



**T.C.
BATMAN ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
BİLGİ TEKNOLOJİLERİ ANA BİLİM DALI**

YÜKSEK LİSANS TEZİ

**DİJİTAL SINAVLARDA YAPAY ZEKA İLE KOPYA TESPİTİ:
IĞDIR ÜNİVERSİTESİ ÖRNEĞİ**

Derya KARA

**Ocak-2025
BATMAN**

**T.C.
BATMAN ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
BİLGİ TEKNOLOJİLERİ ANA BİLİM DALI**

YÜKSEK LİSANS TEZİ

**DİJİTAL SINAVLARDA YAPAY ZEKA İLE KOPYA TESPİTİ:
IĞDIR ÜNİVERSİTESİ ÖRNEĞİ**

Derya KARA

**Danışman
Doç. Dr. Yılmaz KAYA
Dr. Öğr. Üyesi Fevzi DAŞ**

Diğer Jüri Üyeleri

Prof. Dr. Necmettin SEZGİN Doç. Dr. Melih KUNCAN Doç. Dr. Yılmaz KAYA

**Ocak-2025
BATMAN**

TEZ KABUL VE ONAYI

Derya KARA tarafından hazırlanan “DİJİTAL SINAVLARDA YAPAY ZEKA İLE KOPYA TESPİTİ: IĞDIR ÜNİVERSİTESİ ÖRNEĞİ” adlı tez çalışması 02/01/2025 tarihinde aşağıdaki jüri tarafından oy birliği ile Batman Üniversitesi Lisansüstü Eğitim Enstitüsü Bilgi Teknolojileri Ana Bilim Dalı’nda YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

Jüri Üyeleri

İmza

Başkan

Prof. Dr. Necmettin SEZGİN

.....

Danışman

Doç. Dr. Yılmaz KAYA

.....

Danışman

Dr. Öğr. Üyesi Fevzi DAŞ

.....

Üye

Doç. Dr. Melih KUNCAN

.....

Yukarıdaki sonucu onaylarım.

Dr. Öğr. Üyesi Ömer Murat ÖTER
Lisansüstü Eğitim Enstitüsü Müdürü

ETİK BEYANI

Bu tezdeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edildiğini ve tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her türlü ifade ve bilginin kaynağına eksiksiz atıf yapıldığını beyan eder, aksinin ortaya çıkması durumunda her türlü yasal sorumluluğu kabullendiğimi bildiririm.

ETHICAL DECLARATION

I declare that all the information in this thesis has been obtained within the framework of ethical behavior and academic rules, and that the source of any statements and information that do not belong to me in this study prepared in accordance with the thesis writing rules has been fully cited, and I declare that I accept all kinds of legal responsibility in case of any contrary situation.

Derya KARA
02.01.2025

ÖZET

YÜKSEK LİSANS TEZİ

DİJİTAL SINAVLARDA YAPAY ZEKA İLE KOPYA TESPİTİ: IĞDIR ÜNİVERSİTESİ ÖRNEĞİ

Derya KARA

Batman Üniversitesi Lisansüstü Eğitim Enstitüsü

Bilgi Teknolojileri Ana Bilim Dalı

**Danışman: Doç. Dr. Yılmaz KAYA
Dr. Öğr. Üyesi Fevzi DAŞ**

2025, 50 Sayfa

Bu çalışmada, çevrimiçi sınavlarda kopya çekme davranışlarını tespit edebilmek amacıyla öğrencilerin sınav süreçleri analiz edilmiş ve dört farklı yapay zeka modeli Random Forest, XGBoost, K-Nearest Neighbors (KNN) ve Lojistik Regresyon kullanılmıştır. Çalışmanın temel amacı, öğrencilerin sınav esnasında gösterdiği davranışların benzerliklerini değerlendirerek kopya çekme durumunu pozitif veya negatif olarak sınıflandırmaktır. Analiz sonuçlarına göre, Random Forest modeli %66,29 doğruluk oranıyla en yüksek performansı sergilemiş, onu %58,98 doğruluk oranıyla XGBoost takip etmiştir; KNN ve Lojistik Regresyon ise daha düşük doğruluk oranlarıyla sınıflar arasındaki ayrımı başarılı bir şekilde yapamamıştır. ROC eğrileri incelendiğinde, Random Forest ve XGBoost modellerinin pozitif ve negatif sınıfları ayırt etmede oldukça başarılı olduğu; KNN ve Lojistik Regresyon modellerinin ise bu ayrımında daha zayıf performans gösterdiği görülmüştür. Çalışmada ayrıca özniteliklerin modeller üzerindeki etkisi analiz edilmiştir. "Giriş sayısı", "doğru cevap oranı" ve "yanlış cevap oranı" gibi bazı özniteliklerin modellerin başarısında belirgin bir rol oynadığı tespit edilmiştir. Random Forest ve XGBoost modelleri, belirli özniteliklerin çıkarılması durumunda dahi performanslarını koruyabilirken, KNN ve Lojistik Regresyon modelleri bu tür değişimlere daha duyarlı olmuştur. Çalışmada, performansın artırılmasına yönelik öneriler arasında daha karmaşık derin öğrenme tabanlı modellerin denenmesi, öznitelik mühendisliğinin geliştirilmesi ve veri setinin genişletilmesi gibi stratejiler yer almıştır. Bu bağlamda, Random Forest ve XGBoost modelleri çevrimiçi sınav güvenliğini sağlamak ve kopya çekme vakalarını tespit etmek için en güvenilir ve etkili modeller olarak öne çıkmış olup, sınav güvenliğini artırmaya yönelik stratejiler geliştirilmesinde önemli katkı sağlayabilecekleri sonucuna ulaşılmıştır.

Anahtar Kelimeler: Çevrimiçi Sınav, Kopya Tespiti, Sınav Güvenliği, Yapay Zeka

ABSTRACT

MASTER THESIS

CHEAT DETECTION IN DIGITAL EXAMS WITH ARTIFICIAL INTELLIGENCE: İĞDIR UNIVERSITY EXAMPLE

Derya KARA

Batman University Graduate Education Institute

Department of Information Technologies

Advisor: Assoc. Prof. Dr. Yılmaz KAYA

Asst. Prof. Dr. Fevzi DAŞ

2025, 50 Pages

In this study, artificial intelligence models Random Forest, XGBoost, K-Nearest Neighbors (KNN), and Logistic Regression were employed to analyze students' behavior during online exams to detect potential cheating. The primary goal was to classify behaviors as indicative of cheating (positive) or not (negative) by evaluating similarities in student actions during exams. The analysis revealed that Random Forest achieved the highest performance with an accuracy of 66.29%, followed by XGBoost at 58.98%, while KNN and Logistic Regression demonstrated lower accuracy and struggled to distinguish between classes effectively. According to ROC curve analysis, Random Forest and XGBoost excelled in differentiating positive and negative classes, whereas KNN and Logistic Regression showed weaker performance. Additionally, the impact of individual features on the models was examined, with factors such as "number of logins," "correct answer ratio," and "incorrect answer ratio" found to significantly influence model performance. Random Forest and XGBoost remained robust even when certain features were removed, whereas KNN and Logistic Regression models were more sensitive to such changes. Recommendations to enhance performance include exploring more complex deep learning models, refining feature engineering, and expanding the dataset. Ultimately, Random Forest and XGBoost emerged as the most reliable and effective models for ensuring online exam security and detecting cheating, offering valuable insights for developing strategies to prevent academic dishonesty in remote education platforms.

Keywords: Online Exam, Exam Evaluation, Exam Security, Artificial Intelligence

ÖN SÖZ

Uzaktan öğretim uygulama ve araştırma merkezinde devam eden iş hayatım, bu çalışmamın çıkış noktası olmuştur. Uzaktan eğitimde edinmiş olduğum bilgi ve tecrübelerimi bu çalışmada sunmaktayım. Dünyada olduğu gibi ülkemizde de yaygınlaşan uzaktan eğitimin gelişim sürecine bağlı olarak dijital ortamda gerçekleşen sınavların güvenliğine değinmekteyim.

Yüksek lisans tez çalışmamda, araştırma ve planlama aşamasında bana destek olan, tecrübe ve tavsiyeleriyle bana yol gösteren danışmanım Doç. Dr. Yılmaz KAYA' a, uzaktan eğitim bilgi ve tecrübelerini benimle paylaşan, ne zaman danışsam saatlerce kıymetli vaktini bana ayıran müdürüm aynı zamanda II. danışmanım Dr. Öğr. Üyesi Fevzi DAŞ'a ve her daim yanımda olan eşim Serhat KARA ve annem Leman PAPAKÇI' ya çok teşekkür ederim.

Derya KARA
BATMAN-2025

İÇİNDEKİLER

ÖZET	iv
ÖN SÖZ	vi
İÇİNDEKİLER	vii
ÇİZELGELER/TABLolar LİSTESİ	viii
ŞEKİLLER LİSTESİ	ix
1. GİRİŞ	1
1.1. Çalışmanın Amacı ve Önemi	5
1.2. Çalışmanın Problemleri	7
2. KAYNAK ARAŞTIRMASI	8
3. MATERYAL VE YÖNTEM	20
3.1. Materyal	20
3.1.1. Veri toplama ve işleme	22
3.2. Sınıflandırma Metotları	22
3.2.1. Lojistik regresyon metodu	22
3.2.2. Rastgele orman metodu	24
3.2.3. XGBOOST metodu	25
3.2.4. K-En yakın komşu metodu	27
3.2.5. Performans ölçütleri	29
4. BULGULAR	32
5. SONUÇLAR VE ÖNERİLER	45
5.1 Sonuçlar	45
5.2 Öneriler	46
6. KAYNAKLAR	48

ÇİZELGELER/TABLolar LİSTESİ

Tablo 3.1. Alms sınav loglari	20
Tablo 3.2. Öğrenci sınav istatistikleri	20
Tablo 3.2. Öğrenci sınav istatistikleri tablonun devamı	21
Tablo 3.3. Öznitelik alt kümeleri	21
Tablo 3.4. Karışıklık matrisi örneđi	29
Tablo 4.1. Performans deđerleri	34

ŞEKİLLER LİSTESİ

Şekil 4.1. Tanıtıcı istatistikler	32
Şekil 4.2. Korelasyon matrisi	34
Şekil 4.3. Karışıklık matrisleri	38
Şekil 4.4. Roc eğrileri	40

SİMGELER VE KISALTMALAR DİZİNİ

ADASYN	Adaptive Synthetic Sampling Method for Imbalanced Data (Dengesiz Veriler için Uyarlanabilir Sentetik Örneklem Yöntemi)
AI	Artificial Intelligence (Yapay Zeka)
ALMS	Advancity Learning Management System (Öğrenme Yönetim Sistemi)
AUC	Area under the ROC Curve (ROC Eğrisinin Altındaki Alan)
BİT	Bilgi Teknolojileri
HCI	Human-Computer Interaction (İnsan-Bilgisayar Etkileşimi)
IP	Internet Protocol (İnternet Protokolü)
KNN	K-Nearest Neighbours - K-En Yakın Komşular
NLP	Doğal Dil İşleme
ÖYS	Öğrenme Yönetim Sistemini
ROC	Receiver Operating Characteristic (Alıcı işletim karakteristiği)
SMOTE	Synthetic Minority Oversampling Technique (Sentetik Azınlık Aşırı Örneklem Tekniği)
SVM	Support Vector Machine (Destek Vektör Makineleri)
UZEM	Uzaktan Eğitim Merkezi
YÖDAK	Yükseköğretim Kalite Kurulu
YÖK	Yükseköğretim Kurulu
YZ	Yapay Zeka

1. GİRİŞ

Günümüzde teknolojinin hızlı gelişimi, eğitim alanında radikal değişikliklere yol açmıştır. Bu değişikliklerin en belirgin yansımalarından biri, geleneksel sınıf eğitiminin yerini dijitalleşmiş öğrenme yöntemlerinin ve çevrimiçi sınavların almaya başlamasıdır. Özellikle, Covid-19 pandemisi ve 2023 Kahramanmaraş Depremi gibi olağandışı olaylar, eğitim süreçlerinin sürdürülmesi konusundaki zorlukları ön plana çıkarmıştır. Yaşanan afet dönemleri, çevrimiçi öğrenme ve çevrimiçi sınavlar, eğitim faaliyetlerini kesintisiz bir şekilde sürdürebilme ihtiyacını göstermiştir (Telli ve Altun, 2023). Bu ihtiyaçların giderilmesi esnasında, çevrimiçi sınavların avantajları ve dezavantajları aşağıdaki şekilde ortaya çıkmıştır:

Avantajları:

Erişilebilirlik: Çevrimiçi sınavlar, coğrafi konumdan bağımsız olarak öğrencilere erişim sağlar. Bu durum, daha fazla öğrencinin eğitime katılmasına olanak tanır.

Otomasyon ve Hız: Çevrimiçi sınavlar, otomatik değerlendirme özellikleri ile sonuçları hızla üretebilir. Bu, öğrencilere daha hızlı geri bildirim verilmesini sağlar.

Çeşitli Soru Formatları: Çevrimiçi sınavlar, çok çeşitli soru formatlarını destekler, bu da öğrencilerin farklı yetenekleri ve öğrenme stillerini değerlendirmeye olanak tanır.

Dezavantajları:

Etik Olmayan Davranışlar: Çevrimiçi sınavlar, öğrenciler arasında etik olmayan davranışlar, özellikle hile ve kopya çekme riskini artırabilir.

Teknik Sorunlar: Teknik sorunlar, öğrencilerin sınav esnasında sıklıkla karşılaşabileceği sorunlardır. İnternet bağlantısı kesintileri veya sınav platformu sorunları gibi faktörler, öğrencilerin performansını olumsuz etkileyebilir.

Kopya Tespiti Hassasiyeti: Kopya tespiti yazılımları, yanlış pozitif veya yanlış negatif sonuçlar üretebilir. Bu, masum öğrencilerin haksız yere suçlanmasına veya gerçek kopya çekenlerin gözden kaçmasına neden olabilir.

Bu dönüşümün getirdiği avantajlar, çevrimiçi sınavlarda karşılaşılan dezavantajları da beraberinde getirmektedir. Özellikle, sınav güvenliğini sağlamak için kopya tespiti ve hassasiyeti önem arz etmektedir. Öğrencilerin çevrimiçi sınavlar sırasında etik olmayan davranışlar sergileme potansiyeli daima mevcuttur. Bu sebeple,

eđitim kurumları çevrimiçi sınavlar için güvenlik önlemleri ve kopya tespiti mekanizmaları geliřtirmek zorundadır.

Dijitalleşen eğitimde çevrimiçi sınav güvenliđi, öğrencilerin kopya çekmek gibi etik olmayan davranışlarda bulunma potansiyeli nedeniyle geleneksel sınavlara kıyasla daha karmaşık bir konudur. Uzaktan eğitimde sınavlar çevrimiçi ortamda yapıldığı için, öğrencilerin sınavda dürüstlüklerini denetlemesi ve sınavda kopya çekme gibi hilelerin önlenmesi daha zor olabilmektedir. Bu bağlamda, çevrimiçi sınav güvenliđi için çeşitli teknolojik ve organizasyonel önlemler alınmaktadır. Örneđin, otomatik test denetleyicileri, kamera ve mikrofonlar kullanılarak öğrencilerin sınav sırasında etkin bir şekilde izlenmesi, sınav ortamının kaydedilmesi ve kopya çekmeye karşı uyarılar verilmesi gibi önlemler alınmaktadır. Ayrıca, yapay zeka destekli sınav güvenlik sistemleri, öğrencilerin sınav sürecindeki etkileşimlerini izleyerek potansiyel kopya çekme girişimlerini tespit etmeye çalışmaktadır. Ancak, çevrimiçi sınav güvenliđini tamamen sağlamak zorlu bir görevdir ve her zaman kusursuz değildir. Öğrencilerin çeşitli yöntemlerle kopya çekme girişimleri veya sınavlar sırasında teknik sorunlar yaşanması gibi zorluklar günümüzde mevcuttur. Bu nedenle, uzaktan eğitimde çevrimiçi sınav güvenliđi konusunda sürekli olarak yeni yöntemlerin geliştirilmesi ve güncellemelerin yapılması önemlidir.

Uzaktan eğitimde çevrimiçi sınav güvenliđi; eğitim kurumları, öğretmenler ve teknoloji sağlayıcıları için önemli bir öncelik olmaya devam etmektedir. Öğrencilerin sınav sırasında kopya çekme, sahtekarlık veya diđer hilelerle sınav sonuçlarını manipüle etme olasılığı nedeniyle güvenlik önlemleri önemli bir odak noktasıdır. Çevrimiçi sınav güvenliđi; kimlik doğrulama, sınav süreci izleme, otomatik test denetleyicileri, kopya tespiti, sınav düzenlemeleri, eğitimci denetimi gibi yöntemlerle sağlanabilmektedir. Aynı zamanda eğitim kurumları, güvenlik önlemlerini iyileştirmek ve öğrencilerin akademik dürüstlüđünü korumak için uygun stratejiler geliřtirmektedirler. Bununla birlikte teknoloji sağlayıcıları da çevrimiçi sınav güvenliđi için yeni ve gelişmiş çözümler sunmaya çalışmaktadır. Bu alanda kullanılan yapay zeka ve makine öğrenmesi gibi ileri teknolojiler, çevrimiçi sınavların güvenliđinde etkin bir rol oynamakta ve çevrimiçi sınav güvenliđini artırmak için devamlı olarak yeni güvenlik özellikleri geliřtirmektedir.

Güvenlik önlemlerinden olan kopya tespiti, sınav sonuçlarının değerlendirilmesi aşamasında son derece kritik bir rol oynamaktadır. Tespitlerde yapay zeka ve makine öğrenimi algoritmaları, büyük miktarda veri analizi yapabilme ve örüntüleri tespit edebilme yetenekleri ile çevrimiçi sınavlarda kopya tespiti, sahtekarlık analizi ve diđer

güvenlik önlemlerini güçlendirmeye büyük katkı sağlarlar. Örneğin, yapay zeka tabanlı bir kopya tespit sistemi, öğrencilerin sınav sırasında kopya çekme girişimlerini tespit etmek için veri analizi ve benzerlik karşılaştırmaları yapabilir. Bu sistemler, öğrencilerin cevaplarını diğer kaynaklardan kopyalamalarını veya benzer cevaplar vererek sahtekarlık yapmalarını algılayabilirler. Ayrıca, yapay zeka ve makine öğrenmesi yöntemleri kullanarak gerçek zamanlı sınav gözetimi yapılabilir. Bu şekilde, öğrencilerin sınav sırasındaki davranışları ve aktiviteleri izlenir ve şüpheli durumlar otomatik olarak tespit edilir. Bu bağlamda, yapay zeka ve makine öğrenmesi, çevrimiçi sınav güvenliğini artırmak ve öğrencilerin dürüst bir şekilde değerlendirilmesini sağlamak için güçlü bir araçtır. Yapay zeka ve makine öğrenmesi, öğrencilerin sınav sonuçlarını objektif bir şekilde değerlendirmek için kullanılır. Bu, öğrencilerin sınav başarılarını sadece performanslarına dayalı olarak değerlendirilmesini sağlar. Bu da eğitim sisteminin adil ve objektif olmasına katkı sağlar. Öğrenciler, güvenli bir çevrimiçi sınav deneyimi yaşarken, eğitimciler de daha objektif ve adil bir değerlendirme yapabilirler. Bu da öğrencilerin geleceğe daha fazla güvenle ilerlemelerine yardımcı olur.

Öğrencilerin sınav esnasında verdikleri cevaplar, yazdıkları metinler ve diğer sınav materyalleri bu analiz sürecine dahil edilmektedir. Yapay zeka algoritmaları, bu veriler üzerinde çeşitli istatistiksel ve matematiksel yöntemler uygulayarak kopya çekme ihtimali yüksek olan örüntüleri belirlemektedir. Bu algoritmalar, veri setindeki örüntüleri öğrenerek ve çeşitli özelliklerle ilişkilendirerek sınav sırasında olası kopya durumlarını belirlemeye çalışır. Doğruluk, kesinlik, hassasiyet gibi performans metriklerine dayanarak makine öğrenmesi modeli eğitilir ve daha iyi sonuçlar elde etmek için sürekli olarak geliştirilir.

Veri seti, modelin eğitim sürecinde kullanılacak olan temel veri kaynağıdır ve yapay zeka modelinin performansını belirleyen en önemli faktörlerden biridir. Veri seti, farklı türde sınavlar ve çeşitli kopya çekme senaryolarını temsil etmelidir. Öğrencilerin farklı kopya çekme yöntemleri, benzer cevaplar ve kopya gruplarıyla ilgili verilere sahip olmak, yapay zeka modelinin bu senaryolara uyum sağlamasını ve gerçek dünya durumlarını doğru bir şekilde tespit etmesini sağlar. Ayrıca veri seti, çeşitli sınav materyallerini içermelidir. Metin tabanlı sorular, çoktan seçmeli sorular, görsel içerikler ve farklı formatlardaki soruların bulunduğu bir veri seti, modelin farklı türdeki sınav materyallerini analiz etme yeteneğini artırır. Buna ek olarak veri setinde öğrenci çeşitliliği de önemli bir role sahiptir. Farklı yaş gruplarından, farklı öğrenim seviyelerinden ve farklı başarı düzeylerinden öğrencilerin verilerine sahip olmak, yapay zeka modelinin

genellemesini artırır ve farklı öğrenci gruplarını adil bir şekilde değerlendirebilmesini sağlar.

Bu çalışma için kullanılan veri seti, çevrimiçi sınavlarda öğrencilerin gerçek cevaplarına ve sınav sırasındaki davranışlarına dayalıdır. Öğrencilerin kopya çekip çekmediğini doğru bir şekilde tahmin edebilmek için sınavın detayları incelenmiştir. Bu ayrıntılar, sınavın başlama ve bitirme süreleri, kullanılan cihazların bilgileri, IP adresleri, sınavın zorluk seviyesi, sınavın toplam süresi ve öğrencilerin soruları işaretleme hızı gibi farklı senaryo ve davranışları içermektedir. Bu verilerin kullanılması, öğrencilerin çevrimiçi sınavlarda kopya tespiti yapılabilmesi ve sınav güvenliğini artırmak için çeşitli tedbirler alınmasını amaçlamaktadır.

Veri seti, yapay zeka modelinin eğitim sürecinde kullanılacak temel veri kaynağıdır ve modelin başarısını belirleyen en kritik faktörlerden biridir. Ancak, yapay zeka modelleriyle çevrimiçi sınavlarda kopya tespiti yaparken bazen hatalı uyarılarla karşılaşılabilir. Bu, modelin gerçekte kopya çekme olmadığı durumları yanlış bir şekilde kopya olarak algılaması sonucu meydana gelebilir. Bu hatalı uyarılar, öğrenciler ve öğretmenler için olumsuz sonuçlara yol açabilir ve güvenilirliği sarsabilir. Bundan dolayı hatalı uyarıları azaltmak için birkaç önlem alınmalıdır. İlk olarak, veri dengelemesi yapılmalıdır. Özellikle az sayıda kopya içeren veri noktaları veya kopya örnekleri, modelin aşırı duyarlı hale gelmesine ve hatalı uyarılar vermesine neden olacağından doğru karar vermek gerekir.

Sürekli eğitim ve geri bildirim, yapay zeka modelinin kullanılan veri seti ve gerçek zamanlı geri bildirimlerle sürekli olarak eğitilmesi gerektiği anlamına gelir. Modelin eğitimi, yeni kopya tespiti senaryolarıyla güncellenmeli ve performansı düzenli olarak değerlendirilmelidir. Bu sayede model, daha kesin ve az hatalı uyarılar veren bir hale gelir.

Doğrulama veri seti, modelin performansını değerlendirmek için kullanılan bir veri seti olmalıdır ve eğitim veri setinden bağımsız olmalıdır. Bu, modelin eğitim verilerine aşırı uyum sağlamasını ve hatalı uyarıları artırmayı önler. Tüm bu adımlar ve yöntemler, yapay zeka modelinin hatalı uyarıları azaltmasına ve kopya tespiti konusunda daha güvenilir sonuçlar üretmesine yardımcı olur. Hatalı uyarıların minimize edilmesi, öğrenciler ve öğretmenler arasında güvenilir bir çevrimiçi sınav deneyimi sağlar ve yapay zeka yöntemlerinin etkinliğini artırır.

Çalışmada, yapay zeka modeli ile öğrencinin sınav çalışmaları analiz edilerek kopya çekme girişimlerini tespit edilmektedir. Eğer bir öğrencinin sınav hareketleri ve

değerlendirme sonucu başka bir öğrenciyle benzerlik gösteriyorsa bu durumda kopya tespiti olumlu olarak sonuçlanmaktadır. Öğrencinin kopya çektiği tespit edilmediği takdirde, öğrencilerin çalışmalarında herhangi bir kopya çekme veya usulsüz davranışa rastlanmadıysa, bu durumda kopya tespiti olumsuz olarak sonuçlanmaktadır. Yani, öğrencilerin sınav veya çalışma performanslarında herhangi bir kopya tespiti yapılamazsa, bu öğrencilerin çalışmalarının dürüstçe ve doğru bir şekilde tamamlandığına işaret eder. Ancak, şüpheli durumlarda yapay zeka modeli bazı durumlarda kopya çekme şüpheleri taşıyan çalışmaları belirtebilir. Bu durumda, öğretmenler veya sınav gözetmenleri bu şüpheli durumları daha yakından incelemek ve gerektiğinde ek değerlendirmeler yapmak için bilgilendirilirler. Yani, yapay zeka kopya tespiti sürecini hızlandırabilir ve potansiyel ihlalleri tespit etmede yardımcı olabilir. Ancak kesin sonuçlar için insan değerlendirmesi önemlidir çünkü yanlış pozitif veya yanlış negatif sonuçlar olabilir. Bu nedenle, adil bir süreç sağlamak için insan faktörü her zaman dikkate alınmalıdır.

1.1. Çalışmanın Amacı ve Önemi

Çalışmanın temel amacı, iletişim ve bilgi teknolojilerinin ilerlemesi ile birlikte, uzaktan eğitim uygulamalarının giderek yaygınlaştığı bir dönemde, çevrimiçi sınavlarda öğrencilerin kopya çekme olasılığını tespit etmek ve bu tespitlerin sonuçlarına dayalı olarak çevrimiçi sınav güvenliği konusunda öneriler sunmaktır.

Çevrimiçi sınavlarda öğrencilerin kopya çekmesi, eğitim sisteminin temel taşlarından olan kaliteyi doğrudan tehdit etmekte, öğrenciler arasındaki adalet ve eşitlik ilkelerini zedelemekte ve eğitimde dürüstlük ve etik değerlerin sarsılmasına yol açmaktadır. Bu durum, sadece sınav sonuçlarının güvenilirliğini düşürmekle kalmayıp, aynı zamanda öğrencilerin gerçek potansiyellerini yansıtamamalarına neden olarak onların akademik ve kişisel gelişimlerini engellemektedir. Bununla birlikte, haksız avantaj sağlayan öğrencilerle karşılaşan diğer öğrenciler arasında bir güven bunalımı doğmakta, bu da eğitim kurumlarına duyulan güveni azaltmaktadır. Eğitimde bu tür adaletsizliklerin artması, uzun vadede hem bireylerin hem de toplumun başarısını ve gelişimini olumsuz yönde etkileyebilir. Tüm bu nedenlerle, uzaktan eğitim süreçlerinde çevrimiçi sınav güvenliği sorunlarının ele alınması, çözüm yolları geliştirilmesi ve uygulamaya geçirilmesi kritik bir öneme sahiptir. Bu hem eğitim kalitesinin korunması

hem de toplumun etik deęerlere baęlı nitelikli bireyler yetiřtirmesi aısından vazgeilmez bir gerekliliktir.

Uzaktan eęitim platformu geliřtiren firmalara yol gsterici olacak bu alıřma, evrimii sınavlarda kopya ekme vakalarının nerelerde ve nasıl gerekleřtięini tespit ederek, bu firmaların daha etkili gvenlik nlemleri almasına olanak tanıyacaktır. Firmalar, kopya ekmenin hangi yntemlerle yapıldıęını ve hangi alanlarda zafiyetler olduęunu belirlediklerinde, sistemlerini bu aıkları kapatacak řekilde geliřtirebilir ve kullanıcı deneyimini bozmadan daha gvenilir bir sınav ortamı sunabilirler.

ęrencilerin hangi teknikleri kullanarak sınavda kopya ektiklerinin tespit edilmesi (rneęin, ekran paylařımı, bařka bir cihazdan yardım alma, internet zerinden arama yapma) platform geliřtiricilerine hangi gvenlik katmanlarını eklemeleri gerektięi konusunda rehberlik eder. Bu sayede, zellikle kullanılan yazılım ve teknolojilere dair hedeflenen noktalarda iyileřtirmeler yapılabilir.

Platformlar, hangi durumlarda daha fazla gvenlik ihtiyaı olduęunu anlayarak, sınav yapan kurumlara seenekler sunabilir. rneęin, biyometrik doęrulama, yz tanıma, srekli kamera takibi, gz hareketi izleme veya tarayıcı kilitleme gibi zellikler kullanılarak sınav gvenlięi artırılabilir. Bu alıřma, sınav yapan kurumların ihtiyalarına gre platformlara zelleřtirilebilir gvenlik seenekleri sunma gereksinimini de ortaya koyar. Farklı dersler, sınav trleri veya ęrencilerin seviyelerine gre deęiřebilen gvenlik ayarları, esnek ve daha hedefe ynelik zmler saęlar.

Sınav sırasında ęrencilerin davranıřlarının gerek zamanlı olarak izlenmesi ve analiz edilmesi, anormal aktivitelerin anında tespit edilmesine olanak saęlar. Platform geliřtiricileri bu verilerden faydalanarak, kullanıcı dostu ama aynı zamanda gvenli bir sınav deneyimi sunabilir.

Bu alıřma, sınav yapan kurumların sınavları daha gvenli hale getirmek iin platformlardan talep edebilecekleri eřitli gvenlik nlemlerini belirleyebilir. rneęin, sınav esnasında yapılabilecek denetimler, riskli durumların belirlenmesi ve otomatik uyarı sistemleri gibi yenilikler, sınav yapanların iřini kolaylařtırır. Sonu olarak, uzaktan eęitim platformları geliřtiren firmalara, evrimii sınavlarda kopya ekmeyi engellemeye ynelik kapsamlı ve etkili gvenlik araları geliřtirme konusunda bir rehber sunulacaktır. Bu hem eęitim kurumları hem de platform geliřtiricileri iin daha gvenli ve adil bir sınav ortamı saęlamada kritik bir adım olacaktır.

1.2. Çalışmanın Problemleri

Çevrimiçi sınavlarda kopya tespiti yaparken karşılaşılan çeşitli zorluklar vardır. Bu zorluklardan biri, eğitimciler veya öğrencilerin sınav sırasında yaşadığı sistem erişim sorunlarıdır. Örneğin, internet bağlantısının kesilmesi veya zayıf olması, kullanıcı bilgilerinin doğruluğunu etkileyebilir. Bu, sınav başlama ve bitirme süresi, sorular arası geçiş hızı, soruları cevaplama süresi ve cevapların kaydedilme süresi gibi önemli verilerin eksik veya yanlış olması neden olabilmektedir.

Kopya tespiti sürecinde karşılaşılan diğer zorluklar ise, öğrencilerin kimlik doğrulama sorunları yaşamaları, başkalarından yardım almaları, sınav sorularının sızdırılması, teknolojik kısıtlamalar, yüz tanıma ve gizlilik endişeleri gibi faktörlerin çevrimiçi sınav güvenliğini zorlaştırdığını göstermektedir. Bu zorlukların üstesinden gelmek için hem teknik hem de yönetsel çözümlere ihtiyaç vardır.

Dolayısıyla, çevrimiçi sınav güvenliği stratejileri geliştirirken bu tür sorunları göz önünde bulundurmak önemlidir. Teknik olarak güçlü bir altyapı ve veri yönetimi sistemine ek olarak, insan denetimi ve yapay zeka dengesini sağlamak, öğrenci davranış çeşitliliğini anlamak ve yanlış pozitif veya yanlış negatif sonuçları en aza indirmek için hassas bir denge kurmak önemlidir.

2. KAYNAK ARAŞTIRMASI

Dünyada uzaktan eğitim, internetin yaygınlaşması ve teknolojik gelişmelerle birlikte hızla gelişme göstermiştir. Türkiye’de uzaktan eğitim uygulamaları uzun yıllardan beridir farklı yöntemler ile aktif bir şekilde yürütülmektedir. Covid-19 Pandemisi ile birlikte uzaktan eğitim yaygınlaşmış ve özellikle üniversitelerde derslerin ve sınavların önemli bir bölümü online (çevrimiçi) olarak yapılmıştır. Salgın sürecinde etki alanı artan uzaktan eğitim, yükseköğretim kurumlarında olduğu gibi hemen hemen her öğretim kademesinde yerini almış ve alternatif eğitim haline gelmiştir. Uzaktan eğitim ve çevrimiçi sınav güvenliği konularında yapılan çalışmalar aşağıda sıralanmıştır.

Uzaktan eğitim, bilgisayar kullanma imkânı, motivasyon ve algı üzerinde önemli bir etkiye sahiptir. Uzaktan eğitimde başarı, öğrencilerin motivasyon düzeylerine ve uzaktan eğitim algılarına bağlı olarak önemli ölçüde değişebilir. Öğrencilerin bilgisayar kullanma yetenekleri ve teknolojiye olan aşinalıkları, uzaktan eğitim sürecindeki performanslarını etkileyebilir. Karşılaştırmalı yapılan bu örneklemin kapsamını, Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi İşletme Bölümü 3. sınıf öğrencileri (165 kişi), sınav ve anket uygulamaları ile toplanan veriler, ortalama ve testleri oluşturmaktadır. Uzaktan eğitim sürecinde etkileşim, öğrencilerin derinlemesine dersi kavraması ve yanlış anlamaları giderme açısından temel bir öneme sahiptir. Öğrencilerin, öğretmenleri ve diğer öğrencilerle etkileşime geçebilecekleri, sorular sormak ve cevap almak için uygun fırsatlarının olması, öğrenme deneyimini olumlu yönde etkiler. Bu tür etkileşimler, öğrencilerin konuları daha iyi anlamalarına ve gerektiğinde düzeltilmesi gereken yanlış anlamalarını düzeltmelerine yardımcı olabilir. Çalışma sonucunda, uzaktan eğitimde başarı, öğrencilerin bilgisayar kullanma becerileri, motivasyon düzeyleri ve uzaktan eğitim algıları ile doğrudan ilişkili olduğu tespit edilmiş. Ancak, etkileşim, derinlemesine öğrenme ve yanlış anlaşılmanın düzeltilmesi açısından birincil öneme sahiptir ve uzaktan eğitim sürecinin başarısını belirlemede kritik bir faktör olduğu belirtilmiştir (İbicioğlu ve Antalyalı, 2005).

Uzaktan yüksek eğitim programlarının akreditasyon süreci, bu programların kalite standartlarına uygunluğunu değerlendiren ve onaylayan resmi bir değerlendirme sürecidir. Bu süreç, öğrencilere ve topluma güvenilir ve kaliteli eğitim sağlanmasını teşvik etmek ve sağlamak amacıyla uygulanır. Türkiye’de, Yükseköğretim Kurulu (YÖK) ve Yükseköğretim Kalite Kurulu (YÖDAK) gibi kurumlar, üniversitelerin ve

yükseköğretim programlarının akreditasyonunu denetler ve düzenler. Akreditasyon süreci, eğitim programlarının belirli standartlara uygun olup olmadığını değerlendirmek için iç ve dış değerlendirmeleri içerir. Uzaktan yüksek eğitim programlarının akreditasyon süreci, bu programların eğitim kalitesini, öğretim materyallerini, öğrenci destek mekanizmalarını, öğrenci değerlendirme yöntemlerini ve öğrenci başarı düzeyini değerlendirmek için özel olarak uzaktan eğitim özelliklerine uygun kriterler içerir. Bu değerlendirmeler, programların uluslararası standartlara uygunluğunu göstermeye yardımcı olur ve öğrencilere daha güvenilir bir eğitim ortamı sunar. Akreditasyon süreci, eğitim kurumları ve programları için sürekli gelişimi teşvik ederken, öğrencilere de daha kaliteli eğitim alma imkânı sağlar ve işverenlerin eğitim programlarının değerini ve niteliğini anlamalarına yardımcı olur. Yapılan bu çalışmada ardışık karma model, Delfi Tekniği ve odak grup görüşmesinden faydalanılmıştır. Nitel ve nicel veriler toplanarak 2 evreden, çeşitli ölçütlerden ve programlarla örneklemeler kullanılmıştır. Araştırmada elde edilen verilere dayanarak, yeni açılacak uzaktan eğitim programları için bir ön akreditasyon ve mevcut programlar için akreditasyon yenileme sürecini içeren bir çerçeve önerilmiştir. Ayrıca Türkiye’de uzaktan yüksek eğitime özgü akreditasyon sistemine ihtiyaç duyulduğu saptanmıştır (Koçdar, 2011).

Uzaktan ve teknoloji destekli eğitim, öğrencilere daha esnek ve kişiselleştirilmiş bir öğrenme deneyimi sunar. Bu yaklaşım, öğrencilerin kendi hızında ilerlemesine ve öğrenme tarzlarına uygun şekilde içeriklere erişmesine olanak tanır. Aynı zamanda, öğretmenlere de öğrenci ilerlemesi ve performansını daha iyi takip etme imkânı sağlar. Bu tür eğitim yöntemleri, teknolojiye erişimi olan herkese ulaşabilme potansiyeli taşıırken, yine de eşitsizliklerin olduğu bölgelerde veya teknolojik altyapının yetersiz olduğu durumlarda bazı zorluklar yaşanabilir. Ancak genel olarak, uzaktan ve teknoloji destekli eğitim, öğrenme süreçlerini daha etkili ve verimli hale getirme potansiyeline sahiptir. Bu çalışma, uzaktan ve teknoloji destekli eğitimin gelişimini betimleyici bir analiz sunmayı amaçlamaktadır. Çalışmada, uzaktan ve teknoloji destekli eğitimin gelişimindeki önemli adımlara vurgu yapılmıştır ve yeni iletişim teknolojilerinin bu eğitim türünde yeni öğrenim biçimlerinin ortaya çıkmasına olanak sağlayabileceği ifade edilmiştir. Türkiye özelinde ele alındığında, uzaktan eğitim faaliyetlerinin ülkemizde henüz yeni bir olgu olduğu ve son dönemlerde gelişiminin arttığı belirtilmiştir (Çoban, 2013).

Uzaktan eğitimde iletişimsel boyutun dikkatli bir şekilde ele alınması, başarılı ve etkili bir eğitim sürecinin sağlanmasında kritik bir rol oynamaktadır. Uygulamalardaki iletişimsel sorunların çözümü, öğrenci ve öğretmen arasındaki etkileşimin artırılması ve

uzaktan eğitimde kaliteli bir iletişim ağı oluşturulması açısından büyük önem taşımaktadır. “Uzaktan Eğitimde İletişimsel Engeller” konusu araştırılmıştır. Araştırmanın amacı, yükseköğretimde uzaktan eğitim uygulamasında iletişimsel yönden hangi sorunların yaşandığını belirlemektir. Anket uygulaması ile araştırılan çalışmanın kapsamını, İstanbul Aydın Üniversitesi İletişim Fakültesinde uzaktan eğitim dersleri almış olan ikinci sınıf öğrencileri oluşturmaktadır. Araştırma sonucunda, öğrencilerin uzaktan eğitimde iletişime yönelik çeşitli engellerin var olduğu tespit edilmiştir (Elcil ve Şahiner, 2014).

Açık ve uzaktan öğrenme, öğrenmeye hevesli ve yeni insan tipinin ilgisini çekmekte ve bu alana olan ilgi her geçen gün artmaktadır. Ancak bu öğrenme yönteminin bazı dezavantajları da vardır ve öğrencilerin eğitim-öğretim faaliyetlerine katılım ve devamlılığı ile ilgili sorunlar yaşayabilirler. Bu çalışmada öğrenmeye açık insanların eğitim-öğretimdeki devamlılığı ve katılımında yaşadığı sorunları araştırılmıştır. Çalışmada dünyadan çeşitli açık üniversitelerin akademik takvimi ve İstanbul Üniversitesi Açık ve Uzaktan Eğitim Fakültesi akademik takvimi hakkında görüşlerini anket aracılığıyla sunan 11.786 öğrencinin görüşü alınmıştır. Çalışma sonucunda öğrencilerin 3 dönemlik akademik takvim isteğinde bulunduğu ve bu durumun eksikliğinin Türkiye’deki açıktan ve uzaktan öğrenme faaliyetlerinin gözden geçirilmesi gerektiğini ortaya koymuştur (Demir, 2015).

Eğitim, bireylerin kişisel ve zihinsel gelişimi için temel bir unsurdur. Teknolojinin ilerlemesiyle birlikte eğitim kavramı da dönüşmüş ve teknoloji eğitimin vazgeçilmez bir parçası haline gelmiştir. İnternet teknolojisinin eğitimde kullanılmasıyla birlikte "Uzaktan Eğitim" kavramı ortaya çıkmıştır. Uzaktan eğitim, teknoloji sayesinde eğitim süreçlerinin internet üzerinden gerçekleştirilmesini sağlayan bir öğrenme modelidir. Dünya çapında birçok üniversite ve eğitim kurumu uzaktan eğitim programları sunmaktadır. Türkiye’de de uzaktan eğitim uygulamaları hızla yaygınlaşmaktadır ve birçok üniversite uzaktan eğitim programları sunmaktadır. Bu sayede öğrenciler, kendi hızlarına ve tercih ettikleri zamanlara uygun olarak eğitim alabilmektedir. Dünyada ve Türkiye’de uzaktan eğitimin güncel durumu incelenmiştir. Bu incelemede, gelişen teknoloji ile ortaya çıkan uzaktan eğitimin tarihsel süreci ile birlikte, Dünyada ve Türkiye’de seçilmiş olan bazı üniversitelerin uzaktan eğitim veren programlarını durumu analiz edilmiştir. İnceleme sonucunda, uzaktan eğitimin tarihçesinde Dünyada ve Türkiye’de bireylere uzaktan eğitim aracılığıyla eğitim imkânı sağlandığı ve bu imkânların Türkiye’de giderek arttığı vurgulanmıştır (Özbay, 201).

Yükseköğretim kurumları da artan eğitim ihtiyaçlarına web tabanlı uzaktan eğitim imkanları ile destek olmaktadır. Ancak yüz yüze örgün eğitimden farklı ihtiyaçları olan web tabanlı uzaktan eğitim programları, yükseköğretim kurumları, eğitmenler ve öğrenciler gibi paydaşlar için farklı zorluklar ve sorunlar ortaya çıkarmaktadır. Bu çalışmada, yükseköğretim kurumları örneği üzerinden web tabanlı uzaktan eğitim programlarında yaşanan sorunların incelenmesi amaçlanmıştır. Çalışma kapsamını, yükseköğretim kurumlarından öğrenciler, öğretim elemanları, idari personeller ve kullanılan web tabanlı uzaktan eğitim programları oluşturmaktadır. Araştırmanın sonucunda, yükseköğretimde yaygınlaşan web tabanlı uzaktan eğitim programları kullanıcılar için kolaylık sağlarken eğitimde etkileşimden kaynaklanan motivasyon, adaptasyon, alt yapı gibi sorunlarda yaşandığı belirtilmiştir (Bilgiç, 2015).

Çevrimiçi sınavlar, uzaktan eğitim ve çevrimiçi eğitim sistemlerinin bir parçası olarak daha yaygın hale gelmiştir ve birçok avantaj sunmaktadır, özellikle coğrafi sınırlamalara bağlı kalmadan katılım imkânı sağlaması ve otomatik değerlendirme süreci ile zaman ve çaba tasarrufu sağlaması gibi. Ancak güvenlik ve teknik sorunlar gibi zorluklar da çevrimiçi sınavların karşılaştığı bazı dezavantajlar vardır. Manuel sınavlar ise hala birçok eğitim kurumunda kullanılmakta olup, bazı durumlarda özellikle sınavın uygulanması ve değerlendirilmesi konusunda daha fazla güvenlik ve kontrole ihtiyaç olduğunu göstermektedir. İnsan ve Bilgisayar Etkileşimi (HCI) konferansında sunulan bu makale, çevrimiçi sınavlarda öğrenci güvenliği tehditlerini incelemektedir. Çevrimiçi sınav sistemi üzerine kapsamlı bir değerlendirme ve inceleme sunarak, sistemin nasıl çalıştığı, ne tür faydalar sağladığı, hangi algoritma ve teknolojilerin kullanıldığı ve güvenlik önlemlerinin nasıl uygulandığı hakkında bilgi vermektedir. Bu tür çalışmalar, çevrimiçi sınav sistemlerinin geliştirilmesi ve iyileştirilmesi için önemli bir kaynak olabileceği öngörülmüştür (Bobde and Chaudhari & Shahane, 2017).

Uzaktan eğitime yönelik tutumlar, bireylerin uzaktan eğitim yöntemine ve bu yöntemle yapılan eğitim faaliyetlerine karşı sergilediği düşünceleri, inançları ve duyguları ifade etmektedir. Bu tutumlar, uzaktan eğitimin etkili bir eğitim yöntemi olarak kabul edilip edilmediğini, tercih edilip edilmediğini, güvenilir ve kaliteli bir eğitim sağladığına dair inançları ve olumlu/olumsuz hisleri içerebilir. Uzaktan eğitime yönelik tutumlar, farklı bireyler ve topluluklar arasında çeşitlilik gösterebilir ve zamanla değişebilir. Ağrı İbrahim Çeçen Üniversitesinin farklı bölümlerinde eğitim gören öğrencilerin uzaktan eğitim dersine karşı tutumları ele alınmıştır. Betimsel araştırma niteliğinde olan bu araştırmanın kapsamını, 2018-2019 yılında Ağrı İbrahim Çeçen

Üniversitesi'nde öğrenim gören yedi farklı bölüm öğrencilerinden 180 Kadın 260 Erkek toplam 449 kişi oluşturmaktadır. Belirtilen bölüm öğrencilerinin eğitim sisteminde yer alan uzaktan eğitime yönelik alt yapı yetersizliklerinin, uygulama gerektiren bazı derslerde uzaktan eğitimin beklenen verimi sağlayamamasının öğrencilerin uzaktan eğitim programları konusunda aynı tutuma sahip olmaları sonucunu ortaya çıkarmıştır (Bayram ve Peker & Aka & Vural, 2019).

Dijital sınavlar, teknoloji ve internetin yaygınlaşmasıyla giderek daha popüler hale gelmektedir. Bu avantajlar, öğrenci deneyimini geliştirirken, eğitim sağlayıcıların da yönetim ve değerlendirme süreçlerini kolaylaştırır. Ancak dijital sınavların da güvenlik ve teknik sorunlar gibi bazı zorlukları olabilir ve bu yönlerin dikkate alınması önemlidir. Tüm sınavların dijital ortama taşınmasıyla öğrencilerin sınav başarılarına etkisi araştırılmıştır. Tek soruluk anket düzenlenerek yapılan bu karşılaştırmanın kapsamını, Afyon Kocatepe Üniversitesi Uzaktan Eğitim Meslek Yüksek Okulunda 2018-2019 güz döneminden itibaren final ve bütünleme sınavlarına ait sözel ve sayısal olmak üzere 3'er ders oluşturmaktadır. Anket sonucunda, öğrencilerin basılı ortamda gerçekleştirilen sınavı tercih etme oranının %55 olmasına rağmen dijital ortamda gerçekleştirdikleri final sınavlarında 6 dersin 5'inde başarılarının arttığı görülmüştür (Solak ve Ütebay & Yalçın, 2020).

Üniversitelerin pandemi döneminde uzaktan eğitim konusundaki çalışmaları ve bu çalışmaların öğrenci ve öğretmenler üzerindeki etkileri üzerine üniversitelerde uzaktan eğitim konusunda yönetici veya çalışan katılımcıların görüşlerine başvurulmuştur. Katılımcıların büyük çoğunluğunun üniversitelerin UZEM (Uzaktan Eğitim Merkezi) yöneticileri ve çalışanları olduğu ifade edilmiştir. Ancak tüm üniversitelerde UZEM biriminin olmadığı ve bazı üniversitelerde uzaktan eğitim sürecini farklı birimler veya kurulan komisyonlar tarafından yönettiği belirtilmiştir. Covid-19 pandemi döneminde Türkiye'deki üniversitelerin uzaktan eğitim sistemlerinin incelenmiş ve bu incelemede tarama modeli kullanılmıştır. İncelemenin kapsamını, Türkiye'de bulunan 208 üniversite ve uzaktan eğitim sistemleriyle ilgilenen çalışanlar oluşturmaktadır. İnceleme sonucunda uzaktan eğitimde en fazla ALMS ve Moodle sistemlerinin kullanıldığı tespit edilmiştir. Pandemi sürecinde incelemeye katılan sistem kullanıcılarının, öğrenme yönetim sisteminden memnun olduklarını ve uzaktan eğitimin yaygınlaşması gerektiğini ortaya koymuşlardır (Durak ve Çankaya & İzmirli, 2020).

Covid-19 pandemi süreci, üniversitelerin uzaktan eğitim sistemlerinin alt yapısına sahip olmalarının ne kadar önemli olduğunu göstermiştir. Bu dönemde eğitime uzaktan

devam etmek zorunda kalındığında, eğitimin aksamadan ve etkili bir şekilde yürütülmesi için hazır bir uzaktan eğitim sistemine sahip olmak gereklidir. Çalışmada Pandemi sürecinde uzaktan eğitim alan Çanakkale Onsekiz Mart Üniversitesi, Turizm Fakültesi lisans öğrencileri yer almıştır. Amaç Covid19 pandemisi sürecinde internet tabanlı uzaktan eğitim alan öğrencilerin fikirlerini almak ve değerlendirmektir. Üniversitelerin alt yapıyı oluşturması ve Covid-19 gibi durumlar için uzaktan eğitim sistemini hazır bulundurması önerilmektedir. Özellikle Yüksek Öğretim Kurumunun kendi internet tabanlı uzaktan eğitim sisteminin alt yapısını oluşturması önerilmektedir (Buluk ve Eşitti, 2020).

Çevrimiçi gözetim, sınav sırasında öğrencileri izlemek için kullanılan teknolojik bir yöntemdir. Bu gözetim, öğrencinin bilgisayar kamerası ve mikrofonu aracılığıyla canlı olarak ya da kayıt altına alınarak yapılabilir. Çevrimiçi gözetim yazılımları, öğrencinin davranışlarını izleyerek olağandışı aktiviteleri tespit eder ve sınav sırasında hile yapma girişimlerini engellemeye çalışır. Örneğin, öğrencinin başka bir kişiyle konuşup konuşmadığını, ekranda başka bir şey açıp açmadığını veya sınav ortamında yasaklanmış materyaller olup olmadığını tespit edebilir. Bu yazılımların uygulanmasından önce ve sonra yapılan sınav sonuçları karşılaştırılır., Eğer gözetim uygulandıktan sonra sınav performansında bir düşüş gözlenirse, bu durum öğrencilerin gözetim öncesinde kopya çekmiş olabileceğini düşündürür. Ayrıca, gözetim altındaki sınavlarda öğrenci performansındaki değişiklikler, bu tür yazılımların hile yapmayı ne kadar etkili bir şekilde önlediğini gösterir. Bu tür bir değerlendirme, çevrimiçi sınavların güvenilirliğini artırmak ve eğitim sürecinde adaleti sağlamak için önemli bir adımdır. Gözetim yazılımı, öğrencilerin sınavları adil bir şekilde tamamlamasını sağlarken, eğitim kurumlarına da hile yapma girişimlerini tespit etme konusunda güçlü bir araç sunar. “Cheating in online courses: Evidence from online proctoring” makalesinde çevrimiçi sınavlarda kopya çekme davranışları araştırılmış ve çevrimiçi gözetim (proctoring) yazılımının bu tür hileleri önlemede ne kadar etkili olduğunu değerlendirmiştir.

İki farklı çevrimiçi ders üzerinde yürütülen bir çalışmada yüksek riskli sınavlar için çevrimiçi gözetim yazılımı uygulanmıştır. Çalışma, bu gözetim yazılımının kullanılmasından önce ve sonra derslerin yapısı, içeriği ve sınav formatında herhangi bir değişiklik olmaksızın öğrenci performansını karşılaştırmaktadır. Çalışmada, bir yarı deneysel tasarım kullanarak gerçekleştirilmiştir. Yüksek riskli sınavlarda çevrimiçi gözetim sağlamak amacıyla bir webcam kayıt yazılımı uygulanmıştır. Araştırma bulguları, çevrimiçi gözetim uygulandıktan sonra öğrenci performansında genel bir düşüş

olduğunu göstermektedir. Bu düşünüş, çevrimiçi derslerde gözetim öncesinde kopya çekildiğini göstermektedir

Covid-19 salgını, dünya genelinde eğitim sisteminde büyük bir değişimi zorunlu hale getirmiştir. Uzaktan eğitim yöntemleri, öğrencilere bilgilerin sanal ortamda aktarılmasını sağlamıştır. Ancak bu dönemde coğrafya öğretiminin etkililiği hakkında yapılan çalışmaların eksik olduğu görülmüştür. Bu nedenle, coğrafya öğretmenlerinin görüşlerine dayanarak etkili bir ölçek geliştirerek, uzaktan eğitimle gerçekleştirilen coğrafya öğretiminin etkililiğini doğru bir şekilde ölçmeyi ve değerlendirmeyi amaçlayan bir araştırma yapılması önem arz etmektedir. Araştırmanın kapsamını, İstanbul'daki coğrafya öğretmenlerine sorulan 18 ölçek ve örnekleme katılan 356 coğrafya öğretmeni oluşturmaktadır. Çalışma sonucunda, uzaktan eğitimde ölçeğin geçerlilik ve güvenilirliğinin yüksek olduğu bu sayede uzaktan eğitimde farklı bakış açılarının ölçümünde ölçeğin kullanılabileceğini ortaya konulmuştur (İnce ve Yentür & Şahin, 2021).

Eğitim alanında e-öğretim uygulamaları, öğrenci ihtiyaçlarını karşılamak ve eğitim kalitesini artırmak için önemli bir rol oynamaktadır. E-öğretim konusunda yapılan çalışmalarda olumlu ve olumsuz etkilerine ilişkin farklı görüşler olsa da günümüzde birçok öğrenci e-eğitim uygulamalarından yararlanmak istemektedir. Bu talep, yükseköğretim kurumlarının yöneticilerini etkin, verimli ve sürekli stratejiler geliştirmeye yönlendirmiştir. Yükseköğretim kurumları bu alana yönelik etkili ve sürekli stratejiler geliştirmektedirler. Yapılan bu değerlendirmenin kapsamını, yükseköğretim sektörü ve yükseköğretim kurumları oluşturmaktadır. Çalışma “verimli ve etkili e-eğitim stratejilerini nasıl geliştirilebilir?” sorusuna yanıt aramış ve elde edilen bulgular ile Türkiye’de e-öğretim ile ilgili stratejilerin güçlendirilmesine ilişkin yükseköğretim kurumlarına öneriler sunulmuştur (Ateş & Banazılı, 2021).

Covid-19 pandemisi sonrasında ülkemizdeki yükseköğretim kurumlarının uzaktan eğitim uygulamalarına başlamasıyla klasik uzaktan eğitimden farklı olarak, öğrenciler, öğretmenler ve öğretim kurumları hazırlıksız bir şekilde bu olağan dışı uzaktan eğitim süreçleriyle karşılaşmışlardır. Bu nedenle, bu dönemin niteliklerinin anlaşılması ve öğrencilerin "uzaktan eğitim" kavramına yönelik metaforik algılarının belirlenmesi amacıyla bu çalışma yapılmıştır. Nitel araştırma yöntemlerinden olan olgu bilimi (fonomoloji) deseninde tasarlanan çalışmanın kapsamını, 2020-2021 öğretim yılında Kilis 7 Aralık Üniversitesi İslami İlimler Fakültesinin değişik sınıflarında öğrenim gören, uzaktan eğitime köy, ilçe ve il gibi farklı yerleşim birimlerinden katılan 197 kız ve erkek

öğrenci oluşturmaktadır. Analiz sonucunda, 197 katılımcı tarafından uzaktan eğitimle ilgili 167 farklı metafor üretilmiştir. Üretilen metaforlar aracılığıyla öğrenciler çoğunlukla “aldatıcı” ve “yapay” olduğu algısına sahip oldukları uzaktan eğitimi, bu görüşlerinin bir uzantısı olarak “verimsiz” ve “etkisiz” olarak nitelendirmişlerdir (Güngör, 2021).

Covid-19 pandemisi ve benzer süreçlerde örgün eğitim yerine uygulanan uzaktan eğitim modelinde birçok sorun ve fırsatlar tespit edilmiştir. Sorunlar arasında, öğrencilerin eğitim materyallerine erişimde yaşadığı sıkıntılar, internet bağlantısı ve teknoloji altyapısındaki yetersizlikler, öğretmenlerin uzaktan eğitime adaptasyon süreçlerinde yaşadığı zorluklar, etkileşim eksikliği ve motivasyon düşüklüğü gibi konular bulunmaktadır. Aynı zamanda, öğrenci katılımının azalması ve derslerin etkili bir şekilde yönetilememesi gibi zorluklar da sıkça karşılaşılan sorunlar arasında yer almaktadır. Doküman analizi tekniği kullanılan bu çalışmanın kapsamını, dijital eğitim araçları, öğrenciler, öğretmenler kapsamaktadır. Araştırmanın sonucunda, pandemi sürecinde uygulanan uzaktan eğitimde karşılaşılan sorunların internete erişimi sorunu, bilgi iletişim teknolojilerine (BİT) erişimde ve kullanımda ortaya çıkan dijital sorunlar, okuldan uzak kalan öğrencilerin öğrenmelerinde ebeveyn desteğinin düşüklüğü ve öğretmenlerin materyal-eğitsel etkinlik oluşturma ve bunların uygulanması konusunda kendilerini geliştirme ihtiyaçları olduğu tespit edilmiştir (Ülger, 2021).

Ülkemizde uzaktan eğitim sistemi henüz olgunlaşma aşamasındadır. Bu aşamada, alt yapı yetersizliği ve sistemin yeterince tanınmaması gibi bazı sorunlar yaşanmaktadır. Yaşanan sorunlar doğrultusunda ülkemizde yaygın olarak kullanılan uzaktan eğitim platformları ele alınmış ve bu çalışmada eğitimciler ve öğrenciler gözünden hangi kriterlere göre tercihte bulunduklarını tespit edilmeye çalışılmıştır. Analitik Hiyerarşi Prosesi yöntemi kullanılan bu çalışmanın kapsamını, sekiz uzman görüşü, dört ana kritere göre hazırlanan hiyerarşi (kalite, reklam, performans, duyarlılık) ve dört farklı uzaktan eğitim platformu (Perculus, Microsoft Teams, Zoom, ALMS) oluşturmaktadır. Uzmanlar, tanınan, reklamlar ve hizmetleri açısından kendini iyi ifade eden, beklenen kaliteyi sağlayan, performans olarak ihtiyaçları karşılayan ve eğitim hizmetlerine duyarlı olan platformların tercih edilmesi gerektiği konusunda fikir birliğine varmışlardır (Derici, 2021).

“A Secure Online Exams Conceptual Framework For South African Universities” adlı çalışmada Güney Afrika üniversitelerinde gerçekleştirilen çevrimiçi sınavlar için çeşitli güvenlik önlemlerinin alınması gerektiğini ve teknolojik politikalar geliştirerek

üniversitelerin çevrimiçi sınavların güvenliğini artırma konusunda rehberlik etmeyi amaçlamıştır. Çalışmanın kapsamını oluşturan Güney Afrika üniversitelerine çevrimiçi sınavların adil ve güvenilir ortamlarda gerçekleşmesi için öneriler sunmaktadır (Ngqondi ve Maoneke & Mauwa, 202).

e-Sınav sistemi, geleneksel kâğıt ve kalem kullanılan sınavların yerine geçen, bilgisayar veya diğer elektronik cihazlar üzerinden yapılan bir sınav sistemidir. Bu sistem, öğrencilerin sınavlarına dijital platformlar üzerinden katılmalarını sağlar ve genellikle internet bağlantısı gerektiren bir yapıya sahiptir. e-Sınav sistemi, çeşitli avantajlar sunsa da bazı güvenlik sorunları da beraberinde getirebilir. Nicel ve nitel yöntemlerden oluşan karma yöntemler yaklaşımı ile yapılan araştırmanın kapsamını, Türkiye’de bir devlet üniversitesinde önlisans programlarından Turizm ve Otel İşletmeciliği, Bankacılık ve Sigortacılık, Tıbbi Dokümantasyon ve Dokümantasyon, Bilgisayar Programcılığı Uzaktan Eğitim Programı, Adalet Uzaktan Eğitim Programı olmak üzere 5 farklı program oluşturmaktadır. Bu çalışma sonucunda, öğrenciler e-Sınav sistemi ile ilgili genel olarak memnuniyetlerini belirtmiştir. Ancak, bazı sorunların olduğunu da bildirmişlerdir. Soruların içerik ve yapılarına dikkatli bir şekilde kontrol edilmesi gerektiği konusunda öğrencilerin talepleri bulunmaktadır (Adanır, 2010).

Dijital sınavlarda yapay zeka, sınav güvenliğini artırmak ve kopya çekme girişimlerini tespit etmek için kullanılır. Yapay zeka, öğrenci davranışlarını izleyerek anormal aktiviteleri tespit edebilir ve doğal dil işleme (NLP) gibi tekniklerle sınav yanıtlarının özgünlüğünü değerlendirir, böylece akademik bütünlüğü korumaya katkı sağlar. Dijital sınavlar ve yapay zeka teknolojilerinin son yıllarda hızla geliştiği göz önüne alındığında “Machine learning based approach to exam cheating detection” adlı bu makalede dijital sınavlarda yapay zeka teknolojilerinin kopya tespitinde nasıl kullanılabileceğini kapsamlı bir şekilde ele alınmıştır. New South Wales Üniversitesi raporunda, 139 bilim öğrencisinin çalışmalarını tamamlamak için Çinli mesajlaşma sitesi WeChat’ten hayalet yazarlar aracılığıyla sınav sonrası yapılan puan analizi yoluyla kopya vakalarını tespit edilmiş ve anormal puanları belirlenmiştir. Kapsam dahilinde, çeşitli yapay zeka teknikleri ve algoritmalar, öğrenci davranışlarını izlemek ve analiz etmek için kullanılan yöntemler ve kopya tespitine yönelik geliştirilen teknolojiler incelenmiş aynı zamanda, dijital sınavların güvenliğini artırmak amacıyla kullanılan yüz tanıma, doğal dil işleme (NLP) ve makine öğrenimi gibi teknolojilerin uygulanabilirliği de değerlendirilmiştir. Makale, yapay zekanın dijital sınavlarda kopya tespitinde etkili bir araç olduğunu vurgulamıştır. YZ tabanlı sistemlerin, kopya girişimlerini tespit etme

konusunda yüksek doğruluk oranlarına sahip olduğu belirtilmiştir. Bununla birlikte, bu teknolojilerin sürekli olarak geliştirilmesi gerektiği ve özellikle etik ve gizlilik konularında dikkatli olunması gerektiği sonucuna varılmıştır. Gelecekteki araştırmaların, YZ'nin daha da sofistike hale getirilmesine ve dijital sınavlarda daha yaygın bir şekilde uygulanmasına odaklanması gerektiği belirtilmiştir (Kamalov and Sulieman & Santandreu, 2021).

Çevrimiçi öğrenme, eğitim materyallerinin ve iletişimin internet aracılığıyla sağlandığı, öğrencilerin fiziksel olarak derslere katılmadan eğitim alma imkânı bulduğu uzaktan eğitim türüdür. Çevrimiçi öğrenme, esneklik, erişilebilirlik ve etkileşimli öğrenme imkanları sunarak geleneksel eğitim yöntemlerine değerli bir alternatif oluşturmaktadır. Ancak öğrencilerin ihtiyaçlarına uygun şekilde düzenlenmeli ve teknik destek ile öğrenci başarısını artırmak için gerekli önlemler alınmalıdır. "Preventing Cheating in Online Examinations" Bilgi Teknolojileri ve Eğitim (ITRE) dergisinde yayınlanan bu makale, çevrimiçi sınavlarda kopya çekme ve diğer hilelerin nasıl önlenebileceğini araştırmaktadır. Otomatik denetim sistemleri, kimlik doğrulama yöntemleri ve güvenlik politikaları gibi konulara odaklanarak çeşitli önerilerde bulunmuştur (Valızadeh, 2022).

Kopya tespit sürecinde yapılan çalışmaların birçoğunda, öğrencinin sınav boyunca yaptığı eylemler incelenmektedir. Bu inceleme süreçlerinde birçok yapay zekâ modeli sayesinde gerçekleştirilmiştir (Atoum and Chen & Liu & Hsu & Liu 2017), (Raj and Narayanan & Bijlani, 2015), (Ozgen ve Öztürk & Bayraktar & Aksoy, 202), (Motwani and Nagpal & Motwani & Nagdev & Yeole, 2021). Bu çalışmalarda çoğunlukla kamera, mikrofon donanımı, göz takip cihazları kullanılmıştır. Yüksek maliyet ve kaynak gereksinimi, kurulum ve kullanım karmaşıklığı, teknik uyumluluk sorunları, gizlilik ve etik kaygılar, internet ve bağlantı gereksinimleri gibi nedenlerden dolayı, kopya tespitinde yazılım tabanlı çözümler (örneğin, veri analitiği, davranış analizi, anomali tespiti) genellikle daha ölçeklenebilir ve kullanıcı dostu bir alternatif olarak görülmektedir (Dadak ve Uyan & Öztürk, 2022).

Çalışmalarda görüldüğü gibi çevrimiçi sınavların, uzaktan eğitim süreci güvenlik kapsamında odak noktası olmuştur. Çevrimiçi sınav güvenliği sağlanmadığı takdirde sınav değerlendirmelerinde adaletli ve doğru sonuçlara ulaşılamamaktadır. Çevrimiçi sınavlar ile öğrencinin dersleri takip etme durumu, dersleri ne kadar anladığını ve derslerin ne kadar akılda kaldığı sınav sonuçlarına göre ölçüm sonucu elde edilmektedir. Öğretim elemanlarının sınav sonuçlarına göre doğru ölçmede dikkat etmeleri gereken

husus dersleri öğrenme sürecinde öğrencinin performansını ölçmektir. Bunun sonucunun cevabını almak için çevrimiçi sınavların belli periyotlarda yapılmasında öğrencilerin başarısını ölçmek için daha doğru sonuca ulaşmak gerekir. Bu süreç için, öğretim elemanlarının daha fazla sonuç elde edecekleri raporlar ile analiz odaklı imkânlar sunmaktadır. Fakat eğitmenler ve öğrenciler arasındaki etkileşimin derecesi klasik sınav değerlendirme yöntemlerine göre daha düşük olduğundan, öğrencileri tanıma konusunda eğitmenler için dezavantajlar oluşturabilmektedir. Buna bağlı olarak uzaktan eğitimde kalitenin artırılması için önemli bir yeri olan ölçme ve değerlendirme süreci oldukça önemlidir.

Türkiye’de kullanılan bazı uzaktan eğitim uygulamalarında öğrencinin katılmış olduğu aktiviteler ile ilgili çeşitli bilgiler (loglar) tutulmaktadır; ders ilerleme raporu, ders tabanlı sınıf raporu, genel aktivite raporu, özet şube raporu, tamamlama raporu, kişisel giriş raporu, sistem giriş raporu, aktif dönem öğrenci raporu, detaylı kullanıcı raporu, saatlik sınav raporu, sınav sonuçlarının kontrol edildiği sınav kontrol raporu, soru istatistikleri raporu, soru bankasına eklenen soruların soru raporu yaygın olarak tutulmaktadır. Elde edilen bu raporlar sayesinde eğitmenler sınav ölçme-değerlendirme ve başarı durumu analiz işlemlerini daha hızlı gerçekleştirmektedirler. Ayrıca öğrenciyi izlemek ve başarı durumunu geliştirmek için bu raporlarda sunulan bilgiler oldukça önemlidir.

Iğdır Üniversitesi'nde kullanılan ALMS (Advancity Learning Management System), uygulaması çevrimiçi öğrenmeyi desteklemek amacıyla kullanılan ticari bir öğretim yönetim sistemidir. Türkiye’de yaklaşık 60 yükseköğretim kurumu tarafından kullanılan öğrenme yönetim sistemi senkron ve asenkron öğrenme imkânı sunmaktadır. Bu uygulama hem web hem de mobil uygulamalar aracılığıyla öğrencilere çevrimiçi ders materyallerine erişim ve çevrimiçi sınavlara katılım olanağı sağlamaktadır. ALMS (Advancity Learning Management System) gibi öğrenme yönetim sistemleri, genellikle eğitim kurumlarına çeşitli öğrenme ve eğitim ihtiyaçlarını karşılamak için çeşitli özellikler sunar. Bu özellikler, öğrencilerin eğitim materyallerine erişimini kolaylaştırmak, sınavları yönetmek, öğrenci ilerlemesini izlemek ve daha fazlasını yapmak için tasarlanmıştır. İşte genel olarak bir öğrenme yönetim sistemi olarak ALMS'in içeriğinde bulunan özelliklerden bazıları:

Çevrimiçi Ders Materyalleri: ALMS, eğitim kurumlarının öğrencilere sunacakları çevrimiçi ders materyallerini (ders notları, sunumlar, videolar, dokümanlar vb.) yüklemelerine ve bu materyallere erişimini düzenlemelerine olanak tanır.

Çevrimiçi Sınavlar ve Değerlendirme: ALMS, eğitimcilerin ve öğrencilerin çevrimiçi sınavları yönetmelerine ve bu sınavların değerlendirmesini yapmalarına olanak tanır. Sınavlar genellikle çoktan seçmeli sorular, açık uçlu sorular veya diğer soru türlerini içerebilir.

Öğrenci İzleme ve Performans Değerlendirmesi: ALMS, öğrencilerin ilerlemesini izlemek için bir dizi araç sunar. Öğrenci performansını değerlendirmek ve gerektiğinde destek sağlamak için notlar, ilerleme raporları ve değerlendirme araçları kullanılabilir.

Sosyal İş Birliği: Bazı ALMS uygulamaları, öğrencilerin ve öğretmenlerin birbirleriyle iletişim kurmasını ve iş birliği yapmasını kolaylaştırmak için iletişim araçları, forumlar, sohbet odaları vb. içerebilir.

Kullanıcı Yönetimi: ALMS, eğitim kurumlarının öğrenci ve öğretmen hesaplarını yönetmelerine olanak tanır. Kullanıcıları eklemek, silmek ve düzenlemek gibi yönetim işlevleri sunar.

Raporlama: ALMS, eğitim kurumlarına öğrenci performansı, katılım oranları ve diğer veriler hakkında raporlar sunar. Bu raporlar, eğitim programlarını değerlendirmek ve iyileştirmek için kullanılabilir.

Mobil Erişim: Birçok ALMS uygulaması, öğrencilerin ve öğretmenlerin mobil cihazlardan erişebilmeleri için mobil uygulamalar sunar.

Entegrasyon: ALMS, diğer eğitim araçları veya sistemlerle entegre edilebilir. Örneğin, video konferans platformları, içerik yönetim sistemleri veya ödeme sistemleriyle entegrasyon sağlanabilir.

Çevrimiçi öğrenme ve sınav yönetim sisteminde yukarıda belirtildiği gibi çevrimiçi sınavlara ait verileri içeren performans raporları da yer almaktadır. Bu raporlar; oturumu kayıtları, sınav başlatma ve bitirme zamanları, sınav soruları ve cevapları, sınav süresi, açılan sekmeler ve internet erişimi, sınav ilerlemesi ve soru geçişleri gibi bilgileri içermektedir. Elde edilen bu bilgiler ile oluşturulan veri seti yapay zeka teknikleri kullanılarak öğrencilerin kopya çekme eğilimlerini belirlenecektir.

3. MATERYAL VE YÖNTEM

3.1. Materyal

Bu çalışmada kullanılan veri seti, Iğdır Üniversitesi'nin ALMS (Advancity Learning Management System) uygulaması üzerinden elde edilen çevrimiçi sınav sonuçlarına ait verilerden oluşmaktadır. Toplam veri sayısı, seçilen herhangi bir dersin final sınavına katılan yaklaşık 3 bin öğrencinin sınav bilgilerini içermektedir. Verilerin tamamı çevrimiçi sınav kayıtlarından elde edilmiştir. Veri setinde yer alan bazı temel öznitelikler Tablo 3.1’de verilmiştir.

Tablo 3.1. ALMS sınav logları

Kullanıcı Adı	Adı Soyadı
E-posta	Programı
Sınav Başlangıç Zamanı	Sınava Giriş Hakkı
Sınav Bitiş Zamanı	Aktif Kalma Süresi
Sınav Süresi	Katılım Durumu
Toplam Sınav Süresi	Soru Tıklama Tekrar Sayısı
Her Bir Sorunun Doğru Cevabı	IP Bilgisi
Her Bir Soruya Verilen Yanlış Cevap	Tarayıcı Bilgisi
Boş Bırakılan Soru Sayısı	Cihaz Bilgisi
İptal Edilen Sorular	Öğrenci Sınav Notu
Soruların Cevaplanma Süresi	Sınıf Genel Ortalaması

Veri seti için yukarıda belirtilen tüm özellikler yerine, daha az sayıda öznitelik seçmek, çevrimiçi sınavlarda kopya çekme tahmini için hem regresyon hem de sınıflandırma yöntemlerinin tahmin başarımlarını artıracaktır. Ayırıştırılan öznitelikler, kopya çekme tespiti için veri setinde potansiyel verileri temsil etmektedir. Ancak, model performansını iyileştirme, veri karmaşıklığını azaltma, veri saklama ve işleme maliyetini azaltma, anlaşılabilirliği artırma, veri gizliliği gibi konular için hangi özniteliklerin modeliniz için önemli olduğunu belirlemek ve hangi öznitelikleri kullanmanız gerektiğini anlamak için veri analizi ve model eğitimi büyük bir öneme sahiptir. Veri setinde yer alan öznitelikteki bilgilerin alt kümeleri Tablo 3.2’de gösterilmektedir.

Tablo 3.2. Öğrenci sınav istatistikleri

Alan Adı	Veri Tipi	Açıklama
Giriş Sayısı	Int	Öğrencinin toplam giriş sayısı.

Tablo 3.2. Öğrenci sınav istatistikleri tablonun devamı

Alan Adı	Veri Tipi	Açıklama
Yanlış Sayısı/Toplam Soru Sayısı	Float	Yanlış cevaplanan soru sayısının toplam soru sayısına oranı
Doğru Sayısı/Toplam Soru Sayısı	Float	Doğru cevaplanan soru sayısının toplam soru sayısına oranı
Toplam İşlem Sayısı	Int	Toplam işlem sayısı
Soru Başına Düşen İşlem Sayısı	Float	Soru başına düşen ortalama işlem değeri
IP Değişikliği Var Mı?	Boolean	IP değişikliğin olup olmadığı durumu
IP Değişiklik Sayısı	Int	IP değişiklik sayısı
Kopya Durum	Boolean	Kopya çekilip çekilmeme durumu

Tablodaki bilgileri öznitelik ve alt öznitelik olarak ayırdığımızda, nitelikler genel bilgi kategorilerini, öznitelikler ise bu kategorilere ait spesifik verileri temsil eder. Aşağıda bu ayırım yapılmıştır:

Tablo 3.3. Öznitelik alt kümeleri

Öznitelik Kategorisi	Alt Öznitelik	Açıklama
Kullanıcı Bilgileri	Kullanıcı Adı	Kullanıcının adı veya kullanıcı adı
	IP	Kullanıcının IP adresi
	IP Değiştirme Sayısı	Kullanıcının IP değiştirme sayısı
Giriş Bilgileri	İlk Giriş Tarih Saati	Kullanıcının sisteme ilk giriş zamanı
	Son Giriş Tarih Saati	Kullanıcının sisteme son giriş zamanı
	Giriş Sayısı	Kullanıcının sisteme kaç kez giriş yaptığı
	Durumu	Kullanıcının durumu (aktif, pasif vb.)
Performans Bilgileri	Puan	Kullanıcının aldığı toplam puan
	Doğru	Kullanıcının doğru cevapladığı soru sayısı
	Yanlış	Kullanıcının yanlış cevapladığı soru sayısı
	Boş	Kullanıcının boş bıraktığı soru sayısı
	Toplam Soru Sayısı	Kullanıcının toplam soru sayısı
Eylem Bilgileri	Sayfa Kapatma	Kullanıcının sayfa kapatma sayısı
	Sayfa Açma	Kullanıcının sayfa açma sayısı
	Sayfa Yenileme	Kullanıcının sayfa yenileme sayısı
	Toplam İşlem Sayısı	Kullanıcının gerçekleştirdiği toplam işlem sayısı
	Öğrenci İzlediği Yol	Kullanıcının izlediği yol veya işlem adımları
İzleme Bilgileri	Ders	Kullanıcının ilgili olduğu ders
	İşlem Sırası	Kullanıcının işlemlerinin sıralaması
	Öğrenci Adımları	Kullanıcının izlediği adımlar
	Harcanan Dakika	Kullanıcının harcadığı süre (dakika)

Bu veriler, kullanıcının sınav sürecindeki alışılmadık davranışlarını ve etkileşimlerini belirleyerek kopya çekme girişimlerini tespit etmek için kullanılmıştır. Şüpheli aktiviteleri ve anormal davranışları belirlemek, sınav güvenliğini artırmaya ve adil bir değerlendirme süreci sağlamaya yardımcı olacaktır.

3.1.1. Veri toplama ve işleme

Veri toplama ve işleme adımı, çalışmanın en kritik aşamasıdır. Iğdır Üniversitesi'nin ALMS platformundan alınan çevrimiçi sınav sonuçları ve öğrenci aktiviteleri kullanılmıştır. Veriler şu adımlarla hazırlanmıştır:

Etiketleme: Veriler, kopya çekip çekmediğine dair etiketlenmiştir. 'İzlenen Yol' verilerine dayanarak, eğer bir öğrenci sınavı "arkadaşıyla birlikte" ya da "internetten" yaptığı belirtilmişse, bu durumda kopya çektiği kabul edilerek etiketlenmiştir.

Filtreleme: Eksik ve hatalı veriler temizlenmiş ve analiz için uygun hale getirilmiştir. Örneğin, notu -1 olan ya da "İzlenen Yol" bilgisi eksik olan gözlemler dışlanmıştır.

Değişkenler: Öğrenci aktiviteleri (sayfa yenileme, IP adresi değişikliği, sayfa değiştirme, işlem sayısı gibi) bağımsız değişkenler olarak seçilmiş, bağımlı değişken ise öğrencinin kopya çekme durumu olarak belirlenmiştir.

3.2. Sınıflandırma Metotları

3.2.1. Lojistik regresyon metodu

Lojistik regresyon, sınıflandırma problemlerini çözmek için kullanılan istatistiksel bir modeldir. Genellikle iki sınıf (binary) arasında bir ayırım yapmak için kullanılır. Temel olarak, bir olayın olasılığını tahmin etmektedir. Örneğin, bir öğrenci kopya çekmiş mi (1) yoksa çekmemiş mi (0) sorusuna yanıt vermek amacıyla kullanılmıştır.

Temel mantığı ise, doğrusal regresyon gibi bir model kurmaktır, ancak çıktı olarak olasılık verir. Modelin ürettiği olasılık değeri 0 ile 1 arasında yer alır. Bu sayede, "evet" veya "hayır" şeklindeki sınıflandırma problemlerinde kullanılmaktadır. Modelin çıktısı aşağıdaki formülle ifade edilir:

$$P(y = 1 | X) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \dots + \beta_n X_n)}} \quad (3.1)$$

Burada:

- $P(y = 1 | X)$ y'nin 1 olma olasılığı, yani öğrencinin kopya çekme olasılığı.
- $X_1, X_2, \dots, X_n$, bağımsız değişkenlerdir (örneğin, sınav süresi, açılan sekme sayısı).
- $\beta_0, \beta_1, \dots, \beta_n$ ise modelin parametreleridir.

Lojistik regresyon modeli, bağımlı değişkenin ikili (binary) olduğu durumlarda, örneğin bir öğrencinin sınavda kopya çekip çekmediğini belirleme gibi durumlarda, en yaygın kullanılan ve etkili yöntemlerden biri olarak kabul edilir. Bu model, kopya tespiti probleminin doğası gereği sonuçların iki kategoriye (kopya var – kopya yok) ayrıldığı durumlar için oldukça uygundur. Lojistik regresyon, öğrenciye ait veriler (örneğin, sınav süresi, IP adresi, açılan sekme sayısı, sınava giriş tekrar sayısı vb.) kullanılarak öğrencinin kopya çekme olasılığını tahmin eder ve bu olasılığın belirlenen bir eşik değer üstünde olup olmadığını değerlendirerek sınıflandırma yapmaktadır (Alsabhan, 2023).

Lojistik regresyon, bağımlı değişkenin yalnızca iki olasılığa ayrıldığı durumlarda, yani bir öğrencinin kopya çekip çekmediğini belirlemede etkili bir yöntem olarak öne çıkmaktadır. Bu model, bağımsız değişkenlerin kopya çekme olasılığı üzerindeki etkisini belirlemek için etkili bir araç olarak kullanılmaktadır.

Model, öğrencilerin sınav süresi, sınav notları, daha önceki performansları ve diğer ilgili değişkenler gibi faktörleri kullanarak kopya çekme olasılığını tahmin etmektedir. Elde edilen katsayılar, her bir bağımsız değişkenin kopya çekme üzerindeki etkisini net bir şekilde ortaya koymakta ve bu sayede çevrimiçi sınavlar sırasında hangi faktörlerin öğrencilerin davranışlarını etkilediği daha iyi anlaşılabilir.

Elde edilen sonuçları analiz etmek için gerçek yaşam verilerini kullanarak, lojistik regresyonun etkinliğini ve güvenilirliğini ortaya koymaktadır. Bu bağlamda, lojistik regresyon modeli, yüksek doğruluk oranları ile kopya tespitinde güvenilir bir araç olarak kabul edilmektedir. Aynı zamanda, modelin sonuçları, eğitimcilerin sınav güvenliğini artırmak için alabilecekleri stratejik önlemleri daha iyi planlamalarına yardımcı olmaktadır.

3.2.2. Rastgele orman metodu

Rastgele Orman Metodu (Random Forest), makine öğrenmesi ve istatistiksel modelleme alanında sıkça kullanılan güçlü bir sınıflandırma ve regresyon yöntemidir. Rastgele Orman (Random Forest), makine öğrenimi alanında yaygın olarak kullanılan bir denetimli öğrenme yöntemidir ve kopya çekme tespitinde etkili bir araç olarak öne çıkmaktadır. Bu yöntem, birden fazla karar ağacının birleşimiyle oluşturulan bir ansamble modelidir. Her bir karar ağacı, eğitim verileri üzerinde rastgele bir alt küme ile eğitilir ve sonuçta, bu ağaçların tahminleri çoğunluk oyu ile belirlenir. Bu süreç, modelin genelleme yeteneğini artırarak aşırı uyum (overfitting) riskini azaltır ve yüksek doğruluk oranları elde edilmesini sağlar. Gini İndeksi: Rastgele Orman modelinin ağaçlarındaki bölme noktalarını değerlendirmek için kullanılır.

- İki sınıflı bir problemde Gini İndeksi şu şekilde hesaplanır:

$$Gini(D) = 1 - \sum_{i=1}^C p_i^2 \quad (3.2)$$

Burada C , sınıf sayısıdır ve

D : Düğümdeki veri seti

C : Sınıf sayısı (örneğin, kopya var/kopya yok)

p_i : i -inci sınıfın olasılığıdır.

Gini İndeksi, kopya tespitinde karar ağaçları ve makine öğrenimi yöntemlerinde özelliklerin önemini belirlemek, veri setinin homojenliğini değerlendirmek ve model performansını ölçmek için etkili bir araçtır. Bu yöntem, eğitim kurumlarının kopya çekme davranışlarını daha iyi anlamasına ve önleyici tedbirler almasına yardımcı olmaktadır. Rastgele Orman (Random Forest), makine öğreniminde yaygın olarak kullanılan bir denetimli öğrenme algoritmasıdır. Bu model, birden fazla karar ağacının birleşimi ile çalışır ve genellikle sınıflandırma ve regresyon problemlerinde yüksek doğruluk sağlar.

Bu modelde, her bir karar ağacının düğüm bölmelerinde Gini İndeksi hesaplanarak hangi özelliğin en iyi bölmeyi sağladığı belirlenir. Düğümdeki veri seti için Gini İndeksi hesaplanarak, en düşük Gini değeri olan özellik seçilir. Düşük Gini değeri, o düğümdeki verilerin daha homojen olduğunu gösterir; bu da daha iyi bir ayırım sağlar. Eğitim tamamlandıktan sonra, Rastgele Orman modeli test verileri üzerinde tahminler

yapar. Her bir ağaç, verilen özellikler için kopya çekme olasılığını tahmin eder. Nihai tahmin, tüm ağaçların tahminlerinin çoğunluk oyu ile belirlenir:

$$\hat{y} = \underset{j}{\operatorname{argmax}} = \sum_{j=1}^N f_j(x) \quad (3.3)$$

Burada:

$\hat{y}$: Nihai tahmin

$f_j(x)$, i – inci ağaç tarafından yapılan tahmin

N : Toplam karar ağacı sayısı

Kopya çekme tespitinde Rastgele Orman kullanıldığında, çeşitli öğrenci davranış verileri (sınav süreleri, IP adresleri, notlar vb.) kullanılarak model eğitilir. Model, bu veriler aracılığıyla hangi öğrencilerin kopya çekme olasılığının yüksek olduğunu tahmin edebilir. Örneğin, öğrencinin sınav sırasında yaptığı davranışlar, giriş sayısı, sınav süresi gibi faktörler, modelin çıktılarında önemli bir rol oynar. Rastgele Orman ayrıca, hangi özelliklerin kopya çekme üzerinde daha fazla etkili olduğunu belirlemek için özelliklerin önem derecelerini analiz etme yeteneğine de sahiptir.

Çevrimiçi sınavların artan kullanımı ile birlikte akademik dürüstlük sorunlarının önemi artmıştır. Kopya tespiti, eğitim kurumları için kritik bir mesele olup, makine öğrenimi tekniklerinin bu bağlamda sağladığı çözüm potansiyeli dikkat çekmektedir. Özellikle Rastgele Orman algoritması, kopya çekme olasılıklarını tahmin etme konusunda önemli bir araç olarak öne çıkmaktadır. Bu yöntem, önceki çalışmalarda kopya tespitinin zorlukları ve mevcut yöntemlerin sınırlılıkları göz önünde bulundurularak seçilmiştir.

Rastgele Orman algoritması, çevrimiçi sınavlarda öğrenci davranışlarını analiz etmek için kullanılmıştır. Model, belirli bir veri seti üzerinde eğitilmiş ve çeşitli özelliklerin (sınav süreleri, IP adresleri, açılan sekme sayıları vb.) kopya çekme üzerindeki etkileri incelenmiştir. Bu yaklaşım, aşırı uyum riskini azaltarak modelin genelleme yeteneğini artırmıştır.

3.2.3. XGBOOST metodu

XGBoost (Extreme Gradient Boosting), özellikle büyük veri setleri üzerinde yüksek doğruluk ve verimlilik sağlayan bir algoritmadır. Bu algoritmanın temel amacı,

modelin hatalarını minimize etmek ve modelin karmaşıklığını dengelemektir. XGBoost algoritmasında kullanılan matematiksel denklemler, modelin performansını iyileştirmeye yardımcı olan iki ana bileşeni içerir: aşam fonksiyonu ve düzenleme terimi. Amaç fonksiyonu ve düzenleme teriminden (regularization term) oluşur. Bu iki bileşen bir araya gelerek modelin başarısını ve karmaşıklığını dengelemeyi hedefler.

$$Obj = \sum_{i=1}^n L(\hat{y}_i, y_i) + \sum_{k=1}^K \Omega(f_k) \quad (3.4)$$

Burada:

$\hat{y}_i$: Modelin tahmini,

y_i : Gerçek değer,

$(\hat{y}_i, y_i)$: Kayıp fonksiyonu,

$\Omega(f_k)$: Model karmaşıklığı için düzenleme terimi,

K : Toplam karar ağacı sayısıdır.

Kayıp Fonksiyonu (Loss Function): Kayıp fonksiyonu, modelin tahmin ettiği değerler ile gerçek değerler arasındaki farkı ölçer. Regresyon ve sınıflandırma gibi farklı problemler için farklı kayıp fonksiyonları kullanılabilir. Örneğin, sınıflandırma problemlerinde log kayıp (logarithmic loss) kullanılır. Örnek log kayıp fonksiyonu:

$$L(\hat{y}_i, y_i) = -\frac{1}{N} \sum_{i=1}^N (y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)) \quad (3.5)$$

Düzenleme Terimi (Regularization Term): Düzenleme terimi, modelin aşırı karmaşık hale gelmesini ve overfitting yapmasını engeller. XGBoost, karar ağaçlarının yapısını optimize etmek için düzenleme parametreleri kullanır. Düzenleme terimi:

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T \omega_j^2 \quad (3.6)$$

Burada:

T : Ağacın yaprak sayısı,

ω_j : Yaprak ağırlıkları,

γ ve λ : Düzenleme parametreleridir.

3.2.4. K-En yakın komşu metodu

K-En Yakın Komşu (KNN) yöntemi, sınıflandırma problemlerini çözen denetimli öğrenme algoritmaları arasında yer alır. Bu yöntemde, sınıflandırılacak veriler, öğrenme kümesindeki normal davranış verileriyle benzerliklerine göre değerlendirilir ve en yakın olan k komşunun ortalaması alınarak belirlenen eşik değerine göre sınıflara atanır. Burada önemli olan nokta, her sınıfın özelliklerinin önceden açık bir şekilde tanımlanmış olmasıdır. Yöntemin performansı ise, seçilen k komşu sayısı, eşik değerin belirlenmesi, benzerlik ölçüm yöntemi ve öğrenme kümesindeki yeterli sayıda normal davranış örneğinin varlığı gibi faktörlere bağlıdır (Çalışkan ve Soğukpınar, 2008)

K-En Yakın Komşu (KNN) yöntemi, nesneleri sınıflandırmak veya tahmin etmek için n boyutlu bir özellik uzayında en yakın komşularını kullanır. Bu yöntemde, sınıflandırma yapılırken kaç tane en yakın komşunun dikkate alınacağı, k adı verilen bir pozitif tam sayı ile belirlenir. Örneğin, k=1 ise, sınıflandırılan nesne en yakınındaki komşunun sınıfına atanır. Bu yöntem aynı zamanda kesitim analizlerinde de kullanılır.

Komşuların belirlenmesi için, sınıflandırılacak nesne ile eğitim kümesindeki diğer örnekler arasındaki mesafe hesaplanır. Bu mesafeler en kısa olandan en uzağa doğru sıralanır, yani en yakın komşudan en uzak komşuya doğru bir liste oluşur.

KNN'de sınıflandırma yapılırken, verilerin eğitim kümesindeki normal davranışlarına olan benzerlikleri hesaplanır. Daha sonra bu verilerden k tanesinin ortalaması alınır ve belirlenen bir eşik değere göre sınıfa atanır. Yani, sınıflandırılacak yeni veri noktası, daha önce sınıflandırılmış verilerden k tanesine yakın olup olmadığına bakılarak sınıflandırılır. Burada önemli olan, her sınıfın özelliklerinin net bir şekilde tanımlanmış olmasıdır (Köktürk, 2012).

KNN (K-En Yakın Komşu) yöntemi, bir verinin hangi sınıfa ait olduğunu tahmin edebilmek için eğitim setindeki verilerin bağımsız değişkenlerini temsil eden vektörle en yakın konumda bulunan noktaya olan uzaklığın ölçülmesine dayanır. Sınıf belirlemede, genellikle "Öklid uzaklığı" yöntemi tercih edilse de Manhattan ve Minkowski uzaklık ölçüm formülleri de kullanılabilir (Hacıbeyoğlu, 2023).

Öklid uzaklığı, iki veri noktası arasındaki düz bir çizgi mesafesini ölçen en yaygın uzaklık ölçümüdür. İki nokta arasındaki mesafeyi, n boyutlu bir uzayda aşağıdaki formülle hesaplanmaktadır.

$$d(p, q) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \quad (3.7)$$

Burada:

$d(p, q)$: p ve q noktaları arasındaki Öklid uzaklığı

p_i ve q_i : p, q noktalarının i -inci boyuttaki koordinatları

n : Boyut sayısı

Manhattan uzaklığı, iki veri noktası arasındaki mesafeyi, yalnızca dikey ve yatay hareketlerle hesaplayan bir yöntemdir. Bu ölçüm, genellikle "şehir bloğu" mesafesi olarak da adlandırılır.

$$d(p, q) = \sum_{i=1}^n |p_i - q_i| \quad (3.8)$$

Burada:

$d(p, q)$: p ve q noktaları arasındaki Manhattan uzaklığı

p_i ve q_i noktalarının i -inci boyuttaki koordinatları

n : Boyut sayısı

Minkowski uzaklığı, Öklid ve Manhattan uzaklığının genel bir formudur. Bu ölçüm, uzaklık hesaplaması için bir parametre (p) kullanır. Öklid uzaklığı için $p=2$ ve Manhattan uzaklığı için $p=1$ değerini alır.

$$d(p, q) = (\sum_{i=1}^n |p_i - q_i|^p)^{\frac{1}{p}} \quad (3.9)$$

Burada:

$d(p, q)$: p ve q noktaları arasındaki Minkowski uzaklığı

p_i ve q_i : Noktalarının i -inci boyuttaki koordinatları

p : Uzaklık ölçümünün parametresi

n : Boyut sayısı

Bu uzaklık ölçüm yöntemleri, veri analizi, makine öğrenimi ve istatistiksel hesaplamalarda yaygın olarak kullanılır. Uygulama bağlamına göre, farklı durumlarda daha uygun olan uzaklık ölçümünü seçmek önemlidir. Örneğin, Öklid uzaklığı daha doğal mesafeler için uygundur, Manhattan uzaklığı ise dik bir yolda seyahat eden nesneler için daha anlamlıdır. Minkowski uzaklığı ise esnekliği sayesinde farklı mesafe hesaplamaları için kullanılabilir.

3.2.5. Performans ölçütleri

Model performansını değerlendirmek için çeşitli performans ölçütleri kullanılmıştır:

Doğruluk (Accuracy) : Modelin doğru sınıflandırdığı örneklerin oranı.

$$\text{Doğruluk} = \frac{TP+TN}{TP+TN+FP+FN} \quad (3.10)$$

Doğruluk değeri, bir sınıflandırma modelinin doğru tahmin ettiği örneklerin, toplam örnek sayısına oranını temsil etmektedir. Bu değerın hesaplanmasına ilişkin formül, Denklem 10’da verilmiştir.

Burada;

TP : Doğru pozitif (kopya var, model de kopya var demiş)

TN : Doğru negatif (kopya yok, model de kopya yok demiş)

FP : Yanlış pozitif (kopya yok, model kopya var demiş)

FN : Yanlış negatif (kopya var, model kopya yok demiş)

Tablo 3.4. Karışıklık matrisi örneği

	Kopya Çekti	Kopya Çekmedi
Kopya Tespit Edildi	TP (Doğru Pozitif)	FP (Yanlış Pozitif)
Kopya Tespit Edilmedi	FN (Yanlış Negatif)	TN (Doğru Negatif)

Karışıklık Matrisi: Karışıklık Matrisi (confusion matrix), sınıflandırma problemlerinde modelin performansını değerlendirmek için kullanılan bir tablodur. Bu matris, modelin doğru ve yanlış tahminlerini dört farklı kategoride gösterir. Tablo 4’te verilen karışıklık matrisi örneğinde, modelin ne kadar doğru veya hatalı tahminlerde bulunduğunu görsel olarak anlamaya yardımcı olmaktadır. Karışıklık matrisi, sınıflandırma modelinin detaylı performansını analiz etmek ve doğruluk, kesinlik, duyarlılık gibi ölçümleri elde etmek için kullanılmaktadır.

Doğruluk, genel olarak bir modelin ne kadar iyi çalıştığını göstermek için kullanılır. Kopya tespiti gibi bir sınıflandırma probleminde, modelin gerçekte kopya çeken ve çekmeyen öğrencileri ne kadar doğru sınıflandırdığını anlamaya yardımcı olmaktadır. Ancak, doğruluk tek başına yeterli olmayabilir. Özellikle veri setinde sınıf dengesizliği olduğunda yanıltıcı olabilir. Örneğin, 100 öğrenciden sadece 10’u kopya çektiyse ve model tüm öğrencileri kopya çekmedi olarak sınıflandırırsa, %90 doğruluk

elde edebilir. Bu nedenle, kesinlik (precision), duyarlılık (recall) ve F1 skoru gibi diğer ölçütlerle birlikte değerlendirilmesi gerekir.

Hassasiyet (Precision): Sınıflandırma problemlerinde bir modelin doğru pozitif tahminlerinin, tüm pozitif tahminlere oranını ölçen bir performans metriğidir. Yani, modelin yaptığı pozitif tahminlerin ne kadar doğru olduğunu gösterir. Eğer model bir öğrenciyi kopya çekiyor olarak işaretlerse, bu tahminlerin ne kadarının gerçekten doğru olduğunu belirler. Hassasiyetin yüksek olması, modelin kopya tespitinde ne kadar kesin olduğunu ifade eder.

$$\text{Hassasiyet (Precision)} = \frac{TP}{TP+FP} \quad (3.11)$$

Burada:

TP : Doğru pozitif tahminler (gerçekten kopya çeken ve doğru olarak tespit edilen öğrenciler).

FP : Yanlış pozitif tahminler (gerçekten kopya çekmeyen, ancak yanlışlıkla kopya olarak işaretlenen öğrenciler).

Kopya tespitinde hassasiyet, özellikle masum öğrencilerin yanlışlıkla kopya çekiyor olarak işaretlenmesinin önüne geçilmesi açısından kritik bir metrik olarak öne çıkar. Modelin, kopya tespit ederken güvenilir olması ve gereksiz yere çok fazla yanlış alarm vermemesi için yüksek hassasiyet istenir. Ancak, hassasiyetin yüksek olması tek başına yeterli değildir; modelin aynı zamanda duyarlılık (recall) metriklerinde de iyi performans göstermesi gerekir. Bu sayede, model hem kopya çekenleri doğru şekilde tespit eder, hem de gereksiz yanlış pozitif tahminlerden kaçınır.

Duyarlılık (Recall): Sınıflandırma problemlerinde bir modelin, gerçek pozitifleri ne kadar iyi tespit ettiğini gösteren bir performans metriğidir. Yani, modelin pozitif sınıfa ait olan verileri ne kadar başarılı bir şekilde tespit ettiğini gösterir. Yüksek duyarlılık, modelin kopya çeken öğrencileri tespit etme konusunda ne kadar iyi olduğunu belirtir. Kopya çekmiş olan öğrencilerin model tarafından doğru bir şekilde kopya olarak sınıflandırılma oranını ifade eder.

$$\text{Duyarlılık (Recall)} = \frac{TP}{TP+FN} \quad (3.12)$$

Burada:

TP : Doğru pozitifler (gerçekten kopya çekip de doğru bir şekilde tespit edilenler).

FN: Yanlış negatifler (gerçekten kopya çekip de yanlışlıkla kopya çekmediği düşünülenler).

Özgüllük (specificity): Bir sınıflandırma modelinin gerçek negatifleri doğru bir şekilde tanımlama yeteneğini ölçen bir performans metriğidir. Başka bir deyişle, gerçek negatiflerin ne kadarının doğru bir şekilde tahmin edildiğini gösterir.

$$\text{Özgüllük} = \frac{\text{Doğru Negatif}}{\text{Doğru Negatif} + \text{Yanlış Pozitif}} \quad (3.13)$$

Burada:

Doğru Negatif (True Negative – TN): Modelin doğru bir şekilde negatif (örneğin, kopya çekmeyen) olarak sınıflandırdığı örneklerdir.

Yanlış Pozitif (False Positive – FP): Modelin yanlış bir şekilde pozitif (örneğin, kopya çekmediği halde kopya çeken olarak işaretlenen) olarak sınıflandırdığı örneklerdir.

F1-Skoru: Bir modelin performansını değerlendirirken hassasiyet (precision) ve duyarlılık (recall) arasındaki dengeyi ölçen bir metriktir. Özellikle dengesiz veri kümelerinde, yanlış pozitif ve yanlış negatif hataların önemli olduğu durumlarda, F1 skoru daha anlamlı bir ölçüt sunar.

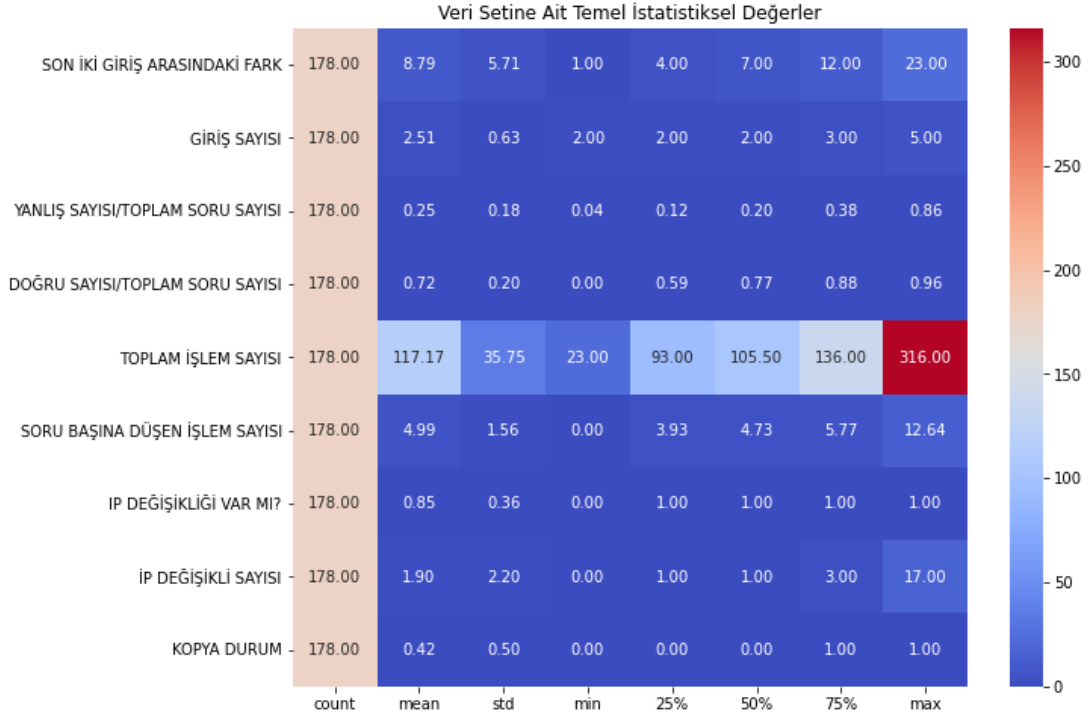
F1 skoru, hassasiyet ve duyarlılığın harmonik ortalaması olarak hesaplanır. F1 skorunun hesaplanmasına ilişkin formül aşağıda şekil 2.14'te verilmektedir.

$$F1 \text{ Skoru} = 2 * \frac{\text{Hassasiyet} * \text{Duyarlılık}}{\text{Hassasiyet} + \text{Duyarlılık}} \quad (3.14)$$

Bu yöntemler sayesinde kopya tespiti daha hassas bir şekilde gerçekleştirilmektedir. Modellerin performansları, kullanılan veri seti üzerinde yapılan çapraz doğrulama ve testler sonucunda analiz edilmiştir.

4. BULGULAR

Bu bölümde kullanılan veri setine yönelik tanıtıcı istatistikler ve makine öğrenmesi yöntemlerinden olan Random Forest, Lojistik regresyon, XGBoost ve KNN yöntemlerine ait sonuçlar verilecektir. Veri setindeki değişkenlere ait tanıtıcı istatistikler aşağıdaki tablo ve Şekil 4.1’de verilmiştir.



Şekil 4.1. Tanıtıcı istatistikler

Veri setindeki değişkenler arasındaki korelasyon Şekil 4.2. de verilmiştir. Korelasyon matrisine göre, bazı değişkenler arasında belirgin ilişkiler gözlemlenmektedir. İşte birkaç dikkat çekici nokta:

1. Yanlış Sayısı/Toplam Soru Sayısı ve Doğru Sayısı/Toplam Soru Sayısı (-0.93): Yanlış cevapların oranı ile doğru cevapların oranı arasında güçlü negatif bir korelasyon bulunmakta, bu da beklenen bir sonuçtur. Bir testte daha fazla doğru cevap alındıkça, yanlış cevapların azalması mantıklı bir sonuçtur.

2. Toplam İşlem Sayısı ve Soru Başına Düşen İşlem Sayısı (0.93): Toplam işlem sayısı ve soru başına düşen işlem sayısı arasında yüksek bir pozitif korelasyon vardır. Bu, toplam işlemlerin artmasıyla, soru başına işlem miktarının da arttığını göstermektedir.

3.IP Değişikliği ve IP Değişikliği Sayısı (0.33): IP değişikliği var mı değişkeni ile IP değişikliği sayısı arasında orta düzeyde bir pozitif ilişki vardır. Bu da IP değişikliklerinin sayısının arttıkça, IP'nin değiştirildiği durumların da arttığını gösterir.

4.Doğru Sayısı/Toplam Soru Sayısı ile Yanlış Sayısı/Toplam Soru Sayısı (-0.93): Doğru ve yanlış cevaplar arasında oldukça güçlü bir negatif korelasyon gözlemlenmektedir. Bu, bir kişinin doğru cevap sayısının artmasıyla yanlış cevap sayısının azalmasıyla ilişkilidir. Bu sonuçlar temel istatistiksel bağlantıları ortaya koymaktadır. Korelasyon 1'e yakınsa güçlü pozitif, -1'e yakınsa güçlü negatif bir ilişki anlamına gelir.

"Kopya Durumu" ile diğer değişkenler arasındaki korelasyonlar genel olarak düşük seviyededir, bu da "Kopya Durumu" değişkeninin diğer değişkenlerle güçlü bir ilişkisi olmadığını gösteriyor. Ancak bazı dikkat çekici noktalar şu şekildedir:

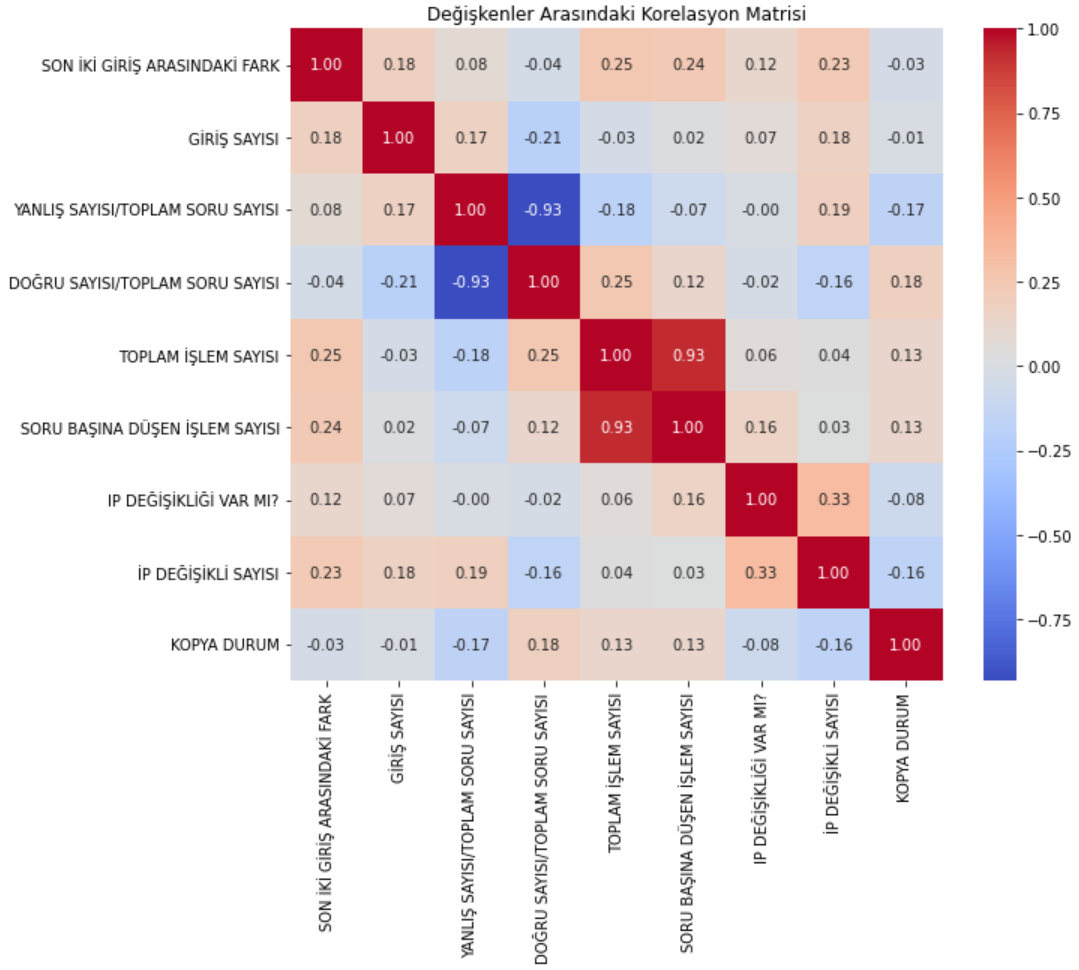
1.Doğru Sayısı/Toplam Soru Sayısı (0.18): Kopya durumu ile doğru cevapların oranı arasında zayıf bir pozitif korelasyon bulunuyor. Bu, kopya durumunun varlığının, doğru cevapların oranını biraz artırabileceğini düşündürebilir. Ancak bu ilişki oldukça zayıf olduğu için anlamlı bir etki yaratmıyor olabilir.

2.Yanlış Sayısı/Toplam Soru Sayısı (-0.17): Kopya durumu ile yanlış cevap oranı arasında zayıf bir negatif korelasyon gözlemleniyor. Kopya durumunun varlığı, yanlış cevapların oranında küçük bir azalmaya işaret edebilir. Bu ilişki zayıf olduğundan, kopya durumu ile yanlış cevap sayısı arasında önemli bir bağ olduğunu söylemek zor.

3.Toplam İşlem Sayısı (0.13): Toplam işlem sayısı ile kopya durumu arasında hafif bir pozitif korelasyon var. Bu durum, kopya yapılan oturumlarda daha fazla işlem yapıldığı anlamına gelebilir. Ancak bu ilişki de zayıf düzeydedir.

4. IP Değişikliği Var mı? (-0.08) ve IP Değişikliği Sayısı (-0.16): Kopya durumu ile IP değişikliği ve IP değişikliği sayısı arasında negatif korelasyonlar var. Bu, kopya yapılan durumlarda IP adresi değiştirilmesinin daha az olabileceğini gösteriyor. Ancak bu da zayıf bir ilişki olup, kesin bir çıkarım yapmak zor.

Sonuç olarak, "Kopya Durumu" ile diğer değişkenler arasındaki korelasyonlar oldukça düşük seviyede. Bu, kopya yapılıp yapılmadığı ile diğer davranışlar veya sonuçlar arasında belirgin bir ilişki olmadığını veya mevcut ilişkilerin çok zayıf olduğunu göstermektedir.



Şekil 4.2. Korelasyon matrisi

Bu tez çalışmasında kullanılan Random Forest, XGboost ve KNN ile elde edilen performans sonuçları Tablo 4.1’de verilmiştir. Bu performans sonuçlarını inceleyerek her bir modelin güçlü ve zayıf yönlerini detaylıca yorumlayalım. Model performanslarını accuracy (doğruluk), precision (kesinlik), ve recall (duyarlılık) metrikleri açısından değerlendireceğiz ve hangi modelin hangi senaryolarda daha iyi olduğunu göreceğiz.

Tablo 4.1. Performans değerleri

Model	accuracy	precision	Recall
Random Forest	0.66292	0.65	0.64
XGBoost	0.58988	0.58	0.58
K-Nearest Neighbors	0.45505	0.43	0.43
Logistic Regression	0.53370	0.49	0.49

Tablo 4.1’e bakıldığında; Random Forest modeli, 0.66292 doğruluk oranıyla en iyi performansı sergileyen modeldir. Bu, modelin veri setindeki genel tahmin başarısının yüksek olduğunu gösterir. Random Forest, ağaç temelli bir yöntem olduğu için çok sayıda karar ağacını kullanarak veri setindeki hataları azaltır ve bu nedenle doğruluk oranı diğer

modellere göre daha yüksektir. Bu model, özellikle özellikler arasındaki karmaşık ilişkileri iyi yakaladığı için güçlü sonuçlar üretir. Precision (Kesinlik) değeri 0.65 ile modelin pozitif sınıfları doğru bir şekilde tahmin edebildiğini gösterir. Pozitif tahmin edilen örneklerin %65'inin gerçekten pozitif olması, modelin yanlış pozitifleri kontrol altında tuttuğunu gösterir. Recall (Duyarlılık) ise 0.64 ile modelin pozitif sınıfları tespit etme yeteneğini ortaya koyar. Bu, pozitif sınıfın çoğunluğunun doğru tahmin edildiği, ancak küçük bir kısmının kaçırıldığı anlamına gelir. Genel olarak, Random Forest dengeli bir modeldir ve veri setinde sınıfları doğru bir şekilde ayırt edebilmiştir.

XGBoost modeli, genellikle karmaşık veri setlerinde güçlü performans gösteren bir modeldir, ancak bu veri setinde doğruluk oranı 0.58988 ile ortalama düzeyde kalmıştır. Bu, modelin sınıflar arasındaki ayrımları yakalamakta zorlandığını gösterir. XGBoost, ağaç temelli bir gradyan artırma algoritması olduğundan, küçük varyasyonları öğrenme eğilimindedir. Ancak bu veri setinde performansının, özellikle overfitting yapmadan sınıfları ayırt etmede yeterli olmadığı görülmektedir. Modelin precision (kesinlik) değeri 0.58 ile, pozitif tahmin edilen örneklerin büyük kısmının doğru olduğunu, ancak yanlış pozitiflerin de dikkate değer bir seviyede olduğunu göstermektedir. Recall (Duyarlılık) değeri de aynı seviyede olup (%58), modelin pozitif sınıfları yakalamada ortalama bir başarı gösterdiğini ifade eder. XGBoost, genellikle veri setlerinde derinlemesine ilişkiyi yakalayabilirken, burada performansı sınırlı kalmış olabilir.

K-Nearest Neighbors (KNN) modeli, 0.45505 doğruluk oranıyla en düşük performansı sergileyen modeldir. KNN, komşuluk ilişkilerine dayalı bir model olduğundan, özellikle veri seti çok boyutlu olduğunda veya sınıflar arasında belirgin ayrımlar olmadığında performansının hızla düştüğü bilinir. Bu düşük doğruluk, modelin sınıflar arasında yeterli ayırım yapamadığını ve veri setinin karmaşıklığı karşısında yetersiz kaldığını gösterir. Precision (Kesinlik) ve recall (duyarlılık) değerleri de sırasıyla 0.43 ile oldukça düşük seviyededir. KNN, pozitif sınıfları tahmin ederken büyük oranda yanlış pozitif üretmiştir. Bu durum, negatif örnekleri pozitif olarak yanlış tahmin ettiğini ve dolayısıyla modelin yanlış pozitiflere karşı savunmasız olduğunu gösterir. Recall da aynı seviyede kalmıştır, yani model pozitif sınıfların çoğunu tespit edememiştir. KNN bu veri seti için uygun bir model olmayabilir, özellikle de veri noktalarının birbirine yakın olduğu veya sınıflar arasında karmaşık ilişkilerin bulunduğu durumlarda.

Logistic Regression modeli, 0.53370 doğruluk oranıyla ortalama bir performans sergilemiştir. Logistic Regression doğrusal bir model olduğu için, sınıflar arasında

doğrusal ilişkilerin olduğu veri setlerinde iyi performans gösterir. Ancak bu veri setinde, sınıflar arasında doğrusal olmayan karmaşık ilişkiler olduğu için modelin sınırlı kaldığı gözlemlenmektedir. Logistic Regression daha basit ve hızlı bir model olsa da bu veri setinde sınıfları ayırt etmekte zorlandığı açıkça görülmektedir. Precision (Kesinlik) değeri 0.49, modelin pozitif sınıfları tahmin ederken yanlış pozitiflerin sıkça yapıldığını gösterir. Yani, negatif sınıflar pozitif olarak yanlış tahmin edilmiştir. Recall (Duyarlılık) değeri de 0.49 ile aynı seviyededir. Bu, modelin pozitif sınıfları tespit etmekte sınırlı kaldığını ve birçok pozitif örneği kaçırdığını gösterir. Logistic Regression, karmaşık veri setlerinde sınıflar arasında yeterince güçlü bir ayırım yapamadığından dolayı düşük performans sergilemiştir.

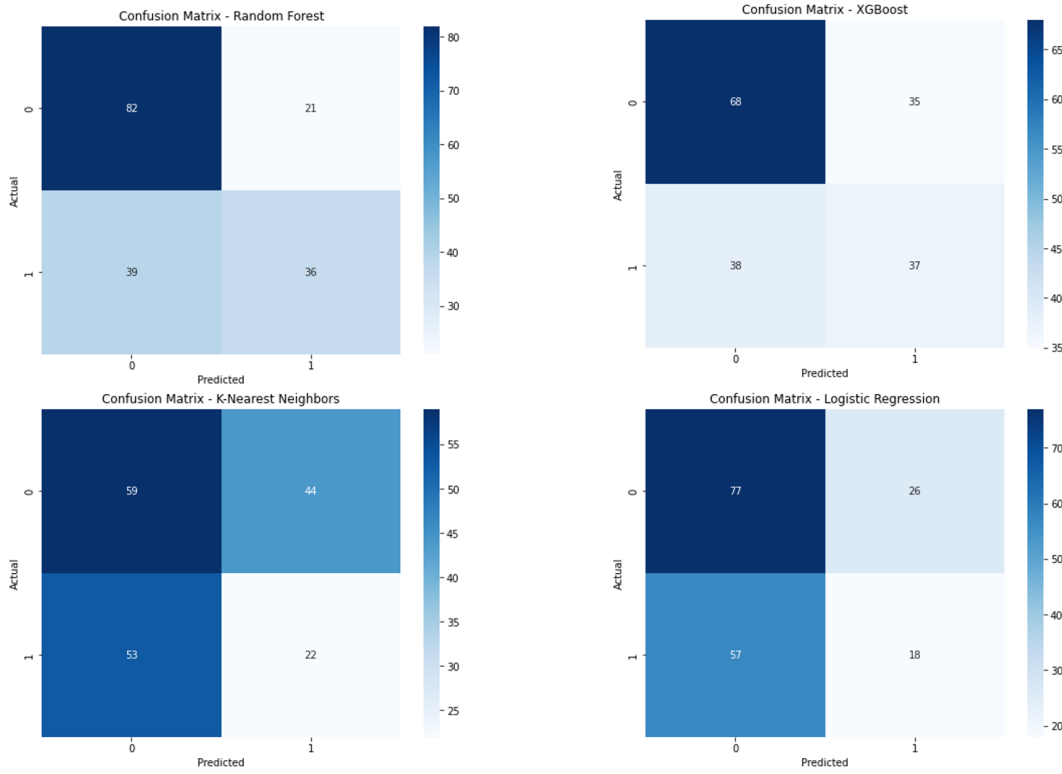
Genel olarak, Random Forest modeli, en dengeli ve en yüksek performansı sunan model olmuştur. Doğruluk, kesinlik ve duyarlılık değerlerinde diğer modellere göre daha başarılıdır. XGBoost ise karmaşık veri setlerinde güçlü bir model olmasına rağmen, bu veri setinde ortalama bir performans göstermiştir. KNN ve Logistic Regression modelleri ise, bu veri setinde sınıflar arasında yeterli ayırım yapamamış ve düşük doğruluk oranları ile zayıf performans sergilemişlerdir. Özellikle KNN, yüksek boyutlu ve karmaşık veri setlerinde daha düşük performans gösterir. Logistic Regression ise doğrusal olmayan ilişkileri yakalayamaması nedeniyle sınıflar arasında başarılı bir ayırım yapamamıştır. Kullanılan 4 modele ait karışıklık matrisleri Şekil 4.3'te verilmiştir. Tüm modeller için elde edilen karışıklık matrisleri ve performans değerlendirmelerini sayısal sonuçlara dayalı olarak bakıldığında;

Random Forest modeli, %61.23 doğruluk, %60.52 precision (kesinlik), ve %61.23 recall (duyarlılık) ile genel olarak dengeli bir performans sergilemiştir. Model, pozitif sınıfları doğru şekilde tespit etmede başarılı olmuş ancak bazı yanlış pozitifler ve negatifler gözlemlenmiştir. Karışıklık matrisine baktığımızda, modelin 45 pozitif sınıfı doğru şekilde tahmin ettiğini (True Positives) ve 75 negatif sınıfı doğru tanımladığını (True Negatives) görüyoruz. Bununla birlikte, 28 negatif sınıf örneğinin yanlış pozitif (False Positives) olarak sınıflandırılması, modelin negatif sınıfları ayırt etmekte zorlandığı bazı durumlar olduğunu göstermektedir. Ayrıca, 30 pozitif örneğin negatif olarak tahmin edilmesi (False Negatives), modelin pozitif sınıfların bir kısmını kaçırdığını gösteriyor. Genel olarak, Random Forest sınıflar arasındaki ayrımı iyi yapabilmiş, özellikle veri setindeki dengesizliğe karşı dayanıklı olmuştur. Ancak modelde pozitif sınıflar için hatalar nispeten daha fazla olduğu için bu sonuçları daha da iyileştirmek mümkündür.

XGBoost modeli, doğruluk ve diğer performans ölçütlerinde Random Forest'a göre biraz daha düşük performans sergilemiştir. %58.98 doğruluk, %58 precision ve %58 recall oranlarına sahip olan model, sınıflar arasında makul bir ayrım yapabilmiş ancak bazı zorluklar yaşamıştır. Özellikle karışıklık matrisinde gözlemlenen yanlış pozitif ve yanlış negatif değerler, modelin pozitif ve negatif sınıfları ayırt etmede sorunlar yaşadığını gösteriyor. Modelin negatif sınıfları pozitif olarak yanlış tahmin etme eğiliminde olduğu gözlemlenmiştir. Buna ek olarak, pozitif sınıfların bir kısmı negatif olarak yanlış sınıflandırılmıştır. XGBoost, genellikle güçlü bir sınıflandırıcı olarak bilinse de bu veri setinde sınıflar arası karmaşıklıkları tam anlamıyla yakalayamamış olabilir. Yanlış pozitif ve negatif oranlarının azaltılması, modelin performansını artırabilir.

K-Nearest Neighbors (KNN) modeli, %45.50 doğruluk oranı ile en düşük performans sergileyen model olmuştur. Model, pozitif ve negatif sınıflar arasında yeterli ayrımı yapamamış ve yüksek sayıda yanlış pozitif (False Positives) ve yanlış negatif (False Negatives) üretmiştir. Precision ve recall oranlarının %43 seviyesinde olması, modelin pozitif sınıfı doğru tahmin etmede ve negatif sınıfı pozitif olarak tahmin etmemede büyük zorluklar yaşadığını gösterir. Karışıklık matrisinde, pozitif sınıfların çoğunluğunun yanlış sınıflandırıldığı ve negatif sınıfların da hatalı olarak pozitif sınıf olarak tahmin edildiği görülmektedir. KNN, yüksek boyutlu ve karmaşık veri setlerinde performansının düşmesiyle bilinir ve bu veri setinde de bu zorlukları yaşamıştır. Modelin düşük performansı, KNN'nin sınıflar arası mesafeye dayalı bir yöntem olduğu ve veri setindeki ayrımların net olmaması durumunda zorlandığını gösterir.

Logistic Regression modeli, doğrusal ilişkilerle çalışmasına rağmen, %53.37 doğruluk oranıyla veri setinde ortalama bir performans sergilemiştir. Model, pozitif ve negatif sınıflar arasındaki ayrımı tam anlamıyla yapamamış ve karışıklık matrisinde gözlemlendiği üzere yanlış pozitif ve yanlış negatif oranları yüksek kalmıştır. Precision (%49) ve recall (%49) değerlerinin düşük olması, modelin pozitif sınıfları tahmin etmede ve negatif sınıfları doğru sınıflandırmada zorlandığını gösterir. Modelin özellikle negatif sınıfları pozitif olarak tahmin etme eğiliminde olduğu ve pozitif sınıfların önemli bir kısmını kaçırdığı gözlemlenmiştir. Logistic Regression, sınıflar arasında doğrusal olmayan ilişkiler bulunan veri setlerinde performansının düşmesiyle bilinir. Bu veri setinde de doğrusal olmayan yapılar olduğundan, model sınıfları doğru bir şekilde ayırt edememiş ve düşük performans sergilemiştir.



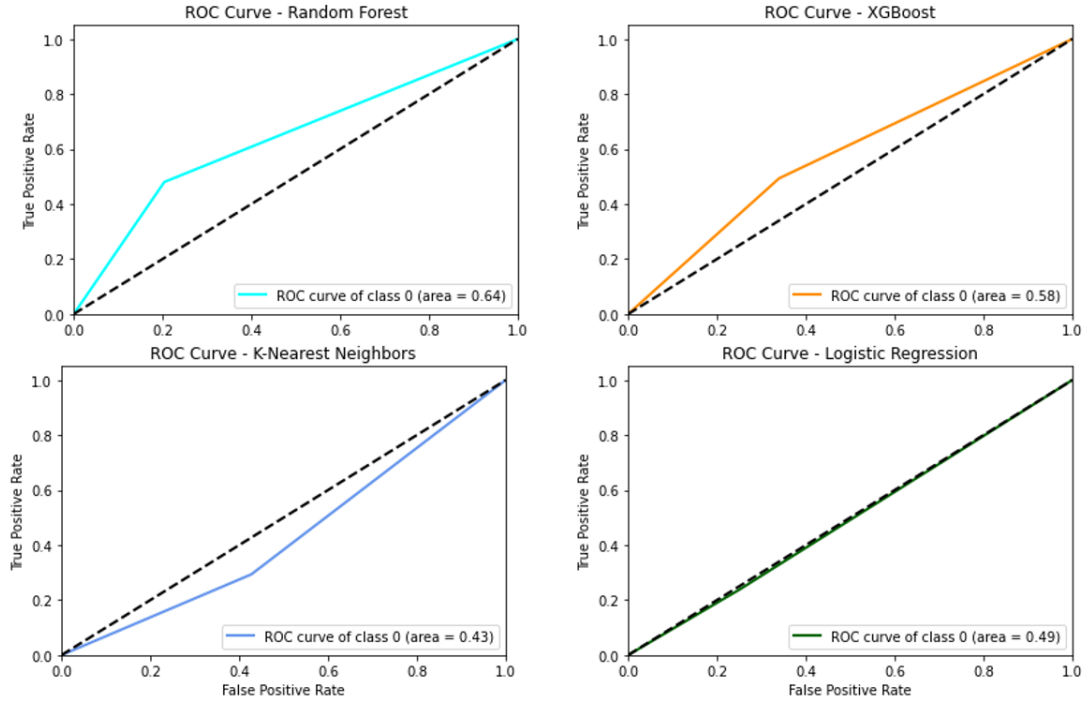
Şekil 4.3. Karışıklık matrisleri

Modellere ait elde edilen ROC eğrileri Şekil 4.4'te verilmiştir. Her bir model için elde edilen ROC eğrileri üzerinden model performansını değerlendirelim. ROC eğrisi, modelin sınıflar arasındaki ayırım yapma yeteneğini gösterir. Eğrinin altındaki alan (AUC - Area Under the Curve), modelin pozitif ve negatif sınıfları ayırt etme başarısının genel bir özetini sunar. ROC eğrisi daha yukarıda ve sola yakınsa, modelin performansı o kadar iyidir.

Random Forest modeli için ROC eğrisi, pozitif ve negatif sınıfları ayırt etme konusunda iyi bir performans sergilemiştir. Eğrinin yukarıda ve sola yakın olduğu gözlemlenmiştir, bu da modelin pozitif sınıfları yanlış pozitif oranını düşük tutarak ayırt edebildiğini gösterir. AUC değeri yüksek olduğunda, modelin genel olarak sınıflar arasında güçlü bir ayırım yaptığını gösterir. Bu durum, Random Forest modelinin sınıflar arasında başarılı bir ayırım yapabildiğini ve özellikle yanlış pozitiflerin düşük olduğunu göstermektedir. Eğrinin doğrusal olmayan doğası, Random Forest'ın pozitif sınıfları yüksek doğrulukla yakaladığını ortaya koyar. Ancak eğrinin bazı bölgelerinde küçük sapmalar olması, modelin bazı pozitif sınıfları kaçırdığını veya yanlış negatif oranının yer yer yükseldiğini gösterir. XGBoost modeli için ROC eğrisi, Random Forest'a göre daha düşük bir AUC değeri sunmuş olabilir. Eğri, özellikle düşük yanlış pozitif oranlarında daha dik başlayarak pozitif sınıfları iyi ayırt etmeye başlamış, ancak modelin pozitif ve

negatif sınıflar arasında tamamen net bir ayrım yapamadığı bölgeler de bulunmaktadır. Bu, pozitif sınıfları yakalama konusunda bazı zorluklarla karşılaştığını ve yanlış pozitif tahminlerin yükseldiği durumlar olduğunu gösterir. AUC değeri, XGBoost modelinin sınıfları ayırıştırma başarısının orta seviyede olduğunu ve yanlış pozitiflerin yer yer arttığını ortaya koymaktadır. Model, ROC eğrisi ile pozitif sınıfları ayırt etmeye çalışırken, bazı negatif sınıfların yanlış pozitif olarak sınıflandırıldığı gözlemlenmiştir. KNN modeli için ROC eğrisi, düşük AUC değerine sahip olabilir ve eğri genellikle 45 dereceye daha yakın bir eğim göstermektedir. Bu durum, modelin pozitif ve negatif sınıfları ayırt etmede başarısız olduğunu ve rastgele tahmin yapmaya yakın bir performans sergilediğini gösterir. KNN'nin sınıflar arasında belirgin bir ayrım yapamadığı ve yanlış pozitif oranının yüksek olduğu bu eğriden anlaşılabılır. ROC eğrisi, modelin birçok pozitif sınıfı kaçırdığını ve negatif sınıfları yanlış pozitif olarak tahmin etme eğiliminde olduğunu göstermektedir. Bu, modelin genel olarak sınıflar arasında çok iyi ayrım yapamadığını ve pozitif sınıfları yakalamakta zayıf kaldığını gösterir. ROC eğrisi üzerindeki düşük performans, modelin sınıflar arasındaki ilişkileri yakalayamadığını açıkça ortaya koyar. Logistic Regression için ROC eğrisi, doğrusal sınırlamalara sahip olmasına rağmen, eğrinin sola yakın bölgelerde olduğu yerlerde pozitif sınıfları ayırt etmede orta düzeyde bir başarı göstermiş olabilir. AUC değeri, modelin sınıflar arasında ayrım yapma becerisinin sınırlı olduğunu gösterir. Eğrinin bazı bölgelerinde, yanlış pozitif oranı yüksek olabilir, bu da negatif sınıfların yanlış bir şekilde pozitif olarak sınıflandırıldığını ortaya koyar. ROC eğrisinde bu sapmalar, Logistic Regression'ın sınıflar arasında güçlü bir ayrım yapamadığını ve pozitif sınıfların bir kısmını kaçırdığını gösterir. Eğri genellikle rastgele tahmin çizgisine yakın olabilir, bu da modelin doğrusal olmayan ilişkilerde zorlandığını ve tahminlerinde kararsız kaldığını işaret eder.

ROC eğrilerine dayanarak, Random Forest ve XGBoost modelleri pozitif ve negatif sınıfları ayırt etmede en başarılı modeller olarak öne çıkarken, KNN ve Logistic Regression daha düşük performans sergilemiştir. ROC eğrisi üzerinden yapılan bu değerlendirme, modellerin yanlış pozitif oranlarını kontrol etme yeteneklerine dayalı olarak genel performanslarını özetlemektedir.



Şekil 4.4. ROC eğrileri

Kopya durumu üzerinde özniteliklerin etkisini gözlemleyebilmek için veri setinden sırası ile değişkenler çıkarılarak kalan öznitelikler ile tüm makine öğrenmesi yöntemler ile Kopya durumu tahmini gerçekleştirilmiştir. Sınıflandırma işlemleri K=5 çapraz geçerlilik yöntemine göre gerçekleştirilmiştir. Elde edilen sonuçlar Tablo 4.2’de verilmiştir.

Dört farklı model için her bir öznitelik çıkarıldıktan sonra kalan özniteliklerle yapılan sınıflandırma işlemi sonucunda elde edilen accuracy (doğruluk), precision (kesinlik) ve recall (duyarlılık) metriklerini göstermektedir.

Lojistik regresyon modelinde, "IP DEĞİŞİKLİĞİ VAR MI?" özelliği çıkarıldığında en yüksek doğruluk oranı olan %57.87 elde edilmiştir. Bu durum, IP değişikliği bilgisinin modelin tahmin yeteneğinde çok önemli olmadığını ve çıkarıldığında performansta belirgin bir düşüş yaşanmadığını gösterir. Ancak "GİRİŞ SAYISI" özelliği çıkarıldığında doğruluk oranı %53.37’ye düşmüş ve modelin performansı zayıflamıştır. Giriş sayısının modelin başarısında kritik bir öznitelik olduğu anlaşılmaktadır. Genel olarak, lojistik regresyon modelinin performansı özniteliklere bağlı olarak değişiklik göstermektedir, ancak bazı özelliklerin çıkarılması (örneğin IP değişikliği) performansı iyileştirebilir.

Random Forest modeli, "GİRİŞ SAYISI" çıkarıldığında en yüksek doğruluk oranını (%65.73) elde etmiştir. Giriş sayısının çıkarılması, modelin karmaşıklığını azaltarak daha başarılı tahminler yapmasına yardımcı olmuş olabilir. Ancak "DOĞRU

SAYISI/TOPLAM SORU SAYISI" çıkarıldığında doğruluk oranı %56.74'e düşmüş, bu da doğru sayısının model için önemli bir bilgi taşıdığını göstermektedir. Random Forest modeli genel olarak yüksek performans sergilemiş ve çıkarılan bazı özniteliklerle daha iyi sonuçlar elde edebilmiştir.

XGBoost modelinde en yüksek doğruluk oranı (%59.55), "YANLIŞ SAYISI/TOPLAM SORU SAYISI" çıkarıldığında elde edilmiştir. Bu durum, modelin yanlış sayısına fazla duyarlı olabileceğini ve bu bilginin çıkarılmasının modelin performansını artırabileceğini göstermektedir. Bununla birlikte, "DOĞRU SAYISI/TOPLAM SORU SAYISI" çıkarıldığında doğruluk %54.49'a düşmüştür, bu da doğru sayısının XGBoost modeli için önemli bir öznitelik olduğunu göstermektedir. XGBoost, genel olarak %54 ile %59 arasında bir doğruluk oranı ile çalışmıştır ve özniteliklerin model performansı üzerinde belirgin bir etkisi olmuştur.

K-Nearest Neighbors (KNN) modeli ise diğer modellere kıyasla daha düşük bir performans sergilemiştir. En yüksek doğruluk oranı olan %55.06, "TOPLAM İŞLEM SAYISI" çıkarıldığında elde edilmiştir, bu da toplam işlem sayısının modelin gereksiz karmaşıklığa neden olabileceğini göstermektedir. Ancak "YANLIŞ SAYISI/TOPLAM SORU SAYISI" çıkarıldığında doğruluk oranı %45.50'ye düşmüş ve modelin performansı önemli ölçüde zayıflamıştır. KNN modelinin genel performansı düşük kalmış ve özellikle çıkarılan özniteliklerle birlikte tutarlı bir başarı sergileyememiştir.

Tablo 4.2. Öznitelik bazlı performans ölçütleri

Model	Removed Feature	Accuracy	Precision	Recall
Lojistik Regresyon	Son İki Giriş Arasındaki Fark	0.573033708	0.55296577	0.573033708
Lojistik Regresyon	Giriş Sayısı	0.533707865	0.507198265	0.533707865
Lojistik Regresyon	Yanlış Sayısı/Toplam Soru Sayısı	0.539325843	0.508380458	0.539325843
Lojistik Regresyon	Doğru Sayısı/Toplam Soru Sayısı	0.550561798	0.522880981	0.550561798
Lojistik Regresyon	Toplam İşlem Sayısı	0.533707865	0.502419507	0.533707865
Lojistik Regresyon	Soru Başına Düşen İşlem Sayısı	0.556179775	0.532667358	0.556179775
Lojistik Regresyon	IP Değişikliği Var Mı?	0.578651685	0.558703794	0.578651685
Lojistik Regresyon	IP Değişikli Sayısı	0.561797753	0.53073157	0.561797753
Random Forest	Son İki Giriş Arasındaki Fark	0.617977528	0.613239602	0.617977528
Random Forest	Giriş Sayısı	0.657303371	0.653185649	0.657303371
Random Forest	Yanlış Sayısı/Toplam Soru Sayısı	0.617977528	0.612302185	0.617977528
Random Forest	Doğru Sayısı/Toplam Soru Sayısı	0.56741573	0.560780862	0.56741573
Random Forest	Toplam İşlem Sayısı	0.617977528	0.612302185	0.617977528

Tablo 4.2. Öznitelik bazlı performans ölçütleri tablonun devamı

Model	Removed Feature	Accuracy	Precision	Recall
Random Forest	Soru Başına Düşen İşlem Sayısı	0.601123596	0.594381406	0.601123596
Random Forest	IP Değişikliği Var Mı?	0.623595506	0.616974541	0.623595506
Random Forest	IP Değişikli Sayısı	0.640449438	0.6339102	0.640449438
XGBoost	Son İki Giriş Arasındaki Fark	0.58988764	0.586387551	0.58988764
XGBoost	Giriş Sayısı	0.573033708	0.576240287	0.573033708
XGBoost	Yanlış Sayısı/Toplam Soru Sayısı	0.595505618	0.591445077	0.595505618
XGBoost	Doğru Sayısı/Toplam Soru Sayısı	0.54494382	0.545771443	0.54494382
XGBoost	Toplam İşlem Sayısı	0.578651685	0.57942278	0.578651685
XGBoost	Soru Başına Düşen İşlem Sayısı	0.578651685	0.576425694	0.578651685
XGBoost	IP Değişikliği Var Mı?	0.578651685	0.575010404	0.578651685
XGBoost	IP Değişikli Sayısı	0.578651685	0.576425694	0.578651685
K-Nearest Neighbors	Son İki Giriş Arasındaki Fark	0.5	0.497248428	0.5
K-Nearest Neighbors	Giriş Sayısı	0.47752809	0.472616075	0.47752809
K-Nearest Neighbors	Yanlış Sayısı/Toplam Soru Sayısı	0.45505618	0.44527488	0.45505618
K-Nearest Neighbors	Doğru Sayısı/Toplam Soru Sayısı	0.45505618	0.44527488	0.45505618
K-Nearest Neighbors	Toplam İşlem Sayısı	0.550561798	0.5489464	0.550561798
K-Nearest Neighbors	Soru Başına Düşen İşlem Sayısı	0.45505618	0.44527488	0.45505618
K-Nearest Neighbors	IP Değişikliği Var Mı?	0.449438202	0.438255788	0.449438202
K-Nearest Neighbors	IP Değişikli Sayısı	0.483146067	0.470826064	0.483146067

Genel olarak, dört farklı modelin performansı çıkarılan özniteliklere göre değişiklik göstermektedir. Random Forest ve XGBoost modelleri en yüksek doğruluk oranlarına sahip olmuş ve çıkarılan özniteliklere rağmen genellikle güçlü performans göstermişlerdir. Özellikle Random Forest, birçok durumda özniteliklerin çıkarılmasına rağmen tutarlı bir başarı elde edebilmiştir ve çıkarılan bazı özniteliklerle modelin karmaşıklığı azalarak daha iyi sonuçlar üretmiştir. XGBoost da benzer şekilde özellikle yanlış sayısına dayalı öznitelikler çıkarıldığında performansını artırmıştır. Lojistik Regresyon modeli, özellikle IP ile ilgili öznitelikler çıkarıldığında daha iyi performans göstermiştir. Bununla birlikte, bazı temel özniteliklerin çıkarılması modelin başarısını olumsuz etkileyerek doğruluk oranını düşürmüştür. Bu, modelin doğrusal ilişkileri iyi işleyebilmesine rağmen bazı özniteliklere bağımlı olabileceğini göstermektedir. K-

Nearest Neighbors (KNN) modeli ise en düşük performansı sergilemiştir. Çıkarılan özniteliklerle modelin başarısı belirgin şekilde azalmış ve özellikle yanlış sayısı gibi öznitelikler çıkarıldığında modelin doğruluğu ciddi şekilde düşmüştür. KNN'nin düşük performansı, modelin veri setindeki karmaşıklığı yeterince iyi ayırt edemediğini ve daha basit veri setlerinde daha etkili olabileceğini göstermektedir.

Sonuç olarak, Random Forest ve XGBoost, öznitelik çıkarılmasına en dayanıklı modeller olarak öne çıkmaktadır. Lojistik Regresyon modeli ortalama bir performans sunarken, KNN özellikle karmaşık veri setlerinde yetersiz kalmıştır. Çıkarılan özniteliklerin model performansı üzerindeki etkisi, her modelin veri setindeki farklı ilişkileri nasıl işlediğine dair önemli ipuçları sunmaktadır.

Tabloya ve modellerin sonuçlarına bakıldığında, en etkili ve performansı en çok etkileyen öznitelikler model bazlı olarak şu şekilde değerlendirilebilir:

Lojistik regresyon modelinde en etkili özniteliklerden biri "GİRİŞ SAYISI" olmuştur. Bu özellik çıkarıldığında modelin doğruluğu %53.37'ye düşmüştür, bu da giriş sayısının model için kritik bir bilgi olduğunu göstermektedir. Ayrıca "IP DEĞİŞİKLİĞİ VAR MI?" özneliğinin çıkarılmasıyla modelin performansı %57.87'ye çıkmış, bu da IP ile ilgili bilgilerin model için çok önemli olmadığını göstermektedir. Lojistik regresyon için giriş sayısı, modelin doğruluğunu belirgin şekilde etkileyen bir özniteliktir. Random Forest modelinde "GİRİŞ SAYISI" çıkarıldığında performansın %65.73 ile en yüksek seviyeye çıktığı görülmüştür. Bu da giriş sayısının modelin karmaşıklığını artırarak bazı durumlarda tahmin doğruluğunu düşürdüğünü ve çıkarılmasının performansı artırdığını gösterir. Buna karşın, "DOĞRU SAYISI/TOPLAM SORU SAYISI" çıkarıldığında doğruluk %56.74'e düşmüş, bu da doğru sayısının model için önemli bir bilgi kaynağı olduğunu göstermektedir. Random Forest modeli, doğru sayısı gibi belirli özniteliklere duyarlı olup, bu özniteliklerin çıkarılması performansı olumsuz etkileyebilir. XGBoost modelinde en etkili özniteliklerden biri "YANLIŞ SAYISI/TOPLAM SORU SAYISI" olmuştur. Bu özellik çıkarıldığında doğruluk %59.55 ile en yüksek seviyeye ulaşmıştır. Yanlış sayısına dayalı bilgilerin çıkarılması modelin performansını artırmış olabilir, bu da modelin fazla hata bilgisine duyarlılığı olduğunu gösterir. Öte yandan, "DOĞRU SAYISI/TOPLAM SORU SAYISI" çıkarıldığında doğruluk %54.49'a düşmüş, bu da doğru sayısının model için kritik bir öznitelik olduğunu ortaya koymaktadır. XGBoost, doğru ve yanlış sayıları gibi metriklere duyarlı bir modeldir ve bu özniteliklerin çıkarılması modelin doğruluğunu doğrudan etkileyebilir. KNN modelinde en etkili özniteliklerden biri "TOPLAM İŞLEM SAYISI" olmuştur. Bu öznitelik çıkarıldığında

modelin doğruluğu %55.06 ile en yüksek seviyeye çıkmıştır. Toplam işlem sayısının çıkarılması, modelin tahmin yeteneğini artırmış olabilir, çünkü bu öznelik KNN için gereksiz karmaşıklık yaratmış olabilir. "YANLIŞ SAYISI/TOPLAM SORU SAYISI" ise modelin performansını en çok düşüren öznelik olmuştur. Bu özellik çıkarıldığında doğruluk %45.50'ye kadar düşmüştür, bu da yanlış sayısının modelin sınıflandırma başarısı üzerinde büyük bir etkisi olduğunu göstermektedir. Modeller genel olarak farklı özneliklere duyarlı olsa da "GİRİŞ SAYISI", "DOĞRU SAYISI/TOPLAM SORU SAYISI", ve "YANLIŞ SAYISI/TOPLAM SORU SAYISI" öznelikleri, çoğu modelde performansı belirgin şekilde etkileyen kritik öznelikler olarak öne çıkmaktadır. Bu özneliklerin çıkarılması, modellerin doğruluğunu büyük ölçüde artırabilir veya düşürebilir. Özellikle doğru ve yanlış sayıları gibi temel öznelikler, sınıflandırma modellerinin performansını doğrudan etkileyen unsurlar olmuştur.

5. SONUÇLAR VE ÖNERİLER

Bu tez çalışmasında, yapay zeka modelleri kullanılarak öğrencilerin sınav süreçleri analiz edilmekte ve kopya çekme girişimleri tespit edilmeye çalışılmıştır. Eğer bir öğrencinin sınav sırasında gösterdiği davranışlar ve sınav sonuçları başka bir öğrenciyle büyük benzerlik gösterirse, kopya çekme durumu pozitif olarak değerlendirilmiştir. Ancak, öğrencinin sınav hareketlerinde herhangi bir kopya çekme veya hileli bir davranış bulunmadığı tespit edilirse, bu durumda kopya çekme tespiti negatif olarak sonuçlandırılmıştır.

5.1 Sonuçlar

Çalışmada elde edilen önemli sonuçlar aşağıda tartışılmıştır.

- Veri setindeki değişkenler arasında yapılan korelasyon analizinde, doğru ve yanlış cevap oranları arasında beklenen güçlü negatif bir ilişki (-0.93) gözlemlendi. Bu ilişki, testte daha fazla doğru cevap verildikçe yanlış cevapların doğal olarak azalmasını açıklar. Toplam işlem sayısı ve soru başına düşen işlem sayısı arasında yüksek bir pozitif korelasyon (0.93) bulunurken, kopya durumu ile diğer değişkenler arasındaki korelasyonlar oldukça zayıf kalmıştır. Örneğin, kopya durumu ile doğru cevap oranı arasında zayıf bir pozitif korelasyon (0.18) bulunmuştur.
- Random Forest, XGBoost, K-Nearest Neighbors (KNN) ve Lojistik Regresyon modelleri kullanılarak elde edilen sonuçlara göre, Random Forest modeli %66.29 doğruluk oranı ile en iyi performansı göstermiştir. Precision (kesinlik) ve recall (duyarlılık) değerlerinde de diğer modellere kıyasla dengeli sonuçlar sunmuştur. XGBoost modeli %58.98 doğruluk oranı ile ortalama düzeyde performans gösterirken, KNN modeli %45.50 doğruluk oranı ile en düşük performansı sergilemiştir. Lojistik regresyon ise %53.37 doğruluk ile KNN'den biraz daha iyi sonuç vermiştir.
- ROC eğrilerine göre, Random Forest modeli en iyi AUC (eğrinin altındaki alan) değerine sahip olup, pozitif ve negatif sınıfları ayırt etmede başarılı olmuştur. XGBoost modelinde ise AUC değeri Random Forest'a göre biraz daha düşük kalmıştır, ancak sınıfları ayırma konusunda makul bir başarı elde etmiştir. KNN

ve Lojistik Regresyon ise sınıflar arasındaki ayrımı net bir şekilde yapamamış ve daha düşük performans sergilemiştir.

- Modeller üzerindeki öznitelik çıkarma analizlerine bakıldığında, özellikle "giriş sayısı" gibi bazı özniteliklerin çıkarılması modellerin performansını önemli ölçüde etkilemiştir. Random Forest modelinde giriş sayısının çıkarılması, modelin karmaşıklığını azaltarak %65.73 doğruluk ile en yüksek performansı sağlamıştır. Aynı zamanda, XGBoost modelinde yanlış sayısı çıkarıldığında doğruluk %59.55 ile en yüksek seviyeye ulaşmıştır.
- Random Forest ve XGBoost modelleri, özniteliklerin çıkarılmasına en dayanıklı modeller olarak öne çıkmışlardır. Özellikle Random Forest, öznitelik çıkarılmasına rağmen tutarlı ve yüksek performans göstermeye devam etmiştir. Diğer yandan, KNN modeli düşük performans sergilemiş, özellikle yanlış sayısına duyarlılığı yüksek olduğu için performansı ciddi şekilde etkilenmiştir.

5.2 Öneriler

- Daha karmaşık modeller kullanılabilir. Mevcut modellerle elde edilen sonuçlar, özellikle veri setindeki karmaşık ilişkileri yeterince yakalayamayan KNN ve lojistik regresyon gibi modellerde düşük performansla sonuçlanmıştır. Bu nedenle, Random Forest ve XGBoost gibi güçlü modellerin yanı sıra, derin öğrenme tabanlı yöntemler (örneğin, yapay sinir ağları veya derin orman modelleri) kullanılarak daha iyi performans elde edilebilir. Ayrıca, bu yöntemlerle sınıf dengesizliklerine daha etkin çözümler bulunabilir.
- Modellerin performansını artırmak için öznitelik mühendisliğinin daha dikkatli ve detaylı yapılması önerilir. Özellikle doğruluk oranları üzerinde etkili olan öznitelikler üzerinde daha fazla çalışma yapılmalı ve yeni özniteliklerin eklenmesiyle modellerin tahmin yeteneği artırılmalıdır. Örneğin, IP değişikliği gibi bazı öznitelikler çıkarıldığında performansın arttığı gözlemlenmiş, bu tür öznitelikler üzerine daha fazla odaklanılabilir.
- XGBoost gibi karmaşık modeller, bazen overfitting (aşırı öğrenme) problemiyle karşılaşabilir. Bu nedenle, model hiperparametrelerinin optimize edilmesi, çapraz doğrulama yöntemleri kullanılarak daha dengeli sonuçlar elde edilmesi

sağlanabilir. Ayrıca, Random Forest modelinde overfitting'i önlemek için daha düşük ağaç sayısı veya derinlik sınırı gibi düzenleme teknikleri kullanılabilir.

- Performansın iyileştirilmesi için daha geniş veri setleriyle çalışılması önerilebilir. Mevcut veri seti üzerinde yapılan sınıflandırmalar, özellikle küçük veri setlerinde sınıf dengesizliği problemlerine yol açabilir. Bu nedenle, veri setinin genişletilmesi, farklı test ortamlarıyla zenginleştirilmesi ve bu verilerle daha iyi genel sonuçlar elde edilmesi sağlanabilir.
- KNN ve Lojistik Regresyon gibi modellerde düşük performans gözlemlenmiş, bu nedenle sınıf dengesizliği olan veri setlerinde SMOTE veya ADASYN gibi dengeleme yöntemleri kullanılabilir. Bu yöntemler, sınıflar arasında denge sağlayarak yanlış pozitif ve yanlış negatif oranlarını düşürebilir.
- Öznitelik çıkarma işlemi, bazı durumlarda model performansını artırırken bazılarında düşürebilir. Bu nedenle, çıkarılan her özneliğin model üzerindeki etkisinin daha detaylı analiz edilmesi ve önemli olan özneliklerin korunmasına özen gösterilmesi gerekmektedir. Özellikle doğru ve yanlış cevap sayıları gibi öznelikler, modellerin performansını doğrudan etkilemiştir.
- KNN modeli, mevcut veri setinde düşük performans sergilemiş olup, bu yöntem yerine destek vektör makineleri (SVM) gibi daha başarılı olabilecek algoritmalar önerilebilir. Yüksek boyutlu ve karmaşık veri setlerinde SVM, sınıflar arasında daha net ayrımlar yapabilir ve daha iyi sonuçlar verebilir.
- Elde edilen sonuçlara göre, özellikle doğru ve yanlış cevap oranları gibi önemli özneliklerin çıkarılmasına karşı modellerin performansı hassas olmuştur. Bu nedenle, modellerin genel performansını artırmak için bu öznelikler üzerinde çalışılarak, modelin doğruluğunu artırmaya yönelik stratejiler geliştirilebilir.

6. KAYNAKLAR

- Adanır, G. A. E-Sınav Sistemi İle İlgili Öğrenci Görüşleri Ve Yaşanan Sorunların Araştırılması. *Organizasyon ve Yönetim Bilimleri Dergisi*, 13(2), 189-207.
- Ateş, H., & Banazılı, A. M. (2021). Yükseköğretim Kurumlarında E-Öğretim Uygulaması: Türkiye İçin Strateji Önerileri. *Assam Uluslararası Hakemli Dergi*, 8(19), 112-122.
- Atoum, Y., Chen, L., Liu, A. X., Hsu, S. D., & Liu, X. (2017). Automated online exam proctoring. *IEEE Transactions on Multimedia*, 19(7), 1609-1624.
- Bayram, M., Peker, A. T., Aka, S. T., & Vural, M. (2019). Üniversite öğrencilerinin uzaktan eğitim dersine karşı tutumlarının incelenmesi. *Gaziantep Üniversitesi Spor Bilimleri Dergisi*, 4(3), 330-345.
- Bilgiç, A. G. D. H. G., & Tüzün, H. (2015). Yükseköğretim kurumları web tabanlı uzaktan eğitim programlarında yaşanan sorunlar. *Açıköğretim Uygulamaları ve Araştırmaları Dergisi*, 1(3), 26-50.
- Bobde, S., Chaudhari, S., Golguri, J., & Shahane, R. (2017). Web based online examination system. *Global Res Develop J Eng*, 2(5), 58-61.
- Buluk, B., & Eşitti, B. (2020). Koronavirüs (COVID-19) sürecinde uzaktan eğitimin turizm lisans öğrencileri tarafından değerlendirilmesi. *Journal of Awareness*, 5(3), 285-298.
- Çalışkan, S. K., & Soğukpınar, İ. (2008). Kxknn: K-means ve k en yakın komşu yöntemleri ile ağlarda nüfuz tespiti. *EMO Yayınları*, 120-24.
- Çoban, S. (2013). Uzaktan ve teknoloji destekli eğitimin gelişimi. *İstanbul: XVI. Türkiye’de İnternet Konferansı Bildiri Kitabı*.
- Çukurova Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 14(2), 325-338.
- Dadak, A., Uyan, U., & Öztürk, M. U. (2022, May). Real time cheating detection pipeline for online exams. In *2022 30th Signal Processing and Communications Applications Conference (SIU)* (pp. 1-4). IEEE.
- Demir, E. (2015). Açık ve uzaktan öğrenmenin katılım ve devamlılık stratejilerinde akademik takvim. *Açıköğretim Uygulamaları ve Araştırmaları Dergisi*, 1(4), 36-58.
- Derici, S. (2021). Uzaktan Eğitim Sistemlerinin Tercih Edilmesini Etkileyen Faktörlerin Belirlenmesi. *Journal of Continuous Vocational Education and Training*, 4(1), 19-23.

- Durak, G., Çankaya, S., & İzmirli, S. (2020). COVID-19 pandemi döneminde Türkiye'deki üniversitelerin uzaktan eğitim sistemlerinin incelenmesi. *Necatibey Eğitim Fakültesi elektronik fen ve matematik eğitimi dergisi*, 14(1), 787-809.
- Elcil, Ş., & Şahiner, D. S. (2014). Uzaktan eğitimde iletişimsel engeller. *Sosyal ve Beşeri Bilimler Dergisi*, 6(1), 21-33.
- Güngör, A. (2021). Yükseköğretim Öğrencilerinin 'Uzaktan Eğitim' Kavramına İlişkin
- Hacıbeyoglu, M., Çelik, M., & Çiçek, Ö. E. K En Yakın Komşu Algoritması ile Binalarda Enerji Verimliliği Tahmini. *Necmettin Erbakan Üniversitesi Fen ve Mühendislik Bilimleri Dergisi*, 5(2), 65-74.
- İbicioğlu, H., & Antalyalı, U. Ö. L. (2005). Uzaktan eğitimin başarısında imkân algı motivasyon ve etkileşim faktörlerinin etkileri: Karşılaştırmalı bir uygulama.
- İnce, Z., Yentür, M., & Şahin, V. (2021). Uzaktan eğitim ile coğrafya öğretiminin etkililiği ölçeği: Ölçek geliştirme, geçerlik ve güvenirlik çalışması. *OPUS International Journal of Society Researches*, 17(Pandemi Özel Sayısı), 3943-3966.
- Kamalov, F., Sulieman, H., & Santandreu Calonge, D. (2021). Machine learning based approach to exam cheating detection. *Plos one*, 16(8), e0254340.
- Koçdar, S. (2011). *Uzman görüşlerine göre Türkiye'de uzaktan eğitim programlarının akreditasyonu* (Doctoral dissertation, Anadolu University (Turkey)).
- Köktürk, F. (2012). K-en yakın komşuluk, yapay sinir ağları ve karar ağaçları yöntemlerinin sınıflandırma başarılarının karşılaştırılması (Master's thesis, Bülent Ecevit Üniversitesi, Sağlık Bilimleri Enstitüsü, Biyoistatistik Anabilim Dalı).
- Metaforik Algıları: Kilis 7 Aralık Üniversitesi İslami İlimler Fakültesi Örneği. *Eskiye*, (45), 693-717.
- Motwani, S., Nagpal, C., Motwani, M., Nagdev, N., & Yeole, A. (2021, May). AI-based proctoring system for online tests. In *Proceedings of the 4th International Conference on Advances in Science & Technology (ICAST2021)*.
- Ngqondi, T., Maoneke, P. B., & Mauwa, H. (2021). A secure online exams conceptual framework for South African universities. *Social Sciences & Humanities Open*, 3(1), 100132.
- Ozgen, A. C., Öztürk, M. U., Bayraktar, U., & Aksoy, S. (2021). An Anti-Cheating System for Online Interviews and Exams. *ASRJETS-Journal*, 83(1), 96-112.
- Özbay, Ö. (2015). Dünyada ve Türkiye'de uzaktan eğitimin güncel durumu. *Uluslararası Eğitim Bilimleri Dergisi*, (5), 376-394.
- Raj, R. V., Narayanan, S. A., & Bijlani, K. (2015, July). Heuristic-based automatic online proctoring system. In *2015 IEEE 15th International Conference on Advanced Learning Technologies* (pp. 458-459). IEEE.

- Solak, H. İ., Ütebay, G., & Yalçın, B. (2020). Uzaktan eğitim öğrencilerinin basılı ve dijital ortamdaki sınav başarılarının karşılaştırılması. *Açıköğretim Uygulamaları ve Araştırmaları Dergisi*, 6(1), 41-52.
- Telli, S. G., & Altun, D. (2023). Türkiye'de deprem sonrası çevrimiçi öğrenmenin vazgeçilmezliği. *Üniversite Araştırmaları Dergisi*, 6(2), 125-136.
- Ülger, K. (2021). Uzaktan eğitim modelinde karşılaşılan sorunlar-fırsatlar ve çözüm önerileri. *Uluslararası Güncel Eğitim Araştırmaları Dergisi*, 7(1), 393-412.
- Valızadeh, M. (2022). Cheating In Online Learning Programs: Learners'perceptions And Solutions. *Turkish Online Journal of Distance Education*, 23(1), 195-209.