



**T.C.
DÜZCE ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ**

**PANDEMİ SÜRECİNDE UZAKTAN EĞİTİMDE SENKRON,
ASENKRON VE HİBRİT YAPILMIŞ DERSLERDE VERİ
MADENCİLİĞİ İLE ÖĞRENCİ PERFORMANS ANALİZİ**

MEHMET YILDIRIM

**YÜKSEK LİSANS TEZİ
ELEKTRİK-ELEKTRONİK VE BİLGİSAYAR MÜHENDİSLİĞİ
ANABİLİM DALI**

**DANIŞMAN
DR. ÖĞR. ÜYESİ SERDAR KIRIŞOĞLU**

DÜZCE, 2022

T.C.
DÜZCE ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

**PANDEMİ SÜRECİNDE UZAKTAN EĞİTİMDE SENKRON,
ASENKRON VE HİBRİT YAPILMIŞ DERSLERDE VERİ
MADENCİLİĞİ İLE ÖĞRENCİ PERFORMANS ANALİZİ**

Mehmet YILDIRIM tarafından hazırlanan tez çalışması aşağıdaki jüri tarafından Düzce Üniversitesi Lisansüstü Eğitim Enstitüsü Elektrik-Elektronik ve Bilgisayar Mühendisliği Anabilim Dalı'nda **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Tez Danışmanı

Dr. Öğr. Üyesi Serdar KIRIŞOĞLU

Düzce Üniversitesi

Jüri Üyeleri

Dr. Öğr. Üyesi Serdar KIRIŞOĞLU

Düzce Üniversitesi

Dr. Öğr. Üyesi Arafat ŞENTÜRK

Düzce Üniversitesi

Dr. Öğr. Üyesi Vedat MARTTİN

Bilecik Şeyh Edebali Üniversitesi

Tez Savunma Tarihi: 29/06/2022

BEYAN

Bu tez çalışmasının kendi çalışmam olduğunu, tezin planlanmasından yazımına kadar bütün aşamalarda etik dışı davranışımın olmadığını, bu tezdeki bütün bilgileri akademik ve etik kurallar içinde elde ettiğimi, bu tez çalışmasıyla elde edilmeyen bütün bilgi ve yorumlara kaynak gösterdiğimi ve bu kaynakları da kaynaklar listesine aldığımı, yine bu tezin çalışılması ve yazımı sırasında patent ve telif haklarını ihlal edici bir davranışımın olmadığını beyan ederim.

29 Haziran 2022

Mehmet YILDIRIM



TEŐEKKÜR

Yüksek lisans öğrenimimde ve bu tezin hazırlanmasında gösterdiği her türlü destek ve yardımdan dolayı çok değerli hocam Dr. Öğr. Üyesi Serdar KIRIŐOĐLU'na en içten dileklerle teşekkür ederim.

Tez çalışmam boyunca değerli katkılarını, bilgisini, desteğini ve tecrübesini esirgemeyen Dr. Gülcan PALO'ya da şükranlarımı sunarım.

Bu çalışmada kullanılan verilerin alınması konusunda yardımlarını esirgemeyen Advancity firmasına, Umut TÜRK'e ve KAYUZEM Müdürü Öğr. Gör. Mustafa KALAY hocama teşekkür ederim.

Ömrümün her anında yanımda yer alan bir an bile beni yalnız bırakmayan, bugünlere gelmemde bana en büyük değeri katan annem Şükran YILDIRIM'a, hiçbir zaman tükenmeyecek saygı ve sevgiyle en özel şükranlarımı sunuyorum.

Yoğun çalışma döneminde yanımda olduğunu hissettiren her daim yükümü hafifleten iyi bir eş olan Demet YILDIRIM'a teşekkür ediyorum.

Aramızdan ayrılışının 8. yıl dönümü ile aynı günde tezimi sunmuş olduğum saygı değer babam Fehmi YILDIRIM'ı sevgi ve özlemlerle anıyorum.

29 Haziran 2022

Mehmet YILDIRIM

İÇİNDEKİLER

Sayfa No

ŞEKİL LİSTESİ.....	vii
ÇİZELGE LİSTESİ.....	viii
KISALTMALAR.....	ix
SİMGELER	x
ÖZET	xi
ABSTRACT	xii
1. GİRİŞ.....	1
2. VERİ MADENCİLİĞİ.....	13
2.1. BİLGİ KEŞFİ SÜRECİ	15
2.1.1. Veri Ön İşleme.....	16
2.1.1.1. Veri Temizleme.....	16
2.1.1.2. Veri Bütünleştirme	16
2.1.1.3. Veri İndirgeme	17
2.1.2. Veri Dönüştürme.....	17
2.1.3. Veri Madenciliği Yöntemini Uygulama	17
2.1.4. Değerlendirme	17
2.2. VERİ MADENCİLİĞİ YÖNTEMLERİ.....	17
2.2.1. Denetimli Öğrenme (Supervised Learning).....	17
2.2.1.1. Sınıflandırma Algoritması.....	19
2.2.1.2. Regresyon Algoritması.....	19
2.2.2. Denetimsiz Öğrenme (Unsupervised Learning)	19
2.2.2.1. Kümeleme.....	20
2.2.2.2. Birliklilik Kuralı.....	20
2.3. VERİ MADENCİLİĞİ İÇİN KULLANILAN YÖNTEMLER	20
2.3.1. Derin Öğrenme	20
2.3.2. k-NN	21
2.3.3. Naive Bayes	23
2.3.4. Gradient Boosted Trees	24
2.3.5. Lojistik Regresyon	24
2.4. KARIŞIKLIK MATRİSİ.....	25
3. METOD VE ANALİZ.....	28
3.1. VERİLER VE ÖZELLİKLERİ.....	29
3.2. ÇALIŞMADA KULLANILAN PROGRAM.....	30
3.3. RAPIDMINER İLE VERİLERİN ANALİZİ.....	30
3.4. VERİ MADENCİLİĞİ YOLUYLA ELDE EDİLEN BULGULAR.....	35
3.5. ÇALIŞMADA KULLANILAN SINIFLANDIRMA YÖNTEMLERİ	37
3.5.1. Derin Öğrenme ile Sınıflandırma	38
3.5.2. k-NN ile Sınıflandırma.....	41
3.5.3. Naive Bayes ile Sınıflandırma	42
3.5.4. GBT ile Sınıflandırma.....	43
3.5.5. Lojistik Regresyon ile Sınıflandırma.....	45

3.5.6. Algoritmaların Karşılaştırılması	46
3.6. VERİ MADENCLİĞİ BAŞARI KRİTERLERİ (KARIŞIKLIK MATRİSİ) 47	
4. SONUÇLAR VE ÖNERİLER.....	50
4.1. SONUÇLAR.....	50
4.2. ÖNERİLER.....	52
5. KAYNAKLAR	53
6. EKLER	59
6.1. EK 1: ARAŞTIRMA İZİN YAZISI.....	59
6.2. EK 2: KVKK TUTANAĞI	60
6.3. EK 3: KAYSERİ ÜNİVERSİTESİ ETİK KURUL BAŞKANLIĞI İZİNİ	61
ÖZGEÇMİŞ.....	62



ŞEKİL LİSTESİ

Sayfa No

Şekil 1.1. Pandemi döneminde üniversitelerde kullanılan ÖYS türleri.....	6
Şekil 1.2. Pandemi döneminde üniversitelerde kullanılan canlı sınıf platformları.....	6
Şekil 2.1. VM tarihi [38].....	13
Şekil 2.2. VM ve disiplinler [2].	15
Şekil 2.3. Bilgi keşfi sürecinin adımları [2].	15
Şekil 2.4. Denetimli öğrenme [49].	18
Şekil 2.5. Denetimsiz öğrenme [49].	19
Şekil 2.6. Kümeleme örneği [55].	20
Şekil 2.7. Sınıf hücresi yapısı [59].	21
Şekil 2.8. Derin öğrenme şeması.	21
Şekil 2.9. k-NN örneği.	22
Şekil 2.10. Farklı k değerine sahip iki sınıfı ayıran farklı sınırlar [64].	23
Şekil 2.11. Naive Bayes örneği.	24
Şekil 2.12. Gradient Boosted Trees örneği [67].	24
Şekil 2.13. Lojistik regresyon şeması [71].	25
Şekil 2.14. Lojistik regresyon modeli [72].	25
Şekil 2.15. İkili sınıflandırıcı karışıklık matrisi.	26
Şekil 3.1. Kayseri Üniversitesi ALMS kullanıcı giriş ekranı.	28
Şekil 3.2. RapidMiner giriş ekranı.	30
Şekil 3.3. RapidMiner programına aktarılmadan önceki veriler.	31
Şekil 3.4. RapidMiner “new process” oluşturma ekranı.	31
Şekil 3.5. RapidMiner veri yükleme ekranı.	32
Şekil 3.6. RapidMiner veri setleri.	32
Şekil 3.7. RapidMiner join operatörü ile veri seti birleştirme ekranı.	33
Şekil 3.8. RapidMiner rol belirleme operatörü ve parametreleri.	33
Şekil 3.9. RapidMiner verileri bölme operatörü ve parametreleri.	34
Şekil 3.10. RapidMiner verileri bölme operatörü ve parametreleri.	34
Şekil 3.11. Çalışmada kullanılan operatörler.	35
Şekil 3.12. Auto model doğruluk oranı.	36
Şekil 3.13. Auto model işlem süresi (ms).	36
Şekil 3.14. Sınıflandırma süreci.	37
Şekil 3.15. Doğrultucu grafiği.	39

ÇİZELGE LİSTESİ

Sayfa No

Çizelge 3.1. Not baremi tanımlama çizelgesi.	29
Çizelge 3.2. Kategorik değişkenlerin sayısal değişkenlere dönüştürülmesi süreci.	38
Çizelge 3.3. Derin öğrenme aktivasyon fonksiyonlarının karşılaştırılması.	39
Çizelge 3.4. Derin öğrenme gizli katman ve nöronların karşılaştırılması.	39
Çizelge 3.5. Derin öğrenme veri kümesi yinelenmesi karşılaştırılması.	40
Çizelge 3.6. Derin öğrenme algoritması için kullanılan parametreler.	41
Çizelge 3.7. Derin öğrenme algoritması karşılaştırması.	41
Çizelge 3.8. k-NN algoritmasının k değerinin karşılaştırılması.	42
Çizelge 3.9. K-NN algoritması karşılaştırması.	42
Çizelge 3.10. Naive Bayes algoritması karşılaştırması.	43
Çizelge 3.11. GBT algoritmasının ağaç sayısı değerinin karşılaştırılması.	43
Çizelge 3.12. GBT algoritmasının minimum bölünmüş iyileştirme değerinin karşılaştırılması.	44
Çizelge 3.13. GBT algoritmasının öğrenme oranı değerinin karşılaştırılması.	44
Çizelge 3.14. GBT algoritması için kullanılan parametreler.	44
Çizelge 3.15. GBT algoritması karşılaştırması.	44
Çizelge 3.16. Lojistik Regresyon algoritması karşılaştırması.	46
Çizelge 3.17. Algoritmaların karşılaştırması.	46
Çizelge 3.18. Algoritmaların çalışma sürelerinin karşılaştırması.	47
Çizelge 3.19. Karışıklık matrisi çizelgesi (Asenkron).	47
Çizelge 3.20. Karışıklık matrisi çizelgesi (Senkron).	48
Çizelge 3.21. Karışıklık matrisi çizelgesi (Hibrit).	48
Çizelge 3.22. Karışıklık matrisi derse katılma yöntemlerinin karşılaştırılması.	49

KISALTMALAR

ALMS	Advancity Öğrenme Yönetim Sistemi
GBT	Gradient Boosted Trees
KAYUZEM	Kayseri Üniversitesinin Uzaktan Eğitim Uygulama ve Araştırma Merkezi
k-NN	k-En Yakın Komşu
KVKK	Kişisel Verileri Koruma Kanunu
Ms-SQL	Microsoft Soft-Structured Query Language
MYO	Meslek Yüksek Okulu
ÖYS	Öğrenme Yönetim Sistemi
REP	Reduced Error Pruning
ROC	Receiver Operating Characteristic
SQL	Veri Tabanı Sorgulama Dili
UZEM	Uzaktan Eğitim Araştırma ve Uygulama Merkezi
VM	Veri Madenciliği
VTBK	Veri Tabanı Bilgi Keşfi
VTYS	Veri Tabanı Yönetim Sistemi
YÖK	Yüksek Öğretim Kurumu
YSA	Yapay Sinir Ağları
WEKA	Waikato Environment for Knowledge Analysis

SİMGELER

Σ

Sigma



ÖZET

PANDEMİ SÜRECİNDE UZAKTAN EĞİTİMDE SENKRON, ASENKRON VE HİBRİT YAPILMIŞ DERSLERDE VERİ MADENCİLİĞİ İLE ÖĞRENCİ PERFORMANSI ANALİZİ

Mehmet YILDIRIM

Düzce Üniversitesi

Lisansüstü Eğitim Enstitüsü, Elektrik-Elektronik ve Bilgisayar Mühendisliği Anabilim
Dalı

Yüksek Lisans Tezi

Danışman: Dr. Öğr. Üyesi Serdar KIRIŞOĞLU

Haziran 2022, 61 sayfa

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 tavsiyesi ile üniversiteler eğitimlerine uzaktan Asenkron, Senkron ve Hibrit yöntemlerini kullanarak devam etmiş, hatta sınavları uzaktan eğitim sisteminde yapmak zorunda kalmışlardır. Bu araştırmada, Kayseri Üniversitesinin Uzaktan Eğitim Uygulama ve Araştırma Merkezi sisteminden alınmış veriler kullanılmıştır. Araştırma kapsamında 8319 işlenmiş veri bulunmaktadır. Bu veriler üzerinde Veri Madenciliği alanında kullanılan RapidMiner programının otomatik modelleme özelliği kullanılarak varsayılan algoritmalarla geleceğe yönelik tahminleme işlemi yapılmıştır. Bu çalışmada RapidMiner programının varsayılan algoritmaları arasından en iyi sonucu veren Derin Öğrenme, Naive Bayes, Gradient Boosted Trees, Lojistik Regresyon algoritmaları, parametreleri değiştirilerek en iyi sonucu almaya yönelik derinlemesine incelenmiştir. Ayrıca bu otomatik modelleme de yer almayan k-En Yakın Komşu algoritması da bu çalışmaya dahil edilmiştir. Bu 5 algoritmanın parametreleri üzerinde değişiklikler yapılarak daha iyi sonuçlar elde edilmeye çalışılmıştır. Öğrenci başarısına göre en iyi tahminleme sonucunu, %73,50 ile 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 sözcükler: Covid-19, Pandemi, RapidMiner, Uzaktan eğitim, Veri madenciliği.

ABSTRACT

STUDENT PERFORMANCE ANALYSIS WITH DATA MINING IN DISTANCE EDUCATION SYNCHRONOUS, ASYNCHRONOUS AND HYBRID COURSES IN THE PANDEMIC PROCESS

Mehmet YILDIRIM

Düzce University

Graduate School of Natural and Applied Sciences, Department of Electrical-
Electronical and Computer Engineering

Master's Thesis

Supervisor: Assist. Prof. Dr. Serdar KIRISOGLU

June 2022, 61 pages

In these days, 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 remotely 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 the RapidMiner program used in Data Mining on these data, future predictions were made with default algorithms. In this study, Deep Learning, Naive Bayes, Gradient Boosted Trees, Logistic Regression algorithms, which give the best results among the default algorithms of RapidMiner program, were examined in depth to get the best result by changing their parameters. In addition, the k-Nearest Neighbor algorithm, which is not included in this automatic modeling, is also included in this 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 with 73.50%. The effect of all of 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 study, 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: Covid-19, Data mining, Distance education, Pandemic, RapidMiner.

1. GİRİŞ

Geçmişten günümüze, teknolojinin gelişmesi verilerin bilgisayarlar ile işlenmesi ve depolanması ile birlikte verilerde hızlı bir artış gözlemlenmiştir. Bunun sonucu olarak VM alanının gün geçtikçe önemi artmıştır. VM'nin tarihteki yerini incelediğimizde, ilk olarak 1950'de bilgisayarların sayımı ile gündeme gelmiş, daha sonra 1960'larda ise veri tabanı ile verilerin depolanması ve veri yönetim sistemlerinin oluşturulması söz konusu olmuştur. İstatistikçiler regresyon analizi, en büyük olasılık kestirim vb. gibi yeni algoritmalar üzerinde çalışmışlardır. Bu gelişmelere paralel olarak veri tabanı sistemleri giderek gelişmiştir. Böylelikle büyük verilerin depolanması ve veri tabanlarında bilgi arama yöntemleri bilgi teknolojileri literatüründe yerini almıştır [1]. 1970'lerde ilişkisel veri tabanı yönetim sistemleri uygulamaları kullanılmış, basit anlamda makine öğrenimi sağlanmıştır. 80'lerde çeşitli sektörlerde VTYS büyük veriler bulunmamaktaydı. Dolayısıyla bu veriler SQL veri tabanı sorgulama dili gibi veri tabanı sorgulama dilleri kullanılarak ulaşılabilir durumdaydı. 90'larda ise veri tabanlarındaki veri miktarları artmış ve literatürde yeni çalışmalar ortaya çıkmıştır [2]. 2000'li yıllarda farklı türlerdeki veriler bulut ortamında depolanmaya başlanmıştır. Çünkü depolanmış verilere farklı veri türleri (sosyal medya paylaşımları, log dosyaları, bloglar, konum bilgisi vb.) dahil edilmiştir. Hacim, hız ve çeşitlilik bakımından artan günümüz verileri büyük veri analitiği yaklaşımına yol açmıştır [3]. VM günümüzde birçok alanda kullanılmakta ve yaygınlaşmaya devam etmektedir. Birçok araştırmacı literatüre katkı sağlamak için VM kullanmıştır. Literatürdeki somut örnekler bakacak olursak:

Sebik ve Bülbül [4], VM'nin yaygın olarak kullanıldığı tıp alanında en yaygın kanser türü olan akciğer kanseri teşhisi hakkında yaptıkları çalışmada, sağlık veri tabanı içerisinden anonim bir şekilde önceden teşhisi konmuş, akciğer teşhisi konmamış hasta ve hasta adayların verilerini kullanmışlardır. Verilerin ön işleme sürecinden sonra toplamda 404 adet veri ve sonucu doğrudan etki eden 9 adet öznelik verisi ile çalışmayı gerçekleştirmişlerdir. Çalışmada açık kaynak kodlu WEKA VM yazılımı ile toplam 10 adet algoritma kullanılmıştır. Bu algoritmalarından en iyi sonucu %91,1 ile Naive Bayes vermiştir. Naive Bayes algoritmasından sonra en iyi sonucu %90,8 ile Bagging algoritması vermiştir. En kötü sonucu ise %80 doğruluk oranı ile ZeroR algoritması

vermiştir. Bu araştırma sayesinde akciğer kanserinin teşhisine yönelik bir uygulamanın geliştirilmesi durumunda, hastalığın teşhis süresinin kısalmasının sağlanacağı, sağlık çalışanlarına fikir verebileceği ve hastalığı önceden teşhis ile insanların yaşam sürelerinin uzamasının mümkün olabileceği anlaşılmıştır.

Mohan ve arkadaşları [5] çalışmalarında, güvenli bir ağ oluşturulması için izinsiz girişlerin tespitinde VM sınıflandırma tekniklerini kullanmışlardır. NSL-KDD veri setinde ön işleme işlemini yaptıktan sonra 42 öznitelik ve toplamda 23866 veri kullanılmıştır. Bu veriler WEKA programında Rastgale Orman, Naive Bayes, Rastgele Ağaç ve J48 algoritmalarıyla işlenmiştir. En iyi sonucu %99,78 doğruluk oranına sahip olan Random Forest algoritması vermiştir. Araştırmacılar, günümüzde önemli bir yere sahip olan ağ saldırılarında, VM sınıflandırma tekniklerinden hangisinin daha faydalı olacağı konusunda literatüre katkı sağlamışlardır.

DrissiTouzani ve arkadaşları [6], Fas'taki trafik kazalarını VM yöntemlerini kullanarak analiz etmişlerdir. Fas'ta ilgili bakanlıktan 2013 ve 2014 yıllarına ait trafik kazası verilerini alınmıştır. Veri seti ön işleme tabi tutulduktan sonra kaza numarası, kaza tarihi, kaza saati, coğrafi bölgesi, karayolu durumu, yüzey durumu, ışık, sertlik, yığılma kodu, yerelleştirme, kavşak türü, çarpışma türü, engeller, ilk şok noktası, hareket (şerit veya hat değişikliği) özelliklerine sahip toplam 1952 adet veri elde edilmiştir. Düzenlenmiş olan veri seti WEKA programında k-ortalama algoritması kullanılarak analiz işlemi gerçekleştirilmiştir. Analiz işlemlerinin sonrasında kazaların çoğunun gündüz meydana geldiğini tespit edilmiştir. Trafik kazalarının azaltılmasına yardımcı olmak için ülkenin ilgili bakanlığına önerilerde bulunulmuştur.

Türkiye de rekabetin fazla olduğu tekstil sektöründe firmalar başarılarını daha üst seviyelere çıkarmak için, teknolojinin de gelişmesiyle birlikte verilerini VM yöntemleriyle analiz etmeye başlamıştır. VM tahminleri ve analizler, geleceğe yönelik piyasa stratejilerini oluşturmada yöneticilere fayda sağlamıştır. Tozak [7], yüksek lisans tezinde, tekstil sektöründe faaliyet gösteren bir işletmenin geçmişe dönük, gömlek ihracatını etkileyen önemli değişkenlerini ve gelecek satışlarını VM teknikleri ile analiz etmiştir. Bu analiz işlemlerinde tekstil sektöründeki bir işletmenin 43 aylık geçmiş satış verilerini kullanmıştır. Bu veriler ile müşterinin satış eğilimini anlayabilmek ve firmanın gelecek satışlarıyla ilgili bilgilere ulaşmak için VM yöntemini kullanmıştır. IBM SPSS, WEKA, RStudio ve Knime VM paket programların da Karar Ağaçları algoritmalarını kullanmış ve paket programların karşılaştırmasını yapmıştır. Bu karşılaştırma sonucunda

en iyi sonucu Knime VM paket programı vermiştir. Dört programın çıktısı incelendiğinde siparişin ait olduğu müşteri, siparişin üretildiği ay ve yıl, siparişin birim fiyatı ve çalışmadaki değişkenlerin sevk adetini etkileme sıralamasının üretildiği fabrika aylık erkek gömlek ihracat miktarı ve ihracat miktarını etkileyen bağımsız değişkenler tespit edilmiştir. Veri analizi sonucunda, sipariş birim fiyatının 8 Euro'dan daha az olduğu zaman siparişin fazla olduğu, müşterilerden gelen siparişlerin ilk üç ayın diğer aylara göre daha az olduğu, en çok D ülkesinde üretim yapıldığı ve en çok sipariş veren müşterilerin tespiti ve bu müşterilere öncelik tanınması gerektiği sonucuna ulaşılmıştır.

Teknolojinin gelişmesi devam ettikçe birçok alanda değişiklikler meydana gelmektedir. Bunlardan biri de Bitcoin kripto para biriminin geliştirilmesiyle dünya ekonomisinde yaşanmaktadır. Topal [8], yüksek lisans tezinde, VM tekniklerini kullanarak bitcoin piyasası hakkında çalışma yapmıştır. 22.03.2017 ile 22.03.2018 tarihleri arasındaki bitcoin verileri bu çalışmada kullanılmıştır. Araştırmada RapidMiner programında YSA, Destek Vektör Makinesi ve Lineer Regresyon algoritmaları kullanılmıştır. Bitcoin verilerinde 0,972 en iyi doğruluk oranıyla en başarılı algoritma Destek Vektör Makinesi yöntemi olmuştur. Araştırmacı bitcoin piyasasını makine öğrenmesi yöntemleri ile modelleyerek, en uygun yöntemin belirlenmesi konusunda literatüre katkı sağlamıştır.

Bakırarar [9], yüksek lisans tezinde, tıp alanındaki hepatit ve meme kanseri veri setlerini kullanarak hibrit VM yaklaşımını önermiştir. Hibrit VM yöntemiyle, algoritmaları birlikte kullanarak zayıflıkların telafi etmeyi hedeflemiştir. Bunun için çalışmasında Karar Ağacı, Destek Vektör Makinesi, Rastgele Orman, Naive Bayes, k-NN, Lojistik Regresyon, k-ortalama ve Yapay Sinir Ağları algoritmalarını kullanmıştır. Araştırmacı kullanmış olduğu algoritmaları birleştirerek hibrit modeller de oluşturmuştur. Hepatit veri setinde ilk başta gerçek değerleri kullanarak verileri değerlendirmiş ve sonrasında verilerin tüm özelliklerini koruyarak örneklem büyüklüğünü 250'den 500'e yükseltmiş. Sonrasında veri setinde dengesiz dağılım olduğu için kümeleme yöntemini kullanarak verileri kümelerine ayırmıştır. Kümelerin performansını ayrı ayrı değerlendirmiş ve bu kümelerin performansını birleştirerek genel veri seti performansına ulaşmıştır. Tüm veri setine ait sonuçları değerlendirdiğinde ise örneklemin büyüklüğü arttıkça performansının da iyileştiğini gözlemlemiştir. Bu çalışmanın aynısını meme kanseri veri setinde de yapmış ve örneklemin büyüklüğü arttıkça performansın iyileştiğini bu veri setinde de gözlemlemiştir. Araştırmacı yaptığı tüm senaryolarda ve gerçek veri setlerinde hibrit modelin daha iyi sonuç verdiğini ortaya koymuştur. Yapılacak diğer çalışmalarda, uygun

VM yöntemlerini birlikte kullanarak, hibrit model yaklaşımının daha iyi modeller elde edilmesini sağlayacağı sonucuna ulaşmıştır.

Şenocak [10], yüksek lisans tezinde, bankacılık alanındaki teknolojik değişimlerle birlikte banka içi denetimlerde uygulanabilecek VM algoritmalar hakkında çalışmıştır. Araştırmasında bireysel ve ticari olmak üzere iki adet veri seti üzerinde çalışmıştır. Bu veri setlerinde Lojistik Regresyon, Naive Bayes, k-NN, Karar Ağacı, Rastgele Orman ve Destek Vektör Makinesi algoritmalarını kullanmıştır. Araştırmanın sonucunda, bireysel müşterilerin ve şirketlere kullanılan kredilerin temerrüde düşme olasılığının hesaplanmasında algoritmalarından oluşturmuş olduğu modellerin faydalı olduğunu tespit etmiştir. Bireysel kredi sürecinde en iyi sonucu Lojistik Regresyon algoritması verirken, ticari kredi sürecinde en iyi sonucu Destek Vektör Makinesi algoritması vermiştir. Literatür ile yaptığı karşılaştırma sonucunda ise, veri seti özelliklerine göre sonuçlarda değişiklik olabileceğinden dolayı farklı algoritmalarla çalışılmasının faydalı olacağını vurgulamıştır. Çalışmada bankacılık sektöründe, banka denetim süreçlerine yapay zekâ teknolojilerinin entegre edilmesi konusunda da tavsiye verilmiştir.

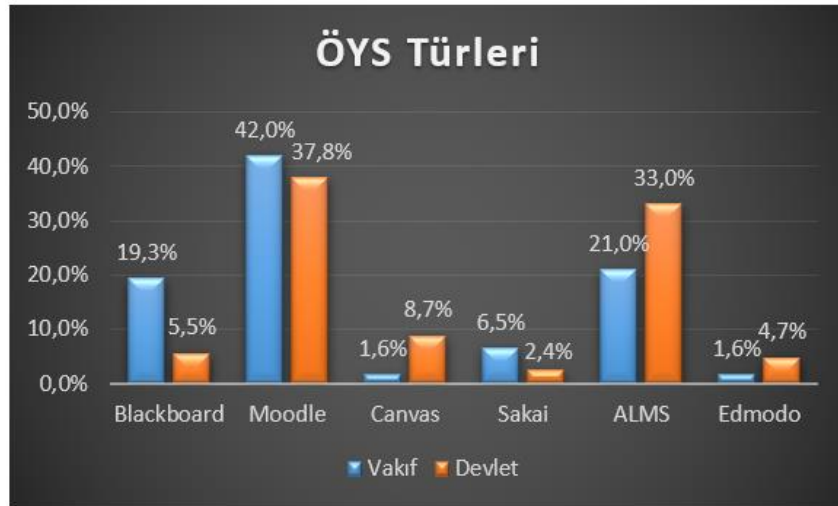
VM yöntemleri sıklıkla perakende sektöründe de kullanılmaktadır. Bu sektördeki kullanım amacı, müşteri kaybının önlenmesi için satış tahminleri yapılarak satışın artırılmasının hedeflenmesidir. Dinçoğlu [11], yüksek lisans tezinde, Migros Ticaret A.Ş.'nin iki farklı mağazasının 2018 ile 2019 satış bilgilerinden oluşan bir veri seti üzerinde VM tekniklerini kullanarak çalışmıştır. Çalışma içerisinde satış trendinin yakalanması için, oluşan enflasyon artışları ve artan fiyatların daha gerçeğe yakın sonuçlar alınabilmesi için, 2019 verilerini kullanarak sonuçları karşılaştırmıştır. Çalışmada IBM SPSS programı kullanılmıştır. Program içerisinde Regresyon ve Zaman Serileri ile satış tahminleri gerçekleştirilmiştir. Kullanılan iki model içerisinde çalışmada kullanılan verilere en uygun modelin Regresyon olduğu tespit edilmiştir.

Aydur [12], yüksek lisans tezinde, havacılık, uzay ve otomotiv sektörleri başta olmak üzere sanayi alanında kullanılan A17075 alüminyum alaşımını, tel elektro erozyon işleme yöntemiyle farklı parametrelerle işlemiş ve oluşan yüzey pürüzlüklerini VM ile tahmin etmiştir. Çalışmada, darbe süresi, voltaj, tel besleme ve basınç değerleri giriş parametreleri olarak alınmıştır. Toplamda 81 farklı yüzey pürüzlülüğü verisini Aşırı Öğrenme Makineleri, West Extreme Learning Machine, Destek Vektör Makinesi ve Qiskit Destek Vektör Regresyonu algoritmalarını kullanarak tahminde bulunmuştur. Tahminleme işleminde en iyi doğruluk oranını West Extreme Learning Machine

algoritması vermiştir. Alüminyumun hassas işlenen yüzeylerin yüzey pürüzlülüğünün belirlenmesi zaman alıcı ve maliyetli olduğunda, bu çalışma yüzey pürüzlülüğünü önceden tahmin ederek zaman, iş gücü ve maliyet konusunda yüksek kazanımların elde edilebileceği bu çalışma ile öngörülmüştür.

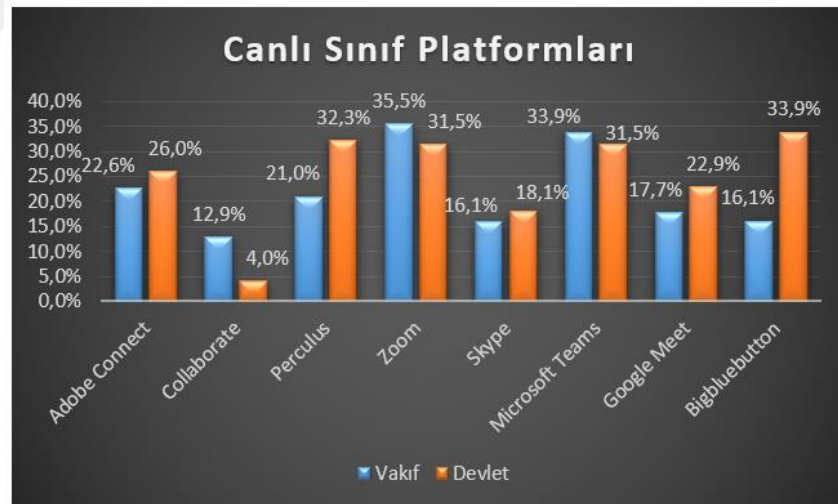
Yukarıdaki çalışmalardan anlaşılacağı üzere VM algoritmaları ile tahminleme işlemleri disiplin farkı olmadan birçok alanda hem bilime katkı sağlamış hem de toplumsal sorunlara çözüm üretmiştir. Bu alanlara bir yenisini daha ekleme fikri ise tüm dünyayı etkisi altına alan Covid-19 salgını ile ortaya çıkmıştır.

2019 yılının sonlarında Çin'in Vuhan Eyaleti'nde Covid-19 hastalığı tespit edilmiş ve dünyadaki diğer ülkelere de bu virüs yayılmıştır [13]. 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 [14]. Vakaların artışıyla birlikte YÖK 16 Mart 2020 tarihinde eğitime 3 hafta ara verileceğini duyurmuştur [15]. 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 [16]. Pandemi koşullarının uzayacağı öngörüldükten sonra YÖK tarafından üniversitelerde 2019 – 2020 bahar yarıyılında eğitim ve öğretimin tamamen uzaktan yapılmasına karar verilmiştir [17]. 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 UZEM sayısı 120'den 140'a çıkmıştır [18]. Bu merkezlerin kullandığı platformlar çeşitlilik göstermiştir. YÖK'ün üniversitelerimizde küresel salgın sürecinde uzaktan eğitim uygulamaları anket sonuçlarına göre, Şekil 1.1'de gösterildiği üzere çoğunlukla kullanılan ÖYS türleri Blackbord, Moodle, Canvas, Sakai, ALMS ve Edmodo'dur [19]. Üniversiteler de en çok kullanılan ÖYS türü Şekil 1.1'de gözüktüğü üzere Moodle olmuştur. Bu çalışmada kullanılan veri seti (Asenkron dersler için) ALMS ÖYS Türü üzerinden temin edilmiştir.



Şekil 1.1. Pandemi döneminde üniversitelerde kullanılan ÖYS türleri.

Aynı anket sonuçlarına göre canlı derslerde (Senkron) çoğunlukla kullanılan Şekil 1.2’de gösterilen canlı sınıf platformları ise Adobe Connect, Collaborate, Perculus, Zoom, Skype, Microsoft Teams, Google Meet ve Bigbluebutton’dır [19]. Üniversiteler de en çok kullanılan canlı ders platformu Şekil 1.2’de görüldüğü üzere Zoom’dur (devlet ve vakıf üniversitelerinin ortalaması alınarak bu sonuca varılmıştır). Bu çalışmada kullanılan veri seti (Senkron dersler için) Perculus canlı sınıf platformu üzerinden temin edilmiştir.



Şekil 1.2. 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 [20]. 2021 – 2022 eğitim ve öğretim yılından itibaren üniversitelerde uzaktan eğitimin pandemi koşullarına göre hibrit olarak devam edilebileceği 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 [21].

Tez çalışmasının amacı, pandemi döneminde KAYUZEM öğrencilerinin, fakülte/meslek yüksek okulu, dersleri, derse katılma yöntemleri (Senkron, Asenkron ve Hibrit) ve başarı durumu kriterleri üzerinden VM yöntemleri uygulanarak, gelecek dönemde öğrencilerin başarı durumlarını tahmin etmektir. Bu tahminleri kullanarak da derse katılma yöntemlerinden hangisinin uzaktan eğitim sürecinde daha güvenilir olduğunu tespit etmektir. Bu amaç doğrultusunda yapılacak olan araştırmanın önemi, aşağıda literatürden çeşitli örneklerle açıklanmaya çalışılmıştır.

Üniversite öğrencilerinin belirli bir derste başarı oranlarının öğrenme yönetim sistemleri hareketliliği ölçülerek başarı düzeyleri arasındaki ilişkinin ortaya konulduğu Özbay ve Ersoy [22] 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.

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 [23], çalışmasında 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 [24]

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.

Ülkemizde çalışmamızla ilgili konular ele alınmış olsa da bu araştırmaya en yakın güncel çalışmalar genellikle yurt dışında yapılmaktadır. Örneğin, Blagojević ve Micić [25], 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 YSA 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.

Villegas-Ch ve arkadaşları [26] 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.

Abe [27], 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.

Yaacob ve arkadaşları [28] 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 VM 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, Karar Ağacı 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.

Byers ve arkadaşları [29] 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.

Ati Jain ve arkadaşları [30] 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.

Hussain [31] 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, Native Bayes, 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 öznelikleri 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 için de aynı yol izlenmiştir.

Chen Tan ve arkadaşları [32] 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 olduğunu 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 [33], 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, Karar Ağacı, 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 [34], çalışmalarında Winneba Üniversitesinin Bilgi ve İletişim Teknolojisi Bölüm öğrencilerinin pandemi döneminde sınav puanlarını etkileyecek olan öznelıklar 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 Rastgele Orman, Rastgele Ağaç, Naive Bayes ve J48 Karar Ağacı algoritmalarını kullanarak tahminleme işlemi yapmışlar. Bu tahminleme işleminde her algoritmanın karışıklık matrisini de hesaplamışlardır. Aynı zamanda algoritmalarda en az performansı gösteren öz nitelikleri 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 Rastgele Orman algoritmaları vermiştir. Bu çalışma bizim çalışmamız gibi öğrencinin performansını arttırmaya yöneliktir. Bizim çalışmamızdan farkı öğrencilerin öz nitelik verilerini kullanarak öğrenci performansını incelemiş olmasıdır. Biz çalışmamızda öğrencilerin performansını derse katılma yöntemlerine göre incelemekteyiz.

Alsammak ve arkadaşları [35] üç tür anket kullanarak 1120 örnek içeren veri toplamış, verileri ön işleme işlemine tabi tutmuşlardır. Veri seti 35 öz nitelikli ve 1000 örnek olarak düzenlemiştir. Düzenlenen bu veri setinde e-öğrenmede öğrenci performansı üzerinden VM algoritmaları WEKA programı yardımıyla uygulamıştır. Uyguladıkları algoritmalar Karar Ağacı, Rastgele Ağaç, Naive Bayes, Rastgele Orman, REP Tree, Torbalama (Bagging) ve k-NN'dir. 35 öz nitelikli yapılan tahminleme sonucunda en iyi sonucu %96,8'le k-NN algoritması vermiştir. Araştırmacılar öz nitelikleri Dereceli Arama (Ranker Search) Metodunu kullanarak en iyi öz nitelikleri seçmiş ve bu öz niteliklere 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. Öz nitelikler 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 [36, 37]. 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 evirildiği dijital dünyada öğrenci başarılarının, öğrenciyi fiziksel olarak gözlemlemek üzerinden yapılmayacağı ya da bunun yeterli olmayacağı yenilikçi gelecek tahminleri arasında yer almaktadır. Bu araştırmanın eğitimin gelecekteki tahayyül somut olarak göstermesi açısından literatüre önemli katkılar sağlayacağı düşünülmektedir. Bu çalışmayı literatürdeki benzer çalışmalardan ayıran en önemli etkenlerden 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.

2. VERİ MADENCİLİĞİ

VM Şekil 2.1’de görüldüğü gibi 1763 yıllarında Bayes teoremi ile ortaya çıkmış ve bilgisayarın icadıyla birlikte hızlı bir gelişme sürecine girmiştir. Bilgisayar teknolojisinin gelişmesiyle, verilerin toplanması, depolanması ve işlenmesinde bir artış meydana gelmiştir. Bu durumun sonucunda, veri toplamanın boyutu ve karmaşıklığı da büyümüştür.



Şekil 2.1. VM tarihi [38].

VM, anlamlı örüntüleri keşfetmek için büyük miktarda verinin araştırılması ve analizidir [39].

VM, büyük veri tabanlarındaki verilerden ilginç, önemsenmeyen, örtük, önceden bilinmeyen ve potansiyel olarak yararlı bilgi veya kalıpları çıkarmaktır [40].

VM, büyük verilerden faydalı bilgiler çıkarabilen derin bir veri analiz yöntemidir [41].

VM, her veriyi analiz eden ve çok sayıda veriden kurallar bulan bir teknolojidir [42].

Rahman M. ve arkadaşları, VM'yi, büyük bir veri kümesinden gizli bilgi veya kalıpları çıkarma prosedürü olarak tanımlamaktadır [43].

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

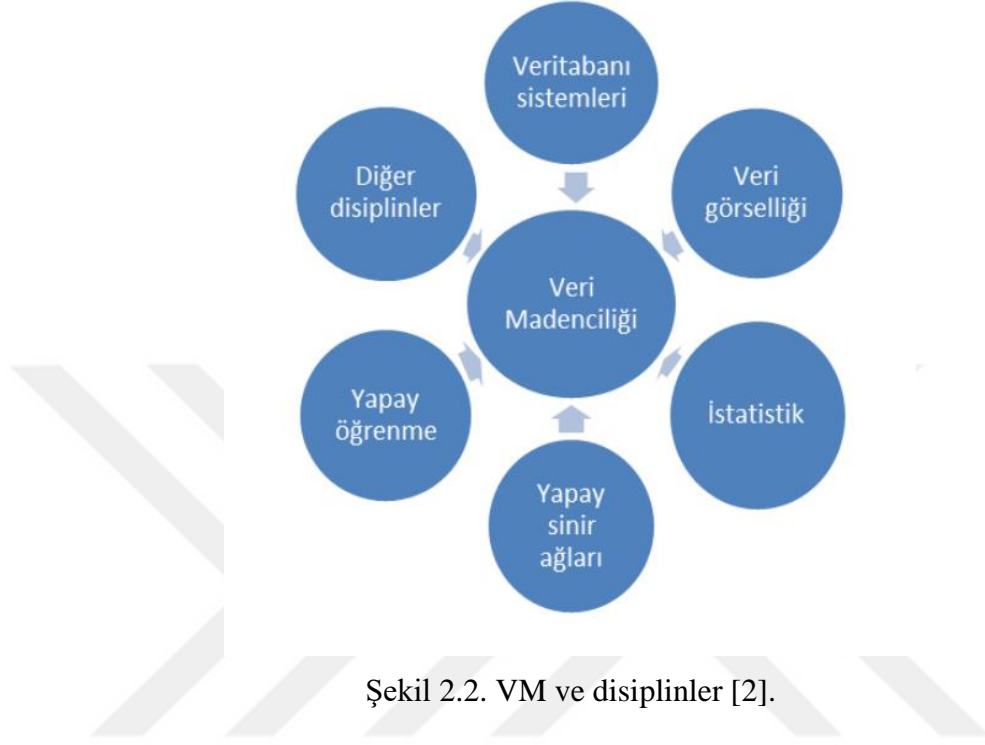
VM birçok uygulama alanında kullanılan, işlenmemiş veriden örtük bilgiye ulaşmayı amaçlayan disiplinler arası bir süreçtir.

VM kullanım alanları:

- Eğitim
- Sigortacılık
- Pazar analizi
- Belgeler arası benzerlik
- Sağlık
- Alışveriş
- Bankacılık ve finans
- Savunma sanayi
- Ticaret
- Müşteri ilişkileri yönetimi
- İnsan kaynakları yönetimi
- Bilgi sistemleri yönetimi
- Mühendislik
- Risk analizi
- Güvenlik ve istihbarat
- Telekomünikasyon

- E-Ticaret

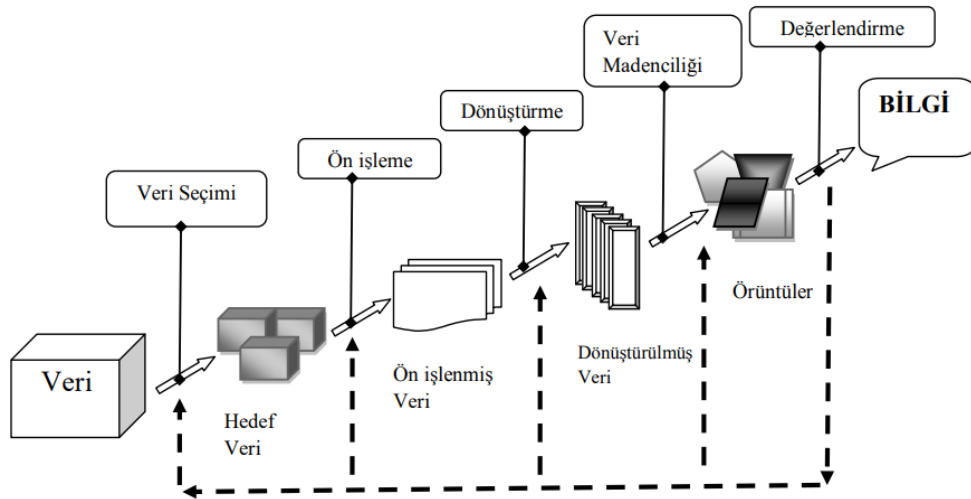
Şekil 2.2’de görüldüğü üzere VM ilişkili olduğu konu başlıkları oldukça çeşitli alanları bir araya getirmektedir.



Şekil 2.2. VM ve disiplinler [2].

2.1. BİLGİ KEŞFİ SÜRECİ

Bilgi keşfi süreci, verilerdeki faydalı ve anlaşır kalıpları çıkarma sürecidir [44]. VM'nin önemli bir adımı olan bilgi keşfi Şekil 2.3’de gösterilmiştir.



Şekil 2.3. Bilgi keşfi sürecinin adımları [2].

Veri Seçimi: Veri tabanından çalışmada kullanılacak olan verinin seçilmesidir.

Veri Ön İşleme: Çalışma için seçilen veriler üzerinde veri bütünleştirme, veri temizleme ve veri indirgeme işlemleri yapılır.

Veri Dönüştürme: Veri ön işlemede işlenmiş verileri, kullanacağımız VM yöntemlerine göre sayısal veya sözel değerlere dönüştürme işlemidir.

Veri Madenciliği Yöntemini Uygulama: Hazırlanmış verilerde VM yöntemlerini kullandığımız bölümdür.

Değerlendirme: VM yöntemlerinden çıkan sonuçları analiz ederek grafiklerle desteklediğimiz alandır.

2.1.1. Veri Ön İşleme

Verilerin kalitesini ve verimliliğini arttırmak için yapılır [45].

2.1.1.1. Veri Temizleme

Veri seti üzerinde yanlış türleri ve aykırı değerler içeren verilere gürültü denir. Veri setinde gürültü söz konusuysa, bu verilerin temizlenmesi gerekmektedir. Bu temizleme işleminde aşağıdaki yöntemler kullanılabilir [46].

- Veri setindeki, eksik veri içeren bölümler silinebilir. Silinme işleminin fazla olması durumunda, veriler eksileceği için çalışmayı olumsuz yönde etkileyebilir.
- Veri setindeki, eksik veri içeren bölümler tamamlanabilir. Bu bölümleri objektif bir şekilde tanımlamak gerekmektedir. Objektif bir şekilde tanımlama yapılmazsa çalışmanın sonucunu olumlu veya olumsuz yönde etkilebilir.
- Veri setindeki, eksik veri içeren bölümlere sabit bir değer girilebilir. Bu duruma en iyi örnek, sayısal olarak “0” değeri, sözel olarak da “bilinmiyor” değeri girilmesidir.
- Veri setindeki diğer veriler VM yöntemleri kullanarak eksik veri içeren bölümler tahminleme yöntemiyle tamamlanabilir.
- Veri setindeki diğer verilerin ortalama değerleri hesaplanarak eksik veri içeren bölümler tamamlanabilir.

2.1.1.2. Veri Bütünleştirme

Bu süreçte birden fazla farklı veri tabanından alınmış veri setlerindeki aynı bölümlerin

içerisindeki farklı türlerdeki verilerin aynı türe dönüştürülmesi gerekir. Örneğin bir öğrencinin başarı durumu bölümü bir veri tabanının da “Başarılı” ve “Başarısız” olarak yazılmış, diğer veri tabanının da “Geçti” ve “Kaldı” olarak yazılmış, başka bir veri tabanının da ise “1” ve “0” olarak yazılmıştır. Böyle bir durumda belirtilen bölümün tek bir veri setinde birleştirdiğimizde veri türünün aynı olması gerekmektedir. Eğer bu bölümde veri bütünleştirme işlemi yapılmazsa, yapılacak olan analizlerde başarılı bir sonuç elde edilemez.

2.1.1.3. *Veri İndirgeme*

Bu süreçte veri setindeki analizi etkilemeyecek veya olumlu yönde etkileyecek olan verinin azaltma işlemi yapılabilir. Burada dikkat edilecek husus, azaltılacak olan verinin çalışmayı etkilememesi veya olumlu yönde bir etki vermesidir.

2.1.2. Veri Dönüştürme

Veri setindeki sayısal veya sözel değerler VM yöntemlerinde direk kullanılamayabilir. Kullanacağımız VM yöntemi sayısal verilerde daha iyi sonuçlar veriyorsa, veri setindeki sözel verileri sayısal verilere dönüştürmemiz gerekmektedir. Bu yaptığımız işlemin tam tersi de olabilir. Bu işlemleri bizim için otomatik olarak yapan VM paket programları da mevcuttur.

2.1.3. Veri Madenciliği Yöntemini Uygulama

Veri seçimi, veri ön işleme ve veri dönüştürme süreçleri gerçekleştirildikten sonra, veri seti VM yöntemleri için hazır hale getirilir. Bu bölümde de çalışmada kullanılacak olan VM yöntemleri veri setimizde uygulanır.

2.1.4. Değerlendirme

Veri setimizin üzerinde VM yöntemleri uygulandıktan sonra çıkan sonuçların değerlendirilmesi bu aşamada yapılır. Değerlendirilen sonuçlar aynı zamanda grafiklerle de desteklenebilir.

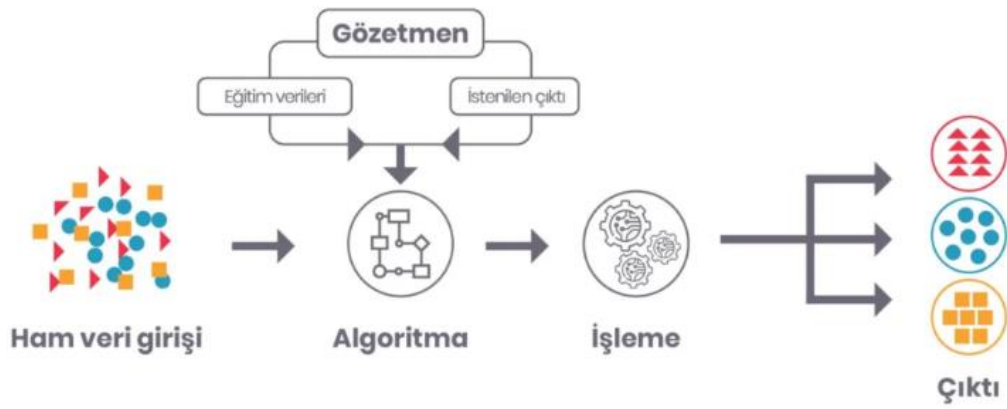
2.2. VERİ MADENCİLİĞİ YÖNTEMLERİ

2.2.1. Denetimli Öğrenme (Supervised Learning)

Denetimli öğrenme, etiketlenmemiş verilerin sınıfını veya yanıt değerini tahmin etmek

için birçok etiketli eğitim verisinden tahmine dayalı bir model öğrenen tipik bir makine öğrenimi yöntemidir [47].

Bir modelin girdileri ve bir dizi çıktı arasındaki bağımlılıkları ve ilişkileri öğrenmesine denetimli öğrenme yardımcı olur. Modelin daha önce görülmemiş özellikleri veya girdileri için hedefi tahmin etmesine yardımcı olur. Denetimli öğrenmeyi iki teknik veya algoritma olarak kategorize edebiliriz. Bunlar sınıflandırma ve regresyon olarak bilinir [48].



Şekil 2.4. Denetimli öğrenme [49].

Denetimli öğrenmede en sık kullanılan algoritmalar;

- K-En Yakın Komşu (K-Nearest Neighbor)
- Karar Ağaçları (Decision Trees)
- Destek Vektör Makineleri (Support Vector Machines)
- Rastgele Orman (Random Forest)
- Naive Bayes
- Yapay Sinir Ağları (Artificial Neural Networks)
- Lojistik Regresyon (Logistic Regression)
- Doğrusal Regresyon (Linear Regression)
- Polinomal Regresyon (Polynomial Regression)
- Çoklu Doğrusal Regresyon (Multiple Linear Regression)

2.2.1.1. Sınıflandırma Algoritması

Sınıflandırma, bir veri setindeki verileri hedef kategorilere veya sınıflara atayan bir VM yöntemidir. Sınıflandırmanın amacı, verilerdeki her bir durum için hedef sınıfı doğru bir şekilde tahmin etmektir [50].

2.2.1.2. Regresyon Algoritması

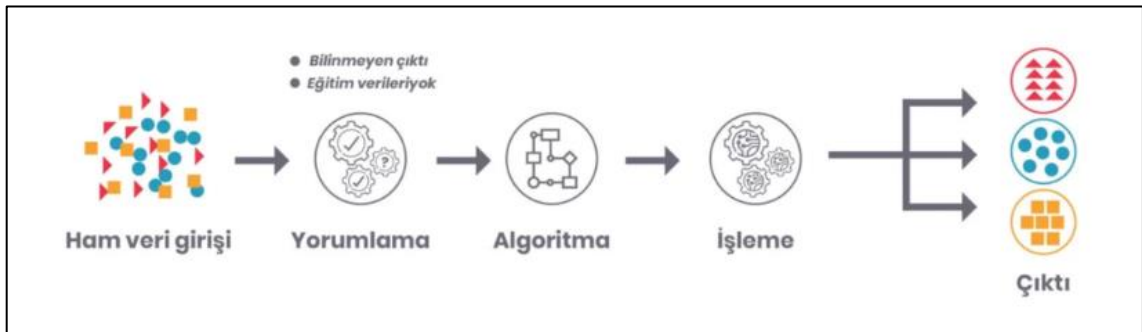
Regresyon, bir bağımlı değişken ile bağımsız değişkenler arasındaki ilişkiyi tahminlemeye çalışan yöntemdir. Sınıflandırma algoritmasından farkı, süreklilik gösteren değerlerin tahmininde kullanılmasıdır [51].

2.2.2. Denetimsiz Öğrenme (Unsupervised Learning)

Denetimsiz öğrenme, herhangi bir ek bilgi olmadan bir veri setinin yapısı hakkında yararlı özellikleri öğrenmeye çalışır [52].

Denetimsiz öğrenme, verileri açıklamaya veya analiz etmeye ya da daha sonraki uygulamalar için faydalı girdiler oluşturmaya çalışır [53].

Denetimli öğrenmenin aksine, örnekler etiketsiz olarak sağlanır ve asıl amaç, örnekler arasındaki ilişkiyi veya verilerin yapısını incelemektir [54]. Denetimsiz öğrenmenin iki önemli konusu, kümeleme ve birliktelik kurallarıdır.



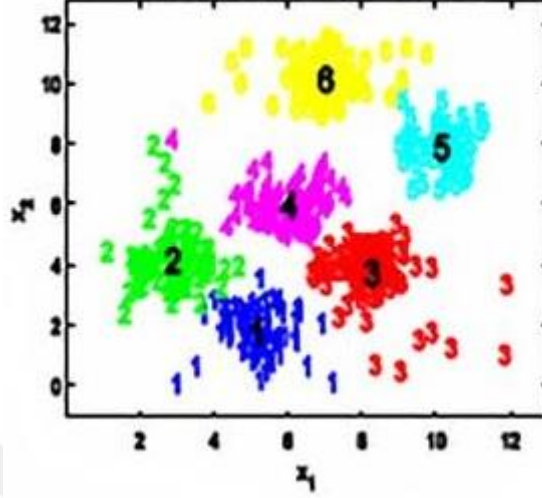
Şekil 2.5. Denetimsiz öğrenme [49].

Denetimsiz öğrenmede en sık kullanılan algoritmalar;

- Hiyerarşik Kümeleme (Hierarchical Clustering)
- K-Ortalama (K-Means)
- Apriori Algoritması
- Gaussian Mixture Modeli
- Gizli Markov Modeli (Hidden Markov Models)

2.2.2.1. Kümeleme

Kümeleme, verilerin benzerlikleri, uzaklıkları veya yakınlıkları gibi kriterlere göre çözümlenerek sınıflara ayrılmasıdır.



Şekil 2.6. Kümeleme örneği [55].

2.2.2.2. Birliktelik Kuralı

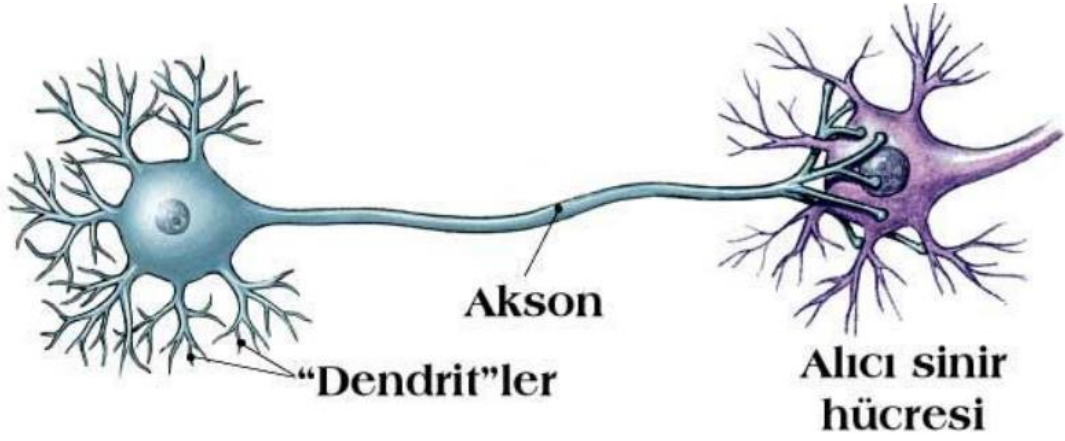
Birliktelik kuralı alışverişlerde aslında çok sık karşımıza çıkmaktadır. E-ticaret web siteleri alışveriş verilerini birliktelik kuralına göre analizler ederler. Bir müşteri alışveriş sitesinden bir ürünü sepetine eklediğinde, aldığı ürünle ilgili veya ilgisiz bir ürünü sistem otomatik olarak “bu ürünü de sepetinize eklemek istemisiniz” diye sorar. İşte bu ürünleri önerme işlemini daha önceki verilere dayanarak birliktelik kuralı sağlamıştır.

2.3. VERİ MADENCİLİĞİ İÇİN KULLANILAN YÖNTEMLER

2.3.1. Derin Öğrenme

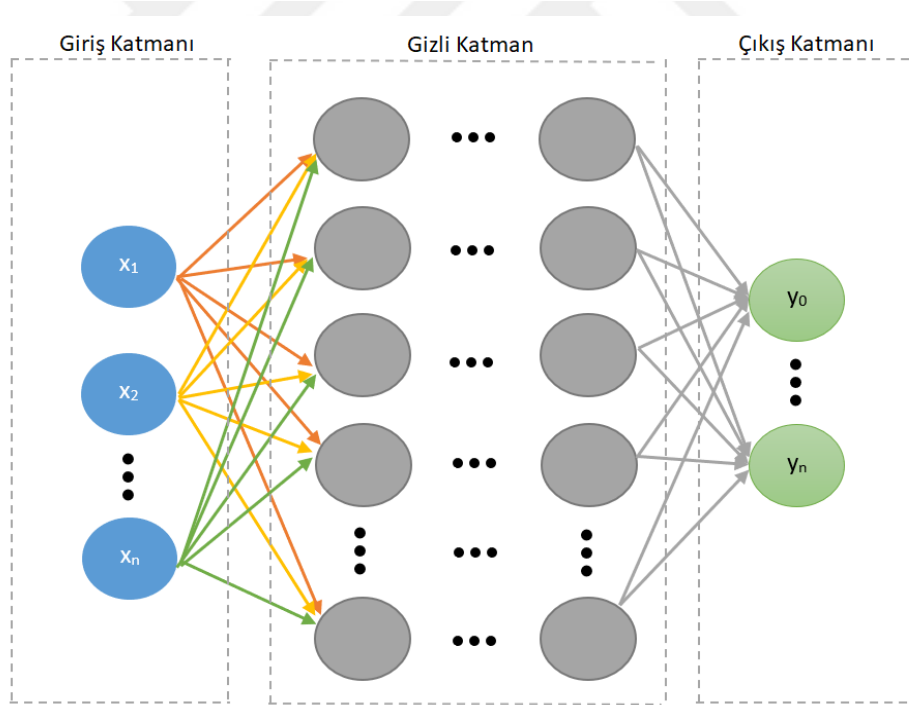
Derin öğrenme, verilen bir veri kümesi ile çıktıları tahmin edecek yapay zekayı eğitmeye olanak sağlar. Yapay zekayı eğitmek için hem denetimli hem de denetimsiz öğrenme kullanılabilir [56]. Temelde çalışma şekli YSA aynıdır. McCulloch ve Pitts 1943 yılındaki ilk denemelerinden insan beyninin sinir ağlarını taklit etmek için matematiksel model oluşturarak, sinir ağlarını birkaç evrimsel aşamaya ilerletmişler [57]. Sonrasında Rosenblatt ve arkadaşları yapay nöronlar için temel olarak kullanılan bir algılayıcı kavramını tanıtmışlardır. Sinir ağları, bir dizi dağıtılmış ve birbirine bağlı hesaplama birimleri veya nöronlar kullanarak bilgiyi iletirken ve işlerken bir beyindeki sinir

hücresinin çalışmasını örnek almıştır [58].



Şekil 2.7. Sınıf hücresi yapısı [59].

Derin öğrenme sınıflandırıcısı Şekil 2.8’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 [60].



Şekil 2.8. Derin öğrenme şeması.

2.3.2. k-NN

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 [61]. Burada k noktasının yakınlığı, denklem 2.1'deki Öklid uzaklık fonksiyonu ile hesaplanır [62].

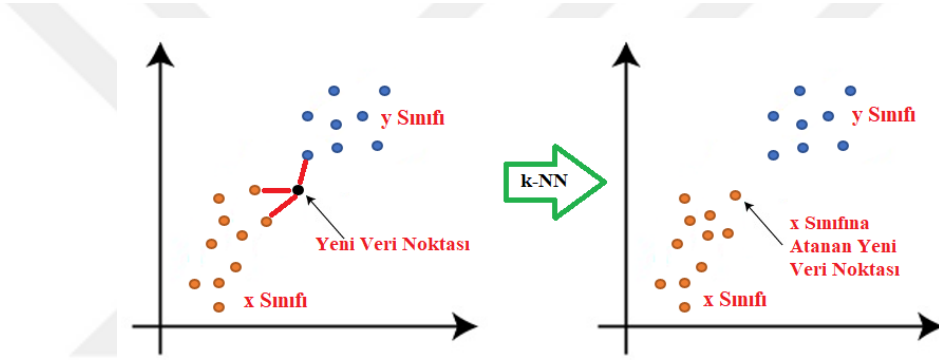
$$dist(P, Q) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2.1)$$

$P = (x_1, x_2, \dots, x_n)$

$Q = (y_1, y_2, \dots, y_n)$

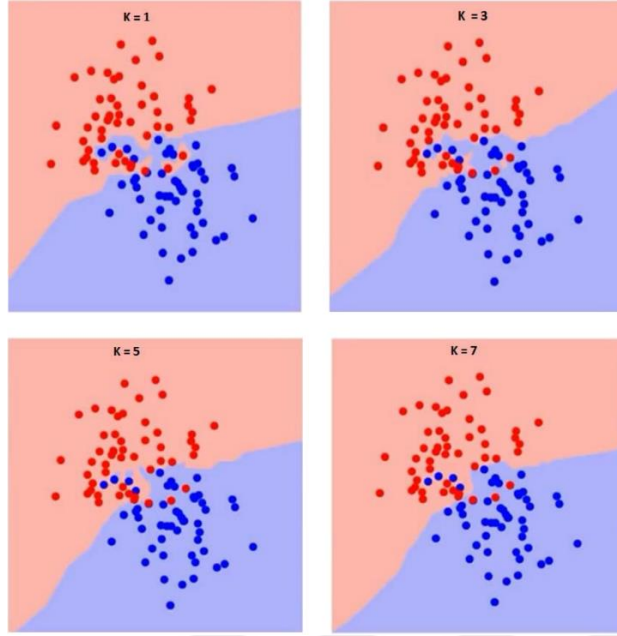
x_n = Verilerin bir değeridir.

y_n = veri tabanının bir değeridir.



Şekil 2.9. k-NN örneği.

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 [63]. k-NN algoritmasında farklı k değerlerine göre iki sınıfı ayıran sınırlar Şekil 2.10'da gösterilmiştir [64].



Şekil 2.10. Farklı k değerine sahip iki sınıfı ayıran farklı sınırlar [64].

2.3.3. Naive Bayes

Naive Bayes, İngiliz matematikçi Thomas Bayes tarafından 1740'larda ortaya atıldığı Bayes teoremine dayanmaktadır [65]. Naive Bayes algoritması, belirli bir veri kümesindeki değerlerin frekansını ve kombinasyonlarını sayarak bir olasılık kümesini hesaplayan basit bir olasılık sınıflandırıcısıdır [66]. Naive Bayes sınıflandırıcısının temeli Bayes teoremine dayanır.

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} \quad (2.2)$$

$P(A | B)$ = B olayı gerçekleştiğinde A olayının gerçekleşme olasılığı

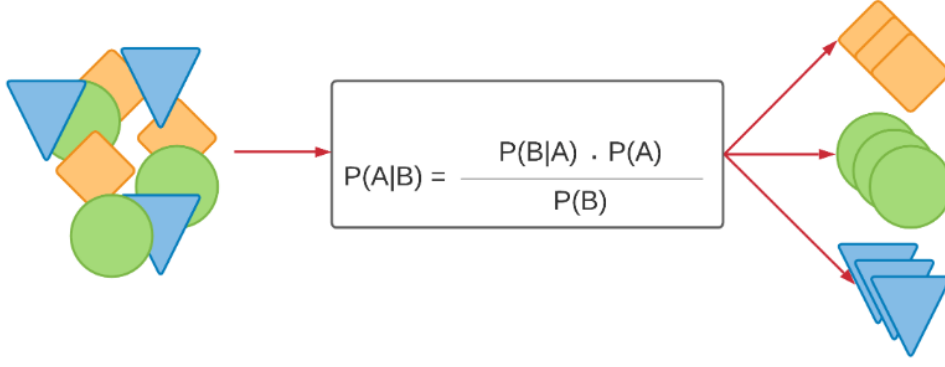
$P(B | A)$ = A olayı gerçekleştiğinde B olayının gerçekleşme olasılığı

$P(A)$ = A olayının gerçekleşme olasılığı

$P(B)$ = B olayının gerçekleşme olasılığı

Naive Bayes algoritması basit bir algoritma olmasına rağmen genellikle karmaşık sınıflandırma yöntemlerinde iyi performans gösterebilir [64].

Naive Bayes algoritmasında yapılan işlemi basit bir örneği Şekil 2.11'de gösterilmektedir.

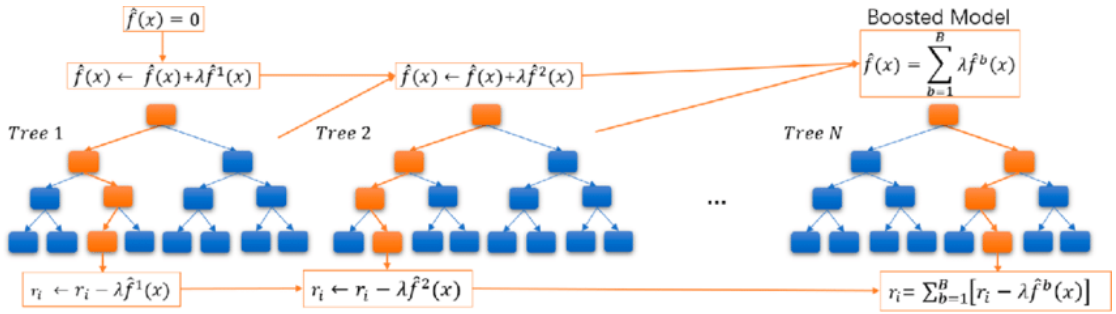


Şekil 2.11. Naive Bayes örneği.

2.3.4. Gradient Boosted Trees

GBT, 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 [67].

GBT'nin optimize edilecek öğrenme hızı, yineleme sayısı, bireysel tahmin edicilerin maksimum derinliği ve bir düğümü bölmek için minimum örnek sayısı dahil olmak üzere birkaç parametresi vardır [68].



Şekil 2.12. Gradient Boosted Trees örneği [67].

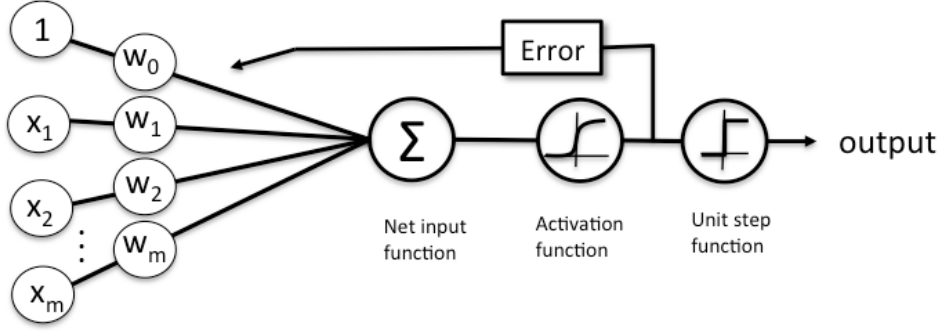
2.3.5. Lojistik Regresyon

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 [69].

Lojistik Regresyon, özellikle sınıflandırma çalışmalarında güçlü bir araçtır ve aynı anda birden fazla açıklayıcı değişkenin analiz edilmesine olanak tanır. Aynı zamanda kafa

karıştırıcı faktörlerin etkisini de azaltır [70].

Lojistik regresyonun Şekil 2.13’de şeması görülmektedir.



Şekil 2.13. Lojistik regresyon şeması [71].

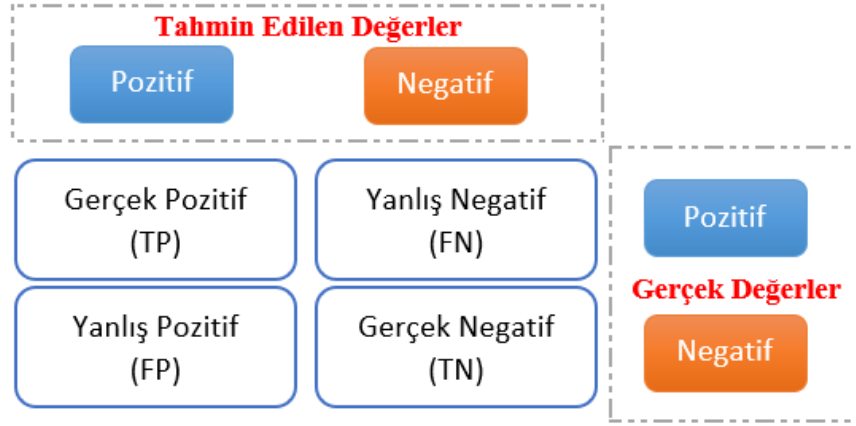
Şekil 2.14’de lojistik regresyonun modeli görülmektedir.

A diagram showing the logistic regression model equation $y_i = \beta_0 + \beta_1 \times x_i + hata(e_i)$. The equation is centered. To the left of the equation is the label 'Bağımlı Değişken' (Dependent Variable) with an arrow pointing to the y_i term. To the right of the equation is the label 'Bağımsız Değişken' (Independent Variable) with an arrow pointing to the x_i term. Above the equation, there are two labels: 'Kesme Terimi (intercept)' with an arrow pointing to β_0 , and 'Katsayı (coef., slope)' with an arrow pointing to β_1 .

Şekil 2.14. Lojistik regresyon modeli [72].

2.4. Karışıklık Matrisi

Karışıklık Matrisi, modellerin tahmin performanslarının belirlenmesi için kullanılır [73]. Diğer bir deyişle tahmin edilen değerleri ve gerçek değerler üzerinde algoritmanın performansını belirlemek için kullanılır. Şekil 2.15’de ikili sınıflandırıcı karışıklık matrisinin bir örneği bulunmaktadır.



Şekil 2.15. İkili sınıflandırıcı karışıklık matrisi.

Karışıklık matrisine ait performans karşılaştırma metrikleri [74]:

Gerçek Pozitif (TP): Pozitif hedefin doğru tahmini

Gerçek Negatif (TN): Negatif hedefin doğru tahmini

Yanlış Pozitif (FP): Negatif hedefin yanlış tahmini

Yanlış Negatif (FN): Pozitif hedefin yanlış tahmini

Doğruluk Oranı (Accuracy Rate): Genel olarak ne kadar doğru tahmin ettiğinin oranını verir.

$$\frac{TP + TN}{TP + TN + FP + FN} \quad (2.3)$$

Yanlış Sınıflandırma Oranı (Misclassification Rate): Genel olarak hata oranını verir.

$$\frac{FP + FN}{TP + TN + FP + FN} \quad (2.4)$$

Gerçek Pozitif Değerlerin Oranı (True Positive Rate veya Recall): Gerçek pozitif değerlerini ne kadar doğru tahmin ettiğinin oranını verir.

$$\frac{TP}{TP + FN} \quad (2.5)$$

Gerçek Negatif Değerlerin Oranı (True Negative Rate veya Specificity): Gerçek negatif değerlerini ne kadar doğru tahmin ettiğinin oranını verir.

$$\frac{TN}{TN + FP} \quad (2.6)$$

Kesinlik Oranı (Precision): Doğru tahminlerin ne kadar kesin olduğu oranı verir.

$$\frac{TP}{TP + FP} \quad (2.7)$$

F Skoru (F Score): Dengesiz sınıflar için yararlı hibrit skorunu verir.

$$\frac{2TP}{2TP + FP + FN} \quad (2.8)$$

3. METOD VE ANALİZ

Bu tez çalışmasında, bir üniversitesinin UZEM'in öğrenci, fakülte, ders ve başarı kriterleri üzerinde VM yöntemleri uygulanarak gelecek dönemde yürütülecek olan uzaktan eğitim modelleri hakkında tahminlerde bulunulmuştur.

Veri hazırlama sürecinde Microsoft Excell, Microsoft Access, Ms-SQL, ALMS ve Rapid Miner Studio programlarından yararlanılmıştır.

KAYUZEM'in 06.04.2020 ve 15.08.2020 tarihleri arasındaki verileri kullanılmıştır.

Kayseri Üniversitesinin kullandığı ALMS sistemi üzerinden veriler Ms-SQL sorgusuyla Microsoft Excell formatında dışa aktarılmıştır. Verilerin alındığı ALMS yazılımına giriş ara yüzü Şekil 3.1'de gösterilmiştir. Üniversite öğrencileri, uzaktan eğitim sistemine şekilde görülen ekrandan öğrenci numarası ve şifresi ile giriş yapabilmektedir.

Şekil 3.1. Kayseri Üniversitesi ALMS kullanıcı giriş ekranı.

Öğrencilere ait veriler Ms-SQL sorgusu ile KAYUZEM tarafından ALMS üzerinden raporlanmıştır. Bu veriler Microsoft Access programında düzenlenmiş olarak araştırmacının 29.12.2020 tarih ve 20247 sayılı izin talebi dilekçesine (EK-1) istinaden 29.01.2021 tarihli tutanak ile KVKK uygun şekilde teslim alınmıştır (EK-2). Verilerin

etik çerçevede kullanılmasının onayı Kayseri Üniversitesi Etik Kurul Başkanlığının iznine tabidir (EK-3).

3.1. VERİLER VE ÖZELLİKLERİ

VTYS 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. Üst başlıklarda öğrencilerin bağlı bulunduğu fakülte ve MYO isimleri “Fakülte/MYO” sütununa, öğrencinin bağlı bulunduğu bölüm isimleri “Bölüm Adı” sütununa, öğrencinin aldığı dersler ‘Ders Adı’ sütununa yazılmıştır. “Derse Katılma Yöntemi” sütunu için öğrencilerin derslere katılma durumları yöntemsel olarak şu şekilde adlandırılmıştır; Bir öğrencinin derse canlı katılım dakikası fazla ise derse katılma yöntemi “Senkron”, öğrenci derse kayıttan izlediye “Asenkron” ve öğrenci derse hem canlı hem kayıttan izlediye “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 alınarak Çizelge 3.1 oluşturulmuştur.

Çizelge 3.1. Not baremi tanımlama çizelgesi.

Başarılı	Başarısız
100-70	69-0

Çizelge 4.1’e göre 100-70 not bareminde harf notu AA, BA, BB, CB, CC olan öğrenciler, “başarılı”, 69-0 not bareminde harf notu DC, DD, FD, FF olan öğrenciler “başarısız” olarak tanımlanmıştır.

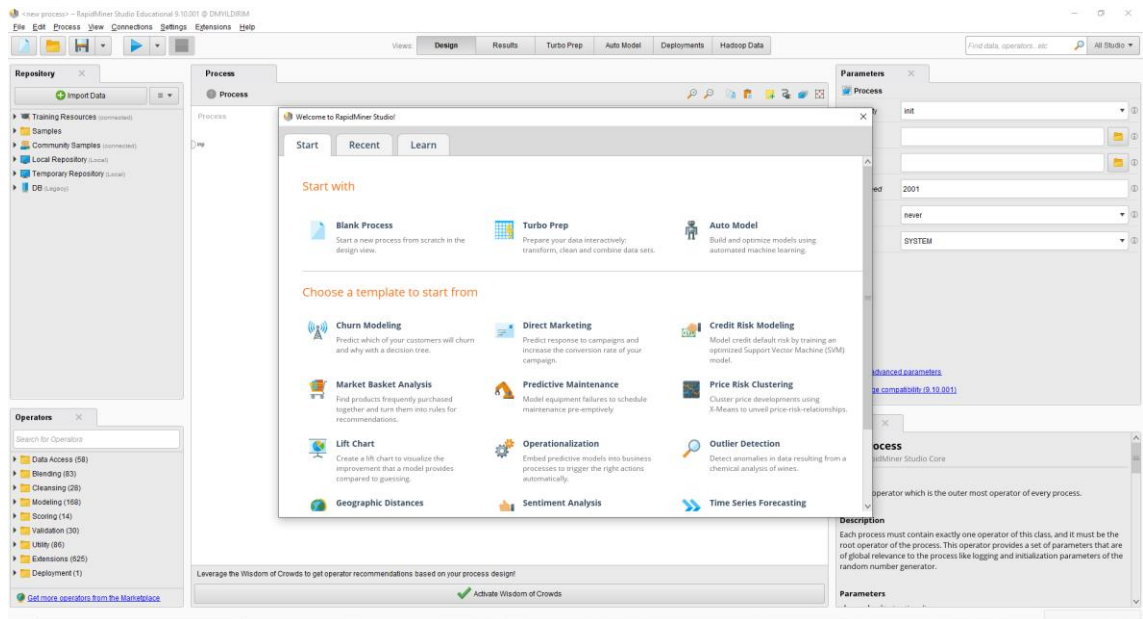
Veri setinde 3 farklı Fakülte ve 1 MYO’dan alınmış toplamda 161 farklı ders ve 1399 öğrencinin başarı durumu bulunmaktadır. Bu alanlarda toplamda 8319 işlenmiş veri vardır. Bu veri alanlarından tahmin işlemine tabi tutulacak kısım öğrencilerin başarı durumu alanıdır. Veri setinde tahminleme öğrencilerin başarı durumu özelliğine göre yapılmıştır.

Düzenlemesi yapılan veriler RapidMiner VM programına aktarılmış ve yukarıda adı geçen üç yöntemin öğrencinin başarı durumu üzerindeki etkisinin analizi yapılmıştır.

Verilerin %80'ni eğitim verisi geri kalan %20'si ise test verisi olarak kullanılmıştır.

3.2. ÇALIŞMADA KULLANILAN PROGRAM

VM algoritmaları ile veri analiz etmek için açık ve kapalı kaynak kodlu en sık kullanılan yazılımlar; SPSS, WEKA, Orange, RapidMiner, SAS, Angoss, Sciptella ETL, PSS Clementine, KNIME gibi yazılımlardır. Verilerin analizinde Rapidminer yazılımı kullanılmıştır. Açık kaynak kodlu bir yazılımdır. Akademik çalışmalarda ücretsiz olarak bütün özellikleriyle birlikte kullanılabilir. Yazılımın giriş ekranı Şekil 3.2'de gösterilmiştir.



Şekil 3.2. RapidMiner giriş ekranı.

3.3. RAPIDMINER İLE VERİLERİN ANALİZİ

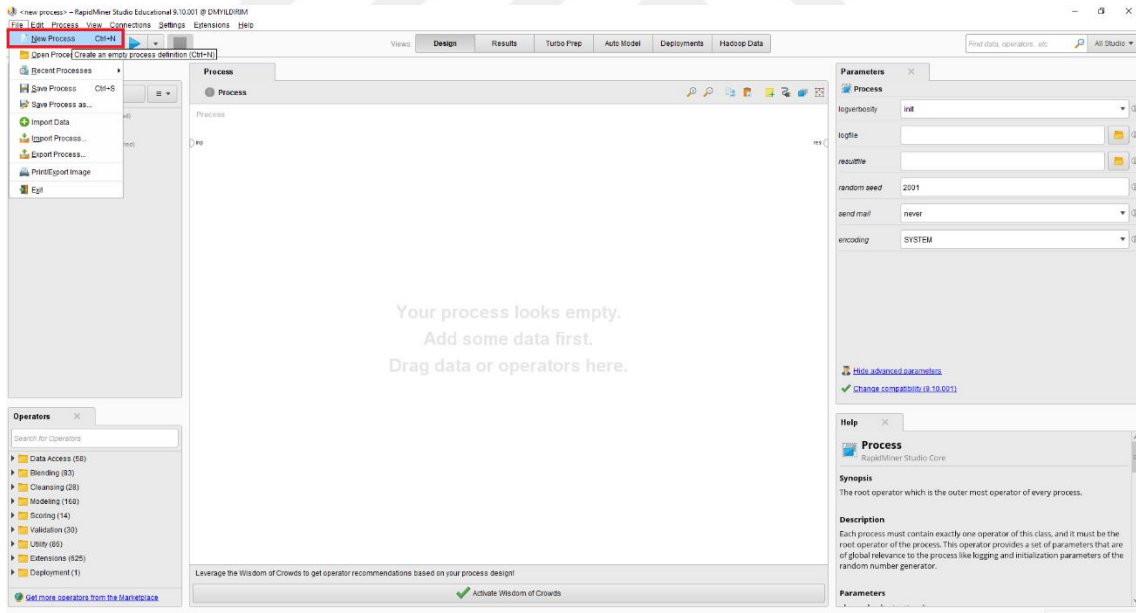
VTBK süreçleri doğrultusunda problem tanımlama, veri hazırlama, veri toplama, birleştirme ve temizleme, seçme işlemlerinden geçirilip hazır hale getirilen veriler, Şekil 3.3'de gösterilmiştir. Fakat kişisel veri alanı olan “Kullanıcı Adı” ve “Adı Soyadı” özelliklerinin entropilerinin çok düşük olmasından dolayı modellerin sonuçlarını etkilemeyeceği için çıkarılmıştır.

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**** ...	Develi İslami İlimler Fakültesi	İslami İlimler	Arapça Okuma Anlama I	Asenkron	Başarılı
10****29	A**** S**** ...	Develi İslami İlimler Fakültesi	İslami İlimler	Arapça Sarf-Nahiv I	Asenkron	Başarılı
10****29	A**** S**** ...	Develi İslami İlimler Fakültesi	İslami İlimler	Eğitim Psikolojisi	Asenkron	Başarılı

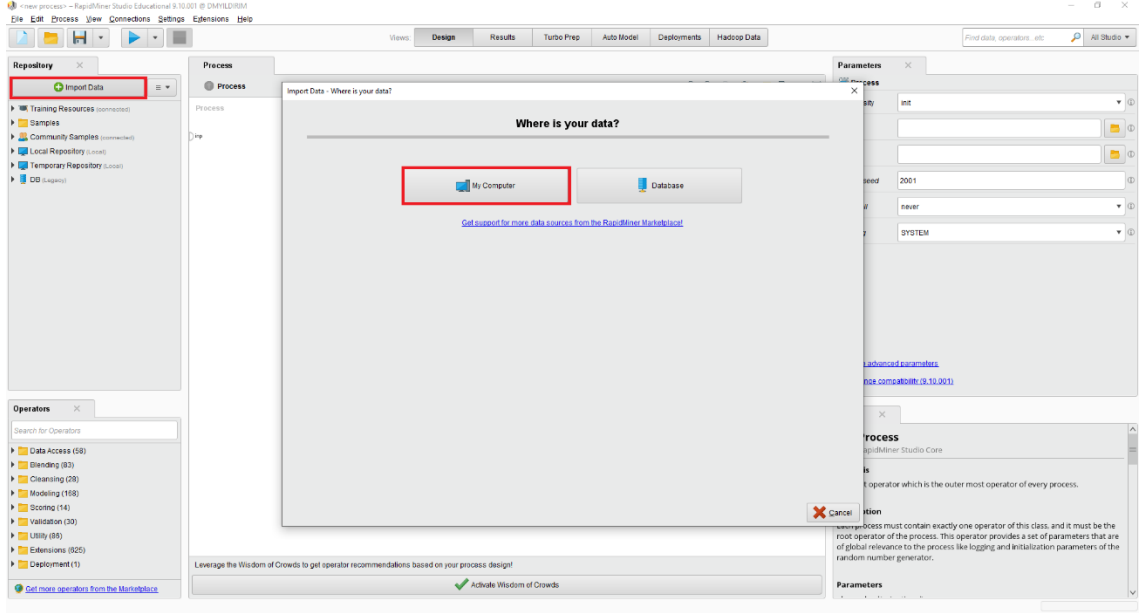
Şekil 3.3. RapidMiner programına aktarılmadan önceki veriler.

RapidMiner programına geçmiş veriden geleceğe yönelik tahmin yapmak üzere aktarılmıştır.

Veri aktarımının ilk adımı olarak “File” menüsünden Şekil 3.4’de gösterildiği üzere “New Process” seçilir. VTBK süreçlerinden geçmiş veriler Microsoft Excell formatında sırasıyla “import data”, “my computer” adımları ile dosya seçilerek yükleme yapılır (Şekil 3.4 ve Şekil 3.5). Çizelge 3.1’de belirtilmiş olan notlara karşılık gelen değerlerden oluşan Microsoft Excell dosyasındaki veri seti RapidMiner’a yüklenmiştir.

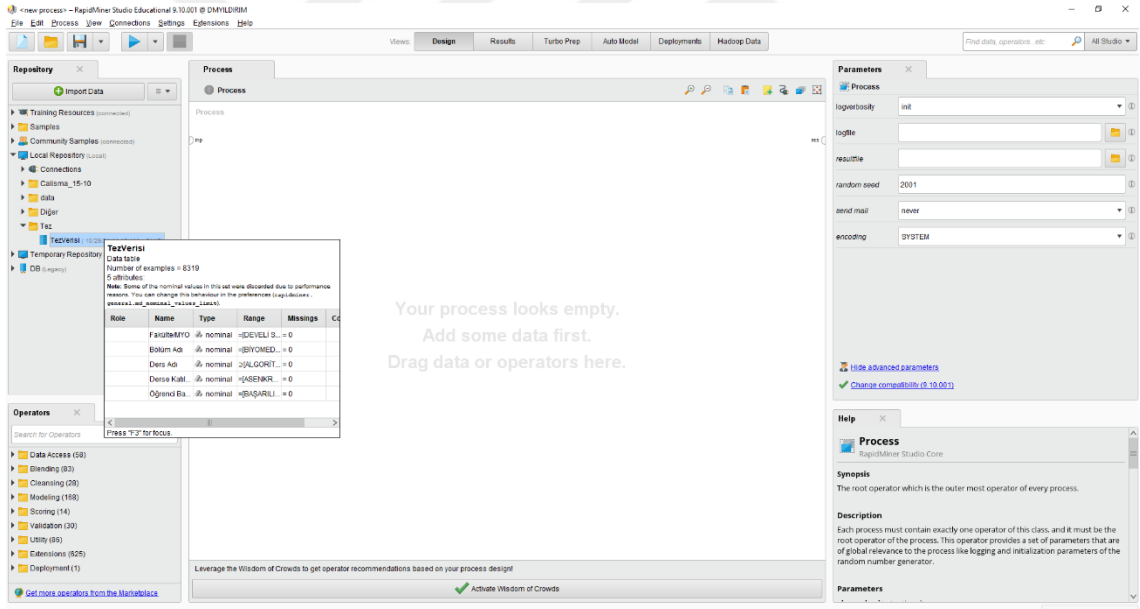


Şekil 3.4. RapidMiner “new process” oluşturma ekranı.



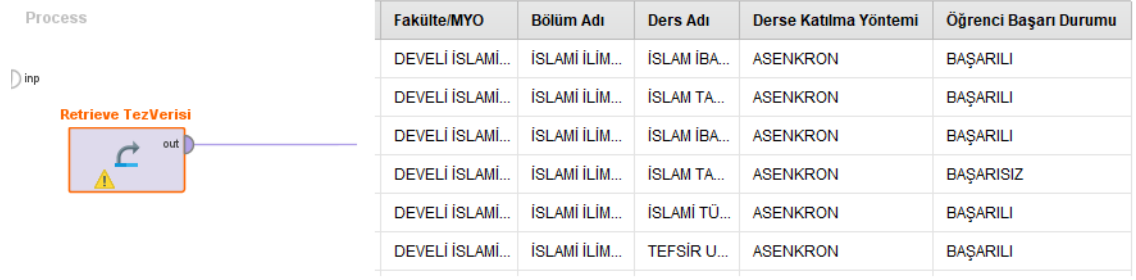
Şekil 3.5. RapidMiner veri yükleme ekranı.

Şekil 3.5’de gösterildiği gibi veri setleri RapidMiner programına yüklenir.



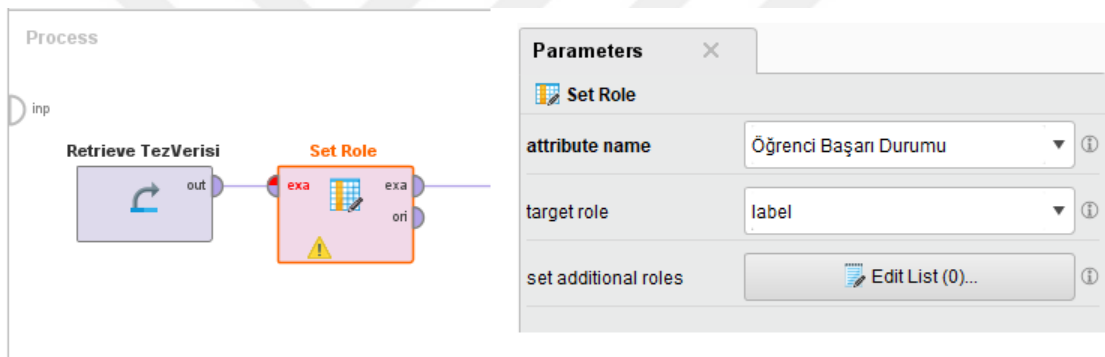
Şekil 3.6. RapidMiner veri setleri.

Yeni oluşturulan Process altına veri seti sürükleyip bırak yöntemiyle yüklenir. Kullanılacak veri setini sürükleyip bırak yöntemi ile Process alanına bırakılır. Şekil 3.7’de kullanılacak veri seti ve içeriğinin bir kısmı gösterilmiştir.



Şekil 3.7. RapidMiner join operatörü ile veri seti birleştirme ekranı.

Veri seti eklendikten sonra Rol Belirleme (Set Role) operatörü Process alanına eklenmiştir. Burada rol belirleme operatörü veri setinde hangi alanın etiketleneceğini yani hangi alanda tahminler yapılacağını belirleyen operatördür. Şekil 3.8’de görüldüğü gibi veri setinden rol belirleme operatörüne bağlama işlemi yapılmış ve rol belirleme operatörünün parametresi olan attribute name alanında “Öğrenci Başarı Durumu” sütunu seçilmiş, Target Role parametresindeki ayarı da label seçilmiştir.



Şekil 3.8. RapidMiner rol belirleme operatörü ve parametreleri.

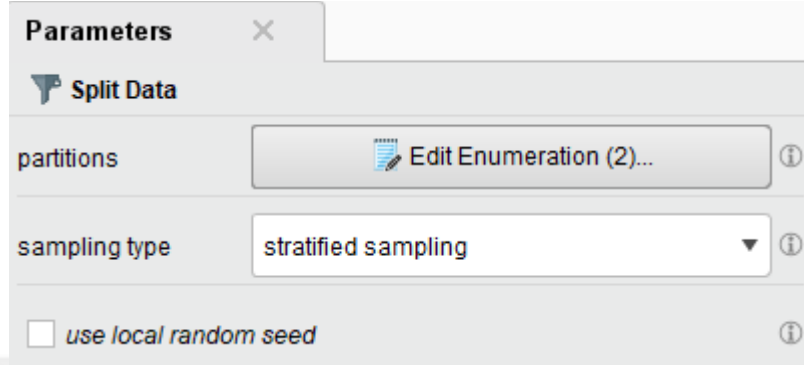
Verileri Bölme (Split Data), verilerin ne kadarının eğitim ve ne kadarının test verisi olarak kullanılacağını belirleyen operatördür. Bu operatör de Process alanına eklenmiş ve rol belirleme operatörü ile bağlantısı gerçekleştirilmiştir. Şekil 3.9’da görünen Verileri Bölme operatörünün parametresinde ilk başta Numaralandırmayı Düzenle (Edit Enumeration) butonuna tıkladığında Şekil 3.10’daki ekran gelmektedir. Burada Giriş Ekleme (Add Entry) butonuna tıklanıp 2 adet ratio oluşturulmuştur. İlk önce eğitim verisi yüzdesi ve ikinci olarak test verisi yüzdesi yazılır. Örneklem Tipi (Sampling Type) bölümündeki seçeneklerin açıklaması aşağıda belirtildiği gibidir ve bu bölümde Tabakalı Örneklem kullanılır.

Doğrusal Örneklem (Linear Sampling): Ardışık örneklere sahip alt kümeler oluşturur. Yani örneklerin sırasını değiştirmeden bölümlere ayırır.

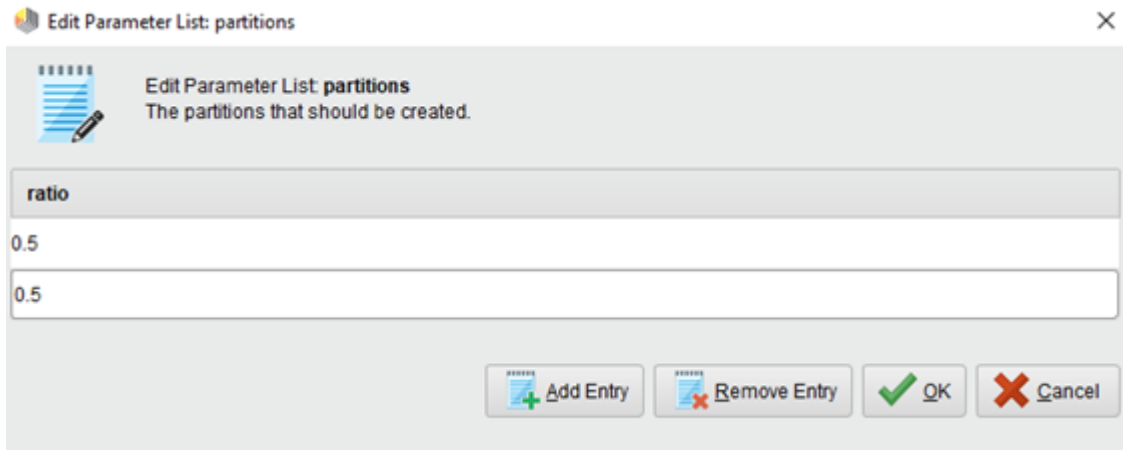
Karıştırılmış Örneklem (Shuffled Sampling): Rastgele alt kümeleri oluşturur.

Tabakalı Örneklem (Stratified Sampling): Rastgele alt kümeleri oluşturur ve alt kümelerdeki sınıf dağılımlarını tüm örnek kümesi ile aynı olmasını sağlar.

Otomatik (Automatic): Veri setine göre yukarıdaki 3 durumdan hangisi iyi bir sonuç veriyorsa onu otomatik olarak seçer ve kullanır.



Şekil 3.9. RapidMiner verileri bölme operatörü ve parametreleri.

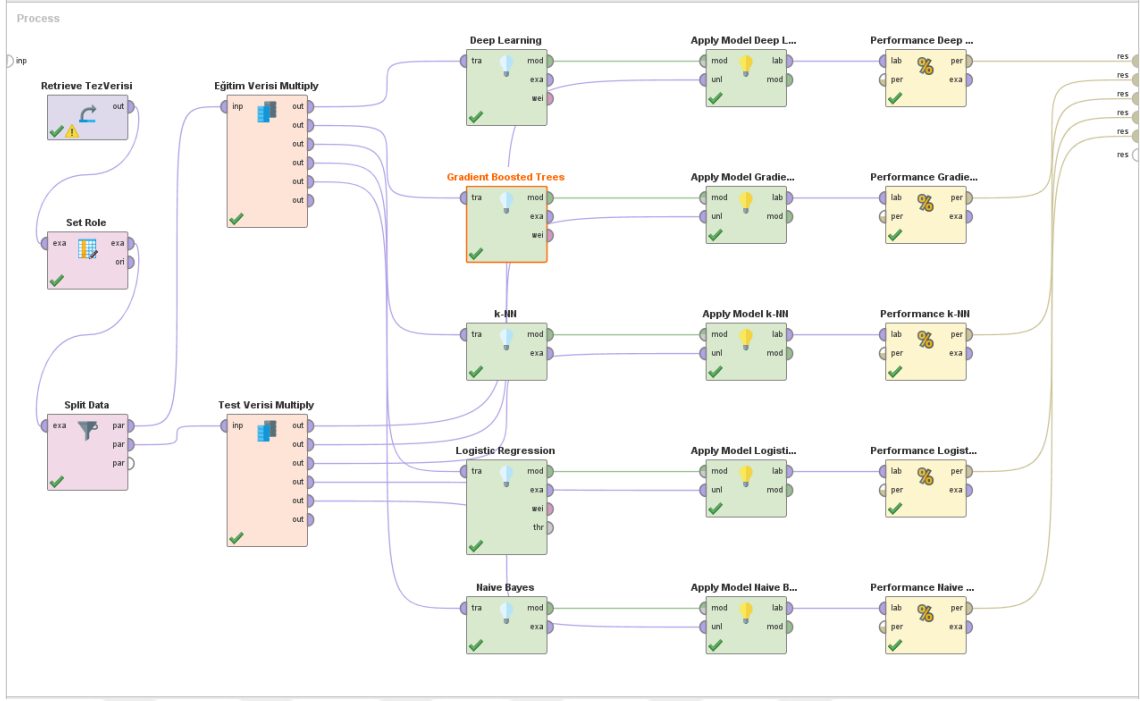


Şekil 3.10. RapidMiner verileri bölme operatörü ve parametreleri.

Verilerin eğitim ve test verisi yüzdeleri Verileri Bölme operatöründe belirlendikten sonra, eğitim verisinin kullanılacağı algoritmaya bağlantı yapılır. Test verisi ve eğitilmiş olan verinin de Modeli Uygula Apply Model operatörüne bağlantısı yapılır. Modeli Uygula operatörü eğitilmiş verileri kullanarak test verilerinin üzerinde tahmin işlemi yapar. Bu operatörün parametrelerinin varsayılan değerleri kullanılır.

Modeli Uygula operatöründen çıkan verilerinin istatistiklerini daha anlaşılır bir şekilde görmek için Performans (Performance) operatörü kullanılmaktadır.

Çalışmada kullanılan operatörler Şekil 3.11’de gösterilmiştir.

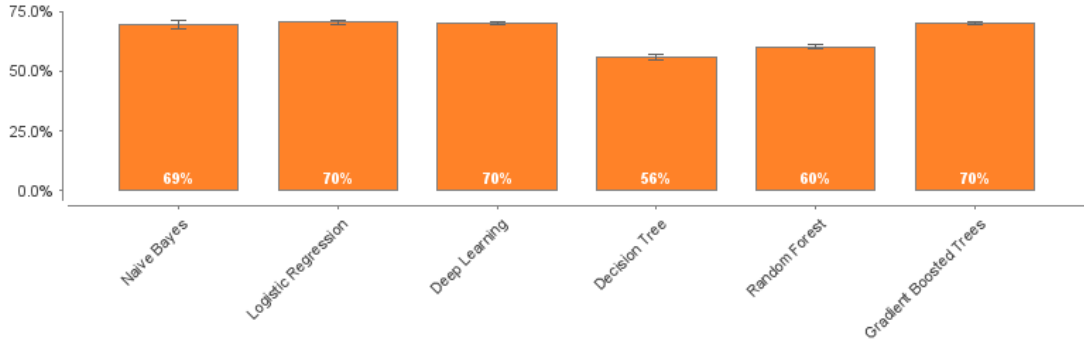


Şekil 3.11. Çalışmada kullanılan operatörler.

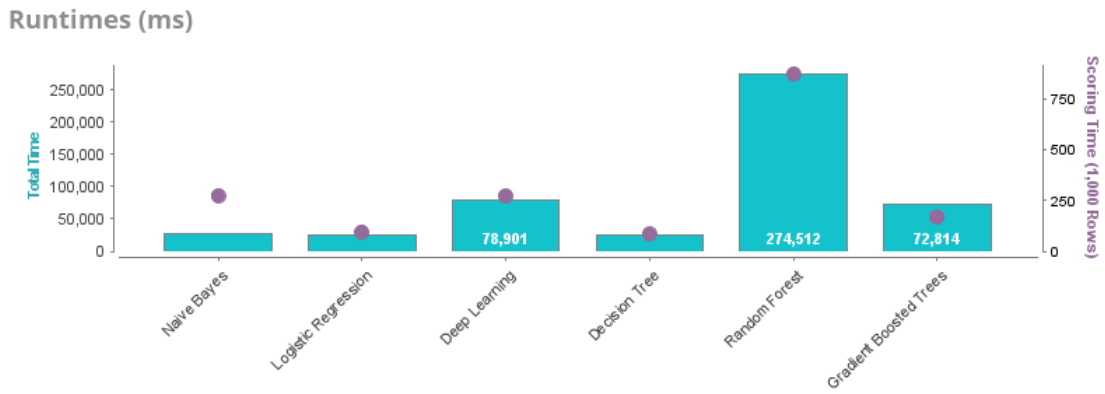
3.4. VERİ MADENCİLİĞİ YOLUYLA ELDE EDİLEN BULGULAR

Kullanılacak olan algoritmaları belirlerken ilk önce Şekil 3.7’de belirtilen verilerin RapidMiner yazılımının bir özelliği olan ve en sık kullanılan algoritmalar ile otomatik olarak tahmin ve analizi yapılır. Bu işlem yapıldıktan sonra tahmin işleminin başarısının hangi algoritmalarda daha iyi olduğu ortaya çıkmıştır. Seçilen algoritmalarda daha iyi sonuçlar alabilmek için algoritmaların parametrelerinde değişikliklere gidilmiştir. Şekil 3.12’de tahmin oranları ve Şekil 3.13’de tahmin için geçen süreler mevcuttur. Bu veriler ışığında otomatik tahminleme analizinde %69 olarak çıkan Naive Bayes, %70 olarak çıkan Lojistik Regresyon (Logistic Regression), %70 olarak çıkan Derin Öğrenme (Deep Learning) ve %70 olarak çıkan GBT algoritmaları daha detaylı olarak incelenmiştir. Karar Ağacı ve Rastgele Orman 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.

¹ Burada yapılan çalışmalarda intel i7 6700HQ işlemci, 16 GB Ram, Nvidia Geforce GTX 960M 4 GB ekran kartı ve Windows 10 işletim sistemi kullanılmıştır. Bu özellikleri dışında bir özelliği olan bilgisayar kullanıldığı takdirde zaman değerinde değişiklikler gözlemlenebilir.



Şekil 3.12. Auto model doğruluk oranı.



Şekil 3.13. Auto model işlem süresi (ms).

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, başarı oranlarının artırılması için parametrelerinde yapılan değişiklikler ve sonuçlar sınıflandırma yöntemleri bölümünde detaylı olarak verilecektir.

Çalışmanın sınıflandırma süreci Şekil 3.14’de gösterilmiştir.

özellikler ikiden fazla kategoriye sahipse kodlama işlemi yapılır. Örneğin Çizelge 3.2’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.

Çizelge 3.2. 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.

3.5.1. Derin Öğrenme ile Sınıflandırma

Derin öğrenme algoritmasında değişiklik yapılan parametrelerin açıklamaları aşağıda verilmiştir. Sonucu etkilemediği görülen bazı parametreler varsayılan olarak bırakılmıştır.

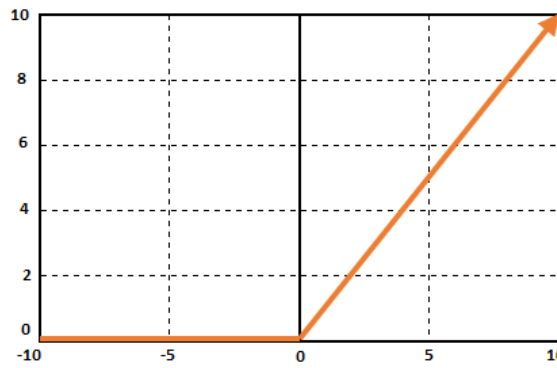
Activation: Yapay sinir ağlarına doğrusal olmayan gerçek dünya özelliklerini tanıtmak için aktivasyon fonksiyonuna ihtiyaç duyulur. Temel olarak basit bir yapay sinir ağında x girdiler, w ağırlıklar olarak tanımlanır ve ağın çıkışına aktarılan değere $f(x)$ yani aktivasyon işlemi uygulanır. Daha sonra bu, nihai çıkış ya da bir başka katmanın girişi olacaktır [75]. Seçim yapılabilecek tanh, Doğrultucu (Rectifier), Maxout ve ExpRectifier fonksiyonları bulunmaktadır. Çizelge 3.3’de aktivasyon fonksiyonlarının

karşılaştırılması mevcuttur. En iyi sonucu veren Doğrultucu fonksiyonu kullanılmıştır. Bütün fonksiyonların çalışma zamanları birbirine çok yakındır. Bu işlem yapılırken veriler rastgele, %80'i eğitim ve %20'si test verisi olarak kullanılmıştır.

Çizelge 3.3. Derin öğrenme aktivasyon fonksiyonlarının karşılaştırılması.

Eğitim Oranı	Test Oranı	Tanh	Doğrultucu	Maxout	ExpRectifier
80%	20%	69,41%	71,33%	69,05%	69,41%

Doğrultucu: X'in giriş değeri olduğu maksimum (0, x) değerini seçer [76].



Şekil 3.15. Doğrultucu grafiği.

Gizli Katman Sayısı (Hidden Layer Sizes): Gizli katman sayısını ve bu katmanlardaki nöronların sayısını ayarlamaya yardımcı olur. Çizelge 3.4'de gizli katmanları ve nöron sayılarını değiştirerek en iyi sonuç bulunmaya çalışılmıştır. Bu işlem yapılırken aktivasyon fonksiyonu Doğrultucu seçilmiş ve veriler rasgele bir şekilde %80'i eğitim ve %20'si test verisi olarak kullanılmıştır.

Çizelge 3.4. Derin öğrenme gizli katman ve nöronların karşılaştırılması.

Eğitim Oranı	Test Oranı	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
80%	20%	71,33%	67,67%	69,05%	70,31%	67,61%

Çizelge 3.4'de görüldüğü üzere en iyi sonucu veren durum; 1 gizli katmandan oluşan ve bu katmanda 30 nöronun bulunmasıdır.

Reproducible (uses 1 thread): Küçük verilerde tahminlerin yeniden üretimini zorlar. Bu parametre seçildiğinde algoritma her yeni çalıştırıldığında farklı sonuçlar elde edilmez. Tahminlemenin yüzdelik oranı değişmez. Bu durumda sabit bir yüzdelik değer verilmesini ve yapılan tahminlemenin net olmasını sağlar. Derin öğrenmenin çalışma hızını yavaşlatır.

Yineleme (Epochs): Veri kümesi kaç kez tekrarlanacaksa, bu parametre ile ayarlanır. Varsayılan ayar burada 10 dur. Çizelge 3.5’de değişik tekrarlanma durumlarına bakılmıştır. Bu durum neticesinde varsayılan değer altında bir değer olduğunda tahmin oranları düşmekte, varsayılan değer üstüne çıkıldığında ise tahmin değerinde herhangi bir değişiklik olmamaktadır. En iyi tahmin değeri varsayılan değer olan 10 olduğu tespit edilmiştir.

Çizelge 3.5. Derin öğrenme veri kümesi yinelenmesi karşılaştırılması.

Eğitim Oranı	Test Oranı	5	7	10	20	100
80%	20%	68,45%	67,25%	71,33%	69,41%	69,41%

Bu çalışmadaki derin öğrenme algoritmasında değişiklik yapılan parametreler Çizelge 3.6’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.

Çizelge 3.6. 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 Yineleme 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 Çizelge 3.7’de verilmiştir.

Çizelge 3.7. 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%	

Çizelgede “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.

3.5.2. k-NN ile Sınıflandırma

k-NN algoritmasında değişiklik yapılan parametrelerin açıklamaları aşağıda verilmiştir.

k: Burada k parametresine verilen değer kadar en yakın komşuya bakar ve baktığı en yakın komşulara göre sınıflandırma yapar. Çalışmamızda k değerinin 1’den başlayarak tek sayı artırımlı olarak modelin başarısının en iyi olduğu değer bulunması hedeflenmiştir. Çizelge 3.8’de denenmiş tüm değerler değil başarının yükseldiğini

gösterebileceğimiz ara değerler verilmiştir. Burada en iyi sonucu veren k değeri 13 değeri olmuştur. k-NN operatörünün diğer parametreleri sonucu değiştirmedığı için varsayılan değerlerde kullanılmıştır. Eğitim verisi oranı %80 ve test verisi oranı %20 kullanılmıştır.

Çizelge 3.8. k-NN algoritmasının k değerinin karşılaştırılması.

Eğitim Oranı	Test Oranı	k = 1	k = 5	k = 13	k = 19	k = 79
80%	20%	61,90%	67,49%	70,13%	68,93%	66,23

K-NN algoritmasının veri seti üzerindeki başarısı Çizelge 3.9’da verilmiştir.

Çizelge 3.9. 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%	

Çizelgede “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 %70,13 olarak hesaplanmıştır. Daha yüksek doğruluk oranına ulaşmak için otomatik modellemeyle belirlediğimiz diğer algoritmalar üzerinde çalışmalara devam edilmiştir.

3.5.3. Naive Bayes ile Sınıflandırma

Naive Bayes algoritmasının parametresinde herhangi bir değişiklik yapılmamıştır. Varsayılan parametre kullanılmıştır. Bu parametrenin açıklaması aşağıda verilmiştir.

Laplace Düzeltmesi (Correction): Naive Bayes algoritmasının eğitim verileri içerisinde bir nitelik değeri vardır ve bu değer belirli bir sınıf bağlamında ortaya çıkmazsa, koşullu olasılığını sıfır olarak ayarlar. Bu sıfır değer diğer olasılıklarla çarptığında diğer değerler de sıfır çıkarak sonucu yanıltabilir. Bu durumun yaşanmaması için, her sayıma bir ekler ve bu sorunu çözmemize yardımcı olur.

Verilerin %80’i eğitim ve %20’si test verisi olarak kullanılmış.

Naive Bayes algoritmasının veri seti üzerindeki başarısı Çizelge 3.10’da verilmiştir.

Çizelge 3.10. 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%	

Çizelgede “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.

3.5.4. GBT ile Sınıflandırma

GBT algoritmasında değişiklik yapılan parametrelerin açıklamaları aşağıda verilmiştir. Parametrelerin varsayılan değerlerinin değiştirilmesi sonucu etkilememiştir.

Ağaç Sayısı (Number of Trees): Bu parametre ile kullanılacak ağaç sayısını belirler. Çizelge 3.11’de ağaç sayısı parametresi değerlerinin karşılaştırması yer almaktadır. Bu karşılaştırma neticesinde ağaç sayısı 25’in altında bir değer verildiğinde tahmin oranı düşmekte, ağaç sayısı 25’in üstüne çıkıldığında ise tahmin değerinde herhangi bir değişiklik olmamaktadır.

Çizelge 3.11. GBT algoritmasının ağaç sayısı değerinin karşılaştırılması.

Eğitim Oranı	Test Oranı	Ağaç = 20	Ağaç = 22	Ağaç = 25	Ağaç = 50	Ağaç = 100
80%	20%	69,05%	69,47%	70,01%	70,01%	70,01%

Minimum Bölünmüş İyileştirme (Min Split Improvement): Bu parametrede bir bölünmenin gerçekleşmesi için karesel hata azaltmada minimum ilgili gelişmeyi belirtir. Çizelge 3.12’de minimum bölünmüş iyileştirme parametresi değerlerinin karşılaştırması yer almaktadır. Bu karşılaştırma neticesinde %70,01 oranına sahip 1.0E-4 değeri kullanılmıştır.

Çizelge 3.12. GBT algoritmasının minimum bölünmüş iyileştirme değerinin karşılaştırılması.

Eğitim Oranı	Test Oranı	1.0E-1	1.0E-4	1.0E-5	1.0E-7	1.0E-10
80%	20%	67,73%	70,01%	69,53%	69,53%	69,53%

Öğrenme Oranı (Learning Rate): Bu parametrede öğrenme oranını belirtir. Çizelge 3.13’de öğrenme oranı parametresinin değerlerinin karşılaştırması yer almaktadır. Bu karşılaştırmanın neticesinde %70,01 oranına sahip 0,191 değeri kullanılmıştır.

Çizelge 3.13. GBT algoritmasının öğrenme oranı değerinin karşılaştırılması.

Eğitim Oranı	Test Oranı	0,01	0,02	0,1	0,15	0,191
80%	20%	69,11%	68,09%	68,81%	68,33%	70,01%

Verilerin %80’i eğitim ve %20’si test verisi olarak kullanılmış. Bu çalışmadaki GBT algoritmasında değişiklik yapılan parametreler Çizelge 3.14’de gösterilmiştir. Sonucu etkilemediği görülen bazı parametreler varsayılan olarak bırakılmıştır.

Çizelge 3.14. GBT 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

GBT algoritmasının veri seti üzerindeki başarısı Çizelge 3.15’de verilmiştir.

Çizelge 3.15. GBT algoritması karşılaştırması.

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

Çizelgede “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 GBT algoritmasındaki toplamdaki başarı oranı ortalaması %70,01 olarak hesaplanmıştır. Daha yüksek doğruluk oranına ulaşmak için diğer algoritmalar üzerinde çalışmalara devam edilmiştir.

3.5.5. Lojistik Regresyon ile Sınıflandırma

Lojistik Regresyon algoritmasında değişiklik yapılan parametrelerin açıklamaları aşağıda verilmiştir. Sonucu etkilemeyen parametreler için varsayılan değerler kullanılmıştır.

Çözücü (Solver): Kullanılacak olan çözümleyicinin seçildiği yerdir. Burada altı adet çözümleyici seçeneği mevcuttur. Bu parametrede varsayılan olarak seçilen otomatik seçeneği daha optimal olmuştur.

Düzenlileştirmeyi Kullanma (Use Regularization): Düzenlileştirme kullanılması gerektiğinde bu parametre kullanılır. Yapılan çalışmada bu parametre açıldıktan sonra karşımıza lambda, lambda arama ve alfa parametreleri gelmektedir. Bu parametrelerde herhangi bir değişiklik yapılmayarak varsayılan değerler kullanılır.

Lambda: Uygulanan düzenlileştirme miktarını kontrol eder.

Lambda Arama: Verilen lambda değerinin maksimum Lambda’dan başlayarak lambda değerlerinin aralığı üzerinden arama yapılıp yapılamayacağını gösteren parametredir.

Alfa (Alpha): L1 (Lasso) ve L2 (Ridge Regresyon) cezaları arasındaki dağılımı kontrol eder.

Verilerin %80’i eğitim ve %20’si test verisi olarak kullanılmış. Lojistik Regresyon algoritmasında parametreler sonucu değiştirmediklerinden dolayı varsayılan değerler kullanılmıştır.

Lojistik Regresyon algoritmasının veri seti üzerindeki başarısı Çizelge 3.16’da verilmiştir.

Çizelge 3.16. Lojistik Regresyon algoritması karşılaştırması.

	Gerçek Başarılı	Gerçek Başarısız	Doğruluk
Tahmin Başarılı	699	208	77,07%
Tahmin Başarısız	233	524	69,22%
Hatırlama	75,00%	71,58%	

Çizelgede “Gerçek Başarılı” sayısı toplamda 932 veriyken modelimiz bunun 699 verisini doğru tahmin etmiştir. 233 adet veriyi ise “Tahmin Başarısız” olarak yanlış tahmin etmiştir. Bu şekilde bütün çizelgedeki başarılı ve başarısız tahminler oranlandığı zaman Lojistik Regresyon algoritmasındaki toplamdaki başarı oranı ortalaması %73,50 olarak hesaplanmıştır.

3.5.6. Algoritmaların Karşılaştırılması

Derin Öğrenme, k-NN, Naive Bayes, GBT ve Lojistik Regresyon algoritmaları karşılaştırıldığında en iyi performansı sırasıyla %73,50 ile Lojistik Regresyon, %71,33 ile Derin Öğrenme, %70,73 ile k-NN, %70,01 ile Naive Bayes ve GBT algoritmaları vermiştir.

Çizelge 3.17’de algoritmaların eğitim ve tahmin için kullanılan oranlarının başarı oranına etkileri verilmiştir.

Çizelge 3.17. Algoritmaların karşılaştırması.

Eğitim İçin Kullanılan Veri / Tahmin İçin Kullanılan Veri	Tahmin Oranları				
	Derin Öğrenme	k-NN	Naive Bayes	GBT	Lojistik Regresyon
%20 / %80	67,24%	65,41%	68,38%	67,84%	59,85%
%40 / %60	64,58%	67,97%	68,79%	69,21%	71,49%
%50 / %50	68,74%	67,37%	68,53%	68,67%	71,12%
%60 / %40	70,36%	68,95%	69,46%	69,34%	72,23%
%80 / %20	71,33%	70,73%	70,01%	70,01%	73,50%

Çizelge 3.17’de en başarılı tahmin oranının %80 eğitim ve %20 test verisiyle olduğu

anlaşlmıřtır

Çizelge 3.18. Algoritmaların alıřma sürelerinin karşılařtırması.

Eđitim İin Kullanılan Veri / Tahmin İin Kullanılan Veri	alıřma Süreleri				
	Derin Öđrenme	k-NN	Naive Bayes	GBT	Lojistik Regresyon
%20 / %80	0 saniyeden daha az	2 saniye	0 saniyeden daha az	0 saniyeden daha az	0 saniyeden daha az
%40 / %60	1 saniye	2 saniye	0 saniyeden daha az	0 saniyeden daha az	0 saniyeden daha az
%50 / %50	1 saniye	3 saniye	0 saniyeden daha az	0 saniyeden daha az	0 saniyeden daha az
%60 / %40	1 saniye	3 saniye	0 saniyeden daha az	0 saniyeden daha az	0 saniyeden daha az
%80 / %20	2 saniye	2 saniye	0 saniyeden daha az	0 saniyeden daha az	0 saniyeden daha az

Çizelge 3.18’de algoritmaların eđitim ve tahmin oranlarına göre alıřma süreleri gözlenmektedir. Tahmin oranları ve alıřma süreleri göz önüne alındığında en iyi deđere ve en iyi alıřma süresine sahip algoritma Lojistik Regresyondur.

3.6. VERİ MADENCLİđİ BAřARI KRİTERLERİ (KARIřIKLIK MATRİSİ)

Lojistik Regresyon algoritmasıyla en iyi tahminlemeye ulařılmıřken, karıřıklık matrisiyle derse katılma yöntemlerinden hangisinin daha verimli ve güvenilir olduđu hesaplanmıřtır.

Çizelge 3.19. Karıřıklık matrisi çizelgesi (Asenkron).

Asenkron	Bařarılı (Tahmin)	Bařarısız (Tahmin)
Bařarılı (Gerek)	298 (TP)	157 (FN)
Bařarısız (Gerek)	105 (FP)	321 (TN)

Çizelge 3.19’da asenkron yönteminin karışıklık matrisinin çizelgesi oluşturulmuştur. Bu çizelgeye göre hesaplamalar yapılarak aşağıdaki değerler elde edilmiştir.

- Doğruluk Oranı: %70,26
- Yanlış Sınıflandırma Oranı: %29,73
- Gerçek Pozitif Değerlerin Oranı: %68,49
- Gerçek Negatif Değerlerin Oranı: %75,35
- Hassasiyet Oranı: %73,94
- F Puanı: %69,46

Çizelge 3.20. Karışıklık matrisi çizelgesi (Senkron).

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

Çizelge 3.20’de senkron yönteminin karışıklık matrisinin çizelgesi oluşturulmuştur. Bu çizelgeye göre hesaplamalar yapılarak aşağıdaki değerler elde edilmiştir.

- Doğruluk Oranı: %67,39
- Yanlış Sınıflandırma Oranı: %26,03
- Gerçek Pozitif Değerlerin Oranı: %66,16
- Gerçek Negatif Değerlerin Oranı: %77,43
- Hassasiyet Oranı: %72,28
- F Puanı: %69,09

Çizelge 3.21. Karışıklık matrisi çizelgesi (Hibrit).

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

Çizelge 3.21’de hibrit yönteminin karışıklık matrisinin çizelgesi oluşturulmuştur. Bu çizelgeye göre hesaplamalar yapılarak aşağıdaki değerler elde edilmiştir.

- Doğruluk Oranı: %83,14
- Yanlış Sınıflandırma Oranı: %16,85
- Gerçek Pozitif Değerlerin Oranı: %97,10
- Gerçek Negatif Değerlerin Oranı: %35
- Hassasiyet Oranı: %83,75
- F Puanı: %89,93

Çizelge 3.22. 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 Puanı	69,46%	69,09%	89,93%

Çizelge 3.22’de 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.

4. SONUÇLAR VE ÖNERİLER

4.1. SONUÇLAR

Dünyada ve ülkemizde veri madenciliği çalışma alanı gün geçtikçe gelişmektedir. Bu gelişmenin önemli örnekleri eğitim-öğretim alanında ortaya çıkmaktadır. Hacım hız ve çeşitlilik bakımından artan günümüz verileri büyük veri analitiği yaklaşımına yol açmıştır [3].

Bütün dünyayı etkisi altına alan covid-19 pandemisiyle birlikte üniversitelerin uzaktan eğitim sistemlerinde biriken verileri odağına alan bu çalışmada, bir üniversitede uzaktan eğitimde verilen derslerin önemli çıktıları olan öğrenci notlarını içeren büyük verileri üzerinde VM metotları ve algoritmaları kullanılmıştır. 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 oluşturulan hipoteze göre yürütülen çalışmada Rapidminer ile tahmin edici ve tanımlayıcı metotlarla inceleme ve analizler yapılmıştır.

Araştırma adımları, ilerleme sırasına göre tezin sunumuna uygun olarak aşağıda verilmiştir.

1. KAYUZEM'in kullandığı ALMS sistemi web tabanlıdır ve Ms-SQL veritabanı kullanmaktadır.
2. ALMS sisteminin 06.04.2020 ve 15.08.2020 tarihleri arasındaki verileri Ms-SQL sorgusuyla Microsoft Excell formatında alınmıştır.
3. Microsoft Excell formatındaki verilerin düzenlenmesi için Microsoft Access veritabanı programına aktarımı yapılmıştır.
4. Alınan verilerde öğrencilerin kayıtlı olduğu Fakülte/MYO tablosu ile diğer öğrenci bilgilerinin bulunduğu tablolar birleştirmiştir. Öğrencilerin yıl sonunda aldığı notlarda düzenlemeler yapılmış ve bu düzenlemeler “Öğrenci Başarı Durumu” sütunu oluşturularak bu sütuna yazılmıştır (öğrencinin notu 70 ve üzeriyse “Başarılı”, 69 ve aşağıdaysa “Başarısız”). Microsoft Access veritabanı programında yapılan değişikliklerden sonra tekrar veriler Microsoft Excell formatına aktarılmıştır.

5. Veri setinde bir öğrencinin derse katılma durumu birden fazla bulunduğu için bu satırlar teke indirgenmiştir. Aynı zamanda öğrencinin derse katılma durumu (Asenkron, Senkron ve Hibrit) veri setine bir özellik olarak eklenmiştir. (Örneğin bir öğrenci bir derse her hafta girmesinden dolayı birden fazla satır oluşmaktadır. Bu durumu tek bir satıra alarak derse katılma yöntemi belirlenmiştir. Öğrenci bütün derslerini canlı izlediye Senkron, dersleri kayıttan izlediye Asenkron ve dersleri hem canlı hem de kayıttan izlediye Hibrit olarak derse katılma yöntemleri özelliği olarak veri setine eklenmiştir.)
6. Fakülte/MYO, Bölüm Adı, Ders Adı, Derse Katılma Yöntemi ve Öğrenci Başarı Durumu sütunları kalacak şekilde verilerdeki diğer sütunların silme işlemi yapılarak VM'nin veri hazırlama süreçleri tamamlanmıştır.
7. Doğru algoritmaların seçimi için veriler RapidMiner programının içerisine aktarımı yapılmış ve varsayılan algoritmalar üzerinde otomatik olarak geleceğe yönelik tahminlerde bulunulmuştur. Elde edilen sonuçlara göre Derin Öğrenme, Naive Bayes, GBT ve Lojistik Regresyon algoritmaları çalışma için seçilmiştir. Otomatize edilmiş modellerin içerisinde olmayan ancak literatürde bu tip veriler üzerindeki başarısını ispat etmiş olan k-NN algoritması da karşılaştırılacak algoritmalar arasına dahil edilmiştir.
8. Öğrencinin başarı durumunda derse katılma yönteminin (yani Senkron, Asenkron ve Hibrit) etkisini tespit etmek için, tahmin sürecinde “Öğrencinin Başarı Durumu” seçilmiştir.
9. Öğrencinin başarı durumunu etkileyen etmenler ise derse katılma yöntemi, Fakülte-MYO, bölüm adı ve ders adı alanları özellik olarak seçilmiştir.

Bu adımlar sonucunda elde edilen bulgular;

- Sonuç olarak ayrıntılı inceleme için seçilen 5 algoritmanın tahminlemede en başarılı model %73,50'lik doğruluk oranı ile Lojistik Regresyon çıkmıştır.
- Lojistik Regresyon modelinde en başarılı oran, veri setinde %80 eğitim ve %20 tahminde bulunma oranları ile elde edilmiştir. Çizelge 3.17'de farklı oranlarda kullanılan modellerin başarı yüzdeleri detaylı olarak verilmiştir.

4.2. ÖNERİLER

KAYUZEM verilerinden yola çıkılarak ortaya konulan tahminler alan literatürü için kanıt niteliği taşımakla beraber; öğrencilerin başarı oranlarının arttırılmasına yönelik önerilerde bulunulmuştur.

Öğrencilerin başarı durumunun daha iyi olması için Lojistik Regresyon algoritmasıyla yapılan tahmin ile Çizelge 3.22'deki karşılaştırmaların sonucunda gelecekte kullanılmaları tavsiye edilen en iyi derse katılma yöntemleri sırasıyla Hibrit, Asenkron ve Senkron olduğu görülmüştür. Bu sebeple hibrit olarak uzaktan eğitim yapılması başarı oranının arttırılması için tavsiye edilebilir.

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 bilgisayarını vb. öznitelikler de çalışmaya dahil edilerek tahminler yapılabilir.

5. KAYNAKLAR

- [1] A. Akçay, "Bilgi ve belge yönetiminde veri madenciliği," Yüksek lisans tezi, Bilgi ve Belge Yönetimi, İstanbul Üniversitesi, İstanbul, Türkiye, 2014.
- [2] S. Savaş, N. Topaloğlu ve M. Yılmaz "Veri madenciliği ve Türkiye'deki uygulama örnekleri," *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi*, s. 21, ss. 1-23, 2012.
- [3] K. Akgün ve M. Bulut Özek , "Eğitsel Veri Madenciliği Yöntemi İle İlgili Yapılmış Çalışmaların İncelenmesi: İçerik Analizi", *Uluslararası Eğitim Bilim ve Teknoloji Dergisi*, c. 6, sayı. 3, ss. 197-213, Ara. 2020.
- [4] N. B. Sebik ve H. İ. Bülbül, "Veri madenciliği modellerinin akciğer kanseri veri seti üzerinde başarılarının incelenmesi," *TÜBAV Bilim Dergisi*, c. 11, s. 3, ss. 1-7, 2018.
- [5] L. Mohan, S. Jain, P. Suyal and A. Kumar, "Data mining Classification Techniques for Intrusion Detection System," *2020 12th International Conference on Computational Intelligence and Communication Networks (CICN)*, 2020, ss. 351-355.
- [6] H. DrissiTouzani, S. Faquir and A. Yahyaouy, "Data mining techniques to analyze traffic accidents data: Case application in Morocco," *2020 Fourth International Conference On Intelligent Computing in Data Sciences (ICDS)*, 2020, ss. 1-4.
- [7] E. Y. Tozak, "Veri madenciliği programları kullanılarak bir tekstil firmasının satış verilerinin değerlendirilmesi," Yüksek lisans tezi, Endüstri Mühendisliği, Karabük Üniversitesi, Karabük, Türkiye, 2021.
- [8] S. Topal, "Veri madenciliği tekniklerinin bitcoin piyasası öngörüsünde kullanımı," Yüksek lisans tezi, İşletme, Giresun Üniversitesi, Giresun, Türkiye, 2021.
- [9] B. Bakırarar, "Veri madenciliğinde hibrit model yaklaşımı," Doktora tezi, Biyoistatistik, Ankara Üniversitesi, Ankara, Türkiye, 2021.
- [10] B. Şenocak, "Data mining and machine learning applications in bank audit," M.S. thesis, Artificial Intelligence, Bahcesehir University, Istanbul, Turkey, 2021.
- [11] P. Dinçoğlu, "Perakende sektöründe veri madenciliği teknikleri ile satış tahmini," Yüksek lisans tezi, Bilgisayar Mühendisliği, Maltepe Üniversitesi, İstanbul, Türkiye, 2022.
- [12] O. Aydur, "Tel elektro erozyon işleme ile kesilen alüminyum alaşımlarının yüzey pürüzlülük değerlerinin makine öğrenmesi yöntemleri ile tahmin edilmesi," Yüksek lisans tezi, Yazılım Mühendisliği, Fırat Üniversitesi, Elazığ, Türkiye, 2022.
- [13] World Health Organization. (2019, Aralık 31). Rolling updates on coronavirus disease (COVID-19). World Health Organization [Online]. Erişim: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/events-as-they-happen>.

- [14] T.C. Sağlık Bakanlığı. (2020, Temmuz 02). Pandemi. T.C. Sağlık Bakanlığı [Online]. Erişim: <https://covid19.saglik.gov.tr/TR-66494/pandemi.html>.
- [15] T.C. Yükseköğretim Kurulu. (2020, 13 Mart). Koronavirüs (Covid-19) Bilgilendirme Notu: 1. T.C. Yükseköğretim Kurulu [Online]. Erişim: <https://covid19.yok.gov.tr/Documents/alinan-kararlar/02-coronavirus-bilgilendirme-notu-1.pdf>.
- [16] T.C. Yükseköğretim Kurulu. (2020, 6 Mart). Basın Açıklaması Yükseköğretim Kurulu Başkanı Prof. Dr. M. A. Yekta Saraç. T.C. Yükseköğretim Kurulu [Online]. Erişim: <https://covid19.yok.gov.tr/Documents/alinan-kararlar/03-uzaktan-egitime-iliskin-alinan-karar.pdf>.
- [17] T.C. Yükseköğretim Kurulu. (2020, 26 Mart). Basın Açıklaması Yükseköğretim Kurulu Başkanı Prof. Dr. M. A. Yekta Saraç. T.C. Yükseköğretim Kurulu [Online]. Erişim: <https://covid19.yok.gov.tr/Documents/alinan-kararlar/04-uzaktan-egitim-ve-yks-ertelenmesine-iliskin.pdf>.
- [18] T.C. Yükseköğretim Kurulu. (2020, 4 Haziran). YÖK'ten Küresel Salgın ile Mücadele Sürecinde Yeni Düzenlemeler-II. T.C. Yükseköğretim Kurulu [Online]. Erişim: <https://covid19.yok.gov.tr/Documents/alinan-kararlar/23-kuresel-salgini-ile-mucadele-gerceklesen-yapisal-duzenlemeler-2.pdf>.
- [19] T.C. Yükseköğretim Kurulu. (2020, 3 Mayıs). Üniversitelerimizde Küresel Salgın Sürecinde Uzaktan Öğretim Uygulamaları Anket Sonuçları. T.C. Yükseköğretim Kurulu [Online]. Erişim: <https://covid19.yok.gov.tr/Documents/AnaSayfa/uzaktan-ogretim-anketi.pdf>.
- [20] T.C. Yükseköğretim Kurulu. (2020, 30 Temmuz). Küresel Salgında Yeni Normalleşme Süreci. T.C. Yükseköğretim Kurulu [Online]. Erişim: <https://www.yok.gov.tr/Documents/Yayinlar/Yayinlarimiz/2020/kuresel-salginda-yeni-normallesme-sureci-2020.pdf>.
- [21] T.C. Yükseköğretim Kurulu. (2021, 3 Eylül). Küresel Salgında Eğitim ve Öğretim Süreçlerine Yönelik Uygulamalar Rehberi. T.C. Yükseköğretim Kurulu [Online]. Erişim: <https://www.yok.gov.tr/Documents/Yayinlar/Yayinlarimiz/2021/K%C3%BCrele1%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>.
- [22] Ö. Özbay ve H. Ersoy, "Öğrenme yönetim sistemi üzerindeki öğrenci hareketliliğinin veri madenciliği yöntemleriyle analizi," *Gazi Üniversitesi Gazi Eğitim Fakültesi Dergisi*, c. 37, s. 2, ss. 523-558, Ağu. 2017.
- [23] 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.
- [24] O. Gushchina, and A. Ochepovsky, "Data mining for the e-learning risk management," *Anadolu University Turkish Online Journal of Distance Education*, c. 20, s. 3, ss. 181-196, 2019.
- [25] M. Blagojević and Ž. Micić, "A web-based intelligent report e-learning system using data mining techniques," *Computers and Electrical Engineering*, c. 39, s. 2, ss. 465 – 474, 2013.

- [26] 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)*, Ecuador, 2017, ss. 1-5, doi: 10.1109/ETCM.2017.8247553.
- [27] 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)*, Japan, 2019, ss. 350-355, doi: 10.1109/DASC/PiCom/CBDCCom/CyberSciTech.2019.00071.
- [28] W. F. W. Yaacob, S. Nasir, "Supervised data mining approach for predicting student performance," *Indonesian Journal of Electrical Engineering and Computer Science*, c. 16, s. 3, ss. 1584 – 1592, 2019.
- [29] 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*, c. 12, s. 2, ss. 2-14, 2020.
- [30] 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)*, India, 2021, ss. 442-447, doi: 10.1109/ICCCIS51004.2021.9397145.
- [31] H. Mohammad, "Application of decision tree algorithm for predicting students' performance via online learning during coronavirus pandemic," *Journal of Theoretical and Applied Information Technology*, c. 99, s. 19, ss. 4546 – 4556, 2021.
- [32] 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*, ss. 1 – 15, 2021.
- [33] L. Masangu, A. Jadhav, R. Ajoodha, "Predicting student academic performance using data mining techniques," *Advances in Science, Technology and Engineering Systems Journal*, c. 6, s. 1, ss. 153 – 163, 2021.
- [34] 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 (ICCM)*, France, 2021, ss. 9-15.
- [35] I. L. H. Alsammak, A. H. Mohammed, I. S. Nasir, "E-learning and covid-19: Predicting student academic performance using data mining algorithms," *Webology*, c. 19, s. 1, ss. 3419 – 3432, 2022.
- [36] Ö. Şen ve G. Kızılcıoğlu, "Covid-19 pandemi sürecinde üniversite öğrencilerinin ve akademisyenlerin uzaktan eğitime yönelik görüşlerinin belirlenmesi", *International Journal of 3D Printing Technologies and Digital Industry*, c. 4, s. 3, ss. 239-252, 2020.
- [37] 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, s. 2, ss. 59-67, 2020.
- [38] Exastax. (2018, 31 Ocak). Veri Madenciliği Tarihi. Exastax [Online]. Erişim:

<https://www.exastax.com.tr/veri-analitigi/veri-madenciliginin-tarihi>.

- [39] K. C. Tan, Q. Yu and J. H. Ang, " A coevolutionary algorithm for rules discovery in data mining," *International Journal of Systems Science*, c. 37, s. 12, ss. 835-864, 2006.
- [40] J. Han, M.Kamber, and J. Pei, *Data Mining Concepts and Techniques*, 3. baskı, Waltham, USA: Elsevier, 2011.
- [41] K. Yu, T. Yuan and Y. Li, "Application of Data Mining Technology in Sports Data Analysis in Colleges and Universities," *2021 International Conference on Information Technology and Contemporary Sports (TCS)*, 2021, ss. 329-332.
- [42] Z. Wang et al., "A Prediction Model of Clinical Diagnosis by The Combination of Traditional Chinese and Western Medicine Based on Data Mining," *2021 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS)*, China, 2021, ss. 751-754.
- [43] M. Rahman, M. I. Ahmed and M. S. Hossain, "Analysis of Student's Achievement through Educational Data Mining," *2021 International Conference on Information Systems and Advanced Technologies (ICISAT)*, Algeria, 2021, ss. 1-5.
- [44] H. Meng, J. Song, W. Hong and S. Li, "Knowledge Discovery from Categorical Data Based on Structured Partial Ordered Attribute Diagram," *2018 Eighth International Conference on Instrumentation & Measurement, Computer, Communication and Control (IMCCC)*, China, 2018, ss. 695-699.
- [45] Z. Yun, "The study of CDM-BSC-based data mining driven fishbone applied for data processing," *2015 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*, China, 2015, ss. 1-5.
- [46] Y. Özkan, *Veri Madencilği Yöntemleri*, 2. baskı, İstanbul, Türkiye: Papatya Yayıncılık Eğitim, 2013, ss.37.
- [47] X. Lu, "Weakly supervised learning with stochastic supervision and knowledge transfer," Ph. D. dissertation, Department of Statistical Science, University College London, London, England, 2020.
- [48] S. K. Masum, "Forecasting from physiological time series through supervised learning," Ph. D. dissertation, Department of ETE, Portsmouth University, Portsmouth, England, 2019.
- [49] H. B. Demir. (2021, 23 Şubat). Makine Öğrenmesi. Mekteb-i Endüstri [Online]. Erişim: <https://mektebiendustri.com/index.php/2021/02/23/makine-ogrenmesi-machine-learning>.
- [50] Oracle. (2022, 20 Nisan). 5 Classification. Oracle [Online]. Erişim: https://docs.oracle.com/cd/E18283_01/datamine.112/e16808/classify.htm#:~:text=Classification%20is%20a%20data%20mining,medium%2C%20or%20high%20credit%20risks.
- [51] A. Pandey and R. Chhikara, "Analysis of Life Expectancy using various Regression Techniques," *2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)*, India, 2020, ss. 209-213.
- [52] R. H. Candidate, "Unsupervised image feature learning for convolutional neural networks," Ph. D. dissertation, Department of Electrical and Electronic

- Engineering, Manchester University, Manchester, England, 2019.
- [53] C. Nash, "Unsupervised learning with neural latent variable models," Ph. D. dissertation, Institute for Adaptive and Neural Computation, Edinburgh University, Edinburgh, Scotland, 2020.
- [54] Z. Liu, "Unsupervised learning with graph theoretical algorithms and its applications to transcriptomic data analysis," Ph. D. dissertation, Anatomy and Genetics, Imperial College London, London, England, 2018.
- [55] K. P. Sinaga and M. Yang, "Unsupervised K-Means Clustering Algorithm," *IEEE Access*, c. 8, ss. 80716-80727, 2020.
- [56] N. Y. Şimşek. (2019, 24 Şubat). Derin Öğrenme (Deep Learning) Nedir ve Nasıl Çalışır? [Online]. Erişim: <https://nyilmazsimsek.medium.com/derin-%C3%B6%C4%9Frenme-deep-learning-nedir-ve-nas%C4%B1-%C3%A7al%C4%B1%C5%9F%C4%B1r-2d7f5850782>.
- [57] L. W. Sommer, *Deep Learning based Vehicle Detection in Aerial Imagery*, Karlsruhe, Germany: KIT Scientific Publishing, 2022.
- [58] M. Lennox, "Deep learning of proteomics data," Ph. D. dissertation, Department for the Economy, Queen's University Belfast, Belfast, Northern Ireland, 2021.
- [59] E. Şanlıtürk, "Makine öğrenme algoritmalarıyla hatalı ürün tahmini," Yüksek lisans tezi, İşletme Mühendisliği, Fen Bilimleri Enstitüsü, İstanbul Teknik Üniversitesi, İstanbul, Türkiye, 2018.
- [60] 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.
- [61] W. F. W. Yaacob, S. Nasir, "Supervised data mining approach for predicting student performance," *Indonesian Journal of Electrical Engineering and Computer Science*, c. 16, s. 3, ss. 1584 – 1592, 2019.
- [62] L. Mandal and N. D. Jana, "A Comparative Study of Naive Bayes and k-NN Algorithm for Multi-class Drug Molecule Classification," *2019 IEEE 16th India Council International Conference (INDICON)*, India, 2019, ss. 1-4.
- [63] S. Kırışoğlu ve A. Yakupoğlu, "Veri madenciliği ile üniversite bilişim teknik servis hizmetleri analizi", *Düzce Üniversitesi Bilim ve Teknoloji Dergisi*, c. 8, s. 1, ss. 326-333, 2020, doi:10.29130/dubited.593830.
- [64] K. Kotan, "Comparative analysis of classification techniques for network anomalies management," M.S. thesis, Electronics And Computer Engineering, Hasan Kalyoncu University, Gaziantep, Turkey, 2019.
- [65] W. Zhao, Y. Lv, X. Guo and J. Huo, "An Investigation on Early Fault Diagnosis Based on Naive Bayes Model," *2022 7th International Conference on Control and Robotics Engineering (ICCRE)*, 2022, China, ss. 32-36.
- [66] M. Saritas and A. Yasar, "Performance analysis of ann and naive bayes classification algorithm for data classification," *IJISAE*, c. 7, s. 2, ss. 88-91, 2019.
- [67] 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*, c. 27, s. 6, ss. 316 – 334, 2020.

- [68] L. Jinshu, W. Yijiang, W. Ganjun, P. Xiaoasheng, L. Taiwei and J. Yuhang, "Gradient Boosting Decision Tree and Random Forest Based Partial Discharge Pattern Recognition of HV Cable," *2018 China International Conference on Electricity Distribution (CICED)*, United Kingdom, 2018, ss. 327-331.
- [69] U. Çelik, E. Akçetin ve M. Gök, *Editörler, RapiMiner ile Uygulamalı Veri Madenciliği*, 1. baskı, İstanbul, Türkiye: Pusula 20 Teknoloji ve Yayıncılık A.Ş., 2017, böl. 6, ss. 81-94.
- [70] S. Sperandei, "Understanding logistic regression analysis," *Biochemia Medica*, c. 24, s. 1, ss.12-18, 2014.
- [71] B. Aksoy ve D. Boztosun, "Diskriminant ve lojistik regresyon yöntemleri kullanılarak finansal başarısızlık tahmini: BIST imalat sektörü örneği," *İstanbul Okan Üniversitesi Finans Politik ve Ekonomik Yorumlar*, s. 646, ss. 9-32, 2018.
- [72] C. Erden, *Python ile Veri Madenciliği*, 1. baskı, İstanbul, Türkiye: KODLAB Yayın Dağıtım Yazılım Ltd. Şti., 2021, böl. 8-9, ss. 185-225.
- [73] 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," *IEEE Access*, c. 8, ss. 90847-90861, 2020.
- [74] A. Amidi and S. Amidi. (2021, 27 Aralık). Machine Learning tips and tricks cheatsheet. Stanford University [Online]. Erişim: <https://stanford.edu/~shervine/teaching/cs-229/cheatsheet-machine-learning-tips-and-tricks>.
- [75] A. Kızrak. (2019, 9 Mayıs). Comparison of Activation Functions for Deep Neural Networks [Online]. Towards Data Science. Erişim: <https://towardsdatascience.com/comparison-of-activation-functions-for-deep-neural-networks-706ac4284c8a>.
- [76] RapidMiner. (2022, 13 Nisan). Deep Learning [Online]. Erişim: https://docs.rapidminer.com/latest/studio/operators/modeling/predictive/neural_networks/deep_learning.html.

6. EKLER

6.1. EK 1: ARAŞTIRMA İZİN YAZISI

Evrak Tarih ve Sayı: 11.01.2021-2642



T.C.
KAYSERİ ÜNİVERSİTESİ REKTÖRLÜĞÜ
Öğrenci İşleri Daire Başkanlığı

Sayı : E-13303056-302.08.01-313
Konu : Araştırma İzin İsteği (Mehmet
YILDIRIM)

DÜZCE ÜNİVERSİTESİ REKTÖRLÜĞÜNE
Düzce Üniversitesi Konuralp Yerleşkesi Konuralp
Merkez/Düzce 81620

İlgi : 29.12.2020 tarihli ve 20247 sayılı yazınız.

İlgide kayıtlı yazınıza istinaden Fen Bilimleri Enstitünüz Elektrik-Elektronik ve Bilgisayar Mühendisliği Anabilim Dalı Yüksek Lisans Programı öğrenciniz Mehmet YILDIRIM'ın Doktor Öğretim Üyesi Serdar KIRIŞOĞLU'nun tez danışmanlığında "Veri Madenciliği ile Pandemi Döneminde Uzaktan Eğitimde Senkron Asenkron Yapılmış Derslerde Öğrenci Performans Analizi" isimli çalışması için, Üniversitemiz Uzaktan Eğitim Uygulama ve Araştırma Merkezi (KAYUZEM)'in 06.04.2020 ile 15.08.2020 tarihleri arasındaki, kullanıcı adı, öğrenci adı ve soyadı alanlarını (6698 sayılı Kişisel Verilerin Korunması Kanunu kapsamında) maskeleyerek kullanması ve bu çalışmaya ilişkin olarak Rektörlüğümüze bağlı Uygulamalı Bilimler Fakültesi, Mühendislik Mimarlık ve Tasarım Fakültesi, Sağlık Bilimleri Fakültesi, Develi Sosyal ve Beşeri Bilimler Fakültesi, Develi İslami İlimler Fakültesi ve Kayseri Teknik Bilimler Meslek Yüksek Okuluna ait verilerin kullanılması Üniversitemiz tarafından uygun görülmüştür.

Bilgilerinizi ve gereğini arz ederim.

Prof.Dr. Semra AKSOYLU
Rektör a.
Rektör Yardımcısı

Bu belge, güvenli elektronik imza ile imzalanmıştır.

Belge Doğrulama Kodu :BE8R5SKD8 Pin Kodu :42702

Belge Takip Adresi :
[https://cbys.kayseri.edu.tr/envision/validate_doc.aspx?](https://cbys.kayseri.edu.tr/envision/validate_doc.aspx?V=BE8A5SKN4)
V=BE8A5SKN4

Adres:Kayseri Üniversitesi Rektörlüğü 15 Temmuz Yerleşkesi Kümeveler No:5 Talas/KAYSERİ
Telefon:0352 504 38 38 Faks:0352 504 38 37
e-Posta:ogrenci@kayseri.edu.tr Elektronik Ağ:https://ogrisl.kayseri.edu.tr/
Kep Adresi:kayseriuni@hs01.kep.tr

Bilgi için: Fatma Kurt
Unvanı: Büro Personeli



Bu belge, 5070 sayılı Elektronik İmza Kanununa göre Güvenli Elektronik İmza ile imzalanmıştır.
Evrak sorgulaması <https://www.turkiye.gov.tr/duzce-universitesi-ebys?V=BE6L59U8D> adresinden yapılabilir. (PIN:03102)

6.2. EK 2: KVKK TUTANAĐI

29.01.2021

TUTANAKTIR

Düzce Üniversitesi Elektrik Elektronik ve Bilgisayar MühendisliĐi Yüksek Lisans programı öğrencisi Mehmet Yıldırım'a ekte verilen izin doğrultusunda, Kayseri Üniversitesi Uzaktan Eğitim Uygulama ve Araştırma Merkezinin 06.04.2020 ile 15.08.2020 tarihleri arasındaki Uygulamalı Bilimler Fakültesi, Mühendislik Mimarlık ve Tasarım Fakültesi, Sağlık Bilimleri Fakültesi, Develi Sosyal ve Beşeri Bilimler Fakültesi, Develi İslami İlimler Fakültesi ve Kayseri Teknik Bilimler Meslek Yüksek Okuluna ait öğrenci verileri Microsoft Excel formatında KVKK uygun olarak teslim edilmiştir.

Ekler _____ :

1. Rektörlük izin yazısı

Teslim Alan
Mehmet YILDIRIM

Teslim Eden
Öğr. Gör. Mustafa KALAY

6.3. EK 3: KAYSERİ ÜNİVERSİTESİ ETİK KURUL BAŞKANLIĞI İZİNİ

BAŞVURU NO: 7	
KAYSERİ ÜNİVERSİTESİ ETİK KURULU PROJE ONAY FORMU	
Projenin Adı	Veri Madenciliği ile Pandemi Döneminde Uzaktan Eğitimde Senkron Asenkron Yapılmış Derslerde Öğrenci Performans Analizi
Projenin Niteliği	Bireysel Araştırma
Proje Araştırmacıları	Mehmet YILDIRIM
Sorumlu Araştırmacının Haberleşme Bilgileri	Cebeci Mah. Akat Sok. No:16/6 Çankaya Ankara

KARAR:

Etik Kurulumuza başvuran **Mehmet YILDIRIM'** ın " **Veri Madenciliği ile Pandemi Döneminde Uzaktan Eğitimde Senkron Asenkron Yapılmış Derslerde Öğrenci Performans Analizi** " adlı Çalışması değerlendirilerek aşağıdaki sonuca ulaşılmıştır.

Proje etik açıdan uygun bulunmuştur.

Projenin etik açıdan geliştirilmesi gerekmektedir.

Proje etik açıdan uygun bulunmamıştır.

29.01.2021

ADI SOYADI	İMZA
Etik Kurul Başkanı	Doç. Dr. Mürsel KARABACAK
Etik Kurul Başkan Yrd.	Doç. Dr. Hüseyin BENLİ
Üye	Doç. Dr. Azize ESMERAY
Üye	Dr. Öğr. Üyesi Ebru SÖNMEZ KARAPINAR
Üye	Dr. Öğr. Üyesi Nihat YILMAZ
Üye	Dr. Öğr. Üyesi Ali DURMUŞ
Üye	Dr. Öğr. Üyesi Elif YILMAZ

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı : Mehmet YILDIRIM

Yabancı Dili : İngilizce

ÖĞRENİM DURUMU

Derece	Alan	Okul/Üniversite	Mezuniyet Yılı
Y. Lisans	Elektrik-Elektronik ve Bilgisayar Müh.	Düzce Üniversitesi	2022
Lisans	Bilgisayar ve Öğretim Teknolojileri Öğretmenliği	Hoca Ahmet Yesevi Üniversitesi	2015
Lise	Bilgisayar Donanım	Kuşadası Anadolu Teknik Lisesi	2007

YAYINLAR

M. Yıldırım, ve S. Kırışođlu, “Uzaktan eğitim verileri üzerinden derin öğrenme veri madenciliđi modeli ile performans analizi,” 4. Uluslararası Mühendislik Arařtırmaları Sempozyumu, Düzce, Türkiye, 2022, ss. 181-186.

S. Kırışođlu, ve M. Yıldırım, “Pandemi sürecinde uzaktan eğitimde senkron, asenkron ve hibrit yapılmış derslerde veri madenciliđi ile öğrenci performans analizi,” Düzce Üniversitesi Bilim ve Teknoloji Dergisi, 2022, (Yayınlanmamış).