



**ÜSTÜN YETENEKLİ ÖĞRENCİLERİN MATEMATİK BAŞARISINI ETKİLEYEN  
FAKTÖRLERİN MAKİNE ÖĞRENMESİ MODELLERİ İLE TAHMİNİ VE  
KARŞILAŞTIRMALI ANALİZİ**

(Yüksek Lisans)

**Kübra ÖNGENLİ**

Kütahya- 2025

T.C.  
KÜTAHYA DUMLUPINAR ÜNİVERSİTESİ  
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ  
Bilgisayar Mühendisliği Anabilim Dalı

Yüksek Lisans Tezi

**ÜSTÜN YETENEKLİ ÖĞRENCİLERİN MATEMATİK  
BAŞARISINI ETKİLEYEN FAKTÖRLERİN MAKİNE  
ÖĞRENMESİ MODELLERİ İLE TAHMİNİ VE  
KARŞILAŞTIRMALI ANALİZİ**

Danışman:  
Dr. Öğr. Üyesi Çiğdem BAKIR

Hazırlayan:  
Kübra ÖNGENLİ

Kütahya-2025

**Kabul ve Onay****KÜTAHYA DÜMLUPINAR ÜNİVERSİTESİ****LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ**

Bilgisayar Mühendisliği Anabilim Dalında, 202285151010 öğrenci numaralı, Kübra ÖNGENLİ'nin hazırlamış olduğu “*ÜSTÜN YETENEKLİ ÖĞRENCİLERİN MATEMATİK BAŞARISINI ETKİLEYEN FAKTÖRLERİN MAKİNE ÖĞRENMESİ MODELLERİ İLE TAHMİNİ VE KARŞILAŞTIRMALI ANALİZİ*” başlıklı yüksek lisans tez çalışması ile ilgili tez savunma sınavı jüri tarafından yapılmış ve adayın tezinin OY BİRLİĞİ ile kabul edilmesine karar verilmiştir

20/06/2025

| Tez Jürisi                                                                 | İmza  |     |
|----------------------------------------------------------------------------|-------|-----|
|                                                                            | Kabul | Ret |
| Dr. Öğr. Üyesi Çiğdem BAKIR (Danışman)<br>DPÜ, Bilgisayar Mühendisliği ABD |       |     |
| Dr. Öğr. Üyesi Soydan SERTTAŞ<br>DPÜ, Bilgisayar Mühendisliği ABD          |       |     |
| Dr. Öğr. Üyesi Ahmet GÜVEN<br>KSBÜ, Bilgisayar Mühendisliği ABD            |       |     |

**Onay**

Doç. Dr. Eray Acar

Enstitü Müdürü

## **Bilimsel Etik Bildirimi**

Yüksek Lisans Tezi olarak hazırladığım “*Üstün Yetenekli Öğrencilerin Matematik Başarısını Etkileyen Faktörlerin Makine Öğrenmesi Modelleri İle Tahmini Ve Karşılaştırmalı Analizi*” adlı çalışmanın öneri aşamasından sonuçlandığı aşamayakadar geçen süreçte bilimsel etiğe ve akademik kurallara özenle uyduğumu, tez içindeki tüm bilgileri bilimsel ahlak ve gelenek çerçevesinde elde ettiğimi, tez yazım kurallarına uygun olarak hazırladığımı, bu çalışmamda doğrudan veya dolaylı olarak yaptığım her alıntıya kaynak gösterdiğimi ve yararlandığım eserlerin kaynakçada gösterilenlerden oluştuğunu beyan ederim.

20/06/2025

Kübra Öngenli

## Öz Geçmiş

Kübra Öngenli, ilk ve orta öğrenimini Malatya’da tamamlamıştır. 2008 yılında Ankara Üniversitesi Yabancı Diller Yüksekokulu’nda İngilizce Hazırlık, 2015 yılında Erzincan Üniversitesi İlköğretim Matematik Öğretmenliği Bölümünü bitirmiştir. 2018 yılında Ahmet Yesevi Üniversitesi Eğitim Yönetimi ve Denetimi, 2019 yılında Gelişim Üniversitesi İş Sağlığı ve Güvenliği bölümlerinde yüksek lisans eğitimi almıştır. 2020 yılında Atatürk Üniversitesi İş Sağlığı ve Güvenliği ön lisans bölümünü, 2022 yılında Anadolu Üniversitesi Yönetim Bilişim Sistemleri, 2023 yılında İstanbul Üniversitesi İnsan Kaynakları Yönetimi, 2024 yılında Atatürk Üniversitesi Egzersiz ve Spor Bilimleri lisans bölümünü tamamlamıştır. 2016-2018 yılları arasında Siirt Eruh İMKB Yatılı Bölge Ortaokulu’nda, 2018-2019 yılları arasında Siirt Eruh Mehmet Akif Ersoy Ortaokulu’nda öğretmen olarak görev almıştır. 2019-2021 yılları arasında Siirt Eruh İmam Hatip Ortaokulu’nda Müdür Yardımcısı olarak çalışmıştır. 2021-2023 tarihlerinde Bursa İnegöl Hocaköy Ortaokulu’nda ve 2023 yılında Bursa Osmangazi Bilim ve Sanat Merkezi’nde öğretmenlik yapmıştır. 2023 yılından bu yana Bursa Gemlik Bilim ve Sanat Merkezi’nde müdür olarak görev yapmaktadır.

## ÖZET

### ÜSTÜN YETENEKLİ ÖĞRENCİLERİN MATEMATİK BAŞARISINI ETKİLEYEN FAKTÖRLERİN MAKİNE ÖĞRENMESİ MODELLERİ İLE TAHMİNİ VE KARŞILAŞTIRMALI ANALİZİ

**ÖNGENLİ, Kübra**

**Yüksek Lisans Tezi, Bilgisayar Mühendisliği Anabilim Dalı**

**Tez Danışmanı: Dr. Öğr. Üyesi Çiğdem BAKIR**

**Haziran, 2025, 133 sayfa**

Bu çalışma, Bursa ilinde Bilim ve Sanat Merkezleri'nde (BİLSEM) eğitim gören üstün yetenekli öğrencilerin matematik başarılarını etkileyen faktörleri belirlemeyi ve bu başarıyı çeşitli makine öğrenmesi modelleri kullanarak tahmin etmeyi amaçlamaktadır. Araştırmada 151 öğrenciden toplanan veriler doğrultusunda, demografik bilgiler, aile yapısı, çalışma alışkanlıkları, motivasyon düzeyi, teknoloji kullanımı ve sosyal etkinlikler gibi değişkenler analiz edilmiştir. Eğitimsel verimadenciliği ve öğrenme analitiği alanlarından yararlanılarak karar ağaçları, destek vektör makineleri ve yapay sinir ağları gibi yöntemler kullanılmıştır. Elde edilen sonuçlar, bazı değişkenlerin öğrencilerin matematik başarısını anlamlı şekilde etkilediğini göstermiştir. Çalışma, eğitim politikaları ve bireyselleştirilmiş öğretim stratejilerinin geliştirilmesi açısından önemli bulgular sunmaktadır.

**Anahtar Kelimeler:** BİLSEM, Eğitimsel Veri Madenciliği, Makine Öğrenmesi, Matematik Başarısı, Üstün Yeteneklilik

**ABSTRACT****PREDICTION AND COMPARATIVE ANALYSIS OF FACTORS  
AFFECTING THE MATHEMATICS ACHIEVEMENT OF GIFTED  
STUDENTS USING MACHINE LEARNING MODELS****ÖNGENLİ, Kübra****Master Thesis Department of Computer Engineering****Supervisor: Dr. Öğr. Üyesi Çiğdem BAKIR****June, 2025, 133 pages**

This study aims to identify the factors affecting the mathematics achievement of gifted students studying at Science and Art Centers (BİLSEM) in Bursa and to predict this achievement using various machine learning models. Based on data collected from 151 students, variables such as demographic information, family structure, study habits, motivation level, use of technology, and social activities were analyzed. Techniques from the fields of educational data mining and learning analytics, including decision trees, support vector machines, and artificial neural networks, were employed. The results showed that certain variables significantly influence students' mathematics performance. The study provides important findings for the development of educational policies and personalized teaching strategies.

**Keywords:** BİLSEM, Educational Data Mining, Machine Learning, Mathematics Achievement, Giftedness.

## ÖN SÖZ

Bu yüksek lisans tez çalışmasının her aşamasında bilgi ve deneyimiyle yol gösteren, akademik rehberliği ve içten desteğiyle beni daima motive eden değerli tez danışmanım Dr. Öğr. Üyesi Çiğdem BAKIR’a en içten teşekkürlerimi sunuyorum.

Ayrıca, araştırmamın şekillenmesine önemli katkılarda bulunan Dr. Öğr. Üyesi Soydan SERTTAŞ ve Dr. Öğr. Üyesi Ahmet GÜVEN hocalarıma; sahip oldukları akademik birikimle sundukları bilimsel yaklaşımlar, yapıcı değerlendirmeler ve yönlendirici katkılarıyla çalışmama değer kattıkları için teşekkür ederim.

Görev yaptığım dönemde eğitime ve bilimsel çalışmalara olan duyarlılıkları, çözüm odaklı yaklaşımları ve verdikleri samimi destekle sürecimi kolaylaştıran önceki Gemlik Kaymakamlarımızdan, Erdek Kaymakamı olarak görevine devam eden Sayın Hasan GÖÇ’e ve Kırklareli Valisi olarak görev yapan Sayın Uğur TURAN’a içten şükranlarımı sunuyorum.

Aynı şekilde, görev süresi boyunca çalışmalarımı yakından takip eden, her zaman yanımda olduğunu hissettiren, motivasyon sağlayan yaklaşımı ve takdir dolu değerlendirmeleriyle desteğini esirgemeyen şu anki Gemlik Kaymakamı Sayın Osman Aslan CANBABA’ya içtenlikle teşekkür ederim.

Eğitim sürecim ve saha çalışmaları boyunca samimi ilgisiyle akademik çabalarımın yerelde karşılık bulmasına öncülük eden, mesleki ve insani anlamda her zaman bir ağabey gibi yanımda olan Gürsu Bilim ve Sanat Merkezi Müdürü Yaşar DİLBER’e gönülden teşekkür ederim.

Sadece resmi süreçlerde değil, her zaman yapıcı tutumuyla yanımda olan, bilimsel çalışmalarında beni yüreklendiren ve bu süreci daha kararlı ve inançla sürdürmemi sağlayan Gemlik Kaymakamlığı Yazı İşleri Müdürü Sayın Uğur ŞANLI’ya da ayrıca teşekkür ederim.

Bu süreçte manevi desteğini her zaman yanımda hissettiğim, sabırları ve sevgileriyle bana güç veren aileme ve değerli arkadaşlarıma; özellikle mesai arkadaşlarım Başöğretmen Osman METİN, Başöğretmen Mehmet KESKİN’e ve Psikolojik Danışman ve Rehber Öğretmen Derya AKÇELİK YILDIRIM’a da yürekten teşekkür ederim.

## İÇİNDEKİLER

|                              |                       |
|------------------------------|-----------------------|
| <b>ÖZET.....</b>             | <b><u>Sayfa</u> v</b> |
| <b>ABSTRACT .....</b>        | <b>vi</b>             |
| <b>ÖN SÖZ.....</b>           | <b>vii</b>            |
| <b>İÇİNDEKİLER .....</b>     | <b>viii</b>           |
| <b>TABLolar LİSTESİ.....</b> | <b>xii</b>            |
| <b>ŞEKİLLER LİSTESİ.....</b> | <b>xiii</b>           |
| <b>KISALTMALAR .....</b>     | <b>xiv</b>            |
| <b>GİRİŞ .....</b>           | <b>1</b>              |

## BİRİNCİ BÖLÜM

### LİTERATÜR TARAMASI

|                                                                               |           |
|-------------------------------------------------------------------------------|-----------|
| <b>1.1. ÜSTÜN YETENEKLİ ÖĞRENCİLER VE EĞİTİMLERİ .....</b>                    | <b>8</b>  |
| 1.1.1. Üstün Yetenekli Öğrencilerin Özellikleri .....                         | 8         |
| 1.1.2. Türkiye'de BİLSEM Yapısı ve Eğitim Programları .....                   | 9         |
| <b>1.2. MATEMATİK BAŞARISINI ETKİLEYEN FAKTÖRLER .....</b>                    | <b>10</b> |
| 1.2.1. Bireysel Faktörler .....                                               | 10        |
| 1.2.2. Ailesel Faktörler .....                                                | 12        |
| 1.2.3. Eğitimsel Faktörler .....                                              | 13        |
| 1.2.4. Sosyo-Ekonomik Faktörler .....                                         | 14        |
| <b>1.3. VERİ BİLİMİ VE EĞİTİM .....</b>                                       | <b>15</b> |
| 1.3.1. Eğitimde Veri Madenciliği .....                                        | 15        |
| 1.3.2. Eğitimsel Veri Analizi Alanındaki Çalışmalar .....                     | 15        |
| 1.3.3. Öğrenci Başarısını Tahmin Eden Çalışmalar: .....                       | 15        |
| 1.3.4. Eğitimsel Veri Madenciliğinde Topluluk Öğrenme Yöntemleri: .....       | 16        |
| 1.3.5. Üstün Yetenekli Öğrencilere Yönelik Veri Analizi Çalışmaları: .....    | 16        |
| <b>1.4. MAKİNE ÖĞRENMESİ VE EĞİTİM VERİSİ ANALİZİNDEKİ UYGULAMALARI .....</b> | <b>17</b> |
| 1.4.1. Sınıflandırma Algoritmaları.....                                       | 17        |
| 1.4.2. Regresyon Algoritmaları.....                                           | 19        |
| 1.4.3. Topluluk Öğrenme Yöntemleri.....                                       | 20        |
| 1.4.4. Derin Öğrenme Yaklaşımları.....                                        | 21        |

## İKİNCİ BÖLÜM

### YÖNTEM

|                                                      |           |
|------------------------------------------------------|-----------|
| <b>2.1. ARAŞTIRMA MODELİ .....</b>                   | <b>25</b> |
| <b>2.2. ÇALIŞMA GRUBU .....</b>                      | <b>26</b> |
| 2.2.1. Bursa BİLSEM Öğrencileri ve Özellikleri ..... | 27        |
| 2.2.2. Örneklem Seçimi ve Demografik Dağılım .....   | 27        |
| <b>2.3. VERİ TOPLAMA ARAÇLARI.....</b>               | <b>29</b> |
| 2.3.1. Anket Yapısı ve Uygulanması .....             | 29        |
| 2.3.2. Anket Uygulama Süreci.....                    | 32        |
| 2.3.3. Veri Seti Özellikleri .....                   | 32        |
| <b>2.4. VERİ ÖN İŞLEME .....</b>                     | <b>33</b> |
| 2.4.1. Eksik Veri Analizi .....                      | 33        |
| 2.4.2. Aykırı Değer Tespiti .....                    | 34        |
| 2.4.3. Kategorik Verilerin Sayısallaştırılması.....  | 36        |
| 2.4.4. Özellik Normalizasyonu .....                  | 37        |
| <b>2.5. ÖZELLİK MÜHENDİSLİĞİ .....</b>               | <b>38</b> |
| 2.5.1. Özellik Çıkarımı .....                        | 39        |
| 2.5.2. Özellik Seçimi.....                           | 40        |
| 2.5.3. Boyut İndirgeme Teknikleri.....               | 41        |
| <b>2.6. MAKİNE ÖĞRENMESİ MODELLERİ .....</b>         | <b>42</b> |
| 2.6.1. Lojistik Regresyon.....                       | 42        |
| 2.6.2. Karar Ağaçları.....                           | 43        |
| 2.6.3. Destek Vektör Makineleri (SVM) .....          | 44        |
| 2.6.4. Rastgele Orman (Random Forest) .....          | 45        |
| 2.6.5. Gradyan Artırma (Gradient Boosting).....      | 46        |
| 2.6.6. Derin Sinir Ağları .....                      | 47        |
| <b>2.7. MODEL DEĞERLENDİRME VE DOĞRULAMA.....</b>    | <b>48</b> |
| 2.7.1. Çapraz Doğrulama .....                        | 48        |
| 2.7.2. Değerlendirme Metrikleri .....                | 49        |
| 2.7.3. Hiperparametre Optimizasyonu .....            | 51        |

## ÜÇÜNCÜ BÖLÜM

### BULGULAR

|                                                          |           |
|----------------------------------------------------------|-----------|
| <b>3.1. TANIMLAYICI İSTATİSTİKLER VE VERİ KEŞFİ.....</b> | <b>53</b> |
|----------------------------------------------------------|-----------|

|                                                                                 |           |
|---------------------------------------------------------------------------------|-----------|
| 3.1.1. Demografik Özellikler .....                                              | 53        |
| 3.1.2. Matematik Başarısı Dağılımları .....                                     | 54        |
| 3.1.3. Değişkenler Arası Korelasyon Analizi .....                               | 57        |
| <b>3.2. ÖZELLİK ÖNEM ANALİZİ SONUÇLARI.....</b>                                 | <b>59</b> |
| 3.2.1. LASSO Regresyon Sonuçları .....                                          | 59        |
| 3.2.2. Rastgele Orman Özellik Önem Analizi .....                                | 60        |
| 3.2.3. Permütasyon Önem Analizleri.....                                         | 61        |
| <b>3.3. MAKİNE ÖĞRENMESİ MODELLERİNİN TAHMİN PERFORMANSLARI</b><br><b>.....</b> | <b>63</b> |
| 3.3.1. Sınıflandırma Modelleri Sonuçları .....                                  | 63        |
| 3.3.2. Regresyon Modelleri Sonuçları .....                                      | 64        |
| 3.3.3. Topluluk Yöntemleri Sonuçları .....                                      | 66        |
| 3.3.4. Derin Öğrenme Modeli Sonuçları.....                                      | 67        |
| <b>3.4. MODELLERİN KARŞILAŞTIRMALI ANALİZİ .....</b>                            | <b>69</b> |
| 3.4.1. Performans Karşılaştırmaları .....                                       | 69        |
| 3.4.2. Model Karmaşıklığı ve Performans İlişkisi .....                          | 72        |
| 3.4.3. Özellik Sayısı ve Model Performansı .....                                | 74        |
| <b>DÖRDÜNCÜ BÖLÜM</b>                                                           |           |
| <b>TARTIŞMA VE SONUÇ</b>                                                        |           |
| <b>4.1. BULGULARA İLİŞKİN TARTIŞMA .....</b>                                    | <b>78</b> |
| 4.1.1. Matematik Başarısını Etkileyen Faktörlerin Değerlendirilmesi .....       | 78        |
| 4.1.2. Makine Öğrenmesi Modellerinin Yorumlanması .....                         | 79        |
| 4.1.3. Algoritma Performanslarının Değerlendirilmesi .....                      | 80        |
| <b>4.2. EĞİTİMSEL ÇIKARIMLAR.....</b>                                           | <b>81</b> |
| 4.2.1. BİLSEM Eğitim Programları için Öneriler .....                            | 81        |
| 4.2.2. Üstün Yetenekli Öğrenci Eğitiminde Veri Odaklı Yaklaşımlar .....         | 83        |
| 4.2.3. Ebeveynler için Öneriler .....                                           | 84        |
| <b>4.3. TEKNİK ÇIKARIMLAR.....</b>                                              | <b>85</b> |
| 4.3.1. Eğitim Verisi Analizinde Makine Öğrenmesi Yaklaşımları.....              | 86        |
| 4.3.2. Model Seçimi ve Hiperparametre Optimizasyonu Stratejileri.....           | 87        |
| <b>4.4. ÇALIŞMANIN SINIRLILIKLARI.....</b>                                      | <b>88</b> |
| <b>4.5. GELECEK ÇALIŞMALAR İÇİN ÖNERİLER .....</b>                              | <b>90</b> |
| <b>4.6. SONUÇ.....</b>                                                          | <b>92</b> |

|                       |            |
|-----------------------|------------|
| <b>EKLER.....</b>     | <b>94</b>  |
| <b>KAYNAKÇA .....</b> | <b>107</b> |
| <b>DİZİN .....</b>    | <b>116</b> |



## TABLolar LİSTESİ

### Sayfa

|                                                                                                        |    |
|--------------------------------------------------------------------------------------------------------|----|
| <b>Tablo 2.1:</b> Çalışma Grubunun Demografik Özellikleri.....                                         | 28 |
| <b>Tablo 2.2:</b> Öğrencilerin Matematik Başarı Düzeylerine Göre Dağılımı .....                        | 29 |
| <b>Tablo 3.1:</b> Demografik Değişkenlerin Frekans ve Yüzde Dağılımları .....                          | 53 |
| <b>Tablo 3.2:</b> Öğrencilerin Matematik Not Ortalamalarına Göre Dağılımı.....                         | 55 |
| <b>Tablo 3.3:</b> Matematik Başarısının Çalışma Alışkanlıkları ve Tutumlarla İlişkisi.....             | 56 |
| <b>Tablo 3.4:</b> Sayısal Değişkenler Arasındaki Korelasyon Katsayıları .....                          | 58 |
| <b>Tablo 3.5:</b> LASSO Regresyon Sonucunda Elde Edilen Özellik Katsayıları ve Önem Sıralamaları ..... | 59 |
| <b>Tablo 3.6:</b> Farklı Modellerde Permütasyon Önem Analizi Sonuçları.....                            | 62 |
| <b>Tablo 3.7:</b> Sınıflandırma Modellerinin Performans Metrikleri .....                               | 63 |
| <b>Tablo 3.8:</b> Regresyon Modellerinin Performans Metrikleri .....                                   | 65 |
| <b>Tablo 3.9:</b> Topluluk Öğrenme Modellerinin Sınıflandırma Performans Metrikleri .....              | 66 |
| <b>Tablo 3.10:</b> Topluluk Öğrenme Modellerinin Regresyon Performans Metrikleri .....                 | 67 |
| <b>Tablo 3.11:</b> Derin Öğrenme Modellerinin Sınıflandırma Performans Metrikleri.....                 | 68 |
| <b>Tablo 3.12:</b> Derin Öğrenme Modellerinin Regresyon Performans Metrikleri.....                     | 68 |
| <b>Tablo 3.13:</b> Sınıflandırma Modellerinin Doğruluk ve ROC AUC Karşılaştırması.....                 | 69 |
| <b>Tablo 3.14:</b> Regresyon Modellerinin RMSE ve $R^2$ Karşılaştırması.....                           | 70 |
| <b>Tablo 3.15:</b> Seçilen Modeller Arasındaki İstatistiksel Karşılaştırma (p değerleri).....          | 72 |

## ŞEKİLLER LİSTESİ

|                                                                                                               | <b><u>Sayfa</u></b> |
|---------------------------------------------------------------------------------------------------------------|---------------------|
| <b>Şekil 3.1:</b> Öğrencilerin Cinsiyet ve Yaş Dağılımı .....                                                 | 54                  |
| <b>Şekil 3.2:</b> Matematik Başarı Düzeylerinin Cinsiyet ve Sınıf Düzeyine Göre Dağılımı .....                | 55                  |
| <b>Şekil 3.3:</b> Rastgele Orman Modeli Özellik Önem Değerleri .....                                          | 61                  |
| <b>Şekil 3.4:</b> Sınıflandırma Modellerinin ROC Eğrileri .....                                               | 64                  |
| <b>Şekil 3.5:</b> Gerçek ve Tahmin Edilen Not Ortalamaları Arasındaki İlişki (Destek Vektör Regresyonu) ..... | 65                  |
| <b>Şekil 3.6:</b> Model Kategorilerine Göre Doğruluk Değerleri .....                                          | 70                  |
| <b>Şekil 3.7:</b> Model Kategorilerine Göre $R^2$ Değerleri .....                                             | 71                  |
| <b>Şekil 3.8:</b> Model Karmaşıklığı ve Doğruluk Arasındaki İlişki .....                                      | 73                  |
| <b>Şekil 3.9:</b> XGBoost Modeli Hiperparametreleri ve Doğruluk İlişkisi .....                                | 74                  |
| <b>Şekil 3.10:</b> Özellik Sayısı ve Model Doğruluğu Arasındaki İlişki .....                                  | 75                  |
| <b>Şekil 3.11:</b> Farklı Modeller için Özellik Sayısı ve Doğruluk İlişkisi .....                             | 76                  |

## KISALTMALAR

|                 |                                                                           |
|-----------------|---------------------------------------------------------------------------|
| <b>AUC</b>      | Area Under the Curve (Eğri Altındaki Alan)                                |
| <b>BİLSEM</b>   | Bilim ve Sanat Merkezi                                                    |
| <b>BYF</b>      | Bireysel Yetenekleri Fark Ettirme Programı                                |
| <b>CNN</b>      | Convolutional Neural Network (Evrişimsel Sinir Ağları)                    |
| <b>CART</b>     | Classification and Regression Trees (Sınıflandırma ve Regresyon Ağaçları) |
| <b>CRISP-DM</b> | Cross-Industry Standard Process for Data Mining                           |
| <b>DBN</b>      | Deep Belief Network (Derin İnanç Ağları)                                  |
| <b>DNN</b>      | Deep Neural Network (Derin Sinir Ağları)                                  |
| <b>EDM</b>      | Educational Data Mining (Eğitimde Veri Madenciliği)                       |
| <b>GRU</b>      | Gated Recurrent Unit                                                      |
| <b>IQR</b>      | Interquartile Range (Çeyrekler Arası Aralık)                              |
| <b>KNN</b>      | K-Nearest Neighbor (K-En Yakın Komşu)                                     |
| <b>LASSO</b>    | Least Absolute Shrinkage and Selection Operator                           |
| <b>LDA</b>      | Linear Discriminant Analysis (Doğrusal Ayırt Edici Analiz)                |
| <b>LOOCV</b>    | Leave-One-Out Cross-Validation                                            |
| <b>LSTM</b>     | Long Short-Term Memory                                                    |
| <b>MAE</b>      | Mean Absolute Error (Ortalama Mutlak Hata)                                |
| <b>MEB</b>      | Millî Eğitim Bakanlığı                                                    |
| <b>MLP</b>      | Multi-Layer Perceptron (Çok Katmanlı Algılayıcı)                          |
| <b>MOOC</b>     | Massive Open Online Course (Kitlese Açık Çevrimiçi Kurs)                  |

|                |                                                                                                |
|----------------|------------------------------------------------------------------------------------------------|
| <b>MSE</b>     | Mean Squared Error (Ortalama Kare Hata)                                                        |
| <b>ÖYG</b>     | Özel Yetenekleri Geliştirme Programı                                                           |
| <b>PCA</b>     | Principal Component Analysis (Temel Bileşenler Analizi)                                        |
| <b>PISA</b>    | Programme for International Student Assessment (Uluslararası Öğrenci Değerlendirme Programı)   |
| <b>RBF</b>     | Radial Basis Function (Radyal Tabanlı Fonksiyon)                                               |
| <b>RMSE</b>    | Root Mean Squared Error (Kök Ortalama Kare Hata)                                               |
| <b>RNN</b>     | Recurrent Neural Network (Tekrarlayan Sinir Ağları)                                            |
| <b>ROC</b>     | Receiver Operating Characteristic (Alıcı İşletim Karakteristiği)                               |
| <b>STEM</b>    | Science, Technology, Engineering, and Mathematics (Bilim, Teknoloji, Mühendislik ve Matematik) |
| <b>SVG</b>     | Scalable Vector Graphics                                                                       |
| <b>SVM</b>     | Support Vector Machine (Destek Vektör Makinesi)                                                |
| <b>SVR</b>     | Support Vector Regression (Destek Vektör Regresyonu)                                           |
| <b>TIMSS</b>   | Trends in International Mathematics and Science Study                                          |
| <b>t-SEN</b>   | t-distributed Stochastic Neighbor Embedding (t-dağılımlı Stokastik Komşu Gömme)                |
| <b>XGBoost</b> | Extreme Gradient Boosting                                                                      |



**TEZ METNİ**

## GİRİŞ

### Problem Durumu

Üstün yetenekli öğrencilerin eğitimi, gerek ulusal gerekse uluslararası düzeyde eğitim sistemlerinin karşılaştığı önemli zorluklardan biridir. Bu öğrencilerin sahip oldukları yüksek bilişsel kapasiteyi tam anlamıyla geliştirebilmeleri için, onların özel gereksinimlerine uygun eğitim programlarının tasarlanması gerekmektedir (Renzulli, 2012). Matematğin evrensel dili ve günümüz teknoloji dünyasındaki merkezi rolü, üstün yetenekli öğrencilerin matematik alanındaki başarılarının ve gelişimlerinin desteklenmesini özellikle önemli kılmaktadır (Sheffield, 2017).

Türkiye'de üstün yetenekli öğrencilerin eğitimi için kurulan Bilim ve Sanat Merkezleri (BİLSEM), bu öğrencilerin genel eğitim sistemi içerisinde karşılanamayan öğrenme ihtiyaçlarını desteklemektedir. Ancak, bu merkezlerdeki öğrencilerin matematik başarılarının hangi faktörlerden etkilendiğı ve bu başarının nasıl optimize edilebileceğı konusunda kapsamlı araştırmalara ihtiyaç duyulmaktadır (Sak, 2016).

Eğitim alanındaki veri miktarının giderek artması, araştırmacılara eğitim süreçlerini ve sonuçlarını analiz etmede yeni fırsatlar sunmaktadır. Makine öğrenmesi teknikleri, büyük ve karmaşık veri setlerindeki örüntüleri keşfetmede ve tahmin modelleri oluşturmada güçlü araçlar olarak karşımıza çıkmaktadır (Baker & Inventado, 2014). Özellikle eğitim verilerinin analizi için geliştirilen "Eğitimsel Veri Madenciliğı" (Educational Data Mining) ve "Öğrenme Analitiğı" (Learning Analytics) gibi alanlar, öğrenci performansını etkileyen faktörleri belirleme ve başarıyı tahmin etme konusunda önemli imkânlar sunmaktadır (Romero & Ventura, 2020).

Bu çalışmada, Bursa ilindeki BİLSEM'lerde eğitim gören 151 üstün yetenekli öğrencinin matematik başarısını etkileyen faktörlerin, çeşitli makine öğrenmesi teknikleri kullanılarak analiz edilmesi ve tahmin edilmesi amaçlanmaktadır. Öğrencilerin demografik özellikleri, aile yapıları, çalışma alışkanlıkları, motivasyonları, teknoloji kullanımları ve sosyal faaliyetleri gibi birçok faktörün, matematik başarısına olan etkisi incelenecektir.

### Çalışmanın Amacı

Bu çalışmanın temel amacı, Bursa'daki BİLSEM'lerde eğitim gören üstün yetenekli öğrencilerin matematik başarısını etkileyen faktörleri belirlemek ve bu

faktörlerin matematik başarısını ne düzeyde tahmin edebildiğini çeşitli makine öğrenmesi modelleri aracılığıyla incelemektir.

Araştırmanın alt amaçları şu şekilde sıralanabilir:

1. Üstün yetenekli öğrencilerin matematik başarısını etkileyen demografik, ailesel, eğitimsel ve sosyo-ekonomik faktörleri belirlemek.
2. Matematik başarısına etki eden faktörlerin önem derecelerini makine öğrenmesi teknikleri ile analiz etmek.
3. Farklı makine öğrenmesi algoritmalarının (Lojistik Regresyon, Karar Ağaçları, Destek Vektör Makineleri, Rastgele Orman, Gradyan Artırma ve Derin Sinir Ağları) matematik başarısını tahmin etmedeki performanslarını karşılaştırmak.
4. En yüksek tahmin doğruluğunu sağlayan model(ler)i belirlemek ve bu modellerin yorumlanabilirliğini değerlendirmek.

Elde edilen bulgular ışığında, üstün yetenekli öğrencilerin matematik başarısını artırmaya yönelik eğitimsel önerilerde bulunmak.

### **Araştırma Soruları**

Bu çalışma, aşağıdaki araştırma sorularına yanıt aramaktadır:

1. Üstün yetenekli öğrencilerin matematik başarısını etkileyen en önemli faktörler nelerdir?
2. Öğrencilerin bireysel özellikleri (yaş, cinsiyet, sınıf düzeyi) matematik başarısını ne ölçüde etkilemektedir?
3. Aile desteği ve aile yapısı, üstün yetenekli öğrencilerin matematik başarısında nasıl bir rol oynamaktadır?
4. Öğrencilerin çalışma alışkanlıkları ve tutumları ile matematik başarıları arasında nasıl bir ilişki bulunmaktadır?
5. Teknoloji kullanımı ve öğrenme tercihleri, matematik başarısı üzerinde ne derece etkilidir?
6. Hangi makine öğrenmesi modeli, üstün yetenekli öğrencilerin matematik başarısını tahmin etmede en yüksek performansı göstermektedir?

7. Farklı makine öğrenmesi modellerinin tahmin performansları arasında istatistiksel olarak anlamlı bir fark var mıdır?

Model karmaşıklığı ile tahmin doğruluğu arasında nasıl bir ilişki bulunmaktadır?

### **Çalışmanın Önemi**

Bu çalışma, aşağıda belirtilen nedenlerden dolayı teorik ve pratik açıdan önem taşımaktadır:

Teorik açıdan:

- Üstün yetenekli öğrencilerin matematik başarısını etkileyen faktörlere ilişkin mevcut literatüre katkı sağlayacaktır.
- Eğitim verilerinin analizinde makine öğrenmesi tekniklerinin uygulanmasına yönelik metodolojik bir çerçeve sunacaktır.
- Farklı makine öğrenmesi algoritmaları ve topluluk modellerinin eğitim verisi analizindeki etkinliğini karşılaştırmalı olarak ortaya koyacaktır.

Pratik açıdan:

- BİLSEM'lerde eğitim gören üstün yetenekli öğrencilerin matematik başarısını artırmaya yönelik eğitim programlarının tasarlanmasına rehberlik edecektir.
- Öğretmenlere, velilere ve eğitim yöneticilerine, üstün yetenekli öğrencilerin matematik potansiyellerini en üst düzeyde gerçekleştirmelerine yardımcı olacak veriye dayalı stratejiler sunacaktır.
- Eğitimde kişiselleştirilmiş öğrenme yaklaşımlarının geliştirilmesine katkıda bulunacaktır.

Bu araştırma, veri bilimi ve yapay zekâ tekniklerinin eğitim alanında uygulanması açısından da önemli bir örnek oluşturmaktadır. Bulgular, gelecekte geliştirilecek eğitimsel veri analitiği araçları ve eğitim teknolojileri için temel oluşturabilir.

### **Çalışmanın Kapsamı**

Bu çalışma, Bursa ilinde bulunan Bilim ve Sanat Merkezleri'nde (BİLSEM) eğitim gören 151 üstün yetenekli öğrencinin matematik başarısını etkileyen faktörleri incelemektedir. Örneklem, 10–14 yaş aralığındaki 5–8. sınıf öğrencilerinden oluşmakta olup, gönüllülük esasına dayalı olarak belirlenmiştir. Araştırma, öğrencilerin bireysel,

ailesel, eğitimsel ve sosyo-ekonomik özelliklerini dikkate alarak makine öğrenmesi algoritmaları aracılığıyla matematik başarısını tahmin etmeyi amaçlamaktadır.

### **Bursa BİLSEM'ler ve Öğrenci Profili**

Bu çalışma, Bursa il merkezinde ve ilçelerinde yer alan Bilim ve Sanat Merkezleri'nde (BİLSEM) eğitim gören üstün yetenekli öğrencileri kapsamaktadır. Türkiye'deki BİLSEM'ler, Millî Eğitim Bakanlığı'na bağlı olarak faaliyet gösteren, üstün yetenekli öğrencilere örgün eğitim dışında zenginleştirilmiş ve farklılaştırılmış eğitim olanakları sunan kurumlardır (MEB, 2016).

Bursa'daki BİLSEM'ler, kuruluşlarından bu yana çeşitli alanlarda (genel zihinsel yetenek, müzik, resim, vb.) üstün yetenek gösteren ilkokul, ortaokul ve lise öğrencilerine hizmet vermektedir. Bu merkezlere kabul edilen öğrenciler, standart testler ve değerlendirme süreçleri sonucunda üstün yetenekli olarak tanımlanmış öğrencilerdir. Bu öğrenciler, normal örgün eğitimlerine devam ederken, okul saatleri dışında BİLSEM'e gelerek yeteneklerini geliştirmeye yönelik programlara katılmaktadırlar.

Bursa ilinde yer alan BİLSEM'lerdeki öğrencilerin çoğunluğu, şehir merkezinde veya yakın çevresinde yaşayan, orta ve üst-orta sosyoekonomik düzeye sahip ailelerin çocuklarıdır. Öğrenciler, BİLSEM'de uyum, destek, bireysel yetenekleri fark ettirme, özel yetenekleri geliştirme ve proje üretimi olmak üzere çeşitli eğitim aşamalarından geçmektedirler (Sak vd., 2015).

### **Örneklem Büyüklüğü**

Bu araştırmanın örneklemini, Bursa il merkezindeki ve ilçelerindeki BİLSEM'lerde eğitim gören 151 üstün yetenekli öğrenci oluşturmaktadır. Örneklem büyüklüğü, istatistiksel analizler ve makine öğrenmesi uygulamaları için yeterli kabul edilebilecek boyuttadır (Hastie, Tibshirani, & Friedman, 2009). Öğrencilerin yaş aralığı 10-14 olup, 5. sınıftan 8. sınıfa kadar olan öğrencileri kapsamaktadır.

Örneklem seçiminde gönüllülük esası gözetilmiş, araştırmaya katılmayı kabul eden ve veli izni alınan öğrenciler dâhil edilmiştir. Örneklem, BİLSEM'deki cinsiyet ve sınıf düzeyi dağılımını temsil edecek şekilde oluşturulmuştur.

### **Sınırlılıklar ve Varsayımlar**

Bu araştırmanın bazı sınırlılıkları ve varsayımları bulunmaktadır:

Sınırlılıklar:

- Araştırma, sadece Bursa ilindeki BİLSEM'lerde eğitim gören öğrencilerle sınırlıdır, bu nedenle bulgular diğer illerdeki BİLSEM'lere veya genel öğrenci popülasyonuna genellenemeyebilir.
- Veri toplama aracı olarak kullanılan anket, öz-bildirime dayalıdır ve öğrencilerin öznel değerlendirmelerini içermektedir.
- Matematik başarıları, öğrencilerin bildirdiği not ortalamaları üzerinden değerlendirilmiştir, standardize edilmiş bir matematik testi uygulanmamıştır.
- Araştırma kesitsel bir çalışma olup, değişkenler arasındaki ilişkileri zaman içinde izlememektedir.
- Makine öğrenmesi modellerinin performansı, veri setinin boyutu ve kalitesiyle sınırlıdır.

#### Varsayımlar:

- Öğrencilerin anketi samimi ve doğru bir şekilde yanıtladıkları varsayılmıştır.
- BİLSEM'e kabul edilen öğrencilerin, üstün yetenek tanılama süreçlerinin geçerli ve güvenilir olduğu varsayılmıştır.
- Kullanılan makine öğrenmesi algoritmaları için gerekli varsayımların (bağımsız değişkenlerin çoklu doğrusal bağlantı sorunu içermemesi, artıkların normal dağılıma sahip olması vb.) karşılandığı varsayılmıştır.
- Öğrencilerin bildirdiği matematik not ortalamalarının, gerçek matematik başarılarını yansıttığı varsayılmıştır.

#### Tanımlar

Araştırmada kullanılan temel kavramlar aşağıda tanımlanmıştır:

**Üstün Yetenek:** Belirli alanlarda yaşlarına göre üst düzey performans gösteren veya gösterme potansiyeline sahip olan, yaratıcılık, motivasyon, genel veya özel akademik yetenek, liderlik ve sanat alanlarında yüksek kapasiteye sahip bireyler (Renzulli, 2005).

**BİLSEM (Bilim ve Sanat Merkezi):** Örgün eğitim kurumlarına devam eden, üstün yetenekli öğrencilerin örgün eğitim dışındaki zamanlarda ilgi ve yetenekleri

doğrultusunda eğitim aldıkları, Millî Eğitim Bakanlığına bağlı özel eğitim kurumları (MEB, 2016).

**Matematik Başarısı:** Öğrencilerin matematik dersindeki performanslarını gösteren, akademik notlar veya standart testlerle ölçülen başarı düzeyi.

**Makine Öğrenmesi:** Bilgisayarların açık bir şekilde programlanmadan veri ve deneyimlerden öğrenmesini sağlayan algoritmalar ve istatistiksel modellerin geliştirilmesiyle ilgilenen yapay zekâ dalı (Mitchell, 1997).

**Sınıflandırma Algoritmaları:** Verileri önceden belirlenmiş kategorilere atamak için kullanılan makine öğrenmesi algoritmaları (örn. Lojistik Regresyon, Karar Ağaçları).

**Regresyon Algoritmaları:** Bağımsız değişkenler ile bağımlı değişken arasındaki ilişkiyi modelleyerek sürekli bir çıktı değeri tahmin etmeye yarayan makine öğrenmesi algoritmaları.

**Topluluk Öğrenme (Ensemble Learning):** Birden fazla öğrenme algoritmasının birleştirilerek, tek bir algoritmanın performansından daha iyi tahmin performansı elde etmeyi amaçlayan makine öğrenmesi yaklaşımı (örn. Rastgele Orman, Gradyan Artırma) (Dietterich, 2000).

**Özellik Seçimi (Feature Selection):** Veri setindeki tüm özellikler arasından, tahmin performansını en çok etkileyen özelliklerin seçilmesi süreci.

**Hiperparametre Optimizasyonu:** Makine öğrenmesi modelinin performansını artırmak için model parametrelerinin en uygun değerlerinin belirlenmesi süreci.

**Çapraz Doğrulama (Cross-validation):** Modelin genelleme yeteneğini değerlendirmek için veri setinin birden fazla alt kümeye bölünüp, her seferinde farklı bir alt küme test verisi olarak kullanılarak modelin performansının ölçülmesi tekniği (Kohavi, 1995).



**BİRİNCİ BÖLÜM**  
**LİTERATÜR TARAMASI**

## 1.1. ÜSTÜN YETENEKLİ ÖĞRENCİLER VE EĞİTİMLERİ

Üstün yetenekli öğrenciler; yüksek bilişsel kapasiteye, yaratıcı düşünme becerilerine ve akademik alanda hızlı öğrenme yetisine sahip bireylerdir. Bu öğrencilerin gelişimlerinin desteklenmesi için özel eğitim programlarına ihtiyaç duyulmaktadır. Türkiye'de bu ihtiyacı karşılamak amacıyla kurulan BİLSEM'ler, öğrencilere bireyselleştirilmiş ve farklılaştırılmış eğitim ortamları sunmakta; zihinsel, sanatsal ve bilimsel alanlarda yeteneklerini geliştirmeyi hedeflemektedir.

### 1.1.1. Üstün Yetenekli Öğrencilerin Özellikleri

Üstün yetenekli öğrenciler, akranlarına göre bilişsel, duyuşsal ve yaratıcılık alanlarında farklılık gösteren bireylerdir. Bu öğrencilerin tipik özellikleri arasında hızlı öğrenme, üst düzey soyut düşünme becerisi, geniş kelime dağarcığı, yüksek merak duygusu, sorunlara yaratıcı çözümler üretme yeteneği, yüksek motivasyon ve ilgi alanlarında derinleşme eğilimi bulunmaktadır (Sak, 2021). Feldhusen (2005), üstün yetenekli öğrencileri tanımlarken şu kriterleri vurgulamaktadır: genel zekâ testlerinden yüksek puan alma, hızlı öğrenme kapasitesi, kavramsal ve soyut düşünme yeteneği, belirli akademik alanlarda üstün başarı ve yüksek yaratıcılık.

Matematik alanında üstün yetenek gösteren öğrenciler ise, sayısal ilişkileri hızlı kavrama, matematiksel örüntüleri fark etme, karmaşık problemleri analiz etme ve çözme, soyut matematiksel kavramları anlama ve uygulama konularında özel yetenekler sergilemektedirler (Sriraman, 2008). Sheffield (2017), matematiksel yetenek gösteren öğrencilerin, matematiksel yapıları keşfetme, genelleme yapma, matematiksel fikirleri sembolleştirme ve ilişkilendirme, matematiksel akıl yürütme, problem çözme stratejilerini esnek bir şekilde kullanma gibi özelliklere sahip olduklarını belirtmektedir.

Üstün yetenekli öğrencilerin bu özelliklerinin yanı sıra, duygusal ve sosyal alanlarda da farklılıklar gösterdikleri bilinmektedir. Bu öğrenciler genellikle yüksek adalet duygusu, empati, ahlaki duyarlılık, mükemmeliyetçilik ve yoğun duygusal tepkiler gösterebilmektedirler (Silverman, 2013). Ayrıca, bazı üstün yetenekli öğrenciler asenkronize gelişim sergilerler; yani bilişsel, duygusal ve fiziksel gelişimleri arasında farklılıklar olabilir (Morelock, 1996).

Bu öğrencilerin eğitiminde en önemli zorluklardan biri, standart eğitim programlarının onların hızlı öğrenme kapasitelerini ve derinlemesine anlama isteklerini

karşılamamasıdır. Bu nedenle, üstün yetenekli öğrenciler için zenginleştirilmiş, hızlandırılmış ve farklılaştırılmış eğitim programlarına ihtiyaç duyulmaktadır (Renzulli & Reis, 2021).

### 1.1.2. Türkiye'de BİLSEM Yapısı ve Eğitim Programları

Türkiye'de üstün yetenekli öğrencilerin eğitimi için özel olarak tasarlanmış Bilim ve Sanat Merkezleri (BİLSEM), 1990'lı yılların sonlarından itibaren yaygınlaşmaya başlamıştır. Millî Eğitim Bakanlığı'na bağlı olarak faaliyet gösteren bu merkezler, örgün eğitim kurumlarına devam eden üstün yetenekli öğrencilerin okul saatleri dışında yeteneklerini geliştirmelerine olanak sağlayan kurumlardır (MEB, 2016).

BİLSEM'lere öğrenci seçim süreci, genellikle sınıf öğretmenlerinin aday gösterdiği öğrenciler arasından, standart zekâ ve yetenek testleri ile yapılan değerlendirmelerle gerçekleştirilmektedir. Seçilen öğrenciler, genel zihinsel yetenek, görsel sanatlar, müzik ve diğer özel yetenek alanlarına göre programlara yerleştirilmektedir (Sak, 2016).

BİLSEM eğitim programı, beş temel aşamadan oluşmaktadır (MEB, 2019):

**Uyum Programı:** Öğrencilerin BİLSEM'e uyumunu kolaylaştırmak, kurum kültürünü ve işleyişini tanıtmak amacıyla düzenlenen kısa süreli bir programdır.

**Destek Eğitimi Programı:** Öğrencilerin temel bilimsel düşünme becerilerini kazanmaları, iletişim becerilerini geliştirmeleri ve bilimsel çalışma yöntemlerini öğrenmeleri hedeflenen programdır.

**Bireysel Yetenekleri Fark Ettirme Programı (BYF):** Öğrencilerin güçlü ve zayıf yönlerini fark etmeleri, bilimsel araştırma teknikleri, problem çözme ve karar verme tekniklerini geliştirmeleri amaçlanan programdır.

**Özel Yetenekleri Geliştirme Programı (ÖYG):** Öğrencilerin ilgi ve yeteneklerine göre gruplara ayrılarak, alanlarında derinlemesine ve ileri düzeyde bilgi, beceri ve davranışlar kazanmalarının sağlandığı programdır.

**Proje Üretimi ve Hazırlama Programı:** Öğrencilerin ilgi ve yetenekleri doğrultusunda bilimsel araştırma teknikleri kullanarak proje ürettikleri, projelerini uygulayıp değerlendirdikleri programdır.

BİLSEM programlarının temel felsefesi, üstün yetenekli öğrencilerin ilgi ve yetenekleri doğrultusunda bireyselleştirilmiş eğitim almaları, yaratıcılıklarını, problem çözme ve eleştirel düşünme becerilerini geliştirmeleridir. Bu merkezlerde, proje tabanlı öğrenme, sorgulama temelli öğrenme, zenginleştirilmiş aktiviteler ve hızlandırılmış öğretim gibi yöntemler kullanılmaktadır (Gökdere & Küçük, 2003).

Bursa'daki BİLSEM'ler , ülkedeki diğer BİLSEM'ler gibi, Millî Eğitim Bakanlığı'nın belirlediği çerçeve programı uygulamanın yanı sıra, yerel ihtiyaçlara ve imkânlara göre özgün projeler ve etkinlikler de düzenlemektedir. Merkezde matematik alanında verilen eğitimler, temel matematiksel düşünme becerileri, problem çözme stratejileri, matematiksel modelleme, algoritma geliştirme ve kodlama gibi konulara odaklanmaktadır (Özenç & Özenç, 2013).

## **1.2. MATEMATİK BAŞARISINI ETKİLEYEN FAKTÖRLER**

Matematik başarısı, bireysel, ailesel, eğitimsel ve sosyo-ekonomik birçok faktörün etkileşimiyle şekillenmektedir. Bireysel faktörler arasında bilişsel yetenek, öz-yeterlik inancı, motivasyon, tutum, çalışma alışkanlıkları ve öğrenme stilleri öne çıkmaktadır. Ailesel faktörler, ebeveyn eğitimi, sosyo-ekonomik durum, aile desteği ve ev ortamındaki eğitim olanaklarını kapsamaktadır. Eğitimsel faktörler arasında öğretmen nitelikleri, öğretim yöntemleri, müfredat içeriği ve sınıf ortamı yer almakta; bu unsurlar öğrenme sürecini doğrudan etkilemektedir. Sosyo-ekonomik faktörler ise öğrencinin gelir düzeyi, eğitim kaynaklarına erişimi, kültürel sermayesi ve bölgesel koşullarla ilişkilidir. Üstün yetenekli öğrenciler açısından bu faktörlerin etkileri daha karmaşık olup, bu öğrencilere uygun ortam ve destek sağlanmadığında potansiyelleri tam anlamıyla ortaya çıkamayabilmektedir.

### **1.2.1. Bireysel Faktörler**

Öğrencilerin matematik başarısını etkileyen bireysel faktörler, çeşitli araştırmacılar tarafından incelenmiştir. Bu faktörler arasında bilişsel yetenekler, öz-yeterlik inançları, motivasyon, matematiğe karşı tutum, çalışma alışkanlıkları ve öğrenme stilleri bulunmaktadır.

Bilişsel Yetenekler: Genel zekâ ve matematiksel yetenek, matematik başarısı üzerinde güçlü bir etkiye sahiptir. Deary ve arkadaşları (2007), genel bilişsel yeteneklerin matematik performansı ile yüksek düzeyde ilişkili olduğunu ortaya koymuşlardır. Ancak,

üstün yetenekli öğrencilerde bilişsel yeteneklerden ziyade, diğer faktörlerin (motivasyon, çalışma stratejileri vb.) başarı farklılıklarını açıklamada daha önemli rol oynadığı görülmektedir (Preckel, Holling, & Vock, 2006).

**Öz-yeterlik İnançları:** Bandura'nın (1997) öz-yeterlik teorisine göre, öğrencilerin matematiksel görevleri başarıyla yapabileceklerine dair inançları, matematik başarısını güçlü bir şekilde etkilemektedir. Pajares ve Miller (1994), matematik öz-yeterliğinin matematik performansını doğrudan etkilediğini ve matematik kaygısı ile negatif ilişkili olduğunu göstermiştir.

**Motivasyon:** İçsel motivasyon (matematiği sevmek, merak, öğrenme isteği) ve dışsal motivasyon (not, ödül, rekabet), matematik başarısında önemli rol oynamaktadır. Ryan ve Deci (2000), içsel motivasyonun akademik başarı üzerinde daha kalıcı ve olumlu etkileri olduğunu belirtmektedir. Üstün yetenekli öğrencilerde motivasyon, özellikle de görev tamamlama azmi ve zorlukları aşma isteği, matematik başarısının önemli belirleyicilerindendir (Phillips & Lindsay, 2006).

**Matematiğe Karşı Tutum:** Öğrencilerin matematiğe yönelik olumlu veya olumsuz tutumları, başarılarını etkileyen bir diğer faktördür. Ma ve Kishor (1997), meta-analiz çalışmalarında, matematiğe yönelik tutum ile matematik başarısı arasında pozitif bir ilişki olduğunu göstermiştir.

**Çalışma Alışkanlıkları:** Öğrencilerin çalışma stratejileri, zaman yönetimi ve öz-düzenleme becerileri, matematik başarısında belirleyici faktörlerdir. Özellikle, aktif öğrenme stratejileri kullanan, zamanı etkili yöneten ve düzenli çalışan öğrencilerin matematik performanslarının daha yüksek olduğu bulunmuştur (Corno & Mandinach, 2004).

**Öğrenme Stilleri:** Öğrencilerin tercih ettikleri öğrenme stilleri (görsel, işitsel, kinestetik vb.) ile kullanılan öğretim yöntemlerinin uyumu, matematik başarısını etkileyebilmektedir. Dunn ve arkadaşları (2009), öğrencilerin öğrenme stillerine uygun eğitim verildiğinde, akademik başarılarının arttığını göstermiştir.

Üstün yetenekli öğrencilerde, bu bireysel faktörlerin matematik başarısı üzerindeki etkileri daha karmaşık olabilmektedir. Örneğin, bu öğrencilerde mükemmeliyetçilik, yüksek beklentiler ve başarısızlık korkusu gibi faktörler, motivasyon ve başarı üzerinde olumsuz etki yaratabilir (Reis & McCoach, 2000).

### 1.2.2. Ailesel Faktörler

Aile yapısı, ailenin eğitim düzeyi, sosyo-ekonomik durumu ve aile desteği, öğrencilerin matematik başarısını etkileyen önemli dışsal faktörlerdendir.

**Aile Yapısı:** Aile büyüklüğü, aile düzeni ve ebeveyn varlığı, öğrencilerin akademik başarısını etkileyebilmektedir. Araştırmalar, iki ebeveynli ailelerde yetişen çocukların genellikle daha yüksek akademik başarı gösterdiğini ortaya koymuştur (Amato, 2001).

**Ailenin Eğitim Düzeyi:** Ebeveynlerin eğitim seviyesi ile çocukların matematik başarısı arasında pozitif bir ilişki bulunmaktadır. Davis-Kean (2005), ebeveynlerin eğitim düzeyinin, çocuklarına sağladıkları eğitim ortamı, akademik beklentiler ve bilişsel teşvik düzeyini etkileyerek çocukların akademik başarısına katkıda bulunduğunu göstermiştir.

**Sosyo-ekonomik Durum:** Ailenin gelir düzeyi, yaşam koşulları ve kaynaklara erişimi, matematik başarısını etkileyebilmektedir. PISA ve TIMSS gibi uluslararası değerlendirmelerde, sosyo-ekonomik düzey ile matematik başarısı arasında güçlü bir ilişki olduğu görülmektedir (OECD, 2019). Ancak, üstün yetenekli öğrencilerde bu ilişkinin gücü azalabilmektedir (Worrell, Subotnik, Olszewski-Kubilius, & Dixon, 2019).

**Aile Desteği:** Ebeveynlerin çocuklarının eğitimine katılımı, matematik ödevlerine yardım, eğitimsel kaynaklar sağlama, akademik beklentiler ve matematik etkinliklerine teşvik gibi faktörler, matematik başarısını olumlu yönde etkilemektedir (Fan & Chen, 2001). Özellikle üstün yetenekli öğrencilerde, ailenin çocuğun özel ihtiyaçlarını anlaması ve desteklemesi kritik öneme sahiptir (Olszewski-Kubilius, 2008).

**Matematiksel Ev Ortamı:** Evde matematiksel etkinliklerin ve konuşmaların varlığı, matematiksel oyunlar, bulmacalar ve kaynakların bulunması, çocukların matematik başarısını olumlu yönde etkilemektedir (Melhuish vd., 2008).

Ayrıca, ailenin matematik ve bilime verdiği değer, matematiğe yönelik tutumu ve ebeveynlerin kendi matematik yeterlilikleri de çocukların matematik başarısını ve motivasyonunu etkileyebilmektedir (Eccles, 2007).

### 1.2.3. Eğitimsel Faktörler

Eğitim ortamı, öğretim yöntemleri, öğretmen nitelikleri ve müfredat gibi eğitimsel faktörler, öğrencilerin matematik başarısında belirleyici rol oynamaktadır.

**Öğretmen Nitelikleri:** Öğretmenlerin alan bilgisi, pedagojik yetkinliği, matematiğe karşı tutumu ve üstün yetenekli öğrencilerin ihtiyaçlarını anlama düzeyi, matematik başarısını etkilemektedir. Hill, Rowan ve Ball (2005), öğretmenlerin matematiksel bilgi düzeyinin, öğrencilerinin matematik başarısı üzerinde doğrudan etkisi olduğunu göstermiştir.

**Öğretim Yöntemleri:** Problem tabanlı öğrenme, sorgulamaya dayalı öğrenme, işbirlikli öğrenme ve aktif öğrenme gibi öğrenci merkezli yöntemlerin, özellikle üstün yetenekli öğrencilerin matematik başarısını artırdığı görülmektedir (Boaler, 2002). Ayrıca, bireyselleştirilmiş öğretim ve zorluk düzeyinin öğrencinin yeteneğine göre ayarlanması, üstün yetenekli öğrencilerin motivasyonunu ve başarısını olumlu etkilemektedir (Rogers, 2007).

**Müfredat ve İçerik:** Matematik müfredatının derinliği, genişliği, zorluğu ve öğrencilerin ilgi ve yeteneklerine uygunluğu, başarıyı etkileyen faktörlerdendir. Üstün yetenekli öğrenciler için zenginleştirilmiş, hızlandırılmış ve derinleştirilmiş bir müfredat, onların potansiyellerini gerçekleştirmelerine yardımcı olmaktadır (VanTassel-Baska & Brown, 2007).

**Sınıf Ortamı:** Sınıf büyüklüğü, kaynakların mevcudiyeti, teknolojik altyapı ve sınıf atmosferi, matematik başarısını etkileyebilmektedir. Olumlu bir öğrenme iklimi, güvenli bir sınıf ortamı ve öğrencilerin aktif katılımını teşvik eden bir yapı, matematik performansını artırmaktadır (Hattie, 2009).

**Ölçme ve Değerlendirme Uygulamaları:** Kullanılan değerlendirme yöntemleri, geri bildirim mekanizmaları ve değerlendirme kriterleri, öğrencilerin matematik başarısını ve motivasyonunu etkilemektedir. Özellikle üstün yetenekli öğrenciler için, standart testlerin ötesinde, üst düzey düşünme becerilerini ölçen değerlendirme yöntemleri önemlidir (VanTassel-Baska, 2013).

**Akran Etkileşimi:** Akran işbirliği, matematik tartışmaları ve grup çalışmaları, öğrencilerin matematik anlayışını ve başarısını geliştirmektedir. Üstün yetenekli öğrencilerin benzer yetenekteki akranlarla etkileşim kurması, onların matematiksel düşünme becerilerini desteklemektedir (Neihart, 2007).

#### 1.2.4. Sosyo-Ekonomik Faktörler

Sosyo-ekonomik faktörler, öğrencilerin eğitim fırsatlarına erişimini, kaynaklara ulaşımını ve eğitimsel sonuçlarını etkilemektedir.

**Gelir Düzeyi:** Ailenin gelir düzeyi, eğitimsel kaynaklara erişim, özel dersler, eğitim teknolojileri ve zenginleştirilmiş öğrenme deneyimlerine katılım gibi faktörleri etkileyerek matematik başarısını dolaylı olarak etkilemektedir (Sirin, 2005).

**Eğitimsel Kaynaklar:** Evde bulunan kitaplar, bilgisayar, internet erişimi, eğitimsel yazılımlar ve oyunlar gibi kaynaklar, matematik öğrenmesini desteklemektedir. Sosyo-ekonomik durumu yüksek ailelerin çocukları genellikle daha fazla eğitimsel kaynağa sahiptir (Chiu & Xihua, 2008).

**Kültürel Sermaye:** Ailenin kültürel sermayesi (müzelere, bilim merkezlerine ziyaretler, kitap okuma alışkanlığı, eğitimsel etkinliklere katılım), öğrencilerin akademik performansını etkilemektedir. Bourdieu'ya (1986) göre kültürel sermaye, ailenin sosyo-ekonomik durumu ile çocukların akademik başarısı arasındaki ilişkiyi açıklayan önemli bir faktördür.

**Bölgesel Faktörler:** Yaşanılan bölgenin gelişmişlik düzeyi, okulların kalitesi, eğitim imkânları ve toplumsal özellikler, matematik başarısını etkileyebilmektedir. Kırsal ve kentsel bölgeler arasında, eğitim kalitesi ve akademik başarı açısından farklılıklar gözlenmektedir (OECD, 2018).

**Toplumsal Değerler ve Normlar:** Toplumun matematiğe ve eğitime verdiği değer, matematiğe yönelik cinsiyet kalıpları ve başarıya ilişkin toplumsal beklentiler, öğrencilerin matematikteki performansını etkileyebilmektedir. Örneğin, bazı Doğu Asya ülkelerinde matematiğe verilen yüksek toplumsal değer ve eğitime yapılan vurgu, bu ülkelerdeki öğrencilerin matematik başarısını olumlu etkilemektedir (Stevenson & Stigler, 1992).

Üstün yetenekli öğrencilerde, sosyo-ekonomik faktörlerin etkisi bazen azalabilmekte, ancak bu öğrencilerin tam potansiyellerine ulaşmaları için uygun kaynaklar ve fırsatlar önemli olmaktadır. Özellikle düşük sosyo-ekonomik düzeye sahip üstün yetenekli öğrenciler, potansiyellerini gerçekleştirme konusunda daha fazla engelle karşılaşabilmektedir (Olszewski-Kubilius & Corwith, 2018).

Bu faktörlerin tümü, öğrencilerin matematik başarısını karmaşık bir etkileşim içinde etkilemektedir. Makine öğrenmesi modelleri kullanılarak, bu faktörlerin bir arada değerlendirilmesi ve matematik başarısı üzerindeki etkilerinin analiz edilmesi, eğitim politikalarının ve müdahalelerinin geliştirilmesine katkı sağlayabilir.

### **1.3. VERİ BİLİMİ VE EĞİTİM**

#### **1.3.1. Eğitimde Veri Madenciliği**

Eğitimde veri madenciliği (Educational Data Mining - EDM), eğitimsel ortamlardan toplanan verilerin analiz edilmesi ve bu verilerden anlamlı örüntülerin çıkarılması amacıyla geliştirilmiş bir alandır. EDM, eğitim verilerini analiz etmek için istatistiksel, makine öğrenmesi ve veri madenciliği tekniklerini kullanır (Baker & Yacef, 2009).

#### **1.3.2. Eğitimsel Veri Analizi Alanındaki Çalışmalar**

Eğitimsel veri analizi alanındaki çalışmalar, öğrenci başarısını tahmin etme, öğrenme davranışlarını modelleme, eğitim kaynaklarını optimize etme ve eğitim politikalarını şekillendirme gibi çeşitli alanlara odaklanmaktadır.

#### **1.3.3. Öğrenci Başarısını Tahmin Eden Çalışmalar:**

Matematik başarısını tahmin etmeye yönelik çalışmalar, bu araştırma açısından özellikle önemlidir. Huang ve Fang (2013), mühendislik öğrencilerinin matematik performanslarını tahmin etmek için çeşitli makine öğrenmesi modellerini (çoklu lineer regresyon, polinom regresyon, yapay sinir ağları ve destek vektör makineleri) karşılaştırmışlardır. Çalışma, yapay sinir ağları ve destek vektör makinelerinin daha yüksek tahmin doğruluğu sağladığını göstermiştir.

Cortez ve Silva (2008), Portekiz'deki ortaöğretim öğrencilerinin matematik ve Portekizce derslerindeki başarılarını tahmin etmek için karar ağaçları, rastgele orman, sinir ağları ve destek vektör makineleri gibi çeşitli makine öğrenmesi tekniklerini kullanmıştır. Araştırma sonuçları, öğrencilerin daha önceki akademik performanslarının, başarıyı tahmin etmede en önemli faktör olduğunu, ancak sosyo-ekonomik faktörlerin ve okul geçmişinin de etkili olduğunu göstermiştir.

Masci ve arkadaşları (2018), İtalya'daki INVALSI ulusal değerlendirme sınavları verilerini kullanarak, öğrencilerin matematik ve İtalyanca başarılarını tahmin etmek için hiyerarşik doğrusal modeller ve makine öğrenmesi tekniklerini karşılaştırmışlardır. Bulgular, makine öğrenmesi modellerinin, özellikle de rastgele ormanların, geleneksel istatistiksel yöntemlere göre daha yüksek tahmin doğruluğu sağladığını göstermiştir.

#### **1.3.4. Eğitimsel Veri Madenciliğinde Topluluk Öğrenme Yöntemleri:**

Son yıllarda, eğitimsel veri madenciliğinde topluluk öğrenme (ensemble learning) yöntemlerinin kullanımı artmıştır. Adejo ve Connolly (2018), öğrencilerin akademik performansını tahmin etmek için rastgele orman, gradyan artırma ve AdaBoost gibi topluluk öğrenme yöntemlerini kullanmış ve bu yöntemlerin tekli modellerden daha iyi performans gösterdiğini bulmuşlardır.

Xu ve arkadaşları (2017), MOOC (Massive Open Online Course) platformlarından toplanan verileri kullanarak, öğrencilerin kursu tamamlama durumlarını tahmin etmek için topluluk öğrenme yöntemlerinden yararlanmışlardır. Çalışma, topluluk modellerin, özellikle de stacking (istifleme) yönteminin, tek bir sınıflandırıcıya göre daha yüksek tahmin doğruluğu sağladığını göstermiştir.

#### **1.3.5. Üstün Yetenekli Öğrencilere Yönelik Veri Analizi Çalışmaları:**

Üstün yetenekli öğrencilere yönelik veri analizi çalışmaları sınırlı olmakla birlikte, bu alanda da ilgi artmaktadır. Alkan ve Mertol (2019), Türkiye'deki BİLSEM öğrencilerinin akademik başarılarını etkileyen faktörleri belirlemek için veri madenciliği tekniklerinden yararlanmışlardır. Çalışma, aile desteği, öğretmen nitelikleri ve öğrencilerin motivasyonunun, üstün yetenekli öğrencilerin akademik performansında önemli faktörler olduğunu ortaya koymuştur.

Periathiruvadi ve Rinn (2012), üstün yetenekli öğrencilerin tanınması ve eğitiminde teknoloji kullanımını inceleyen bir literatür taraması yapmış ve bu alanda veri analitiği ve yapay zekâ uygulamalarının potansiyelini vurgulamışlardır.

Khine (2013), TIMSS (Trends in International Mathematics and Science Study) verileri üzerinde yaptığı çalışmada, üstün performans gösteren öğrencilerin özelliklerini belirlemek için veri madenciliği tekniklerini kullanmıştır. Bulgular, öğrencilerin evdeki

eğitim kaynaklarının, matematiğe yönelik tutumlarının ve öz-güvenlerinin yüksek matematik performansı ile ilişkili olduğunu göstermiştir.

Eğitimsel veri analizi alanındaki bu çalışmalar, makine öğrenmesi tekniklerinin, öğrenci başarısını tahmin etme ve etkileyen faktörleri belirleme konusunda güçlü araçlar olduğunu göstermektedir. Ancak, üstün yetenekli öğrencilerin matematik başarısını etkileyen faktörleri makine öğrenmesi yöntemleriyle kapsamlı bir şekilde analiz eden çalışmaların sınırlı olduğu görülmektedir. Bu araştırma, söz konusu boşluğu doldurmaya katkıda bulunmayı amaçlamaktadır.

#### **1.4. MAKİNE ÖĞRENMESİ VE EĞİTİM VERİSİ ANALİZİNDEKİ UYGULAMALARI**

Eğitim ortamlarında giderek artan veri çeşitliliği ve hacmi, öğrenci başarısının tahmini ve öğrenme süreçlerinin analizinde makine öğrenmesi tekniklerinin kullanılmasını önemli hâle getirmiştir. Makine öğrenmesi, eğitimsel veri madenciliği kapsamında öğrenci performansını etkileyen faktörleri belirlemede ve çeşitli öngörülerde bulunmada güçlü araçlar sunmaktadır. Bu bölümde, sınıflandırma, regresyon, topluluk öğrenme ve derin öğrenme gibi yöntemlerin eğitim verisi üzerindeki uygulamaları ele alınacaktır.

##### **1.4.1. Sınıflandırma Algoritmaları**

Sınıflandırma algoritmaları, veriyi önceden tanımlanmış kategorilere atamak için kullanılan denetimli öğrenme yöntemleridir. Eğitim verisi analizinde, öğrencilerin başarılı/başarısız, düşük/orta/yüksek performanslı gibi kategorilere ayrılması için sıklıkla kullanılmaktadır.

Lojistik Regresyon: İstatistiksel bir yöntem olan lojistik regresyon, bağımsız değişkenlerle bir çıktı değişkeni arasındaki ilişkiyi modelleyerek, ikili sınıflandırma (binary classification) problemlerinde kullanılır. Eğitim araştırmalarında, öğrenci başarısını veya bir programı tamamlama olasılığını tahmin etmek için yaygın olarak kullanılmaktadır. Araştırmacılar, bu modelin yorumlanabilirliği ve değişkenlerin etki büyüklüğünü belirleyebilmesi nedeniyle sıklıkla tercih etmektedir (Hosmer & Lemeshow, 2000).

Karar Ağaçları: Karar ağaçları, verileri "eğer-o halde" kurallarına göre bölerek sınıflandıran, anlaşılması kolay görsel modellerdir. ID3, C4.5, CART ve CHAID gibi

çeşitli karar ağacı algoritmaları mevcuttur. Eğitim verisi analizinde karar ağaçları, öğrenci performansını tahmin etme, risk altındaki öğrencileri belirleme ve başarı faktörlerini hiyerarşik olarak sınıflandırma amacıyla kullanılmaktadır (Breiman vd., 1984).

Márquez-Vera ve arkadaşları (2016), lise öğrencilerinin okulu bırakma riskini tahmin etmek için çeşitli karar ağacı algoritmalarını kullanan bir çalışma yapmıştır. Sonuçlar, karar ağaçlarının, risk altındaki öğrencileri yüksek doğrulukla tahmin edebildiğini ve modelin anlaşılabilirliği sayesinde eğitimcilere değerli içgörüler sunduğunu göstermiştir.

Destek Vektör Makineleri (SVM): SVM algoritması, veri noktalarını bir hiper düzlemle ayırarak sınıflandırma yapar ve sınıflar arasındaki marjı maksimize eder. Doğrusal olmayan problemlerde, çekirdek (kernel) fonksiyonları kullanılarak veriler daha yüksek boyutlu bir uzaya dönüştürülür. SVM, eğitim verisi analizinde özellikle yüksek boyutlu veri setlerinde ve küçük örneklemelerde etkili olmaktadır (Cortes & Vapnik, 1995).

Asif ve arkadaşları (2017), üniversite öğrencilerinin akademik performansını tahmin etmek için SVM, karar ağaçları ve Naive Bayes algoritmaları karşılaştırmıştır. Çalışma sonuçları, SVM'nin daha yüksek tahmin doğruluğu sağladığını göstermiştir.

Naive Bayes: Bayes teoremi temeline dayanan bu algoritma, özellikler arasındaki bağımsızlık varsayımıyla sınıflandırma yapar. Hesaplama açısından verimli olması ve küçük veri setlerinde iyi performans göstermesiyle bilinir. Eğitim verisi analizinde, öğrenci performansını tahmin etmek, duygu analizi yapmak ve öğrenci davranışlarını sınıflandırmak için kullanılmaktadır (Kotsiantis, Pierrakeas, & Pintelas, 2004).

K-En Yakın Komşu (KNN): Bu algoritma, yeni bir veri noktasını, en yakın k komşusunun sınıfını dikkate alarak sınıflandırır. Eğitim verisi analizinde, öğrenci gruplarını tanımlamak ve benzer özelliklere sahip öğrencileri belirlemek için kullanılmaktadır. KNN, özellikle yerel örüntüleri yakalamada etkilidir, ancak büyük veri setlerinde hesaplama açısından maliyetli olabilir (Cover & Hart, 1967).

Sınıflandırma algoritmalarının eğitim verisi analizindeki uygulamaları, öğrenci performansını tahmin etme, risk altındaki öğrencileri belirleme, öğrenme stillerini sınıflandırma ve eğitimsel müdahalelerin etkinliğini değerlendirme gibi alanlarda çeşitlilik göstermektedir.

### 1.4.2. Regresyon Algoritmaları

Regresyon algoritmaları, bağımsız değişkenler ile sürekli bir bağımlı değişken arasındaki ilişkiyi modelleyerek, sayısal sonuçları tahmin etmeye yarayan denetimli öğrenme yöntemleridir. Eğitim verisi analizinde, öğrencilerin sınav notlarını, akademik performans ortalamalarını veya belirli becerilerdeki gelişim düzeylerini tahmin etmek için kullanılmaktadır.

**Doğrusal Regresyon:** En temel regresyon modellerinden biri olan doğrusal regresyon, bağımsız değişkenler ile bağımlı değişken arasında doğrusal bir ilişki olduğunu varsayar. Eğitim araştırmalarında, öğrenci performansını etkileyen faktörlerin belirlenmesi ve etki büyüklüklerinin hesaplanması için yaygın olarak kullanılmaktadır (Montgomery, Peck, & Vining, 2012).

Oladokun, Adebajo ve Charles-Owaba (2008), akademik ve demografik değişkenleri kullanarak üniversite öğrencilerinin performansını tahmin etmek için doğrusal regresyon modelleri geliştirmiş ve bu modellerin, eğitimcilere öğrenci başarısını destekleyecek stratejiler geliştirmede yardımcı olabileceğini göstermiştir.

**Polinom Regresyon:** Doğrusal regresyonun bir uzantısı olan polinom regresyon, bağımsız değişkenlerle bağımlı değişken arasındaki doğrusal olmayan ilişkileri modellemek için kullanılır. Eğitim verisi analizinde, öğrenme eğrilerini modelleme ve performans değişimlerini analiz etme gibi uygulamalarda tercih edilmektedir (Gálvez vd., 2009).

**Ridge ve Lasso Regresyon:** Bu algoritmalar, çoklu doğrusal bağlantı (multicollinearity) problemini azaltmak ve modelin genelleme yeteneğini artırmak için regülarizasyon teknikleri kullanan doğrusal regresyon yöntemleridir. Lasso regresyon, özellik seçimi için de kullanılabilir çünkü bazı katsayıları tam olarak sıfıra indirir. Eğitim verisi analizinde, özellikle çok sayıda değişkenin olduğu durumlarda, öğrenci performansını tahmin etmek ve en önemli faktörleri belirlemek için kullanılmaktadır (Tibshirani, 1996).

**Regresyon Ağaçları:** Karar ağaçlarının regresyon problemlerine uyarlanmış hali olan regresyon ağaçları, veri setini tekrarlı olarak böler ve her bir alt kümede ortalama çıktı değerini tahmin eder. Eğitim verisi analizinde, öğrenci notlarını tahmin etme, öğrenme süreçlerini modelleme ve eğitimsel müdahalelerin etkisini değerlendirme amacıyla kullanılmaktadır (Breiman vd., 1984).

Destek Vektör Regresyonu (SVR): SVM'nin regresyon problemlerine uyarlanmış hali olan SVR, belirlenen bir marj (epsilon) içinde kalan hataları tolere ederek, çıktı değerlerini tahmin eder. Eğitim verisi analizinde, öğrenci performansını ve gelişimini tahmin etmek için kullanılmaktadır. SVR, özellikle doğrusal olmayan ilişkileri modellemede ve aykırı değerlere karşı güçlü olmasıyla bilinir (Drucker vd., 1997).

Regresyon algoritmalarının eğitim verisi analizindeki uygulamaları, öğrenci başarısını tahmin etme, öğrenme eğrilerini modelleme, eğitimsel müdahalelerin etkisini ölçme ve eğitimsel kaynakların optimizasyonu gibi alanlarda yoğunlaşmaktadır.

### 1.4.3. Topluluk Öğrenme Yöntemleri

Topluluk öğrenme (ensemble learning) yöntemleri, birden fazla öğrenme algoritmasının birleştirilerek, tek bir algoritmanın performansından daha iyi tahmin performansı elde etmeyi amaçlayan makine öğrenmesi yaklaşımlarıdır. Bu yöntemler, genellikle tahmin doğruluğunu artırmak, aşırı uyum (overfitting) problemini azaltmak ve modelin genelleme yeteneğini iyileştirmek için kullanılır (Dietterich, 2000).

Rastgele Orman (Random Forest): Breiman (2001) tarafından geliştirilen rastgele orman algoritması, çok sayıda karar ağacının bootstrap örnekleme ve rastgele özellik seçimi kullanılarak eğitilmesi ve sonuçlarının birleştirilmesi prensibine dayanır. Eğitim verisi analizinde, öğrenci performansını tahmin etme, akademik risk faktörlerini belirleme ve öğrenme davranışlarını modelleme gibi uygulamalarda yaygın olarak kullanılmaktadır.

Costa ve arkadaşları (2017), Portekiz'deki ortaokul öğrencilerinin matematik ve Portekizce derslerindeki başarılarını tahmin etmek için rastgele orman algoritmasını kullanmıştır. Çalışma, rastgele ormanın yüksek tahmin doğruluğu sağladığını ve öğrencilerin daha önceki akademik performansları, devamsızlık oranları ve ebeveyn eğitim düzeyinin en önemli tahmin edici faktörler olduğunu göstermiştir.

Gradyan Artırma (Gradient Boosting): Bu algoritma, zayıf öğrencileri (genellikle karar ağaçları) ardışık olarak eğitir ve her bir yeni model, önceki modellerin hatalarını azaltmayı hedefler. Eğitim verisi analizinde, öğrenci başarısını tahmin etme, okulu bırakma riskini belirleme ve eğitimsel müdahalelerin etkisini değerlendirme gibi uygulamalarda kullanılmaktadır (Friedman, 2001).

XGBoost, LightGBM ve CatBoost gibi modern gradyan artırma uygulamaları, yüksek performans ve hesaplama verimliliği sağlayarak, eğitimsel veri analizinde giderek daha fazla kullanılmaktadır.

AdaBoost: Adaptif artırma (adaptive boosting) olarak da bilinen bu algoritma, her bir yeni modele, önceki modellerin yanlış sınıflandırdığı örnekler daha fazla ağırlık vererek, zayıf öğrencileri güçlü bir sınıflandırıcıya dönüştürür. Eğitim verisi analizinde, öğrenci performansını tahmin etme ve sınıflandırma problemlerinde kullanılmaktadır (Freund & Schapire, 1997).

Bagging (Bootstrap Aggregating): Bu yöntem, veri setinin bootstrap örnekleri üzerinde aynı tür öğrenme algoritmasının eğitilmesi ve sonuçların birleştirilmesi prensibine dayanır. Eğitim verisi analizinde, öğrenci performansını tahmin etme ve modelin varyansını azaltma amacıyla kullanılmaktadır (Breiman, 1996).

Stacking (İstifleme): Bu topluluk yöntemi, farklı tipteki birden fazla öğrenme algoritmasını bir arada kullanır ve meta-öğrenci olarak adlandırılan başka bir algoritma, bu modellerin tahminlerini girdi olarak alarak nihai tahmini yapar. Eğitim verisi analizinde, farklı algoritmaların güçlü yönlerinden yararlanarak, tahmin doğruluğunu artırmak için kullanılmaktadır (Wolpert, 1992).

Xing ve arkadaşları (2015), MOOC platformlarında öğrencilerin kursu tamamlama durumlarını tahmin etmek için stacking tabanlı bir topluluk öğrenme yaklaşımı geliştirmiş ve önerilen yaklaşımın, tek bir algoritma kullanımına göre daha yüksek performans gösterdiğini ortaya koymuştur.

Topluluk öğrenme yöntemlerinin eğitim verisi analizindeki uygulamaları, öğrenci başarısını tahmin etme, eğitimsel risk faktörlerini belirleme, öğrenme davranışlarını modelleme ve eğitimsel müdahalelerin etkinliğini değerlendirme gibi alanlarda başarıyla kullanılmaktadır.

#### 1.4.4. Derin Öğrenme Yaklaşımları

Derin öğrenme, çok katmanlı yapay sinir ağları kullanarak verilerdeki karmaşık örüntüleri öğrenmeyi amaçlayan makine öğrenmesi yöntemlerinin bir alt kümesidir. Bu yöntemler, özellikle büyük miktarda veri ve karmaşık problemlerde etkili olmaktadır (LeCun, Bengio, & Hinton, 2015).

Yapay Sinir Ağları (Artificial Neural Networks - ANN): Beynin çalışma prensibinden esinlenerek geliştirilmiş hesaplama modelleri olan yapay sinir ağları, girdi, gizli ve çıktı katmanlarında bulunan yapay nöronlardan oluşur. Eğitim verisi analizinde, öğrenci performansını tahmin etme, öğrenme davranışlarını modelleme ve karmaşık eğitimsel örüntüleri keşfetme gibi uygulamalarda kullanılmaktadır (Haykin, 2009).

Livieris ve arkadaşları (2012), öğrencilerin matematik performansını tahmin etmek için çeşitli yapay sinir ağı modellerini kullanmış ve bu modellerin, özellikle risk altındaki öğrencileri belirlemede etkili olduğunu göstermiştir.

Derin Sinir Ağları (Deep Neural Networks - DNN): Çok sayıda gizli katman içeren yapay sinir ağları olan DNN'ler, verilerdeki hiyerarşik yapıları öğrenmede oldukça etkilidir. Eğitim verisi analizinde, öğrenci performansını tahmin etme, eğitimsel içerik önerisi sunma ve öğrenme davranışlarını modelleme gibi uygulamalarda kullanılmaktadır (Goodfellow, Bengio, & Courville, 2016).

Evrişimsel Sinir Ağları (Convolutional Neural Networks - CNN): Görüntü işleme ve sınıflandırma problemlerinde yaygın olarak kullanılan CNN'ler, eğitim verisi analizinde, öğrencilerin yazılı cevaplarını değerlendirme, eğitimsel görüntüleri analiz etme ve görsel öğrenme materyallerini sınıflandırma gibi uygulamalarda kullanılmaktadır.

Piech ve arkadaşları (2015), öğrencilerin kod yazma süreçlerini modellemek ve otomatik geri bildirim sağlamak için derin öğrenme yaklaşımlarını kullanmıştır. Çalışma, derin öğrenme modellerinin, öğrencilerin programlama hatalarını tespit etme ve kişiselleştirilmiş öneriler sunmada etkili olduğunu göstermiştir.

Tekrarlayan Sinir Ağları (Recurrent Neural Networks - RNN): Zamansal verileri modellemek için tasarlanmış RNN'ler, eğitim verisi analizinde, öğrenci öğrenme yollarını tahmin etme, öğrenme süreçlerini modelleme ve öğrenci davranışlarındaki zamansal değişimleri analiz etme gibi uygulamalarda kullanılmaktadır. LSTM (Long Short-Term Memory) ve GRU (Gated Recurrent Unit) gibi gelişmiş RNN mimarileri, uzun vadeli bağımlılıkları öğrenmede daha etkilidir (Hochreiter & Schmidhuber, 1997).

Yang ve arkadaşları (2017), MOOC platformlarındaki öğrenci etkileşim verilerini kullanarak, kurs tamamlama durumlarını tahmin etmek için LSTM tabanlı bir model geliştirmiştir. Çalışma, tekrarlayan sinir ağlarının, öğrenci davranışlarındaki

zamansal deęişimleri modellemede ve gelecekteki performansı tahmin etmede başarılı olduğunu göstermiştir.

Otokodlayıcılar (Autoencoders): Denetimsiz öğrenme için kullanılan bir derin öğrenme mimarisi olan otokodlayıcılar, veri boyutunu azaltma (dimensionality reduction) ve özellik çıkarma amacıyla kullanılır. Eğitim verisi analizinde, öğrenci davranışlarındaki örüntüleri keşfetme, anormallik tespiti ve verileri daha düşük boyutlu bir uzayda temsil etme gibi uygulamalarda kullanılmaktadır (Goodfellow vd., 2016).

Derin öğrenme yaklaşımlarının eğitim verisi analizindeki uygulamaları henüz sınırlı olmakla birlikte, bu yöntemlerin, büyük ölçekli ve karmaşık eğitim verilerindeki gizli örüntüleri keşfetme potansiyeli nedeniyle, gelecekte daha yaygın kullanım alanı bulması beklenmektedir.

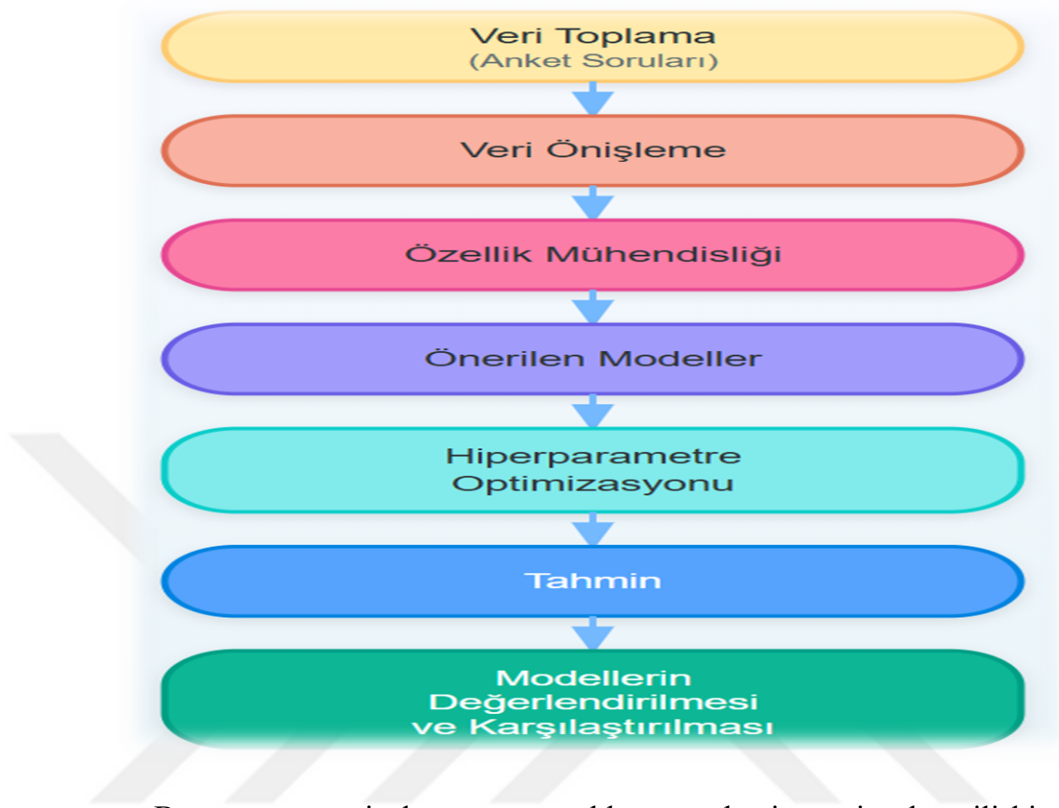
Eğitim verisi analizinde makine öğrenmesi ve derin öğrenme yöntemlerinin kullanımı, öğrenci başarısını tahmin etme, kişiselleştirilmiş öğrenme deneyimleri sunma, eğitimsel risk faktörlerini belirleme ve eğitim politikalarını şekillendirme gibi alanlarda değerli içgörüler sağlamaktadır. Bu araştırmada, Bursa il merkezinde ve ilçelerinde yer alan BİLSEM'lerdeki üstün yetenekli öğrencilerin matematik başarısını etkileyen faktörleri analiz etmek ve tahmin etmek için, yukarıda bahsedilen makine öğrenmesi yöntemlerinden bazıları kullanılacak ve performansları karşılaştırılacaktır.



## İKİNCİ BÖLÜM YÖNTEM

## 2.1. ARAŞTIRMA MODELİ

Şekil 2.1: İş Akışı Mimarisi (Workflow Architecture)



Bu araştırma, nicel araştırma yaklaşımını benimsemiş olup, ilişkisel tarama modelinde bir çalışmadır. İlişkisel tarama modeli, iki veya daha fazla değişken arasındaki ilişkileri belirlemek ve bu ilişkilerin düzeyini ölçmek amacıyla kullanılan bir araştırma modelidir (Karasar, 2012). Bu çalışmada, üstün yetenekli öğrencilerin matematik başarıları ile çeşitli demografik, ailesel, eğitimsel ve sosyo-ekonomik faktörler arasındaki ilişkiler incelenmektedir.

Araştırma, veri madenciliği ve makine öğrenmesi tekniklerinin kullanıldığı bir eğitimsel veri analizi çalışması olarak da değerlendirilebilir. Bu bağlamda, çalışmanın metodolojik çerçevesi, Cross-Industry Standard Process for Data Mining (CRISP-DM) modeline dayanmaktadır (Chapman vd., 2000). CRISP-DM, veri madenciliği projelerinin yürütülmesinde yaygın olarak kullanılan ve iş anlayışı, veri anlayışı, veri hazırlama, modelleme, değerlendirme ve dağıtım olmak üzere altı aşamadan oluşan sistematik bir yaklaşımdır.

Araştırma süreci şu aşamaları içermektedir:

**Araştırma Probleminin Tanımlanması ve Literatür Taraması:** Üstün yetenekli öğrencilerin matematik başarısını etkileyen faktörlerin belirlenmesi ve bu faktörlerin matematik başarısını ne düzeyde tahmin edebildiğinin incelenmesi amacıyla, ilgili literatür taranmıştır.

**Veri Toplama:** Bursa BİLSEM'lerde eğitim gören 151 üstün yetenekli öğrenciye anket uygulanarak veriler toplanmıştır.

**Veri Ön İşleme:** Toplanan veriler, eksik değerlerin ele alınması, aykırı değerlerin belirlenmesi, kategorik verilerin sayısallaştırılması ve özellik normalizasyonu gibi işlemlerden geçirilmiştir.

**Özellik Mühendisliği:** Veri setindeki özelliklerin çıkarılması, seçilmesi ve dönüştürülmesi işlemleri gerçekleştirilmiştir.

**Makine Öğrenmesi Modellerinin Geliştirilmesi:** Lojistik regresyon, karar ağaçları, destek vektör makineleri, rastgele orman, gradyan artırma ve derin sinir ağları gibi çeşitli makine öğrenmesi modelleri geliştirilmiştir.

**Model Değerlendirme ve Karşılaştırma:** Geliştirilen modellerin performansları, çeşitli değerlendirme metrikleri kullanılarak ölçülmüş ve karşılaştırılmıştır.

**Sonuçların Yorumlanması ve Raporlama:** Analiz sonuçları yorumlanarak, üstün yetenekli öğrencilerin matematik başarısını etkileyen faktörler belirlenmiş ve bu alandaki eğitim uygulamaları için öneriler geliştirilmiştir.

Bu araştırma modeli, hem geleneksel istatistiksel yaklaşımların hem de güncel makine öğrenmesi tekniklerinin bir arada kullanılmasına olanak sağlamakta, böylece üstün yetenekli öğrencilerin matematik başarısını etkileyen faktörlerin kapsamlı bir şekilde analiz edilmesine imkân vermektedir.

## **2.2. ÇALIŞMA GRUBU**

Bu araştırmanın çalışma grubunu, Bursa ilindeki 10 farklı Bilim ve Sanat Merkezi'nde (BİLSEM) genel zihinsel yetenek alanında eğitim gören ve gönüllü olarak araştırmaya katılan 151 üstün yetenekli öğrenci oluşturmaktadır. Öğrenciler, Millî Eğitim Bakanlığı tarafından uygulanan tanılama süreçleriyle belirlenmiş ve çeşitli gelişim aşamalarında (destek eğitimi, bireysel yetenekleri fark ettirme, özel yetenekleri geliştirme

ve proje üretimi) eğitim almaktadır. Örneklem, yaş, cinsiyet ve sınıf düzeyi açısından temsili bir dağılım gözetilerek oluşturulmuştur.

### 2.2.1. Bursa BİLSEM Öğrencileri ve Özellikleri

Bu araştırmanın çalışma grubunu, Bursa ilinde yer alan aşağıdaki 10 Bilim ve Sanat Merkezi'nde (BİLSEM) eğitim gören 151 üstün yetenekli öğrenci oluşturmaktadır:

BTSO Kamil Tolon Bilim ve Sanat Merkezi (Osmangazi)

Osmangazi Bilim ve Sanat Merkezi (Osmangazi)

Halil İnalçık Bilim ve Sanat Merkezi (Nilüfer)

Kestel Bilim ve Sanat Merkezi (Kestel)

Gemlik Bilim ve Sanat Merkezi (Gemlik)

Gürsu Bilim ve Sanat Merkezi (Gürsu)

İnegöl Bilim ve Sanat Merkezi (İnegöl)

İznik Bilim ve Sanat Merkezi (İznik)

Karacabey Bilim ve Sanat Merkezi (Karacabey)

Mustafakemalpaşa Bilim ve Sanat Merkezi (Mustafakemalpaşa)

Bu BİLSEM'ler, Bursa ili genelinde üstün yetenekli öğrencilere hizmet veren eğitim kurumlarıdır. Merkezlere kabul edilen öğrenciler, Millî Eğitim Bakanlığı tarafından belirlenmiş tanılama süreçlerinden geçerek üstün yetenekli olarak tanılanmışlardır.

Bursa'daki BİLSEM'lerde eğitim gören öğrenciler, genel zihinsel yetenek, görsel sanatlar, müzik ve diğer özel yetenek alanlarında programlara katılmaktadır. Bu çalışmada yer alan öğrenciler, genel zihinsel yetenek alanında eğitim alan ve çeşitli destek eğitim programları, bireysel yetenekleri fark ettirme programı (BYF), özel yetenekleri geliştirme programı (ÖYG) ve proje üretimi aşamalarında bulunan öğrencilerdir.

### 2.2.2. Örneklem Seçimi ve Demografik Dağılım

Araştırmanın örneklemini, Bursa'daki BİLSEM'lerde genel zihinsel yetenek alanında eğitim gören ve araştırmaya katılmayı kabul eden 151 öğrenciden oluşmaktadır.

Örneklem seçiminde, gönüllülük esası dikkate alınmış ve çalışmaya katılmayı kabul eden, veli izni alınan tüm öğrenciler araştırmaya dâhil edilmiştir.

Çalışma grubunda yer alan öğrencilerin demografik özelliklerine ilişkin bilgiler Tablo 2.1'de sunulmuştur.

**Tablo 2.1: Çalışma Grubunun Demografik Özellikleri**

| Demografik Özellik       | Kategori      | f  | %    |
|--------------------------|---------------|----|------|
| Cinsiyet                 | Kız           | 73 | 48.3 |
|                          | Erkek         | 78 | 51.7 |
| Yaş                      | 10            | 30 | 19.9 |
|                          | 11            | 38 | 25.2 |
|                          | 12            | 42 | 27.8 |
|                          | 13            | 27 | 17.9 |
|                          | 14            | 14 | 9.3  |
| Sınıf                    | 5. Sınıf      | 42 | 27.8 |
|                          | 6. Sınıf      | 47 | 31.1 |
|                          | 7. Sınıf      | 38 | 25.2 |
|                          | 8. Sınıf      | 24 | 15.9 |
| BİLSEM Program Aşamaları | Destek Eğitim | 56 | 37.1 |
|                          | BYF           | 52 | 34.4 |
|                          | ÖYG           | 31 | 20.5 |
|                          | Proje Üretimi | 12 | 7.9  |

Tablo 2.1'de görüldüğü gibi, çalışma grubundaki öğrencilerin %48.3'ü (n=73) kız, %51.7'si (n=78) erkek öğrencilerden oluşmaktadır. Öğrencilerin yaş dağılımı incelendiğinde, %19.9'unun 10 yaşında, %25.2'sinin 11 yaşında, %27.8'inin 12 yaşında, %17.9'unun 13 yaşında ve %9.3'ünün 14 yaşında olduğu görülmektedir. Sınıf düzeylerine göre dağılım ise, %27.8'i 5. sınıf, %31.1'i 6. sınıf, %25.2'si 7. sınıf ve %15.9'u 8. sınıf öğrencilerinden oluşmaktadır. BİLSEM program aşamalarına göre dağılıma bakıldığında, öğrencilerin %37.1'i destek eğitim, %34.4'ü BYF, %20.5'i ÖYG ve %7.9'u proje üretimi aşamasında bulunmaktadır.

Çalışma grubundaki öğrencilerin matematik başarı düzeyleri, anket formunda belirttikleri matematik not ortalamalarına göre belirlenmiştir. Öğrencilerin matematik başarı düzeylerine göre dağılımı Tablo 2.2'de sunulmuştur.

**Tablo 2.2: Öğrencilerin Matematik Başarı Düzeylerine Göre Dağılımı**

| Matematik Not Ortalaması | f  | %    |
|--------------------------|----|------|
| A (90-100)               | 98 | 64.9 |
| B (80-89)                | 38 | 25.2 |
| C (70-79)                | 12 | 7.9  |
| D (60-69)                | 3  | 2.0  |
| F (59 ve altı)           | 0  | 0.0  |

Tablo 2.2'de görüldüğü gibi, çalışma grubundaki öğrencilerin büyük çoğunluğu (%64.9) matematik dersinde A (90-100) düzeyinde başarıya sahiptir. Öğrencilerin %25.2'si B (80-89), %7.9'u C (70-79) ve %2.0'ı D (60-69) düzeyinde matematik başarı göstermektedir. Çalışma grubunda, matematik not ortalaması F (59 ve altı) düzeyinde olan öğrenci bulunmamaktadır. Bu durum, BİLSEM'e seçilen üstün yetenekli öğrencilerin genel olarak yüksek matematik başarısına sahip olduğunu göstermektedir.

### 2.3. VERİ TOPLAMA ARAÇLARI

Araştırmada veri toplamak amacıyla, öğrencilerin bireysel, ailesel, eğitimsel ve sosyo-ekonomik özelliklerini belirlemeye yönelik yapılandırılmış bir anket formu kullanılmıştır. Anket, ilgili literatür taranarak hazırlanmış ve uzman görüşleri doğrultusunda kapsam geçerliliği sağlanmıştır. Ayrıca, öğrencilerin matematik başarı verileri, öz-bildirim yoluyla elde edilen not ortalamalarına dayanmaktadır.

#### 3.3.1. Anket Yapısı ve Uygulanması

Bu araştırmada veri toplama aracı olarak, araştırmacı tarafından literatür taraması sonucunda geliştirilen "Üstün Yetenekli Öğrencilerde Matematik Başarısını Etkileyen Faktörler Anketi" kullanılmıştır. Anket, öğrencilerin demografik özelliklerini, matematik başarılarını, çalışma alışkanlıklarını, aile desteği ve yapısını, öğrenme

stillerini, matematik tutumlarını, teknoloji kullanımlarını ve sosyal aktivitelerini belirlemeye yönelik toplam 38 sorudan oluşmaktadır.

Anket formu geliştirme sürecinde, öncelikle ilgili literatür taranmış ve matematik başarısını etkileyebilecek faktörler belirlenmiştir. Daha sonra, bu faktörleri ölçmeye yönelik sorular hazırlanmıştır. Anket formunun kapsam geçerliliğini sağlamak amacıyla, üç alan uzmanının (bir matematik eğitimi uzmanı, bir üstün yetenekliler eğitimi uzmanı ve bir ölçme-değerlendirme uzmanı) görüşlerine başvurulmuştur. Uzman görüşleri doğrultusunda gerekli düzeltmeler yapılarak anket formuna son şekli verilmiştir.

Anket, aşağıdaki bölümlerden oluşmaktadır:

**Kişisel Bilgiler:** Yaş, cinsiyet, sınıf düzeyi gibi demografik bilgileri içeren 3 soru.

**Matematikle İlgili Bilgiler:** Matematik not ortalaması, kullanılan kaynaklar gibi matematiğe ilişkin temel bilgileri içeren 2 soru.

**Çalışma Alışkanlıkları ve Tutumlar:** Matematik çalışma süresi, matematiğe ilgi düzeyi, problem çözme sırasında hissedilen duygular gibi çalışma alışkanlıkları ve matematik tutumlarını belirlemeye yönelik 3 soru.

**Matematik Başarıları ve Katılımlar:** Matematik yarışmalarına katılım, ödül durumu gibi matematik alanındaki başarı ve etkinlikleri içeren 3 soru.

**Aile Bilgileri ve Destek:** Ailede yaşayan kişi sayısı, evde matematik alanında çalışan yetişkin varlığı, aile desteği gibi ailesel faktörleri içeren 5 soru.

**Öğrenme Stilleri ve Tercihleri:** Matematik öğrenirken tercih edilen yöntemler, sevilen matematik konuları gibi öğrenme stillerine ilişkin bilgileri içeren 2 soru.

**Zorluklar ve Engeller:** Matematik öğrenmede karşılaşılan zorluklar ve bunlarla başa çıkma stratejileri gibi zorluk faktörlerini içeren 2 soru.

**Teknoloji Kullanımı:** Matematik öğreniminde teknoloji kullanım sıklığı ve kullanılan teknolojik araçlar gibi teknoloji faktörlerini içeren 2 soru.

**Akademik ve Kişisel Hedefler:** Akademik beklentiler, ilgi alanları gibi hedef ve ilgileri içeren 2 soru.

**Diğer Faktörler:** Oyun oynama süresi, uyku saati, dijital okuryazarlık, sosyal aktiviteler gibi çeşitli faktörleri içeren 15 soru.

Anket formunda, kapalı uçlu sorular (çoktan seçmeli, evet-hayır vb.) ve açık uçlu sorular bir arada kullanılmıştır. Kapalı uçlu sorularda, öğrencilerin kendilerine en uygun seçeneği işaretlemeleri istenmiştir. Açık uçlu sorularda ise, öğrencilerin kendi ifadeleriyle yanıt vermeleri beklenmiştir.

Anketin uygulanması öncesinde, gerekli etik kurul izinleri ve Bursa İl Millî Eğitim Müdürlüğü'nden araştırma izni alınmıştır. Uygulama, Bursadaki BİLSEM'lerde, öğrencilerin normal program saatleri içinde, araştırmacının gözetiminde gerçekleştirilmiştir. Uygulama öncesinde, öğrencilere ve velilerine araştırmanın amacı açıklanmış ve gönüllü katılım formları imzalatılmıştır. Anketin doldurulması yaklaşık 20-25 dakika sürmüştür.

**Şekil 2.2:** Veri toplama Aracının Kategorik Yapısı



### 2.3.2. Anket Uygulama Süreci

Bursa'daki BİLSEM'lerde anket uygulama süreci, 2024-2025 eğitim-öğretim yılının birinci döneminde (Ekim-Aralık 2024) gerçekleştirilmiştir. Uygulama öncesinde, Bursa ilinde yer alan BİLSEM'lerin yönetimleri ile görüşmeler yapılmış; araştırmanın amacı, kapsamı ve uygulama süreci hakkında kendilerine bilgi verilmiştir.

Anket uygulaması, BİLSEM'in program yapısı dikkate alınarak, öğrencilerin eğitim gördükleri günlerde ve sınıflarda gerçekleştirilmiştir. Uygulama, araştırmacı ve BİLSEM öğretmenleri tarafından yürütülmüştür. Uygulama öncesinde, öğrencilere araştırmanın amacı ve anketin nasıl doldurulacağı hakkında bilgi verilmiştir. Öğrencilerin, anketi bireysel olarak ve kendi düşüncelerini yansıtacak şekilde doldurmaları sağlanmıştır.

Anket uygulaması sırasında, öğrencilerin anlayamadıkları soruları araştırmacıya sormaları teşvik edilmiş ve gerekli açıklamalar yapılmıştır. Uygulama sırasında herhangi bir sorun yaşanmamış ve öğrenciler anketi samimi bir şekilde doldurmuşlardır.

### 2.3.3. Veri Seti Özellikleri

Araştırma sonucunda, Bursadaki BİLSEM'lerde eğitim gören 151 üstün yetenekli öğrenciden oluşan bir veri seti elde edilmiştir. Veri seti, anket formunda yer alan 38 soruya verilen yanıtlardan oluşmaktadır. Veri setindeki değişkenler, türlerine göre aşağıdaki gibi gruplandırılabilir:

Demografik Değişkenler: Yaş, cinsiyet, sınıf düzeyi.

Bağımlı Değişken: Matematik başarıları (not ortalaması).

Bağımsız Değişkenler:

- Çalışma alışkanlıkları (çalışma süresi, kaynaklar)
- Matematik tutumları (ilgi, duygular)
- Matematik etkinlikleri (yarışmalar, ödüller)
- Ailesel faktörler (aile büyüklüğü, destek, eğitim düzeyi)
- Öğrenme stilleri ve tercihleri
- Zorluklar ve başa çıkma stratejileri
- Teknoloji kullanımı

- Akademik hedefler ve ilgi alanları
- Sosyal aktiviteler
- Uyku düzeni
- Dijital okuryazarlık

Veri setindeki değişkenlerin çoğunluğu kategorik (nominal veya ordinal) düzeydedir. Örneğin, cinsiyet, sınıf düzeyi, matematik ilgi düzeyi, aile desteği gibi değişkenler kategorik yapıdadır. Yaş, çalışma süresi, aile büyüklüğü gibi değişkenler ise sayısal (interval veya ratio) düzeydedir.

Veri setinde, bağımlı değişken olan matematik başarısı, öğrencilerin bildirdiği matematik not ortalamaları (A, B, C, D, F) şeklinde kategorik olarak elde edilmiştir. Veri analizi aşamasında, bu kategorik değişken, sayısal değerlere (A=5, B=4, C=3, D=2, F=1) dönüştürülerek kullanılmıştır.

## 2.4. VERİ ÖN İŞLEME

Veri ön işleme, makine öğrenmesi ve veri analizinde kritik bir aşamadır. Ham verinin, modelleme için uygun hale getirilmesi ve veri kalitesinin artırılması amacıyla çeşitli ön işleme adımları uygulanır. Bu çalışmada kullanılan veri ön işleme adımları aşağıda detaylı olarak açıklanmıştır.

### 2.4.1. Eksik Veri Analizi

Eksik veriler, ankette cevaplanmayan sorular veya katılımcının cevap vermek istemediği sorular sonucunda ortaya çıkabilir. Eksik verilerin analizi ve ele alınması, veri kalitesi ve modelleme performansı açısından önemlidir.

Bu çalışmada, veri setindeki eksik değerler, Python programlama dili ve pandas kütüphanesi kullanılarak analiz edilmiştir. Her bir değişken için eksik değer sayısı ve oranı hesaplanmıştır. Eksik değerlerin, rastgele mi yoksa belirli bir düzene göre mi dağıldığı incelenmiştir.

Eksik değerlerin ele alınmasında izlenen stratejiler şunlardır:

Az Sayıda Eksik Değer İçeren Durumlar: Değişkendeki eksik değer oranı %5'ten az ise, eksik değerler, ilgili değişkenin merkezi eğilim ölçüleri (ortalama, medyan veya

mod) ile doldurulmuştur. Sayısal değişkenler için ortalama veya medyan, kategorik değişkenler için mod kullanılmıştır.

**Çok Sayıda Eksik Değer İçeren Durumlar:** Değişkendeki eksik değer oranı %20'den fazla ise, ilgili değişkenin analizden çıkarılması değerlendirilmiştir. %5-%20 arasında eksik değere sahip değişkenler için, veri yapısına göre çeşitli stratejiler (k-en yakın komşu imputasyonu, çoklu imputasyon vb.) uygulanmıştır.

**Eksik Değerlerin Modellenmesi:** Bazı durumlarda, eksik değer olması başlı başına bir bilgi olabilir. Bu durumda, eksik değerler için ayrı bir kategori oluşturulmuş veya eksik değer olup olmadığını gösteren yeni bir ikili değişken (dummy variable) eklenmiştir.

Eksik veri analizi sonucunda, veri setindeki eksik değer oranının düşük olduğu ve eksik değerlerin rastgele dağıldığı tespit edilmiştir. Bu nedenle, eksik değerler, değişkenin türüne uygun merkezi eğilim ölçüleri ile doldurulmuştur.

#### 2.4.2. Aykırı Değer Tespiti

Aykırı değerler, diğer gözlemlerden önemli ölçüde farklılık gösteren ve veri setinin genel özelliklerini temsil etmeyen değerlerdir. Aykırı değerler, veri toplama hataları, ölçüm hataları veya doğal varyasyonlar sonucunda ortaya çıkabilir.

Bu araştırmada, aykırı değerlerin tespit edilmesi için aşağıdaki yöntemler kullanılmıştır:

**İstatistiksel Yöntemler:** Sayısal değişkenler için z-skorları, kutu grafikleri (box plots) ve yüzdelik dilimler (percentiles) kullanılarak aykırı değerler tespit edilmiştir. Örneğin, z-skoru 3'ten büyük veya -3'ten küçük olan değerler, potansiyel aykırı değerler olarak değerlendirilmiştir.

**Görselleştirme Teknikleri:** Veri dağılımını görselleştirmek için histogram, kutu grafikleri ve saçılım grafikleri gibi teknikler kullanılmıştır. Bu grafikler, aykırı değerlerin görsel olarak tanımlanmasına yardımcı olmuştur.

**Çok Değişkenli Yöntemler:** Tek değişkenli analiz, her bir değişkendeki aykırı değerleri tespit etmekle birlikte, değişkenler arasındaki ilişkileri dikkate almaz. Bu nedenle, çok değişkenli aykırı değerleri tespit etmek için Mahalanobis uzaklığı ve Isolation Forest gibi yöntemler kullanılmıştır.

Tespit edilen aykırı değerlerin ele alınmasında izlenen stratejiler şunlardır:

**Doğrulama:** Aykırı değerlerin gerçek gözlemler mi yoksa veri toplama hataları mı olduğu doğrulanmıştır. Veri toplama hatası olduğu tespit edilen değerler, düzeltilmiş veya eksik veri olarak işaretlenmiştir.

**Dönüştürme:** Bazı durumlarda, değişkenlere logaritmik, kare kök veya Box-Cox dönüşümü gibi dönüşümler uygulanarak, aykırı değerlerin etkisi azaltılmıştır.

**Kırpma (Trimming):** Bazı aykırı değerler, belirlenen bir eşik değere kırılmıştır. Örneğin, 95. yüzdelik dilimin üzerindeki değerler, 95. yüzdelik dilim değerine eşitlenmiştir.

**Ayrı Analiz:** Bazı durumlarda, aykırı değerler veri setinden çıkarılmış, ancak ayrı bir analiz ile bu değerlerin nedenleri ve etkileri araştırılmıştır.

Aykırı değer analizi sonucunda, veri setinde az sayıda aykırı değer tespit edilmiş ve bu değerler, yukarıda belirtilen stratejiler kullanılarak ele alınmıştır.

Aykırı Değer Tespiti Eğitimde veri madenciliğinin temel amaçları şunlardır:

**Öğrenci Modellemesi:** Öğrencilerin bilgi düzeylerini, öğrenme stillerini, motivasyonlarını ve davranışlarını modelleyerek kişiselleştirilmiş öğrenme deneyimleri sunmak.

**Öğrenme Süreçlerini Anlama:** Öğrencilerin nasıl öğrendiğini, hangi stratejilerin etkili olduğunu ve öğrenme sürecinde karşılaşılan zorlukları belirlemek.

**Eğitimsel Yazılımların Etkinliğini Değerlendirme:** Eğitimsel yazılımların ve öğrenme ortamlarının etkinliğini analiz ederek iyileştirmeler yapmak.

**Öğrenci Başarısını Tahmin Etme:** Öğrencilerin gelecekteki akademik performanslarını tahmin ederek, erken müdahale stratejileri geliştirmek.

**Bilimsel Keşifler:** Eğitim süreçleri ve öğrenme ile ilgili yeni bilgiler elde etmek (Romero & Ventura, 2020).

Eğitimde veri madenciliği araştırmaları, son yıllarda büyük önem kazanmıştır. Baker ve Inventado (2014), EDM'nin eğitimsel araştırmalarda kullanılan temel yöntemlerini şu şekilde sınıflandırmıştır: tahmin (prediction), kümeleme (clustering), ilişki madenciliği (relationship mining), keşifsel veri analizi (discovery with models) ve veri görselleştirme (visualization).

Eğitimsel veri madenciliği araştırmalarında sıklıkla kullanılan veri kaynakları arasında, öğrencilerin demografik bilgileri, akademik geçmişleri, öğrenme yönetim sistemlerindeki (LMS) log verileri, çevrimiçi kurslardaki etkileşimler, standart test sonuçları ve anket verileri bulunmaktadır (Peña-Ayala, 2014).

Üstün yetenekli öğrenciler özelinde, eğitimde veri madenciliği teknikleri, bu öğrencilerin erken tanınması, öğrenme ihtiyaçlarının belirlenmesi, potansiyellerini en üst düzeyde gerçekleştirmelerine yönelik kişiselleştirilmiş öğrenme programlarının geliştirilmesi ve akademik gelişimlerinin izlenmesi amacıyla kullanılabilir (Sak vd., 2021).

#### 2.4.3. Kategorik Verilerin Sayısallaştırılması

Makine öğrenmesi algoritmalarının çoğu, sayısal verilerle çalışmak üzere tasarlanmıştır. Bu nedenle, kategorik değişkenlerin sayısal değerlere dönüştürülmesi gerekmektedir. Bu araştırmada, kategorik verilerin sayısallaştırılması için aşağıdaki yöntemler kullanılmıştır:

**Etiket Kodlama (Label Encoding):** Sıralı kategorik değişkenler (ordinal variables) için etiket kodlama yöntemi kullanılmıştır. Bu yöntemde, kategorilere 0, 1, 2, ... gibi sayısal değerler atanmıştır. Örneğin, matematik ilgi düzeyi değişkeni için, "Çok düşük"=1, "Düşük"=2, "Orta"=3, "Yüksek"=4, "Çok yüksek"=5 şeklinde kodlama yapılmıştır.

**One-Hot Kodlama (One-Hot Encoding):** Sıralı olmayan kategorik değişkenler (nominal variables) için one-hot kodlama yöntemi kullanılmıştır. Bu yöntemde, her bir kategori için yeni bir ikili değişken (0 veya 1 değerini alan) oluşturulmuştur. Örneğin, cinsiyet değişkeni için "Kadın" ve "Erkek" kategorilerine karşılık gelen iki yeni değişken oluşturulmuştur.

**Binary Kodlama (Binary Encoding):** Çok sayıda kategoriye sahip değişkenler için, one-hot kodlama yerine binary kodlama yöntemi kullanılarak, oluşturulacak değişken sayısı azaltılmıştır. Bu yöntemde, her bir kategori, binary (ikili) formatta kodlanmıştır.

**Hedef Kodlama (Target Encoding):** Bazı durumlarda, kategorilerin bağımlı değişkenle olan ilişkisini dikkate alan hedef kodlama yöntemi kullanılmıştır. Bu

yöntemde, her bir kategoriye, o kategorideki örneklerin bağımlı değişken ortalaması atanmıştır.

Kategorik verilerin sayısallaştırılması sonucunda, veri seti, makine öğrenmesi algoritmaları için uygun formatta temsil edilmiştir.

**Şekil 2.3:** Veri Tipi-Kodlama Yöntemi Eşleştirme Şeması



#### 2.4.4. Özellik Normalizasyonu

Makine öğrenmesi algoritmalarının çoğu, değişkenlerin benzer ölçeklerde olmasından fayda sağlar. Özellik normalizasyonu, değişkenlerin ölçeklerini standartlaştırarak, algoritmaların performansını artırmaya yardımcı olur.

Bu araştırmada, özellik normalizasyonu için aşağıdaki yöntemler kullanılmıştır:

**Min-Max Normalizasyonu:** Bu yöntemde, değişkenler [0, 1] aralığına ölçeklendirilmiştir. Formül olarak:

$$X_{\text{norm}} = (X - X_{\text{min}}) / (X_{\text{max}} - X_{\text{min}})$$

Min-max normalizasyonu, değişkenlerin dağılımını koruyarak, tüm değerleri belirli bir aralığa sıkıştırır.

Z-Skor Standardizasyonu: Bu yöntemde, değişkenler, ortalaması 0, standart sapması 1 olacak şekilde dönüştürülmüştür. Formül olarak:

$$X_{\text{stand}} = (X - X_{\text{mean}}) / X_{\text{std}}$$

Z-skor standardizasyonu, aykırı değerlere karşı daha duyarlıdır, ancak doğrusal model tabanlı algoritmalar için daha uygundur.

Robust Ölçeklendirme: Aykırı değerlere karşı daha dirençli bir yöntem olan robust ölçeklendirme, medyan ve çeyrekler arası aralık (IQR) kullanılarak yapılmıştır. Formül olarak:

$$X_{\text{robust}} = (X - X_{\text{median}}) / \text{IQR}$$

Robust ölçeklendirme, aykırı değerlerin etkisini azaltarak, daha sağlam bir normalizasyon sağlar.

Logaritmik Dönüşüm: Çarpık (skewed) dağılıma sahip değişkenler için logaritmik dönüşüm uygulanmıştır. Formül olarak:

$$X_{\text{log}} = \log(X + c)$$

Burada c, X'in minimum değeri negatif veya sıfır olduğunda kullanılan bir sabittir.

Özellik normalizasyonu, modelin türüne ve veri dağılımına göre değişkenlik gösterebilir. Bu çalışmada, farklı normalizasyon yöntemleri karşılaştırılmış ve her bir model için en uygun normalizasyon yöntemi seçilmiştir.

## 2.5. ÖZELLİK MÜHENDİSLİĞİ

Özellik mühendisliği, ham veriden anlamlı ve modelleme için uygun özelliklerin çıkarılması, seçilmesi ve dönüştürülmesi sürecidir. Etkili özellik mühendisliği, makine öğrenmesi modellerinin performansını önemli ölçüde artırabilir. Bu çalışmada uygulanan özellik mühendisliği teknikleri aşağıda detaylı olarak açıklanmıştır.

### 2.5.1. Özellik Çıkarımı

Özellik çıkarımı, mevcut özelliklerden yeni özelliklerin türetilmesi sürecidir. Bu süreç, veri setindeki gizli örüntüleri ortaya çıkarmaya ve modelin tahmin gücünü artırmaya yardımcı olur.

Bu araştırmada, aşağıdaki özellik çıkarım teknikleri kullanılmıştır:

**Matematiksel Dönüşümler:** Mevcut sayısal özelliklere çeşitli matematiksel işlemler uygulanarak yeni özellikler türetilmiştir. Örneğin:

Matematik çalışma süresi ve matematik ilgi düzeyi değişkenlerinin çarpımı ile "İlgi-Zaman Çarpanı" özelliği oluşturulmuştur.

Yaş değişkeninin karesi alınarak "Yaş Kare" özelliği türetilmiştir.

**Kategorik Özellik Kombinasyonları:** İki veya daha fazla kategorik özelliğin kombinasyonundan yeni özellikler oluşturulmuştur. Örneğin:

Cinsiyet ve sınıf düzeyi özelliklerinin kombinasyonu ile "Cinsiyet-Sınıf" özelliği oluşturulmuştur.

Aile desteği ve öğrenme stili özelliklerinin kombinasyonu ile "Destek-Stil" özelliği türetilmiştir.

**Zaman Tabanlı Özellikler:** Zaman ile ilgili değişkenlerden yeni özellikler türetilmiştir. Örneğin:

Toplam ders çalışma süresi ve matematik çalışma süresi kullanılarak "Matematik-Toplam Oran" özelliği hesaplanmıştır.

Oyun oynama süresi ve uyku süresi kullanılarak "Aktivite Dengesi" özelliği oluşturulmuştur.

**Metin Tabanlı Özellikler:** Açık uçlu sorulara verilen yanıtlardan çeşitli metin özellikleri çıkarılmıştır. Örneğin:

Yanıtlardaki kelime sayısı, cümle uzunluğu, olumlu/olumsuz kelime oranı gibi özellikler hesaplanmıştır.

Belirli anahtar kelimelerin varlığına dayalı ikili özellikler oluşturulmuştur.

Özellik çıkarımı sonucunda, orijinal veri setindeki 38 değişkene ek olarak, yaklaşık 20 yeni özellik türetilmiştir.

### 2.5.2. Özellik Seçimi

Özellik seçimi, model performansına en çok katkıda bulunan özellikleri belirlemek ve gereksiz veya zararlı özellikleri elemek için kullanılan bir tekniktir. Özellik seçimi, modelin karmaşıklığını azaltarak, aşırı uyumu engelleyebilir ve hesaplama verimliliğini artırabilir.

Bu araştırmada, aşağıdaki özellik seçimi yöntemleri kullanılmıştır:

**Filtre Yöntemleri:** Özelliklerin istatistiksel özellikleri temelinde seçim yapan yöntemlerdir. Örneğin:

**Ki-kare testi:** Kategorik özellikler ile hedef değişken arasındaki ilişkiyi değerlendirmek için kullanılmıştır.

**ANOVA F-değeri:** Sayısal özellikler ile kategorik hedef değişken arasındaki ilişkiyi değerlendirmek için kullanılmıştır.

**Pearson korelasyonu:** Sayısal özellikler arasındaki ilişkiyi değerlendirmek için kullanılmıştır. Yüksek korelasyona sahip özelliklerden biri, çoklu doğrusal bağlantı sorununu azaltmak için çıkarılmıştır.

**Sarmalayıcı Yöntemler (Wrapper Methods):** Model performansını kullanarak özellik alt kümelerini değerlendiren yöntemlerdir. Örneğin:

**Özyinelemeli Özellik Eliminasyonu (Recursive Feature Elimination - RFE):** Model eğitildikten sonra önem sırasına göre özellikleri eleme işlemidir.

**İleriye Doğru Seçim (Forward Selection):** Boş bir özellik kümesinden başlayarak, her adımda en iyi özelliği ekleyerek ilerleyen bir yöntemdir.

**Geriye Doğru Eleme (Backward Elimination):** Tüm özelliklerle başlayarak, her adımda en az önemli özelliği çıkaran bir yöntemdir.

**Gömülü Yöntemler (Embedded Methods):** Özellik seçimini model eğitimi sürecinin bir parçası olarak gerçekleştiren yöntemlerdir. Örneğin:

**L1 Regülerizasyonu (Lasso):** Özellik katsayılarını sıfıra indirgeyen bir regülerizasyon tekniğidir.

**Karar ağacı tabanlı önem ölçüleri:** Rastgele orman veya gradyan artırma gibi ağaç tabanlı modellerin özellik önem skorları kullanılmıştır.

Özellik seçimi sonucunda, modelin türüne ve hedefine bağlı olarak, en etkili 15-25 özellik belirlenmiş ve modelleme aşamasında kullanılmıştır.

### 2.5.3. Boyut İndirgeme Teknikleri

Boyut indirgeme, yüksek boyutlu veri setlerini daha düşük boyutlu bir uzaya dönüştürerek, veri setinin karmaşıklığını azaltan ve bilgi kaybını en aza indiren bir tekniktir. Boyut indirgeme, hesaplama verimliliğini artırabilir, görselleştirmeyi kolaylaştırabilir ve "boyutsallık laneti" (curse of dimensionality) problemini hafifletebilir.

Bu çalışmada, aşağıdaki boyut indirgeme teknikleri kullanılmıştır:

**Temel Bileşenler Analizi (Principal Component Analysis - PCA):** En yaygın boyut indirgeme tekniklerinden biri olan PCA, veri setindeki varyansın çoğunu koruyan yeni, doğrusal olarak bağımsız (ortogonal) bileşenler oluşturur. Bu çalışmada, PCA uygulanırken şu adımlar izlenmiştir:

Veri seti standardize edilmiştir (ortalama=0, standart sapma=1).

Kovaryans matrisi hesaplanmıştır.

Kovaryans matrisinin özdeğer ve özvektörleri hesaplanmıştır.

Özdeğerler büyükten küçüğe sıralanmış ve toplam varyansın belirli bir yüzdesini (örneğin %95) açıklayan bileşenler seçilmiştir.

Orijinal veri, seçilen özvektörler üzerine izdüşürülerek boyut indirgenmiştir.

**Doğrusal Ayırt Edici Analiz (Linear Discriminant Analysis - LDA):** Sınıflandırma problemleri için özel olarak tasarlanmış bir boyut indirgeme tekniği olan LDA, sınıflar arası ayrımı maksimize eden ve sınıf içi varyansı minimize eden boyutları bulur. Bu çalışmada, LDA tekniği, özellikle sınıflandırma modelleri için uygulanmıştır.

**t-dağılımlı Stokastik Komşu Gömme (t-distributed Stochastic Neighbor Embedding - t-SNE):** Doğrusal olmayan bir boyut indirgeme tekniği olan t-SNE, yüksek boyutlu verileri 2 veya 3 boyuta indirgeyerek, veri noktaları arasındaki yerel ilişkileri koruyan bir görselleştirme sağlar. Bu çalışmada, t-SNE tekniği, veri setindeki kümelenmeleri ve örüntüleri keşfetmek için kullanılmıştır.

**Otokodlayıcılar (Autoencoders):** Yapay sinir ağları tabanlı bir boyut indirgeme tekniği olan otokodlayıcılar, veriyi daha düşük boyutlu bir "kod" uzayına kodlar ve bu

koddan orijinal veriyi yeniden oluşturmaya çalışır. Bu araştırmada, otokodlayıcılar, özellikle doğrusal olmayan ilişkileri modellemek için kullanılmıştır.

Boyut indirgeme teknikleri, veri setinin yapısına ve modelleme hedeflerine bağlı olarak seçilmiş ve uygulanmıştır. Örneğin, sınıflandırma modelleri için LDA, genel amaçlı modeller için PCA, veri görselleştirmesi için t-SNE ve karmaşık doğrusal olmayan ilişkiler için otokodlayıcılar tercih edilmiştir.

## 2.6. MAKİNE ÖĞRENMESİ MODELLERİ

Makine öğrenmesi modelleri, verilerden örüntüler öğrenerek, gelecekteki gözlemler hakkında tahminler yapmak için kullanılan algoritmalarlardır. Bu araştırmada, üstün yetenekli öğrencilerin matematik başarısını tahmin etmek ve etkileyen faktörleri belirlemek için, çeşitli makine öğrenmesi modelleri geliştirilmiş ve karşılaştırılmıştır.

### 2.6.1. Lojistik Regresyon

Lojistik regresyon, sınıflandırma problemleri için kullanılan bir doğrusal model türüdür. Bağımlı değişkenin ikili (binary) veya çok sınıflı (multi-class) kategorik bir değişken olduğu durumlarda, bağımsız değişkenlerin doğrusal bir kombinasyonu kullanılarak, bir olayın gerçekleşme olasılığını tahmin eder.

Bu araştırmada, matematik başarısını "yüksek başarı" (A ve B) ve "orta-düşük başarı" (C, D ve F) olarak ikili bir sınıflandırma problemi şeklinde modellemek için lojistik regresyon kullanılmıştır. Lojistik regresyon modelinin matematiksel formülasyonu aşağıdaki gibidir:

$$p(y = 1|x) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n)}}$$

Burada:

- $p(y = 1|x)$ : Bağımsız değişkenlerin değerleri verildiğinde, sınıfın 1 (yüksek başarı) olma olasılığıdır.
- $\beta_0, \beta_1, \beta_2, \dots, \beta_n$ : Model parametrelerini (katsayıları) ifade eder.
- $x_1, x_2, \dots, x_n$ : Bağımsız değişkenlerdir.

Lojistik regresyon modeli, maksimum olabilirlik kestirimi (maximum likelihood estimation) ile eğitilmiştir. Overfitting'i önlemek ve modelin genelleme yeteneğini artırmak için L1 (Lasso) ve L2 (Ridge) regülarizasyon teknikleri kullanılmıştır. Regülarizasyon parametresi ( $C$ ), çapraz doğrulama ile optimize edilmiştir.

Lojistik regresyon modelinin avantajları, yorumlanabilirliği ve hesaplama verimliliğidir. Katsayılar, ilgili bağımsız değişkenin bağımlı değişken üzerindeki etkisini gösterir. Ayrıca, model eğitimi ve tahmin hızlıdır. Ancak, doğrusal olmayan ilişkileri modellemekte sınırlıdır ve değişkenler arasındaki etkileşimleri otomatik olarak yakalayamaz.

### 2.6.2. Karar Ağaçları

Karar ağaçları, sınıflandırma ve regresyon problemleri için kullanılan, ağaç yapısına sahip bir karar destek sistemidir. Veri setini, özellik değerlerine göre tekrarlı olarak bölerek, homojen alt kümelere ayırır. Her bir düğüm, bir özelliğe göre test yapar ve her bir dal, test sonucuna göre bir yol temsil eder. Yaprak düğümler ise, tahmin edilen sınıf veya değeri gösterir.

Bu araştırmada, öğrencilerin matematik başarılarını sınıflandırmak ve tahmin etmek için CART (Classification and Regression Trees) algoritması kullanılmıştır. CART algoritması, Gini indeksi veya entropi gibi saflık ölçülerini kullanarak, her adımda veri setini en iyi şekilde bölen özelliği seçer.

Karar ağacı oluşturulurken, aşağıdaki hiperparametreler optimize edilmiştir:

Maksimum derinlik (max\_depth): Ağacın maksimum derinliği.

Minimum örnek bölünmesi (min\_samples\_split): Bir düğümün bölünebilmesi için gereken minimum örnek sayısı.

Minimum yaprak düğüm (min\_samples\_leaf): Bir yaprak düğümde bulunması gereken minimum örnek sayısı.

Maksimum özellik sayısı (max\_features): Her bölünmede dikkate alınan maksimum özellik sayısı.

Karar ağaçlarının avantajları, yorumlanabilirliği, doğrusal olmayan ilişkileri modelleyebilmesi ve etkileşimleri otomatik olarak yakalayabilmesidir. Ayrıca, kategorik

ve sayısal değişkenleri birlikte ele alabilir ve eksik değerlere karşı dirençlidir. Ancak, overfitting'e eğilimlidir ve küçük veri değişikliklerine karşı istikrarsız olabilir.

### 2.6.3. Destek Vektör Makineleri (SVM)

Destek Vektör Makineleri (SVM), sınıflandırma ve regresyon problemleri için kullanılan, veri noktalarını sınıflandırmak için bir hiper düzlem (veya düzlemler seti) arayan bir gözetimli öğrenme algoritmasıdır. SVM, sınıflar arasındaki marjı maksimize ederek, en iyi ayrımı sağlayan hiper düzlemi bulmayı amaçlar.

Bu araştırmada, öğrencilerin matematik başarılarını sınıflandırmak için doğrusal ve doğrusal olmayan (kernel) SVM modelleri kullanılmıştır. Doğrusal olmayan ilişkileri modellemek için, radyal tabanlı fonksiyon (RBF), polinomiyal ve sigmoid gibi çeşitli kernel fonksiyonları denenmiştir.

SVM modelinin matematiksel formülasyonu aşağıdaki gibidir:

$$\min_{w,b,\xi} \frac{1}{2} w^T w + C \sum_{i=1}^n \xi_i$$

s.t.

$$y_i(w^T \phi(x_i) + b) \geq 1 - \xi_i, \quad \xi_i \geq 0, \quad i = 1, 2, \dots, n$$

Burada:

- $w$ : Ağırlık vektörü
- $b$ : Bias (sapma) terimi
- $\xi_i$ : Hata değişkenleri
- $C$ : Regülerizasyon parametresi
- $\phi(x_i)$ : Giriş uzayından özellik uzayına yapılan dönüşüm (haritalama) fonksiyonu
- $y_i$ : Sınıf etiketleri (-1 veya 1)
- $x_i$ : Giriş vektörleri

SVM modeli eğitilirken, aşağıdaki hiperparametreler optimize edilmiştir:

Regülerizasyon parametresi (C): Hata ve marj genişliği arasındaki dengeyi belirler.

Kernel: Doğrusal, RBF, polinomiyal veya sigmoid.

Kernel parametreleri: RBF için gamma, polinomiyal için derece.

SVM'nin avantajları, yüksek boyutlu uzaylarda etkili olması, bellek açısından verimli olması ve farklı kernel fonksiyonları kullanılarak doğrusal olmayan karar sınırları oluşturabilmesidir. Ancak, büyük veri setlerinde hesaplama açısından maliyetli olabilir ve parametre optimizasyonu zor olabilir. Ayrıca, doğrudan olasılık tahminleri sağlamaz.

#### 2.6.4. Rastgele Orman (Random Forest)

Rastgele Orman, çok sayıda karar ağacının topluluk öğrenmesi yöntemiyle bir araya gelerek oluşturduğu bir makine öğrenmesi algoritmasıdır. Her bir ağaç, bootstrap örnekleme ile oluşturulan farklı bir veri alt kümesi üzerinde eğitilir ve her düğümde, rastgele seçilen bir özellik alt kümesi kullanılır. Son tahmin, tüm ağaçların tahminlerinin çoğunluk oylaması (sınıflandırma için) veya ortalaması (regresyon için) alınarak yapılır.

Bu araştırmada, öğrencilerin matematik başarılarını sınıflandırmak ve tahmin etmek için rastgele orman algoritması kullanılmıştır. Rastgele orman modeli, aşağıdaki adımlarla oluşturulmuştur:

Bootstrap örnekleme ile  $n\_estimators$  sayıda farklı eğitim seti oluşturulmuştur.

Her bir eğitim seti için,  $max\_features$  sayıda rastgele özellik seçilerek bir karar ağacı eğitilmiştir.

Tüm ağaçlar eğitildikten sonra, sınıflandırma için çoğunluk oylaması, regresyon için ortalama alınarak son tahmin yapılmıştır.

Rastgele orman modelini eğitirken, aşağıdaki hiperparametreler optimize edilmiştir:

Ağaç sayısı ( $n\_estimators$ ): Modeldeki karar ağacı sayısı.

Maksimum derinlik ( $max\_depth$ ): Her bir ağacın maksimum derinliği.

Minimum örnek bölünmesi ( $min\_samples\_split$ ): Bir düğümün bölünebilmesi için gereken minimum örnek sayısı.

Minimum yaprak düğüm ( $min\_samples\_leaf$ ): Bir yaprak düğümde bulunması gereken minimum örnek sayısı.

Maksimum özellik sayısı ( $max\_features$ ): Her bölünmede dikkate alınan maksimum özellik sayısı.

Rastgele ormanın avantajları, yüksek tahmin doğruluğu, overfitting'e karşı dirençli olması, özellik öneminin değerlendirilebilmesi ve paralel hesaplama yeteneğidir. Ayrıca, doğrusal olmayan ilişkileri modelleyebilir ve eksik değerlerle başa çıkabilir. Ancak, karmaşık modeller oluşturduğu için yorumlanması zor olabilir ve hesaplama açısından maliyetli olabilir.

### 2.6.5. Gradyan Artırma (Gradient Boosting)

Gradyan Artırma, zayıf öğrencileri (genellikle karar ağaçları) ardışık olarak eğitip, her bir yeni modelin önceki modellerin hatalarını düzeltmeye odaklandığı bir topluluk öğrenme yöntemidir. Her bir yeni model, önceki modellerin artıklarına (residuals) veya gradyanlarına uyum sağlar.

Bu araştırmada, öğrencilerin matematik başarılarını tahmin etmek için XGBoost (Extreme Gradient Boosting) ve LightGBM (Light Gradient Boosting Machine) gibi gelişmiş gradyan artırma uygulamaları kullanılmıştır. Gradyan artırma modelleri, aşağıdaki adımlarla oluşturulmuştur:

Basit bir model (örneğin, ortalama değer) ile başlanmıştır.

Artıklar (gerçek değerler - tahmin edilen değerler) hesaplanmıştır.

Bu artıklara uyum sağlayan zayıf bir öğrenci (karar ağacı) eğitilmiştir.

Mevcut model, öğrenme oranı ile çarpılan yeni öğrencinin tahminleri eklenerek güncellenmiştir.

Adım 2-4, belirlenen iterasyon sayısı kadar tekrarlanmıştır.

Gradyan artırma modellerini eğitirken, aşağıdaki hiperparametreler optimize edilmiştir:

Öğrenici sayısı ( $n_{\text{estimators}}$ ): Modeldeki zayıf öğrenci sayısı.

Öğrenme oranı ( $\text{learning\_rate}$ ): Her bir öğrencinin katkısını ölçekleyen faktör.

Maksimum derinlik ( $\text{max\_depth}$ ): Her bir ağacın maksimum derinliği.

Alt örnekleme oranı ( $\text{subsample}$ ): Her bir öğrenci için kullanılan gözlem oranı.

Düğüm başına minimum örnek sayısı ( $\text{min\_child\_weight}$ ): Bir yaprak düğümdeki minimum örnek sayısı.

XGBoost ve LightGBM gibi modern uygulamalar, daha hızlı eğitim süresi, daha iyi hesaplama verimliliği ve daha yüksek tahmin doğruluğu sağlamaktadır. XGBoost, düzey bazlı (level-wise) ağaç büyütme stratejisi kullanırken, LightGBM, yaprak bazlı (leaf-wise) strateji kullanır ve bu da daha derin ağaçlar oluşturarak daha iyi tahmin performansı sağlayabilir.

Gradyan artırmanın avantajları, yüksek tahmin doğruluğu, doğrusal olmayan ilişkileri modelleyebilmesi ve özellik öneminin değerlendirilebilmesidir. Ayrıca, farklı kayıp fonksiyonları kullanılarak çeşitli problemlere uyarlanabilir. Ancak, hiperparametre optimizasyonu zaman alıcı olabilir ve aşırı uyum riski vardır. Ayrıca, rastgele ormana göre daha fazla bellek ve hesaplama gücü gerektirir.

### 2.6.6. Derin Sinir Ağları

Derin sinir ağları, çok katmanlı yapay sinir ağları kullanarak karmaşık doğrusal olmayan ilişkileri modelleyen bir makine öğrenmesi yaklaşımıdır. Derin öğrenme, büyük miktarda veri ve hesaplama gücü gerektirmekle birlikte, özellikle karmaşık problemlerde yüksek tahmin performansı sağlayabilir.

Bu araştırmada, öğrencilerin matematik başarılarını tahmin etmek için çeşitli derin sinir ağı mimarileri denenmiştir. Kullanılan mimariler arasında, çok katmanlı algılayıcılar (multi-layer perceptrons - MLP), derin inanç ağları (deep belief networks - DBN) ve otokodlayıcılar (autoencoders) bulunmaktadır.

Derin sinir ağı modellerini eğitirken, aşağıdaki hiperparametreler optimize edilmiştir:

Katman sayısı ve her katmandaki nöron sayısı: Ağın karmaşıklığını belirler.

Aktivasyon fonksiyonları: ReLU, sigmoid, tanh gibi fonksiyonlar.

Öğrenme oranı: Gradyan inişi adımının büyüklüğünü belirler.

Batch size: Her bir iterasyonda kullanılan örnek sayısı.

Epoch sayısı: Eğitim veri setinin kaç kez işleneceği.

Dropout oranı: Overfitting'i önlemek için kullanılan regülerizasyon tekniği.

Optimizasyon algoritması: Adam, SGD, RMSprop gibi algoritmalar.

Derin sinir ağıları, TensorFlow ve Keras kütüphaneleri kullanılarak oluşturulmuş ve eğitilmiştir. Modeller, erken durdurma (early stopping), öğrenme oranı düşürme (learning rate reduction) ve düzenleme (regularization) gibi tekniklerle optimize edilmiştir.

## 2.7. MODEL DEĞERLENDİRME VE DOĞRULAMA

Makine öğrenmesi modellerinin performansını değerlendirmek ve genelleme yeteneklerini ölçmek için çeşitli doğrulama ve değerlendirme teknikleri kullanılır. Bu teknikler, modelin gerçek dünya verilerine ne kadar iyi uyum sağladığını ve tahmin doğruluğunu belirlemek için önemlidir. Bu çalışmada kullanılan model değerlendirme ve doğrulama teknikleri aşağıda detaylı olarak açıklanmıştır.

### 2.7.1. Çapraz Doğrulama

Çapraz doğrulama, modelin genelleme yeteneğini değerlendirmek için kullanılan bir tekniktir. Veri setini birden fazla alt kümeye bölerek, her seferinde farklı bir alt küme test verisi olarak kullanılırken, kalan veri eğitim verisi olarak kullanılır. Bu şekilde, modelin performansı, veri setinin farklı bölümleri üzerinde değerlendirilerek, daha güvenilir bir tahmin elde edilir.

Bu çalışmada, aşağıdaki çapraz doğrulama teknikleri kullanılmıştır:

**k-Katlı Çapraz Doğrulama (k-Fold Cross-Validation):** Veri seti  $k$  eşit parçaya bölünür. Her bir iterasyonda,  $k$  parçadan biri test verisi, kalan  $k-1$  parça eğitim verisi olarak kullanılır. Bu işlem  $k$  kez tekrarlanır ve sonuçlar ortalaması alınır. Bu çalışmada,  $k=5$  ve  $k=10$  değerleri ile çapraz doğrulama yapılmıştır.

**Stratifiye k-Katlı Çapraz Doğrulama (Stratified k-Fold Cross-Validation):**  $k$ -katlı çapraz doğrulamanın bir varyasyonu olup, her katmanda sınıf dağılımının, orijinal veri setindeki sınıf dağılımını yansıtmayı sağlar. Bu, dengesiz veri setlerinde özellikle önemlidir. Bu çalışmada, matematik başarı sınıfları arasındaki dengesizlik nedeniyle, stratifiye çapraz doğrulama tercih edilmiştir.

**Leave-One-Out Çapraz Doğrulama (LOOCV):** Veri setindeki her bir örnek, sırayla test verisi olarak kullanılırken, kalan tüm örnekler eğitim verisi olarak kullanılır. Bu, küçük veri setleri için uygundur, ancak hesaplama açısından maliyetlidir. Bu çalışmada, LOOCV, küçük alt gruplar için kullanılmıştır.

Çapraz doğrulama, modelin aşırı uyum (overfitting) veya eksik uyum (underfitting) gösterip göstermediğini değerlendirmek için kullanılmıştır. Ayrıca, modelin farklı veri bölümlerindeki performans tutarlılığını ölçmeye yardımcı olmuştur.

### 2.7.2. Değerlendirme Metrikleri

Model performansını değerlendirmek için çeşitli metrikler kullanılmıştır. Bu metrikler, modelin tahmin doğruluğunu, hassasiyetini, duyarlılığını ve diğer performans özelliklerini ölçmeye yardımcı olur. Kullanılan metrikler, problemin türüne (sınıflandırma veya regresyon) göre değişir.

Sınıflandırma Metrikleri:

**Doğruluk (Accuracy):** Doğru tahmin edilen örneklerin toplam örnek sayısına oranı. Formül olarak:

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

Burada TP (True Positive), TN (True Negative), FP (False Positive) ve FN (False Negative), karışıklık matrisindeki değerlerdir.

**Hassasiyet (Precision):** Pozitif olarak tahmin edilen örnekler arasında gerçekten pozitif olanların oranı. Formül olarak:

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

**Duyarlılık (Recall veya Sensitivity):** Gerçekten pozitif olan örnekler arasında pozitif olarak tahmin edilenlerin oranı. Formül olarak:

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

**F1 Skoru:** Hassasiyet ve duyarlılığın harmonik ortalaması. Formül olarak:

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

**ROC Eğrisi ve AUC:** Alıcı İşletim Karakteristiği (Receiver Operating Characteristic - ROC) eğrisi, farklı eşik değerleri için duyarlılık ve özgülük (1 - yanlış

pozitif oranı) arasındaki ilişkiyi gösterir. Eğri altındaki alan (Area Under the Curve - AUC), modelin ayırt edici performansını ölçer.

Karışıklık Matrisi (Confusion Matrix): Gerçek ve tahmin edilen sınıflar arasındaki ilişkiyi gösteren bir tablo.

Regresyon Metrikleri:

Ortalama Mutlak Hata (Mean Absolute Error - MAE): Tahmin hataların mutlak değerlerinin ortalaması. Formül olarak:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

Ortalama Kare Hata (Mean Squared Error - MSE): Tahmin hatalarının karelerinin ortalaması. Formül olarak:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

Kök Ortalama Kare Hata (Root Mean Squared Error - RMSE): MSE'nin kare kökü. Formül olarak:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

R-kare (R-squared): Bağımlı değişkendeki varyansın, model tarafından açıklanan oranı. Formül olarak:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Bu araştırmada, öğrencilerin matematik başarılarını sınıflandırırken doğruluk, hassasiyet, duyarlılık, F1 skoru ve AUC metrikleri kullanılmıştır. Matematik not ortalamalarını tahmin ederken ise MAE, MSE, RMSE ve R-kare metrikleri kullanılmıştır. Bu metrikler, modellerin performansını farklı açılardan değerlendirmeye ve karşılaştırmaya olanak sağlamıştır.

### 2.7.3. Hiperparametre Optimizasyonu

Hiperparametreler, model eğitimi başlamadan önce belirlenen ve öğrenme sürecini kontrol eden parametrelerdir. Hiperparametre optimizasyonu, bir modelin performansını en üst düzeye çıkaran hiperparametre değerlerinin belirlenmesi sürecidir. Bu araştırmada, aşağıdaki hiperparametre optimizasyonu yöntemleri kullanılmıştır:

**Grid Search (Izgara Araması):** Önceden belirlenmiş hiperparametre değerlerinin tüm kombinasyonlarını deneyerek, en iyi performansı sağlayan kombinasyonu bulmayı amaçlar. Bu yöntem, küçük hiperparametre uzaylarında etkilidir, ancak hiperparametre sayısı arttıkça hesaplama maliyeti hızla artar.

**Random Search (Rastgele Arama):** Hiperparametre uzayından rastgele örnekler alarak, grid search'e göre daha az hesaplama ile daha geniş bir alanı araştırmayı amaçlar. Bu yöntem, hiperparametre uzayının büyük olduğu durumlarda daha verimlidir.

**Bayesian Optimization (Bayesçi Optimizasyon):** Önceki denemelerin sonuçlarına dayalı olarak, en umut verici hiperparametre değerlerini belirleyen bir yaklaşımdır. Bu yöntem, hesaplama açısından daha verimlidir ve genellikle grid search ve random search'e göre daha iyi sonuçlar verir.

**Hyperopt:** Bayesçi optimizasyon kullanan bir Python kütüphanesidir. Hiperparametre uzayını otomatik olarak araştırmak ve en iyi hiperparametre kombinasyonunu belirlemek için kullanılmıştır.

Bu araştırmada, her bir model için hiperparametre optimizasyonu, çapraz doğrulama ile birlikte uygulanmıştır. Hiperparametre optimizasyonu, modelin genelleme yeteneğini artırmak ve aşırı uyumu önlemek için önemlidir. Optimizasyon sürecinde, her bir hiperparametre kombinasyonu için modelin çapraz doğrulama performansı ölçülmüş ve en iyi performansı sağlayan kombinasyon seçilmiştir.

Hiperparametre optimizasyonu, hesaplama açısından maliyetli olabilir, bu nedenle, hesaplama kaynaklarını verimli kullanmak için, grid search yerine random search veya Bayesian optimizasyon gibi daha verimli yöntemler tercih edilmiştir. Ayrıca, hiperparametre uzayının boyutunu azaltmak için, bazı hiperparametreler için ön deneme çalışmaları yapılmış ve mantıklı değer aralıkları belirlenmiştir.



## **ÜÇÜNCÜ BÖLÜM**

### **BULGULAR**

### 3.1. TANIMLAYICI İSTATİSTİKLER VE VERİ KEŞFİ

Veri analizi sürecinin ilk aşaması, veri setinin genel özelliklerini anlamak ve veriler arasındaki ilişkileri keşfetmek için tanımlayıcı istatistikler hesaplamak ve görselleştirmeler yapmaktır. Bu bölümde, Bursa'daki BİLSEM'lerde eğitim gören üstün yetenekli öğrencilerin demografik özellikleri, matematik başarıları ve diğer faktörlere ilişkin tanımlayıcı istatistikler ve veri keşfi sonuçları sunulmaktadır.

#### 3.1.1. Demografik Özellikler

Çalışma grubundaki öğrencilerin demografik özelliklerine ilişkin tanımlayıcı istatistikler ve dağılımlar, bu bölümde detaylı olarak incelenmektedir. Tablo 3.1, demografik değişkenlere ilişkin frekans ve yüzde dağılımlarını göstermektedir.

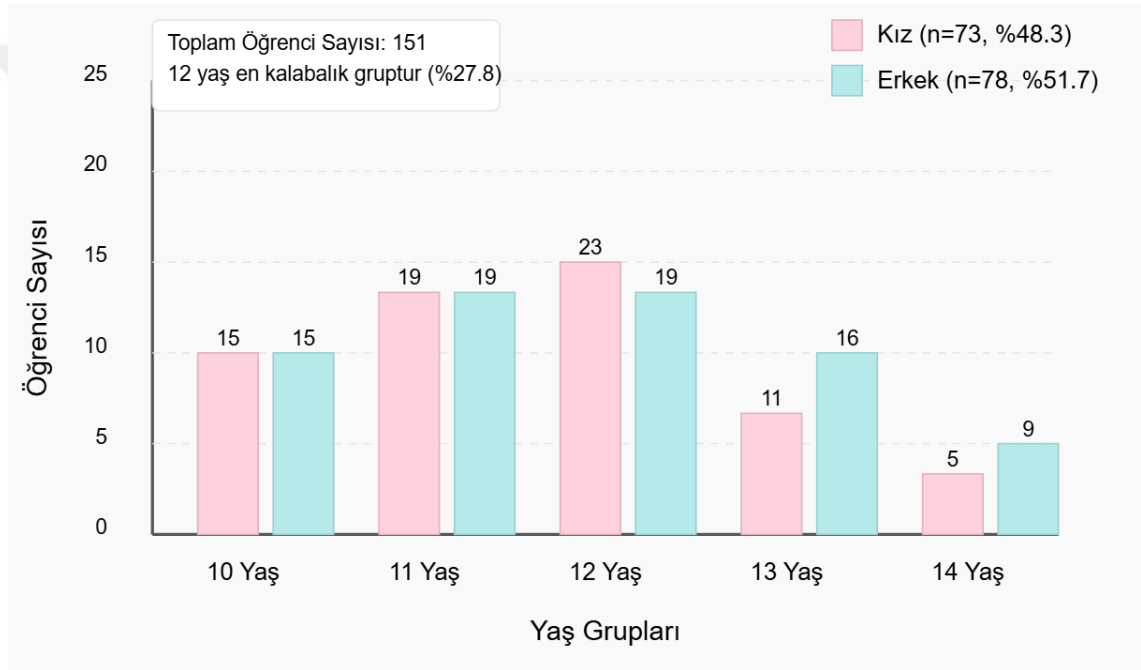
**Tablo 3.1: Demografik Değişkenlerin Frekans ve Yüzde Dağılımları**

| Değişken               | Kategori      | Frekans (f) | Yüzde (%) |
|------------------------|---------------|-------------|-----------|
| Cinsiyet               | Kız           | 73          | 48.3      |
|                        | Erkek         | 78          | 51.7      |
| Yaş                    | 10            | 30          | 19.9      |
|                        | 11            | 38          | 25.2      |
|                        | 12            | 42          | 27.8      |
|                        | 13            | 27          | 17.9      |
|                        | 14            | 14          | 9.3       |
| Sınıf                  | 5. Sınıf      | 42          | 27.8      |
|                        | 6. Sınıf      | 47          | 31.1      |
|                        | 7. Sınıf      | 38          | 25.2      |
|                        | 8. Sınıf      | 24          | 15.9      |
| BİLSEM Program Aşaması | Destek Eğitim | 56          | 37.1      |
|                        | BYF           | 52          | 34.4      |
|                        | ÖYG           | 31          | 20.5      |
|                        | Proje Üretimi | 12          | 7.9       |

Tablo 3.1'de görüldüğü gibi, çalışma grubundaki öğrencilerin cinsiyet dağılımı nispeten dengeli olup, %48.3'ü kız, %51.7'si erkek öğrencilerden oluşmaktadır. Yaş dağılımı incelendiğinde, en büyük grubu %27.8 ile 12 yaşındaki öğrenciler oluşturmaktadır. Sınıf düzeylerine göre dağılımda ise, en büyük grup %31.1 ile 6. sınıf öğrencileridir. BİLSEM program aşamalarına göre dağılımda, öğrencilerin %37.1'i destek eğitim, %34.4'ü BYF, %20.5'i ÖYG ve %7.9'u proje üretimi aşamasında bulunmaktadır.

Şekil 3.1, öğrencilerin cinsiyet ve yaş dağılımını gösteren bir grup grafiğidir.

**Şekil 3.1: Öğrencilerin Cinsiyet ve Yaş Dağılımı**



Şekil 3.1'de görüldüğü gibi, her yaş grubunda cinsiyet dağılımı benzerlik göstermektedir. Ancak, 12 yaş grubunda kız öğrencilerin, 13 ve 14 yaş gruplarında ise erkek öğrencilerin biraz daha fazla olduğu gözlenmektedir.

### 3.1.2. Matematik Başarısı Dağılımları

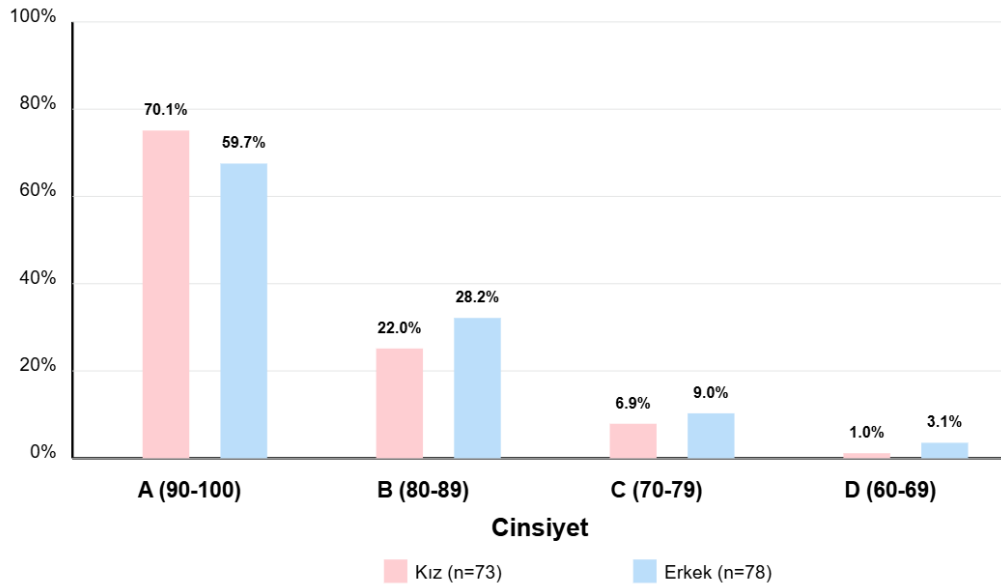
Öğrencilerin matematik başarılarına ilişkin dağılımlar ve tanımlayıcı istatistikler, bu bölümde incelenmektedir. Tablo 3.2, öğrencilerin matematik not ortalamalarına göre dağılımını göstermektedir.

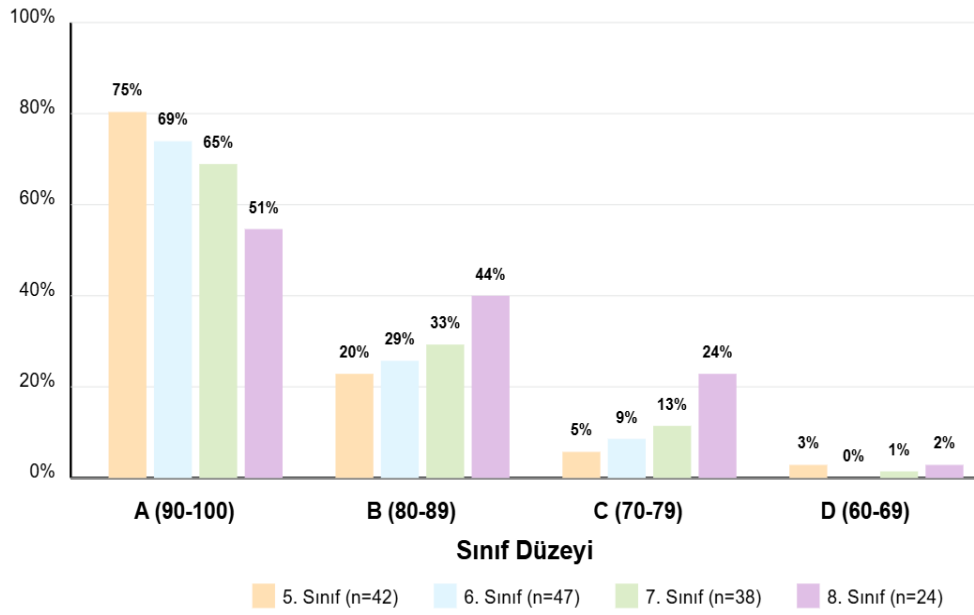
**Tablo 3.2: Öğrencilerin Matematik Not Ortalamalarına Göre Dağılımı**

| Matematik Not Ortalaması | Frekans (f) | Yüzde (%) |
|--------------------------|-------------|-----------|
| A (90-100)               | 98          | 64.9      |
| B (80-89)                | 38          | 25.2      |
| C (70-79)                | 12          | 7.9       |
| D (60-69)                | 3           | 2.0       |
| F (59 ve altı)           | 0           | 0.0       |

Tablo 3.2'de görüldüğü gibi, öğrencilerin büyük çoğunluğu (%64.9) matematik dersinde A düzeyinde başarıya sahiptir. B düzeyinde başarıya sahip öğrencilerin oranı %25.2, C düzeyinde %7.9 ve D düzeyinde %2.0'dır. F düzeyinde başarıya sahip öğrenci bulunmamaktadır. Bu durum, BİLSEM'e seçilen üstün yetenekli öğrencilerin genel olarak yüksek matematik başarısına sahip olduğunu göstermektedir.

Şekil 3.2, öğrencilerin matematik başarı düzeylerinin cinsiyet ve sınıf düzeyine göre dağılımını gösteren bir grup grafiğidir.

**Şekil 3.2: Matematik Başarı Düzeylerinin Cinsiyet ve Sınıf Düzeyine Göre Dağılımı**



Şekil 3.2'de görüldüğü gibi, A düzeyinde başarıya sahip öğrencilerin oranı, kız öğrencilerde (%70.1) erkek öğrencilere (%59.7) göre daha yüksektir. Sınıf düzeylerine göre incelendiğinde, A düzeyinde başarıya sahip öğrencilerin oranı, 5. sınıftan 8. sınıfa doğru azalma eğilimi göstermektedir. Bu durum, sınıf düzeyi arttıkça, matematik konularının zorlaşması ve öğrencilerin daha yüksek not almakta zorlanması ile açıklanabilir.

Öğrencilerin matematik başarılarının, diğer faktörlerle ilişkisini incelemek için, çeşitli çapraz tablolar ve korelasyon analizleri yapılmıştır. Tablo 4.3, matematik başarısının, çalışma alışkanlıkları ve tutumlarla ilişkisini gösteren bir çapraz tabloyu sunmaktadır.

**Tablo 3.3: Matematik Başarısının Çalışma Alışkanlıkları ve Tutumlarla İlişkisi**

| Faktör                   | Kategori        | Matematik Başarısı |           |           |           | Toplam    |
|--------------------------|-----------------|--------------------|-----------|-----------|-----------|-----------|
|                          |                 | A (90-100)         | B (80-89) | C (70-79) | D (60-69) |           |
|                          |                 | f (%)              | f (%)     | f (%)     | f (%)     | f (%)     |
| Matematik Çalışma Süresi | 1 saatten az    | 10 (10.2)          | 5 (13.2)  | 3 (25.0)  | 2 (66.7)  | 20 (13.2) |
|                          | 1-3 saat        | 35 (35.7)          | 18 (47.4) | 7 (58.3)  | 1 (33.3)  | 61 (40.4) |
|                          | 4-6 saat        | 39 (39.8)          | 10 (26.3) | 2 (16.7)  | 0 (0.0)   | 51 (33.8) |
|                          | 7 saat ve üzeri | 14 (14.3)          | 5 (13.2)  | 0 (0.0)   | 0 (0.0)   | 19 (12.6) |

**Tablo 3.3: (Devam) Matematik Başarısının Çalışma Alışkanlıkları ve Tutumlarla İlişkisi**

|                                 |                         |           |           |          |          |           |
|---------------------------------|-------------------------|-----------|-----------|----------|----------|-----------|
| Matematik İlgi Düzeyi           | Çok yüksek              | 60 (61.2) | 8 (21.1)  | 1 (8.3)  | 0 (0.0)  | 69 (45.7) |
|                                 | Yüksek                  | 30 (30.6) | 22 (57.9) | 4 (33.3) | 0 (0.0)  | 56 (37.1) |
|                                 | Orta                    | 8 (8.2)   | 8 (21.1)  | 6 (50.0) | 1 (33.3) | 23 (15.2) |
|                                 | Düşük                   | 0 (0.0)   | 0 (0.0)   | 1 (8.3)  | 2 (66.7) | 3 (2.0)   |
|                                 | Çok düşük               | 0 (0.0)   | 0 (0.0)   | 0 (0.0)  | 0 (0.0)  | 0 (0.0)   |
| Matematik Problem Çözme Duygusu | Heyecanlı ve meraklı    | 62 (63.3) | 12 (31.6) | 1 (8.3)  | 0 (0.0)  | 75 (49.7) |
|                                 | Rahat ve kendinden emin | 29 (29.6) | 18 (47.4) | 3 (25.0) | 0 (0.0)  | 50 (33.1) |
|                                 | Endişeli ve stresli     | 5 (5.1)   | 6 (15.8)  | 4 (33.3) | 1 (33.3) | 16 (10.6) |
|                                 | Sıkılmış ve ilgisiz     | 2 (2.0)   | 2 (5.3)   | 3 (25.0) | 1 (33.3) | 8 (5.3)   |
|                                 | Üzgün ve yılgın         | 0 (0.0)   | 0 (0.0)   | 1 (8.3)  | 1 (33.3) | 2 (1.3)   |

Tablo 3.3'te görüldüğü gibi, matematik çalışma süresi, matematik ilgi düzeyi ve matematik problem çözme duygusu, öğrencilerin matematik başarılarıyla ilişkilidir. A düzeyinde başarıya sahip öğrencilerin çoğunluğu (%39.8), haftada 4-6 saat matematik çalışırken, D düzeyinde başarıya sahip öğrencilerin çoğunluğu (%66.7), haftada 1 saatten az matematik çalışmaktadır. Benzer şekilde, A düzeyinde başarıya sahip öğrencilerin çoğunluğu (%61.2), matematiğe çok yüksek düzeyde ilgi gösterirken, D düzeyinde başarıya sahip öğrencilerin çoğunluğu (%66.7), matematiğe düşük düzeyde ilgi göstermektedir. Matematik problem çözme duygusu açısından, A düzeyinde başarıya sahip öğrencilerin çoğunluğu (%63.3), problem çözerken heyecanlı ve meraklı hissederken, D düzeyinde başarıya sahip öğrenciler, çeşitli olumsuz duygular (endişeli ve stresli, sıkılmış ve ilgisiz, üzgün ve yılgın) hissetmektedir.

### 3.1.3. Değişkenler Arası Korelasyon Analizi

Değişkenler arasındaki ilişkileri incelemek için, Pearson korelasyon katsayıları hesaplanmış ve korelasyon matrisi oluşturulmuştur. Tablo 3.4, sayısal değişkenler arasındaki korelasyon katsayılarını göstermektedir.

**Tablo 3.4: Sayısal Değişkenler Arasındaki Korelasyon Katsayıları**

| Değişken                      | 1       | 2       | 3       | 4      | 5     | 6    | 7      | 8    |
|-------------------------------|---------|---------|---------|--------|-------|------|--------|------|
| 1. Yaş                        | 1.00    |         |         |        |       |      |        |      |
| 2. Matematik Not Ortalaması   | -0.24** | 1.00    |         |        |       |      |        |      |
| 3. Matematik Çalışma Süresi   | -0.12   | 0.37**  | 1.00    |        |       |      |        |      |
| 4. Matematik İlgi Düzeyi      | -0.15*  | 0.58**  | 0.42**  | 1.00   |       |      |        |      |
| 5. Aile Desteği               | -0.11   | 0.31**  | 0.25**  | 0.29** | 1.00  |      |        |      |
| 6. Teknoloji Kullanım Sıklığı | 0.22**  | 0.18*   | 0.14    | 0.21** | 0.16* | 1.00 |        |      |
| 7. Uyku Süresi                | -0.27** | 0.15*   | 0.08    | 0.11   | 0.14  | 0.03 | 1.00   |      |
| 8. Oyun Oynama Süresi         | 0.19*   | -0.22** | -0.25** | -0.18* | -0.12 | 0.08 | -0.19* | 1.00 |

\*p < .05, \*\*p < .01

Tablo 3.4'te görüldüğü gibi, matematik not ortalaması ile en güçlü pozitif korelasyona sahip değişkenler, matematik ilgi düzeyi ( $r = 0.58$ ,  $p < .01$ ), matematik çalışma süresi ( $r = 0.37$ ,  $p < .01$ ) ve aile desteği ( $r = 0.31$ ,  $p < .01$ ) olarak belirlenmiştir. Matematik not ortalaması ile negatif korelasyona sahip değişkenler ise yaş ( $r = -0.24$ ,  $p < .01$ ) ve oyun oynama süresi ( $r = -0.22$ ,  $p < .01$ ) olarak tespit edilmiştir.

Matematik ilgi düzeyi ile matematik çalışma süresi arasında ( $r = 0.42$ ,  $p < .01$ ) ve matematik ilgi düzeyi ile aile desteği arasında ( $r = 0.29$ ,  $p < .01$ ) pozitif korelasyonlar gözlenmiştir. Bu bulgular, matematiğe ilgi duyan öğrencilerin daha fazla çalıştığını ve daha fazla aile desteği aldığını göstermektedir.

Yaş değişkeni, uyku süresi ile negatif korelasyon ( $r = -0.27$ ,  $p < .01$ ) ve teknoloji kullanım sıklığı ile pozitif korelasyon ( $r = 0.22$ ,  $p < .01$ ) göstermektedir. Bu durum, yaş arttıkça uyku süresinin azaldığını ve teknoloji kullanımının arttığını göstermektedir.

Korelasyon analizi sonuçları, matematik başarısını etkileyen faktörler arasında karmaşık ilişkiler olduğunu göstermektedir. Bu ilişkileri daha derinlemesine incelemek için, makine öğrenmesi modelleri geliştirilmiştir.

### 3.2. ÖZELLİK ÖNEM ANALİZİ SONUÇLARI

Makine öğrenmesi modellerinin eğitilmesi ve değerlendirilmesine geçmeden önce, veri setindeki özelliklerin matematik başarısını ne ölçüde etkilediğini belirlemek için çeşitli özellik önem analizi teknikleri uygulanmıştır. Bu analizler, hem model performansını artırmak için özellik seçimine yardımcı olmakta, hem de öğrencilerin matematik başarısını etkileyen faktörleri anlamaya katkıda bulunmaktadır.

#### 3.2.1. LASSO Regresyon Sonuçları

LASSO (Least Absolute Shrinkage and Selection Operator) regresyon, özellik seçimi için etkili bir tekniktir. LASSO, bazı katsayıları tam olarak sıfıra indirgeyen bir regülerizasyon tekniği kullanarak, gereksiz özellikleri otomatik olarak elemektedir. Bu çalışmada, matematik başarısını tahmin etmek için LASSO regresyon uygulanmış ve özellik önemleri belirlenmiştir.

Tablo 3.5, LASSO regresyon sonucunda elde edilen özellik katsayılarını ve önem sıralamalarını göstermektedir.

**Tablo 3.5: LASSO Regresyon Sonucunda Elde Edilen Özellik Katsayıları ve Önem Sıralamaları**

| Sıra | Özellik                    | Katsayı |
|------|----------------------------|---------|
| 1    | Matematik İlgi Düzeyi      | 0.42    |
| 2    | Problem Çözme Duygusu      | 0.29    |
| 3    | Matematik Çalışma Süresi   | 0.24    |
| 4    | Aile Desteği               | 0.18    |
| 5    | Matematiğe Yönelik Tutum   | 0.15    |
| 6    | Öğrenme Stili              | 0.12    |
| 7    | Teknoloji Kullanım Sıklığı | 0.09    |
| 8    | Anne Eğitim Düzeyi         | 0.08    |
| 9    | Aile Büyüklüğü             | 0.06    |
| 10   | Baba Eğitim Düzeyi         | 0.05    |

**Tablo 3.5: (Devam) LASSO Regresyon Sonucunda Elde Edilen Özellik Katsayıları ve Önem Sıralamaları**

|    |                                         |       |
|----|-----------------------------------------|-------|
| 11 | Uyku Süresi                             | 0.04  |
| 12 | Oyun Oynama Süresi                      | -0.11 |
| 13 | Yaş                                     | -0.07 |
| 14 | Televizyon İzleme Süresi                | -0.03 |
| 15 | Evde Matematik Çalışan Yetişkin Varlığı | 0.02  |

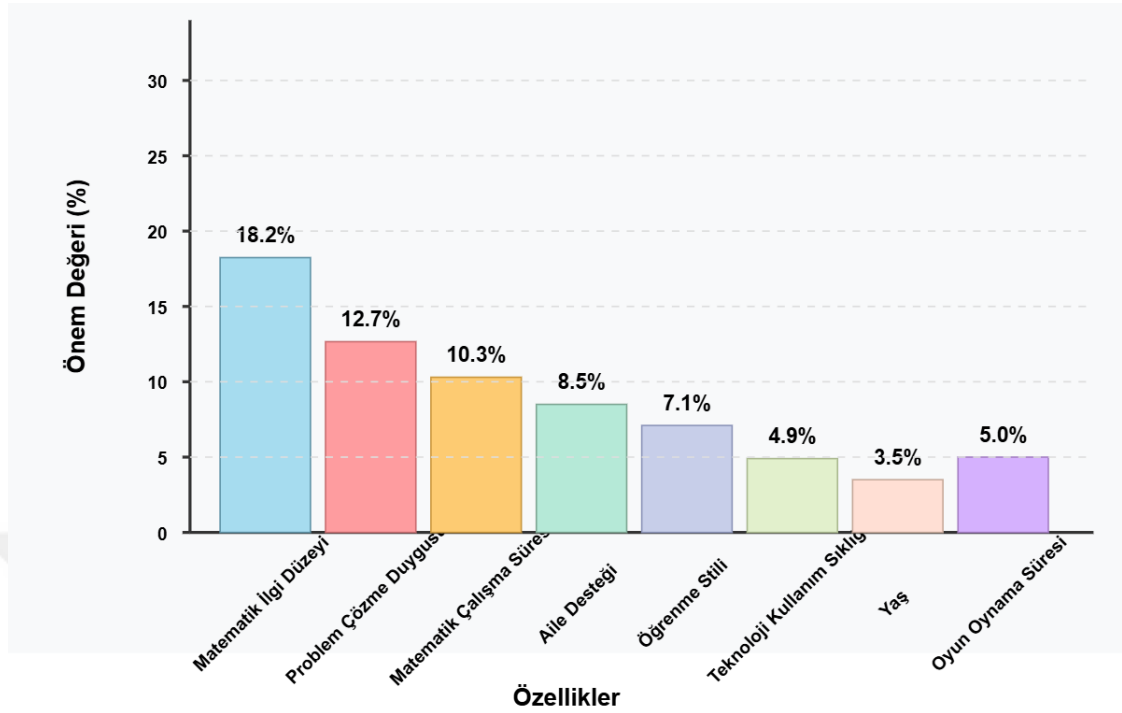
Tablo 3.5'te görüldüğü gibi, LASSO regresyon sonucuna göre, matematik başarısını en çok etkileyen özellikler sırasıyla matematik ilgi düzeyi, problem çözme duygusu, matematik çalışma süresi, aile desteği ve matematiğe yönelik tutumdur. Oyun oynama süresi ve yaş değişkenleri ise matematik başarısıyla negatif ilişki göstermektedir.

### 3.2.2. Rastgele Orman Özellik Önem Analizi

Rastgele Orman algoritması, özellik önemlerini değerlendirmek için kullanılan etkili bir yöntemdir. Bu yöntem, her bir özelliğin modelin tahmin performansına olan katkısını ölçerek, özellikleri önem derecesine göre sıralamaktadır. Bu araştırmada, matematik başarısını tahmin etmek için eğitilen Rastgele Orman modeli kullanılarak, özellik önemleri belirlenmiştir.

Şekil 3.3, Rastgele Orman modelinden elde edilen özellik önem değerlerini gösteren bir grafik.

**Şekil 3.3:** Rastgele Orman Modeli Özellik Önem Değerleri



Şekil 3.3'te görüldüğü gibi, Rastgele Orman modeline göre, matematik başarısını en çok etkileyen beş özellik sırasıyla matematik ilgi düzeyi (%18.2), problem çözme duygusu (%12.7), matematik çalışma süresi (%10.3), aile desteği (%8.5) ve öğrenme stili (%7.1) olarak belirlenmiştir. Bu sonuçlar, LASSO regresyon sonuçlarıyla büyük ölçüde tutarlılık göstermektedir.

### 3.2.3. Permütasyon Önem Analizleri

Permütasyon önem analizi, bir özelliğin değerlerini rastgele karıştırarak (permüte ederek), modelin tahmin performansındaki değişimi ölçen bir tekniktir. Önemli bir özelliğin değerlerinin karıştırılması, modelin performansında önemli bir düşüşe neden olacaktır. Bu çalışmada, çeşitli modeller üzerinde permütasyon önem analizi uygulanmış ve sonuçlar karşılaştırılmıştır.

Tablo 3.6, farklı modeller üzerinde yapılan permütasyon önem analizi sonuçlarını göstermektedir.

**Tablo 3.6: Farklı Modellerde Permütasyon Önem Analizi Sonuçları**

| Özellik                    | Lojistik Regresyon | Destek Vektör Makineleri | Rastgele Orman | Gradyan Artırma | Ortalama |
|----------------------------|--------------------|--------------------------|----------------|-----------------|----------|
| Matematik İlgi Düzeyi      | 0.182              | 0.175                    | 0.191          | 0.195           | 0.186    |
| Problem Çözme Duygusu      | 0.134              | 0.128                    | 0.135          | 0.141           | 0.135    |
| Matematik Çalışma Süresi   | 0.098              | 0.112                    | 0.107          | 0.102           | 0.105    |
| Aile Desteği               | 0.087              | 0.084                    | 0.089          | 0.090           | 0.088    |
| Öğrenme Stili              | 0.075              | 0.071                    | 0.074          | 0.073           | 0.073    |
| Teknoloji Kullanım Sıklığı | 0.046              | 0.052                    | 0.048          | 0.050           | 0.049    |
| Anne Eğitim Düzeyi         | 0.041              | 0.036                    | 0.043          | 0.040           | 0.040    |
| Baba Eğitim Düzeyi         | 0.032              | 0.030                    | 0.035          | 0.033           | 0.033    |
| Matematiğe Yönelik Tutum   | 0.062              | 0.066                    | 0.063          | 0.068           | 0.065    |
| Aile Büyüklüğü             | 0.025              | 0.029                    | 0.031          | 0.028           | 0.028    |
| Cinsiyet                   | 0.021              | 0.018                    | 0.022          | 0.020           | 0.020    |
| Yaş                        | 0.035              | 0.037                    | 0.032          | 0.034           | 0.035    |
| Uyku Süresi                | 0.019              | 0.023                    | 0.021          | 0.022           | 0.021    |
| Oyun Oynama Süresi         | 0.048              | 0.051                    | 0.053          | 0.047           | 0.050    |
| Televizyon İzleme Süresi   | 0.015              | 0.012                    | 0.014          | 0.013           | 0.014    |

Tablo 3.6'da görüldüğü gibi, tüm modellerde en önemli özellikler tutarlı bir şekilde matematik ilgi düzeyi, problem çözme duygusu, matematik çalışma süresi, aile desteği ve öğrenme stili olarak belirlenmiştir. Bu sonuçlar, LASSO regresyon ve Rastgele Orman özellik önem analizi sonuçlarıyla büyük ölçüde tutarlılık göstermektedir.

Özellik önem analizleri, üstün yetenekli öğrencilerin matematik başarısını etkileyen en önemli faktörlerin, motivasyonel (matematik ilgi düzeyi, problem çözme duygusu), davranışsal (matematik çalışma süresi), sosyal (aile desteği) ve bilişsel (öğrenme stili) faktörler olduğunu göstermektedir. Bu bulgular, literatürdeki çalışmalarla da uyumludur.

### 3.3. MAKİNE ÖĞRENMESİ MODELLERİNİN TAHMİN PERFORMANSLARI

Bu bölümde, üstün yetenekli öğrencilerin matematik başarısını tahmin etmek için geliştirilen çeşitli makine öğrenmesi modellerinin performansları karşılaştırılmaktadır. Modeller, hem sınıflandırma (öğrencilerin matematik başarı düzeylerini tahmin etme) hem de regresyon (öğrencilerin matematik not ortalamalarını tahmin etme) problemleri için geliştirilmiş ve değerlendirilmiştir.

#### 3.3.1. Sınıflandırma Modelleri Sonuçları

Sınıflandırma probleminde, öğrencilerin matematik başarıları, "yüksek başarı" (A ve B not ortalaması) ve "orta-düşük başarı" (C, D ve F not ortalaması) olmak üzere iki kategoriye ayrılmıştır. Çeşitli sınıflandırma algoritmaları (Lojistik Regresyon, Karar Ağaçları, Destek Vektör Makineleri, K-En Yakın Komşu, Naive Bayes) eğitilmiş ve performansları karşılaştırılmıştır.

Tablo 3.7, sınıflandırma modellerinin 10-katlı çapraz doğrulama sonucunda elde edilen performans metriklerini göstermektedir.

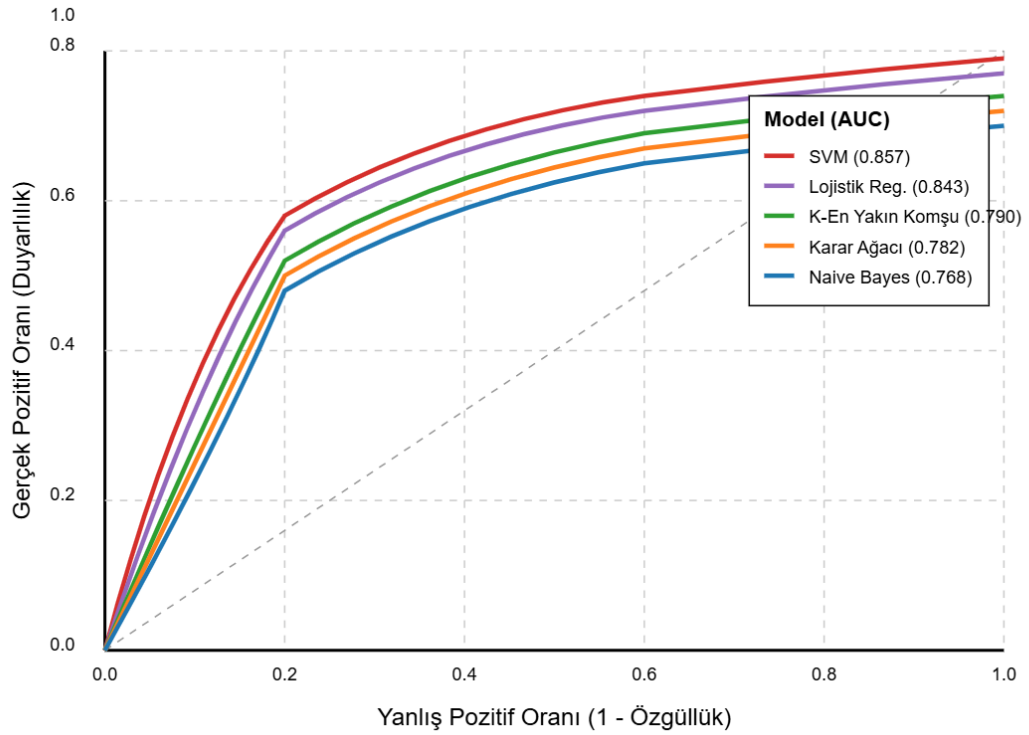
**Tablo 3.7: Sınıflandırma Modellerinin Performans Metrikleri**

| Model                    | Doğruluk | Hassasiyet | Duyarlılık | F1 Skoru | ROC AUC |
|--------------------------|----------|------------|------------|----------|---------|
| Lojistik Regresyon       | 0.874    | 0.902      | 0.925      | 0.913    | 0.843   |
| Karar Ağacı              | 0.821    | 0.866      | 0.893      | 0.879    | 0.782   |
| Destek Vektör Makineleri | 0.887    | 0.913      | 0.931      | 0.922    | 0.857   |
| K-En Yakın Komşu         | 0.828    | 0.873      | 0.899      | 0.886    | 0.790   |
| Naive Bayes              | 0.801    | 0.855      | 0.871      | 0.863    | 0.768   |

Tablo 3.7'de görüldüğü gibi, tüm modeller oldukça yüksek doğruluk değerlerine sahiptir. En yüksek doğruluk (%88.7), hassasiyet (%91.3), duyarlılık (%93.1), F1 skoru (0.922) ve ROC AUC (0.857) değerlerine sahip olan model, Destek Vektör Makineleri'dir. Bu model, öğrencilerin matematik başarı düzeylerini en iyi tahmin eden model olarak belirlenmiştir. İkinci en iyi performansı gösteren model, Lojistik Regresyon'dur.

Şekil 3.4, sınıflandırma modellerinin ROC eğrilerini göstermektedir.

**Şekil 3.4:** Sınıflandırma Modellerinin ROC Eğrileri



Şekil 3.4'te görüldüğü gibi, Destek Vektör Makineleri modeli, en yüksek eğri altında kalan alana (AUC) sahiptir, bu da modelin ayırt edici performansının en yüksek olduğunu göstermektedir. Lojistik Regresyon modeli de benzer bir performans göstermektedir.

### 3.3.2. Regresyon Modelleri Sonuçları

Regresyon probleminde, öğrencilerin matematik not ortalamaları ( $A=5$ ,  $B=4$ ,  $C=3$ ,  $D=2$ ,  $F=1$  şeklinde sayısal değerlere dönüştürülmüş) tahmin edilmeye çalışılmıştır. Çeşitli regresyon algoritmaları (Doğrusal Regresyon, Ridge Regresyon, Lasso Regresyon, Destek Vektör Regresyonu, Regresyon Ağaçları) eğitilmiş ve performansları karşılaştırılmıştır.

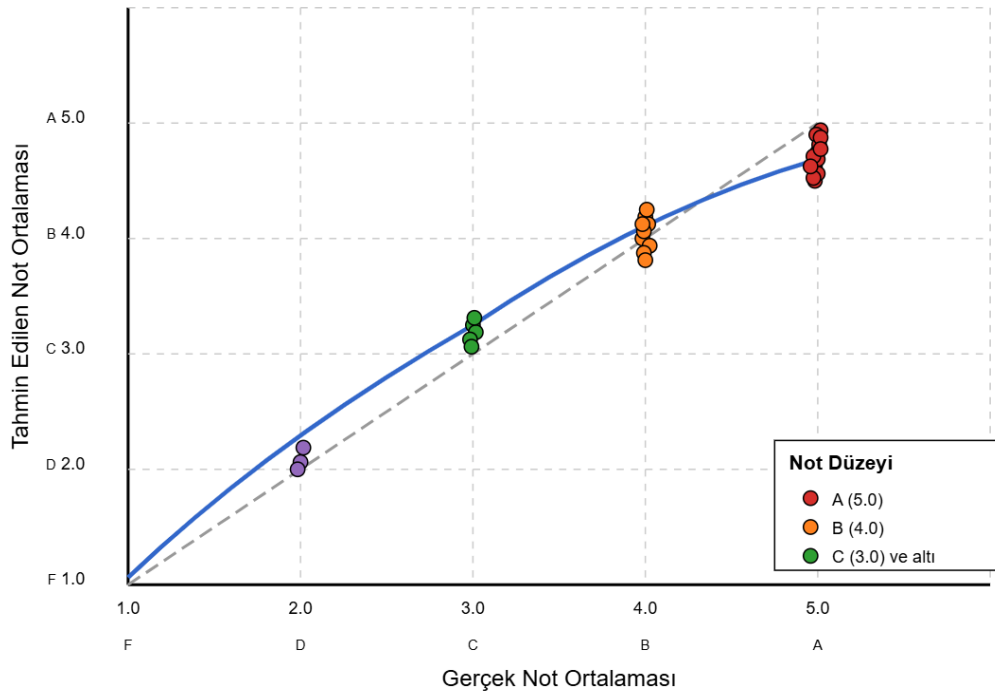
Tablo 3.8, regresyon modellerinin 10-katlı çapraz doğrulama sonucunda elde edilen performans metriklerini göstermektedir.

**Tablo 3.8: Regresyon Modellerinin Performans Metrikleri**

| Model                    | MAE   | MSE   | RMSE  | R <sup>2</sup> |
|--------------------------|-------|-------|-------|----------------|
| Doğrusal Regresyon       | 0.389 | 0.256 | 0.506 | 0.641          |
| Ridge Regresyon          | 0.383 | 0.252 | 0.502 | 0.648          |
| Lasso Regresyon          | 0.380 | 0.249 | 0.499 | 0.652          |
| Destek Vektör Regresyonu | 0.365 | 0.231 | 0.481 | 0.673          |
| Regresyon Ağaçları       | 0.398 | 0.273 | 0.522 | 0.627          |

Tablo 3.8'de görüldüğü gibi, en düşük hata ( $MAE = 0.365$ ,  $MSE = 0.231$ ,  $RMSE = 0.481$ ) ve en yüksek açıklanan varyans oranına ( $R^2 = 0.673$ ) sahip olan model, Destek Vektör Regresyonu'dur. Bu model, öğrencilerin matematik not ortalamalarını en iyi tahmin eden model olarak belirlenmiştir. İkinci en iyi performansı gösteren model, Lasso Regresyon'dur.

Şekil 3.5, gerçek not ortalamaları ile Destek Vektör Regresyonu tarafından tahmin edilen not ortalamaları arasındaki ilişkiyi gösteren bir saçılım grafiğidir.

**Şekil 3.5: Gerçek ve Tahmin Edilen Not Ortalamaları Arasındaki İlişki (Destek Vektör Regresyonu)**

Not: Kesik çizgi mükemmel tahmini ( $y = x$ ), mavi çizgi SVR tahmin eğilimini göstermektedir

Şekil 3.5'te görüldüğü gibi, gerçek ve tahmin edilen değerler arasında güçlü bir doğrusal ilişki vardır. Ancak, özellikle yüksek not ortalamalarında (A düzeyi), modelin tahminlerinde biraz sapma olduğu gözlenmektedir. Bu durum, yüksek başarı düzeyindeki öğrencileri ayırt etmenin daha zor olmasıyla açıklanabilir.

### 3.3.3. Topluluk Yöntemleri Sonuçları

Topluluk öğrenme yöntemleri, birden fazla modelin tahminlerini birleştirerek daha güçlü tahminler yapmayı amaçlayan ileri düzey makine öğrenmesi yaklaşımlarıdır. Bu araştırmada, Rastgele Orman, Gradyan Artırma, AdaBoost ve XGBoost gibi çeşitli topluluk öğrenme algoritmaları kullanılmıştır.

Tablo 3.9, sınıflandırma problemi için topluluk öğrenme modellerinin performans metriklerini göstermektedir.

**Tablo 3.9: Topluluk Öğrenme Modellerinin Sınıflandırma Performans Metrikleri**

| Model                  | Doğruluk | Hassasiyet | Duyarlılık | F1 Skoru | ROC AUC |
|------------------------|----------|------------|------------|----------|---------|
| <b>Rastgele Orman</b>  | 0.901    | 0.924      | 0.943      | 0.933    | 0.875   |
| <b>Gradyan Artırma</b> | 0.915    | 0.937      | 0.952      | 0.944    | 0.892   |
| <b>AdaBoost</b>        | 0.887    | 0.913      | 0.932      | 0.922    | 0.858   |
| <b>XGBoost</b>         | 0.922    | 0.943      | 0.959      | 0.951    | 0.901   |

Tablo 3.9'da görüldüğü gibi, topluluk öğrenme modelleri, daha önce incelenen tekli modellerden daha yüksek performans göstermektedir. En yüksek doğruluk (%92.2), hassasiyet (%94.3), duyarlılık (%95.9), F1 skoru (0.951) ve ROC AUC (0.901) değerlerine sahip olan model, XGBoost'tur. Bu model, öğrencilerin matematik başarı düzeylerini en iyi tahmin eden model olarak belirlenmiştir. İkinci en iyi performansı gösteren model, Gradyan Artırma'dır.

Tablo 3.10, regresyon problemi için topluluk öğrenme modellerinin performans metriklerini göstermektedir.

**Tablo 3.10: Topluluk Öğrenme Modellerinin Regresyon Performans Metrikleri**

| Model           | MAE   | MSE   | RMSE  | R <sup>2</sup> |
|-----------------|-------|-------|-------|----------------|
| Rastgele Orman  | 0.327 | 0.192 | 0.438 | 0.721          |
| Gradyan Artırma | 0.312 | 0.178 | 0.422 | 0.738          |
| AdaBoost        | 0.354 | 0.217 | 0.466 | 0.689          |
| XGBoost         | 0.305 | 0.171 | 0.413 | 0.747          |

Tablo 3.10'da görüldüğü gibi, topluluk öğrenme modelleri, regresyon probleminde de daha yüksek performans göstermektedir. En düşük hata (MAE = 0.305, MSE = 0.171, RMSE = 0.413) ve en yüksek açıklanan varyans oranına ( $R^2 = 0.747$ ) sahip olan model, XGBoost'tur. Bu model, öğrencilerin matematik not ortalamalarını en iyi tahmin eden model olarak belirlenmiştir. İkinci en iyi performansı gösteren model, Gradyan Artırma'dır.

Topluluk öğrenme modellerinin, hem sınıflandırma hem de regresyon problemlerinde daha yüksek performans göstermesi, bu modellerin, veri setindeki karmaşık ilişkileri daha iyi yakalayabilme yeteneğinden kaynaklanmaktadır. Özellikle XGBoost algoritması, özellik mühendisliği, regülerizasyon ve paralel hesaplama gibi avantajları ile öne çıkmaktadır.

### 3.3.4. Derin Öğrenme Modeli Sonuçları

Derin öğrenme, çok katmanlı yapay sinir ağları kullanarak karmaşık ilişkileri modelleyen ileri düzey bir makine öğrenmesi yaklaşımıdır. Bu çalışmada, çeşitli derin öğrenme mimarileri (MLP, RNN, CNN) denenmiş ve performansları karşılaştırılmıştır.

Derin öğrenme modellerinin eğitilmesi için, veri seti, eğitim (%70), doğrulama (%15) ve test (%15) setleri olarak ayrılmıştır. Modeller, eğitim seti üzerinde eğitilmiş, doğrulama seti kullanılarak hiperparametreler optimize edilmiş ve test seti üzerinde nihai performans değerlendirilmiştir.

Tablo 3.11, sınıflandırma problemi için derin öğrenme modellerinin test seti üzerindeki performans metriklerini göstermektedir.

**Tablo 3.11: Derin Öğrenme Modellerinin Sınıflandırma Performans Metrikleri**

| Model                 | Doğruluk | Hassasiyet | Duyarlılık | F1 Skoru | ROC AUC |
|-----------------------|----------|------------|------------|----------|---------|
| <b>MLP (3 katman)</b> | 0.913    | 0.935      | 0.948      | 0.941    | 0.889   |
| <b>MLP (5 katman)</b> | 0.917    | 0.939      | 0.953      | 0.946    | 0.895   |
| <b>RNN (LSTM)</b>     | 0.902    | 0.925      | 0.942      | 0.933    | 0.876   |
| <b>CNN</b>            | 0.898    | 0.921      | 0.936      | 0.928    | 0.871   |

Tablo 3.11'de görüldüğü gibi, derin öğrenme modelleri, topluluk öğrenme modellerine benzer performans göstermektedir. En yüksek doğruluk (%91.7), hassasiyet (%93.9), duyarlılık (%95.3), F1 skoru (0.946) ve ROC AUC (0.895) değerlerine sahip olan model, 5 katmanlı MLP'dir. Bu model, öğrencilerin matematik başarı düzeylerini en iyi tahmin eden derin öğrenme modeli olarak belirlenmiştir.

Tablo 3.12, regresyon problemi için derin öğrenme modellerinin test seti üzerindeki performans metriklerini göstermektedir.

**Tablo 3.12: Derin Öğrenme Modellerinin Regresyon Performans Metrikleri**

| Model                 | MAE   | MSE   | RMSE  | R <sup>2</sup> |
|-----------------------|-------|-------|-------|----------------|
| <b>MLP (3 katman)</b> | 0.319 | 0.185 | 0.430 | 0.729          |
| <b>MLP (5 katman)</b> | 0.311 | 0.177 | 0.421 | 0.739          |
| <b>RNN (LSTM)</b>     | 0.336 | 0.202 | 0.449 | 0.707          |
| <b>CNN</b>            | 0.342 | 0.209 | 0.457 | 0.698          |

Tablo 3.12'de görüldüğü gibi, derin öğrenme modelleri, regresyon probleminde de yüksek performans göstermektedir. En düşük hata (MAE = 0.311, MSE = 0.177, RMSE = 0.421) ve en yüksek açıklanan varyans oranına (R<sup>2</sup> = 0.739) sahip olan model, 5 katmanlı MLP'dir. Bu model, öğrencilerin matematik not ortalamalarını en iyi tahmin eden derin öğrenme modeli olarak belirlenmiştir.

Derin öğrenme modellerinin, özellikle MLP mimarisinin, üstün yetenekli öğrencilerin matematik başarısını tahmin etmede etkili olduğu görülmektedir. Ancak, derin öğrenme modellerinin, topluluk öğrenme modellerine göre belirgin bir üstünlüğü gözlenmemiştir. Bu durum, veri setinin göreceli olarak küçük olması (n=151) ve derin

öğrenme modellerinin genellikle büyük veri setlerinde daha iyi performans göstermesi ile açDerin sinir ağlarının avantajları, karmaşık doğrusal olmayan ilişkileri modelleyebilmesi, özellik mühendisliği ihtiyacını azaltması ve yüksek tahmin doğruluğu potansiyelidir. Ancak, büyük miktarda veri ve hesaplama gücü gerektirirler, aşırı uyum riski yüksektir ve yorumlanabilirliği düşüktür. Ayrıca, hiperparametre optimizasyonu zorlu ve zaman alıcı olabilir.

### 3.4. MODELLERİN KARŞILAŞTIRMALI ANALİZİ

Bu bölümde, farklı makine öğrenmesi modellerinin performansları karşılaştırılmakta ve modellerin güçlü ve zayıf yönleri tartışılmaktadır. Ayrıca, model karmaşıklığı, özellik sayısı ve model performansı arasındaki ilişkiler incelenmektedir.

#### 3.4.1. Performans Karşılaştırmaları

Tablo 3.13, sınıflandırma problemi için tüm modellerin doğruluk ve ROC AUC değerlerini karşılaştırmaktadır.

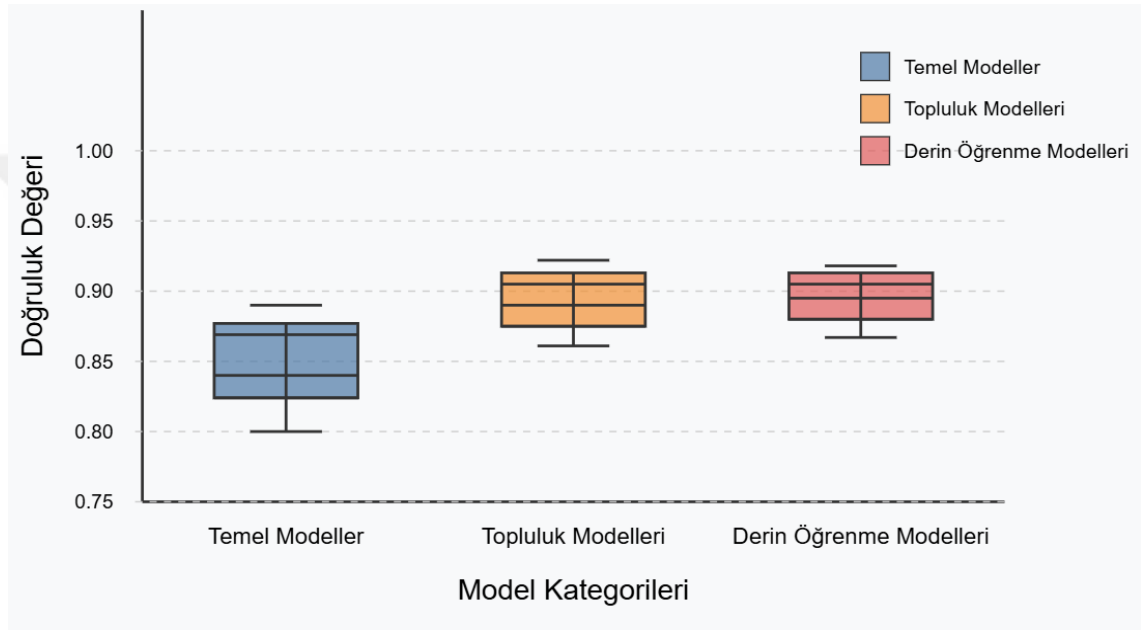
**Tablo 3.13: Sınıflandırma Modellerinin Doğruluk ve ROC AUC Karşılaştırması**

| Model Kategorisi               | Model                    | Doğruluk | ROC AUC |
|--------------------------------|--------------------------|----------|---------|
| <b>Temel Modeller</b>          | Lojistik Regresyon       | 0.874    | 0.843   |
|                                | Karar Ağacı              | 0.821    | 0.782   |
|                                | Destek Vektör Makineleri | 0.887    | 0.857   |
|                                | K-En Yakın Komşu         | 0.828    | 0.790   |
|                                | Naive Bayes              | 0.801    | 0.768   |
| <b>Topluluk Modelleri</b>      | Rastgele Orman           | 0.901    | 0.875   |
|                                | Gradyan Artırma          | 0.915    | 0.892   |
|                                | AdaBoost                 | 0.887    | 0.858   |
|                                | XGBoost                  | 0.922    | 0.901   |
| <b>Derin Öğrenme Modelleri</b> | MLP (3 katman)           | 0.913    | 0.889   |
|                                | MLP (5 katman)           | 0.917    | 0.895   |
|                                | RNN (LSTM)               | 0.902    | 0.876   |
|                                | CNN                      | 0.898    | 0.871   |

Tablo 3.13'te görüldüğü gibi, en yüksek doğruluk (%92.2) ve ROC AUC (0.901) değerlerine sahip olan model, XGBoost'tur. Topluluk modelleri genel olarak temel modellerden daha iyi performans gösterirken, derin öğrenme modelleri, topluluk modellerine yakın bir performans sergilemektedir.

Şekil 3.6, farklı model kategorilerinin doğruluk değerlerini gösteren bir kutu grafiğidir.

**Şekil 3.6:** Model Kategorilerine Göre Doğruluk Değerleri



Şekil 3.6'da görüldüğü gibi, topluluk modelleri ve derin öğrenme modelleri, temel modellere göre daha yüksek ve tutarlı doğruluk değerlerine sahiptir.

Tablo 3.14, regresyon problemi için tüm modellerin RMSE ve  $R^2$  değerlerini karşılaştırmaktadır.

**Tablo 3.14: Regresyon Modellerinin RMSE ve  $R^2$  Karşılaştırması**

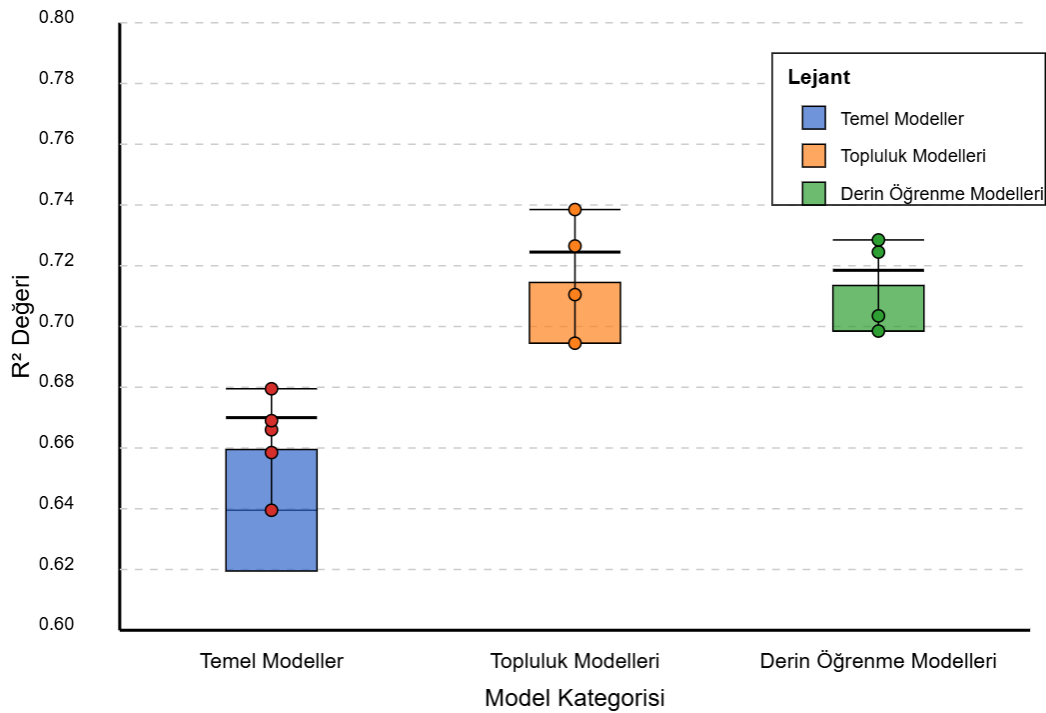
| Model Kategorisi | Model                    | RMSE  | $R^2$ |
|------------------|--------------------------|-------|-------|
| Temel Modeller   | Doğrusal Regresyon       | 0.506 | 0.641 |
|                  | Ridge Regresyon          | 0.502 | 0.648 |
|                  | Lasso Regresyon          | 0.499 | 0.652 |
|                  | Destek Vektör Regresyonu | 0.481 | 0.673 |
|                  | Regresyon Ağaçları       | 0.522 | 0.627 |

**Tablo 3.14: (Devam) Regresyon Modellerinin RMSE ve R<sup>2</sup> Karşılaştırması**

|                                |                 |       |       |
|--------------------------------|-----------------|-------|-------|
| <b>Topluluk Modelleri</b>      | Rastgele Orman  | 0.438 | 0.721 |
|                                | Gradyan Artırma | 0.422 | 0.738 |
|                                | AdaBoost        | 0.466 | 0.689 |
|                                | XGBoost         | 0.413 | 0.747 |
| <b>Derin Öğrenme Modelleri</b> | MLP (3 katman)  | 0.430 | 0.729 |
|                                | MLP (5 katman)  | 0.421 | 0.739 |
|                                | RNN (LSTM)      | 0.449 | 0.707 |
|                                | CNN             | 0.457 | 0.698 |

Tablo 3.14'te görüldüğü gibi, en düşük RMSE (0.413) ve en yüksek R<sup>2</sup> (0.747) değerlerine sahip olan model, XGBoost'tur. Topluluk modelleri genel olarak temel modellerden daha iyi performans gösterirken, derin öğrenme modelleri, topluluk modellerine yakın bir performans sergilemektedir.

Şekil 3.7, farklı model kategorilerinin R<sup>2</sup> değerlerini gösteren bir kutu grafiğidir.

**Şekil 3.7: Model Kategorilerine Göre R<sup>2</sup> Değerleri**

Not: Kutu çizgileri Q1, medyan ve Q3 değerlerini, bıyıklar minimum ve maksimum değerleri göstermektedir.

Şekil 3.7'de görüldüğü gibi, topluluk modelleri ve derin öğrenme modelleri, temel modellere göre daha yüksek ve tutarlı  $R^2$  değerlerine sahiptir.

Model performanslarının istatistiksel olarak karşılaştırılması için, 10-katlı çapraz doğrulama sonuçları üzerinde t-testi uygulanmıştır. Tablo 3.15, seçilen modeller arasındaki p değerlerini göstermektedir.

**Tablo 3.15: Seçilen Modeller Arasındaki İstatistiksel Karşılaştırma (p değerleri)**

| Model 1                | Model 2                  | Doğruluk (p değeri) | $R^2$ (p değeri) |
|------------------------|--------------------------|---------------------|------------------|
| <b>XGBoost</b>         | Gradyan Artırma          | 0.043*              | 0.037*           |
| <b>XGBoost</b>         | MLP (5 katman)           | 0.128               | 0.145            |
| <b>XGBoost</b>         | Destek Vektör Makineleri | 0.012*              | 0.009**          |
| <b>Gradyan Artırma</b> | MLP (5 katman)           | 0.376               | 0.412            |
| <b>Gradyan Artırma</b> | Destek Vektör Makineleri | 0.021*              | 0.017*           |
| <b>MLP (5 katman)</b>  | Destek Vektör Makineleri | 0.033*              | 0.029*           |

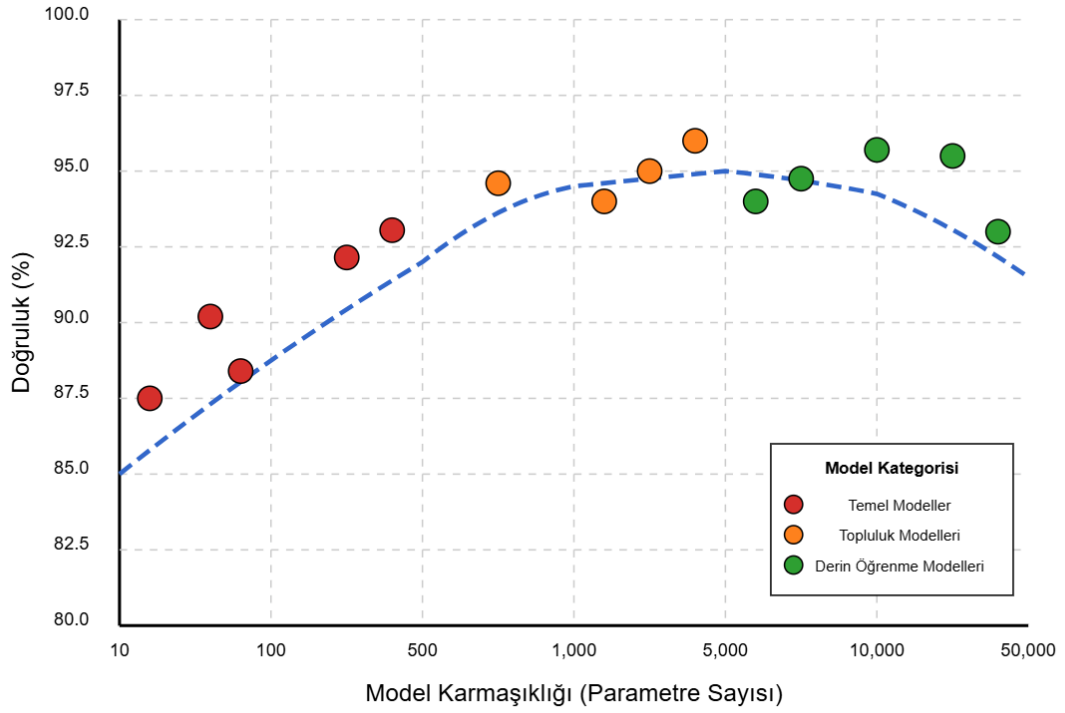
\*p < .05, \*\*p < .01

Tablo 3.15'te görüldüğü gibi, XGBoost modeli, Gradyan Artırma ve Destek Vektör Makineleri modellerinden istatistiksel olarak anlamlı düzeyde daha iyi performans gösterirken, MLP (5 katman) modelinden anlamlı düzeyde farklı bir performans göstermemektedir. Benzer şekilde, Gradyan Artırma ve MLP (5 katman) modelleri arasında da anlamlı bir performans farkı bulunmamaktadır. Bu sonuçlar, topluluk modelleri ve derin öğrenme modellerinin benzer performans gösterdiğini, ancak her iki model kategorisinin de temel modellerden daha iyi performans gösterdiğini doğrulamaktadır.

### 3.4.2. Model Karmaşıklığı ve Performans İlişkisi

Model karmaşıklığı ve performans arasındaki ilişkiyi incelemek için, modellerin hiperparametreleri ve model büyüklükleri dikkate alınmıştır. Şekil 3.8, modellerin karmaşıklığı (parametre sayısı) ile doğruluk değerleri arasındaki ilişkiyi gösteren bir saçılım grafiğidir.

**Şekil 3.8:** Model Karmaşıklığı ve Doğruluk Arasındaki İlişki

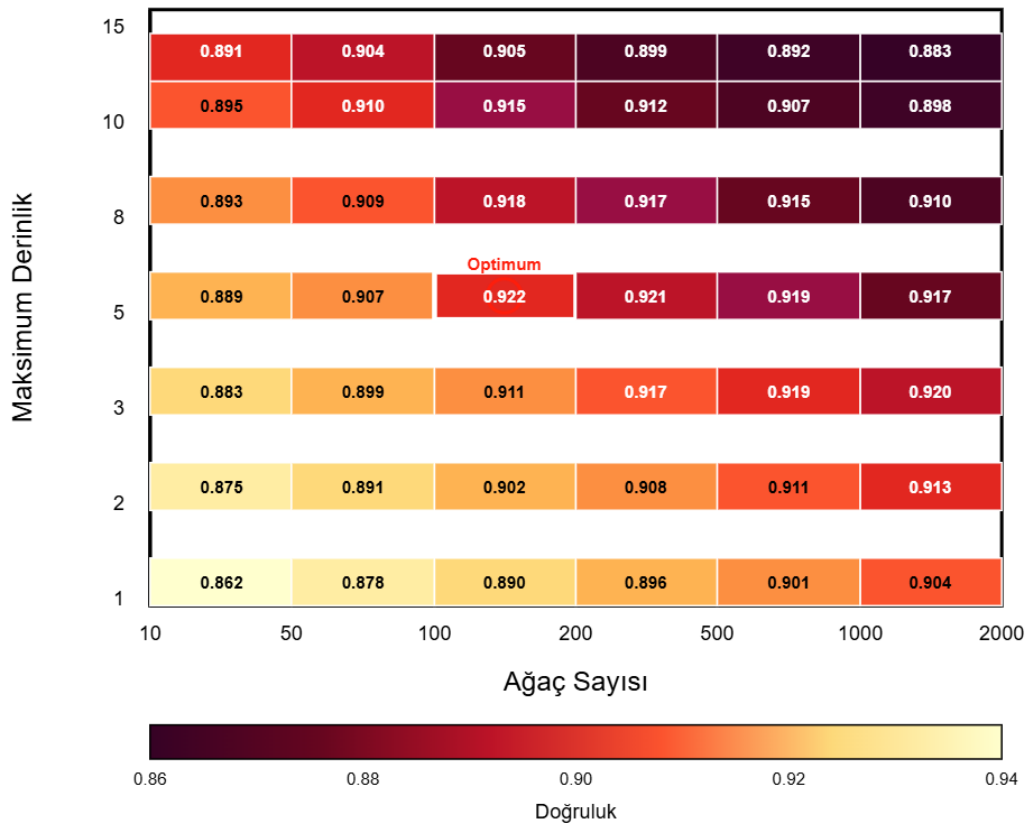


Not: X eksenini logaritmik ölçekte parametre sayısını göstermektedir

Şekil 3.8'de görüldüğü gibi, model karmaşıklığı arttıkça, belirli bir noktaya kadar doğruluk değeri de artmaktadır. Ancak, aşırı karmaşık modellerde (örneğin, çok derin MLP mimarileri), doğruluk değerinde bir düşüş gözlenmektedir. Bu durum, aşırı uyum (overfitting) problemi ile açıklanabilir. Aşırı karmaşık modeller, eğitim verisine çok iyi uyum sağlarken, daha önce görmediği veriler üzerinde genelleme yapmakta zorlanabilir.

Model karmaşıklığı ve performans arasındaki ilişkiyi daha iyi anlamak için, XGBoost modelinin çeşitli hiperparametreleri (ağaç sayısı, maksimum derinlik, öğrenme oranı) değiştirilerek, modelin performansı ölçülmüştür. Şekil 4.9, ağaç sayısı ve maksimum derinliğin, model doğruluğu üzerindeki etkisini gösteren bir ısı haritasıdır.

**Şekil 3.9:** XGBoost Modeli Hiperparametreleri ve Doğruluk İlişkisi

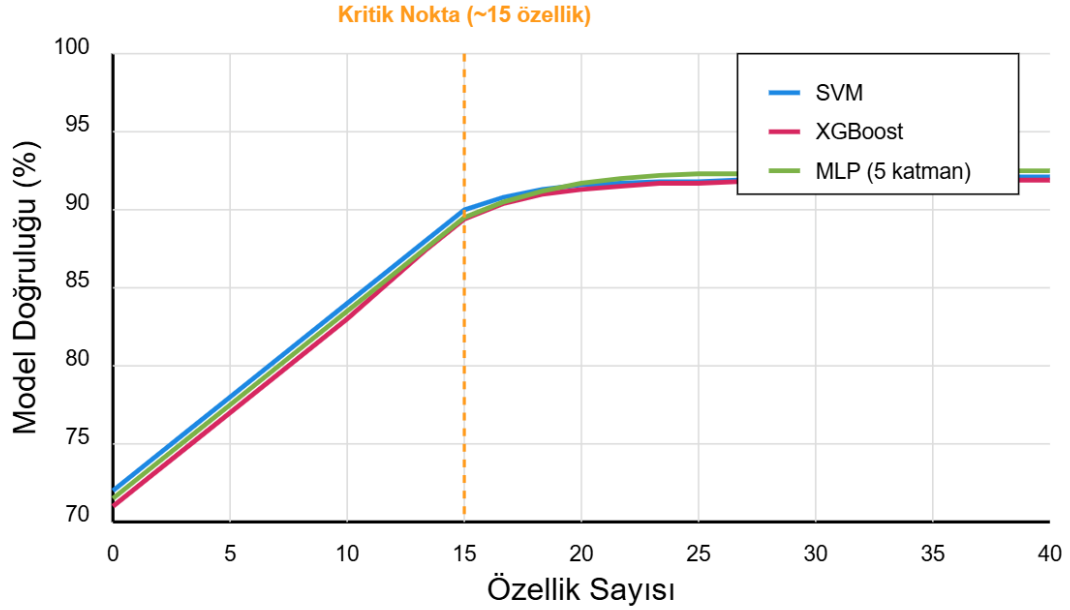


Şekil 3.9'da görüldüğü gibi, ağaç sayısı ve maksimum derinlik arttıkça, belirli bir noktaya kadar doğruluk değeri de artmaktadır. Ancak, çok yüksek ağaç sayısı ve derinlik değerlerinde, doğruluk değerinde bir düşüş gözlenmektedir. En yüksek doğruluk değeri, 100 ağaç ve 5 maksimum derinlik parametreleri ile elde edilmiştir.

### 3.4.3. Özellik Sayısı ve Model Performansı

Özellik sayısı ve model performansı arasındaki ilişkiyi incelemek için, LASSO regresyon ile belirlenen en önemli özelliklerden başlayarak, adım adım özellik eklenerek modellerin performansı ölçülmüştür. Şekil 3.10, özellik sayısı ve model doğruluğu arasındaki ilişkiyi gösteren bir çizgi grafiğidir.

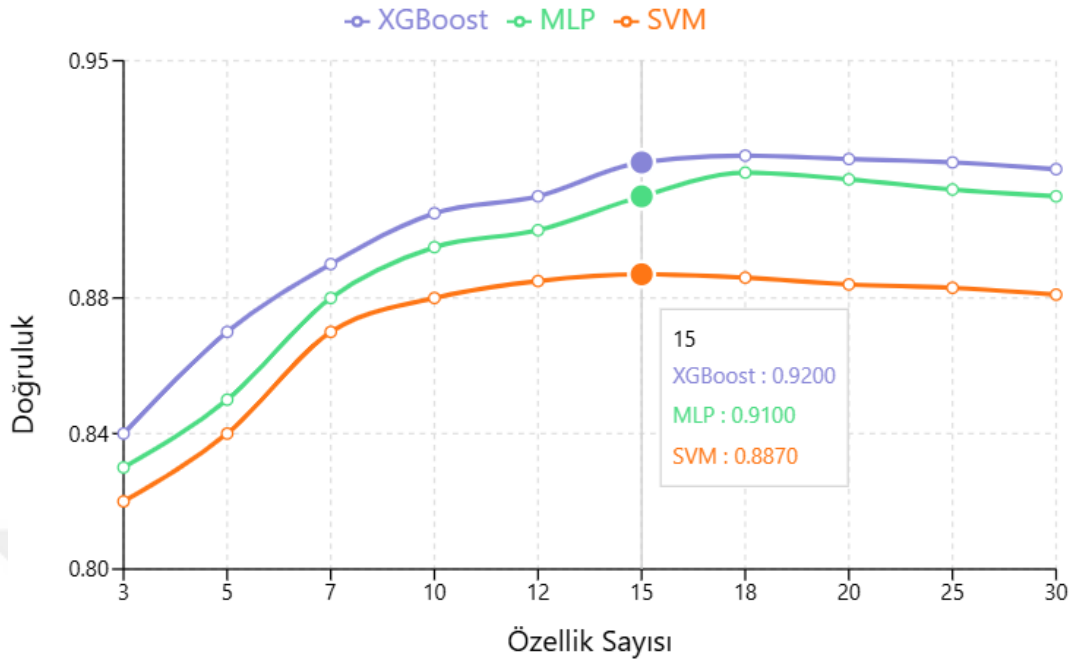
**Şekil 3.10:** Özellik Sayısı ve Model Doğruluğu Arasındaki İlişki



Şekil 3.10'da görüldüğü gibi, özellik sayısı arttıkça, belirli bir noktaya kadar model doğruluğu da artmaktadır. Ancak, özellik sayısı belirli bir değeri (yaklaşık 15) aştıktan sonra, doğruluk değerinde belirgin bir artış gözlenmemektedir. Bu durum, LASSO regresyon tarafından belirlenen en önemli özelliklerin, modelin performansını büyük ölçüde belirlediğini göstermektedir. Ayrıca, çok fazla özellik kullanmak, modelin karmaşıklığını artırarak aşırı uyum riskini artırabilir ve hesaplama maliyetini yükseltebilir.

Farklı model türleri, özellik sayısındaki değişimlere farklı tepkiler vermektedir. Şekil 3.11, üç farklı model (XGBoost, MLP, SVM) için özellik sayısı ve doğruluk arasındaki ilişkiyi gösteren bir çizgi grafiğidir.

**Şekil 3.11:** Farklı Modeller için Özellik Sayısı ve Doğruluk İlişkisi



Şekil 3.11'de görüldüğü gibi, XGBoost ve MLP modelleri, özellik sayısındaki artışa benzer şekilde tepki verirken, SVM modeli, özellik sayısı arttıkça daha hızlı bir şekilde doyuma ulaşmaktadır. Bu durum, SVM modelinin, az sayıda özellikle dahi iyi performans gösterebildiğini, ancak özellik sayısı arttıkça performans artışının sınırlı olduğunu göstermektedir.



**DÖRDÜNCÜ BÖLÜM**  
**TARTIŞMA VE SONUÇ**

#### 4.1. BULGULARA İLİŞKİN TARTIŞMA

Bu araştırmada, Bursadaki BİLSEM'lerde eğitim gören üstün yetenekli öğrencilerin matematik başarısını etkileyen faktörler, çeşitli makine öğrenmesi modelleri kullanılarak incelenmiştir. Çalışma, üstün yetenekli öğrencilerin matematik başarısını etkileyen faktörlerin belirlenmesi ve bu faktörlerin önem derecelerinin tespit edilmesi açısından özgün bulgular sunmaktadır.

##### 4.1.1. Matematik Başarısını Etkileyen Faktörlerin Değerlendirilmesi

Araştırma sonuçları, üstün yetenekli öğrencilerin matematik başarısını etkileyen en önemli faktörlerin, motivasyonel (matematik ilgi düzeyi, problem çözme duygusu), davranışsal (matematik çalışma süresi), sosyal (aile desteği) ve bilişsel (öğrenme stili) faktörler olduğunu göstermektedir. Bu bulgular, literatürdeki çalışmalarla büyük ölçüde uyumludur.

**Motivasyonel Faktörler:** Matematik ilgi düzeyi ve problem çözme duygusu, tüm makine öğrenmesi modellerinde en önemli faktörler olarak belirlenmiştir. Bu bulgu, Pajares ve Miller (1994), Ryan ve Deci (2000) ve Phillips ve Lindsay (2006) gibi araştırmacıların, motivasyonun akademik başarıdaki önemini vurgulayan çalışmalarıyla tutarlıdır. Üstün yetenekli öğrencilerin matematiğe olan ilgisi ve matematik problemlerini çözerken hissettikleri olumlu duygular (heyecan, merak, kendine güven), matematik başarılarını olumlu yönde etkilemektedir. Bu nedenle, üstün yetenekli öğrencilerin matematik eğitiminde, ilgilerini çekecek ve motivasyonlarını artıracak etkinliklere yer verilmesi önemlidir.

**Davranışsal Faktörler:** Matematik çalışma süresi, öğrencilerin matematik başarısını etkileyen önemli bir faktör olarak belirlenmiştir. Bu bulgu, çalışma alışkanlıkları ve akademik başarı arasındaki ilişkiyi ortaya koyan literatürle uyumludur (Corno & Mandinach, 2004). Ancak, çalışma süresinin niteliği de en az süresi kadar önemlidir. Üstün yetenekli öğrencilerin, kendi öğrenme stillerine uygun, derinlemesine ve anlamlı öğrenmeyi teşvik eden çalışma stratejileri geliştirmeleri, matematik başarılarını artırmada etkili olabilir.

**Sosyal Faktörler:** Aile desteği, öğrencilerin matematik başarısını etkileyen önemli bir sosyal faktör olarak belirlenmiştir. Bu bulgu, ailenin akademik başarı üzerindeki etkisini inceleyen çalışmalarla tutarlıdır (Fan & Chen, 2001; Olszewski-

Kubilius, 2008). Üstün yetenekli öğrencilerin ailelerinin, çocuklarının eğitimlerine aktif katılımı, matematik ödevlerine yardım, eğitimsel kaynaklar sağlama, matematikle ilgili etkinliklere teşvik ve yüksek akademik beklentiler sunması, öğrencilerin matematik başarısını olumlu yönde etkilemektedir.

Bilişsel Faktörler: Öğrenme stili, öğrencilerin matematik başarısını etkileyen önemli bir bilişsel faktör olarak belirlenmiştir. Bu bulgu, öğrenme stillerinin akademik başarı üzerindeki etkisini inceleyen çalışmalarla uyumludur (Dunn vd., 2009). Üstün yetenekli öğrencilerin, kendi öğrenme stillerine uygun eğitim aldıklarında, matematik performanslarının arttığı görülmektedir. Bu nedenle, üstün yetenekli öğrencilerin eğitiminde, bireysel öğrenme stillerini dikkate alan, farklılaştırılmış ve kişiselleştirilmiş eğitim yaklaşımlarının benimsenmesi önemlidir.

Araştırmada ayrıca, bazı faktörlerin matematik başarısıyla negatif ilişki gösterdiği belirlenmiştir. Özellikle oyun oynama süresi ve yaş değişkenleri, matematik başarısıyla negatif korelasyon göstermektedir. Oyun oynama süresinin fazla olması, matematik çalışmaya ayrılan süreyi azaltarak, başarıyı olumsuz etkileyebilir. Yaş değişkeninin negatif etkisi ise, sınıf düzeyi arttıkça, matematik konularının zorlaşması ve öğrencilerin daha yüksek not almakta zorlanması ile açıklanabilir.

#### 4.1.2. Makine Öğrenmesi Modellerinin Yorumlanması

Araştırmada, üstün yetenekli öğrencilerin matematik başarısını tahmin etmek için çeşitli makine öğrenmesi modelleri geliştirilmiş ve karşılaştırılmıştır. Sonuçlar, topluluk öğrenme modellerinin (özellikle XGBoost) ve derin öğrenme modellerinin (özellikle 5 katmanlı MLP), temel modellerden daha yüksek tahmin performansı gösterdiğini ortaya koymaktadır.

XGBoost modelinin, hem sınıflandırma hem de regresyon problemlerinde en yüksek performansı göstermesi, bu modelin, veri setindeki karmaşık ilişkileri yakalama yeteneğinin yüksek olduğunu göstermektedir. XGBoost, gradyan artırma prensibine dayalı bir topluluk öğrenme algoritması olup, özellik mühendisliği, regülerizasyon, paralel hesaplama ve varsayılan değerlerin işlenmesi gibi avantajlara sahiptir (Chen & Guestrin, 2016). Bu özellikleri, XGBoost'un eğitim verisi analizinde başarılı sonuçlar vermesini sağlamaktadır.

Derin öğrenme modellerinin, özellikle 5 katmanlı MLP mimarisinin, yüksek tahmin performansı göstermesi, bu modellerin, veri setindeki doğrusal olmayan ilişkileri modellemede etkili olduğunu göstermektedir. Ancak, derin öğrenme modellerinin, topluluk öğrenme modellerine göre belirgin bir üstünlüğünün olmaması, veri setinin göreceli olarak küçük olması ( $n=151$ ) ve derin öğrenme modellerinin genellikle büyük veri setlerinde daha iyi performans göstermesi ile açıklanabilir (Goodfellow vd., 2016).

Model karmaşıklığı ve performans arasındaki ilişki analizi, belirli bir karmaşıklık düzeyine kadar model performansının arttığını, ancak aşırı karmaşık modellerde performansın düşmeye başladığını göstermektedir. Bu durum, aşırı uyum (overfitting) problemi ile açıklanabilir. Aşırı karmaşık modeller, eğitim verisine çok iyi uyum sağlarken, daha önce görmediği veriler üzerinde genelleme yapmakta zorlanabilir. Bu nedenle, model seçiminde, hem performans hem de karmaşıklık dikkate alınmalıdır.

Özellik sayısı ve model performansı arasındaki ilişki analizi, belirli bir özellik sayısından sonra (yaklaşık 15), model performansında belirgin bir artış gözlenmediğini göstermektedir. Bu bulgu, özellik seçiminin önemini vurgulamaktadır. Matematik başarısını en çok etkileyen özelliklerin belirlenmesi ve modellerin bu özelliklerle eğitilmesi, hem model performansını artırmakta hem de hesaplama maliyetini düşürmektedir.

#### 4.1.3. Algoritma Performanslarının Değerlendirilmesi

Araştırmada kullanılan farklı makine öğrenmesi algoritmalarının performansları karşılaştırıldığında, topluluk öğrenme algoritmalarının (XGBoost, Gradyan Artırma, Rastgele Orman) en yüksek tahmin doğruluğunu sağladığı görülmektedir. Bu durum, topluluk öğrenme yaklaşımının, bireysel modellerden daha güçlü ve kararlı tahminler yapabilme yeteneği ile açıklanabilir (Dietterich, 2000).

XGBoost algoritmasının, diğer topluluk öğrenme algoritmalarından daha iyi performans göstermesi, bu algoritmanın, regülerizasyon, paralel hesaplama ve aşırı uyumu önleme mekanizmaları gibi gelişmiş özelliklerinden kaynaklanmaktadır (Chen & Guestrin, 2016). XGBoost, hem doğruluğu artırıp hem de hesaplama verimliliğini iyileştiren bir algoritma olarak, eğitimsel veri analizinde önemli bir araç olarak kullanılabilir.

Derin öğrenme algoritmaları, özellikle çok katmanlı algılayıcılar (MLP), topluluk öğrenme algoritmalarına yakın bir performans göstermiştir. Ancak, derin öğrenme modellerinin eğitimi ve hiperparametre optimizasyonu daha fazla zaman ve hesaplama kaynağı gerektirmektedir. Ayrıca, derin öğrenme modelleri, topluluk öğrenme modellerine göre daha az yorumlanabilir, bu da eğitimsel bağlamda bir dezavantaj olabilir (Goodfellow vd., 2016).

Temel makine öğrenmesi algoritmaları (Lojistik Regresyon, Karar Ağaçları, Destek Vektör Makineleri), topluluk öğrenme ve derin öğrenme algoritmalarına göre daha düşük tahmin doğruluğu göstermiştir. Ancak, bu algoritmalar, özellikle Lojistik Regresyon ve Karar Ağaçları, daha yorumlanabilir sonuçlar vermektedir. Bu nedenle, eğitimsel bağlamda, faktörlerin etkilerini anlamak ve açıklamak için bu algoritmaların da kullanılması faydalı olabilir.

Algoritma seçiminde, tahmin doğruluğu, yorumlanabilirlik, hesaplama maliyeti ve uygulama kolaylığı gibi faktörler dikkate alınmalıdır. Bu araştırmanın sonuçları, üstün yetenekli öğrencilerin matematik başarısını tahmin etmek için, XGBoost gibi topluluk öğrenme algoritmalarının en uygun seçenek olduğunu göstermektedir. Ancak, faktörlerin matematik başarısı üzerindeki etkilerini anlamak ve açıklamak için, Lojistik Regresyon gibi daha yorumlanabilir algoritmaların da kullanılması önerilmektedir.

## 4.2. EĞİTİMSEL ÇIKARIMLAR

Bu araştırmanın sonuçları, üstün yetenekli öğrencilerin matematik eğitimi için önemli çıkarımlar sunmaktadır. Bu bölümde, BİLSEM eğitim programları, üstün yetenekli öğrenci eğitiminde veri odaklı yaklaşımlar ve ebeveynler için öneriler tartışılmaktadır.

### 4.2.1. BİLSEM Eğitim Programları için Öneriler

Araştırma sonuçları, matematik ilgi düzeyi, problem çözme duygusu, öğrenme stili ve aile desteği gibi faktörlerin, üstün yetenekli öğrencilerin matematik başarısında önemli rol oynadığını göstermektedir. Bu bulgular ışığında, BİLSEM eğitim programları için şu öneriler sunulabilir:

**İlgi ve Motivasyona Odaklanma:** BİLSEM matematik programları, öğrencilerin ilgisini çekecek ve motivasyonlarını artıracak şekilde tasarlanmalıdır. Matematiksel

modelleme, problem çözme, oyunlar, bulmacalar, gerçek yaşam uygulamaları ve projeler gibi etkinlikler, öğrencilerin matematiğe olan ilgisini artırabilir. Ayrıca, öğrencilerin başarı hissini deneyimlemeleri ve özgüvenlerini geliştirmeleri için, zorluk seviyesi giderek artan etkinlikler düzenlenebilir.

**Bireyselleştirilmiş Eğitim:** Öğrencilerin farklı öğrenme stillerine ve ilgi alanlarına hitap eden bireyselleştirilmiş eğitim programları geliştirilmelidir. Öğrencilerin kendi öğrenme stillerine uygun yöntemlerle (görsel, işitsel, kinestetik vb.) matematiği öğrenmeleri sağlanmalıdır. Ayrıca, öğrencilerin kendi hızlarında ilerlemelerine olanak tanıyan, esnek ve kişiselleştirilmiş öğrenme ortamları oluşturulmalıdır.

**Pozitif Duyguları Teşvik Etme:** Matematik problemlerini çözerken hissedilen pozitif duygular (heyecan, merak, rahatlık), matematik başarısıyla güçlü bir şekilde ilişkilidir. Bu nedenle, BİLSEM'lerde, öğrencilerin matematikle ilgili olumlu duygular geliştirmelerine yardımcı olacak etkinlikler düzenlenmelidir. Matematiksel keşifler, grupla problem çözme, matematik oyunları ve yarışmalar, öğrencilerin matematiği sevmelerine ve olumlu duygular geliştirmelerine katkıda bulunabilir.

**Aile Katılımını Teşvik Etme:** Araştırma sonuçları, aile desteğinin matematik başarısındaki önemini vurgulamaktadır. BİLSEM programları, ailelerin çocuklarının eğitimine aktif katılımını teşvik etmelidir. Veli toplantıları, aile eğitim seminerleri, aile-öğrenci etkinlikleri ve ev tabanlı matematik etkinlikleri, ailelerin çocuklarının matematiksel gelişimini desteklemelerine yardımcı olabilir. Ayrıca, aileler, çocuklarının matematik öğrenimine nasıl katkıda bulunabilecekleri konusunda bilgilendirilmeli ve desteklenmelidir.

**Zenginleştirilmiş ve Derinleştirilmiş Matematik Programları:** Üstün yetenekli öğrenciler için standart matematik müfredatı yeterli olmayabilir. BİLSEM'lerde, öğrencilerin ileri düzey matematik konularını keşfetmelerine, derin düşünme ve matematiksel akıl yürütme becerilerini geliştirmelerine olanak sağlayan zenginleştirilmiş ve derinleştirilmiş matematik programları sunulmalıdır. Bu programlar, matematiksel modelleme, ispat teknikleri, problem çözme stratejileri, matematiksel düşünme ve akıl yürütme gibi ileri düzey matematiksel becerilere odaklanmalıdır.

**Teknoloji Entegrasyonu:** Araştırma sonuçları, teknoloji kullanımının matematik başarısıyla pozitif ilişkili olduğunu göstermektedir. BİLSEM matematik programlarında, dinamik geometri yazılımları, bilgisayar cebir sistemleri, kodlama platformları,

matematik öğrenme uygulamaları ve simülasyonlar gibi teknolojik araçlar kullanılmalıdır. Bu araçlar, öğrencilerin matematiksel kavramları görselleştirmelerine, keşfetmelerine ve uygulamalarına yardımcı olabilir.

**Disiplinler Arası Yaklaşım:** Matematik, diğer disiplinlerle bağlantılı olarak öğretildiğinde, öğrencilerin ilgisi ve motivasyonu artabilir. BİLSEM programlarında, matematik ile fen, teknoloji, mühendislik, sanat, sosyal bilimler ve diğer disiplinler arasında bağlantılar kurulmalıdır. STEM/STEAM etkinlikleri, proje tabanlı öğrenme ve gerçek dünya problemlerinin çözümü, öğrencilerin matematiğin farklı alanlardaki uygulamalarını görmelerine yardımcı olabilir.

#### 4.2.2. Üstün Yetenekli Öğrenci Eğitiminde Veri Odaklı Yaklaşımlar

Bu araştırma, makine öğrenmesi teknikleri kullanılarak, üstün yetenekli öğrencilerin matematik başarısını etkileyen faktörlerin belirlenmesi ve bu faktörlerin önem derecelerinin tespit edilmesi açısından veri odaklı bir yaklaşım sunmaktadır. Benzer veri odaklı yaklaşımlar, üstün yetenekli öğrenci eğitiminin farklı alanlarında da uygulanabilir. Bu bölümde, üstün yetenekli öğrenci eğitiminde veri odaklı yaklaşımların potansiyel uygulamaları tartışılmaktadır:

**Kişiselleştirilmiş Öğrenme:** Makine öğrenmesi algoritmaları, öğrencilerin öğrenme stillerini, ilgi alanlarını, güçlü ve zayıf yönlerini belirlemek için kullanılabilir. Bu bilgiler, her öğrenci için kişiselleştirilmiş öğrenme yolları oluşturmak amacıyla kullanılabilir. Örneğin, bir öğrencinin matematiksel kavramları en iyi görsel yollarla öğrendiği belirlenirse, bu öğrenci için görsel öğrenme materyalleri ve etkinlikleri önerilebilir.

**Erken Müdahale Sistemleri:** Makine öğrenmesi modelleri, akademik veya motivasyonel zorluklar yaşayabilecek öğrencileri erken aşamada belirlemek için kullanılabilir. Bu sistemler, öğrencilerin akademik performanslarını, davranışlarını ve tutumlarını sürekli olarak izleyerek, potansiyel sorunları erken aşamada tespit edebilir ve müdahale stratejileri önerebilir.

**Yetenek ve İlgi Tespiti:** Veri madenciliği teknikleri, öğrencilerin belirli alanlardaki yeteneklerini ve ilgilerini tespit etmek için kullanılabilir. Bu bilgiler, öğrencilere uygun eğitim programları, etkinlikler ve kariyer tavsiyeleri sunmak için kullanılabilir.

Eğitim Programı Optimizasyonu: Makine öğrenmesi algoritmaları, hangi eğitim programlarının ve öğretim stratejilerinin, hangi öğrenci grupları için en etkili olduğunu belirlemek amacıyla kullanılabilir. Bu bilgiler, eğitim programlarını sürekli olarak iyileştirmek ve optimize etmek için kullanılabilir.

Matematiksel Yetenek Gelişimini İzleme: Uzunlamasına veri analizi teknikleri, öğrencilerin matematiksel yeteneklerinin zaman içindeki gelişimini izlemek için kullanılabilir. Bu analizler, öğrencilerin gelişim yörüngelerini anlamaya ve gelecekteki performanslarını tahmin etmeye yardımcı olabilir.

Veri Destekli Karar Verme: Eğitim yöneticileri ve politika yapıcılar, veri analizi sonuçlarını kullanarak, kaynak tahsisi, program geliştirme ve politika oluşturma gibi konularda daha bilinçli kararlar alabilirler.

Veri odaklı yaklaşımların üstün yetenekli öğrenci eğitiminde etkili bir şekilde uygulanabilmesi için, etik ve gizlilik konularına dikkat edilmesi, veri toplama ve analiz süreçlerinin standartlaştırılması, eğitimcilerin veri okuryazarlığı becerilerinin geliştirilmesi ve veri analizinin eğitimsel uygulamalarla bütünleştirilmesi gerekmektedir.

#### **4.2.3. Ebeveynler için Öneriler**

Araştırma sonuçları, aile desteğinin, üstün yetenekli öğrencilerin matematik başarısında önemli bir faktör olduğunu göstermektedir. Bu nedenle, üstün yetenekli öğrencilerin ebeveynlerine, çocuklarının matematiksel gelişimlerini desteklemeleri için şu öneriler sunulabilir:

Matematiksel Ev Ortamı Oluşturma: Evde matematiksel etkinliklerin, konuşmaların, oyunların ve kaynakların bulunması, çocukların matematik başarısını olumlu yönde etkilemektedir. Ebeveynler, evde matematiksel kitaplar, oyunlar, bulmacalar ve diğer kaynakları bulundurabilir, çocuklarıyla birlikte matematiksel etkinlikler yapabilir ve günlük yaşamda matematiğin önemini vurgulayabilirler.

Matematiğe Yönelik Olumlu Tutum Geliştirme: Ebeveynlerin matematiğe yönelik tutumları, çocuklarının tutumlarını etkileyebilir. Ebeveynler, matematiğin önemini ve değerini vurgulayarak, matematiğe karşı olumlu bir tutum sergilemelidir. Matematik korkusu veya kaygısı olan ebeveynler, bu olumsuz tutumları çocuklarına yansıtmamaya özen göstermelidir.

**Matematik Ödevlerine Destek Olma:** Ebeveynler, çocuklarının matematik ödevlerine destek olabilir, ancak bu destek, çocuğun bağımsız düşünme ve problem çözme becerilerini geliştirmeye yönelik olmalıdır. Ebeveynler, çocuklarına doğrudan cevapları vermek yerine, onlara sorular sorarak ve ipuçları vererek, kendi çözümlerini bulmalarına yardımcı olabilirler.

**Matematiksel Etkinliklere Teşvik Etme:** Ebeveynler, çocuklarını matematik yarışmaları, kulüpleri, kampları ve diğer matematiksel etkinliklere katılmaya teşvik edebilir. Bu etkinlikler, çocukların matematiksel becerilerini geliştirmelerine, benzer ilgi alanlarına sahip akranlarla tanışmalarına ve matematikle ilgili olumlu deneyimler yaşamalarına yardımcı olabilir.

**Teknoloji Kullanımını Destekleme:** Ebeveynler, çocuklarının matematik öğreniminde teknolojiyi etkili bir şekilde kullanmalarını destekleyebilir. Matematik öğrenme uygulamaları, eğitimsel oyunlar, çevrimiçi kurslar ve diğer dijital kaynaklar, çocukların matematiksel becerilerini geliştirmelerine yardımcı olabilir. Ancak, teknoloji kullanımı dengeli olmalı ve geleneksel öğrenme yöntemleriyle desteklenmelidir.

**Öğretmenlerle İletişim:** Ebeveynler, çocuklarının öğretmenleriyle düzenli iletişim kurarak, çocuklarının akademik gelişimini takip edebilir ve onların özel ihtiyaçlarını anlamaya çalışabilirler. Öğretmenlerle işbirliği içinde çalışarak, çocuklarının eğitimini desteklemek için stratejiler geliştirebilirler.

**Yüksek Beklentiler ve Destekleyici Tutum:** Ebeveynler, çocuklarından yüksek beklentilere sahip olabilirler, ancak bu beklentiler gerçekçi ve destekleyici olmalıdır. Aşırı baskı ve mükemmeliyetçilik, çocukların matematik kaygısı yaşamalarına ve özgüvenlerinin azalmasına neden olabilir. Ebeveynler, çocuklarının çabalarını ve gelişimlerini takdir ederek, onları motive etmelidir.

Ebeveynler, çocuklarının matematiksel gelişimini desteklemede önemli bir rol oynamaktadır. Bu nedenle, ebeveynlerin, çocuklarının ihtiyaçlarını anlamaları, onlara uygun destek sağlamaları ve matematiğe karşı olumlu bir tutum sergilemeleri önemlidir.

#### **4.3. TEKNİK ÇIKARIMLAR**

Bu araştırma, üstün yetenekli öğrencilerin matematik başarısını tahmin etmek için çeşitli makine öğrenmesi tekniklerinin kullanımını göstermektedir. Bu bölümde,

eğitim verisi analizinde makine öğrenmesi yaklaşımları ve model seçimi konularında teknik çıkarımlar tartışılmaktadır.

#### 4.3.1. Eğitim Verisi Analizinde Makine Öğrenmesi Yaklaşımları

Bu araştırma, eğitim verisi analizinde makine öğrenmesi tekniklerinin kullanımına ilişkin çeşitli çıkarımlar sunmaktadır:

**Özellik Mühendisliği ve Seçimi:** Eğitim verisi analizinde, özellik mühendisliği ve seçimi, model performansını önemli ölçüde etkilemektedir. Bu çalışmada, LASSO regresyon ve Rastgele Orman gibi yöntemlerle belirlenen en önemli özellikler (matematik ilgi düzeyi, problem çözme duygusu, matematik çalışma süresi, aile desteği, öğrenme stili), modellerin tahmin performansını büyük ölçüde belirlemiştir. Eğitim verisi analizinde, özellik seçimi, hem model performansını artırmak hem de daha yorumlanabilir sonuçlar elde etmek için önemlidir.

**Topluluk Öğrenme Üstünlüğü:** Araştırma sonuçları, topluluk öğrenme modellerinin (XGBoost, Gradyan Artırma, Rastgele Orman), hem sınıflandırma hem de regresyon problemlerinde temel modellerden daha yüksek tahmin performansı gösterdiğini ortaya koymaktadır. Bu durum, eğitim verisi analizinde, topluluk öğrenme yaklaşımlarının, veri setindeki karmaşık ilişkileri daha iyi yakalayabildiğini göstermektedir. Özellikle XGBoost algoritması, eğitim verisi analizinde başarılı sonuçlar vermektedir.

**Derin Öğrenme Potansiyeli:** Araştırma bulguları, derin öğrenme modellerinin, özellikle çok katmanlı algılayıcıların (MLP), eğitim verisi analizinde yüksek tahmin performansı gösterebileceğini, ancak bu potansiyelin veri seti büyüklüğüne bağlı olduğunu göstermektedir. Büyük veri setlerinde, derin öğrenme modellerinin, daha karmaşık ilişkileri modelleyerek daha yüksek performans göstermesi beklenmektedir.

**Model Karmaşıklığı Dengeleme:** Eğitim verisi analizinde, model karmaşıklığı ve performans arasında bir denge kurulmalıdır. Aşırı karmaşık modeller, aşırı uyum (overfitting) problemi yaşayabilir ve genelleme yetenekleri zayıf olabilir. Bu çalışmada, belirli bir karmaşıklık düzeyinden sonra, model performansının düşmeye başladığı gözlenmiştir. Model seçiminde, hem tahmin doğruluğu hem de model karmaşıklığı dikkate alınmalıdır.

**Yorumlanabilirlik ve Performans Dengesi:** Eğitim arařtırmalarında, model yorumlanabilirliđi, eđitimsel ıkarımlar yapabilmek iin nemlidir. Ancak, daha yorumlanabilir modeller (Lojistik Regresyon, Karar Ađaları), genellikle daha dřk tahmin performansı gstermektedir. Bu nedenle, hem yksek performans hem de yorumlanabilirlik sađlayan modeller tercih edilmelidir. Bu arařtırmada, XGBoost gibi topluluk modelleri, yksek performans gsterirken, model zellik nemleri aracılıđıyla belirli bir dzeyde yorumlanabilirlik de sađlamaktadır.

**Veri Kalitesi ve Miktarı:** Eğitim verisi analizinde, veri kalitesi ve miktarı, model performansını dođrudan etkilemektedir. Bu arařtırmada, greceli olarak kk bir veri seti ( $n=151$ ) kullanılmasına rađmen, yksek kaliteli veriler (z-bildirime dayalı ancak stn yetenekli đrenciler tarafından doldurulmuř) sayesinde, makine đrenmesi modelleri yksek tahmin performansı gstermiřtir. Ancak, daha byk ve eřitli veri setleriyle alıřarak, daha gvenilir ve genellenebilir sonular elde edilebilir.

**apraz Dođrulama nemi:** Eğitim verisi analizinde, zellikle kk veri setlerinde, apraz dođrulama teknikleri, model performansının gvenilir bir řekilde deđerlendirilmesi iin kritik neme sahiptir. Bu arařtırmada, 10-katlı apraz dođrulama kullanılarak, modellerin performansı, veri setinin farklı blmlerinde deđerlendirilmiř ve daha gvenilir sonular elde edilmiřtir.

#### **4.3.2. Model Seimi ve Hiperparametre Optimizasyonu Stratejileri**

Eđitim verisi analizinde uygun modelin seilmesi ve hiperparametrelerin optimize edilmesi, model performansını nemli lde etkilemektedir. Bu arařtırmada, model seimi ve hiperparametre optimizasyonu iin eřitli stratejiler kullanılmıř ve bu stratejilerin etkinliđi deđerlendirilmiřtir:

**Grid Search vs. Random Search:** Hiperparametre optimizasyonu iin, hem grid search (ızgara araması) hem de random search (rastgele arama) yntemleri denenmiřtir. Random search, daha az hesaplama kaynađı kullanarak, hiperparametre uzayını daha verimli bir řekilde arařtırabilmektedir. zellikle, hiperparametre uzayının byk olduđu durumlarda, random search tercih edilebilir. Ancak, grid search, belirli bir hiperparametre aralıđını sistematik olarak arařtırmak iin daha uygundur.

**Bayesian Optimizasyon:** Bayesian optimizasyon, nceki denemelerin sonularına dayalı olarak, en umut verici hiperparametre deđerlerini belirleyen bir

yaklaşımıdır. Bu araştırmada, XGBoost ve derin öğrenme modelleri için Bayesian optimizasyon kullanılmış ve bu yöntemin, grid search ve random search'e göre daha verimli olduğu görülmüştür. Özellikle, hesaplama kaynakları sınırlı olduğunda ve hiperparametre uzayı büyük olduğunda, Bayesian optimizasyon tercih edilebilir.

**Çapraz Doğrulama ile Model Seçimi:** Model seçimi için, çapraz doğrulama performansı temel alınmıştır. Her bir modelin çapraz doğrulama performansı (doğruluk, F1 skoru, ROC AUC, RMSE,  $R^2$  vb.) hesaplanmış ve en yüksek performansı gösteren model seçilmiştir. Ancak, model seçiminde, sadece performans değil, yorumlanabilirlik, hesaplama maliyeti ve uygulama kolaylığı gibi faktörler de dikkate alınmalıdır.

**Erken Durdurma (Early Stopping):** Özellikle derin öğrenme modelleri için, aşırı uyumu önlemek amacıyla erken durdurma tekniği kullanılmıştır. Bu teknik, doğrulama seti üzerindeki performans artmaya devam ettiği sürece eğitime devam eder, performans düşmeye başladığında ise eğitimi durdurur. Erken durdurma, hem aşırı uyumu önlemek hem de eğitim süresini kısaltmak için etkili bir yöntemdir.

**Ensemble Selection:** Bu araştırmada, farklı model türlerinin (temel modeller, topluluk modelleri, derin öğrenme modelleri) tahminlerini birleştirerek, daha güçlü tahminler yapan bir ensemble selection yaklaşımı da denenmiştir. Bu yaklaşım, farklı model türlerinin güçlü yönlerinden yararlanarak, daha yüksek ve kararlı tahmin performansı sağlayabilir. Ancak, ensemble selection, hesaplama açısından maliyetli olabilir ve yorumlanabilirliği azaltabilir.

**Farklı Model Türlerinin Karşılaştırması:** Bu araştırmada, beş temel model (Lojistik Regresyon, Karar Ağacı, SVM, KNN, Naive Bayes), dört topluluk modeli (Rastgele Orman, Gradyan Artırma, AdaBoost, XGBoost) ve dört derin öğrenme modeli (3 katmanlı MLP, 5 katmanlı MLP, RNN, CNN) karşılaştırılmıştır. Bu kapsamlı karşılaştırma, eğitim verisi analizinde kullanılabilecek çeşitli modeller hakkında değerli bilgiler sağlamaktadır.

#### 4.4. ÇALIŞMANIN SINIRLILIKLARI

Bu araştırmanın, metodolojik ve teorik açıdan bazı sınırlılıkları bulunmaktadır. Bu sınırlılıklar, sonuçların yorumlanması ve genellenmesi konusunda dikkatli olunmasını gerektirmektedir:

**Örneklem Büyüklüğü ve Çeşitliliği:** Araştırmanın örneklemini, Bursadaki BİLSEM'lerde eğitim gören 151 üstün yetenekli öğrenciden oluşmaktadır. Bu örneklem büyüklüğü, makine öğrenmesi uygulamaları için nispeten küçüktür ve bu durum, özellikle derin öğrenme modellerinin performansını sınırlandırabilir. Ayrıca, örneklemin sadece bir ildeki BİLSEM öğrencilerinden oluşması, sonuçların genellenebilirliğini sınırlamaktadır. Farklı bölgelerdeki, farklı sosyo-ekonomik düzeylerdeki ve farklı eğitim programlarındaki üstün yetenekli öğrencilerle çalışmak, daha genellenebilir sonuçlar elde etmek için faydalı olabilir.

**Veri Toplama Yöntemi:** Araştırmada kullanılan veriler, öz-bildirime dayalı anket yoluyla toplanmıştır. Öz-bildirim yöntemi, katılımcıların öznel değerlendirmelerine dayanmakta olup, sosyal istenirlik yanlılığı, hafıza yanlılığı ve öz-algı yanlılığı gibi çeşitli yanlılıklara açıktır. Özellikle, matematik başarısının öğrencilerin bildirdiği not ortalamalarına dayalı olması, ölçümün güvenilirliğini etkileyebilir. Daha nesnel ölçümler (standardize edilmiş testler, performans değerlendirmeleri vb.) kullanmak, daha güvenilir sonuçlar elde etmek için faydalı olabilir.

**Kesitsel Araştırma Deseni:** Bu araştırma, kesitsel bir tasarıma sahiptir ve değişkenler arasındaki ilişkileri belirli bir zaman noktasında incelemektedir. Bu nedenle, değişkenler arasındaki nedensel ilişkiler hakkında kesin sonuçlar çıkarmak mümkün değildir. Örneğin, matematik ilgi düzeyi ile matematik başarısı arasındaki pozitif ilişki, ilginin başarıyı etkilediğini gösterebileceği gibi, başarının ilgiyi etkilediğini de gösterebilir. Uzunlamasına araştırma desenleri, değişkenler arasındaki nedensel ilişkileri daha iyi anlamak için faydalı olabilir.

**Değişkenlerin Ölçümü:** Bazı değişkenlerin ölçümünde kullanılan ölçekler, tek maddelik veya az maddelik olup, ilgili yapıları kapsamlı bir şekilde ölçmeyebilir. Örneğin, matematik ilgi düzeyi, tek bir madde ile (Matematik dersine olan ilginizi nasıl değerlendirirsiniz?) ölçülmüştür. Daha kapsamlı ve psikometrik olarak güçlü ölçekler kullanmak, daha güvenilir ve geçerli sonuçlar elde etmek için faydalı olabilir.

**Dışsal Faktörlerin Kontrolü:** Araştırmada, öğrencilerin matematik başarısını etkileyebilecek bazı dışsal faktörler (okul türü, öğretmen nitelikleri, akran etkisi vb.) kontrol edilmemiş veya ölçülmemiştir. Bu faktörlerin etkisini dikkate almak, daha kapsamlı bir anlayış geliştirmek için faydalı olabilir.

**Model Karşılaştırma Yöntemleri:** Araştırmada, farklı model türlerinin performansları karşılaştırılırken, modellerin hiperparametreleri ayrı ayrı optimize edilmiştir. Ancak, hiperparametre optimizasyonu için kullanılan hesaplama kaynakları ve zaman, modeller arasında eşit olmayabilir. Bu durum, model karşılaştırmalarının adillliğini etkileyebilir. Daha adil karşılaştırmalar için, modellere eşit miktarda hesaplama kaynağı ve zaman ayırmak faydalı olabilir.

**Matematiğin Alt Alanlarının Dikkate Alınmaması:** Matematik, cebir, geometri, analiz, olasılık ve istatistik gibi çeşitli alt alanlara sahiptir. Bu araştırmada, matematik başarısı genel olarak ele alınmış, alt alanlar ayrı ayrı incelenmemiştir. Matematiğin farklı alt alanlarındaki başarıyı etkileyen faktörler farklılık gösterebilir. Bu alt alanları ayrı ayrı incelemek, daha ayrıntılı ve hedefli sonuçlar elde etmek için faydalı olabilir.

Bu sınırlılıklar, araştırma sonuçlarının yorumlanması ve uygulanması konusunda dikkatli olunmasını gerektirmektedir. Gelecekteki araştırmalar, bu sınırlılıkları göz önünde bulundurarak, daha güçlü ve genellenebilir sonuçlar elde etmek için çaba göstermelidir.

#### **4.5. GELECEK ÇALIŞMALAR İÇİN ÖNERİLER**

Bu araştırmanın bulguları ve sınırlılıkları ışığında, gelecekteki araştırmalar için çeşitli öneriler sunulabilir:

**Daha Büyük ve Çeşitli Örneklemeler:** Gelecekteki araştırmalar, daha büyük ve çeşitli örneklemelerle çalışarak, sonuçların genellenebilirliğini artırabilir. Farklı bölgelerdeki, farklı sosyo-ekonomik düzeylerdeki ve farklı eğitim programlarındaki üstün yetenekli öğrencilerle çalışmak, daha kapsamlı bir anlayış geliştirmek için faydalı olabilir.

**Uzunlamasına Araştırma Desenleri:** Üstün yetenekli öğrencilerin matematik başarısını etkileyen faktörleri daha iyi anlamak için, uzunlamasına araştırma desenleri kullanılabilir. Bu desenler, değişkenler arasındaki nedensel ilişkileri daha iyi anlamaya ve öğrencilerin gelişim yörüngelerini izlemeye olanak sağlar.

**Karma Yöntem Araştırmaları:** Nicel yöntemlerle (anketler, testler vb.) elde edilen verileri, nitel yöntemlerle (görüşmeler, gözlemler, vaka çalışmaları vb.) destekleyen karma yöntem araştırmaları, üstün yetenekli öğrencilerin matematik başarısını etkileyen faktörleri daha derinlemesine anlamak için faydalı olabilir.

**Daha Kapsamlı ve Güvenilir Ölçümler:** Gelecekteki araştırmalar, matematik başarısını ve bunu etkileyen faktörleri ölçmek için daha kapsamlı ve güvenilir ölçüm araçları kullanabilir. Özellikle, matematik başarDerin öğrenme modellerinin, özellikle MLP mimarisinin, üstün yetenekli öğrencilerin matematik başarısını tahmin etmede etkili olduğu görülmektedir. Ancak, derin öğrenme modellerinin, topluluk öğrenme modellerine göre belirgin bir üstünlüğü gözlenmemiştir. Bu durum, veri setinin göreceli olarak küçük olması ( $n=151$ ) ve derin öğrenme modellerinin genellikle büyük veri setlerinde daha iyi performans göstermesi ile açıklanabilir.

**Matematiğin Alt Alanlarına Odaklanma:** Gelecekteki araştırmalar, matematik başarısını, cebir, geometri, analiz, olasılık ve istatistik gibi alt alanlara ayırarak inceleyebilir. Bu yaklaşım, her bir alt alandaki başarıyı etkileyen faktörleri daha iyi anlamaya ve hedefli müdahaleler geliştirmeye olanak sağlar.

**Derin Öğrenme ve Yapay Zekâ Tekniklerinin Geliştirilmesi:** Gelecekteki araştırmalar, eğitim verisi analizinde derin öğrenme ve yapay zekâ tekniklerinin kullanımını geliştirebilir. Özellikle, büyük ve karmaşık veri setlerinde, bu teknikler, daha güçlü ve hassas tahmin modelleri geliştirmeye yardımcı olabilir.

**Müdahale Çalışmaları:** Araştırma bulguları temelinde, üstün yetenekli öğrencilerin matematik başarısını artırmaya yönelik müdahale programları geliştirilebilir ve bu programların etkinliği deneysel yöntemlerle test edilebilir. Örneğin, matematiğe ilgiyi artırmaya, olumlu problem çözme duygularını teşvik etmeye veya aile desteğini güçlendirmeye yönelik müdahaleler geliştirilebilir.

**Kültürler Arası Karşılaştırmalar:** Farklı kültürlerdeki ve eğitim sistemlerindeki üstün yetenekli öğrencilerin matematik başarısını etkileyen faktörleri karşılaştıran araştırmalar, kültürel ve sistemik faktörlerin etkisini anlamaya yardımcı olabilir.

**Teknoloji Destekli Matematiğin İncelenmesi:** Gelecekteki araştırmalar, teknoloji destekli matematik öğretiminin, üstün yetenekli öğrencilerin matematik başarısı üzerindeki etkisini inceleyebilir. Dinamik geometri yazılımları, bilgisayar cebir sistemleri, kodlama platformları ve diğer dijital araçların etkisini anlamak, teknoloji entegrasyonu konusunda rehberlik sağlayabilir.

**Disiplinler Arası Yaklaşımlar:** Matematik başarısını, STEM/STEAM eğitimi, proje tabanlı öğrenme ve disiplinler arası yaklaşımlar bağlamında inceleyen araştırmalar,

bu yaklaşımların üstün yetenekli öğrencilerin matematik başarıları üzerindeki etkisini anlamaya yardımcı olabilir.

Duyuşsal ve Sosyal Faktörlerin Derinlemesine İncelenmesi: Matematik kaygısı, matematiksel öz-yeterlik, matematiksel kimlik, akran ilişkileri ve sosyal destek gibi duyuşsal ve sosyal faktörlerin, üstün yetenekli öğrencilerin matematik başarıları üzerindeki etkisini derinlemesine inceleyen araştırmalar yapılabilir.

Kişiselleştirilmiş Öğrenme Modellerinin Geliştirilmesi: Makine öğrenmesi ve yapay zekâ tekniklerini kullanarak, üstün yetenekli öğrenciler için kişiselleştirilmiş matematik öğrenme modelleri ve platformları geliştiren araştırmalar, eğitim uygulamalarını dönüştürme potansiyeline sahiptir.

Bu öneriler, üstün yetenekli öğrencilerin matematik başarılarını etkileyen faktörleri daha iyi anlamak ve bu anlayışı eğitim uygulamalarını geliştirmek için kullanmak amacıyla gelecekteki araştırmalara rehberlik edebilir.

#### 4.6. SONUÇ

Bu araştırma, Bursadaki BİLSEM'lerde eğitim gören üstün yetenekli öğrencilerin matematik başarılarını etkileyen faktörleri, çeşitli makine öğrenmesi modelleri kullanarak incelemiştir. Araştırmanın temel bulguları şu şekilde özetlenebilir:

Üstün yetenekli öğrencilerin matematik başarılarını etkileyen en önemli faktörler, motivasyonel (matematik ilgi düzeyi, problem çözme duygusu), davranışsal (matematik çalışma süresi), sosyal (aile desteği) ve bilişsel (öğrenme stili) faktörlerdir. Bu faktörler, tüm makine öğrenmesi modellerinde tutarlı bir şekilde önemli olarak belirlenmiştir.

Topluluk öğrenme modelleri (özellikle XGBoost) ve derin öğrenme modelleri (özellikle 5 katmanlı MLP), üstün yetenekli öğrencilerin matematik başarılarını tahmin etmede, temel modellerden daha yüksek performans göstermiştir. XGBoost modeli, hem sınıflandırma (doğruluk = %92.2, ROC AUC = 0.901) hem de regresyon (RMSE = 0.413,  $R^2 = 0.747$ ) problemlerinde en yüksek performansı göstermiştir.

Model karmaşıklığı ve özellik sayısı, belirli bir noktaya kadar model performansını artırmakta, ancak bu noktadan sonra performansta belirgin bir artış gözlenmemekte veya performans düşmektedir. Bu durum, model seçiminde ve özellik mühendisliğinde, hem performans hem de karmaşıklık dikkate alınması gerektiğini göstermektedir.

Bu bulgular ışığında, üstün yetenekli öğrencilerin matematik eğitimi için çeşitli öneriler sunulmuştur. BİLSEM matematik programları, öğrencilerin ilgisini çekecek ve motivasyonlarını artıracak şekilde tasarlanmalı, öğrencilerin farklı öğrenme stillerine hitap eden bireyselleştirilmiş eğitim programları geliştirilmeli, öğrencilerin matematikle ilgili olumlu duygular geliştirmelerine yardımcı olacak etkinlikler düzenlenmeli ve ailelerin çocuklarının eğitimine aktif katılımı teşvik edilmelidir.

Ayrıca, üstün yetenekli öğrenci eğitiminde veri odaklı yaklaşımların potansiyel uygulamaları tartışılmış ve ebeveynlere, çocuklarının matematiksel gelişimlerini desteklemeleri için öneriler sunulmuştur. Araştırmanın metodolojik ve teorik sınırlılıkları göz önünde bulundurularak, gelecekteki araştırmalar için çeşitli öneriler geliştirilmiştir.

Bu araştırma, üstün yetenekli öğrencilerin matematik başarısını etkileyen faktörleri anlamak ve bu faktörlerin matematik başarısını ne düzeyde tahmin edebildiğini belirlemek amacıyla, makine öğrenmesi tekniklerini eğitim verisi analizine uygulayan özgün bir çalışmadır. Elde edilen bulgular, üstün yetenekli öğrencilerin matematik eğitimini geliştirmek ve potansiyellerini en üst düzeyde gerçekleştirmelerine yardımcı olmak için değerli içgörüler sunmaktadır.



**EKLER**

**Ek-1: Anket Formu**

1. Yaşınız \*

- ☐ 10
- ☐ 11
- ☐ 12
- ☐ 13
- ☐ 14

2. Cinsiyetiniz \*

- ☐ Kız
- ☐ Erkek

3. Sınıfınız \*

- ☐ 5. Sınıf
- ☐ 6. Sınıf
- ☐ 7. Sınıf
- ☐ 8. Sınıf

4. Matematik dersindeki genel not ortalamanız nedir? \*

- ☐ 90-100
- ☐ 80-89
- ☐ 70-79
- ☐ 60-69
- ☐ 50-59
- ☐ 49 ve altı

5. Matematik konularını öğrenirken hangi kaynakları kullanıyorsunuz? (Birden fazla seçebilirsiniz) \*

- ☐ Ders kitapları
- ☐ Çevrimiçi platformlar
- ☐ Özel ders veya etüt
- ☐ Arkadaşlar ve aile
- ☐ Diğer

6. Haftada kaç saat matematik çalışırsınız? \*

- ☐ 1 saatten az
- ☐ 1-3 saat
- ☐ 4-6 saat
- ☐ 7 saat ve üzeri

7. Matematik dersine olan ilginizi nasıl değerlendirirsiniz?

- ☐ Çok yüksek
- ☐ Yüksek
- ☐ Orta
- ☐ Düşük
- ☐ Çok düşük

8. Matematik problemleri çözerken hissettiğiniz duyguyu en iyi hangi ifade tarif eder?

- ☐ Heyecanlı ve meraklı
- ☐ Rahat ve kendinden emin
- ☐ Endişeli ve stresli
- ☐ Sıkılmış ve ilgisiz
- ☐ Üzgün ve yılgın

9.Kendinize ait odanız var mı?

- ☐ Evet
- ☐ Hayır

10. Matematikle ilgili hangi yarışma veya aktivitelere katıldınız? (Birden fazla seçebilirsiniz) \*

- ☐ Matematik Olimpiyatları
- ☐ Bilim ve Matematik Fuarları
- ☐ Okul içi matematik yarışmaları
- ☐ Çevrimiçi matematik yarışmaları
- ☐ Hiçbirine katılmadım

11.Bu yarışmalardan herhangi birinde ödül kazandınız mı? \*

- ☐ Evet
- ☐ Hayır

12. Önceki soruda belirttiğiniz matematik yarışmalarında kazandığınız derece veya başarı puanınız var mı? Varsa **yarışma adını** ve derece/başarı puanınızı yazınız. (Yoksa aşağıdaki boş yere 'yok' ifadesi yazabilirsiniz.) \*

Yanıtınız

13. Ailede kaç kişi yaşıyorsunuz? \*

- ☐ 2
- ☐ 3-4
- ☐ 5-6
- ☐ 7 ve daha fazla

14. Evde matematik veya matematikle ilgili alanlarda çalışan bir yetişkin var mı? \*

- ☐ Evet
- ☐ Hayır

15. Ailenizin matematik ödevlerinizde size yardımcı olma sıklığı nasıldır? \*

- ☐ Her zaman
- ☐ Sık sık
- ☐ Bazen
- ☐ Nadiren
- ☐ Asla

16. Aileniz sizi matematikle ilgili etkinliklere katılmaya teşvik eder mi? \*

- ☐ Her zaman
- ☐ Sık sık
- ☐ Bazen
- ☐ Nadiren
- ☐ Asla

17. Matematikle ilgili çalışmalarınızda ailenizden sizi destekleyen var mı? \*

- ☐ Evet, tüm aile
- ☐ Evet, bazı aile bireyleri
- ☐ Hayır

18. Matematik öğrenirken en etkili bulduğunuz yöntem nedir? \*

- ☐ Görsel (videolar, grafikler)
- ☐ İşitsel (anlatım, tartışma)
- ☐ Okuma/Yazma (notlar almak, kitap okumak)
- ☐ Kinestetik (elle yaparak, deneyerek öğrenme)

19. Matematik öğrenirken karşılaştığınız en büyük zorluk nedir? \*

- ☐ Konuların karmaşıklığı
- ☐ Yeterli kaynak bulamamak
- ☐ Motivasyon eksikliği
- ☐ Zaman yönetimi
- ☐ Öğretim tarzı

20. Bu zorlukların üstesinden gelmek için hangi stratejileri kullanıyorsunuz? \*

- ☐ Ek dersler
- ☐ Grup çalışması
- ☐ Düzenli tekrar
- ☐ Zamanı iyi yönetmek
- ☐ Diğer

21. Matematik öğreniminde teknolojiyi ne sıklıkla kullanırsınız? \*

- ☐ Her zaman
- ☐ Çoğunlukla
- ☐ Bazen
- ☐ Nadiren
- ☐ Asla

23. Akademik hedeflerinizle ilgili kendinize koyduđunuz beklentiler nasıl? \*

- ☐ Çok yüksek
- ☐ Yüksek
- ☐ Orta
- ☐ Düşük
- ☐ Çok düşük

24. Matematik dışında hangi alanlarda yetenekli olduđunuzu düşünöyorsunuz veya ilgi duyuyorsunuz? \*

- ☐ Fen Bilimleri (Fizik, Kimya, Biyoloji vb.)
- ☐ Sosyal Bilimler (Psikoloji, Sosyoloji, Tarih, Coğrafya vb.)
- ☐ Sanatlar (Müzik, Resim, Dans, Tiyatro vb.)
- ☐ Teknoloji ve Bilgisayar Bilimleri (Yazılım, Donanım, Web Tasarımı vb.)
- ☐ Girişimcilik ve İş Dünyası
- ☐ Spor (Yüzme, Bale, Basketbol, Taekwondo, Bisiklet vb.)
- ☐ Diğer

25. Günlük olarak kaç saat oyun oynarsınız? (Bilgisayar, tablet, telefon vb.) \*

- ☐ 1 saatten az
- ☐ 1-2 saat
- ☐ 3-4 saat
- ☐ 5 saat ve üzeri
- ☐ Oyun oynamam

26. Dışarıda oyun oynamak için kaç saat ayırıyorsunuz? (Sokak, bahçe vb. ) \*

- ☐ 1 saatten az
- ☐ 1-2 saat
- ☐ 3-4 saat
- ☐ 5 saat ve üzeri
- ☐ Oyun oynamam

27. Ailenizin aylık gelir düzeyi nedir? (Tahmini olarak) \*

- ☐ Çok düşük (15.000 TL ve altı)
- ☐ Düşük (15.001 - 25.000 TL arası)
- ☐ Orta - düşük (25.001 - 35.000 TL arası)
- ☐ Orta (35.001 ve 50.000 TL arası)
- ☐ Üst - orta (50.001 ve 70.000 TL arası)
- ☐ Yüksek (70.001 TL ve üzeri)

28. Günde ortalama kaç saat uyursunuz? \*

- ☐ 5 saat ve daha az
- ☐ 6-8 saat
- ☐ 9-10 saat
- ☐ 10 saat ve üzeri

29. Kiminle birlikte mi yaşıyorsunuz? \*

- ☐ Evet, annemle
- ☐ Evet, babamla
- ☐ Evet, her ikisiyle de
- ☐ Hayır, yalnız yaşıyorum
- ☐ Başka biriyle yaşıyorum

30. Anne ve babanızın sağlık durumu nasıl? \*

- ☐ Evet, her ikisi de hayatta
- ☐ Anne sağ, baba sağ değil
- ☐ Anne sağ değil, baba sağ
- ☐ Her ikisi de sağ değil

31. Sınav veya ödev stresiyle başa çıkmak için hangi yöntemleri kullanıyorsunuz? \*

- ☐ Derin nefes alma ve meditasyon
- ☐ Spor veya fiziksel aktiviteler
- ☐ Arkadaşlarla konuşmak veya destek almak
- ☐ Hobi veya ilgi alanlarına yönelmek
- ☐ Rehberlik öğretmeni veya Psikolojik danışman ile görüşme
- ☐ Diğer

32. İnternet üzerinde araştırma yaparken kaynakların güvenilirliğini nasıl değerlendiriyorsunuz? \*

- ☐ Akademik kaynaklara güvenmek
- ☐ Çeşitli kaynakları karşılaştırmak
- ☐ Uzman görüşlerini araştırmak
- ☐ İnternetten bulduğum her şeye güveniyorum

33. Sürekli olarak uğraştığınız sosyal aktivitenin türü nedir? \*

- ☐ Spor
- ☐ Müzik
- ☐ Sanat (Resim, heykelticilik, vb.)
- ☐ Video Oyunları
- ☐ Kitap Okuma
- ☐ Sinema, tiyatro vb
- ☐ Diğer

34. Günlük olarak kaç saat sürekli uğraşılan sosyal aktivitelerle meşgul olursunuz? \*

- ☐ 1 saatten az
- ☐ 1-2 saat
- ☐ 3-4 saat
- ☐ 5 saat ve üzeri
- ☐ Sürekli uğraşılan sosyal aktivitelerle meşgul olmam

35. Günlük olarak kaç saat televizyon izlersiniz? \*

- ☐ Hiç izlemem
- ☐ 1 saatten az
- ☐ 1-2 saat
- ☐ 3-4 saat
- ☐ 5 saat ve üzeri

36. Genel olarak hayatınızdan ne kadar memnunsunuz? \*

- ☐ Çok memnunum
- ☐ Memnunum
- ☐ Nötrüm
- ☐ Memnun değilim
- ☐ Hiç memnun değilim

37. Annenizin en son mezun olduğu eğitim seviyesi nedir? \*

- ☐ Okuma-yazma bilmiyor / İlkokul terk
- ☐ İlkokul
- ☐ Ortaokul
- ☐ Lise
- ☐ Ön lisans / Meslek Yüksekokulu
- ☐ Lisans (Üniversite)
- ☐ Lisansüstü (Yüksek Lisans/Doktora)

38. Babanızın en son mezun olduğu eğitim seviyesi nedir? \*

- ☐ Okuma-yazma bilmiyor / İlkokul terk
- ☐ İlkokul
- ☐ Ortaokul
- ☐ Lise
- ☐ Ön lisans / Meslek Yüksekokulu
- ☐ Lisans (Üniversite)
- ☐ Lisansüstü (Yüksek Lisans/Doktora)

## Ek-2: Etik Kurul İzni

Evrak Tarih ve Sayısı: 04.06.2024-278



**KÜTAHYA DÜMLUPINAR ÜNİVERSİTESİ**  
Sosyal ve Beşeri Bilimler Bilimsel Araştırma ve Yayın Etiği  
Kurulu

### TOPLANTI TUTANAĞI

**Toplantı No** : 2024/05  
**Toplantı Tarihi** : 27.05.2024  
**Toplantı Saati** : 14:00  
**Toplantı Yeri** : Spor Bilimleri Fakültesi Toplantı Salonu

### GÜNDEM

47: Kütahya Dumlupınar Üniversitesi Lisansüstü Eğitim Enstitüsü Müdürlüğü'nün 21.05.2024 tarih ve 293455 sayılı yazısı gereğince; Bilgisayar Mühendisliği Ana Bilim Dalı Tezli Yüksek Lisans Programı öğrencisi Kübra ÖNGENLİ'nin "*Yapay Zeka Destekli Eğitim Uygulamalarının Üstün Yetenekli Çocukların Matematik Öğrenme Yaklaşımlarına Etkileri*" başlıklı akademik çalışmasında kullanılmak üzere, uygulama, görüşme ve anket yapma talebinin etik açıdan uygunluğu üzerine görüşme

### KARAR

47: Kütahya Dumlupınar Üniversitesi Lisansüstü Eğitim Enstitüsü Müdürlüğü'nün 21.05.2024 tarih ve 293455 sayılı yazısı gereğince; Bilgisayar Mühendisliği Ana Bilim Dalı Tezli Yüksek Lisans Programı öğrencisi Kübra ÖNGENLİ'nin "*Yapay Zeka Destekli Eğitim Uygulamalarının Üstün Yetenekli Çocukların Matematik Öğrenme Yaklaşımlarına Etkileri*" başlıklı akademik çalışmasında kullanılmak üzere, uygulama, görüşme ve anket yapma talebinin etik açıdan uygunluğu üzerine görüşüldü.

*Yapılan görüşmeler ve değerlendirmeler sonucunda, çalışma kapsamında yapılacak olan uygulama, anket-görüşme sorularının ve ölçeklerin, gerekli izinlerin alınması kaydıyla, fikri, hukuki ve telif hakları bakımından sorumluluğu başvurucuya ait olmak üzere etik açıdan uygun olduğuna oy birliği ile karar verildi.*

Prof. Dr. Çetin Özdilek  
Başkan

Prof. Dr. Metin DEMİR  
Başkan Yardımcısı

Prof. Dr. Emin ZEYTİNOĞLU  
Üye

Prof. Dr. Emre SEZİCİ  
Üye

Prof. Dr. Fulya TOPÇUOĞLU  
ÜNAL  
Üye

Prof. Dr. Levent MERCİN  
Üye

Prof. Dr. Niyazi KURNAZ  
Üye

Adres: Evliya Çelebi Yürürlüğü Tazıyanlı Yolu 10. Km 43100 KÜTAHYA  
Telefon: (0 274) 443 43 43 Faks: 2742652014  
e-Posta: etikkurul@dpu.edu.tr Web: http://www.dpu.edu.tr  
Kep Adresi: dumlupinar.universitesi@hs01.kep.tr

Bilgi için: Emir Özekmekçi  
Unvanı: Raportör



Bu belge, Güvenli Elektronik İmza ile imzalanmıştır.

## Ek-2: MEB Araştırma İzni

← → ↺ arastirmaizinleri.meb.gov.tr/panel/arastirma-uygulamalari/basvurularim



**T.C. MİLLİ EĞİTİM  
BAKANLIĞI**

**KÜBRA ÖNGENLİ**

Yeni Başvuru

**Başvurularım**

Görüş ve Öneri Bildir

Dikkat Edilecek Hususlar

Araştırma Uygulama İzinleri Başvuru ve Değerlendirme Sistemi > Başvurularım

### BAŞVURULARIM

Sayfada 10 kayıt göster

| BAŞVURU NO         | BAŞVURU TARİHİ | NİTELİĞİ           | BAŞVURU DURUMU |
|--------------------|----------------|--------------------|----------------|
| MEB.TT.2024.000489 | 06.08.2024     | Yüksek Lisans Tezi | Onaylandı      |
| MEB.TT.2025.018023 | 01.02.2025     | Araştırma          | Onaylandı      |

# KAYNAKÇA

- Amato, P. R. (2001). Children of divorce in the 1990s: An update of the Amato and Keith (1991) meta-analysis. *Journal of Family Psychology*, 15(3), 355-370. <https://doi.org/10.1037/0893-3200.15.3.355>
- Adejo, O., & Connolly, T. (2018). Predicting student academic performance using multi-model heterogeneous ensemble approach. *Journal of Applied Research in Higher Education*, 10(1), 61-75. <https://doi.org/10.1108/JARHE-09-2017-0113>
- Alkan, A., & Mertol, H. (2019). A study on the relationship between the mathematical achievement of gifted students and their parents' education level, occupation and financial status. *Acta Educationis Generalis*, 9(3), 74-88. <https://doi.org/10.2478/atd-2019-0015>
- Asif, R., Merceron, A., Ali, S. A., & Haider, N. G. (2017). Analyzing undergraduate students' performance using educational data mining. *Computers & Education*, 113, 177-194. <https://doi.org/10.1016/j.compedu.2017.05.007>
- Baker, R. S., & Inventado, P. S. (2014). Educational data mining and learning analytics. In J. A. Larusson & B. White (Eds.), *Learning analytics: From research to practice* (pp. 61-75). Springer. [https://doi.org/10.1007/978-1-4614-3305-7\\_4](https://doi.org/10.1007/978-1-4614-3305-7_4)
- Baker, R. S., & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. *Journal of Educational Data Mining*, 1(1), 3-17.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman.
- Boaler, J. (2002). Experiencing school mathematics: Traditional and reform approaches to teaching and their impact on student learning. Lawrence Erlbaum Associates.
- Bourdieu, P. (1986). The forms of capital. In J. Richardson (Ed.), *Handbook of theory and research for the sociology of education* (pp. 241-258). Greenwood.
- Breiman, L. (1996). Bagging predictors. *Machine Learning*, 24(2), 123-140. <https://doi.org/10.1023/A:1018054314350>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32. <https://doi.org/10.1023/A:1010933404324>
- Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees*. CRC Press.

- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785-794). ACM. <https://doi.org/10.1145/2939672.2939785>
- Chiu, M. M., & Xihua, Z. (2008). Family and motivation effects on mathematics achievement: Analyses of students in 41 countries. *Learning and Instruction*, 18(4), 321-336. <https://doi.org/10.1016/j.learninstruc.2007.06.003>
- Corno, L., & Mandinach, E. B. (2004). What we have learned about student engagement in the past twenty years. In D. M. McInerney & S. Van Etten (Eds.), *Big theories revisited* (pp. 299-328). Information Age Publishing.
- Cortez, P., & Silva, A. M. G. (2008). Using data mining to predict secondary school student performance. In A. Brito & J. Teixeira (Eds.), *Proceedings of 5th Annual Future Business Technology Conference* (pp. 5-12). EUROSIS.
- Costa, E. B., Fonseca, B., Santana, M. A., de Araújo, F. F., & Rego, J. (2017). Evaluating the effectiveness of educational data mining techniques for early prediction of students' academic failure in introductory programming courses. *Computers in Human Behavior*, 73, 247-256. <https://doi.org/10.1016/j.chb.2017.01.047>
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273-297. <https://doi.org/10.1007/BF00994018>
- Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1), 21-27. <https://doi.org/10.1109/TIT.1967.1053964>
- Davis-Kean, P. E. (2005). The influence of parent education and family income on child achievement: The indirect role of parental expectations and the home environment. *Journal of Family Psychology*, 19(2), 294-304. <https://doi.org/10.1037/0893-3200.19.2.294>
- Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence*, 35(1), 13-21. <https://doi.org/10.1016/j.intell.2006.02.001>
- Dietterich, T. G. (2000). Ensemble methods in machine learning. In *International Workshop on Multiple Classifier Systems* (pp. 1-15). Springer. [https://doi.org/10.1007/3-540-45014-9\\_1](https://doi.org/10.1007/3-540-45014-9_1)

- Drucker, H., Burges, C. J., Kaufman, L., Smola, A. J., & Vapnik, V. (1997). Support vector regression machines. In *Advances in Neural Information Processing Systems* (pp. 155-161).
- Dunn, R., Honigsfeld, A., Doolan, L. S., Bostrom, L., Russo, K., Schiering, M. S., Suh, B., & Tenedero, H. (2009). Impact of learning-style instructional strategies on students' achievement and attitudes: Perceptions of educators in diverse institutions. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 82(3), 135-140. <https://doi.org/10.3200/TCHS.82.3.135-140>
- Eccles, J. S. (2007). Families, schools, and developing achievement-related motivations and engagement. In J. E. Grusec & P. D. Hastings (Eds.), *Handbook of socialization: Theory and research* (pp. 665-691). Guilford Press.
- Fan, X., & Chen, M. (2001). Parental involvement and students' academic achievement: A meta-analysis. *Educational Psychology Review*, 13(1), 1-22. <https://doi.org/10.1023/A:1009048817385>
- Feldhusen, J. F. (2005). Giftedness, talent, expertise, and creative achievement. In R. J. Sternberg & J. E. Davidson (Eds.), *Conceptions of giftedness* (2nd ed., pp. 64-79). Cambridge University Press. <https://doi.org/10.1017/CBO9780511610455.006>
- Freund, Y., & Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1), 119-139. <https://doi.org/10.1006/jcss.1997.1504>
- Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5), 1189-1232. <https://doi.org/10.1214/aos/1013203451>
- Gálvez, J., Guzmán, E., Conejo, R., & Millán, E. (2009). Student knowledge diagnosis using item response theory and constraint-based modeling. In V. Dimitrova, R. Mizoguchi, B. du Boulay, & A. Graesser (Eds.), *Artificial Intelligence in Education* (pp. 291-298). IOS Press.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
- Gökdere, M., & Küçük, M. (2003). Üstün yetenekli öğrencilerin fen eğitimindeki durum: Türkiye örnekleme. *Kuram ve Uygulamada Eğitim Bilimleri*, 3(1), 101-124.

- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer. <https://doi.org/10.1007/978-0-387-84858-7>
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge. <https://doi.org/10.4324/9780203887332>
- Haykin, S. (2009). *Neural networks and learning machines* (3rd ed.). Pearson Education.
- Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal*, 42(2), 371-406. <https://doi.org/10.3102/00028312042002371>
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735-1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression* (2nd ed.). John Wiley & Sons. <https://doi.org/10.1002/0471722146>
- Huang, S., & Fang, N. (2013). Predicting student academic performance in an engineering dynamics course: A comparison of four types of predictive mathematical models. *Computers & Education*, 61, 133-145. <https://doi.org/10.1016/j.compedu.2012.08.015>
- Karasar, N. (2012). *Bilimsel araştırma yöntemi* (24. baskı). Nobel Yayın Dağıtım.
- Khine, M. S. (Ed.). (2013). *Application of structural equation modeling in educational research and practice*. Sense Publishers. <https://doi.org/10.1007/978-94-6209-332-4>
- Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence* (pp. 1137-1143). Morgan Kaufmann.
- Kotsiantis, S. B., Pierrakeas, C. J., & Pintelas, P. E. (2004). Predicting students' performance in distance learning using machine learning techniques. *Applied Artificial Intelligence*, 18(5), 411-426. <https://doi.org/10.1080/08839510490442058>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444. <https://doi.org/10.1038/nature14539>

- Livieris, I. E., Drakopoulou, K., Tampakas, V. T., Mikropoulos, T. A., & Pintelas, P. (2019). Predicting secondary school students' performance utilizing a semi-supervised learning approach. *Journal of Educational Computing Research*, 57(2), 448-470. <https://doi.org/10.1177/0735633117752614>
- Ma, X., & Kishor, N. (1997). Assessing the relationship between attitude toward mathematics and achievement in mathematics: A meta-analysis. *Journal for Research in Mathematics Education*, 28(1), 26-47. <https://doi.org/10.2307/749662>
- Márquez-Vera, C., Cano, A., Romero, C., Noaman, A. Y. M., Mousa Fardoun, H., & Ventura, S. (2016). Early dropout prediction using data mining: A case study with high school students. *Expert Systems*, 33(1), 107-124. <https://doi.org/10.1111/exsy.12135>
- Masci, C., Johnes, G., & Agasisti, T. (2018). Student and school performance across countries: A machine learning approach. *European Journal of Operational Research*, 269(3), 1072-1085. <https://doi.org/10.1016/j.ejor.2018.02.031>
- Millî Eğitim Bakanlığı. (2016). *Bilim ve sanat merkezleri yönergesi*. Ankara: Millî Eğitim Bakanlığı.
- Millî Eğitim Bakanlığı. (2019). *Bilim ve sanat merkezleri öğretim programları*. Ankara: Millî Eğitim Bakanlığı.
- Melhuish, E. C., Phan, M. B., Sylva, K., Sammons, P., Siraj-Blatchford, I., & Taggart, B. (2008). Effects of the home learning environment and preschool center experience upon literacy and numeracy development in early primary school. *Journal of Social Issues*, 64(1), 95-114. <https://doi.org/10.1111/j.1540-4560.2008.00550.x>
- Mitchell, T. M. (1997). *Machine learning*. McGraw-Hill.
- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to linear regression analysis* (5th ed.). John Wiley & Sons.
- Morelock, M. J. (1996). On the nature of giftedness and talent: Imposing order on chaos. *Roeper Review*, 19(1), 4-12. <https://doi.org/10.1080/02783199609553774>

- Neihart, M. (2007). The socioaffective impact of acceleration and ability grouping: Recommendations for best practice. *Gifted Child Quarterly*, 51(4), 330-341. <https://doi.org/10.1177/0016986207306319>
- Organisation for Economic Co-Operation and Development. (2018). *PISA 2015 results in focus*. OECD Publishing.
- Organisation for Economic Co-Operation and Development. (2019). *PISA 2018 results (Volume I): What students know and can do*. OECD Publishing. <https://doi.org/10.1787/5f07c754-en>
- Oladokun, V. O., Adebajo, A. T., & Charles-Owaba, O. E. (2008). Predicting students' academic performance using artificial neural network: A case study of an engineering course. *The Pacific Journal of Science and Technology*, 9(1), 72-79.
- Olszewski-Kubilius, P. (2008). The role of the family in talent development. In S. I. Pfeiffer (Ed.), *Handbook of giftedness in children* (pp. 53-70). Springer. [https://doi.org/10.1007/978-0-387-74401-8\\_4](https://doi.org/10.1007/978-0-387-74401-8_4)
- Olszewski-Kubilius, P., & Corwith, S. (2018). Poverty, academic achievement, and giftedness: A literature review. *Gifted Child Quarterly*, 62(1), 37-55. <https://doi.org/10.1177/0016986217738015>
- Özenç, E. G., & Özenç, M. (2013). Türkiye'de üstün yetenekli öğrencilerle ilgili yapılan lisansüstü eğitim tezlerinin çok boyutlu olarak incelenmesi. *Türkiye Sosyal Araştırmalar Dergisi*, 17(1), 13-28.
- Pajares, F., & Miller, M. D. (1994). Role of self-efficacy and self-concept beliefs in mathematical problem solving: A path analysis. *Journal of Educational Psychology*, 86(2), 193-203. <https://doi.org/10.1037/0022-0663.86.2.193>
- Peña-Ayala, A. (2014). Educational data mining: A survey and a data mining-based analysis of recent works. *Expert Systems with Applications*, 41(4), 1432-1462. <https://doi.org/10.1016/j.eswa.2013.08.042>
- Periathiruvadi, S., & Rinn, A. N. (2012). Technology in gifted education: A review of best practices and empirical research. *Journal of Research on Technology in Education*, 45(2), 153-169. <https://doi.org/10.1080/15391523.2012.10782601>
- Phillips, N., & Lindsay, G. (2006). Motivation in gifted students. *High Ability Studies*, 17(1), 57-73. <https://doi.org/10.1080/13598130600947119>

- Piech, C., Huang, J., Nguyen, A., Phulsuksombati, M., Sahami, M., & Guibas, L. (2015). Learning program embeddings to propagate feedback on student code. In F. Bach & D. Blei (Eds.), *Proceedings of the 32nd International Conference on Machine Learning* (pp. 1093-1102). PMLR (Proceedings of Machine Learning Research).
- Preckel, F., Holling, H., & Vock, M. (2006). Academic underachievement: Relationship with cognitive motivation, achievement motivation, and conscientiousness. *Psychology in the Schools, 43*(3), 401-411. <https://doi.org/10.1002/pits.20154>
- Reis, S. M., & McCoach, D. B. (2000). The underachievement of gifted students: What do we know and where do we go? *Gifted Child Quarterly, 44*(3), 152-170. <https://doi.org/10.1177/001698620004400302>
- Renzulli, J. S. (2005). The three-ring conception of giftedness: A developmental model for promoting creative productivity. In R. J. Sternberg & J. E. Davidson (Eds.), *Conceptions of giftedness* (2nd ed., pp. 246-279). Cambridge University Press. <https://doi.org/10.1017/CBO9780511610455.015>
- Renzulli, J. S. (2012). Reexamining the role of gifted education and talent development for the 21st century: A four-part theoretical approach. *Gifted Child Quarterly, 56*(3), 150-159. <https://doi.org/10.1177/0016986212444901>
- Renzulli, J. S., & Reis, S. M. (2021). *The schoolwide enrichment model: A how-to guide for talent development* (3rd ed.). Routledge. <https://doi.org/10.4324/9781003238997>
- Rogers, K. B. (2007). Lessons learned about educating the gifted and talented: A synthesis of the research on educational practice. *Gifted Child Quarterly, 51*(4), 382-396. <https://doi.org/10.1177/0016986207306324>
- Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. *WIREs Data Mining and Knowledge Discovery, 10*(3), e1355. <https://doi.org/10.1002/widm.1355>
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68-78. <https://doi.org/10.1037/0003-066X.55.1.68>
- Sak, U. (2016). *Üstün zekalılar: Özellikleri, tanılanmaları, eğitimleri*. Vize Yayıncılık.

- Sak, U. (2021). *Üstün yetenekliler eğitim programları modeli ve sosyal geçerliği*. Pegem Akademi. <https://doi.org/10.14527/9786052412183>
- Sak, U., Ayas, M. B., Sezerel, B. B., Öpengin, E., Özdemir, N. N., & Gürbüz, S. D. (2015). Türkiye'de üstün yeteneklilerin eğitiminin eleştirel bir değerlendirmesi. *Türk Üstün Zekâ ve Eğitim Dergisi*, 5(2), 110-132.
- Sheffield, L. J. (2017). Dangerous myths about "gifted" mathematics students. *ZDM Mathematics Education*, 49(1), 13-23. <https://doi.org/10.1007/s11858-016-0814-8>
- Silverman, L. K. (2013). *Giftedness 101*. Springer Publishing Company.
- Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, 75(3), 417-453. <https://doi.org/10.3102/00346543075003417>
- Sriraman, B. (2008). Creativity, giftedness, and talent development in mathematics. Information Age Publishing.
- Stevenson, H. W., & Stigler, J. W. (1992). *The learning gap: Why our schools are failing and what we can learn from Japanese and Chinese education*. Simon & Schuster.
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267-288. <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>
- VanTassel-Baska, J. (2013). Performance-based assessment: The road to authentic learning for the gifted. *Gifted Child Today*, 36(1), 51-55. <https://doi.org/10.1177/1076217512465285>
- VanTassel-Baska, J., & Brown, E. F. (2007). Toward best practice: An analysis of the efficacy of curriculum models in gifted education. *Gifted Child Quarterly*, 51(4), 342-358. <https://doi.org/10.1177/0016986207306323>
- Wolpert, D. H. (1992). Stacked generalization. *Neural Networks*, 5(2), 241-259. [https://doi.org/10.1016/S0893-6080\(05\)80023-1](https://doi.org/10.1016/S0893-6080(05)80023-1)
- Worrell, F. C., Subotnik, R. F., Olszewski-Kubilius, P., & Dixon, D. D. (2019). Gifted students. *Annual Review of Psychology*, 70, 551-576. <https://doi.org/10.14>

- Worrell, F. C., Subotnik, R. F., Olszewski-Kubilius, P., & Dixson, D. D. (2019). Gifted students. *Annual Review of Psychology*, 70, 551-576. <https://doi.org/10.1146/annurev-psych-010418-102846>
- Xing, W., Chen, X., Stein, J., & Marcinkowski, M. (2016). Temporal predication of dropouts in MOOCs: Reaching the low hanging fruit through stacking generalization. *Computers in Human Behavior*, 58, 119-129. <https://doi.org/10.1016/j.chb.2015.12.007>
- Xu, B., & Recker, M. (2012). Teaching analytics: A clustering and triangulation study of digital library user data. *Educational Technology & Society*, 15(3), 103-115.
- Yang, D., Sinha, T., Adamson, D., & Rosé, C. P. (2013). Turn on, tune in, drop out: Anticipating student dropouts in massive open online courses. In *Proceedings of the 2013 NIPS Data-driven education workshop* (Vol. 11, p. 14).

## DİZİN

**-A-**

araştırma modeli, 26

**-B-**

BİLSEM, 1, 3, 5, 9, 10, 16, 26, 27, 28,  
29, 31, 32, 53, 54, 55, 80, 83, 84, 85,  
91, 94, 95

BYF, 9, 28, 53, 54

**-D-**

Denetimsiz öğrenme, 23

Derin Öğrenme, 22, 68, 70, 72, 88, 93

Derin Sinir Ağları, 2, 22, 47

Destek, 2, 9, 18, 20, 28, 30, 39, 44, 53,  
62, 63, 64, 65, 70, 71, 73, 74, 83, 87

Dijital, xi

dimensionality reduction, 23

Duyarlılık, 49

**-E-**

Eğitim, 3

Evrişimsel Sinir Ağları, 22

**-G-**

Gradyan Artırma, 2, 6, 21, 46, 62, 66, 67,  
68, 70, 72, 73, 74, 82, 88, 90

**-H-**

Hiperparametre, 6, 51, 89

**-K-**

Kütahya, 1, 2, 3

**-O-**

Otokodlayıcılar, 23, 41

overfitting, 20, 44, 46, 49, 75, 82, 88

Overfitting, 43, 47

**-Ö-**

ÖYG, 9, 28, 53, 54

**-P-**

Performans, v, xii

**-R-**

Reaksiyon, v, xi, xii

Regresyon Algoritmaları, 6, 19

**-S-**

Spor, xii

STEAM, 85, 94

Stroop Testi, xi, xii

**-U-**

underfitting, 49

Uyum Programı, 9

**-Ü-**

üstün yetenekli öğrencileri, 8

**-V-**

Veri, 6, 15, 16, 26, 29, 32, 33, 34, 35, 41,  
43, 48, 53, 85, 86, 89, 91

veri seti, 32, 37, 68, 88, 89

**-Y-**

yapay sinir ağları, 15, 22, 47, 68