





Düzce Üniversitesi Bilim ve Teknoloji Dergisi

Araştırma Makalesi

Pandemi Sürecinde Uzaktan Eğitimde Senkron, Asenkron Ve Hibrit Yapılmış Derslerde Veri Madenciliği İle Öğrenci Performans Analizi

 Serdar KIRIŞOĞLU ^{a,*},  Mehmet YILDIRIM ^b

^a Bilgisayar Mühendisliği Bölümü, Mühendislik Fakültesi, Düzce Üniversitesi, Düzce, TÜRKİYE

^b Elektrik-Elektronik ve Bilgisayar Mühendisliği Bölümü, Fen Bilimleri Enstitüsü, Düzce Üniversitesi, Düzce, TÜRKİYE

* Sorumlu yazarın e-posta adresi: serdarkirisoglu@duzce.edu.tr

DOI: 10.29130/dubited.1067122

ÖZ

Günümüzde üniversite öğrencilerinin eğitime ve eğitim materyallerine internetten erişim oranları oldukça artmıştır. Eğitimde internetin kullanılması ve ders materyallerine erişimin artmasına bağlı olarak analiz edilebilecek veri setinde artış meydana gelmiştir. Bu veri setlerinden bir tanesi de (planlı veya acilen) uzaktan eğitim sürecine geçen üniversitelerin, uzaktan eğitim sistemlerinde biriken ödev, sınav, proje, performans, devam notları ve benzeridir. Yeni Korona Virüs (Covid-19) pandemisinde Yüksek Öğretim Kurumu'nun (YÖK) tavsiyesi ile üniversiteler eğitimlerine uzaktan Asenkron, Senkron ve Hibrit yöntemlerini kullanarak devam etmiş, hatta sınavları uzaktan eğitim sisteminde yapmak zorunda kalmışlardır. Bu çalışmada, Kayseri Üniversitesinin Uzaktan Eğitim Uygulama ve Araştırma Merkezi (KAYUZEM) sisteminden alınmış veriler kullanılmıştır. Araştırma kapsamında 8319 işlenmiş veri bulunmaktadır. Bu veriler üzerinde Veri Madenciliği (VM) alanında kullanılan RapidMiner programının otomatik modelleme özelliği kullanılarak varsayılan algoritmalarla geleceğe yönelik tahminleme işlemi yapılmıştır. Varsayılan algoritmalar arasından en iyi sonucu veren Derin Öğrenme, Naive Bayes, Gradient Boosted Trees, Lojistik Regresyon kullanılmış ve bu otomatik modelleme de yer almayan k-En Yakın Komşu (k-NN) algoritması da çalışmaya dahil edilmiştir. Bu 5 algoritmanın parametreleri üzerinde değişiklikler yapılarak daha iyi sonuçlar elde edilmeye çalışılmıştır. Öğrenci başarısına göre en iyi tahminleme sonucunu, Lojistik Regresyon ile kurulan model vermiştir. Derse katılma yöntemlerinin tümünün (Senkron, Asenkron ve Hibrit) öğrenci başarısına etkisi Karışıklık Matrisi yöntemiyle karşılaştırılmıştır ve en güvenilir yöntemin Hibrit olduğu görülmüştür. Bu çalışma ile üniversitelerde derse katılma yöntemlerinden hangisinin, öğrenciler açısından daha güvenilir olacağına yönelik çıkarımlarda bulunulmuştur. Dolayısıyla yapılan çıkarımlar ile birlikte bir dahaki akademik dönem için öğrenci başarı düzeyinin artması yönünde, derse katılma yöntemlerinden hangisinin daha güvenilir olduğu konusunda tahminleme mümkün olmuştur.

Anahtar Kelimeler: Veri madenciliği, Uzaktan eğitim, Pandemi, Covid-19, RapidMiner

Student Performance Analysis With Data Mining In Distance Education Synchronous, Asynchronous And Hybrid Courses In The Pandemic Process

ABSTRACT

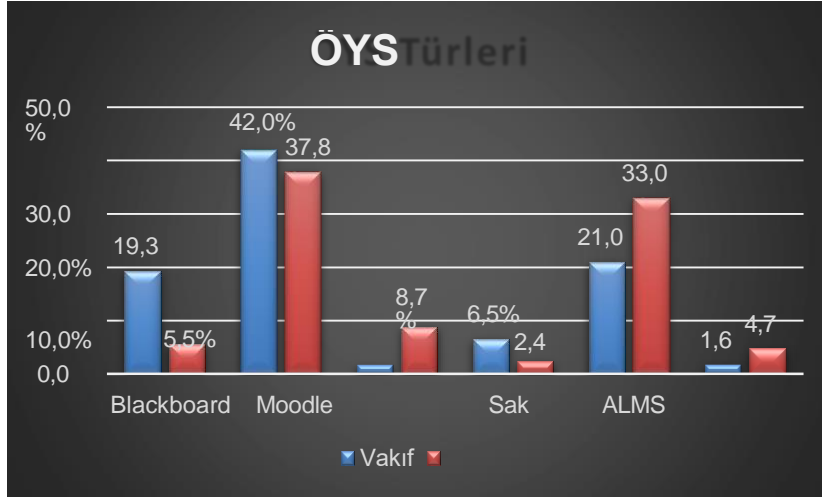
Today, the rate of university students' access to education and training materials from the internet has increased

considerably. There has been an increase in the data set that can be analyzed due to the use of the internet in education and the increase in access to course materials. One of these data sets (planned or urgently) is the homework, exam, project, performance, attendance grades and the like accumulated in the distance education systems of the universities that started the distance education process. In the new Corona Virus pandemic, with the recommendation of the Higher Education Council, universities continued their education by using Asynchronous, Synchronous and Hybrid methods and even had to take the exams in the distance education system. In this research, data obtained from the Distance Education Application and Research Center system of Kayseri University was used. There are 8319 processed data within the scope of the research. Using the automatic modeling feature of RapidMiner program, which is used in Data Mining field, on these data, predictions for the future were made with default algorithms. Deep Learning, Naive Bayes, Gradient Boosted Trees, Logistic Regression, which gave the best results among the default algorithms, were used and the k-NN algorithm, which is not included in this automatic modeling, was also included in the study. It has been tried to obtain better results by making changes on the parameters of these 5 algorithms. The model established with Logistic Regression gave the best estimation result according to student success. The effect of all the participation methods (Synchronous, Asynchronous and Hybrid) on student achievement was compared with the Confusion Matrix method and it was seen that the most reliable method was Hybrid. With this research, inferences were made about which of the methods of attending lectures in universities would be more reliable for students. Therefore, with the inferences made, it has been possible to predict which of the methods of participation in the course is more reliable in order to increase the level of student achievement for the next academic term.

Keywords: Data mining, Distance education, Pandemic, Covid-19, RapidMiner

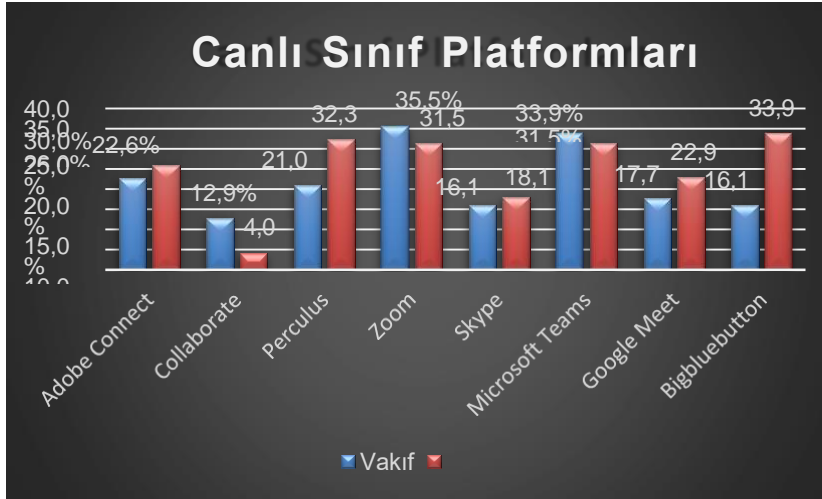
I. GİRİŞ

2019 yılının sonlarında Çin'in Vuhan Eyaleti'nde Covid-19 hastalığı tespit edilmiş ve dünyadaki diğer ülkelere de bu virüs yayılmıştır [1]. Ülkemizde ilk Covid-19 virüsü vakası 11 Mart 2020 tarihinde görüldükten sonra Dünya Sağlık Örgütü pandemi ilan etmiş ve bu tarihten itibaren vaka sayıları belirli bir oranda artış göstermiştir [2]. Vakaların artışıyla birlikte YÖK 16 Mart 2020 tarihinde eğitime 3 hafta ara verileceğini duyurmuştur [3]. Eğitim ve öğretim süreçlerinin kesintiye uğramaması için 18.03.2020 tarihinde YÖK tarafından yapılan açıklamada mevzuatta uzaktan öğretim ile ilgili değişikliğe gidilmiş ve "Uygulama konusuna gelince; üniversitelerimiz senkron (eş zamanlı) ya da asenkron (eş zamanlı olmaksızın) uzaktan öğretim yöntemlerini kullanabilir. İçinde bulunulan ortamda senkron uygulamaların güçlüğü göz önüne alındığında üniversitelerimizin özellikle önlisans ve lisans programlarında asenkron uygulamalara ağırlıklı yöneleceği tahmin edilmekte olup daha çok etkileşim içeren senkron uygulamalar da YÖK tarafından teşvik edilecektir." kararı alınmıştır [4]. Pandemi koşullarının uzayacağı öngörüldükten sonra YÖK tarafından üniversitelerde 2019 – 2020 bahar yarıyılında eğitim ve öğretimin tamamen uzaktan yapılmasına karar verilmiştir [5]. Daha sonra yapılan çalışmalar pandeminin seyrine göre değişiklik göstermiştir. Örneğin; 04.06.2020 tarihli karar ile birlikte derslerin uzaktan eğitimde verilme oranı %30'dan %40'a çıkmıştır. Bunun yanında üniversitelerdeki mevcut Uzaktan Eğitim Araştırma ve Uygulama Merkezi (UZEM) sayısı 120'den 140'a çıkmıştır [6]. Bu merkezlerin kullandığı platformlar çeşitlilik göstermiştir. YÖK'ün üniversitelerimizde küresel salgın sürecinde uzaktan eğitim uygulamaları anket sonuçlarına göre, Şekil 1'de gösterildiği üzere çoğunlukla kullanılan Öğrenme Yönetim Sistemi (ÖYS) türleri Blackbord, Moodle, Canvas, Sakai, Advancity Learning Management System (ALMS) ve Edmodo'dur [7]. Üniversiteler de en çok kullanılan ÖYS türü Şekil 1'de gözüktüğü üzere Moodle olmuştur. Bu çalışmada kullanılan veri seti (Asenkron dersler için) ALMS ÖYS Türü üzerinden temin edilmiştir.



Şekil 1. Pandemi Döneminde Üniversitelerde Kullanılan ÖYS Türleri.

Aynı anket sonuçlarına göre canlı derslerde (Senkron) çoğunlukla kullanılan Şekil 2’de gösterilen canlı sınıf platformları ise Adobe Connect, Collaborate, Perculus, Zoom, Skype, Microsoft Teams, Google Meet ve Bigbluebutton’dır [7]. Üniversiteler de en çok kullanılan canlı ders platformu Şekil 2’de görüldüğü üzere Zoom’dur. Bu çalışmada kullanılan veri seti (Senkron dersler için) Perculus canlı sınıf platformu üzerinden temin edilmiştir.



Şekil 2. Pandemi Döneminde Üniversitelerde Kullanılan Canlı Sınıf Platformları.

YÖK 30.07.2020 tarihinde “Küresel Salgında Yeni Normalleşme Süreci” rehberi yayınlamış ve salgının bölgesel seyrine göre yapılacak olan uygulamalara yönelik hususlarda yetkiyi üniversitelerin ilgili kurullarına bırakmıştır. Salgının devam ettiği yeni normal süreçte harmanlanmış (hibrit) öğrenme yöntemini tavsiye etmiştir. Dolayısıyla uzaktan öğretimde kullanılan senkron, asenkron (açık ders malzemelerinin önden verilmesi) ve yüz yüze eğitimin bir arada yapılabileceğini yani hibrit yöntemini tavsiye etmiştir [8]. 2021 – 2022 eğitim ve öğretim yılından itibaren üniversitelerde uzaktan eğitimin pandemi koşullarına göre hibrit olarak devam edebileceği ve bunun yine üniversitelerinin kendi kurul kararlarına bırakıldığı YÖK tarafından Ağustos 2021 de yayınlanan Küresel Salgında Eğitim ve Öğretim Süreçlerine Yönelik Uygulamalar Rehberi’nde belirtilmiştir [9].

Pandemi sürecinde sadece uzaktan eğitimin verilme kararı alındığı 2019 – 2020 bahar yarıyılında KAYUZEM biriminin kullandığı ÖYS türlerinden ALMS ve canlı sınıf platformu Perculus de VM çalışabilecek kadar verinin biriktiği tespit edilmiştir. Bu verilere bakılarak Senkron, Asenkron ve

Hibrit uzaktan öğrenme yöntemlerinin öğrencilerin başarılarındaki etkisi bu araştırmanın konusunu oluşturmaktadır. Bu yönüyle de VM alanında yeni bir yol açacağı öngörülmektedir.

Veri tabanı bilgi keşfi kısaca VTBK süreci olarak tanımlanan, bir veri tabanından faydalı bilginin seçimini, ön değerlendirilmesini, dönüştürülmesini, veri kazıma işlemini ve sentez-yorum adımlarını içeren süreçtir [10]. Bir üniversitede uzaktan eğitimde verilen derslerin önemli çıktıları olan öğrenci notlarını içeren büyük veriler içerisinde veri sürecinin yürütülebilmesi için VM metotları ve algoritmaları kullanılmaktadır. Bu çalışmanın metod ve analiz bölümünde benzer bir süreç işlemiştir.

VM, büyük veri yığınları içerisinde modeller ve araçlar kullanarak faydalı bilgiye ulaşma sürecidir. Bu faydalı bilgilere belirli kurallar ve modeller uygulayarak sınıflandırma yapılabilir. Aynı zamanda bu bilgilerin analizlerini yorumlayarak ve değerlendirerek herhangi bir kullanıcıya ve/veya kuruluşa yarar sağlayabiliriz.

Son yıllarda VM konusunda yapılan çalışmalardan bir tanesi de Türkiye’deki bir üniversite de 3 öğretim üyesinin verdiği derslerle ilgili öğrenci memnuniyetlerini ölçen bir anketteki verileri kullanan bir çalışmadır. Can ve arkadaşları [11], çalışmada da öğretim üyelerinin başarı değişkenlerine göre oluşturulan iki ayrı tahmin modeli kullanılmış ve beklentileri karşılayan bulgular elde edilmiştir. Bu çalışma ile öğrencilerin derse katılımları ve derslerin zorluk dereceleri öğretim üyelerinin başarılarını aynı yönde etkilediği görülmüştür. Bunun yanında VM sayesinde farklı tutum davranışların da başarı algısına olan etkisi de incelenebilmiştir. Görüldüğü gibi bu tür çalışmalar üniversitenin hem öğrenci hem de öğretim üyelerinin eğitimdeki başarı oranlarının daha ayrıntılı değerlendirilmesini veya ölçülmesini sağlamaktadır. Buda bu tür teknolojilerin getirdiği en önemli faydalardan birisidir.

VM konusu Türkiye’deki çalışmalarda şimdiye kadar genelde sınıflamalar ve risk tahminleri üzerinden ilerlemiştir. Ancak 2019 yılında Gushchina ve Ochepovsky [12] tarafından yapılan bir çalışmada uzaktan eğitimdeki öğrenme çıktıları üzerindeki tahmini etkilerinin, VM ile ölçülebildiği ortaya konulmuştur. Burada söz konusu tahmini etkileri azaltmada eğitimi iyileştirmeye yardımcı olmuştur. Çalışmada öncelikle veriler üzerinden risk yönetimi planlanmış ve tüm dinamikler hesaplanmıştır. Çalışmanın sonucunda matematiksel yöntemlerin yetersiz kaldığı yerlerde VM yöntemlerini kullanarak daha iyi sonuç alınabileceği kanısına varılmıştır.

Üniversite öğrencilerinin belirli bir dersteki başarı oranlarının öğrenme yönetim sistemleri hareketliliği ölçülerek başarı düzeyleri arasındaki ilişkinin ortaya konulduğu Özbay ve Ersoy [13] tarafından ele alınan makale bu araştırma için çıkış noktası olmuştur. Özbay ve Ersoy’un bu çalışmalarında, bu tür ölçümlerde kullanılacak olan algoritmalarda doğru çıktı için hangi oranların kullanılacağı ve hangi gruplandırmaların yapılacağı hakkında bilgi verilmiştir. VM de kullanılacak algoritmaların kriter seçiminde bu çalışma yol gösterici olmuştur.

Ülkemizde çalışmamızla ilgili konular ele alınmış olsa da bu araştırmaya en yakın güncel çalışmalar genellikle yurt dışında yapılmaktadır. Örneğin, Abe [14], üniversitenin eğitimsel büyük verilerini VM ve makine öğrenmesinin uygulamaları ile ele almıştır. Ciddi bir sorun teşkil eden mezuniyet erteleme veya okul terki durumlarının daha önceden tahmin edilmesi ve azaltılması için faydalı yöntemler önermiştir. Benzer çalışmalardan yola çıkılarak mevcut araştırmamızda hangi alanların boşluklar bıraktığını ve nasıl doldurulabileceği bu araştırmanın sorularından bazılarıdır.

Ati Jain ve arkadaşları [15] çalışmalarında, pandemi süreci nedeniyle yüz yüze eğitimin sahip olduğu avantajlardan uzaklaşan çevrimiçi ve uzaktan eğitimin, video ve ses analizi yaparak daha başarılı bir şekilde kullanılabileceğini göstermişlerdir. Çevrimiçi yapılan öğrenme ve öğretme etkinliklerinde, öğrencilerin gerçek zamanlı olarak yüz ifadeleri ve seslerini kullanarak analiz yapmışlar, bu sisteme yenilikçi bir bakış açısı kazandırmışlardır. Bu yeni ve geliştirici yaklaşımla, kalem ve kâğıt yönteminin normal geri bildirim prosedürünü iyileştiren ve ona bir çözüm olarak hareket eden çeşitli öğretim, eğitim, koçluk ve danışma hizmetlerinde kalitenin iyileştirilmesine yardımcı olacağını aktarmışlardır. Konsantrasyon İndeksi kullanılarak tüm kullanıcılara öğrenme ve öğretme tekniklerini geliştirmelerinde etkili olacağını açıklamışlardır. Bu çalışmayla derin öğrenme algoritmasında

eğitimdeki kaliteyi belirleyen verilerin de işlenebileceği ve çeşitli verilerden elde edilen tahminlerden yararlanılabileceği görülmüştür.

Benzer şekilde, Villegas-Ch ve arkadaşları [16] eğitim kalitesinin artırılması amacıyla, veri algoritmalarının yönetilerek her öğrencinin ihtiyacına uygun bir eğitim sağlayacak şekilde değerlendirilmesinin yapıldığı, başka bir çalışmada da benzer şekilde VM kullanmıştır. Bu çalışmada da öğrenme yönetim sistemleri gibi platformlarda öğrencilerin eğilimleri, başarı tahminleri ve gelişim göstergelerinin sadece veriler üzerinden izlenebileceği ortaya konmuştur.

Byers ve arkadaşları [17] çalışmalarında geleneksel sınıflar ile modern sınıflarda verilen eğitime göre değişiklik gösteren ders başarı verilerini kullanmışlardır. Byers ve arkadaşlarının bu çalışması geleneksel veya uzaktan öğrenme ortamlarındaki değişkenlerin etkilerini hesaplayarak farklı sınıf düzeylerinin öğrenci notlarını etkilediğini ortaya koymuştur. Araştırmada farklı öğrenme alanlarından biri olan uzaktan eğitim sistemindeki verilerin gelecekteki eğitim yöntemlerini de etkileyeceği ön görülmektedir.

Blagojević ve Micić [18], çalışmalarında öğrenim çıktılarının analizinde en doğru sonucu elde etmek için birden fazla parametrenin doğru işlenmesi gerektiğini anlatmışlardır. Doğru analiz için üç farklı denetimli makine öğrenimi tekniği kullanılmış, her tekniğin hesaplama gereksinimleri farklı girdilerle test edilmiş, tahminlerin gerçekleştirilmesi için yapay sinir ağı modeli her durumda ve ihtiyaç duyulan sürede değerlendirilmiştir. Öğrenci performans tahminleri ve sonucunda da yüksek riskli öğrencilerin kursa devam etmeme ihtimallerinin tespiti bu yolla sağlanmıştır.

Bir önceki çalışmada olduğu gibi Hussain [19] çalışmasında “Sınırlı bir veri kümesi boyutu kullanarak bir öğrencinin notunu adil ve anlamlı bir doğruluk oranıyla tahmin etmek için en iyi makine öğrenimi sınıflandırma modeli nedir?” ve “Öğrencilerin notlarını tahmin etmek için bir sınıflandırma modelinin tasarımına yardımcı olabilecek en önemli özellikler nelerdir?” sorularını araştırmış. Bunun için Qassim Üniversitesinin verileri kullanmış ve 17 adet öznitelik verisi ön işleme işlemini yaparak, 13 adet öznitelik (ağırlıklı not ortalaması, kayıtlı dersler, lise notu, öğretim üyesi etkisi, e-öğrenme yöntemi etkisi, değerlendirme etkisi, donanım ve internet hızı, çevrimiçi öğrenme aracı kullanılabilirliği, e-öğrenme kursu uygunluğu, ikamet yeri, kurs adı, öğrenci yaşı ve öğrenci seviyesi) verisini dikkate almıştır. Weka programında BayesNet-D, NativeBayes, J48 algoritmalarını kullanmış. Bu algoritmalar arasında en iyi sonucu J48 algoritması %95,0617 oranını vermiş. Sonuç olarak araştırmacı bu çalışmada, öğrencilerin notlarını tahmin etmek için öznitelikleri belirlemiş ve sınırlı bir veri kümesi boyutu kullanarak bir öğrencinin notunu adil ve anlamlı bir doğruluk oranıyla tahmin etmek için en iyi algoritmayı belirlemiştir. Bu her iki çalışmada görüldüğü üzere tahminlemelerin doğru bir şekilde elde edilmesi için ilk başta ön işleme işlemlerinin detaylı bir şekilde yapılması ve sonrasında algoritmaların parametreleri en iyi sonucu verecek şekilde ayarlanması gerekmektedir. Bu araştırma içinde aynı yol izlenmiştir.

Yaacob ve arkadaşları [20] çalışmalarında öğrenci performansını tahmin etmek için sınıflandırma algoritmalarını kullanarak tahmine dayalı modeller geliştirmişler. Bunun için ilk başta CRISP-DM (Veri madenciliği için sektörler arası standart süreleri) modelini kullanmışlar. İlk adım olarak, öğrencilerin performansını tahmin etme iş problemini veri madenciliği problemine dönüştürmeyi içeren iş faaliyetlerini ve problemlerini anlamışlar. İkinci adımda, verilerin toplanmasını yapmışlar. Üçüncü adımda, toplanan verileri ön işleme işlemine tabi tutmuşlar. Son olarak dördüncü adımda da k- NN, Naive Bayes, Decision Tree ve Lojistik Regresyon algoritmaları ile tahminler yapmışlar ve bu tahminlerin ROC eğrisinde değerlendirmişler. Çalışmanın sonunda öğrenci performansını en iyi tahmin eden algoritma %85,4 ile Naive Bayes olmuştur. Bu yönüyle çalışma bu araştırmadaki veri işleme adımları açısından benzerlik göstermektedir.

Chen Tan ve arkadaşları [21] bu makalede, veri madenciliği kullanan çevrimiçi eğitim sistemlerinde öğretme ve e-öğrenmenin teknik yönlerini tespit etmek amacıyla yeni bir tahmin modeli sunmuşlardır. Öğretim görevlileri ve öğrenciler arasındaki öğretim ve eğitimin davranışsal yönlerinin önemli olması sebebiyle, çevrimiçi eğitim sistemlerinde deneyim kalitesinin tahmini kritik bir konu olduğuna

vurgulamışlardır. Çevrimiçi eğitim sistemlerinde deneyim kalitesinin verimli faktörlerini tespit etmek için birliktelik kuralları madenciliği ve denetimli teknikler uygulamışlardır. Önerilen tahmin modelinin, çevrimiçi eğitim sistemlerinde öğrenciler için öğretim ve e-öğrenmenin davranışsal yönlerini tahmin etmek için uygun doğruluk, kesinlik ve hatırlama faktörlerini karşıladığını açıklamışlardır. Çevrimiçi eğitim sistemlerinin performansını değerlendirmek için Deneyim Kalitesi tabanlı yeni bir tahmin modeli önermişlerdir.

Başka bir çalışmada, Lonia Masangu ve arkadaşları [22], öğrencilerin çevrimiçi öğrenme etkinlikleri ve demografik bilgilerinden oluşan verilerden öğrenci performansını tahmin etmeyi amaçlamışlardır. Öğrencilerin akademik performansını tahmin etmek için toplam 480 kayıttan oluşan veri setini, Perceptron Sınıflandırıcı, Destek Vektör Makinesi, Decision Tree, Lojistik Regresyon ve Random Forest olmak üzere beş makine öğrenme tekniği ile değerlendirmişlerdir. Destek vektör makinesi algoritmasının verileri %70.8 doğrulukla en iyi şekilde ele alarak, öğrencinin akademik performansını tahmin etmek için en uygun olduğunu kanıtlamışlardır. Elde ettikleri sonuçlarla, öğrencilerin devamsızlıklarının akademik performanslarını etkilediğini, diğer yandan öğrenci sınıf notlarının ise akademik performansı etkilemediğini ortaya koymuşlardır.

Delali Kwasi Dake ve arkadaşları [23], çalışmalarında Winneba Üniversitesinin Bilgi ve İletişim Teknolojisi Bölüm öğrencilerinin pandemi döneminde sınav puanlarını etkileyecek olan öznitelikler cinsiyeti, yaşı, oturduğu bölge, evinin internet bağlantısı, evindeki öğrenme ortamı, kişisel bilgisayar, evde akademik çalışma imkânı, kardeşlerinin rahatsız ediciliği, Covid-19 ve ruh hali, evde genel öğretim görevlisi desteği ve genel aile desteği olarak belirlemişler. Google form kullanarak 637 adet öğrenci verisi örneği toplamışlar. Toplamış oldukları verileri ön işleme işlemi yaptıktan sonra çalışmalarında 536 adet veriyi kullanılabilir duruma getirmişler. Bu verilerle Weka programında Random Forest, Random Tree, Naive Bayes ve J48 Decision Tree algoritmalarını kullanarak tahminleme işlemi yapmışlar. Bu tahminleme işleminde her algoritmanın karışıklık matrisini de hesaplamışlardır. Aynı zamanda algoritmalarda en az performansı gösteren öznitelikleri kaldırılarak tekrardan hesaplamaları yapılmış ve algoritmaların performanslarını bu şekilde arttırmaya çalışmışlardır. Algoritmaların performansını daha fazla değerlendirmek için ROC eğrisini kullanmışlar. Sonuç olarak çalışmalarında en iyi sonucu Naive Bayes ve Random Forest algoritmaları vermiştir. Bu çalışma bizim çalışmamız gibi öğrencinin performansını arttırmaya yöneliktir. Bizim çalışmamızdan farklı öğrencilerin öznitelik verilerini kullanarak öğrenci performansını incelemiş olmasıdır. Biz çalışmamızda öğrencilerin performansını derse katılma yöntemlerine göre incelemekteyiz.

Alsammak ve arkadaşları [24] üç tür anket kullanarak 1120 örnek içeren veri toplamış, verileri ön işleme işlemine tabi tutmuşlardır. Veri seti 35 öznitelikli ve 1000 örnek olarak düzenlemiştir. Düzenlenen bu veri setinde e-öğrenmede öğrenci performansı üzerinden veri madenciliği algoritmaları WEKA programı yardımıyla uygulamıştır. Uyguladıkları algoritmalar Decision Tree, Random Tree, Naive Bayes, Random Forest, REP Tree, Bagging ve k-NN'dir. 35 öznitelikle yapılan tahminleme sonucunda en iyi sonucu %96,8'le k-NN algoritması vermiştir. Araştırmacılar öznitelikleri Ranker Search Metodunu kullanarak en iyi öznitelikleri seçmiş ve bu özniteliklere göre tahminleme işleminde en iyi sonucu %77 ile yine k-NN algoritması vermiştir. Aynı zamanda algoritmaların parametrelerinde yapılan değişikliklerle birlikte yine en iyi sonucu %100'le k-NN algoritması vermiştir. Öğrencilerin e- öğrenmede olumlu memnuniyetleri %61, istenmeyen memnuniyetleri %39 olarak tespit edilmiştir. Öznitelikler arasında en önemli nitelikler; öğrencinin annesinin vasfı, öğrencinin babasının vasfı, ebeveynin birlikte yaşamaları, aile çalışması desteği, geçmiş sınıf başarısızlıkları, haftalık çalışma süresi, ekstra ücretli dersler, müfredat dışı etkinlikler, okul sonrası serbest zamanı, sağlık durumu, Daha yüksek çalışma istenmesi, zamanında mezun olma olasılığı, giriş sınav puanı, yıllık değerlendirme ve final notu olarak belirlenmiştir. Bu çalışma ile öğrencilerin e-öğrenmeden ve özel olarak e-öğrenme sisteminden daha az memnun olduğu, ideal eğitim tekniğinin, durum normale döner dönmez başlanması gereken kombinasyon eğitimi olduğu tespit edilmiştir.

Pandemi döneminde uzaktan eğitimin avantajları ve dezavantajları konusu üzerinde yapılan çalışmalara bakıldığında, genel olarak; öğrencilerin sürekli bilgiye erişim sağlayabilmeleri, hastalıktan

korunmaları, mekandan bağımsız dersleri takip edebilmeleri ve asenkron ders materyallerine istedikleri zaman ulaşabilmeleri uzaktan eğitim sisteminin avantajları arasında yer alsa da, öğrencilerin uzaktan eğitim sistemlerinde teknik sorunlar yaşaması, internet altyapısından ve teknolojiye sahip olma durumlarından dolayı fırsat eşitsizliğinin meydana gelmesi ve öğrencilerde duyuşsal eksikliklerin yaşanması uzaktan eğitim sisteminin dezavantajları arasında yer almaktadır [25, 26]. Her ne kadar uzaktan eğitimde genellikle dezavantaj olarak teknoloji konulu sorunlardan bahsedilse de geleceğe yönelik tahminlerde bulunmak, yukarıda bahsedilen çalışmalarda olduğu gibi ancak bu teknolojilerle mümkündür.

Bu çalışmalardan da anlaşılacağı üzere öğrencileri ilgilendiren çeşitli verilerin kullanıldığı veri kümeleriyle tahminlerde bulunulması ve bunun üzerinden eğitim kalitesinin artırılmasına yönelik tavsiyeler ve çözüm önerileri üretilmesi mümkündür.

Özellikle Pandemi zamanlarında eğitimde fırsat eşitliği yaratmak adına her şeyin uzaktan ve çevrimiçi yapılmaya evrildiği dijital dünyada öğrenci başarılarının, öğrenciyi fiziksel olarak gözlemek üzerinden yapılmayacağı ya da bunun yeterli olmayacağı fütüristlerin gelecek tahminleri arasında yer almaktadır. Bu araştırmanın eğitimin gelecekteki tahaülünü somut olarak göstermesi açısından literatüre önemli katkılar sağlayacağı düşünülmektedir. Bu çalışmayı literatürdeki benzer çalışmalardan ayıran en önemli etken birincisi veri seti, ikincisi ise veri setindeki pandemi dönemi ile önem kazanmış olan uzaktan derse katılma yöntemlerinin (Senkron, Asenkron, Hibrit) öğrenci başarısı üzerindeki etkilerinin geleceğe yönelik VM modellerinin tahmin başarısıdır.

II. METOT VE ANALİZ

Bu çalışmada, bir üniversitesinin UZEM'in öğrenci, fakülte, ders ve başarı kriterleri üzerinde VM yöntemleri uygulanarak, gelecek dönemde öğrencilerin en başarılı olabileceği uzaktan eğitim modelleri hakkında tahminlerde bulunulmuştur.

Araştırmada KAYUZEM'in 06.04.2020 ve 15.08.2020 tarihleri arasındaki verileri kullanılmıştır.

VTYS den alınan verilerin araştırmaya katkısı göz önünde bulundurularak çeşitli elemeler ve düzenleme işlemleri yapılmıştır. Bu düzenleme işleminde ilk olarak verilerin Fakülte/MYO Adı, Bölüm Adı, Ders Adı, Ders İşleme Yöntemi, Öğrenci Başarı Durumu Başlıkları altında toplanmıştır. Bir öğrencinin derse canlı katılım dakikası fazla ise derse katılma yöntemi "Senkron", öğrenci dersi kayıttan izlediyse "Asenkron" ve öğrenci dersi hem canlı hem kayıttan izlediyse "Hibrit" olarak tanımlanmıştır. "Öğrenci Başarı Durumu" sütunu için öğrencilerin bahar dönemi sonunda aldığı notların verilere dönüştürülmesinde Kayseri Üniversitesinin ön lisans ve lisans eğitimi ve öğretim yönetmeliğindeki not baremi baz alınmıştır. Çalışmada kullanılan not baremi Tablo 1'de görülmektedir.

Tablo 1. Not baremi tanımlama çizelgesi

Başarılı	Başarısız
100-70	69-0

Düzenlemesi yapılan veriler RapidMiner VM programına aktarılmış ve yukarıda adı geçen üç yöntemin öğrencinin başarı durumu üzerindeki etkisinin analizi yapılmıştır. Veri setinde toplamda 8319 işlenmiş veri bulunmaktadır.

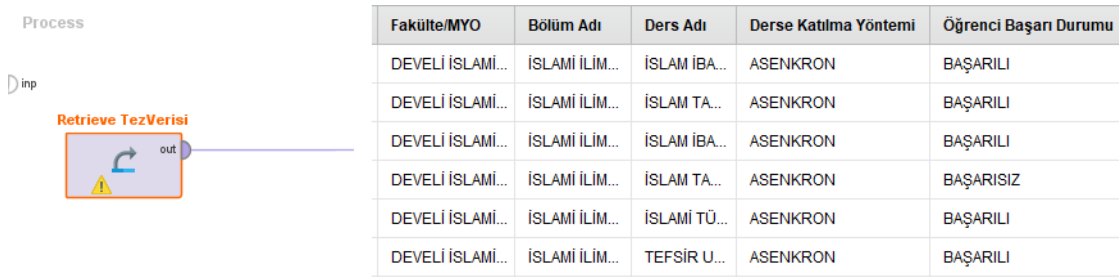
Verilerin analizinde Rapidminer yazılımı kullanılmıştır. Açık kaynak kodlu bir yazılımdır. Akademik çalışmalarda ücretsiz olarak bütün özellikleriyle birlikte kullanılabilir.

VTBK süreçleri doğrultusunda problem tanımlama, veri hazırlama, veri toplama, birleştirme ve temizleme, seçme işlemlerinden geçirilip hazır hale getirilen veriler, Tablo 2’de gösterilmiştir. Fakat kişisel veri alanı olan “Kullanıcı Adı” ve “Adı Soyadı” özelliklerinin entropilerinin çok düşük olmasından dolayı modellerin sonuçlarını etkilemeyeceği için çıkarılmıştır.

Tablo 2. RapidMiner programına aktarılmadan önceki veri seti.

Kullanıcı Adı	Adı ve Soyadı	Fakülte ve MYO	Bölüm	Ders Adı	Derse Katılma Yöntemi	Öğrenci Başarı Durumu
10*****64	A***** B*****	Develi İslami İlimler Fakültesi	İslami İlimler	İslam Tarihi I	Asenkron	Başarılı
10*****64	A***** B*****	Develi İslami İlimler Fakültesi	İslami İlimler	Yabancı Dil I Yaz Okulu	Senkron	Başarılı
10*****67	A***** K*****	Develi İslami İlimler Fakültesi	İslami İlimler	İslam Tarihi I	Hibrit	Başarılı
10*****67	A***** K*****	Develi İslami İlimler Fakültesi	İslami İlimler	Arapça Okuma Anlama I	Senkron	Başarılı
10*****67	A***** K*****	Develi İslami İlimler Fakültesi	İslami İlimler	Tefsir Usulü	Senkron	Başarısız
10*****29	A***** S***** D*****	Develi İslami İlimler Fakültesi	İslami İlimler	Arapça Okuma Anlama I	Asenkron	Başarılı
10*****29	A***** S***** D*****	Develi İslami İlimler Fakültesi	İslami İlimler	Arapça Sarf-Nahiv I	Asenkron	Başarılı
10*****29	A***** S***** D*****	Develi İslami İlimler Fakültesi	İslami İlimler	Eğitim Psikolojisi	Asenkron	Başarılı

RapidMiner programına geçmiş veriden geleceğe yönelik tahmin yapmak üzere aktarılmıştır. Aktarılan veri seti Şekil 3’de gösterilmiştir.

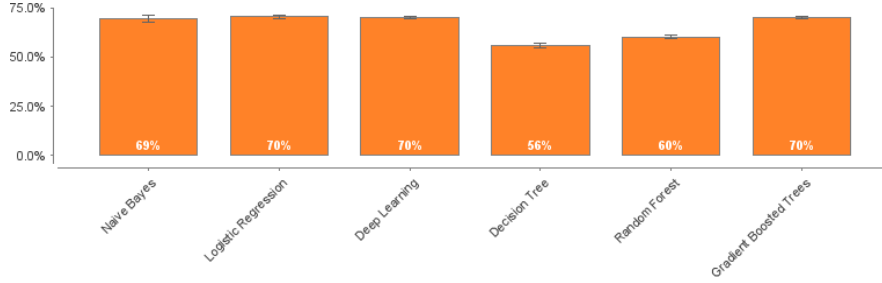


Fakülte/MYO	Bölüm Adı	Ders Adı	Derse Katılma Yöntemi	Öğrenci Başarı Durumu
DEVELİ İSLAMİ...	İSLAMİ İLİM...	İSLAM İBA...	ASENKRON	BAŞARILI
DEVELİ İSLAMİ...	İSLAMİ İLİM...	İSLAM TA...	ASENKRON	BAŞARILI
DEVELİ İSLAMİ...	İSLAMİ İLİM...	İSLAM İBA...	ASENKRON	BAŞARILI
DEVELİ İSLAMİ...	İSLAMİ İLİM...	İSLAM TA...	ASENKRON	BAŞARISIZ
DEVELİ İSLAMİ...	İSLAMİ İLİM...	İSLAMİ TÜ...	ASENKRON	BAŞARILI
DEVELİ İSLAMİ...	İSLAMİ İLİM...	TEFSİR U...	ASENKRON	BAŞARILI

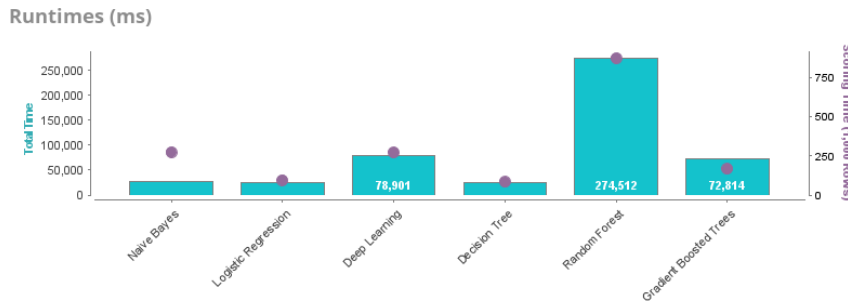
Şekil 3. RapidMiner programına aktarılan veri seti.

Aktarılan veri seti üzerinde, RapidMiner yazılımının bir özelliği olan ve en sık kullanılan algoritmalar ile otomatik olarak tahmin ve analizi yapılmıştır. Yapılan otomatik tahminleme analizinde doğruluk oranı Şekil 4’de gösterildiği gibi %69 olarak çıkan Naive Bayes, %70 olarak çıkan Logistic Regression, %70 olarak çıkan Deep Learning ve %70 olarak çıkan Gradient Boosted Trees algoritmaları daha detaylı olarak incelenmiştir. Decision tree ve Random Forest algoritmaları işlem

hızı ve doğruluk oranları çalışmamız için belirlediğimiz eşik değerinin altında kaldığı için çalışmaya dahil edilmemiştir. Otomatize edilmiş modellerin içerisinde olmayan ancak literatürde bu tip veriler üzerindeki başarısını ispat etmiş olan k-NN algoritması da karşılaştırılacak algoritmalar arasına dahil edilmiştir.



Şekil 4. Auto Model doğruluk oranı.

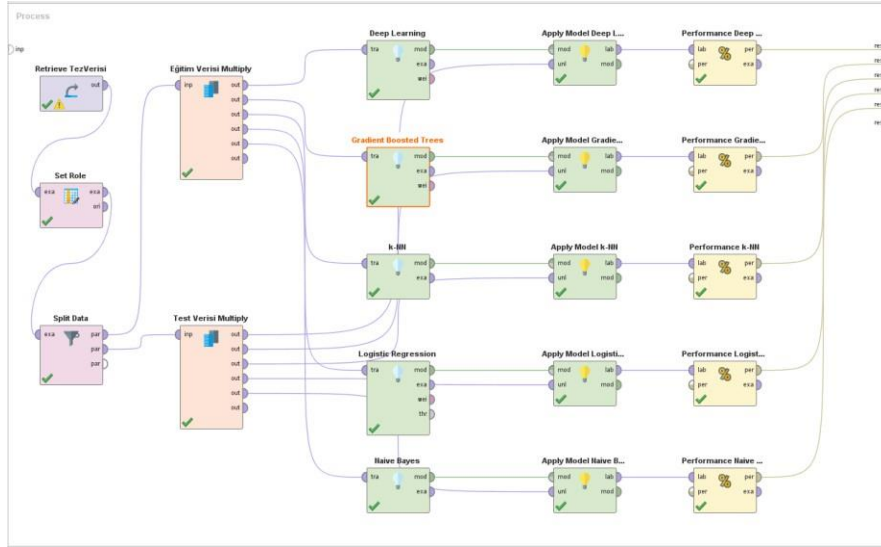


Şekil 5. Auto Model işlem süresi (ms).

Şekil 6’da oluşturulmuş modellerin ve operatörlerin yardımcı program üzerindeki grafiksel arayüzü gösterilmektedir. Veri seti solda, model sonuçları sağda olacak şekilde veriler soldan sağa doğru işlenerek yönlendirilmektedir. Burada kullanılan modüller ve görevleri kısaca aşağıdaki gibidir;

1. Retrieve: Modelde veri setini temsil eder. Çalışmanın ileri aşamasında veri setiyle ilgili detaylı bilgi verilecektir.
2. Set Role: Veri setinde tahmin yapılacak olan alan (Öğrenci Başarı Durumu) etiketlenmiştir.
3. Split Data: Verilerin ne kadarının eğitim ve test verisi olacağını belirler. Bu belirleme işlemi yapılırken Split Data operatörünün sampling type parametresinden “stratified sampling” seçilmiştir. Çalışmanın ileri aşamasında verilerin ne kadarının test ve ne kadarının eğitim verisi olduğuyula ilgili bilgi verilmiştir.
Stratified Sampling (Tabakalı Örnekleme): Rastgele alt kümeleri oluşturur ve alt kümelerdeki sınıf dağılımlarını tüm örnek kümesi ile aynı olmasını sağlar.
4. Multiply: Birden fazla algoritmaya bağlantı yapılabilmesi için kullanılmıştır.
5. Deep Learning (Derin Öğrenme), Gradient Boosted Trees, k-NN, Logistic Regression (Lojistik Regresyon), Naive Bayes: Algoritmalar ile ilgili detaylı bilgi verilecektir.
6. Apply Model: Eğitilmiş veriye göre test verisinin tahmin işlemini yapar.
7. Performance: Tahmin edilmiş olan verinin anlamlı bir duruma gelmesine yardımcı olur.

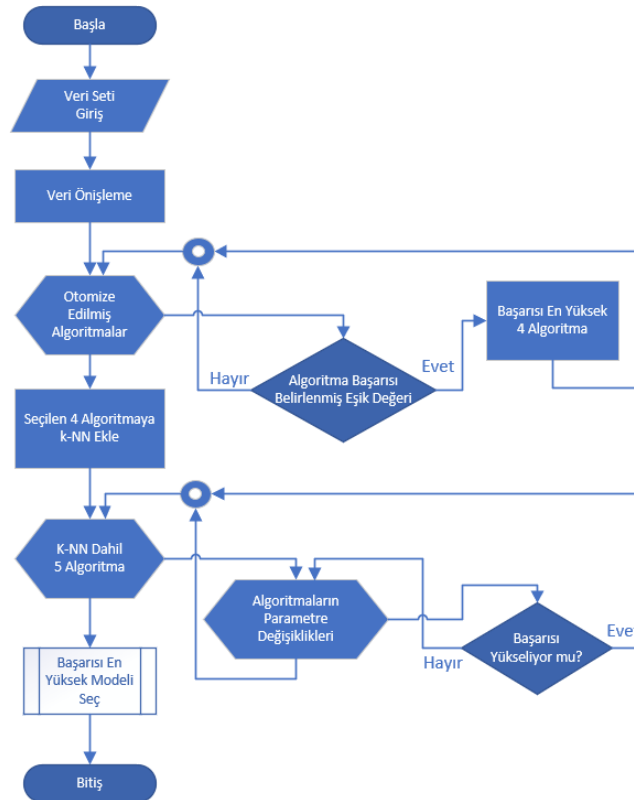
Rapidminer programının içerisine aktarılan veri seti otomatik model ile işleme tabi tutulduktan sonra içlerinden en yüksek başarı yüzdesine sahip 4 algoritma ve bu algoritmalar arasında olmayan k-NN algoritması ile aşağıdaki şekilde verilen model oluşturulmuştur.



Şekil 6. Çalışmada kullanılan operatörler.

Bu modeldeki kullanılan algoritmaların başarı oranlarının artırılması için parametrelerinde yapılan değişiklikler ve sonuçlar sınıflandırma yöntemleri bölümünde detaylı olarak verilecektir.

Çalışmanın sınıflandırma süreci Şekil 7’de gösterilmiştir.



Şekil 7. Sınıflandırma süreci.

Şekil 7’deki süreçte verileri önleme sürecinde veri hazırlama, birleştirme, temizleme vb. işlemleri yapılır. Sonrasında RapidMiner programının özelliği olan otomatik modelleme işleminin sonunda en iyi çıkan algoritmalar belirlenir. Algoritmaların içerisine iyi bir sonuç vereceğini düşündüğümüz k-NN algoritması da eklenir. Bu algoritmalarından maksimum doğruluk oranını almak için parametrelerinde

değişiklikler yapılır. Algoritmaların en yüksek doğruluk oranlarını belirledikten sonra aralarında en yüksek model seçilir.

A. ÇALIŞMADA KULLANILAN SINIFLANDIRMA YÖNTEMLERİ

RapidMiner programında otomatize edilmiş modellerden en iyi sonuç veren 4 algoritma ve içlerinde bulunmayan k-NN (literatürde buna benzer veri setlerinde daha önce kullanılmış ve iyi sonuçlar verdiği bildirildiği için) algoritmalarının veri seti üzerindeki başarısı ile beraber bu bölümde detaylı olarak incelenmiştir. Algoritmalarda detaylı işleme yapılırken varsayılan olarak verilerin %80'ni eğitim verisi geri kalan %20'si ise test verisi olarak kullanılmıştır.

RapidMiner programı algoritmaların özelliklerine göre verilerin tiplerinde kodlama (encoding) işlemini otomatik olarak gerçekleştirmektedir. Yani kategorik değişkenler sayısal değişkenlere RapidMiner programı ile otomatik olarak dönüşmektedir. Oluşturulacak model sayısal verilerle çalışması gerektiği takdirde normalizasyon (değerlerin 0 ile 1 arasına çekmek) yapılması gerekmektedir. Örneğin Tablo 3'deki cinsiyet sütunundaki erkek ve kadın verileri veri ön işleme aşamasında iki kategoriye sahip olduğu için veriler 1 ve 0'a dönüştürülebilir. Ancak veri setinde kullanılacak olan özellikler ikiden fazla kategoriye sahipse kodlama işlemi yapılır. Örneğin Tablo 3'deki kişinin yaşadığı şehir bu sefer her kategori farklı bir kolon oluşturulacak şekilde kodlanır. Naive Bayes algoritması olasılıklar üzerine hesap yaptığı için kategorik verilerle çalışır. Bu nedenle bu çalışmada kategorik veriden dönüşüm Naive Bayes algoritmasını kullanan diğer modellerde RapidMiner programı tarafından otomatik olarak yapılmıştır.

Tablo 3. Kategorik değişkenlerin sayısal değişkenlere dönüştürülmesi süreci.

Cinsiyeti	Şehir	Cinsiyet1	Şehir1	Şehir2	Şehir3
Erkek	Ankara	1	1	0	0
Kadın	İstanbul	0	0	1	0
Kadın	İzmir	0	0	0	1

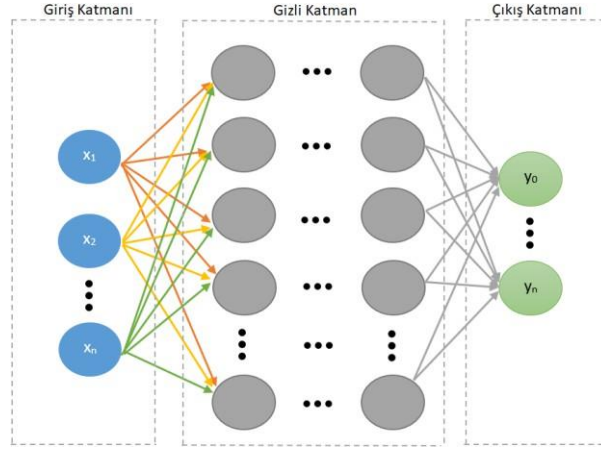
Bu çalışmada giriş bilgileri olarak Fakülte/MYO, Bölüm Adı, Ders Adı, Derslere Katılma Yöntemleri kullanılmaktadır. Bu kategorik bilgiler sayısal bilgilere dönüştürüldüğünde;

- Fakülte/MYO bilgileri için toplamda 4 kolon oluşur.
- Bölüm Adı bilgileri için toplamda 23 kolon oluşur.
- Ders Adı bilgileri için toplamda 143 kolon oluşur.
- Derse katılma Yöntemleri için toplamda 3 kolon oluşur.

Algoritma sayısal bir veri kullanıyorsa, bu çalışmada toplamda 173 giriş kolonu oluşur. RapidMiner programı bu kodlama işlemini otomatik olarak gerçekleştirmektedir.

A. 1. Derin Öğrenme ile Sınıflandırma

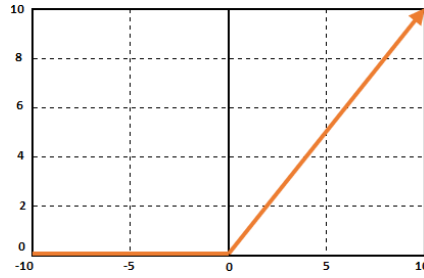
Derin Öğrenme bir makine öğrenme yöntemidir. Verilen bir veri kümesi ile çıktıları tahmin edecek yapay zekayı eğitmeye olanak sağlar. Yapay zekayı eğitmek için hem denetimli hem de denimsiz öğrenme kullanılabilir [27]. Derin öğrenme sınıflandırıcısı Şekil 8'de gözüktüğü gibi 3 tip (Giriş, Gizli ve Çıkış) katman içerir. Her katman en az bir adet birbirine bağlı düğüm içerir. Veri setinde sınıflandırıcı karmaşık yapıyı algılar ve önceki katmanları hesaplamak için kendi iç parametrelerini değiştirir [28].



Şekil 8. Derin Öğrenme Şeması.

Yapay sinir ağlarına doğrusal olmayan gerçek dünya özelliklerini tanıtmak için aktivasyon fonksiyonuna ihtiyaç duyulur. Temel olarak basit bir yapay sinir ağında x girdiler, w ağırlıklar olarak tanımlanır ve ağına çıkışına aktarılan değere $f(x)$ yani aktivasyon işlemi uygulanır. Daha sonra bu, nihai çıkış ya da bir başka katmanın girişi olacaktır [29]. Bu çalışmada aktivasyon fonksiyonu olarak Rectifier kullanılmıştır.

Rectifier: X 'in giriş değeri olduğu maksimum $(0, x)$ değerini seçer [30].



Şekil 9. Rectifier grafiği.

Hidden Layer Sizes: Gizli katman sayısını ve bu katmanlardaki nöronların sayısını ayarlamaya yardımcı olur. Tablo 4'de gizli katman ve nöron sayılarını değiştirerek en iyi sonuç 1 Gizli katmanda 30 nöron olarak bulunmuştur.

Tablo 4. Derin öğrenme gizli katman ve nöronların karşılaştırılması.

1 Gizli Katman		2 Gizli Katman		3 Gizli Katman
30 Nöron	50 Nöron	50 x 50 Nöron	100 x 100 Nöron	50 x 50 x 50 Nöron
71,33%	67,67%	69,05%	70,31%	67,61%

RapidMiner programının Derin Öğrenme operatörünün Epochs parametresinde veri kümesinin kaç kez tekrarlanacağını ayarlanması yapılır. Tablo 5'de değişik tekrarlar durumlarına bakılmıştır. En iyi tahmin değerinin 10 olduğu tespit edilmiştir.

Tablo 5. Derin öğrenme veri kümesi yinelenmesi karşılaştırılması.

Tekrar Sayısı	Tahmin Oranı
5	68,45%
7	67,25%
10	71,33%
20	69,41%
100	69,41%

Bu çalışmadaki derin öğrenme algoritmasında değişiklik yapılan parametreler Tablo 6’da gösterilmiştir. Sonucu etkilemediği görülen bazı parametreler varsayılan olarak bırakılmıştır. RapidMiner programı verileri otomatik olarak sayısal değere dönüştürdüğü için derin öğrenmedeki giriş katmanı 173, gizli katman sayısı 30 nöron ve çıkış katmanı 1 adettir.

Tablo 6. Derin öğrenme algoritması için kullanılan parametreler.

Özellik Adı	Değer
173 Adet Giriş	Fakülte/MYO (4), Bölüm Adı (23), Ders Adı (143), Derse Katılma Yöntemi (3)
1 Adet Çıkış	Öğrenci Başarı Durumu
Eğitim için Veri Setinin Tekrarlanma Sayısı	10
Gizli Katmanlarda Kullanılan Aktivasyon Fonksiyonu	Rectifier
Gizli Katman Sayısı	1
Gizli Katmanlarda Kullanılan Nöron Sayısı	30

Derin Öğrenme algoritmasının veri seti üzerindeki başarısı Tablo 7’de verilmiştir.

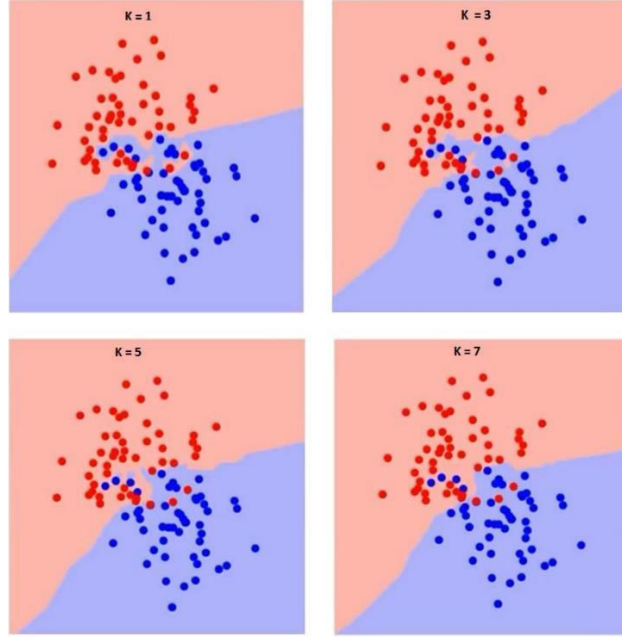
Tablo 7. Derin öğrenme algoritması karşılaştırması.

	Gerçek Başarılı	Gerçek Başarısız	Doğruluk
Tahmin Başarılı	602	147	80,37%
Tahmin Başarısız	330	585	63,93%
Hatırlama	64,59%	79,92%	

Tabloda “Gerçek Başarılı” sayısı toplamda 932 veriyken modelimiz bunun 602 verisini doğru tahmin etmiştir. 330 adet veriyi ise “Tahmin Başarısız” olarak yanlış tahmin etmiştir. Bu şekilde bütün çizelgedeki başarılı ve başarısız tahminler oranlandığı zaman Derin Öğrenme algoritmasındaki toplamdaki başarı oranı ortalaması %71,33 olarak hesaplanmıştır. Daha yüksek doğruluk oranına ulaşmak için diğer algoritmalar üzerinde çalışmalara devam edilmiştir.

A. 2. k-NN ile Sınıflandırma

k-NN algoritması, öznitelik alanındaki yakın eğitim örneklerine dayalı olarak nesnelere sınıflandırır. Yakınlık, öklid uzaklığı adı verilen bir uzaklık metriği cinsinden tanımlanır. Böylece, nesne komşusunun çoğunluk oyu ile sınıflandırılır ve nesne en yakın k komşusu arasında en yaygın olan sınıfa atanır [20]. k değerinin tek sayı verilmesine dikkat edilmesi gerekir. k değerine çift sayı verildiğinde komşulukların çift sayıda ve aynı uzaklıkta olmasından dolayı hangi sınıfta olacağı belirlenemez [31]. k-NN algoritmasında farklı k değerlerine göre iki sınıfı ayıran sınırlar Şekil 10’da gösterilmiştir [32].



Şekil 10. Farklı k değerine sahip iki sınıfı ayıran farklı sınırlar [32].

Tablo 8’de k parametresi değerlerinin karşılaştırması yer almaktadır. Burada en iyi sonucu veren k değeri 13’tür.

Tablo 8. k-NN algoritmasının k değerinin karşılaştırılması.

k Değeri	Tahmin Oranı
1	61,90%
5	67,49%
13	70,13%
19	68,93%
79	66,23%

K-NN algoritmasının veri seti üzerindeki başarısı Tablo 9’da verilmiştir.

Tablo 9. K-NN algoritması karşılaştırması.

	Gerçek Başarılı	Gerçek Başarısız	Doğruluk
Tahmin Başarılı	678	233	74,42%
Tahmin Başarısız	254	499	66,27%
Hatırlama	72,75%	68,17%	

Tabloda “Gerçek Başarılı” sayısı toplamda 932 veriyken modelimiz bunun 678 verisini doğru tahmin etmiştir. 254 adet veriyi ise “Tahmin Başarısız” olarak yanlış tahmin etmiştir. Bu şekilde bütün çizelgedeki başarılı ve başarısız tahminler oranlandığı zaman k-NN algoritmasının toplamdaki başarı oranı ortalaması %70,33 olarak hesaplanmıştır. Daha yüksek doğruluk oranına ulaşmak için otomatik modellemeyle belirlediğimiz diğer algoritmalar üzerinde çalışmalara devam edilmiştir.

A. 3. Naive Bayes ile Sınıflandırma

Naive Bayes algoritması, belirli bir veri kümesindeki değerlerin frekansını ve kombinasyonlarını sayarak bir olasılık kümesini hesaplayan basit bir olasılık sınıflandırıcısıdır [33]. Naive

Bayes sınıflandırıcısının temeli Bayes teoremine¹ dayanır. Naive Bayes algoritması basit bir algoritma olmasına rağmen genellikle karmaşık sınıflandırma yöntemlerinde iyi performans gösterebilir [32]. Çalışma için Naive Bayes algoritmasının parametresinde herhangi bir değişiklik yapılmamıştır. Varsayılan parametre kullanılmıştır. Naive Bayes algoritması kategorik verilerle çalışır. Bu veri seti için herhangi bir kategorik – numerik dönüşümüne ihtiyaç duyulmamıştır.

Naive Bayes algoritmasının veri seti üzerindeki başarısı Tablo 10’da verilmiştir.

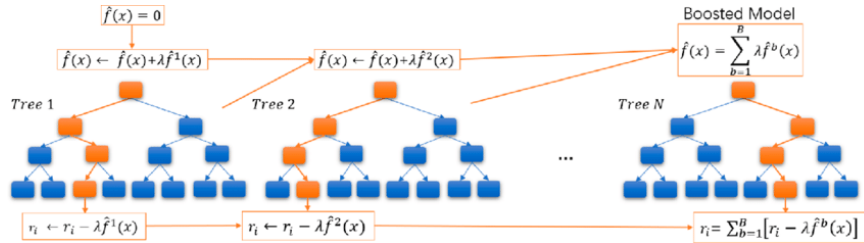
Tablo 10. Naive Bayes algoritması karşılaştırması.

	Gerçek Başarılı	Gerçek Başarısız	Doğruluk
Tahmin Başarılı	622	189	76,70%
Tahmin Başarısız	310	543	63,66%
Hatırlama	66,74%	74,18%	

Tabloda “Gerçek Başarılı” sayısı toplamda 932 veriyken modelimiz bunun 622 verisini doğru tahmin etmiştir. 310 adet veriyi ise “Tahmin Başarısız” olarak yanlış tahmin etmiştir. Bu şekilde bütün çizelgedeki başarılı ve başarısız tahminler oranlandığı zaman Naive Bayes algoritmasındaki toplamdaki başarı oranı ortalaması %70,01 olarak hesaplanmıştır. Daha yüksek doğruluk oranına ulaşmak için diğer algoritmalar üzerinde çalışmalara devam edilmiştir.

A. 4. Gradient Boosted Trees ile Sınıflandırma

Gradient Boosted Trees, bir grup zayıf tahmin modelleri biçimindeki regresyon ve sınıflandırma sorunları için bir makine öğrenimi yöntemidir. Bu yöntem, modelleri aşamalı olarak oluşturur ve isteğe bağlı türevlenebilir bir kayıp fonksiyonunun optimizasyonuna izin vererek modelleri geliştirir [34].



Şekil 11. Gradient Boosted Trees algoritması örneği [34].

Gradient Boosted Trees algoritmasında değişiklik yapılan parametrelerin açıklamaları aşağıda verilmiştir.

Number of trees: Bu parametre ile kullanılacak ağaç sayısını belirler.

Min Split improvement: Bu parametrede bir bölünmenin gerçekleşmesi için karesel hata azaltmada minimum ilgili gelişmeyi belirtir.

Learning rate: Bu parametrede öğrenme oranı belirlenir. Her döngüde hata sonucu learning rate oranı kadar tekrar hesaplama yapılır.

Bu çalışmadaki Gradient Boosted Trees algoritmasında değişiklik yapılan parametreler Tablo 11’de

¹ Bayes Teoremi: $P(A|B) = \frac{P(A) \cdot P(B|A)}{P(B)}$ [32]

gösterilmiştir. Sonucu etkilemediği görülen bazı parametreler varsayılan olarak bırakılmıştır.

Tablo 11. Gradient Boosted Trees algoritması için kullanılan parametreler.

Özellik Adı	Değer
Ağaç Sayısı	25
Minimum Bölünmüş İyileştirme Değeri	1.0E-4
Eğitim için Veri Setinin Tekrarlanma Sayısı	10
Öğrenme Oranı	0,191

Gradient Boosted Trees algoritmasının veri seti üzerindeki başarısı Tablo 12’de verilmiştir.

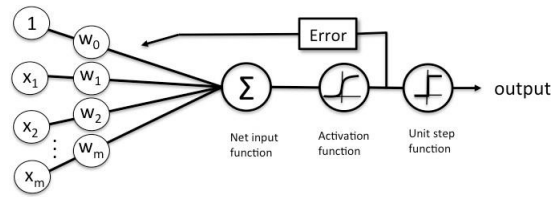
Tablo 12. Gradient Boosted Trees algoritması karşılaştırması.

	Gerçek Başarılı	Gerçek Başarısız	Doğruluk
Tahmin Başarılı	591	158	78,91%
Tahmin Başarısız	341	574	62,73%
Hatırlama	63,41%	78,42%	

Tabloda “Gerçek Başarılı” sayısı toplamda 932 veriyken modelimiz bunun 591 verisini doğru tahmin etmiştir. 341 adet veriyi ise “Tahmin Başarısız” olarak yanlış tahmin etmiştir. Bu şekilde bütün çizelgedeki başarılı ve başarısız tahminler oranlandığı zaman Gradient Boosted Trees algoritmasındaki toplamdaki başarı oranı ortalaması %70,01 olarak hesaplanmıştır. Daha yüksek doğruluk oranına ulaşmak için diğer algoritmalar üzerinde çalışmalara devam edilmiştir.

A. 5. Lojistik Regresyon ile Sınıflandırma

Lojistik Regresyon bağımlı ve bağımsız değişkenler arasındaki ilişkileri araştıran sınıflandırma algoritmasıdır. Lojistik Regresyon analizi ile verilerin hangi sınıfa ait olduğunu tahmin eden bir regresyon işlevi oluşturulur. Yani bağımlı değişkenlerin sürekli çıkış değerleri yerine sınıf üyelikleri tahmin edilir. Bu sürecin hepsi lojistik regresyon olarak tanımlanabilir [35]. Lojistik regresyonun Şekil 12’de şeması görülmektedir.



Şekil 12. Lojistik Regresyon Şeması [36].

Şekil 13’de lojistik regresyonun modeli görülmektedir.

$$y_i = \beta_0 + \beta_1 \times x_i + hata(e_i)$$

Bağımlı Değişken y_i ← Kesme Terimi (intercept) β_0 + Katsayı (coef., slope) $\beta_1 \times$ Bağımsız Değişken x_i + hata(e_i)

Şekil 13. Lojistik Regresyon modeli [37].

Lojistik Regresyon algoritmasında parametreler sonucu deęiřtirmedięinden dolayı varsayılan deęerler kullanılmıřtır.

Lojistik Regresyon algoritmasının veri seti üzerindeki bařarıřı Tablo 13’de verilmiřtir.

Tablo 13. Lojistik Regresyon algoritması karřılařtırması.

	Gerçek Bařarılı	Gerçek Bařarıřız	Doęruluk
Tahmin Bařarılı	699	208	77,07%
Tahmin Bařarıřız	233	524	69,22%
Hatırlama	75,00%	71,58%	

Tabloda ‘‘Gerçek Bařarılı’’ sayısı toplamda 932 veriyken modelimiz bunun 699 verisini doęru tahmin etmiřtir. 233 adet veriyi ise ‘‘Tahmin Bařarıřız’’ olarak yanlıř tahmin etmiřtir. Bu řekilde bütn izelgedeki bařarılı ve bařarıřız tahminler oranlandıęı zaman Lojistik Regresyon algoritmasındaki toplamdaki bařarı oranı ortalaması %73,50 olarak hesaplanmıřtır.

III. TARTIřMA

A. ALGORİTMALARIN KARřILAřTIRILMASI ve VM’DE BAřARI KRİTERLERİ

Literatrdeki alıřmaların oęunluęunda đrenci performansını farklı aıllardan hesaplamalar yer almaktadır. Bizim alıřmamızda derse katılma yntemleri đrenci performans tahminlerine gre yorumlanmıřtır. Derse katılma yntemlerine gre đrenci performansı literatrde rastlanmamıřtır. Aynı zamanda literatrde oęunlukla Weka programının kullanıldıęı grřmřtir. alıřmamızda aık kaynak kodlu ve akademiye cretsiz olan RapidMiner programı kullanılmıřtır.

Derin đrenme, k-NN, Naive Bayes, Gradient Boosted Trees ve Lojistik Regresyon algoritmaları karřılařtırıldıęında en iyi performansı sırasıyla %73,50 ile Lojistik Regresyon, %71,33 ile Derin đrenme, %70,73 ile k-NN, %70,01 ile Naive Bayes ve Gradient Boosted Trees algoritmaları vermiřtir.

Tablo 14’de algoritmaların eđitim ve tahmin iin kullanılan oranların bařarı oranına ve alıřma sresine etkileri verilmiřtir. Literatr genelinde yapılan alıřmalarda algoritmaların alıřma sreleri belirtilmemiřtir. alıřmamızda bu srelerde belirtilmiřtir.

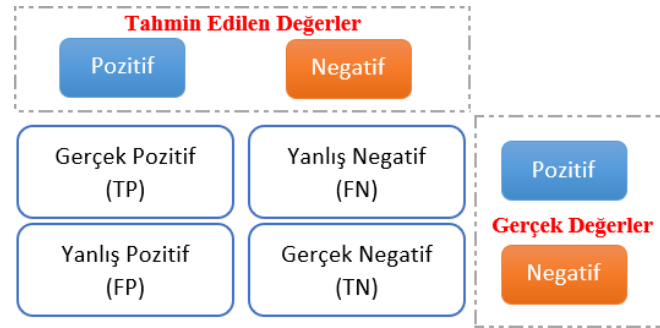
Tablo 14. Algoritmaların karřılařtırması.

Eđitim Verisi / Test Verisi	Tahmin Oranları / alıřma Sresi				
	Derin đrenme	k-NN	Naive Bayes	Gradient Boosted Trees	Lojistik Regresyon
%20 / %80	67,24% / 0 saniyeden az	65,41% / 2 saniye	68,38% / 0 saniyeden az	67,84% / 0 saniyeden az	59,85% / 0 saniyeden az
%40 / %60	64,58% / 1 saniye	67,97% / 2 saniye	68,79% / 0 saniyeden az	69,21% / 0 saniyeden az	71,49% / 0 saniyeden az
%50 / %50	68,74% / 1 saniye	67,37% / 3 saniye	68,53% / 0 saniyeden az	68,67% / 0 saniyeden az	71,12% / 0 saniyeden az
%60 / %40	70,36% / 1 saniye	68,95% / 3 saniye	69,46% / 0 saniyeden az	69,34% / 0 saniyeden az	72,23% / 0 saniyeden az
%80 / %20	71,33% / 2 saniye	70,73% / 2 saniye	70,01% / 0 saniyeden az	70,01% / 0 saniyeden az	73,50% / 0 saniyeden az

Tablo 14’de en başarılı tahmin oranının %80 eğitim ve %20 test verisiyle olduğu anlaşılmıştır. Doğruluk oranı ve çalışma süresine bakıldığında en iyi performansı Lojistik Regresyon algoritması vermiştir.

Çoğu literatürde olduğu gibi çalışmamızda da modellerin performansını belirlemek için karışıklık matrisi baz alınmıştır. Derse katılma yöntemlerine göre en iyi sonucun hesaplanması ise en yüksek model başarısına sahip Lojistik Regresyon algoritmasındaki sonuçların Senkron, Asenkron ve Hibrit için filtrelenerek bu filtreler sonucu elde edilen veriler üzerinde Karışıklık Matrisi hesaplanmıştır. Yani en başta veri setinde Asenkron, Senkron ve Hibrit yöntemler birlikte modellere verilmiş, tahmin performansı en iyi model olan Lojistik Regresyon sonucu bu veriler, bu üç özellik için ayrı ayrı Karışıklık Matrisi hesabı yapılmıştır. Bunların sonuçları Tablo 16,17,18,19’da detaylı bir şekilde verilmiştir.

Karışıklık Matrisi, modellerin tahmin performanslarının belirlenmesi için kullanılır [38]. Diğer bir deyişle tahmin edilen değerleri ve gerçek değerler üzerinde algoritmanın performansını belirlemek için kullanılır. Şekil 14’de ikili sınıflandırıcı karışıklık matrisinin bir örneği bulunmaktadır. Bu çalışmada ikili sınıflandırıcı karışıklık matrisi kullanılmıştır.



Şekil 14. İkili Sınıflandırıcı Karışıklık Matrisi.

Şekil 12’de gösterilen ikili sınıflandırıcı karışıklık matrisi verileri elde edildikten sonra, sınıflandırma modellerinin performansını belirlemek için Tablo 15’de gösterilen metrikler kullanılmıştır.

Tablo 15. Performans karşılaştırma metrikleri [39].

Metrik	Formül	Açıklama
Doğruluk Oranı (Accuracy Rate)	$\frac{TP + TN}{TP + TN + FP + FN}$	Genel olarak ne kadar doğru tahmin ettiğinin oranını verir.
Yanlış Sınıflandırma Oranı (Misclassification Rate)	$\frac{FP + FN}{TP + TN + FP + FN}$	Genel olarak hata oranını verir.
Gerçek Pozitif Değerlerin Oranı (True Positive Rate veya Recall)	$\frac{TP}{TP + FN}$	Gerçek pozitif değerlerini ne kadar doğru tahmin ettiğinin oranını verir.
Gerçek Negatif Değerlerin Oranı (True Negative Rate veya Specificity)	$\frac{TN}{TN + FP}$	Gerçek negatif değerlerini ne kadar doğru tahmin ettiğinin oranını verir.
Kesinlik Oranı (Precision)	$\frac{TP}{TP + FP}$	Doğru tahminlerin ne kadar kesin olduğu oranını verir.
F Skoru (F Score)	$\frac{2TP}{2TP + FP + FN}$	Dengesiz sınıflar için yararlı hibrit sokurunu verir.

Lojistik Regresyon algoritmasıyla en iyi tahmin değerlerine ulaşılmışken, karışıklık matrisiyle derse katılma yöntemlerinden hangisinin daha verimli olduğu hesaplanmıştır.

Tablo 16. Karışıklık matrisi tablosu (Asenkron).

Asenkron	Başarılı (Tahmin)	Başarısız (Tahmin)
Başarılı (Gerçek)	298 (TP)	157 (FN)
Başarısız (Gerçek)	105 (FP)	321 (TN)

Tablo 16’da asenkron yönteminin karışıklık matrisinin tablosu oluşturulmuştur. Bu tabloya göre hesaplamalar yapılmış ve Tablo 19’da verilmiştir.

Tablo 17. Karışıklık matrisi tablosu (Senkron).

Senkron	Başarılı (Tahmin)	Başarısız (Tahmin)
Başarılı (Gerçek)	133 (TP)	68 (FN)
Başarısız (Gerçek)	51 (FP)	175 (TN)

Tablo 17’de senkron yönteminin karışıklık matrisinin tablosu oluşturulmuştur. Bu tabloya göre hesaplamalar yapılmış ve Tablo 19’da verilmiştir.

Tablo 18. Karışıklık matrisi tablosu (Hibrit).

Hibrit	Başarılı (Tahmin)	Başarısız (Tahmin)
Başarılı (Gerçek)	268 (TP)	8 (FN)
Başarısız (Gerçek)	52 (FP)	28 (TN)

Tablo 18’de hibrit yönteminin karışıklık matrisinin tablosu oluşturulmuştur. Bu tabloya göre hesaplamalar yapılmış ve Tablo 19’da verilmiştir.

Tablo 19. Karışıklık matrisi derse katılma yöntemlerinin karşılaştırılması.

Karışıklık Matrisi Hesaplamaları	Derse Katılma Yöntemleri		
	Asenkron	Senkron	Hibrit
Doğruluk Oranı	70,26%	67,39%	83,14%
Yanlış Sınıflandırma Oranı	29,73%	26,03%	16,85%
Gerçek Pozitif Değerlerin Oranı	68,49%	66,16%	97,10%
Gerçek Negatif Değerlerin Oranı	75,35%	77,43%	35,00%
Hassasiyet Oranı	73,94%	72,28%	83,75%
F Skoru	69,46%	69,09%	89,93%

Tablo 19’da derse katılma yöntemlerinin karışıklık matrisi hesaplamalarının karşılaştırması yapılmıştır. Bu karşılaştırmadaki değerler göz önüne alındığında, en iyi sonucu veren derse katılma yöntemi sırasıyla Hibrit, Asenkron ve Senkron olarak görülmektedir. Bir dersten gerçekte başarılı iken tahminde başarılı olma durumunun en yüksek performansı (öğrenci başarı durumu) Hibrit yöntem ile işlenmiş derslerde alınmıştır. Bir dersten gerçekte başarısız iken tahminde başarısız olma durumunun en yüksek performansı (öğrenci başarı durumu) Senkron yöntem ile işlenmiş derslerde alınmıştır.

IV. SONUC

Bütün dünyayı etkisi altına alan Covid-19 pandemisiyle birlikte üniversitelerin uzaktan eğitim sistemlerinde biriken verileri odağına alan bu çalışmada, bir üniversiteye ait uzaktan eğitim sisteminden elde edilmiş veriler üzerinden öğrenciler için gelecek dönemlerde en güvenilir olabilecek uzaktan eğitim modeli, VM yoluyla tahmin edilmeye çalışılmıştır. Öğrencilerin uzaktan eğitim sisteminde derse Senkron, Asenkron ve Hibrit katılma biçimlerinin öğrenci başarısına etkilerinin VM

yoluyla tahmin edilip edilmeyeceği sorusu araştırmanın temel problemidir. Bu problem çerçevesinde yürütülen araştırmada RapidMiner yardımcı programı ile tahmin edici metotlarla inceleme ve analizler yapılmıştır.

Araştırmanın sonucunda VM modellerinden Lojistik Regresyon algoritması kullanılarak elde edilen sonuçlar doğrultusunda, üniversite öğrencilerinin başarı durumları tahmin edilmek istendiğinde uzaktan eğitim sürecinde en güvenilir sonucu veren derse katılma yöntemi olan “Hibrit” olarak seçilmesi tavsiye edilmektedir.

Gelecekte yapılacak çalışmalar için, üniversitelerin Fakülte, MYO, Bölüm ve ders bazlı olarak bu VM modelleri uygulanarak çalışmalar yapılabilir. Bu alanda yapılacak somut çalışmalar şu şekilde sıralanabilir;

1. Çalışmalar ders bazlı yapılarak hangi yöntemin en güvenilir olduğunun tahmin edilmesi
2. Çalışmalar bölüm bazlı yapılarak hangi yöntemin en güvenilir olduğunun tahmin edilmesi
3. Çalışmalar Fakülte/MYO bazlı yapılarak hangi yöntemin en güvenilir olduğunun tahmin edilmesi
4. Yukarıdaki maddeleri tekrar baz alarak öğrencilerin cinsiyeti, yaşı, evinin internet bağlantısı, oturduğu bölge, evindeki öğrenme ortamı, kişisel bilgisayar vb. öznitelikler de çalışmaya dahil edilerek tahminler yapılabilir.

IV. KAYNAKLAR

[1] World Health Organization. (2021, November 26). *Rolling updates on coronavirus disease (COVID-19)*. World Health Organization [Online]. Available: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/events-as-they-happen>.

[2] T.C. Sağlık Bakanlığı. (2021, 26 Kasım). *Pandemi*. T.C. Sağlık Bakanlığı [Çevrimiçi]. Erişim: <https://covid19.saglik.gov.tr/TR-66494/pandemi.html>.

[3] T.C. Yükseköğretim Kurulu. (2020, 13 Mart). *Koronavirüs (Covid-19) Bilgilendirme Notu: 1*. T.C. Yükseköğretim Kurulu [Çevrimiçi]. Erişim: <https://covid19.yok.gov.tr/Documents/alinan-kararlar/02-coronavirus-bilgilendirme-notu-1.pdf>.

[4] T.C. Yükseköğretim Kurulu. (2020, 6 Mart). *Basın Açıklaması Yükseköğretim Kurulu Başkanı Prof. Dr. M. A. Yekta Saraç* [Çevrimiçi]. Erişim: <https://covid19.yok.gov.tr/Documents/alinan-kararlar/03-uzaktan-egitime-iliskin-alinan-karar.pdf>.

[5] T.C. Yükseköğretim Kurulu. (2020, 26 Mart). *Basın Açıklaması Yükseköğretim Kurulu Başkanı Prof. Dr. M. A. Yekta Saraç* [Çevrimiçi]. Erişim: <https://covid19.yok.gov.tr/Documents/alinan-kararlar/04-uzaktan-egitim-ve-yks-ertelenmesine-iliskin.pdf>.

[6] T.C. Yükseköğretim Kurulu. (2020, 4 Haziran). *YÖK'ten Küresel Salgın ile Mücadele Sürecinde Yeni Düzenlemeler-II* [Çevrimiçi]. Erişim: <https://covid19.yok.gov.tr/Documents/alinan-kararlar/23-kuresel-salgin-ile-mucadele-gerceklesen-yapisal-duzenlemeler-2.pdf>.

[7] T.C. Yükseköğretim Kurulu. (2020, 3 Mayıs). *Üniversitelerimizde Küresel Salgın Sürecinde Uzaktan Öğretim Uygulamaları Anket Sonuçları* [Çevrimiçi]. Erişim: <https://covid19.yok.gov.tr/Documents/AnaSayfa/uzaktan-ogretim-anketi.pdf>.

- [8] T.C. Yükseköğretim Kurulu. (2020, 30 Temmuz). *Küresel Salgında Yeni Normalleşme Süreci* [Çevrimiçi]. Erişim: <https://www.yok.gov.tr/Documents/Yayinlar/Yayinlarimiz/2020/kuresel-salginda-yeni-normallesme-sureci-2020.pdf>.
- [9] T.C. Yükseköğretim Kurulu. (2021, 3 Eylül). *Küresel Salgında Eğitim ve Öğretim Süreçlerine Yönelik Uygulamalar Rehberi* [Çevrimiçi]. Erişim: <https://www.yok.gov.tr/Documents/Yayinlar/Yayinlarimiz/2021/K%C3%BCresel%20Salg%C4%B1nda%20E%C4%9Fitim%20ve%20%C3%96%C4%9Fretim%20S%C3%BCre%C3%A7lerine%20Y%C3%B6nelik%20Uygulamalar%20K%C4%B1lavuzu%202021.pdf>.
- [10] A. Koyuncuğil ve N. Özgülbaş, "Veri Madenciliği: Tıp ve Sağlık Hizmetlerinde Kullanımı ve Uygulamaları", *Bilişim Teknolojileri Dergisi*, c. 2, sayı. 2, Tem. 2010.
- [11] Ş. Can, T. Özdil ve C. Yılmaz, "Üniversite Eğitiminde Öğretim Üyesi Tutumunun Ders Başarısına Etkisinin Veri Madenciliği Teknikleriyle Araştırılması," *Manisa Celal Bayar Üniversitesi Yönetim ve Ekonomi Dergisi*, c. 26, s. 2, ss. 609-623, 2019.
- [12] O. Gushchina, and A. Ochepovsky, "Data Mining for the e-Learning Risk Management," *Anadolu University Turkish Online Journal of Distance Education*, vol. 20, no. 3, pp. 181-196, 2019.
- [13] Ö. Özbay ve H. Ersoy, "Öğrenme Yönetim Sistemi Üzerindeki Öğrenci Hareketliliğinin Veri Madenciliği Yöntemleriyle Analizi," *Gazi Üniversitesi Gazi Eğitim Fakültesi Dergisi*, c. 37, sayı. 2, ss. 523-558, Ağu. 2017.
- [14] K. Abe, "Data Mining and Machine Learning Applications for Educational Big Data in the University," *2019 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCCom/CyberSciTech)*, 2019, pp. 350- 355, doi: 10.1109/DASC/PiCom/CBDCCom/CyberSciTech.2019.00071.
- [15] A. Jain and H. Ram Sah, "Student's Feedback by emotion and speech recognition through Deep Learning," *2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)*, 2021, pp. 442-447, doi: 10.1109/ICCCIS51004.2021.9397145.
- [16] W. Villegas-Ch, S. Luján-Mora, D. Buenaño-Fernandez and M. Román-Cañizares, "Analysis of web-based learning systems by data mining," *2017 IEEE Second Ecuador Technical Chapters Meeting (ETCM)*, Salinas, 2017, pp. 1-5, doi: 10.1109/ETCM.2017.8247553.
- [17] H. C. Hung, I. F. Liu, C. T. Liang and Y. S. Su, "Applying educational data mining to explore students' learning patterns in the flipped learning approach for coding education," *Symmetry*, vol. 12, no. 2, 2020.
- [18] M. Blagojević and Ž. Micić, "A web-based intelligent report e-learning system using data mining techniques," *Computers and Electrical Engineering*, vol. 39, no. 2, pp. 465 – 474, 2013.
- [19] H. Mohammad, "Application of decision tree algorithm for predicting students' performance via online learning during coronavirus pandemic," *Journal of Theoretical and Applied Information Technology*, vol. 99, no. 19, pp. 4546 – 4556, 2021.
- [20] W. F. W. Yaacob, S. Nasir, "Supervised data mining approach for predicting student performance," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 16, no. 3, pp. 1584 – 1592, 2019.
- [21] C. Tan, J. Lin, "A new QoE-based prediction model for evaluating virtual education systems with COVID-19 side effects using data mining." *Soft Computing*, pp. 1 – 15, 2021.

- [22] L. Masangu, A. Jadhav, R. Ajoodha, "Predicting student academic performance using data mining techniques," *Advances in Science, Technology and Engineering Systems Journal*, vol. 6, no. 1, pp. 153 – 163, 2021.
- [23] D. K. Dake, D. D. Essel and J. E. Agbodaze, "Using Machine Learning to Predict Students' Academic Performance During Covid-19," *2021 International Conference on Computing, Computational Modelling and Applications (ICCMA)*, 2021, pp. 9-15, doi: 10.1109/ICCMA53594.2021.00010.
- [24] I. L. H. Alsammak, A. H. Mohammed, I. S. Nasir, "E-learning and covid-19: Predicting student academic performance using data mining algorithms," *Webology*, vol. 19, no. 1, pp. 3419 – 3432, 2022.
- [25] Ö. Şen ve G. Kızılcıoğlu, "COVID-19 PANDEMİ SÜRECİNDE ÜNİVERSİTE ÖĞRENCİLERİNİN VE AKADEMİSYENLERİN UZAKTAN EĞİTİME YÖNELİK GÖRÜŞLERİNİN BELİRLENMESİ", *International Journal of 3D Printing Technologies and Digital Industry*, c. 4, sayı. 3, ss. 239-252, Ara. 2020, doi:10.46519/ij3dptdi.830913.
- [26] M. Keskin ve D. Özer Kaya, "COVID-19 Sürecinde Öğrencilerin Web Tabanlı Uzaktan Eğitime Yönelik Geri Bildirimlerinin Değerlendirilmesi", *İzmir Kâtip Çelebi Üniversitesi Sağlık Bilimleri Fakültesi Dergisi*, c. 5, sayı. 2, ss. 59-67, Haz. 2020.
- [27] N. Y. Şimşek. (2019 24 Şubat). *Derin Öğrenme (Deep Learning) Nedir ve Nasıl Çalışır?*. nyilmazsimsek.medium.com. [Çevrimiçi]. Erişim: <https://nyilmazsimsek.medium.com/derin-%C3%B6%C4%9Frenme-deep-learning-nedir-ve-nas%C4%B1l-%C3%A7al%C4%B1%C5%9F%C4%B1r-2d7f5850782>.
- [28] B. Kotan, "Network monitoring system using machine learning comparative analysis of classification techniques for network traffic monitoring," *M.S. thesis, Electronics And Computer Engineering, Hasan Kalyoncu University, Gaziantep, Turkey*, 2019.
- [29] A. Kızrak. (2019, May 9). *Comparison of Activation Functions for Deep Neural Networks* [Online]. Towards Data Science. Available: <https://towardsdatascience.com/comparison-of-activation-functions-for-deep-neural-networks-706ac4284c8a>.
- [30] RapidMiner. (2022, April 13). *Deep Learning* [Online]. Available: https://docs.rapidminer.com/latest/studio/operators/modeling/predictive/neural_nets/deep_learning.html.
- [31] S. Kırıçoğlu ve A. Yakupoğlu, "Veri Madenciliği ile Üniversite Bilişim Teknik Servis Hizmetleri Analizi", *Düzce Üniversitesi Bilim ve Teknoloji Dergisi*, c. 8, sayı. 1, ss. 326-333, Oca. 2020, doi:10.29130/dubited.593830.
- [32] K. Kotan, "Comparative analysis of classification techniques for network anomalies management," *M.S. thesis, Electronics And Computer Engineering, Hasan Kalyoncu University, Gaziantep, Turkey*, 2019.
- [33] M. Saritas and A. Yasar, "Performance Analysis of ANN and Naive Bayes Classification Algorithm for Data Classification," *IJISAE*, vol. 7, no. 2, pp. 88-91, Jun. 2019.
- [34] Z. Jiao, P. Hu, H. Xu, Q. Wang, "Machine learning and deep learning in chemical health and safety: a systematic review of techniques and applications," *ACS Chemical Health & Safety*, vol. 27, no. 6, pp. 316 – 334, 2020.

- [35] U. Çelik, E. Akçetin ve M. Gök, Editörler, *RapıMiner ile Uygulamalı Veri Madenciliği*, 1. baskı, İstanbul, Türkiye: Pusula 20 Teknoloji ve Yayıncılık A.Ş., 2017, böl. 6, ss. 81-94.
- [36] B. Aksoy ve D. Boztosun, "Diskriminant ve Lojistik Regresyon Yöntemleri Kullanılarak Finansal Başarısızlık Tahmini: BIST İmalat Sektörü Örneği," *İstanbul Okan Üniversitesi Finans Politik ve Ekonomik Yorumlar*, s. 646, ss. 9-32, 2018.
- [37] C. Erden, *Python ile Veri Madenciliği*, 1. baskı, İstanbul, Türkiye: KODLAB Yayın Dağıtım Yazılım Ltd. Şti., 2021, böl. 8-9, ss. 185-225.
- [38] M. Hasnain, M. F. Pasha, I. Ghani, M. Imran, M. Y. Alzahrani and R. Budiarto, "Evaluating Trust Prediction and Confusion Matrix Measures for Web Services Ranking," in *IEEE Access*, vol. 8, pp. 90847-90861, 2020, doi: 10.1109/ACCESS.2020.2994222.
- [39] A. Amidi and S. Amidi. (2021, December 27). *Machine Learning tips and tricks cheatsheet. Stanford University* [Online]. Available: <https://stanford.edu/~shervine/teaching/cs-229/cheatsheet-machine-learning-tips-and-tricks>.