



İSTANBUL
TOPKAPI
ÜNİVERSİTESİ

LİSANSÜSTÜ
EĞİTİM ENSTİTÜSÜ

**MAKİNE ÖĞRENMESİ YÖNTEMLERİ
KULLANILARAK ORTAOKUL ÖĞRENCİLERİNİN
BAŞARILARININ DEĞERLENDİRİLMESİ**

YÜKSEK LİSANS TEZİ

ZEYNEP GÖKMEN

Yazılım Mühendisliği Anabilim Dalı

Yapay Zekâ Bilim Dalı

**MAKİNE ÖĞRENMESİ YÖNTEMLERİ
KULLANILARAK ORTAOKUL ÖĞRENCİLERİNİN
BAŞARILARININ DEĞERLENDİRİLMESİ**

YÜKSEK LİSANS TEZİ

ZEYNEP GÖKMEN

22221001002

Yazılım Mühendisliği Anabilim Dalı

Yapay Zekâ Bilim Dalı

Tez Danışmanı: Dr. Öğr. Üyesi Ali Haydar Eser

ARALIK, 2023

09/01/2024

YÜKSEK LİSANS TEZ ONAY FORMU

Yazılım Mühendisliği Anabilim Dalı, Yapay Zekâ (Tezli) Yüksek Lisans Programı 22221001002 numaralı öğrencisi Zeynep GÖKMEN'in "**MAKİNE ÖĞRENMESİ YÖNTEMLERİ KULLANILARAK ORTAOKUL ÖĞRENCİLERİNİN BAŞARILARININ DEĞERLENDİRİLMESİ**" konulu Yüksek Lisans tezi Enstitümüz Yönetim Kurulunun 20/12/2023 tarihli ve 2023/22 sayılı Yönetim Kurulu kararıyla oluşturulan jüri tarafından oybirliği / oyçokluğu ile 09.01.2024 tarihinde kabul edilmiştir.

	<u>Unvan</u>	<u>Adı Soyadı</u>	<u>Üniversite</u>	<u>İmza</u>
ASIL ÜYELER				
Danışman	Dr. Öğr. Üyesi	Ali Haydar ESER	İstanbul Topkapı Üniversitesi	
1. Üye	Dr. Öğr. Üyesi	Buket İŞLER	İstanbul Topkapı Üniversitesi	
2. Üye	Dr. Öğr. Üyesi	Alper FİDAN	İstanbul Aydın Üniversitesi	
YEDEK ÜYE				
1. Üye	Dr. Öğr. Üyesi	Yüksel BAL	İstanbul Topkapı Üniversitesi	
2. Üye	Dr. Öğr. Üyesi	Şükrü Mustafa KAYA	İstanbul Aydın Üniversitesi	

ONAY

Prof. Dr. Özlem KUNDAY
Enstitü Müdürü

(*) Oybirliği/Oyçokluğu hâli yazı ile yazılacaktır.

(**) Kabul / Ret veya Düzeltme kararı hâli yazı ile yazılacaktır.

AKADEMİK DÜRÜSTLÜK BEYANI

Yüksek Lisans Tezi olarak sunduğum “Makine Öğrenmesi Yöntemleri Kullanılarak Ortaokul Öğrencilerinin Başarılarının Değerlendirilmesi” başlıklı çalışmanın, bilimsel ahlak ve geleneklere uygun olarak tarafımdan yazıldığını, yararlandığım eserlerin tamamının kaynaklarda gösterildiğini ve çalışmam içerisinde kullanıldıkları her yerde atıf yapıldığını belirtir ve onurumla doğrularım.

14.12.2023

Zeynep Gökmen

TEŞEKKÜR

Öncelikle bir kız çocuğuna Karadeniz'in ücra bir köyünden çıkıp öğretmen olma şansı veren ve akademik çalışmaları ile bilime katkıda bulunma imkânı sunan Ulu Önderimiz Mustafa Kemal Atatürk'e teşekkürlerimi borç bilirim. Tez çalışmam boyunca gerek bilgi birikimi gerek motivasyon desteği ile yardımlarını esirgemeyen değerli tez danışmanım Dr. Öğretim Üyesi Ali Haydar Eser'e ve eğitim hayatımın her aşamasında beni maddi manevi her zaman destekleyen aileme sonsuz teşekkür ederim.

Zeynep Gökmen

İÇİNDEKİLER

KABUL VE ONAY	ii
AKADEMİK DÜRÜSTLÜK BEYANI	iii
TEŞEKKÜR	iv
İÇİNDEKİLER	v
KISALTMALAR	vii
ŞEKİLLER LİSTESİ.....	viii
TABLolar LİSTESİ.....	ix
ÖZET.....	x
ABSTRACT	xi
1. GİRİŞ.....	1
2. LİTERATÜR TARAMASI	3
2.1. Çalışmanın Diğer Çalışmalardan Farkı	11
3. MATERYAL VE METOT	12
3.1. Materyal	12
3.1.1. Veri seti (data set)	12
3.2. Metot	14
3.2.1. Makine öğrenmesi algoritmaları	14
3.2.1.1. Random forest classifier (rastgele orman sınıflandırıcısı)	14
3.2.1.2. Gradient boosting classifier (gradyan artırma sınıflandırıcısı)	15
3.2.1.3. Support vector machine (destek vektör makinesi)	15
4. UYGULAMA VE BULGULAR.....	17
4.1. Uygulama	17
4.1.1. Veri ön işleme	17
4.1.2. Nitelik seçimi	17
4.1.3. Değerlendirme metrikleri	18
4.1.3.1. Doğruluk (accuracy):	18
4.1.3.2. Kesinlik (precision):	18
4.1.3.3. Duyarlılık (recall):	18
4.1.3.4. F1-score:	18
4.1.4. Sınıflandırma algoritmalarının modellenmesi	19
4.2. Bulgular	25
4.2.1. Modellerin karşılaştırılması	25

5. SONUÇ VE ÖNERİLER	27
5.1. Sonuç	27
5.2. Öneriler.....	28
KAYNAKÇA	30
ÖZGEÇMİŞ.....	34



KISALTMALAR

RF	: Random Forest
GBC	: Gradient Boosting Classifier
SVM	: Support Vector Machine
TP	: Gerçek pozitif
FP	: Yanlış pozitif
TN	: Doğru negatif
FN	: Yanlış negatif
RO	: Rastgele Orman
GAS	: Gradyan Artırma Sınıflandırıcısı
DVM	: Destek Vektör Makinesi
KNN	: K-En Yakın Komşu
DT	: Karar Ağaçları
CART	: Classification And Regression Tree
AutoML	: Automated machine learning
PISA	: Uluslararası Öğrenci Değerlendirme Programı
TIMSS	: Uluslararası Matematik ve Fen Eğilimleri Araştırması

ŞEKİLLER LİSTESİ

	<u>Sayfa</u>
Şekil 3.1: Random Forest Classifier Yapısı.....	15
Şekil 3.2: Destek Vektör Makinesi Yapısı	16



TABLÖLAR LİSTESİ

	<u>Sayfa</u>
Tablo 2.1: Literatür Taraması.....	5
Tablo 3.1: Veri Seti Nitelikleri ve Açıklamaları	12
Tablo 4.1: Random Forest Performans Sonuçları	20
Tablo 4.2: Random Forest Başarıya Etki Eden Değişkenler.....	20
Tablo 4.3: Gradient Boosting Performans Sonuçları	22
Tablo 4.4: Gradient Boosting Başarıya Etki Eden Değişkenler.....	22
Tablo 4.5: Support Vector Performans Sonuçları	24
Tablo 4.6: Support Vector Özelliklerin Ağırlıkları	25
Tablo 4.7: Modellerin Karşılaştırılması	25

ÖZET

MAKİNE ÖĞRENMESİ YÖNTEMLERİ KULLANILARAK ORTAOKUL ÖĞRENCİLERİNİN BAŞARILARININ DEĞERLENDİRİLMESİ

Eğitimin hemen her kademesinde hedef öğrencilerin başarılı olmasıdır. Öğrenci başarısını arttırmak için ise başarıya etki eden unsurların tespit edilmesi ve ihtiyaç duyulan alanlarda öğrencilere destek verilmesi önem arz etmektedir. Bu çalışmada makine öğrenmesi yöntemleri kullanılarak öğrencilerin ders başarısına yüksek oranda etki eden değişkenler yardımıyla öğrencilerin yılsonu başarı durumlarını tahmin eden makine öğrenmesi modelleri geliştirilmiş ve bu modeller arası karşılaştırmalar yapılmıştır. Bu amaçla İstanbul Esenler’de resmî bir ortaokulda eğitim gören 307 7.sınıf öğrencisine başarıya etki edebileceği düşünülen soruların yer aldığı 31 soruluk bir anket uygulanmıştır. Anket cevapları temel alınarak ve öğrencilerin bir önceki dönem yılsonu notları kullanılarak makine öğrenimi algoritmaları yardımı ile başarı tahmini yapılmıştır. Başarı tahmininde Random Forest, Gradient Boosting, Support Vector Machine olmak üzere 3 farklı makine öğrenmesi algoritması kullanılmıştır. Çalışma sonucunda ankette yer alan ödevlerini düzenli yapıp yapmama, üniversite okumayı isteyip istememe, ders sırasında not alma, anne-babanın eğitim düzeyi değişkenlerinin yılsonu başarısını oldukça etkilediği görülmüştür. Başarı tahmininde en başarılı modelin %90 oranı ile Gradient Boosting algoritması kullanılarak geliştirilen model olduğu gözlemlenirken en başarısız modelin ise %82 oranı ile Support Vector Machine algoritması ile geliştirilen model olduğu gözlemlenmiştir. Daha erken bir aşamada başarı tahmini, kurumların öğrencilerin genel performansını artırabilen etkili öğrenme yaklaşımlarını göz önünde bulundurarak net kayıt yönergeleri oluşturmalarına ve zayıf performanslardan kaçınmasına olanak sağlayacaktır.

Anahtar kelimeler: *yapay zekâ ve eğitim, kişiselleştirilmiş eğitim, makine öğrenmesi, akıllı eğitim*

ABSTRACT

EVALUATING THE ACHIEVEMENT OF SECONDARY SCHOOL STUDENTS BY USING MACHINE LEARNING METHODS

The goal at almost every level of education is for students to be successful. In order to increase student success, it is important to identify the factors that affect success and provide support to students in areas where they are needed. In this study, using machine learning methods, machine learning models that predict students' end-of-year success status with the help of variables that have a high impact on students' course success were developed and comparisons were made between these models. For this purpose, a 31-question survey was applied to 307 7th grade students studying at a public secondary school in Esenler, Istanbul, including questions that were thought to affect success. Success prediction was made with the help of machine learning algorithms, based on the survey answers and using the students' end-of-year grades from the previous semester. Three different machine learning algorithms were used to predict success: Random Forest, Gradient Boosting and Support Vector Machine. As a result of the study, it was seen that the variables included in the survey, whether they do their homework regularly or not, whether they want to go to university or not, taking notes during class, and the education level of their parents, significantly affect the end-of-year success. It was observed that the most successful model in predicting success was the model developed using the Gradient Boosting algorithm with a rate of 90%, while the least successful model was the model developed with the Support Vector Machine algorithm with a rate of 82%. Predicting success at an earlier stage will enable institutions to establish clear enrollment guidelines and avoid poor performances, taking into account effective learning approaches that can improve students' overall performance.

Keywords: *artificial intelligence and education, personalised education, machine learning, smart education*

1. GİRİŞ

Eğitim, kişilerin topluma faydalı bir birey olabilmeleri için gerekli bilgi, beceri ve anlayışların kazandırıldığı planlı bir süreçtir (TDK, 2023). Eğitim sürecinde yaşanan aksaklıklar büyük ölçüde toplumun gidişatını etkilemektedir. Son yıllarda ülkemizde yapılan gerek ulusal sınavlarda gerekse PISA (Uluslararası Öğrenci Değerlendirme Programı) ve TIMSS (Uluslararası Matematik ve Fen Eğilimleri Araştırması) kapsamında yapılan uluslararası sınavlarda öğrencilerimizin sınav ortalamalarının oldukça düşük olduğu görülmektedir. Türkiye uluslararası sınavlarda averajın altında kalarak son sıralarda yer almaktadır (Acar, 2012; Sarier, 2021). Bu sebeple öğrencilerimizin başarısını artırmak amacıyla pek çok proje ve çalışma yapılmaktadır (Fırat, 2019; Mahmut ve diğerleri, 2020). Öğrencilerin ders başarısının artması için öncelikle başarıya etki eden değişkenlerin belirlenmesi ve bu değişkenlerin iyileştirilmesi önem arz etmektedir. Başarıyı etkileyen pek çok etmen vardır. Aile yaşantısı, anne babanın eğitim durumu, sosyoekonomik durum, çevre, okul, öğrencinin derse katılımı, öğrencinin başarılı olacağına inancı bu etmenlere örnektir (Dursun ve Yüksel, 2004; Guven ve Cabakcor, 2013).

Makine öğrenmesi algoritmaları, veriler arasındaki kompleks ilişkileri en ince detaylarına kadar kavrayabilen, verilere dayalı yordamalar yapabilen, gerekli durumlarda karar verebilen algoritmalarlardır. Bu algoritmalar ile eğitim öğretim dönemi öncesinde performans tahmini yapılması potansiyel başarısızlığa karşı erken önlem alma imkânı tanıyacaktır. Hemen her alanda kullanılan makine öğrenmesi algoritmaları akademik başarının erken tahmininde de sıklıkla kullanılan yöntemlerden olmuştur (Agarwal, 2013).

Makine öğrenmesi algoritmaları eğitimde pek çok alanda kullanılmaktadır.

- Öğrenme Stillerinin Tespiti,
- Kişiselleştirilmiş Eğitim (Akıllı Eğitim Sistemleri),
- İçerik Geliştirme,
- Ölçme Sistemleri,

- Diyalog Tabanlı Eğitim Sistemleri,
- Eğitim Yönetimi,
- Akademik Başarı Tahmini bunlardan bazılarıdır.

Eğitimde veri madenciliği ve yapay zekâ algoritmalarının kullanılması;

- Çeşitli analizlerle öğrenci düzeylerini belirlemede öğretmenlere yardımcı olma,
- Öğrencilerin zayıf veya güçlü yönlerine yönelik ilgi çekici materyal ve içerikler sunma,
- Öğrencilerin performansını takip ederek değerlendirmelerde bulunma,
- Dersi planlama aşamasında öğretmene destek olma,
- Kişiselleştirilmiş eğitim platformları ile öğrencilerin hızına ve öğrenme stiline uygun içerikler sunma, gibi pek çok katkısı ile eğitimin kalitesini arttırmakta ve öğrencilerin daha verimli öğrenme deneyimleri kazanmasını sağlamaktadır.

Başarıya etki eden etmenlerin eğitim dönemi öncesinde tespit edilmesi ve düşük performans göstermesi beklenen öğrencilere ihtiyaç duyduğu konularda destek verilmesi öğrencilerin başarısını arttıracaktır (Zeineddine ve diğerleri, 2021).

Bu çalışmada eğitim öğretim dönemi başında öğrencilere bir anket yapılmış ve ankette öğrenci başarısına etki edeceği düşünülen unsurlara yer verilmiştir. Yapılan anket cevapları temel alınarak makine öğrenmesi algoritmaları ile modeller oluşturulmuş ve ders başarısına yüksek oranda etki eden değişkenler belirlenmiştir. Ders başarısına yüksek oranda etki eden değişkenler yardımı ile de öğrencilerin başarı durumlarına yönelik tahminde bulunulmuştur. Başarı tahmininde e-okuldan alınan bir önceki dönem yılsonu genel başarı ortalamaları kullanılmıştır. Eğitim öğretim yılı başında öğrencilerin genel başarısının tahmin edilmesi düşük başarı göstermesi beklenen öğrencilere yönelik rehberlik çalışmaları yapılarak olası başarısızlığın engellenmesine imkân tanıyacaktır.

2. LİTERATÜR TARAMASI

Makine öğrenmesi yöntemleri kullanılarak akademik başarının tahmin edilmesi konusunda yapılan literatür taramasında yurt içinde ve yurt dışında bu konuda pek çok çalışma yapıldığı görülmüştür.

Irmak (2023), tez çalışmasında metin madenciliği yöntemleri ile uzaktan eğitimdeki sorunları tespit etmiştir. Veri seti olarak Türkiye’de bir devlet üniversitesinde görev yapan akademisyenlerin ve eğitim gören öğrencilerin uygulamalar üzerindeki yazışmaları kullanılmıştır. Çalışmada 6 farklı regresyon modeli kullanılmıştır. Sorun tespitinde en başarılı modelin yapay sinir ağı olduğu görülmüştür (Irmak ve Ecevit Satı, 2022).

Hussain ve Khan (2023), makine öğrenmesi kullanılarak ortaokul ve ortaöğretim düzeyindeki öğrencilerin akademik performansının tahmin edilmesi konulu çalışmada denetimli makine öğrenimi yaklaşımlarını kullanılarak öğrencilerin notlarına ve not tahminlerine odaklanmıştır. Bu araştırma çalışmasında kullanılan veri seti Peşaver, Khyber Pakhtunkhwa Ortaokul ve Ortaöğretim Kurulu'ndan (BISE) elde edilmiştir. Çalışmada 4 farklı algoritma kullanılmış ve çalışma sonunda not tahmininde en başarılı algoritmanın % 96,6 oranı ile Genetik Tabanlı Karar Ağacı algoritması olduğu görülmüştür (Hussain ve Khan, 2023).

Kye (2023), tez çalışmasında ABD üniversiteye kayıt tahmini modelleme için sınıflandırma performansının karşılaştırmalı analizi konusunda çalışmıştır. Çalışmada Midwest kentsel bölgesindeki kâr amacı gütmeyen bir üniversitenin 4 yıllık bir programına kabul edilen ve her biri 50 ilişkili özelliğe sahip 108.798 öğrenciye ilişkin yedi yıllık veri seti kullanmıştır. Öğrencilerin üniversiteye kayıt kararlarını tahmin etmek için 4 farklı regresyon modeli kullanılmıştır. Kayıt için üniversiteye gelen öğrencilerin verilerinden yola çıkarak kaydolma ihtimali düşük olanlar ve kesin kayıt olacaklar tahmin edilmiştir. Mevcut çalışmada destek vektör makinesi ve yapay sinir ağı modelleri üniversite için uygulanacak en iyi tahmin modelleri olarak tanımlanmıştır (Kye, 2023).

Baasha ve ark. (2022), yapay zekâ modelleri kullanılarak lisansüstü akademik performansın değerlendirilmesi konulu çalışmada Malezya'da özel bir üniversitenin lisansüstü eğitim programında eğitim gören 635 öğrencinin gerçek verileri veri seti olarak kullanılmıştır. Çeşitli regresyon algoritmaları yardımıyla öğrencilerin başarıları tahmin edilmiştir. Çalışmada 6 farklı regresyon algoritması kullanılmıştır. Çalışma sonunda yapay sinir ağları algoritmasının maksimum tahmin doğruluğu ile %89 oranında başarılı tahminler yaptığı görülmüştür (Baashar ve diğerleri, 2022).

Bastem (2021), tez çalışmasında makine öğrenimi ve istatistiksel yöntemler kullanarak öğrencilerin programlama dersindeki kazanım bilgileri ile başarı tahmini konusuna yoğunlaşmıştır. Çalışmada Kaggle'ın sitesinden indirilen Portekiz'de bulunan iki farklı okuldaki öğrencilere ait veri setini kullanılmıştır. 3 farklı regresyon algoritması ile başarı tahmini yapılmış ve en başarılı tahmin yapan modelin lojistik regresyon olduğu görülmüştür (Bastem, 2021).

Abbasoğlu (2020), ortaokul öğrencilerinin akademik başarılarının eğitsel veri madenciliği yöntemleri ile tahmini konulu çalışmasında Yalova ilinde bulunan 4 farklı resmî ortaokulda eğitim gören öğrencilere ait sosyodemografik özelliklerin e-okul yönetim bilgi sisteminden alınmasıyla veri setini oluşturmuştur. 8 farklı regresyon algoritması kullanılarak başarı tahmini yapılmış ve en başarılı tahmin yapan algoritmanın lojistik regresyon olduğu görülmüştür (Bastem, 2021).

Wakelam ve ark. (2020), minimum kullanılabilir niteliklere sahip küçük gruplarda öğrenci performansı tahmini potansiyeli konulu çalışmasında makine öğrenmesi algoritmalarının küçük guruplarda da işe yaradığını kanıtlamıştır. 23 kişilik üniversite son sınıf öğrenci grubu üzerinde yürütülen çalışmada Karar Ağacı, K-En Yakın Komşular ve Rastgele Orman modelleri kullanılmıştır. En başarılı tahminleri rastgele orman algoritmasının gerçekleştirdiği görülmüştür (Wakelam ve diğerleri, 2020).

Lau (2019), yapay sinir ağları kullanılarak öğrenci akademik performansının modellenmesi, tahmin edilmesi ve sınıflandırılması konulu çalışmasında Çin'deki bir üniversitede eğitim gören 175 kız ve 825 erkek olmak üzere toplam 1.000 lisans öğrencisine ait bilgileri veri seti olarak kullanmıştır. Öğrenci performansı, dört yıllık öğrenim süresi boyunca standartlaştırılmış genel not ortalaması kullanılarak

belirlenmiştir. Çalışmada yapay sinir ağları kullanılarak makine öğrenmesi gerçekleştirilmiş ve başarılı sonuçlar elde edilmiştir (Lau ve diğerleri, 2019).

Gök (2017), makine öğrenmesi yöntemleri ile akademik başarının tahmin edilmesi konulu çalışmasında Yalova’da bulunan 3 farklı ortaokuldaki öğrencilere uygulanan 24 soruluk bir anket sonucunda veri setini oluşturmuştur. Anket yardımıyla elde edilen veriler ve Matematik, Türkçe ve dönem sonu ortalamaları kullanılarak makine öğrenmesi gerçekleştirilmiştir. Notlar Milli Eğitim Bakanlığı puan ölçeğine göre 100’lük sistemin bölümlenmesi ile oluşturulmuştur. Çalışmada kullanılan 6 farklı regresyon modeli arasında puan tahmininde en iyi modelin Rastgele Orman Regresyonu not tahmininde ise Lojistik Sınıflandırma Algoritmasının en iyi başarıyı gösterdiği görülmüştür (Gök, 2017).

Makine öğrenmesi yöntemleri kullanılarak başarı tahmini konusunda yapılan akademik çalışmalar Tablo 2.1’de özetlenmiştir.

Tablo 2.1: Literatür Taraması

Yazar (Yıl)	Çalışma	Yöntemler	Performans Odağı
Irmak ve Ecevit Satı, (2022)	Yapay Zekâ Yöntemleri ile Uzaktan Eğitimdeki Sorunların Tespiti Ve Öğrencilerin Akademik Performanslarının Tahmin Edilmesi	Destek Vektör Makineleri, Yapay Sinir Ağları, Lojistik Regresyon, Random Forest, K-En Yakın Komşu ve Naive (Basit) Bayes Sınıflandırıcı	Ön Lisans, Lisans
Christou ve diğerleri, (2023)	Makine Öğrenimini Kullanan Yüksek Öğrenim Öğrencileri İçin Performans ve Erken Düşüş Tahmini	Levenberg Marquardt Geri Yayılım, Broyden, Fletcher, Goldfarb ve Shanno Geri Yayılım, Stokastik Gradyan İniş, RBF Tabanlı Sinir Ağı, Karar Ağacı, Bayes Ağı, Destek Vektör Makinesi Ve Sinir Ağı Yapısı	Lisans

Tablo 2.2 (Devamı): Literatür Taraması

Yazar (Yıl)	Çalışma	Yöntemler	Performans Odağı
Hussain ve Khan, (2023)	Makine Öğrenmesi Kullanılarak Ortaokul ve Ortaöğretim Düzeyindeki Öğrencilerin Akademik Performansının Tahmin Edilmesi	Karar Ağacı, Genetik Tabanlı Karar Ağacı, K-En Yakın Komşular algoritması ve Genetik Tabanlı K-En Yakın Komşular	Ortaokul ve Ortaöğretim
Kye, (2023)	ABD Üniversiteye Kayıt Tahmini Modelleme İçin Sınıflandırma Performansının Karşılaştırmalı Analizi	Yapay Sinir Ağı, Karar Ağacı, Destek Vektör Makinesi, Lojistik Regresyon	Ortaöğretim
Alkan, (2023)	Akademik Başarının Eğitsel Veri Madenciliği Ve Makine Öğrenmesi Teknikleri Kullanılarak Tahminlenmesi	Lojistik Regresyon, Linear Diskriminant Analizi, Rastgele Orman, K-En Yakın Komşu, Karar Ağaçları, Destek Vektör Makineleri ve Naive Bayes	Açık Üniversite
Zilyas, (2023)	Makine Öğrenmesi Yöntemi İle Eğitim Başarısının Tahmini Modeli	K-En Yakın Komşu, Rastgele Orman, Lineer Regresyon, Bagged Trees, Gradient Boosting Regressor ve Karar Ağaçları (DT)	Ortaokul
Civelek, (2023)	Makine Öğrenmesi Yöntemleri İle Öğrenci Mezuniyet Notu Tahmini	Yapay Sinir Ağları, K-En Yakın Komşu Algoritması, Lineer Regresyon, Destek Vektör Makineleri ve Karar Ağaçları	Lisans
Bingül, (2023)	Yapay Zekâ Destekli Makine Öğrenmesi Algoritmaları İle Eğitsel Veri Madenciliği Alanında Öğrencilerin Akademik Performanslarının Tahmin Edilmesi	Random Forest, K-Nearest Neighbors, Destek Vektör Makineleri, Logistic Regression Naive Bayes, Yapay Sinir Ağları, Karar Ağaçları	Ortaöğretim

Tablo 2.1 (Devamı): Literatür Taraması

Yazar (Yıl)	Çalışma	Yöntemler	Performans Odağı
Baashar ve diğerleri, (2022)	Yapay Zekâ Modelleri Kullanılarak Lisansüstü Akademik Performansın Değerlendirilmesi	Sinir Ağları, En Küçük Kareler regresyonu, Destek Vektör Makineleri, Karar Ağaçları, Gauss Süreç Regresyonu, Topluluk Modelleri	Lisansüstü
(Çelik, 2022)	Makine Öğrenmesi ile Lise Öğrencilerine Mesleki Alan Seçim Rehberi	Rasgele Orman, K-En Yakın Komşu, Destek Vektör Makineleri, Gradyan Arttırma Makineleri, Aşırı Gradyan Arttırma, Hafif Gradyan Arttırma Makineleri, CatBoost, Yapay Sinir Ağları, Sınıflandırma ve Regresyon Ağaçları, Doğrusal Ayrım Analizi, Hiyerarşik Kümeleme	Ortaöğretim
Kaya, (2022)	Makine Öğrenmesi Algoritmaları Kullanarak Öğrencilerin Akademik Başarısını Etkileyen Faktörlerin Tespit Edilmesi	Karar Ağacı, Naive Bayes Çok Katmanlı Algılayıcı, Azaltılmış, En Yakın Komşu Hata Budaması, Destek Vektör Makineleri, Yapay Sinir Ağları, Sıralı Minimal Optimizasyon, Rastgele Orman, Karar Tabloları, Sinir Ağı	Ortaokul, Ortaöğretim, Lisans
Hasib ve diğerleri, (2022)	Ortaokul Öğrencisi Performansını Tahmin Etmeye Yönelik Makine Öğrenimi ve Açıklanabilir Yapay Zekâ Yaklaşımı	Lojistik Regresyon, En Yakın Komşu, Destek Vektör Makineleri, Xgboost ve Naive Bayes	Ortaokul

Tablo 2.1 (Devamı): Literatür Taraması

Yazar (Yıl)	Çalışma	Yöntemler	Performans Odağı
Bujang ve diğerleri, (2021)	Denetimli Makine Öğrenimi ile Öğrencilerin Not Tahmini İçin Kestirimci Analitik Modeli	Karar Ağacı, Rastgele Orman, Destek Vektör Makinesi, Lojistik Regresyon	Lisans
Yağcı, (2022)	Eğitsel Veri Madenciliği: Makine Öğrenimi Algoritmalarını Kullanarak Öğrencilerin Akademik Performansının Tahmini	Rastgele Ormanlar, En Yakın Komşu, Destek Vektör Makineleri, Lojistik Regresyon, Naive Bayes Ve K-En Yakın Komşu	Lisans
Koç ve Akın, (2022)	Liseye Giriş Sınavı Başarı Oranlarının Makine Öğrenmesi ve Beta Regresyon Modellerini Kullanarak Tahmin Edilmesi	Destek Vektör Regresyonu, Ortaokul Rastgele Orman, Karar Ağacı Ve Beta Regresyon	
Yavuzarslan ve Erol, (2022)	Öğrenme Yönetim Sistemi Log Kayıtlarının Akademik Başarı Tahmininde Kullanılması	KNN, Naive Bayes, SVM, CART ve C5.0	Lisans
Bastem, (2021)	Makine Öğrenme Algoritmaları Kullanılarak Yapay Zekâ Yoluyla Öğrenci Akademik Performans Tahmini	Karar Ağacı, Rastgele Orman ve Lojistik Regresyon	Ortaöğretim, Lisans
Aghalarova ve Bozkurt Keser, (2021)	Önerilen Yapay Sinir Ağı Algoritması ile Ortaokul Öğrencilerin Akademik Performansının Tahmini	AutoML, KNN ve SVM	Ortaokul

Tablo 2.1 (Devamı): Literatür Taraması

Yazar (Yıl)	Çalışma	Yöntemler	Performans Odağı
Kanchana ve diğerleri, (2021)	Öğrencilerin Akademik Performansının Erken Tahminine Yönelik Veri Madenciliği Yaklaşımı	SVM, Naive Bayes ve Karar Ağacı	Lisans
Buraimoh ve diğerleri, (2021)	Makine Öğrenimi Tekniklerinin Öğrenci Başarısının Tahminine Uygulanması	Doğrusal Diskriminant Analizi, Lojistik Regresyon, Sınıflandırma ve Regresyon Ağacı, K-En Yakın Komşu, Naive Bayes, SVM	-
Abbasoğlu, (2020)	Ortaokul Öğrencilerinin Akademik Başarılarının Eğitsel Veri Madenciliği Yöntemleri ile Tahmini	Lojistik Regresyon, Doğrusal Destek Vektör Makineleri, Doğrusal Olmayan Destek Vektör Makineleri, Rastgele Orman, Naive Bayes, Bagging, K-En Yakın Komşu, Yapay Sinir Ağları	Ortaokul
Selvi, (2020)	Bilecik İlinde İlköğretimden Liseye Geçiş Sınavlarında Makine Öğrenmesi Yöntemleri ile Öğrenci Başarısının Tahmini	Karar Ağaçları, Yapay Sinir Ağları, Bayes Sınıflandırması, Destek Vektör Makineleri, K-En Yakın Komşu, Genetik Algoritmalar	Ortaokul ve Ortaöğretim
Haridas ve diğerleri, (2020)	Akıllı Bir Özel Ders Sisteminden Okul Performansını Ve Erken Başarısızlık Riskini Tahmin Etmek	Karar Ağacı, Lojistik Regresyon, Sinir Ağları, SVM, KNN ve Rastgele Orman	-
Wakelam ve diğerleri, (2020)	Minimum Kullanılabilir Niteliklere Sahip Küçük Gruplarda Öğrenci Performansı Tahmini Potansiyeli	Karar Ağacı, K-En Yakın Komşu ve Rastgele Orman	Lisans

Tablo 2.1 (Devamı): Literatür Taraması

Yazar (Yıl) Çalışma	Yöntemler	Performans Odağı	
Hashim ve diğerleri, (2020)	Denetimli Makine Öğrenimi Algoritmalarına Dayalı Öğrenci Performansı Tahmin Modeli	Karar Ağacı, Naive Bayes, Lojistik Regresyon, Destek Vektör Makinesi, K-En Yakın Komşu, Sıralı Minimal Optimizasyon ve Sinir Ağları	Lisans
Lau ve diğerleri, (2019)	Yapay Sinir Ağları Kullanılarak Öğrenci Akademik Performansının Modellenmesi, Tahmin Edilmesi ve Sınıflandırılması	Yapay Sinir Ağları	Lisans
Tomasevic ve diğerleri, (2020)	Öğrenci Sınav Performansı Tahmini İçin Denetimli Veri Madenciliği Tekniklerine Genel Bakış ve Karşılaştırma	K-En Yakın Komşular, Destek Vektör Makineleri, Yapay Sinir Ağları, Karar Ağaçları, Bayesian Regresyon, Naif Bayes ve Lojistik Regresyon	Açık Üniversite
Ulutan, (2018)	Teog Fen Bilgisi Başarısını Etkileyen Değişkenlerin Çok Düzeyli Regresyon Modeli İle İncelenmesi	Çok Düzeyli Regresyon Analizi	Ortaokul
Miguéis ve diğerleri, (2018)	Öğrencilerin akademik performanslarına göre erken segmentasyonu: Tahmine dayalı bir modelleme yaklaşımı	Rastgele Ormanlar, Karar Ağaçları, Destek Vektör Makineleri, Naif Bayes, Torbalı Ağaçlar ve Uyarlanabilir Güçlendirme Ağaçlar	Lisans
Gök, (2017)	Makine Öğrenmesi Yöntemleri İle Akademik Başarının Tahmin Edilmesi	Naive Bayes K-En Yakın Komşular, Doğrusal Destek Vektör Makineleri, Radyal Tabanlı DVM ve Rastgele Orman, Lojistik	Ortaokul

Tablo 2.1 (Devamı): Literatür Taraması

Yazar (Yıl)	Çalışma	Yöntemler	Performans Odağı
Aydemir, (2017)	Veri Madenciliği Yöntemleri Kullanarak Meslek Yüksek Okulu Öğrencilerinin Akademik Başarı Tahmini	Yapay Sinir Ağları, Genetik Algoritmalar, Bayes Sınıflandırması, Karar Ağaçları, K-En Yakın Komşu, Destek Vektör Makineleri	Ön Lisans

2.1. Çalışmanın Diğer Çalışmalardan Farkı

- Veri seti daha önce kullanılmamış araştırmacı tarafından elde edilen orijinal bir veri setidir.
- Anket yardımı ile elde edilen başarıya etki eden değişkenler diğer çalışmalarla kıyaslandığında farklılıklar görülmektedir.
- Öğrencilerin akademik başarısını tahmin etmeye yönelik yapılan çalışmalarda genellikle çalışma grubu olarak lisans, ön lisans ve lisansüstü öğrencilerin tercih edildiği görülmektedir. Bu çalışmada çalışma grubu olarak ortaokul öğrencileri tercih edilmiştir.

3. MATERYAL VE METOT

Bu çalışmada öğrencilerden anket yoluyla toplanan orijinal veriler işlenerek veri yapıları ve formatı dönüştürüldü. Daha sonra farklı tahmin modelleri üretmek için python kullanılarak çeşitli sınıflandırma algoritmaları uygulandı. Son olarak model performansı, tahmine dayalı karar vermenin verimliliği ve doğruluğuna göre değerlendirildi.

3.1. Materyal

3.1.1. Veri seti (data set)

Bu çalışmada kullanılan veri seti İstanbul Esenler’de bulunan bir ortaokuldaki 7.sınıf öğrencilerine uygulanan 31 soruluk anket sonucunda elde edilmiştir. Öğrencilere yapılan anket ile ilgili gerekli izinler alınmıştır. Eğitim veri seti 32 nitelik ve 307 örnekten oluşmaktadır. Veri seti, çalışmaya konu olan öğrencilerin 6. sınıf yılsonu başarı ortalaması kullanılarak başarı tahmini yapılması üzerine oluşturulmuş yapıdadır. Çalışmadaki bağımlı değişken öğrencilerin yılsonu başarı ortalaması olarak belirlenmiştir. Çalışmada kullanılan veri setine ait nitelikler Tablo 3.1’ de gösterilmiştir.

Tablo 3.1: Veri Seti Nitelikleri ve Açıklamaları

Özellik	Açıklama
Cinsiyet	Öğrenci cinsiyeti
Yaş	Öğrenci yaşı
Ev Durum	Öğrencinin yaşadığı evin kira olup olmaması
Oda	Öğrencinin ders çalışmak için kendine ait bir alanın olup olmaması
Anne-Baba Durum	Öğrencinin anne babasının hayatta olup olmaması ve birliktelik durumu
Yaşam	Öğrencinin kimle birlikte yaşadığı

Tablo 3.1 Devamı: Veri Seti Nitelikleri ve Açıklamaları

Özellik	Açıklama
Kardeş	Öğrenci kardeş sayısı
Aeğitim	Annenin eğitim durumu
Beğitim	Babanın eğitim durumu
Agelir	Ailenin gelir düzeyi
Öçalışma	Öğrencinin okul saatleri dışında her hangi bir işte çalışma durumu
Kişi Sayısı	Öğrenci ile aynı evde yaşayan kişi sayısı
Uzaklık	Öğrencinin okula ulaşmak için harcadığı süre
H.İçi Ders Çalışma	Öğrencinin okul dışında hafta içi günlük ortalama ders çalışma süresi
H.sonu Ders Çalışma	Öğrencinin okul dışında hafta sonu günlük ortalama ders çalışma süresi
Öğrenme Yöntemi	Öğrencinin öğrenmek için tercih ettiği yöntem
Destekleme ve Yetiştirme	Öğrencinin hafta sonu yapılan destekleme ve yetiştirme kurslarına katılıp katılmadığı
Kurslarına Katılım	Öğrencinin üniversite okumak isteyip istememesi
Üniversite Okuma İsteği	Öğrencinin yaşadığı evde internet erişiminin olup olmaması
İnternet Erişimi	
Teknoloji ile Geçirilen Vakit	Öğrencinin günlük televizyon/ tablet/ bilgisayar veya telefon ile geçirdiği süre (Oyun ve eğlence amacıyla)
Okul Öncesi Eğitim	Öğrencinin okul öncesi eğitim alıp almadığı
Öz Yeterlilik	Öğrencinin çalışırsa başarılı olacağına inancı
Zihinsel Yeterlilik	Öğrencinin zihinsel olarak kendini nasıl hissettiği
Tekrar	Öğrencinin okuldan sonra o gün işlenen dersleri tekrar etme durumu
Ödev	Öğrencinin ödevlerini düzenli yapma durumu
Aile Başarı İnancı	Öğrencinin ailesinin öğrencinin başarılı olacağına inancı
Not Alma	Öğrencinin ders sırasında not alma durumu
Odaklanma	Öğrencinin ders sırasında derse odaklanma durumu

Tablo 3.1 Devamı: Veri Seti Nitelikleri ve Açıklamaları

Özellik	Açıklama
Devamsızlık	Öğrencinin okuldaki devamsızlık durumu
Sınav Kaygısı	Öğrencinin sınavlarda başarısız olma kaygısı
Uyku Saati	Öğrencinin günlük ortalama uyku saati
Başarı	Öğrencinin yılsonu genel başarı ortalaması (70 ve üstü “başarılı”, 70’in altı “başarısız”)

3.2. Metot

Bu tez çalışmasında Python programlama dili ve makine öğrenmesi algoritmaları kullanılarak öğrencilerin akademik başarıları tahmin edilecektir. Başarı tahmininde sınıflandırma algoritmaları kullanılacaktır. Kullanılan algoritmalar:

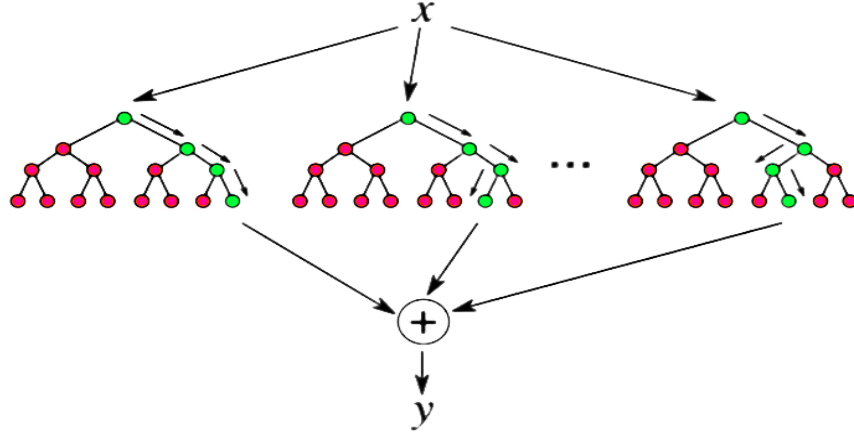
- Random Forest Classifier (Rastgele Orman Sınıflandırıcısı)
- Gradient Boosting Classifier (Gradyan Artırma Sınıflandırıcısı)
- Support Vector Machine (Destek Vektör Makinesi)

3.2.1. Makine öğrenmesi algoritmaları

Makine öğrenmesi algoritmalarının eğitim sektöründe kullanımı giderek artmaktadır. Bu çalışmada öğrencilerin başarılarını etkileyen değişkenler yardımı ile Random Forest (RF), Gradient Boosting Classifier (GBC) ve Support Vector Machine (SVM) olmak üzere 3 farklı algoritma kullanılarak başarı tahmini yapılmıştır.

3.2.1.1. Random forest classifier (rastgele orman sınıflandırıcısı)

Rastgele Orman modelleme sonuçları için rastgele seçilen eğitim veri kümeleri ve tahmin değişkenlerinin rastgele alt kümeleri kullanılarak birçok sınıflandırma ve regresyon ağacı oluşturulur (Breiman, 2001). Her ağaçtan elde edilen sonuçlar, her gözlem için bir tahmin vermek amacıyla toplanır. Bu nedenle, rastgele orman, ağaç modellerinin bazı yararlı niteliklerini (örneğin, tahmin ediciler ve sonuç arasındaki ilişkileri yorumlama yeteneği) korurken, tek bir karar ağacı modeline kıyasla genellikle daha yüksek doğruluk sağlar (Speiser ve diğerleri, 2019). Şekil 3.1’de Random Forest Classifier yapısı gösterilmiştir.



Şekil 3.1: Random Forest Classifier Yapısı

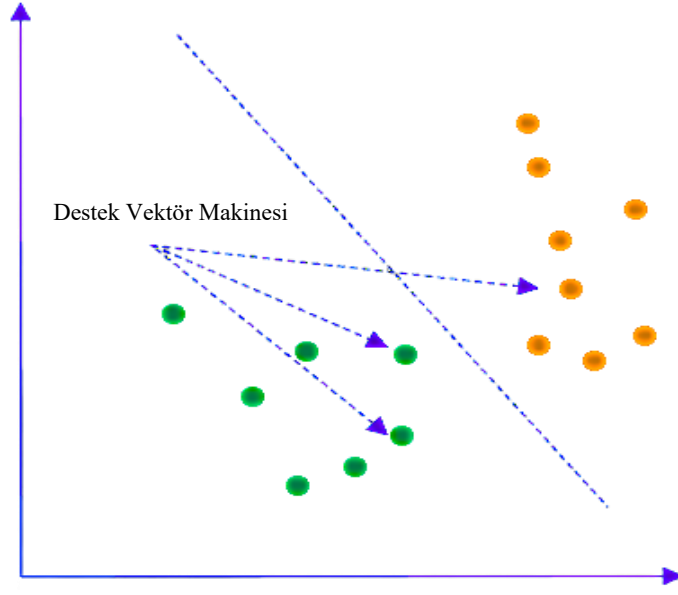
Kaynak: (Zilyas, 2023)

3.2.1.2. Gradient boosting classifier (gradyan artırma sınıflandırıcısı)

Gradyan Artırma Sınıflandırıcısı yanıt değişkeni daha doğru bir tahmin sağlamak için art arda yeni modellere uyarlar (Natekin ve Knoll, 2013). Yinelenen bir şekilde tahmin modellerini belirli kurallarla birleştirerek güçlü tahmin modellerini bulmaya çalışır. İlk yineleme anında tahminler üreten f_1 işlevi oluşturulmaktadır. Tahminler ile hedefler arasındaki değer farkı hesaplanarak farklılıklar için bir f_2 işlevi oluşturulmaktadır. İkinci kez yineleme yapıldığında f_1 ve f_2 işlevleri ile yeniden tahminler ve hedefler arasındaki fark birleştirilir. Bu işlem tekrarlanarak f_1 işlevinin başarısı sürekli eklenerek tahminler ve hedefler arasındaki farklılık sıfıra indirilmeye çalışılır (Friedman, 2001; Sönmez Sarıkaya, 2020).

3.2.1.3. Support vector machine (destek vektör makinesi)

Destek Vektör Makinesi (DVM), çok boyutlu özellik vektörlerinin ikili sınıflandırmasını gerçekleştirmek için tasarlanan bir algoritmadır. DVM başlangıçta iki sınıflı doğrusal verilerin sınıflandırılmasında kullanılırken daha sonra geliştirilerek çok sınıflı doğrusal olmayan verilerin sınıflandırılmasında kullanılmaya başlanmıştır. Çalışma prensibi olarak sınıfları birbirinden ayıran en uygun fonksiyonun tahmin edilmesi, sınıfların birbirinden ayrılabilmesi için en uygun karar fonksiyonlarının hiper düzlemde tanımlanması temeline dayanmaktadır (Kavzoğlu ve Çölkesen, 2010; Hearst ve diğerleri, 1998). Şekil 3.2’de Destek Vektör Makinesi yapısı gösterilmiştir.



Şekil 3.2: Destek Vektör Makinesi Yapısı
Kaynak: (Civelek, 2023)

4. UYGULAMA VE BULGULAR

4.1. Uygulama

Millî Eğitim Bakanlığı Okul Öncesi Eğitim ve İlköğretim Kurumları Yönetmeliği Sekizinci Bölüm Madde 53’de kapsamında dönem sonu tüm derslerin ağırlıklı ortalaması 70-84.99 olanlar “Teşekkür”, 85-100 olanlar “Takdir” belgesi ile ödüllendirilmektedir (MEB, 2014). Bu madde kapsamında veri setimizdeki bağımlı değişken olan tüm dersler ağırlıklı yılsonu başarı ortalaması 70 ve üzeri olan öğrenciler “başarılı”, yılsonu başarı ortalaması 70’in altında olan öğrenciler “başarısız” olarak nitelendirilerek sınıflandırma yapılmıştır.

4.1.1. Veri ön işleme

Yaptığımız çalışmada ankette sorulan soruların cevapları genellikle kategorik nitelik taşımaktadır. Fakat makine öğrenmesi algoritmaları kategorik yapıları veriler üzerinde doğrudan çalışmamaktadır. Bu sebeple string özellik taşıyan cevaplar One Hot Encoding yöntemi ile sayısal veriye dönüştürülmüştür.

Bu çalışmada öğrenci başarısına etki eden çevresel etmenler araştırıldığından özel öğrenme güçlüğü olan kaynaştırma öğrencileri çalışmadan çıkarılmıştır.

4.1.2. Nitelik seçimi

Makine öğrenmesinde model karmaşıklığının azaltılması, eğitim süresinin kısaltılması, model performansının ve anlaşılabilirliğinin artırılması gibi nedenlerle özellik seçimi yapılır. Özellik seçiminde filtreleme yöntemleri, gömülü yöntemler gibi birçok yöntem kullanılır. Bu çalışmada, veri setindeki tüm özellikler kullanılarak başarı tahmini yapılmıştır.

4.1.3. Değerlendirme metrikleri

Sınıflandırma sonucunda her bir algoritma için doğruluk (accuracy), kesinlik (precision), duyarlılık (recall), F1-Score değerlerine göre kullanılan algoritmaların başarı oranları hesaplanmış ve kıyaslamalar yapılmıştır.

4.1.3.1. Doğruluk (accuracy): Doğru yapılan tahminlerin sayısının tüm yapılan tahminlerin sayısına bölünmesiyle bulunur. Doğruluk formülü aşağıda gösterilmiştir (Coşkun ve Baykal, 2011).

$$\text{Doğruluk} = \frac{TP + TN}{TP + FP + TN + FN}$$

TP (Gerçek pozitif): Başarılı öğrenciye başarılı demektir.

FP (Yanlış pozitif): Başarısız öğrenciye başarılı demektir.

TN (Doğru Negatif): Başarısız öğrenciye başarısız demektir.

FN (Yanlış negatif): Başarılı öğrenciye başarısız demektir.

4.1.3.2. Kesinlik (precision): Pozitif tahmin edilen örneklerden kaçının doğru tahmin edildiğinin oranlanmasıdır. Kesinlik formülü aşağıda gösterilmiştir (Coşkun ve Baykal, 2011).

$$\text{Kesinlik} = \frac{TP}{TP + FP}$$

4.1.3.3. Duyarlılık (recall): Pozitif tahmin edilmesi gereken örneklerin oranlanıp ne kadarının doğru tahmin edildiğini gösterir. Duyarlılık formülü aşağıda gösterilmiştir (Coşkun ve Baykal, 2011).

$$\text{Duyarlılık} = \frac{TP}{TP + FN}$$

4.1.3.4. F1-score: F1 skoru duyarlılık ve kesinlik değerlerini birleştirip tek sayıya indirgemektedir. F1-Score formülü aşağıda gösterilmiştir (Coşkun ve Baykal, 2011).

$$F_1 - \text{Score} = \frac{2 \times P \times R}{P + R}$$

P: Kesinlik (Precision)

R: Duyarlılık (Recall)

F_1 skoru, elimizde birden çok modelin doğruluk ve kesinlik değerleri olduğunda değerler arasından en iyi modelin seçilmesi için kullanılır. F_1 skoru en çok 1 değerini alabilirken en düşük değer olarak da 0 değerini alabilmektedir.

4.1.4. Sınıflandırma algoritmalarının modellenmesi

Bu çalışmada 7.sınıf öğrencilerinin yılsonu başarılarının tahmin edilmesi amacıyla Random Forest Classifier (Rastgele Orman Sınıflandırıcısı), Gradient Boosting Classifier (Gradyan Artırma Sınıflandırıcısı), Support Vector Machine (Destek Vektör Makinesi) algoritmaları kullanılmıştır.

4.1.4.1. Random forest classifier (rastgele orman sınıflandırıcısı) modelinin performans ölçütü

Öncelikle One Hot Encoding yöntemi ile sayısal veriye dönüştürülen veri seti eğitim ve test seti olmak üzere rasgele olarak iki gruba ayrılmıştır.

Random Forest Model Eğitimi (kodları)

```
model = RandomForestClassifier(random_state=42)
model.fit(X_train, y_train)
```

Komutları ile Rastgele Orman algoritması ile makine öğrenmesi gerçekleştirilmiştir. Algoritmanın performans değerleri aşağıdaki komutlar yardımı ile hesaplanmıştır.

Random Forest Performansı Değerlendirmesi (kodları)

```
accuracy = accuracy_score(y_test, y_pred)
precision = precision_score(y_test, y_pred)
recall = recall_score(y_test, y_pred)
f1 = f1_score(y_test, y_pred)
```

Random Forest Performans Metriklerinin Gösterimi

```
print(f'Model Accuracy: {accuracy}')
print(f'Model Precision: {precision}')
print(f'Model Recall: {recall}')
print(f'Model F1-Score: {f1}')
```

Random Forest algoritması ile gerçekleştirilen makine öğrenmesi sonuçları Tablo 4.1’de verilmiştir.

Tablo 4.1: Random Forest Performans Sonuçları

Değerlendirme Metriği	Değer
Model Accuracy (Doğruluk):	0.8548
Model Precision (Kesinlik):	0.8780
Model Recall (Duyarlılık) :	0.9000
Model F1-Score (F ₁ Sonucu) :	0.8888

Random Forest algoritmasına göre başarıya etki eden değişkenlerin önem sırasına alınması ve sıralanması

```
feature_importances = model.feature_importances_  
sorted_feature_importance = sorted(feature_importance_dict.items(),  
key=lambda x: x[1], reverse=True)
```

Random Forest algoritması ile oluşturulan modelde öğrencilerin başarısına en çok etki eden ilk 10 değişken Tablo 4.2’ de listelenmiştir.

Tablo 4.2: Random Forest Başarıya Etki Eden Değişkenler

Sıra No	Değişken	Değer
1	Ödevlerin düzenli yapımı	0.0449
2	Üniversite okumayı isteme (hedef belirleme)	0.0383
3	Sınıfta derse konsantre olabilme	0.0279
4	Ders sırasında not alma	0.0226
5	Annenin eğitim durumu	0.0192
6	Öğrenmek için tercih edilen yöntem Her ikisi(Görsel/ Yazılı)	0.0188

Tablo 4.2 Devamı: Random Forest Başarıya Etki Eden Değişkenler

Sıra No	Değişken	Değer
7	Evinizin durumu (kendimize ait)	0.0162
8	Ailenin başarıya inancı (ailem başarılı olacağıma inanır)	0.0156
9	Anaokuluna gitme durumu (Evet)	0.0149
10	Sınav kaygısı (Orta)	0.0145

4.1.4.2. Gradient boosting classifier (gradyan artırma sınıflandırıcısı) modelinin performans ölçütü

One Hot Encoding yöntemi ile sayısal veriye dönüştürülen veri seti eğitim ve test seti olmak üzere rasgele olarak iki gruba ayrılmıştır.

Gradient Boosting Model Eğitimi

```
model = GradientBoostingClassifier(random_state=42)
model.fit(X_train, y_train)
```

Komutları ile Gradient Boosting algoritması ile makine öğrenmesi gerçekleştirilmiştir. Algoritmanın performans değerleri aşağıdaki komutlar yardımı ile hesaplanmıştır.

Gradient Boosting Performansı Değerlendirmesi

```
accuracy = accuracy_score(y_test, y_pred)
precision = precision_score(y_test, y_pred)
recall = recall_score(y_test, y_pred)
f1 = f1_score(y_test, y_pred)
```

Gradient Boosting Performans Metriklerinin Gösterimi

```
print(f'Model Doğruluğu: {accuracy}')
print(f'Model Precision: {precision}')
print(f'Model Recall: {recall}')
print(f'Model F1-Score: {f1}')
```

Gradient Boosting algoritması ile gerçekleştirilen makine öğrenmesi sonuçları Tablo 4.3’de verilmiştir.

Tablo 4.3: Gradient Boosting Performans Sonuçları

Değerlendirme Metriği	Değer
Model Accuracy (Doğruluk):	0.9032
Model Precision (Kesinlik):	0.9047
Model Recall (Duyarlılık) :	0.9500
Model F1-Score (F ₁ Sonucu) :	0.9268

Gradient Boosting algoritmasına göre başarıya etki eden değişkenlerin önem sırasına alınması ve sıralanması

```
feature_importances = model.feature_importances_  
sorted_feature_importance = sorted(feature_importance_dict.items(),  
key=lambda x: x[1], reverse=True)
```

Gradient Boosting algoritması ile oluşturulan modelde öğrencilerin başarısına en çok etki eden ilk 10 değişken Tablo 4.4’ de listelenmiştir.

Tablo 4.4: Gradient Boosting Başarıya Etki Eden Değişkenler

Sıra No	Değişken	Değer
1	Ödevlerin düzenli yapımı	0.1674
2	Annenin eğitim durumu	0.0555
3	Üniversite okumayı isteme (hedef belirleme)	0.0523
4	Ders sırasında not alma	0.0365
5	Anaokuluna gitme durumu (Evet)	0.0356
6	Yaş (13)	0.0269

Tablo 4.4 Devamı: Gradient Boosting Başarıya Etki Eden Değişkenler

Sıra No	Değişken	Değer
7	Öğrenmek için tercih edilen yöntem Her ikisi(Görsel/ Yazılı)	0.0263
8	Zihinsel olarak kendini nasıl hissediyorsun? (Çok iyi)	0.0256
9	Sınıfta derse odaklanabilme (Genellikle)	0.0234
10	Evinizin durumu (kira)	0.0227

4.1.4.3. Support vector machine (destek vektör makinesi) modelinin performans ölçütü

One Hot Encoding yöntemi ile sayısal veriye dönüştürülen veri seti eğitim ve test seti olmak üzere rasgele olarak iki gruba ayrılmıştır.

Support Vector Model Eğitimi

```
model = SVC(kernel='linear', random_state=42)
model.fit(X_train, y_train)
```

Komutları ile Support Vector algoritması ile makine öğrenmesi gerçekleştirilmiştir. Algoritmanın performans değerleri aşağıdaki komutlar yardımı ile hesaplanmıştır.

Support Vector Performansı Değerlendirmesi:

```
accuracy = accuracy_score(y_test, y_pred)
precision = precision_score(y_test, y_pred)
recall = recall_score(y_test, y_pred)
f1 = f1_score(y_test, y_pred)
```

Support Vector Performans metriklerinin gösterimi

```
print(f'Model Doğruluğu: {accuracy}')
print(f'Model Precision: {precision}')
print(f'Model Recall: {recall}')
print(f'Model F1-Score: {f1}')
```

Support Vector algoritması ile gerçekleştirilen makine öğrenmesi sonuçları Tablo 4.5’ de verilmiştir.

Tablo 4.5: Support Vector Performans Sonuçları

Değerlendirme Metriği	Değer
Model Accuracy (Doğruluk):	0.8225
Model Precision (Kesinlik):	0.8717
Model Recall (Duyarlılık) :	0.8500
Model F₁-Score (F ₁ Sonucu) :	0.8607

Support Vector Machine Ağırlık Büyüklüğü (Magnitude): Support Vector Machine algoritmasının modellenmesinde bir özelliğin ağırlığı ne kadar büyük olursa o özellik tahminde o kadar etkili olmaktadır. Özelliğin ağırlığı büyük olduğunda modelin tahminine daha fazla katkı sunmaktadır.

Support Vector Machine Ağırlık İşaretleri (Sign): Belirli bir özellik artınca tahminin artması ya da azalması hakkında bilgi vermektedir. Pozitif ağırlık özellik arttıkça tahminin artması anlamına gelirken; negatif ağırlık özellik arttıkça tahminin azalması anlamına gelmektedir.

Support Vector algoritması ağırlık büyüklüğünün hesaplanması ve sıralanması ilgili komutlar aşağıda verilmiştir.

```
if hasattr(model, 'coef_'):
    feature_weights = model.coef_.flatten()
    sorted_feature_weight = sorted(feature_weight_dict.items(), key=lambda x:
    x[1], reverse=True)
```

Özelliklerin ağırlıkları hesaplanarak tahmine en çok etki eden ilk 10 özellik ve etki oranları aşağıda Tablo 4.6’ de belirtilmiştir.

Tablo 4.6: Support Vector Özelliklerin Ağırlıkları

	Özellik	Etki Oranı
1	Okul dışında hafta içi ders çalışma süresi	1.0900
2	Kardeş sayısı	1.0802
3	Ödev yapımı	0.8625
4	Babanın eğitim durumu	0.8204
5	Ders sırasında not alma durumu	0.6988
6	Kardeş sayısı	0.6712
7	Okul dışında hafta sonu ders çalışma süresi	0.6435
8	Sınıfta derse konsantre olabilme	0.6421
9	Aynı evde yaşanan kişi sayısı	0.6364
10	Sınavlarda başarısız olma kaygısı	0.5676

4.2. Bulgular

4.2.1. Modellerin karşılaştırılması

Rastgele Orman, Gradyan Artırma Sınıflandırıcısı ve Destek Vektör Makinesi algoritmaları kullanılarak modeller ayrı ayrı eğitilmiş ve performans sonuçları aşağıdaki Tablo 4.7’ de açıklanmıştır.

Tablo 4.7: Modellerin Karşılaştırılması

Model	Doğruluk	Kesinlik	Duyarlılık	F ₁ Sonucu
RO	0.85	0.87	0.9	0.88
GAS	0.90	0.90	0.95	0.92
DVM	0.82	0.87	0.85	0.86

Tablo 4.7’ de görüldüğü gibi Gradyan Artırma Sınıflandırıcısı modeli en yüksek doğruluk, kesinlik, duyarlılık ve F₁ Score değerleri ile en başarılı model olmuştur.

Rastgele Orman ve Gradyan Artırma Sınıflandırıcısı modellerinde başarıya en çok etki eden ilk 10 değişken arasında Ödevlerin düzenli yapımı, Üniversite okumayı isteme (hedef belirleme), Sınıfta derse konsantre olabilme, Ders sırasında not alma, Annenin eğitim durumu, Öğrenmek için tercih edilen yöntem (Her ikisi(Görsel/ Yazılı)), Evinizin durumu (kendimize ait), Anaokuluna gitme durumu (Evet) değişkenlerinin ortak olduğu görülmektedir.



5. SONUÇ VE ÖNERİLER

5.1. Sonuç

Resmî veya özel tüm eğitim kurumlarının amacı öğrenci başarısını arttırmaktır. Öğrenci başarısının artması içinse başarıya etki eden faktörlerin tespit edilmesi ve risk grubunda yer alan öğrencilerin ihtiyaç duyduğu alanda desteklenmesi gerekmektedir. Ülkemizin de katıldığı PISA ve TIMSS gibi uluslararası sınavlarda öğrencilerimizin başarıları ortalamanın altında kalmaktadır. Bu sebeple öğrenci performansını iyileştirmeye yönelik daha fazla çalışmaya ihtiyaç duyulmaktadır. Bu çalışma velilere, öğrencilere ve öğretmenlere sınava girmeden önce akademik başarı projeksiyonu elde edebilme imkânı tanımaktadır.

Çalışmada kullanılmak üzere İstanbul Esenler’de resmî bir devlet okulunda 7.sınıfta eğitim gören öğrencilere bir anket uygulanarak veri seti elde edilmiştir. Bu veri seti başarı tahmini yapılabilmesi ve başarıya etki eden değişkenlerin belirlenebilmesi için kullanılmıştır. Anket yardımı ile öğrencilerin başarısına etki edebileceği düşünülen yaş, cinsiyet, anne babanın eğitim durumu, ödev yapımı, hafta içi ve hafta sonu öğrencilerin okul dışında ders çalışma süreleri gibi 31 farklı özellik elde edilmiştir. Anket sonucu elde edilen veri seti kategoriler içerdiğinden One Hot Encoding yöntemi kullanılarak sayısallaştırılmıştır. Encode edilen veri seti python programlama dili kullanılarak 3 farklı algoritma ile modellenerek makine öğrenmesi gerçekleştirilmiştir. Kullanılan algoritmalar Random Forest, Gradient Boosting, Support Vector Machine algoritmalarıdır. Modellerde bağımlı değişken olarak gerekli izinler alınarak e-okul üzerinden elde edilen öğrencilerin 6.sınıf yılsonu başarı ortalamaları kullanılmıştır. Gerçekleştirilen makine öğrenmesi sonucunda %90 doğruluk oranı ile en başarılı modelin Gradient Boosting algoritması kullanılarak oluşturulan model olduğu görülmüştür. En başarısız modelin ise %82 doğruluk oranı ile Support Vector Machine algoritması kullanılarak oluşturulan model olduğu görülmüştür. Random Forest algoritması kullanılarak oluşturulan model ile ise %85 doğruluk oranı sağlamıştır.

Modellerde ortak olarak başarıya en çok etki eden ortak değişkenlerin:

- Ödevlerin düzenli yapımı,
- Üniversite okumayı isteme(hedef belirleme),
- Sınıfta derse konsantre olabilme,
- Ders sırasında not alma,
- Annenin eğitim durumu,
- Öğrenmek için tercih edilen yöntem (Her ikisi(Görsel/ Yazılı)),
- Evinizin durumu (kendimize ait),
- Anaokuluna gitme durumu (Evet) olduğu görülmüştür.

Uygulanan modellere göre öğrenci başarısına en çok etki eden değişkenlerin çoğunun değiştirilebilir değişkenler olduğu görülmektedir. Bu çalışma sonunda aileler bilinçlendirilerek ve öğrencilere gerekli rehberlik çalışmaları yapılarak öğrencilerin başarılarının arttırılabileceği sonucuna ulaşılmıştır.

5.2. Öneriler

Bu tez çalışması sonucunda şu önerilerde bulunulabilir:

- Bu çalışmada başarıya etki eden değişkenlerin belirlenmesi ve başarı tahmininde bağımlı değişken olarak öğrencilerin bir önceki dönem yılsonu genel başarı ortalamaları kullanılmıştır. Fakat her dersin kazanımı ve çalışma metodu farklı olduğundan her ders için kazanım temelli başarı tahmini yaptırılabilir.
- Multi disiplinler yaklaşım benimsenerek bir derste performans başarısının diğer derslere etkisi makine öğrenmesi metodları ile araştırılarak kazanımlar çıkan sonuçlara göre düzenlenebilir.
- Başarıya etki eden değişkenlerin belirlenmesi ve eğitim öğretim dönemi öncesi başarı tahmini yapılması velilere ve öğretmenlere akademik başarıyı arttırabilmek için gerekli önlemlerin alınmasında rehber olabilir.
- Makine öğrenmesi modelleri oluşturularak öğrencilerin öğrenme stilleri tespit edilebilir ve öğrencilerin yeteneklerine uygun öğrenme ortamları tasarlanabilir.

- Özel yetenekli olan ve özel öğrenme güçlüğü yaşayan öğrencilerin yetenekleri makine öğrenmesi modelleri ile belirlenerek yeteneklerine uygun öğrenme platformları ve öğrenme materyalleri geliştirilebilir.
- Farklı makine öğrenmesi algoritmaları ve farklı veri setleri kullanılarak tahmin başarısı kıyaslanabilir.



KAYNAKÇA

- Abbasoğlu, B. (2020). Ortaokul öğrencilerinin akademik başarılarının eğitsel veri madenciliği yöntemleri ile tahmini. *Veri Bilimi*, 3(1), 1–10.
- Acar, T. (2012). Türkiye'nin PISA 2009 sonuçlarına göre OECD'ye üye ve aday ülkeler arasındaki yeri. *Kuram ve Uygulamada Eğitim Bilimleri*, 12(4), 2561–2572.
- Agarwal, S. (2013). Data mining: Data mining concepts and techniques. *2013 International Conference on Machine Intelligence and Research Advancement*, 203–207.
- Aghalarova, S., & Bozkurt Keser, S. (2021). Önerilen Yapay Sinir Ağı Algoritması ile Ortaokul Öğrencilerin Akademik Performansının Tahmini. *Veri Bilimi*, 4(2), 19–32.
- Alkan, A. (2023). *Akademik başarının eğitsel veri madenciliği ve makine öğrenmesi teknikleri kullanılarak tahminlenmesi*.
- Aydemir, B. (2017). *Veri madenciliği yöntemleri kullanarak meslek yüksek okulu öğrencilerinin akademik başarı tahmini*. (Tez No. 486838) [Yüksek lisans tezi, Pamukkale Üniversitesi]. Yükseköğretim Kurulu Ulusal Tez Merkezi.
- Baashar, Y., Hamed, Y., Alkaws, G., Fernando Capretz, L., Alhussian, H., Alwadain, A., & Al-amri, R. (2022). Evaluation of postgraduate academic performance using artificial intelligence models. *Alexandria Engineering Journal*, 61(12), 9867–9878. <https://doi.org/10.1016/j.aej.2022.03.021>
- Bastem, H. N. (2021). *Student academic performance prediction via artificial intelligence using machine learning algorithms*. (Tez No. 703366) [Yüksek lisans tezi, Çankaya Üniversitesi]. Yükseköğretim Kurulu Ulusal Tez Merkezi.
- Bingül, F. (2023). *Yapay zekâ destekli makine öğrenmesi algoritmaları ile eğitsel veri madenciliği alanında öğrenci akademik performanslarının tahmin edilmesi*. (Tez No.784774) [Yüksek lisans tezi, İstanbul Gedik Üniversitesi]. Yükseköğretim Kurulu Ulusal Tez Merkezi.
- Bujang, S. D. A., Selamat, A., & Krejcar, O. (2021). A predictive analytics model for students grade prediction by supervised machine learning. *IOP Conference Series: Materials Science and Engineering*, 1051(1), 1051. <https://doi.org/10.1088/1757-899X/1051/1/012005>
- Buraimoh, E., Ajoodha, R., & Padayachee, K. (2021). Application of machine learning techniques to the prediction of student success. *2021 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS)*, 1–6.

- Çelik, E. F. (2022). *Makine öğrenmesi ile lise öğrencilerine mesleki alan seçim rehberi*. (Tez No. 732376) [Yüksek lisans tezi, Manisa Celal Bayar Üniversitesi]. Yükseköğretim Kurulu Ulusal Tez Merkezi.
- Christou, V., Tsoulos, I., Loupas, V., Tzallas, A. T., Gogos, C., Karvelis, P. S., Antoniadis, N., Glavas, E., & Giannakeas, N. (2023). Performance and early drop prediction for higher education students using machine learning. *Expert Systems with Applications*, 225, 120079.
- Civelek, S. (2023). *Makine öğrenmesi yöntemleri ile öğrenci mezuniyet notu tahmini*. (Tez No. 809753) [Yüksek lisans tezi, Bolu Abant İzzet Baysal Üniversitesi]. Yükseköğretim Kurulu Ulusal Tez Merkezi.
- Coşkun, C., & Baykal, A. (2011). Veri madenciliğinde sınıflandırma algoritmalarının bir örnek üzerinde karşılaştırılması. *Akademik Bilişim*, 2011, 1–8.
- Duralioğlu, Ö. (2022). *Makine öğrenimi ve istatistiksel yöntemler kullanarak öğrencilerin programlama dersindeki kazanım bilgileri ile başarı tahmini*. (Tez No. 766340) [Yüksek lisans tezi, Sakarya Üniversitesi]. Yükseköğretim Kurulu Ulusal Tez Merkezi.
- Dursun, Ş., & Yüksel, D. (2004). Öğrencilerin matematikte başarısını etkileyen faktörler matematik öğretmenlerinin görüşleri bakımından. *Gazi Üniversitesi Gazi Eğitim Fakültesi Dergisi*, 24(2), 217–230.
- Fırat, İ. (2019). *Türkiye’de matematik okuryazarlık ile ilgili 2020 yılına kadar yapılan çalışmaların doküman analizi yöntemiyle incelenmesi* [Yüksek Lisans Tezi]. [YÖK Ulusal Tez Merkezi].
- Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, 1189–1232.
- Gök, M. (2017). Makine öğrenmesi yöntemleri ile akademik başarının tahmin edilmesi. *Gazi University Journal Of Science Part C: Design And Technology*, 5(3), 139–148.
- Güven, B., & Cabakcor, B. O. (2013). Factors influencing mathematical problem-solving achievement of seventh grade Turkish students. *Learning and Individual Differences*, 23, 131–137. <https://doi.org/10.1016/j.lindif.2012.10.003>
- Haridas, M., Gutjahr, G., Raman, R., Ramaraju, R., & Nedungadi, P. (2020). Predicting school performance and early risk of failure from an intelligent tutoring system. *Education and Information Technologies*, 25, 3995–4013.
- Hashim, A. S., Awadh, W. A., & Hamoud, A. K. (2020). Student performance prediction model based on supervised machine learning algorithms. *IOP Conference Series: Materials Science and Engineering*, 928(3), 032019. <https://doi.org/10.1088/1757-899X/928/3/032019>
- Hasib, K. M., Rahman, F., Hasnat, R., & Alam, M. G. R. (2022). A machine learning and explainable ai approach for predicting secondary school student performance. *2022 IEEE 12th Annual Computing and*

- Communication Workshop and Conference (CCWC)*, 399-405).
<https://doi.org/10.1109/CCWC54503.2022.9720806>
- Hearst, M. A., Dumais, S. T., Osuna, E., Platt, J., & Scholkopf, B. (1998). Support vector machines. *IEEE Intelligent Systems and Their Applications*, 13(4), 18–28.
- Hussain, S., & Khan, M. Q. (2023). Student-performulator: Predicting students' academic performance at secondary and intermediate level using machine learning. *Annals of Data Science*, 10(3), 637–655.
<https://doi.org/10.1007/s40745-021-00341-0>
- Irmak, H., & Ecevit Satı, Z. (2022). Covid-19 pandemisi döneminde öğrencilerin uzaktan eğitimde yaşadıkları sorunların kelime bulutu ve gizli anlamsal dizinleme yöntemleriyle tespit edilmesi. *Uluslararası Yönetim Bilişim Sistemleri ve Bilgisayar Bilimleri Dergisi*, 6(2), 162–170.
<https://doi.org/10.33461/uybisbbd.1214562>
- Kanchana, J. D., Amarasinghe, G., Nanayakkara, V., & Perera, A. S. (2021). A data mining approach for early prediction of academic performance of students. *2021 IEEE International Conference on Engineering, Technology & Education (TALE)*, 1–8.
<https://doi.org/10.1109/TALE52509.2021.9678558>
- Kavzoğlu, T., & Çölkesen, İ. (2010). Destek vektör makineleri ile uydu görüntülerinin sınıflandırılmasında kernel fonksiyonlarının etkilerinin incelenmesi. *Harita Dergisi*, 144(7), 73–82.
- Kaya, F. H. (2022). *Makine öğrenmesi algoritmaları kullanarak öğrencilerin akademik başarısını etkileyen faktörlerin tespit edilmesi*.
- Koç, T., & Akın, P. (2022). Estimation of high school entrance examination success rates using machine learning and beta regression models. *Journal of Intelligent Systems: Theory and Applications*, 5(1), 9–15.
<https://doi.org/10.38016/jista.922663>
- Kye, A. (2023). *Comparative analysis of classification performance for u.s. college enrollment predictive modeling using four machine learning algorithms (logistic regression, decision tree, support vector machine, artificial neural network)*. (Publication No. 10169573) [Doctoral dissertation, Loyola University Chicago]. ProQuest Dissertations & Theses Global.
- Lau, E. T., Sun, L., & Yang, Q. (2019). Modelling, prediction and classification of student academic performance using artificial neural networks. *SN Applied Sciences*, 1, 1–10.
- Mahmut, Ö., Gençoğlu, C., & Eren, S. (2020). Türkiye’de eğitimde eşitsizlikleri azaltmak için uygulanan politikalar. *Ondokuz Mayıs University Journal of Education Faculty*, 39(2), 294–312.
- MEB, M. E. B. (2014). *Millî Eğitim Bakanlığı okul öncesi eğitim ve ilköğretim kurumları yönetmeliği*.

- Miguéis, V. L., Freitas, A., Garcia, P. J. V., & Silva, A. (2018). Early segmentation of students according to their academic performance: A predictive modelling approach. *Decision Support Systems*, 115, 36–51.
- Natekin, A., & Knoll, A. (2013). Gradient boosting machines, a tutorial. *Frontiers in Neurorobotics*, 7, 21.
- Sarier, Y. (2021). Pisa Uygulamalarında Türkiye'nin Performansı ve Öğrenci Başarısını Yordayan Değişkenler. *Türkiye Sosyal Araştırmalar Dergisi*, 25(3), 905–926.
- Selvi, A. (2020). *Bilecik ilinde ilköğretimden liseye geçiş sınavlarında makine öğrenmesi yöntemleri ile öğrenci başarısının tahmini*. (Tez No. 637536) [Yüksek lisans tezi, Bilecik Şeyh Edebali Üniversitesi]. Yükseköğretim Kurulu Ulusal Tez Merkezi.
- Sönmez Sarıkaya, B. (2020). AES Algoritmasına Yapılan Zaman Odaklı Önbellek Saldırılarının Makine Öğrenmesi ile Tespiti. *Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi*, 13(1), 57–68.
- Speiser, J. L., Miller, M. E., Tooze, J., & Ip, E. (2019). A comparison of random forest variable selection methods for classification prediction modeling. *Expert Systems with Applications*, 134, 93–101.
- TDK. (2023). “Eğitim”, “Güncel Türkçe Sözlük”, *Türk Dil Kurumu*. Erişim: 20 Eylül 2023.
- Tomasevic, N., Gvozdenovic, N., & Vranes, S. (2020). An overview and comparison of supervised data mining techniques for student exam performance prediction. *Computers & Education*, 143, 103676.
- Ulutun, E. (2018). *TEOG fen bilgisi başarısını etkileyen değişkenlerin çok düzeyli regresyon modeli ile incelenmesi*. (Tez No. 517251) [Yüksek lisans tezi, Hacettepe Üniversitesi]. Yükseköğretim Kurulu Ulusal Tez Merkezi.
- Wakelam, E., Jefferies, A., Davey, N., & Sun, Y. (2020). The potential for student performance prediction in small cohorts with minimal available attributes. *British Journal of Educational Technology*, 51(2), 347–370. <https://doi.org/10.1111/bjet.12836>
- Yağcı, M. (2022). Educational data mining: prediction of students' academic performance using machine learning algorithms. *Smart Learning Environments*, 9(1), 11.
- Yavuzarslan, M., & Erol, Ç. (2022). Öğrenme yönetim sistemi log kayıtlarının akademik başarı tahmininde kullanılması. *Bilişim Teknolojileri Dergisi*, 15(2), 199–207. <https://doi.org/10.17671/gazibtd.837884>
- Zeineddine, H., Braendle, U., & Farah, A. (2021). Enhancing prediction of student success: Automated machine learning approach. *Computers & Electrical Engineering*, 89, 106903. <https://doi.org/10.1016/j.compeleceng.2020.106903>
- Zilyas, D. (2023). *Makine öğrenmesi yöntemi ile eğitim başarısının tahmini modeli*. (Tez No. 808254) [Yüksek lisans tezi, İstanbul Beykent Üniversitesi]. Yükseköğretim Kurulu Ulusal Tez Merkezi.

ÖZGEÇMİŞ

Zeynep Gökmen

Eğitim Derecesi	Eğitim Kurumu	Mezuniyet
Lisans	Gaziosmanpaşa Üniversitesi / Bilgisayar ve Öğretim Teknolojileri Öğretmenliği	2014
Lise	Gerze Teknik Lise	2009

İş Deneyimi

Kurum	Yıl
Milli Eğitim Bakanlığı	2014 - 2015
Şafak Okulları	2017 – 2019
Milli Eğitim Bakanlığı	2019 - Halen

Daha Önce Yapılan Yayınlar

Yayın	Konferans	Sayfalar
Yapay Zekâ Uygulamalarının Eğitime Etkileri Üzerine Bir İnceleme	8th International Aegean Congress on Innovation Technologies & Engineering	122-139