



T.C.

BARTIN ÜNİVERSİTESİ

LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

BİLİŞİM SİSTEMLERİ VE TEKNOLOJİLERİ ANABİLİM DALI

YÜKSEK LİSANS TEZİ

PISA 2022 ÖĞRENCİ ANKET VERİLERİ İLE OKUMA BAŞARISI
VE MATEMATİK BAŞARISININ TAHMİN EDİLMESİ

SEHER YILMAZ

DANIŞMAN

PROF. DR. FATMA GİZEM KARAOĞLAN YILMAZ

BARTIN-2024



T.C.

BARTIN ÜNİVERSİTESİ

LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

BİLİŞİM SİSTEMLERİ VE TEKNOLOJİLERİ ANABİLİM DALI

**PISA 2022 ÖĞRENCİ ANKET VERİLERİ İLE OKUMA BAŞARISI VE
MATEMATİK BAŞARISININ TAHMİN EDİLMESİ**

YÜKSEK LİSANS TEZİ

SEHER YILMAZ

JÜRİ ÜYELERİ

Danışman : Prof. Dr. Fatma Gizem KARAOĞLAN YILMAZ

Üye : Doç. Dr. Yusuf Ziya OLPAK

Üye : Dr. Öğr. Üyesi Şeyma ÇAĞLAR ÖZHAN

BARTIN-2024

BEYANNAME

Bartın Üniversitesi Lisansüstü Eğitim Enstitüsü tez yazım kılavuzuna göre Prof. Dr. Fatma Gizem KARAOĞLAN YILMAZ danışmanlığında hazırlamış olduğum “PISA 2022 ÖĞRENCİ ANKET VERİLERİ İLE OKUMA BAŞARISI VE MATEMATİK BAŞARISININ TAHMİN EDİLMESİ” başlıklı yüksek lisans tezimin bilimsel etik değerlere ve kurallara uygun, özgün bir çalışma olduğunu, aksinin tespit edilmesi halinde her türlü yasal yaptırımını kabul edeceğimi beyan ederim.

16.07.2024

Seher YILMAZ



ÖN SÖZ

Çalışmam süresince her türlü yol gösterici olan, olumlu tavrıyla beni cesaretlendiren, beraber çalışmaktan çok mutluluk duyduğum değerli danışman hocam Fatma Gizem KARAOĞLAN YILMAZ'a teşekkürlerimi sunuyorum.

Çok çeşitli zorluklarla beni büyüten, eğitim sürecimde ve zor zamanlarımda hep yanımda olan değerli anneme ve babama, sen yaparsın diyerek beni her zaman cesaretlendiren sevgili eşime, sonsuz ilham ve mutluluk kaynağım biricik oğlum Ozan'ıma teşekkür ederim, varlığınıza minnettarım.

.

Seher YILMAZ

ÖZET

Yüksek Lisans Tezi

PISA 2022 ÖĞRENCİ ANKET VERİLERİ İLE OKUMA BAŞARISI VE MATEMATİK BAŞARISININ TAHMİN EDİLMESİ

Seher YILMAZ

Bartın Üniversitesi

Lisansüstü Eğitim Enstitüsü

Bilişim Sistemleri ve Teknolojileri Anabilim Dalı

Tez Danışmanı: Prof. Dr. Fatma Gizem KARAOĞLAN YILMAZ

Bartın-2024, sayfa: 92

Yapay sinir ağları, insan beyninin öğrenme işlevini taklit eden ve öğrenme sürecini örneklerle gerçekleştiren bilgisayar sistemleridir. Bu ağlar, yapay sinir hücrelerinden oluşan birbirine bağlı işlem birimlerinden oluşur ve her bağlantının bir ağırlık değeri vardır. Ağın bilgisi bu ağırlık değerlerinde saklanır ve ağa yayılmıştır. Geleneksel hesaplama yöntemlerinden farklı olarak, yapay sinir ağları adaptif, eksik bilgiyle çalışabilen ve belirsizlik altında karar verebilen sistemlerdir. Çok katmanlı algılayıcı ağları, günümüzde yapay sinir ağlarının en yaygın kullanılan modelidir. Bu ağlar, özellikle mühendislik problemlerinin çözümünde geçerli sonuçlar verebilmektedir. PISA araştırması, OECD tarafından üç yılda bir düzenlenir ve 15 yaşındaki öğrencilerin temel bilgi ve becerilerini ölçer. Türkiye, 2003 yılından beri PISA'ya katılmaktadır ve 2022'de matematik okuryazarlığına odaklanılmıştır. PISA 2022 verilerine göre, öğrencilerin sosyo-ekonomik durumları, aile eğitim durumu, ebeveyn tutumları ile akademik başarıları arasındaki ilişki çeşitli anket soruları ve yanıtlar aracılığıyla incelenmektedir.

Yapılan çalışmanın amacı, yukarıda belirtilen sosyo-ekonomik durumun ve ebeveyn tutumları arasındaki farklılıkların öğrenci başarısını ne derecede tahmin edebildiğini ölçmek, ayrıca Rapid Miner derin öğrenme ağı ve Python çok katmanlı yapay derin öğrenme ağı

arasındaki tahmin başarısını karşılaştırmaktır. Bu amaçla OECD web sitesinde 5 Aralık 2023 tarihinde yayınlanan SPSS formatındaki veriler indirilmiş, öğrencilere uygulanan öğrenci anketleri ve ebeveynlere uygulanan ebeveyn anketleri incelenmiş, ekonomik konularla ilgili etiketleri ve ebeveyn tutumları ile ilgili etiketleri içeren değişkenler seçilmiştir. Çalışmanın bağımlı değişkenlerini oluşturan matematik okuryazarlığı ve okuma becerisi puanları, 10 adet makul okuma becerisi puanı ve 10 adet makul matematik okuryazarlığı puanı şeklinde yayınlanmıştır. Öncelikle yayınlanan 10 adet puanın ortalaması alınmış sonrasında ise Milli Eğitim Bakanlığı PISA 2024 raporunda belirtilen şekilde 0-6 arasında kodlanmıştır. Ayrıca seçilen bağımsız değişkenlerin öğrenci başarısını ne derecede ölçtüğünü görebilmek için ortalaması alınan matematik becerisi ve okuma becerisi puanları OECD ülkelerinin ortalama puan durumlarına göre (0-başarısız) ve (1-başarılı) olacak şekilde kodlanmıştır.

Seçilen bağımsız değişkenler, (0-6) ve (0-1) arasında kodlanan bağımlı değişkenler matematik becerisi puanlarını ve okuma becerisi puanlarını tahmin etmesi için, Rapid Miner programında ve Python programlama dili kullanılarak oluşturulan yapay sinir ağına aktarılmış ve bağımsız değişkenlerin bağımlı değişkenleri tahmin etme oranı tespit edilmiştir. Yapılan analiz sonucunda ekonomik değişkenlerin matematik becerisi ve okuma becerisi puanlarını tahmin başarıları %39-%43, aile eğitim durumu ve ebeveyn tutumları değişkenlerinin matematik becerisi ve okuma becerisi puanlarını tahmin başarıları %38-%42 olarak bulunmuştur. Ekonomik değişkenlerin öğrencilerin matematik becerisi ve okuma becerisi başarı durumlarını tahmin etme oranı %77-%78, aile eğitim durumu ve ebeveyn tutumları değişkenlerinin matematik becerisi ve okuma becerisi başarı durumları tahmin etme oranı %78-%77 olarak bulunmuştur. Bu bağlamda ekonomik unsurların ve ebeveyn tutumlarının öğrenci başarısının önemli bir yordayıcısı olduğu bulunmuştur. Ayrıca Rapid Miner programında oluşturulan derin öğrenme ağı Python programı ve Keras kütüphanesiyle oluşturulan derin öğrenme ağına oranla %10 oranında daha başarılı tahminler yapmıştır.

Anahtar Kelimeler: Derin öğrenme, Rapid Miner, PISA 2022, yapay sinir ağları.

ABSTRACT

M.Sc. Thesis

PREDICTING READING ACHIEVEMENT AND MATHEMATICS ACHIEVEMENT WITH PISA 2022 STUDENT SURVEY DATA

Seher YILMAZ

Bartın University

Graduate School

Department of Information Systems and Technologies

Thesis Advisor: Prof. Dr. Fatma Gizem KARAOĞLAN YILMAZ

Bartın-2024, pp: 92

Artificial neural networks are computer systems that mimic the learning function of the human brain and perform the learning process with examples. These networks consist of interconnected processing units consisting of artificial nerve cells, and each connection has a weight value. The network's information is stored in these weight values and spread throughout the network. Unlike traditional computing methods, artificial neural networks are adaptive systems that can work with incomplete information and make decisions under uncertainty. Multilayer perceptron networks are the most widely used model of artificial neural networks today. These networks can provide valid results, especially in solving engineering problems. The PISA survey is organized by the OECD every three years and measures the basic knowledge and skills of 15-year-old students. Turkey has been participating in PISA since 2003, and the focus in 2022 is on mathematical literacy. According to PISA 2022 data, the relationship between students' socioeconomic status, family education status, parental attitudes and academic achievement is examined through various survey questions and answers.

The aim of the study is to measure the extent to which the differences between the socioeconomic status and parental attitudes mentioned above can predict student success,

and also to compare the prediction success between the Rapid Miner deep learning network and the Python multilayer artificial deep learning network. For this purpose, the SPSS format data published on the OECD website on December 5, 2023 were downloaded, the student surveys applied to students and the parent surveys applied to parents were examined, and variables containing labels related to economic issues and labels related to parental attitudes were selected. The mathematical literacy and reading skill scores that constitute the dependent variables of the study were published as 10 reasonable reading skill scores and 10 reasonable mathematical literacy scores. First, the average of the 10 published scores was taken and then coded between 0-6 as specified in the Ministry of National Education PISA 2024 report. In addition, in order to see to what extent the selected independent variables measure student success, the averaged mathematical skill and reading skill scores were coded as (0-unsuccessful) and (1-successful) according to the average scores of the OECD countries.

The selected independent variables, dependent variables coded between (0-6) and (0-1), were transferred to the artificial neural network created in the Rapid Miner program and Python programming language to predict the scores of mathematics skills and reading skills, and the rate of independent variables predicting the dependent variables was determined. As a result of the analysis, the success rate of economic variables in predicting mathematics skills and reading skills scores was found to be 39%-43%, and the success rate of family education status and parental attitudes variables in predicting mathematics skills and reading skills scores was found to be 38%-42%. The rate of economic variables predicting students' mathematics skills and reading skills success was found to be 77%-78%, and the rate of family education status and parental attitudes variables predicting mathematics skills and reading skills success was found to be 78%-77%. In this context, it was found that economic elements and parental attitudes were important predictors of student success. In addition, the deep learning network created in the Rapid Miner program made 10% more successful predictions compared to the deep learning network created with the Python program and Keras library.

Keywords: Artificial neural networks, Deepth learning, PISA 2022, Rapid Miner.

İÇİNDEKİLER

BEYANNAME	ii
ÖN SÖZ	iii
ÖZET	iv
ABSTRACT	vi
İÇİNDEKİLER.....	viii
ŞEKİLLER DİZİNİ.....	xi
TABLolar DİZİNİ.....	xiii
EKLER DİZİNİ	xiv
SİMGELER VE KISALTMALAR DİZİNİ..	xv
1. GİRİŞ.....	1
2. LİTERATÜR ÖZETİ.....	5
2.1. PISA Nedir?.....	5
2.2. Yapay Zeka.....	5
2.3. Makine Öğrenmesi.....	7
2.3.1 Makine Öğrenmesi Teknikleri.....	10
2.3.1.1 Denetimli Öğrenme.....	11
2.3.1.2 Denetimsiz Öğrenme.....	12
2.3.1.3 Yarı Denetimli Öğrenme.....	13
2.3.1.4 Takviyeli Öğrenme.....	13
2.4. Yapay Sinir Ağları.....	14
2.4.1 Yapay Sinir Ağlarında Backpropagation (Geriye Yayılım).....	19
2.4.2 Yapay Sinir Ağları Mimarileri.....	21
2.4.2.1 Çok Katmanlı Algılayıcı ((Multi Layer Perceptron).....	21
2.4.2.2. Evrişimli Sinir Ağı (Convolutional Neural Networks)...	24
2.4.2.3 Tekrarlayan Sinir Ağları (Recurrent Neural Networks)	
Ve Uzun Kısa Süreli Bellek Tekrarlayan Sinir Ağı (Long	
Short-Term Memory Networks).....	25
2.4.2.4 Radyal Temel Fonksiyonu(Radyal Basis Function).....	27
2.4.3 Yapay Sinir Ağı Aktivasyon Fonksiyonları.....	28
2.4.3.1 Rectified Linear Unit (ReLU).....	29
2.4.3.2 Sigmoid Fonksiyonu.....	29

2.4.3.3 Tanh Fonksiyonu.....	29
2.4.3.4 SoftMax.....	30
2.4.4 Yapay Sinir Ağı Optimizasyon Fonksiyonları.....	30
2.4.4.1 Gradyan Tabanlı Optimize Ediciler.....	31
2.4.4.2 RMSProp (Kök Ortalama Kare Yayılmı).....	32
2.4.4.3 Adam (Uyarlanabilir Moment Tahmini).....	32
2.5. Derin Öğrenme.....	32
3. MATERYAL VE YÖNTEM.....	36
3.1. Araştırmanın Modeli.....	36
3.2. Katılımcılar.....	36
3.3. Veri Toplama Araçları.....	36
3.4. Veri Toplama Süreci.....	37
3.4.1 Çalışmanın Bağımsız Değişkenleri.....	39
3.4.2 Ekonomik Durumu Gösteren Değişkenlerdeki Ayırık Verilerin Temizlenmesi.....	40
3.4.3 Anne Ve Baba Eğitim Durumları Ve Ebeveyn Tutumları Gösteren Değişkenlerdeki Ayırık Verilerin Silinmesi.....	45
3.5. Ekonomik Değişkenlerin Rapid Miner Programına Aktarılması ve Modelin Oluşturulması.....	46
3.6. Ebeveyn Eğitim Durumları Ve Ebeveyn Tutumları ile İlgili Olan Değişkenlerin Rapid Miner Programına Aktarılması.....	49
3.7. Python ve Çok Katmanlı Yapay Sinir Ağları Kullanılarak PVMATH_KODLAMA, PVREAD_KODLAMA, PVMATH_1_2, PVREAD_1_2 Değişkenlerinin Tahmin Edilmesi.....	50
3.7.1 Kütüphanelerin İmport (Dahil) Edilmesi.....	50
3.7.2 Veri Setinin Yüklenmesi ve DataFrame'e (Çok Boyutlu Dizi) Dönüştürülmesi.....	51
3.7.3 Verilerin Test ve Eğitim Olarak Bölünmesi.....	53
3.7.4 Verilerin Kategorik Değişkenlere Dönüştürülmesi.....	53
3.7.5 Modelin Oluşturulması.....	54
3.7.6 Kodun Derlenmesi.....	59
3.7.7 Modelin Eğitilmesi.....	59
3.7.7.1 Epoch Sayısı.....	60

3.7.7.2 Batch Size.....	60
3.7.8 Eğitilen Modelin Test Edilmesi.....	61
4. BULGULAR VE TARTIŞMA.....	62
4.1. Rapid Miner Programı ile Ekonomik Değişkenler Kullanılarak PVMATH_KODLAMA ve PVREAD_KODLAMA Değişkenlerinin Tahmin Edilmesi.....	62
4.2. Rapid Miner Programı ile Ekonomik Değişkenler Kullanılarak PVMATH_1_ ve PVREAD_1_2 Değişkenlerinin Tahmin Edilmesi.....	63
4.3. Rapid Miner Programı ile Anne Baba Eğitim Durumu Ve Ebeveyn Tutumları Değişkenleri ile PVMATH_KODLAMA ve PVREAD_KODLAMA Değişkenlerinin Tahmin Edilmesi.....	64
4.4. Rapid Miner Programı ile Anne Baba Eğitim Durumu Ve Ebeveyn Tutumları Değişkenleri İle PVMATH_1_2 Ve PVREAD_1_2 Değişkenlerinin Tahmin Edilmesi.....	65
4.5. Python Ve Çok Katmanlı Yapay Sinir Ağları Kullanılarak Ekonomik Değişkenler İle PVMATH_KODLAMA ve PVREAD_KODLAMA Değişkenlerinin Tahmin Edilmesi.....	67
4.6. Python ve Çok Katmanlı Yapay Sinir Ağları Kullanılarak Ekonomik Değişkenler İle PVMATH_1_2 ve PVREAD_1_2 Değişkenlerinin Tahmin Edilmesi.....	67
4.7. Python ve Çok Katmanlı Yapay Sinir Ağları Kullanılarak Anne Baba Eğitim Durumu ve Ebeveyn Tutumları Değişkenleri İle PVMATH_KODLAMA ve PVREAD_KODLAMA Değişkenlerinin Tahmin Edilmesi.....	68
4.8. Python ve Çok Katmanlı Yapay Sinir Ağları Kullanılarak Anne Baba Eğitim Durumu Ve Ebeveyn Tutumları Değişkenleri ile PVMATH_1_2 ve PVREAD_1_2 Değişkenlerinin Tahmin Edilmesi.....	68
4.9. Rapid Miner İle Python Derin Öğrenme Ağının Performans Karşılaştırması.....	69
5. SONUÇLAR.....	72
6. ÖNERİLER.....	74
KAYNAKLAR.....	76
EKLER.....	82

ŞEKİLLER DİZİNİ

Şekil No	Sayfa No
2.1: Veri bilimi teknikleri	6
2.2: Klasik programlama (A) ve makine öğrenimi (B) paradigması.....	8
2.3: Makine öğrenmesi uygulama alanları.....	9
2.4: Makine öğrenmesi uygulamaları.....	10
2.5: Denetimli öğrenme.....	11
2.6: Yapay sinir ağı (a) ve bir nöron'un (b) temel şekli.....	14
2.7: Biyolojik sinir hücresi.....	15
2.8: Yapay sinir ağı modeli.....	16
2.9: İleri beslemeli yapay sinir ağı.....	17
2.10: Geri beslemeli yapay sinir ağı.....	17
2.11: Yapay bir nöron.....	18
2.12: Yapay sinir ağının mimarisi.....	19
2.13: Hatanın geriye yayılımının öğrenilmesindeki ağırlık düzeltmelerinin sırası.....	20
2.14: Yapay sinir ağları mimarileri.....	21
2.15: 2 Gizli katmandan oluşan çok katmanlı algılayıcı.....	22
2.16: Çok katmanlı algılayıcı için hata yüzeyi.....	23
2.17: LeNet-5 ağı.....	25
2.18: Standart tekrarlayan sigma hücresinin şeması.....	26
2.19: LSTM hücresi.....	26
2.20: RBF ağı yapısı.....	27
2.21: Basit sinir ağı ve derin sinir ağı.....	33
2.22: Basit evrişimli sinir ağı (CNN).....	34
3.1: PISA 2022' ye katılan ülkelerin matematik becerisi puan sıralamaları.....	38
3.2: PISA 2022' ye katılan ülkelerin okuma becerisi puan sıralamaları.....	39
3.3: ST250D07JA değişkeninin kodlanmış hali.....	43
3.4: Ekonomik değişkenler ile PVMATH_KODLAMA, PVREAD_KODLAMA, PVMATH_1_2, PVREAD_1_2 puanlarının tahmin edilmesi Rapid Miner modeli....	47
3.5: Ebeveyn eğitim durum ve ebeveyn tutumları değişkenleri ile PVMATH_KODLAMA, PVREAD_KODLAMA, PVMATH_1_2, PVREAD_1_2 puanlarının tahmin edilmesi Rapid Miner modeli.....	49
3.6: Veri setinde bulunan ilk 5 satırın görüntülenmesi.....	51

3.7: X çok boyutlu dizisi.....	52
3.8: Y çok boyutlu dizileri.....	52
3.9: Kategorilendirilmiş tahmin edilecek puanlar.....	54
3.10: Ekonomik değişkenlerinin matematik becerisi ve okuma becerisi puanını tahmin ederken oluşturulan Python ve YSA modeli (0-6 arası tahmin)	55
3.11: Ekonomik değişkenlerinin matematik becerisi ve okuma becerisi puanını tahmin ederken oluşturulan Python ve YSA modeli (0-1 arası tahmin)	56
3.12: Aile eğitim durumu ve ebeveyn tutumları değişkenlerinin matematik becerisi puanını tahmin ederken oluşturulan Python ve YSA modeli (0-6 arası tahmin)	57
3.13: Aile eğitim durumu ve ebeveyn tutumları değişkenlerinin okuma becerisi puanını tahmin ederken oluşturulan Python ve YSA modeli(0-6 arası tahmin).....	57
3.14: Aile eğitim durumu ve ebeveyn tutumları değişkenlerinin matematik becerisi puanını tahmin ederken oluşturulan Python ve YSA modeli (0-1 arası tahmin).....	58
3.15: Aile eğitim durumu ve ebeveyn tutumları değişkenlerinin okuma becerisi puanını tahmin ederken oluşturulan Python ve YSA modeli (0-1 arası tahmin)	58
3.16: Modelin eğitimi ve hiper parametreleri.....	59
4.1: Modelin sınıflandırma başarısı.....	67
4.2: Modelin performans tablosu.....	67
4.3: Modelin tahmin başarısı.....	68
4.4: Modelin tahmin başarısı.....	69

TABLÖLER DİZİNİ

Tablo	Sayfa
No	No
2.1: Geleneksel bilgi işlem ve YSA arasındaki yaklaşım farklılıkları	16
3.1: Ayrık veriler için değer etiket tablosu.....	40
3.2: Ekonomik değişkenleri içeren verileri bulunduran ülkeler ve PISA sıralamaları.....	41
3.3: Matematik becerisi puanı kodlaması.....	45
3.4: Okuma becerisi puanı kodlaması.....	45
3.5: PVMATH_1_2 ve PVREAD_1_2 değişkenlerinin oluşturulması.....	45
3.6: ISCED 2011 seviyeleri.....	45
3.7: Derin öğrenme hiper parametreleri.....	48
3.8: Modelin hiper parametreleri.....	50
3.9: Eğitim veri setinin train, validation ve test olarak ayrılması.....	60
4.1: Modelin performans tablosu.....	62
4.2: Modelin performans tablosu.....	63
4.3: Modelin performans tablosu.....	63
4.4: Modelin performans tablosu.....	64
4.5: Modelin performans tablosu.....	64
4.6: Modelin performans tablosu.....	65
4.7: Modelin performans tablosu.....	66
4.8: Modelin performans tablosu.....	66
4.9: Rapid Miner ve Python arasındaki başarı kıyaslaması (0-6 kategorili).....	70
4.10: Rapid Miner ve Python arasındaki başarı kıyaslaması (0-1 kategorili).....	70

EKLER DİZİNİ

Ek	Sayfa
<u>No</u>	<u>No</u>
EK 1. Analizlerde kullanılan ekonomik değişkenler.	82
EK 2. Analizlerde kullanılan aile eğitim durumu ve ebeveyn tutumları değişkenleri.	84
EK 3. Matematik okuryazarlığı yeterlik düzeylerinin özeti.	87
EK 4. Okuma becerileri yeterlik düzeylerinin özeti.	89
EK 5. ST250D07JA değişkenine ait veriler.	90



SİMGELER VE KISALTMALAR DİZİNİ

X	: Yapay sinir ağı girdisi
Y	: Yapay sinir ağı çıktı bilgisi
ϕ_k	: Gizli katmandaki bir aktivasyon fonksiyonu
m^2	: Öğrenme oranı
J	: Maliyet fonksiyonu
Θ	: Güncellenecek parametre

KISALTMALAR

ADAM	: Adaptive Moment Estimation (Uyarlanabilir Moment Tahmini)
AF	: Aktivasyon Fonksiyonu
CCN	: Konvolüsyonel Sinir Ağları (Convolutional Neural Networks)
GA	: Genetik Algoritmalar
ML	: Machine Learning (Makine Öğrenmesi)
MLP	: Multi Layer Perceptron(Çok Katmanlı Algılayıcı)
NN	: Sinir Ağı(Neural Network)
OECD	: Organisation for Economic Co-Operation and Development (Ekonomik İşbirliği ve Kalkınma Örgütü'dür.
PISA	: Uluslararası Öğrenci Değerlendirme Programı
PVMATH_ KODLAMA	: 10 adet makul matematik beceri puanı ortalaması
PVREAD_ KODLAMA	: 10 adet makul okuma beceri puanı ortalaması
ReLU	: Rectified Linear Unit Fonksiyonu
RNN	: Recurrent Neural Network (Tekrarlayan Sinir Ağı)
SGD	: Skolastik Gradyan İnişi (Stochastic Gradient Descent)
YSA	: Yapay Sinir Ağı

1. GİRİŞ

İnsanlık tarihinde doğayla iç içe yaşayarak, çeşitli ihtiyaçlarını karşılamak için doğal kaynakları kullanmış ve doğal afetlerle başa çıkmak için çeşitli çözümler geliştirmiştir. Doğadaki olayları anlamak ve sebep-sonuç ilişkilerini kavramak için düşünce ve hislerini kullanarak çalışmıştır. Bu çalışmalar, zamanla bilgi birikimi ve teknolojinin ilerlemesiyle daha kapsamlı hale gelmiştir. Özellikle 1950'lerden sonra bilgisayar teknolojisinin hızla gelişmesiyle, son birkaç on yılda büyük ilerlemeler kaydedilmiştir. Bazı geliştirilen yöntemler, canlı organizmaların işleyişinden esinlenerek ortaya çıkmıştır. Bu yöntemlere örnek olarak Yapay Sinir Ağları (YSA) ve Genetik Algoritmalar (GA) verilebilir (Şen, 2004).

Yapay sinir ağları, insan beyninin öğrenme işlevini gerçekleştiren bilgisayar sistemleridir ve örneklerle öğrenme sürecini taklit ederler. Bu ağlar, yapay sinir hücrelerinden oluşan birbirine bağlı işlem birimlerinden oluşur ve her bağlantıda bir ağırlık değeri bulunur. Ağın bilgisi, bu ağırlık değerlerinde saklıdır ve ağa yayılmıştır. Yapay Sinir Ağları, geleneksel hesaplama yöntemlerinden farklı bir yaklaşım sunar ve adaptif, eksik bilgiyle çalışabilen ve belirsizlik altında karar verebilen özelliklere sahiptir. Bu yöntem, hemen hemen her alanda başarılı uygulamalarıyla dikkat çeker. Ağın yapısının belirlenmesi ve parametrelerinin seçimi konusunda belirli bir standart olmamasına rağmen, yapay sinir ağları tahmin, sınıflandırma, örüntü tanıma, veri filtreleme, veri ilişkilendirme, veri yorumlama gibi alanlarda güçlü teknikler olarak kabul edilirler. Günlük yaşamda, optimum rotaların belirlenmesi, görüntü ve ses analizinde, tıpta hastalıkların tespitinde, trafik kontrol sistemleri gibi alanlarda birçok başarılı uygulamaları vardır. Çok katmanlı algılayıcı ağları, günümüzde yapay sinir ağlarının en yaygın kullanılan modelidir. Bu ağlar, özellikle mühendislik problemlerinin çözümünde geçerli sonuçlar verebilmektedir.

Girdi ve çıktı katmanlarında kaç proses elemanının olacağı, problem yapısına bağlı olarak belirlenir. Ara katman sayısı ve her ara katmandaki eleman sayısı için kesin bir yöntem bulunmamaktadır; bu, deneme-yanılma yöntemiyle belirlenir. Bilgi akışı, girdi katmanından ara katmana ve ardından çıktı katmanına doğru ilerler.

Bulut'un (2016) çalışmasında lise öğrencilerine meslek seçimlerinde karar verirken yapay bir karar destek sistemiyle yardımcı olmayı amaçlayan çok katmanlı algılayıcı kullanılarak oluşturulan bir sistem oluşturulmuştur. İşeri ve Arıman'ın (2019) çalışmalarında Orta Karadeniz Kıyı şeridi deniz ve ırmak sedimanlarında ağır metal konsantrasyonunun Çok Katmanlı Algılayıcı Ağ (MLP) ve diğer regresyon modelleri kullanılarak tahminlenebilmesi gerçekleştirilmiştir. İban ve Şentürk (2021) İyonosfer Parametrelerinin Çok Katmanlı Algılayıcılar ile Kestirimi çalışmalarında iyonosferin f0F2, hmF2 ve TEC parametrelerini tahmin etmek için çok katmanlı algılayıcı regresyonunun potansiyel kullanımını göstermeyi amaçlamışlardır.

Çok katmanlı algılayıcıların kullanıldığı PISA verilerini içeren çok sayıda çalışma bulunmaktadır. İnal ve Turabik'in (2017) çalışmasında PISA 2012'ye katılan 15 yaş grubu Türk öğrencilere ilişkin veriler ışığında; öğrencilerin matematik başarılarını etkileyen bazı faktörlerin yordama gücünün belirlenmesi çok katmanlı algılayıcı kullanılmış ve yapay sinir ağı matematik başarısını oldukça yüksek bir oranda doğru olarak sınıflandırmıştır. Güre, Kayri ve Erdoğan (2020) çalışmalarında PISA 2015 matematik okuryazarlığını etkileyen faktörleri Çok Katmanlı Algılayıcı Yapay Sinir Ağları ve Rastgele Orman yöntemlerini kullanarak incelemişlerdir. Koyuncu (2020) PISA (Uluslararası Öğrenci Değerlendirme Programı) 2003 ve 2012'de yer alan matematiğe özgü eğilim değişkenlerinin yıllar itibarıyla matematik performansını etkileme düzeylerini çok katmanlı algılayıcı ve radyal tabanlı fonksiyon teknikleri ile çoklu doğrusal regresyon kullanarak incelemiştir.

OECD üye ülkeleri ve ortakları, COVID-19 salgını nedeniyle yaşanan aksaklıkları yansıtmak amacıyla PISA 2021 değerlendirmesini 2022'ye, PISA 2024 değerlendirmesini ise 2025'e erteleme kararı almıştır. PISA 2022'ye 81 ülke ve ekonomi katıldı (OECD, 2024). Türkiye, 2003 yılından bu yana bu programa katılmaktadır. Bu araştırmada, matematik, fen ve okuma becerilerini değerlendirilmekte, ancak her uygulamada bu alanlardan biri öne çıkmakta ve derinlemesine incelenmektedir. PISA 2022'de odaklanılan alan matematik okuryazarlığıdır.

PISA 2022'ye katılan 81 ülkenin matematik ortalaması 438 ve 37 OECD ülkesinin ortalaması ise 472'dir. Okuma becerileri alanında ise PISA 2022'ye katılan 81 ülkenin ortalaması 435 ve 37 OECD ülkesinin ortalaması ise 476'dır (MEB, 2024).

PISA verilerine göre, öğrencilerin sosyoekonomik durumları ile akademik başarıları arasındaki ilişki çeşitli anket soruları ve yanıtlar aracılığıyla incelenmektedir. MEB'in (2024) raporunda, Türkiye'de öğrencilerin matematik performanslarındaki farklılıkların %12.6'sı, okuma becerilerindeki farklılıkların %9.8'i sosyoekonomik durumlarından kaynaklandığı belirtilmektedir. OECD ülkelerinde ise matematik performans farklılıklarının %15.5'i ve okuma becerileri performansının %12.6'sı sosyoekonomik durumdan etkilenmektedir. Bu veriler, Milli Eğitim Bakanlığı'nın yayınladığı PISA 2022 raporunda belirtilmiştir (MEB, 2024).

Türkiye'deki 15 yaş grubundaki öğrencilerin %67'si, anne babalarıyla her gün yemeklerini birlikte yediklerini, Türkiye örneklemini için PISA 2022'ye katılan öğrencilerin yarısı ebeveynlerinin okulda neler yaptıkları, gün içindeki aktiviteleriyle alakalı olarak sohbet ettiklerini belirtmişlerdir. Her 10 öğrenciden bir veya iki tanesi ise bu tür deneyimleri hiç yaşamadıklarını ya da nadiren yaşadıklarını dile getirmiştir (MEB, 2024).

PISA 2022'nin ilk sonuçları 5 Aralık 2023'te açıklanmıştır. Bu sebeple literatür incelendiğinde PISA 2022 verileri kullanılarak yapılan çalışma henüz az sayıdadır. Çalışma verilerinin yeni olması, ayrıca sosyoekonomik faktörler ve aile eğitim durumu-ebeveyn tutumları değişkenlerinin matematik okuryazarlığı ve okuma becerisi başarısı tahmininde kullanılmasının, geçmişte yapılan çalışmalar ile bağlantı kurulabilmesi açısından önem arz etmektedir.

Yapılan çalışmanın amacı, yukarıda belirtilen sosyoekonomik durumun ve ebeveyn tutumları arasındaki farklılıkların öğrenci başarısını tahmin becerisini ölçmek, ayrıca Rapid Miner derin öğrenme ağı ve Python çok katmanlı yapay derin öğrenme ağı arasındaki tahmin başarısını karşılaştırmaktır. Bu amaçla OECD web sitesinde 5 Aralık 2023 tarihinde yayınlanan SPSS formatındaki veriler indirilmiş, öğrencilere uygulanan öğrenci anketleri ve ebeveynlere uygulanan ebeveyn anketleri incelenmiş, ekonomik konularla ilgili etiketleri ve ebeveyn tutumları ile ilgili etiketleri içeren değişkenler seçilmiştir. Ekonomik koşullar ile ilgili 18 bağımsız değişken, ebeveyn tutumlarına aile eğitim durumu da eklenmiş ve toplam 33 adet bağımsız değişken de bu alan için seçilmiştir. Çalışmanın bağımlı değişkenlerini oluşturan matematik okuryazarlığı ve okuma becerisi puanları, 10 adet makul okuma becerisi puanı ve 10 adet makul matematik okuryazarlığı puanı şeklinde yayınlanmıştır.

Öncelikle yayınlanan 10 adet puanın ortalaması alınmış sonrasında ise Milli Eğitim Bakanlığı PISA 2024 raporunda belirtilen (Tablo 3.6-3.7) şekilde 0-6 arasında kodlanmıştır. Ayrıca seçilen bağımsız değişkenlerin PISA 2022'ye katılan öğrenci puanlarını başarılı ve başarısız olarak sınıflandırması için PISA Türkiye 2024 raporunda da belirtildiği üzere OECD ortalamasının üzerinde puan alanlar başarılı (1), altında puan alanlar ise başarısız (0) olarak kodlanmıştır. Buna göre PISA 2022'de 37 OECD ülkesinin ortalama okuma becerisi puanı 476, ortalama matematik becerisi puanı ise 472'dir.

Seçilen bağımsız değişkenler ve (0-6), (0-1) arasında kodlanan bağımlı değişkenler matematik becerisi puanlarını ve okuma becerisi puanlarını tahmin etmesi için, Rapid Miner programında ve Python programlama dili kullanılarak oluşturulan yapay sinir ağına aktarılmış ve bağımsız değişkenlerin bağımlı değişkenleri tahmin etme oranı tespit edilmiştir. Yapılan analiz sonucunda ekonomik değişkenlerin matematik becerisi ve okuma becerisi puanlarını tahmin başarıları %39-%43, aile eğitim durumu ve ebeveyn tutumları değişkenlerinin matematik becerisi ve okuma becerisi puanlarını tahmin başarıları %38-%42 olarak bulunmuştur. Ekonomik değişkenlerin öğrencilerin matematik becerisi ve okuma becerisi başarı durumlarını tahmin etme oranı %77-%78, aile eğitim durumu ve ebeveyn tutumları değişkenlerinin matematik becerisi ve okuma becerisi başarı durumları tahmin etme oranı %78-%77 olarak bulunmuştur. Bu bağlamda ekonomik unsurlar ve ebeveyn tutumları öğrenci başarısının önemli bir yordayıcısıdır. Ayrıca Rapid Miner programında oluşturulan derin öğrenme ağı Python programı ve Keras kütüphanesiyle oluşturulan derin öğrenme ağına oranla %10 oranında daha başarılı tahminler yapmıştır.

2. LİTERATÜR ÖZETİ

Bu bölümde PISA ve yapay sinir ağları hakkında genel bilgilere yer verilmiştir.

2.1. PISA Nedir?

PISA, Ekonomik İşbirliği ve Kalkınma Teşkilatı (OECD) tarafından her üç yılda bir, 15 yaş grubundaki öğrencilerin edindikleri bilgi ve becerileri değerlendirmek amacıyla yürütülen bir araştırmadır. Bu araştırmanın temel amacı, öğrencilerin okulda öğrendikleri bilgi ve becerileri günlük yaşamda nasıl kullandıklarını ölçmektir. Ayrıca gençlerin öğrenme isteklerini, derslerdeki performanslarını ve öğrenme ortamları ile ilgili tercihlerini daha açık bir şekilde anlamak da hedeflenmektedir. PISA'da, zorunlu eğitimin sonunda örgün eğitim almaya devam eden 15 yaş grubundaki öğrencilerin matematik, fen bilimleri ve okuma becerileri gibi konu alanlarının yanı sıra motivasyonları, öğrenme biçimleri, okul ortamları ve aileleri ile ilgili veriler de toplanmaktadır. Bu araştırmada kullanılan "okuryazarlık" kavramı, öğrencilerin yazılı kaynakları bulma, kullanma, kabul etme ve değerlendirme becerilerini ifade etmektedir. PISA'nın uygulanması ve sonuçların değerlendirilmesi, OECD eğitim direktörlüğü'ne bağlı olan PISA yönetim kurulu tarafından gerçekleştirilmektedir. Sonuçlar, ulusal raporlar hâlinde düzenlenerek eğitim politikalarının geliştirilmesinde ve eğitim alanındaki araştırmalarda kaynak olarak kullanılmaktadır. Ülkemiz de OECD üyesi olarak, eğitim düzeyinin belirlenmesi ve geliştirilmesi amacıyla PISA araştırmasına katılmaktadır.

2.2. Yapay Zeka

Yapay zekanın babası olan olarak anılan Alan Turing, 1950 yılında yayınlanan “Bilgisayar Makineleri ve Zekâsı” adlı makalesinde akıllı bir makine kavramına odaklandı. Bir makinenin akıllı olup olmadığını anlayabilmek için de Turing Testi denilen bir ölçme aracı geliştirdi. Turing Testi; makinenin bir şeyi gerçekten bilip bilmediğini, kendisinin farkında olup olmadığını anlamaktan da daha ziyade bir makinenin büyük miktarda bilgiyi işleyebildiğini, konuşmayı yorumlayabildiğini ve insanlarla iletişim kurabildiğini göstermeyi amaç edinmektedir. Alan Turing aslında bir makinenin bu testi geçebilmesinin yüzyılın sonuna kadar mümkün olmayacağına inanıyordu. Evet, bu yapay zekanın başarısızlıkla sonuçlanacak pek çok öngörüsünden biriydi (Taulli, 2019). Turing’ in temel

iddiası bir insanın konuşmasını kusursuz şekilde taklit edebilecek makineler yapılabilecek olmasıydı. Turing'in makalesi yapay zekâ alanındaki diğer makalelerden daha fazla yorum ve tartışma yarattı. Herhangi bir web tarayıcısına Turing Testi yazıldığında binlerce sonuç elde edilmesinin nedeni kısmen insan davranışlarını taklit edebilen ve insan gibi düşünebilen makineler üretme tutkusudur (French, 2020). Alan Turing makalesinde 50 yıl içinde bu testin geçilebileceğini belirtmiş ve 2014 yılında Eugene Goostman adlı program Turing testini başarıyla geçmiştir (Fancher, 2016).

Mehrotra'ya (2019) göre yapay zekanın üç farklı seviyesi vardır. Buna göre makine belirli bir görevi insandan daha iyi gerçekleştirebildiğinde yapay zekanın dar olduğu söylenir. Yapay zekâ, herhangi bir entelektüel görevi insanla aynı doğruluk düzeyinde gerçekleştirebildiğinde genel duruma ulaşır. Güçlü yapay zekâ ise birçok görevde insanları yenebilen bir teknoloji olarak tanımlanmaktadır. Yapay zekâ, akıllı makinelerin ve yazılımların incelenmesi ve geliştirilmesiyle ilgilenen bir teknoloji ve bilgisayar biliminin bir dalıdır; bir makinenin akıllı bir insan gibi düşünüp hareket etmesini sağlayan bilimdir (Mehrotra, 2019).



Şekil 2.1: Veri bilimi teknikleri (Choi vd., 2020)

Şekil 2.1’de de görüldüğü üzere makine öğrenmesi (Machine Learning) ve derin öğrenme (Deep Learning) yapay zekanın bir alt alanıdır ve yapay zekâ normalde insanlar tarafından gerçekleştirilen özel görevlerin otomatikleştirilmesine odaklanan bir alandır. Örneğin, sembolik yapay zekâ olarak bilinen alt alan, belirli bir ilgi alanındaki olası her senaryo için kuralların sabit kodlanmasına (yani açıkça yazılmasına) odaklanır. Programcıların yazdığı bu kurallar bütünü insanların bu problem durumuyla ilgili olarak bildikleri önsel bilgilerinden gelmektedir. Örneğin bir odadaki ısı seviyesinin bir program yardımıyla kontrol etmek istendiğinde, ısı seviyesi belirli bir sıcaklığın üzerine çıktığında klimanın otomatik olarak açılmasını sağlayan bir algoritma yazılabilir. Yapay zekâ, açıkça tanımlanmış mantıksal problemleri çözme konusunda becerikli olabilir, ancak konuşma tanıma veya görüntü sınıflandırma gibi daha karmaşık desen tanıma gerektiren görevlerde sıklıkla başarısız olabilir. Bu tür karmaşık görevler, genellikle makine öğrenimi ve derin öğrenme yöntemlerinin iyi performans sergilediği alanlardır (Choi vd., 2020).

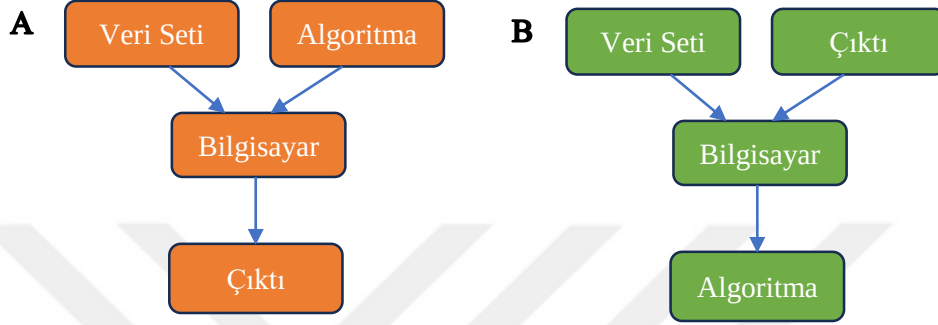
Yapay zeka; yapay zeka, makine öğrenmesi ve derin öğrenme olarak üç alt alana ayrılabilir (Mehrotra, 2019).

2.3. Makine Öğrenmesi (Machine Learning-ML)

Yapay zekanın bir alt dalı olarak tanımlanabilen makine öğrenmesi, bilgisayarlara açık bir şekilde programlanmadan deneyimler yoluyla öğrenme yeteneği veren bir çalışma alanı şeklinde tanımlanmıştır (Samuel, 1959). Bir algoritmanın bilinen özellikler kullanılarak açık bir şekilde kodlanabildiği klasik programlamanın aksine, makine öğrenmesi ilk ilkelerden türetilen yeni veya farklı özellik ve ağırlık kombinasyonlarını kullanabilen bir algoritma oluşturmak için veri alt kümelerini kullanır (Choi vd., 2019). Makine öğrenmesi, sistemin önceden belirlenmiş talimatlarla değil, veri veya deneyimlerden yararlanarak kendiliğinden öğrenme yeteneği kazandığı bir disiplindir.

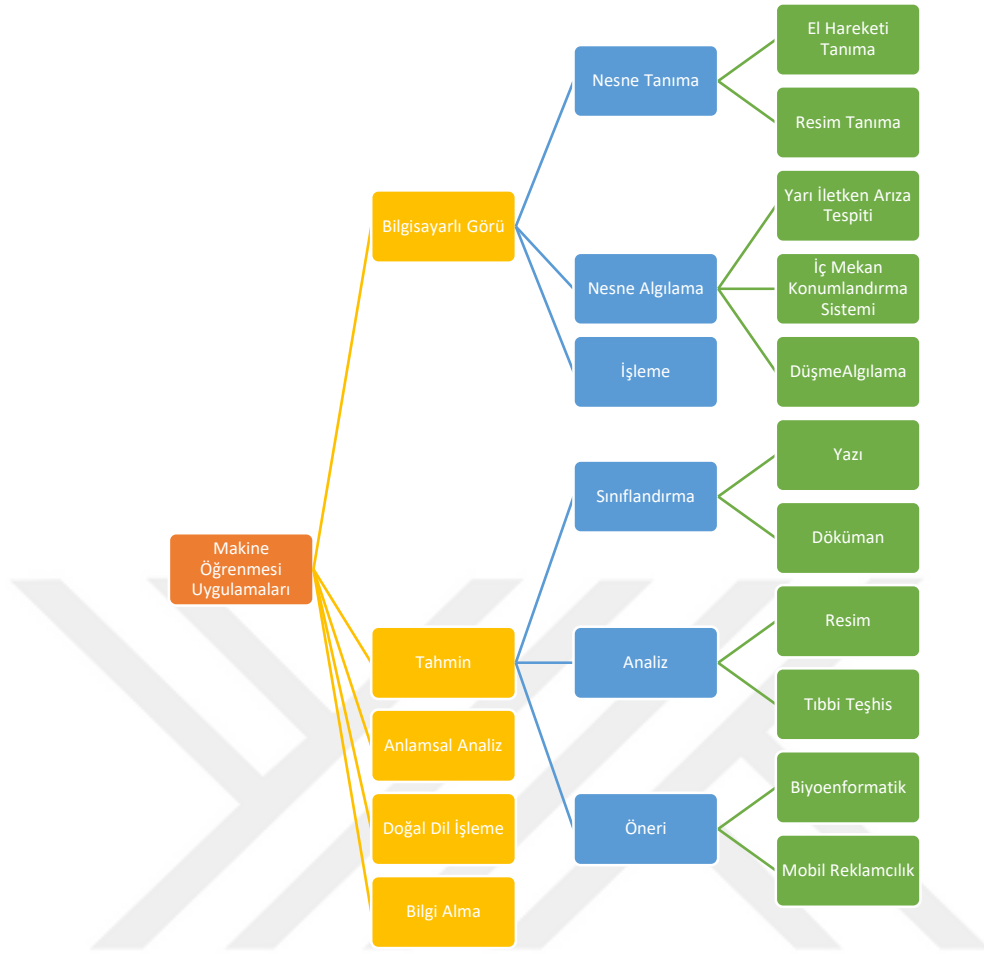
Makine öğrenmesi fikrinin öncüsü kabul edilen Arthur Samuel deneyimlerden öğrenen bir dama oyuncusu geliştirdi. Bu çalışma, makine öğreniminin en eski ve en etkili örneklerinden biridir (Wiederhold ve McCarthy, 1992). Dama oyunu ile ilgili yapılan çalışmalar, makine öğreniminin nasıl işlediğini gösterdi; yani bir bilgisayarın, açıkça programlanmadan verileri işleyerek öğrenebileceğini ve kendi oyun deneyimleri ve farklı oyuncularla oynadığı maçlar

sonrasında başarılı hamleleri gözlemleyerek kendisini geliştirebileceğini ortaya koydu. Program, bu gözlemlere dayanarak stratejilerini ayarladı ve başarılı hareketleri programlamasına entegre etti. Bu, özellikle olasılık analizi ve ileri düzey istatistik kavramlarının kullanımıyla mümkün oldu. Bu sayede, bir bilgisayar doğru tahminler yapabilecek şekilde eğitilebilir hale geldi (Taulli, 2019).



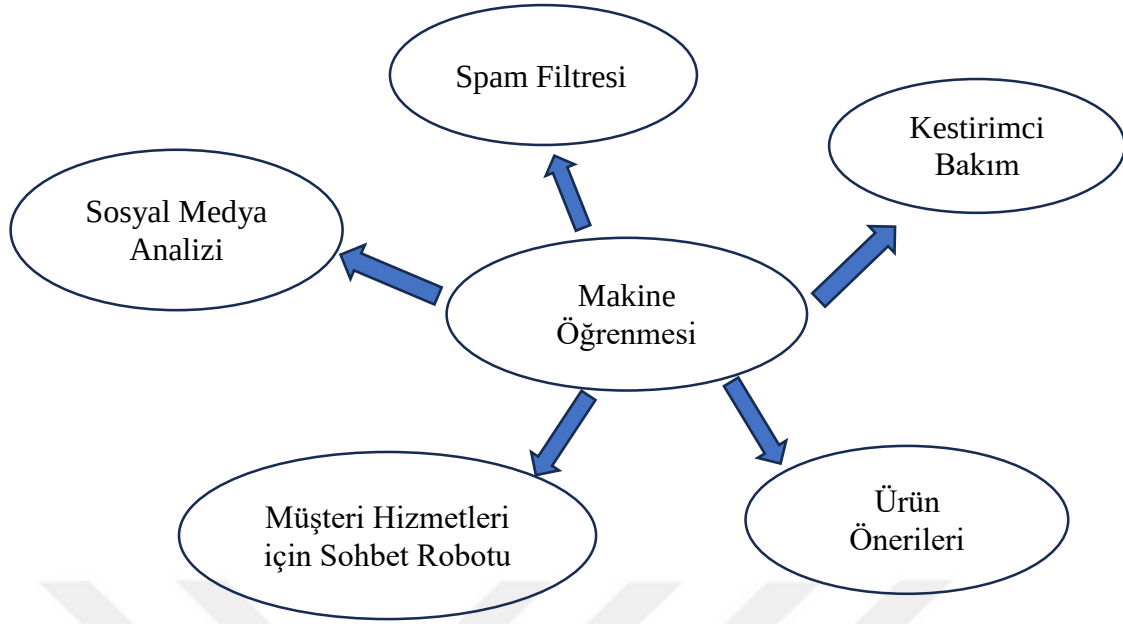
Şekil 2.2: Klasik programlama (A) ve makine öğrenimi (B) paradigması (Choi vd., 2019)

Şekil 2.2’de klasik programlama ve makine öğrenimi arasındaki fark gösterilmektedir. (A) Klasik programlamada bilgisayara bir veri seti ve bir algoritma sağlanır. Algoritma, bilgisayara çıktı oluşturmak için veri kümesi üzerinde nasıl çalışacağını bildirir. (B) Makine öğreniminde, bilgisayara bir veri kümesi ve ilgili çıktılar sağlanır. Bilgisayar, ikisi arasındaki ilişkiyi tanımlayan bir algoritmayı öğrenir ve üretir. Bu algoritma gelecekteki veri kümeleri hakkında çıkarım yapmak için kullanılabilir.



Şekil 2.3: Makine öğrenmesi uygulama alanları (Shinde ve Shah, 2018)

Şekil 2.3’te makine öğrenmesi uygulamaları gösterilmiştir. Makine öğrenmesi, genellikle kullanılan dört farklı öğrenme yöntemiyle çeşitli görevleri çözebilir: denetimli, denetimsiz, yarı denetimli ve takviyeli öğrenme (Choi vd., 2019).



Şekil 2.4: Makine öğrenmesi uygulamaları

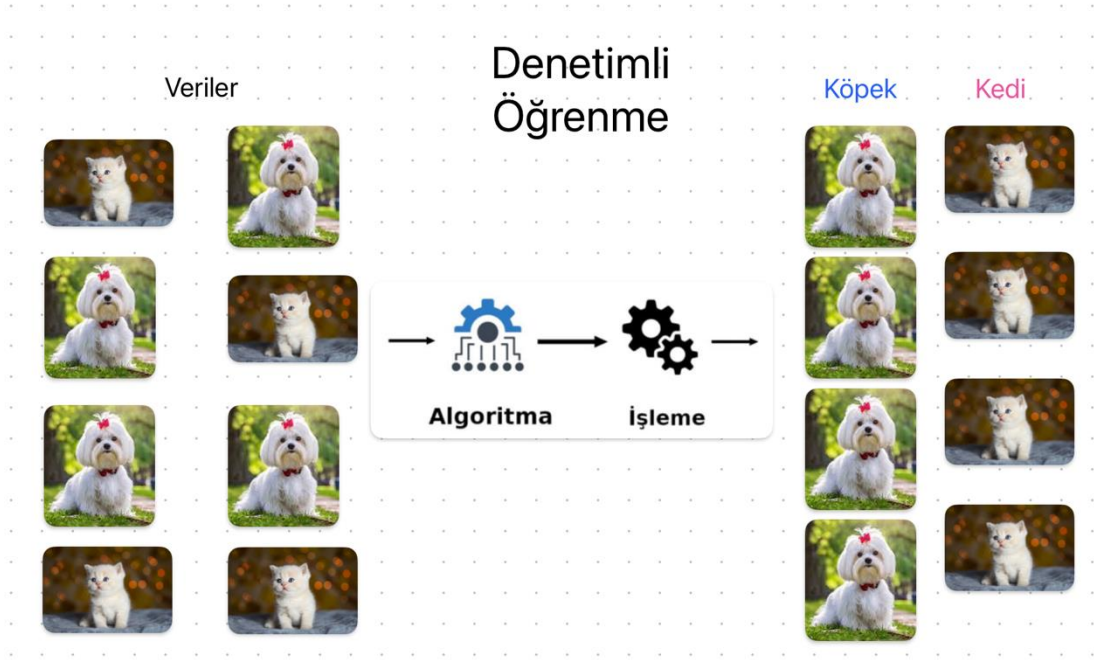
Şekil 2.4’te yaygın olarak kullanılan bazı makine öğrenmesi uygulamaları gösterilmiştir. Kestirimci bakım; ekipmanın ne zaman arızalanabileceğini tahmin etmek için sensörleri izler. Bu yalnızca maliyetleri azaltmaya yardımcı olmakla kalmaz, aynı zamanda arıza süresini azaltır ve güvenliği artırır. Çalışanları işe almak; birçok özgeçmiş genellikle çeşitlilik gösterdiğinden bu sıkıcı bir süreç olabilir. Bu, kalifiye adayları gözden kaçırmamanın kolay olduğu anlamına gelir. Ancak makine öğrenimi büyük ölçüde yardımcı olabilir. Müşteri deneyimi; Günümüzde müşteriler kişiselleştirilmiş bir deneyim istemektedir. Makine öğrenimi sayesinde bir şirket, iç gözü elde etmek, yani gerçekten neyin işe yaradığını öğrenmek için verilerinden yararlanabilmektedir. Müşteri hizmetleri; Geçtiğimiz birkaç yılda, müşterilerle etkileşimleri otomatikleştirmek için makine öğrenimini kullanan sohbet robotlarında büyüme görüldü (Taulli, 2019).

2.3.1 Makine Öğrenmesi Teknikleri

Makine öğrenimi algoritmaları, geçmiş verileri kullanarak veri modellerini çıkararak çalışır. Denetimli ve denetimsiz makine öğrenimi, makine öğrenimi algoritmalarının iki ana kategorisidir.

2.3.1.1 Denetimli Öğrenme

Denetimli öğrenmede, bir nöron ağından gelen bilgi, başka bir ağdaki bağlantı modelini etkilemek için bir öğretici sinyal görevi görür. Sonuç olarak, bu ikinci ağ, öğretici sinyal tarafından belirlenen belirli bir hedefe veya dönüşüme ulaşacak şekilde bilgiyi işlemeyi öğrenir (Knudsen, 1994). Denetimli öğrenme (supervised learning), epidemiyolojide sıkça kullanılan modelleme yöntemlerine benzer: Her gözlem için sonuç değeri (bağımlı değişken) bilinir. Makine öğreniminde bu sonuç değeri "etiket" olarak adlandırılır. Sonuç değerlerinin bulunduğu verilere "etiketli veri" denir (Bi vd., 2019).



Şekil 2.5: Denetimli öğrenme

Şekil 2.5'te denetimli öğrenmenin eğitim ve verileri sınıflandırma aşamaları gösterilmiştir. Bu bağlamda denetimli öğrenmede makineye bir etiketlenmiş bir veri örneği verilir ve makine bu veriler ile eğitilir ve yeni verilen girdi verileri üzerinde çıktılar üretmesi beklenir. Yani, çiçek görsellerinden oluşan etiketli bir veri seti, modele hangi fotoğrafların manolya, sümbül, akasya ve nergislerden oluştuğunu söyleyebilir. Denetimli öğrenme modeli, yeni bir resim gösterildiğinde, bu resmi eğitim örnekleriyle karşılaştırarak doğru etiketi tahmin etmeye çalışır.

Denetimli öğrenme genel olarak iki tür probleme ayrılabilir:

- Sınıflandırma algoritmaları, test verilerinin belirli kategorilere doğru şekilde atanmasını hedefler, elmaları portakallardan ayırmak gibi. Gerçek dünyada, denetimli öğrenme algoritmaları rampa ölçümlü trafik ışıklarını kırmızı ve yeşil olmak üzere iki farklı duruma sınıflandırmak için kullanılır (Wang, Cui ve Ke, 2023).
- Regresyon, bağımlı ve bağımsız değişkenler arasındaki ilişkiyi anlamak için bir algoritma kullanan başka bir tür denetimli öğrenme yöntemidir. Regresyon modelleri, belirli bir yol bölümündeki trafik hızı gibi farklı veri noktalarına dayalı sayısal değerleri tahmin etmekte başarılı sonuçlar verebilmektedir (Wang vd., 2023).

Denetimli öğrenmenin önemli noktalarından biri, büyük miktarda verinin olması gerektiğidir. Bu, modelin iyileştirilmesine ve daha doğru sonuçlar üretilmesine yardımcı olur. Mevcut verilerin çoğunun etiketlenemediği durumlarda ve ayrıca veri kümesi de çok büyük ise verilerin etiketlenmesi sorun olabilir. Bu bağlamda verilerin etiketlenebilmesi için ImageNet gibi, görsel nesne tanıma yazılım araştırmalarında kullanılmak üzere tasarlanmış büyük bir görsel veritabanı kullanılmaktadır (Taulli, 2019).

2.3.1.2 Denetimsiz Öğrenme

Her zaman tam olarak etiketlenmiş veri kümelerine ulaşmak kolay değildir. Denetimsiz öğrenme, etiketlenmemiş veri kümelerini analiz etmek ve kümelemek için makine öğrenimi algoritmalarını kullanır. Bu algoritmalar, insan müdahalesine gerek kalmadan verilerdeki gizli kalıpları keşfeder bu nedenle denetlenmezler (Wang vd., 2023).

Denetimsiz öğrenme modelleri genellikle aşağıdaki gibi tanımlanabilecek üç ana görev olan kümeleme, ilişkilendirme ve boyutluluğun azaltılması için kullanılır. Kümeleme, etiketlenmemiş verileri benzerliklerine veya farklılıklarına göre gruplamaya yönelik bir tekniktir. Bu yöntem, seyahat davranışı kategorizasyonunda başarılı olabilmektedir (Wang vd., 2023). Boyut azaltma, belirli bir veri kümesindeki özelliklerin (veya boyutların) sayısı çok yüksek olduğunda kullanılan bir öğrenme tekniğidir. Veri girişi sayısını yönetilebilir bir boyuta indirirken aynı zamanda veri bütünlüğünü de korur (Wang vd., 2023).

Çoğu zaman bu teknik, örneğin AutoEncoders'ın (otomatik kodlayıcıların) resim kalitesini artırmak için görsel verilerdeki gürültüyü ortadan kaldırdığı durumlarda olduğu gibi veri ön işleme aşamasında kullanılır (Wang vd., 2023). AutoEncoders'ın amaçlarından birisi de verinin içinden önemli giriş etiketleri çıkarmaktır.

2.3.1.3 Yarı Denetimli Öğrenme

Yarı denetimli öğrenmede eğitim verileri az miktarda etiketlenmiş ve daha büyük miktarda etiketlenmemiş veri olarak ikiye ayrılır. Bu durumda model, etiketlenmemiş veriler hakkında çıkarımlar yapmak için etiketli verileri girdi olarak kullanır ve normal denetimli öğrenme modellerinden daha doğru sonuçlar sağlar. Etiketli veriler esas olarak algoritmanın verileri anlayabilmesi için anlamlı etiketlere sahip olan bilgilerdir; etiketlenmemiş veriler ise bu bilgilerden yoksundur. Bu kombinasyonu kullanarak, makine öğrenimi algoritmaları etiketlenmemiş verileri etiketlemeyi öğrenir (Wang vd., 2023).

Veri etiketleme, insan müdahalesi gerektiren zaman alıcı bir süreçtir ve büyük veri kümeleri ve örüntü tespiti ile uğraşırken makine öğrenimi aynı süreci yürütmek için kullanılır ve kodlamayı katılan herkes için kolaylaştırır bu nedenle de denetimsiz öğrenme günümüzde çok önemli bir teknik haline gelmiştir. Denetimli öğrenmede, bir veri bilimci, örnek yoluyla öğrenebilmesi için sisteme kedi olarak etiketlenen kedilerin fotoğrafları gibi etiketlenmiş veriler sunar. Denetimsiz öğrenmede, bir veri bilimci yalnızca fotoğraf verir ve verileri inceleyip bunların kedi görseli olup olmadığını belirlemek sisteme kalmıştır (Naeem vd., 2023). Bunun yanı sıra denetimsiz öğrenmeye, etkili bir ağ koruması aracı olarak Lu ve Ding'in (2021) geliştirdikleri ağ kötü amaçlı trafik algılama sistemi örnek verilebilir.

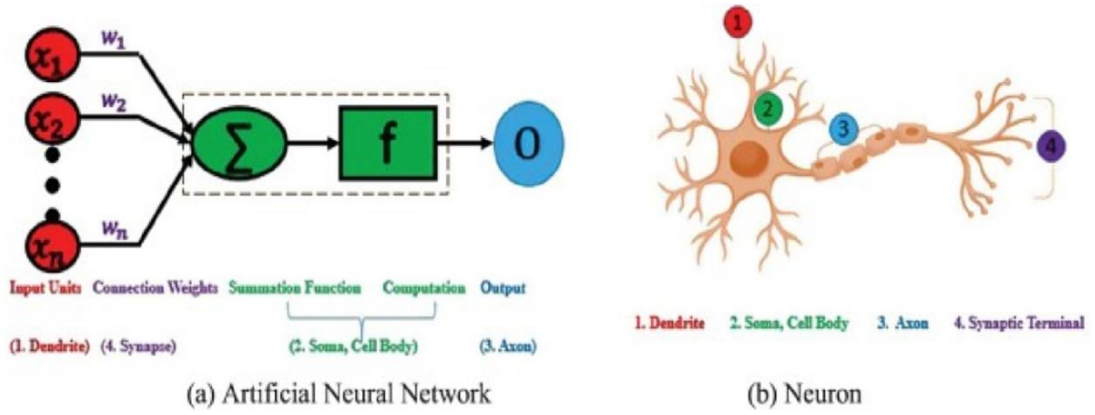
2.3.1.4 Takviyeli Öğrenme

Makine öğreniminin bir çeşidi olan takviyeli öğrenme, bir sistemin bulunduğu ortamı algılayıp kendi başına kararlar alarak hedefine ulaşabilmesi için doğru kararları nasıl öğreneceğini gösterir (Kaelbling vd., 1996). Yeni bir spor yapmaya başlandığında, bu sporu yapan kişilerin gözlemlenmesi ve nasıl yapıldığının anlaşılmasına çalışılması önemlidir. Örneğin, futbol oynarken top kaybedildiğinde takım arkadaşları bu durumdan hoşnut olmazlar, ancak doğru hamleler yapılarak gol atıldığında herkes memnuniyet duyar. Bu

şekilde, futbol öğrenme süreci deneme yanılma yoluyla gelişir. Oyunlar, takviyeli öğrenme için idealdir çünkü net kuralları, puanları ve çeşitli kısıtlamaları içerirler. Bir model oluşturulurken, milyonlarca simülasyonla test edilebilir; bu da sisteminin hızla daha akıllı hale gelmesi anlamına gelir. Bu sayede bir program, Go veya satranç gibi oyunlarda dünya şampiyonunu yenmeyi öğrenebilir. Takviyeli öğrenme robotikte de etkili bir şekilde kullanılmaktadır. Bir alan içinde gezinme yeteneği önemlidir ve bunun için çevreyi birçok farklı noktada değerlendirmek gerekir. Örneğin, bir robotun mutfığa gitmek istediği bir durumda, mobilyaların etrafından ve diğer engellerden geçmesi gerekebilir. Eğer bir şeye çarparsa, negatif bir takviye eylemi gerçekleşir. Bu sayede robot, olumsuz deneyimlerden öğrenerek daha etkili bir şekilde hareket etmeyi öğrenir (Taulli, 2019).

2.4. Yapay Sinir Ağları

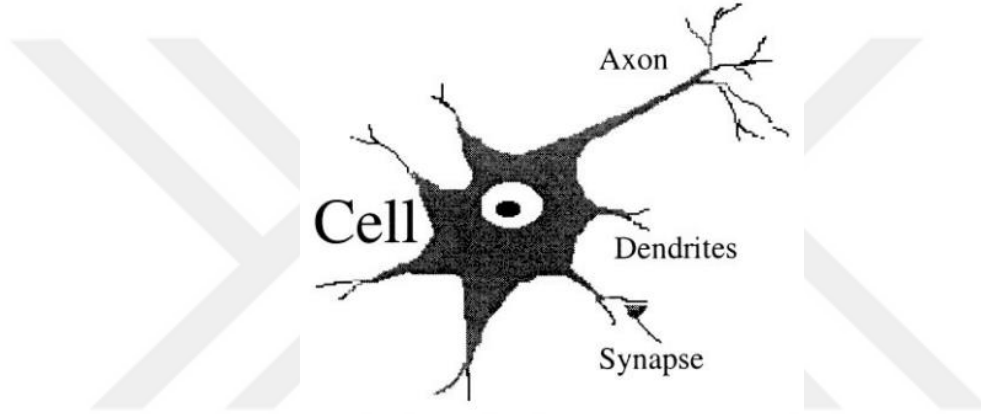
YSA'lar olarak da bilinen yapay sinir ağları, biyolojik olarak insan beynindeki nöronlardan ilham alır ve özel hesaplama görevlerini gerçekleştirmek üzere yapılandırılmış yapay nöronlardan oluşur.



Şekil 2.6: Yapay sinir ağı (a) ve bir nöron'un (b) temel şekli (Tyagi ve Abraham, 2022)

YSA'nın temel modeli, Şekil 2.6'da gösterilen biyolojik bir nörona benzer. İnsan beyni, diğer nöronlarla birlikte paralel çalışan birimlerin muazzam bir şekilde birbirine bağlanması nedeniyle son derece karmaşık bir sistem oluşturan milyarlarca nöronlardan oluşur. Bu nöronlar yaklaşık 10.000 diğer nöronlardan sinyal alır (Tyagi ve Abraham, 2022).

Yapay sinir ağı (YSA)'lar, insan beyninin sayısallaştırılmış modelleridir ve insan beyninin bilgi işleme şeklini taklit etmek için tasarlanmış bilgisayar programlarıdır. YSA'lar, programlama yoluyla değil, insanlar gibi, uygun öğrenme örnekleriyle deneyim yoluyla öğrenirler veya eğitilirler. Beyin, mükemmel bir desen tanıma aracıdır. Bir kaleme baktığımızda, beynimizdeki belirli bir bölgedeki nöronlar daha önce benzer bir girdi deseniyle karşılaşmış ve bu deseni "kalem" nesnesiyle ilişkilendirmeyi öğrenmiştir. Beynimiz, birbirine bağlı milyarlarca nöron içerdiğinden, neredeyse sonsuz çeşitlilikte girdi kalıplarını öğrenip tanıyabilir. Ortalama bir beyin, her biri diğer nöronlarla 1000-10.000 bağlantıya sahip olan 100 milyar nöron içerir.



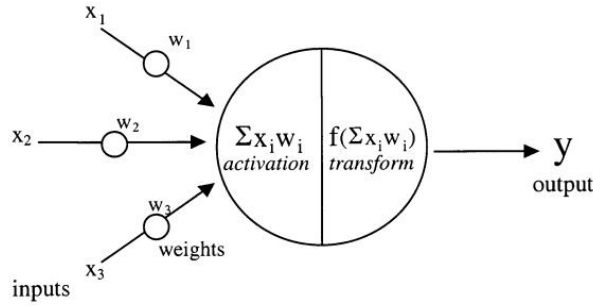
Şekil 2.7: Biyolojik sinir hücresi (Kustrin ve Beresford, 2000).

Şekil 2.7’de biyolojik yapısı gösterilen nöronlar, hücre aktivitesini kontrol eden bir çekirdek, çok sayıda ince iplik, bilgi taşıyan dendritler ve sinyali taşıyan daha uzun bir iplik olan bir akson içeren bir hücre gövdesinden oluşur. İmpuls, akson boyunca sinapslara doğru ilerler. Sinapslarda, bir nörondan diğerine geçiş yapar. Nöronlar tamamen birbirine bağlı bir ağ oluşturur ve dürtü alımında ve iletiminde haberci gibi hareket ederler. Sonuç olarak, öğrenme, tahmin etme ve zekâ yeteneğine sahip akıllı bir beyin ortaya çıkar (Kustrin ve Beresford, 2000).

Tablo 2.1: Geleneksel bilgi işlem ve YSA arasındaki yaklaşım farklılıkları (Kustrin ve Beresford, 2000).

Karakteristik Özellikler	Geleneksel bilgi işlem (Uzman Sistemler Dahil)	Yapay Sinir Ağları (YSA)
Öğrenme Metodu	Kurallar (Didaktik Yöntem)	Örnekler ile (Sokratik Yöntem)
Fonksiyonlar İşlem Şekli	Mantıksal Sıralı	Algısal Desen Paralel

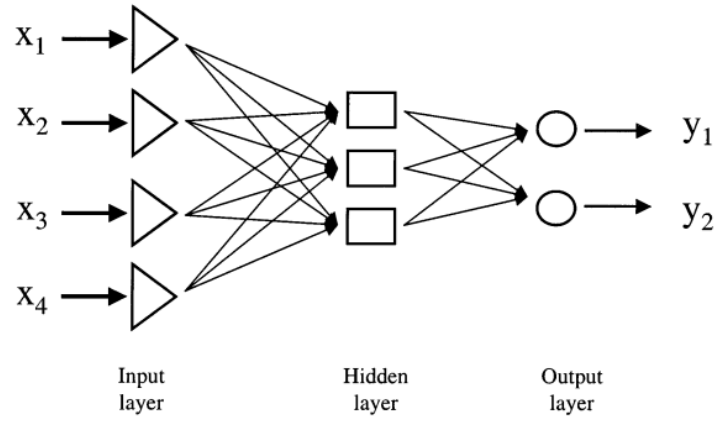
Tablo 2.1’de geleneksel programlama ve yapay sinir ağları modelleri arasındaki farklılıklar görülmektedir.



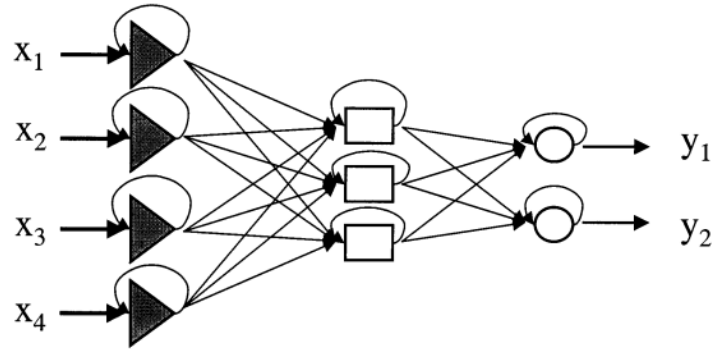
Şekil 2.8: Yapay Sinir Ağı Modeli

Şekil 2.8’de şekli görünen yapay sinir ağlarının bileşenlerinden biri olan yapay nöron, biyolojik nöronun işlevlerini taklit etmek için tasarlanmıştır. Bu yapay nöronlar, giriş sinyallerini bağlantı ağırlıklarıyla çarparak, toplarlar ve ardından bir transfer fonksiyonundan geçirerek nöron için çıktı üretirler. Aktivasyon fonksiyonu, bir nöronun aldığı girdilere (input) göre çıkışını (output) belirleyen matematiksel bir fonksiyondur. Yapay sinir ağları, biyolojiden esinlenerek yapılandırılmış hesaplama modelleridir. Bu yapay nöronlar birbirine bağlanan yüzlerce PE’den oluşur ve bilgi işleme elemanları olarak bilinir. Her bir PE’nin, ağırlıklı girişleri, transfer fonksiyonu ve bir çıkışı bulunur. PE’ler, girdileri ve çıktıları dengeleyen denklemlerdir ve bağlantı ağırlıkları sistemin hafızasını temsil eder, bu yüzden YSA’lara bağlantıcı modeller denir. Tek bir nöron bazı basit bilgi işleme işlevlerini gerçekleştirebilir, ancak sinirsel hesaplamaların gücü, nöronların bir ağda birbirine bağlanmasından gelir. Yapay sinir ağlarının "zekâsı" hala tartışmalıdır. Bu ağlar genellikle birkaç yüz veya birkaç bin PE’ye sahip iken, insan beyni milyarlarca nöron içerir. Dolayısıyla, yapay ağlar hala insan beyninin karmaşıklığının çok gerisindedir. Ancak

YSA'lar, büyük miktarda veriyi işleyebilir ve bazen etkileyici doğrulukta tahminler yapabilir. Bununla birlikte, bu onları "insan" anlamında zeki yapmaz; bu nedenle, bilgisayar zekâsı terimi bu sistemleri daha iyi tanımlar. Farklı türde birçok sinir ağı tasarlanmış olup, yeni tipleri sürekli olarak geliştirilmektedir, ancak hepsi temelde aktarımla açıklanabilir. Nöronlarının işlevleri öğrenme kurallarına ve bağlantı formülüne bağlıdır (Kustrin ve Beresford, 2000).



Şekil 2.9: İleri beslemeli yapay sinir ağı



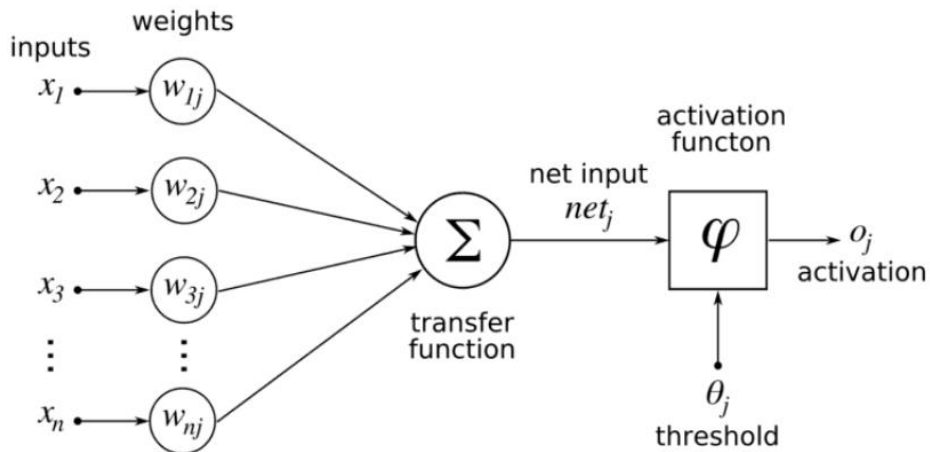
Şekil 2.10: Geri beslemeli yapay sinir ağı

Bir katmandan elde edilen çıktının önceki katmanın girişine veya aynı katmanın kendisine geri gönderildiği bağlantı türüne geri bildirim denir. Bir ağda ileri ve geri besleme olmak üzere iki tür mimari tanımlanabilir. İleri beslemeli mimaride (Şekil 2.9), çıkıştan giriş nöronlarına geri bağlantı yoktur ve bu nedenle önceki çıkış değerlerinin kaydını tutmaz. Geri bildirim mimarisinde (Şekil 2.10) ise çıkıştan giriş nöronlarına bağlantılar bulunur. Her bir nöronun, eğitim hatasını en aza indirmeye çalışırken ek bir serbestlik derecesi sağlayacak

bir girdi olarak ek bir ağırlığı vardır. Bu tür bir ağ, önceki durumun hafızasını tutar, böylece bir sonraki durum yalnızca giriş sinyallerine değil, aynı zamanda ağın önceki durumlarına da bağlıdır.

Yapay zekâ araştırmalarının başlangıcından kısa bir süre sonra, en eski sinir ağı 1940'larda geliştirildi. İlk sinir ağı matematiksel modelini öneren ufuk açıcı bir makale olan "Sinir Aktivitesinde Düşüncelere Ait Bir Mantıksal Hesap" (A Logical Calculus of Ideas Immanent in Nervous Activity) 1943'te yayınlandı. Bu modelin birimi, basit bir biçimlendirilmiş nörondan oluşan McCulloch-Pitts nöronu olarak adlandırılır. Bu matematiksel fonksiyon, biyolojik nöronların modeli olarak tasarlanmış ve yapay bir sinir ağındaki temel birimleri oluşturmuştur. Yapay bir nöronun şekli Şekil 2.11'de gösterilmektedir (Bhardwaj, Di ve Wei, 2018).

İlk yapay sinir ağı modelleri, yalnızca küçük bir sanal nöron kümesinden oluşuyordu ve bunları bağlamak için rastgele ağırlıklar kullanılıyordu. Bu ağırlıklar, her bir simüle edilmiş nöronun kendi aralarında nasıl bilgi aktardığını belirliyordu, yani her bir nöronun 0 ile 1 arasında bir değerle nasıl tepki verdiğini belirlemekteydi. Bu matematiksel gösterimle, sinir çıktısı, görüntüden bir kenar veya şekil veya belirli bir enerji seviyesi içerebilir.

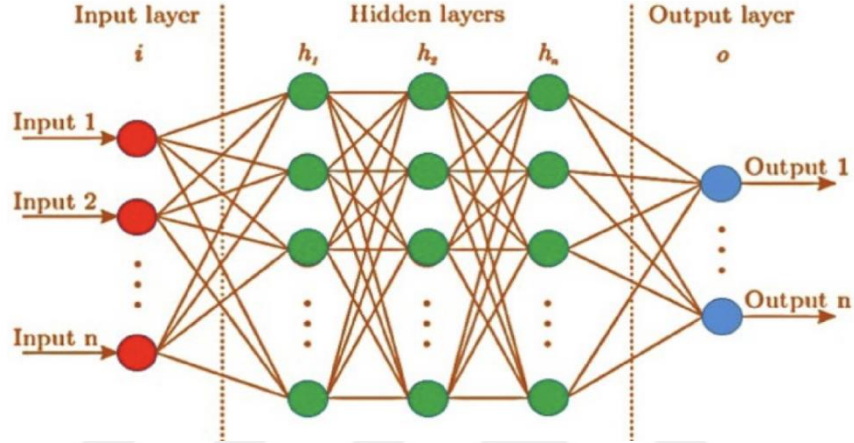


Şekil 2.11: Yapay bir nöron (Bhardwaj vd., 2018)

Şekil 2.11, bir yapay nöronun matematiksel olarak formüle edilmiş bir gösterimini sunmaktadır; burada girdi dendritlere karşılık gelir, bir aktivasyon fonksiyonu, eşikle ulaşıldığında nöronun ateşlenip ateşlenmediğini kontrol eder ve çıktı aksona karşılık gelir.

Ancak, ilk sinir ağı aynı anda yalnızca çok sınırlı sayıda nöronu simüle edebiliyordu, bu nedenle bu basit mimari kullanılarak çok fazla desen tanınmamaktaydı (Bhardwaj vd., 2018).

Yapay Sinir ağı üç katmandan oluşur (Şekil 2.12).



Şekil 2.12: Yapay sinir ağının mimarisi

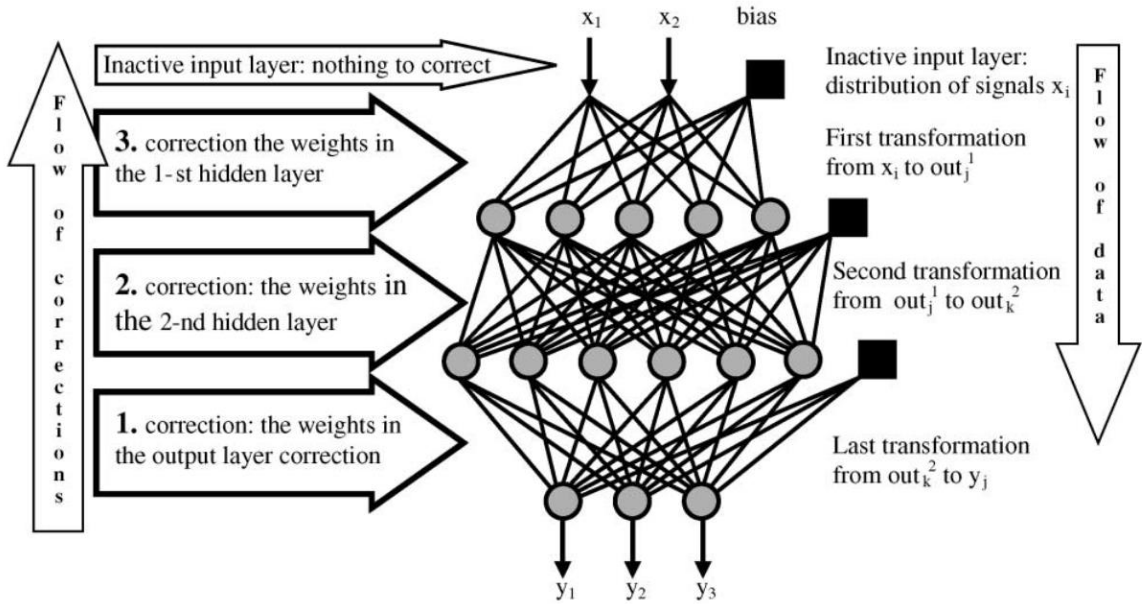
Şekil 2.12’ye göre giriş katmanı; ağ tarafından işlenmesi gereken verileri alır. Gizli katman, giriş katmanı ile çıkış katmanı arasındadır ve birden fazla katmandan oluşabilir. Bu katman giriş katmanından alınan bilgileri işler. Çıkış Katmanı gizli katmandan alınan dönüştürülmüş bilgilere dayanarak nihai çıktıyı verir (Tyagi ve Abraham, 2022).

2.4.1 Yapay Sinir Ağlarında Backpropagation (Geriye Yayılım)

Yapay sinir ağlarının ağırlıklarını ayarlama süreci, birçok modelde karşılaşılan önemli bir zorluktur. Geleneksel yöntemler, zaman alıcı olabilen mutasyon algoritmalarını veya rastgele değerler kullanmayı içeriyordu. Ancak araştırmacılar, bu yöntemlerin sınırlılıklarını göz önünde bulundurarak alternatifler aradılar. Geri yayılım gibi teknikler, daha önce var olan ancak etkinliği sorgulanan yöntemler arasındaydı. Ancak David Rumelhart, Geoffrey Hinton ve Ronald Williams, geri yayılımın potansiyelini anlayarak bu alanda ilerleme sağladılar. 1986’da yayınladıkları "Geriye Yayılan Hatalarla Temsilleri Öğrenmek" başlıklı makale, yapay zekâ alanında önemli bir dönüm noktası oldu. Bu çalışma, geri yayılımın daha hızlı ve güçlü yapay sinir ağlarına olanak sağlayabileceğini açıkça gösterdi. Geri yayılım

süreci matematiksel olarak karmaşık olabilir, ancak temelde, hataları tanımlayıp düzeltmek ve ardından modeli optimize etmek için küçük ayarlamaları içerir. Örneğin, bir girişin çıkışının 0,6 olduğunu varsayalım. Bu durumda, hatanın 0,4 (1,0 eksi 0,6) olduğunu görebiliriz; bu, beklenen çıktının ortalamasının altında olduğunu gösterir. Ancak, eğitimi geriye yayarak yeni bir çıktı elde edebiliriz, örneğin 0,65. Eğitim, çıktının 1'e daha da yaklaşması için devam edecektir. Bu süreç, istenen çıktıya ulaşana kadar tekrarlanır (Taulli, 2019).

Geriye yayılan hata öğrenmesi, denetimli bir öğrenme yöntemidir. Bu yöntemde, bir dizi girdi ve hedef çifti $\{X; Y\}$ mevcuttur: Her X girdisi, m boyutlu bir nesneyi temsil ederken, Y hedef çıktısı ise n boyutlu bir yanıtı ifade eder. Örneğin, bir ürünün m bileşen analizi, bir bileşiğin m yoğunluk spektrumu veya bir kimyasal prosesin m deneysel koşulları olabilir ve bu girdilere karşılık n özellikli bir ürün, bileşik veya reaksiyon elde etmek istenir.



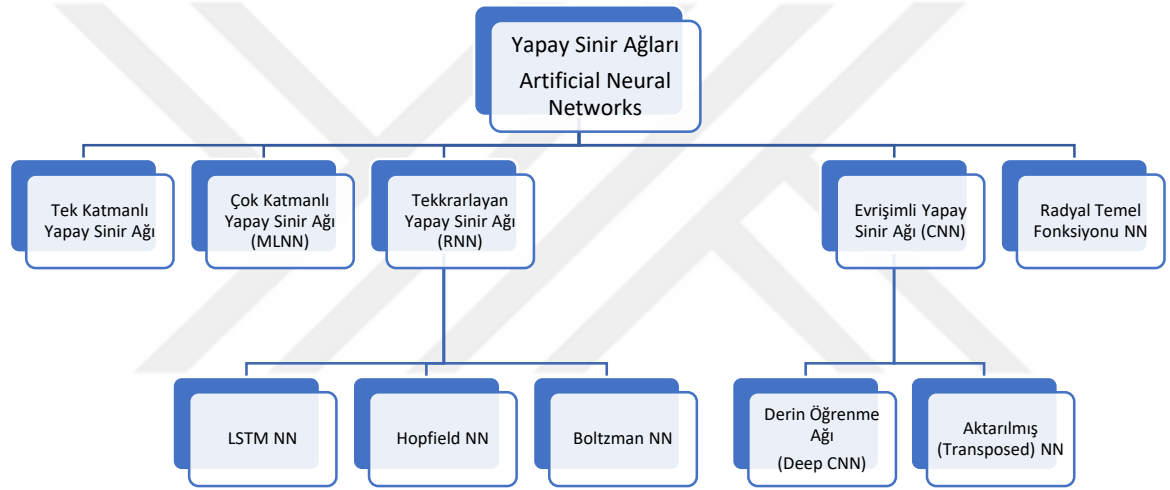
Şekil 2.13: Hatanın geriye yayılımının öğrenilmesindeki ağırlık düzeltmelerinin sırası (Zupan, 2003)

Şekil 2.13'te gösterildiği üzere geriye yayılan hata stratejisiyle model oluşturulurken, her X ve Y çifti ($X; Y$) arasındaki karşılaştırmaya dayanır. Bu yöntem, bilinen gerçek çıktı (Y) ile her girdi/çıktı çifti arasındaki hata ($X; Y$) üzerinde odaklanır ve öğrenme sırasında nöronlardaki ağırlıklar buna göre düzeltilir. Bu düzeltmeler, çıkıştan giriş katmanına doğru, yani girdilerden başlayarak geriye doğru yayılan sinyallerle gerçekleşir. Özetle, her

katmandaki her nöronun ağırlıkları, eğitim sırasında belirli bir formüle göre ayarlanır (Zupan, 2003).

2.4.2 Yapay Sinir Ağları Mimarileri

Yapay sinir ağı mimarileri kullanım amaçlarına göre farklılık gösterir. Örneğin bir görüntünün analizini yapmayı sağlayan CNN mimarisi, Konuşma tahmini çalışmaları için RNN mimarisi, el yazısı tanıma uygulamaları için LSTM mimarileri kullanılır.



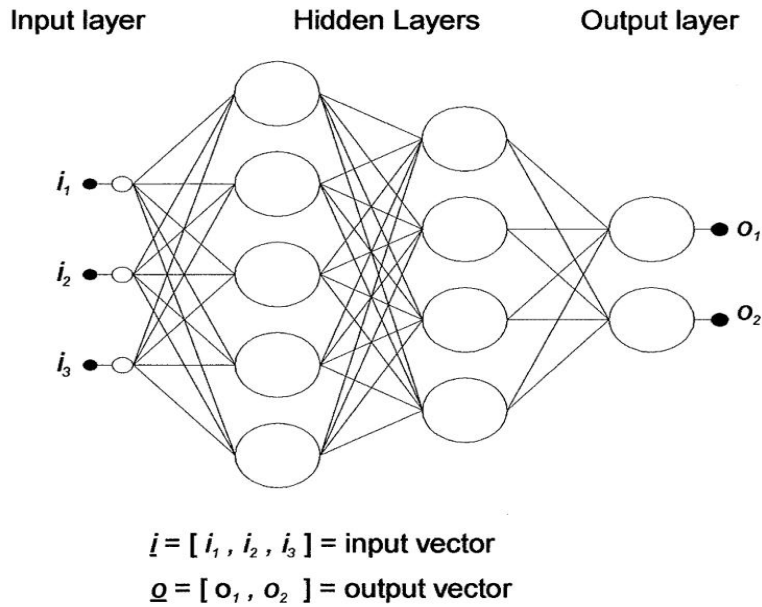
Şekil 2.14: Yapay sinir ağları mimarileri

Şekil 2.14’te yapay sinir ağı mimarileri gösterilmiştir.

2.4.2.1 Çok Katmanlı Algılayıcı (Multi Layer Perceptron)

Katmanlar halinde düzenlenmiş bir algılayıcı grubu olan sinir ağları, genellikle ileri yönde işleyen ve nöronlar arasında çoklu bağlantılara sahip derin sinir ağları olarak isimlendirilebilir. Bu ağlar, herhangi bir doğrusal olmayan fonksiyonu öğrenebilecek yetenekte oldukları için evrensel fonksiyon yaklaşımıcıları olarak da tanımlanır. Sinir ağlarının bu özellikleri, geniş bir uygulama alanına sahip olmalarını sağlar ve birçok karmaşık problemin modellenmesinde etkili bir şekilde kullanılabilirler (Tyagi ve Abraham, 2022).

Çevresel modelleme, genellikle çeşitli yaklaşımların bir kombinasyonunu gerektirir. Hangi yaklaşımın en uygun olduğunu belirlemek, ele alınan problemin karmaşıklığına ve anlaşılma düzeyine bağlıdır. Yeterli veri ve hesaplama kaynakları varsa ve problemin sağlam bir teorik temeli varsa, tam bir sayısal model en arzu edilen çözüm olabilir. Ancak, genellikle problemin karmaşıklığı arttıkça, teorik anlayış azalır ve sistemler arasındaki etkileşimlerin net olmaması nedeniyle istatistiksel yaklaşımlara ihtiyaç duyulur. Son zamanlarda, sınır ağlarının ve özellikle çok katmanlı algılayıcının, geleneksel istatistiksel tekniklere etkili alternatifler olduğu gösterilmiştir (Schalkoff, 1992). Hornik vd.'ne (1989) göre çok katmanlı algılayıcı, çeşitli pürüzsüz ve ölçülebilir fonksiyonlar ile eğitilebilir ve veri dağılımı hakkında önceden bir varsayımda bulunmadan kullanılabilir. Bu nedenle, doğrusal olmayan fonksiyonları modelleyebilir ve yeni verilerle genelleme yapmak üzere eğitilebilir. Bu özellikler, çok katmanlı algılayıcıyı sayısal modellerin geliştirilmesinde ve istatistiksel yaklaşımlar arasında seçim yaparken çekici bir alternatif haline getirir. Atmosfer bilimlerinde, çok katmanlı algılayıcının birçok uygulaması bulunmaktadır (Gardner ve Dorling, 1998).

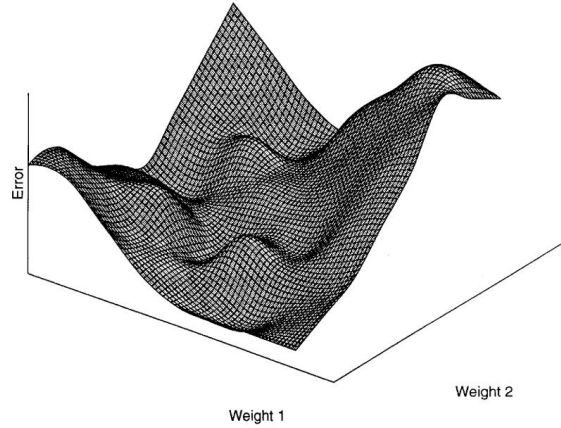


Şekil 2.15: 2 Gizli katmandan oluşan çok katmanlı algılayıcı

Multi layer perceptron (MLP), Şekil 2.15'te gösterildiği gibi, bir giriş vektöründen bir çıkış vektörüne doğrusal olmayan bir işlemeyi gerçekleştiren bir modeldir. Bu model, basit

nöronlar veya düğümler arasındaki bağlantıdan oluşur. Her düğüm, aldığı girdilerin ağırlıklı toplamını hesaplar ve bu toplamı bir aktivasyon fonksiyonundan geçirerek çıkış üretir. Bu, MLP' nin doğrusal olmayan fonksiyonlara yaklaşmasını sağlar. Ayrıca, her düğüm, bir sonraki katmandaki düğümlere girdi olarak çıktısını ileterek bilgiyi işler. Bu nedenle, MLP, ileri beslemeli bir sinir ağı olarak bilinir. Mimari değişken olsa da genellikle birkaç nöron katmanından oluşur. Giriş katmanı, yalnızca giriş vektörünü ağı aktarır ve herhangi bir hesaplama yapmaz. Giriş ve çıkış vektörleri, çok katmanlı algılayıcının giriş ve çıkışlarını ifade eder. MLP, tamamen bağlı olarak tanımlanır, yani her düğüm bir önceki ve bir sonraki katmandaki her düğüme bağlıdır (Gardner ve Dorling, 1998).

Çok katmanlı algılayıcılar, uygun bağlantı ağırlıkları ve transfer fonksiyonları seti seçilerek giriş ve çıkış vektörleri arasındaki herhangi bir düzgün, ölçülebilir fonksiyona yaklaşabilmektedir (Hornik, Stinchcombe ve White, 1989). Bu algılayıcıların öğrenme yeteneği, eğitim süreciyle geliştirilir ve bu süreç, bir dizi girdi ve bunlara karşılık gelen çıktı vektörlerini içeren bir eğitim veri seti gerektirir. Eğitim sırasında, ağı tekrar tekrar eğitim verileriyle beslenir ve ağırlıklar, istenen giriş-çıkış ilişkisi elde edilene kadar ayarlanır. Ağırlıkları olası tüm değerlere göre değiştirerek ve hataları üç boyutlu uzayda çizerek,



Şekil 2.16: Çok katmanlı algılayıcı için hata yüzeyi

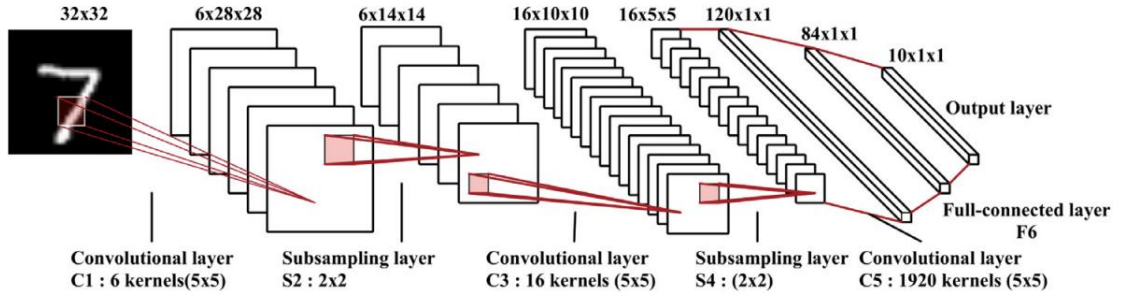
Şekil 2.16'da gösterilene benzer bir çizim elde edilebilir. Bu öğrenme süreci denetimli bir şekilde gerçekleşir; yani, her bir giriş vektörü için algılayıcıdan elde edilen çıktı, istenen çıktı ile karşılaştırılır ve aradaki fark hata sinyali olarak tanımlanır. Eğitim, bu hata sinyalinin büyüklüğünü kullanarak ağıdaki ağırlıkların ne dereceye kadar ayarlanması gerektiğini

belirler, böylece algılayıcı genel hatasını azaltır. Çok katmanlı bir algılayıcıyı eğitmek için çeşitli algoritmalar kullanılabilir ve uygun temsili eğitim verileriyle eğitildikten sonra, yeni ve görünmeyen giriş verilerine genelleme yapılır (Gardner ve Dorling, 1998).

2.4.2.2 Evrişimli Sinir Ağı (Convolutional Neural Networks)

Evrişimli sinir ağları, çok katmanlı algılayıcının biyolojik olarak esinlenmiş türleridir ve görüntü tanıma ve sınıflandırma gibi alanlarda oldukça etkilidir. Dönüşüm ağları, yüzler, nesneler ve trafik işaretleri gibi nesneleri tanımlamanın yanı sıra robotlarda ve otonom araçlarda görüşü güçlendirmede başarılı bir şekilde kullanılmıştır. Convolutional Neural Networks (CNN)'ler, bitişik katmanlardaki nöronlar arasında yerel bir bağlantı modeli kullanarak mekânsal olarak yerel korelasyondan faydalanırlar. Başka bir deyişle, gizli birimlerin bir katmanında, girdiler, bir önceki katmandaki birimlerin bir alt kümesinden gelir ve bu birimler birbirine bitişik alıcı alanlara sahiptir. LeNet, 1988'de Yann LeCun tarafından önerilen ilk evrişimli sinir ağlarından biriydi. Esas olarak posta kodlarını, rakamları vb. okumak gibi karakter tanıma görevleri için kullanılıyordu. 2012 yılında Alex Krizhevsky ve Hinton, derin öğrenmede yeniden canlanma çağını başlatan CNN'i kullanarak sınıflandırma hatasını %26'dan %15'e düşürerek inanılmaz bir gelişmeyle imageNet yarışmasını kazandılar (Bhardwaj vd., 2018).

Evrişimli sinir ağları, özellikleri otomatik olarak çıkarabilme yeteneği ile bilinir. Görüntü tanıma ve sınıflandırma gibi görevlerde etkili olan CNN'ler, görüntünün mekânsal özelliklerini kullanarak çalışır. Ayrıca, CNN'ler parametre paylaşımını benimseyerek işlevlerini gerçekleştirirler (Tyagi ve Abraham, 2022). Bu parametre paylaşımı sayesinde, ağ, her bir özelliği belirlemek için her bir görüntü parçası üzerinde ayrı ayrı çalışmak yerine, aynı parametreleri kullanarak verimli bir şekilde işlem yapabilir. CNN'nin tasarımı ve bakımı, genellikle karmaşıktır. Bu ağlar, genellikle daha fazla sayıda gizli katman ve büyük miktarda giriş verisi kullanıldığı için işlemleri yavaşlatır (Tyagi ve Abraham, 2022). Bu durum, özellikle büyük veri setleri veya karmaşık görsel veri türleri üzerinde çalışırken belirgin hale gelir. Bu nedenle, CNN'lerin etkin bir şekilde kullanılması ve optimize edilmesi, zaman ve kaynak gerektirmektedir.

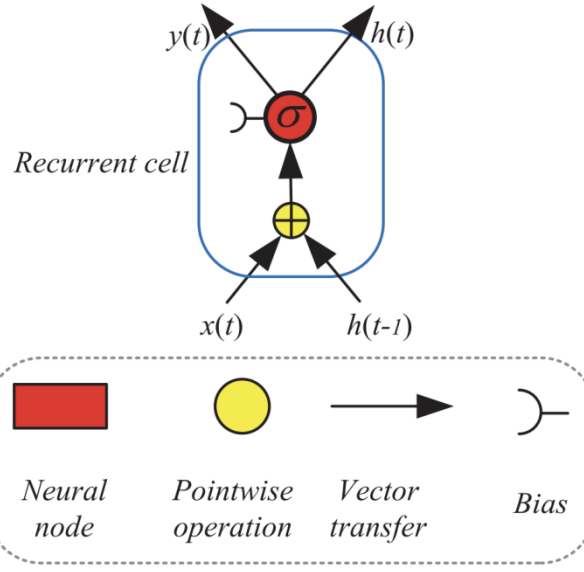


Şekil 2.17: LeNet-5 ağı (Gu vd., 2018)

Literatürde birçok farklı CNN (Convolutional Neural Network) mimarisi vardır. Ancak, temel bileşenleri oldukça benzerdir. Şekil 2.17’de gösterilen LeNet-5’i örnek olarak alırsak, üç tür katmandan oluşur: evrişimsel katman, havuzlama katmanı ve tam bağlantılı katman. Evrişimsel katman, girdilerin özellik temsillerini öğrenmeyi amaçlar. Evrişim katmanı, farklı özellik haritalarını hesaplamak için kullanılan birkaç evrişim çekirdeğinden oluşur. Özellikle, bir özellik haritasındaki her nöron, önceki katmandaki komşu nöronların bir bölgesine bağlanır. Bu tür bir komşuluk, önceki katmandaki nöronun alıcı alanı olarak adlandırılır. Yeni özellik haritası, girdinin öğrenilmiş bir çekirdek ile önce evriştirilmesi ve ardından evriştirilen sonuçlara eleman bazında bir doğrusal olmayan aktivasyon fonksiyonunun uygulanmasıyla elde edilir. Havuzlama katmanı, özellik haritalarının çözünürlüğünü azaltarak kayma-invaryansı elde etmeyi amaçlar. Genellikle iki evrişim katmanı arasına yerleştirilir. Bir havuzlama katmanının her bir özellik haritası, önceki evrişim katmanının karşılık gelen özellik haritasına bağlıdır (Gu vd., 2018). Birkaç evrişim ve havuzlama katmanından sonra, yüksek seviyeli çıkarım yapmayı amaçlayan bir veya daha fazla tam bağlantılı katman olabilir (Simonyan ve Zisserman, 2014).

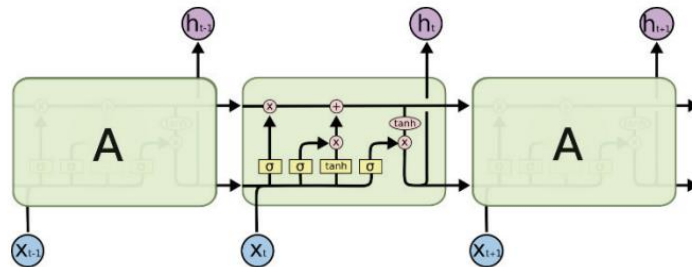
2.4.2.3 Tekrarlayan Sinir Ağları (Recurrent Neural Networks) ve Uzun Kısa Süreli Bellek Tekrarlayan Sinir Ağı (Long Short-Term Memory Networks)

Tekrarlayan bir sinir ağı fonksiyonu yalnızca girişi değil aynı zamanda zaman içindeki önceki girişleri de işler. Bunun bir örneği, bir mesajlaşma uygulamasına karakter girdiğinizde meydana gelen durumdur. Siz yazmaya başladığınızda sistem kelimeleri tahmin edecektir. Yani "O" seçeneğine dokunursanız bilgisayar "O", "Merhaba" ve "İşte" önerecektir. Recurrent Neural Networks (RNN) aslında karmaşık algoritmalara dayalı olarak birbirini besleyen bir sinir ağları dizisidir (Taulli, 2019).



Şekil 2.18: Standart tekrarlayan sigma hücresinin şeması (Yu vd., 2019)

Şekil 2.18’de şeması verilen RNN’ler genellikle sigma hücreleri ve tanh hücreleri gibi standart tekrarlayan hücrelerden oluşan ağlardır. RNN’lerde tekrarlayan katmanlar veya gizli katmanlar, durumları hem geçmiş durumlardan hem de geri besleme bağlantılarıyla mevcut girişten etkilenen tekrarlayan hücrelerden oluşur. Tekrarlayan katmanlar, farklı RNN’ler oluşturmak için çeşitli mimarilerde düzenlenebilir. Bu nedenle, RNN’ler temel olarak tekrarlayan hücre ve ağ mimarisiyle ayırt edilir. Farklı hücreler ve iç bağlantılar, RNN’lerin farklı kapasitelere sahip olmasını sağlar. RNN mimarisinin tipik özelliği, RNN’nin geçmiş durumlara ve mevcut giriş verilerine dayalı olarak mevcut durumu güncelleme kapasitesine sahip olmasını sağlayan döngüsel bir bağlantıdır (Yu vd., 2019).



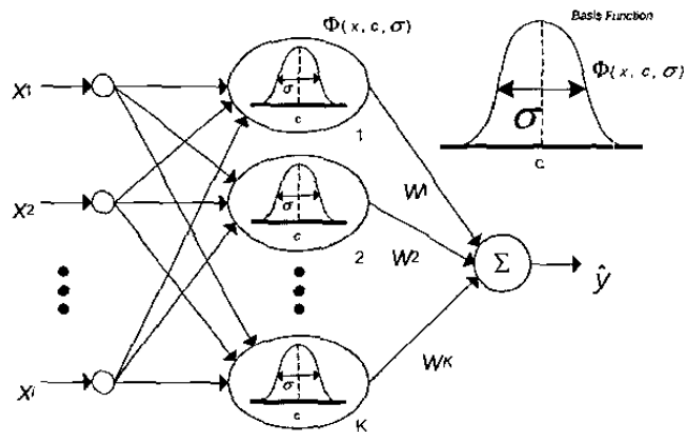
Şekil 2.19: LSTM hücresi (Prakash, Jalal ve Pathak, 2023)

Şekil 2.19'da grafiksel gösterimi verilen Long Short-Term Memory Networks (LSTM) hücreleri, uzun vadeli bağımlılıkları tanıyabilen benzersiz bir RNN sınıfıdır. LSTM'ler kasıtlı olarak uzun vadeli bağımlılık sorununu önlemek için oluşturulmuştur. Pratik olarak, uzun süreli hafıza tutmayı varsayılan mod haline getirecek şekilde hareket ederler. Bu nedenle zaman serisi tahmini için idealdirler (Prakash vd., 2023).

LSTM'lerin dezavantajları ise değerlendirme sırasında verileri hücreden hücreye taşıma sorunuyla karşı karşıya olmaları, hücrelerin yüksek bellek bant genişliğine ihtiyaç duyması nedeniyle öğrenme sürecinin uzun olması, aşırı uyum karşısında savunmasız olmalarıdır. LSTM, jest tanıma, konuşma tanıma, metin tahmini ve konuşma işlemede kullanılır (Tyagi ve Abraham, 2022).

2.4.2.4 Radyal Temel Fonksiyonu (Radial Basis Function)

Radyal Basis (RBF) fonksiyonları, işlev yaklaşımı, örüntü tanıma ve zaman serisi tahmini gibi problemleri çözmek için kullanılır. Bu ağların basit yapısı, aşamalı öğrenmeyi sağlar ve eğitim süresini kısaltır, bu da bu tür ağların birçok pratik probleme uygulanmasını kolaylaştırır. Ayarlanabilir parametreleri arasında, alıcı alan merkezleri (temel fonksiyonların konumu), genişlik (yayılma), alıcı alanın şekli ve doğrusal çıktı ağırlıkları bulunmaktadır (Panchapakesan vd., 2002).



Şekil 2.20: RBF ağı yapısı (Arisariyawong ve Charoenseang, 2002)

Şekil 2.20’de gösterilen RBF sinir ağı yapısı genellikle iki katmanlı bir şekilde tasarlanır. İlk katmanda, belirli bir merkez kümesi için giriş alanını yeni bir alana haritalayan sabit, doğrusal olmayan bir dönüşüm gerçekleştirilir. Her bir ϕ_k terimi, gizli katmandaki bir birimin aktivasyon fonksiyonunu temsil eder. Daha sonra çıkış katmanı, bu yeni alandaki aktivasyonların doğrusal bir kombinasyonunu uygular. Bu sayede, RBF sinir ağı, karmaşık ilişkileri yakalamak için giriş verilerini uygun bir şekilde dönüştürür ve çıktıları elde eder (Arisariyawong ve Charoenseang, 2002).

2.4.3 Yapay Sinir Ağı Aktivasyon Fonksiyonları

Bir YSA’daki aktivasyon fonksiyonu (AF), girdi ve sapmaların ağırlıklı toplamının bir miktar gradyan kullanarak bir çıktıya aktarılmasını tanımlar. AF, doğrusal giriş sinyali doğrusal olmayan bir çıkış sinyaline dönüştürür, bu da özellikle yüksek dereceli hesaplamalar için öğrenme sürecini geliştirir. AF, sınıflandırmalar veya tahminler gibi farklı hesaplamalar gerçekleştirir. Bir ağda AF’ nin konumu da önemlidir. AF, gizli katmanın sonrasına yerleştirilirse daha fazla yayılma için doğrusal bir veriyi doğrusal olmayan bir veriye dönüştürür ve çıktı katmanına yerleştirilirse tahminler üzerinde çalışır. Farklı görev alanlarında AF seçimi de önemlidir çünkü ağın çıktısını kontrol eder ve böylece sonuçları iyileştirir. Aktivasyon fonksiyonları doğrusal ve doğrusal olmayan aktivasyon fonksiyonları olarak iki türe ayrılabilir. Doğrusal(linear) aktivasyon fonksiyonu, bir yapay sinir ağında çıkışın doğrudan girişle orantılı olduğu bir ilişki kurar, yani $y=A_x$ şeklinde ifade edilir. Ancak, bu tür bir aktivasyon fonksiyonu ikili aktivasyon için uygun değildir. Çünkü doğrusal aktivasyon fonksiyonuna sahip bir ağ, türevinin sabit ve girişe bağımlı olmadığı için geri yayılım algoritmasını kullanamaz. Doğrusal olmayan (non-linear) aktivasyon fonksiyonları, sinir ağlarında yaygın olarak kullanılır çünkü bunlar, girdi verisinin karmaşıklığını ele almak ve daha karmaşık ilişkileri modellemek için gereklidir (Tyagi ve Abraham, 2022). Bu aktivasyon fonksiyonları birçok sinir ağı mimarisinde verilerin işlenmesine yardımcı olur ve ağın genel performansını artırır. Sigmoid, Tanh, ReLU, Leaky ReLU, Softmax başlıca nonlinear aktivasyon fonksiyonlarıdır.

Bir yapay sinir ağındaki aktivasyon fonksiyonunun seçimi, modelin daha iyi performans gösterebilmesi ve daha doğru sonuçlar verebilmesi için göz önünde bulundurulması gereken önemli hiper parametrelerden biridir (Sharma, Sharma ve Athaiya, 2017).

2.4.3.1 Rectified Linear Unit (ReLU)

Rectifier Linear Unit (ReLU) yani doğrultulmuş lineer birim, sigmoid ve tanh gibi diğer aktivasyon fonksiyonlarına kıyasla hesaplaması daha basit ve etkilidir. Doğrusal ve doymamış bir formu olduğu için, ReLU'nun yakınsama hızını altı kat arttırdığı kanıtlanmıştır. Ayrıca, pahalı üstel işlemler içeren tanh veya sigmoid fonksiyonlarından farklı olarak, ReLU aktivasyonunu sıfır noktasında basitçe eşikleyerek elde edilebilir. Bu nedenle, son birkaç yılda oldukça popüler hale gelmiştir ve günümüzde neredeyse tüm derin öğrenme modelleri tarafından tercih edilmektedir. ReLU'nun önemli bir avantajı da kaybolan gradyan sorununu önlemesi veya düzeltmesidir (Bhardwaj vd., 2018).

ReLU fonksiyonu doğrusal olmayan bir yapıya sahiptir. Sıfır odaklı değildir; negatif girdiler için sıfır değerini döndürürken, pozitif girdileri olduğu gibi kullanır. Hesaplaması oldukça basittir ve biyolojik nöronlara daha yakın bir işleyiş gösterir (Aydın, 2021).

2.4.3.2. Sigmoid Fonksiyonu

Bir sigmoid fonksiyonu kendine özgü bir S şekline sahiptir ve herhangi bir gerçek giriş değeri için türevlenebilir bir gerçek fonksiyondur. Aralığı 0 ile 1 arasındadır. Ancak, sigmoid fonksiyonu kullanıldığında, gradyanın kaybolması veya gradyanın doygunluğu gibi sorunlarla karşılaşılabilir. Ayrıca, sigmoid fonksiyonunun yavaş bir yakınsama sağladığı bilinmektedir. Bu nedenle, pratikte sigmoid fonksiyonu yerine ReLU tercih edilir hale gelmiştir (Bhardwaj vd., 2018).

2.4.3.3 Tanh Fonksiyonu

Tanh fonksiyonu sigmoid fonksiyonuna benzer bir yapıya sahiptir. Ancak, sigmoid fonksiyonu girdileri 0 ile 1 arasında değerler alırken, tanh fonksiyonunun girişleri -1 ile 1 arasında bulunur (Aydın, 2021).

2.4.3.4 SoftMax

Softmax fonksiyonu, birden fazla sigmoid fonksiyonunun birleşimidir. Bir sigmoid fonksiyonunun 0 ila 1 aralığında değerler döndürdüğünü bildiğimiz için, bunlar belirli bir sınıfın veri noktalarının olasılıkları olarak ele alınabilir. Softmax fonksiyonu, ikili sınıflandırma için kullanılan sigmoid fonksiyonlarının aksine, çok kategorili sınıflandırma problemleri için kullanılabilir (Sharma vd., 2017).

2.4.4 Yapay Sinir Ağı Optimizasyon Fonksiyonları

ConvNet (Görsel veri analizi ve görüntü tanıma gibi görevlerde kullanılan derin sinir ağı) gibi derin sinir ağlarının eğitimi büyük ölçüde zorlu bir dışbükey olmayan ve yüksek boyutlu optimizasyon problemidir. Alan sorununa ve veri kümelerine bağlı olarak, derin öğrenme araştırmacısının amacı, kayıp fonksiyonunu en aza indirmek için hiper parametre ayarlaması yoluyla daha iyi ve daha hızlı sonuçlar üretecek bir model uygulamaktır. Derin sinir ağlarında ağırlıkları en iyi şekilde ayarlamak, doğru model sınıflandırması veya tahmini üretmenin anahtarıdır. Ancak ağırlıkları ayarlamak, en düşük kayıp fonksiyonuna (gradient descent) sahip ağırlıkları ve eğim inişi (gradient descent) sırasında eğimin değişim yönünü (geri yayılım) ayarlamaya bağlıdır. Yerel minimum tuzağı sorununu önlemek için, istenen mutlak küresel minimum değeri elde etmenin yanı sıra, boyutluluk sorunu gibi eğim inişi ile ilişkili diğer zorlukların yanı sıra, son yıllarda birçok optimizasyon algoritması önerilmiştir; eğilim çoğunlukla otomatik olarak çalışan ve hiper parametrelerin çok az ayarlanmasına izin veren uyarlanabilir tahminlere dayanmaktadır. Dışbükey olmayan problemler, sinir ağları için en yaygın durumdur, bu nedenle bu ağlarda küresel optimumları bulmaya çalışan optimizasyon stratejisi seçmek, yüksek boyutlu arama alanında çok sayıda parametrenin tahmin edilmesi süreci nedeniyle genellikle zordur. Uygunsuz bir optimizasyon tekniği, ağı eğitimi sırasında herhangi bir iyileştirme olmaksızın yerel minimumlarda kalmasına neden olabilir. Bu nedenle, davranışlarını daha iyi anlamak için kullanılan modele ve veri setine bağlı olarak optimize edicilerin performansını analiz etmek önemlidir (Dogo vd., 2018).

2.4.4.1 Gradyan Tabanlı Optimize Ediciler

Gradyan inişi (gradient descent), optimizasyon gerçekleştirmek için en popüler algoritmalarından biri olarak öne çıkmakta ve makine öğrenimi görevlerini optimize etmenin en yaygın yolu olmaktadır. Stokastik versiyonu son yıllarda dikkat çekmiştir ve bu özellikle derin sinir ağlarını optimize etmek için geçerlidir. Derin sinir ağlarında, tek bir örnek veya bir örnek grubunun izlediği gradyan, hesaplama kaynaklarını tasarruf etmek kullanılır (Lu, 2022).

Stokastik Gradient Descent (SGD), gradyan inişi optimizasyonunun bir çeşididir ve hedef fonksiyonun minimum veya maksimum değerini bulmak için kullanılan yinelemeli bir yöntemdir. Bu stokastik yaklaşım, her yinelemede sadece bir örnek veya küçük bir örnek grubu üzerinde çalışarak, fonksiyonun en düşük veya en yüksek değerlerini daha verimli bir şekilde keşfetmeye çalışır. Bu sayede, özellikle büyük veri kümelerinde daha hızlı ve hesaplama açısından daha verimli bir optimizasyon sağlar (Zinkevich vd., 2010). Eğitim veri kümelerinin sayısı yüksek olduğunda, SGD her entegrasyonda tüm eğitim örneklerini kullanmadığı için hızlı çalışır, bu da hesaplama miktarını azaltır ve hesaplama hızını artırır. Dahası, SGD her parametrenin gradyanının birinci ve ikinci dereceden matrislerinin tahminini kayıp fonksiyonuna göre dinamik olarak ayarlayabilir. Böylece, modelin yerel optimuma yakınsama riski azaltılabilir. Bu avantajları göz önünde bulundurarak, SGD kullanımının hesaplama maliyetinin üstesinden gelebileceğini söylenebilir (Yang ve Yang, 2018).

Yıllar geçtikçe stokastik gradyan tabanlı optimizasyon, bilgisayarlı görü ve otomatik konuşma tanıma işleme dahil olmak üzere bilim ve mühendisliğin çeşitli alanlarında temel bir yöntem olarak ortaya çıkmıştır (Krizhevsky, Sutskever ve Hinton, 2012). Bu yöntemler, derin sinir ağlarını (DNN) eğitmek için kullanılır. Derin sinir ağları, belirli bir görev için geliştirildiğinde, ağın eğitimiyle ilgili bazı hiper-parametreler belirlenir. Bu hiper-parametreler genellikle deneme-yanılma yöntemiyle seçilir. Her olası hiper-parametre kombinasyonu için yeni bir ağ sıfırdan eğitilir ve değerlendirilir. Tek bir ağ yapısını eğitmek için donanım ve yazılım alanında önemli ilerlemeler kaydedilmiş olsa da birçok farklı yapı denemek hala çok zaman alır. Bu yüzden, hiper-parametrelere duyarlı olmayan bir stokastik optimize ediciye ihtiyaç duyulur (Lu, 2022).

Çok büyük ve oldukça fazla tekrar içeren eğitim kümelerine sahip büyük sinir ağları için, mini toplu öğrenme (Mini batch gradient descent) yöntemini kullanmak neredeyse her zaman en verimli yaklaşımdır. Karmaşık yöntemlere uyum sağlarken mini kümelerin (mini-batch) oldukça büyük olması gerekebilir. Büyük mini-kümeler hesaplama açısından daha verimlidir (Hinton, Srivastava ve Swersky, 2012).

2.4.4.2 RMSProp (Kök Ortalama Kare Yayılmı)

Bu optimize algortimasında, gradyanların hareketli (indirimli) ortalaması korunur ve bu ortalamayı tahmin etmek için değişkenin köküne bölünmüş olan degradelerin karesi kullanılır. Bu işlem, RMSprop'un uygulamasında kullanılan bir yöntemdir. Ancak bu versiyon, Nesterov momentumunu değil, düz momentumu kullanır. Bu yaklaşım, gradyanların hareketli ortalamasını güncelleyerek ve gradyanların karelerinin hareketli ortalamasını koruyarak gradyanları düzgün bir şekilde düzenler. Bu, modelin daha iyi ve istikrarlı bir şekilde öğrenmesine olanak tanır (TensorFlow, 2024).

2.4.4.3 Adam (Uyarlanabilir Moment Tahmini)

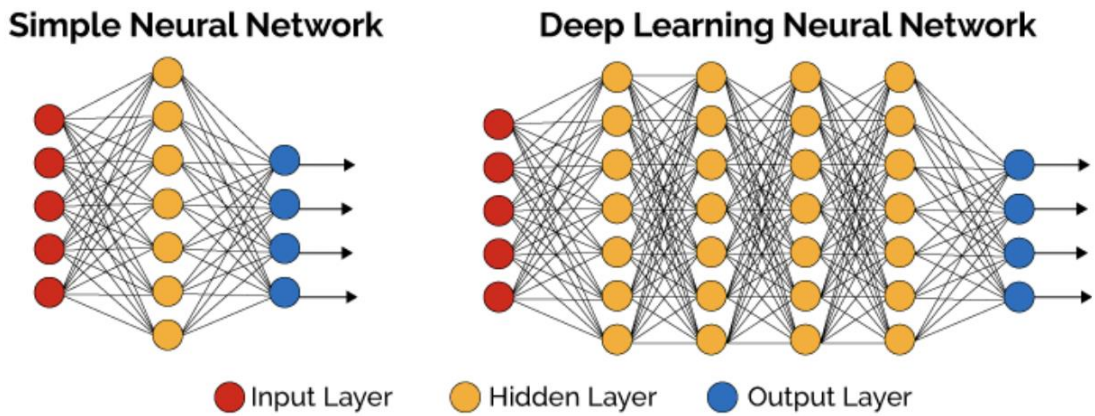
Tabanını düşük dereceli momentlerin uyarlanabilir tahminlerine dayandıran Adam, birinci dereceden gradyan tabanlı optimizasyon için bir algoritmadır. Bu yöntem basit uygulanabilirlik, hesaplama verimliliği ve düşük bellek gereksinimi gibi avantajlar sunar. Ayrıca, gradyanların çapraz yeniden ölçeklendirilmesine karşı değişmezdir ve büyük veri veya parametre setlerine sahip problemlerde etkili bir şekilde çalışır. Adam, durağan olmayan hedefler ve çok gürültülü veya seyrek eğimlere sahip problemlerde de etkili bir şekilde kullanılabilir. Hiper parametrelerin sezgisel yorumları bulunmakla birlikte genellikle az ayarlama gerektirir. Ayrıca, algoritmanın teorik yakınsaması da analiz edilmiştir. Ampirik sonuçlar, Adam'ın diğer stokastik optimizasyon yöntemlerine kıyasla olumlu sonuçlar verdiğini ve pratik uygulamalarda etkili olduğunu göstermektedir (Kingma ve Ba, 2014).

2.5. Derin Öğrenme

Derin öğrenme ile makine öğrenimi arasında genellikle karışıklık yaşanır. Her ikisi de oldukça karmaşıktır ve birçok benzerliğe sahiptir. Ancak, temel farklar mevcuttur. Örneğin,

bir atı bulmak için bir algoritma oluşturmak istediğimizi düşünelim ve binlerce hayvanın fotoğrafına sahip olduğumuzu varsayalım. Makine öğrenimi, bu fotoğrafları tek başına analiz edemez; bunun yerine verilerin etiketlenmesi gerekir. Bu işlem, genellikle denetimli öğrenme olarak bilinen bir süreç aracılığıyla gerçekleşir. Makine öğrenimi algoritması, atları tanıyacak şekilde eğitilir. Ancak, bu yöntemin sınırlamaları vardır. Özellikle, atın özelliklerini (şekil, toynaklar, renk ve yükseklik gibi) tanımlayan özellik çıkarma adı verilen bir işlem gereklidir. Bu işlem, modelin doğruluğunu etkileyebilecek aykırı değerleri veya istisnaları dikkate almak zorundadır. Bununla birlikte, özellik çıkarma süreci büyük miktarda veriyi göz ardı etme riski taşır ve bazen karmaşık hale gelebilir. Örneğin, bilgisayar virüslerinin imza yapısını takip etmek ve bu yapıları tahmin etmek zor olabilir. Bu nedenle, siber güvenlik yazılımları genellikle virüsün verdiği zararın ardından imza toplamakla ilgilidir. Ancak derin öğrenme, bu tür sorunları çözmek için etkili bir yaklaşımdır. Bu yöntem, tüm verileri detaylı bir şekilde analiz eder ve ardından insan beynini taklit eden bir sinir ağı kullanarak ilişkileri bulur. Derin öğrenme, makine öğreniminin alt dalıdır ve büyük miktarda veriyi işleyerek insanların genellikle fark edemediği ilişkileri ve kalıpları keşfetmek için kullanılır (Taulli, 2019).

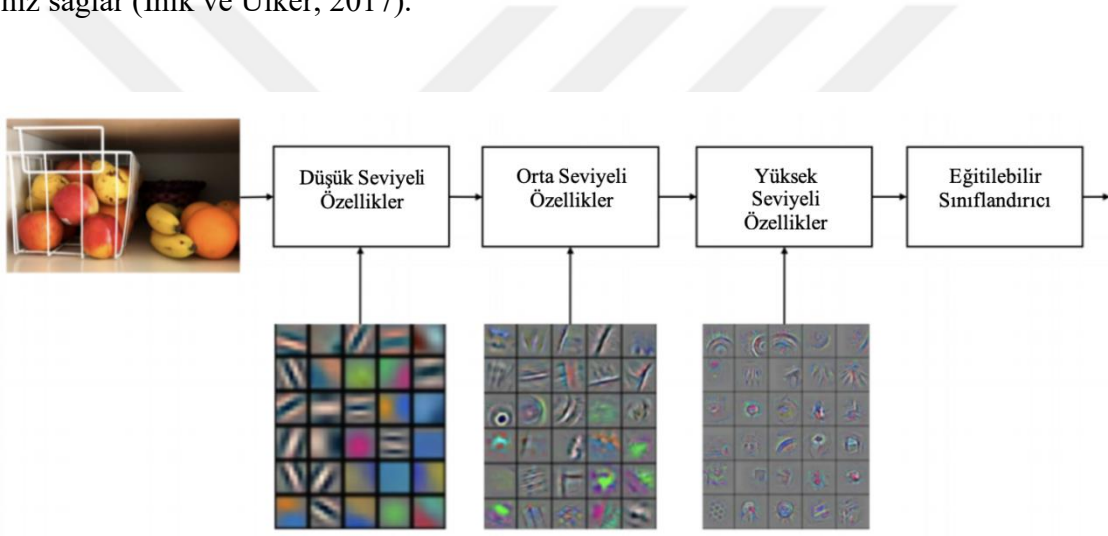
Derin öğrenme, Makine Öğreniminin en heyecan verici ve güçlü dalıdır. Derin öğrenmenin “derin” kısmı teknik bir terimdir ve “sinir ağlarının” “ağ” kısmındaki katman veya segment sayısını ifade eder (Swetha, 2020).



Sekil 2.21: Basit sinir ağı ve derin sinir ağı (Swetha, 2020)

Şekil 2.21 basit bir sinir ağı ile derin sinir ağı arasındaki farkı göstermektedir. Basit sinir ağında hidden layer (gizli katman) bir iken derin öğrenme ağında gizli katmanların sayısı artmaktadır.

Klasik makine öğrenme teknikleriyle bir model tanımlamak veya bir makine öğrenimi sistemi kurmak için ilk adım, özellik vektörünün çıkarılmasıdır. Ancak, bu işlem hem uzmanlık gerektirir hem de oldukça zaman alıcıdır. Bu nedenle, bu teknikler ham veriyi ön işlem yapmadan veya uzman yardımı olmadan işleyemezler. Derin öğrenme, geleneksel makine öğrenimi ve görüntü işleme tekniklerinin aksine, öğrenme işlemini ham veri üzerinde gerçekleştirdiğinden uzun süren veri işleme sürecini ortadan kaldırarak işlemlerde hız sağlar (İnik ve Ülker, 2017).



Şekil 2.22: Basit evrişimli sinir ağı (CNN) (Bhardwaj vd., 2018)

Derin öğrenme, Şekil 2.22’de de gösterildiği gibi görüntü işleme, makine öğrenmesi, ses tanıma, nesne takibi gibi birçok alanda performansı yüksek olan evrişimli sinir ağlarının bir yöntemidir (Turkoglu, Hanbay ve Sengur, 2022).

Derin öğrenmenin popülerliğinde önemli etkenlerden biri, bol miktarda veriye sahip olunmasıdır. Bununla birlikte, bu algoritmaların başarılı olabilmesi için gerekli kaynakların artık daha erişilebilir hale gelmesi de dikkate değer bir gelişmedir. Dijitalleşmenin ivme kazanmasıyla birlikte, bilgisayarlar aracılığıyla gerçekleştirilen faaliyetler daha kapsamlı bir şekilde kaydedilmekte ve bu da makine öğrenimi uygulamaları için zengin bir veri havuzu oluşturmaktadır. Derin öğrenmenin popülerliğinde bir diğer etken ise, günümüzde daha

büyük ve karmaşık modelleri çalıştırmak için yeterli hesaplama kaynaklarının bulunmasıdır. Yapay sinir ağlarında kullanılan bellek ve işlemci kapasitelerinin artmasıyla, daha derin ve karmaşık ağ yapılarına olan ilgi artmıştır. Ancak, bu büyük veri setlerinin ve karmaşık modellerin eğitimi için gereken bellek ve hesaplama kaynaklarının sağlanması önemli bir zorluktur. Örneğin, bir derin öğrenme modelinin eğitimi için gereken bellek miktarı, kullanılan görüntü sayısı ve boyutlarına bağlı olarak oldukça yüksek olabilir. Bu nedenle, büyük veri setlerinin kullanılması durumunda bellek kaynaklarının sağlanması büyük önem taşır (İnik ve Ülker, 2017).



3. MATERYAL VE YÖNTEM

Bu bölümde çalışmanın modeli, katılımcılar, veri toplama süreci, çalışmanın bağımlı ve bağımsız değişkenleri ile analiz süreci ile ilgili açıklamalara yer verilmiştir.

3.1. Araştırmanın Modeli

Bu çalışma, PISA 2022 okuma ve matematik becerisi puanlarının belirli bağımsız değişkenlerle ilişkisini incelemeye yönelik olarak yapılmıştır. Bu bağlamda, çalışmanın modeli ilişkisel tarama modelidir. Bu model birden çok bağımsız değişken ve bağımlı değişkenler arasındaki neden-sonuç ilişkisini belirleyerek analiz ettiği için araştırmanın modeli ilişkisel tarama modelidir (Karasar, 2006).

3.2. Katılımcılar

PISA 2022 araştırmasına 81 ülke katılmıştır. Çalışmanın bağımsız değişkenleri iki ana kategoriye ayrılmıştır (EK 1, EK 2). SPSS formatında indirilen öğrenci anket formları ve ebeveynlere uygulanan ebeveyn anketi içinden sırasıyla EK 1'de gösterilen 18 ekonomik değişken ve EK 2'de gösterilen 33 adet aile eğitim durumu ve ebeveyn tutumları etiketlerini içeren bağımsız değişken seçilmiştir. SPSS formatında indirilen tüm veri dosyası içinden bağımsız değişkenler ayrılmış, (99, 98, 97, 96, 95 ve #BOŞ!) olarak gösterilen çok sayıda ayrı veri olduğu gözlenmiştir. Tüm ayrı veriler temizlendiğinde Tablo 3.2'de gösterilen 42 ülke verisi çalışmaya katılmıştır.

3.3. Veri Toplama Araçları

Çalışmanın verileri OECD web sitesinden "<https://www.oecd.org/pisa/data/2022database>" öğrenci anketi veri dosyası SPSS formatında indirilmiştir. Çalışma için kullanılacak olan değişkenler için OECD'nin web sayfasında bulunan öğrenci anketi dosyası <https://www.oecd.org/pisa/data/2022database/pisa-2022-national-questionnaires.htm> ve https://www.oecd.org/pisa/data/2022database/CY8_202111_QST_MS_PAQ_NoNotes.pdf adresindeki ebeveyn anket dosyası indirilmiş ve incelenmiştir. Çalışmanın değişkenlerinin

ailenin ekonomik durumunu gösteren deęiřkenler (EK 1) ve anne baba eęitim durumu ve ebeveyn tutumları (EK 2) ana bařlıkları altında 2 grupta toplanmıřtır.

3.4. Veri Toplama Sreci

Kresel bir izleme arařtırması olan PISA, OECD tarafından  yılda bir dzenlenir ve genellikle 15 yařındaki ęrencilerin temel bilgi ve becerilerini lmektedir. PISA Trkiye 2022 n raporu incelenmiř ve ncelikle ařaęıdaki tablolar elde edilmiřtir.



Ölke Sıralaması	OECD Ölkeleri Sıralaması	Ölke	Ortalama Puan	%95 Güven Aralığında Ortalama Puan	Ölke Sıralaması	OECD Ölkeleri Sıralaması	Ölke	Ortalama Puan	%95 Güven Aralığında Ortalama Puan
1		Singapur	575	572 - 577	42		Sırbistan	440	434 - 446
2		Makao (Çin)	552	550 - 554	43		Birleşik Arap Emirlikleri	431	429 - 433
3		Tayvan (Çin)	547	540 - 554	44	33	Yunanistan	430	426 - 435
4		Hong Kong (Çin)	540	534 - 546	45		Romanya	428	420 - 436
5	1	Japonya	536	530 - 541	46		Kazakistan	425	422 - 429
6	2	Güney Kore	527	520 - 535	47		Moğolistan	425	420 - 430
7	3	Estonya	510	506 - 514	48		Güney Kıbrıs	418	416 - 421
8	4	İsviçre	508	504 - 512	49		Bulgaristan	417	411 - 424
9	5	Kanada	497	494 - 500	50		Moldova	414	410 - 419
10	6	Hollanda	493	485 - 500	51		Katar	414	412 - 416
11	7	İrlanda	492	488 - 496	52	34	Şili	412	408 - 416
12	8	Belçika	489	485 - 494	53		Uruguay	409	405 - 413
13	9	Danimarka	489	485 - 493	54		Malezya	409	404 - 413
14	10	Birleşik Krallık	489	485 - 493	55		Karadağ	406	403 - 408
15	11	Polonya	489	485 - 493	56		Azerbaycan (Bakü)	397	392 - 402
16	12	Avusturya	487	483 - 492	57	35	Meksika	395	391 - 399
17	13	Avustralya	487	484 - 491	58		Tayland	394	389 - 399
18	14	Çek Cumhuriyeti	487	486 - 494	59		Peru	391	387 - 396
19	15	Slovenya	485	482 - 487	60		Gürcistan	390	385 - 395
20	16	Finlandiya	484	480 - 488	61		Suudi Arabistan	389	385 - 392
21	17	Letonya	483	479 - 487	62		Kuzey Makedonya	389	387 - 390
22	18	İsveç	482	478 - 486	63	36	Kosta Rika	385	381 - 388
23	19	Yeni Zelanda	479	475 - 483	64	37	Kolombiya	383	377 - 389
24	20	Litvanya	475	472 - 479	65		Brezilya	379	376 - 382
25	21	Almanya	475	469 - 481	66		Arjantin	378	373 - 382
26	22	Fransa	474	469 - 479	67		Jamaika	377	371 - 384
27	23	İspanya	473	470 - 476	68		Arnavutluk	368	364 - 372
28	24	Macaristan	473	468 - 478	69		Filistin	366	362 - 369
29	25	Portekiz	472	467 - 477	70		Endonezya	366	361 - 370
30	26	İtalya	471	465 - 477	71		Fas	365	358 - 371
31		Vietnam	469	462 - 477	72		Özbekistan	364	360 - 368
32	27	Norveç	468	464 - 472	73		Ürdün	361	357 - 365
33		Malta	466	463 - 469	74		Panama	357	351 - 362
34	28	Amerika Birleşik Devletleri	465	457 - 473	75		Kosova	355	353 - 357
35	29	Slovakya	464	458 - 470	76		Filipinler	355	350 - 360
36		Hırvatistan	463	458 - 468	77		Guatemala	344	340 - 349
37	30	İzlanda	459	456 - 462	78		El Salvador	343	340 - 347
38	31	İsrail	458	451 - 464	79		Dominik Cumhuriyeti	339	336 - 342
39	32	Türkiye	453	450 - 456	80		Paraguay	338	333 - 342
40		Brunei	442	440 - 444	81		Kamboçya	336	331 - 342
41		Ukrayna	441	433 - 449					

Ölkeler, matematik alanındaki ortalama puanlarına göre azalan sırada listelenmiştir.
OECD ölkeleri mavî renk ile vurgulanmıştır.

Şekil 3.1: PISA 2022’ ye katılan ölkelerin matematik becerisi puan sıralamaları (MEB, 2024)

Şekil 3.1’e göre PISA 2022 sınavına 81 ölkeler katılım sağlamıştır. Tablolar Milli Eğitim Bakanlığı Türkiye ön raporundan elde edilmiştir. Türkiye matematik becerisi puanları kategorisinde tüm ölkeler arasında 39., OECD ölkeleri bazında 32. sırada yer almıştır. Şekil 3.2 incelendiğinde okuma becerileri puanları kategorisinde ise tüm ölkeler arasında 36. OECD ölkeleri arasında ise 30. sırada yer almıştır.

Ülke Sıralaması	OECD Ülkeleri Sıralaması	Ülke	Ortalama Puan	%95 Güven Aralığında Ortalama Puan
1		Singapur	543	539 - 546
2	1	İrlanda	516	511 - 521
3	2	Japonya	516	510 - 522
4	3	Güney Kore	515	508 - 523
5		Tayvan (Çin)	515	509 - 522
6	4	Estonya	511	506 - 516
7		Makao (Çin)	510	508 - 513
8	5	Kanada	507	503 - 511
9	6	Amerika Birleşik Devletleri	504	495 - 512
10	7	Yeni Zelanda	501	497 - 505
11		Hong Kong (Çin)	500	494 - 505
12	8	Avustralya	498	494 - 502
13	9	Birleşik Krallık	494	490 - 499
14	10	Finlandiya	490	486 - 495
15	11	Danimarka	489	484 - 494
16	12	Polonya	489	483 - 494
17	13	Çek Cumhuriyeti	489	484 - 493
18	14	İsveç	487	482 - 492
19	15	İsviçre	483	479 - 488
20	16	İtalya	482	476 - 487
21	17	Avusturya	480	475 - 486
22	18	Almanya	480	473 - 487
23	19	Belçika	479	474 - 484
24	20	Portekiz	477	471 - 482
25	21	Norveç	477	472 - 482
26		Hırvatistan	475	471 - 480
27	22	Letonya	475	470 - 479
28	23	İspanya	474	471 - 478
29	24	Fransa	474	468 - 480
30	25	İsrail	474	467 - 481
31	26	Macaristan	473	467 - 479
32	27	Litvanya	472	468 - 476
33	28	Slovenya	469	465 - 472
34		Vietnam	462	454 - 470
35	29	Hollanda	459	451 - 468
36	30	Türkiye	456	452 - 460
37	31	Şili	448	443 - 453
38	32	Slovakya	447	441 - 453
39		Malta	445	442 - 449
40		Sırbistan	440	435 - 446
41	33	Yunanistan	438	433 - 444

Ülke Sıralaması	OECD Ülkeleri Sıralaması	Ülke	Ortalama Puan	%95 Güven Aralığında Ortalama Puan
42	34	İzlanda	436	432 - 440
43		Uruguay	430	426 - 435
44		Brunei	429	427 - 432
45		Romanya	428	421 - 436
46		Ukrayna	428	420 - 435
47		Katar	419	416 - 422
48		Birleşik Arap Emirlikleri	417	415 - 420
49	35	Meksika	415	410 - 421
50	36	Kosta Rika	415	410 - 420
51		Moldova	411	406 - 416
52		Brezilya	410	406 - 414
53		Jamaika	410	401 - 418
54	37	Kolombiya	409	401 - 416
55		Peru	408	403 - 414
56		Karadağ	405	402 - 408
57		Bulgaristan	404	398 - 411
58		Arjantin	401	396 - 406
59		Panama	392	385 - 399
60		Malezya	388	383 - 393
61		Kazakistan	386	383 - 390
62		Suudi Arabistan	383	379 - 386
63		Güney Kıbrıs	381	379 - 383
64		Tayland	379	373 - 384
65		Moğolistan	378	374 - 383
66		Guatemala	374	369 - 379
67		Gürcistan	374	369 - 378
68		Paraguay	373	368 - 378
69		Azerbaycan (Bakü)	365	360 - 370
70		El Salvador	365	359 - 370
71		Endonezya	359	353 - 364
72		Kuzey Makedonya	359	357 - 360
73		Arnavutluk	358	355 - 362
74		Dominik Cumhuriyeti	351	347 - 356
75		Filistin	349	345 - 353
76		Filipinler	347	340 - 353
77		Kosova	342	340 - 344
78		Ürdün	342	337 - 347
79		Fas	339	332 - 347
80		Özbekistan	336	332 - 339
81		Kamboçya	329	325 - 333

Ülkeler, okuma becerileri alanındaki ortalama puanlarına göre azalan sırada listelenmiştir.
OECD Ülkeleri mavi renk ile vurgulanmıştır.

Şekil 3.2: PISA 2022’ ye katılan ülkelerin okuma becerisi puan sıralamaları (MEB, 2024)

3.4.1 Çalışmanın Bağımsız Değişkenleri

Çalışmanın bağımsız değişkenleri OECD web sitesinden indirilen öğrencilerin anketlere verdikleri cevaplar içinden (EK 1: 18 adet ekonomik etiketleri içeren bağımsız değişken) ve (EK 2: 33 adet aile eğitim durumu ve ebeveyn tutumları değişkenleri) seçilmiştir.

EK 2 incelendiğinde OECD web sitesinden indirilen PISA 2022'ye katılan öğrencilerin anketlere verdikleri cevaplar içinden 33 adet aile eğitim durumu ve ebeveyn tutumları etiketleri içeren bağımsız değişken seçilmiştir. Bu değişkenlerden, ST256Q01JA01, ST256Q02JA01, ST256Q03JA01, ST256Q06JA01, ST256Q07JA01, ST256Q08JA01, ST256Q09JA01, ST256Q01JA10 değişkenleri evdeki kitap sayısı ile ilgili oldukları için matematik becerisi puanını tahmininden çıkarılmıştır. Ayrıca PA003Q18WA, PA197Q01WA değişkenleri de matematik becerisi ile ilgili oldukları için okuma becerisi tahmininden çıkarılmıştır. Böylece matematik becerisi tahmininde kullanılacak 27, okuma becerisi tahmininde kullanılacak 33 adet değişken elde edilmiştir.

3.4.2 Ekonomik Durumu Gösteren Değişkenlerdeki Ayırık Verilerin Temizlenmesi

OECD web sitesinden SPSS formatındaki veriler ekonomik değişkenler ayrıldıktan sonra xlsx dosya biçiminde olan EXCEL dosyası şeklinde kaydedilmiştir. Ayırık veriler temizlenmeden önce 613745 satır veri olduğu görülmüştür. Veriler incelendiğinde 99, 98, 97, 95 kodları ile ifade edilen ayırık verilerin çok fazla olduğu görülmüştür. Ayırık veriler Bul-Değiştir yöntemi ile bulunarak boşluk ile değiştirilmiş, sonrasında ise boşluk içeren tüm satırlar silinmiştir. Tüm veriler temizlendiğinde Tablo 3.2'deki ülkelere ait veriler kalmıştır.

Tablo 3.1: Ayırık veriler için değer etiket tablosu

Değer	Açıklama
95	Valid Skip
97	Not Applicable
98	Invalid
99	No Response
9999999	Not Applicable
9999998	Invalid
9999997	Missing
#BOŞ!	Boş

99, 98, 97, 95 ve #BOŞ! kodlarıyla belirtilen ayırık verilerle ilgili bilgi Tablo 3.1'de gösterilmiştir.

Tablo 3.2: Ekonomik deęiřkenleri ieren verileri bulunduran lkeler ve PISA sıralamaları

lke Kodu	Aıklama	PISA 2022 lke Sıralaması (Matematik Bařarısı)
76	Brezilya	65
96	Brunei Sultanlıęı	40
100	Bulgaristan	49
116	Kamboya	81
152	řili	52
170	Kolombiya	64
214	Dominik Cumhuriyeti	79
222	El Salvador	78
233	Estonya	7
246	Finlandiya	16
250	Fransa	26
275	Filistin	69
320	Guatemala	77
344	Hong Kong(in)	4
348	Macaristan	28
360	Endonezya	70
388	Jamaika	67
398	Kazakistan	46
400	rdn	73
410	Gney Kore	6
428	Letonya	21
446	Makao(in)	2
458	Malezya	54
484	Meksika	57
496	Moęolistan	47
504	Fas	71
528	Hollanda	10
554	Yeni Zellanda	23
591	Panama	74

604	Peru	59
608	Filipinler	76
634	Katar	51
642	Romanya	45
688	Sırbistan	42
704	Vietnam	31
705	Slovenya	19
764	Tayland	58
784	Birleşik Arap Emirlikleri	43
792	Türkiye	39
901	Ukrayna	41
826	Birleşik Krallık	14
860	Özbekistan	72

Ayrık veriler temizlendiğinde Tablo 3.2’de gösterilen ülkelere ait veriler bulunmaktadır. PISA 2022 çalışmasına 81 ülke katılmıştır fakat anket verilerindeki ayrık veriler silindiğinde Tablo 3.2’de gösterilen 42 ülkeye ait veri kalmıştır. Bu bağlamda Tablo 3.1’de kodlarla gösterilen ayrık veriler temizlendiğinde 65032 satır veri elde edilmiştir.

EK 5, ST250D07JA değişkenine ait verileri göstermektedir. Veriler incelendiğinde her ülke için lüks kabul edilecek yeterliliklerin farklı farklı olduğu görülmüştür. Bu nedenle de her ülke için bu yeterliliklerle ilgili kodlamalar farklılıklar göstermektedir.

	D	E	F	G	H
1	T250Q02J	ST250Q03J	ST250Q04JA	ST250D07JA	ST250D07JA_kategori
2	2	2		1 0760002	2
3	1	1		1 0760001	1
4	2	2		1 0760002	2
5	1	1		1 0760001	1
6	1	2		1 0760001	1
7	1	1		1 0760001	1
8	2	2		1 0760002	2
9	1	2		1 0760001	1
10	1	1		1 0760001	1
11	2	1		1 0760002	2
12	1	1		1 0760001	1
13	2	1		1 0760002	2
14	2	2		1 0760002	2
15	1	1		1 0760001	1
16	2	1		1 0760002	2
17	1	2		1 0760002	2
18	1	1		1 0760001	1
19	2	2		1 0760001	1
20	1	1		1 0760001	1
21	1	2		1 0760001	1
22	2	2		1 0760002	2
23	1	1		1 0760002	2
24	1	1		1 0760001	1
25	2	2		1 0760002	2
26	2	2		1 0760002	2
27	1	1		1 0760001	1
28	2	1		1 0760002	2
29	2	2		1 0760002	2
30	1	1		1 0760001	1
31	1	1		1 0760001	1
32	2	2		2 0760002	2
33	1	1		1 0760002	2
34	1	2		1 0760001	1
35	1	1		1 0760001	1
36	1	1		1 0760001	1
37	1	1		1 0760001	1

Şekil 3.3: ST250D07JA değişkeninin kodlanmış hali

EK 5’te de görüldüğü üzere ülkelere göre çeşitli kodlarla gösterilen ve zenginlik seviyesini gösteren soruların ve kodlamaların birbirinden farklı olduğu görülmüştür. Bu nedenle tüm ülkeler için ST250D07JA değişkeni verilerini aynı kodlamaya dönüştürmek için her bir hücrede bulunan veriye göre =SAĞDAN(G2;1) formülü kullanılarak verilerde düzenlemeye gidilmiş ve şekil 3.3’te gösterilen tablo elde edilmiştir.

PISA 2022 sınavında her öğrenciye ait 10 adet makul okuma başarıları puanı ve 10 adet makul matematik başarıları puanı yayınlanmıştır. Bu puanların ortalamasının hesaplanması ile PVMATH_ORTALAMA ve PVREAD_ORTALAMA değişkenleri elde edilmiştir.

PVMATH_ORTALAMA ve PVREAD_ORTALAMA değişkenleri, MEB’in PISA 2022 Türkiye raporunda yer alan (EK 3, EK 4) ‘deki tablolardaki puan aralıkları göz önünde bulundurularak yeniden kodlanmış ve PVMATH_KODLAMA ve PVREAD_KODLAMA değişkenleri oluşturulmuştur.

Tablo 3.3: Matematik becerisi puanı kodlaması

Puan Aralığı	Değer
0-233	0
234-358	1
359-420	2
421-482	3
483-545	4
546-607	5
608 ve sonrası	6

10 adet makul matematik becerisi puanının ortalaması alınarak oluşturulan PVMATH_ORTALAMA değişkeni Tablo 3.3'te yer alan aralıklara göre tekrar kodlanmış ve çalışmanın bağımlı değişkenini oluşturmuştur.

Tablo 3.4: Okuma becerisi puanı kodlaması

Puan Aralığı	Değer
0-189	0
190-335	1
336-407	2
408-480	3
481-553	4
554-626	5
627 ve sonrası	6

10 adet makul okuma becerisi puanının ortalaması alınarak oluşturulan PVREAD_ORTALAMA değişkeni Tablo 3.4'te yer alan aralıklara göre tekrar kodlanmış ve çalışmanın diğer bağımlı değişkenini oluşturmuştur.

Tablo 3.5: PVMATH_1_2 ve PVREAD_1_2 değişkenlerinin oluşturulması

Değişken Adı	Değişken Açıklaması	Puan Aralığı	Değer	Durum
PVMATH_1_2	Matematik Becerisi Puanı	0-471	0	Başarısız
PVMATH_1_2	Matematik Becerisi Puanı	472-Sonrası	1	Başarılı
PVREAD_1_2	Okuma Becerisi Puanı	0-475	0	Başarısız
PVREAD_1_2	Okuma Becerisi Puanı	476-Sonrası	1	Başarılı

Tablo 3.5'e göre seçilen bağımsız değişkenlerin PISA 2022'ye katılan öğrenci puanlarını başarılı ve başarısız olarak sınıflandırması için PISA Türkiye 2024 raporunda da belirtildiği üzere OECD ortalamasının üzerinde puan alanlar başarılı (1), altında puan alanlar ise

başarısız (0) olarak kodlanmıştır. Buna göre PISA 2022’de 37 OECD ülkesinin ortalama okuma becerisi puanı 476, ortalama matematik becerisi puanı ise 472’dir. Bu bağlamda PVREAD_1_2 ve PVMATH_1_2 bağımlı değişkenleri de çalışmaya eklenmiştir.

3.4.3 Anne ve Baba Eğitim Durumları ve Ebeveyn Tutumları Gösteren Değişkenlerdeki Ayrık Verilerin Silinmesi

Anne ve baba eğitim durumlarını ve ebeveyn tutumları ile ilgili seçilen değişkenlerden PA003Q02TA, PA003Q03TA, PA003Q05IA, PA003Q18WA, PA003Q13JA, PA003Q15JA, PA003Q16JA, PA003Q17JA verilerinin analize uygun olmadığı görülmüştür. SPSS ortamındaki veriler EXCEL programına aktarılmış ve 613745 satır veri olduğu görülmüştür. Anne ve baba eğitim durumlarını gösteren veriler ve ISCED seviyeleri incelendiğinde ST006Q05JA(Does your mother have any of the following qualifications: [ISCED level 4]) ve ST008Q05JA(Does your father have any of the following qualifications: [ISCED level 4]) ülkeler bazında 97 koduyla gösterildiği görülmüştür. Bu seviye (ISCED 4) Türkiye’de de olmadığı için veriler 97 (Not Applicable) olarak gösterilmiştir. Bu nedenle ST006Q05JA ve ST008Q05JA değişkenleri ve verileri çalışmadan çıkarılmıştır.

Tablo 3.6: ISCED 2011 seviyeleri

Seviye 0	Erken Çocukluk Dönemi Eğitimi
Seviye 1	Temel Eğitim
Seviye 2	Alt Düzey Ortaöğretim Eğitimi
Seviye 3	Üst Düzey Ortaöğretim Eğitimi
Seviye 4	Ortaöğretim Sonrası Yükseköğretim Derecesinde Olmayan Eğitim
Seviye 5	Kısa Dönem Yükseköğretim Eğitimi
Seviye 6	Lisans veya Eşdeğeri
Seviye 7	Yüksek Lisans veya Eşdeğeri Seviye
Seviye 8	Doktora veya Eşdeğeri Seviye

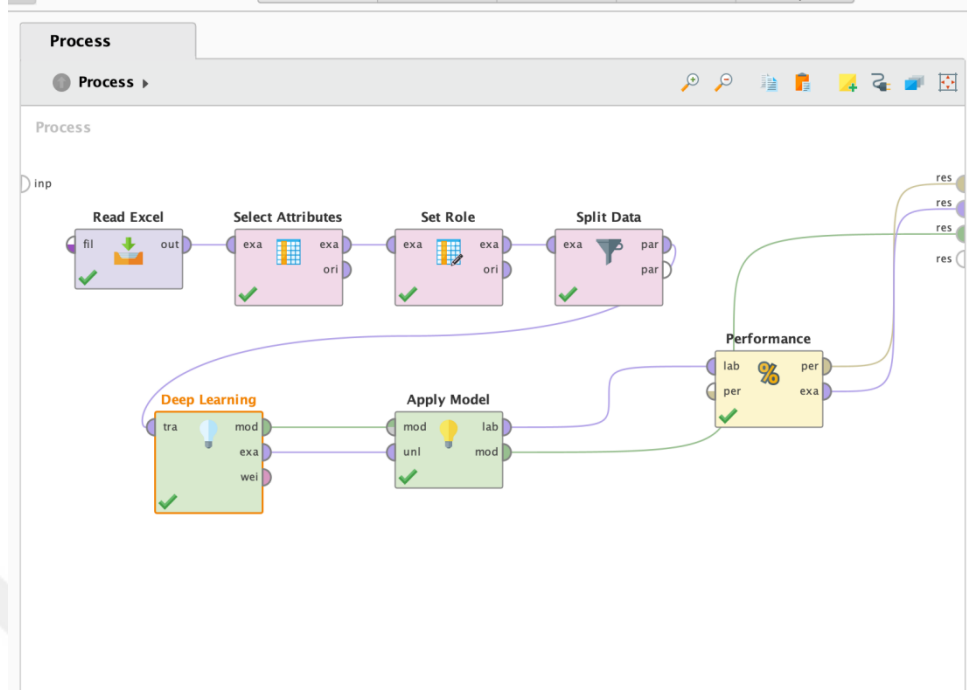
Tablo 3.6’da gösterilen ISCED seviyeleri UNESCO tarafından oluşturulan Uluslararası Standart Eğitim Sınıflaması (ISCED), eğitim alanında ülkeler arasında karşılaştırılabilir istatistiklerin toplanması, derlenmesi ve analiz edilmesi için bir çerçeve sunmaktadır.

ISCED, eğitim programlarının sahip olması gereken özellikleri alanlarına ve seviyelerine göre sınıflandıran bir sistemdir. İlk olarak 1976'da yayınlanan ISCED, daha sonraki yıllarda revize edilmiş ve günümüze kadar birçok ülke tarafından eğitim politikalarının belirlenmesi ve istatistik verilerinin toplanması gibi amaçlarla kullanılmıştır. Bu revizyonlar, uluslararası anlaşmalar çerçevesinde UNESCO'ya üye ülkeler tarafından resmen kabul edilmiştir (YÖK, 2024).

Anne baba eğitim durumları ve ebeveyn tutumlarını gösteren değişkenler ve verilerinden oluşan EXCEL dosyası incelenmiş ve 97, 98, 99 ve #BOŞ! kodlarıyla tanımlanan ayırık verilerin yine çok fazla olduğu görülmüştür. Bu verileri içeren tüm satırları sildiğimizde PISA 2022'ye katılan tüm ülkelerden 48542 satır veri elde edilmiştir.

3.5. Ekonomik Değişkenlerin Rapid Miner Programına Aktarılması ve Modelin Oluşturulması

Rapid Miner, daha önceki adı YALE (Yet Another Learning Environment) olan bir grup yazılımcı tarafından 2001 yılında geliştirilen bir yazılımdır. Veri ön işleme, istatistiksel modelleme, tahmin analizi ve optimizasyon için kullanılmaktadır (Yadav, Malik ve Chandel, 2015). Rapid Miner, Gartner Inc.'in 2020 yılı değerlendirmesinde veri hazırlama, makine öğrenimi ve model operasyonlarını birleştiren kendisini "Veri Bilimi ve Makine Öğrenimi Platformları için Magic Quadrant" içinde konumlandıran, bir veri bilimi platformudur (Apaza-Pinto vd., 2022).



Şekil 3.4: Ekonomik değişkenler ile PVMATH_KODLAMA, PVREAD_KODLAMA, PVMATH_1_2, PVREAD_1_2 puanlarının tahmin edilmesi Rapid Miner modeli

Rapid Miner ve derin öğrenme yöntemi kullanılarak oluşturulan model şekil 3.4'te gösterilmektedir. Rapid Miner programında oluşturulan modelin EK 1'deki ailenin ekonomik durumunu gösteren değişkenlerin PVMATH_KODLAMA, PVREAD_KODLAMA, PVMATH_1_2, PVREAD_1_2 bağımlı değişkenlerini tahmin etmesi istenmiştir. Öğrenme modeli olarak derin öğrenme kullanılmıştır. Modelde Set Role nesnesinde ekonomik değişkenler regular ve modelin tahmin etmesi istenen PVMATH_KODLAMA, PVREAD_KODLAMA, PVMATH_1_2, PVREAD_1_2 değişkenleri label olarak seçilmiştir. Verileri eğitim ve test verisi olarak seçmek için Split Data nesnesi kullanılmış ve örnekleme yöntemi olarak shuffled sampling yöntemi seçilmiştir. Split Data (bölünmüş veri operatörü), bir örnek kümesini (veri setini) alır ve bu veri setini birden fazla alt kümeye böler. Bu alt kümeler, operatörün çıkış bağlantı noktaları üzerinden iletilir. Bu işlemin amacı, bir veri setini belirli kurallara göre parçalara ayırmak ve her parçayı ayrı ayrı işlemektir. Bölünmüş veri operatörü, bir veri setini birden fazla alt kümeye ayırarak diğer örnekleme ve filtreleme operatörlerinden ayrılır. Diğer operatörler genellikle veri setinin belirli bir bölümünü seçer veya belirli kriterlere göre filtrelerken, bölünmüş veri operatörü, veri setini birden fazla bölümde sunar. Bu sayede, aynı veri seti

üzerinde farklı işlemler yapılabilir ve sonuçlar karşılaştırılabilir (Rapidminer, 2024). Derin öğrenmede kullanılan diğer bazı önemli parametreler Tablo 3.7’de gösterilmektedir.

Tablo 3.7: Derin öğrenme hiper parametreleri

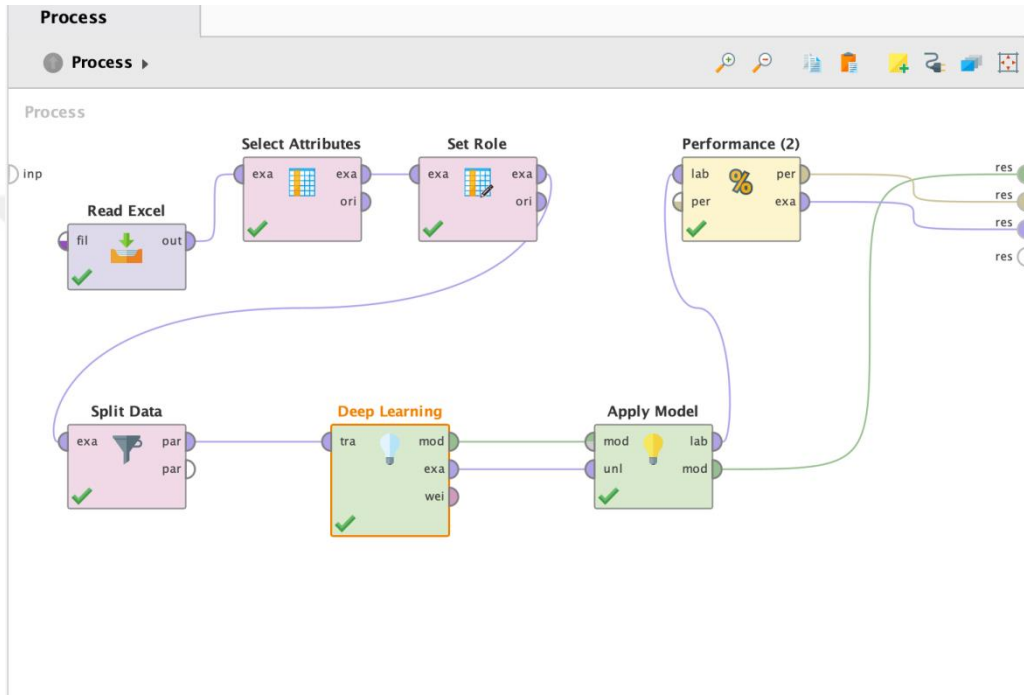
Activation	Rectifier (ReLU)
Epochs	5000
Train samples per iteration	-1
Learning rate	0.01
Learning rate annealing	1.0E-4
Learning rate decay	0.5
Momentum start	0.9
Momentum stable	0.99

Tablo 3.7’de gösterilen hiper parametreleri gösterilen model 5000 tur sayısında eğitilmiştir. Modelin aktivasyon fonksiyonu olarak ReLU kullanılmıştır. ReLU (Rectified Linear Unit) fonksiyonu, iyi bir seyrekliğe, hızlı yakınsama hızına ve basit hesaplama sahiptir. Bununla birlikte, ReLU'nun negatif değerlerdeki gradyanı her zaman sıfır olduğundan, bu durum bazı nöronların eğitim sırasında "ölü" hale gelmesine neden olabilir (Dolezel, Skrabanek ve Gago, 2016). Bu, negatif girişler alan nöronların artık aktivasyon göstermemesi ve dolayısıyla öğrenmeye katkı sağlamaması anlamına gelir. ReLU aktivasyon fonksiyonunun kullanılması, sigmoid ve tanh fonksiyonlarında karşılaşılan "gradyanların genişlemesi ve kaybolması" sorunlarını etkili bir şekilde çözer (He vd., 2015).

ReLU aktivasyon fonksiyonu kullanıldığında, değerler ya 0 olur (bazı gradyanların etkisiz kalmasına neden olur) ya da 1 olur. ReLU sinir ağına dahil edildiğinde, aynı zamanda yüksek derecede seyrek bir yapı da oluşturmuş olunur (Wen vd., 2016). Bu, sinir ağında aktivasyon matrisinin birçok sıfır içerdiği anlamına gelmektedir. Aktivasyonların belirli bir yüzdesi (örneğin %50) sıfır olduğunda, bu duruma sinir ağı seyrek denir. Seyreklik, zaman ve bellek kullanımı açısından verimliliği artırabilir çünkü sıfır değerler genellikle daha az bellek kullanır ve daha düşük hesaplama maliyeti gerektirir (Kim, Khan ve Kyung, 2019).

3.6 Ebeveyn Eğitim Durumları ve Ebeveyn Tutumları ile İlgili Olan Değişkenlerin Rapid Miner Programına Aktarılması

81 ülkenin anket verileriyle oluşturulan Parent questionnaire anket sorularına verilen cevaplardan oluşturulan ebeveyn tutum değişkenleri Rapid Miner programına aktarılmış ve şekil 3.5’te gösterilen model oluşturulmuştur.



Şekil 3.5: Ebeveyn eğitim durum ve ebeveyn tutumları değişkenleri ile PVMATH_KODLAMA, PVREAD_KODLAMA, PVMATH_1_2, PVREAD_1_2 puanlarının tahmin edilmesi Rapid Miner modeli

Şekil 3.5’te gösterilen modelde verileri eğitim ve test olarak bölmek için Split Data nesnesi kullanılmış olup eğitim ve test verilerinin seçimi için shuffled sampling yöntemi kullanılmıştır. PVMATH_KODLAMA, PVREAD_KODLAMA, PVMATH_1_2, PVREAD_1_2 değişkenleri label, ebeveyn eğitim durumu ve tutumları ile ilgili seçilen değişkenler ise regular olarak tanımlanmıştır.

Tablo 3.8: Modelin hiper parametreleri

Activation	Rectifier (ReLU)
Epochs	5000
Train samples per iteration	-1
Learning rate	0.01
Learning rate annealing	1.0E-4
Learning rate decay	0.5
Momentum start	0.9
Momentum stable	0.99

Öğrenme yöntemi olarak derin öğrenme kullanılmış olup modelin bazı önemli parametreleri Tablo 3.8’de hiper parametreleri gösterilen model 5000 turda eğitilmiştir.

3.7 Python ve Çok Katmanlı Yapay Sinir Ağları Kullanılarak PVMATH_KODLAMA, PVREAD_KODLAMA, PVMATH_1_2, PVREAD_1_2 Değişkenlerinin Tahmin Edilmesi

Bu çalışmada çok katmanlı algılayıcı yapay sinir ağı modelinin geliştirilmesi için Python programlama platformu kullanılmıştır. Python, geniş standart kütüphaneye sahip, ücretsiz, üst düzey genel amaçlı açık kaynaklı bir dildir. Modeli çok sayıda yinelemeli ve çeşitli verilerle geliştirirken Python ortamının üretken, esnek, dinamik ve ücretsiz açık kaynak yapısı nedeniyle seçilmiştir (Radhakrishnan, Nair ve Isaac, 2022).

3.7.1 Kütüphanelerin İmport (Dahil) Edilmesi

Python’da bulunan pandas kütüphanesi verilerin seçilmesi, gruplandırılması, analizi, dönüştürülmesi ve birleştirilmesi için gerekli olan araçları bulunduran bir kütüphanedir. Pandas kütüphanesi Series ve DataFrame veri yapılarını sunar. Series, tek boyutlu dizi tipindeki verileri, DataFrame ise çok boyutlu dizi tipindeki verileri temsil eder.

```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
from sklearn.model_selection import train_test_split
from scipy import stats
```

```
from sklearn.metrics import explained_variance_score, mean_absolute_error,
mean_squared_error
from sklearn.metrics import median_absolute_error, r2_score
```

3.7.2 Veri Setinin Yüklenmesi ve DataFrame'e (Veri Çerçevesi) Dönüştürülmesi

DataFrame, büyük veri setleri üzerinde kolayca işlem yapmanızı sağlayan, sütun eklemek, çıkarmak, filtrelemek, gruplamak ve dönüştürmek gibi işlemler hızlı ve etkili bir şekilde gerçekleştirilebilen pandas kütüphanesinin bir nesnesidir. DataFrame, CSV, EXCEL, SQL, JSON gibi çeşitli dosya formatlarından veri okumayı ve bu formatlara veri yazmayı destekler, bu özelliği ile veriyi içe ve dışa aktarmayı çok basit hale getirmektedir. Etiketlere veya pozisyonlara dayalı olarak veri seçimi ve dilimleme yapılır ayrıca Pandas, NumPy gibi diğer Python kütüphaneleriyle sorunsuz bir şekilde entegre olmaktadır. Aşağıdaki kod bloğu ile verilerin bulunduğu EXCEL dosyası projeye eklenmektedir.

```
veriseti=pd.read_excel("/Users/seheryilmaz/desktop/ekonomik_pvmath_kodlama.xlsx")
print(veriseti.head(5))
```

	ST250Q01JA	ST250Q02JA	ST250Q03JA	ST250Q04JA	ST250D07JA_kategori	\
0	1	2	2	1	2	
1	1	1	1	1	1	
2	1	2	2	1	2	
3	1	1	1	1	1	
4	1	1	2	1	1	

	ST251Q03JA	ST251Q04JA	ST251Q07JA	ST254Q03JA	ST258Q01JA	ST259Q01JA	\
0	2	2	2	1	1	5	
1	3	3	2	2	1	7	
2	3	3	1	1	1	6	
3	3	4	4	2	1	6	
4	3	3	1	2	1	6	

	ST259Q02JA	ST294Q01JA	ST294Q03JA	ST294Q04JA	ST349Q01JA	ST352Q01JA	\
0	7	2	1	1	2	2	
1	9	6	3	4	2	2	
2	8	1	4	1	2	4	
3	4	6	1	1	3	2	
4	9	6	6	1	2	1	

	ST352Q05JA	PVMATH_KODLAMA
0	2	2
1	2	3
2	4	3
3	1	5
4	1	4

Şekil 3.6: Veri setinde bulunan ilk 5 satırın görüntülenmesi

print(veriseti.head(5)) kod bloğu ile Şekil 3.6'daki ekran çıktısı elde edilir. Bu kodlama ile veri setinde bulunan ilk 5 satırın görüntülenmesi sağlanır.

X=pd.DataFrame(veriseti.iloc[:, :-1])

Y=pd.DataFrame(veriseti.iloc[:, -1])

Yukarıdaki kod bloğu ile (Şekil 3.7-3.8) ekran çıktıları elde edilir.

In [5]: X

Out[5]:

	ST250Q01JA	ST250Q02JA	ST250Q03JA	ST250Q04JA	ST250D07JA_kategori	ST251Q03JA	ST251Q04JA	ST251Q07JA	ST254Q03JA	ST258Q01JA	ST258Q02JA
0	1	2	2	1	2	2	2	2	1	1	
1	1	1	1	1	1	3	3	2	2	1	
2	1	2	2	1	2	3	3	1	1	1	
3	1	1	1	1	1	3	4	4	2	1	
4	1	1	2	1	1	3	3	1	2	1	
...	...	...	...	...	...	...	...	...	...	...	
64997	1	1	2	1	1	1	4	1	3	5	
64998	2	2	1	1	1	2	2	1	1	1	
64999	1	2	2	2	1	4	4	4	1	1	
65000	1	2	2	1	2	2	2	1	1	2	
65001	1	2	2	1	1	2	3	4	1	1	

65002 rows x 18 columns

Şekil 3.7: X çok boyutlu dizisi

Out[26]:

	PVMATH_KODLAMA
0	1
1	1
2	2
3	4
4	2
...	...
64997	1
64998	1
64999	3
65000	1
65001	A

65002 rows x 1 columns

Out[5]:

	PVMATH_1_2
0	0
1	0
2	0
3	1
4	0
...	...
64997	0
64998	0
64999	0
65000	0
65001	B

65002 rows x 1 columns

Şekil 3.8: Y çok boyutlu dizileri (A ve B)

(Şekil 3.7-3.8) 'de DataFrame tipindeki X ve Y çok boyutlu dizilerinin içeriği görünmektedir. Buna göre modelin PVMATH_KODLAMA, PVREAD_KODLAMA, PVMATH_1_2, PVREAD_1_2 puanlarını tahmininde kullanacağı 18 adet bağımsız değişken X adında bir

DataFrame’de, tahmin edilmesi istenen PVMATH_KODLAMA, PVREAD_KODLAMA ve PVMATH_1_2, PVREAD_1_2 bağımlı değişkenleri ise Y adında bir DataFrame’e aktarılmıştır.

3.7.3 Verilerin Test ve Eğitim Olarak Bölünmesi

Oluşturulan modelde veriyi eğitim ve test verisi olarak bölmek için scikit-learn kütüphanesine ait train_test_split metodu kullanılmıştır. Test_size = 0.1 parametresi ile toplam verinin %90’lık kısmının eğitim geriye kalan %10’luk kısmının ise test verisi olarak ayrılmıştır. Modelin eğitimi için, ayrılan %90’lık veri bölümünü kullanır. Modelin eğitiminden sonra ürettiği tahminler ile gerçek değerlerin karşılaştırılmasında ise ayrılan %10’luk kısım olan test verileri kullanılır. Verilerin eğitim ve test olarak ayrılmasında aşağıdaki kod bloğu kullanılmıştır.

```
X_train,X_test,Y_train,Y_test=train_test_split(X,Y,test_size=0.1, random_state=1)
```

random_state=1 parametresi ise verilerimizin eğitim ve test olarak ayrılırken rastgele olarak ayrılacağı fakat, 1 değeri ile her bölümlenme (split) yapıldığında aynı rastgele sayıları üreteceği ve aynı random (rastgele) şekilde bölüneceğini belirtir.

3.7.4 Verilerin Kategorik Değişkenlere Dönüştürülmesi

Oluşturulan modelde tahmin edilmesi istenen puanlar PVMATH_KODLAMA, PVREAD_KODLAMA 0-6 arasında, PVMATH_1_2, PVREAD_1_2 değişkenleri ise 0-1 arasında değerler almaktadır. Bu bağlamda çoklu sınıflandırma yapıldığı için bu değerleri kategorik olarak etiketlemek gereklidir.

Bu işlemi Keras kütüphanesinde yapmak için aşağıdaki kodlama kullanılmış ve Şekil 3.9’daki tablo elde edilmiştir.

```
import tensorflow as tf
Y_train=tf.keras.utils.to_categorical(Y_train)
Y_test=tf.keras.utils.to_categorical(Y_test)
```

<pre>In [33]: Y_train[:10,:]</pre> <pre>Out[33]: array([[0., 0., 0., 0., 1., 0., 0.], [0., 0., 0., 0., 0., 1., 0.], [0., 0., 1., 0., 0., 0., 0.], [0., 0., 1., 0., 0., 0., 0.], [0., 0., 1., 0., 0., 0., 0.], [0., 0., 0., 1., 0., 0., 0.], [0., 0., 0., 1., 0., 0., 0.], [0., 0., 0., 1., 0., 0., 0.], [0., 0., 0., 1., 0., 0., 0.], [0., 0., 0., 0., 0., 0., 1.], [0., 0., 0., 0., 0., 1., 0.]])</pre>	<pre>In [12]: Y_train[:10,:]</pre> <pre>Out[12]: array([[1., 0.], [0., 1.], [1., 0.], [1., 0.], [1., 0.], [1., 0.], [1., 0.], [1., 0.], [1., 0.], [0., 1.], [0., 1.]])</pre>
--	--

Şekil 3.9: Kategorilendirilmiş tahmin edilecek puanlar

Şekil 3.9’da kategorilendirilmiş eğitim ve test verileri görülmektedir. Bu yöntemle çıkış katmanından (0,1,2,3,4,5,6) ve (0,1) yerine 0 veya 1 değeri elde edilmektedir. Bu sayede doğru rakama karşılık gelen etiketin indeksi 1 iken diğer tüm etiketler için bu indeks 0 değerini alır. Bu şekilde etiketlerimiz bir vektöre dönüştürülmüş olmaktadır.

3.7.5 Modelin Oluşturulması

Yapay sinir ağları oluşturmak için kullanılan bir Python kütüphanesi olan Keras, modellerin inşası, eğitimi ve değerlendirilmesi için gerekli olan tüm araçlara sahiptir. Keras, yüksek seviyeli bir arayüz sunar. Yapay sinir ağı modellerinin inşası, eğitimi ve değerlendirilmesinde kullanılan TensorFlow, Theano ve CNTK gibi çeşitli derin öğrenme çerçeveleri ile çalışabilir. Keras, hızlı prototipleme, kolay genişletilebilme ve kullanım kolaylığı gibi avantajlarıyla tanınır. Keras API’sinde bir model türü olan Sıralı (Sequential), ardışık katmanlardan oluşur ve her katman girişten çıkışa doğru sıralı olarak tanımlanır.

```

In [36]: #Ağın oluşturulması
import keras
from keras.models import Sequential
from keras.layers import Dense
siniflandirici=Sequential()

In [37]: #Model
siniflandirici.add(Dense(units=128, activation="relu", input_dim=18))
siniflandirici.add(Dense(units=256, activation="relu"))
siniflandirici.add(Dense(units=128, activation="relu"))
siniflandirici.add(Dense(units=64, activation="relu"))
siniflandirici.add(Dense(units=128, activation="relu"))
siniflandirici.add(Dense(units=64, activation="relu"))
siniflandirici.add(Dense(units=32, activation="relu"))
siniflandirici.add(Dense(units=16, activation="relu"))
siniflandirici.add(Dropout(0.5))

siniflandirici.add(Dense(units=7, activation="softmax"))

```

Şekil 3.10: Ekonomik değişkenlerinin matematik becerisi ve okuma becerisi puanını tahmin ederken oluşturulan Python ve YSA modeli (0-6 arası tahmin)

Şekil 3.10’da oluşturulan modelde sırasıyla 18 giriş değişkenini girdi olarak alan bir giriş katmanı, sırasıyla 128, 256, 128, 64, 128, 64, 32, 16 nöronu bulunan 8 gizli katman ve 7 çıkış veren bir çıktı katmanı bulunmaktadır. Girdi katmanı, ekonomik değişkenlerle PVMATH_KODLAMA ve PVREAD_KODLAMA değişkenlerini tahmin ederken 18 giriş değişkeni almaktadır. Modelde aktivasyon fonksiyonu olarak ReLU kullanılmıştır. Modelin çıktı katmanında Softmax fonksiyonu kullanılmıştır. Bu fonksiyon genellikle bir ağın çıktısını, tahmin edilen çıktı sınıfları üzerinden bir olasılık dağılımına göre normalleştirmek için bir sinir ağının son aktivasyon fonksiyonu olarak kullanılır (Wikipedia, 2024). Softmax, çok sayıda sınıf içeren durumlarda kullanılır ve her bir sınıf için olasılık değerleri hesaplayarak kapsamlı bir sınıflandırma sağlar. Bu yöntem, birden fazla kategorinin olduğu senaryolarda her bir kategoriye ait olasılıkların belirlenmesini mümkün kılar. Çalışmada 7 kategorili (0-6) bir tahmin işlemi yapılacağından çıktı katmanında Softmax kullanılmıştır.

```

In [17]: #Ağın oluşturulması
import keras
from keras.models import Sequential
from keras.layers import Dense
siniflandirici=Sequential()

In [18]: #Model
siniflandirici.add(Dense(units=128, activation="relu", input_dim=18))
siniflandirici.add(Dense(units=256, activation="relu"))
siniflandirici.add(Dense(units=128, activation="relu"))
siniflandirici.add(Dense(units=64, activation="relu"))
siniflandirici.add(Dense(units=128, activation="relu"))
siniflandirici.add(Dense(units=64, activation="relu"))
siniflandirici.add(Dense(units=32, activation="relu"))
siniflandirici.add(Dense(units=16, activation="relu"))
siniflandirici.add(Dropout(0.5))

siniflandirici.add(Dense(units=2, activation="sigmoid"))

```

Şekil 3.11: Ekonomik değişkenlerinin matematik becerisi ve okuma becerisi puanını tahmin ederken oluşturulan Python ve YSA modeli (0-1 arası tahmin)

Şekil 3.11’de oluşturulan modelde sırasıyla 18 giriş değişkenini girdi olarak alan bir giriş katmanı, sırasıyla 128, 256, 128, 64, 128, 64, 32, 16 nöronu bulunan 8 gizli katman ve 2 çıkış veren bir çıktı katmanı bulunmaktadır. Girdi katmanı, ekonomik değişkenlerle PVMATH_1_2, PVREAD_1_2 değişkenlerini tahmin ederken 18 giriş değişkeni almaktadır. Oluşturulan modelin çıktı katmanında Sigmoid fonksiyonu kullanılmıştır. Sigmoid fonksiyonu, bir sayı alır ve bu sayıyı 0 ile 1 arasında bir değere dönüştürür. Bu fonksiyonun grafiği "S" şeklindedir: Büyük negatif sayılar için sonuç 0'a, büyük pozitif sayılar için ise sonuç 1'e yaklaşır. Sigmoid fonksiyonunun çıktıları, olasılık olarak yorumlanabilir, bu da onu iki sınıf arasında ayırım yapmayı gerektiren problemler için ideal hale getirir (datacamp, 2024). Çalışmamızda iki kategorili bir tahmin işlemi yapılacağı için Sigmoid fonksiyonu kullanılmıştır.

```

In [40]: #Ağın oluşturulması
import keras
from keras.models import Sequential
from keras.layers import Dense
siniflandirici=Sequential()

In [41]: #Model
siniflandirici.add(Dense(units=128, activation="relu", input_dim=27))
siniflandirici.add(Dense(units=256, activation="relu"))
siniflandirici.add(Dense(units=128, activation="relu"))
siniflandirici.add(Dense(units=64, activation="relu"))
siniflandirici.add(Dense(units=128, activation="relu"))
siniflandirici.add(Dense(units=64, activation="relu"))
siniflandirici.add(Dense(units=32, activation="relu"))
siniflandirici.add(Dense(units=16, activation="relu"))
siniflandirici.add(Dropout(0.5))

siniflandirici.add(Dense(units=7, activation="softmax"))

```

Şekil 3.12: Aile eğitim durumu ve ebeveyn tutumları değişkenlerinin matematik becerisi puanını tahmin ederken oluşturulan Python ve YSA modeli (0-6 arası tahmin)

```

In [16]: #Ağın oluşturulması
import keras
from keras.models import Sequential
from keras.layers import Dense
siniflandirici=Sequential()

In [20]: #Model
siniflandirici.add(Dense(units=128, activation="relu", input_dim=33))
siniflandirici.add(Dense(units=256, activation="relu"))
siniflandirici.add(Dense(units=128, activation="relu"))
siniflandirici.add(Dense(units=64, activation="relu"))
siniflandirici.add(Dense(units=128, activation="relu"))
siniflandirici.add(Dense(units=64, activation="relu"))
siniflandirici.add(Dense(units=32, activation="relu"))
siniflandirici.add(Dense(units=16, activation="relu"))
siniflandirici.add(Dropout(0.5))

siniflandirici.add(Dense(units=7, activation="softmax"))

```

Şekil 3.13: Aile eğitim durumu ve ebeveyn tutumları değişkenlerinin okuma becerisi puanını tahmin ederken oluşturulan Python ve YSA modeli (0-6 arası tahmin)

Şekil (3.12-3.13)'de oluşturulan modellerde sırasıyla 27 ve 33 giriş değişkenini girdi olarak alan bir giriş katmanı, sırasıyla 128, 256, 128, 64,128, 64, 32, 16 nöronu bulunan 8 gizli katman ve 7 çıkış veren bir çıktı katmanı bulunmaktadır. Girdi katmanı, anne baba eğitim durumu ve ebeveyn tutumları değişkenleri ile PVMATH_KODLAMA bağımlı değişkeninin tahmin edilmesinde 27, PVREAD_KODLAMA bağımlı değişkeninin tahmin edilmesinde 33 giriş değişkeni almaktadır.

```

In [37]: #Ağın oluşturulması
import keras
from keras.models import Sequential
from keras.layers import Dense
siniflandirici=Sequential()

In [38]: #Model
siniflandirici.add(Dense(units=128, activation="relu", input_dim=27))
siniflandirici.add(Dense(units=256, activation="relu"))
siniflandirici.add(Dense(units=128, activation="relu"))
siniflandirici.add(Dense(units=64, activation="relu"))
siniflandirici.add(Dense(units=128, activation="relu"))
siniflandirici.add(Dense(units=64, activation="relu"))
siniflandirici.add(Dense(units=32, activation="relu"))
siniflandirici.add(Dense(units=16, activation="relu"))
siniflandirici.add(Dropout(0.5))

siniflandirici.add(Dense(units=2, activation="sigmoid"))

```

Şekil 3.14: Aile eğitim durumu ve ebeveyn tutumları değişkenlerinin matematik becerisi puanını tahmin ederken oluşturulan Python ve YSA modeli (0-1 arası tahmin)

```

In [98]: #Ağın oluşturulması
import keras
from keras.models import Sequential
from keras.layers import Dense
siniflandirici=Sequential()

In [99]: #Model
siniflandirici.add(Dense(units=128, activation="relu", input_dim=33))
siniflandirici.add(Dense(units=256, activation="relu"))
siniflandirici.add(Dense(units=128, activation="relu"))
siniflandirici.add(Dense(units=64, activation="relu"))
siniflandirici.add(Dense(units=128, activation="relu"))
siniflandirici.add(Dense(units=64, activation="relu"))
siniflandirici.add(Dense(units=32, activation="relu"))
siniflandirici.add(Dense(units=16, activation="relu"))
siniflandirici.add(Dropout(0.5))

siniflandirici.add(Dense(units=2, activation="sigmoid"))

```

Şekil 3.15: Aile eğitim durumu ve ebeveyn tutumları değişkenlerinin okuma becerisi puanını tahmin ederken oluşturulan Python ve YSA modeli (0-1 arası tahmin)

Şekil (3.14-3.15)'te oluşturulan modellerde sırasıyla 27 ve 33 giriş değişkenini girdi olarak alan bir giriş katmanı, sırasıyla 128, 256, 128, 64,128, 64, 32, 16 nöronu bulunan 8 gizli katman ve 2 çıkış veren bir çıktı katmanı bulunmaktadır. Girdi katmanı, anne baba eğitim durumu ve ebeveyn tutumları değişkenleri ile PVMATH_1_2 bağımlı değişkeninin tahmin edilmesinde 27, PVREAD_1_2 bağımlı değişkeninin tahmin edilmesinde 33 giriş değişkeni almaktadır.

3.7.6 Kodun Derlenmesi

Aşağıdaki kodlarda model için optimizasyon algoritması olarak Adam (Adaptive Moment Estimation) algoritması kullanılmıştır. Her eğitimden sonra elde edilen değerler ile gerçek değerler arasındaki hata farkının hesaplanması için "categorical_crossentropy" algoritması ve "metrics" parametresi ile de her eğitim sonrası sonuçları değerlendirmek için "accuracy" parametresi kullanılmıştır.

```
siniflandirici.compile(optimizer="adam", loss="categorical_crossentropy",  
metrics=["accuracy"])
```

3.7.7 Modelin Eğitilmesi

Model eğitimi sırasında, veriler genellikle aynı anda tamamıyla değil, parçalar halinde eğitime dahil edilir. Bu parçalar, "batch_size" olarak adlandırılır.

```
In [40]: #Eğitim  
history = siniflandirici.fit(X_train,Y_train,  
                             epochs=1000,  
                             verbose=1,  
                             validation_split=0.1,  
                             batch_size=1024)
```

```
Epoch 1/1000  
52/52 [=====] - 1s 6ms/step - loss: 1.8355 - accuracy: 0.2424 - val_loss: 1.6988 - val_a  
ccuracy: 0.2953  
Epoch 2/1000  
52/52 [=====] - 0s 5ms/step - loss: 1.7483 - accuracy: 0.2609 - val_loss: 1.6384 - val_a  
ccuracy: 0.3307  
Epoch 3/1000  
52/52 [=====] - 0s 6ms/step - loss: 1.6984 - accuracy: 0.2716 - val_loss: 1.6034 - val_a  
ccuracy: 0.3394  
Epoch 4/1000  
52/52 [=====] - 0s 6ms/step - loss: 1.6633 - accuracy: 0.2820 - val_loss: 1.5620 - val_a  
ccuracy: 0.3435  
Epoch 5/1000  
52/52 [=====] - 0s 6ms/step - loss: 1.6416 - accuracy: 0.2877 - val_loss: 1.5497 - val_a  
ccuracy: 0.3413  
Epoch 6/1000  
52/52 [=====] - 0s 6ms/step - loss: 1.6284 - accuracy: 0.2864 - val_loss: 1.5295 - val_a  
ccuracy: 0.3418  
Epoch 7/1000  
52/52 [=====] - 0s 6ms/step - loss: 1.6057 - accuracy: 0.2910 - val_loss: 1.5205 - val_a  
ccuracy: 0.3418
```

Şekil 3.16: Modelin eğitimi ve hiper parametreleri

Şekil 3.16'daki model 1000 epochs, 1024'lük mini yığınlar kullanılarak eğitilmiştir. Ayrıca, modelin performansı ve kayıp değeri, doğrulama veri setinden gözlemlenmiştir. Bu amaçla, validation_split parametresi ile veri setinin %10'u ayrılmıştır.

Tablo 3.9: Eğitim veri setinin train, validation ve test olarak ayrılması

	Train (Eğitim)						Validation (Doğrulama)		Test	
Index	1	2	3	4	5	6	7	8	9	10
Girdi_1	x	x	x	x	x	x	x	x	x	x
Girdi_2	x	x	x	x	x	x	x	x	x	x
Girdi_n	x	x	x	x	x	x	x	x	x	x
Hedef	y	y	y	y	y	y	y	y	y	y
	%81						%9		%10	

Tablo 3.9’da gösterildiği üzere validation bölümü, eğitim veri setinden ayrılarak seçilir. Eğitim veri seti üzerinde model seçimi yapılırken, validation bölümünde ise modelin iyileştirilmesi amaçlanır. Bu amaçla, hiper parametre ayarlamaları gibi yöntemler kullanılarak en uygun katsayılar veya ağırlıklar bulunmaya çalışılır. Özellikle büyük veri setleri üzerinde sürekli eğitim yapmak mümkün olmadığından, eğitim setinden küçük bir bölüm alınarak validation seti oluşturulur. Bu set, validation_split parametresiyle belirlenir.

3.7.7.1 Epoch Sayısı

Epoch sayısı, eğitim sırasında tüm eğitim verilerinin ağı kaç kez gösterildiğini belirtir. Eğitim ilerledikçe, eğitim doğruluğu artabilir; ancak, validation doğruluğu düşmeye başlayana kadar epoch sayısını artırmak önemlidir. En yüksek doğruluğa ulaşıldığında, epoch sayısı ideal olarak belirlenmiş olur.

3.7.7.2 Batch Size

Batch_size, veri setinin küçük gruplara bölünerek öğrenme işleminin bu küçük gruplar üzerinde gerçekleştirilmesini sağlar. Bu şekilde, birden fazla girdinin parçalar halinde işlenmesi, "mini-batch" olarak adlandırılır. Model tasarlanırken belirlenen mini-batch parametresi, modelin aynı anda işleyeceği veri miktarını belirtir. Batch size genellikle 32 ile başlar ve eğitim setinin büyüklüğüne göre değeri ikinin katları şeklinde artar.

3.7.8 Eğitilen Modelin Test Edilmesi

Modelimizi eğittikten sonra test veri seti üzerinde performansını görmek için `siniflandirici.evaluate(X_test, Y_test)` metodunu kullanmamız gerekmektedir.



4. BULGULAR VE TARTIŞMA

Bu bölümde toplanan veriler yapay sinir ağları kullanılarak analiz edilerek yorumlanmıştır.

4.1. Rapid Miner Programı ile Ekonomik Değişkenler Kullanılarak PVMATH_KODLAMA ve PVREAD_KODLAMA Değişkenlerinin Tahmin Edilmesi

Şekil 3.4' te Rapid Miner modeli görülen ekonomik değişkenler ile PVMATH_KODLAMA değişkeninin tahmin edilmesinde model 25642 PVMATH_KODLAMA verisini doğru olarak tahmin etmiştir. Modelin performans tablosu tablo 4.1'de gösterilmiştir.

Tablo 4.1: Modelin performans tablosu accuracy (doğruluk): %39.45

	true 1	true 2	true 4	true 3	true 6	true 5	true 0	class precision
pred. 1	5497	2696	378	1220	42	139	75	54.71%
pred. 2	3377	7572	1869	4179	253	671	6	42.24%
pred. 4	284	1206	4029	2705	1611	2649	0	32.27%
pred. 3	1361	3788	4168	6515	789	2021	2	34.94%
pred. 6	25	127	740	364	1131	856	0	34.88%
pred. 5	41	174	653	363	526	894	0	33.72%
pred. 0	0	1	0	1	0	0	4	66.67%
class recall	51.93%	48.65%	34.04%	42.45%	25.99%	12.37%	4.60%	

Tablo 4.1'e göre EK 1'de gösterilen 18 adet ekonomik değişken PISA 2022'ye katılan öğrencilerin matematik becerisi puanını tahmin etmede %39.45 oranında başarılı olmuştur.

Şekil 3.4' te Rapid Miner modeli görülen EK 1'de gösterilen ekonomik değişkenler ile PVREAD_KODLAMA değişkeninin tahmin edilmesinde model 28134 PVREAD_KODLAMA verisini doğru olarak tahmin etmiştir. Modelin performans tablosu tablo 4.2'de gösterilmiştir.

Tablo 4.2: Modelin performans tablosu accuracy (doğruluk): %43.28

	true 2	true 4	true 5	true 3	true 1	true 6	true 0	class precision
pred. 2	7103	1407	272	4173	3029	22	14	44.34%
pred. 4	1607	7352	3753	4415	422	952	0	39.74%
pred. 5	280	1961	2673	936	47	972	0	38.91%
pred. 3	3786	3756	1120	7352	1046	133	1	42.76%
pred. 1	1737	170	22	664	3574	4	56	57.40%
pred. 6	1	44	45	19	2	79	0	41.58%
pred. 0	0	0	0	0	0	0	1	100.00%
class recall	48.94%	50.05%	33.90%	41.87%	44.01%	3.65%	1.39%	

Tablo 4.2'ye göre EK 1'de gösterilen 18 adet ekonomik değişken PISA 2022'ye katılan öğrencilerin okuma becerisi puanını tahmin etmede %43.28 oranında başarılı olmuştur.

4.2. Rapid Miner Programı ile Ekonomik Değişkenler Kullanılarak PVMATH_1_ ve PVREAD_1_2 Değişkenlerinin Tahmin Edilmesi

Şekil 3.4'te Rapid Miner modeli görülen EK 1'de gösterilen ekonomik değişkenler ile PVMATH_1_2 değişkeninin tahmin edilmesinde model 65002 veriden 50102 PVMATH_1_2 verisini doğru olarak tahmin etmiştir. Modelin performans tablosu tablo 4.3'te gösterilmiştir.

Tablo 4.3: Modelin performans tablosu accuracy (doğruluk): %77.08

	true 0	true 1	class precision
pred. 0	29125	4733	86.02%
pred. 1	10167	20977	67.35%
class recall	67.88%	85.38%	

Tablo 4.3'e göre EK 1'de gösterilen 18 adet ekonomik değişken PISA 2022'ye katılan öğrencilerin matematik becerisi alanında başarılı-başarısız olarak tahmin etmede %77.08 oranında başarılı olmuştur. Yani sosyo-ekonomik koşulların, öğrencilerin matematik becerisi başarı durumlarının doğrudan önemli bir yordayıcısı olduğu görülmektedir.

Şekil 3.4'te Rapid Miner modeli görülen EK 1'de gösterilen ekonomik değişkenler ile PVREAD_1_2 değişkeninin tahmin edilmesinde model 65002 veriden 51173

PVREAD_1_2 verisini doğru olarak tahmin etmiştir. Modelin performans tablosu tablo 4.4'te gösterilmiştir.

Tablo 4.4: Modelin performans tablosu accuracy (doğruluk): %78.73

	true 0	true 1	class precision
pred. 0	29913	4397	87.18%
pred. 1	9432	21260	69.27%
class recall	72.00%	86.70%	

Tablo 4.4'e göre seçilen EK 1'de gösterilen 18 adet ekonomik değişken PISA 2022'ye katılan öğrencilerin okuma becerisi alanında başarılı-başarısız olarak tahmin etmede %78.73 oranında başarılı olmuştur. Yani sosyo-ekonomik koşulların, öğrencilerin okuma becerisi başarı durumlarının doğrudan önemli bir yordayıcısı olduğu görülmektedir.

4.3. Rapid Miner Programı ile Anne Baba Eğitim Durumu ve Ebeveyn Tutumları Değişkenleri ile PVMATH_KODLAMA ve PVREAD_KODLAMA Değişkenlerinin Tahmin Edilmesi

EK 2'deki anne baba eğitim durumu ve ebeveyn tutumları değişkenlerinin Şekil 3.5'teki modele uygulanması ile oluşturulan bu modelde 48541 adet veri eğitim ve test olarak ayrılırken örnekleme yöntemi olarak shuffled sampling yöntemi kullanılmıştır. Model 18911 adet PVMATH_KODLAMA puanını doğru olarak tahmin etmiştir. Modelin performans tablosu Tablo 4.5'te gösterilmiştir.

Tablo 4.5: Modelin performans tablosu accuracy (doğruluk): %38.96

	true 5	true 6	true 4	true 3	true 2	true 1	true 0	class precision
pred. 5	2088	1033	1168	647	374	117	0	38.47%
pred. 6	925	1811	698	403	161	41	0	44.84%
pred. 4	2897	1839	4567	2915	1484	461	0	32.25%
pred. 3	1183	708	2446	4248	2226	861	1	36.39%
pred. 2	413	195	900	1696	3324	1420	3	41.81%
pred. 1	132	59	252	580	1388	2873	4	54.33%
pred. 0	0	0	0	0	0	0	0	0.00%
class recall	27.34%	32.08%	45.53%	40.50%	37.11%	49.77%	0.00%	

Tablo 4.5'e göre EK 2'de gösterilen 27 adet aile eğitim durumu ve ebeveyn tutumları etiketlerini içeren bağımsız değişken PISA 2022'ye katılan öğrencilerin matematik becerisi puanını tahmin etmede %38.96 oranında başarılı olmuştur.

EK 2'deki anne baba eğitim durumu ve ebeveyn tutumları değişkenlerinin Şekil 3.5'teki modele uygulanması ile oluşturulan bu modelde 48541 adet veri eğitim ve test olarak ayrılırken örnekleme yöntemi olarak shuffled sampling yöntemi kullanılmıştır. Model 20602 adet PVREAD_KODLAMA puanını doğru olarak tahmin etmiştir. Modelin performans tablosu Tablo 4.6'da gösterilmiştir.

Tablo 4.6: Modelin performans tablosu accuracy (doğruluk): %42.44

	true 4	true 5	true 3	true 2	true 6	true 1	true 0	class precision
pred. 4	7740	4182	3947	1698	902	369	1	41.08%
pred. 5	2127	3594	745	192	1421	42	0	44.26%
pred. 3	3316	1347	6166	2966	195	1267	4	40.40%
pred. 2	507	116	957	2270	12	1039	1	46.31%
pred. 6	36	76	15	4	135	0	0	50.75%
pred. 1	41	9	145	255	1	694	3	60.45%
pred. 0	0	0	1	0	0	0	3	75.00%
class recall	56.22%	38.55%	51.49%	30.74%	5.06%	20.35%	25.00%	

Tablo 4.6'ya göre EK 2'de gösterilen 33 adet aile eğitim durumu ve ebeveyn tutumları etiketlerini içeren bağımsız değişken PISA 2022'ye katılan öğrencilerin okuma becerisi puanını tahmin etmede %42.44 oranında başarılı olmuştur.

4.4. Rapid Miner Programı ile Anne Baba Eğitim Durumu ve Ebeveyn Tutumları Değişkenleri ile PVMATH_1_2 ve PVREAD_1_2 Değişkenlerinin Tahmin Edilmesi

EK 2'deki anne baba eğitim durumu ve ebeveyn tutumları değişkenlerinin Şekil 3.5'teki modele uygulanması ile oluşturulan bu modelde 48541 adet veri eğitim ve test olarak ayrılırken örnekleme yöntemi olarak shuffled sampling yöntemi kullanılmıştır. Model 37914 adet PVMATH_1_2 puanını doğru olarak tahmin etmiştir. Modelin performans tablosu Tablo 4.7'de gösterilmiştir.

Tablo 4.7: Modelin performans tablosu accuracy (doğruluk):%78.11

	true 1	true 0	class precision
pred. 1	18299	3868	82.55%
pred. 0	6759	19615	74.37%
class recall	73.03%	83.53%	

Tablo 4.7'ye göre EK 2'de gösterilen 27 adet aile eğitim durumu ve ebeveyn tutumları değişkenleri PISA 2022'ye katılan öğrencilerin matematik becerisi alanında başarılı-başarısız olarak tahmin etmede %78.11 oranında başarılı olmuştur. Yani aile eğitim durumu ve ebeveyn tutumları değişkenlerinin, öğrencilerin matematik becerisi başarı durumlarının doğrudan önemli bir yordayıcısı olduğu görülmektedir.

EK 2'deki anne baba eğitim durumu ve ebeveyn tutumları değişkenlerinin Şekil 3.5'teki modele uygulanması ile oluşturulan bu modelde 48541 adet veri eğitim ve test olarak ayrılırken örnekleme yöntemi olarak shuffled sampling yöntemi kullanılmıştır. Model 37493 adet PVREAD_1_2 puanını doğru olarak tahmin etmiştir. Modelin performans tablosu Tablo 4.8'de gösterilmiştir.

Tablo 4.8: Modelin performans tablosu accuracy (doğruluk): %77.24

	true 1	true 0	class precision
pred. 1	18474	2997	86.04%
pred. 0	8051	19019	70.26%
class recall	69.65%	86.39%	

Tablo 4.8'e göre EK 2'de gösterilen 33 adet aile eğitim durumu ve ebeveyn tutumları değişkenleri PISA 2022'ye katılan öğrencilerin okuma becerisi alanında başarılı-başarısız olarak tahmin etmede %77.24 oranında başarılı olmuştur. Yani aile eğitim durumu ve ebeveyn tutumları değişkenlerinin, öğrencilerin okuma becerisi başarı durumlarının doğrudan önemli bir yordayıcısı olduğu görülmektedir.

4.5. Python ve Çok Katmanlı Yapay Sinir Ağları Kullanılarak Ekonomik Değişkenler ile PVMATH_KODLAMA ve PVREAD_KODLAMA Değişkenlerinin Tahmin Edilmesi

Python ve Keras kütüphanesi kullanılarak çok katmanlı yapay sinir ağı oluşturulmuş ve EK 1'de gösterilen 18 adet ekonomik değişkenin PISA 2022'ye katılan öğrencilerin PVMATH_KODLAMA (Matematik becerisi puanı) ve PVREAD_KODLAMA (Okuma becerisi puanı)'nı tahmin etmesi istenmiştir. Veri setinde bulunan 65002 adet verinin 58502 tanesi eğitim, 6501 tanesi test verisi olarak kullanılmış olup model 1000 epochs (tur)'da eğitilmiştir. Modelin sınıflandırma başarısı Şekil 4.1'de görülmektedir.

```
In [41]: siniflandirici.evaluate(X_test, Y_test)
204/204 [=====] -
Out[41]: [84.60359191894531, 0.2841101288795471]
```

```
In [37]: siniflandirici.evaluate(X_test, Y_test)
204/204 [=====]
Out[37]: [4.041178226470947, 0.3447161912918091]
```

Şekil 4.1: Modelin sınıflandırma başarısı

Şekil 4.1 incelendiğinde yapay sinir ağının matematik becerisi puanını tahmin etme oranı %28, okuma becerisi puanını tahmin etme oranı %34 olmuştur.

4.6. Python ve Çok Katmanlı Yapay Sinir Ağları Kullanılarak Ekonomik Değişkenler ile PVMATH_1_2 ve PVREAD_1_2 Değişkenlerinin Tahmin Edilmesi

Python ve Keras kütüphanesi oluşturulan yapay sinir ağının, EK 1'de gösterilen 18 adet ekonomik değişken kullanılarak, PISA 2022'ye katılan öğrencilerin PVMATH_1_2 (Matematik becerisi) ve PVREAD_1_2 (Okuma becerisi) alanında başarılı (1) ve başarısız (0) olarak tahmin etmesi istenmiştir. Veri setinde bulunan 65002 adet verinin 58502 tanesi eğitim, 6501 tanesi test verisi olarak kullanılmış olup model 1000 epochs (tur)'da eğitilmiştir. Modelin sınıflandırma başarısı Şekil 4.2'de görülmektedir.

```
In [20]: siniflandirici.evaluate(X_test, Y_test)
204/204 [=====]
Out[20]: [5.086200714111328, 0.6645131707191467]
```

```
In [20]: siniflandirici.evaluate(X_test, Y_test)
204/204 [=====]
Out[20]: [6.147380352020264, 0.6858944892883301]
```

Şekil 4.2: Modelin performans tablosu

Şekil 4.2 incelendiğinde yapay sinir ağının matematik becerisi alanı için %66, okuma becerisi alanı için %68 oranında doğru olarak tahmin etmiştir.

4.7. Python ve Çok Katmanlı Yapay Sinir Ağları Kullanılarak Anne Baba Eğitim Durumu ve Ebeveyn Tutumları Değişkenleri ile PVMATH_KODLAMA ve PVREAD_KODLAMA Değişkenlerinin Tahmin Edilmesi

Python ve Keras kütüphanesi kullanılarak çok katmanlı yapay sinir ağı oluşturulmuş ve EK 2'de gösterilen 33 adet aile eğitim durumu ve ebeveyn tutumu değişkeni içinden Tablo 3.2'de de görüldüğü üzere 27 tanesi PISA 2022'ye katılan öğrencilerin PVMATH_KODLAMA (Matematik becerisi puanı) ve 33 tanesi PVREAD_KODLAMA (Okuma becerisi puanı)'nı tahmin etmede kullanılmıştır. Veri setinde bulunan 48541 adet verinin 43686 tanesi eğitim 4815 tanesi test verisi olarak kullanılmış olup model 1000 epochs (tur)'da eğitilmiştir. Modelin sınıflandırma başarısı Şekil 4.3'te görülmektedir.

```
In [159]: siniflandirici.evaluate(X_test, Y_test)  In [89]: siniflandirici.evaluate(X_test, Y_test)
152/152 [=====]                                152/152 [=====]
Out[159]: [16.04941177368164, 0.27929967641830444] Out[89]: [24.59345054626465, 0.3132852613925934]
```

Şekil 4.3: Modelin sınıflandırma başarısı

Şekil 4.3 incelendiğinde yapay sinir ağının matematik becerisini tahmin etme oranı %27, okuma becerisini tahmin etme oranı %31 olmuştur.

4.8. Python ve Çok Katmanlı Yapay Sinir Ağları Kullanılarak Anne Baba Eğitim Durumu ve Ebeveyn Tutumları Değişkenleri ile PVMATH_1_2 ve PVREAD_1_2 Değişkenlerinin Tahmin Edilmesi

Python ve Keras kütüphanesi oluşturulan yapay sinir ağının, EK 2'de gösterilen 33 adet anne baba eğitim durumu ve ebeveyn tutumları etiketlerini içeren değişken kullanılarak, PISA 2022'ye katılan öğrencilerin PVMATH_1_2 (matematik becerisi) ve PVREAD_1_2 (okuma becerisi) alanında başarılı (1) ve başarısız (0) olarak tahmin etmesi istenmiştir. Veri setinde bulunan 48541 adet verinin 43686 tanesi eğitim, 4855 tanesi test verisi olarak kullanılmış

olup model 1000 epochs (tur)'da eğitilmiştir. Modelin sınıflandırma başarısı Şekil 4.4'te görülmektedir.

```
In [20]: siniflandirici.evaluate(X_test, Y_test) In [22]: siniflandirici.evaluate(X_test, Y_test)
152/152 [=====] 152/152 [=====]
Out[20]: [8.430556297302246, 0.6508753895759583] Out[22]: [2.2250616550445557, 0.6459320187568665]
```

Şekil 4.4: Modelin sınıflandırma başarısı

Şekil 4.4 incelendiğinde yapay sinir ağının matematik becerisini tahmin etme oranı %65, okuma becerisini tahmin etme oranı %64 olmuştur.

4.9. Rapid Miner ile Python Derin Öğrenme Ağının Performans Karşılaştırması

Literatür incelendiğinde Rapid Miner ve Python derin öğrenme ağı ile ilgili yapılmış çok sayıda çalışmaya rastlanmıştır ve Rapid Miner ile derin öğrenme yönteminin daha başarılı sonuçlar verdiği görülmüştür. Swetha (2020) araştırmasında, yoğun bakımda en fazla risk altında olan hastaları tahmin etmek için en gelişmiş derin öğrenmeyi geleneksel makine öğrenimi algoritmalarıyla karşılaştırmıştır. Python-Keras, R, Rapid Miner-Derin öğrenme ile oluşturulan derin sinir ağı sınıflandırıcılarının doğruluk, kesinlik, geri çağırma, özgüllük, AUC ve eğitim süresi açısından Lojistik Regresyon, GLM, Naive Bayes, Rastgele Orman ve Karar Ağacı gibi geleneksel yöntemlerden daha iyi performans gösterdiği gösterilmiştir. Apaza-Pinto vd.'nin (2022) fotovoltaik güneş enerjisiyle çalışan ekipmanın pillerinin şarj durumunu tahmin etmeyi amaçlayan çalışmalarında Rapid Miner ve Derin Öğrenme ağının %90 Python ve Tensorflow kütüphanesiyle oluşturulan yapay sinir ağının ise %84 başarıma ulaştığını göstermektedir.

Tablo 4.9: Rapid Miner ve Python arasındaki başarı kıyaslaması (0-6 kategorili)

	Rapid Miner	Python
Ekonomik değişkenler kullanılarak matematik becerisinin tahmini	%39.45	%28.41
Ekonomik değişkenler kullanılarak okuma becerisinin tahmini	%43.28	%34.47
Anne baba eğitim durumları ve ebeveyn tutumları değişkenlerinin matematik becerisinin tahmini	%38.96	%27.92
Anne baba eğitim durumları ve ebeveyn tutumları değişkenlerinin okuma becerisinin tahmini	%42.44	%31.32

Tablo 4.9'a göre Rapid Miner ve Python yapay sinir ağı arasında EK1'de gösterilen ekonomik değişkenler kullanılarak matematik becerisi tahmininde %11'lik fark, okuma becerisi tahmininde ise %9'luk bir fark bulunmaktadır. EK 2'de gösterilen aile eğitim durumları ebeveyn tutumları değişkenleri kullanılarak matematik becerisi tahmin etmede ise %11, okuma becerisi tahmininde ise %11 fark olduğu görülmektedir. Bu bağlamda Rapid Miner programı ve derin öğrenmenin daha başarılı olduğu söylenebilir. Yapılan çalışmayı destekler nitelikte olan Baharun vd.'nin (2022) çalışmalarında Rapid Miner ve Python gibi makine öğrenimi yazılım araçları, otomatik modelleme için farklı yaklaşımlar karşılaştırılmıştır. Sonuçlar, Rapid Miner ve Rastgele Orman Regresörlerinin, tüm otomatik modelleme sürecinde yüksek doğruluk puanları elde etme konusunda daha başarılı olduğunu göstermiştir.

Tablo 4.10: Rapid Miner ve Python arasındaki başarı kıyaslaması (0-1 kategorili)

	Rapid Miner	Python
Ekonomik değişkenler kullanılarak matematik becerisinin tahmini	%77.08	%66.45
Ekonomik değişkenler kullanılarak okuma becerisinin tahmini	%78.73	%68.58
Anne baba eğitim durumları ve ebeveyn tutumları değişkenlerinin matematik becerisinin tahmini	%78.11	%65.08
Anne baba eğitim durumları ve ebeveyn tutumları değişkenlerinin okuma becerisinin tahmini	%77.24	%64.59

Tablo 4.10'a göre Rapid Miner ve Python yapay sinir ağı arasında EK 1'de gösterilen ekonomik değişkenler kullanılarak matematik becerisi tahmininde %11'lik fark, okuma becerisi tahmininde ise %10'luk bir fark bulunmaktadır. EK 2'de gösterilen aile eğitim

durumları ebeveyn tutumları deęiřkenleri kullanılarak matematik becerisi tahmin etmede ise %13, okuma becerisi tahmininde ise %13'lük fark olduęu görölmektedir. Bu bağlamda Rapid Miner programı ve derin öğrenmenin daha başarılı olduęu söylenebilir.



5. SONUÇLAR

Yapılan çalışma ile ekonomik değişkenlerin öğrencilerin matematik becerisi %38, okuma becerisi puanlarını %43 oranında etkilediği bulunmuştur. Cingöz ve Gür (2020) tarafından yapılan araştırmada, öğrencilerin ekonomik statüleri, toplumsal pozisyonları ve kültürel durumları yükseldikçe sınav puanlarının da arttığını görülmektedir. Çalışmada öğrencilerin ekonomik, sosyal ve kültürel statü özelliklerinin TEOG puanlarına %38 oranında etki ettiği ortaya konmuştur. Araştırmada, öğrencilerin sosyoekonomik statüsündeki her bir birimlik artışın, TEOG puanlarında 60.562 puanlık bir yükselmeye neden olacağı tespit edilmiş ve bu bulgular, öğrencilerin başarısındaki sosyoekonomik faktörlerin önemini vurgulamaktadır. Ayrıca farklı sınav türlerinde ve sınıf kademelerinde sosyoekonomik faktörlerin öğrenci başarısında önemli bir etken olduğu görülmektedir.

Uluslararası Matematik ve Fen Bilgisi Çalışması-TIMSS, ilköğretim 4. ve 8. Sınıf öğrencilerinin matematik ve fen bilgisi alanındaki başarı seviyelerini çeşitli değişkenler bakımından ölçen uluslararası bir çalışmadır. Öztürk ve Uçar'ın (2010) çalışması, TIMSS 2007 verileri üzerine yapılmış olup ve Türkiye ve Tayvan'daki öğrencileri çeşitli değişkenler üzerinden karşılaştırmış ve sosyoekonomik faktörlerin ülkelerin eğitim seviyesini belirlemede önemli etkileri olduğunu ortaya koymuştur. Genellikle ekonomik açıdan güçlü olan ülkelerin eğitim başarısı yüksekken, zayıf ekonomiye sahip ülkelerin başarısı düşüktür. Tayvan'da kişi başına düşen gelir 13235 dolar iken Türkiye'de bu rakam 3130 dolardır (Martin, 2008). Bu durum, Tayvan'ın TIMSS verilerinde Türkiye'ye göre daha başarılı olmasının bir göstergesi olabilir. Ayrıca, eğitim seviyesi düşük olan ailelerin çocuklarının eğitime olan katkısının sınırlı olabileceği ve bu nedenle eğitilmiş ailelerin çocuklarını yetiştirmenin daha bilinçli olabileceği de vurgulanmıştır. Martin'e (2008) göre Tayvan'daki üniversite mezunu ailelerin oranının Türkiye'nin 2 katı daha yüksek olması, bu ülkedeki öğrencilerin başarısının bir diğer nedeni olmaktadır.

Yapılan çalışma Öztürk ve Uçar'ın (2010) çalışmalarını destekler niteliktedir. Buna göre ekonomik değişkenlerin matematik başarısını (PVMATH_1_2) tahmin etme oranı %77 iken okuma becerisi puanını (PVREAD_1_2) tahmin etme oranı %78 olarak bulunmuştur. Yine aile eğitim durumu ve ebeveyn tutumları değişkenlerinin matematik başarısını (PVMATH_1_2) tahmin etme oranı %78, okuma becerisi (PVREAD_1_2) başarı durumunu

tahmin etme oranı %77 olarak bulunmuş ve hem ekonomik değişkenlerin hem de aile eğitim durumu ve ebeveyn tutumlarının öğrencilerin başarı durumlarını yordayan önemli faktörler oldukları görülmektedir.

Aksu, Aksu ve Sarıcaoğlu'nun (2023) çalışmalarında PISA 2015 sınavı verilerinden, öğrencilerin matematik becerisi başarı düzeyleri veri madenciliği ve yapay sinir ağıları yöntemleri kullanılarak tahmin edilmiş ve matematik okuryazarlığını etkileyen değişkenler belirlenmiştir. Buna göre Norveç, Singapur, Amerika ve Türkiye örneklemelerinde incelenen 15 bağımsız değişkenden matematik becerisini etkileyen ön önemli değişkenin sosyo-ekonomik durum indeksi, Japonya örnekleminde birinci sırada matematik öğrenme süresi ikinci sırada ise baba eğitim durumu olduğu görülmüştür. Bu çalışmada da yine benzer bir sonuç elde edildiği belirlenmiş olup PISA 2015-2022 arasında halen öğrencilerin sosyo-ekonomik düzeylerinin başarıyı yordayan en önemli unsur olduğu belirlenmiştir.

Yapılan çalışma ile aile eğitim durumu ve ebeveyn tutumları değişkenlerin öğrencilerin matematik becerisi (PVMATH_KODLAMA) %38, okuma becerisi (PVREAD_KODLAMA) puanlarını %42 oranında tahmin ettiği bulunmuştur. Kemer (2022) yüksek lisans tez çalışmasında sadece anne baba eğitim seviyesinin PISA 2018 öğrenci başarısına etkisini araştırmış, anne, baba ve en yüksek eğitim düzeyine sahip ebeveynin eğitim düzeyi arttıkça öğrencinin akademik başarısının bazı sapmalar olsa da arttığı bulmuştur. Bu bağlamda yıllar bazında PISA sonuçları ve yapılan çalışmalar incelendiğinde anne baba eğitim durumunun öğrenci başarısını önemli ölçüde yordadığı görülmektedir.

6. ÖNERİLER

PISA 2022 sınavına 81 ülke katılmıştır. Ayrık veriler temizlendiğinde 42 ülkenin verisi kalmıştır. OECD web sitesinden SPSS formatında indirilen verilerden gerekli bağımsız değişkenler seçildikten sonra 613745 satır EXCEL verisi görülmüşken, ayrık veriler temizlendiğinde ekonomik değişkenler için 65003 satır EXCEL verisi, aile eğitim durumu ve ebeveyn tutumları değişkenleri için ise 48542 satır EXCEL verisi kalmıştır. Çok katmanlı derin sinir ağının daha yüksek başarı oranlarına ulaşabilmesi için daha fazla eğitim verisine ihtiyaç duyulmaktadır. PISA bir bakıma anket soruları ile öğrencilerin, öğretmenlerin, ebeveynlerin anket sorularına verdikleri cevaplarla öğrenci başarısı arasındaki ilişkiyi incelemektedir. Bu bağlamda, öğrencilerin anketlere katılımının teşvik ederek, ekonomik koşulların, ebeveyn tutumlarının, okul araç gereçlerinin tam olup olmamasının ve başka faktörlerin öğrenci başarısına hangi derecede etki ettiğini tam manasıyla ölçülmesi sağlanabilir.

Daha fazla eğitim verisi elde edebilmek için farklı yapay sinir ağı mimarileri kullanılabilir. Bu amaçla, ayrık verilerden oluşan boşluklar hibrit bir yapay sinir ağı ile doldurulup elde edilen verilerle ağ tekrardan eğitilebilir. Bu şekilde derin öğrenme ağının başarısı bir miktar daha artabilir. Fakat öğrenci ve ailesi içinde yaşanan o günlük maddi ve manevi sorunlar, okul ve öğretmenler ile ilgili yaşanan olumsuz durumlar anket sorularına verilen cevaplarda değişimlere yol açabilir. Bu gibi parametrelerin de öğrenci başarısını bir miktar etkileyeceği göz önünde bulundurulmalıdır.

Değişik yıllarda PISA verileri kullanılarak yapılan çalışmalar incelendiğinde öğrencilerin başarılarını etkileyen en önemli faktörlerden birinin sosyo-ekonomik durum değişkeni olduğu görülmüştür. Bu bakımdan farklı gelir grubundaki öğrencilerin aynı eğitim araç-gereçlerine, beslenme, giyim, barınma bakımından da aynı şartlarda eşit eğitim görmelerinin önem arz etmekte olduğunu yapılan çalışmaların sonuçlarından görmekteyiz. Okul ve eğitim sistemine ilişkin politikalar geliştirilirken bu bulgunun özellikle üzerinde durulması, sosyo-ekonomik açıdan zayıf öğrencilere ilişkin çalışmalar yapılması gerekmektedir.

Yapılan çalışmada aile eğitim durumları ve ebeveyn tutumları değişkenleri birlikte ele alınmış ve öğrencilerin matematik becerisi ve okuma becerileri tahmin edilmiştir. İleriki

alışmalarda aile eđitim durumlarını gsteren deđiřkenler ve ebeveyn tutumlarını etiketlerini ieren deđiřkenler ayrı ayrı ele alınarak bu deđiřkenlerin matematik ve okuma becerisi tahmininde kullanılabilir. đrenci bařarısını artırmak iin yapılacak alışmalarda bu bulgunun dikkate alınması faydalı olacaktır.



KAYNAKLAR

- Agatonovic-Kustrin, S. ve Beresford, R. (2000). Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. *Journal of pharmaceutical and biomedical analysis*, 22(5), 717-727.
- Aksu, N., Aksu, G. ve Saracaloglu, S. (2022). Prediction of the factors affecting PISA mathematics literacy of students from different countries by using data mining methods. *International Electronic Journal of Elementary Education*, 14(5), 613-629.
- Apaza-Pinto, A., Esquicha-Tejada, J., López-Casaperalta, P. ve Sulla-Torres, J. (2022). Supervised Machine Learning Techniques for the Prediction of the State of Charge of Batteries in Photovoltaic Systems in the Mining Sector. *IEEE Access*, 10, 134307-134317.
- Arisariyawong, S. ve Charoenseang, S. (2002, December). Dynamic self-organized learning for optimizing the complexity growth of radial basis function neural networks. In *2002 IEEE International Conference on Industrial Technology*, 1, 655-660.
- AYDIN, Ahmet, "Derin Öğrenme", Bilgisayar Bilimlerinde Teorik Ve Uygulamalı Araştırmalar, Efe Akademik Yayıncılık, İstanbul 2021, 105-127.
- Baharun, N., Razi, N. F. M., Masrom, S., Yusri, N. A. M. ve Rahman, A. S. A. (2022). Auto modelling for machine learning: a comparison implementation between rapid miner and python. *Int. J. Emerg. Technol. Adv. Eng*, 12(5), 15-27.
- Bhardwaj, A., Di, W. ve Wei, J. (2018). *Deep Learning Essentials: Your hands-on guide to the fundamentals of deep learning and neural network modeling*. Packt Publishing Ltd.
- Bi, Q., Goodman, K. E., Kaminsky, J. ve Lessler, J. (2019). What is machine learning? A primer for the epidemiologist. *American journal of epidemiology*, 188(12), 2222-2239.
- Bulut, F. (2016). Çok Katmanlı Algılayıcılar İle Doğru Meslek Tercihi. *Anadolu University Journal of Science and Technology A-Applied Sciences and Engineering*, 17(1), 97-109.
- Choi, R. Y., Coyner, A. S., Kalpathy-Cramer, J., Chiang, M. F. ve Campbell, J. P. (2020). Introduction to machine learning, neural networks, and deep learning. *Translational vision science & technology*, 9(2), 14-14.
- Cingöz, Z. K. ve Gür, B. S. (2020). Ekonomik, sosyal ve kültürel statünün akademik başarıya etkisi PISA 2015 ve TEOG 2017 sonuçlarının karşılaştırması. *İnsan ve Toplum*, 10(4), 247-288.
- Datacamp. "Introduction to Activation Functions in Neural Networks". Erişim:07.06.2024. <https://www.datacamp.com/tutorial/introduction-to-activation-functions-in-neural-networks>.

- Dogo, E. M., Afolabi, O. J., Nwulu, N. I., Twala, B. ve Aigbavboa, C. O. (2018, December). A comparative analysis of gradient descent-based optimization algorithms on convolutional neural networks. In *2018 international conference on computational techniques, electronics and mechanical systems (CTEMS)*, 92-99.
- Dolezel, P., Skrabanek, P. ve Gago, L. (2016). Weight initialization possibilities for feedforward neural network with linear saturated activation functions. *IFAC-PapersOnLine*, 49(25), 49-54.
- Fancher, P. (2016). Composing artificial intelligence: Performing Whiteness and masculinity. *Present Tense*, 6(1), 1-8.
- French, R. M. (2000). The Turing Test: the first 50 years. *Trends in cognitive sciences*, 4(3), 115-122.
- Gardner, M. W. ve Dorling, S. R. (1998). Artificial neural networks (the multilayer perceptron)-a review of applications in the atmospheric sciences. *Atmospheric environment*, 32(14-15).
- Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., ... ve Chen, T. (2018). Recent advances in convolutional neural networks. *Pattern recognition*, 77, 354-377.
- Güre, Ö. B., Kayri, M. ve Erdoğan, F. (2020). PISA 2015 matematik okuryazarlığını etkileyen faktörlerin eğitsel veri madenciliği ile çözümlenmesi. *Eğitim ve Bilim*, 45(202).
- He, K., Zhang, X., Ren, S. ve Sun, J. (2015). Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proceedings of the IEEE international conference on computer vision* (pp. 1026-1034).
- Hinton, G., Srivastava, N. ve Swersky, K. (2012). Neural networks for machine learning lecture 6a overview of mini-batch gradient descent. *Cited on*, 14(8), 2.
- Hornik, K., Stinchcombe, M. and White, H. (1989) Multilayer feedforward networks are universal approximators. *Neural Networks* 2, 359-366.
- İban, M. C. ve Şentürk, E. (2021). İyonosfer Parametrelerinin Çok Katmanlı Algılayıcılar ile Kestirimi. *El-Cezeri*, 8(3), 1480-1494.
- İnal, H. ve Turabik. (2017). T. Matematik Başarısını Etkileyen Bazı Faktörlerin Yordama Gücünün Yapay Sinir Ağları İle Belirlenmesi. *Uşak Üniversitesi Eğitim Araştırmaları Dergisi*, 3(1), 23-50.
- İşeri, İ. ve Arıman, S. (2019). Sedimandaki Ağır Metal Konsantrasyonunun Çoklu Değişken Regresyon Modelleri ve Çok Katmanlı Algılayıcı Ağ Modeli ile Tahmini. *Avrupa Bilim ve Teknoloji Dergisi*, 389-397.
- Kabalci, E. (2017). Yapay Sinir Ağları. *Artificial Neural Networks*.

- Kaelbling, L. P., Littman, M. L. ve Moore, A. W. (1996). Reinforcement learning: A survey. *Journal of artificial intelligence research*, 4, 237-285.
- Karasar, N. (2006). *Bilimsel araştırma yöntemi*. Ankara: Nobel Yayın Dağıtım.
- KEMER, Yeliz, "PISA 2018 sonuçlarına göre Türkiye'deki aile eğitim düzeylerinin öğrenci akademik başarısına etkisinin incelenmesi", (Yüksek Lisans Tezi, Akdeniz Üniversitesi Eğitim Bilimleri Enstitüsü), Antalya 2022.
- Kim, H., Khan, M. U. K. ve Kyung, C. M. (2019). Efficient neural network compression. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 12569-12577).
- Knudsen, E. I. (1994). Supervised learning in the brain. *The Journal of Neuroscience*, 14(7), 3985.
- Koyuncu, İ. (2020). Investigation of Mathematics-Specific Trend Variables in PISA Studies with Neural Networks and Linear Regression. *Journal of Curriculum and Teaching*, 9(4), 40-54.
- Krizhevsky, A., Sutskever, I. ve Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25.
- Lu, J. (2022). Gradient descent, stochastic optimization, and other tales. *arXiv preprint arXiv:2205.00832*.
- Lu, W. ve Ding, Y. (2021, August). A network malicious traffic detection method based on semi-supervised deep learning. In *2021 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)* (pp. 1-6). IEEE.
- Martin, M.O., Mullis, I.V.S., Foy, P., Olson, J. E., Erberger, E., Preuschoff, C., Galia, J. (2008). TIMSS 2007 International Science Report: Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Mehrotra, D. (2019). *Basics of artificial intelligence & machine learning*. Notion Press.
- MEB. "PISA 2022 Türkiye Raporu". Erişim: 11.05.2024. <https://pisa.meb.gov.tr/www/pisa-2022-turkiye-raporu-yayimlandi/icerik/11>.
- MEB. "PISA Nedir?". Erişim: 01.05.2024. <https://pisa.meb.gov.tr/www/pisa-nedir/icerik/4>
- Naeem, S., Ali, A., Anam, S. ve Ahmed, M. M. (2023). An unsupervised machine learning algorithms: Comprehensive review. *International Journal of Computing and Digital Systems*.
- OECD. "What is PISA". Erişim:12.05.2024. <https://www.oecd.org/pisa/>

OECD. "PISA 2022 Database". Eriřim: 12.05.2024.

<https://www.oecd.org/pisa/data/2022database/>

Onan, A. ve Korukoęlu, S. (2016). Makine ęrenmesi yęntemlerinin gęrüş madencilięinde kullanılması ęzerine bir literatür arařtırması. *Pamukkale University Journal of Engineering Sciences*, 22(2).

Öztemel, E. (2003). *Yapay sinir aęlari*. (1. Baskı). İstanbul: PapatyaYayıncılık.

Özkan, İ ve Ülker, E. (2017). Derin ęrenme ve gęrüntü analizinde kullanılan derin ęrenme modelleri. *Gaziosmanpařa Bilimsel Arařtırma Dergisi*, 6(3), 85-104.

Öztürk, D. ve Uçar, S. (2010). Timss verileri kullanılarak tayvan ve türkiye'deki 8 sınıf ęrencilerinin fen başarısına etki eden faktörlerin belirlenmesi ve karřılařtırılması. *Çukurova Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 19(3), 241-256.

Panchapakesan, C., Palaniswami, M., Ralph, D. ve Manzie, C. (2002). Effects of moving the center's in an RBF network. *IEEE Transactions on Neural Networks*, 13(6), 1299-1307.

Prakash, S., Jalal, A. S. ve Pathak, P. (2023, March). Forecasting covid-19 pandemic using prophet, lstm, hybrid gru-lstm, cnn-lstm, bi-lstm and stacked-lstm for india. In *2023 6th International Conference on Information Systems and Computer Networks (ISCON)* (pp. 1-6). IEEE.

Radhakrishnan, S., Nair, S. G. ve Isaac, J. (2022). Multilayer perceptron neural network model development for mechanical ventilator parameters prediction by real time system learning. *Biomedical signal processing and control*, 71, 103170.

Rapidminer. "Split Data". Eriřim: 08.06.2024.

https://docs.rapidminer.com/latest/studio/operators/blending/examples/sampling/split_data.html

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

Savalia, S. ve Emamian, V. (2018). Cardiac arrhythmia classification by multi-layer perceptron and convolution neural networks. *Bioengineering*, 5(2), 35.

Schalkoff, R. (1992) *Pattern Recognition: Statistical. Structural and Neural Approaches*. Wiley, New York.

Sharma, S., Sharma, S. ve Athaiya, A. (2017). Activation functions in neural networks. *Towards Data Sci*, 6(12), 310-316.

Shinde, P. P. ve Shah, S. (2018). A review of machine learning and deep learning applications. In *2018 Fourth international conference on computing communication control and automation (ICCUBEA)* (pp. 1-6). IEEE.

- Simonyan, K. ve Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- Swetha, J. (2020). *Prediction of patient's health risk in critical care using a deep neural network* (Doctoral dissertation, Dublin Business School).
- Şen, Z. (2004). *Yapay sinir ağları*. İstanbul: Su Vakfı.
- Taulli, T. ve Oni, M. (2019). *Artificial intelligence basics*. Berkeley CA: Apress.
- TensorFlow. "Python". Erişim: 05.05.2024.
https://www.tensorflow.org/api_docs/python/tf/keras/optimizers/RMSprop
- Turkoglu, M., Hanbay, D. ve Sengur, A. (2022). Multi-model LSTM-based convolutional neural networks for detection of apple diseases and pests. *Journal of Ambient Intelligence and Humanized Computing*, 13(7), 3335-3345.
- Tyagi, A. K. ve Abraham, A. (Eds.). (2022). *Recurrent neural networks: Concepts and applications*. Boca Raton: CRC Press.
- Wang, Y., Cui, Z. ve Ke, R. (2023). *Machine Learning for Transportation Research and Applications*. Elsevier.
- Wen, W., Wu, C., Wang, Y., Chen, Y. ve Li, H. (2016). Learning structured sparsity in deep neural networks. *Advances in neural information processing systems*, 29, 2074-2082.
- Wiederhold, G. ve McCarthy, J. (1992). Arthur Samuel: Pioneer in machine learning. *IBM Journal of Research and Development*, 36(3), 329-331.
- Wikipedia. "Softmax Function". Erişim: 07.06.2024
https://en.wikipedia.org/wiki/Softmax_function
- YÖK. "Uluslararası İlişkiler Daire Başkanlığı". Erişim:13.05.2024.
<https://uluslararasi.yok.gov.tr/Sayfalar/avrupa-yuksekokretim-alani-ile-uyum-projesi/uluslararasi-standart-egitim-siniflamasi-isced.aspx>
- Veena, S., Shankari, T., Sowmiya, S., Varsha, M., Thiruverkadu, C. ve Thiruverkadu, C. (2020). A Survey on Tools Used For Machine Learning. *International Journal of Engineering Applied Sciences and Technology*, 4(9), 116-119.
- Yadav, A. K., Malik, H. ve Chandel, S. S. (2015). Application of rapid miner in ANN based prediction of solar radiation for assessment of solar energy resource potential of 76 sites in Northwestern India. *Renewable and Sustainable Energy Reviews*, 52, 1093-1106.
- Yang, J. ve Yang, G. (2018). Modified convolutional neural network based on dropout and the stochastic gradient descent optimizer. *Algorithms*, 11(3), 28.

- Yu, Y., Si, X., Hu, C. ve Zhang, J. (2019). A review of recurrent neural networks: LSTM cells and network architectures. *Neural computation*, 31(7), 1235-1270.
- Zhou, Z. H. (2018). A brief introduction to weakly supervised learning. *National science review*, 5(1), 44-53.
- Zinkevich, M., Weimer, M., Li, L. ve Smola, A. (2010). Parallelized stochastic gradient descent. *Advances in neural information processing systems*, 23.
- Zupan, J. (2003). Basics of artificial neural networks. *Data handling in science and technology*, 23, 199-229.



EKLER

EK 1: Ailenin Ekonomik Durumunu Gösteren Değişkenler

Değişken Açıklaması	Değişken Adı
Evinizde aşağıdakilerden hangileri bulunmaktadır?	
Kendinize ait oda.	ST250Q01JA
Okulda yapılan çalışmalar için kullanabileceğiniz bir bilgisayar (dizüstü, masaüstü bilgisayar ve tablet bilgisayar).	ST250Q02JA
Eğitim yazılımı.	ST250Q03JA
İnternet erişimi olan kendinize ait cep telefonu (ör. akıllı telefon).	ST250Q04JA
Ev işlerine yardımcı personel (yarı zamanlı veya tam zamanlı)	ST250D07JA
Evinizde aşağıdakilerden kaç tane bulunmaktadır?	
Küvetli veya duşlu oda	ST251Q03JA
Alafranga tuvaleti olan odalar	ST251Q04JA
Sanat eserleri (ör. tablolar, heykeller)	ST251Q07JA
Evinizde aşağıdaki dijital cihazlardan kaç tane vardır?	
Dizüstü bilgisayar veya notebook	ST254Q03JA
Son 30 gün içinde yiyecek almaya paranız yetmediği için ne sıklıkla aç kaldınız?	
Hiç veya hemen hemen hiç	ST258Q01JA01
Haftada yaklaşık bir kez	ST258Q01JA02
Haftada 2-3 kez	ST258Q01JA03
Haftada 4-5 kez	ST258Q01JA04
Her gün ya da hemen hemen her gün	ST258Q01JA05
Aşağıdaki ölçek, Türkiye'deki toplumsal yapıyı temsil etmektedir. Ölçeğin en tepesinde (10 değeri) ekonomik durumları en iyi olan insanlar bulunmaktadır. Bunlar en fazla para kazanan, en iyi eğitimi alan ve en saygın işlerde çalışan insanlardır. Ölçeğin en altında ise (1 değeri) ekonomik durumları en kötü olan insanlar bulunmaktadır. Bunlar en az para kazanan, eğitim almayan, işleri olmayan veya en az saygı duyulan işlerde çalışan insanlardır.	
Şu anda ailenizin ölçeğin neresinde bulunduğunu düşünüyorsunuz ?	259Q01JA01-259Q01JA010
30 yaşınıza geldiğinizde ölçeğin neresinde olacağınızı düşünüyorsunuz?	ST259Q02JA -ST259Q02JA10
Normal bir okul haftasında, haftanın kaç günü okula gitmeden önce aşağıdakilerin her birini yapıyorsunuz?	
0 Gün-1 Gün-2 Gün-3 Gün-4 Gün- 5 Gün ve Daha Fazla	
Kahvaltı yapma	ST294Q01JA0- ST294Q01JA06
Ev işi yapma veya aile üyelerine bakma	ST294Q03JA1- ST294Q03JA6

Para kazanmak için çalışma	ST294Q04JA01- ST294Q04JA06
Okulunuzun Korona virüs nedeniyle kapalı olduğu süre boyunca derslerle ilgili çalışmalarınız için aşağıdaki dijital cihazlardan en çok hangisini kullandınız?	
Kendi dizüstü bilgisayarımı, masaüstü bilgisayarımı veya tabletimi	ST349Q01JA01
Kendi akıllı telefonumu	ST349Q01JA02
Diğer aile fertleri tarafından da kullanılan bir dijital cihazı	ST349Q01JA03
Okulumun hediye ettiği veya ödünç verdiği bir dijital cihazı	ST349Q01JA04
Derslerle ilgili çalışmalarda kullanacağım bir dijital cihazım yoktu.	ST349Q01JA05
Okulunuzun koronavirüs nedeniyle kapalı olduğu süre boyunca derslerle ilgili çalışmalarınızı tamamlarken aşağıdaki sorunları ne sıklıkla yaşadınız? Hiçbir zaman- Birkaç kez- Haftada yaklaşık bir ya da iki kez- Her gün ya da hemen hemen her gün	
İhtiyacım olduğunda dijital bir cihaza erişimle ilgili sorunlar	ST352Q01JA01- ST352Q01JA04
Evde sorumluluklarım olduğundan çalışmak için zaman bulmakla ilgili sorunlar	ST352Q05JA01- ST352Q05JA04

EK 2: Anne Baba Eğitim Durumu ve Ebeveyn Tutumları Değişkenleri

Değişken Açıklaması	Değişken Adı				
Evinizde aşağıdaki kitap türlerinden kaç tane bulunmaktadır.					
	Yok	1-5 arası	6-10 arası	10'dan fazla	Bilmiyorum
Dini Kitaplar	ST256Q01JA01	ST256Q01JA02	ST256Q01JA03	ST256Q01JA04	ST256Q01JA05
Edebiyat Klasikleri	ST256Q02JA01	ST256Q02JA02	ST256Q02JA03	ST256Q02JA04	ST256Q02JA05
Modern Edebiyat	ST256Q03JA01	ST256Q03JA02	ST256Q03JA03	ST256Q03JA04	ST256Q03JA05
Bilim Kitapları	ST256Q06JA01	ST256Q06JA02	ST256Q06JA03	ST256Q06JA04	ST256Q06JA05
Sanat, müzik veya tasarım konulu kitaplar	ST256Q07JA01	ST256Q07JA02	ST256Q07JA03	ST256Q07JA04	ST256Q07JA05
Teknik içerikli kitaplar	ST256Q08JA01	ST256Q08JA02	ST256Q08JA03	ST256Q08JA04	ST256Q08JA05
Sözlükler	ST256Q09JA01	ST256Q09JA02	ST256Q09JA03	ST256Q09JA04	ST256Q09JA05
Okul Çalışmaları için kaynak kitaplar	ST256Q10JA01	ST256Q10JA02	ST256Q10JA03	ST256Q10JA04	ST256Q10JA05
Anneniz, aşağıdaki eğitim kademelerinden en son hangisini bitirmiştir?					
Lise	ST005Q01JA01				
Ortaokul	ST005Q01JA03				
İlkokul	ST005Q01JA04				
İlkokulu tamamlamadı.	ST005Q01JA05				
Anneniz aşağıdaki diplomalardan herhangi birine sahip midir?					
	Evet		Hayır		
Doktora	ST006C01JA01		ST006C01JA02		
Yüksek Lisans	ST006C02JA01		ST006C02JA02		
Lisans	ST006C03JA01		ST006C03JA02		
Önlisans	ST006C04JA01		ST006C04JA02		
Babanız, aşağıdaki eğitim kademelerinden en son hangisini bitirmiştir?					

Lise	ST007Q01JA01	
Ortaokul	ST007Q01JA03	
İlkokul	ST007Q01JA04	
İlk okulu tamamlamadı	ST007Q01JA05	
Babanız aşağıdaki diplomalardan herhangi birine sahip midir?		
	Evet	Hayır
Doktora	ST008C01JA01	ST008C01JA02
Yüksek Lisans	ST008C02JA01	ST008C02JA02
Lisans	ST008C03JA01	ST008C03JA02
Önlisans	ST008C04JA01	ST008C04JA02
Siz veya evinizdeki bir başkası aşağıdakileri ne sıklıkla yapıyorsunuz?		
Çocuğumun okuldaki başarısını tartışma	PA003Q01TA	
Akşam yemeğini toplu olarak bir masada yemek.	PA003Q02TA	
Çocuğumla konuşarak zaman geçirmek	PA003Q03TA	
Çocuğumla sosyal ve politik konuları tartışırım	PA003Q05IA	
Çocuğuma matematik ödevlerine yardım ediyorum	PA003Q18WA	
Çocuğuma okuldaki diğer öğrencilerle arasının nasıl olduğunu sorma	PA003Q13JA	
Çocuğumun okulda öğrendikleriyle ilgilenmek	PA003Q15JA	
Çocuğumla gelecekteki eğitimi hakkında konuşmak	PA003Q16JA	
Çocuğuma o gün okulda ne yaptığını sormak	PA003Q17JA	
Ailenizde siz dahil matematikle ilgili bir kariyerde çalışan var mı	PA197Q01WA	
Geçen akademik yıl boyunca okulla ilgili aşağıdaki faaliyetlerden birine katıldınız mı?		
Çocuğumun davranışlarını kendi inisiyatifimle öğretmeniyile tartıştım	PA008Q01TA	
Çocuğumun gelişimini kendi inisiyatifimle bir öğretmenle tartıştım.	PA008Q03TA	
Çocuğumun öğretmenleriyle evde ödevlere nasıl yardım edebileceğimi konuştum	PA008Q09NA	
Çocuğumun öğretmenleriyle ebeveynlik, aile desteği ve çocuğumun gelişimi ile ilgili fikir alışverişinde bulundum.	PA008Q10NA	
Geçen akademik yıl boyunca çocuğunuzun okulundaki etkinliklere katılımınız aşağıdaki nedenlerden biri yüzünden engellendi mi?		
Katılmamın çocuğumun gelişimi açısından önemli olmadığını düşünüyorum	PA009Q09NA	
Çocuğum katılmamı istemiyor.	PA009Q11NA	

Evde konuştuđunuz dil(ler) de dahil olmak üzere kaç dili başkalarıyla konuşabilecek kadar iyi konuşuyorsunuz?	PA166Q01HA
---	------------



EK 3: Matematik okuryazarlığı yeterlik düzeylerinin özeti (MEB,2024)

Düzy	Alt Puan Limiti	Yeterlik Düzeyinde Bulunan Öğrencilerin Davranışları
6	669	Bu düzeydeki öğrenciler; soyut problemler üzerinde çalışabilir, çözüm geliştirmek için yaratıcılık ve esnek düşünme becerisi sergileyebilir. Örneğin verilen görevde belirtilmeyen bir işlemin alışılmışın dışında bir bağlamda ne zaman kullanılabileceğini veya matematikle ilgili bir kavramın tam olarak anlaşılmasının yapacakları gerekçelendirmenin bir unsuru olarak ne zaman gerekli olduğunu fark edebilir. Çözümlerinin bir bölümü olarak simülasyonları veya elektronik tabloları etkin bir şekilde kullanmak da dâhil olmak üzere farklı bilgi kaynakları ve bilgi gösterimleri arasında bağlantı kurabilir. Bu düzeydeki öğrenciler eleştirel düşünme yeteneğine sahiptir ve akıl yürütme becerilerini açık bir şekilde ifade etmek için kullandıkları simgesel ve biçimsel matematik işlemleri ve ilişkileri konusunda yetkinlik sahibidir. Bu öğrenciler, buldukları çözüm ve başlangıçtaki durumla ilgili olarak yaptıklarının uygunluğu üzerine derinlemesine düşünebilir.
5	607	Bu düzeydeki öğrenciler, var olan kısıtlamaları belirleyerek veya yeni kısıtlamalar getirerek ve varsayımları tanımlayarak karmaşık durumlar için modeller geliştirebilir ve bu modellerle çalışabilir. Daha zorlu görevlerle başa çıkmak için sistematik, iyi planlanmış problem çözme stratejileri uygulayabilir. Örneğin bir deneyin nasıl geliştirileceğine karar verebilir, en uygun işlemi kurgulayabilir veya görevde verilmeyen daha karmaşık görsellerle çalışabilir. Bu öğrenciler, çözümleri genellikle görevde açıkça belirtilmeyen matematiksel bilgileri bir araya getirmeyi gerektiren problemleri çözme konusunda ileri düzeyde beceri gösterir. Bu düzeydeki öğrenciler çalışmaları üzerinde derinlemesine düşünür ve matematiksel sonuçları gerçek dünya bağlamında değerlendirir.
4	545	Bu düzeydeki öğrenciler, bazen iki değişken içeren karmaşık somut durumlarda açıkça tanımlanmış modellerle etkili bir şekilde çalışabilir ve daha karmaşık bir bilgi işlemsel düşünme yaklaşımı kullanarak oluşturdukları tanımlanmamış modellerle çalışma becerisi gösterebilir. Bu öğrenciler, verilen bilgilerden hesaplama yapmak mümkün olmadığında nitel yargılarda bulunarak bir sonucun akla uygunluğunu değerlendirmek gibi eleştirel düşünme boyutlarıyla meşgul olmaya başlar. Simgesel veya grafiksel gösterimler de dâhil olmak üzere farklı bilgi gösterimlerini seçip bütünleştirebilir ve bunları gerçek yaşam durumlarının özellikleriyle doğrudan ilişkilendirebilir. Bu düzeyde öğrenciler yorumlarına, akıl yürütmelerine ve metodolojilerine dayalı açıklamalar ve argümanlar oluşturabilir ve bunları sunabilir.
3	482	Bu düzeydeki öğrenciler, bilindik kavramların anlaşılmasında aşamalı karar verme becerisi veya esneklik gerektiren stratejiler de dâhil olmak üzere çözüm stratejileri geliştirebilir. Bu öğrenciler, çözüm stratejilerini geliştirmek için bilgi işlemsel düşünme becerilerini kullanmaya başlar. Tamamı problem cümlesinde açıkça belirtilmeyen farklı ancak standart hesaplamaların yapılmasını gerektiren birkaç görevi çözebilir. Uzamsal görselleştirmeyi bir çözüm stratejisinin bileşeni olarak kullanabilir veya göreve uygun verileri toplamak için bir simülasyonun nasıl kullanılacağını belirleyebilir. Bu düzeydeki öğrenciler; farklı bilgi kaynaklarına dayalı gösterimleri yorumlayabilir ve kullanabilir, iki yönlü bir tablo kullanarak koşullu karar verme de dâhil olmak üzere doğrudan bu gösterimlerden yola çıkarak akıl yürütebilir. Genellikle yüzdeleri, kesirleri ve ondalık sayıları kullanma ve orantısal ilişkilerle çalışma becerisi sergiler.
2	420	Bu düzeydeki öğrenciler, çözüm stratejilerinin bir bileşeni olarak tek değişken içeren basit simülasyonlar kullanmak da dâhil olmak üzere problemleri çözmek için basit stratejiler tasarlamaları gereken durumları fark edebilir. İki yönlü tablolar, grafikler veya üç boyutlu nesnelerin iki boyutlu gösterimleri gibi biraz daha karmaşık gösterim biçimlerini kullanan bir veya daha fazla kaynaktan ilgili bilgileri çıkarabilir. Bu öğrenciler, fonksiyonel ilişkiler konusunda temel bir anlayışa sahiptir ve basit oranlar içeren problemleri çözebilir. Bu öğrencilerin sonuçlar hakkında gerçekçi yorumlar yapabilme becerileri vardır.
1a	358	Bu düzeydeki öğrenciler, gerekli tüm bilgilerin mevcut olduğu ve soruların açıkça tanımlandığı basit bağlamları içeren soruları yanıtlayabilir. Bilgiler çeşitli basit

		formatlarda sunulabilir ve öğrencilerin ilgili bilgileri bulup çıkarmak için aynı anda iki kaynakla çalışması gerekebilir. Bazen bir problemi çözmek için rutin bir işlemin birden fazla tekrarlanması gerektirebilen açık durumlarda direkt olarak verilen yönergelere göre basit, rutin işlemleri gerçekleştirebilir. Öğrenciler anlaşılır olan veya çok az bilgi sentezi gerektiren eylemleri gerçekleştirebilir ancak tüm durumlarda bu eylemler verilen uyarıcıları net bir şekilde takip eder. Bu düzeydeki öğrenciler, çoğunlukla tam sayıları içeren problemleri çözmek için temel algoritmaları, formülleri, işlemleri veya kuralları kullanabilir.
1b	295	Bu düzeydeki öğrenciler; ihtiyaç duyulan tüm bilgilerin basit bir gösterimle (ör. tablo veya grafik) açıkça verildiği, anlaşılması kolay bağlamları içeren sorulara yanıt verebilir, gerektiğinde bazı bilgilerin gereksiz olduğunu ve sorulan belirli bir soru açısından göz ardı edilebileceğini fark edebilir. Kısa, söz dizimsel olarak basit bir metinle tanımlanmış, açıkça belirtilmiş yönergeleri takip ederek tam sayılarla basit hesaplamalar yapabilir.
1c	233	Bu düzeydeki öğrenciler; ilgili tüm bilgilerin basit, tanıdık bir formatta (ör. küçük bir tablo veya resim) açıkça verildiği ve çok kısa, söz dizimsel olarak basit bir metinde tanımlandığı, anlaşılması kolay bağlamları içeren sorulara yanıt verebilir. Tek bir adımı veya işlemi açıklayan net bir yönergeyi takip edebilir.

EK 4: Okuma becerileri yeterlik düzeylerinin özeti (MEB,2024)

Düzyey	Alt Puan Limiti	Yeterlik Düzeyinde Bulunan Öğrencilerin Davranışları
6	698	Bu düzeydeki öğrenciler, istenilen bilginin metin içerisinde saklı olduğu uzun ve soyut metinleri anlayabilir. Bilginin nasıl kullanılacağına karar vermek için çeşitli ölçütler kullanabilir, bilgilerin benzer ve zıt yönlerini karşılaştırabilir ve bu bilgileri bir araya getirebilir. Dış ölçütler kullanarak metnin kaynağı hakkında derinlemesine düşünebilir. Bilginin kaynağı ve geçerliğiyle ilgili ipuçları aracılığıyla metinler arasındaki uyumsuzlukları belirleyebilir. Metinler arası tutarsızlıkları çözebilir, farklı metinlerdeki bilgilerin benzer ve zıt yönlerini karşılaştırabilir.
5	626	Bu düzeydeki öğrenciler, uzun metinlerde konuyla ilgili saklı bilgileri bularak bu metinleri kavrayabilir. Kapsamlı metinlere yönelik derin bir anlayış göstererek farklı akıl yürütme yöntemlerini kullanabilir. Çeşitli metin veya kaynaklarda yer alan bilgiler ile soru arasında ilişkiyi kurarak soruları cevaplayabilir. Belli bilgilerden yola çıkarak hipotezler oluşturabilir veya var olan hipotezlere ilişkin değerlendirme yapabilir. Karışık ve soyut ifadelerin olduğu durumlarda gerçek ile algı arasındaki ayrımı yapabilir. Çeşitli ipuçlarından yola çıkarak verilen bilginin kaynağının veya içeriğinin objektifliğini değerlendirebilir.
4	553	Bu düzeydeki öğrenciler, bir veya birden çok metinde yer alan uzun paragrafları anlayabilir. Metni bir bütün olarak ele alarak dile dayalı farklılıkları yorumlayabilir. Çeşitli kaynaklardan yola çıkarak farklı bakış açılarını karşılaştırabilir ve sonuç çıkarabilir. Metne yerleştirilmiş bilgileri arayabilir, bu bilgilere ulaşabilir ve bir araya getirebilir. Bilginin uygunluğunu değerlendirmek için çıkarımlar yapabilir. Metnin dikkat çekici özelliklerinden yola çıkarak yazarların kendi fikirlerini aktarmak için kullandıkları ifadeleri belirleyebilir. Çeşitli metinlerde açıkça ifade edilen iddiaları kıyaslayabilir ve kriterlere göre bilgi kaynağının güvenilirliğini değerlendirebilir.
3	480	Bu düzeydeki öğrenciler, açık şekilde sunulmadığı durumlarda da metnin genel anlamını ifade edilebilir. Bilgileri derleyebilir, basit ve ileri düzeyde çıkarımlar yapabilir. Çok uzun olmayan metinlerin ana düşüncesini belirlemek, ilişkileri anlamak, bir kelime veya ifadenin anlamını çıkarmak için metnin farklı bölümlerindeki bilgileri bir araya getirebilir. Çeşitli kriterlere göre elde edilen bilgiler arasındaki ilişkileri anlayabilir. Bu düzeydeki öğrenciler, bir veya birkaç metin üzerine derinlemesine düşünebilir. Sunulan bilgilerden yola çıkarak farklı yazarların bakış açılarını kıyaslayabilir.
2	407	Bu düzeydeki öğrenciler, orta uzunluktaki metinlerin ana düşüncesini belirleyebilir. Gerekli bilgilerin açıkça verilmediği durumlarda metnin belirli bir bölümündeki ilişkileri anlayabilir ve bir bölümünden anlam çıkarabilir. Birkaç sayfalık metin içerisinde konu ile ilgili sayfayı bulabilir ve konuyla ilgili düzeylerini artırabilir. Gerekli bilgilerin açıkça verildiği durumlarda orta uzunluktaki metinlerin genel amacı ve bu metinlerdeki belirli detaylar üzerine derinlemesine düşünebilir. İddiaları karşılaştırabilir ve bu iddiaları destekleyen düşünceleri belirleyebilir.
1a	335	Bu düzeydeki öğrenciler, cümlelerin veya kısa paragrafların gerçek anlamını kavrayabilir. Aşına oldukları konularda yazılmış metinlerin ana fikrini ve yazarın amacını anlayabilir. Metinde verilen bilgiler arasında veya metinde verilen bilgiler ile kendi sahip olduğu bilgiler arasında basit bağlantılar kurabilir. Birkaç sayfalık metinler içinde ilgili sayfayı bulabilir ve kısa metinlerde yer alan bilgilere ulaşabilir. Bilgilerin açıkça sunulması durumunda metinlerin genel amacı ve metindeki nispeten önemli bilgiler üzerinde (metnin ana fikri gibi) derinlemesine düşünebilir.
1b	262	Bu düzeydeki öğrenciler, basit cümlelerin gerçek anlamlarını değerlendirebilir. Soru ve/veya metindeki bilgiler arasında basit bağlantılar kurarak metinlerin gerçek anlamını yorumlayabilir. Tek bir cümle, kısa bir metin veya basit bir listede istenen bilgileri tarayabilir ve bulabilir. Açık bir şekilde istendiğinde birkaç sayfalık metin içerisinde ilgili sayfayı bulabilir.
1c	189	Bu düzeydeki öğrenciler, kısa ve basit cümlelerin anlamını kavrayabilir. Sınırlı bir süre içinde açık, basit ve somut amaçlar için okuma yapabilir.

EK 5: ST250D07JA değişkenine ait veriler

0360001 = "Australia :A home theatre-Yes"	0360002 = "Australia :A home theatre-No"
0760001 = "Brazil :A desk to study at-Yes"	0760002 = "Brazil :A desk to study at- No"
0960001 = "Brunei Darussalam :Video or online games (e.g. used with game consoles such as a PlayStation 4®)-Yes"	0960002 = "Brunei Darussalam :Video or online games (e.g. used with game consoles such as a PlayStation 4®)- No"
1000001 = "Bulgaria :air conditioner-Yes"	1000002 = "Bulgaria :air conditioner- No"
1160001 = "Cambodia :Which of the following are in your [home]: A dictionary-Yes"	1160002 = "Cambodia :Which of the following are in your [home]: A dictionary- No "
1240001 = "Canada :Your own sports equipment-Yes"	1240002 = "Canada :Your own sports equipment- No"
1520001 = "Chile :Scanner-Yes"	1520002 = "Chile :Scanner- No"
1700001 = "Colombia :fridge-Yes"	1700002 = "Colombia :fridge- No"
1960001 = "Cyprus :A home security alarm system-Yes"	1960002 = "Cyprus :A home security alarm system- No"
2080001 = "Denmark :Your own headphones-Yes"	2080002 = "Denmark :Your own headphones- No "
2140001 = "Dominican Republic :your own car-Yes"	2140002 = "Dominican Republic :your own car- No "
2220001 = "El Salvador :Microwave-Yes"	2220002 = "El Salvador :Microwave- No"
2330001 = "Estonia :Your own table for studying-Yes"	2330002 = "Estonia :Your own table for studying- No"
2460001 = "Finland :garage or car port-Yes"	2460002 = "Finland :garage or car port- No"
2500001 = "France :An action camera (e.g. GoPro)-Yes"	2500002 = "France :An action camera (e.g. GoPro)- No"
2750001 = "Palestinian Authority :digital books-Yes"	2750002 = "Palestinian Authority :digital books- No"
2760001 = "Germany :A quiet place to study-Yes"	2760002 = "Germany :A quiet place to study- No"
3000001 = "Greece :Digital games (e.g Playstation4®)-Yes"	3000002 = "Greece :Digital games (e.g Playstation4®)- No"
3200001 = "Guatemala :Which of the following are in your [home]: Blender-Yes"	3200002 = "Guatemala :Which of the following are in your [home]: Blender- No"
3440001 = "Hong Kong (China) :Newspaper or Educational Magazine (e.g. National Geographic Magazine)-Yes"	3440002 = "Hong Kong (China) :Newspaper or Educational Magazine (e.g. National Geographic Magazine)- No"
3480001 = "Hungary :Dishwasher in kitchen-Yes"	3480002 = "Hungary :Dishwasher in kitchen- No"
3600001 = "Indonesia :Your own Tablet/iPad-Yes"	3600002 = "Indonesia :Your own Tablet/iPad- No"
3760001 = "Israel :Membership to the theater, gym, or swimming pool-Yes"	3760002 = "Israel :Membership to the theater, gym, or swimming pool- No"
3880001 = "Jamaica :Portable Wi-fi-Yes"	3880002 = "Jamaica :Portable Wi-fi- No"
3920001 = "Japan :passport-Yes"	3920002 = "Japan :passport- No"

3980001 = "Kazakhstan :Digital photo camera-Yes"	39800012= "Kazakhstan :Digital photo camera- No"
4000001 = "Jordan :digital books-Yes"	4000002 = "Jordan :digital books- No"
4100001 = "Korea :Massage chair-Yes"	4100002 = "Korea :Massage chair- No"
4280001 = "Latvia :Scooter-Yes"	4280002 = "Latvia :Scooter- No"
4460001 = "Macao (China) :Electric massage chair-Yes"	4460002 = "Macao (China) :Electric massage chair- No"
4580001 = "Malaysia :Refrigerator-Yes"	4580002 = "Malaysia :Refrigerator- No"
4840001 = "Mexico :A quiet place to study-Yes"	4840002 = "Mexico :A quiet place to study- No"
4960001 = "Mongolia :Wall painting-Yes"	4960002 = "Mongolia :Wall painting- No"
5040001 = "Morocco :Electric water heater tank-Yes"	5040002 = "Morocco :Electric water heater tank- No"
5280001 = "Netherlands :Your own tablet (e.g. iPad®, Samsung Galaxy®)-Yes"	5280002 = "Netherlands :Your own tablet (e.g. iPad®, Samsung Galaxy®)- No"
5540001 = "New Zealand :Your own musical instrument (e.g. guitar, keyboard)-Yes"	5540002 = "New Zealand :Your own musical instrument (e.g. guitar, keyboard)- No"
5910001 = "Panama :Video games-Yes"	5910002 = "Panama :Video games- No"
6040001 = "Peru :Bike-Yes"	6040002 = "Peru :Bike- No"
6080001 = "Philippines :Air-conditioning unit-Yes"	6080002 = "Philippines :Air-conditioning unit- No"
6340001 = "Qatar :Cinema-Yes"	6340002 = "Qatar :Cinema- No"
6420001 = "Romania :Cable/Satellite TV-Yes"	6420002 = "Romania :Cable/Satellite TV- No"
7020001 = "Singapore :Domestic helper (e.g. full/part-time maid)-Yes"	7020002 = "Singapore :Domestic helper (e.g. full/part-time maid)- No"
7040001 = "Viet Nam :Which of the following are in your [home]: a set of chair and table for learning-Yes"	7040002 = "Viet Nam :Which of the following are in your [home]: a set of chair and table for learning- No"
7050001 = "Slovenia :Sports equipment (e.g. skis, bike, tennis racket, etc.)-Yes"	7050002 = "Slovenia :Sports equipment (e.g. skis, bike, tennis racket, etc.)- No"
7240001 = "Spain :pay TV (Movistar+, Orange TV)-Yes"	7240002= "Spain :pay TV (Movistar+, Orange TV)- No"
7640001 = "Thailand :Air Purifier-Yes"	7640002 = "Thailand :Air Purifier- No"
7840001 = "United Arab Emirates :Cinema-Yes"	7840002 = "United Arab Emirates :Cinema- No"
7920001 = "Türkiye :Helper for houseworks (part time or full time)-Yes"	7920002 = "Türkiye :Helper for houseworks (part time or full time)- No"
8040001 = "Ukraine :Reference books-Yes"	8040002 = "Ukraine :Reference books- No"
8270001 = "United Kingdom (Excl. Scotland) :A smart speaker such as Amazon Echo or Google Home-Yes"	8270002 = "United Kingdom (Excl. Scotland) :A smart speaker such as Amazon Echo or Google Home- No"
8280001 = "United Kingdom (Scotland) :Your own smartwatch-Yes"	8280002 = "United Kingdom (Scotland) :Your own smartwatch- No"

8600001 = "Uzbekistan:Book-shelf-Yes"	8600002= "Uzbekistan :Book-shelf- No"
---------------------------------------	---------------------------------------

