	24278900	5340	4546,61		
Toplam	28560835	5345			

Tablo 12’de ilgili çoklu regresyon modelindeki değişkenler arasındaki ilişkiyi belirlemek için F-testi sonuçları görülmektedir. F-test istatistiği sonucunda $F=188,35$ ve $p= 0,000$ bulunmuştur. Tahmin edilen regresyon denklemi değişkenler arasındaki ilişkiyi açıklıyor yorumu yapılabilir.

Tahminleme için alternatif yöntem olarak geliştirilen genetik algoritma yönteminde kromozom sayısı (n) 50, çaprazlama oranı (ço) 0,5 mutasyon oranı (mo) 0,02 ve maksimum iterasyon sayısı (its) 600 olarak belirlenmiştir. Jong (1975)’ e göre

popülasyon büyüklüğü 50-100 arasında olmalıdır. Grefenstte (1986) mutasyon oranı 0,01 olmalıdır. Çaprazlama oranı ise 0,5 veya daha fazla olmalıdır (Jong,1975). Yapılan literatür taramasına göre mutasyon oranının düşük, çaprazlama oranının ise yüksek seçilmesi gerektiği görülmüştür.

Tablo 13. *Matematik Puanına İlişkin Oluşturulan Başlangıç Popülasyonu*

Kromozom Numarası	1.Gen β_1	2.Gen β_2	3.Gen β_3	4.Gen β_4	5.Gen β_5	6.Gen β_0
1	-15,7263404	6,917352979	13,32067892	33,30900818	43,17205852	321,1114024
2	-12,0093709	6,082949762	13,73612454	31,63027475	44,32946768	319,1462237
3	-12,7704350	2,706416818	12,28831558	36,46503844	42,350763	334,050563
4	-13,3314930	6,833245852	13,03925201	30,98717947	38,84415766	331,173227
5	-15,7263404	4,291924618	13,42607314	33,30900818	42,59173331	317,4790715
6	-13,9591424	3,01764483	13,32067892	35,75677076	43,07603905	326,1929559
7	-14,4834539	5,371736616	15,07169799	35,87175529	42,85657131	313,7975283
8	-17,5436257	3,628684028	11,2459055	30,54301028	42,85657131	313,4782451
9	-14,7816648	5,776398691	11,0614504	32,89304932	40,44733239	312,6514617
10	-16,2000168	9,841130716	15,07169799	35,87175529	38,59026963	313,7975283
11	-13,3314930	8,17245986	15,59416013	33,1588021	45,59032486	324,2157963
12	-11,0723157	4,505289279	11,37195366	35,06111685	40,29801041	316,9390759
13	-13,1828300	5,155200184	13,87022761	34,55926452	40,55632417	326,151188
14	-12,8414376	3,736078481	13,40855832	32,2899692	40,79418165	337,8183464
15	-17,5436257	6,082949762	11,2459055	31,63027475	44,32946768	337,2476269
16	-15,0565446	6,917352979	11,09111116	33,30900818	40,00866757	330,5660416
17	-15,7263404	6,917352979	13,32067892	35,34466332	45,38155535	319,7489018
18	-12,8414376	3,736078481	11,76651934	32,2899692	39,60126723	337,8183464
19	-12,8414376	3,736078481	13,40855832	34,98308741	45,3319286	337,8183464
20	-12,7704350	2,706416818	12,28831558	36,46503844	42,350763	334,050563
21	-15,7446874	4,223697178	16,99689138	29,38274565	39,22273012	320,3020189
22	-14,4567270	7,285245167	16,49805748	33,68465328	41,09101245	316,6157442
23	-13,9153593	8,17245986	13,32067892	33,1588021	43,06284282	321,1114024

24	-18,6395706	3,156159203	16,77463982	34,98308741	45,3319286	322,8370962
25	-12,8414376	5,734533667	13,40855832	32,2899692	40,79418165	331,4213932
26	-11,9773734	8,1132186	11,93940797	31,81608282	40,3880122	327,1364861
27	-11,3088673	4,693216649	11,18172041	35,90591591	42,61125046	336,7576253
28	-11,2842170	5,371736616	15,07169799	34,98308741	43,16106971	313,7975283
29	-13,9153593	8,17245986	13,83192554	28,81051871	41,70500776	325,2223307
30	-16,5882207	2,547996316	11,75377256	30,51531571	45,4915811	319,277494
31	-15,7263404	8,349869857	13,32067892	33,30900818	43,17205852	321,1114024
32	-17,3833742	8,1132186	15,58110325	30,68548529	42,61125046	325,2829187
33	-15,7263404	6,917352979	15,28567083	36,77875592	43,17205852	321,1114024
34	-11,2842170	2,706416818	16,11307146	36,46503844	42,350763	322,8370962
35	-12,5185129	5,003704137	12,98101041	32,6109128	45,12015428	333,2851012
36	-12,0093709	6,082949762	13,73612454	31,63027475	44,32946768	319,1462237
37	-15,7263404	6,917352979	13,42607314	35,56171537	42,59173331	317,4790715
38	-11,9613038	5,562546066	14,65379791	32,89304932	37,8424478	334,4338383
39	-15,7263404	6,917352979	16,49805748	31,86090922	41,09101245	321,1114024
40	-16,2000168	6,082949762	16,76250244	28,82583483	38,59026963	319,1462237
41	-13,9591424	8,483598434	11,0614504	31,67027768	40,44733239	312,6514617
42	-15,6260005	2,590327058	16,23796388	34,26242058	39,30295534	317,6082866
43	-15,0565446	7,530295221	11,74932722	30,78872281	40,00866757	330,5660416
44	-17,3833742	8,1132186	11,93940797	30,68548529	40,3880122	327,5844605
45	-18,3214662	4,10652827	14,8685817	29,78398575	37,29806978	333,1255215
46	-11,3088673	8,507566137	15,58110325	35,90591591	42,61125046	325,2829187
47	-14,4567270	3,761900864	16,49805748	31,86090922	41,09101245	316,6157442
48	-12,0093709	6,082949762	13,73612454	31,63027475	44,32946768	319,1462237
49	-11,6619399	4,291924618	13,42607314	35,56171537	42,59173331	317,4790715
50	-15,7263404	6,917352979	13,32067892	33,30900818	43,17205852	321,1114024

Tablo 13'te matematik puanı tahminlemesi için oluşturulacak modele ait başlangıç popülasyonu sunulmuştur. Başlangıç popülasyonu 50 bireyden oluşmuş ve

her bireyde 6 gen bulunmaktadır. Genler tahminleme için çözüm aradığımız parametrelerdir.

Her kromozom bir uygunluk değerine sahiptir. Uygunluk değeri amaç fonksiyonu yardımıyla bulunmaktadır. Araştırmada kullanılan amaç fonksiyonu gözlenen matematik puanı ile tahminlenen matematik puanı arasındaki mutlak sapmalardır.

Tablo 14. *Başlangıç Popülasyonuna Ait Uygunluk Değerleri (Matematik)*

Kromozom Numarası	Amaç Fonksiyonu
1	54,1244336
2	54,13486221
3	54,1503744
4	54,57578977
5	54,50647486
6	54,10994243
7	54,12551198
8	57,20063987
9	56,05027893
10	54,32431283
11	56,16243174
12	55,24509563
13	54,20730661
14	54,72627331
15	54,58246896
16	54,18598009
17	54,1365958
18	54,25673629
19	55,28660644
20	54,1503744
21	54,2319757

22	54,49901123
23	54,33289225
24	54,31568397
25	54,59502206
26	54,39732332
27	54,47988459
28	54,12285336
29	54,70004691
30	55,62860829
31	54,30789979
32	55,38362363
33	54,62834165
34	54,23966317
35	54,80614483
36	54,13486221
37	54,09352296
38	55,31096977
39	54,7511174
40	54,30498643
41	54,86972593
42	54,32570272
43	54,34168741
44	54,25800853
45	54,37248666
46	56,5597776
47	54,16903447
48	54,13486221
49	54,29474688
50	54,1244336

50 tane kromozomdan elde edilen amaç fonksiyonu değerleri Tablo 14'te gösterilmiştir. Araştırmadaki amacımız hatayı en aza indirmek olduğu için amaç fonksiyonunu yani uygunluk değerleri yüksek olan bireyler bir sonraki nesle aktarılmaz. Bu araştırmada $\rho=0,5$, $m=0,02$ olarak seçilmiş ve bir sonraki nesle seçilme yöntemi olarak ise rulet çarkı yöntemi kullanılmıştır. Rulet çarkı seçim metotları arasında en çok kullanılan metotlardan biridir (Yeo ve Agyei, 1998).

Matematik puanı için oluşturulacak regresyon modeli Matlab programı yardımıyla belirlenen genetik işlemler ile 60 iterasyon sonucunda elde edilmiştir. 60. iterasyonda genetik algoritma en iyi çözüme ulaşmış ve iterasyonu durdurmuştur.

Tablo 15. *Genetik Algoritma Yöntemi İle Regresyonda Model Parametre Tahminleri*

β_0	β_1	β_2	β_3	β_4	β_5
321,1114	-15,8239	5,3717	13,7361	36,6798	43,3378

Son iterasyonda elde edilen parametre tahminleri Tablo 15'te görülmektedir. Bu parametreler için elde edilen uygunluk değeri 54,0767 olarak hesaplanmıştır.

Matematik başarısı için oluşturulan regresyon modeli ;

$Y_{mat_i} = 321,1114 - 15,8239 X_{1i} + 5,3717 X_{2i} + 13,7361 X_{3i} + 36,6798 X_{4i} + 43,3378 X_{5i}$ şeklindedir.

Hata yüzdesi mutlak hatanın gerçek değere oranı olarak bulunur. Matematik puanı için veri setinin ortalaması 419,86'dır. Genetik algoritma yöntemiyle bulunan minimum mutlak sapma ise 54,0767 olarak ölçülmüştür. Buna göre

$$\text{Hata yüzdesi} = \frac{54,0767}{419,86} = \%13 \text{ olarak hesaplanmıştır.}$$

4.2. İkinci Alt Probleme İlişkin Bulgular

İkinci alt problemimiz “*Fen Bilimleri Yeterliğine ilişkin olarak genetik algoritma ile yapılan başarı yordamasının performansı nasıldır?*” şeklindedir.

Uygulama verisi PISA 2015 uygulamasına katılan 5346 öğrencinin verdiği cevaplardan oluşmaktadır. Amaç tahmin yapabilmek için bir model oluşturmaktır. Tahminleme için kullanılan regresyon modelinde parametreler EKK yöntemine alternatif olacak genetik algoritma yöntemiyle yapılmıştır. Araştırmada bağımlı değişken (Y) öğrencinin fen puanı olarak belirlenmiştir. Öğrencinin başarısını etkilediği düşünülen bağımsız değişkenler ise cinsiyet (X_1), okula ilgi ortalaması (X_2), aile ilgisi ortalaması (X_3), baba eğitim durumu (X_4) ve sınıf (X_5) olarak ifade edilmiştir.

Uygulama verisinin öncelikle regresyon modeline uygunluğu test edilmeye çalışılmıştır. Çözümleme esnasında SPSS 15.0 paket programı ve Matlab programından yararlanılmıştır.

SPSS 15.0 paket programı kullanılarak veri setine EKK regresyon yöntemi uygulanmış ve veri setinin çoklu doğrusal regresyon varsayımlarını karşılayıp karşılamadığı belirlenmiştir.

1.varsayım bağımlı değişken olan fen puanlarının normal dağılıp dağılmadığıdır. Yapılan normallik testi sonucunda fen puanına ilişkin çarpıklık değeri 0,25 iken basıklık değeri -0,44 olduğu görülmüştür. Çarpıklık ve basıklık katsayıları -1 ile +1 arasında bulunduğu için fen puanlarının normal dağıldığı söylenebilir. Analizlerdeki temel amaç sapmaların normalden aşırı fazla olmamasıdır.

2.varsayım bağımsız değişkenler arasındaki korelasyon sıfır ya da sıfıra yakın olmalıdır. Yani bağımsız değişkenler arasında yüksek ilişkiden kaçınılmalıdır.

Tablo 16. *Korelasyon Matrisi*

Bağımsız Değişkenler	X₁	X₂	X₃	X₄	X₅
X ₁	1,00	-0,09	-0,05	0,04	0,08
X ₂	-0,09	1,00	-0,02	-0,32	-0,04
X ₃	-0,05	-0,02	1,00	-0,09	-0,06
X ₄	0,04	-0,32	-0,01	1,00	-0,09
X ₅	-0,08	-0,04	-0,06	-0,09	1,00

Tablo 16’da matris olarak gösterilen bağımsız değişkenler arasındaki korelasyonlar incelendiğinde hiçbir değişken arasında güçlü bir ilişki olmadığı anlaşılmaktadır. Bu sonuç ise bağımsız değişkenler arasında çoklu bağlantı sorununun olmadığı ortaya çıkmaktadır.

3.varsayım ise oluşturulacak regresyon modelinden varyansların homojen dağılmasıdır.

Tablo 17. *Varyans Homojenliği Testi*

Model	Kareler toplamı	Sd	Kareler ortalaması	F	p
Regresyon	4178322	5	835664	185,29	0,000
Artık	24083586	5340	4510,04		
Toplam	28261908	5345			

Tablo 17’de ilgili çoklu regresyon modelindeki değişkenler arasındaki ilişkiyi belirlemek için F-testi sonuçları görülmektedir. F-test istatistiği sonucunda $F=185,29$ ve $p=0,000$ bulunmuştur. Tahmin edilen regresyon denklemi değişkenler arasındaki ilişkiyi açıklıyor yorumu yapılabilir.

Tablo 18. *Fen Puanına İlişkin Oluşturulan Başlangıç Popülasyonu*

Kromozom Numarası	1.Gen β_1	2.Gen β_2	3.Gen β_3	4.Gen β_4	5.Gen β_5	6.Gen β_0
1	0,47664333	4,47551376	13,24729383	29,27355473	44,43492603	322,3131046
2	-2,24357951	9,266268401	12,27190566	30,47388	44,73617727	315,8307791
3	-5,73161695	9,513824465	12,4114893	29,75207365	40,86064248	335,6136611
4	0,476643333	3,049116484	14,72952891	34,69156756	46,91423439	322,7085863
5	-0,41254921	4,47551376	14,58742289	34,78698467	44,43492603	322,3131046
6	-0,41254921	6,752531361	14,58742289	27	44,40097675	313
7	-2,62948839	7,32169917	15,0933893	34,20452343	44,91467828	324,0316435
8	-2,62948833	3,049116484	15,0933893	29,9017173	46,91423439	322,7085863
9	-5,73161695	6,606436041	15,26879411	29,75207365	40,30714369	327,2262709
10	0,476643333	3,049116484	13,24729383	29,27355473	44,43492603	322,3131046
11	-2,48938114	9,266268401	15,78636047	30,34264098	44,73617727	316,9715422
12	-6,35742986	6,757910551	13,04153783	32,55862505	45,12313925	338,2945399
13	-2,24357951	2,381133963	15,96170979	30,08124051	43,74658816	325,6468798
14	-2,24357951	9,266268401	12,27190566	27,50628062	45,62348981	328,5049733

15	-6,67777215	3,049116484	14,72952891	34,69156756	46,91423439	322,7085863
16	-6,67777215	6,886214423	14,76101126	34,69156756	46,91423439	322,7085863
17	-5,12412993	6,886214423	14,76101126	35,02663862	44,98206894	336,2802922
18	-3,57082029	9,513824465	11,90948056	34,56432803	40,86064248	336,7855882
19	-6,35742986	8,917012024	12,8700835	28,86322097	47,1966226	338,2945399
20	-2,24357951	8,19967443	12,27190566	30,47388	44,73617727	315,8307791
21	-0,30608058	4,407726566	10,72343187	34,91499448	43,89388345	316,4489853
22	0,476643333	3,049116484	13,24729383	29,27355473	44,43492603	322,3131046
23	-6,27582595	3,540889672	14,43885287	29,81873963	46,20738138	316,9555866
24	-6,67777218	9,270320145	10,76298622	27,17610662	41,03095417	332,5602532
25	-6,67777215	3,049116484	14,72952891	34,54423889	46,91423439	322,7085863
26	-5,73161695	3,080720982	12,4114893	29,75207365	39,65739033	319,9125452
27	-6,67777215	4,009088311	13,89975187	27,23555034	46,91423439	338,0897156
28	-3,00485708	8,829955608	13,96655035	34,54423889	39,85032671	316,9555866
29	0,476643333	4,857789391	13,24729383	29,27355473	41,46810922	322,3131046
30	-0,41254921	6,752531361	12,35038439	34,78698467	40,86064248	335,6136611
31	0,476643333	3,049116484	14,72952891	29,27355473	46,91423439	322,3131046
32	-5,26834342	5,491819079	13,58167962	29,51507838	44,90752673	321,1839684
33	-0,71097717	9,532096329	11,92625139	32,39887463	47,47663946	318,7506364
34	-2,62948833	7,32169917	15,0933893	29,9017173	43,89388345	324,0316435
35	-6,67777215	3,049116484	14,72952891	34,69156756	46,91423439	322,7085863
36	-6,67777215	4,407726566	13,91039794	34,20452343	43,89388345	316,4489853
37	-3,00485708	8,126829686	13,96655035	27,42209162	43,86316229	322,9777904
38	-1,68654039	8,19967443	11,87244238	27,78960496	41,5152831	327,9438612
39	-6,67777215	3,049116484	14,72952891	33,05380964	46,91423439	322,7085863
40	-2,82154769	2,896956692	12,4114893	29,75207365	43,86316229	314,8164679
41	-3,27302302	3,967224135	10,02818339	35,6853915	41,19013032	316,7499829
42	-0,71097717	7,759774076	13,82928581	33,44770687	44,67802777	334,9571172
43	-4,55765158	9,27288832	15,38517548	31,27441209	43,46073805	332,9655299
44	-2,62948833	7,32169917	15,0933893	29,9017173	44,91467828	334,8514618
45	0,476643333	4,47551376	13,24729383	29,27355473	44,43492603	322,3131046
46	-0,30608058	4,407726566	13,91039794	34,20452343	43,89388345	316,4489853
47	0,290576504	3,233445598	11,57409759	27,49575087	46,1466065	327,5583156
48	-4,07436288	8,126829686	11,30061893	32,54448562	43,86316229	316,0037931
49	-5,69656394	4,009088311	13,89975187	27,23555034	39,24468174	338,0897156
50						

Tablo 18’de fen puanı tahminlemesi için oluşturulacak modele ait başlangıç popülasyonu sunulmuştur. Başlangıç popülasyonu 50 bireyden oluşmuş ve her bireyde 6 gen bulunmaktadır. Genler tahminleme için çözüm aradığımız parametrelerdir.

Her kromozom bir uygunluk değerine sahiptir. Uygunluk değeri amaç fonksiyonu yardımıyla bulunmaktadır. Araştırmada kullanılan amaç fonksiyonu gözlenen fen puanı ile tahminlenen fen puanı arasındaki mutlak sapmalardır.

Tablo 19. *Başlangıç Popülasyonuna Ait Uygunluk Değerleri (Fen)*

Kromozom Numarası	Amaç Fonksiyonu
1	54,11340697
2	54,1578686
3	54,1590452
4	55,94417607
5	54,19202236
6	54,23932764
7	54,17704302
8	55,30472788
9	54,14653202
10	54,81683101
11	54,28878971
12	55,20230353
13	56,0411986
14	54,26363498
15	55,25590404
16	54,11340697
17	54,72933519
18	57,19361891
19	56,32605288
20	57,32341
21	54,14493359
22	55,28292243
23	54,28878971
24	54,40766796
25	54,62173844
26	54,11350951
27	55,76820945
28	55,1928364
29	54,28808277
30	54,22116512
31	55,29529825
32	54,15471366
33	54,14327926
34	54,41164665
35	55,03721059
36	54,11340697

37	54,43478395
38	54,64435585
39	54,4843581
40	54,11913114
41	55,89350769
42	56,61708835
43	57,07478955
44	58,04825594
45	57,31445814
46	54,1578686
47	54,20593651
48	54,28116888
49	54,32440003
50	54,70359429

50 tane kromozomdan elde edilen amaç fonksiyonu değerleri Tablo 19’da gösterilmiştir. Araştırmadaki amacımız hatayı en aza indirmek olduğu için amaç fonksiyonunu yani uygunluk değerleri yüksek olan bireyler bir sonraki nesle aktarılmaz. Bu araştırmada $\rho=0,5$, $m_0=0,02$ olarak seçilmiş ve bir sonraki nesle seçilme yöntemi olarak ise rulet çarkı yöntemi kullanılmıştır.

Fen puanı için oluşturulacak regresyon modeli Matlab programı yardımıyla belirlenen genetik işlemler ile 99 iterasyon sonucunda elde edilmiştir. 99. iterasyonda genetik algoritma en iyi çözüme ulaşmış ve iterasyonu durdurmuştur.

Tablo 20. *Genetik Algoritma Yöntemi İle Regresyonda Model Parametre Tahminleri*

β_0	β_1	β_2	β_3	β_4	β_5
322,3131	-3,7945	3,1001	14,3183	34,7185	46,9142

Son iterasyonda elde edilen parametre tahminleri Tablo 20’de görülmektedir. Bu parametreler için elde edilen uygunluk değeri 54,0909 olarak hesaplanmıştır.

Fen başarısı için oluşturulan regresyon modeli ;

$Y_{fen_i} = 322,3131 - 3,7945X_{1i} + 3,1001X_{2i} + 14,3183X_{3i} + 34,7185X_{4i} + 46,9142X_{5i}$ şeklindedir.

Hata yüzdesi mutlak hatanın gerçek değere oranı olarak bulunur. Fen puanı için veri setinin ortalaması 425,96. Genetik algoritma yöntemiyle bulunan minimum mutlak sapma ise 54,0909 olarak ölçülmüştür. Buna göre

$$\text{Hata yüzdesi} = 54,0909/425,96 = \%13 \text{ olarak hesaplanmıştır.}$$

4.3. Üçüncü Alt Probleme İlişkin Bulgular

Üçüncü alt problemimiz “*Okuma Yeterliğine ilişkin olarak genetik algoritma ile yapılan başarı yordamasının performansı nasıldır?*” şeklindedir.

Uygulama verisi PISA 2015 uygulamasına katılan 5346 öğrencinin verdiği cevaplardan oluşmaktadır. Amaç tahmin yapabilmek için bir model oluşturmaktır. Tahminleme için kullanılan regresyon modelinde parametreler EKK yöntemine alternatif olacak genetik algoritma yöntemiyle yapılmıştır. Araştırmada bağımlı değişken (Y) öğrencinin okuma puanı olarak belirlenmiştir. Öğrencinin başarısını etkilediği düşünülen bağımsız değişkenler ise cinsiyet (X_1), okula ilgi ortalaması (X_2), aile ilgisi ortalaması (X_3), baba eğitim durumu (X_4) ve sınıf (X_5) olarak ifade edilmiştir.

SPSS 15.0 paket programı kullanılarak veri setine EKK regresyon yöntemi uygulanmış ve veri setinin çoklu doğrusal regresyon varsayımlarını karşılayıp karşılamadığı belirlenmiştir.

1.varsayım bağımlı değişken olan okuma puanlarının normal dağılıp dağılmadığıdır. Yapılan normallik testi sonucunda okuma puanına ilişkin çarpıklık değeri -0,26 iken basıklık değeri -0,22 olduğu görülmüştür. Çarpıklık ve basıklık katsayıları -1 ile +1 arasında bulunduğu için okuma puanlarının normal dağıldığı söylenebilir. Analizlerdeki temel amaç sapmaların normalden aşırı fazla olmamasıdır.

2.varsayım bağımsız değişkenler arasındaki korelasyon sıfır ya da sıfıra yakın olmalıdır. Yani bağımsız değişkenler arasında yüksek ilişkiden kaçınılmalıdır.

Tablo 21. Korelasyon Matrisi

Bağımsız Değişkenler	X₁	X₂	X₃	X₄	X₅
X ₁	1,00	-0,09	-0,05	0,04	0,08
X ₂	-0,09	1,00	-0,02	-0,32	-0,04
X ₃	-0,05	-0,02	1,00	-0,09	-0,06
X ₄	0,04	-0,32	-0,01	1,00	-0,09
X ₅	-0,08	-0,04	-0,06	-0,09	1,00

Tablo 21’de matris olarak gösterilen bağımsız değişkenler arasındaki korelasyonlar incelendiğinde hiçbir değişken arasında güçlü bir ilişki olmadığı anlaşılmaktadır. Bu sonuç ise bağımsız değişkenler arasında çoklu bağlantı sorununun olmadığı ortaya çıkmaktadır.

3. varsayım ise oluşturulacak regresyon modelinden varyansların homojen dağılmasıdır.

Tablo 22. Varyans Homojenliği Testi

Model	Kareler toplamı	Sd	Kareler ortalaması	F	p
Regresyon	5741727	5	1148345	270,37	0,000
Artık	22680516	5340	4247,29		
Toplam	28422243	5345			

Tablo 22’de ilgili çoklu regresyon modelindeki değişkenler arasındaki ilişkiyi belirlemek için F-testi sonuçları görülmektedir. F-test istatistiği sonucunda $F=270,3$ ve $p= 0,000$ bulunmuştur. Tahmin edilen regresyon denklemi değişkenler arasındaki ilişkiyi açıklıyor yorumu yapılabilir.

Tablo 23. *Okuma Puanına İlişkin Oluşturulan Başlangıç Popülasyonu*

Kromozom Numarası	1.Gen β_1	2.Gen β_2	3.Gen β_3	4.Gen β_4	5.Gen β_5	6.Gen β_0
1	15,22798575	9,986912416	14,99389238	36,50308471	50,30422412	295,670143
2	15,53679584	12,05760183	14,81641135	34,1809789	49,40094202	294,5005609
3	18,43762141	7,297458466	15,43971515	35,35864937	48,20672555	306,4670162
4	18,67491865	8,008670864	15,43949692	30,75714525	51,35968726	301,4762931
5	15,53679584	12,05760183	16,4108482	34,1809789	49,40094202	294,5005609
6	14,3694935	10,92835741	17,47822752	33,84203927	44,05675886	304,618056
7	16,91811517	10,04765641	14,50292827	32,38112469	45,36234011	291,1859162
8	16,13781616	8,489720336	19,21216823	28,38721421	47,12963945	294,3543484
9	15,21538388	12,64662902	16,58848296	35,39074636	51,31041678	284,6690465
10	19,79303445	8,034257432	14,81641135	28,53801589	45,04897868	298,9931988
11	20,71910828	10,68214201	17,92447459	31,03947469	45,36234011	286,5730617
12	20,65733559	8,040760924	16,70324959	29,0008248	48,60227577	284,467807
13	14,30089388	8,489720336	14,29792658	33,79886317	50,59877913	301,9045968
14	15,53679584	10,54257313	14,81641135	34,1809789	49,40094202	297,2466662
15	18,43762141	9,786310791	15,55922242	34,58550147	48,20672555	306,4670162
16	13,27556864	8,489720336	14,29792658	31,31636137	47,12963945	294,5117906
17	15,22798575	7,94819975	14,99389238	31,17842714	50,30422412	295,670143
18	13,77705425	10,39778887	15,09108217	30,66688725	48,35406467	286,9452457
19	15,22798575	8,040760924	14,99389238	29,0008248	50,30422412	291,8878804
20	16,88300519	10,04765641	19,21216823	28,38721421	48,28340234	291,1859162
21	18,42988124	8,008670864	14,02780534	29,52091027	51,14392872	300,0306875
22	20,76474225	12,01013647	18,48890956	31,32322103	47,95140709	304,2730626
23	16,05246766	6,606834317	15,43971515	35,35864937	48,20672555	297,4416597
24	18,17050408	8,489720336	14,29792658	33,79886317	43,2629825	294,5117906
25	16,13781616	7,572762003	19,21216823	28,38721421	50,30422412	295,670143
26	14,13509071	13,12722602	19,48002417	30,17522157	45,71121697	282,9434717
27	18,94505974	9,986912416	14,99389238	36,50308471	50,30422412	298,3439337
28	16,13781616	9,986912416	14,99389238	28,38721421	44,66334688	294,3543484
29	14,26090465	7,790495516	17,07949724	36,10048462	45,36234011	297,6015022
30	15,53679584	12,05760183	14,81641135	34,1809789	49,40094202	294,5005609
31	20,65733559	8,040760924	15,43971515	35,35864937	48,60227577	306,4670162
32	19,06192104	12,64662902	18,64946279	35,39074636	45,32258226	284,6690465
33	15,53679584	12,05760183	14,81641135	34,1809789	49,40094202	294,5005609
34	18,24592559	8,060066033	14,46905317	32,92308003	45,03329603	299,4686458
35	15,21538388	7,790495516	16,58848296	36,10048462	50,8834853	297,6015022
36	20,65733559	8,040760924	14,81641135	34,1809789	48,60227577	291,8878804
37	20,60177639	12,02983275	15,5828175	29,6516004	44,99572061	293,9980535
38	15,53679584	8,286712151	14,87323388	34,1809789	48,4255878	300,6018565
39	20,66005468	9,986912416	15,57782771	32,40327375	46,15654393	287,7898597
40	16,88300519	10,04765641	17,07949724	32,38112469	45,36234011	286,5730617
41	15,53679584	10,22826508	19,41629666	31,40748444	51,66779685	294,5005609

42	19,33765864	7,108995543	17,23005461	29,18775963	50,59877913	301,9045968
43	16,91811517	13,47208547	17,07949724	32,38112469	45,36234011	286,5730617
44	17,37505215	13,67795167	17,61189165	32,41777683	47,7954303	295,0283958
45	18,42988124	7,94819975	14,02780534	31,17842714	45,79992058	300,0306875
46	16,91811517	13,47208547	17,07949724	32,38112469	45,36234011	286,5730617
47	15,21538388	12,64662902	16,58848296	32,05831336	45,32258226	290,3540763
48	16,13781616	7,572762003	19,21216823	28,38721421	44,66334688	294,3543484
49	20,67593941	7,194352044	19,9768083	36,47845532	44,75287861	304,4297838
50	15,22798575	9,986912416	14,99389238	36,50308471	50,30422412	295,670143

Tablo 23'te okuma puanı tahminlemesi için oluşturulacak modele ait başlangıç popülasyonu sunulmuştur. Başlangıç popülasyonu 50 bireyden oluşmuş ve her bireyde 6 gen bulunmaktadır. Genler tahminleme için çözüm aradığımız parametrelerdir.

Her kromozom bir uygunluk değerine sahiptir. Uygunluk değeri amaç fonksiyonu yardımıyla bulunmaktadır. Araştırmada kullanılan amaç fonksiyonu gözlenen fen puanı ile tahminlenen okuma puanı arasındaki mutlak sapmalardır.

Tablo 24. *Başlangıç Popülasyonuna Ait Uygunluk Değerleri (Okuma)*

Kromozom Numarası	Amaç Fonksiyonu
1	52,512598
2	52,51380116
3	52,52585833
4	52,49392766
5	52,66719644
6	53,3050283
7	53,29140614
8	52,55083367
9	52,52295592
10	52,93677444
11	52,57734543
12	53,53732445
13	52,52505734
14	52,49449368

15	53,00600763
16	53,6321567
17	52,90872618
18	53,46110485
19	53,45684444
20	52,62841035
21	52,68952274
22	55,55699027
23	52,9276625
24	53,59994812
25	52,57728591
26	52,59792322
27	52,54408499
28	53,03009702
29	52,55323314
30	52,51380116
31	52,73311969
32	52,64284376
33	52,51380116
34	52,89893901
35	52,47368521
36	53,12492563
37	52,57376099
38	52,55028948
39	53,01726047
40	52,86098386
41	53,2452231
42	52,60142337
43	52,55182804
44	53,58844238
45	53,01019643
46	52,55182804

47	52,53231834
48	52,56589213
49	53,79117606
50	52,512598
	52,512598

50 tane kromozomdan elde edilen amaç fonksiyonu değerleri Tablo 24'te gösterilmiştir. Araştırmadaki amacımız hatayı en aza indirmek olduğu için amaç fonksiyonunu yani uygunluk değerleri yüksek olan bireyler bir sonraki nesle aktarılmaz. Bu araştırmada $\rho=0,5$, $m=0,02$ olarak seçilmiş ve bir sonraki nesle seçilme yöntemi olarak ise rulet çarkı yöntemi kullanılmıştır.

Okuma puanı için oluşturulacak regresyon modeli Matlab .. programı yardımıyla belirlenen genetik işlemler ile 57 iterasyon sonucunda elde edilmiştir. 57. iterasyonda genetik algoritma en iyi çözüme ulaşmış ve iterasyonu durdurmuştur.

Tablo 25. Genetik Algoritma Yöntemi İle Regresyonda Model Parametre Tahminleri

β_0	β_1	β_2	β_3	β_4	β_5
297,6015	15,2399	8,0086	16,5884	34,7829	50,8834

Son iterasyonda elde edilen parametre tahminleri Tablo 25'te görülmektedir. Bu parametreler için elde edilen uygunluk değeri 52,4677 olarak hesaplanmıştır.

Okuma başarısı için oluşturulan regresyon modeli ;

$Y_{okuma_i} = 322,3131 - 3,7945X_{1i} + 3,1001X_{2i} + 14,3183X_{3i} + 34,7185X_{4i} + 46,9142X_{5i}$ şeklindedir.

Hata yüzdesi mutlak hatanın gerçek değere oranı olarak bulunur. Okuma puanı için veri setinin ortalaması 430,03. Genetik algoritma yöntemiyle bulunan minimum mutlak sapma ise 52,4677 olarak ölçülmüştür. Buna göre

Hata yüzdesi = $52,4677/430,03 = \%12$ olarak hesaplanmıştır.

5.BÖLÜM

SONUÇ, TARTIŞMA VE ÖNERİLER

5.1. Sonuç ve Tartışma

Yapay zeka gelişen teknoloji sayesinde hayatın her alanında kullanılan bir yazılımsal ve donanımsal sistemdir. Yapay zekanın alt modellerinden veya yöntemlerinden olan genetik algoritmalar ise birçok farklı alanda karşımıza çıkmaktadır. Genetik algoritmalar sayesinde optimizasyon ve tahminleme işlemleri kolaylıkla yapılmaktadır. Meydana gelen teknolojik ve bilimsel gelişmelerle beraber genetik algoritma uygulamalarının araştırmalarda kullanımının arttığı görülmektedir. Birçok alanda kullanılmasına rağmen eğitim alanında genetik algoritma yöntemini kullanarak yapılan çalışma sayısı yok denecek kadar azdır. Bilimsel gelişmelere ayak uydurabilmek için özellikle eğitim alanında bu tür çalışmalara ihtiyaç duyulduğu düşünülmektedir. Ülkelerin eğitim politikalarını düzenlemesi ve geliştirmesi için, öğrenci başarılarının yordanması ve sınıflanması gibi işlemlerin geleceğe ışık tutması, önlemler alınması gibi durumlarda eğitime ve öğrenci başarısını arttırmaya katkı sağlayabileceği öngörülmektedir. Özellikle ülkemizde yapay zeka modellerinden yapay sinir ağları, veri madenciliği ile ilgili eğitim alanında çalışmaların sayısı genetik algoritma yöntemi ile yapılan çalışma sayısına göre oldukça fazladır. Eğitimde bu tür tahminleme yöntemlerinin kullanılması gelecek bireyler açısından önemli ve değerlidir. Öğrencilerin başarılarını etkileyen değişkenler göz önüne alınarak başarıyı arttırıcı önlemler alınabilir.

Bu çalışmada PISA 2015 uygulamasına katılan 15 yaş grubu öğrencilerin sonuçlarından elde edilen veriler yardımıyla çoklu doğrusal regresyon modeli geliştirmek için parametre tahminlemesi yapılmıştır. Parametre tahmini yapılırken regresyon analizinde klasik ve yaygın olarak kullanılan EKK yöntemine alternatif yöntem olarak genetik algoritma yöntemi kullanılmıştır. Genetik algoritma yönteminin tahminleme konusundaki performansını ortaya koymak amaçlanmıştır. Çalışmada model geliştirme yöntemi olarak genetik algoritma uygulaması Matlab programı aracılığıyla yapılmıştır. Sonuçların ortaya konulması ve değerlendirilmesi açısından EKK sonuçlarıyla kıyaslama yapılmıştır.

Çalışmanın bu bölümünde, araştırmaya ilişkin sonuçlar alt problemler yardımıyla ortaya konuldu ve sonuçlar literatürden çeşitli araştırmalarla desteklenmiştir.

Araştırmanın problemi PISA 2015 başarılarının tahminlenmesinde genetik algoritma yönteminin performansının değerlendirilmesidir. Performansın değerlendirilmesi hususunda çoklu regresyon modeli geliştirilmiş modele en uygun parametreler genetik algoritma yardımıyla bulunmuştur.

Çalışmanın alt problemlerini PISA uygulamasında ölçülen Matematik, Fen ve Okuma Becerilerine ilişkin elde edilen puanların tahminlenmesinde kullanılacak modeller belirlenmiştir. Üç farklı puan tipi olduğu için çalışmada üç farklı model geliştirilmiştir.

Çalışmanın sonucunda elde edilen sonuçlar maddeler halinde verilmiştir.

1. Birinci model Matematik Becerilerine ilişkin başarıyı tahminlemek amacıyla oluşturulmuştur. Model oluşturulurken matematik başarı puan ortalaması bağımlı değişken olarak belirlenirken öğrenci anketleri aracılığıyla belirlenen cinsiyet, okula ilgi ortalaması, aile ilgisi ortalaması, sınıf ve baba eğitim durumu ise bağımsız değişkenler olarak belirlenmiştir. Burada amaç tahmin yapılırken hatayı en aza indirmektir. Genetik algoritma yönteminde sonuca ulaşabilmek için amaç fonksiyonu kullanılır. Bu çalışmada amaç fonksiyonu olarak tahmin edilen puan ile gözlenen puan arasındaki mutlak sapmayı veren formül kullanılmıştır. Genetik algoritmayla minimize işlemi uygulanıp mutlak sapmanın en küçük olacağı değer bulunulmaya çalışılmıştır. Yapılan işlemler sonucunda hata sapma miktarı 54,08 olarak ölçülmüştür ve hata yüzdesi % 13 bulunmuştur.

2. Fen Becerilerine ilişkin başarıyı yordamak amacıyla oluşturulan modelde bağımsız değişkenler matematik modelinde olduğu gibi seçilirken bağımlı değişken veri setinden elde edilen fen başarı puan ortalaması olarak seçilmiştir. Amaç fonksiyonu doğrultusunda hata sapma miktarı 54,08 olarak ölçülmüş ve hata yüzdesi %13 bulunmuştur.

3. Son olarak Okuma Becerilerine ilişkin başarı tahmininde geliştirilen modelde bağımlı değişken PISA 2015 verilerinden elde edilen Okuma puan ortalaması olarak

belirlenmiştir. Hata sapma miktarı 52,47 olarak ölçülmüştür. Hata yüzdesi ise % 12 bulunmuştur.

4. Geliştirilen modellerde hata yüzdeleri beklenenden fazla çıkmıştır. Bunun nedeni aynı özelliklere sahip olan öğrencilerin puanları arasındaki farkın fazla olmasıdır. Yani öğrenciler seçilen değişkenler bağlamında benzer özelliklere sahipken aldıkları puanlar arasındaki fark birbirine yakın değildir. Oysaki benzer özelliklere sahip bireylerin başarı puanlarının da birbirine yakın olması beklenmektedir. Veri setindeki dağılımdan kaynaklı genetik algoritma yöntemi istenilen düzeyde performans gösterememiştir. Veri setindeki bu uyuşmazlığın nedeni öğrencilerin bireysel farklılıklarından kaynaklı olduğu düşünülmektedir. Öğrencilerin uygulamaya katıldığı andaki motivasyonları, ortam gibi sebeplerden dolayı bu farklılık ortaya çıkmış olabilir.

Literatür incelendiğinde genetik algoritmada parametre tahmini yapılan çalışmalarda algoritma performansının oldukça iyi olduğu sonuçlarına varılmış ve alternatif yöntem olarak kullanılabileceği dile getirilmiştir (Aydın, 1998; Altunkaynak ve Esin, 2004; Roeva, 2005; Akyol, 2006; Karakoca, 2009).

Şen, Yılmaz ve Geban'ın 2018 yılında yaptıkları çalışmada öğrencilerin bireysel farklılıkları öğrenme-öğretme sürecinde dikkate alınması gereken özelliklerdendir. Bunun nedeni olarak ise öğrencilerin seçtikleri öğrenme yaklaşımları ve öğretmeye verdikleri cevaplarının bireysel özelliklerindeki farka göre değişmesi olarak düşünmüşlerdir. Özgen ve Alkan (2014) yaptıkları araştırmada öğrencilerin öğrenme stillerine uygun etkinliklerin akademik başarıya etkisini araştırmış ve başarıyı arttırdığı sonucuna varmıştır. Topkaya ve Çelik (2009) "Eğitimde Bireysel Farklılıklar" adlı kitabı incelemişler ve öğrencilerin öğrenme süreçlerinde farklılık yaratan birçok değişkenin olduğunu dile getirmişlerdir. Bireysel farklılıkların eğitime yansımaları, zeka ve yetenekler, ilgiler, öğrenmede güdülenme, öğrencinin kişilik yapısı, inançları gibi bölümlerden oluştuğunu açıklamışlardır. Öğrencilerinin akademik eğitimleri boyunca bu farklılıkların başarılarını etkileyeceğini bu yüzden tek tip öğretim yönteminden kaçınılması gerektiğini savunmuşlardır.

5.2. Öneriler

5.2.1. Araştırmacılara Yönelik Öneriler

1. Aynı veri setinde farklı değişkenler seçilerek genetik algoritma yönteminin performansı değerlendirilebilir.

2. Genetik algoritma yönteminde kullanılan çaprazlama oranı, mutasyon oranı, popülasyon büyüklüğü gibi parametreler değiştirilerek farklı uygulamalar ile çalışma tekrarlanabilir.

3. Genetik algoritma yöntemi ile klasik yöntemler karşılaştırılabilir.

4. Genetik algoritma yöntemi uygulanırken Matlab paket programı kullanılmıştır. Genetik algoritma yönteminin uygulanabileceği farklı programlar üzerinden model değerlendirilebilir.

5.2.2. Uygulamaya Yönelik Öneriler

1. Veri setinde aynı özellikteki bireyler arasındaki puan farkı fazla olduğu için model istenilen düzeyde performans gösterememiştir. Uygulama yapılırken daha tutarlı veri grubuyla çalışılması daha tutarlı sonuçlar elde edileceği düşünülmektedir.

2. Model geliştirilirken modeli aynı veri seti üzerinde eğitmek ve test etmek yerine veri setinden uygun şekilde örneklem seçilip eğitim ve test etmede farklı örneklem kullanılması hata yüzdesinin daha aza indirgenmesini sağlamada önemlidir.

3. Değişken seçimi yapılırken öğrencilerin bireysel farklılıklarının ön plana çıktığı değişkenler seçilmesi daha tutarlı sonuçlar elde edilmesini sağlar.

4. PISA gibi uluslararası düzeyde ya da ulusal düzeyde yapılan çalışmalarda anketlerde verilen sorularda öğrencilerin bireysel farklılıkları ön plana çıkaracak öğrenme stilleri, motivasyonları, ders ve sınav anında güdülenmeleriyle ilgili soruların sayısının artırılması gerekmektedir.

5. Çalışmada PISA verileri kullanılarak uluslararası düzeyde yapılan uygulama değerlendirilmeye çalışılmıştır. Ülkemizde yapılan ABİDE (Akademik Becerilerin İncelenmesi ve Değerlendirilmesi) uygulamasından elde edilen sonuçlar için tahmin

modeli geliřtirmede standart regresyon yöntemleri yerine genetik algoritma yöntemi kullanılarak analiz yapılması önerilmektedir.



KAYNAKÇA

- Akyol, A. P. (2006). *Doğrusal olmayan ekonometrik modellerin genetik algoritma ile parametre tahmini*. Ankara: Gazi Üniversitesi Sosyal Bilimler Enstitüsü: Yüksek lisans tezi.
- Altunkaynak, B., Esin, A. (2004). Doğrusal olmayan regresyonda parametre tahmini için genetik algoritma yöntemi. *Gazi Üniversitesi Fen Bilimleri Dergisi*, 17(2). 43-51.
- Aydın, Ö. M. (1998). *Hedef programlama olarak modellenmiş minmad regresyon problemine genetik algoritma yaklaşımı*. Ankara: Ankara Üniversitesi Fen Bilimleri Enstitüsü: Yüksek lisans tezi.
- Bayata, Ö. (2012). *Genetik algoritmaların ders çizelgeleme probleminde kullanımı ve eğitim kurumlarında uygulanması*. Ankara: Gazi Üniversitesi Fen Bilimler Enstitüsü: Doktora tezi.
- Biroğul, S. (2005). *Genetik algoritma yaklaşımıyla atölye çizelgeleme*. Ankara: Gazi Üniversitesi Fen Bilimler Enstitüsü: Yüksek lisans tezi.
- Bolat, B., Erol, K.O., İmrak, C.E. (2004). Mühendislik uygulamalarında genetik algoritma ve operatörlerin işlevleri. *Mühendislik ve Fen Bilimleri Dergisi*, 4: 264-271.
- Büyüköztürk, Ş. (2015). *Sosyal bilimler için veri analizi el kitabı*. Ankara: Pegem Akademi 21. Baskı.
- Coşkuntuncel, O. (2010). Sosyal bilimlerde yanlı regresyon tahmin edicilerinin kullanılması. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi* 1(2). 100-108.
- Çetin, M.K. (2006). *Hisse senedi yatırım kararlarında genetik algoritma kullanımı*. Antalya: Akdeniz Üniversitesi Sosyal Bilimler Enstitüsü: Doktora tezi.
- Daban, F. ve Özdemir, E. (2004). Eğitimde verimliliği arttıran ders programlarının hazırlanması için genetik algoritma kullanımı. *Eğitim Bilimleri ve Uygulama*, 3(6). 245-257.
- Davis, L. (1991). *Handbook of genetic algorithms*. New York: Van Nostrand Reinhold.

- Demir, E., Akkuş, Ö. (2015). An introductory study on “ how the genetic algorithm Works in the parameter estimation of binary logit model?”. *International Journal of Sciences: Basic and Applied Research*, 19(2). 162-180.
- Dilaver, D. (2015). *Genetik algoritmalar yardımıyla iş atölye çizelgelemesi üzerine bir çalışma*. İzmir: Dokuz Eylül Üniversitesi: Yüksek lisans tezi.
- Doğan, M. (2016). *LASSO tahminlerinin genetik algoritma yaklaşımı ile bulunması*. Samsun: Ondokuz Mayıs Üniversitesi Fen Bilimleri Enstitüsü: Yüksek lisans tezi.
- Edwards, A.L. (1984). *An introduction to linear regression and correlation*. 206. New York: W. H. Freeman and Company.
- Erar, M.A. (2013). *Doğrusal regresyon analizine giriş*. Ankara: Nobel Akademik Yayıncılık.
- Erginer, A. (2006). *Avrupa birliği eğitim sistemleri ve Türkiye eğitim sistemiyle karşılaştırmalar*. Ankara: Pegem A. Yayıncılık
- Gen, M., Cheng, R. (2000). *Genetic algorithms and engineering design and automation*. New York: John Wiley & Sons, Inc.
- Genel, H. (2004). *Genetik algoritmalarla portföy organizasyonu*. Ankara: Ankara Üniversitesi Sosyal Bilimler Enstitüsü: Yüksek lisans tezi.
- Gerşil, M. ve Palamutçuoğlu, T. (2013). Ders çizelgeleme probleminin melez genetik algoritmalar ile performans analizi. *Niğde Üniversitesi İİFB Dergisi*, 6(1). 242-262.
- Ghareb, A.S., Bakar, A.A. ve Hamdan, A.R. (2016) Hybrid feature selection based on enhanced genetic algorithm for text categorization. *Expert Systems with Applications*, 49:31-47.
- Grefenstette, J. (1986). Optimisation of control parameters for genetic algorithms. *Systems, Man and Cybernetics, IEEE Transactions on (0018-9472)*, 16 (1). 122-128.
- Goldberg, D. E. (1989). *Genetic algorithms in search, optimization, and machine learning*. Newyork: Adisson-Wesley.

- İşçioğlu, F., Kırkık, Ş., Deveci, F. (2019). Türkiye gıda sektörü BIST-100 endeksi bildirimlerine ait çoklu doğrusal regresyon analizi. *Türk Tarım-Gıda Bilim ve Teknoloji Dergisi*, 7(2). 325-332.
- Fraenkel, J. R., Wallen, N. E., Hyun, H. (2011). *How to design and evaluate research in education*. McGraw-Hill.
- Jong, K. A. (1975). *An analysis of the behavior of a class of genetic adaptive systems*. Michigan Üniversitesi: Computer and communication Sciences Department: Dissertation.
- Kabra, R. R. ve Bichkar, R. S. (2014). Students' performance prediction using genetic algoritihm. *International Journal of Computer Engineering and Applications* 4(3). 19-29.
- Karakoca, A. (2009). *Çok değişkenli lineer olmayan modellerde genetik algoritma*. Konya: Selçuk Üniversitesi Fen Bilimleri Enstitüsü: Doktora tezi.
- Kaya, O. (2010). *Kombinatoriyal optimizasyon problemlerinin bir sınıfının genetik algoritmalar ile çözümü üzerine*. İzmir: Ege Üniversitesi Fen Bilimleri Enstitüsü: Yüksek lisans tezi.
- Kayım, H. (1985). *İstatistiksel ön tahmin yöntemleri*. Ankara: Hacettepe Üniversitesi İktisadi ve İdari Bilimler Yayınları.
- Koza, J. R. (1995). Two ways of discovering the size and shape of a computer program to solve a problem., *Proceedings of the Sixth International Conference on Genetic Algorithm*, 287-294.
- Kurt, M., Semetay, C. (2001). *Genetik algoritma ve uygulama alanları*. [Çevrim-içi: https://www.mmo.org.tr/sites/default/files/e7757b1e12abc7_ek.pdf], Erişim tarihi: 18 Mayıs 2019.
- Lakshmi, T. M., Martin, A. ve Venkatesan, V. P. (2013). An analysis of students performance using genetic algorithm. *Journal of Computer Sciences and Applications* 1(4). 75-79.
- Melanie, M. (1999). *An introduction to genetic algorithm*. Cambridge: the MIT Press. 5th edition.
- MEB (2016). Uluslararası öğrenci değerlendirme programı (PISA). [Çevrim-içi: http://pisa.meb.gov.tr/wp-content/uploads/2014/11/PISA2015_UlusalRapor.pdf], Erişim tarihi: 19 Mayıs 2019.

- MEB (2016). TIMSS uluslararası matematik ve fen eğilimleri araştırması. [Çevrim-içi: http://timss.meb.gov.tr/wp-content/uploads/TIMSS_2015_Ulusal_Rapor.pdf], Erişim Tarihi: 25 Mayıs 2019.
- Metin, H. M. (2012). *Genetik algoritmaları kullanarak bir sinir ağının eğitilmesi ve doğrusal olmayan modellerle uygulanması*. Muğla: Sıtkı Koçman Üniversitesi Fen Bilimleri Enstitüsü: Yüksek lisans tezi.
- Michalewicz, Z. (1996). *Genetic algorithm+data structures= evolution programs*, New York: Springer 3rd edition.
- Moreno, J., Ovalle, D. A. ve Vicari, R. M. (2012). A genetic algorithm approach for group formation in collaborative learning considering multiple students characteristics. *Computer & Education* 58, 560-569.
- Orhunbilge N. (2017). *Uygulamalı regresyon ve korelasyon analizi*. İstanbul: Nobel Yayın Dağıtım.
- Özgen, K., Alkan, H. (2014). Yapılandırmacı öğrenme yaklaşımı kapsamında, öğrencilerin öğrenme stillerine uygun öğrenme etkinliklerinin akademik başarı ve tutuma etkileri: fonksiyon ve türev kavramı örnekleme. *Turkish Journal of Computer and Mathematics Education*, 5(1). 1-38.
- Öztürkler, H., ve Altan Ş. (2008). A genetic algorithm approach to parameter estimation in nonlinear econometric models. *Dumlupınar Üniversitesi: Sosyal Bilimler Dergisi* 20, 67-76.
- Paksoy, S. (2007) *Genetik algoritma ile proje çizelgeleme*. . Adana: Çukurova Üniversitesi Sosyal Bilimler Enstitüsü: Doktora tezi.
- Reeves, C.R., Rowe, J.E. (2002). *Genetic algorithm principles and perspectives a guide to GA theory*. Usa: Kluwer Academic Pres.
- Roeva, O. (2005). Genetic algorithms for a parameter estimation of a fermentation process model : a comparison. *Bulgarian Academy of Science* 3, 19-28.
- Sivanandam, S.N. ve Deepa, S.N. (2008) *Introduction to Genetic Algorithms*, New York: Springer.
- Şen, Ş., Yılmaz, A., Geban, Ö. (2018). Self-regulated learning skills: Adaptation of scale. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 9(4). 339-355.

- Şenol, H. (2014). *Koroner arter hastalığının risk faktörlerinin irdelenmesinde alternatif bir yaklaşım: genetik algoritmalar*. Denizli: Pamukkale Üniversitesi Sağlık Bilimleri Enstitüsü: Yüksek lisans tezi.
- Şıklar E. (2000). *Regresyon analizine giriş*. Eskişehir: Anadolu Üniversitesi Fen Fakültesi Yayınları, 133s.
- Taştan, S. (2015). Genetik algoritmalar ile arama modellerinin belirlenmesi. *Cumhuriyet Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 16(2). 161-171.
- Tate, D. M. ve Smith, A. E. (1992). A genetic approach to the quadratic assignment problem. *Computers & Operations Research*.
- Topkaya Zehir, E., Çelik, H. (2009). Eğitimde bireysel farklılıklar. *Eğitimde Kuram ve Uygulama Dergisi*, 5(2). 316-321.
- Yeo, M. F., Agyei, E. O. (1998). Optimising engineering problems using genetic algorithms. *Engineering Computations*, 15 (2): 268-280.
- Yıldız Aybek, H.S. (2016). *Öğrenci başarısının yapay sinir ağları ile kestirilmesi: anadolu üniversitesi açık öğretim sistemi örneği*. Eskişehir: Sosyal Bilimler Enstitüsü: Yüksek lisans tezi.
- Yiğit, V. (2011). Genetik algoritma ile Türkiye net elektrik enerjisi tüketiminin 2020 yılına kadar tahmini. *International Journal of Engineering Research and Development* 3(2). 37-41.

ÖZ GEÇMİŞ

Kişisel Bilgiler

- Adı Soyadı : Yasemin YETKİN
- Doğum Yeri ve Tarihi :Denizli/ 15.02.1991
- Lisans Öğrenimi : Boğaziçi Üniversitesi, İlköğretim Matematik Öğretmenliği
- Yüksek Lisans Öğrenimi : Yüzüncü Yıl Üniversitesi Eğitim Bilimleri Enstitüsü Eğitimde Ölçme ve Değerlendirme
- Bildiği Yabancı Diller : İngilizce
- Bilimsel Faaliyetleri :
- Stajlar : -Şehit Mehmet Emin Yurdakul Ortaokulu, Mart 2011- Mayıs 2011
- Mecidiye Ortaokulu, Nisan 2013- Mayıs 2013
- Özel Asfa Halil Necati Ortaokulu Ekim 2013- Ocak 2014
- Türkan Şoray Ortaokulu Mart 2014- Mayıs 2014
- Projeler :
- Çalıştığı Kurumlar : Mehmetçik Selen Ortaokulu 2014-

İletişim

E-Posta Adresi : ysmn.ytkn@gmail.com



VAN YÜZÜNCÜ YIL ÜNİVERSİTESİ
Eğitim Bilimler Enstitüsü

LİSANSÜSTÜ TEZ ORJİNALLİK RAPORU

VAN YÜZÜNCÜ YIL ÜNİVERSİTESİ
Eğitim Bilimler Enstitüsü

08/07/2019

Tez Başlığı / Konusu

PISA Başarısını Tahmin Etmede Genetik Algoritma Yaklaşımı

Yukarıda başlığı/konusu belirlenen tez çalışmamın Kapak sayfası, Giriş, Ana bölümler ve Sonuç bölümlerinden oluşan toplam 57 sayfalık kısmına ilişkin, 08/07/2019 tarihinde şahsım/tez danışmanım tarafından Turnitin intihal tespit programından aşağıda belirtilen filtreleme uygulanarak alınmış olan orijinallik raporuna göre, tezimin benzerlik oranı %13 (On üç) dür.

Uygulanan Filtreler Aşağıda Verilmiştir:

- Kabul ve onay sayfası hariç,
- Teşekkür hariç,
- İçindekiler hariç,
- Simge ve kısaltmalar hariç,
- Gereç ve yöntemler hariç,
- Kaynakça hariç,
- Alıntılar hariç,
- Tezden çıkan yayınlar hariç,
- 7 kelimedenden daha az örtüşme içeren metin kısımları hariç (Limit match size to 7 words)

Van Yüzüncü Yıl Üniversitesi Lisansüstü Tez Orijinallik Raporu Alınması ve Kullanılmasına İlişkin Yönergeyi İnceledim ve bu yönergede belirtilen azami benzerlik oranlarına göre tez çalışmamın herhangi bir intihal içemediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

Gereğini bilgilerinize arz ederim.

08/07/2019

Yasemin YETKİN

Adı Soyadı : Yasemin YETKİN

Öğrenci No : 15940001051

Anabilim

Dalı : Eğitim Bilimleri Anabilim Dalı

Programı : Ölçme ve Değerlendirme

Statüsü : Y. Lisans Doktora

DANIŞMAN

Dr. Öğr. Üyesi Gürol ZIRHLIOĞLU

08/07/2019

ENSTİTÜ ONAYI
UYGUNDUR

