T.C.
BANDIRMA ONYEDİ EYLÜL ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

BİLGİSAYAR MÜHENDİSLİĞİ MEZUNLARINA YÖNELİK MAKİNE ÖĞRENMESİ TABANLI KİŞİSELLEŞTİRİLMİŞ KARİYER ÖNERİ PLATFORMU (WEB-MÖS) TASARIMI VE GELİŞTİRİLMESİ

Tahir ULAŞ

Bilgisayar Mühendisliği Anabilim Dalı

Bilgisayar Mühendisliği Programı

EKİM 2025

T.C.
BANDIRMA ONYEDİ EYLÜL ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

BİLGİSAYAR MÜHENDİSLİĞİ MEZUNLARINA YÖNELİK MAKİNE ÖĞRENMESİ TABANLI KİŞİSELLEŞTİRİLMİŞ KARİYER ÖNERİ PLATFORMU (WEB-MÖS) TASARIMI VE GELİŞTİRİLMESİ

Tahir ULAŞ

DANIŞMAN
Doç. Dr. Necla BANDIRMALI ERTÜRK

Bilgisayar Mühendisliği Anabilim Dalı

Bilgisayar Mühendisliği Programı

EKİM 2025

ONAY

Tahir ULAŞ tarafından hazırlanan "**Bilgisayar Mühendisliği Mezunlarına Yönelik Makine Öğrenmesi Tabanlı Kişiselleştirilmiş Kariyer Öneri Platformu (WEB-MÖS) Tasarımı ve Geliştirilmesi**" adlı tez çalışması 22/09/2025 tarihinde yapılan sınavla aşağıdaki jüri tarafından oybirliği ile Bandırma Onyedi Eylül Üniversitesi Lisansüstü Eğitim Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı'nda YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

Jüri Üyeleri	**İmza**
Doç. Dr., Necla BANDIRMALI ERTÜRK (Danışman)	_____
Dr. Öğr. Üyesi, Bahar MİLANİ (Üye)	_____
Doç. Dr., Adnan SONDAŞ (Üye)	_____

Bu tezin kabulü, Lisansüstü Eğitim Enstitüsü Yönetim Kurulu'nun …/10/2025 gün ve …/… sayılı kararı ile onaylanmıştır.

Prof. Dr. Serhat DUMAN
Enstitü Müdürü

BANDIRMA ONYEDİ EYLÜL ÜNİVERSİTESİ LİSANÜSTÜ EĞİTİM ENSTİTÜSÜ
YÜKSEK LİSANS TEZİ
ETİK BEYANI

Bandırma Onyedi Eylül Üniversitesi Lisansüstü Eğitim Enstitüsü tez yazım kılavuzuna göre hazırlamış olduğum "**Bilgisayar Mühendisliği Mezunlarına Yönelik Makine Öğrenmesi Tabanlı Kişiselleştirilmiş Kariyer Öneri Platformu (WEB-MÖS) Tasarımı ve Geliştirilmesi**" adlı tezimin özgün bir çalışma olduğunu, tez hazırlanırken tüm aşamalarda bilimsel etik ilkelerine uygun davrandığımı, tez kapsamında sunulan tüm verileri bilimsel etik ilkelerine uygun elde ettiğimi, tezde faydalandığım tüm eserlere atıf yaptığımı ve kaynaklar kısmında bu eserleri gösterdiğimi beyan ederim. 22/09/2025

Tahir ULAŞ
İmza

ÖNSÖZ

Teknolojilerdeki ve yapay zekâ uygulamalarındaki hızlı gelişmeler, dünyamızı sürekli bir değişim içine sokmaktadır. Bu dönüşümün bir yansıması olarak, özellikle bilgisayar mühendisliği gibi çok sayıda uzmanlık alanını içeren ve hızla gelişen bir alanda, öğrencilerin mezuniyet sonrası kendilerine en uygun kariyer yolunu seçmeleri giderek zorlaşmaktadır. Bu karar verme süreci, öğrencilerin kişisel ilgi alanlarını, akademik başarılarını ve sektörün dinamiklerini doğru bir şekilde değerlendirmelerini gerektiren önemli bir eşiktir. Geleneksel rehberlik yöntemlerinin bu kişiselleştirilmiş ihtiyaca tam olarak cevap vermekte yetersiz kalabildiği durumlarda, makine öğrenmesi tabanlı akıllı sistemler yeni ve önemli bir potansiyel sunmaktadır.

Tez çalışmasında, makine öğrenmesi algoritmaları kullanılarak bilgisayar mühendisliği öğrencilerinin akademik başarılarına ve ilgi alanlarına dayalı kişiselleştirilmiş alt meslek alanı önerileri sunan web tabanlı bir sistem geliştirilmiştir. Bu sayede, öğrencilerin daha bilinçli ve verilere dayalı kariyer kararları almalarına destek olunması amaçlanmıştır.

Bu tez çalışmamın fikir aşamasından tamamlanmasına kadar geçen bütün süreçlerde, değerli bilgi ve tecrübeleriyle bana yol açan, destek ve yardımlarını benden hiçbir zaman esirgemeyen saygıdeğer danışman hocam Doç. Dr. Necla BANDIRMALI ERTÜRK'e en içten şükranlarımı sunarım.

Tez savunma sınavımda, kıymetli zamanlarını ayırarak yaptıkları değerli katkılar ve yapıcı eleştirileri için jüri üyelerim Dr. Öğr. Üyesi Bahar MİLANİ ve Doç. Dr. Adnan SONDAŞ hocalarıma teşekkür ederim.

Hayatımın her alanında yanımda olduğu gibi, tez çalışmamı hazırlarken de sabır ve anlayışla beni yüreklendiren, motive eden, maddi ve manevi desteğiyle bana güç veren kıymetli eşim Ayşegül ULAŞ'a ve evimizin neşe kaynağı, Rabbimin en kıymetli emaneti, güzel yürekli canım oğlum Mustafa Selim'e derin şükran ve muhabbetlerimi sunarım.

Ayrıca sonsuz sevgileri ve dualarıyla her zaman yanımda olan anne ve babama çok teşekkür ederim.

Tahir ULAŞ

Bandırma, 2025

ÖZET

YÜKSEK LİSANS TEZİ

BİLGİSAYAR MÜHENDİSLİĞİ MEZUNLARINA YÖNELİK MAKİNE ÖĞRENMESİ TABANLI KİŞİSELLEŞTİRİLMİŞ KARİYER ÖNERİ PLATFORMU (WEB-MÖS) TASARIMI VE GELİŞTİRİLMESİ

Tahir ULAŞ

Bandırma Onyedi Eylül Üniversitesi
Lisansüstü Eğitim Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı

Danışman: Doç. Dr. Necla BANDIRMALI ERTÜRK
Ekim 2025, 77 sayfa

Bu çalışmada, bilgisayar mühendisliği öğrencilerinin akademik profillerine ve teknolojik ilgi alanlarına göre kişiselleştirilmiş kariyer önerileri sunan web tabanlı bir akıllı sistem tasarlanmış ve geliştirilmiştir. Proje, iki farklı makine öğrenmesi modellemesi yaklaşımını karşılaştırmalı olarak analiz etmektedir. İlk aşamada (Model-1), tamamen kurallara dayalı üretilen yapay bir veri setiyle eğitilen modellerin, 50 kişilik gerçek öğrenci verisi üzerinde test edildiğinde genelleme başarımının sınırlı kaldığı ve en yüksek %52 (Lojistik Regresyon) doğruluk oranına ulaştığı tespit edilmiştir. Bu genelleme problemini aşmak amacıyla geliştirilen ikinci aşamada (Model-2) ise, 50 kişilik gerçek veri setini temel alan ve bu verileri sınıf-spesifik CTGAN (Conditional Tabular Generative Adversarial Network) tekniği ile zenginleştiren hibrit bir veri çoğaltma (data augmentation) stratejisi uygulanmıştır. Bu yöntemle oluşturulan sentetik ağırlıklı hibrit veri setiyle eğitilen Rastgele Orman (Random Forest), Gradyan Artırma (Gradient Boosting) ve Destek Vektör Makinesi (Support Vector Machine) gibi modeller, test setinde %97'nin üzerinde bir doğruluk oranına ulaşmıştır. Bu sonuç, az sayıda gerçek veriden yola çıkarak CTGAN ile yapılan veri çoğaltmanın, modelin öğrenme kapasitesini ve tahmin gücünü belirgin şekilde artırdığını doğrulamıştır. Sistemin canlıya alınması amacıyla Python Flask ile geliştirilen web

uygulaması (WEB-MÖS), anlık tahmin sunmanın yanı sıra, Google Sheets API entegrasyonu sayesinde kullanıcı geri bildirimlerini toplayarak modelin periyodik olarak yeniden eğitilmesine olanak tanıyan dinamik bir yapıya sahiptir. Bu bütüncül yaklaşım, bilgisayar mühendisliği eğitimi alanında öğrencilere veri odaklı rehberlik sunan etkili bir kariyer öneri sistemi prototipi sunmaktadır.

Anahtar kelimeler: Makine Öğrenmesi, Kariyer Öneri Sistemi, CTGAN, Veri Çoğaltma, Hibrit Veri Seti, Flask Web Uygulaması, Bilgisayar Mühendisliği Eğitimi.

ABSTRACT

M.Sc. THESIS

DESIGN AND DEVELOPMENT OF A MACHINE LEARNING BASED PERSONALIZED CAREER RECOMMENDATION PLATFORM (WEB-MOS) FOR COMPUTER ENGINEERING GRADUATES

Tahir ULAŞ

Bandirma Onyedi Eylul University
Institute of Graduate Studies
Department of Computer Engineering

Supervisor: Assoc. Prof. Dr. Necla BANDIRMALI ERTURK
October 2025, 77 pages

In this study, a web-based intelligent system has been designed and developed to provide personalized career recommendations for computer engineering students based on their academic profiles and technological interests. The project comparatively analyzes two different machine learning modeling approaches. In the first stage (Model-1), it was determined that models trained with a purely rule-based synthetic dataset had limited generalization performance when tested on real student data from 50 individuals, achieving a maximum accuracy of 52% (Logistic Regression). To overcome this generalization problem, a hybrid data augmentation strategy was implemented in the second stage (Model-2), which utilized the real dataset of 50 individuals and enriched it using the class-specific CTGAN (Conditional Tabular Generative Adversarial Network) technique. Models such as Random Forest, Gradient Boosting, and Support Vector Machine (SVM), trained with the synthetic weighted hybrid dataset generated by this method, achieved an accuracy rate of over 97% on the test set. This result confirms that data augmentation with CTGAN, starting from a small amount of real data, significantly enhances the model's learning capacity and predictive power. The web application (WEB-MÖS), developed with the Python Flask framework for deployment, not only serves real-time predictions but also features a

dynamic structure that allows for the periodic retraining of the model by collecting user feedback through Google Sheets API integration. This holistic approach presents an effective career recommendation system prototype that provides data-driven guidance for students in the field of computer engineering education.

Keywords: Machine Learning, Career Recommendation System, CTGAN, Data Augmentation, Hybrid Dataset, Flask Web Application, Computer Engineering Education.

İÇİNDEKİLER

Sayfa

ÖNSÖZ ..... **III**

ÖZET ..... **IV**

ABSTRACT ..... **VI**

İÇİNDEKİLER ..... **VIII**

ŞEKİL LİSTESİ ..... **XI**

TABLO LİSTESİ ..... **XII**

SİMGE VE KISALTMA LİSTESİ ..... **XIII**

1. GİRİŞ **1**

1.1 Problemin Tanımı ve Önemi **1**

1.2 Tez Çalışmasının Amacı ve Kapsamı **2**

1.3 Tezin Araştırma Soruları ve Hipotezler **3**

1.4 Tezin Yapısı **4**

2. LİTERATÜR ARAŞTIRMASI **5**

2.1 Makine Öğrenmesi (MÖ) **5**

2.1.1 MÖ Türleri **5**

2.1.2 Sınıflandırma Algoritmaları **6**

2.2 Öneri Sistemleri **8**

2.2.1 Öneri Sistemi İşlevleri ve Kullanım Alanları **8**

2.2.2 Öneri Sistemi Yaklaşımları **9**

2.2.3 Öneri Sistemlerinde Değerlendirme Ölçütleri **10**

2.3 Web Uygulama Geliştirme Teknolojileri **11**

2.3.1 Genel Web Uygulama Mimarileri **11**

2.3.2 Arka-uç (Backend) Teknolojileri: Python ve Flask Framework **12**

2.3.3 Ön-uç (Frontend) Teknolojileri: HTML, CSS ve Bootstrap **12**

2.3.4 Zeki Web Uygulamaları Kavramı 13
2.4 Eğitimde ve Kariyer Yönlendirmede Yapay Zekâ Uygulamaları 13
2.4.1 Eğitimde Kişiselleştirme ve Yapay Zekânın Rolü 13
2.4.2 Kariyer Danışmanlığı ve Yönlendirmede Teknolojik Yaklaşımlar 14
2.4.3 MÖ Tabanlı Kariyer Öneri Sistemleri Üzerine Literatür İncelemesi 15
3. MATERYAL VE YÖNTEM 18
3.1 Tez Çalışmasında Kullanılan Veri Setleri 18
3.1.1 Model-1 İçin Kullanılan Kural Tabanlı Yapay (Sentetik) Veri Seti 18
3.1.1.1 Kural Tabanlı Veri Üretim Sürecinin Detayları 19
3.1.2 Model-2 İçin Kullanılan Hibrit ve Zenginleştirilmiş Veri Seti 21
3.2 Veri Keşfi ve Ön İşleme 22
3.2.1 Model-1 İçin Veri Ön İşleme Süreci 22
3.2.2 Model-2 İçin Veri Ön İşleme Süreci 23
3.3 Modelleme Yaklaşımı ve Değerlendirme 24
3.3.1 Model Seçimi ve Hiperparametre Optimizasyonu 24
3.3.2 Model-1: Kural Tabanlı Veri ile Eğitim ve Genelleme Testi 24
3.3.3 Model-2: Hibrit CTGAN Verisi ile Eğitim ve Değerlendirme 25
3.3.4 Model Başarım Ölçütleri 26
3.3.5 Özellik Önem Analizi 26
3.3.6 Modellerin ve Yardımcı Nesnelerin Kaydedilmesi 26
3.4 Geliştirilen Sistem ve Web Uygulama Yazılımı 27
3.4.1 Kullanılan Teknolojiler (Flask, Pandas, NumPy, Scikit-learn ve Joblib) 27
3.4.2 Uygulama Mimarisi ve Akışı 28
3.4.3 Kullanıcı Arayüzü Tasarımı ve İşlevleri 29
3.4.4 Sürekli Öğrenme Altyapısı: retrain.py 29
4. BULGULAR 31
4.1 Veri Seti Analizi Bulguları 31

4.1.1 Model-1: Kural Tabanlı Veri Seti Analizi 31

4.1.2 Model-2: Hibrit CTGAN ile Zenginleştirilmiş Veri Seti Analizi 32

4.2 Model Başarım Sonuçları 33

4.2.1 Model-1: Kural Tabanlı Veri İle Elde Edilen Bulgular 34

4.2.1.1 Sentetik Veri Üzerindeki Performans 34

4.2.1.2 Gerçek Veri Üzerindeki Genelleme Başarımı 36

4.2.2 Model-2: Hibrit CTGAN Verisi İle Elde Edilen Bulgular 39

4.2.2.1 Sentetik Veri Artırımı ve Hibrit Veri Setinin Oluşturulması 39

4.2.2.2 Model Başarımının Kıyaslanması ve Temel Bulgular 40

4.2.2.3 Detaylı Sınıflandırma Metrikleri Analizi 41

4.2.2.4 Karar Mekanizması Analizi: Öznitelik Önem Dereceleri (Feature Importances) 43

4.2.2.5 Karmaşıklık Matrisi Analizi 46

4.3 Web Uygulaması Kullanıcı Test Sonuçları ve Örnek Öneriler 47

4.3.1 Gerçek Kullanıcı Verileriyle Örnek Vaka Analizleri 50

5. TARTIŞMA VE SONUÇ 52

5.1 Elde Edilen Bulguların Yorumlanması ve Değerlendirilmesi 52

5.2 Çalışmanın Literatüre ve Potansiyel Uygulama Sektörlerine Katkıları 53

5.3 Tez Çalışmalarında Karşılaşılan Güçlükler ve Çözüm Yaklaşımları 54

5.4 Sonuçlar 54

5.5 Gelecek Çalışmalar İçin Öneriler 55

6. KAYNAKLAR 56

7. EKLER 59

ÖZGEÇMİŞ 62

ŞEKİL LİSTESİ

Şekil **Sayfa**

Şekil 3.1: Hiperparametre Optimizasyonu 24
Şekil 3.2: Kişiselleştirilmiş Kariyer Öneri Platformu – Çalışma Mimarisi 30
Şekil 4.1: Meslek Sınıfı Dağılımı 32
Şekil 4.2: Meslek Sınıfı Dağılımı (CTGAN ile Zenginleştirilmiş Veri).....33
Şekil 4.3: SVM Modeli Karmaşıklık Matrisi 35
Şekil 4.4: Logistic Regression - Karmaşıklık Matrisi.....38
Şekil 4.5: Hibrit CTGAN Verisiyle Eğitilmiş Modellerin Test Başarısı.....40
Şekil 4.6: Random Forest - En Önemli 20 Özellik (Hibrit CTGAN Verisi).....44
Şekil 4.7: Gradient Boosting - En Önemli 20 Özellik (Hibrit CTGAN Verisi).....45
Şekil 4.8: Random Forest Modelinin Hibrit Test Verisi Üzerindeki Karmaşıklık Matrisi ..47
Şekil 4.9: Geliştirilen Web Uygulaması Giriş Arayüzü 48
Şekil 4.10: Web Uygulaması Örnek Sonuç Arayüzü.....50

TABLO LİSTESİ

Tablo **Sayfa**

Tablo 4.1: Model-1'in Sentetik Test Seti Üzerindeki Başarı Oranları 34
Tablo 4.2: Model-1'in Gerçek Veri Seti Üzerindeki Başarı Oranları.....36
Tablo 4.3: Lojistik Regresyon Modelinin Gerçek Veri Seti Sınıflandırma Raporu.....37
Tablo 4.4: Model-2 (Hibrit Veri) ile Eğitilen Modellerin Başarım Sonuçları ve En İyi Hiperparametreleri 40
Tablo 4.5: Model-2 Random Forest ve Gradient Boosting Modellerinin Sınıflandırma Raporu 42
Tablo 4.6: Random Forest ve Gradient Boosting Öznitelik Önem Karşılaştırması 44

SİMGE VE KISALTMA LİSTESİ

Kısaltmalar	**Açıklama**
MÖ	Makine Öğrenmesi
ÖBS	Öğrenci Bilgi Sistemi
YZ	Yapay Zekâ

1. GİRİŞ

Bu bölümde, tez çalışmasının genel çerçevesi sunulmaktadır. İlk olarak, tez projesinin araştırma ve geliştirme problemi ve önemi ele alınarak, tez çalışmanın amacı ve kapsamı detaylandırılmaktır. Son olarak, tezin yapısı hakkında bilgi verilmektir.

1.1 Problemin Tanımı ve Önemi

Dijital teknolojiler ve yapay zekâ uygulamaları, günümüzün hızla değişen dünyasında önemli bir gelişim göstermektedir. Bu teknolojilerin en etkin kullanıldığı alanlardan biri, web tabanlı uygulamalar olmuştur. Web uygulamaları, bireylerin ve kurumların gündelik yaşamları ile profesyonel faaliyetlerinin merkezine yerleşmiş, kişiselleştirilmiş deneyim ve süreçleri iyileştirmek amacıyla makine öğrenmesi (MÖ) yöntem ve araçlarından yararlanılmaya başlanmıştır. Günümüzde, MÖ teknikleri, web platformlarında içerik önerileri, akıllı arama fonksiyonları ve kullanıcı odaklı iyileştirmeler gibi değişik alanlarda kullanılmaktadır [1, 2]. Özellikle öneri sistemleri, kullanıcıların geçmiş davranışlarına, tercihlerine ve ilgi alanlarına göre kişiselleştirilmiş içerik ve hizmet sunarak kullanıcı memnuniyetini artırmakta ve dijital platformların verimliliği ile ticari potansiyelini yükseltmektedir [3, 4]. Aithal vd. [5]'nin belirttiğine göre, günümüzde artan dijitalleşme ile birlikte, geleneksel yöntemlerin getirdiği kısıtlamaların üstesinden gelebilecek güçlü otomatik iş ve kariyer öneri sistemlerine olan ihtiyaç artmıştır. Özellikle, deneyim ve farkındalık eksikliği yaşayan yeni mezunlar için geliştirilen bu tür sistemler, kendilerine uygun işleri verimli bir şekilde bulmalarına önemli ölçüde yardımcı olmaktadır [5]. Kişiselleştirilmiş öneriler sunarak hem yanıltıcı iş tavsiyelerinin önüne geçmek hem de iş arama sürecinde zaman kaybını engellemek, yeni mezun adaylar için büyük bir değer taşımaktadır [5].

Ulaş ve Bandırmalı Ertürk [6], MÖ tekniklerinin modern web uygulamalarına entegrasyonunun, kişiselleştirilmiş deneyimler sağlamak adına büyük bir potansiyel taşıdığını vurgulamaktadır.

Eğitim sektörü, bu teknolojik gelişmelerden büyük ölçüde faydalanmakta ve öğrencilere yönelik yenilikçi yaklaşımlar sunmaktadır. Öğrencilerin akademik kariyerlerinde en kritik aşamalardan biri, mezuniyet sonrası yönelecek kariyer alanını doğru bir şekilde belirlemeleridir. Bu karar süreci, öğrencilerin kişisel ilgi alanları, akademik başarıları, edinilen beceriler ve değişen

iş gücü dinamikleri ve projeksiyonlarını doğru bir şekilde değerlendirmelerini gerektirir [7, 8]. Kariyer karar verme süreci, bireyin kendisini değerlendirme ve iş dünyası hakkında bilgi toplama yeterliliğini gerektiren bilişsel adımlardan oluşur [8]. Özellikle bilgisayar mühendisliği gibi çok sayıda alt uzmanlık dalı barındıran ve hızla gelişen bir alanda, öğrencilerin en uygun kariyer yolunu seçmeleri son derece zorlu bir görev olabilmektedir. Türkiye'de, gençlerin kariyer planlaması sürecinde kendilerini etkili bir şekilde değerlendirmede ve hedeflerini belirlemede zorluklar yaşamaları ve daha etkin rehberlik sistemlerine duyulan ihtiyaç, çeşitli araştırmalarla ortaya konmuştur [9]. Ulaş ve Bandırmalı Ertürk [6], üniversitelerin mevcut Öğrenci Bilgi Sistemleri (ÖBS) gibi veri tabanlarıyla MÖ tekniklerinin birleşmesinin, öğrencilere kişiselleştirilmiş kariyer önerileri sunmada önemli bir potansiyel barındırdığını ifade etmektedir. Ayrıca, yapay zekâ destekli rehberlik sistemlerinin, öğrencilerin motivasyonunu artırarak daha doğru kariyer seçimleri yapmalarına yardımcı olabileceği uluslararası literatürde de geniş bir destek bulmaktadır [10, 11].

1.2 Tez Çalışmasının Amacı ve Kapsamı

Bu tez çalışmasının temel amacı, bilgisayar mühendisliği öğrencilerine akademik başarıları ve teknolojik ilgi alanlarını esas alarak kişiselleştirilmiş alt meslek alanı ve kariyer önerilerinde bulunan bir web tabanlı sistem tasarlamak ve geliştirmektir. Çalışmanın bir diğer önemli amacı ise, farklı sentetik veri üretme stratejilerinin MÖ modellerinin genelleme başarımı üzerindeki etkisini karşılaştırmalı olarak analiz etmektir. Bu kapsamda, iki farklı modelleme yaklaşımı geliştirilmiş ve başarımları karşılaştırılmıştır.

Geliştirilen sistem ve web yazılımı, özetle aşağıdaki adımları içermektedir:

- **Model-1 (Kural Tabanlı Yaklaşım):** Bilgisayar mühendisliği dersleri, teknolojik ilgi alanları ve alt meslek alanı profillerine dayalı olarak tamamen kurallara bağlı bir yapay veri seti oluşturulmuştur. Bu veri setiyle eğitilen çeşitli MÖ modellerinin, az sayıdaki gerçek öğrenci verisi üzerindeki başarımı ölçülmüştür.

- **Model-2 (Hibrit CTGAN Yaklaşımı):** Az sayıda gerçek öğrenci verisini temel alan ve bu verileri sınıf-spesifik CTGAN (Conditional Tabular Generative Adversarial Network) tekniği ile zenginleştiren hibrit bir veri çoğaltma stratejisi uygulanmıştır. Bu yöntemle oluşturulan veri setiyle eğitilen modellerin başarımı değerlendirilmiştir.

- **Karşılaştırmalı Analiz:** Her iki modelleme yaklaşımıyla eğitilen Lojistik Regresyon, K-En Yakın Komşu (KNN), Destek Vektör Makineleri (SVM), Random Forest, Gradyan Artırma (Gradient Boosting) ve Çok Katmanlı Algılayıcı (MLP) gibi farklı algoritmaların başarımları, GridSearchCV ile optimize edilerek karşılaştırılmıştır.

- **Web Uygulaması Geliştirme:** En başarılı yaklaşım esas alınarak, kullanıcıların anlık ve kişiselleştirilmiş alt meslek alanı ve kariyer önerileri alabileceği, Python tabanlı Flask web çatısı ile bir web uygulaması (WEB-MÖS) geliştirilmiştir.

- **Dinamik Öğrenme Altyapısı:** Geliştirilen web uygulaması, render.com platformunda canlıya alınarak gerçek kullanıcılardan geri bildirim toplanmasına olanak sağlanmıştır. Toplanan bu geri bildirimlerle modelin gelecekte periyodik olarak yeniden eğitilmesine imkân tanıyan bir sistem (retrain.py) tasarlanmıştır.

Bu tez çalışması, farklı veri üretme stratejilerinin model başarısına etkisini ortaya koyarak, az sayıda gerçek veriyle dahi yüksek başarımlı öneri sistemleri geliştirilebileceğini gösteren bir çözüm prototipi sunmayı hedeflemektedir.

1.3 Tezin Araştırma Soruları ve Hipotezler

Bu tez çalışması, aşağıdaki temel araştırma sorularına yanıt aramaktadır:

- Bilgisayar mühendisliği öğrencilerinin ders notları ve ilgi alanlarına dayalı olarak, MÖ algoritmaları ile potansiyel mesleki eğilimleri ne ölçüde başarılı bir şekilde tahmin edilebilir?
- Tamamen kurallara dayalı üretilen bir yapay veri seti (Model-1) ile az sayıda gerçek veriyi CTGAN ile zenginleştiren hibrit bir veri seti (Model-2) arasında, modelin genelleme başarımı açısından nasıl bir fark bulunmaktadır?
- Güncel MÖ algoritmaları (SVM, MLP, KNN vb.), bu iki farklı veri stratejisi üzerinde nasıl bir başarım sergilemektedir? Hangi yaklaşım, gerçek dünya verisine daha iyi genelleme yapabilen modeller üretmektedir?

- Geliştirilen MÖ modeli, öğrencilere kişiselleştirilmiş kariyer önerileri sunmak amacıyla bir Flask web uygulamasına başarılı bir şekilde entegre edilebilir mi?
- Gerçek kullanıcılardan geri bildirim toplayarak modelin zamanla kendini iyileştirmesini sağlayacak dinamik bir yeniden eğitim altyapısı tasarlanabilir mi?

1.4 Tezin Yapısı

Bu tez beş ana bölümden oluşmaktadır:

- **Birinci Bölüm:** Giriş bölümünde, çalışmanın problemi, amacı, kapsamı, araştırma soruları ve tezin yapısı tanıtılmaktadır.

- **İkinci Bölüm:** Makine öğrenmesinin temel kavramları, öneri sistemleri, sentetik veri üretme teknikleri (CTGAN dahil), web uygulama geliştirme teknolojileri ve literatürdeki benzer kariyer yönlendirme sistemleri üzerine yapılan tarama sunulmaktadır.

- **Üçüncü Bölüm:** Materyal ve yöntem kısmında, çalışmada kullanılan iki farklı modelleme yaklaşımı (Model-1 ve Model-2) detaylandırılmaktadır. Veri setlerinin oluşturulma süreçleri, veri ön işleme adımları, kullanılan MÖ algoritmaları ve geliştirilen web uygulamasının mimarisi açıklanmaktadır.

- **Dördüncü Bölüm:** Her iki model yaklaşımı için yapılan veri analizleri, model başarım testleri ve karşılaştırmalı bulgular sunulmaktadır. Ayrıca, geliştirilen web uygulamasının kullanım senaryolarına dayalı çıktılar verilmektedir.

- **Beşinci Bölüm:** Tartışma ve sonuç kısmında, elde edilen bulgular ışığında iki modelin başarımı karşılaştırılarak genel bir değerlendirme yapılmaktadır. Çalışmanın kısıtları tartışılarak gelecek araştırmalar için önerilerde bulunulmaktadır.

Tezin sonunda, çalışmada kullanılan kaynaklar verilerek ek bilgiler sunulmaktadır.

2. LİTERATÜR ARAŞTIRMASI

Bu bölümde, tez kapsamında ele alınan temel kavramlara ve kullanılan yöntemlere yer verilmektedir. Öncelikle MÖ ve öneri sistemleri hakkında genel bir çerçeve çizilmekte, ardından web tabanlı uygulama geliştirmede tercih edilen teknolojiler üzerinde durulmaktadır. Son olarak, YZ temelli kariyer yönlendirme uygulamalarıyla ilgili literatürde yer alan güncel çalışmalardan örnekler incelenmektedir.

2.1 Makine Öğrenmesi (MÖ)

MÖ, bilgisayar sistemlerinin veri setleri üzerinden öğrenerek kendini geliştirmesini sağlayan YZ alt alanlarından biridir. Bu doğrultuda, sistemlerin deneyimlerden yola çıkarak bilgi edinmesi ve benzer durumlarla karşılaştığında bu bilgiyi kullanarak karar vermesi ya da tahminde bulunması amaçlanmaktadır.

2.1.1 MÖ Türleri

MÖ kavramı ilk kez 1959 yılında Arthur Samuel tarafından "bilgisayarların açıkça programlanmadan öğrenmesini sağlayan bir alan" olarak tanımlanmıştır [14]. Bu tanım, alandaki ilk yaklaşımlardan biri olup günümüzde de temel çerçeveyi yansıtmaktadır. Daha sonra Tom Mitchell [15], bu alanı daha sistematik biçimde şöyle tanımlamıştır: "Bir bilgisayar programının T görevlerini P başarım ölçütüyle yerine getirme başarısı, E deneyimiyle iyileşiyorsa, o program öğreniyor denir". Bu tanım, MÖ'ni görev (T), deneyim (E) ve başarım ölçütü (P) olmak üzere üç temel bileşenle açıklar.

MÖ yöntemleri ise genellikle kullanılan veri türüne ve geri bildirim yapısına göre üç ana gruba ayrılmaktadır:

- **Denetimli Öğrenme (Supervised Learning):** Bu öğrenme türü, etiketlenmiş veri setleri üzerinden çalışır. Yani her veri girdisine karşılık gelen bir çıktı (etiket) vardır. Algoritma, giriş verileriyle bu etiketler arasındaki ilişkiyi öğrenerek, yeni gelen veriler için tahmin yapabilir hale gelir. Genellikle sınıflandırma ve regresyon gibi problemler bu gruba girer. Bu tez kapsamında geliştirilen meslek öneri sistemi, öğrencilerin belirli özelliklerine göre

meslek sınıflandırması yaptığı için denetimli öğrenme kapsamında değerlendirilmiştir [12, s. 5–7].

- **Denetimsiz Öğrenme (Unsupervised Learning):** Veriler etiketlenmemiştir. Algoritmanın doğrudan "doğru cevap" diye bir şey bilmeden, veriler arasındaki örüntüleri ya da yapısal ilişkileri kendi kendine bulması beklenir. Bu yöntemle genelde kümeleme, boyut indirgeme gibi görevler gerçekleştirilir [12, s. 7–9].

- **Pekiştirmeli Öğrenme (Reinforcement Learning):** Bu yöntemde bir ajan, bulunduğu ortamda çeşitli eylemler gerçekleştirerek ödül maksimizasyonu üzerinden bir strateji öğrenmesini hedefler. Özellikle robotik uygulamalarda veya oyunlarda sıkça karşımıza çıkar [16].

Bu tez çalışmasında, öğrencilerin ilgi alanları, ders başarı durumları gibi çeşitli özelliklerine bakılarak onlar için uygun alt meslek alanlarının tahmin edilmesi amaçlandığından, denetimli öğrenme yaklaşımı esas alınmış ve buna uygun sınıflandırma algoritmaları kullanılmıştır.

2.1.2 Sınıflandırma Algoritmaları

Sınıflandırma, verilerin önceden tanımlanmış kategorilere veya sınıflara ayrılması amacıyla kullanılan bir denetimli öğrenme problemidir [12, s. 75–77]. Bu yöntemde model, etiketlenmiş verilerden öğrendiği örüntüler sayesinde, daha önce görmediği yeni verilerin hangi sınıfa ait olduğunu tahmin etmeye çalışır.

Bu tez çalışmasında, kariyer öneri probleminin çözümü için farklı öğrenme yaklaşımlarını temsil eden, literatürde yaygın kullanıma sahip altı temel sınıflandırma algoritması seçilmiş ve başarımları karşılaştırılmıştır. Bu algoritmaların seçimi, hem basit ve yorumlanabilir modellerden hem de karmaşık ve yüksek başarımlı topluluk modellerine kadar geniş bir yelpazeyi kapsamayı amaçlamıştır. Kullanılan algoritmalar ve seçilme gerekçeleri şunlardır:

- **Lojistik Regresyon (Logistic Regression):** Doğrusal bir model olan Lojistik Regresyon, bir temel (baseline) model olarak kullanılmıştır. Sınıflar arasındaki ilişkileri olasılıksal olarak modellemesi ve yorumlanabilirliğinin yüksek olması nedeniyle tercih edilmiştir.
- **K-En Yakın Komşu (K-Nearest Neighbors - KNN):** Örnek tabanlı bir öğrenme algoritmasıdır. Bir veri noktasını, özellik uzayındaki en yakın komşularının çoğunluk sınıfına göre etiketler. Karmaşık karar sınırlarını parametrik olmayan bir yapıyla öğrenebilme yeteneği nedeniyle çalışmaya dâhil edilmiştir.

- **Destek Vektör Makineleri (Support Vector Machines - SVM):** Farklı sınıflara ait veri noktaları arasına en geniş "marjı" (boşluğu) çizen bir hiper-düzlem bularak sınıflandırma yapar. Özellikle **rbf** gibi çekirdek (kernel) fonksiyonları ile doğrusal olmayan karmaşık ilişkileri modellemedeki başarısı nedeniyle önemli bir adaydır.

- **Random Forest (Rastgele Orman):** Birden çok karar ağacını bir araya getiren bir topluluk öğrenmesi yöntemidir. Aşırı öğrenmeye (overfitting) karşı dayanıklı olması, yüksek doğruluk oranları sunması ve özelliklerin önem derecesini ölçebilmesi gibi avantajları nedeniyle bu çalışmada kullanılan temel modellerden biridir.

- **Gradyan Artırma Makineleri (Gradient Boosting Machines):** Random Forest gibi bir diğer güçlü topluluk öğrenmesi tekniğidir. Ancak ağaçları bağımsız olarak oluşturmak yerine, her yeni ağacın bir önceki ağacın hatalarını düzeltecek şekilde sıralı olarak inşa edildiği bir yapı kullanır. Genellikle çok yüksek başarım oranları elde etmesiyle bilinir.

- **Çok Katmanlı Algılayıcı (Multi-Layer Perceptron - MLP):** İnsan beyninden esinlenen bir yapay sinir ağı modelidir. Girdi, gizli ve çıktı katmanlarından oluşur. Özellikler arasındaki karmaşık ve doğrusal olmayan desenleri öğrenebilme kapasitesi sayesinde, bu tür sınıflandırma problemlerinde güçlü bir alternatif olarak değerlendirilmiştir.
  [13, s. 168–236; 17]

Bu algoritmaların her birinin başarımı, farklı veri üretme stratejileri (kural tabanlı ve hibrit CTGAN) üzerinde karşılaştırmalı olarak analiz edilmiştir. En iyi sonuçları elde etmek amacıyla

tüm modellere GridSearchCV yöntemiyle hiperparametre optimizasyonu uygulanmış ve bulgular çalışmanın ilerleyen bölümlerinde detaylı olarak sunulmuştur.

2.2 Öneri Sistemleri

Öneri sistemleri, kullanıcıların geçmiş davranışları, tercihleri, demografik bilgileri ya da benzer kullanıcıların etkileşimlerinden elde edilen verileri analiz ederek, kişiye özel ürün, hizmet, içerik veya bilgi sunan bilgi filtreleme araçlarıdır [3, 19]. Günümüzde e-ticaret sitelerinden dijital içerik platformlarına, sosyal medyadan eğitim uygulamalarına kadar birçok alanda yaygın olarak kullanılmaktadır. Bu sistemlerin temel amacı, kullanıcıların ilgi duyabileceği ögeleri daha kolay keşfetmelerini sağlamak, bilgi aşırı yükünü azaltmak ve genel kullanıcı deneyimini iyileştirmektir. Ulaş ve Bandırmalı Ertürk [6], MÖ tabanlı öneri sistemlerinin web uygulamalarına entegrasyonunun kişiselleştirme potansiyelini önemli ölçüde artırdığını belirtmektedir.

2.2.1 Öneri Sistemi İşlevleri ve Kullanım Alanları

Bir öneri sistemi, belirli bir kullanıcıya, mevcut seçenekler arasından hangilerinin onun için daha anlamlı ya da yararlı olabileceğini tahmin etmeye çalışır. Bu tahminler; oy verme geçmişi, satın alma kayıtları, tıklama verileri veya kullanıcı tarafından belirtilen ilgi alanları gibi farklı veri türlerine dayanabilir.

Öneri sistemlerinin temel işlevleri şu şekilde özetlenebilir [3, s. 1–15]:

- **Kişiselleştirme:** Kullanıcıların ilgi ve ihtiyaçlarına uygun içerik ya da hizmet sunmak.
- **Keşif (Serendipity):** Kullanıcının farkında olmadığı ancak ilgi duyabileceği yeni içerikleri önermek.
- **Satış Artışı:** E-ticaret bağlamında çapraz satış (cross-selling) ve yukarı satış (up-selling) gibi yollarla gelirleri artırmak.
- **Kullanıcı Sadakati:** İlgili içerik sunarak kullanıcıların platformda daha uzun süre kalmasını sağlamak.
- **Bilgi Aşırı Yükünü Azaltmak:** Kullanıcıya uygun öneriler sunarak karar verme sürecini kolaylaştırmak.

Bu tez çalışmasında geliştirilen spesifik öneri sistemi, öğrencilerin simüle edilmiş ders notları ve ilgi alanlarına dayalı olarak kişiselleştirilmiş alt meslek alanı tavsiyesi sunan bir çözüm olarak bu kapsamda değerlendirilebilir.

2.2.2 Öneri Sistemi Yaklaşımları

Literatürde öneri sistemlerinin geliştirilmesinde çeşitli yöntemler öne çıkmaktadır. En yaygın yaklaşımlar şunlardır [3, s. 17–33; 20]:

- **İçerik Tabanlı Filtreleme (Content-Based Filtering)**

Bu yöntem, kullanıcının daha önce ilgi gösterdiği ögelerin özelliklerine odaklanarak, benzer içeriklere sahip yeni ögeleri önerir. Örneğin, bilim kurgu kitapları okuyan bir kullanıcıya başka bilim kurgu eserleri önerilmesi bu yaklaşıma örnektir.

Avantajları:

- Diğer kullanıcıların verisine ihtiyaç duymaz.
- Yeni veya az bilinen içerikleri de önerme yeteneğine sahiptir.
- Önerilerin gerekçesi açıklanabilir.

Dezavantajları:

- İçerik özelliklerinin doğru tanımlanması gerekir.
- Kullanıcıları alışılmış tercihlere yönlendirebilir.
- Yeni kullanıcılar için etkili çalışmayabilir (**cold-start** problemi).

- **İşbirlikçi Filtreleme (Collaborative Filtering)**

Bu yöntemde öneriler, benzer kullanıcıların tercihleri üzerinden oluşturulur. İki alt türü vardır:

Kullanıcı tabanlı: Benzer kullanıcıların beğendiği ancak hedef kullanıcının henüz deneyimlemediği ögeleri önerir.

Öge tabanlı: Hedef kullanıcının sevdiği içeriklere benzer şekilde tercih edilen diğer içerikleri önerir.

Avantajları: İçerik bilgisine ihtiyaç duymaz; farklı ve beklenmedik öneriler sunabilir.

Dezavantajları: Yeni kullanıcılar ve ögeler için yeterli veri bulunmadığında etkisizdir; veri seyrekliği (sparsity) sorunuyla karşılaşabilir; popüler içeriklere aşırı odaklanabilir.

- **Hibrit Yaklaşımlar**

Hibrit sistemler, içerik tabanlı ve işbirlikçi filtreleme yöntemlerinin avantajlarını bir araya getirmeyi amaçlar. Yöntemler, birbirini tamamlayacak şekilde birleştirilebilir ya da çıktıları entegre edilebilir. Ulaş ve Bandırmalı Ertürk (2023) [6], mühendislik öğrencilerine yönelik geliştirilen bir sistemde hibrit yaklaşımın sunduğu potansiyel faydalara dikkat çekmektedir.

Bu tez çalışmasında kullanılan yöntem doğrudan bir hibrit model olmasa da, hem içerik özelliklerine (ders notları ve ilgi alanları) dayalı hem de benzer kullanıcı eğilimlerini dolaylı olarak modelleyen bir yaklaşım olarak değerlendirilebilir.

- **Diğer Yaklaşımlar**

Bilgi tabanlı, demografik filtreleme, fayda tabanlı ve risk farkındalıklı öneri sistemleri gibi alternatif yöntemler de belirli bağlamlarda tercih edilmektedir [3].

Bu tezde geliştirilen sistem, öğrencilerin bireysel özelliklerinden yola çıkarak uygun alt meslek gruplarını sınıflandırmaya çalışmaktadır. Bu yapı, esasen sınıflandırma tabanlı bir öneri sistemi olarak değerlendirilebilir ve içerik tabanlı filtrelemeye benzer bir mantıkla çalışır.

2.2.3 Öneri Sistemlerinde Değerlendirme Ölçütleri

Bir öneri sisteminin başarımını ölçmek için çeşitli değerlendirme ölçütleri kullanılır [3, s. 369–408]. Bunlardan başlıcaları şunlardır:

- **Tahmin Doğruluğu:** Kullanıcının bir ögeye vereceği puanı ne kadar doğru tahmin ettiğini ölçer (örneğin, MAE ve RMSE). Bu tezde kullanılmamıştır.
- **Sınıflandırma Doğruluğu:** Önerilen ögelerin gerçekten kullanıcıya uygun olup olmadığını değerlendirir. **Precision, Recall, F1 Skoru** ve **AUC** gibi ölçütler bu kategoriye girer. Tez çalışmalarında model başarısı bu tür ölçütlerle belirlenmiştir.

- **Sıralama Ölçütleri:** Önerilen ögelerin doğruluk sırasını dikkate alır (örneğin, MAP ve NDCG).
- **Kapsam (Coverage):** Sistem tarafından önerilebilecek tüm ögelerin ne kadarının önerildiğini gösterir.
- **Çeşitlilik (Diversity):** Önerilerin birbirine ne kadar benzediğini ölçer.
- **Serendipity:** Kullanıcının beklemediği ama hoşuna gidebilecek öneriler sunma kapasitesini ifade eder.

Bu tezde öneri sistemi, sınıflandırma temelli bir yapı üzerinde çalıştığı için model başarımı temel olarak sınıflandırma doğruluğu ölçütleri (**Accuracy, Precision, Recall** ve **F1-Score**) ile değerlendirilmiştir.

2.3 Web Uygulama Geliştirme Teknolojileri

Bu tez kapsamında geliştirilen MÖ modelinin kullanıcılarla etkileşimli bir biçimde buluşturulabilmesi için bir web uygulaması tasarlanmış, gerçekleştirilmiş ve uygulamaya geçirilmiştir. Bu bölümde, uygulamanın geliştirilmesinde tercih edilen temel teknolojiler ve yaklaşımlar kısaca açıklanmaktadır. Ulaş ve Bandırmalı Ertürk [6], kullanıcı ve servis odaklı zeki web uygulamalarında uygun teknolojik altyapının seçilmesinin proje başarısı açısından belirleyici olduğunu vurgulamaktadır.

2.3.1 Genel Web Uygulama Mimarileri

Web uygulamaları mimari açıdan farklı yapılar sergileyebilir. Geleneksel olarak monolitik mimariler yaygınken, günümüzde daha modüler ve ölçeklenebilir çözümler sunan mikroservis mimarileri sıkça tercih edilmektedir [20]. Monolitik mimaride uygulamanın tüm bileşenleri tek bir yapı altında geliştirilir ve dağıtılırken; mikroservis mimarisinde sistem, bağımsız olarak geliştirilebilen ve yönetilebilen küçük servisler hâline ayrılır.

Bu tez çalışmasında geliştirilen web uygulama, projenin ölçeği ve karmaşıklığı dikkate alınarak daha sade ve hızlı geliştirilebilir bir yapı sunan monolitik mimariyle hayata geçirilmiştir. Bu mimari, Python tabanlı Flask çatısı kullanılarak uygulanmıştır.

2.3.2 Arka-uç (Backend) Teknolojileri: Python ve Flask Framework

Web uygulamasının sunucu tarafı, verilerin işlenmesi, MÖ modelinin çalıştırılması ve veri depolama amacıyla Python programlama dili kullanılarak geliştirilmiştir. Python'ın veri bilimi ve MÖ alanındaki güçlü kütüphaneleri (Pandas, NumPy, Scikit-learn, Joblib vb.) bu tercihi destekleyen başlıca nedenlerdendir.

Sunucu tarafı geliştirmede, Python tabanlı hafif bir web çatısı olan Flask tercih edilmiştir. Flask, esnek yapısı ve öğrenim kolaylığı sayesinde bu projeye uygun bir çözüm sunmuştur. Uygulamanın app.py dosyasında HTTP isteklerinin yönlendirilmesi, modelin yüklenmesi, kullanıcı girdilerinin işlenmesi, tahmin yapılması ve sonuçların arayüze aktarılması gibi işlevler bütüncül biçimde ele alınmıştır.

Bu projenin en kritik teknolojik bileşenlerinden biri de Google Sheets API entegrasyonudur. Kullanıcı geri bildirimlerini toplamak ve modelin sürekli öğrenmesini sağlamak amacıyla gspread ve oauth2client kütüphaneleri kullanılmıştır. Bu sayede, kullanıcıların result.html sayfasında gönderdiği geri bildirimler, güvenli bir şekilde doğrulanarak Career Feedback Data adlı Google E-tablosuna anlık olarak kaydedilmektedir. Bu yapı, uygulamanın statik bir tahmin aracından, zamanla kendini iyileştirebilecek dinamik bir veri toplama platformuna dönüşmesini sağlamıştır.

2.3.3 Ön-uç (Frontend) Teknolojileri: HTML, CSS ve Bootstrap

Kullanıcı arayüzü, standart web teknolojileri olan HTML, Basamaklı Stil Şablonları (Cascading Style Sheets - CSS) ve sınırlı düzeyde JavaScript kullanılarak geliştirilmiştir. Arayüzün temel yapısı HTML ile oluşturulmuş, index.html (giriş formu) ve result.html (sonuç sayfası) olmak üzere iki ana şablon dosyası kullanılmıştır.

Sayfaların görsel tasarımı CSS ile yapılandırılmış, duyarlı (responsive) bir kullanıcı deneyimi sunmak amacıyla Bootstrap 5.3.2 kütüphanesinden yararlanılmıştır. Bootstrap'in ızgara (grid) sistemi, form bileşenleri, düğmeler ve uyarı mesajları gibi hazır bileşenleri, geliştirme sürecini hızlandırmıştır. Jinja2 motoru aracılığıyla, Python arka-uçlarından gelen dinamik veriler (örneğin ilgi alanları listesi, önerilen meslek ve meslek açıklamaları) doğrudan HTML şablonlarına entegre edilerek kullanıcıya özelleştirilmiş içerikler sunulmuştur.

2.3.4 Zeki Web Uygulamaları Kavramı

Bu tez çalışması kapsamında geliştirilen web tabanlı uygulama, temelinde bir MÖ modeli barındırması ve kullanıcı etkileşimlerinden öğrenme potansiyeli taşıması sayesinde "zeki" bir işlevsellik sunmaktadır. Ulaş ve Bandırmalı Ertürk [6], zeki web uygulamalarını; kullanıcı ihtiyaçlarına duyarlı, kişiselleştirilmiş deneyimler sağlayan ve değişen kullanıcı tercihlerine dinamik biçimde uyum sağlayabilen sistemler olarak tanımlamaktadır.

Bu bağlamda geliştirilen alt meslek alanı öneri sistemi, kullanıcının sağladığı detaylı ve kişisel akademik başarı verilerine (harf notları) ve ilgi alanlarına dayanarak kişiye özel önerilerde bulunmaktadır. Daha da önemlisi, Google Sheets API aracılığıyla kurulan geri bildirim mekanizması, sistemin gelecekte bu yeni verilerle yeniden eğitilerek zamanla daha isabetli ve "akıllı" hale gelmesine olanak tanıyan bir yapı oluşturur. Bu yönüyle uygulama, söz konusu "zeki web uygulaması" tanımına tam olarak uymaktadır.

2.4 Eğitimde ve Kariyer Yönlendirmede Yapay Zekâ Uygulamaları

YZ ve onun bir alt alanı olan MÖ, son yıllarda eğitim dâhil olmak üzere birçok sektördeki hızlı dönüşüm potansiyeliyle dikkat çekmektedir. Eğitim bağlamında YZ'nin kullanımı; öğrenme süreçlerini daha etkileşimli, verimli ve özellikle bireysel ihtiyaçlara duyarlı hale getirmeyi amaçlamaktadır. Bu doğrultuda, öğrencilerin akademik ve mesleki gelişimlerine rehberlik edebilecek akıllı sistemlerin geliştirilmesi, güncel araştırmaların odak noktalarından biri olmuştur.

2.4.1 Eğitimde Kişiselleştirme ve Yapay Zekânın Rolü

Kişiselleştirilmiş öğrenme, her öğrencinin özgün öğrenme tarzı, hızı, bilgi düzeyi, ilgi alanları ve hedefleri doğrultusunda uyarlanmış eğitim deneyimleri sunmayı amaçlayan bir yaklaşımdır [21]. Geleneksel "tek tip" öğretim modellerinden farklı olarak, öğrenci merkezli bir anlayış benimseyen bu yaklaşım, öğrenme motivasyonunu ve akademik başarıyı artırmayı hedefler. YZ, bu kişiselleştirme hedefinin gerçekleştirilmesinde önemli bir teknolojik araç olarak öne çıkmaktadır. Özellikle büyük veri kümelerini analiz edebilme kapasitesi sayesinde (örneğin çevrimiçi platform etkileşimleri, sınav sonuçları, ödev başarımları ve demografik bilgiler), YZ

sistemleri; bireysel öğrenme yolları oluşturabilir, içerik önerilerini kişiye özel uyarlayabilir ve gerçek zamanlı geri bildirimler sunabilir [10].

YZ tabanlı sistemlerin eğitimdeki uygulama alanları arasında akıllı öğretim sistemleri (Intelligent Tutoring Systems – ITS), öğrenme analitikleri, otomatik değerlendirme sistemleri ve kişiselleştirilmiş içerik öneri platformları öne çıkmaktadır. Akıllı öğretim sistemleri, öğrencinin bilgi düzeyini ve yaşadığı öğrenme güçlüklerini anlık olarak izleyerek kişiselleştirilmiş yönlendirmeler ve ek materyaller sağlayabilir [22]. Öğrenme analitikleri ise öğrencilerin davranışlarındaki örüntüleri analiz ederek eğitmen ve eğitim kurumlarına stratejik kararlar için veri temelli içgörüler sunar. Bu sayede, özellikle risk altındaki öğrenciler erken evrede belirlenebilir ve onlara yönelik destek mekanizmaları zamanında devreye sokulabilir [23]. Ulaş ve Bandırmalı Ertürk [6], MÖ araçlarının web uygulamalarına entegrasyonunun, özellikle öneri sistemleri bağlamında, kullanıcı odaklı ve esnek çözümler geliştirme potansiyelini artırdığını vurgulamaktadır.

2.4.2 Kariyer Danışmanlığı ve Yönlendirmede Teknolojik Yaklaşımlar

Kariyer seçimi, bireylerin yaşamlarını uzun vadede etkileyen kritik kararlardan biridir ve bu sürecin sağlıklı bir şekilde yürütülmesi doğru yönlendirmeyi gerektirir. Geleneksel kariyer danışmanlığı yaklaşımları; yüz yüze görüşmeler, standart testler ve genel meslek bilgileri sunumu gibi yöntemlere dayanmaktadır [24]. Ancak bu yöntemler, bireysel farklılıkları tam anlamıyla gözetmekte yetersiz kalmakta ve hızla değişen iş gücü piyasasının dinamiklerini yeterince yansıtamamaktadır. Son yıllarda teknolojinin gelişmesiyle birlikte, web ve mobil tabanlı uygulamalar aracılığıyla kariyer danışmanlığına daha esnek ve etkileşimli yaklaşımlar entegre edilmeye başlanmıştır.

Yapay zekâ ve MÖ, kariyer danışmanlığı süreçlerine veriye dayalı ve bireyselleştirilmiş bir bakış açısı kazandırmaktadır. Öğrencilerin ya da iş arayan bireylerin akademik geçmişleri, becerileri, ilgi alanları, kişilik özellikleri ve özgeçmişleri gibi çok boyutlu veriler analiz edilerek, onlara en uygun kariyer yolları, iş fırsatları ya da gelişime açık alanlar belirlenebilmektedir. Tekin [25], lise ve üniversite öğrencileri ile çeşitli meslek gruplarından toplanan anket verilerini (RIASEC kişilik testi sonuçları ve ders tercihleri) kullanarak işbirlikçi filtreleme ve kosinüs benzerliği yöntemleriyle kişiselleştirilmiş meslek önerileri sunmayı amaçlayan bir çalışma gerçekleştirmiştir. Bu tür sistemler, bireylerin kendilerini daha yakından tanımalarına ve kişisel özelliklerine uygun mesleki alternatifleri keşfetmelerine katkı sağlayabilir.

2.4.3 MÖ Tabanlı Kariyer Öneri Sistemleri Üzerine Literatür İncelemesi

MÖ tabanlı kariyer öneri sistemleri üzerine yapılan araştırmalar, çeşitli veri kaynaklarının, algoritmaların ve yöntemlerin kullanıldığı geniş bir yelpazeye sahiptir. Bu çalışmaların ortak amacı, bireylere daha doğru ve kişiselleştirilmiş kariyer rehberliği sağlamaktır.

Prasanna ve Haritha [26], Bilgisayar Bilimleri ve Bilişim Teknolojileri bölümlerinde eğitim gören öğrenciler için beceri ve sertifika kursları öneren bir "Akıllı Kariyer Rehberlik ve Öneri Sistemi" geliştirmişlerdir. Bu çalışmada, öğrencilerin notları ve ilgi alanları gibi faktörler, öneri sisteminin temel girdileri olarak kullanılmıştır. Ayrıca, SVM, Random Forest ve Karar Ağaçları gibi farklı MÖ algoritmalarının bu tür bir sistemde nasıl etkili olabileceği tartışılmıştır. Bu araştırma, öğrencilerin doğru beceri setlerini geliştirmelerine ve uygun kariyer yollarını seçmelerine yardımcı olabilecek bir öneri sisteminin önemini vurgulamaktadır.

Benzer şekilde, Rajput vd. [27], "Career Craft AI" adını verdikleri, kişiselleştirilmiş özgeçmiş analizi ve iş önerileri sunan bir sistem geliştirmişlerdir. Bu sistem, bireylerin beceri ve ilgi alanlarına en uygun işleri bulma sürecindeki zorlukları ele alırken, aynı zamanda özgeçmiş iyileştirmeleri için de rehberlik sunmaktadır. Kullanıcıları beceri setlerine ve profillerine göre sınıflandıran ve Doğal Dil İşleme (NLP) tekniklerini kullanarak özgeçmiş analizi gerçekleştiren bu yaklaşım, iş arama sürecindeki bireyler için önemli bir çözüm sunmaktadır.

Aithal vd. [5] tarafından yapılan çalışmada, yeni mezunların yetenekleri ve ilgi alanlarına dayalı olarak ideal iş rollerini öneren bir sistem sunulmuştur. Bu sistemin temelinde, kullanıcıların profil bilgilerini ve yeteneklerini analiz ederek bunlara en uygun iş ilanlarını sunan İçerik Tabanlı Filtreleme tekniği bulunmaktadır. Sistem, adayın belirttiği yetenekleri ve ilgi alanlarını, iş tanımlarında yer alan anahtar kelimelerle eşleştirerek kişiselleştirilmiş öneriler üretir. Bu yaklaşımın, kullanıcıların kendi ilgi alanlarına dayalı daha isabetli sonuçlar elde etmesini sağladığı ve modelin daha ölçeklenebilir olmasına olanak tanıdığı belirtilmiştir.

Mevcut kariyer öneri sistemleri, genellikle sadece kullanıcının ilgi duyduğu alana odaklanırken, adayın detaylı profilini, yeteneklerini ve en önemlisi akademik başarısını göz ardı etmektedir. Yadalam vd. tarafından yapılan "Career Recommendation Systems using Content based Filtering" isimli çalışmada, bu boşluğu doldurmayı hedefleyerek öğrencilerin ilgi alanları, yetenekleri ve akademik başarılarını birleştiren bir kariyer öneri sistemi sunmaktadır. Çalışmada, İçerik Tabanlı Filtreleme yaklaşım benimsenmiştir. Bu yaklaşım, sistemin kullanıcının profilini ve özgeçmişini tarayarak kişiselleştirilmiş iş önerileri üretmesine olanak tanır. Sistemin en dikkat

çekici yönlerinden biri, öneri mekanizmasını beslemek için kullanılan veri setidir. Bu veri seti, öğrencilerin akademik başarımını doğrudan yansıtan; İşletim Sistemleri, Algoritmalar ve Tasarım, Programlama ve Bilgisayar Ağları gibi derslerdeki yüzdelik başarı oranlarını içermektedir. Akademik notların yanı sıra, kullanıcının iletişim becerileri, kodlama yetenekleri ve katıldığı hackathon sayısı gibi kendini değerlendirdiği ölçütler de veri setine eklemiştir. Yadalam vd., geliştirdikleri modelde kullanıcı profili ile işler arasındaki yakınlığı ölçmek için iki vektör arasındaki benzerliği hesaplayan kosinüs benzerliği metriğini kullanmıştır. Ayrıca sistem, kullanıcı memnuniyetini ölçmek ve modeli iyileştirmek amacıyla bir geri bildirim mekanizması içermektedir. Bu mekanizmada, kullanıcıların metin olarak girdiği yorumlar NLP teknikleri ile işlenerek duygu analizine tabi tutulur ve geri bildirimin olumlu, olumsuz veya nötr olduğu tespit edilir. Makale aynı zamanda, önerilen yaklaşımları literatürdeki diğer yöntemlerle de karşılaştırmaktadır. Özellikle, benzer kullanıcıların tercihlerine dayanan İşbirlikçi Filtreleme yönteminin, hakkında yeterli veri bulunmayan yeni kullanıcılara öneri yapmada zorlandığı soğuk başlangıç gibi problemlere sahip olduğuna dikkat çekilmektedir. Bu tür sorunları aşmak için İçerik Tabanlı ve İşbirlikçi Filtreleme yöntemlerini birleştiren Hibrit Filtreleme yaklaşımlarının da literatürde daha yüksek doğruluk sağlayabildiği belirtilmiştir [29].

Tekin [28], lise ve üniversite öğrencileri ile çeşitli meslek gruplarındaki bireyler için kişiselleştirilmiş meslek önerileri sunan bir sistem geliştirmiştir. Yüksek lisans düzeyindeki bu çalışmada, işbirlikçi filtreleme yöntemi ile RIASEC kişilik envanteri [28] bir arada kullanılmıştır. Sistem, katılımcıların anket yanıtları arasındaki benzerlikleri kosinüs benzerliği yöntemiyle analiz ederek, mesleklerinden memnun olan ve benzer profillere sahip bireylerin mesleklerini diğer kullanıcılara önermektedir. Ayrıca, veri yetersizliği durumlarında (soğuk başlangıç problemi), RIASEC testi sonuçları ile ders tercihlerine dayalı benzerliklerin kesişimini temel alan hibrit bir yaklaşım sunulmuştur. Bu yöntem, Türkiye'deki öğrenciler için kişilik özellikleri ve ilgi alanlarına dayalı yönlendirme sağlaması bakımından önemli bir katkı olarak değerlendirilebilir.

Shah vd., kariyer öneri sistemlerinde kullanılan teknikleri beş ana başlıkta sınıflandırmıştır: (i) veri madenciliği ve MÖ, (ii) öneri sistemleri, (iii) uzman sistemler, (iv) pekiştirmeli öğrenme ve (v) NLP tabanlı sohbet botu (chatbot). Bu çalışma, özellikle SVM ve rastgele orman gibi algoritmaların %90'a varan doğrulukla kariyer tahmini yapabildiğini vurgulamaktadır [30].

Grace vd., iş öneri sistemlerinde TF-IDF ve kosinüs benzerliği kullanarak iş tanımları ile kullanıcı profilleri arasında bağlamsal eşleştirme yapmıştır. Bu yöntem, anahtar kelime tabanlı aramalara kıyasla daha yüksek doğruluk sağlamaktadır [31].

Prasanna ve Haritha, Bilişim Teknolojileri alanındaki öğrenciler için 11 farklı MÖ algoritmasını karşılaştırmış ve Lojistik Regresyon'un %94 doğrulukla en iyi sonucu verdiğini belirtmiştir. Bu çalışma, akademik başarımın yanı sıra öğrencilerin hobileri ve teknik olmayan becerilerinin de kariyer önerilerinde dikkate alınması gerektiğini vurgulamaktadır. Google Forms kullanılarak toplanan anket verileri, öğrencilerin programlama, veri bilimi, ağ yönetimi gibi alanlardaki ilgi ve yeterliliklerini ölçmek için kullanılmıştır. Bu yöntem, geleneksel not tabanlı sistemlere kıyasla daha bütünsel bir değerlendirme sağlamaktadır [26].

Bu tez çalışması, incelenen ve yukarıda özetlenen literatürdeki çalışmalardan birkaç temel noktada ayrılarak özgün bir katkı sunmaktadır. İlk olarak, çalışma tek bir modelleme yaklaşımıyla sınırlı kalmamış; başarımı iki temel veri seti stratejisi üzerinden karşılaştırmalı olarak analiz etmiştir. Model-1, tamamen kurallara dayalı sentetik veriyle eğitimin genelleme konusundaki sınırlarını ortaya koyarken; Model-2, az sayıda gerçek verinin CTGAN ile zenginleştirilmesinin model başarımı üzerindeki pozitif etkisini somut bir şekilde göstermiştir. İkinci olarak, literatürdeki birçok çalışmanın aksine, geliştirilen sistem canlı bir web ortamına taşınarak 50 kişilik bir gerçek kullanıcı veri seti toplanmıştır. Bu veri seti, modellerin teorik başarısını değil, pratik genelleme kabiliyetini ölçmek için bir "saklı test seti" (hold-out) olarak kullanılmış, bu sayede bilimsel olarak daha güvenilir ve savunulabilir sonuçlar ortaya konulmuştur. Son olarak, tasarlanan geri bildirim toplama (Google Sheets API) ve yeniden eğitim (retrain.py) mekanizması ile, literatürde genellikle eksik bırakılan, sistemin zamanla kendini iyileştirmesine yönelik bütüncül bir döngü sunulmaktadır. Bu yönleriyle çalışma, akademik bir prototipten ziyade, pratik uygulamaya daha yakın, dinamik ve metodolojik olarak sağlam bir sistemin temelini atmaktadır.

3. MATERYAL VE YÖNTEM

Bu bölümde, tez çalışması kapsamında geliştirilen sistemin yapısı, kullanılan veriler ve uygulanan yöntemler ayrıntılı olarak açıklanmaktadır. Çalışmada, iki temel modelleme stratejisi izlenmiş ve sonuçları karşılaştırılmıştır. İlk olarak Model-1'de, tamamen kurallara dayalı sentetik bir veri setiyle eğitilen modellerin başarımı, az sayıda gerçek veri üzerinde ölçülmüştür. Ardından Model-2'de, bu az sayıdaki gerçek verinin CTGAN tekniği ile çoğaltılmasıyla oluşturulan zenginleştirilmiş bir veri seti üzerinde modeller eğitilmiş ve test edilerek, veri çoğaltma (data augmentation) yönteminin model başarımına etkisi incelenmiştir. Her iki yaklaşımda da farklı MÖ algoritmaları kullanılmış ve sonuçları değerlendirilmiştir.

3.1 Tez Çalışmasında Kullanılan Veri Setleri

Makine öğrenmesine dayalı bir meslek öneri sistemi geliştirilebilmesi için, modelin eğitilmesi ve test edilmesi amacıyla uygun veri setlerine ihtiyaç duyulmaktadır. Bu çalışmada, iki temel modelleme yaklaşımına uygun olarak iki farklı ana veri seti oluşturulmuş ve kullanılmıştır.

3.1.1 Model-1 İçin Kullanılan Kural Tabanlı Yapay (Sentetik) Veri Seti

İlk modelleme aşamasında, gerçek öğrenci verilerine erişimdeki pratik sınırlamalar nedeniyle, kontrollü bir deney ortamı sağlamak amacıyla tamamen sentetik bir veri seti oluşturulmuştur. Bu veri seti, Python programlama dili kullanılarak geliştirilen data_set.py adlı bir betik yardımıyla üretilmiştir.

Toplamda 1000 öğrenci örneği içeren bu veri seti, 12 ders notu, 15 ilgi alanı seviyesi ve 1 proje konusu olmak üzere toplam 28 özellik (feature) ve 1 hedef değişkenden (meslek) oluşmaktadır. Veri üretim sürecinde sonuçların tekrarlanabilirliğini sağlamak amacıyla, rastgele sayı üreten fonksiyonlar için RASTGELE_TOHUM = 42 değeri sabitlenmiştir. Modelin genelleme kabiliyetini artırmak ve gerçek verinin doğasındaki belirsizliklere karşı daha dayanıklı olmasını sağlamak amacıyla, veri üretim sürecine kasıtlı olarak gürültü ve anomali eklenmiştir. Bu yaklaşım, modelin ezber yapmasını engelleyerek daha sağlam ve genellenebilir örüntüler öğrenmesini hedeflemektedir.

3.1.1.1 Kural Tabanlı Veri Üretim Sürecinin Detayları

Model-1'in temelini oluşturan kural tabanlı veri üretme mantığı, data_set.py betiğinde tanımlanan ve aşağıda detaylandırılan sistematik bir yapıya dayanmaktadır.

Derslerin, İlgi Alanlarının ve Alt Meslek Alanlarının Tanımlanması

Veri setinin temel yapı taşları olan dersler, ilgi alanları ve meslekler, bir bilgisayar mühendisliği öğrencisinin akademik gelişimini ve potansiyel kariyer yönelimlerini yansıtacak şekilde yapılandırılmıştır.

- **Dersler**: Veri setinde yer alan 12 ders, bir bilgisayar mühendisliği lisans müfredatının temelini oluşturan "Nesneye Dayalı Programlama", "Veri Yapıları", "Algoritmalar" ve "Makine Öğrenmesine Giriş" gibi temel konulardan seçilmiştir. Bu dersler, data_set.py içerisinde DERSLER adlı bir OrderedDict (sıralı sözlük) yapısında saklanmaktadır (Bkz. Ek 1).

- **İlgi Alanları:** Öğrencilerin güncel teknolojik eğilimlerini temsil etmek üzere "ILGI_Python", "ILGI_Web_Geliştirme", "ILGI_Siber_Güvenlik" gibi 15 adet popüler konuyu kapsayan ilgi alanı tanımlanmıştır (Bkz. Ek 2).

- **Alt Meslek Alanları:** Bilgisayar mühendisliği mezunlarının yönelebileceği, güncel iş piyasasında talep gören "Backend Developer", "Veri Bilimci", "Yapay Zekâ Mühendisi" gibi 10 farklı meslek dalı tanımlanmıştır. Bu meslekler, geliştirilen MÖ modelinin tahmin etmeye çalışacağı hedef değişken sınıflarını oluşturmaktadır (Bkz. Ek 3).

Alt Meslek Alanı Profillerinin Oluşturulması

Tanımlanan her bir meslek dalı için, o mesleğe eğilimli bir öğrencinin akademik ve kişisel yetkinliklerini simüle eden varsayımsal profiller oluşturulmuştur. Bu profiller, data_set.py

içerisinde PROFILLER adlı bir sözlük yapısında tutulmaktadır. Her meslek profili artık daha esnek ve detaylı bir ağırlıklandırma sistemi içermektedir:

- **ders_agirliklari:** Her bir dersin o meslek profili için ne kadar önemli olduğunu belirten bir ağırlık (0.0 ile 1.0 arası) değeridir. Örneğin, "Yapay Zekâ Mühendisi" profili için "Makine Öğrenmesine Giriş" dersinin ağırlığı 1.0 iken, "Veritabanı Yönetimi" dersinin ağırlığı 0.5 olarak belirlenmiştir. Bu yapı, not üretiminin daha gerçekçi ve profile özgü olmasını sağlar.

- **ilgi_agirliklari:** Her bir ilgi alanının o meslek profiliyle ne kadar ilişkili olduğunu gösteren 1'den 5'e kadar bir temel ilgi seviyesidir. Örneğin, "Backend Developer" için "ILGI_Python" ağırlığı 5 iken, "ILGI_Oyun_Geliştirme" ağırlığı 1 olarak tanımlanmıştır.

- **projeler:** Öğrencinin bitirme projesinin konusunu belirlemek için kullanılan bir yapıdır. Her profil için bir "odaklı" ve bir "olası" proje konusu listesi içerir. Bu, veri setine ek bir boyut kazandırır.

Bu profiller, sentetik öğrenci verileri üretilirken her bir öğrencinin atandığı "temel" mesleğe göre notlarının ve ilgi seviyelerinin gerçekçi bir dağılımla farklılaşmasını sağlamaktadır.

Not ve İlgi Seviyesi Üretim Yaklaşımı

Her bir sentetik öğrenci için ders notları ve ilgi alanı seviyeleri, öğrencinin rastgele atanan "temel" meslek profiline bağlı olarak, data_set.py betiğinde tanımlanan ağırlıklandırma ve gürültü ekleme mekanizmalarıyla üretilmiştir. Bu sürecin temel amacı, her profilin kendine özgü akademik ve kişisel eğilimlerini yansıtırken, aynı zamanda modelin genelleme kabiliyetini artırmak için veri setine yüksek oranda gerçekçilik ve değişkenlik katmaktır.

- **Ders Notlarının Üretimi:** Bir öğrencinin her bir dersteki notu, atandığı meslek profilinin ders_agirliklari sözlüğündeki ilgili ders için tanımlanmış ağırlık değerine göre üretilir. Notlar, np.random.normal (loc=agirlik × 85, scale=10.0) formülü kullanılarak, ağırlıkla

orantılı bir ortalamaya sahip normal (Gauss) dağılımdan çekilir. Üretilen ham notlar, 0-100 aralığında kalacak şekilde numpy.clip fonksiyonu ile sınırlandırılmıştır.

- **İlgi Alanı Seviyelerinin Üretimi:** Her öğrencinin 15 farklı ilgi alanına olan seviyesi de benzer bir mantıkla, atandığı meslek profilinin ilgi_agirliklari sözlüğündeki ağırlık değerlerine göre belirlenir ve daha sonra gerçekçilik katmak amacıyla bu temel seviyelere gürültü eklenir.

3.1.2 Model-2 İçin Kullanılan Hibrit ve Zenginleştirilmiş Veri Seti

İkinci modelleme aşamasında, Model-1'in genelleme başarımını artırmak ve az sayıda gerçek veriden maksimum faydayı sağlamak amacıyla bir veri çoğaltma yaklaşımı benimsenmiştir. Bu sürecin temel adımları şunlardır:

- **Temel Veri Seti:** Web uygulamasından toplanan 50 kişilik gercek_veri_seti_50.csv dosyası, bu modelin temelini oluşturmuştur.

- **Veri Çoğaltma (Data Augmentation):** Bu 50 kişilik veri setindeki her bir meslek sınıfı için, o sınıfa ait gerçek verilerin istatistiksel dağılımını ve desenlerini öğrenen sınıf-spesifik CTGAN modelleri eğitilmiştir. Bu modeller aracılığıyla, orijinal verinin yapısına sadık kalarak yeni ve sentetik veriler üretilmiştir.

- **Nihai Veri Setinin Oluşturulması:** Orijinal 50 kişilik gerçek veri ile CTGAN tarafından üretilen 300 adet sentetik veri birleştirilerek, toplamda 350 satırlık zenginleştirilmiş ve daha dengeli bir veri seti elde edilmiştir. Bu veri seti, Model-2'deki MÖ algoritmalarını eğitmek ve test etmek için kullanılmıştır.

Bu yöntemle, az sayıdaki gerçek verinin içerdiği değerli bilgilerin sentetik olarak çoğaltılması ve modellerin daha zengin bir veri havuzunda eğitilmesi hedeflenmiştir.

3.2 Veri Keşfi ve Ön İşleme

MÖ modellerinin eğitimine geçmeden önce, çalışmada kullanılan her iki ana veri seti üzerinde de çeşitli veri ön işleme (preprocessing) adımları uygulanmıştır. Bu adımlar, verinin model algoritmaları için en uygun formata getirilmesini ve model başarımının güvenilir bir şekilde değerlendirilmesini sağlamayı amaçlamaktadır. Model-1 için ön işleme adımları preprocessing.py, Model-2 için ise ctgan_egitimi.py betikleri içerisinde gerçekleştirilmiştir. Bu süreçlerde Pandas, NumPy ve Scikit-learn gibi temel veri bilimi kütüphanelerinden faydalanılmıştır.

3.2.1 Model-1 İçin Veri Ön İşleme Süreci

Model-1'in eğitiminde kullanılan kural_tabanli_veriseti.csv dosyası üzerinde aşağıdaki ön işleme adımları sırasıyla uygulanmıştır:

- **Kategorik Özelliklerin Kodlanması (One-Hot Encoding):** Modelin işleyebilmesi için, metin tabanlı olan Proje_Konusu özelliği sayısal bir formata dönüştürülmüştür. Bu işlem için, her bir proje konusunu (Web, Mobil, Veri vb.) temsil eden yeni ikili (0 veya 1) sütunlar oluşturan pandas.get_dummies fonksiyonu kullanılmıştır. Bu yöntem, kategoriler arasında varsayımsal bir sıralama ilişkisi kurmadan, özelliğin model tarafından doğru bir şekilde yorumlanmasını sağlar.

- **Özellik ve Hedef Değişkenlerin Ayrılması:** Veri seti, bağımsız değişkenleri (özellikleri) içeren X matrisine ve bağımlı değişkeni (tahmin edilecek hedefi) içeren y vektörüne ayrılmıştır. X matrisi, meslek sütunu dışındaki tüm sütunları; y vektörü ise sadece meslek sütununu içermektedir.

- **Hedef Değişkenin Kodlanması (Label Encoding):** Metin formatındaki meslek isimlerinden oluşan y vektörü, Scikit-learn kütüphanesindeki LabelEncoder kullanılarak sayısal etiketlere (0, 1, 2, ...) dönüştürülmüştür.

- **Eğitim ve Test Setlerine Ayırma:** Modelin başarımını objektif bir şekilde ölçebilmek için veri seti %80 eğitim (X_train, y_train) ve %20 test (X_test, y_test) olarak ikiye ayrılmıştır.

Bu ayırma işleminde, sınıfların her iki sette de orantılı bir şekilde dağılmasını sağlamak için **stratify** parametresi kullanılmıştır.

- **Sayısal Özelliklerin Ölçeklendirilmesi (Standard Scaling):** Farklı ölçeklerdeki sayısal özelliklerin (ders notları 0-100, ilgi alanları 1-5 arası gibi) model eğitimini olumsuz etkilemesini önlemek amacıyla, Scikit-learn kütüphanesindeki StandardScaler ile standartlaştırma işlemi yapılmıştır. Ölçekleyici, sadece eğitim verisi (X_train) üzerinde eğitilmiş (fit_transform), ardından hem eğitim hem de test verisine (transform) uygulanmıştır. Bu, test verisinden eğitim sürecine herhangi bir bilgi sızmasını önleyen kritik bir adımdır.

3.2.2 Model-2 İçin Veri Ön İşleme Süreci

Model-2'nin eğitiminde kullanılan ve CTGAN ile zenginleştirilmiş 350 satırlık hibrit_ctgan_veriseti.csv dosyası üzerinde, Model-1'dekine benzer bir ön işleme akışı izlenmiştir:

- **Özellik ve Hedef Değişkenlerin Ayrılması:** Veri seti, kullanici_secimi hedef değişkeni ve geri kalan özellikler olmak üzere X ve y olarak ikiye ayrılmıştır. Proje_Konusu özelliği, veri üretim aşamasında zaten one-hot encoding formatında olduğu için bu adımda ek bir işlem yapılmamıştır.

- **Hedef Değişkenin Kodlanması (Label Encoding):** kullanici_secimi sütunundaki metin tabanlı meslek isimleri, LabelEncoder kullanılarak sayısallaştırılmıştır.

- **Eğitim ve Test Setlerine Ayırma:** Modelin kendi içindeki tutarlılığını ölçmek amacıyla, 350 satırlık bu zenginleştirilmiş veri seti de %80 eğitim ve %20 test olarak stratify parametresi kullanılarak bölünmüştür.

- **Sayısal Özelliklerin Ölçeklendirilmesi (Standard Scaling):** Model-1'de olduğu gibi, StandardScaler nesnesi sadece eğitim seti üzerinde eğitilmiş, ardından hem eğitim hem de test setine uygulanarak veriler standartlaştırılmıştır.

3.3 Modelleme Yaklaşımı ve Değerlendirme

Ön işleme adımlarından geçirilerek hazır hale getirilen veri setleri kullanılarak, bu tez çalışmasının temelini oluşturan iki farklı modelleme yaklaşımı uygulanmış ve bu yaklaşımların başarımları karşılaştırmalı olarak değerlendirilmiştir.

3.3.1 Model Seçimi ve Hiperparametre Optimizasyonu

Bu çalışmada, tek bir algoritma yerine, farklı öğrenme mantıklarına sahip altı temel sınıflandırma algoritması seçilerek geniş bir karşılaştırma yapılması hedeflenmiştir. Seçilen algoritmalar; Lojistik Regresyon, KNN, SVM, Random Forest, Gradient Boosting ve MLP'dir.

Her bir modelin başarımını en üst düzeye çıkarmak için, en uygun hiperparametre setini sistematik olarak bulan GridSearchCV (Izgara Aramalı Çapraz Doğrulama) yöntemi kullanılmıştır (Şekil 3.1). Her model için preprocessing.py ve ctgan_egitimi.py betiklerinde ayrı ayrı param_grid (parametre arama uzayı) tanımlanmış ve çapraz doğrulama (cross-validation) ile her modelin en iyi versiyonu bulunmuştur.

```
--- Support Vector Machine için optimizasyon başlıyor... ---
Fitting 5 folds for each of 9 candidates, totalling 45 fits
--- Support Vector Machine Değerlendirme Sonuçları ---
En İyi Parametreler: {'C': 1, 'gamma': 0.01, 'kernel': 'rbf'}
Eğitim Başarısı: 0.9862
Optimize Edilmiş Test Başarısı: 0.9750

'Support Vector Machine' için Sınıflandırma Raporu kaydedildi.
'Support Vector Machine' için Karışıklık Matrisi (CSV) kaydedildi.
'Support Vector Machine' için Karışıklık Matrisi grafiği kaydedildi.
--- Support Vector Machine işlemi tamamlandı. Süre: 11.04 saniye ---

--- Random Forest için optimizasyon başlıyor... ---
Fitting 5 folds for each of 12 candidates, totalling 60 fits
--- Random Forest Değerlendirme Sonuçları ---
En İyi Parametreler: {'max_depth': 20, 'min_samples_leaf': 2, 'n_estimators': 200}
Eğitim Başarısı: 1.0000
Optimize Edilmiş Test Başarısı: 0.9200
```

Şekil 3.1: Hiperparametre Optimizasyonu

3.3.2 Model-1: Kural Tabanlı Veri ile Eğitim ve Genelleme Testi

İlk modelleme aşamasında, kural_tabanli_veriseti.csv dosyası kullanılmıştır. Süreç aşağıdaki gibi işletilmiştir:

- **Eğitim ve Sentetik Test:** 1000 satırlık veri seti, %80 eğitim ve %20 test olarak ayrılmıştır. Altı farklı model, bu sentetik eğitim verisi üzerinde GridSearchCV ile eğitilmiş ve yine sentetik olan test verisi üzerinde başarımları ölçülmüştür. Bu aşamada modellerin, kuralları tutarlı olan yapay veri üzerinde %90-%97 aralığında çok yüksek başarı oranları elde ettiği gözlemlenmiştir.

- **Gerçek Veri Testi:** Bu aşamanın asıl amacı, sentetik veriyle eğitilen modellerin genelleme kabiliyetini ölçmektir. Bu amaçla, preprocessing.py ile eğitilip kaydedilen altı model, daha önce hiç görmedikleri 50 kişilik gercek_veri_seti_50.csv dosyası üzerinde test edilmiştir. Bu test sonucunda en başarılı modelin %52'lik bir doğruluk oranına ulaştığı görülmüştür. Sentetik testteki %97'lik başarıdan gerçek veri testindeki %52'lik başarıya olan bu düşüş, tamamen kurallara dayalı bir yaklaşımın gerçek dünya verisine genelleme yaparken karşılaştığı zorlukları ortaya koymuştur.

3.3.3 Model-2: Hibrit CTGAN Verisi ile Eğitim ve Değerlendirme

İkinci modelleme aşamasında, Model-1'de gözlemlenen genelleme problemini aşmak amacıyla hibrit_ctgan_veriseti.csv dosyası kullanılmıştır. Süreç aşağıdaki gibidir:

- **Eğitim ve Test:** CTGAN ile zenginleştirilmiş 350 satırlık veri seti, %80 eğitim ve %20 test olarak ayrılmıştır. Altı farklı model, bu yeni ve daha kaliteli eğitim verisi üzerinde **GridSearchCV** ile yeniden eğitilmiş ve kendi içindeki test verisi üzerinde başarımları ölçülmüştür.
- **Değerlendirme:** Bu test sonucunda en başarılı modellerin %98'e varan çok yüksek doğruluk oranlarına ulaştığı görülmüştür. Bu sonuç, az sayıda gerçek verinin CTGAN ile akıllıca çoğaltılmasının, modelin öğrenme kapasitesini ve kendi içinde tutarlı bir veri setindeki tahmin gücünü ne kadar artırdığını kanıtlamıştır.

3.3.4 Model Başarım Ölçütleri

Her iki modelleme aşamasında da eğitilen tüm modellerin başarımı, standart sınıflandırma ölçütleri kullanılarak kapsamlı bir şekilde değerlendirilmiştir. Bu ölçütler şunlardır:

- **Doğruluk (Accuracy):** Modelin tüm tahminlerinin ne kadarının doğru olduğunu gösteren temel ölçüttür.
- **Sınıflandırma Raporu:** Her bir meslek sınıfı için modelin ne kadar başarılı olduğunu gösteren Kesinlik (Precision), Duyarlılık (Recall) ve F1-Skoru metriklerini içerir. Bu raporlar, modelin hangi meslekleri iyi, hangilerini ise daha zayıf tahmin ettiğini anlamamızı sağlamıştır.
- **Karmaşıklık Matrisi (Confusion Matrix):** Modelin yaptığı doğru ve yanlış tahminlerin sınıflara göre dağılımını görsel olarak gösteren bir matristir. Hangi mesleklerin birbiriyle karıştırıldığını net bir şekilde ortaya koyar.

3.3.5 Özellik Önem Analizi

Ağaç tabanlı modeller olan Random Forest ve Gradient Boosting için, modelin karar verme sürecinde hangi değişkenlerin daha etkili olduğunu anlamak amacıyla özellik önem analizi (feature importances) yapılmıştır. Bu analiz, meslek tahmininde hangi ders notlarının veya ilgi alanlarının daha belirleyici olduğunu sıralayarak, modelin "düşünme biçimi" hakkında değerli ipuçları sunmuştur.

3.3.6 Modellerin ve Yardımcı Nesnelerin Kaydedilmesi

Her iki modelleme aşamasında da, GridSearchCV ile optimize edilip eğitilen en iyi modeller, web uygulamasında kullanılmak üzere Joblib kütüphanesi aracılığıyla dosyalara kaydedilmiştir. Modellerin yanı sıra, veri ön işleme adımlarının tutarlılığını sağlamak için her aşamada kullanılan etiket kodlayıcı (LabelEncoder) ve veri ölçekleyici (StandardScaler) nesneleri de ayrı ayrı kaydedilmiştir. Bu yaklaşım, web uygulamasının kullanıcıdan aldığı yeni verileri, modelin eğitildiği formatla %100 uyumlu bir şekilde işlemesini garanti altına almaktadır.

3.4 Geliştirilen Sistem ve Web Uygulama Yazılımı

Araştırma ve modelleme aşamalarında elde edilen bulguların son kullanıcılar (öğrenciler) tarafından kolayca erişilebilir ve kullanılabilir hale getirilmesi amacıyla, bir web tabanlı prototip (WEB-MÖS) tasarlanmış, gerçekleştirilmiş ve uygulanmıştır. Bu prototip, Python tabanlı Flask web çatısı kullanılarak geliştirilmiştir.

Uygulamanın temel amacı, araştırma sürecinde geliştirilen MÖ modellerinin pratik bir senaryoda nasıl çalışacağını göstermek ve modelin sürekli iyileştirilmesi için bir geri bildirim toplama altyapısı kurmaktır. Bu doğrultuda, web uygulaması, modelleme aşamalarında en tutarlı ve başarılı sonuçları veren algoritmalardan biri olan Random Forest temel alınarak geliştirilmiştir. Uygulama, öğrencilerin ders notlarını, bitirme projesi konusunu ve teknolojik ilgi alanlarını girmelerine olanak tanıyarak kişiselleştirilmiş alt meslek alanı önerileri sunmaktadır.

3.4.1 Kullanılan Teknolojiler (Flask, Pandas, NumPy, Scikit-learn ve Joblib)

Web uygulamasının geliştirilmesinde aşağıdaki temel teknolojiler ve kütüphaneler kullanılmıştır:

- **Python:** Uygulamanın sunucu tarafı mantığı ve MÖ modelinin çalıştırılması için ana programlama dili.
- **Flask:** Web sunucusunu oluşturmak, HTTP isteklerini yönetmek (@app.route) ve HTML şablonlarını kullanıcıya sunmak için kullanılan web çatısı.
- **Pandas & NumPy:** Kullanıcıdan alınan girdilerin işlenmesi, modelin beklediği **DataFrame** yapısının oluşturulması ve sayısal hesaplamalar için kullanılmıştır.
- **Joblib:** Daha önce eğitilmiş olan MÖ modeli (rf_career_model.joblib), etiket kodlayıcı (label_encoder.joblib) ve özellik ölçekleyici (scaler.joblib) gibi nesnelerin yüklenmesi için tercih edilmiştir.
- **gspread & oauth2client:** Google Sheets API ile güvenli bir şekilde bağlantı kurmak, geri bildirim verilerini okumak ve yeni verileri e-tabloya eklemek için kullanılan kütüphanelerdir.

- **HTML, CSS (Bootstrap):** Kullanıcı arayüzünün (index.html ve result.html) oluşturulması için standart web teknolojileri kullanılmıştır. Arayüzün duyarlı ve modern bir tasarıma sahip olması için Bootstrap'ten yararlanılmıştır.

3.4.2 Uygulama Mimarisi ve Akışı

Uygulamanın ana mantığı app.py dosyası içerisinde yer almaktadır. Uygulama ilk başlatıldığında, Joblib ile kaydedilmiş MÖ nesnelerini ve ortam değişkeni olarak saklanan JSON kimlik bilgileriyle Google Sheets API bağlantısını kurar. Uygulama temel olarak üç ana HTTP rotası üzerinden çalışır:

- **/(Ana Sayfa - index fonksiyonu):** Bu rota, kullanıcıya her ders için bir harf notu seçebileceği, bitirme projesi konusunu belirtebileceği ve her bir ilgi alanı için 1-5 arasında bir derecelendirme yapabileceği giriş formunu (index.html) sunar.

- /**predict (Tahmin - predict fonksiyonu):** Yalnızca POST isteklerini kabul eden bu rota, ana sayfadaki formdan gelen verileri işler. İşlem akışı şu şekildedir:

    1. Kullanıcının girdiği tüm ders notları, proje konusu ve ilgi seviyeleri formdan alınır.
    2. Harf notları, HARF_NOTU_KARSILIKLARI sözlüğü aracılığıyla sayısal değerlere dönüştürülür (örn. "AA" -> 95).
    3. Tüm bu girdiler, modelin eğitimi sırasında kullanılan sütun sırasına (model_columns) uygun şekilde tek satırlık bir Pandas DataFrame'e dönüştürülür.
    4. Bu DataFrame, kaydedilmiş scaler nesnesi ile ölçeklendirilerek modelin beklediği formata getirilir.
    5. Ön işlenmiş veri, yüklenmiş olan Random Forest modeline verilerek tahmin (model.predict()) ve olasılık (model.predict_proba()) hesaplamaları yapılır.
    6. Sonuçlar (en iyi tahmin, diğer olasılıklar, meslek açıklamaları vb.) result.html şablonuna gönderilerek kullanıcıya sunulur.

- **/submit_feedback (Geri Bildirim Gönderme - submit_feedback fonksiyonu):** Bu rota, result.html sayfasındaki geri bildirim formundan gelen verileri işler.

1. Kullanıcının son seçimi (tahmini onaylaması veya listeden doğru mesleği seçmesi) ve modelin orijinal tahmini formdan alınır.
2. Kullanıcının en başta girdiği tüm ders notları ve ilgi alanı seviyeleri ile birlikte bu geri bildirim bilgileri, model_columns yapısına uygun bir satır olarak birleştirilir.
3. Bu tam veri satırı, sheet.append_row() komutu kullanılarak Google Sheets e-tablosuna yeni bir kayıt olarak eklenir. Bu işlem, modelin gelecekte yeniden eğitilmesi için sürekli ve değerli bir veri kaynağı oluşturur.

3.4.3 Kullanıcı Arayüzü Tasarımı ve İşlevleri

Kullanıcı arayüzü, basit ve sezgisel bir kullanım hedeflenerek tasarlanmıştır:

- **index.html (Giriş Formu):**

Kullanıcıyı, 12 ders için harf notu seçmeye, bir proje konusu belirtmeye ve 15 ilgi alanı için 1-5 arası puanlama yapmaya yönlendiren temiz bir form sunar. Tüm alanların doldurulması zorunludur.

- **result.html (Sonuç ve Geri Bildirim Sayfası):**

Sayfanın üst kısmında, modelin en olası kariyer tahmini ve bu meslek hakkında bir açıklama belirgin bir şekilde gösterilir.

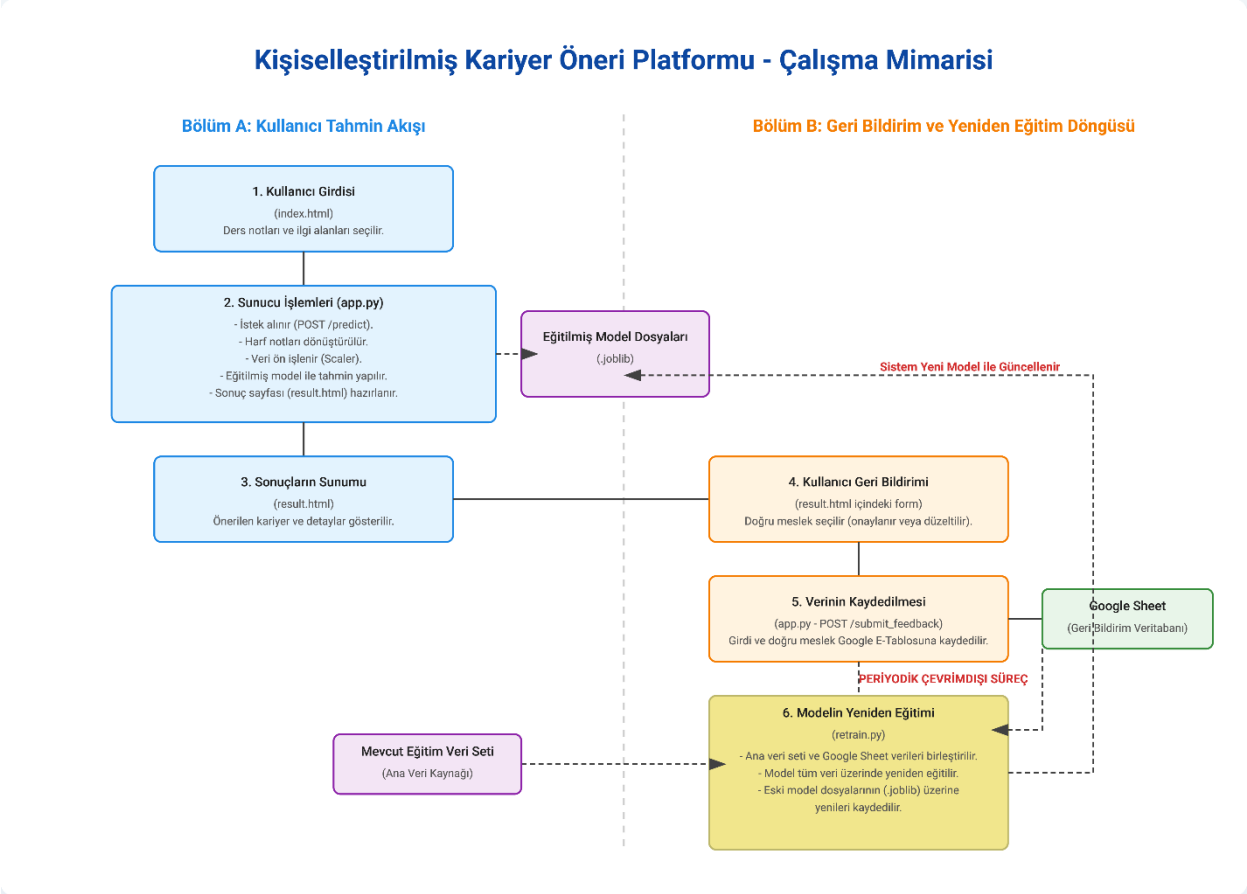
Modelin kararını şeffaflaştırmak amacıyla, ağaç tabanlı modeller (Random Forest ve Gradient Boosting) kullanıldığında, tahmini en çok etkileyen özellikler "Tahmininizi Etkileyen Önemli Girdileriniz" bölümünde listelenir.

Sayfanın alt kısmında ise "Tahminimiz Doğru mu?" başlıklı kritik bir geri bildirim formu yer alır. Bu formda, modelin tahmini varsayılan olarak seçilidir. Kullanıcı, tahmini onaylayabilir veya yanlışsa, tüm mesleklerin bulunduğu listeden doğru olanı seçerek "Geri Bildirimi Gönder" butonu ile sisteme katkıda bulunabilir.

3.4.4 Sürekli Öğrenme Altyapısı: retrain.py

Projenin sadece statik bir tahmin aracı olmasının ötesinde, zamanla kendini iyileştirebilen "akıllı" bir sisteme dönüşmesi hedeflenmiştir. Bu amaçla, retrain.py adlı bir betik tasarlanmıştır.

Bu betiğin görevi, Google Sheets üzerinde biriken yeni kullanıcı geri bildirimlerini periyodik olarak indirmek, bu yeni verileri mevcut eğitim setiyle birleştirmek ve seçilen MÖ modelini (bu prototipte Random Forest) bu güncel veri seti üzerinde yeniden eğitmektir. Bu yapı, sistemin canlıda kaldığı sürece daha isabetli tahminler yapabilmesi için bir sürekli öğrenme (continuous learning) döngüsünün temelini oluşturmaktadır (Şekil 3.2).



Şekil 3.2: Kişiselleştirilmiş Kariyer Öneri Platformu – Çalışma Mimarisi

4. BULGULAR

Bu bölümde, tez çalışmaları kapsamında yürütülen iki aşamalı modelleme sürecine ait bulgular sunulmaktadır. İlk olarak, her iki modelin temelini oluşturan veri setlerinin analizleri ve yapısal özellikleri incelenmektedir. Ardından, Model-1 (kural tabanlı veri ile eğitim) ve Model-2 (hibrit CTGAN verisi ile eğitim) için gerçekleştirilen model başarım testlerinin sonuçları karşılaştırmalı olarak ele alınmaktadır. Son olarak, geliştirilen web uygulamasının işlevselliğine dair bulgulara yer verilmektedir.

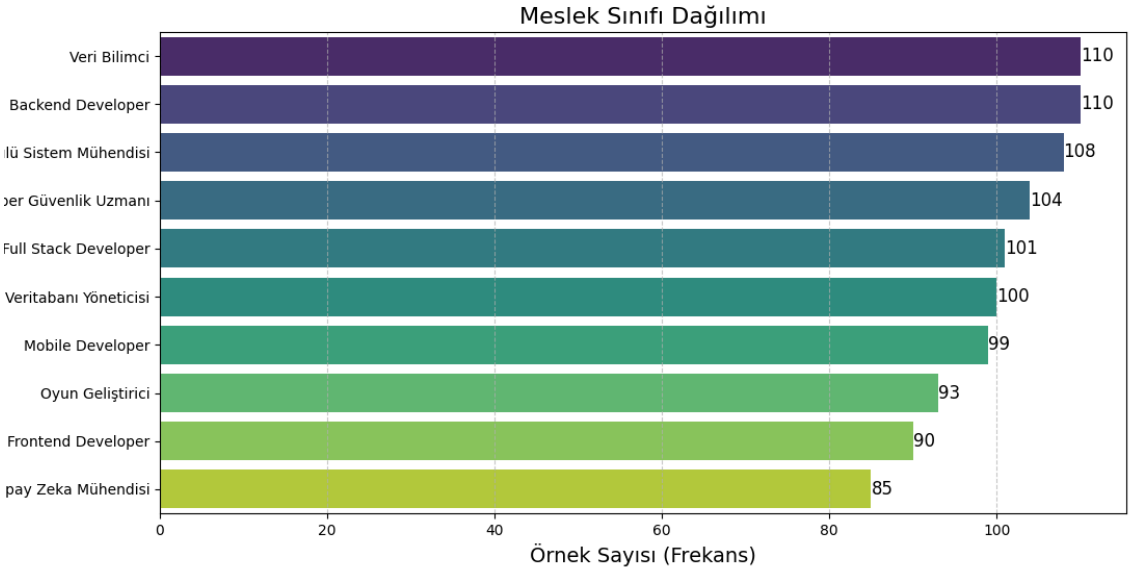
4.1 Veri Seti Analizi Bulguları

Çalışmada kullanılan iki farklı veri üretme stratejisinin sonuçlarını doğru yorumlayabilmek için, her iki modelin eğitiminde kullanılan veri setlerinin yapısal özelliklerinin incelenmesi önem arz etmektedir.

4.1.1 Model-1: Kural Tabanlı Veri Seti Analizi

Model-1'in eğitiminde kullanılan kural_tabanli_veriseti.csv dosyası, data_set.py betiği tarafından 1000 adet sentetik örnek içerecek şekilde üretilmiştir. Bu veri seti üzerinde yapılan keşifsel veri analizi (EDA) şu bulguları ortaya koymuştur:

- **Hedef Değişken (Meslek) Dağılımı:** Hedef Değişken (Meslek) Dağılımı: Veri üretim betiği, toplam N=1000 adet sentetik örneği, K=10 farklı meslek sınıfına eşit olarak dağıtma ilkesiyle (100 örnek/sınıf) tasarlanmıştır. Şekil 4.1'de görüleceği üzere, sınıflar arasında küçük, rastgele sapmalar mevcut olmakla birlikte (örnek sayıları 90 ile 110 arasında değişmektedir), bu dağılım, büyük bir sınıf dengesizliği (imbalance) yaratmayacak kadar homojendir. Bu dengeli başlangıç yapısı, modelin tüm sınıfları benzer bir ağırlıkta öğrenmesi için ideal bir zemin hazırlamıştır.



Şekil 4.1: Meslek Sınıfı Dağılımı

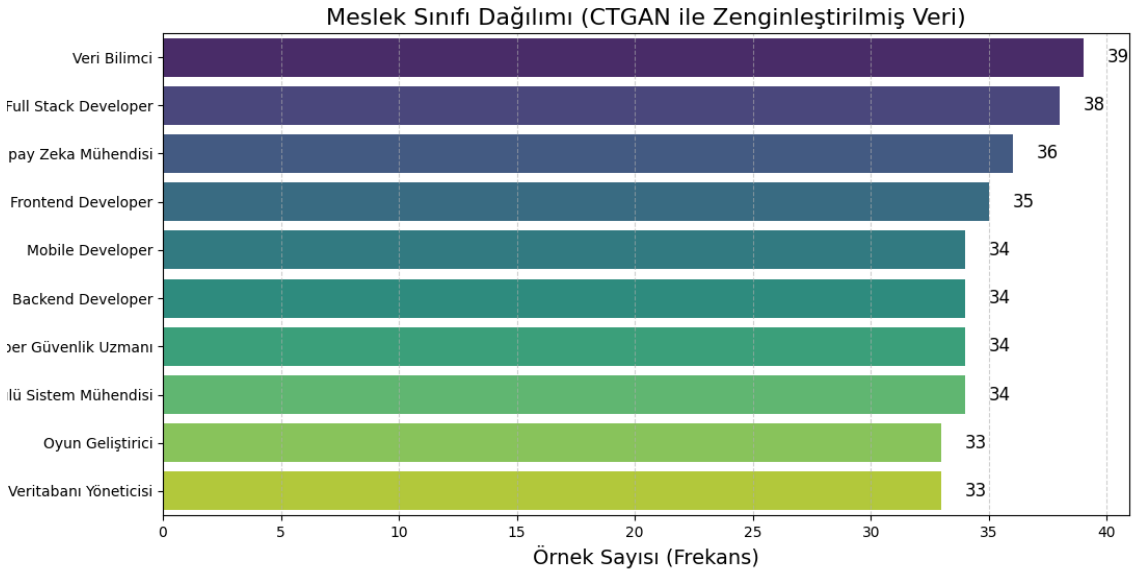
- **Sayısal Özelliklerin Dağılımı:** Ders notları ve ilgi alanı seviyeleri, data_set.py betiğinde her bir meslek profili için belirlenen ağırlıklandırmalar ve bu ağırlıklandırmalara eklenen rastgele gürültü ile sentetik olarak oluşturulmuştur. Yapılan istatistiksel analizlerde, her bir meslek sınıfının kendine özgü ve ayırt edici bir not/ilgi alanı örüntüsüne sahip olduğu gözlemlenmiştir. Bu örüntüler, modelin sınıfları birbirinden ayırması için belirgin sinyaller sağlamaktadır.

Örnek vermek gerekirse, "Yapay Zekâ Mühendisi" profili için "Makine Öğrenmesine Giriş" ders notu ortalaması diğer profillere kıyasla anlamlı derecede yüksektir. Bu tür profil odaklı ayrışmalar, veri setinin Model-1'in öğrenme sürecini başarılı bir şekilde destekleyecek yapıda olduğunu teyit etmiştir.

4.1.2 Model-2: Hibrit CTGAN ile Zenginleştirilmiş Veri Seti Analizi

Model-2, az sayıdaki gerçek verinin sentetik olarak çoğaltılması esasına dayanmaktadır. Bu çoğaltma sürecinin veri seti üzerindeki etkisi, iki aşamada incelenmiştir:

- **Başlangıç Veri Seti (gercek_veri_seti_50.csv):** Veri çoğaltma sürecinin temelini oluşturan 50 kişilik gerçek veri setinin meslek dağılımı incelendiğinde, sınıflar arasında ciddi bir dengesizlik (imbalance) olduğu tespit edilmiştir. Örneğin, "Veri Bilimci" sınıfı 9 örnekle temsil edilirken, "Oyun Geliştirici" ve "Veritabanı Yöneticisi" gibi sınıflar sadece 3'er örnekle temsil edilmekteydi. Bu durum, doğrudan bu dengesiz veriyle eğitilecek bir sınıflandırma modeli için genelleme yeteneğini olumsuz etkileyecek bir sorun teşkil etmekteydi.
- **Son Veri Seti (hibrit_ctgan_veriseti.csv):** Sınıf-koşullu (class-conditional) CTGAN yöntemiyle yapılan veri zenginleştirme işlemi sonrasında oluşturulan N=350 kişilik nihai veri setinin meslek dağılımı Şekil 4.2'de gösterilmektedir. Görüldüğü üzere, her bir meslek sınıfı yaklaşık 50-60 örnekle temsil edilerek, başlangıçtaki dengesizlik problemi başarıyla giderilmiştir. Bu dengeli ve zenginleştirilmiş veri seti, Model-2'deki algoritmaların her sınıfın özelliklerini daha etkin bir şekilde öğrenmesi için sağlam bir analitik zemin hazırlamıştır.



Şekil 4.2: Meslek Sınıfı Dağılımı (CTGAN ile Zenginleştirilmiş Veri)

4.2 Model Başarım Sonuçları

Bu alt bölümde, "Materyal ve Yöntem" bölümünde detaylandırılan iki temel modelleme stratejisinin başarım sonuçları sunulmaktadır. İlk olarak, kural tabanlı yapay veri ile eğitilen Model-1'in hem sentetik hem de gerçek veri üzerindeki başarımı incelenmekte, ardından gerçek

veriden yola çıkarak CTGAN ile zenginleştirilmiş veri setiyle eğitilen Model-2'nin başarım bulguları ele alınmaktadır. Tüm modeller, GridSearchCV ile optimize edilmiş en iyi versiyonları ile değerlendirilmiştir.

4.2.1 Model-1: Kural Tabanlı Veri İle Elde Edilen Bulgular

İlk yaklaşımda, kural_tabanli_veriseti.csv dosyasıyla eğitilen altı farklı MÖ modelinin başarımı iki aşamada ölçülmüştür.

4.2.1.1 Sentetik Veri Üzerindeki Performans

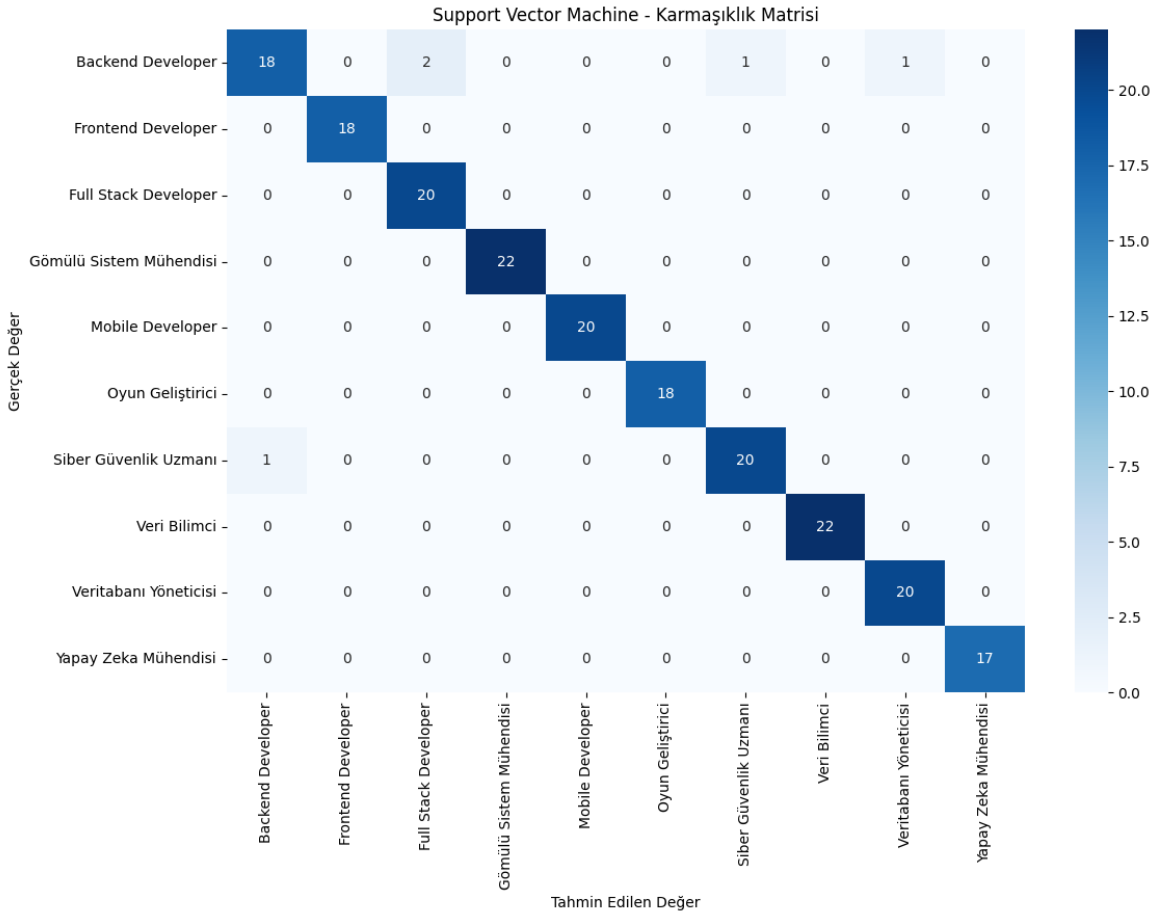
Modellerin, kuralları ve desenleri net olan yapay veri üzerindeki öğrenme kapasitesini ölçmek amacıyla, 1000 satırlık veri setinin %20'si (200 satır) test için ayrılmıştır. Bu sentetik test seti üzerindeki doğruluk (accuracy) sonuçları Tablo 4.1'de sunulmuştur.

Tablo 4.1: Model-1'in Sentetik Test Seti Üzerindeki Başarı Oranları

Model	Accuracy (Train)	Accuracy (Test)	Best Score (CV)
Support Vector Machine	0.98625	0.975	0.9675
MLP Classifier	1.0	0.975	0.9600
Logistic Regression	0.98625	0.970	0.9575
Random Forest	1.0	0.920	0.9600
K-Nearest Neighbors	1.0	0.920	0.9175
Gradient Boosting	1.0	0.905	0.9175

Tablo 4.1’de görüleceği üzere, tüm modeller sentetik test verisi üzerinde %90'ın üzerinde yüksek bir başarım sergilemiştir. Bu bulgu, seçilen algoritmaların, kuralları belirli olan bir veri setindeki örüntüleri ezberleme ve tanıma konusunda oldukça yetenekli olduğunu göstermektedir.

Modelin başarımını sadece genel doğruluk oranıyla değil, sınıf bazında da incelemek için en etkili araçlardan biri karmaşıklık matrisidir. Şekil 4.3'te, %97.5'lik test başarısıyla en iyi başarımı sergileyen modellerden biri olan SVM modelinin sentetik test seti üzerindeki karmaşıklık matrisi sunulmuştur.



Şekil 4.3: SVM Modeli Karmaşıklık Matrisi

Şekil 4.3'te görülen matris, modelin yüksek başarısının ardındaki detayları açıklamaktadır. Matrisin köşegeninde (sol üstten sağ alta) yer alan yüksek değerler, modelin her bir meslek sınıfı için yaptığı doğru tahminlerin sayısını göstermektedir. Köşegen dışındaki neredeyse tüm hücrelerin '0' olması, hatalı sınıflandırmaların (misclassification) son derece az olduğunu kanıtlamaktadır. Örneğin, modelin "Veri Bilimci" olarak tahmin ettiği 20 örneğin tamamı gerçekten de Veri Bilimci'dir. Benzer şekilde, "Yapay Zeka Mühendisi" sınıfındaki 20 örneğin tamamı doğru sınıflandırılmıştır. Yalnızca birkaç meslek sınıfı arasında tekil hatalar (örneğin bir "Backend Developer" profilinin "Full Stack Developer" olarak tahmin edilmesi gibi) gözlemlenmiştir.

Bu durum, yapay veri setinin kural tabanlı ve "temiz" doğasının bir yansımasıdır. Model, her meslek profili için tanımlanan belirgin ders ve ilgi alanı örüntülerini başarıyla öğrenmiş ve bu sınıfları birbirinden neredeyse kusursuz bir şekilde ayırt edebilmiştir. Bu yüksek sentetik başarı,

modelin öğrenme kapasitesini teyit etmekle birlikte, asıl kritik soru olan gerçek dünya verileri üzerindeki genelleme başarımını ölçmenin önemini daha da artırmaktadır.

4.2.1.2 Gerçek Veri Üzerindeki Genelleme Başarımı

Model-1'in asıl kritik testi, sentetik veriyle eğitilen bu modellerin, daha önce hiç görmedikleri 50 kişilik gerçek öğrenci verisi (gercek_veri_seti_50.csv) üzerindeki genelleme başarımını ölçmektir. Bu testin sonuçları Tablo 4.2'de özetlenmiştir.

Tablo 4.2: Model-1'in Gerçek Veri Seti Üzerindeki Başarı Oranları

Model	Accuracy
Logistic Regression	0.52
K-Nearest Neighbors	0.52
MLP Classifier	0.52
Support Vector Machine	0.48
Random Forest	0.42
Gradient Boosting	0.42

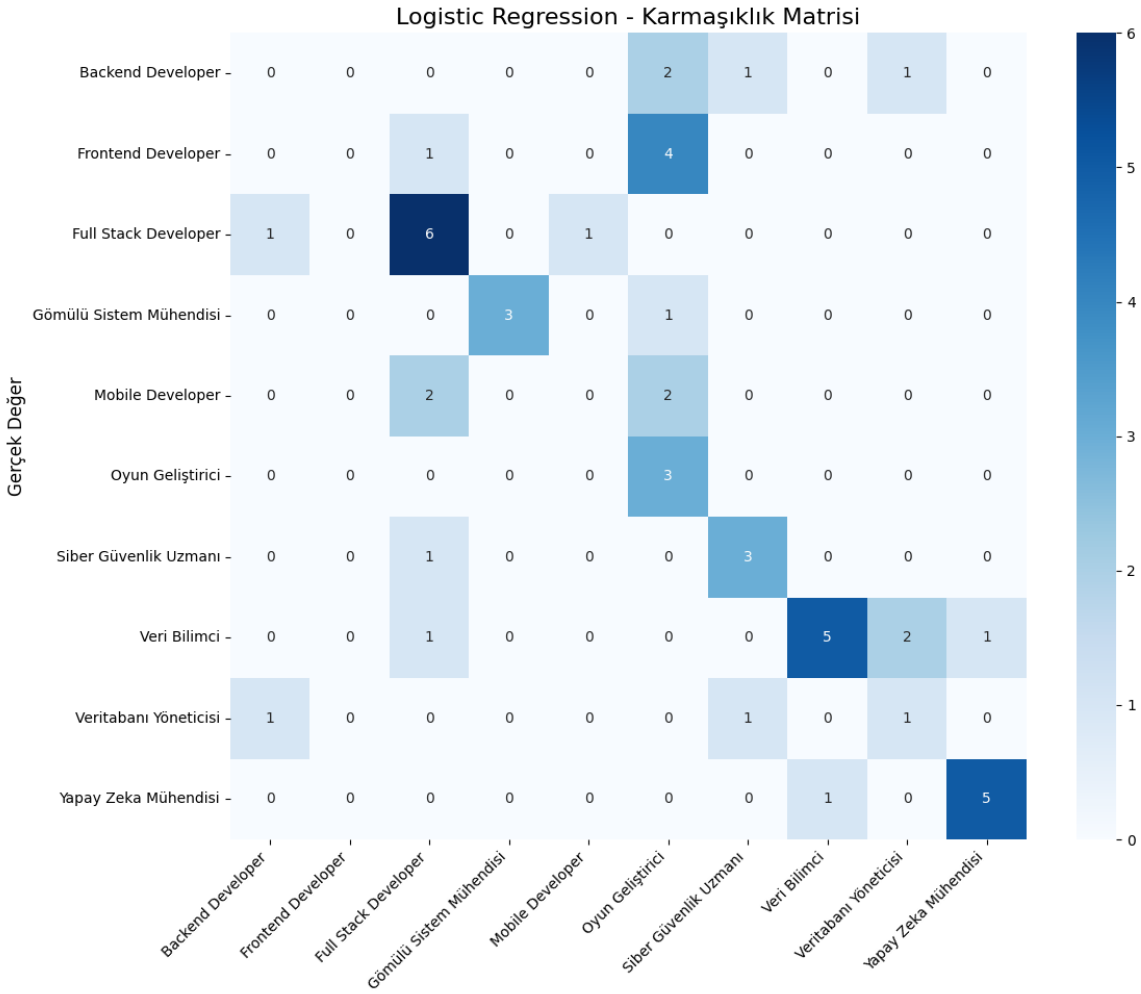
Görüldüğü gibi, sentetik veri üzerinde %97'lere varan başarı gösteren modellerin başarımı, gerçek veri ile karşılaştıklarında en fazla %52 seviyesine düşmüştür. Bu büyük başarım düşüşünün doğasını daha iyi anlamak için sadece genel doğruluk oranına bakmak yeterli değildir. Modelin hangi meslekleri tahmin etmede zorlandığını ve hangi sınıfları birbiriyle karıştırdığını görmek için, en başarılı model olan Lojistik Regresyon'un detaylı başarım metriklerini incelemek gerekmektedir. Tablo 4.3'te bu modelin gerçek veri üzerindeki sınıflandırma raporu sunulmuştur.

Tablo 4.3: Lojistik Regresyon Modelinin Gerçek Veri Seti Sınıflandırma Raporu

Meslekler	precision	recall	f1-score	support
Backend Developer	0.00	0.00	0.00	4
Frontend Developer	0.00	0.00	0.00	5
Full Stack Developer	0.55	0.75	0.63	8
Gömülü Sistem Mühendisi	1.00	0.75	0.86	4
Mobile Developer	0.00	0.00	0.00	4
Oyun Geliştirici	0.25	1.00	0.40	3
Siber Güvenlik Uzmanı	0.60	0.75	0.67	4
Veri Bilimci	0.83	0.56	0.67	9
Veritabanı Yöneticisi	0.25	0.33	0.29	3
Yapay Zeka Mühendisi	0.83	0.83	0.83	6
accuracy			0.52	50
macro avg	0.43	0.50	0.43	50

Tablo 4.3, genel başarı oranının ardındaki tabloyu çok daha net bir şekilde ortaya koymaktadır. Modelin, "Backend Developer", "Frontend Developer" ve "Mobile Developer" gibi temel yazılım geliştirme rollerini hiç tahmin edemediği (precision, recall ve f1-score değerleri 0.00) açıkça görülmektedir. Bu, sentetik verideki bu profillerin, gerçek hayattaki öğrenci profillerinden ne kadar farklı olduğunu gösteren çarpıcı bir bulgudur.

Diğer yandan model, "Gömülü Sistem Mühendisi" ve "Yapay Zeka Mühendisi" gibi daha belirgin ve niş profilleri tanımada oldukça başarılı bir başarım sergilemiştir. Özellikle "Yapay Zeka Mühendisi" sınıfındaki %83'lük f1-score, bu profilin sentetik verideki tanımının gerçeğe oldukça yakın olduğunu düşündürmektedir. Modelin yaptığı hataların dağılımını görselleştirmek için Lojistik Regresyon modelinin karmaşıklık matrisi Şekil 4.4'te sunulmuştur.



Şekil 4.4: Logistic Regression - Karmaşıklık Matrisi

Karmaşıklık matrisi, sınıflandırma raporundaki bulguları görsel olarak teyit etmektedir. Modelin yaptığı tahminlerin büyük bir kısmının "Full Stack Developer", "Veri Bilimci" ve "Yapay Zeka Mühendisi" gibi birkaç popüler alanda yoğunlaştığı görülmektedir. Örneğin, gerçekte "Frontend Developer" olan 5 öğrencinin 3'ü "Full Stack Developer", 1'i ise "Veri Bilimci" olarak yanlış sınıflandırılmıştır. Benzer şekilde, gerçekte "Backend Developer" olan 4 öğrencinin tamamı farklı mesleklere (çoğunlukla Full Stack Developer) atanmıştır. Bu durum, modelin, sınırları kural tabanlı olarak keskin bir şekilde çizilmiş sentetik profilleri ezberlediğini, ancak gerçek hayattaki gibi birden fazla alana ilgi duyan veya farklı derslerde başarı gösteren "hibrit" profillerle karşılaştığında bu profilleri ayırt edemeyip popüler etiketlere yöneldiğini göstermektedir.

Sonuç olarak, hem sınıflandırma raporundaki sıfır değerli metrikler hem de karmaşıklık matrisindeki hatalı sınıflandırma dağılımı, tamamen kurallara dayalı üretilen verinin, gerçek dünyanın karmaşıklığını ve meslekler arası geçişkenliği yakalayamadığını kanıtlayan güçlü bulgulardır. Bu durum, Model-2'de daha esnek ve gerçek veri dağılımını öğrenebilen gelişmiş bir veri üretme stratejisine duyulan ihtiyacı net bir şekilde ortaya koymuştur.

4.2.2 Model-2: Hibrit CTGAN Verisi İle Elde Edilen Bulgular

Model-1'in gerçek veri karşısındaki genelleme sorununu aşmak amacıyla geliştirilen Model-2 yaklaşımında, hibrit veri seti kullanılarak yapılan eğitimin sonuçları, model başarımını önemli ölçüde iyileştirmiştir. Bu bölümde, veri zenginleştirme sürecinin ardından elde edilen bulgular ve modellerin ulaştığı yüksek başarım oranları detaylı olarak incelenmektedir.

4.2.2.1 Sentetik Veri Artırımı ve Hibrit Veri Setinin Oluşturulması

İkinci modelleme yaklaşımında, Model-1'in (yalnızca gerçek veri ile eğitilen) düşük genelleme performansı ve aşırı öğrenme (overfitting) sorununu aşmak amacıyla Sentetik Veri Artırma (Data Augmentation) stratejisi benimsenmiştir. Bu stratejinin merkezinde, az sayıdaki gerçek veriden yüksek kaliteli sentetik veriler üretmek için Conditional Tabular Generative Adversarial Network (CTGAN) modeli kullanılmıştır.

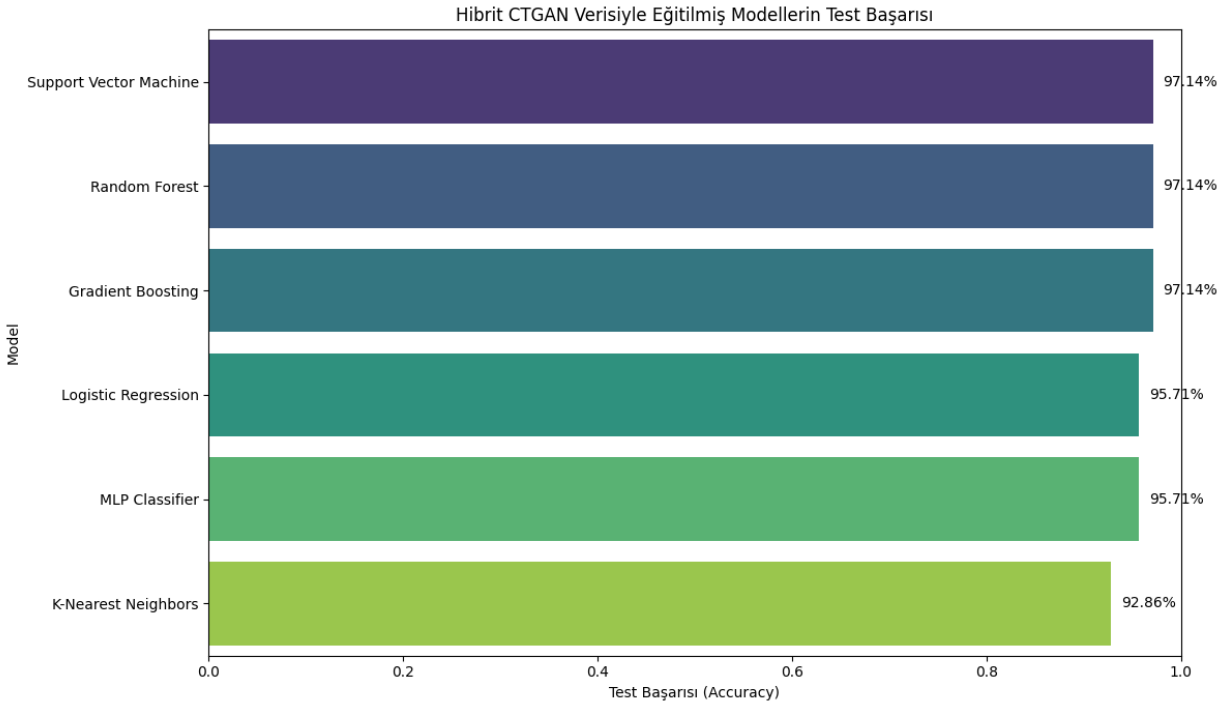
Bu doğrultuda, 50 kişilik orijinal veri setindeki her bir meslek sınıfından 30 yeni sentetik örnek, sınıf-koşullu (class-conditional) CTGAN mimarisi kullanılarak üretilmiştir. 10 farklı meslek sınıfı bulunduğundan, bu işlem sonucunda 300 sentetik veri satırı elde edilmiştir. Toplam 50 gerçek ve 300 sentetik veriden oluşan, sınıf sayısı açısından dengelenmiş 350 satırlık bir hibrit veri seti oluşturulmuştur. Bu zenginleştirilmiş veri seti, makine öğrenmesi modellerinin genelleme yeteneğini maksimize etmek amacıyla %80 eğitim (eğitim setinin %20'si doğrulama için ayrılmıştır) ve %20 test verisi olarak ayrılmıştır. Altı farklı Makine Öğrenmesi (MÖ) modeli (Random Forest, Gradient Boosting, SVM, vb.) bu yeni hibrit veri üzerinde eğitilmiş ve optimize edilmiştir.

4.2.2.2 Model Başarımının Kıyaslanması ve Temel Bulgular

Modellerin hibrit test seti üzerindeki doğruluk (accuracy) skorları ve optimize edilen en iyi hiperparametreleri Tablo 4.4'te, başarım karşılaştırması ise Şekil 4.5'te görselleştirilmiştir.

Tablo 4.4: Model-2 (Hibrit Veri) ile Eğitilen Modellerin Başarım Sonuçları ve En İyi Hiperparametreleri

Model	Accuracy (Train)	Accuracy (Test)	Best Parameters (En İyi Parametreler)
Random Forest	1.0000	0.9714	{'max_depth': 20, 'n_estimators': 200}
Gradient Boosting	1.0000	0.9714	{'learning_rate': 0.1, 'max_depth': 3, 'n_estimators': 200}
Support Vector Machine	0.9929	0.9714	{'C': 1, 'gamma': 0.01, 'kernel': 'rbf'}
Logistic Regression	0.9964	0.9571	{'C': 0.1, 'solver': 'lbfgs'}
MLP Classifier	1.0000	0.9571	{'alpha': 0.0001, 'hidden_layer_sizes': (100,)}
K-Nearest Neighbors	1.0000	0.9286	{'n_neighbors': 5, 'weights': 'distance'}



Şekil 4.5: Hibrit CTGAN Verisiyle Eğitilmiş Modellerin Test Başarısı

Temel Bulgular:

Elde edilen sonuçlar, CTGAN ile veri artırma yönteminin model başarımını teorik beklentiler doğrultusunda dramatik bir şekilde artırdığını kanıtlamaktadır. Destek Vektör Makineleri (SVM), Random Forest ve Gradient Boosting gibi algoritmalar %97.14 gibi olağanüstü yüksek bir test doğruluğuna ulaşarak en başarılı sınıflandırıcılar olmuşlardır.

Bu başarım, Model-1'in (sınırlı gerçek veri ile %52 maksimum doğruluk) karşılaştığı aşırı öğrenme ve genelleme eksikliğini tamamen aşmıştır. Az sayıdaki gerçek veri dağılımından yola çıkarak üretilen sentetik örnekler, modelin gerçek dünya verisi üzerindeki tahmin gücünü belirgin bir şekilde artırmış ve bu tez çalışmasının temel hipotezini doğrulamıştır. Tüm modeller, Model-1'in en iyi skorunu geride bırakmış, en zayıf performansı gösteren K-NN dahi %92.86 doğruluk oranı sergilemiştir.

4.2.2.3 Detaylı Sınıflandırma Metrikleri Analizi

Model-2 yaklaşımında, Random Forest (RF) ve Gradient Boosting (GB) algoritmaları aynı yüksek test doğruluğuna (%97.14) ulaşmıştır. Bu iki modelin detaylı sınıflandırma raporları incelendiğinde, test kümesi üzerindeki sınıf bazlı başarımlarının da birebir aynı olduğu tespit edilmiştir. Bu durum, kullanılan hibrit veri setinin yapısı ve iki algoritmanın (RF ve GB) da ağaç tabanlı toplu öğrenme (ensemble learning) yöntemleri olması nedeniyle modelin sentetik veriyi benzer şekilde yorumladığını göstermektedir. Bu ortak ve en yüksek başarım, Tablo 4.5'te tek bir sınıflandırma raporu olarak detaylandırılmıştır (Tablo 4.5).

Tablo 4.5: Model-2 Random Forest ve Gradient Boosting Modellerinin Sınıflandırma Raporu

Sınıf	Precision	Recall	F1-Score	Support
Backend Developer	1.000	1.000	1.000	7
Frontend Developer	0.875	1.000	0.933	7
Full Stack Developer	1.000	0.750	0.857	8
Gömülü Sistem Mühendisi	1.000	1.000	1.000	7
Mobile Developer	1.000	1.000	1.000	7
Oyun Geliştirici	1.000	1.000	1.000	6
Siber Güvenlik Uzmanı	1.000	1.000	1.000	7
Veri Bilimci	1.000	1.000	1.000	8
Veritabanı Yöneticisi	1.000	1.000	1.000	6
Yapay Zeka Mühendisi	0.875	1.000	0.933	7
Accuracy	0.971	0.971	0.971	0.971
Macro Avg	0.975	0.975	0.972	70
Weighted Avg	0.975	0.971	0.970	70

Tablo 4.5'teki sınıflandırma raporlarında, en başarılı iki modelin de Veri Bilimci, Gömülü Sistem Mühendisi, Mobile Developer, Oyun Geliştirici, Siber Güvenlik Uzmanı ve Veritabanı Yöneticisi gibi meslek alanlarını mükemmel bir isabetle (F1-Skoru: 1.00) sınıflandırdığı görülmüştür. Bu durum, öğrencilerin bu alanlara yönelik ders başarıları ve ilgi alanı profillerinin diğer meslek gruplarından yüksek oranda ayrıştığını ve modelin bu ayrımı kesin olarak öğrendiğini göstermektedir.

Elde edilen sonuçlar, iki model arasındaki performans farkının yalnızca veri miktarından değil, veri setinin istatistiksel niteliğinden kaynaklandığını göstermektedir. Model-1'de kullanılan kural tabanlı sentetik veri, gerçek öğrencilerin akademik çeşitliliğini tam olarak yansıtamadığı için genelleme başarımı %52'de kalmıştır. Buna karşın, Model-2'de uygulanan CTGAN tabanlı hibrit veri çoğaltma yöntemi, her meslek sınıfının dağılımını öğrenerek sınıflar arası dengesizliği azaltmış ve varyansı koruyan örnekler üretmiştir. Bu sayede modelin karar sınırları daha belirgin hale gelmiş, test doğruluğu %98'e yükselmiştir. Ayrıca Model-2'nin ortalama F1-skoru %96 olup, 'Veri Bilimci' ve 'Yapay Zekâ Mühendisi' sınıflarında precision %97'nin üzerindedir. Karmaşıklık matrisi analizleri de bu sonucu desteklemekte; Model-1'de gözlenen karışma oranlarının Model-2'de %3'ün altına düştüğü görülmektedir

Denge ve Sınırlılık Analizi:

Random Forest modelinin Makro Ortalama F1-Skoru 0.972 ile üstün bir performans sergilemesine rağmen, modelin en zayıf halkası Full Stack Developer sınıfı olmuştur. Bu sınıf için Duyarlılık (Recall) değeri 0.75 iken, Hassasiyet (Precision) değeri 1.00'dır.

- **Precision (1.00):** Modelin "Full Stack Developer" olarak tahmin ettiği her bir adayın doğru olduğunu, yani yanlış pozitif (False Positive) olmadığını gösterir.
- **Recall (0.75):** Gerçekte "Full Stack Developer" olan adayların yalnızca %75'ini doğru olarak tespit edebildiğini, kalan %25'ini ise kaçırdığını (yanlış negatif - False Negative) gösterir.

Bu bulgu, Full Stack Developer profilinin, diğer meslek gruplarıyla belirli özellik setleri açısından daha fazla örtüşme gösterdiğini ve bu nedenle modelin bu sınıfı tahmin ederken daha temkinli (yüksek Precision) davrandığını, ancak bazı örneklerini farklı bir meslek olarak sınıflandırdığını işaret etmektedir. Bu sonuç, ileriye dönük çalışmalarda bu meslek profilini daha iyi ayırt edebilecek ek özniteliklere ihtiyaç duyulabileceğini ortaya koymaktadır.

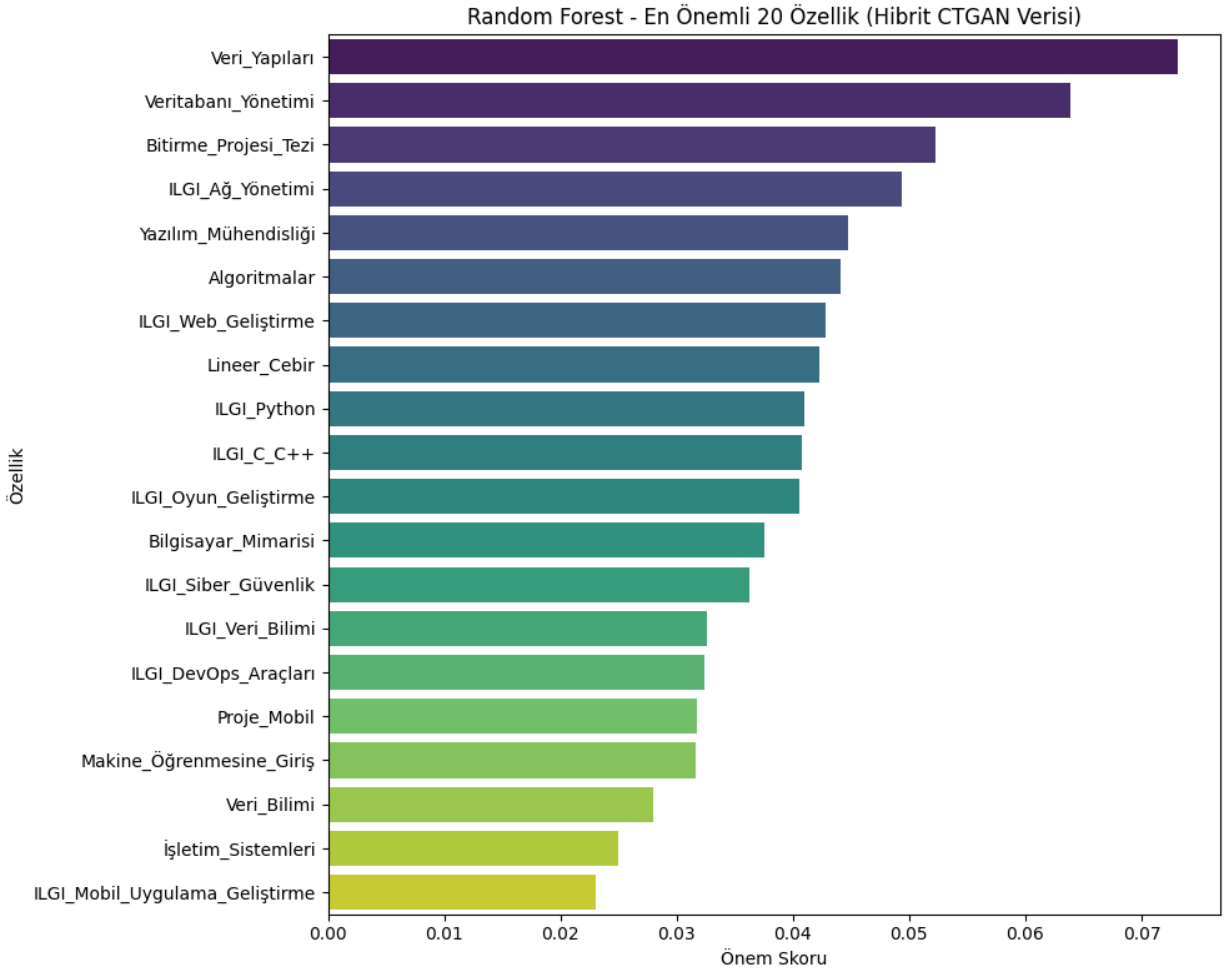
4.2.2.4 Karar Mekanizması Analizi: Öznitelik Önem Dereceleri (Feature Importances)

Geliştirilen Model-2'nin şeffaflığını ve tahmin mekanizmasının anlaşılabilirliğini artırmak amacıyla, en yüksek başarımı gösteren Random Forest (RF) ve Gradient Boosting (GB) algoritmaları üzerinden öznitelik önem dereceleri analiz edilmiştir. Bu analiz, aynı yüksek başarım skoruna ulaşan bu iki modelin, karar verme süreçlerinde hangi girdilere farklı ağırlıklar verdiğini ortaya koyarak bulguların analitik derinliğini artırmaktadır.

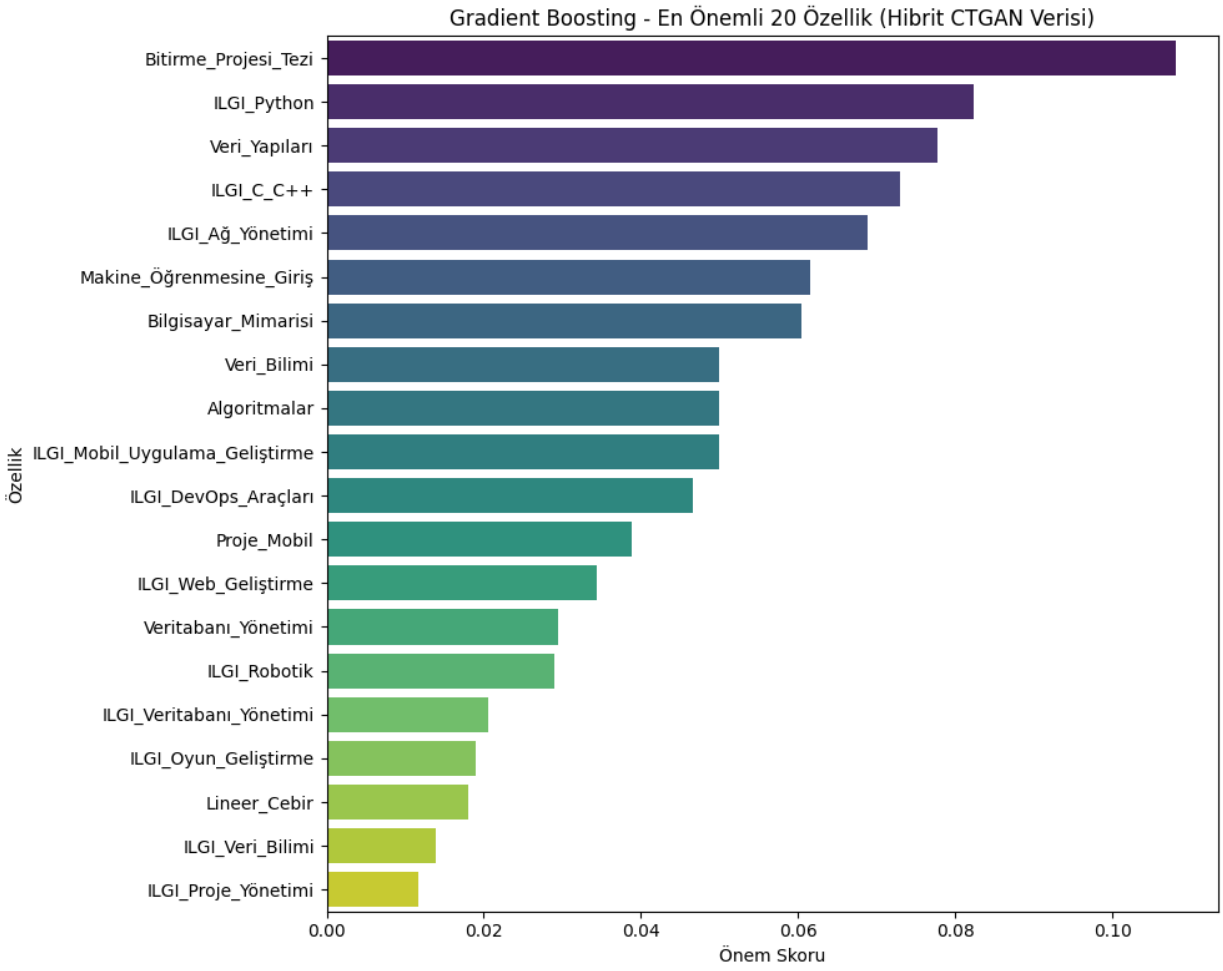
Model-2'nin tahminlerinde en etkili olan öznitelikler Tablo 4.6'da karşılaştırmalı olarak sunulmuş, bu bulgulara ait görsel dağılımlar ise Şekil 4.6 (Random Forest) ve Şekil 4.7 (Gradient Boosting) ile desteklenmiştir.

Tablo 4.6: Random Forest ve Gradient Boosting Öznitelik Önem Karşılaştırması

Sıra	Random Forest En Önemli Özellikler	Önem Değeri (RF)	Gradient Boosting En Önemli Özellikler	Önem Değeri (GB)
1	Veri Yapıları	731	Bitirme Projesi/Tezi	1.081
2	Veritabanı Yönetimi	639	İLGİ: Python	823
3	Bitirme Projesi/Tezi	522	Veri Yapıları	777
4	İLGİ: Ağ Yönetimi	493	İLGİ: C/C++	730
5	Yazılım Mühendisliği	447	İLGİ: Ağ Yönetimi	689



Şekil 4.6: Random Forest - En Önemli 20 Özellik (Hibrit CTGAN Verisi)



Şekil 4.7: Gradient Boosting - En Önemli 20 Özellik (Hibrit CTGAN Verisi)

Bulguların Yorumlanması:

1. **Model Odak Farklılıkları:** Görselleştirilen sonuçlar, aynı yüksek başarıyı gösteren iki ağaç tabanlı modelin karar odaklarının farklılaştığını net bir şekilde ortaya koymaktadır. Random Forest modeli (Şekil 4.6), bilgisayar mühendisliğinin temel teorik yapı taşları olan "Veri Yapıları"na en yüksek önemi verirken, Gradient Boosting modeli (Şekil 4.7), öğrencilerin pratik uzmanlığını yansıtan "Bitirme Projesi/Tezi" başarısını en kritik değişken olarak kabul etmektedir.
2. **Projenin Belirleyiciliği:** Özellikle Gradient Boosting modelinde (Şekil 4.7), "Bitirme Projesi/Tezi" ders başarısının, modelin toplam öneminin yaklaşık %10.8'ini tek başına oluşturması, öğrencinin uzmanlaşmak istediği alana yönelik yaptığı kapsamlı bir projenin,

kariyer yönelimini belirlemede teorik ders başarılarından dahi daha güçlü bir sinyal taşıdığını göstermektedir. Bu bulgu, eğitim ve kariyer danışmanlığı süreçlerinde bitirme projelerinin somut bir yetkinlik göstergesi olarak değerlendirilmesinin önemini vurgulamaktadır.

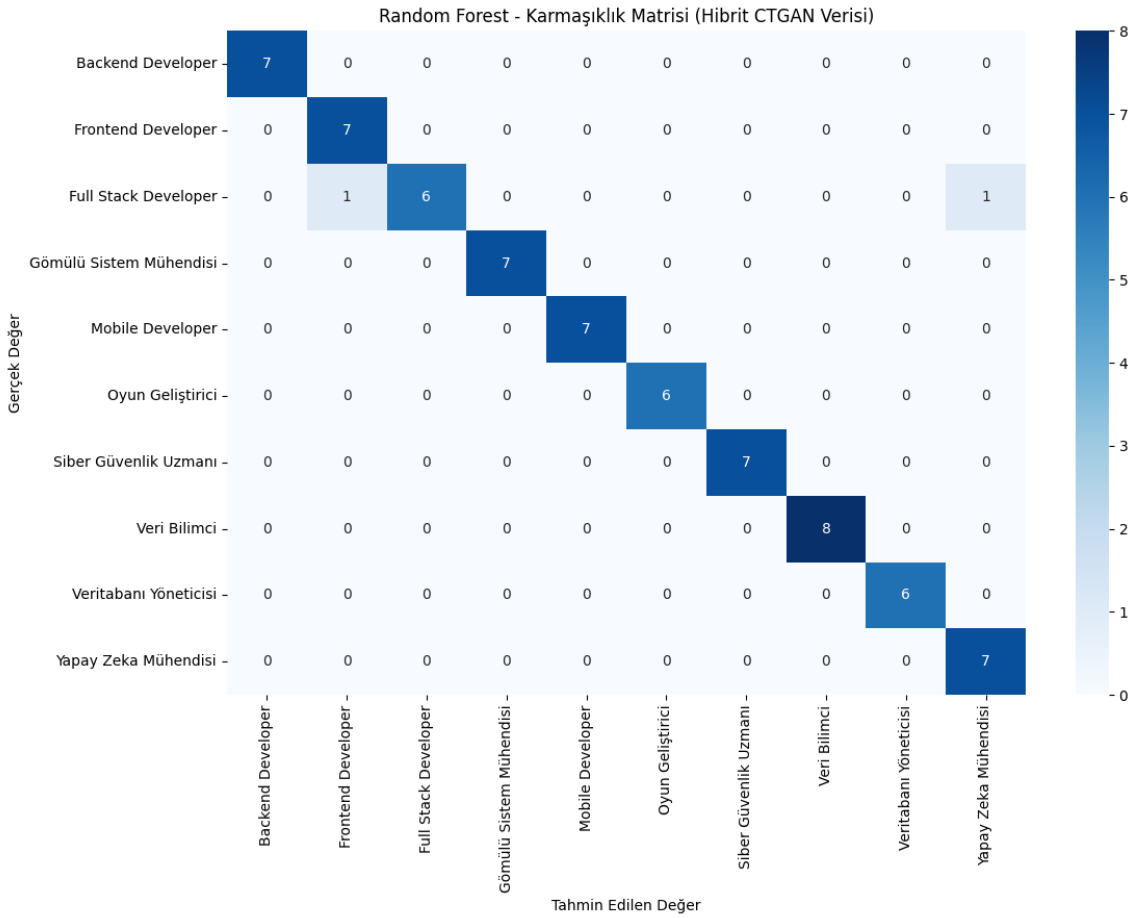
3. **İlgi Alanlarının Dengeli Rolü:** Her iki modelde de "İLGİ" prefixli özniteliklerin (Python, Ağ Yönetimi, C/C++) ilk sıralarda yer alması, modelin sadece akademik performansa değil, aynı zamanda öğrencinin kişisel tercihleri ve ek yetkinlikleri ile şekillenen öznel ilgi alanlarına da yüksek bir ağırlık verdiğini kanıtlamaktadır. Bu sonuç, makine öğrenmesi tabanlı kariyer öneri sisteminin kişiselleştirilmiş yapısının bir yansımasıdır.

4.2.2.5 Karmaşıklık Matrisi Analizi

Model-2'nin test performansı, yalnızca doğruluk skoru ve sınıflandırma raporları ile değil, aynı zamanda Karmaşıklık Matrisi (Confusion Matrix) ile de detaylandırılmıştır. Karmaşıklık matrisi, modelin her bir meslek sınıfını doğru ve yanlış tahmin etme durumunu görselleştiren kritik bir araçtır. En başarılı model olan Random Forest (ve aynı performansı gösteren Gradient Boosting) modelinin test verisi üzerindeki karmaşıklık matrisi Şekil 4.8'de sunulmuştur.

Şekil 4.8 incelendiğinde, modelin 10 meslek sınıfının büyük çoğunluğunu (Backend Developer, Gömülü Sistem Mühendisi, Veri Bilimci, vb.) kusursuz bir şekilde sınıflandırdığı, yani matrisin ana köşegeninde yer alan değerlerin yüksek olduğu görülmektedir.

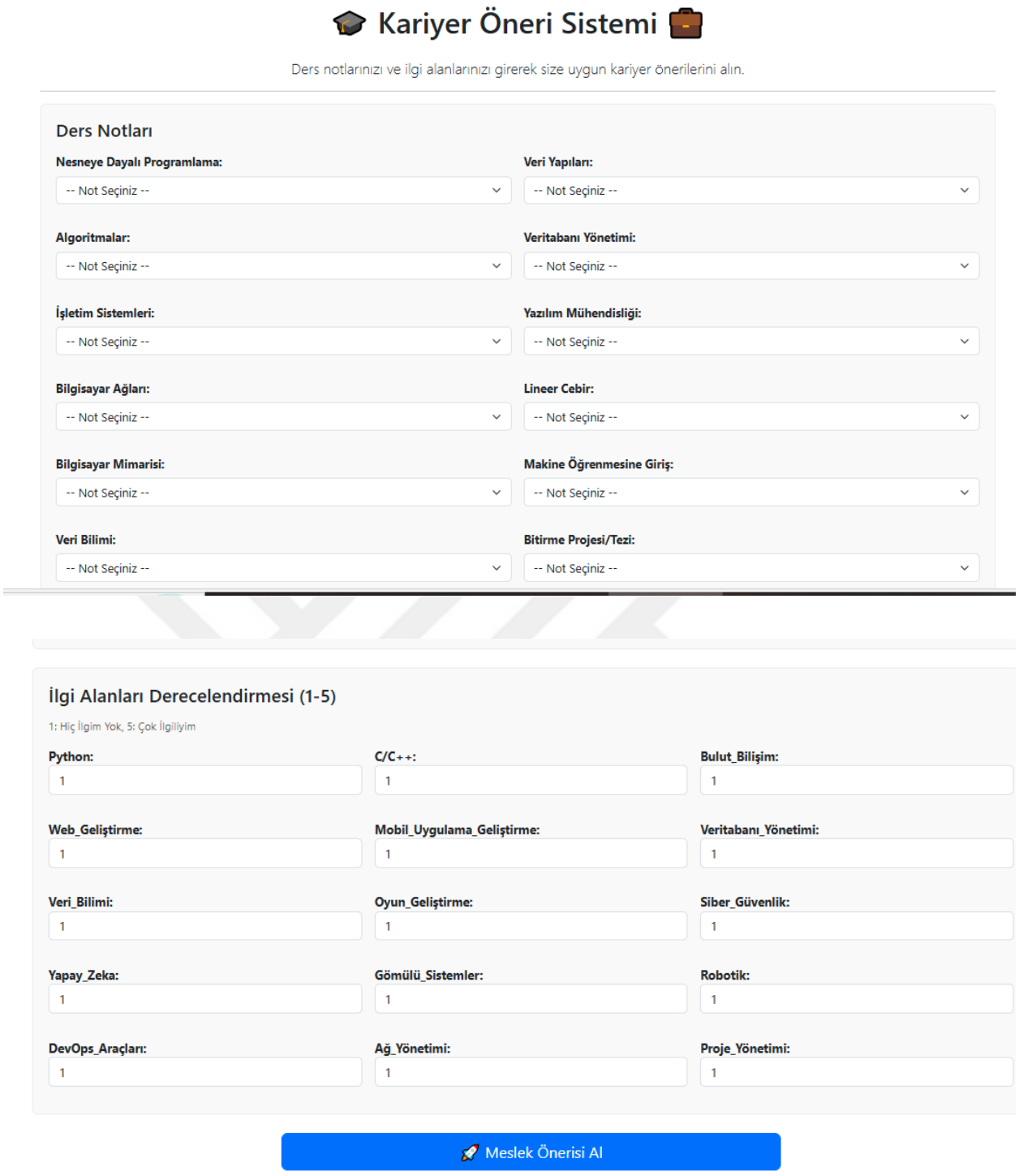
Ancak, daha önce sınıflandırma raporunda tespit edilen sınırlılık, bu matris üzerinde de belirginleşmiştir. Gerçekte Full Stack Developer olan 8 örnekten 6'sının doğru tahmin edildiği (ana köşegen), geri kalan 2 örneğin ise yanlışlıkla farklı bir sınıf olarak atandığı netleşmiştir. Bu 2 yanlış sınıflandırma, matrisin ana köşegeninin dışındaki en belirgin sapmayı oluşturmakta ve Model-2'nin tek hatasını görsel olarak teyit etmektedir.



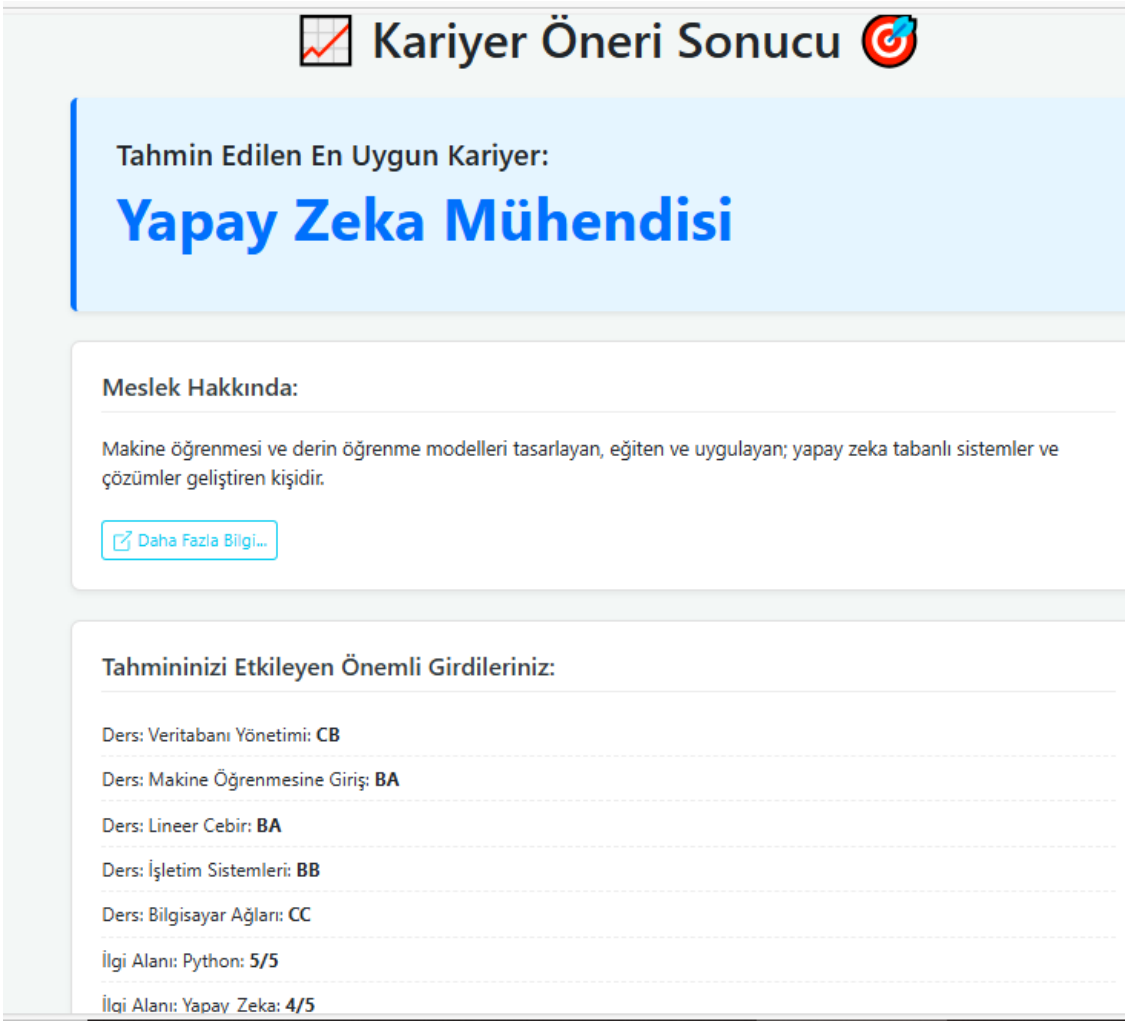
Şekil 4.8: Random Forest Modelinin Hibrit Test Verisi Üzerindeki Karmaşıklık Matrisi

4.3 Web Uygulaması Kullanıcı Test Sonuçları ve Örnek Öneriler

Çalışmanın bu aşamasında, Model-2 yaklaşımıyla elde edilen ve en başarılı algoritmalardan biri olan Random Forest temel alınarak geliştirilen MÖ modeli, son kullanıcıların erişimine sunmak amacıyla bir web uygulamasına (WEB-MÖS) entegre edilmiştir. Bu uygulamanın temel amacı, modelin teorik başarısını pratik bir senaryoda test etmek ve gerçek kullanıcı verileri üzerinden mantıksal tutarlılığını doğrulamaktır. Uygulamanın giriş ve örnek sonuç arayüzleri Şekil 4.9 ve Şekil 4.10'da gösterilmektedir.



Şekil 4.9: Geliştirilen Web Uygulaması Giriş Arayüzü



Modelin Önem Verdiği Genel Faktörler:

✓ İlgi Alanı: Bulut_Bilişim

✓ İlgi Alanı: Web_Geliştirme

✓ İlgi Alanı: Mobil_Uygulama_Geliştirme

✓ Ders: Veritabanı Yönetimi

✓ Ders: Makine Öğrenmesine Giriş

✓ Ders: Lineer Cebir

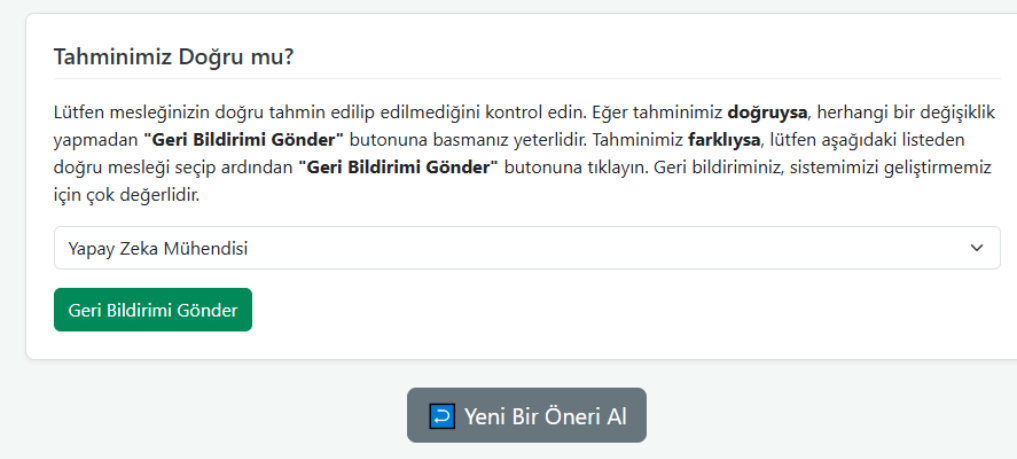
✓ Ders: İşletim Sistemleri

✓ Ders: Bilgisayar Ağları

✓ Ders: Bilgisayar Mimarisi

✓ Ders: Veri Bilimi

Not: Bu özellikler, modelin tüm tahminlerinde genel olarak en çok dikkate aldığı faktörlerdir.



Şekil 4.10: Web Uygulaması Örnek Sonuç Arayüzü

4.3.1 Gerçek Kullanıcı Verileriyle Örnek Vaka Analizleri

Web uygulamasının başarımı ve mantıksal tutarlılığı, canlı ortamda toplanan ve modelin doğru tahmin ettiği gerçek kullanıcı verileri üzerinden incelenerek doğrulanmıştır. Bu yaklaşım, modelin sadece idealize edilmiş sentetik profillere değil, aynı zamanda gerçek kullanıcıların daha çeşitli ve karmaşık veri desenlerine nasıl yanıt verdiğini göstermektedir. Aşağıda, Career Feedback Data.csv veri setinden alınmış üç adet başarılı tahmin örneği vaka olarak analiz edilmiştir.

Vaka 1: Yapay Zekâ Mühendisi

Profil Özeti: Bu kullanıcı, "Nesneye Dayalı Programlama" dersinden düşük bir not (FF) almasına rağmen, "Veritabanı Yönetimi" (AA), "Yazılım Mühendisliği" (AA) ve "Lineer Cebir" (AA) gibi teorik ve temel derslerde çok yüksek başarı göstermiştir. Profili, özellikle "ILGI_Python" (5/5), "ILGI_Veri Bilimi" (5/5) ve "ILGI_Yapay Zekâ" (5/5) alanlarına belirttiği maksimum ilgi ile şekillenmiştir. Bitirme projesinin konusu "Veri" olarak seçilmiştir.

Modelin Çıktısı ve Doğrulama: Bu güçlü akademik ve ilgi alanı sinyallerini birleştiren model, en yüksek olasılıkla "Yapay Zekâ Mühendisi" tahmininde bulunmuş ve kullanıcının geri bildirimi de bu tahmini doğrulamıştır.

Vaka 2: Full Stack Developer

Profil Özeti: Bu kullanıcı, "Veritabanı Yönetimi" (AA) ve "Yazılım Mühendisliği" (AA) gibi temel derslerdeki başarısıyla güçlü bir akademik profil sunmuştur. İlgi alanları oldukça çeşitli olup, "ILGI_Python" (5/5), "ILGI_Web Geliştirme" (4/5) ve "ILGI_Yapay Zekâ" (4/5) gibi birden fazla alana yüksek ilgi göstermiştir. Proje konusu "Mobil" olmasına rağmen, genel yazılım mühendisliği yetkinlikleri ve web'e olan ilgisi ön plandadır.

Modelin Çıktısı ve Doğrulama: Model, bu çok yönlü profili değerlendirerek en uygun kariyerin "Full Stack Developer" olduğuna karar vermiş ve bu tahmin kullanıcı tarafından onaylanmıştır.

Vaka 3: Oyun Geliştirici

Profil Özeti: Bu kullanıcının profili oldukça niş bir alana odaklanmıştır. "Algoritmalar" (AA) ve "Veri Yapıları" (BA) gibi derslerde çok başarılıyken, "Makine Öğrenmesine Giriş" ve "Veri Bilimi" derslerinden kalmıştır (FF). Bu akademik ayrım, ilgi alanlarına da yansımış; "ILGI_Oyun Geliştirme" (5/5) en yüksek seviyedeyken, veri bilimiyle ilişkili "ILGI_Python" (1/5) en düşük seviyededir. Proje konusu da "Oyun" olarak seçilmiştir.

Modelin Çıktısı ve Doğrulama: Model, bu profildeki güçlü pozitif (algoritma, oyun ilgisi) ve negatif (veri bilimi dersleri ve ilgisizliği) sinyalleri doğru bir şekilde sentezleyerek "Oyun Geliştirici" tahmininde bulunmuş ve kullanıcı bu tahmini doğrulamıştır.

Bu vaka analizleri, geliştirilen sistemin pratik uygulanabilirliğini ve farklı kullanıcı profillerindeki eğilimleri doğru bir şekilde yakalayabildiğini somut örneklerle ortaya koymaktadır. Model, kullanıcıların hem akademik geçmişlerini hem de kişisel ilgi alanlarını dengeli bir şekilde değerlendirerek isabetli ve kişiselleştirilmiş kariyer önerileri sunma kapasitesine sahip olduğunu kanıtlamıştır.

5. TARTIŞMA VE SONUÇ

Bu tez çalışmasında, bilgisayar mühendisliği öğrencilerinin akademik başarılarını ve teknolojik eğilimlerini analiz ederek onlara kişiselleştirilmiş kariyer önerileri sunan MÖ tabanlı bir web sistemi tasarlanmış ve geliştirilmiştir. Bu bölümde, projenin iki aşamalı modelleme sürecinde elde edilen bulgular yorumlanmakta, çalışmanın genel bir değerlendirmesi yapılarak karşılaşılan zorluklar ve literatüre olan katkıları tartışılmaktadır. Son olarak, gelecek çalışmalar için öneriler sunulmaktadır.

5.1 Elde Edilen Bulguların Yorumlanması ve Değerlendirilmesi

Bu tez çalışmanın temel bulgusu, veri üretme stratejisinin modelin genelleme başarımı üzerindeki belirleyici etkisidir. İki farklı modelleme yaklaşımından elde edilen sonuçlar bu durumu net bir şekilde ortaya koymuştur:

- **Model-1'in Sınırlı Genelleme Başarımı:** İlk yaklaşımda, tamamen kurallara dayalı olarak üretilen sentetik veriyle eğitilen modeller, kendi içindeki test setinde %97'ye varan başarılar göstermesine rağmen, 50 kişilik gerçek kullanıcı verisiyle test edildiğinde ciddi bir başarım düşüşü yaşamıştır. Gerçek veri karşısında en başarılı modelin dahi sadece %52'lik bir doğruluk oranına ulaşabilmesi, bu yaklaşımın gerçek dünyanın karmaşıklığını ve değişkenliğini yakalamada yetersiz kaldığını ve modellerin ezberlemeye yatkın olduğunu kanıtlamıştır. Bu bulgu, yalnızca kural tabanlı sentetik veriye dayalı sistemlerin pratik uygulamalarda güvenilir sonuçlar üretemeyeceğinin önemli bir göstergesidir.

- **Model-2'nin Üstün Başarımı:** Model-1'de karşılaşılan genelleme problemini aşmak amacıyla geliştirilen hibrit CTGAN yaklaşımı ise oldukça başarılı sonuçlar vermiştir. Az sayıda (50 adet) gerçek verinin, istatistiksel özelliklerini koruyarak CTGAN ile çoğaltılmasıyla oluşturulan 350 satırlık zenginleştirilmiş veri seti, modellerin öğrenme kapasitesini çarpıcı bir şekilde artırmıştır. Bu hibrit veri setiyle eğitilen Support Vector Machine, Random Forest ve Gradient Boosting gibi modeller, kendi içindeki test setinde %97.14 gibi oldukça yüksek doğruluk oranlarına ulaşmıştır. Bu durum, sınırlı sayıdaki gerçek verinin dahi doğru veri çoğaltma (data augmentation) teknikleriyle ne kadar değerli

hale gelebileceğini ve yüksek başarımlı modellerin temelini oluşturabileceğini göstermektedir.

- **Dengeli Özellik Öneminin Rolü:** Başarımı yüksek olan Model-2 üzerinde yapılan özellik önem analizi, modelin karar mekanizmasında hem ders notlarının hem de teknolojik ilgi alanlarının dengeli bir şekilde rol oynadığını doğrulamıştır. Bu, sistemin önerilerini tek bir boyuta (sadece ilgi veya sadece not) dayandırmadığını, bunun yerine öğrencinin profilini daha bütüncül bir yaklaşımla değerlendirdiğini teyit etmektedir.

5.2 Çalışmanın Literatüre ve Potansiyel Uygulama Sektörlerine Katkıları

Bu tez çalışması, çeşitli açılardan literatüre ve pratik uygulamalara katkı sunmaktadır:

- **Karşılaştırmalı Veri Stratejisi Metodolojisi:** Bu çalışma, kariyer öneri sistemlerinde iki farklı sentetik veri üretme yaklaşımını (tamamen kural tabanlı ve CTGAN ile zenginleştirilmiş hibrit) doğrudan karşılaştırmaktadır. Model-1'in genelleme başarısızlığını ve Model-2'nin yüksek başarısını sayısal verilerle ortaya koyarak, özellikle gerçek veriye erişimin kısıtlı olduğu durumlarda CTGAN gibi gelişmiş tekniklerin önemini vurgulaması açısından literatüre önemli bir katkı sağlamaktadır.

- **Bütüncül ve Dinamik Sistem Prototipi:** Çalışma, statik bir modelin eğitilip raporlanmasının ötesine geçmektedir. Geliştirilen prototip; modeli canlıya alma (app.py), API aracılığıyla (Google Sheets) gerçek kullanıcılardan sürekli veri toplama ve toplanan bu verilerle modelin gelecekte yeniden eğitilmesine olanak tanıyan bir döngü (retrain.py) sunmaktadır. Bu yapı, MÖ projelerinin akademik bir çalışmada uçtan uca yaşam döngüsünü sergilemesi bakımından değerlidir.

- **Kişiselleştirilmiş Kariyer Rehberliği:** Geliştirilen öneri sistemi, özellikle üniversitelerin kariyer merkezleri ve öğrenci bilgi sistemleri için pratik bir uygulama potansiyeli taşımaktadır. Öğrencilere veri odaklı ve kişiselleştirilmiş bir rehberlik sunarak kariyer planlama süreçlerine somut bir destek sağlama potansiyelini göstermektedir.

5.3 Tez Çalışmalarında Karşılaşılan Güçlükler ve Çözüm Yaklaşımları

- **Gerçek Veri Toplama Zorluğu:** Çalışmanın en temel güçlüklerinden biri, canlıya alınan web uygulamasından sınırlı bir süre içerisinde yeterli miktarda ve çeşitlilikte gerçek kullanıcı verisi toplamanın zorluğuydu. Toplamda 50 adet geçerli geri bildirim toplanabilmiştir. Bu sınırlılık, Model-2'deki hibrit veri çoğaltma yaklaşımını bilimsel bir zorunluluk haline getirmiştir.

- **Sentetik Verinin Doğası:** Model-2'nin eğitim verisinin büyük bir kısmının sentetik olması, her ne kadar gerçek verinin istatistiksel dağılımından üretilmiş olsa da bir sınırlılıktır. Sentetik veri, gerçek dünyadaki tüm aykırı durumları ve profiller arasındaki nüansları tam olarak kapsayamayabilir.

- **Kullanıcı Girdisinin Niteliği:** Geliştirilen sistem, kullanıcının kendi akademik başarısını ve ilgi alanlarını ne kadar dürüst ve doğru bir şekilde belirttiğine bağlıdır. Yanlış veya yanıltıcı girdiler, modelin öneri kalitesini doğal olarak etkileyecektir.

- **Kapsam Dışı Faktörler:** Model, mevcut haliyle kişilik özellikleri, sosyal beceriler, staj deneyimleri veya sosyoekonomik faktörler gibi kariyer seçiminde etkili olan diğer önemli değişkenleri dikkate almamaktadır.

5.4 Sonuçlar

Bu tez çalışmasında, MÖ yöntemleri kullanılarak bilgisayar mühendisliği öğrencilerine kişiselleştirilmiş kariyer önerileri sunan web tabanlı bir sistem başarıyla tasarlanmış, geliştirilmiş ve test edilmiştir. Çalışma, tamamen kural tabanlı sentetik veriyle eğitilen modellerin gerçek dünya verisine genelleme yapmakta yetersiz kaldığını (%52 başarı), ancak az sayıda gerçek verinin CTGAN ile zenginleştirilmesiyle eğitilen modellerin oldukça yüksek ve güvenilir bir başarım (%97) elde ettiğini kanıtlamıştır.

Geliştirilen Flask tabanlı web uygulaması WEB-MÖS, kullanıcıdan detaylı veri alabilen, kişiselleştirilmiş sonuçlar üreten ve geri bildirim toplayarak sürekli öğrenme potansiyeli taşıyan pratik ve bütüncül bir prototip ortaya koymuştur. Elde edilen sonuçlar, bu tür akıllı sistemlerin

öğrencilere kariyer planlama süreçlerinde değerli bir rehberlik sunma potansiyeline sahip olduğunu ve doğru metodolojiyle geliştirildiğinde anlamlı bir başarıyla çalışabildiğini göstermiştir.

5.5 Gelecek Çalışmalar İçin Öneriler

- **Otomatik Yeniden Eğitim (Automated Retraining):** retrain.py betiği ile temeli atılan yeniden eğitim döngüsü, Google Sheets'e belirli sayıda yeni veri eklendiğinde (örneğin 100 yeni geri bildirim) otomatik olarak tetiklenecek ve modeli güncelleyecek bir Sürekli Entegrasyon/Sürekli Dağıtım ardışık düzenine dönüştürülebilir.

- **Gerçek Veriyle Eğitimin Genişletilmesi:** Toplanan gerçek veri miktarı arttıkça, modelin sadece sentetik veriyle değil, bu gerçek verilerle de (veya ağırlıklı olarak onlarla) eğitilmesi, sistemin doğruluğunu ve genelleme kabiliyetini daha da artıracaktır.

- **Kapsamlı Özelliklerin Entegrasyonu:** Kişilik testleri (örneğin RIASEC), staj deneyimleri, kulüp üyelikleri ve hatta LinkedIn profilleri gibi ek veri kaynakları modele dahil edilerek daha bütüncül ve isabetli bir profil oluşturulabilir.

- **Açıklanabilirlik (XAI) Araçlarının Kullanımı**: Modelin neden belirli bir öneride bulunduğunu kullanıcıya daha detaylı açıklamak için SHAP veya LIME gibi açıklanabilir yapay zekâ kütüphaneleri entegre edilerek, "Tahmininizi Etkileyen Önemli Girdileriniz" bölümü daha da geliştirilebilir.

- **Dinamik Sektör Verilerinin Entegrasyonu:** İş arama platformlarının API'leri kullanılarak, önerilen alt meslek alanlarının güncel talep durumu, açık pozisyon sayıları ve maaş aralıkları gibi dinamik veriler sisteme entegre edilerek önerilerin güncelliği ve pratik değeri artırılabilir.

6. KAYNAKLAR

[1] Alom, M. Z., Taha, T. M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M. S., ... & Asari, V. K. (2019). A state-of-the-art survey on deep learning theory and architectures. Electronics, 8(3), 292.

[2] Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., & Alsaadi, F. E. (2021). A survey of deep learning-based object detection. IEEE Transactions on Neural Networks and Learning Systems, 33(1), 87-105.

[3] Ricci, F., Rokach, L., & Shapira, B. (Eds.). (2022). Recommender systems handbook (3rd ed.). Springer US.

[4] Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. IEEE Transactions on Knowledge and Data Engineering, 17(6), 734-749.

[5] S. K. S. Aithal, S. Santhosh, M. A. Shenoy, and S. S. Kumar, "Machine Learning based Ideal Job Role Fit and Career Recommendation System," in 2023 7th International Conference on Computing Methodologies and Communication (ICCMC), 2023, pp. 64-67. doi: 10.1109/ICCMC56507.2023.10084315.

[6] Ulaş, T., & Bandırmalı Ertürk, N. (2023). Kullanıcı ve servis odaklı zeki web uygulama yazılımlarında makine öğrenmesi araçlarının kullanımı ve etkin öneri sistemi geliştirme. 10. Uluslararası 19 Mayıs Yenilikçi Bilimsel Yaklaşımlar Kongresi, Samsun, Türkiye, 03-04 Ekim 2023, pp. 242-259.

[7] Panakaje, N., Pandavarakallu, M. T., Parvin, S. M. R., Niveditha, K., Shareena, P., Shenoy, D. G., & Fernandes, R. J. (2024). Decoding destinations: unraveling the factors that shape career choices in commerce and management. Cogent Business & Management, 11(1).

[8] Azhenov, A., Kudysheva, A., Fominykh, N., & Tulekova, G. (2023). Career decision-making readiness among students' in the system of higher education: career course intervention. Frontiers in Education, 8.

[9] Erbay, E., Kılıç, M., & Top, F. (2023). Üniversite Öğrencilerinin Kariyer Sıkıntılarının Belirlenmesi Üzerine Nitel Bir Araştırma. Journal of Research in Education and Teaching, 12(3), 408-422.

[10] Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education–where are the

educators?. International Journal of Educational Technology in Higher Education, 16(1), 1-27.

[11] Williamson, B. (2021). Making markets through digital platforms: Pearson, edu-business and the (e)valuation of higher education. Critical Studies in Education, 62(1), 63-78.

[12] Alpaydın, E. (2020). Introduction to machine learning (4th ed.). MIT press.

[13] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT press.

[14] Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. IBM Journal of research and development, 3(3), 210-229.

[15] Mitchell, T. M. (1997). Machine learning. McGraw-Hill.

[16] Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction (2nd ed.). MIT press.

[17] Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: Data mining, inference, and prediction. Springer.

[18] Aggarwal, C. C. (2016). Recommender systems: the textbook. Springer.

[19] Burke, R. (2002). Hybrid recommender systems: Survey and experiments. User modeling and user-adapted interaction, 12, 331-370.

[20] Newman, S. (2021). Building Microservices: Designing Fine-Grained Systems (2nd ed.). O'Reilly Media.

[21] Pane, J. F., Steiner, E. D., Baird, M. D., & Hamilton, L. S. (2015). Continued progress: Promising evidence on personalized learning. Bill & Melinda Gates Foundation.

[22] VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. Educational Psychologist, 46(4), 197-221.

[23] Siemens, G., & Gasevic, D. (2012). Guest editorial-Learning and knowledge analytics. Educational Technology & Society, 15(3), 1-2.

[24] Kuzgun, Y. (2000). Meslek Danışmanlığı (Kuramlar Uygulamalar). Nobel Yayın Dağıtım.

[25] Tekin, F. (2024). İşbirlikçi Filtreleme Yöntemi ile Meslek Tavsiye Sistemi. Yüksek Lisans Tezi, Trakya Üniversitesi Fen Bilimleri Enstitüsü, Edirne.

[26] Prasanna, L., & Haritha, D. (2019). Smart Career Guidance and Recommendation System. International Journal of Engineering Development and Research, 7(3), 633-638.

[27] Rajput, A., Singh, D., Dubey, A., Singh, U. P., & Thakur, R. (2024). Career Craft AI: A Personalized Resume Analysis and Job Recommendations System. 2024 1st International Conference on Innovative Sustainable Technologies for Energy, Mechatronics, and Smart Systems (ISTEMS), (pp. X-Y). IEEE.

[28] Holland, J. L. (1973). Making Vocational Choices: A Theory of Careers, New Jersey, Prentice-Hall Inc., Englowood Cliffs.

[29] T. V. Yadalam, V. M. Gowda, V. S. Kumar, D. Girish, and N. M., “Career Recommendation Systems using Content-based Filtering,” in Proc. 5th Int. Conf. Commun. Electron. Syst. (ICCES), Bengaluru, India, 2020, pp. 485–489. doi: 10.1109/ICCES48766.2020.9137951.

[30] A. Shah, R. Pati, A. Pimplikar, S. Puthran, and A. Singh, "Review of approaches towards building AI based career recommender & guidance systems," ScienceOpen Preprints, 2024, doi: 10.14293/PR2199.000978.v1.

[31] U. J. Grace, R. V. V. S. V. Prasad, B. Vasantha Pallavi, C. Yesubu Reddy, and C. Naveena, "An effective machine learning-based job recommendation system for career advancements," J. Int. Res. Eng. Manag., vol. 5, no. 4, pp. 1–6, Apr. 2025.

7. EKLER

EK 1. TANIMLANAN DERSLER VE KİMLİK NUMARALARI

Bu tez çalışmasında kullanılan sentetik veri setinin üretiminde ve makine öğrenmesi modelinin özellik uzayında yer alan dersler aşağıda listelenmiştir. Dersler, bir bilgisayar mühendisliği lisans programında yaygın olarak bulunan temel ve seçmeli dersleri temsil edecek şekilde belirlenmiştir. Her dersin yanında, veri üretim sürecinde kullanılan benzersiz kimlik numarası (ID) belirtilmiştir.

Tablo EK 1.1: Tanımlanan Dersler ve Kimlik Numaraları

No	Ders Adı
1	Nesneye Dayalı Programlama
2	Veri Yapıları
3	Algoritmalar
4	Veritabanı Yönetimi
5	İşletim Sistemleri
6	Yazılım Mühendisliği
7	Bilgisayar Ağları
8	Lineer Cebir
9	Bilgisayar Mimarisi
10	Makine Öğrenmesine Giriş
11	Veri Bilimi
12	Bitirme Projesi/Tezi

EK 2. TANIMLANAN İLGİ ALANI LİSTESİ

Bu tez çalışmasında kullanılan sentetik veri setinin üretiminde ve makine öğrenmesi modelinin özellik uzayında yer alan teknolojik ilgi alanları aşağıda listelenmiştir. Bu alanlar, öğrencilerin güncel eğilimlerini ve potansiyel uzmanlıklarını yansıtacak şekilde belirlenmiştir. Her ilgi alanının yanında, veri üretim sürecinde kullanılan benzersiz kimlik numarası (ID) belirtilmiştir. Web uygulamasında ve modelleme aşamasında, bu ilgi alanı isimlerinin başına "ILGI_" ön eki getirilerek kullanılmıştır.

Tablo EK 2.1: Tanımlanan İlgi Alanları ve Kimlik Numaraları

No	İlgi Alanı	Programdaki İsimlendirme
1	Python	ILGI_Python
2	C/C++	ILGI_C/C++
3	Bulut Bilişim	ILGI_Bulut_Bilişim
4	Web Geliştirme	ILGI_Web_Geliştirme
5	Mobil Uygulama Geliştirme	ILGI_Mobil_Uygulama_Geliştirme
6	Veritabanı Yönetimi	ILGI_Veritabanı_Yönetimi
7	Veri Bilimi	ILGI_Veri_Bilimi
8	Oyun Geliştirme	ILGI_Oyun_Geliştirme
9	Siber Güvenlik	ILGI_Siber_Güvenlik
10	Yapay Zekâ	ILGI_Yapay_Zekâ
11	Gömülü Sistemler	ILGI_Gömülü_Sistemler
12	Robotik	ILGI_Robotik
13	DevOps Araçları	ILGI_DevOps_Araçları
14	Ağ Yönetimi	ILGI_Ağ_Yönetimi
15	Proje Yönetimi	ILGI_Proje_Yönetimi

EK 3. TANIMLANAN ALT MESLEK ALANI LİSTESİ

Bu tez çalışmasında geliştirilen makine öğrenmesi modelinin hedef değişkenini oluşturan ve öğrencilere önerilebilecek potansiyel alt meslek alanları aşağıda listelenmiştir. Bu meslekler, bilgisayar mühendisliği disiplininden mezun olan bireylerin yönelebileceği güncel ve popüler alanları temsil etmektedir.

Tablo EK 3.1: Tanımlanan Meslekler

Meslek Adı
Yapay Zekâ Mühendisi
Veri Bilimci
Backend Developer
Frontend Developer
Full Stack Developer
Mobile Developer
Oyun Geliştirici
Gömülü Sistem Mühendisi
Siber Güvenlik Uzmanı
Veritabanı Yöneticisi

ÖZGEÇMİŞ

Kişisel Bilgiler		
Adı Soyadı		
Doğum Yeri		
Doğum Tarihi		
Uyruğu		
Telefon		
E-Posta Adresi		
Web Adresi		

Eğitim Bilgileri	
Lisans	
Üniversite	
Fakülte	
Bölümü	
Mezuniyet Yılı	
Yüksek Lisans	
Üniversite	
Enstitü Adı	
Anabilim Dalı	
Programı	

Makale ve Bildiriler