

Makine Öğrenmesi Algoritmaları Kullanılarak Öğrencilerin Yükseköğretime Devam Etme Durumunun Tahmin Edilmesi

M. Hanefi CALP

Yönetim Bilişim Sistemleri, İİBF, Ankara Hacı Bayram Veli Üniversitesi, Türkiye

hanefi.calp@hbv.edu.tr

Özet – Bu çalışmada, okul dışı ölçütler dikkate alınarak öğrencilerin yükseköğretime devam etme durumunun veya potansiyelinin tahmin edilmesi amaçlanmıştır. Bu noktada, öğrencilerin okul dışı ölçütleri olarak; öğrencilerin evinin konumunun, aile büyüklüğünün, ebeveynlerin beraberlik durumunun, haftalık ders çalışma süresinin, özel ders alma durumunun, okul sonrası boş zamanın, arkadaşlarıyla dışarıda buluşma sıklığının, okul günlerinde aldığı alkol miktarı olarak belirlenmiştir. Bu ölçütler hakkındaki veriler incelenerek öğrencilerin devamsızlık sayısına ve final notlarına olan etkisi incelenmiştir. Yöntem olarak; K-En Yakın Komşu, Doğrusal Regresyon, Karar Ağacı ve Naive Bayes makine öğrenmesi tekniklerinden yararlanılmıştır. Kullanılan bu yöntemlerin sonuçları karşılaştırılmış ve en başarılı yöntem tespit edilmiştir. Elde edilen sonuçlara göre, başarı performansı en yüksek olan yöntemin %98,60'lık bir oranla Naive Bayes olduğu ortaya çıkmıştır.

Anahtar Kelimeler – Makine Öğrenmesi, K-En Yakın Komşu, Doğrusal Regresyon, Karar Ağacı, Naive Bayes, Yükseköğretim

I. GİRİŞ

Eğitim birçok insanın ve ülkenin geleceğini önemli ölçüde etkileyen bir kavramdır. Bununla birlikte Eğitim kavramı; bilgi, kültür, yetenek ve becerilerin kuşaklar arası aktarılması ve bireylerde bilinçli bir davranışın oluşturulması hali olarak tanımlanır. Bir başka ifadeyle Eğitim; bireyin kendini veya özünü tanıması, kendini keşfetmesi ve toplumun faydası için kendini geliştirmesi, topluma uyum sağlaması ve sosyalleşmesi olarak da tanımlanabilir [1,2]. Bu bağlamda, okulda her öğrenciye aynı eğitim verildiği halde öğrencilerin başarıları farklılık göstermektedir. Bunun nedeni akademik başarının sadece okulda verilen eğitime değil aynı zamanda okul dışında yapılan faaliyetlere de bağlı olmasıdır. Çocuk gelişimi sürecinde, hem ev hem de okul ortamının çok önemli pozitif katkısı olduğu gibi destekleyici bir çevre niteliği de taşıdığı söylenebilir [3].

Sosyal bir varlık olan İnsan, içinde bulunduğu çevreden doğal olarak etkilenmekle birlikte aynı zamanda sözkonusu çevreyi de etkiler. Yaşam sürülen mahalle, köy veya şehirler insanların

sosyalleşmesi ve böylece gelişmesi bakımından oldukça etkili ve önemlidir [2].

Öğrencilerin demografik veya okullarından kaynaklanan birçok durum, onların akademik başarı seviyelerinde etkilidir. Eğitim süreçlerinde, öğrencinin yaşantısı esas olarak kabul edilirse, bu çerçevedeki öğrencilere ait veriler makine öğrenmesi yöntemleri kullanılarak onların başarıları analiz edilebilir [4]. Özellikle günümüzde, yapay zekâ, makine öğrenmesi, robot teknolojileri, nesnelerin interneti ve hizmet otomasyon konularında büyük gelişmeler yaşanmaktadır. Eğitim başta olmak üzere tıp, ulaşım, sosyal hizmet, askeri, turizm, tarım gibi birçok alanda sözkonusu makine öğrenmesi yöntemleri kullanılmaktadır [5].

Makine öğrenmesi yönteminde, algoritmalar veri yapıları öncesinden kendilerine bir kez öğretildikten sonra tüm yeni durumlara uyum sağlayabilecek yeterlilikte olup, aynı zamanda gelecek durumlarla alakalı da tahmin yeteneğine sahip olabilmektedirler [6].

Bu çalışmada, makine öğrenmesi yöntemlerinden K-En Yakın Komşu, Doğrusal Regresyon, Karar Ağacı ve Naive Bayes yöntemleri kullanılarak öğrencilerle ilgili birtakım verilerle öğrencilerin yükseköğretime devam edip etmeyeceğinin belirlenmesi amaçlanmıştır. Her bir faktörün yükseköğretime devam edip etmeyeceğine etkisi yukarıda sayılan makine öğrenmesi yöntemleri ile incelenmiş, sonrasında bu yöntemlerden hangisinin daha doğru sonuç verdiğini ortaya koyabilmek için alınan sonuçlar karşılaştırılarak analiz edilmiştir.

II. LİTERATÜR TARAMASI

Bu bölümde, makine öğrenmesi algoritmalarının farklı alanlarda kullanımı üzerine kısa bir literatür özeti verilmiştir.

Kardaş ve Güvenir (2020), yaptıkları çalışmada sınav, ödev ve projelerin dönem sonunda yapılan sınavlara etkisini makine öğrenmesi yöntemleri ile incelemiştir. Yöntem olarak; Lojistik Regresyon, Destek Vektör Makineleri, Lineer Regresyon, K-En Yakın Komşu ve Lineer Destek Vektör Makineleri kullanmışlardır. Büyük boyutta veri kullanarak elde ettikleri sonuçları karşılaştırmış ve analiz etmişlerdir. Deneysel sonuçlar, Lineer Regresyon modelinin biraz daha geliştirilmesi gerektiğini, sınıflandırma işlemlerinin ise başarılı sonuçlar ürettiğini ortaya koymuştur. Sonuç itibarıyla, dönem sonu sınav sonucunun tespiti için önerilen lineer modelin performansı düşük, ancak sınıflandırma başarısı yüksek olduğu ortaya çıkmıştır [1].

Demirel (2019), Karar Ağacı algoritmaları ile çocuk işçiliğine neden olan çeşitli faktörler belirlemiştir. Bu çalışmada; karar ağacı algoritmalarından CART ve CHAID ile 114 ülke üzerinde çocuk işçiliği konusunu etkileyen faktörlerin belirlenmesi amaçlanmıştır. Bu süreçte, öncelikle literatür taraması yapılmış daha sonra toplam 10 farklı değişken dikkate alınarak regresyon ağaçları oluşturulmuş ve sonuçları analiz edilmiştir [7].

Korkmaz (2020), Yapay Sinir Ağları ve Destek Vektör Makineleri yöntemlerini ve karşılaştırma amacıyla da istatistiğe dayalı yöntem olan Lojistik Regresyon kullanarak kurumsal bir bankadan temin edilen müşterilere ait verilere uygulamışlardır. Çalışmada amaç, kredi başvurusu yapan bir müşterinin risk durumunun belirlenmesidir. Yani hangi risk kategorisine girip girmediğinin tespit edilmesidir. Söz konusu riske

göre de başvurunun kabul edilip edilmeyeceği belirlenecektir [8].

Başar ve Genç (2020), ülke bazında gerçekleşen suç oranlarını tespit etmek amacıyla suç endeksi tahminini gerçekleştiren bir model önermişlerdir. Çalışmada, hesaplanan suç endeksi kullanılmış ve 106 ülkeye ait güvenilirlik düzeyleri tahmin edilmiştir. Modellerin oluşturulması sürecinde, yapay sinir ağları, lojistik regresyon analizi ve çok kriterli karar verme yöntemlerinden biri olan MOORA yöntemi kullanılmıştır. Çalışmadan elde edilen sonuçlara göre, yapay sinir ağları yöntemi kullanılarak ülkelerin güvenilirlik düzeylerine ait yapılan tahminlerin doğru bir şekilde sınıflandırma oranının veya başarı performansının diğer yöntemlere göre daha yüksek olduğu ortaya çıkmıştır [9].

Yılmaz ve diğerleri (2019), inşaat bütçesinin etkin kullanımı amacıyla firmalara verilen ihale bedellerini olumsuz etkileyen faktörlerin belirlenmesini incelemiştir. Burada yapay sinir ağları ve lineer regresyon yöntemleri kullanılarak bina bakım-onarım ödeneklerinin etkin kullanımını sağlamak amaçlanmıştır. Veriler, regresyon-korelasyon yöntemi kullanılarak SPSS programı ile analiz edilmiştir. Elde edilen sonuçlar, MatlabR2013b programında “nntool” modülü yardımıyla test edilmiş olup karşılaştırmalar yapılarak analiz edilmiştir [10].

III. YÖNTEM VE TEKNİKLER

A. Çalışmanın Amacı

Bu çalışmanın amacı, okul dışı ölçütler dikkate alınarak öğrencilerin yükseköğretime devam durumunun belirlenmesidir. Bu süreçte, öğrencilerin okul dışı faaliyetleri (öğrencilerin evinin konumunun, aile büyüklüğünün, ebeveynlerin beraberlik durumunun, haftalık ders çalışma süresinin, özel ders alma durumunun, okul sonrası boş zamanın, arkadaşlarıyla dışarıda buluşma sıklığının, okul günlerinde aldığı alkol miktarının) hakkındaki verileri incelenerek devamsızlık sayısına ve final notlarına olan etkisi belirlenmiştir.

B. Yöntem ve Teknikler

Çalışmanın yöntemi sürecinde; Doğrusal Regresyon, K-En Yakın Komşu, Karar Ağacı ve Naive Bayes yöntemleri kullanılmış olup bu yöntemlerin sonuçları karşılaştırılmış ve en etkili

yöntem tespit edilmiştir. Bu bölümde, çalışmada kullanılan yöntemlere kısaca yer verilmiştir.

i. Karar Ağacı

Sınıflama ve regresyon modellerinden biri olan Karar Ağaçları, genelden özele bir ağaç yapısına sahip olup eğitilmiş veri ile oluşturulmaktadır. Oluşturulan ağaç yapısı, örneklem içindeki tüm verileri kapsayan bir kök düğümle başlamaktadır. Karar ağacı yönteminde veriler, öğrenme ve sınıflama aşaması olmak üzere iki şekilde sınıflandırılmaktadır. Öğrenme aşaması öncesinde bilinen eğitim verisi, modeli oluşturmak için sınıflandırma yöntemi kullanılır. Sözkonusu model, sınıflama kuralları ile belirtilmektedir. İkinci aşama olan sınıflama aşamasında ise, test verisi karar ağacının doğru yapılandırılıp yapılandırılmadığını test etmek için kullanılmaktadır [7].

ii. K-En Yakın Komşu

Bu algoritmada bir problemin çözümü sınıflandırma mantığı ile gerçekleştirilir. Örnek tabanlı algoritmalar sınıfında yer almaktadır. Öğrenme süreci, veri içerisindeki eğitim seti kullanılarak yapılır. Eğitim işlemi ise, en yakın k adet veriyi (varsayılan), daha önceden belirlenmiş uzaklık ölçütü dikkate alınarak benzerlik düzeylerinin belirlenmesi sonucunda gerçekleştirilir. Sınıflandırmak için algoritmaya yeni bir veri girildiğinde, eğitilmiş veri seti içerisinde bulunan k adet en yakın merkez komşunun sınıf etiketlerine bakılır. Daha sonra sınıf etiketlerinin yoğunluk düzeyine göre sözkonusu veri ilgili kümeye eklenir [11].

iii. Doğrusal Regresyon

Doğrusal Regresyon algoritması, bağımlı değişkenle bağımsız değişken arasındaki ilişkiyi ortaya koymayı amaçlayan en uygun modelin tespit edilmesi olarak tanımlanabilir. Sözkonusu analiz, temelde olasılık tabanlı olup; bir olayın gerçekleşme olasılığının gerçekleşmeme olasılığına oranı ile hesaplanır [12].

iv. Naive Bayes

Naive Bayes algoritması, makine öğrenme yönteminin danışmanlı öğrenme altyapısına dayanır. Verisinde bulunan örnek verilerin hangi sınıfa ait olduğu önceden bellidir. Bu yöntem, istatistiksel bir yöntem olup her özelliğin başarı sonucuna veya performansına etkisinin olasılığıyla hesaplanır. Bununla birlikte, sözkonusu algoritma

genel olarak var olan verisetinin, birleşik olasılıklar ile sınıflandırılması sürecinde kullanılır [14].

C. Verilerin Tanımlanması

Bu bölümde, veriler hakkında ayrıntılı bilgiye yer verilmiştir. Tablo 1’de herbir veri ve açıklamaları bulunmaktadır.

Tablo 1. Veriler ve Açıklamaları

Sütun No	Sütun Adı ve Açıklaması	Sütun İçeriği ve Açıklaması	Sütun İçeriğinin Türü
1	<i>school</i> → Öğrencinin okulunun adı	<i>GP</i> → Gabriel Pereira Lisesi <i>MS</i> → Mousinho da Siveria Lisesi	Kategorik
2	<i>sex</i> → Öğrencinin cinsiyeti	<i>F</i> → Kız <i>M</i> → Erkek	Kategorik
3	<i>age</i> → Öğrencinin yaşı	15 – 22 arası	Sayısal
4	<i>address</i> → Öğrencinin evinin bulunduğu konum	<i>U</i> → Kentsel bölge <i>R</i> → Kırsal bölge	Kategorik
5	<i>famsize</i> → Öğrencinin ailesindeki birey sayısı	<i>LE3</i> → 3 veya daha az <i>GT3</i> → 3’ten az	Kategorik
6	<i>Pstatus</i> → Öğrencinin ebeveynleriyle birlikteliği	<i>T</i> → Beraber yaşıyorlar <i>A</i> → Ayrı yaşıyorlar	Kategorik
7	<i>Medu</i> → Öğrencinin annesinin eğitim durumu	<i>0</i> → Hiç öğrenim görmemiş <i>1</i> → İlköğretim (4. sınıfa kadar) <i>2</i> → Ortaokul (5. – 9. sınıf arası) <i>3</i> → Lise <i>4</i> → Yüksek öğrenim	Kategorik
8	<i>Fedu</i> → Öğrencinin babasının eğitim durumu	<i>0</i> → Hiç öğrenim görmemiş <i>1</i> → İlköğretim (4. sınıfa kadar) <i>2</i> → Ortaokul (5. – 9. sınıf arası) <i>3</i> → Lise <i>4</i> → Yüksek öğrenim	Kategorik
9	<i>Mjob</i> → Öğrencinin annesinin işi	<i>teacher</i> → Öğretmen <i>health</i> → Sağlık çalışanı <i>services</i> → Hizmet sektöründe <i>at_home</i> → İşsiz <i>other</i> → Başka bir alanda çalışmakta	Kategorik
10	<i>Fjob</i> → Öğrencinin babasının işi	<i>teacher</i> → Öğretmen <i>health</i> → Sağlık çalışanı <i>services</i> → Hizmet sektöründe <i>at_home</i> → İşsiz <i>other</i> → Başka bir alanda çalışmakta	Kategorik
11	<i>reason</i> → Öğrencinin okulunu seçme	<i>course</i> → Ders programından dolayı	Kategorik

	nedeni	<i>home</i> → Evine yakın olduğu için <i>reputation</i> → Okul hakkında olumlu şeyler duyduğu için <i>other</i> → Başka bir sebepten ötürü	
12	<i>guardian</i> → Öğrencinin velisi	<i>mother</i> → Annesi <i>father</i> → Babası <i>other</i> → Başkası	Kategorik
13	<i>traveltime</i> → Öğrencinin okula gidiş süresi	1→ 15 dakikadan az 2→ 15 – 30 dakika 3→ 30 – 60 dakika 4→ 1 saat veya üzeri	Kategorik
14	<i>studyttime</i> → Öğrencinin haftalık ders çalışma süresi	1→ 2 saatten az 2→ 2 – 5 saat 3→ 5 – 10 saat 4→ 10 saat veya üzeri	Kategorik
15	<i>failures</i> → Öğrencinin kaldığı ders sayısı	0→ Hiç yok 1→ 1 tane 2→ 2 tane 3→ 3 tane ve üzeri	Kategorik
16	<i>schoolsupsup</i> → Öğrencinin eğitim yardımı alma durumu	<i>yes</i> → Yardım alıyor <i>no</i> → Yardım almıyor	Kategorik
17	<i>famsupsup</i> → Öğrencinin ailesinin sosyal hizmet yardımı alma durumu	<i>yes</i> → Yardım alıyor <i>no</i> → Yardım almıyor	Kategorik
18	<i>paid</i> → Öğrencinin ücretli özel ders alma durumu	<i>yes</i> → Ders alıyor <i>no</i> → Ders almıyor	Kategorik
19	<i>activities</i> → Öğrencinin okul sonrası etkinliklere katılımı	<i>yes</i> → Katılıyor <i>no</i> → Katılmıyor	Kategorik
20	<i>nursery</i> → Öğrencinin okul öncesi eğitim geçmişi	<i>yes</i> → Eğitim almış <i>no</i> → Eğitim almamış	Kategorik
21	<i>higher</i> → Öğrencinin yükseköğretimle devamı	<i>yes</i> → Yüksek öğrenim görmek istiyor <i>no</i> → Yüksek öğrenim görmek istemiyor	Kategorik
22	<i>internet</i> → Öğrencinin evde internete erişme durumu	<i>yes</i> → İnternet var <i>no</i> → İnternet yok	Kategorik
23	<i>romantic</i> → Öğrencinin romantik ilişki durumu	<i>yes</i> → İlişkisi var <i>no</i> → İlişkisi yok	Kategorik
24	<i>famrel</i> → Öğrencinin ailesiyle olan yakınlığı	1→ Hiç yakın değil 2→ Az yakınlıkta 3→ Orta yakınlıkta 4→ Yakın 5→ Çok yakın	Kategorik
25	<i>freetime</i> → Öğrencinin okul sonrasındaki boş zamanı	1→ Hiç yok 2→ Az derecede 3→ Orta derecede 4→ Yeterli derecede 5→ Çok fazla var	Kategorik
26	<i>goout</i> → Öğrencinin	1→ Hiç	Kategorik

	arkadaşlarıyla dışarıda buluşma sıklığı	2→ Az derecede 3→ Orta derecede 4→ Yeterli derecede 5→ Çok fazla	
27	<i>Dalc</i> → Öğrencinin okul günlerinde aldığı alkol miktarı	1→ Hiç 2→ Az miktarda 3→ Orta miktarda 4→ Fazla miktarda 5→ Aşırı miktarda	Kategorik
28	<i>Walc</i> → Öğrencinin hafta sonlarında aldığı alkol miktarı	1→ Hiç 2→ Az miktarda 3→ Orta miktarda 4→ Fazla miktarda 5→ Aşırı miktarda	Kategorik
29	<i>health</i> → Öğrencinin sağlık durumu	1→ Çok kötü 2→ Kötü 3→ Ortalama 4→ İyi 5→ Çok iyi	Kategorik
30	<i>absences</i> → Öğrencinin devamsızlık sayısı	0 – 93 gün arası	Sayısal
31	<i>G1</i> → Öğrencinin 1. dönem notu	0 – 20 puan arası	Sayısal
32	<i>G2</i> → Öğrencinin 2. dönem notu	0 – 20 puan arası	Sayısal
33	<i>G3</i> → Öğrencinin final notu	0 – 20 puan arası	Sayısal

Veri setindeki verilerin fazla ve detaylı olması sebebiyle önışlemeden geçirilmiştir. Ön işleme tabi tutulan verilerin veri madenciliği teknikleri ile analiz edilmesi amaçlanmış, bu kapsamda veri setindeki özelliklere ait değerler oluşturulmuştur. Veriseti değişkenleri proje konusuna göre incelendiğinde bazı niteliklerin veri analizinde kullanılmayacağına karar verilmiş olup veriseti 6 adet niteliğe düşürülmüştür. Nihai olarak belirlenen verisetinde değişkenlerden bazıları boş olarak belirtilmiştir. Literatürde ve uygulamalarda, bir veriseti için boş hücre varsa, o hücre “kayıp veri” olarak tanımlanır. Boş hanelerin veri setinden çıkartılması karar verilmiştir.

IV. BULGULAR VE TARTIŞMA

Deneyler açık kaynak kodlu Weka ve Colab Research kullanılarak gerçekleştirilmiştir. Karar Ağacı, K-En Yakın Komşu, Naive Bayes ve Doğrusal Regresyon algoritmaları kendilerine has birçok farklı parametresi değiştirilerek test edilmiştir. Test sonucunda oluşturulan modeller, verisetinin sınıflandırma başarısının hangi ölçüde etkilendiği bakımından incelenmiştir.

Algoritmalar kullanılırken yapılan denemelerden sonra en yüksek doğruluk oranları elde edildiği için eğitim sınıfında %75, test sınıfında da %25’lik veri alınmıştır.

Veriler kullanılarak her bir yöntemle bir model geliştirilmiştir. Modellerden elde edilen bulgular karşılaştırılmış ve analiz edilmiştir. Sözkonusu bulgulara göre; Naive Bayes algoritması ile %98,60 gibi yüksek bir başarı oranı yakalandığı görülmüştür. K-en yakın komşu algoritması ile %95,95 gibi yüksek bir başarı oranı yakalandığı görülmüştür. Ayrıca, Kesinlik değeri 0,96; F Skor değeri ise 0,979 olarak elde edilmiştir. Doğrusal regresyon algoritması ile %96,39 gibi yüksek bir başarı oranı yakalandığı görülmüştür. Ayrıca, Kesinlik değeri 0,964; F Skor değeri ise 0,982 olarak elde edilmiştir. Karar ağacı algoritması ile %95,76 gibi yüksek bir başarı oranı yakalandığı görülmüştür.

Tablo 2. Yöntemlerden elde edilen bulgular

Ölçüt	Naive Bayes	K-En Yakın Komşu	Doğrusal Regresyon	Karar Ağacı
Doğruluk	98,60	95,95	96,39	93,76
Kesinlik	0,976	0,960	0,964	0,928
F Skor	0,988	0,979	0,982	0,958

Tablo 2 incelenerek, K-En Yakın Komşu, Doğrusal Regresyon, Karar Ağacı ve Naive Bayes yöntemlerinin başarı oranları kıyaslandığı bu çalışmada; yükseköğretime devam eğilimi tahmininde Naive Bayes yöntemi ile yapılan performans testlerinin diğer yöntemlere göre daha başarılı sonuçlar verdiği görülmüştür.

Bu çalışma için kullanılabilir en iyi makine öğrenme yönteminin Naive Bayes olduğu ortaya çıkmıştır. Yöntemlerden üçü birbirine yakın sonuçlar verirken Naive Bayes yöntemi 2-3 puan farkıyla (ki önemli sayılabilecek bir fark) diğerlerinden ayrılmaktadır.

Karar ağacı yönteminin ise %93,76 doğruluk oranıyla en düşük sonucu veren yöntem olduğu görülmüştür.

Önerilen modeller kullanılarak elde edilen test sonuçlarının yüzdesi Tablo 3'te verilmiştir.

Tablo 3. Modellere göre test sonuçlarının yüzdesi

Önerilen Model	Doğru (%)	Yanlış (%)
NB	99,41	1,59
KNN	95,80	4,20
LR	96,26	3,73
KA	92,57	7,42

Tablo 3 incelendiğinde, Tablo 2'de bertilen performans başarı sonuçlarıyla büyük ölçüde örtüştüğü görülmektedir. Tablo 3 teki veriler, Naive Bayes yönteminin en başarılı sonuçlar ürettiğini ortaya koymuştur.

V. SONUÇ VE ÖNERİLER

Okul öncesi eğitim geçmişi, eğitim yardımı alma durumu, evde internete erişme durumu, okul sonrası etkinliklere katılımı, ücretli özel ders alma durumu gibi faktörler yükseköğretime devam etme eğilimini etkileyebilmektedir. Makine öğrenmesi yöntemlerinin icra edildiği diğer alanlardaki örnekler dikkate alındığında, sözkonusu yöntemlerin yükseköğretime devam davranışı konusunda da başarılı bir tahmin performansı gösterme potansiyeline sahip olacağı söylenebilir. Bu bağlamda, gerçekleştirilen çalışmada makine öğrenmesi yöntemlerinden K-En Yakın Komşu, Doğrusal Regresyon, Karar Ağacı ve Naive Bayes kullanılarak öğrencilerin yükseköğretime devam davranışı belirlenmiştir. Bu arada, çalışmanın yönteminde yararlanılan makine öğrenmesi yöntemleri hakkında özet niteliğinde bilgilere yer verilmiştir. Çalışmada kullanılan veriler, Gabriel Pereira ve Mousinho da Silveira ortaöğretim kurumlarında yapılan anketlerden elde edilmiştir.

Kullanılan makine öğrenmesi yöntemlerine göre elde edilen sonuçlar karşılaştırılmış ve analiz edilmiştir. Böylece yükseköğretime devam faktörlerini, makine öğrenme yöntemlerinden en iyi performans elde edilen yöntemin bulunması sağlanmıştır. Bu verilerden yola çıkarak okul öncesi eğitim geçmişi, eğitim yardımı alma durumu, evde internete erişme durumu, okul sonrası etkinliklere katılımı, ücretli özel ders alma durumu gibi faktörlerin yükseköğretime devam etme eğilimine yüksek oranda etki ettikleri tespit edilmiştir.

KAYNAKLAR

- [1] Kardaş, K., & Güvenir, H. A. (2020). Kısa Sınavların, Ödevlerin ve Projelerin Dönem Sonu Sınavına Olan Etkilerinin Farklı Makine Öğrenmesi Teknikleri ile Araştırılması. EMO Bilimsel Dergi, 10(1), 21-27.
- [2] Kahraman, S. (2016). Boşanma Sonrasında Genç-Ebeveyn İlişkilerinin Sürdürülmesi ve Öğrencinin Akademik Başarısına Etkisi (Afyon İli Örneği) (Doctoral Dissertation, Selçuk Üniversitesi Sosyal Bilimleri Enstitüsü).

- [3] Erdoğan Ç. & Demirkasımoğlu, Nihan. (2010). Ailelerin Eğitim Sürecine Katılımına İlişkin Öğretmen ve Yönetici Görüşleri. *Kuram ve Uygulamada Eğitim Yönetimi*, 2010, 3:3:399-431.
- [4] Gök. Murat. (2017). Makine Öğren Yöntemleri İle Akademik Başarının Tahmin Edilmesi. *Gazi Üniversitesi Fen Bilimleri Dergisi Part C: Tasarım ve Teknoloji*, 5(3), 139-148.
- [5] Yalçın Kayıkçı, M., & Kutluk Bozkurt, A. (2018). Dijital Çağda Z ve Alpha Kuşağı, Yapay Zeka Uygulamaları ve Turizme Yansımaları.
- [6] Bayır, A., Özdemir, Ş., & Gülseçen, S. Türkiye'deki Seçmen Eğilimlerinin C4. 5 Karar Ağacı Algoritması İle Belirlenmesi. *Yönetim Bilişim Sistemleri Dergisi*, 2(2), 223-233.
- [7] Demirel, Ş., & Yakut, S. G. (2019). Karar Ağacı Algoritmaları ve Çocuk İşçiliği Üzerine Bir Uygulama. *Sosyal Bilimler Araştırma Dergisi*, 8(4), 52-65.
- [8] Korkmaz, G. Yapay Zekâ Yöntemleriyle Sınıflandırma ve Finans Sektöründe Bir Uygulama. *Akademik Yaklaşımlar Dergisi*, 11(2), 91-109.
- [9] Başar, Ö. D., & Genç, E. G. (2020). Ülkelerin Güvenli Olmalarının Tahmininde Lojistik Regresyon, Yapay Sinir Ağları ve Moora Yöntemlerinin Karşılaştırılması. *Journal Of Life Economics*, 7(2), 123-134.
- [10] Yılmaz, M., Kanıt, R., Erdal, M., Yıldız, S., & Bakış, A. (2016). Bina Bakım Onarım Ödeneklerinin Etkin Kullanımı Maksadıyla İhale Bedelini Etkileyen Faktörlerin Yapay Sinir Ağları Ve Lineer Regresyon Yöntemleri ile Belirlenmesi. *Politeknik Dergisi*, 19(4), 461-470.
- [11] Pala, M. A., Çimen, M. E., Boyraz, Ö. F., Yıldız, M. Z., & Boz, A. F. (2019). Meme Kanserin Teşhis Edilmesinde Karar Ağacı ve KNN Algoritmalarının Karşılaştırmalı Başarım Analizi. *Academic Perspective Procedia*, 2(3), 544-552.
- [12] Çakın, E., & Özdemir, A. (2019). Veri Zarflama Analizi Temelli Yapay Sinir Ağları ve Lojistik Regresyon Analizi ile Teknoloji Geliştirme Bölgelerinin Etkinliklerinin Tahminlenmesi. *Hacettepe Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 37(2), 271-293.
- [13] Filiz, E., Karaboga, H. A., & Akogul, S. (2017). BIST-50 Endeksi Değişim Değerlerinin Sınıflandırılmasında Makine Öğrenmesi Yöntemleri ve Yapay Sinir Ağları Kullanımı. *Journal of The Cukurova University Institute Of Social Sciences*, 26(1).