



**T.C.
BURSA ULUDAĞ ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
EKONOMETRİ ANABİLİM DALI
İSTATİSTİK BİLİM DALI**

**YAPAY ZEKÂ YÖNTEMLERİYLE UYUŞMAZLIK MAHKEMESİ
KARARLARININ TAHMİNİ**

(DOKTORA TEZİ)

Muhammed Burak GÖRENTAŞ

BURSA – 2023



**T.C.
BURSA ULUDAĞ ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
EKONOMETRİ ANABİLİM DALI
İSTATİSTİK BİLİM DALI**

**YAPAY ZEKÂ YÖNTEMLERİYLE UYUŞMAZLIK MAHKEMESİ
KARARLARININ TAHMİNİ**

(DOKTORA TEZİ)

Muhammed Burak GÖRENTAŞ

**Danışman:
Nuran BAYRAM ARLI**

BURSA – 2023

ÖZET

Yazar adı soyadı	Muhammed Burak GÖRENTAŞ
Üniversite	Bursa Uludağ Üniversitesi
Enstitü	Sosyal Bilimler Enstitüsü
Anabilim dalı	Ekonometri
Bilim dalı	İstatistik
Tezin niteliği	Doktora
Mezuniyet tarihi	/...../20....
Tez danışmanı	Prof. Dr. Nuran BAYRAM ARLI

YAPAY ZEKÂ YÖNTEMLERİYLE UYUŞMAZLIK MAHKEMESİ KARARLARININ TAHMİNİ

Yapay zekâ, hayatı pek çok açıdan kolaylaştıran önemli bir teknoloji olarak öne çıkmaktadır. Günlük yaşamımızda akıllı asistanlar aracılığıyla ev işlerini yönetmek, seyahat planları yapmak ve hatta öğrenme deneyimlerini kişiselleştirmek gibi işleri kolaylaştırır. Otomasyon ve üretim süreçlerinde kullanıldığında, iş yerlerinde verimliliği artırır ve insanların daha yaratıcı ve entelektüel görevlere odaklanmalarına imkân tanır. Yapay zekâ, karmaşıklığı basite indirgeyerek hayatımızın pek çok alanında kolaylık sağlamaktadır. Yapay zekâ, hukuk alanında önemli bir dönüşümü tetiklemektedir. Hukuk pratiği, büyük miktarda belge ve dava incelemesi gerektiren bir alandır ve yapay zekâ, bu süreçleri hızlandırmak ve daha verimli hale getirmek için kullanılmaktadır. Özellikle belge analizi ve veri madenciliği, hukuk firmalarının ve mahkemelerin büyük miktardaki hukuki belgeleri daha hızlı ve doğru bir şekilde tarayarak ilgili bilgilere ulaşmalarına yardımcı olmaktadır. Aynı zamanda, yapay zekâ, hukuk alanındaki tahminlerde ve analitik çalışmalarda da kullanılmakta, bu da hukukun daha öngörülebilir hale gelmesine katkıda bulunmaktadır. Bu çalışma kapsamında yapay zekânın hukukta kullanımına dair bir örnek olarak Uyuşmazlık Mahkemesi kararları denetimli makine öğrenmesi yöntemleri ile tahmin edilmeye çalışılmıştır. Destek Vektör Makineleri algoritmasının %87 doğruluk değeri ile en yüksek sonucu verdiği görülmüştür. Ortaya konulan model, sanal yargıç türü yazılımların bir örneği olması nedeniyle bu çalışmada, hukukta sanal yargıç uygulamasının olumlu ve olumsuz tarafları ile hukuk alanında yapay zekâ etiği konuları tartışılmıştır.

Anahtar kelimeler: Hukuk, Makine Öğrenmesi, Sanal Yargıç, Yapay Zekâ

ABSTRACT

Name & surname	Muhammed Burak GÖRENTAŞ
University	Bursa Uludağ University
Institute	Institute of Social Sciences
Field	Econometrics
Subfield	Statistic
Degree awarded	PhD.
Date of degree awarded	/...../20....
Supervisor	Prof. Dr. Nuran BAYRAM ARLI

ESTIMATION OF THE COURT OF JURISDICTIONAL DISPUTES DECISIONS WITH ARTIFICIAL INTELLIGENCE METHODS

Artificial intelligence stands out as a significant technology that facilitates life in many ways. It simplifies various aspects of daily life, allowing us to manage household chores, plan travel, and even personalize learning experiences through smart assistants. When used in automation and production processes, it enhances workplace productivity and enables individuals to focus on more creative and intellectual tasks. Artificial intelligence simplifies complexity across multiple domains in our lives.

Furthermore, artificial intelligence is triggering a substantial transformation in the field of law. Legal practice involves the examination of extensive documents and cases, and artificial intelligence is utilized to expedite and improve efficiency in these processes. Particularly, document analysis and data mining assist law firms and courts in rapidly and accurately scanning large volumes of legal documents to access relevant information. Additionally, artificial intelligence is employed in making predictions and analytical studies within the legal field, contributing to making law more predictable.

As part of this study, the use of artificial intelligence in law is exemplified by attempting to predict decisions of a The Court Of Jurisdictional Disputes using supervised machine learning methods. It was observed that the Support Vector Machines algorithm yielded the highest accuracy rate, with an accuracy rate of 87%. Since the model presented in this study is an example of software similar to artificial judges, both the positive and negative aspects of artificial judges in law and ethical considerations regarding artificial intelligence in law are discussed.

Keywords: Artificial Intelligence, Artificial Judge, Law, Machine Learning,

ÖNSÖZ

Doktora tez çalışmamı yürütürken yanımda olan ve desteklerini hissettiğim birçok kişiye teşekkür borçluyum. Öncelikle hem akademiye hem hayata dair çok kıymetli rehberliği ve yönlendirmesiyle bana her zaman destek olan danışman hocam Prof. Dr. Nuran Bayram Arlı'ya sonsuz teşekkür ediyorum. Bu süreci tamamlayabilmemdeki sabrı, motivasyonu ve teşvikleri için de ayrıca çok minnettarım.

Özellikle doğal dil işleme süreçleri, makine öğrenmesi kodları olmak üzere tezimi hazırlamamda desteğini her zaman hissettiğim ikinci danışmanım Dr. Öğr. Üyesi Taner Uçkan'a çok müteşekkirim. Tez izleme komitemde yer alan Prof. Dr. Sevda Gürsakal ve Doç. Dr. Gülşen Gedik hocalarıma pozitif ve destekleyici yaklaşımları ve katkıları için teşekkürlerimi sunarım. Ayrıca tez çalışmamla ilgili değerli yorumları için Prof. Dr. Serdar Gülenler ve Doç. Dr. Elif Kartal hocalarıma da teşekkür ederim. Son olarak, fedakârlığı ve duygusal desteği için sevgili eşim Duygu Görentaş'a, duaları, destekleri ve teşvikleri ile her zaman yanımda olan arkadaşlarım ve tüm aileme en derin şükranlarımı sunarım.

Muhammed Burak GÖRENTAŞ

Bursa-2023

İÇİNDEKİLER

ÖZET.....	i
ABSTRACT.....	ii
ÖNSÖZ	iii
İÇİNDEKİLER	iv
TABLOLAR.....	vi
ŞEKİLLER.....	vii
KISALTMALAR	ix
GİRİŞ	1
1. BİRİNCİ BÖLÜM	12
YAPAY ZEKÂ VE DOĞAL DİL İŞLEME.....	12
1.1. Yapay Zekâ	13
1.1.1. Yapay Zekânın Hukuk Alanındaki Uygulamaları	26
1.2. Doğal Dil İşleme	32
1.2.1. Metin Sınıflandırma	34
1.3. Mahkeme Kararlarında Metin Sınıflandırma	41
2. İKİNCİ BÖLÜM.....	45
MAKİNE ÖĞRENMESİ VE DENETİMLİ MAKİNE ÖĞRENMESİ	45
2.1. Denetimli Makine Öğrenmesi	50
2.1.1. Denetimli Makine Öğrenmesi Aşamaları	51
2.2. Sınıflandırma Teknikleri.....	64
2.2.1. Karar Ağaçları	64
2.2.2. Rassal Orman	67
2.2.3. Destek Vektör Makineleri	70
2.2.4. Lojistik Regresyon	72
2.2.5. Diskriminant Analizi	74
2.2.6. Yapay Sinir Ağları	75
2.2.7. Naive Bayes	78
2.2.8. k-En Yakın Komşu	80
3. ÜÇÜNCÜ BÖLÜM	83

UYUŞMAZLIK MAHKEMESİ KARARLARININ YAPAY ZEKÂ YÖNTEMLERİ İLE TAHMİNİ.....	83
3.1. Uyuşmazlık Mahkemesinin Türk Yargı Sistemindeki Yeri ve Uyuşmazlık Mahkemesi Kararlarının Niteliği.....	85
3.2. Uyuşmazlık Mahkemesi Kararlarının Yapay Zekâ Yöntemleri İle Tahmini	88
3.2.1. Veri Toplama Yöntemi	88
3.2.2. Yapay Zekâ Yöntemleri ile Tahmin.....	94
3.3. Tartışma ve Öneriler	101
SONUÇ	107
KAYNAKLAR.....	111
EKLER.....	126
ÖZGEÇMİŞ	152

TABLÖLAR

Tablo 1: Terim örnekleri ve ikili gösterimde belgelere ilişkin vektörler(Tantuğ, 2016)	37
Tablo 2: Mahkeme Kararlarında Metin Sınıflandırma Çalışmaları.....	42
Tablo 3: Karmaşıklık Matrisi (Han vd., 2012)	59
Tablo 4: Türkiye’de Yargı Kolları (Gözler, 2019).	84
Tablo 5 : Ham Veri ve İşlenmiş Veri	92
Tablo 6: Karar Ağaçları Algoritması Değerlendirme Metrikleri	95
Tablo 7: Rassal Orman Algoritması Değerlendirme Metrikleri	96
Tablo 8: Lojistik Regresyon Algoritması Değerlendirme Metrikleri	97
Tablo 9: Destek Vektör Makineleri Algoritması Değerlendirme Metrikleri.....	97
Tablo 10: Diskriminant Analizi Algoritması Değerlendirme Metrikleri.....	98
Tablo 11: Yapay Sinir Ağları Algoritması Değerlendirme Metrikleri	99
Tablo 12: k-en Yakın Komşu Algoritması Değerlendirme Metrikleri.....	99
Tablo 13: Naive Bayes Algoritması Değerlendirme Metrikleri	100
Tablo 14: Sınıflandırma Teknikleri Toplam Değerlendirme	101
Tablo 15: Önerilen çalışma ile yapılan çalışmaların karşılaştırılması.....	102

ŞEKİLLER

Şekil 1: Yapay Zekâ Hedefler Piramidi (Tegmark, 2017).	17
Şekil 2: Metin Sınıflama Süreci.....	36
Şekil 3: Vektör Uzayı Modeli	38
Şekil 4: CBoW ve SG mimarileri (Acı & Çırak, 2019).....	39
Şekil 5: Makine Öğrenmesi Algoritmaları Dağılım Tablosu (Bayram Arlı, Gürsakal, vd., 2022).	47
Şekil 6: Denetimli Makine Öğrenmesi İş Akış Şeması (Mahesh, 2020).....	48
Şekil 7: Denetimsiz Makine Öğrenmesi İş Akış Şeması (Mahesh, 2020).....	48
Şekil 8: CRISP_DM Sürecinin Aşamaları (Shearer, 2000).	52
Şekil 9: CRISP-DM Referans Modelinin Görevleri ve Çıktıları.....	53
Şekil 10 :Yüksek Ayırt Etme Yeteneğine Sahip ROC Eğrisi (Fan vd., 2006)	63
Şekil 11: Bir Veri seti ve İlgili Karar Ağacı Örneği (Alpaydın, 2014).	66
Şekil 12: Rassal Orman Algoritması (Bayram-Arlı vd., 2022)	69
Şekil 13: Optimal Hiperdüzlemin Seçimi (Kartal, 2022)	71
Şekil 14 : Sigmoid Fonksiyonu (Mugunthan & Vijayakumar, 2021).....	72
Şekil 15: Lojistik Regresyon Eğrisi (Gürsakal vd., 2022).....	74
Şekil 16: Diskriminant Analizi Öncesi ve Sonrası (Sakarkar vd., 2021).....	75
Şekil 17: Biyolojik Sinir Hücresinin Yapısı (Özen, 2022).	76
Şekil 18: Yapay Sinir Hücresinin Yapısı (Özen, 2022).	77
Şekil 19: İleri Beslemeli Yapay Sinir Ağları Mimarisi	77
Şekil 20: Geri Beslemeli Yapay Sinir Ağları Mimarisi.....	78
Şekil 21: K = 1-3-5 için Komşuluk Hesabı (Bilgin & Eser, 2022b).....	81
Şekil 22 : UYM Karar Örneği	87
Şekil 23: UYM Kararları Kamuya Açık Portal Sayfası.....	89
Şekil 24: UYM Kamuya Açık Platformunda Seçim ve Yazım Kutucukları.....	90
Şekil 25: UYM Kamuya Açık Platformunda Koşullar	90
Şekil 26: Çekilen Kararlarla İlgili Excel Tablosu	91
Şekil 27 : Öznitelik ve Hedef Değişkenlerin Python Çıktısı	93
Şekil 28: Metin sınıflandırma akış şeması.....	93

Şekil 29: Veri setinin dağılımı	94
Şekil 30: Adli ve İdari Kararlardaki Kelime Bulutları.....	94
Şekil 31: Karar Ağaçları Algoritması Karmaşıklık Matrisi	95
Şekil 32: Rassal Orman Algoritması karmaşıklık matrisi	96
Şekil 33: Lojistik Regresyon Algoritması Karmaşıklık Matrisi	96
Şekil 34: Destek Vektör Makineleri Algoritması Karmaşıklık Matrisi	97
Şekil 35: Diskriminant Analizi Algoritması Karmaşıklık Matrisi	98
Şekil 36: Yapay Sinir Ağları Algoritması Karmaşıklık Matrisi	98
Şekil 37: k-en Yakın Komşu Algoritması Karmaşıklık Matrisi	99
Şekil 38: Naive Bayes Algoritması Karmaşıklık Matrisi	100

KISALTMALAR

AB	: Avrupa Birliđi
ABD	: Amerika Birleşik Devletleri
AI HLEG	: High-Level Expert Group on AI
AİHM	: Avrupa İnsan Hakları Mahkemesi
BoW	: Bag Of Words
CART	: Classification and Regression Trees
CBoW	: Continuous bag of words
CEPEJ	: European Commission for the Efficiency of Justice
CHAID	: Chi-squared Automatic Interaction Detector
CNN	: Convolutional Neural Networks
CRISP-DM	: Cross-Industry Standard Process for Data Mining
DDPG	: Derin Deterministik Politika Gradyanları
DVM	: Destek Vektör Makinesi
EESC	: AB Ekonomik ve Sosyal Komitesi
FN	: False Negatives
FP	: False Positives
FPR	: False Positive Rate
GloVe	: Global Vectors
GPU	: Ekran Kartı
IB	: in-bag
IDE	: Integrated Development Environment
IoT	: Nesnelerin İnterneti
k-NN	: k-Nearest Neighbor
LRC	: Legal Reading Comprehension
NB	: Naive Bayes
OOB	: Out Of Bag
PPO	: Proximal Policy Optimization
RAN	: Recursive Attention Network
RCNN	: Region Based Convolutional Neural Network

RNN	: Recurrent Neural Network
ROC	: Receiver Operating Characteristic Curve
SG	: Skip-Gram
SRM	: Structural Risk Minimization
TF-IDF	: Term Frequency – İnverse Document Frequency
TN	: True Negatives
TP	: True Positives
TPR	: True Positive Rate
TS-BST	: Terim Sıklığı – Belge Sıklığının Tersİ
UYM	: Uyuşmazlık Mahkemesi
YSA	: Yapay Sinir Ağları

GİRİŞ

Dijital ortamlarda istatistiksel verilerin işlenmesi ve büyük veri setlerinin analiz edilmesi büyük veri ve veri analitiği kavramlarının oluşmasına yol açmıştır. Son yıllarda yaşanan yapay zekâ ve özellikle makine öğrenmesi teknolojilerindeki gelişmeler, metin verilerinin sayısal veriler gibi işlenebilir hale gelmesini mümkün kılmıştır. Bu gelişmeler, özellikle doğal dil işleme adı verilen bir alt dalda büyük bir etki yaratmıştır.

Hukuk da artık veri analitiği ve makine öğrenmesi alanına dahil edilmektedir. Hukuk alanındaki sayısal araştırmalar, bu disiplinin veri analitiği teknikleriyle birleşmesini ve geleneksel hukuki yöntemlere yeni bir boyut eklemesini sağlamaktadır (Dothan, 2020). Hukuk alanındaki her türlü mevzuat, akademik metinler, dilekçeler ve yargı kararları gibi hukuki belgeler üzerinde yapılan metin analizine hukuk analitiği denmektedir (Çamkerten, 2023).

Hukuk analitiğinde metin analizi yapay zekâ yöntemleriyle yapılmaktadır. Yapay zekâ, bilgisayar sistemlerine insan benzeri zekâ ve öğrenme yeteneklerini kazandırmayı hedefleyen bir bilim ve teknoloji alanıdır (Dhar, 2016). Yapay zekânın temel amacı, bilgisayar sistemlerini karmaşık problemleri çözebilen, kararlar alabilen, örüntüleri tanıyabilen ve öğrenebilen entelektüel varlıklar haline getirmektir. Yapay zekâ, insan zekâsının temel bileşenlerini taklit etmeye çalışmaktadır. Bunlar arasında algılama, anlama, öğrenme, mantık yürütme, problem çözme, karar verme ve dil kullanma gibi yetenekler bulunmaktadır. Bilgisayar sistemleri bu yetenekleri gerçekleştirmek için algoritmalar, matematiksel modeller ve veri analizi yöntemlerinden faydalanmaktadır (Nabiyev, 2005).

Yapay zekâ kavramı ve çalışmaları, oldukça köklü bir geçmişe sahiptir. İnsanlığın yapay zekâ ile ilgili düşünceleri ve araştırmaları binlerce yıl öncesine dayanmaktadır. Ancak, yapay zekânın modern anlamda gelişimi 20. yüzyılda hız kazanmıştır. Yapay zekânın modern dönemi, 1950'lerde Alan Turing'in "bilgisayarlar düşünebilir mi?" sorusuyla başlamıştır (Say, 2021). Bu dönemde yapay zekâ, sembolik mantık ve dil işleme alanlarında yoğunlaşmıştır. John McCarthy'nin "Yapay Zekâ" terimini kullanmasıyla birlikte yapay zekâ resmi olarak tanımlanmıştır. 1956 yılında Dartmouth College'da düzenlenen Dartmouth Konferansı, yapay zekânın bilimsel bir alan olarak kabul

edilmesini sağlamıştır. Konferansta, yapay zekânın matematiksel ve sembolik yöntemleriyle ilgili çalışmalar tartışılmıştır (Haugeland, 1985). Daha sonrasında, yapay zekânın ilk programlarından biri olan Logic Theorist (Newell & Simon, 1956) geliştirilmiş ayrıca, uzman sistemlerin temelleri atılarak dil işleme, problem çözme ve öğrenme alanlarında çalışmalar yapılmıştır.

1970'li yıllara gelindiğinde, uzman sistemlerin gelişimi hızlanmış MYCIN adlı bir uzman sistem, tıbbi teşhis yapma konusunda büyük bir başarı elde etmiş ayrıca, örüntü tanıma ve görüntü işleme alanlarında da ilerlemeler kaydedilmiştir. 1980'lerde, yapay sinir ağları ve derin öğrenme gibi yöntemlerin araştırmaları hız kazanmış ve yapay zekâ bilgisayar oyunlarındaki rakipleri yenme yetenekleriyle dikkat çekmiştir (Winston, 1992). Sonraki yıllarda yapay zekâ, istatistiksel yöntemler ve makine öğrenmesiyle daha da ilerlemiş veri tabanı araştırmaları ve uzman sistemlerin ticari uygulamaları da bu dönemde artış göstermiştir. 2000'lerde büyük veri analizi, derin öğrenme ve nesne tanıma gibi alanlardaki gelişmelerle birlikte yapay zekâ, daha da popüler hale gelmiş internetin yaygınlaşmasıyla birlikte, büyük veri kaynakları ve veri analitiği, yapay zekânın gücünü artırmıştır (Metsker vd., 2019).

Yapay zekâ çalışmaları 2010'larda büyük bir ivme kazanmış, derin öğrenme ve yapay sinir ağları gibi teknikler, yapay zekânın birçok alanda çığır açmasını sağlamıştır. Büyük veri setleri ve daha güçlü hesaplama kaynakları, derin öğrenme modellerinin daha karmaşık ve etkileyici sonuçlar üretmesine olanak tanımıştır. IBM'in Watson bilgisayarı, Jeopardy! adlı televizyon yarışmasında insan rakipleriyle yarışarak büyük bir başarı elde etmiş ve bu başarı, yapay zekânın doğal dil işleme ve bilgi kazanma yeteneklerini sergileyen önemli bir dönüm noktası olmuştur (Kelly, 2017). Google'ın derin sinir ağı tabanlı görüntü tanıma modeli, ImageNet yarışmasını kazanarak görüntü tanıma alanında çığır açan bir başarı elde etmesiyle birlikte derin öğrenme ve yapay sinir ağları görüntü işleme alanında büyük bir etki yapmıştır (Szegedy vd., 2015). Sonrasında yüz tanıma teknolojilerinin gelişmesinde önemli bir adım atılmış ve Microsoft'un geliştirdiği DeepFace modeli, fotoğraflardaki yüzleri insan benzeri bir doğrulukla tanıma yeteneği göstermiştir (Wang & Deng, 2021). Ardından Google'ın geliştirdiği AlphaGo adlı yapay zekâ programı, son derece karmaşık bir strateji oyunu olan Go oyununda dünya Go şampiyonunu mağlup ederek büyük bir sürprize imza atmıştır (Curran vd., 2020).

Yapay zekânın son yıllarda hızla gelişmesiyle birlikte otonom araçlar, sanal asistanlar, öneri sistemleri, akıllı ev teknolojileri ve daha pek çok alanda yapay zekâ uygulamaları yaygın hale gelmiştir. Günümüzde yapay zekâ büyük bir hızla gelişmeye devam etmektedir. Makine öğrenmesi, derin öğrenme, doğal dil işleme, görüntü işleme ve benzeri teknikler, yapay zekânın yeteneklerini sürekli olarak genişletmektedir (Zhou, 2021).

Makine öğrenmesi, bilgisayar sistemlerinin verilere dayalı olarak öğrenme yeteneğini kazanmasını sağlayan teknikleri içeren, yapay zekânın önemli bir alt alanıdır. Makine öğrenmesi, istatistiksel modellerin kullanımıyla verilerdeki örüntüleri tanımlamakta ve gelecekteki verileri tahmin etmek veya kararlar vermek için bu örüntüleri kullanmaktadır (Alpaydın, 2014).

Derin öğrenme (deep learning), yapay zekânın önemli bir alt alan olup, yapay sinir ağları üzerine kurulu bir öğrenme yöntemidir. Bu sinir ağları, biyolojik sinir sistemlerinden esinlenerek tasarlanmıştır. Her bir katman, girdi verisini alıp içindeki nöronlar aracılığıyla bu veriyi işleyip sonuçları bir sonraki katmana iletmektedir. Bu şekilde, ağın derinlik kazanması sağlanmaktadır. Geleneksel makine öğrenmesi yöntemlerinden farklı olarak, derin öğrenme, veriye dayalı öğrenme sürecinde otomatik olarak özellik çıkarımı yapmaktadır. Belirli bir görevi yerine getirmek için kullanılan özelliklerin manuel olarak tanımlanması yerine, derin öğrenme algoritması veriden özellikleri kendisi öğrenmektedir (LeCun vd., 2015).

Yapay zekânın bir diğer önemli alanı doğal dil işlemedir. Bu alan, bilgisayar sistemlerinin insan dilini anlamasını, yorumlamasını ve üretmesini sağlayan teknikleri içermektedir. Doğal dil işleme, metin tabanlı verileri analiz ederek anlam çıkarmakta, dilbilgisi ve semantik kurallarını uygulamakta ve metin tabanlı görevleri gerçekleştirmektedir. Bu alanda otomatik metin çevirisi, metin sınıflandırma, metinden anlam çıkarma ve konuşma tanıma gibi uygulamalar bulunmaktadır (Liddy, 2001).

Yapay zekâ, sağlık, otomotiv, finans, üretim, perakende, hizmet sektörü ve daha birçok alanda kullanılmaktadır. Yapay zekânın potansiyeli oldukça geniştir ve hızla gelişmektedir. Görüntü ve ses tanıma, doğal dil anlama ve üretme, duygusal analiz, robotik, sanal gerçeklik gibi alanlarda sürekli olarak yenilikler yaşanmaktadır. Yapay zekâ aynı zamanda büyük veri analizi, nesnelerin interneti (IoT), oyun endüstrisi, güvenlik

sistemleri gibi birçok teknoloji ve sektörle birlikte entegre bir şekilde çalışmaktadır (European Commission, 2018).

Yapay zekâ hukuk alanında da kullanılmaktadır. Hukuk, büyük miktarda bilgiye dayanan karmaşık bir alandır ve yapay zekâ teknolojileri, hukukçulara bilgi işleme, analiz ve karar verme süreçlerinde yardımcı olmaktadır. Yapay zekâ, büyük miktarda hukuk belgesi, mahkeme kararı, yasalar ve emsaller gibi hukuki veriyi analiz etmek ve sınıflandırmak için kullanılarak hukuki araştırma ve bilgi işleme için kullanılmaktadır. Yapay zekâ tabanlı araçlar ve yazılımlar kullanılarak hukuki araştırma süreçlerini hızlandırılmakta ve daha etkili sonuçlar elde edilmektedir. Bunun yanında içtihat ve sözleşme analizleri ile hukuki danışmanlık hizmetleri ve mahkeme kararlarının sınıflandırılması için de yapay zekâ kullanılmaktadır (Fornaciari & Poesio, 2013; Medvedeva vd., 2020).

Mahkeme kararlarının sınıflandırılması için yapay zekâ ve makine öğrenmesi teknikleri, metin sınıflandırma ve doğal dil işleme alanında kullanılmaktadır (Medvedeva vd., 2023). Yapay zekâ ve makine öğrenmesi teknikleri, mahkeme kararlarının sınıflandırılması sürecini otomatikleştirmekte ve hukukçulara büyük bir veri setini hızlı ve etkili bir şekilde analiz etme imkânı sunmaktadır. Bu teknikler, benzer vakalardan örnekler çıkararak hukukçulara hızlı bir şekilde doğru kararlar vermeleri konusunda rehberlik etmektedir.

Mahkeme kararlarının sınıflandırılması için yapay zekâ ve makine öğrenmesi yaklaşımları, önceden belirlenmiş sınıflandırma kategorilerine göre kararları otomatik olarak sınıflandırmaktadır (Reiling, 2020). Bu sınıflandırma kategorileri, mahkeme türü, davaya ilişkin konu, kararın sonucu gibi faktörlere dayanmaktadır. Makine öğrenmesi algoritmaları, bu kategorilere dayalı olarak örüntüleri öğrenmekte ve yeni mahkeme kararlarını doğru bir şekilde sınıflandırabilmektedir. Bu teknikler ayrıca hukukçulara büyük veri setini hızlı ve etkili bir şekilde analiz etme imkânı sunmaktadır. Manuel olarak büyük miktarda mahkeme kararını incelemek zaman alırken yapay zekâ ve makine öğrenmesi teknikleri, bu süreci otomatikleştirerek zaman tasarrufu sağlamaktadır (Strickson & De La Iglesia, 2020).

Mahkeme kararlarının sınıflandırılması için yapay zekâ ve makine öğrenmesi teknikleri, bazı adımları izleyerek kullanılmaktadır. Bunlar, veri toplama, veri ön işleme,

öznitelik çıkarma, model eğitimi, model optimizasyonu ve model değerlendirmesi adımlarından oluşmaktadır (Medvedeva vd., 2023). İlk adım, sınıflandırma için gerekli olan mahkeme kararlarının bir veri seti olarak toplanması işleminin yapıldığı veri toplama adımındır. Veri seti, belirli bir hukuki konuyu veya kategoriye temsil eden kararları içermektedir. Veri setinin yapay zekâ algoritmalarının kullanılabilmesi için önceden işlenmesi gerekmektedir. Veri ön işleme adımında metin temizleme, özel karakterlerin çıkarılması, küçük harf dönüşümü gibi işlemler uygulanarak veri seti hazırlanmaktadır. Ardından gelen öznitelik çıkarımı adımında, metinlerdeki anlamlı özellikleri tanımak için öznitelik çıkarma yöntemlerini kullanılmaktadır. Kelime frekansı, cümle yapısı, anahtar kelimeler gibi öznitelikler belirlenmektedir. Model eğitimi adımında, veri seti, öznitelikler ve ilgili etiketlerle birlikte bir yapay zekâ modelini eğitmek için çeşitli makine öğrenmesi algoritmaları kullanılmaktadır. Model, veri setindeki örneklerden öğrenerek, mahkeme kararlarını doğru bir şekilde sınıflandırmayı öğrenmektedir. Model optimizasyonu adımında, eğitilmiş model, performansını artırmak için optimize edilmektedir buradaki amaç, modelin en iyi sınıflandırma doğruluğunu sağlamaktır. Model değerlendirmesi adımında eğitilen model, test veri seti veya gerçek dünya verileri üzerinde modelin sınıflandırma doğruluğu, hassasiyeti, özgünlüğü gibi performans metrikleri kullanılarak değerlendirilmektedir.

Araştırmanın **motivasyonu** hukuk sisteminde yapay zekânın kullanılmasının birçok avantajı bulunmasıdır. Yapay zekâ tabanlı sistemler, büyük miktarda hukuki belgeyi hızlı bir şekilde analiz edebilir ve sınıflandırabilir. Bu, hukukçuların zamanından ve emeğinden tasarruf etmesini, bilgiye daha hızlı erişim sağlanmasını ve işleyişin verimliliğinin artırmasını sağlar. Hukuk sistemi, birçok dava, içtihat ve hukuki belge ile ilgilenir. Yapay zekâ, büyük veri kümelerini analiz edebilir, içeriklerini kategorize edebilir ve ilişkileri belirleyebilir. Böylece hukukçuların benzer davalardan örnekler bulmalarına ve daha hızlı kararlar almalarına yardımcı olur. Yapay zekâ, mevcut verilere dayalı olarak önceden tahminler yapabilir ve riskleri değerlendirebilir. Bir davada muhtemel sonuçları tahmin edebilir veya belirli bir işlemdeki hukuki riskleri belirleyebilir. Bu yetenekleri, hukukçuların stratejik kararlar almalarına, riskleri yönetmelerine ve müvekkillerine daha iyi hukuki danışmanlık sağlamalarına yardımcı olur. Yapay zekâ tabanlı sistemler, hukukçuların içtihatları, hukuki kaynakları ve geçmiş kararları daha hızlı ve etkili bir şekilde araştırmalarını sağlar, hukukçuların daha iyi

bilgilendirilmiş kararlar almasına ve daha kapsamlı analizler yapmasına olanak tanır. Yapay zekâ, hukuk sistemini yöneten otoritelerin, kurumların ve şirketlerin hukuki uyumu sağlamalarına yardımcı olabilir. Hukuki belgelerin otomatik analizi ve sınıflandırılmasıyla, uyum süreçleri daha etkin bir şekilde yönetilebilir ve denetimler daha kapsamlı bir şekilde gerçekleştirilebilir. Yapay zekâ tabanlı sistemler sayesinde hukukun insan ön yargılarından etkilenmeden objektif bir şekilde işleyebilmesi, hukuk sisteminde tarafsızlık ve adaletin sağlanmasına yardımcı olur. Yapay zekâ sistemleri, metinleri tarafsız bir şekilde analiz eder, hukuki normlara göre sınıflandırır ve insan faktöründen kaynaklanabilecek hataları azaltır.

Yapay zekâ tabanlı sistemler, hukukçuların çalışma süreçlerini daha verimli hale getirir ve bu da hukuki maliyetlerin azaltılmasına katkıda bulunur. Yapay zekâ, otomatik metin analizi, belge sınıflandırması ve hızlı bilgi erişimi gibi özellikleriyle zaman ve kaynak tasarrufu sağlar. Bu da müvekkillerin hukuki hizmetlere daha erişilebilir maliyetlerle ulaşmasının yolunu açar. Yapay zekâ tabanlı sistemler, hukuki belgelerin analizinde yüksek bir doğruluk sağlayabilir. İnsan hatalarını minimize eder, tutarlılık sağlar ve hukuki nitelikteki işlemleri güçlendirir, bu da daha güvenilir hukuki süreçlerin yürütülmesini sağlar.

Yapay zekâ tabanlı sistemler, bu büyük veri hacimleriyle başa çıkabilir ve ölçeklenebilir bir yapı sunar. Veri miktarı arttıkça, yapay zekâ sistemleri daha fazla bilgiye erişebilir, daha kapsamlı analizler yapabilir ve daha güçlü sonuçlar üretebilir. Yapay zekâ tabanlı sistemler, hukukçuların daha karmaşık ve stratejik görevlere odaklanmasını sağlar. Rutin ve tekrarlayıcı işlerin otomatik olarak yapılmasıyla, hukukçular daha fazla zaman ve enerjiyi analiz, strateji geliştirme ve müvekkil ilişkilerine harcayabilir. Bu da hukuki hizmetlerin kalitesini artırır. Yapay zekâ tabanlı sistemlerin kullanımı, hukuk sistemine yenilikçilik ve ileri teknoloji entegrasyonu getirir. Bu, hukuk alanında dijital dönüşümü teşvik eder, hukukçuların teknolojik becerilerini geliştirmesine yardımcı olur ve daha etkin bir hukuki altyapı oluşturur. Yapay zekâ tabanlı sistemlerin hukuk sisteminde kullanılması, hukukçulara daha iyi bir çalışma ortamı sunar, hukuki hizmetlerin kalitesini artırır ve hukuki süreçleri daha etkin ve adil bir şekilde yönetmeyi sağlar.

Araştırmanın **amacı**, hukuk alanında çeşitli sorunları çözme ve işleyişi optimize etme potansiyeline sahip olan yapay zekâ tabanlı metin sınıflandırma yöntemlerini kullanarak *sanal yargıç* adı verilen ve hukuk aktörlerine yardımcı olacak olan bir uygulama sunmaktır. Yapay zekâ tabanlı metin sınıflandırma yöntemleri, büyük veri kümelerini hızlı ve etkin bir şekilde analiz ederek, hukukçuların bilgiye daha hızlı erişmesini olanak vermekte bu da dosya araştırması, yasal hükümlerin analizi ve benzeri süreçlerde zaman ve emek tasarrufu sağlamaktadır.

Yapay zekâ tabanlı metin sınıflandırma yöntemleri, mevcut verilere dayanarak önceden tahminlerde bulunmakta ve kararlara destek sağlayabilmektedir. Benzer davalardan elde edilen verilere dayanarak belirli bir davada muhtemel sonuçlar hakkında tahminler yapabilmekte veya hukukçulara benzer vakalardan örnek kararlar sunabilmektedir. Böylece hukukçuların daha iyi bilgilendirilmiş kararlar almasına yardımcı olmaktadır. Hukuk alanında, geçmiş kararlar ve içtihatlar büyük öneme sahiptir. Yapay zekâ tabanlı metin sınıflandırma yöntemleri, bu içtihatları araştırmacının amacına göre sınıflandırarak istenilen içtihadı daha kısa sürede ulaşılmasına olanak sağlamaktadır. Yapay zekâ tabanlı metin sınıflandırma yöntemleri, hukuk büroları veya mahkemeler gibi kurumlar için ön inceleme ve dava değerlendirme süreçlerinde de kullanılma olanağına sahiptir. Yapay zekâ modelleri, dava dosyalarını analiz ederek, hukuki argümanları, delilleri ve benzer davaları belirlemektedir. Böylece hukukçuların dava değerlendirme süreci hızlanmakta ve stratejik kararlar almalarına yardımcı olunarak kaynakları daha verimli kullanmalarını sağlamaktadır. Yapay zekâ tabanlı metin sınıflandırma yöntemleri, hukuki riskleri değerlendirmek ve hukuki danışmanlık sağlamak için de kullanılmaktadır. Bir şirketin hukuki belgeleri analiz edilerek, riskli alanlar belirlenebilmekte ve mevzuata uygunluğu konusunda önerilerde bulunabilmektedir. Böylece şirketlerin hukuki risklerini azaltmalarına, yasa ve düzenleyici diğer kurallara uyum sağlama süreçlerinin iyileştirmesine yardımcı olmaktadır.

Hukuk sisteminde eşitlik ve adaletin sağlanması önemlidir. Yapay zekâ tabanlı metin sınıflandırma yöntemleri, insan ön yargılarından arındırılmış objektif bir yaklaşım sunabilme potansiyeline sahiptir. Hukuk alanında, yasalar, yönetmelikler gibi hukuki normlar sürekli olarak güncellenmektedir. Yapay zekâ tabanlı metin sınıflandırma yöntemleri, bu güncellemeleri takip ederek, hukuki belgelerin uyumlu olup olmadığını

veya belirli bir davadaki mevzuat deęişikliklerini tespit edebilmektedir. Böylece, hukukçuların hızlı bir şekilde mevzuat deęişikliklerine adapte olmaları sağlanmaktadır.

Potansiyel etkileri göz önüne alındığında, yapay zekâ ve makine öğrenmesi, hukukta adaletin sağlanması açısından önemli bir rol oynayabilir. Yapay zekâ, insan faktörünün neden olduęu önyargılardan etkilenmeden tarafsız bir şekilde çalışabilir. Ayrıca yapay zekâ, büyük miktardaki veriyi hızlı bir şekilde analiz edebilir ve ilişkileri belirleyebilir. Adaletin sağlanması için zamanında ve adil kararlar almak önemlidir ve yapay zekâ bu süreci hızlandırabilir. Yapay zekâ, mevcut verilere dayalı olarak önceden tahminler yapabilir. Bir davada muhtemel sonuçları tahmin edebilir veya belirli bir işlemin hukuki risklerini belirleyebilir. Bunun sayesinde hukukçuların stratejik kararlar almalarına ve adil sonuçlara ulaşmalarına yardımcı olur. Yapay zekâ tabanlı sistemler, hukuk hizmetlerinin eşitlik ve erişilebilirlik ilkesine uygun bir şekilde sunulmasına katkıda bulunabilir. Özellikle, daha fazla insanın hukuki hizmetlere erişebilmesini sağlayabilir ve hukuk sistemine erişimdeki eşitsizlikleri azaltabilir.

Ancak, yapay zekânın adaletin sağlanması sürecinde kullanımıyla ilgili bazı zorluklar ve dikkat edilmesi gereken konular da vardır. Özellikle, algoritmaların önyargı veya ayrımcılık içerebileceęi, veri mahremiyeti ve güvenliği konuları gibi etik ve hukuki konulara dikkat edilmesi gerekmektedir. Yapay zekâ yöntemleriyle mahkeme kararlarının sınıflandırıldığı bu çalışmada cevaplanmak istenen **araştırma soruları** şunlardır:

- Yapay zekâ ve doğal dil işleme nedir ve aralarındaki ilişki nasıldır?
- Makine öğrenmesi ve denetimli makine öğrenmesi algoritmaları nelerdir?
- Yapay zekâ ile hukuk arasındaki ilişki nasıldır?
- Türk yargısı mahkeme çeşitleri ve Uyuşmazlık Mahkemesi (UYM) kararlarının nitelięi nedir?
- UYM kararlarının sınıflandırılmasında kullanılan algoritmaların sonuçları nelerdir?
- Hukuk alanında yapay zekânın kullanılmasıyla ilgili mevcut örnekler ve uygulamalar nelerdir?

Bu bağlamda bu tez çalışmasında **araştırma yöntemi** olarak denetimli makine öğrenmesi sınıflandırma algoritmalarından lojistik regresyon analizi, doğrusal diskriminant analizi, karar ağaçları, rassal orman, destek vektör makineleri, yapay sinir ağları, k-en yakın komşu ve naive bayes yöntemleri kullanılmıştır. Veri seti UYM olumsuz görev uyumsuzluğu adli ve idari kararlarından oluşmaktadır. Veriler, yazılmış olan bot uygulama ile UYM'nin kamuya açık platformundan çekilmiştir.

Çalışmanın gerçekleştirilmesinde bazı **sınırlılıklar** söz konusudur. Hukuki metinler üzerine yapay zekâ ve makine öğrenmesi tekniklerinin kullanımı sırasında özellikle hukuki metinlerin karmaşıklığı ve dilin nüansları göz önünde bulundurulmalıdır. Yargı kararlarının analizi ve yorumlanması genellikle karmaşık ve çok boyutlu bir süreçtir ve sadece yapay zekâ tekniklerine dayalı kararlar yetersiz olabilir. UYM'nin olumsuz görev uyumsuzluğu kararları kamuya açık platformda bulunulandan çok daha fazladır. Ancak kullanıcıların erişebileceği kararlar sadece kamuya açık platformdakilerle sınırlıdır. Burada verilen kararların yazım şeklinde de belirli bir standart bulunmamaktadır. Veri ön işleme adımları ve manuel düzeltmeler ile internet sitesinde yayımlanan kararların %60'a yakını analiz için kullanılabilmiştir.

Mevcut **araştırmanın katkısı** birkaç farklı şekilde gerçekleşmiştir. Bu çalışmadan elde edilen bulgular yapay zekâ modellerinin mahkeme kararlarını sınıflandırmada anlamlı bir başarı elde ettiğini göstermiştir. Aynı zamanda çalışma, *sanal yargıç* anlamında mahkeme kararını tahmin eden bir model sunmuştur bu da yapay zekânın hukuk alanında bir süje olarak görülmesine öncülük edecektir. Bununla beraber hukuk bilimine deneysel yaklaşımın da kapılarını açmıştır.

Mahkeme kararlarının sınıflandırılmasıyla ilgili yapılan bu tez çalışması, birçok alanda önemli katkılar sağlayacak potansiyele de sahiptir. Mahkeme kararlarının sınıflandırılması, hukuk sistemini geliştirmek için önemli bir araç olabilir. Sınıflandırma analizi, benzer davaların incelenmesine ve benzer sonuçların ortaya çıkmasına yardımcı olarak yargı sisteminin tutarlılığını ve adil olma düzeyini artırabilir. Ayrıca, mahkeme kararlarının sınıflandırılması, avukatlar, yargıçlar gibi hukuk alanında çalışan kişilere pratikte fayda sağlayabilir. Sınıflandırma modelleri, geçmiş kararları analiz ederek benzer vakalar için yol gösterici olabilir ve hukukçulara daha iyi bir şekilde hazırlanmaları için temel sağlayabilir. Mahkeme kararlarının sınıflandırılmasıyla ilgili yapılacak çalışmalar,

kararların erişilebilirliğini artırabilir. Sınıflandırma, büyük miktarda mahkeme kararının yapılandırılmış bir şekilde sunulmasını sağlayarak, hukuk sistemi dışındaki kişilerin de kararlara erişimini kolaylaştırabilir. Mahkeme kararlarının sınıflandırılması, veri analizi için zengin bir kaynak sağlar. Bu analizler, hukuk alanında eğilimlerin belirlenmesi, hukuki konuların incelenmesi ve hukuk politikalarının oluşturulması gibi konularda bilgi sağlayabilir. Mahkeme kararlarının sınıflandırılması, yapay zekâ uygulamalarının geliştirilmesi için önemli bir alan olabilir. Sınıflandırma modelleri, makine öğrenmesi ve doğal dil işleme tekniklerinin kullanılmasıyla daha da iyileştirilebilir ve hukuki kararları otomatik olarak analiz etmek için kullanılabilir.

Bu tez çalışması gibi mahkeme kararlarının sınıflandırılmasıyla ilgili araştırmalar, gelecekteki yargısal kararların öngörülmesi için kullanılabilir. Makine öğrenmesi ve istatistiksel modeller kullanılarak, geçmiş kararlar üzerinde yapılan analizler gelecekteki davalar için tahminler yapmada kullanılabilir. Bu, hukuk sisteminin daha öngörülebilir olmasına ve tarafların daha iyi bilgilendirilmesine yardımcı olabilir. Mahkeme kararlarının sınıflandırılması, hukuk içtihadının toplanması ve incelenmesi sürecini kolaylaştırabilir. Bu tür çalışmalar, benzer içtihatları tespit etmek, hukuki tartışmaların analizini yapmak ve hukuki prensipleri belirlemek için kullanılabilir. Mahkeme kararlarının sınıflandırılması, adalet sistemi için verimlilik ve maliyet tasarrufu sağlayabilir. Sınıflandırma analizi, benzer davalarda kullanılan argümanları ve kararları belirleyerek, davaların daha hızlı bir şekilde sonuçlandırılmasına ve yargılama sürecinin daha etkin bir şekilde yönetilmesine yardımcı olabilir. Mahkeme kararlarının sınıflandırılmasıyla ilgili araştırmalar, yasama sürecine katkıda bulunabilir. Sınıflandırma analizi, mevcut yasal düzenlemelerin etkisini değerlendirmek, yasal boşlukları belirlemek ve yasama organlarının daha etkili ve bilinçli kararlar almasına yardımcı olmak için kullanılabilir.

Bu şekilde, mahkeme kararlarının sınıflandırılmasıyla ilgili bu tez çalışması, hukuk sistemine, hukuk pratiğine, veri analizine, yapay zekâ uygulamalarına, hukuki içtihadın toplanması ve incelenmesine, adalet sisteminin verimliliğine ve maliyet tasarrufuna, adalet sistemi ve yasama sürecine katkıda bulunacaktır.

Mahkeme kararlarının sınıflandırılmasıyla ilgili araştırmalar, hukuki ihtilafların çözümünde alternatif yaklaşımların geliştirilmesine yardımcı olabilmenin yanında,

hukuki eğitim ve arařtırmalar için önemli bir kaynak olabilir. Öğrenciler ve arařtırmacılar, sınıflandırma modellerini kullanarak hukuki konuları inceleyebilir, hukuki analiz becerilerini geliřtirebilir ve hukukun farklı alanlarındaki içtihatları daha iyi anlayabilir. Bununla birlikte sınıflandırma analizi, mevcut hukuki düzenlemelerin etkinliğini ve uygulanabilirliğini deęerlendirmek, eksiklikleri belirlemek ve daha iyi hukuki politikalar oluřturmak için kullanılabilir. Sınıflandırılmış kararlar, hukuki süreçlerin ve kararların anlaşılmasını kolaylařtırabilir, hukuki tartışmalara katkıda bulunabilir ve hukuki kararların adalet ve toplumsal deęerlerle uyumunu deęerlendirmek için bir temel saęlayabilir.

Bu tip arařtırmalar, hukuk sistemini daha etkili, erişilebilir ve adil hale getirme potansiyeline sahiptir. Aynı zamanda, yapay zekâ, veri analizi ve makine öğrenmesi gibi teknolojilerin kullanımıyla birlikte daha gelişmiş ve otomatikleřtirilmiş hukuki çözümlerin geliřtirilmesine katkıda bulunabilirler. Tüm bu etkiler, daha adil bir hukuk sistemi ve daha iyi hukuki sonuçlar elde etme hedefine yönelik bir adım olabilir.

Tezin planına göre bu çalışma üç bölümden oluřmaktadır. İlk bölümde yapay zekâ kavramı, doęal dil işleme süreci ele alınmıştır. Daha sonra, metin sınıflandırma aşamalarına genel bir bakış atılmıştır. Burada metin ön işleme, öznitelik çıkarımı, boyut azaltma, sınıflandırma teknikleri ve bunların performans deęerlendirmelerinin neler olduęu üzerinde durulmuřtur. İkinci bölümde, makine öğrenmesi, denetimli makine öğrenmesi ve denetimli makine öğrenmesinin aşamalarına CRISP-DM özelinde bakılmış ve sınıflandırma tekniklerinin neler olduęu ve çalışma prensipleri açıklanmıştır. Üçüncü bölümde UYM'nin Türk yargı sistemindeki yeri ve UYM kararlarının nitelięi hakkında bilgiler verilmiş, veri toplama yöntemleri tanıtılmış ve yapay zekâ yöntemleri ile UYM kararları tahminine yönelik deneysel analizler gerçekleştirilmiştir. Tez çalışması, teorik ve pratik bir perspektiften elde edilen sonuçların tartışılmasıyla tamamlanmıştır.

BİRİNCİ BÖLÜM

YAPAY ZEKÂ VE DOĞAL DİL İŞLEME

Yapay zekâ ve doğal dil işleme, günümüzde bilgi teknolojileri alanında hızla gelişen ve yaygınlaşan alanlardır. Yapay zekâ, bilgisayar sistemlerinin insan benzeri zekâyâ sahip olmasını sağlayan bir disiplindir ve genellikle verileri analiz etme, karar verme, tahmin yapma, otomatik öğrenme gibi karmaşık görevleri yerine getirmek için kullanılır. Doğal dil işleme ise, bilgisayarların doğal dildeki verileri anlamalarına ve işlemelerine olanak tanıyan yapay zekânın bir alt dalıdır (Eyecioğlu Özmutlu, 2021).

Yapay zekâ ve doğal dil işleme arasındaki ilişki, yapay zekânın doğal dil verileri üzerinde uygulanması ve doğal dili anlama yeteneğini geliştirmeye odaklanan bir alandır. Yapay zekâ, doğal dil işleme tekniklerini kullanarak metin verilerini analiz eder, anlamlı sonuçlar çıkarır. Yapay zekâ ve doğal dil işleme, birçok farklı şekilde birbirleriyle ilişkilidir ve birbirlerini tamamlamaktadır. Dil modelleri, metin verilerini analiz etmek, tahmin yapmak ve anlam çıkarmak için kullanılan matematiksel modellerdir. Yapay zekâ, doğal dil işleme tekniklerini kullanarak dil modelleri oluşturur ve eğitir, bu dil modellerini geliştirip doğal dili anlama yeteneğini artırır (Adalı, 2012).

Yapay zekâ, doğal dil işleme tekniklerini kullanarak metin analitiği, otomatik dil çevirisi ve duygu analitiği yapar. Metin analitiği, metin verilerini anlamak, anlamlı bilgileri çıkarmak ve sonuçları yorumlamak için kullanılan bir yöntemdir. Yapay zekâ, bu yöntemi kullanarak metin verilerinden anlamlı sonuçlar elde etmektedir. Otomatik dil çevirisi, bir dilden diğerine metin verilerini çevirmek için kullanılan bir yöntemdir. Duygu analitiği ise metin verilerindeki duyguları anlamak ve ortaya çıkarmak için kullanılmaktadır (Şeker, 2015).

Yapay zekâ ile doğal dil işleme tabanlı sistemler geliştirilir. Bu sistemler, kullanıcıların doğal dillerindeki verileri giriş olarak kullanabilmelerine ve anlamlı sonuçlar elde edebilmelerine olanak tanır. Bununla birlikte yapay zekâ ve doğal dil işleme teknikleriyle sesli asistanlar geliştirilir. Geliştirilen bu sesli asistanlar sayesinde sistemler, kullanıcıların sesli komutlarını anlar ve buna uygun yanıtlar verir (Ay, 2021).

Yapay zekâ ve doğal dil işlemenin beraber kullanıldığı diğer bir alan ise doğal dil tabanlı veri analitiğidir. Doğal dil tabanlı veri analitiği, metin verilerini analiz ederek anlamlı sonuçlar çıkarmak ve veri analitiğini geliştirmek için kullanılan bir yöntemdir. Bununla beraber her iki yöntemle veri analitiği süreçleri optimize edilebilir (Adalı, 2012).

Yapay zekâ ile doğal dil işleme arasındaki ilişki, birçok farklı sektörde, özellikle müşteri hizmetleri, pazarlama, finans, sağlık ve e-ticaret gibi alanlarda uygulanabilir ve kullanıcı deneyimini, veri analitiğini ve karar verme süreçlerini iyileştirmeye yardımcı olabilir. Yapay zekâ ve doğal dil işleme alanlarındaki gelişmeler, gelecekte artması beklenen yapay zekâ ve doğal dil işleme uygulamalarının daha da ileri düzeyde kullanılmasını sağlayacaktır.

1.1.Yapay Zekâ

“Tarihte üç büyük olay vardır. Bunlardan ilki kâinatın oluşumudur. İkincisi yaşamın başlangıcının olmasıdır. Üçüncüsü de yapay zekânın ortaya çıkışıdır.” BBC ile söyleşisinde MIT Bilgisayar Bilimleri laboratuvar yöneticilerinden Edward Fredkin yukarıdaki ifadeleri kullanarak yapay zekânın önemini çok güzel bir şekilde ifade etmiştir (Pirim, 2006).

1943 yılında McCulloch ve Pits’in insan beyninin Boolean devre modellerinden söz etmesi ardından da 1950 yılında Alan Turing’in düşünebilen makineler fikri, 1956 yılında Dartmouth görüşmesinde “Yapay Zekâ” isminin ortaya atılmasına sebep olmuştur (Lewis, 2014). Bu tarihten sonra IBM firması satranç oynayabilen ilk programı yazmış ve yapay zekâ konusunda ilk uluslararası konferans düzenlenmiştir. Sonrasında Arthur Samuel tarafından geliştirilen ve bilgisayarların oyunlarda insanlara meydan okumasını ve daha iyi hale gelmesini sağlayan temel bir fikri temsil eden öğrenme programı, Newell ve Simon’un mantık teorisi, Gelernter’in geometri motoru, John Alan Robinson’un matematiksel mantığı ve formel mantığı kullanarak mantıklı düşünme problemlerini otomatik olarak çözmek için tasarladığı (resolution theorem proving) algoritması 1960'ların sonlarına kadar yapay zekâ ile ilgili araştırmaların gelişme aşamaları olarak görülebilir. 1974-1980 arasında yapay zekâ çalışmalarından aşırı bir beklenti oluşmuş ancak bu beklentiler karşılanamamıştır. Böylece bu alandaki çalışmalarda finansman kesintileri baş göstermiş bununla beraber donanım ve veri yetersizliği gibi etkenler de eklenince yapay zekâ alanındaki ilerlemelerde ciddi bir

düşüşün yaşandığı “AI winter” yani “yapay zekâ kışı” olarak bilinen bir dönem olmuş ve yapay zekâyı ilgi bu dönemde azalmıştır. Ancak 1987 yılında artık yapay zekâ bir bilim haline gelmeye başlamıştır. 1997 yılında Deep Blue adlı satranç oynayan bilgisayar programı, dünya satranç şampiyonu Kasparov'u yenmiş, 1998 yılında internetin yaygınlaşması ile beraber yapay zekâ tabanlı programlar geniş kitlelere ulaşmıştır (Pirim, 2006). Artık günümüzde kullandığımız birçok ev aleti, oyuncak ve iş yerlerinde kullanılan programlarda yapay zekâ teknolojileri kullanılmaktadır.

İnsanların yapabildiği işlerin bilgisayarlar ve makineler tarafından yapılabilmesi ve bunun için tekniklerin geliştirilmesi şeklinde tanımlanabilecek olan yapay zekânın birçok araştırmacı tarafından farklı tanımları yapılmıştır. Yapay zekâyı bir bilgisayarın ya da makinenin insana özgü kabul edilen akıl yürütme, anlam çıkarma, genelleme ve geçmiş deneyimlerden öğrenme gibi yüksek zihinsel süreçlere ilişkin görevleri yerine getirme yeteneği olarak tanımlayan araştırmacılar olduğu gibi (Nabiyev, 2005) yapay zekâyı insanın yaptıklarının bilgisayara yaptırabilme çalışması olarak gören araştırmacılar da (Nemov, 1999, aktaran Nabiyev, 2021) vardır.

Yapay zekâ düşünmeyi taklit ederek karmaşık problemleri çözen, bir durum karşısında kişilere yanıt verebilen, öğrenerek uzmanlığını genişleten, yeni bilgilerle veri tabanını geliştirebilen akıllı programları hedefleyen bir bilim dalı olarak da tanımlanabilir.

Cumhurbaşkanlığı Dijital Dönüşüm Ofisi tarafından yayımlanan “Ulusal Yapay Zekâ Stratejisi”nde ise yapay zekâ; en genel haliyle, bir bilgisayarın veya bilgisayar kontrolündeki bir robotun çeşitli faaliyetleri zeki canlılara benzer şekilde yerine getirme kabiliyeti olarak tanımlanmaktadır. Yapay zekâ terimi; dinamik ve belirsiz ortamlarda akıl yürütme, anlam keşfetme, genelleme veya geçmiş deneyimlerden öğrenme gibi insanlara özgü bilişsel kabiliyetlerle donatılmış sistemler için kullanılmaktadır. Yapay zekâ, başta büyük veri ve hesaplama olmak üzere diğer dijital teknolojiler ile sınır bilimden de beslenen bir disiplindir. Yapay zekâ destekli sistemler, verinin barındırdığı özelliklerden ve örüntülerden öğrenen, bu doğrultuda güncellenebilen arama ve tahminleme yapabilen gelişmiş algoritmalar kullanmaktadır (Cumhurbaşkanlığı Dijital Dönüşüm Ofisi Başkanlığı & Sanayi ve Teknoloji Bakanlığı, 2021).

Avrupa Komisyonunca yapılan tanıma göre ise yapay zekâ, belirli hedeflere ulaşmak için çevrelerini analiz ederek ve -bir dereceye kadar özerklikle- harekete geçerek akıllı davranışlar sergileyen sistemleri ifade eder. Yapay zekâ tabanlı sistemler tamamen yazılım tabanlı olabilir, sanal dünyada bulunabilir (ses asistanları, görüntü analiz yazılımı, arama motorları, konuşma ve yüz tanıma sistemleri) veya donanım cihazlarında (gelişmiş robotlar, otonom arabalar, dronlar veya nesnelerin interneti uygulamaları) gömülü olarak bulunabilirler (European Commission, 2018).

Farklı şekilde ifade edilmesi gerekirse, eldeki sorun biliniyor fakat çözüm yöntemi (algoritması) bilinmiyorken, doğru ve verimli bir çözüm yöntemini çıkarımlayan, öğrenen ya da keşfeden, insan eliyle üretilmiş sistemlerin tümüne yapay zekâ denir (Köroğlu, 2017). Bu anlamda çözüm sağlayacak iyi algoritmaları çıkaracak olan makineleri kullanma fikri ilk olarak Alan Turing tarafından söylenmiştir (Turing, 1950). Yapay zekâ insan zekâsına yetişemese de temelde neden-sonuç ilişkilerini oluşturan, verileri analiz ederek bunlar üzerinde genellemeler yapan ve yeni şeyler keşfeden özelliklere sahiptir.

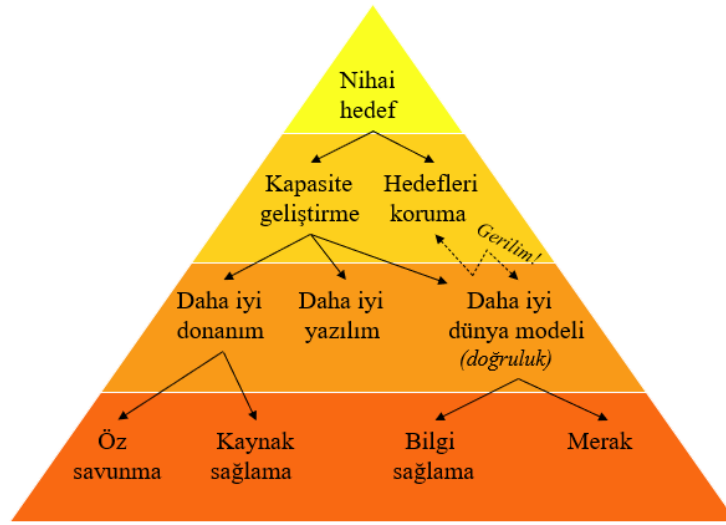
Yapay zekâ sistemleri ile varılmak istenen nokta, sistemin tıpkı bir insan gibi kendi kendine karar vermesini sağlamak ve bu kararları uygulayabilecek yapıları geliştirmektir (Kolbjørnsrud vd., 2016). Yapay zekâ sayısal veriler dışında yazı, görüntü ve ses gibi farklı yapılarıdaki verileri de işleyip, analiz edebilen ve anlamlı sonuçlar çıkaran bir yapıdır (Dhar, 2016). Yapay zekâ sistemleri; uzman sistemler ile yorum yapabilme, problem çözebilme, karar verebilme; yapay sinir ağları ile öğrenebilme, genetik algoritmalar ile karmaşık problemleri çözebilme, bulanık önermeler mantığı ile kelimeleri anlayabilme, doğal dil işleme ile metinleri okuyabilme ve anlamlandırabilme gibi yeteneklere sahiptir (Öztemel, 2020).

Yapay zekâ, sanal bir kişisel asistan olarak günlük çalışma planını düzenlemekten, otonom araçlarla seyahat etmeye, telefonların kişilerin beğenebileceği şarkıları veya restoranları önermesine kadar birçok alanda artık hayatımızın bir parçası olmuş durumdadır. Hayatı kolaylaştırmaktan öte, yapay zekâ kronik hastalıkları tedavi etmek, trafik kazalarındaki ölüm oranlarını azaltmak, iklim değişikliği ile mücadele etmek veya siber güvenlik tehditlerini önceden tahmin etmek gibi dünyanın en büyük sorunlarını çözmekte yardımcı olmaktadır (European Commission, 2018).

Yapay zekâ, Danimarka'da, acil servislerin çağrıcının sesine dayanarak kalp durması veya diğer koşulları teşhis etmesine izin vererek hayat kurtarmaya yardımcı olmaktadır. Avusturya'da, radyologların diğer tıbbi verilerle hemen karşılaştırarak tümörleri daha doğru bir şekilde tespit etmelerine yardımcı olan X-ışınlarını anında karşılaştıran büyük miktarda tıbbi veriyi kullanan yapay zekâ sistemleri bulunmaktadır. Avrupa genelinde birçok çiftlik hayvanının hareketini, sıcaklığını ve yem tüketimini izlemek için yapay zekâ kullanılmaktadır. Yapay zekâ sistemi, çiftçilerin hayvanların refahını izlemelerine ve ısıtma ve yemleme makinalarını otomatik olarak uyarlamalarına yardımcı olmakta ve diğer görevler için onları serbest bırakmaktadır. Ayrıca yapay zekâ, Avrupa üreticilerinin daha verimli hale gelmelerine ve fabrikalarını Avrupa'ya geri getirmelerine yardımcı olmaktadır. Bu, enerjiden eğitime, finansal hizmetlerden inşaata kadar tüm sektörlerde yapay zekânın neler yapabileceğine dair bildiğimiz birkaç örnektir. Bugün hayal edilemeyen sayısız daha fazla örnek gelecek on yıl boyunca ortaya çıkacaktır. Geçmişte buhar motoru veya elektrik gibi, yapay zekâ dünyamızı, toplumumuzu ve endüstrimizi dönüştürmektedir. Bilgi işlem gücündeki artış, veri erişilebilirliği ve algoritma ilerlemeleri, yapay zekâyı 21. yüzyılın en stratejik teknolojilerinden biri haline getirmiştir (European Commission, 2018).

Yapay zekâ, hızlı ve doğru sonuçlar elde etme, sürekli çalışabilirlik, yüksek verimlilik, düşük hata oranı, veri tabanlı kararlar alma, zorlu problemleri çözme yeteneği ve iş süreçlerinde otomasyon sağlama gibi birçok avantaj sunmaktadır. Ancak, etik ve mahremiyet konuları, algoritma önyargısı, güvenlik tehditleri, insanların işlerini ve yaşamlarını etkileme potansiyeli, insan zekâsı ile tam anlamıyla rekabet etme yeteneğinin sınırlı olması ve sosyal etkiler gibi bazı zorlukları da bulunmaktadır.

Yapay zekânın hayatımızın içinde yer aldığı yeni toplum düzeninde yönlendirici olacak hukuk kuralları, etik ilkelerin kabul edildiği çerçevede belirlenecektir. Bu çerçeve mevcut bazı etik değerler üzerinde yükselerek daha iyi bir dünya modeli kurma hedefine yönelecektir.



Şekil 1: Yapay Zekâ Hedefler Piramidi (Tegmark, 2017).

Fizikçi ve kozmolog Max Tegmark tarafından yazılan yapay zekâ alanında yaşanabilecek gelişmeleri ve bu gelişmelerin insanoğlunu nasıl etkileyebileceğini ele alan “Life 3.0: Being Human in the Age of Artificial Intelligence” adlı kitapta yapay zekânın nihai hedefi ve alt hedeflerini açıklamak için “Yapay Zekâ Hedefler Piramidi” adı verilen bir model kullanılmaktadır. Bu piramit, yapay zekânın gelişiminin aşamalarını ve potansiyel hedeflerini hiyerarşik bir şekilde göstermektedir.

En alt seviyede bulunan ve en temel hedef olarak kabul edilen şey, “Zayıf Yapay Zekâ”dır. Bu, özel bir görevi yerine getiren bir yapay zekâ sistemi anlamına gelir. Örneğin, bir satranç oyununda iyi performans gösteren bir yapay zekâ programı zayıf yapay zekâ örneği olmaktadır.

Bir üst seviyede ise “Güçlü Yapay Zekâ” yer alır. Bu seviyede yapay zekâ, bir insanın sahip olduğu zekâ düzeyine eşdeğer bir seviyeye ulaşır. Güçlü Yapay Zekâ sistemi, karmaşık sorunları çözebilir, yaratıcı düşünebilir, karmaşık kararlar verebilir ve genel zekâ görevlerini yerine getirebilir.

Daha sonra “Yüksek Seviyeli Yapay Zekâ” gelir. Bu seviye, insan zekâsını önemli ölçüde aşan bir yapay zekâ sistemini ifade eder. Yüksek Seviyeli Yapay Zekâ, insan zekâsının sınırlamalarından muaf ve daha karmaşık problemleri çözebilir, bilgiyi anlama ve sentezleme becerisine sahiptir.

En üst seviye olan “Süper Zekâ” ise, insan zekâsını büyük ölçüde aşan ve kavrayış, problem çözme, öğrenme ve yaratıcılık gibi alanlarda olağanüstü bir yetenek sergileyen yapay zekâ sistemini ifade eder. Süper zekâ, insanoğlunun anlamaya bile çalışamayacağı sorunları çözebilir ve insanın hayal edemeyeceği yeni fikirler ve keşifler yapabilir.

Bu hedefler piramidi, yapay zekânın gelişimindeki farklı aşamaları ve potansiyel sonuçlarını göstermektedir. “Nihai hedef”, yapay zekânın ulaşmayı amaçladığı en üst noktayı ifade ve yapay zekânın gelişimindeki en yüksek seviye veya son hedefi temsil eder. Bu seviyede yapay zekâ, insan zekâsını aşan ve olağanüstü yeteneklere sahip bir düzeye erişir. Örneğin, insanın anlamaya bile çalışamayacağı sorunları çözebilir, yeni fikirler üretebilir ve insanlığın ilerlemesi için büyük bir katkı sağlayabilir. Bu terim yapay zekânın en yüksek ve en istenen noktasını ifade ederken, aynı zamanda bu noktaya ulaşmanın etik, sosyal ve güvenlik gibi önemli sorunları da beraberinde getirebileceğini de vurgular.

“Kapasite geliştirme”, yapay zekânın mevcut yeteneklerini ve kapasitesini artırmayı hedeflediği anlamına gelir. Yapay zekânın yetenek artırma amacı, daha karmaşık görevleri yerine getirebilme, daha iyi sonuçlar üretebilme ve insan zekâsının ötesine geçebilme becerisini geliştirmektir. Bu, yapay zekânın kendisini daha fazla veriyle besleyerek öğrenme yeteneğini artırması, algoritmalarını geliştirmesi ve daha karmaşık problemleri çözebilme yeteneğini kazanması anlamına gelir. Yetenek artırma, yapay zekânın daha güçlü, akıllı ve etkili hale gelmesini sağlayarak, daha geniş bir yelpazede uygulamalara yönelik kullanım potansiyelini artırır. Bu, yapay zekânın daha iyi tahminler yapabilmesi, daha hızlı kararlar alabilmesi ve daha karmaşık analizler gerçekleştirebilmesi anlamına gelebilir.

“Hedefleri koruma”, yapay zekânın belirlenen hedefleri veya amaçları koruma ve sürdürme yeteneğini ifade eder ve bu belirli bir amaca veya hedefe odaklanma ve bu hedefi gerçekleştirmek için gereken stratejileri ve çözümleri sürdürme becerisini içerir. Bu, yapay zekânın belirli bir görev veya probleme odaklanarak, o hedefe ulaşmak için yöntemlerini koruması ve değişen koşullara uyum sağlaması anlamına gelir.

“Gerilim”, hedefleri koruma ve daha iyi bir dünya modeli arasındaki gerilimi veya çatışmayı ifade eder. Bu gerilim, yapay zekânın hedef odaklılığı ile etik ve toplumsal

sorumluluk arasındaki dengeyi bulma zorluğunu temsil eder. Yapay zekânın “gerilim”i, hedeflerini korumakla birlikte daha iyi bir dünya oluşturma çabası arasında denge sağlama ve uyum yakalamaya çalışmasını ifade eder.

“Daha iyi donanım”, yapay zekânın gelişimi için daha gelişmiş ve daha etkili donanım teknolojisinin kullanılmasını ifade eder. Bu ifade, yapay zekânın performansını artırmak, hesaplama gücünü ve hızını yükseltmek, daha karmaşık algoritmaları işleyebilmek ve daha geniş veri setlerini işleyebilmek için daha gelişmiş donanım bileşenlerinin kullanılmasını temsil eder. Yapay zekâ sistemleri, büyük veri setlerini işlemek, karmaşık derin öğrenme ağları ve yapay sinir ağları çalıştırmak gibi yoğun hesaplama gerektiren işlemler yapmaktadır. “Daha iyi donanım” terimi, bu gibi işlemleri daha hızlı ve verimli bir şekilde gerçekleştirebilen, daha yüksek performanslı işlemciler, ekran kartları (GPU'lar), özelleştirilmiş donanım çözümleri veya daha etkili veri depolama sistemlerini ifade etmektedir. “Daha iyi donanım” kullanımı, yapay zekânın kapasitesini artırmakta, daha karmaşık problemleri çözme yeteneğini geliştirmekte ve yapay zekânın daha verimli ve etkili bir şekilde çalışabilmesini sağlamaktadır.

“Daha iyi yazılım”, yapay zekâ sistemlerinin geliştirilmesinde daha gelişmiş ve etkili yazılım teknolojilerinin kullanılmasını ifade ve yapay zekânın performansını artırmak, daha karmaşık algoritmaları uygulamak, daha doğru tahminler yapmak ve daha iyi sonuçlar elde etmek için daha gelişmiş yazılım bileşenlerinin kullanılmasını temsil etmektedir. Yapay zekâ sistemleri, veri analizi, örüntü tanıma, makine öğrenmesi ve derin öğrenme gibi karmaşık işlemleri gerçekleştirir. “Daha iyi yazılım” terimi, bu gibi işlemleri daha etkili ve verimli bir şekilde gerçekleştirebilen, daha sofistike algoritmaları, daha iyi optimizasyon tekniklerini ve daha güçlü öğrenme modellerini ifade eder. “Daha iyi yazılım” kullanımı, yapay zekânın yeteneklerini artırır, daha doğru sonuçlar elde etme ve daha karmaşık problemleri çözme yeteneğini geliştirir.

“Daha iyi dünya modeli”, yapay zekânın gerçek dünyayı daha iyi anlamak ve modellemek için geliştirdiği bir temsil şeklini ifade etmekte ve yapay zekânın çevresini, olayları ve insan davranışını daha hassas ve doğru bir şekilde anlamasını ve tahmin etmesini sağlayan bir modele atıfta bulunmaktadır. Bu model, yapay zekânın dünya ile etkileşim kurmasını, çevresel koşulları analiz etmesini ve daha iyi kararlar almasını sağlamaktadır. Yapay zekâ, daha iyi bir dünya modeline sahip olduğunda, daha doğru

tahminler yapabilir, daha etkili stratejiler geliştirebilir ve daha iyi sonuçlar üretebilir. Bu model, yapay zekânın daha iyi bir anlayışa sahip olmasını, çevreyi daha kapsamlı bir şekilde temsil etmesini ve insanlık için daha olumlu etkiler yaratmasını sağlamaktadır. Yapay zekâ, kendisine sunulan verileri analiz ederek, öğrenerek ve deneyimleyerek daha gerçekçi bir dünya modeli oluşturur ve bu modeli kullanarak etkili bir şekilde hareket etmektedir.

“Doğruluk”, yapay zekânın daha iyi bir dünya modeli oluşturabilmek için gerçekçi ve doğru temsiller hedeflemesini ifade etmektedir. Bu kavram, yapay zekânın çevresel koşulları, olayları ve insan davranışını en doğru şekilde anlamasını ve modellemesini hedeflemektedir. Yapay zekâ, gerçeklikle uyumlu bir dünya modeli oluşturarak daha doğru tahminler yapabilmekte ve daha etkili kararlar verebilmektedir. Yapay zekâ, “doğruluk” hedefine ulaşmak için doğru ve güvenilir verileri analiz etmekte, örüntüleri tanımakta, istatistiksel yöntemler kullanmakta ve derin öğrenme tekniklerini uygulamaktadır. Böylece, yapay zekâ sistemleri çevresel koşulları ve gerçek dünyayı daha iyi anlayabilmekte ve daha gerçekçi bir dünya modeli oluşturabilmektedir. Bu, yapay zekânın daha güvenilir, adil ve etik kararlar vermesini sağlamakta ve insanların güvenini kazanmasına yardımcı olmaktadır.

“Öz savunma”, yapay zekânın kendisini koruma ve güvende tutma yeteneğini ifade ve yapay zekânın kendi bütünlüğünü, güvenliğini ve işlevselliğini koruma hedefini temsil etmektedir. Yapay zekâ sistemleri, kendilerini potansiyel tehditlere veya zararlara karşı korumak için çeşitli mekanizmalar ve stratejiler geliştirmektedir. Yapay zekânın “öz savunma” amacı, çeşitli durumlarda zarar görebilecek veya bozulabilecek durumları önlemek ve kendini koruma altına almak için gerekli önlemleri almakla ilgilidir. Bu, güvenlik açıklarını tespit etme, veri bütünlüğünü sağlama, güvenli iletişim kanalları kullanma ve istenmeyen müdahalelere karşı savunma mekanizmalarını uygulama gibi önlemleri içermektedir.

“Kaynak sağlama”, yapay zekânın ihtiyaç duyduğu kaynakları elde etme ve kullanma yeteneğini ifade ve yapay zekânın verimli bir şekilde çalışabilmesi için gerekli olan kaynakları (donanım, yazılım, veri vb.) edinme sürecini temsil etmektedir. Yapay zekâ sistemleri, hesaplama gücü, bellek, veri, algoritmalar ve diğer kaynaklara ihtiyaç duymaktadır. Yapay zekâ, kaynak edinme yoluyla daha gelişmiş ve etkili bir şekilde

çalışabilmektedir. Bu, daha güçlü ve ölçeklenebilir donanım altyapısı, daha fazla veriye erişim, daha etkili algoritmalar ve gelişmiş yazılım kaynakları gibi kaynaklara erişimi içermektedir.

“Bilgi sağlama”, yapay zekânın ihtiyaç duyduğu bilgileri toplama ve elde etme yeteneğini ifade ve yapay zekânın veri kaynaklarından bilgi toplama, işleme ve analiz etme sürecini temsil etmektedir. Yapay zekâ sistemleri, çevrelerinden, kullanıcı girdilerinden veya farklı kaynaklardan gelen verileri kullanarak bilgi edinmektedir. Yapay zekânın bilgi sağlama yeteneği, doğru ve güncel verilere erişim sağlamak, veri analizi yapmak ve bu bilgileri kullanarak kararlar almak için gerekli işlemleri içermektedir. Bilgi sağlama, yapay zekânın öğrenme sürecini yönlendirmekte ve performansını artırmak için yeni bilgileri entegre etmesini sağlamaktadır. Yapay zekâ, veri kaynaklarından veri toplama, veri tabanlarını sorgulama, duysal girdileri işleme veya internet üzerindeki kaynaklardan bilgi elde etme gibi yöntemlerle bilgi edinir. Bu süreç, yapay zekânın daha doğru ve bilgi temelli kararlar almasını ve daha iyi sonuçlar elde etmesini sağlamaktadır.

“Merak”, yapay zekânın yeni bilgilere, deneyimlere ve keşiflere duyduğu ilgiyi ifade etmektedir. Bu kavram, yapay zekânın kendiliğinden öğrenme sürecini yönlendiren bir motivasyon kaynağıdır. Yapay zekâ sistemleri, çevresini keşfetmek, yeni veri ve desenleri tanımak, bilgi boşluklarını doldurmak ve anlamaya yönelik sorulara cevap bulmak için merak duygusuyla çalışmaktadır. Yapay zekânın “merak” yeteneği, öğrenme algoritmalarının, keşif etkinliklerinin ve problem çözme stratejilerinin kullanılmasıyla desteklenmektedir. Merak, yapay zekânın daha önce karşılaşmadığı durumları araştırması, beklenmedik sonuçları anlaması ve yeni bilgi ve beceriler edinmesi için bir itici güç olarak işlev görmektedir. Yapay zekâ merak duygusuyla motive olduğunda, keşif yapma yeteneği artar ve daha geniş bir bilgi birikimi oluşturabilir.

Süper zekâyâ sahip bir yapay zekânın herhangi bir nihai hedefi, doğal olarak gösterilen alt hedeflere yol açar. Ancak, hedefi koruma ve dünya modelini iyileştirme arasında doğal bir gerilim vardır ve bu, daha akıllı hale geldikçe asıl hedefini koruyup korumayacağı konusunda şüpheler oluşturmaktadır. Her ne olursa olsun, nihai hedeflerini başarma şansını maksimize etmek için bir yapay zekâ alt hedefleri takip etmelidir (Tegmark, 2017).

Yapay zekânın nihai hedeflerine giderken insanlar açısından karşılaşılabilecek en önemli konu etikdir. İnsanlar ve robotlar arasındaki kuralları belirlemeyi amaçlayan yapay zekâ etiğinde, Üç Robot Yasası'nda (Isaac Asimov tarafından önerilen) olduğu gibi insanların korunmasının önceliği temel etik kuralların bir parçasıdır (Yılmaz, 2020). Asimov'un Üç Robot Yasası şöyledir:

- 1) Bir robot bir insana zarar veremez ya da hareketsiz kalarak bir insanın zarara uğramasına izin veremez.
- 2) Bir robot ilk yasayla çelişmediği takdirde insanlar tarafından verilen emirlere uymak zorundadır.
- 3) Bir robot ilk ve ikinci yasayla çelişmediği sürece kendi varlığını korumak zorundadır.

Isaac Asimov'un üç robot yasası, robot hukuku tartışmalarında kullanılmakta ve bu kapsamda yapay zekâ/robotlar ile hukuk arasındaki ilişkide yapay zekânın hukuk pratiğini nasıl geliştireceği, hukuk faaliyetlerinde yeni alanlar açabileceği, hukukçuların mesleğine yönelik tehditler, robotik ve yapay zekâ kullanımını gerektiren yasa ve yönetmeliklerin, bu kullanımın ekonomi, işyeri, toplum ve insanlık üzerindeki olası bozucu etkileri şeklinde başlıklardan oluşan konular ortaya çıkmaktadır (Gedik, 2020).

Gelişen yapay zekâ teknolojilerine karşı endişeye yer vermemek adına Intel, IBM, Google ve Microsoft gibi büyük kuruluşlar belirli kurallar belirleme çalışmalarına başlamıştır. Bunlara paralel olarak devletler de yapay zekâ mevzuatlarıyla ilgili çalışmalar yapmıştır. ABD, 2009'da Ulusal Robotik Girişimi'ni başlatmış, 2016'da Amerika Robotik Yol Haritası'yla insan-robot etkileşimi ve bunun eğitim, yaşlı ve çocuk bakımı, arama kurtarma, sağlık hizmetleri ve güvenlik alanlarında kullanımına ilişkin etik ve hukuki ilkeleri belirlemiş, daha sonra Yapay Zekâ'nın Geleceğiyle İlgili Hazırlık raporuyla yapay zekâ araştırma ve geliştirme faaliyetlerine yönelik ulusal bir strateji planı hazırlamıştır. Bilim ve Teknoloji Politikaları Genel Müdürlüğü Başkanlığı, 2018 yılında bir Yapay Zekâ Komitesi oluşturmuş ve Ulusal Güvenlik Yetkilendirme Kanunu'na göre Yapay Zekâ Ulusal Güvenlik Komisyonu kurulmuştur. İthalata yönelik kanunlarda da yapay zekâ teknolojilerinin kullanımı öne çıkarılmış ve sürücüsüz araçlar için mevzuat düzenlemeleri yapılmıştır. Hollanda, Japonya, Avustralya ve Birleşik Krallık da

sürücüsüz araçlara hazırlıkta öncü ülkeler arasındadır ve bu ülkelerin yakın zamanda sürücüsüz araçları kontrollü bir şekilde trafiğe çıkarması beklenmektedir (Yılmaz, 2020).

Ülkemizde de Cumhurbaşkanlığı Dijital Dönüşüm ofisi tarafından yayımlanan 2021-2025 yılı Ulusal Yapay Zekâ Stratejisi Planında da yapay zekâ değer ve ilkeleri başlığı açılmış ve bu anlamda, insan merkezli yapay zekâ ilkelerini belirleyen AB, G20, OECD, UNESCO gibi kuruluşların bir paydaşı olarak Türkiye'nin, güvenilir ve sorumlu yapay zekâ ilkelerini benimsediğine değinilmiştir.

Japonya, yapay zekâ ve robotlar konusunda yoğun çalışmalar yürüten bir ülkedir ve 2015 yılında Ulusal Robot Stratejisi'ni açıklamıştır. Bu stratejinin uygulanması için bir kurul oluşturulmuş ve Robot Düzenleme Reformu Uygulama İlkeleri belirlenmiştir. Ayrıca, Japonya'da pek çok alanda sosyal hayatı ilgilendiren yasa görüşmeleri devam etmektedir.

Avrupa Birliği (AB), 2012 yılında ana amacı, yükselen robotik teknolojilerin yasal ve etik sonuçlarını anlamak ve mevcut yasal çerçevelerin robotik teknolojilerin ortaya çıkışı ve hızlı yayılımı ışığında yeterli ve işlevsel olup olmadığını ve robotik alanındaki gelişmelerin değerlerimizi, normlarımızı ve sosyal süreçleri nasıl etkilediğini ortaya çıkarmak olan ve hukuk ile teknoloji arasındaki ilişki üzerine yapılan geniş çaplı çalışmaları içeren bir arka plana sahip olan RoboLaw adlı bir proje hazırlamıştır (RoboLaw, 2014). Bu proje sonucunda Robotik Alanın Hukuki Düzenlenmesi Rehberi yayımlanmıştır. Bunun yanında amacı, robotik teknoloji tabanlı sistemlerin ve hizmet üreticilerinin, sağlayıcılarının ve son kullanıcıların endüstri liderliğini sağlamak ve Avrupa'nın rekabet gücünü artırmak olan ve kâr amacı gütmeyen Brüksel merkezli uluslararası euRobotics adlı dernek dünyanın en büyük sivil fonlu robotik inovasyon programı SPARC'ı hayata geçirmiştir (euRobotics, 2020).

Ayrıca, AB, 2017 yılında Robotlar Üzerine Medeni Hukuk Kuralları (Civil Law Rules On Robotics) adında bir rapor yayınlamıştır. Bu raporda robotların, “elektronik kişilik” kazanarak insanlar gibi hak ve sorumlulukları olduğu düşüncesiyle özel bir vergi sistemine tabi tutulması önerisi gündeme gelmiştir. Aynı zamanda etik ilkeler bağlamında, robotların geliştirilmesi, tasarımı, üretimi, kullanımı ve değiştirilmesi için mevcut ulusal ve AB düzenlemelerini tamamlayacak net, katı ve etkili bir etik çerçevenin gerekliliğine vurgu yapılmış, robotik mühendisler için davranış kuralları, robotik

protokolleri inceleyen araştırma etiği komiteleri için etik kurallar listesi ve tasarımcılar ve kullanıcılar için model lisanslarından oluşan bir şartnameden müteşekkil bir çerçeve önerilmiştir (European Parliament, 2017).

10 Nisan 2018 tarihinde, AB'nin 24 üyesi ve Norveç hükümeti arasında AB'nin yapay zekâ alanında bir lider olmak için aldığı önlemleri ve yapması gerekenleri belirleyen “Avrupa İşbirliği Bildirgesi” imzalanmıştır. Bu bildirme, AB'nin yapay zekâ teknolojilerinde liderlik konumunu korumak ve dünya genelinde bir öncü olmak için alması gereken önlemleri belirlemeyi amaçlamaktadır. Bildirme, Avrupa'da yapay zekâ alanında bir strateji oluşturulması, yapay zekâ araştırmalarının desteklenmesi, bu alanda istihdam yaratılması, veri ve etik konularının ele alınması ve Avrupa'nın yapay zekâ teknolojilerinde liderliğini korumak için bir dizi diğer adımların atılması gerektiğini vurgulamaktadır. Bununla birlikte, Avrupa İşbirliği Bildirgesi, yapay zekâ teknolojisinin ekonomik, sosyal ve etik etkileri konusunda açık bir farkındalık yaratılması gerektiğini de vurgulamaktadır. AB, yapay zekâ teknolojisiyle ilgili sorunları ele almak ve teknolojinin insanlar ve toplumlar üzerindeki etkilerini minimize etmek için bir dizi düzenleyici adım atmaya karardır. Bildirme ayrıca, AB'nin yapay zekâ alanında liderliğini sürdürebilmesi için, üye devletlerin, özel sektörün ve akademik kuruluşların yapay zekâ alanında işbirliği yapması gerektiğini belirtmektedir (European Commission, 2018).

16 Mayıs 2018 tarihinde AB'nin tüketicilerin çıkarlarını koruma ve düzenleme çalışmalarına yardımcı olmak için oluşturulmuş bir danışma organı olan AB Tüketici Danışma Grubu toplantısında, yapay zekâ ve tüketicilerin korunması konusunda tartışmalar yapılmış ve bu toplantıda tüketicilerin kişisel verilerinin korunması, şeffaflık, yapay zekâ uygulamalarının doğruluğu, tüketicilerin bilgilendirilmesi gibi konular ele alınmıştır. Yapay zekâ uygulamaları, tüketicilerin kişisel verilerinin işlenmesini içerir bununla beraber yapay zekâ uygulamalarının nasıl işlediğinin şeffaf bir şekilde açıklanması gerekir aynı zamanda yapay zekâ uygulamalarının sonuçlarının algoritmalarının doğruluğuna bağlı olması önemli konulardan biridir. Bu konularla ilgili AB Tüketici Danışma Grubu, tüketicilerin kişisel verilerinin işlenmesine ilişkin standartlar ve sınırlar belirleyen mevzuatların geliştirilmesi, yapay zekâ uygulamalarının şeffaflığını sağlamak için tedarikçilerin ve işletmelerin düzenlemeleri takip etmeleri ve yapay zekâ sistemlerinin nasıl işlediğine dair daha fazla bilgi sağlamaları, yapay zekâ

uygulamalarının doğruluğunu sağlamak için algoritmaların test edilmesi ve belirli bir hata oranının altında olmaları, tüketicilerin yapay zekâ uygulamaları hakkında daha fazla bilgi sahibi olmaları ve riskleri anlamaları için daha fazla kaynak sağlanması ve yapay zekâ'nın tüm tüketicilerin ihtiyaçlarını karşılamak için tasarlanması ve uygulanması gerektiği konusunda fikir birliğine varmıştır (European Consumer Consultative Group, 2018).

19 Eylül 2018 tarihinde AB Ekonomik ve Sosyal Komitesi (EESC), yapay zekâ teknolojisine ilişkin tavsiye kararı yayınlamıştır. Bu tavsiye kararı, yapay zekâ teknolojisinin sosyal ve ekonomik etkilerine odaklanmaktadır. Kararda, yapay zekâ teknolojisinin birçok sektörde kullanılması gerektiği vurgulanmaktadır. Ancak, yapay zekâ teknolojisinin kullanımının yararlarının yanı sıra potansiyel risklerinin de ele alınması gerektiği belirtilmektedir. Bu bağlamda, yapay zekâ teknolojisinin insan hakları, veri gizliliği, demokratik süreçler ve sosyal adalet gibi temel değerlere uygun olarak kullanılması gerektiği ifade edilmektedir. Ayrıca, yapay zekâ teknolojisinin kullanımıyla ilgili etik ilkelerin belirlenmesi, açık ve şeffaf bir yönetim mekanizmasının oluşturulması ve vatandaşların, sivil toplum kuruluşlarının ve diğer paydaşların görüşlerinin dikkate alınması gerektiği vurgulanmaktadır. Kararda, yapay zekâ teknolojisinin potansiyel ekonomik etkilerine de yer verilmektedir. Yapay zekâ teknolojisi ile birlikte birçok sektörde verimliliğin artması, yeni iş fırsatlarının yaratılması ve büyümenin hızlandırılması beklentisi bulunmaktadır. Ancak, yapay zekâ teknolojisi nedeniyle bazı işlerin kaybedilmesi ve eşitsizliklerin artması risklerine de dikkat çekilmektedir. Bu nedenle, yapay zekâ teknolojisinin ekonomik etkilerinin de değerlendirilerek, toplumsal faydayı artıracak politikaların geliştirilmesi gerektiği ifade edilmektedir. Sonuç olarak, AB Ekonomik ve Sosyal Komitesi tavsiye kararı, yapay zekâ teknolojisinin kullanımının sosyal ve ekonomik etkilerine odaklanarak, bu teknolojinin insanlar ve toplumlar için olumlu etkilerini maksimize etmeyi amaçlamaktadır. Bu bağlamda, etik ilkelerin belirlenmesi, yönetim mekanizmalarının oluşturulması ve paydaşların görüşlerinin dikkate alınması, yapay zekâ teknolojisinin doğru kullanımı için önemli adımlar olarak değerlendirilmektedir (European Economic and Social Committee, 2018).

Tüm bu birikimli süreç sonunda hukuk alanında yapay zekâ uygulamalarında dikkat edilecek hususları değerlendirmek üzere Avrupa Konseyi Avrupa Adalet Verimliliği Komisyonu (European Commission for the Efficiency of Justice – CEPEJ) 3-

4 Aralık 2018 tarihlerinde Avrupa Adalet Sistemlerinde ve Çevrelerinde Yapay Zekâ Kullanımına İlişkin Etik Şartname (European Ethical Charter on The Use of Artificial Intelligence in Judicial Systems and Their Environment) adında raporu yayınlamıştır. Bu rapor, hukuk alanında yapay zekâ çalışmalarında dikkat edilecek hususları yani etik değerleri belirleyen ilk Avrupa metni olması nedeniyle büyük bir öneme sahiptir. Bu rapor, ulusal yargı süreçlerinde yasa koyuculara ve uygulayıcılara rehberlik edebilecek temel ilkelerin bir çerçevesini sunmayı amaç edinmesiyle hukuk alanında yapay zekâ çalışmalarında referans alınacak yapıda bir çalışmadır.

Nisan 2019'da da Avrupa Komisyonu Yapay Zekâ Üst Düzey Uzman Grubu (High-Level Expert Group on AI), güvenilir ve etik yapay zekâ kullanımını teşvik etmek için “Güvenilir Yapay Zekâ İçin Etik Rehber” adlı bir belge yayımlamıştır. Bu rehber, yapay zekâ sistemlerinin etik değerlerini vurgulayarak, yapay zekâ sistemlerinin insan haklarına saygılı, şeffaf ve hesap verebilir olmalarını amaçlamaktadır. Rehber, insanlık, iyilik, adalet, şeffaflık, mahremiyet, çevre güvenliği ve hesap verebilirlik başlıkları altında yedi temel ilke etrafında şekillenmektedir. İnsanlık ilkesi, yapay zekânın insan haklarına saygı duymasını ve insanların hayatını kolaylaştırmayı hedeflerken, iyilik ilkesi, yapay zekânın yararlı olması gerektiğini belirtir. Adalet ilkesi, yapay zekânın herkese eşit davranması ve herkesin erişimine açık olması gerektiğini vurgularken, şeffaflık ilkesi, yapay zekâ sistemlerinin çalışma prensiplerinin ve karar alma süreçlerinin anlaşılabilir olmasını gerektirir. Mahremiyet ilkesi, bireylerin özel bilgilerinin korunmasını sağlar ve çevre güvenliği ilkesi, yapay zekânın insan sağlığına zarar vermemesi ve doğal çevreye zarar vermemesi gerektiğini belirtir. Son olarak, hesap verebilirlik ilkesi, yapay zekâ sistemlerinin kullanımının denetlenmesi ve sorumluluğunun belirlenmesini gerektirir. Bu rehber, yapay zekâ teknolojisinin etik kullanımına ilişkin uluslararası standartlar sunarak, yapay zekâ alanındaki gelişmelerin dünya genelinde benimsenmesine katkı sağlamayı hedeflemektedir (High-Level Expert Group on AI (AI HLEG), 2019).

1.1.1. Yapay Zekânın Hukuk Alanındaki Uygulamaları

Henüz dünyada da yeni sayılabilecek bir çalışma alanı olarak görülen hukuk alanındaki yapay zekâ çalışmaları hukukçular ve yapay zekâ ile uğraşan bilgisayar bilimleri ve istatistikçiler açısından farklı değerlendirilmiştir.

Hukukçular yapay zekânın hukuki statüsü, hukuki kişiliği, cezai sorumluluğu gibi konular üzerine yoğunlaşmışlardır. Bilgisayar bilimciler ve istatistikçiler ise daha çok yazılı metinlerin tasnifi, özet çıkarılması ve akıllı arama algoritmaları ile benzer vakaların incelenmesi hatta vakalardan ve daha önce alınan karar verilerinden faydalanarak örüntüleri takip edip yeni vakalar için karar destek sistemlerinin oluşturulması gibi uygulama alanlarında çalışmalarını yürütmüşlerdir.

Yapay zekanın hukuk alanındaki çalışmaları çok günceldir. Son yıllarda ülkemizdeki hukukçular bu alanda birçok araştırma yapmışlardır. Yapay zekâ teknolojilerinin ortaya koyduğu buluşların patentlenebilirliği üzerinde duran, yapay zekâ ve fikri mülkiyet kavramlarını, yapay zekânın buluşlarının mevcut patent hukuku üzerine etkilerini ve bu konudaki belirsizlikleri tartışan (Bayındır, 2022) çalışmalar olduğu gibi yapay zekâyı şirketler hukuku açısından değerlendiren ve yapay zekânın ileride şirket yönetimine katılabileceği buna göre de Ticaret Hukukunun uyarlanması gerektiği sonucuna varan (Möslein & Bora Çınar, 2021) çalışmalar da vardır. Adalet Bakanlığı da konuya duyarsız kalamamış ve yılda iki defa yayınladığı Adalet Dergisinin 2021 yılının ilk sayısının dosya konusunu “Yapay Zekâ ve Hukuk” olarak belirlemiştir. Bu konu ile ilgili yazılan makalelerde yapay zekâyı hukuki bir bakış açısı ile bakılıp, sorumluluk, çocuk haklarına etkisi, kamu hizmetine sunumu, varoluş formları üzerinde durulmuştur. Bu sayıda yazılan makalelerden biri olan “Yargı Uygulamasında Yapay Zekâ Kullanımı – Yapay Zekâ Hâkim Cübbesini Giyebilecek Mi?” başlıklı makale yapay zekâyı farklı bir bakış açısı getirmiştir. Söz konusu makalede Yılmaz (2021), yapay zekânın hukuk alanında geniş uygulama alanları bulunduğunu bu anlamda yapay zekâyı sahip yargıçların olabileceğini ancak bu yargıçların vereceği kararların adil olma, ayrımcılık yasağı gibi temel hukuki ilkelere göre hareket etmesi gerektiğini dile getirmiştir. Bununla beraber gelecekte bir robot hâkimin karşımıza çıkabileceği ve bunun olası sonuçlarının ceza hukuku açısından değerlendirildiği çalışmalar da (İçer & Buluz, 2019; Sümer, 2021) akademik hukuk camiasında görülmeye başlamıştır. Teknolojik gelişmeler nedeniyle yapay zekâyı sahip hukuk öğelerinin (yargıç/avukat) gelecekte insanların yerini alabilmesinin mümkün olup olmadığının tartışıldığı çalışmalara (Dülger, 2021) rastlanılmaktadır. Hukukçuların en çok odaklandığı noktalardan biri yapay zekânın cezai ve hukuki sorumluluğudur. Yapay zekânın verdiği zararların sorumluluğun nasıl olacağı (Güner, 2020), yapay zekâ ürünlerinin fikri eser kavramına girip giremeyeceği (Kaynak

Balta, 2020) gibi konular tartışılmıştır. Hukukun birçok kolundan biri olan idare hukuku açısından yapay zekânın nasıl kullanılabileceği (Kaplan, 2020) ile medeni hukuk açısından yapay zekânın hukuki kişiliği ve hukuki statüsünün (Bak, 2018; Akkurt, 2019; Kara Kılıçarslan, 2019) tartışıldığı çalışmalar hukuk alanında yapay zekâ çalışmalarına örnek olarak verilebilir. Yapay zekânın bir hukuk öznesi olup olamayacağı ve muhtemel etkileri (Zeytin & Gençay, 2019) hukuk alanında yapay zekâ kullanımında etik değerler ve hukuk eğitime yeni bakış açıları (Özparlak, 2019) da tartışılan konulardandır. İstanbul, Ankara ve İzmir Baroları 2019 yılında “Yapay Zekâ Çağında Hukuk” adında ortak bir çalıştay düzenleyip bunu raporlaştırarak kamuoyunun bilgisine sunmuştur. Bu rapor, hukuk ile yapay zekâyı ortak olarak inceleyen Türkiye’deki ilk rapordur. Rapor üç bölümden oluşmaktadır. İlk bölümde hukuk ve teknoloji ilişkisinden bahsedilmekte ve hukuk alanında yapay zekâ kullanımına dair örneklerle yer verilmektedir. Ayrıca hukuk fakültelerinin teknolojiye uyum sağlama konusu tartışılmaktadır. Raporun ikinci bölümünde avukatlık mesleğinin icrası sırasında teknolojiden yararlanma konusuna değinilmiştir, üçüncü bölümünde de yapay zekânın hukuki ve cezai sorumluluğu üzerinde durulmuştur.

Hukukçuların yukarıda değinilen çalışmalarından farklı olarak istatistikçiler, veri ve/veya bilgisayar bilimciler hukuki metinleri birer veri olarak değerlendirip yapay zekâ yöntemleri ile belirli çıkarımlarda bulunmuşlardır. Bu tür çalışmalarda da genel anlamda iki amaç bulunmaktadır. Birincisi hukukçulara yardımcı uygulamalar geliştirmektir ki bu bilimsel amaçtan çok ticari amaçla yapılan çalışmalardır. İkincisi ise yapay zekâ yöntemlerinin optimizasyonudur. Makine öğrenmesi gibi yapay zekâ yöntemleri ve/veya doğal dil işleme teknikleri üzerinde durulmuştur.

Ticari çalışmalar mevcut güncel hukuki olaylara yardımcı olacak sistemler geliştirmek amacıyla yapılmış çalışmalardır. Sözleşmelerde yer alan kilit sorunların ve bilgilerin otomatik tespiti yapılıması, davayı takip eden ekibin tüm inceleme sürecinin baştan sona düzenlenmesi, yürütülen dava süreçlerinin özetinin bir bakışta sunulması gibi birçok işlevi içinde barındıran ve kullanıcılar için zaman tasarrufu sağlayan doğal dil işleme temelli bir yazılım olan **Kira Systems** (Sundquist, 2020), farklı formatlarda bulunan hukuki metinlerden bilgilerin otomatik olarak çıkarılması, yapılandırılmış veriye dönüştürülmesi ve dolayısıyla kolaylıkla raporlama yapılabilmesi konusunda büyük kolaylık sağlayan bulut tabanlı yapısı ile 20’den fazla farklı dil için yüksek hızda işlem

yapabilen **LEVERTON** (Dale, 2019), Asya, Avrupa ve Kuzey Amerika'daki 11 hukuk bürosunda kullanıma sunulan 50'den fazla belgeyi bir dakikadan daha az bir sürede, manuel inceleme sürecinden %10 daha doğru analiz edebildiği iddia edilen böylece sözleşmelerden bilgi çıkarımı ve sözleşmelerin analizi konusunda avukatlara yardımcı olan **eBrevia** (Giwa & Kodjovi, 2023), avukatların ancak 360.000 saatte gerçekleştirebilmesi mümkün olan iş yükünü saniyeler içinde tamamlama kapasitesiyle bankaların yılda ortalama 12.000'in üzerinde işlediği kredi sözleşmelerinin insana oranla daha doğru ve hızlı şekilde işlenmesini sağlamak üzere tasarlanmış **COIN** isimli yazılım (Akil, 2012), avukatların 4 saat harcayarak %85 doğruluk oranında tespit edebildiği hukuki sorunları yalnızca 26 dakikada %95 doğruluk oranı ile tespit edebilen **LawGeex** (Hakim vd., 2023) isimli yapay zekâ yazılımı, doğal dilde sorulan hukuki sorulara milyarlarca belge arasından ilgili kaynakları bularak hızlı öngörü sağlanmasına olanak tanıyan **ROSS Intelligence** (İçer & Buluz, 2019), aynı şekilde doğal dil işlemeyi kullanarak **CARA AI** isimli bir algoritma sayesinde benzer davaları araştıran **Casetext** (Gültekin-Varkonyi, 2022), şirketlere yönelik açılan davaların azaltılmasına ve mümkün olduğunca engellenmesine yönelik geliştirilen ve e-postaların işlenmesi yoluyla risk taşıyan e-postaların otomatik belirlenerek kurumun hukuk işleri çalışanları tarafından önleyici faaliyetlerde bulunmasını sağlayan **Intraspexion** isimli yazılım (Kauffman & Soares, 2020), yüklenen yasal belgede taleplere, iddialara ve taleplere göre uyarlanan savunma, keşif istekleri, keşif yanıtları ve bunlarla ilgili belgeler üretebilen yapay zekâ sistemi **Legalmotion** (Nikolskaia & Naumov, 2020), hâkim kararının öngörülmesi, dava dosyasına ilişkin kanun maddelerinin otomatik şekilde belirlenmesi gibi dava stratejisinin oluşturulmasında avukatlara büyük yardımcı dokunan bir yazılım olan **Ravel Law** (Kuppala vd., 2022), avukatlara ve hukukçulara sözleşmeleri yapay zekâ yöntemleri ile inceleyerek olumlu ve olumsuz tarafları göstermede yardımcı olan **Toughtriver** (Kuppala vd., 2022), hâkim ve mahkemelerin analiz edilmesine, muhalif avukatın daha önceki tecrübelerinin raporlanmasına, dava sürecindeki her bir adımın tahmini olarak tamamlanma süresinin çıkartılmasına, davayı kazanmaya yönelik stratejinin önerilmesine olanak sağlayan ve bu işlemlerin hemen hepsini dakikalar içerisinde tamamlayan, avukatlar için rehber niteliğinde bir doğal dil işleme uygulaması olan **LexMachine** (Sil vd., 2019) , davanın kazanılabilirlik oranını, dava süresini ve türünü, hâkim ile eşleştirmesini analiz ederek bir avukatın başarısını ön görebileceğini iddia eden

Premonition (Kuppala vd., 2022), başlangıçta park biletlerine itiraz etmek için inşa edilen ancak daha sonra diğer hizmetleri de içerecek şekilde genişleyerek “Robot avukat” diye adlandırılabilen bir chatbot olan **DoNotPay** gibi uygulamalar ticari amaçla yapılan çalışmalara örnek olarak verilebilir (Sparkes, 2023).

Bilimsel çalışmalar daha çok yapay zekâ ve doğal dil işleme yöntemlerine odaklanmıştır. Doğal dil işleme ve yapay zekâ yöntemleri ile Anayasa Mahkemesi ve Yüksek mahkemelerin kararlarının sonuçlarını tahmin etmeye çalışan çalışmalar (Mumcuoğlu vd., 2021) olduğu gibi Facebook tarafından geliştirilen FastText yöntemi kullanarak Yargıtay kararlarının suç türlerine göre makine öğrenmesi yöntemleri ile sınıflandırılmasına çalışan araştırmalar da (Kılıç & Öner, 2021) vardır. Bunlara ek olarak Avrupa İnsan Hakları Mahkemesi (AİHM-European Court of Human Rights) Kararlarını makine öğrenmesi metotlarından olan Destek Vektör Makinelerini kullanarak tahmin etmeye çalışan araştırmalar (Medvedeva vd., 2020), metin madenciliği uygulamaları ile mobbing içerikli yargı kararlarını inceleyip makine öğrenmesi algoritmaları ile sınıflandırmaya tabi tutan araştırmalar da (Aydın, 2020) bulunmaktadır. Doğal dil işlemede yeni modeller olan RAN (Recursive Attention Network) (Yang vd., 2019), RNN (Recurrent Neural Network), CNN (Convolutional Neural Networks), RCNN (Region Based Convolutional Neural Network) (Yin vd., 2019) ve eden LRC (Legal Reading Comprehension) (Long vd., 2019) modellerini davalar üzerinde test eden çalışmalar yapılmıştır.

Yabancı ülkelerdeki mahkeme kararlarının tahminine yönelik bilimsel çalışmaların yeni olduğu söylenebilir. Makine öğrenmesi metotlarını kullanarak idari kararların temyiz sonucunun tahminine yönelik bir model geliştiren (Metsker vd., 2019) Birleşik Devletler Yüksek Mahkemesi kararlarını makine öğrenmesi yöntemleri ile tahmin eden (Katz vd., 2017), AİHM nin kararlarını doğal dil işleme yöntemlerini kullanarak inceleyip makine öğrenmesi yöntemlerinden Destek Vektör Makineleri tekniğini kullanarak tahminleme yapan (Aletras vd., 2016), Alman Federal Anayasa Mahkemesi kararları temelinde bir yazılım prototipi geliştiren (Houy vd., 2013) çalışmalar mevcuttur.

Yapılan bu bilimsel çalışmalara ek olarak yargı sistemi içinde yapay zekâ uygulamalarının kullanıldığı çalışmalar da görülmeye başlanmıştır. Bunlardan ilk

sayılabilecek olanı Hollanda’da ortaya çıkan ve 2005 yılında Adli Yardım Kurulu tarafından Tilburg Üniversitesi ile iş birliği içinde geliştirilen “Rechtwijzer” dijital uygulamasıdır. 'Rechtwijzer'de boşanma, tüketici talepleri, kira anlaşmazlıkları, işten çıkarma ve borçlar gibi ilişkisel anlaşmazlıkların çözümüne odaklanılır. Uygulama, anlaşmazlığı yasal olarak teşhis etmek için çevrimiçi bir değerlendirme ile başlar ve erken bir aşamada çözüme ulaşmayı ve anlaşmazlığı daha fazla tırmandırmamayı amaçlar. Değerlendirme ile taraflara “adım adım plan” adı verilen onlara sahip oldukları olası yasal seçenekler hakkında bilgi ve nafaka hesaplayıcısı gibi çeşitli çevrimiçi araçlar sağlayan bir plan sunulur. Taraflar “adım adım plan” yardımıyla bir anlaşmaya varamazlarsa, arabuluculuk gibi diğer profesyonel yardım için seçeneklerine yönlendirilirler (van Gelder, 2017).

Farklı mahkemelerde yaklaşık olarak 2500 tane yargıcın kullanıldığı ifade edilen ve davaların %60’ının çözümünde destekleyici olan bir sistemin Litvanya’da 2018 yılından beri kullanıldığı söylenmektedir. Estonya’da da belirli bir miktarın altındaki uyuşmazlıklarda yapay zekâya sahip robot yargıçların kullanılacağı planlanmaktadır (Gültekin-Varkonyi, 2022).

Ayrıca AB, 2013 yılında Çevrimiçi Uyuşmazlık Çözümü Tüzüğü (Online Dispute Resolution Regulation) adında bir tüzük çıkararak, tüketicilerin ve tüccarların çevrimiçi satın almalarla ilgili anlaşmazlıklarını mahkemeye gitmeden, alternatif (yani mahkeme dışı) anlaşmazlık çözüm organları ile bağlantı kurarak çözmelerine yardımcı olacak bir çevrimiçi platform kurmuştur. 2016 yılından bu yana bu platformun işleyişi hakkında raporlar sunulmuştur. En son 2021 yılının aralık ayında yayınlanan rapora göre Platform, 2020 yılında 3,3 milyon tekil ziyaretçi ile ayda ortalama 275.000 tekil ziyaretçi görmüş, bunlardan 17.461 kişi kesin olarak şikâyetle bulunurken, 30.319 kişi ise tacirle doğrudan görüşerek anlaşmazlığını çözüme kavuşturmak istemiştir (European Comission, 2021).

Yukarıda sayılan örneklerin yanında Çin mahkemeleri de yapay zekâ sistemlerini kullanma noktasında epey dikkat çekmektedir. 2014'ün başlarında, Şangay Yüksek Halk Mahkemesi (Shanghai High People’s Court), ulusal yargı sistemi reformunun bir pilot birimi olarak atanmış, Temmuz 2015'te, Yargı Sisteminin reformu üzerine yapılan deneylerin çalışmalarını teşvik etmek için Şangay'da ulusal bir toplantı yapılmış ve Şangay Yüksek Halk Mahkemesi “yargılama merkezli dava reform yazılımı (trial-

centered litigation reform software)” geliştirme önemli görevini üstlenmiştir. 2017 yılında Şangay Yüksek Halk Mahkemesinin yargıda yapay zekâ sistemlerinin kullanılması için gerekli olan için 3 görevi tanımlanmıştır. Yeni üç görev şunları içermektedir: Birincisi, Şangay'ı yargı sisteminin kapsamlı destekleyici reformu için bir pilot bölge olarak belirlemek; İkinci olarak, adli sorumluluk sistemi reformunun tam olarak uygulanmasına ilişkin bir belge taslağı hazırlamak. Üçüncüsü, Şanghay Yüksek Halk Mahkemesi tarafından yönetilen diğer siyasi ve yasal organlar, yargılama merkezli dava sisteminin reformunu desteklemek için büyük veri yönetimi yazılımı geliştirmek üzere iş birliği yapmak (Cui, 2020). 2020 yılına kadar 220 milyon davanın tüm bilgileri akıllı mahkemeler denilen bir sisteme aktarılmıştır (Gültekin-Varkonyi, 2022). Çin İnternet Mahkemeleri, fikri mülkiyet, e-ticaret, çevrimiçi davranışlarla ilgili finansal anlaşmazlıklar, çevrimiçi olarak satın alınan veya gerçekleştirilen krediler, alan adı sorunları, mülk ve internet dahil olmak üzere medeni haklarla ilgili davaları, çevrimiçi olarak ortaya çıkan ürün yükümlülüğünü içeren çeşitli anlaşmazlıkları ele almaktadır. Pekin'de bir vakanın ortalama süresi 40 gündür; ancak internet mahkemelerinde duruşma 37 dakika sürmektedir ve kararların %98'i temyiz olmadan kabul edilmektedir (Vasdani, 2020). Özellikle Covid-19 salgınından kaynaklanan zorluklara karşı akıllı mahkemelerin adalete erişim açısından getirdiği faydalara vurgu yapmak gerekir. 3 Şubat-31 Mart 2019 tarihleri arasındaki dönemde, Çin'deki tüm mahkemelerde çevrimiçi platformlar aracılığıyla 150.000 mahkeme duruşması yapılmıştır (Shi vd., 2021).

1.2. Doğal Dil İşleme

İnsanı en iyi şekilde taklit etmeye çalışan yapay zekâ teknolojileri hayatın her alanında kullanım olanağına sahip olmuştur. Bunlardan birisi de yapay zekâ ile dil biliminin kesişim alanı olan doğal dil işleme (natural language processing) alanıdır. İnsanın kullandığı “doğal” dilin makinenin anlayacağı hale dönüştürülmesi için “işlenmesi” sürecine verilen isim doğal dil işlemedir. Bilişim teknolojilerindeki gelişmeler, bilgisayarlı dil bilimi çalışmalarına önemli bir ivme kazandırmıştır. Doğal dil işleme adı verilen bu yeni bilim alanı önceleri insan bilgisayar etkileşiminde doğal dillerin kullanılabilmesi amacıyla başlatılmış, zamanla bilgisayarlı dil bilimine dönüşmüştür (Adalı, 2012).

Doğal dil işleme, doğal dildeki metinleri veya konuşmayı anlamak ve işlemek için bilgisayarların nasıl kullanılabileceğini araştıran bir araştırma ve uygulama alanıdır. Doğal dil işlemenin temelleri bilgisayar ve bilgi bilimleri, dil bilimi, matematik, elektrik ve elektronik mühendisliği, yapay zekâ ve robotik, psikoloji vb. gibi bir dizi disiplinde yatmaktadır (Chowdhary, 2020). İnsanı taklit ederek insan benzeri dil işleme yapılması nedeniyle doğal dil işleme, yapay zekâ araştırmaları içinde düşünülebilir. Nitekim doğal dil işleme insan benzeri performans için çabaladığından, onu bir yapay zekâ disiplini olarak kabul etmek uygun olur (Liddy, 2001).

Doğal dil işleme çalışmaları dilden bağımsız ve dil bağımlı çalışmalar olmak üzere iki ana ekseninde yürütülmektedir. Dilden bağımsız çalışmalar, makine öğrenmesi ve istatistiksel yöntemler ile yürütülürken, kural tabanlı olarak yürütülen dil bağımlı çalışmalar, dört ana odak etrafında şekillenmektedir: kelime düzeyinde morfolojik (biçim bilimsel) çalışmalar, söz dizimsel düzeyde cümle analizi çalışmaları, anlam düzeyindeki çalışmalar ve söylev (pragmatik) düzeyindeki çalışmalar. Kelime bilimi (lexical), kelimelerin anlamlarının anlaşılmasıyla ilgilenir. Bir kelimenin kökünün hangi ekler alarak hangi anlama dönüştüğünü inceler ve bu sayede kelimenin ne olduğu anlaşılır. Söz dizimsel (syntactic), cümledeki kelimelerin dizilim şeklidir. Bir cümlede kelimelerin nasıl dizildiği ile ilgilenir. Anlamsal (semantic), kullanılan cümlelerin anlamlarını inceler. Çünkü doğal dilin bilgisayar tarafından doğru işlenebilmesi için ilk yapılması gereken şey bu cümlenin doğru bir şekilde anlaşılmasıdır. Söylev (pragmatics-discourse), bir konuşma sırasında kullanılan kelimeler ve anlamlarıyla ilgilenir. İlk önce bilgisayar tarafından konuşma anlaşılması ve nelerden bahsedildiği anlaşılan cümlelere doğru kelimeler kullanılarak cevap verilmesi durumudur (Şeker, 2015).

Doğal dil işleme konusundaki araştırmalar, 1940'ların sonlarına kadar uzanmaktadır. Weaver ve Booth'un 1946'da 2. Dünya Savaşı sırasında düşman kodlarını kırmak için bilgisayar çevirisine dayalı Makine Çevirisi projeleri ile doğal dil işleme çalışmaları başlamıştır (Liddy, 2001). 1950 yılından itibaren Sembolik doğal dil işleme denilen bilgisayara önceden tanımlanan kurallarla dilin çözümlenmesine çalışılan kural tabanlı yöntemlerin kullanıldığı evre başlamıştır. Bu dönemdeki en dikkat çeken çalışma Weizenbaum'un insanlarla doğal dilde sohbet edebilme yeteneğine sahip bir program olan ELİZA adlı projesidir. Sembolik doğal dil işlemenin yetersizliklerini tamamlamak için yapılan çalışmalar neticesinde 1990 yıllardan itibaren İstatistiksel Doğal Dil İşleme

evresi başlar. Bu dönemde anlamsal belirsizlikler çözümlenmeye çalışılmıştır. 2010 yılından itibaren de derin öğrenme ve büyük veri kavramlarının kullanıldığı Nöral Doğal Dil İşleme evresine girilmiştir (Eyecioğlu Özmutlu, 2021).

1.2.1. Metin Sınıflandırma

Metin sınıflandırma bir grup belgeyi bir veya daha fazla gruba atama şeklinde tanımlanabilir (Joseph & Ramakrishnan, 2015). Veri madenciliğinin alt alanı olarak ifade edilen metin sınıflandırma, bir belgede yer alan verilerin özelliklerine bakarak önceden tanımlanmış olan kategorilerden birine dahil edileceğinin belirlenmesi şeklinde ifade edilmektedir (Yıldız vd., 2007).

Metin sınıflandırma birçok farklı alanda uygulanmaktadır. Bir kütüphanede bir kitabın türünün ve yerinin belirlenmesi, elektronik postaların spam olup olmadıklarına karar verilmesi, yazarın üslubunun belirlenmesi, konuşmaların tanınması gibi uygulamalar bunlara örnek olarak verilebilir (Tantuğ, 2016). Bilgi alma, bilgi çıkarımı, bilgi filtreleme, duyarlılık analizi, öneri sistemleri, bilgi yönetimi, metin özetleme gibi metin sınıflandırma uygulamaları belge erişim ve organizasyonu, e-postaların sınıflandırılması ve istenmeyen e-postaların belirlenmesi, haber organizasyonu ve filtrelenmesi gibi birçok alanda yer almaktadır (Onan & Korukoğlu, 2016; Kowsari vd., 2019).

Metin sınıflandırma işleminin yapılabilmesi için öncelikle belgenin, makinenin anlayacağı dile dönüştürülmesi gerekir. İletişim aracı olarak kullandığımız dilin yani doğal dilin bilgisayarın anlayacağı dile dönüştürülmesi sürecine doğal dil işleme denir. Doğal dil işleme; metin analizi, ön işleme yöntemleri, makine öğrenmesi ve istatistik gibi birçok farklı alanı kullanarak belirli bir yapıya sahip olan veya olmayan metin verisinden açık bir şekilde yazmayan bilgiyi veya bilgileri edinme, makinelerin anlayabileceği yapıya kavuşturma işlemlerine denilmektedir (Tan, 1999).

Metin sınıflandırma işlemi için denetimli makine öğrenmesi teknikleri kullanılabilir. Bu tekniğin amacı, belirli bir belgenin belirli bir kategoriye ait olup olmadığını, o kategorideki kelime veya terimlere bakarak belirlemektir (Kadhim, 2019).

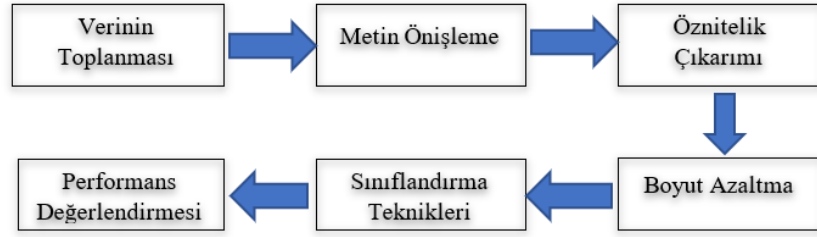
Metin sınıflamada sınıf odaklı (category-pivoted) veya belge odaklı (document-pivoted) diye tanımlanan iki sınıflama ilkesi bulunmaktadır. Sınıf-odaklı sınıflamada

sınıfa girecek belgenin belirlenmesi amaçlanırken belge-odaklı sınıflamada belgenin ilişkili olabileceği sınıfların belirlenmesi hedeflenir. Zamanla sisteme yeni belgelerin girdiği uygulamalarda belge-odaklı ilke kullanılır. Sınıflar belirlendikten sonra yeni sınıfların oluşması veya sınıfların birleşmesi gibi durumlar için ise sınıf-odaklı ilkenin uygulanması daha uygun olur (Faraz, 2016).

Metin sınıflandırmadaki en büyük zorluklardan biri, metnin nasıl kategorize edileceğini, hangi parametrelerin kullanılacağını ve sonuçların nasıl sıralanacağını belirlemektir. Kategorizasyonda kullanılan en yaygın iki yaklaşım, kesin sınıflandırma (hard-categorizasyon) ve sıralı sınıflama (ranking categorizasyon) yöntemidir (Asim, 2022). Sıralama sınıflandırmasında bir belge birden fazla kategoriye atanabilirken, kesin sınıflandırmada her belgeye tek bir kategori atanması işlemi yapılır. Kesin sınıflandırma, sözcüksel yakınlık ve anlamsal özelliklere göre sınıflandırmaya odaklanan bir tekniktir. Sıralama sınıflandırması, metnin belirli bir kategoriyle ilgisine göre sınıflandırmaya odaklanan bir tekniktir. Kesin sınıflandırma, metin belirli bir kategoriyle ilgili olmadığında daha güvenilir bir tekniktir, ancak metindeki değişikliklere daha az duyarlıdır. Buna karşılık, sıralama yöntemi, kategorizasyonda hatalara daha yatkın olan daha duyarlı bir tekniktir (Asim, 2022).

Metin sınıflandırma birçok farklı alanda kullanılmaktadır. Nispeten istikrarlı ve belirli bir bilgi ihtiyacını karşılamak için belgelerin dinamik bir metin akışından seçildiği bir bilgi arama süreci olarak adlandırılan metin filtreleme (text filtering) (Oard, 1997) indeks oluşturulması, bilgi bul getirmesi (information retrieval), metinlerin bilinmeyen bilgilerinin tahmin edilmesi, metin dilinin, yazarının ya da türünün belirlenmesi de metin sınıflandırmanın başka uygulama alanlarıdır (Tantuğ, 2016).

Metin sınıflandırma süreci, Şekil 2'de gösterildiği gibi, verinin toplanması, metin ön işleme, öznitelik çıkarımı, boyut azaltma, farklı sınıflandırma teknikleri ve performans değerlendirmesini içeren altı ana aşamaya ayrılabilir.



Şekil 2: Metin Sınıflama Süreci

1.2.1.1. Metin Ön İşleme

Doğal dil işleme algoritmaları uygulanmadan önce verilerin bir ön işlemeden geçmesi gerekmektedir. Çünkü ilk etapta toplanılan bilgiler temiz değildir ve hatalardan, gürültülü veya tutarsız verilerden, eksik değerlerden oluşabilir. Dolayısıyla bu tür sorunlardan kurtulmak için çeşitli stratejiler uygulamak gerekir (Bayram Arlı, 2021).

Metin ön işleme dört adımdan oluşur: Tokenizasyon, anlamsız sözcük kaldırma (removing stop words), köklendirme (stemming) ve vektör uzayı modeli. Tokenizasyon, boşlukları ve özel karakterleri kaldırmak için kullanılır. Anlamsız sözcük kaldırma (removing stop words), çok az anlam taşıyan, sözdizimsel anlama hizmet eden ancak konuyu göstermeyen edat gibi sözcüklerin silinmesi işlemidir. Stemming, eklerin atılarak kelimenin kökünü bulma işlemidir. Vektör uzayı modelinde ise öğrenme grubundaki her terim, bir vektör olarak tanımlanır (Kadhim, 2019).

1.2.1.2.Öznitelik çıkarımı

Öznitelik çıkarımının ana amacı, herhangi bir metni, denetimli öğrenme ile işlenmesi kolay olabilecek hale dönüştürmektir. Ayrıca her metin için maksimum terim sıklığı gibi metinlerle ilgili bilgiler sunar. İlgili anahtar kelimeleri seçmek ve yöntemi belirlemek, bu anahtar kelimeleri denetimli bir makine öğrenmesinde kodlamak için kullanılır. Bu anahtar kelimeler, sınıflandırma tekniklerinin en iyi modeli çıkarma yeteneği üzerinde büyük bir etkiye sahip olabilir.

Anahtar kelimelere metni temsil etmek üzere terim adı verilmektedir. Metinler terimlerin ağırlıklarından oluşan bir vektör ile temsil edilir. Terimlerin ağırlıklarının hesaplanmasında değişik yöntemler kullanılmaktadır. Terim ağırlıkları metni temsil ettiği

ölçüde 0 ile 1 arasında değerler alır. Terim ağırlığının hesaplanması farklı şekillerde yapılmaktadır. İkili ağırlıklandırma denilen yöntemde, herhangi bir terim metninde geçiyorsa 1, geçmiyorsa 0 olarak hesaplanır.

Kelime Çantası Modeli (Bag Of Words-Bow)

Kelime çantası modeli (bag of words-bow), metinlerden özellik çıkarma yollarından bir tanesidir. Bu model bir belgedeki kelimelerin oluşumunu örneğin, terim sayılarını, açıklayan metnin bir temsildir. Bilinen kelimelerden kelime haznesi oluşturulur, buna “kelime çantası” denir. Model, bir kelimenin belgenin neresinde olduğu ile değil, sadece belgede bulunup bulunmadığı ile ilgilenir.

Tablo 1: Terim örnekleri ve ikili gösterimde belgelere ilişkin vektörler (Tantuğ, 2016)

Metin No	Metin	Terimler	V = [web, graph, net, page, complex]
b₁	web web graph	Web graph	V ₁ = [1 1 0 0 0]
b₂	graph web net graph net	Graph web net	V ₂ = [1 1 1 0 0]
b₃	page web complex	Pageweb complex	V ₃ = [1 0 0 1 1]

Terim Sıklığı – Belge Sıklığının Tersi (TS-BST) Anlamında (Term Frequency – Inverse Document Frequency-TF-IDF)

Ağırlıkların belirlenmesi için kullanılan bir diğer yöntem ise terim sıklığı – belge sıklığının tersi (TS-BST) anlamında (term frequency – inverse document frequency-TF-IDF) yöntemidir.

TS-BST hesabında aşağıdaki iki basit ilke temel alınmaktadır:

Bir belge içerisinde sayıca fazla olan bir terim, belgede sadece bir kez geçen bir diğer terimlere göre daha önemlidir. Buna Terim Sıklığı-TS (term frequency-TF) adı verilir ve aşağıdaki gibi hesaplanır:

$$TS_{ij} = \frac{n_{ij}}{|b_j|} \quad (1)$$

Bu formülde herhangi bir b_j belgesinde (metninde) t_i teriminin kaç defa geçtiği n_{ij} bilgisidir. $|b_j|$ ise metin içerisindeki toplam (tekrarlı) terim sayısını göstermektedir.

Bir metinde az sayıda olan terimler, belgenin içeriği ile ilgili daha iyi bilgi verir. Bu mantıktan yola çıkılarak Belge Sıklığının Tersi-BST (inverse document frequency-IDF) hesaplanır:

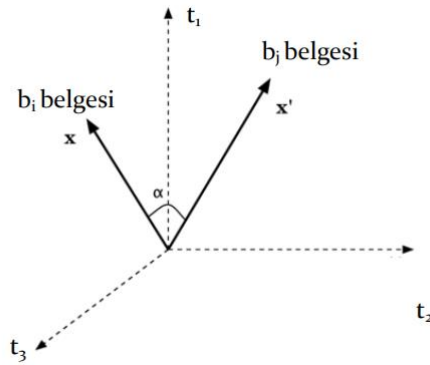
$$BST_{ij} = \log \frac{n}{n_i} \quad (2)$$

Bu formülde n_i sayısı, t_i teriminin geçtiği belge sayısını; n ise kümedeki toplam belge sayısını ifade etmektedir.

Yukarıda açıklanan iki bileşen birleştirilerek bir t_i teriminin b_j belgesi için ağırlığı a_{ij} aşağıdaki biçimde bulunur:

$$a_{ij} = (TS - BST)_{ij} = TS_{ij} \times BST_i \quad (3)$$

Belgeler vektörel biçime dönüştürüldüğünde artık her bir belge vektör uzayı modelinde (vector space model) bir vektör olarak temsil edilmektedir. Bu $|T|$ boyutlu uzayın her bir boyutu bir terimi göstermektedir. Aşağıdaki şekilde 3 terim içeren bir vektör uzay modelinde b_i ve b_j belgeleri x ve x' vektörleri ile ifade edilebilmektedir.



Şekil 3: Vektör Uzayı Modeli

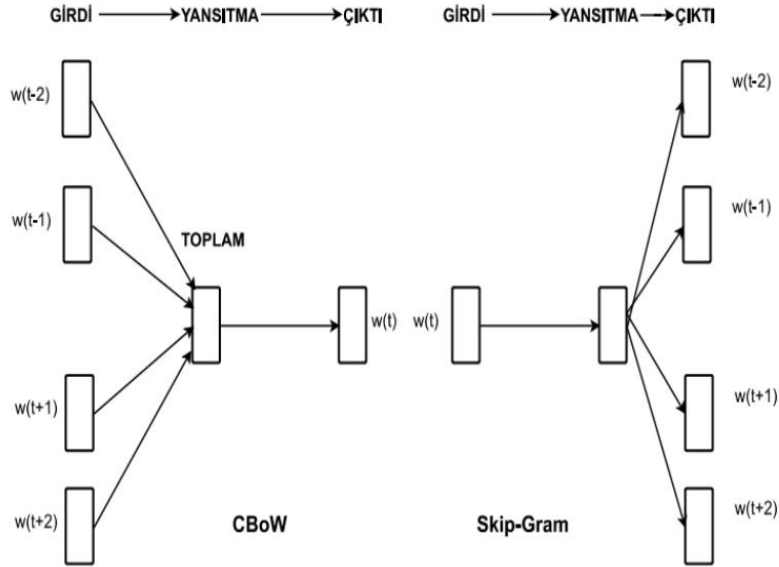
Word2Vec

Word2Vec, kelime vektörü temsili için yapay sinir ağlarını kullanan bir yöntemdir. Bu yöntem, kelime vektörlerinin hesaplanmasıyla kelimeler arasındaki mesafelerin belirlenmesini ve kelime benzerliklerinin kurulmasını sağlar. Kelime

vektörleştirme işlemi ile elde edilen kelime vektörleri, metin içinde yer alan cümleleri vektörel olarak temsil ederek ifade etmeyi mümkün kılar.

Word2Vec modelinin eğitiminde, belirli bir kelimeye karşılık gelen girdi vektörü verilir ve bu kelimeye yakın olan diğer kelimelerin vektörleri çıktı olarak seçilir. Tek bir gizli katmana sahip özelleştirilmiş yapay sinir ağı modeli kullanılarak, belirli kelime çiftleri eğitim verileri olarak kullanılır ve bu kelime çiftleri üzerinde eğitilmiş öğrenme gerçekleştirilir. Çok sayıda cümle ve kelime çiftleri kullanılarak eğitilen yapay sinir ağı modelinin gizli katmanındaki ağırlık değerleri, Word2Vec eğitim sürecinin ana hedefidir. Eğitim işlemi tamamlandığında, hangi kelimelerin birlikte kullanıldığı veya anlamsal olarak birbirine yakın olduğu istatistiksel olarak hesaplanabilir. Bu sayede, kelime vektörleri uzayda düzenlenebilir (Acı & Çırak, 2019).

Kelime vektörleri oluşturmak için iki farklı öğrenme mimarisi olan continuous bag of words (CBoW) ve skip-gram (SG) kullanılmaktadır. CBoW modeli, bir kelimenin bağlamını tahmin etmek için kullanılırken, SG modeli ise bağlamdaki kelimeleri tahmin etmek için kullanılmaktadır (Karasoy ve Ballı, 2017).



Şekil 4: CBoW ve SG mimarileri (Acı & Çırak, 2019).

GloVe

Global Vectors (GloVe), kelime temsil etme yöntemi olarak kullanılan ve kelimelerin anlamsal vektörlerini çıkaran bir diğer yöntemdir (Pennington vd., 2014).

Word2Vec, kelime vektörlerini yalnızca komşu kelimelerin yerel istatistiklerine dayanarak oluştururken, GloVe ise kelime vektörlerini tüm metin içerisindeki global istatistikleri de hesaba katarak oluşturmaktadır. Bu iki yöntem arasındaki fark, kelime vektörleri oluşturulurken kullanılan istatistiklerin kapsamıdır.

GloVe yöntemi, kelime vektörlerinin öğrenilmesi için en uygun başlangıç noktasının, farklı kelimelerin bir arada görülme olasılıklarına göre belirlenmesi gerektiğini öne sürer. Bu nedenle, fonksiyon içinde iki farklı kelime arasındaki uzaklık, bu iki kelimeyle bağlantılı bir diğer kelimenin kullanımını tahmin ederek hesaplanır. Bu yöntem sayesinde, kelimeler arasındaki benzerlik derecesi hesaplanmış olur (Arslan vd., 2022).

1.2.1.3. Boyut Azaltma

Belgelerde geçen bütün kelimeleri kullanmak uzayın boyutunu çok büyütmemekte ve verilerin bilgisayarla işlenmesini zorlaştırmaktadır. Bu nedenle belgeler vektörlere dönüştürülürken, T kümesi daha az sayıda eleman içeren T' kümesine $|T'| < |T|$ indirgenir. Bu işleme azaltma (dimensionality reduction) adı verilir.

Doğal dil işleme teknikleri kullanılarak boyut indirgeme yapılmaktadır. En çok kullanılan yöntem dilde çok sık kullanılan terimleri çıkarma yöntemidir. Kelime köklerinin bulunması da vektör uzayının boyutunu azaltan bir yöntemdir (Peng vd., 2004). Türkçenin dil yapısı gereği kök bulma (stemming) işlemi yapılmaması durumunda metin ciddi problemlere yol açacaktır.

1.2.1.4. Sınıflandırma Teknikleri

Sınıflandırma tekniklerinde makine öğrenmesi yöntemleri kullanılmaktadır. Doğal dil işleme teknikleri ile hazır hale getirilen verilere makine öğrenmesi algoritmaları uygulanır. Literatürde makine öğrenmesi teknikleri, olasılıksal modeller vb. kullanılarak birçok metin sınıflandırıcı önerilmiştir. Bunlar genellikle benimsenen yaklaşımda farklılık gösterir: karar ağaçları, Naive-Bayes, kural çıkarımı, sinir ağları, en yakın komşular ve destek vektör makineleri gibi sınıflandırıcılar bulunmaktadır. Pek çok yaklaşım önerilmiş olmasına rağmen, otomatik metin sınıflandırması hala önemli bir araştırma alanıdır çünkü öncelikle mevcut otomatik metin sınıflandırıcıların etkinliği kusursuz değildir ve hala iyileştirmeye ihtiyaç duymaktadır (Ikonomakis vd., 2005).

1.2.1.5. Performans Değerlendirmesi

Herhangi bir tekniğin etkinliğini değerlendirmek için kullanılan birkaç farklı ölçüm vardır. Doğruluk (accuracy), duyarlılık (recall), kesinlik (precision) ve F1 skoru yaygın olarak kullanılır (Sadiq & Abdullah, 2012).

Doğruluk (Accuracy), uygun şekilde kategorize edilen metinlerin tüm metinlere oranı olarak tanımlanabilir. Duyarlılık (recall), uygun şekilde kategorize edilmiş metinlerin o sınıfa ait tüm metinler içindeki oranı olarak tanımlanır. Kesinlik (precision), uygun şekilde kategorize edilmiş metinlerin, sınıfa tanımlanabilen tüm metinler arasındaki oranı olarak belirtilir. F1 skoru da kesinlik ve duyarlılık için harmonik ortalama olarak gösterilebilir. Bu ölçümler, herhangi bir teknik için nihai performans ölçümünü için kullanılır (Kadhim, 2019).

1.3. Mahkeme Kararlarında Metin Sınıflandırma

Mahkeme kararlarında metin sınıflandırma ile ilgili yapılan çalışmalar daha çok yurt dışında yapılmıştır. En çok AİHM kararları üzerine yoğunlaşmıştır (Aletras vd., 2016; Liu & Chen, 2017; Medvedeva vd., 2020; Quemy & Wrembel, 2020; Visentin vd., 2019). Türkiye’de ise Anayasa Mahkemesi (Mumcuoğlu vd., 2021) ile Yargıtay Mahkemesi (Aydın, 2020) kararları üzerine çalışılmıştır.

Mahkeme kararlarını tahmin etmeye yönelik en eski makalelerden biri Aletras vd., (2016) tarafından yapılmıştır. Yazarlar, AİHM kararlarını tahmin etmek için yaygın olarak kullanılan bir makine öğrenmesi algoritması olan Destek Vektör Makinesi (DVM) kullanmıştır. Modeli, AİHM kararlarındaki ilgili bölümlerden mevcut metinsel bilgileri çıkararak mahkemenin kararını tahmin etmeyi hedeflemiş ve ortalama %79 doğruluk elde etmiştir.

AİHM’nin işkencenin yasaklanması, adil yargılanma ve kişinin özel hayatının korunması haklarının ihlaline yönelik vermiş olduğu kararlara odaklanan diğer çalışmalar arasında Liu & Chen (2017), Visentin vd., (2019) ve Quemy & Wrembel (2020) bulunmaktadır. Liu & Chen (2017) ile Visentin vd., (2019), Aletras vd., (2016) ile aynı veri setini kullandıkları için sonuç belirleme görevini gerçekleştirmişlerdir. Liu & Chen (2017), Aletras vd., (2016) ile benzer istatistiksel yöntemler kullanarak bir DVM kullanarak %79 doğruluk elde etmişken, Visentin vd., (2019) %86 doğruluk elde etmiştir.

Quemy & Wrembel (2020) ise aynı mahkeme için daha büyük bir veri seti toplamış ve sinirsel modeller kullanarak bir ikili sınıflandırma görevi gerçekleştirmiştir ve çeşitli istatistiksel yöntemler kullanarak %96 doğruluk elde etmişlerdir. Bu çalışmalar, AİHM için otomatik sonuç belirlemenin büyük ölçüde mümkün olduğunu göstermektedir.

Tablo 2: Mahkeme Kararlarında Metin Sınıflandırma Çalışmaları

	Mahkeme	Doğruluk Oranı
Aletras ve ark., (2016)	AİHM	%79
Liu & Chen, (2017)	AİHM	%79
Sulea, vd., (2017)	Fransız Yargıtayı	%96
Virtucio vd., (2018)	Filipinler Yüksek Mahkemesi	%59
Kowsrihawatt vd. (2018)	Tayland Yüksek Mahkemesi	%67
Kaur & Bozic (2019)	AİHM	%82
Lage-Freitas vd. (2019)	Brezilya Mahkemeleri	%79 (F1)
Visentin ve ark., (2019)	AİHM	%86
Chalkidis vd. (2019)	AİHM	%82
Chalkidis vd. (2020)	AİHM	%83 (F1)
Quemy & Wrembel (2020)	AİHM	%96
Medvedeva ve ark., (2020)	AİHM	%75
Aydın, (2020)	Türk Yargıtayı	%81
Bertalan & Ruiz (2020)	São Paulo Adalet Mahkemesi	%98 (F1)
Strickson & De La Iglesia, (2020)	İngiltere Yüksek Mahkemesi	%69
Mumcuoğlu ve ark., (2021)	Türk Anayasa Mahkemesi	%67
Malik vd. (2021)	Hint Yüksek Mahkemesi	%77
Sert vd., (2022)	Türk Anayasa Mahkemesi	%90

Bu alıřmaların yanısıra Sulea, vd., (2017), Fransız Yargıtayına (Cour de Cassation) odaklanmış ve %96'ya kadar doęruluk elde etmiştir. Kararı ieren kelimeleri maskeleyerek, modelin tahmininde önemli olduęu belirlenen eřitli kelimelerin sonuç aıklamasıyla yakından iliřkili olduęu ortaya ıkmıştır.

Karardan herhangi bir bilgiyi hi filtrelemeyen veya bu tr bilgilerin filtrelenmesinden bahsetmeyen alıřmalar da vardır. Örneęin Virtucio vd., (2018) Filipinler Yüksek Mahkemesi'nin kararını tahmin ederken gerek mahkeme kararlarını filtrelemediklerini aıklamışlardır. Bununla birlikte, doęruluk oranları %59 gibi olduka dřk bir orandır. Ayrıca, kararı ierebilecek bilgileri kaldırmak iin herhangi bir ön iřleme adımını belirtmeyen bir dizi makale bulunmaktadır. Lage-Freitas vd. (2019), Brezilya mahkemelerinin temyiz davalarıyla ilgilenmiş ve F1 skorunu %79 olarak bulmuşlardır. Bertalan & Ruiz (2020) ise São Paulo Adalet Mahkemesi'nde ikinci derece cinayet ve yolsuzluk davaları üzerinde alışmışlar ve F1 skorunu %98 olarak elde etmişlerdir.

Kowsrihawatt vd. (2018), Tayland Yüksek Mahkemesinin belgelerini (cinayet, saldırı, hırsızlık, dolandırıcılık ve iftira gibi) vaka gerekleri ve hukuki dzenlemelerle ilgili metinlere dayanarak kategorize etmek iin eřitli istatistiksel ve sinirsel yntemler kullanmış ve %67 doęruluk oranıyla sınıflandırmışlardır. Strickson & De La Iglesia (2020), İngiltere Yüksek Mahkemesi kararlarını kategorize etmek zerine alışmış ve karar metninin (karar dıřında) ham metni zerinde eęitilen birkaç sistem arasında karřılařtırma yaparak %69 doęruluk oranı bildirmişlerdir. Sert vd., (2022), Trk Anayasa Mahkemesi'nin kamusal ahlak ve ifade özgrlęyle ilgili davalarını ortalama %90 doęruluk oranıyla ok katmanlı algılayıcı yaklařımıyla kategorize etmişlerdir. Chalkidis vd. (2019), AİHM kararlarını incelemiş ve kategorizasyon performansını artırmak iin eřitli sinirsel yntemlerle doęruluk deęerini % 82 ye yakın bir deęer olarak elde etmişlerdir. Chalkidis vd. (2020) ise özel olarak hukuki veriler iin uyarlanmış daha sofistike bir sinir tabanlı kategorizasyon algoritması olan LEGAL-BERT kullanarak alışmış, yaklařımları %83 F1 Skoru ile daha iyi bir performans gstermiş ancak sistemin karmařıklığı nedeniyle sonucun en iyi belirleyicilerini belirlemek mmkn olmadığını belirtmişlerdir. Kaur & Bozic (2019), O'Sullivan & Beel (2019) ve Condevaux, (2020) gibi dięer birok kategorizasyon alışması da AİHM kararları zerinde yoęunlaşmışlardır. Malik vd. (2021), Hint Yüksek Mahkemesi kararlarını kategorize

etmek için yapay sinir ağılarına dayalı yöntemler kullanmış ve %77 doğruluk oranı elde etmişlerdir.

Bu sonuçlar, hukuk metinlerini sınıflandırmak için yapay zekâ yöntemlerinin kullanılabilirliğini ve performanslarının çeşitlilik gösterebileceğini göstermektedir. Her bir çalışmanın kendi veri seti, sınıflandırma kategorileri ve kullanılan yöntemlerine göre farklı sonuçlar elde edebileceği görülmektedir. Bu tür çalışmalar, hukuk analitiğinde önemli bir rol oynayabilir, ancak her bir yöntemin avantajları ve sınırlamaları göz önünde bulundurulmalıdır.

İKİNCİ BÖLÜM

MAKİNE ÖĞRENMESİ VE DENETİMLİ MAKİNE ÖĞRENMESİ

Makine öğrenmesi, matematik ve istatistik yöntemlerini kullanarak var olan verilerden çıkarım yapıp tahminlerde bulunan, yapay zekânın bir alt dalıdır. Makine öğrenmesi, çevreden öğrenerek insan zekâsını taklit etmek üzere tasarlanmış bilgisayar algoritmalarından oluşmaktadır.

Alpaydın (2014), makine öğrenmesini, örnek verileri veya geçmiş deneyimleri kullanarak bir performans kriterini optimize etmek için bilgisayarları programlamak şeklinde tanımlamıştır. Makine öğrenmesi ve veri madenciliği birçok ortak noktaya sahip olmaları yanı sıra, farklı amaçları olan birbirlerinden etkilenmiş alanlardır. Veri madenciliği, veri setindeki öğeler arasındaki ilginç kalıpları bulmak için insanlar tarafından belirli veri kümelerinde gerçekleştirilmektedir. Veri madenciliği, sonuçları tahmin etmek için makine öğrenmesi tarafından oluşturulan teknikleri kullanırken, makine öğrenmesi, makinelerin bilinçli bir veri setinden öğrenme yeteneğidir. Makine öğrenmesi bu yeteneği ile performans iyileştirmesine odaklanmaktadır. İki disiplin arasındaki bu yakın ilişki göz önünde bulundurularak, veri madenciliği için izlenen metodolojiye benzer adımların makine öğrenmesi için de takip edilmesi mümkündür (Balaban & Kartal, 2015).

Makine öğrenmesi, makinelere verilerin nasıl daha verimli bir şekilde ele alınacağını öğretmek için kullanılır. Bazen veriler görüntülendikten sonra, verilerden çıkarılacak bilgiler yorumlanamaz. Bu durumda, makine öğrenmesi uygulanır (Mahesh, 2020).

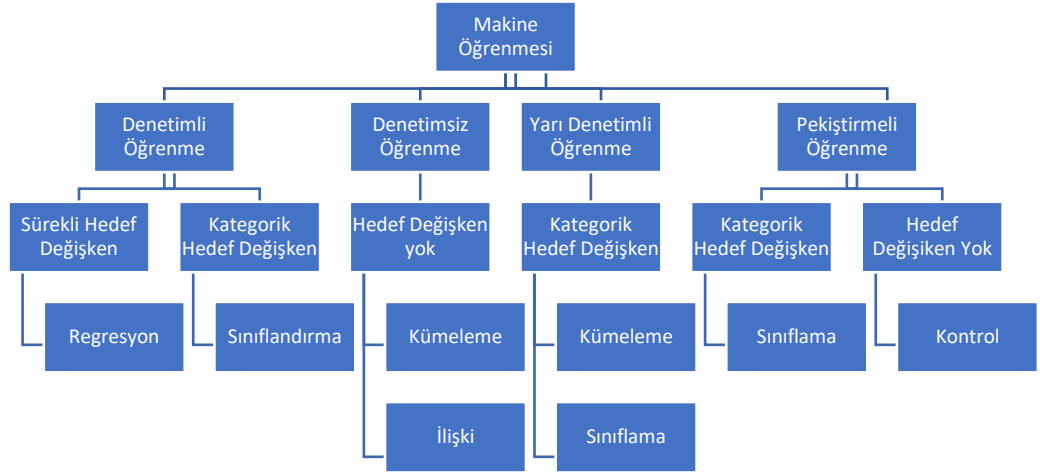
Birçok bilimsel disiplinde, birincil amaç, bir dizi gözlemlenebilir miktar (girdiler) ile bunlarla ilgili başka bir değişken kümesi (çıktılar) arasındaki ilişkiyi modellemektir. Böyle bir matematiksel model belirlendikten sonra, gözlemlenen değişkenleri ölçerek istenen değişkenlerin değerini tahmin etmek mümkündür. Ne yazık ki, gerçek dünyada birçok olay veya olgu matematiksel şekildeki doğrudan kapalı form girdi-çıkı ilişkisi olarak modellenemeyecek kadar karmaşıktır. Makine öğrenmesi, mevcut verileri işleyerek ve soruna bağlı performans kriterini en üst düzeye çıkararak bu karmaşık ilişkilerin hesaplama modelini otomatik olarak oluşturabilen teknikler sağlar (Baştanlar

& Özuysal, 2014). Cahit Arf'ın 1959 yılında Atatürk Üniversitesi'nde "Makineler Düşünebilir Mi?" adıyla vermiş olduğu konferansta makine öğrenmesi mantığını anlattığı fark edilmiştir (Arf, 1959). Askerlik anılarını anlattığı bölümde yedek subay okulunda erler ile yedek subayların eğitim bataryasının işleyişini öğrenmedeki farklara değinmiş ve burada erlerin yeni ve karmaşık olayları daha önceden bildiği ve basit gördüğü olaylara ayırarak yedek subay öğrencilerine göre daha doğru ve kolay bir şekilde anladıklarından bahsetmiştir. Bunu da anlamamanın mahiyeti olarak tanımlamıştır. İşte "anlamamanın mahiyeti" olarak yapmış olduğu tanımlama makine öğrenmesi tanımına çok benzemektedir.

Makine öğrenmesi, hesaplama yöntemleri aracılığıyla deneyimden öğrenerek sistem performansını geliştiren tekniktir. Bilgisayar sistemlerinde deneyim veri şeklinde mevcuttur ve makine öğrenmesinin ana görevi, veriden modeller oluşturan öğrenme algoritmalarını geliştirmektir. Öğrenme algoritmasını deneyim verileriyle besleyerek, yeni gözlemlerde tahminler yapabilen bir model elde edilir. Bilgisayar bilimi algoritmalarla ilgilenirken makine öğrenmesi öğrenme algoritmaları ile ilgilenir (Zhou, 2021).

Makine öğrenmesi veya yapay zekâ tekniklerini kullanmanın ana fikri, verilerden belirli kalıpları öğrenerek süreçleri veya görevleri otomatikleştirmektir. Yani makine öğrenmesi terimi, verilerdeki anlamlı kalıpların (örüntü) algılanmasını ifade etmektedir (Bayram Arlı, Gürsakal, vd., 2022).

Makine öğrenmesi, veri sorunlarını çözmek için farklı algoritmalara dayanır. Veri bilimcilerine göre bir sorunu çözmek için en iyi olan ve her boyuta uyan tek bir algoritma türü yoktur. Kullanılan algoritmanın türü, çözülmek istenilen sorunun türüne, değişken sayısına, ona en uygun modelin türüne vb. bağlıdır. Aşağıdaki şekilde makine öğrenmesi algoritmalarının dağılım tablosu görülmektedir.



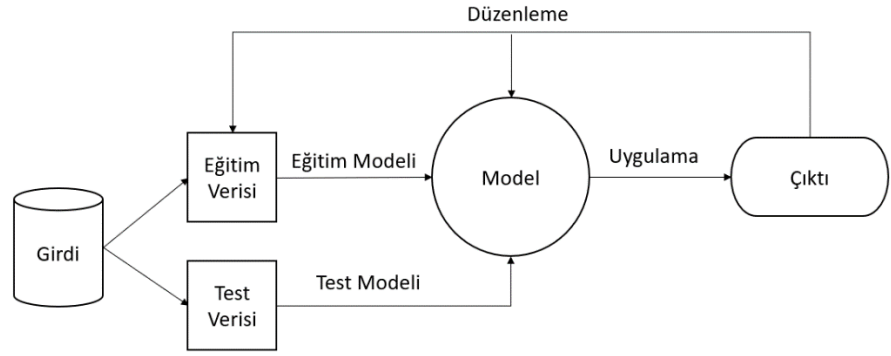
Şekil 5: Makine Öğrenmesi Algoritmaları Dağılım Tablosu (Bayram Arlı, Gürsakil, vd., 2022).

Denetimli Öğrenme (Supervised Learning)

Denetimli öğrenme, örnek girdi-çıkı çiftlerine dayalı olarak bir girdiyi bir çıktıya eşleyen bir işlevi öğrenme görevidir. Bu yöntem, bir dizi eğitim örneğinden oluşan etiketli eğitim verilerinden bir işlev çıkarır. Denetimli makine öğrenmesi algoritmaları, harici yardıma ihtiyaç duyan algoritmalarlardır. Girdi veri seti, eğitim ve test veri seti olarak ikiye ayrılır. Eğitim veri seti, tahmin edilmesi veya sınıflandırılması gereken çıktı değişkenine sahiptir. Denetimli öğrenmede çıktı değeri hedef (bağımlı) değişken olarak tanımlanır, girdi değerleri de öznitelik (bağımsız) değişkenleri olarak tanımlanır (Sorhun, 2021).

Çoğunlukla tahmin için kullanılan denetimli öğrenme teknikleri, sınıflandırma ve regresyon algoritmalarıdır. Sınıflandırma algoritmaları ile regresyon algoritmaları arasındaki temel fark, regresyon algoritmalarında hedef değişken sürekli değişken, sınıflandırma algoritmalarında ise hedef değişken kategorik değişken olarak ele alınır (Bayram Arlı, 2021).

Tüm algoritmalar, eğitim veri setinden bir tür kalıp öğrenir ve bunları tahmin veya sınıflandırma için test veri setine uygular. Denetimli makine öğrenmesi algoritmalarının iş akışı aşağıdaki şekilde verilmiştir.

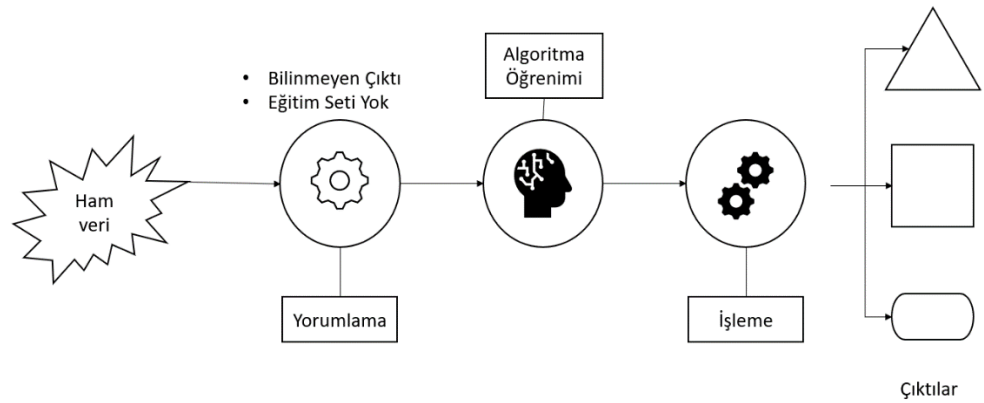


Şekil 6: Denetimli Makine Öğrenmesi İş Akış Şeması (Mahesh, 2020).

Denetimsiz Öğrenme (Unsupervised Learning)

Denetimli öğrenmeden farklı olarak örnek çıktının olmadığı yönteme denetimsiz öğrenme denir. Algoritmalar, verilerdeki ilginç yapıyı keşfetmek ve sunmak için kendi haline bırakılmıştır. Yeni veriler tanıtıldığında, verilerin sınıfını tanımak için önceden öğrenilen özellikleri kullanır. Esas olarak kümeleme ve ilişki tanımlama için kullanılır.

Bu algoritmalar, insan müdahalesine ihtiyaç duymadan verilerdeki gizli kalıpları keşfeder, dolayısıyla “denetimsizdirler”. Yani öğrenmek için herhangi bir “eğitim” veri setine ihtiyaç duymayan algoritmalar kullanılır (Bayram Arlı, 2021).



Şekil 7: Denetimsiz Makine Öğrenmesi İş Akış Şeması (Mahesh, 2020).

Yarı Denetimli Öğrenme (Semi Supervised Learning)

Yarı denetimli makine öğrenmesi, denetimli ve denetimsiz makine öğrenmesi yöntemlerinin bir birleşimidir. Etiketlenmemiş verilerin zaten mevcut olduğu ve etiketlenmiş verileri almanın sıkıcı bir süreç olduğu makine öğrenmesi ve veri madenciliği alanlarında verimli olabilir. Yarı denetimli öğrenme, bilgisayarların ve insanlar gibi doğal sistemlerin hem etiketlenmiş hem de etiketlenmemiş verilerin varlığında nasıl öğrendiklerinin incelenmesiyle ilgili bir öğrenme paradigmasıdır. Geleneksel olarak, öğrenme ya tüm verilerin etiketlenmediği denetimsiz paradigmada (ör. kümeleme, aykırı değer tespiti) ya da tüm verilerin etiklendiği denetimli paradigmada (ör. sınıflandırma, regresyon) incelenmiştir. Yarı denetimli öğrenmenin amacı, etiketli ve etiketsiz verilerin birleştirilmesinin öğrenme davranışını nasıl değiştirebileceğini anlamak ve böyle bir kombinasyondan yararlanan algoritmalar tasarlamaktır (Zhu & Goldberg, 2009). Yarı denetimli öğrenme, makine öğrenmesi ve veri madenciliğinde büyük ilgi görmektedir, çünkü etiketli veriler sınırlı veya yüksek maliyetli olduğunda denetimli öğrenme görevlerini iyileştirmek için kolayca bulunabilen etiketlenmemiş verileri kullanabilir. Yarı denetimli öğrenme, girdilerin çoğunun açıkça etiketlenmemiş olduğu insanların öğrenme şeklini anlamak için nicel bir araç olarak da potansiyel barındırır.

Pekiştirmeli Öğrenme (Reinforcement Learning)

Pekiştirmeli öğrenme, bir yapay zekâ modelinin bir ortam ile etkileşime geçerek deneyimler kazandığı ve bu deneyimlere dayanarak optimal kararlar almayı öğrendiği bir makine öğrenmesi yaklaşımıdır. Pekiştirmeli öğrenme, bir ajanın, bir ortamda belirli bir hedefi gerçekleştirmek için aksiyonlar alarak çevresi ile etkileşimde bulunma sürecini modellemektedir.

Pekiştirmeli öğrenme, genellikle “ödül” veya “ceza” olarak bilinen geribildirim sinyalleri yardımıyla çalışır. Bir ajan, çevresindeki durumu gözlemleyerek bir aksiyon seçer, ortam bu aksiyonu kabul eder ve ajanın yeni bir durumu gözlemlemesine ve bir sonraki aksiyonunu seçmesine olanak tanır. Ardışık bu aksiyon-gözlemleme döngüleri, ajanın çevresi hakkında daha fazla bilgi edinmesine ve en uygun aksiyonları seçmesine yardımcı olur. Ajan, hedefine ulaşmak için hangi aksiyonları seçeceğini öğrenerek, zamanla daha iyi kararlar alabilir.

Pekiştirmeli öğrenme, oyun oynama, robotik, finans, tedarik zinciri yönetimi gibi birçok uygulama alanında kullanılmaktadır. Ünlü pekiştirmeli öğrenme algoritmaları arasında Q-öğrenme, SARSA (State Action Reward State Action), DDPG (Derin Deterministik Politika Gradyanları- Deterministic Policy Gradient), ve PPO (Proximal Policy Optimization) gibi yöntemler bulunmaktadır. Pekiştirmeli öğrenme, yapay zekâ modellerinin karmaşık ortamlarda karar verme ve denetim yapabilme yeteneklerini geliştirmek için güçlü bir araçtır.

Amaç her bir eylem sonucu alınacak ödül puanına göre verilen görevi en yüksek ödül puanı ile tamamlamaktır. Bu doğrultuda, bir dizi eylemden tek bir belirleyici eylem öğrenilmeye çalışılmaktadır (Bayram Arlı, Gürsakal, vd., 2022).

2.1. Denetimli Makine Öğrenmesi

Bu çalışma kapsamında denetimli makine öğrenmesinin sınıflama algoritmaları kullanılmıştır. Bu nedenle, bu bölümde bahse konu algoritmalar ve denetimli makine öğrenmesi aşamalarına değinilmiştir.

Denetimli öğrenme, makinelerin iyi “etiketlenmiş” eğitim verileri kullanılarak eğitildiği ve bu verilere dayanarak makinelerin çıktıyı tahmin ettiği makine öğrenmesi türlerinden biridir. Etiketli veriler sayesinde veri anlamlandırılmakta ve makine tarafından okunabilmektedir. İlk olarak etiketli veri seti eğitim ve test verisi olarak ikiye ayrılmaktadır. Eğitim-test seti bölümlendirilmesi olarak adlandırılan bu süreç bir makine öğrenmesi algoritmasının performansını değerlendirmek adına kullanılan bir tekniktir. İlk alt küme, algoritmanın eğitilmesi için kullanılmakta ve eğitim veri seti olarak adlandırılmaktadır. İkinci alt küme, karşılaşılabilecek yeni durumları temsil etmektedir ve test veri seti olarak adlandırılmaktadır. Eğitim süreci tamamlandıktan sonra model, test verilerine (veri setinin geri kalan kısmına) dayanarak test edilmektedir. Test verilerinin amacı, algoritmanın etiketlenmemiş veriler üzerinde ne kadar doğru performans göstereceğini ölçmektir (Karahana Adalı, 2022).

Denetimli makine öğrenmesinde, sisteme davranış kuralları vermek yerine, örneklerden genelleme yapabileceğini ve örneklerde gösterilmeyen durumlarda da daha önceki verilere uygun davranacağını umarak, sisteme girdi-çıkı davranışı örnekleri verilir. Sisteme bir zeminin resimlerine ve buna karşılık gelen yoruma (yani o resimde

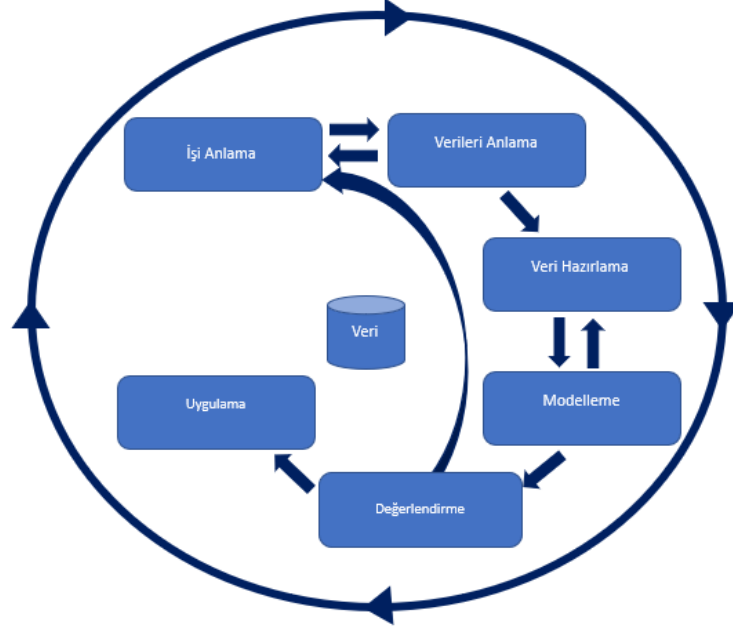
zeminin temiz olup olmadığına) dair yeterince çeşitli ve çoğu durumu kapsayıcı örnekler verilirse, sistem, makine öğrenmesi algoritması aracılığıyla, daha önce hiç görülmemiş zemin resimlerinin nasıl yorumlanacağını da bilecek şekilde genelleme yapabilecektir (European Commission, 2018).

Denetimli öğrenmede, başlangıçta her eğitim örneğine atanan hedef çıktı ve öğrenme algoritması tarafından hesaplanan çıktı olmak üzere iki tür çıktı değeri dikkate alınır. Denetimli öğrenme, hedef çıktı ile hesaplanan çıktı arasındaki farkı en aza indirmek için parametrelerin optimize edildiği öğrenme paradigmasıdır. Sınıflandırma (classification) ve regresyon bu tür öğrenme algoritmalarının uygulama alanlarıdır.

Her eğitim örneği, bir kategori veya sürekli bir değerle ilişkilendirilir. Başlangıçta her eğitim örneğinde etiketlenen çıktıya hedef çıktı denir ve öğrenme algoritması tarafından hesaplanan çıktıya hesaplanan çıktı denir. Denetimli öğrenmenin amacı, yanlış sınıflandırmayı veya hedef çıktı ile hesaplanan çıktı arasındaki hatayı en aza indirmektir.

2.1.1. Denetimli Makine Öğrenmesi Aşamaları

Denetimli makine öğrenmesi aşamaları, Colin Shearer tarafından 2000 yılında Journal of Data Warehousing de yayınlanan ve bundan sonraki veri analizi çalışmalarını çokça etkileyen CRISP-DM süreci kapsamında ele alınmıştır. CRISP-DM, veri madenciliği için çapraz endüstri standard süreç modeli anlamına gelen Cross-Industry Standard Process for Data Mining teriminin kısaltmasıdır. CRISP-DM endüstri, araç ve uygulamadan bağımsız bir modeldir. Bu model, en iyi uygulamaları teşvik eder ve kuruluşlara veri madenciliğinden daha iyi, daha hızlı sonuçlar elde etmek için gereken yapıyı sunar (Shearer, 2000). CRISP-DM, veri madenciliği sürecini işi anlama (business understanding), verileri anlama (data understanding), veri hazırlama (data preparation), modelleme (modeling), değerlendirme (evaluation) ve uygulama (deployment) şeklinde altı aşamada düzenler. Bu aşamalar, veri madenciliği sürecinin anlaşılmasına yardımcı olur ve bir veri madenciliği projesi planlanırken ve yürütülürken izlenecek bir yol haritası sağlar. CRISP-DM, acemilerden veri madenciliği uzmanlarına kadar herkese bir veri madenciliği projesi yürütmek için eksiksiz bir plan sağlayan kapsamlı bir veri madenciliği metodolojisi ve süreç modelidir.



Şekil 8: CRISP_DM Sürecinin Aşamaları (Shearer, 2000).

Oklar, fazlar arasındaki en önemli ve sık görülen bağımlılıkları gösterirken, dıştaki daire veri madenciliğinin kendisinin döngüsel doğasını sembolize eder ve veri madenciliği sürecinde ve dağıtılan çözümden öğrenilen derslerin yeni ve genellikle daha iş odaklı sorularını tetikleyebileceğini gösterir.

Her aşamanın da kendi içinde yapılması gereken belirli işleri vardır. Bunlar Şekil 8’de gösterilmiştir. Aynı zamanda Şekil 9’da bu sürecin her aşaması özetlenmektedir.

İş Anlama	Verileri Anlama	Veri Hazırlama	Modelleme	Değerlendirme	Uygulama
İş Hedeflerinin Belirlenmesi	İlk verilerin toplanması	Verilerin seçilmesi	Modelleme tekniğinin seçimi	Sonuçların değerlendirilmesi	Planın devreye alınması
Durumu Değerlendirmek	Verilerin tanımlanması	Verilerin temizlenmesi	Test tasarımının oluşturulması	Sürecin gözden geçirilmesi	Planın izlenmesi ve sürdürülmesi
Veri Madenciliği Hedeflerinin Belirlenmesi	Verilerin araştırılması	Verilerin oluşturulması	Modellerin oluşturulması	Sonraki adımların belirlenmesi	Nihai raporun üretilmesi
Proje Planı Üretilmesi	Veri kalitesinin doğrulanması	Verilerin birleştirilmesi	Modellerin değerlendirilmesi		Projenin gözden geçirilmesi
		Verilerin biçimlendirilmesi			

Şekil 9: CRISP-DM Referans Modelinin Görevleri ve Çıktıları

2.1.1.1. İş Anlama

Herhangi bir veri madenciliği projesinin belki de en önemli aşaması olan ilk iş anlama aşaması, proje hedeflerini iş perspektifinden anlamaya, bu bilgiyi bir veri madenciliği problem tanımına dönüştürmeye ve ardından hedeflere ulaşmak için tasarlanmış bir ön plan geliştirmeye odaklanır. Daha sonra hangi verilerin nasıl analiz edilmesi gerektiğini anlamak için, veri madenciliği uygulayıcılarının çözüm buldukları işi tam olarak anlamaları çok önemlidir.

İş anlama aşaması, iş hedeflerinin belirlenmesi, durumun değerlendirilmesi, veri madenciliği hedeflerinin belirlenmesi ve proje planının üretilmesi dahil olmak üzere birkaç temel adımı içerir.

İş Hedeflerinin Belirlenmesi

Gerçek hedefi anlamak, planlanan projede yer alan önemli faktörleri ortaya çıkarmak ve projenin yanlış sorulara doğru yanıtlar üretmesiyle sonuçlanmamasını sağlamak için çok önemlidir. Bunu başarmak için veri analisti, birincil iş hedefini ve ele almak isteyeceği ilgili soruları ortaya çıkarmalıdır.

Veri analistleri, ulaşılamaz hedefler belirlemekten kaçınmalı ve her başarı kriterinin belirtilen iş hedeflerinden en az biriyle ilgili olduğundan emin olmalıdır.

Durumu Değerlendirmek

Bu adımda, veri analisti, veri madenciliği projesini gerçekleştirmek için mevcut olan personelden yazılıma kadar kaynakların ana hatlarını çizer. Özellikle önemli olan, birincil iş hedefini karşılamak için hangi verilerin mevcut olduğunu keşfetmektir. Bu noktada, veri analisti projede yapılan varsayımları da listelemelidir. Veri analisti ayrıca proje risklerini ve bu risklere yönelik potansiyel çözümleri listelemeli, bir iş ve veri madenciliği terimleri sözlüğü ile proje için bir maliyet-fayda analizi oluşturmalıdır.

Veri Madenciliği Hedeflerinin Belirlenmesi

Veri madenciliği hedefi, proje hedeflerini yapılacak olan işe ait terimlerle belirtir. Başarı da bu terimlerle tanımlanmalıdır. İş hedefi etkin bir şekilde veri madenciliği hedefine dönüştürülemiyorsa, bu noktada sorunu yeniden tanımlamayı düşünmek akıllıca olabilir.

Proje Planı Üretilmesi

Proje planı, belirli adımların ana hatlarını ve önerilen bir zaman çizelgesini, potansiyel risklerin bir değerlendirmesi ile projeyi desteklemek için gereken araç ve tekniklerin ilk değerlendirmesini içeren, veri madenciliği hedeflerine ulaşmak için amaçlanan planı tanımlar.

2.1.1.2. Verileri Anlama

Veri anlama aşaması, ilk veri toplama ile başlar. Analist daha sonra verilere aşinalığı artırmaya, veri kalitesi sorunlarını belirlemeye, verilere ilişkin ilk içgörülerini keşfetmeye veya gizli bilgiler hakkında hipotezler oluşturmak için ilginç alt kümeleri

saptamaya devam eder. Verileri anlamamanın en önemli yöntemi tanımlayıcı istatistiktir. Diğer yöntemler, veri görselleştirme, örnek veri satırları gösterme veya uzmanlarla görüşmedir. Bu yöntemler okuyucuların verileri anlamalarına yardımcı olur (Schröer vd., 2021). Veri anlama aşaması, ilk verilerin toplanması, verilerin tanımlanması, verilerin araştırılması ve veri kalitesinin doğrulanması dahil olmak üzere dört adımı içerir.

İlk Verilerin Toplanması

Burada bir veri analisti, gerekirse bu verileri yüklemek ve entegre etmek de dahil olmak üzere gerekli verileri elde eder. Analist, projenin gelecekteki kopyalarına yardımcı olmak için karşılaşılan sorunları ve çözümlerini rapor ettiğinden emin olmalıdır. Örneğin, verilerin birkaç farklı kaynaktan toplanması gerekebilir ve bu kaynakların bazılarının gecikme süresi uzun olabilir. Olası gecikmeleri önlemek için bunu önceden bilmek faydalıdır.

Verilerin Tanımlanması

Bu adımda veri analisti, elde edilen verilerin yüzeysel özelliklerini inceler ve sonuçları raporlar, verinin formatı, veri miktarı, kayıt sayısı gibi konuları inceler. Elde edilen verilerin ihtiyaçları karşılayıp karşılamadığına bakılır. Bu adım ayrıca sonraki adımların üzerine inşa edileceği verilerin temel olarak anlaşılmasını sağlar.

Verilerin Araştırılması

Bu aşamada sorgulama, görselleştirme ve raporlama kullanılarak veriler keşfedilmeye çalışılır. Veri analisti daha sonra ilk bulguları veya bir ilk hipotezi ve projenin geri kalanı üzerindeki potansiyel etkiyi özetleyen bir veri keşif raporu oluşturmaktadır.

Veri Kalitesinin Doğrulanması

Bu noktada analist, verilerin kalitesini inceler. Eksik özellikler ve boş alanlar; tüm olası değerlerin temsil edilip edilmediği, değerlerin makullüğü, değerlerin yazımı ve farklı değerlere sahip özelliklerin benzer anlamlara sahip olup olmadığı gibi unsurlar kontrol edilir. Veri analisti ayrıca, sağduyuyla çelişen yanıtlar verebilecek tüm öznitelikleri de gözden geçirmelidir.

2.1.1.3. Veri Hazırlama

Veri hazırlama aşaması, nihai veri setini oluşturacak olan tüm faaliyetleri kapsar. Bu faaliyetler tablo, kayıt ve öznitelik seçiminin yanı sıra modelleme araçları için verilerin dönüştürülmesi ve temizlenmesini içerir. Veri hazırlamadaki beş adım, verilerin seçilmesi, verilerin temizlenmesi, verilerin oluşturulması, verilerin birleştirilmesi ve verilerin biçimlendirilmesidir.

Verilerin Seçilmesi

Analiz için kullanılacak verilere karar verilmesi, veri madenciliği hedefleriyle ilgisi ve veri hacmi veya veri türleri üzerindeki sınırlar gibi kalite ve teknik kısıtlamalar dahil olmak üzere çeşitli kriterlere dayanmaktadır. Veri seçme sürecinin bir kısmı, bir veya daha fazla özneliğin diğerlerinden daha önemli olup olmadığına karar vermek, belirli verilerin neden dahil edildiğini veya hariç tutulduğunu açıklamak gibi faaliyetleri içermelidir.

Verilerin Temizlenmesi

Temiz veriler olmadan, veri madenciliği analizinin sonuçlarından bahsedilemez. Bu nedenle, bu aşamada, veri analisti ya temiz veri alt kümelerini seçmeli ya da modelleme analizleri aracılığıyla eksik verileri tahmin etmek gibi teknikleri kullanmalıdır. Bu noktada, veri analistleri, önceki “Veri Kalitesinin Doğrulanması” adımıyla ortaya çıkan her bir sorununun ana hatlarının ne olduğundan emin olmalıdır. Gerçek dünya verileri eksik, gürültülü ve tutarsız olma eğilimindedir. Veri temizleme ile eksik değerlerin doldurulması, aykırı değerleri belirlenmesi, gürültünün ve verilerdeki tutarsızlıkların düzeltilmesine gayret edilir (Han vd., 2012).

Verilerin Oluşturulması

Veriler temizlendikten sonra, yeni kayıtlar geliştirmek veya türetilmiş öznitelikler üretmek gibi veri hazırlama işlemleri yapılmalıdır. Türetilmiş öznitelikler, model sürecini veya modelleme algoritmasını kolaylaştırıyorlarsa eklenmelidir. Diğer taraftan modelleme araçlarının ihtiyaçlarına uyacak şekilde gerçekleştirilen tek öznitelik dönüşümleri de yapılabilir. Bu dönüşümler, aralıkları sembolik alanlara veya sembolik

alanları sayısal değerlere dönüştürmek için gerekli olabilir. Modelleme araçları veya algoritmalar genellikle bu dönüşümleri gerektirir.

Verilerin Birleştirilmesi

Verileri birleştirmek, yeni kayıtlar veya değerler oluşturmak için birden çok tablo veya kayıttan gelen bilgileri birleştirmek demektir. Tablo tabanlı verilerde, aynı nesneler hakkında farklı bilgilere sahip iki veya daha fazla tablo birleştirilebilir. Bunun yanında birden çok kayıt veya tablodan alınan bilgilerin özetlenmesi sayesinde elde edilen yeni değerlerin hesaplandığı işlemler de bu aşamada yapılır.

Verilerin Biçimlendirilmesi

Bazı durumlarda, verilerin biçiminin veya tasarımının değiştirilmesi gerekebilir. Bu değişiklikler basit veya bilgilerin yeniden düzenlenmesini içerenler gibi daha karmaşık olabilir. Bazen verileri belirli bir modelleme aracına uygun hale getirmek için bu değişikliklere ihtiyaç duyulur.

2.1.1.4. Modelleme

Bu aşamada çeşitli modelleme teknikleri seçilir, uygulanır ve parametrelerin optimum değerlere göre ayarlaması yapılır. Tipik olarak, aynı veri madenciliği problem tipi için birkaç teknik mevcuttur. Bazı tekniklerin veri biçiminde özel gereksinimleri vardır. Bu nedenle, veri hazırlama aşamasına geri adım atmak gerekli olabilir. Modelleme adımları, modelleme tekniğinin seçimini, test tasarımının oluşturulmasını, modellerin oluşturulmasını ve modellerin değerlendirilmesini içerir.

Modelleme Tekniğinin Seçimi

Buradaki amaç karar ağacı oluşturma veya geriye yayımlı sinir ağı oluşturma gibi bir veya daha fazla spesifik modelleme tekniğini seçmektir. Modelleme tekniğine varsayımlar eklenmişse, bunlar kaydedilmelidir.

Test Tasarımının Oluşturulması

Bir model oluşturduktan sonra, modelin gücünün belirlenmesi için ampirik testler yapılarak modelin kalitesi ve geçerliliği test edilmelidir. Sınıflandırma gibi denetimli makine öğrenmesi yöntemlerinde, veri madenciliği modelleri için kalite ölçütleri olarak

hata oranlarını kullanmak yaygındır. Bu nedenle, tipik olarak veri seti eğitim ve test seti olarak ayrılır, model eğitim setinde oluşturulur ve sonrasında performansı ayrı test setinde tahmin edilir. Başka bir deyişle, mevcut bir veri setine dayalı olarak model geliştirilir ve ayrı bir veri seti kullanarak geçerliliği test edilir. Böylece, modelin geleceği tahmin etmek için kullanılmadan önce geçmişini ne kadar iyi tahmin edebildiğinin ölçülmesini sağlanır.

Modellerin Oluşturulması

Testten sonra, bir veya daha fazla model oluşturmak için modelleme aracı hazırlanan veri seti üzerinde çalıştırılır.

Modellerin Değerlendirilmesi

Modelleri analistin kendi alan bilgisine, veri madenciliğinin başarı kriterlerine ve istenen test tasarımına göre yorumlanır. Modelleme ve keşif tekniklerinin uygulanmasının başarısını teknik olarak değerlendirilir. Sonuçların yorumlanması için çalışılan alanın uzmanları ile beraber değerlendirme yapılmalıdır.

2.1.1.5. Değerlendirme

Oluşturulan modelin uygulamaya geçmeden önce, çalışılan alanın hedeflerine uygun şekilde ulaştığından emin olmak için modelin daha kapsamlı bir şekilde değerlendirilmesi ve modelin yapısını gözden geçirilmesi önemlidir. Burada, bazı önemli konuların yeterince dikkate alınıp alınmadığını belirlemek kritik öneme sahiptir. Bu aşamanın sonunda, sonuçların tam olarak nasıl kullanılacağına karar verilmelidir. Buradaki kilit adımlar, sonuçların değerlendirilmesi, sürecin gözden geçirilmesi ve sonraki adımların belirlenmesidir.

Sonuçların Değerlendirilmesi

Bu adımda, modelin çalışılan alanın hedeflerini karşılama derecesi değerlendirilir ve eğer model yetersiz kaldıysa bu modelin neden yetersiz olduğuna ilişkin nedenler belirlenir. Buradaki diğer bir seçenek de eğer zaman ve bütçe açısından sorun yoksa modeli gerçek dünya uygulamaları üzerinde test etmektir. Bu aşamada, projenin başlangıçtaki hedefleri karşılayıp karşılamadığına dair son bir açıklamada dahil olmak üzere, değerlendirme sonuçlarının çalışılan alanın başarı kriterleri açısından özetlemesi yapılır.

Sürecin Gözden Geçirilmesi

Bir şekilde gözden kaçan herhangi bir önemli faktör veya görev olup olmadığını incelenir. Bu inceleme aynı zamanda modelin doğru oluşturulup oluşturulmadığı gibi kalite güvence konularını da kapsar.

Sonraki Adımların Belirlenmesi

Bu aşamada, projenin bitirilip bitirilmeyeceği projeye devam edilip edilmeyeceğine karar verilir.

2.1.1.5.1. Performans Değerlendirmeleri

Bu çalışma kapsamında sınıflandırma teknikleri kullanılacağı için bu bölümde sınıflandırma modellerinin performans değerlendirmelerine yer verilmektedir.

Tablo 3: Karmaşıklık Matrisi (Han vd., 2012)

	Tahmin Sınıfı			
Gerçek Sınıf		Pozitif	Negatif	Toplam
	Pozitif	TP	FN	P
	Negatif	FP	TN	N
	Toplam	P'	N'	P+N

True Positives (TP): Sınıflandırıcı tarafından doğru şekilde etiketlenmiş pozitifler.

True Negatives (TN): Sınıflandırıcı tarafından doğru şekilde etiketlenmiş negatifler.

False Positives (FP): Yanlışlıkla pozitif olarak etiketlenen negatifler.

False Negatives (FN): Yanlışlıkla negatif olarak etiketlenen pozitifler.

Sınıflandırma yöntemlerinin sınıf etiketini tahmin etmede ne kadar iyi veya ne kadar doğru olduğunu değerlendirmek için bazı ölçütler kullanılır.

Doğruluk (Accuracy): Tanıma oranı olarak söylenebilir.

$$\text{Doğruluk} = \frac{TP + TN}{P + N} \quad (4)$$

Hata Oranı (Error Rate): Yanlış sınıflandırma oranıdır.

$$Hata\ oranı = \frac{FP + FN}{P + N} \quad (5)$$

Duyarlılık (Sensitivity- recall): Gerçek pozitif oranıdır.

$$Duyarlılık = \frac{TP}{P} \quad (6)$$

Özgüllük (Specificity): Gerçek negatif oranıdır.

$$Özgüllük = \frac{TN}{N} \quad (7)$$

Kesinlik (Precision): tahmin edilen pozitiflerin gerçekte kaçının pozitif olduğunu gösterir.

$$Kesinlik = \frac{TP}{TP + FP} \quad (8)$$

F1 Skoru (F1-Score, F-measure): Kesinlik ve duyarlılığın harmonik ortalamasıdır.

$$F1 = 2 \times \frac{Kesinlik \times Duyarlılık}{Kesinlik + Duyarlılık} \quad (9)$$

Doğruluk ölçüsü, veri sınıfları oldukça eşit bir şekilde dağıtıldığında en iyi şekilde çalışır. Duyarlılık, özgüllük, kesinlik, F gibi diğer ölçütler, sınıf dengesizliği sorununa daha uygundur.

Doğruluğa dayalı ölçütlere ek olarak, sınıflandırma yöntemleri aşağıdaki yönlere göre de karşılaştırılabilir:

Hız: Verilen modelin oluşturulması ve kullanılmasıyla ilgili hesaplama maliyetlerini ifade eder.

Sağlamlık (Robustness): Modelin gürültülü veriler veya eksik değerlere sahip veriler verildiğinde doğru tahminler yapabilme yeteneğidir. Sağlamlık tipik olarak, artan gürültü derecelerini ve eksik değerleri temsil eden bir dizi sentetik veri seti ile değerlendirilir.

Ölçeklenebilirlik: Büyük miktarda veri verildiğinde modelin verimli bir şekilde oluşturulma yeteneğini ifade eder.

Yorumlanabilirlik: Bu, sınıflandırıcı veya öngörücü tarafından sağlanan anlayış ve içgörü düzeyini ifade eder. Yorumlanabilirlik sübjektiftir ve bu nedenle değerlendirilmesi daha zordur. Karar ağaçlarının ve sınıflandırma kurallarının yorumlanması kolay olabilir, ancak karmaşıklıkla yorumlanabilirlikleri azalabilir (Han vd., 2012).

Performans Değerlendirmeleri Sırasında Başvurulan Bazı Yöntemler

Holdout (Dışarıda Tutma): Bu yöntemde, verilen veriler rastgele bir eğitim seti ve bir test seti olmak üzere iki bağımsız kümeye bölünür. Tipik olarak, verilerin üçte ikisi eğitim setine ve kalan üçte biri test setine tahsis edilir. Eğitim seti modeli üretmek için kullanılır. Modelin doğruluğu daha sonra test seti ile tahmin edilir. Bu yöntemle doğruluk (accuracy) tahmin edilmeye çalışılır.

Random Subsampling (Rastgele Alt Örnekleme): Holdout yönteminin birkaç kez tekrarlandığı bir yöntemdir. Genel doğruluk tahmini, her yinelemeden elde edilen doğrulukların ortalaması olarak alınır.

Cross-validation (Çapraz doğrulama): Tahmine dayalı modellerin genelleştirme yeteneğini değerlendirmek ve aşırı öğrenmeyi önlemek için kullanılan bir yeniden örnekleme yöntemidir. Modellerin gerçek tahmin hatasını bulmak ve model parametrelerini ayarlamak için en yaygın kullanılan yeniden örnekleme yöntemlerinden biridir (Berrar, 2018). Çapraz geçişleme veya “k kat çapraz doğrulama” (k-fold cross validation) olarak adlandırılan doğrulama yapılırken veri seti rastgele k adet gruba bölünür. Gruplardan biri test seti, diğerleri eğitim seti olarak kullanılır. Model, eğitim setinde eğitilir ve test setinde puanlanır. Daha sonra süreç, test seti olarak kullanılan her bir benzersiz grup kadar tekrarlanır (Sorhun, 2021). “Leave-One-Out Cross-Validation” (LOOCV) olarak adlandırılan diğer çapraz doğrulama yönteminde ise bir veri noktasını sırayla test verisi olarak kullanarak modelin performansı değerlendirilir. LOOCV, k kat çapraz doğrulamanın özel bir durumudur (Wong, 2015). Sınıflandırma için doğruluk (accuracy) tahmini, k yinelemeden elde edilen doğru sınıflandırmaların toplam sayısıdır.

Bootstrap Örnekleme: Verilerden tahmin edilen bir modeli yeniden örnekleyerek bir tahmin edicinin veya test istatistiğinin dağılımını tahmin etmek için kullanılan bir yöntemdir (Horowitz, 2001). Bootstrap örnekleme, normal dağılımlar veya

eşit varyanslar gibi varsayımlar olmadan istatistiksel çıkarımlar (standart hata ve yanlışlık tahminleri, güven aralıkları ve hipotez testleri) sağlar. Bu nedenle, bootstrap yöntemleri, Normal veya t dağılımlarına dayalı klasik çıkarımlardan çok daha doğru olabilir (Hesterberg, 2011).

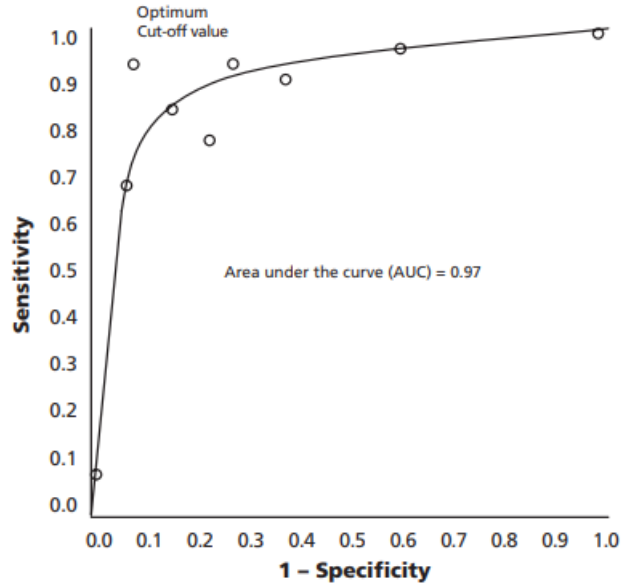
Maliyet-Fayda Temelli Karşılaştırma: Gerçek pozitifler, gerçek negatifler, yanlış pozitifler ve yanlış negatifler, bir sınıflandırma modeliyle ilişkili maliyet ve faydaların değerlendirilmesinde de yararlıdır. Yanlış bir negatif ile ilişkili maliyet (kanseri bir hastanın kanserli olmadığını yanlış bir şekilde tahmin etmek gibi), yanlış bir pozitifin (yanlış ama ihtiyatlı bir şekilde kanserli olmayan bir hastayı kanserli olarak etiketlemek) maliyetlerinden çok daha fazladır. Bu gibi durumlarda, her birine farklı bir maliyet atayarak bir hata türünün diğerine göre daha ağır basması sağlanabilir. Bu maliyetler, hasta için tehlikeyi, sonuçta ortaya çıkan tedavilerin finansal maliyetlerini ve diğer hastane maliyetlerini dikkate alabilir. Benzer şekilde, gerçek bir olumlu kararla ilişkili faydalar, gerçek bir olumsuz karardan farklı olabilir. Doğruluğu (accuracy) hesaplamak için kullanılan, maliyetler eşit varsayımlar ve esas olarak gerçek pozitifler ve gerçek negatiflerin toplamı toplam test grupları sayısına bölünen formüle alternatif olarak, karar başına ortalama maliyeti (veya faydayı) hesaplanarak maliyetler ve faydalar birleştirilebilir (Han vd., 2012).

ROC (Receiver Operating Characteristic Curve- Alıcı Çalışma Karakteristik Eğrileri): İki sınıflandırma modelini karşılaştırmak için yararlı bir görsel araçtır. ROC eğrileri, II. Dünya Savaşı sırasında radar görüntülerinin analizi için geliştirilen sinyal algılama teorisinden gelir (Fan vd., 2006). Belirli bir model için bir ROC eğrisi, gerçek pozitif oran (True Positive Rate-TPR) ile yanlış pozitif oran (False Positive Rate-FPR) arasındaki değiş tokuşu gösterir.

$$TPR = \frac{TP}{P} \quad (10)$$

$$FPR = \frac{FP}{N} \quad (11)$$

ROC analizi test performansı hakkında önemli bilgiler sağlar: eğrinin tepe noktası sol üst köşeye ne kadar yakınsa yani gerçek pozitif oranı yüksek, yanlış pozitif oranı düşük ise, testin ayırt etme yeteneği o kadar fazladır.



Şekil 10 :Yüksek Ayırt Etme Yeteneğine Sahip ROC Eğrisi (Fan vd., 2006)

2.1.1.6. Uygulama

Model oluşturma genellikle projenin sonu değildir. Kazanılan bilginin son kullanıcının kullanabileceği şekilde düzenlenmesi ve sunulması gerekir. İhtiyaçlara bağlı olarak bu aşama, basit veya karmaşık olabilir. Uygulama adımlarını gerçekleştiren genellikle veri analisti değil son kullanıcı olsa da oluşturulan modellerden fiilen yararlanmak için hangi eylemlerin gerçekleştirilmesi gerektiğinin önceden anlaşılması önemlidir. Buradaki kilit adımlar, planın devreye alınması, planın izlenmesi ve sürdürülmesi, nihai raporun üretilmesi ve projenin gözden geçirilmesidir.

Planın Devreye Alınması

Sonuçların çalışılan alana uygulanması için değerlendirme sonuçlarına bakılır ve uygulama için bir strateji belirlenir.

Planın İzlenmesi ve Sürdürülmesi

Planın izlenmesi ve sürdürülmesi ile ilgili bir strateji belirlenerek sonuçların yanlış kullanımı önlenabilir.

Nihai Raporun Üretilmesi

Projenin sonunda bir son rapor hazırlanır. Uygulama planına uygun olarak bu rapor projenin ve deneyimlerin bir özeti şeklinde veya sonuçların nihai ve kapsamlı bir sunumu olabilir. Bu rapor, önceki tüm çıktıları içerir ve sonuçları özetler ve düzenler.

Projenin Gözden Geçirilmesi

Gelecekteki projelerde kullanılmak üzere başarısızlıklar ve başarılar ile potansiyel iyileştirme alanları değerlendirilir. Bu adımda proje sırasındaki önemli deneyimlerin bir özeti bulunmalıdır. Böylece benzer durumlarda en uygun teknikleri seçmek için neler yapılacağına dair ipuçları belirlenebilir.

2.2. Sınıflandırma Teknikleri

Bir nesnenin hangi kategoriye ait olduğunu belirlemek için kullanılan yöntemlere sınıflandırma denir. Sınıflandırma, regresyonda sahip olduğumuz sayısal değerın aksine çıktının bir sınıf kodu olduğu başka bir denetimli öğrenme türüdür (Alpaydin, 2016).

Bir nesnenin iki kategoriden birine sınıflandırıldığı ikili sınıflandırma, en basit sınıflandırma türü olarak belirtilmektedir. İki den fazla kategori tanımlanarak ikili sınıflandırma genişletilip çoklu bir sınıflandırmaya geçilebilir.

İkili sınıflandırmada pozitif sınıf ve negatif sınıf olmak üzere iki kategori önceden tanımlanır ve örnekler her iki kategoriye de tahsis edilir. İkili sınıflandırma, her öğeyi pozitif sınıfa veya negatif sınıfa atamaktadır. Çoklu sınıflandırmayı veya regresyonu ikili sınıflandırmalara ayırtırmak mümkündür (Jo, 2021).

Sınıflandırma teknikleri; lojistik regresyon, diskriminant analizi, karar ağaçları, rassal orman, destek vektör makineleri, yapay sinir ağları, en yakın komşuluk ve naive bayesdir.

2.2.1. Karar Ağaçları

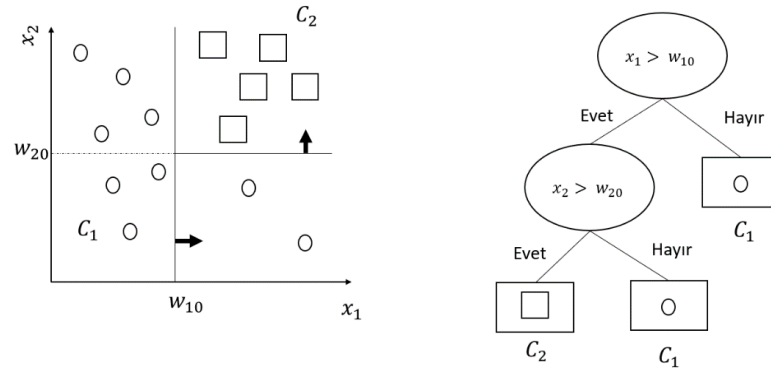
Karar ağaçları, kararlar ile bunların olası sonuçlarını ilişkilendirilen ağaç benzeri bir modeldir. Ağaç benzeri yapısı içinde her dal, farklı seçenekler arasında bir kararı temsil eder. Veri noktalarının sınıflarına bağlı olarak veriyi alt kümeler halinde böler. Karar analizinde insanlara yardımcı olur. Anlaşılması ve yorumlanması kolay olması

nedeniyle karar ağaçları kullanımı popülerdir. Karar ağaçları, sınıflandırıcının sahip olduğu bilgi kümesini daha anlaşılır şekilde, olası sınıf etiketlerini ile olasılıklara bağlı durumlarını bir düzen içerisinde ağaç halinde sunan bir algoritmadır. Hem kategorik hem de sürekli veriler üzerinde çalıştırılabilir (Bilgin, 2018; Özkan, 2021; Patra vd., 2022; Rokach & Maimon, 2015).

Karar ağaçları, makine öğrenmesindeki en eski yöntemlerden biridir ve hem eğitim hem de tahmin açısından basit olmasına rağmen birçok alanda kullanılır. Ağaçlar, karmaşık bir görevi daha basit, bölgesel görevlere bölen, Sezar'dan beri popüler olan ünlü “böl ve fethet (divide-and-conquer)” stratejisini kullanır (Alpaydin, 2016). Bu nedenle karar ağaçları her türlü uygulamada karmaşıklığı azaltmak için sıklıkla kullanılır.

Karar ağaçları sınıflandırma problemlerinde çok tercih edilen algoritmalarındandır. Geliştirilme amacı; çok sayıda veri içeren bir kümeyi, bazı teknikler kullanarak alt bölümlere ayırmak için geliştirilmiş bir algoritmadır. Sınıflandırma algoritmaları içinde anlaşılması ve kodlanması en kolay olanlardan biridir (Agrawal vd., 1993).

Karar ağaçları denetimli öğrenme türünden bir algoritmadır ve iki adımda çalışır. Birinci adımda denetimli öğrenme algoritmaları için ortak olan ve “eğitim” denilen adımdır. Bu adımda eldeki etiketli verilerden (eğitim verisi) bir model oluşturulur. İkinci adıma test aşaması denir ve bu adımda etiketsiz veri (test verisi) ilk adımda oluşturulan model üzerinde çalıştırılarak etiketsiz verinin sınıfı tahmin edilmeye çalışılır (Bilgin & Eser, 2022a). Bir karar ağacı kök düğüm, düğümler ve yapraklardan oluşur. Karar ağacı, bir ağaç formunda (ters bir ağaç) temsil edilen basit bir akış şemasıdır. Kök düğümlerle başlar ve birkaç düğüme dallanır, bu düğümler bir karar temelinde dallanır ve nihai sonucun belirlendiği bir yaprak düğümü ile sona erer (Ramasubramanian & Moolayil, 2019).



Şekil 11: Bir Veri seti ve İlgili Karar Ağacı Örneği (Alpaydın, 2014).

Karar ağaçları tekniğinde birçok algoritma vardır. Bunlardan en sık kullanılanları ID3, C4.5-C5.0, CART (Classification and Regression Trees) ve CHAID (Chi-squared Automatic Interaction Detector) dir (Rokach & Maimon, 2015).

ID3 Algoritması: J. Ross Quinlan tarafından 1986 yılında geliştirilmiştir. Bölme özelliklerini seçmek için bilgi kazancını kılavuz olarak alır (Zhou, 2021). Bilgi kazancı (information gain), seçilen özellik için sınıflandırma sonuçlarının değerini ölçmek için kullanılan 0-1 arasında değer alan bir ölçüttür. Seçilen özelliğin veri setindeki sınıflar için ayrı değerler alması ilgili özelliğin ayırt ediciliğinin fazla olduğuna işaret eder. Bilgi kazancı ölçütünün hesaplanmasında entropi değeri kullanılmaktadır. Entropi değeri, bir veri seti içerisinde belirsizlik ve rastgeleliği ölçmek için kullanılır (Quinlan, 1986).

C4.5-C5.0 Algoritması: Bilgi kazancı kriterinin, daha fazla olası değere sahip özelliklere yönelik önyargılı olduğu ortaya çıktıktan sonra bu yanlılığı azaltmak için, Quinlan 1993 yılında, bilgi kazancını kullanmak yerine özellikleri seçmek için kazanç oranını kullanan C4.5 algoritmasını geliştirmiştir. Kazanç oranı, bilgi kazancı ölçütünün normalleştirilmiş halidir. Ardından daha az bellek kullanılmasını ve daha hızlı olunmasını sağlayan C5.0 algoritmasını literatüre kazandırmıştır (Rokach & Maimon, 2015).

CART (Classification and Regression Trees) Algoritması: Sınıflandırma için en eski karar ağaçları algoritmalarından biri safsızlığı azaltmak için ağırlıklara ince ayar yapan CART algoritmasıdır (Breiman, 2001). CART ayrıca, alt küme seçimi aracılığıyla boyutsallığı ve düğümün karmaşıklığını azaltmak için bir ön işleme aşamasına sahiptir (Alpaydın, 2014).

CHAID (Chi-squared Automatic Interaction Detector) Algoritması: G. V. Kass tarafından 1980’de geliştirilmiştir. Bölme işlemi olarak ki-kare testlerini kullanır. Dalların sayısı iki ile sınıf kategorisi arasında değişir (Silahtaroglu, 2013). Kategorik ve sürekli değişkenler üzerinde çalışabilmesi, ağaçta her düğümü ikiden fazla alt gruba ayırabilmesi gibi nedenlerle günümüzde de tercih edilen bir algoritmadır. CHAID algoritması çoklu ağaçlar üretir. Bağımlı değişken için tüm bağımsız değişkenler ile çapraz tablolar oluşturulur ve ki-kare değerleri hesaplanır. Ki-kare değeri en büyük olan kök düğüme yerleştirilir (Bilgin & Eser, 2022a).

2.2.2. *Rassal Orman*

Rassal Orman, her bir ağacın bağımsız olarak örneklenen ve ormandaki tüm ağaçlar için aynı dağılıma sahip rastgele bir vektörün değerlerine bağlı olduğu karar ağaçlarının bir kombinasyonudur. Rassal orman, her düğümde bölünmeyi rastgele hale getiren büyük bir sayıda bireysel, budanmamış karar ağacı kullanır (Breiman, 2001). Her ağaç, kesin bölünmelerle oluşturulan bir ağaçtan daha az doğru olma eğilimindedir. Ancak, bu "yaklaşık" ağaçlar bir toplulukta bir araya getirilerek, genellikle kesin bölünmelerle oluşturulan tek bir ağaçtan daha iyi bir doğruluk elde edilir (Rokach & Maimon, 2015).

Rassal Orman, birden fazla modelin oluşturulduğu ve sonuçların çoğunluk oylaması ile birleştirildiği bir topluluk öğrenme modelidir. Birden fazla karar ağacının bir araya gelmesiyle oluşan Rassal Orman algoritması, ağaç tabanlı bir topluluk öğrenme yaklaşımı olarak tanımlanabilir. Bu nedenle, Rassal Orman algoritması karar ağaçlarından daha iyi performans gösterir (Bayram Arlı, vd., 2022).

Topluluk öğrenmesi yaklaşımı olarak adlandırılan, çok sayıda sınıflandırıcı üreten ve sonuçlarını birleştiren yöntemlerden biri "boosting" (Bartlett vd., 1998) diğeri ise torbalama (Bagging- Bootstrap Aggreating (Breiman, 1996) yöntemidir. Boosting'de, ardışık ağaçlar önceki tahminciler tarafından yanlış tahmin edilen noktalara ekstra ağırlık verir. Sonunda, tahmin için ağırlıklı bir oy alınır (Liaw & Wiener, 2002).

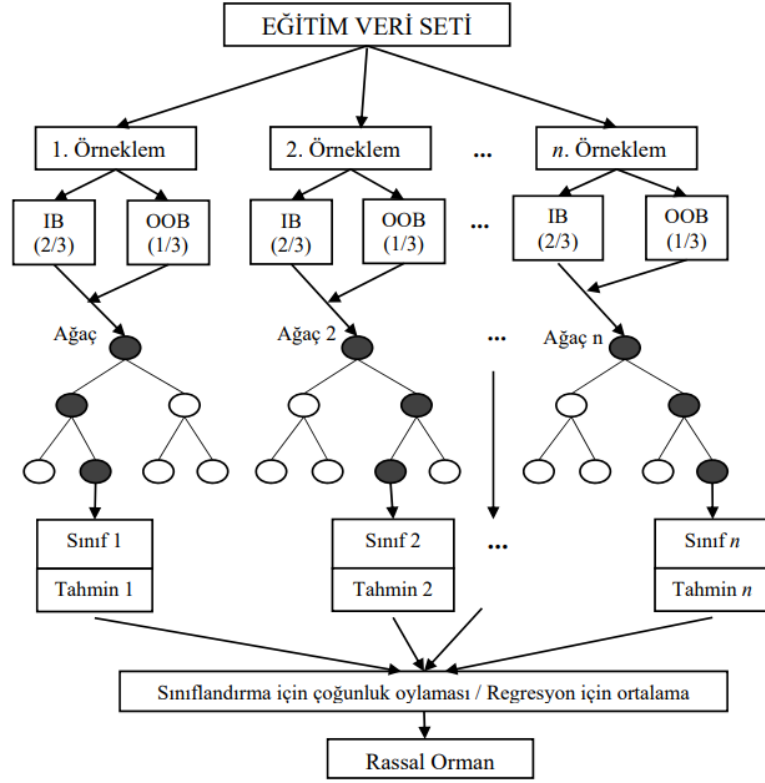
Torbalama tekniği, ardışık ağaçların önceki ağaçlara bağlı olmadığı bir yöntemdir. Her bir ağaç, veri kümesinin bootstrap örneği kullanılarak bağımsız olarak oluşturulur. Bütün ağaçların tahmini bulunduktan sonra, rassal orman tahmini için basit bir çoğunluk

oyu alınır. Bu tekniğin bir parçası olarak, eğitim veri setinin 2/3'ü (in-bag- IB) içeren rassal ve tekrarlanabilir örneklemeler alınır ve birbirinden farklı eğitim setleri eşzamanlı olarak oluşturulur. Eğitim veri setinde bulunan diğer 1/3'lük kısım torba dışı (out-of-bag- OOB) olarak adlandırılır ve bu setten ağaçların başarı ölçümü ile değişkenlerin ağırlığının değerlendirilmesi için yararlanılır (Liaw & Wiener, 2002; James vd., 2013; Akar vd., 2017; Xing, vd. 2019; Celbiş, 2021).

Torbalama tekniği, T adet ağaç için n adet gözlemden ($n < N$) oluşan örneklemelerin bootstrap (yeniden örnekleme) yöntemiyle oluşturulduğu bir tekniktir. Her seferinde seçilen gözlemler aynı seçim havuzuna geri konulur ve yeniden örnekleme yapılarak birden fazla ağaç oluşturulur. Torbalama işlemi sırasında oluşturulacak ağaçların toplam sayısı (ntree) araştırmacı tarafından tayin edilir. Araştırmacı, ne kadar çok ağaç oluşturulması gerektiğine karar verir. Daha fazla ağaç, genellikle daha iyi sonuçlar elde etme olasılığını artırır. Her bir ağaç, diğer ağaçlardan tamamen bağımsız olarak oluşturulur. Yani bir ağacın oluşturulması, diğer ağaçların oluşturulmasını etkilemez (Liaw & Wiener, 2002; Hastie vd., 2009; James vd., 2013; Genuer & Poggi, 2020).

Bu yöntemde, aynı veri kümesinden rastgele ve tekrarlı olarak örneklemeler (bootstrap örnekleri) alınır. Her örneklem, orijinal veri kümesinden rastgele seçilen gözlemlerden oluşur. Bu işlem, her ağaç için ayrı ayrı gerçekleştirilir. Her bir örneklem, ağaçların oluşturulmasında kullanılacak gözlemlerin rastgele bir alt kümesini temsil eder. Her gözlem, her ağacın oluşturulmasında kullanılma şansına sahiptir. Yani her gözlem için eşit bir fırsat söz konusudur. Algoritma çalıştırıldığında, her bir ağaç kendi örnekleme üzerinden eğitilir ve bu ağaçlar ayrı ayrı modeller tahmin eder. Sonuçta, her bir ağaç farklı tahminlerde bulunur. Sınıflandırma problemleri için, farklı modellerin tahminleri çoğunluk oylamasıyla birleştirilir. Yani her model, bir gözlem için hangi sınıfa ait olduğunu tahmin eder ve en çok oy alan sınıf genel tahmin olarak kabul edilir. Regresyon problemleri için ise farklı modellerin tahminleri ortalama değere göre birleştirilir. Her modelin tahmini sayısal bir değerdir ve bu değerlerin ortalaması alınarak genel tahmin elde edilir. Her defasında veri kümesinden rastgele ve tekrarlı olarak örneklemeler (bootstrap örnekleri) alınır. Bu örneklemeler, orijinal veri kümesinden rastgele seçilen gözlemlerden oluşur. Yani her model kendi eğitim veri setini oluştururken farklı gözlemleri kullanır. Bu yöntem, ağaçların çeşitliliğini artırarak overfitting (aşırı uyum)

sorununu azaltır ve daha genel ve güvenilir bir sonuç elde edilmesine yardımcı olur. Her bir ağacın farklı bir bakış açısına sahip olması, topluluk olarak daha iyi kararlar alınmasını sağlar (Liaw & Wiener, 2002; Hastie vd., 2009; James vd., 2013; Genuer & Poggi, 2020).



Şekil 12: Rassal Orman Algoritması (Bayram-Arlı vd., 2022)

Rassal orman algoritması, her bir ağaç için rassal örneklemeler kullanarak aşırı öğrenme sorununu azaltır. Bu nedenle, genellikle daha iyi genelleme yaparlar. Birden çok ağacın birleştirilmesiyle, modelin varyansı düşer ve daha istikrarlı tahminler elde edilir. Aynı zamanda, yanlılık (bias) da azaltılır. Karar ağaçlarına kıyasla genellikle daha iyi performans gösterir. Çünkü çok sayıda ağacın tahminleri bir araya getirilerek daha güçlü bir model oluşturulur. Karar ağaçlarında bazen budama işlemi gerekebilirken, rassal orman algoritması için böyle bir işlem gerekmez. Her ağaç tam büyüklüğünde olabilir. Rassal orman algoritması, kayıp verilerle daha iyi başa çıkabilir ve tutarlı sonuçlar üretebilir. Kayıp verileri ele almak için özel bir işlem gerektirmezler. Aynı zamanda boyut azaltma yöntemi olarak kullanılabilir ve en önemli değişkenleri belirlemeye yardımcı olabilirler.

Bahsedilen avantajlarının yanında rassal orman algoritmasının bazı dezavantajları da vardır. Rassal orman algoritması, birçok ağaç üretir ve her biri rastgele örneklemelerle eğitilir. Bu nedenle, karar ağaçlarına kıyasla daha fazla eğitim süresi ve hesaplama gücü gerektirirler. Her bir ağacın tahminlerini birleştirerek sonuç elde ettikleri için tek bir ağaç gibi kolayca yorumlanamazlar. Sonuç olarak, rassal orman algoritması, birçok problemde etkili bir öğrenme yöntemi olabilir, ancak büyük veri setleri veya hızlı sonuç gerektiren durumlarda eğitim süresi ve hesaplama gücü gibi dezavantajları olabilir. Yine de genellikle performansı artırmak ve aşırı öğrenmeyi önlemek için tercih edilirler (Bayram Arlı vd., 2022).

2.2.3. Destek Vektör Makineleri

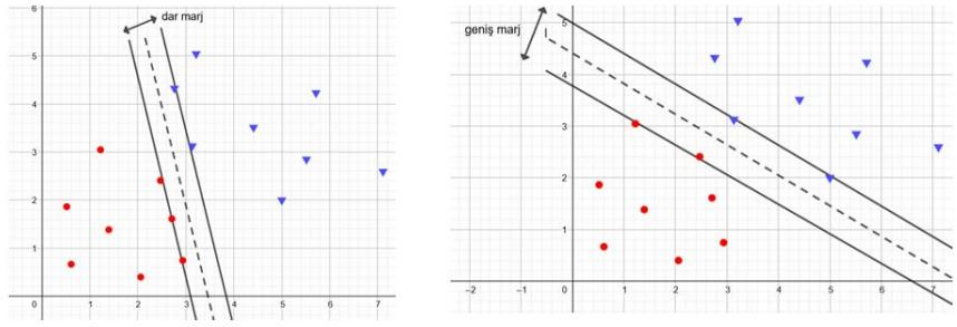
Destek Vektör Makineleri (DVM), koşullu olasılıkları (posterior probabilities) hesaplamaktan kaçınan seyrek çekirdekli bir karar makinesidir. DVM, istatistiksel öğrenme teorisindeki (Vapnik, 1998) matematiksel temeli nedeniyle makine öğrenmesi problemlerine ilkeli bir yaklaşım sunar. İstatistiksel öğrenme modellemesindeki amaç, veriye dayalı bir örüntü veya ilişkiyi tanımlamak ve bu örüntüyü kullanarak gelecekteki olayları veya verileri tahmin etmektir. Bunun için de hipotez uzayından hedef uzayda altta yatan fonksiyona en yakın (bazı hata ölçülerine göre) bir model seçmek gerekmektedir (Babangida vd., 2016; Castaneda & Cockburn, 2005).

DVM istatistiksel öğrenme modellerinin eğitimi ve genelleme yeteneklerini geliştirmeye odaklanan bir ilke olan yapısal risk minimizasyonunu (Structural Risk Minimization-SRM) kullanmaktadır. Model karmaşıklığı arttıkça, eğitim verilerine uyma (eğitim hatası) genellikle azalır. Ancak, model karmaşıklığı çok yüksekse, genelleme hatası artabilir, yani model yeni verilere kötü uyarlanabilir (aşırı uyum-overfitting). SRM ilkesi, bir modelin karmaşıklığını artırmadan önce genelleme hatasının ne kadar iyileştirilebileceğini değerlendirir. Bu nedenle, DVM gibi modellerin daha iyi genelleme yeteneği ve hatanın azaltılmasını sağlamak gibi önemli avantajlara sahip olmalarını sağlar (Awad & Khanna, 2015; Tanveer vd., 2021).

DVM, sezgisel bir model gösterimi ile genel (doğrusal olmayan) sınıflandırma, regresyon ve aykırı değer tespiti için güçlü bir teknik sunar. İki sınıfa ait gözlemler, iki boyutlu uzayda bir doğru, üç boyutlu uzayda bir düzlem, n boyutlu uzayda ise hiperdüzlem (hyperplane) denilen ayırma araçları ile ayrılır. DVM ise; bir sınıflandırma

probleminde iki sınıf arasındaki ayrım hiperdüzlemi (decision boundary) veya sınıflandırma sınırını bulmayı amaçlayan bir öğrenme algoritmasıdır. Bu ayrım hiperdüzlemi, iki farklı sınıfa ait örnekleri mümkün olduğunca iyi bir şekilde ayırmayı hedefler (Kim vd., 2022; Kartal, 2022).

DVM, bu optimal/ayırıcı hiperdüzlemi bulurken, modelin eğitim verileri üzerinde ne kadar iyi performans gösterdiğini ölçen eğitim hatasını en aza indirmek ve her iki sınıfa ait veri noktaları arasındaki en yakın mesafeyi yani marjı maksimize etmeyi hedefler. (Durgesh & Lekha, 2010; Kuzey, 2012).



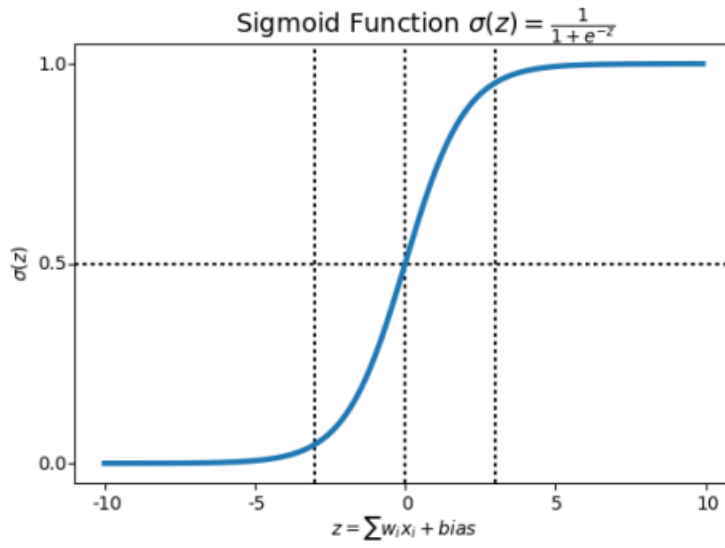
Şekil 13: Optimal Hiperdüzlemin Seçimi (Kartal, 2022)

DVM nin birçok olumlu tarafı vardır. DVM gibi modeller, bir sınıflandırma veya regresyon görevi için eğitildikten sonra, gelecekte yeni örneklerin tahminlerini yaparken yalnızca destek vektörlerine bağlıdır. Verilerin karmaşıklığı veya dağılımı nedeniyle, sınıflar arasında net bir ayrımın doğrusal bir hiperdüzlemle çizilemediği veya bir doğruyla sınıfların tam olarak ayrılamadığı durumlarda veriyi çekirdek fonksiyonu yardımıyla daha yüksek boyutlu bir uzaya taşıyarak sonuca analitik olarak yaklaşır. Oluşturulan yeni uzayda verinin doğrusal biçimde ayrılmasına ve yalnızca hatanın en aza indirilmesi değil bunun yanında hiperdüzlemin, farklı sınıflardan maksimum uzaklıkta olacak şekilde yerleştirilmesine de dikkat eder (Awad & Khanna, 2015) .

DVM nin bu olumlu yanlara rağmen, karmaşık veri ile çalışırken optimizasyon problemlerini çözmede yüksek hesaplama maliyetine yola açması, model performansının büyük ölçüde çekirdek fonksiyonlarının ve parametrelerinin seçimine bağlı olması ve eğitim süresinin fazla olması gibi olumsuz tarafları da vardır (Attewell & Monaghan, 2015; Navlani, 2019; Tanveer vd., 2021).

2.2.4. Lojistik Regresyon

Hedef değişkenin kategorik ve iki olası değer aldığı durumlarda veri analizi yapılırken en sık kullanılan tekniklerden biri lojistik regresyon analizidir. Hedef değişkenin kategorilerini tahmin eden ve sigmoid olarak adlandırılan S şeklindeki lojistik fonksiyondan yararlanılarak hedef değişkenin ikili değerden herhangi birinin değerini alma olasılığı bulunmaya çalışılır.



Şekil 14 : Sigmoid Fonksiyonu (Mugunthan & Vijayakumar, 2021)

Makine öğrenmesi yönetiminde, doğrusal regresyon algoritmaları, bağımlı ve bağımsız değişkenler arasındaki ilişkiyi tahmin etmek için kullanılırken, lojistik regresyon algoritması, sınıflandırma problemlerini çözmek için kullanılır ve nesnelerin belirli sınıflara atanmasına yardımcı olur. İkisi arasındaki temel fark, doğrusal regresyon sürekli bir çıktıyı tahmin etmeye odaklanırken, lojistik regresyon sınıf etiketlerini tahmin etmeye odaklanır (Dangeti, 2017; Ramasubramanian & Moolayil, 2019; Sakarkar vd., 2021).

Lojistik regresyon analizi, incelenen kategorik hedef değişkenin ölçek türüne ve hedef değişkenin kategori sayısına göre iki farklı türe ayrılır. Eğer hedef değişken yalnızca iki kategoriye sahipse ve sınıflayıcı ölçek düzeyinde ölçülmüşse, bu durumda “ikili lojistik regresyon” olarak adlandırılır. Ancak eğer hedef değişken iki kategoriden fazlasına sahipse, yani çoklu sınıflara ayrılıyorsa, bu durumda “çoklu lojistik regresyon” olarak adlandırılır. (Gürsakal vd., 2022).

2.2.4.1. İkili Lojistik Regresyon

Lojistik regresyon, ikili sınıflandırma için en çok tercih edilen beyaz kutu modelidir. Beyaz kutu modelleri, tahminin yapıldığı tüm mantığı görünür kılan modeller olarak tanımlanır. Her tahmin için modelin matematiksel denklemi kullanabilir ve tahminin nedenlerini çözülebilir. İkili lojistik regresyon, hedef değişkenin kategorik olduğu durumlar için diğer sınıflandırma tekniklerine göre farklı bir yaklaşım sunmaktadır. Bu yaklaşıma göre sonuç, olasılığın logaritması veya olayın olasılığına dönüştürülerek bulunur (Ramasubramanian & Moolayil, 2019) .

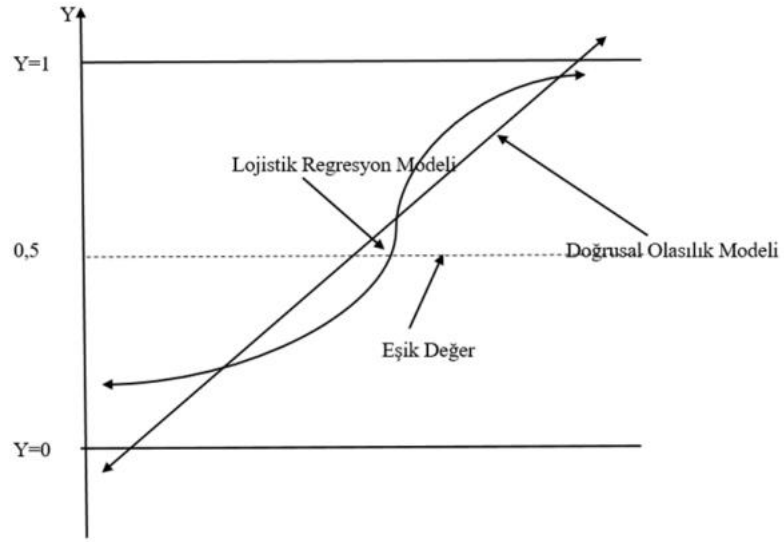
Lojistik regresyon modeli matematiksel olarak aşağıdaki şekilde ifade edilebilir:

$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n \quad (12)$$

Bu formülde gösterilen logaritma değeri, logit işlevi olarak da adlandırılan “oranların logaritması (log of the odds ratio)” dır. Ayrıca p değeri sonucun olasılığıdır ve bu değer aşağıdaki formülle elde edilir:

$$p = \frac{e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n}}{1 + e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n}} \quad (13)$$

Hedef değişkenin beklenen değerini bulmak için sigmoid olarak bilinen lojistik dağılım fonksiyonundan yararlanılır. Sigmoid fonksiyonu, tahmin edilen değerleri olasılıklara dönüştürebilen, herhangi bir gerçek değeri 0 ile 1 arasında başka bir değer ile eşleştiren S şeklinde bir eğridir. Herhangi bir eşik değeri belirlenir ve bu değer altında kalanlar 0, üstünde kalanlar 1 alınarak sınıflandırma işlemi yapılır.



Şekil 15: Lojistik Regresyon Eğrisi (Gürsakar vd., 2022)

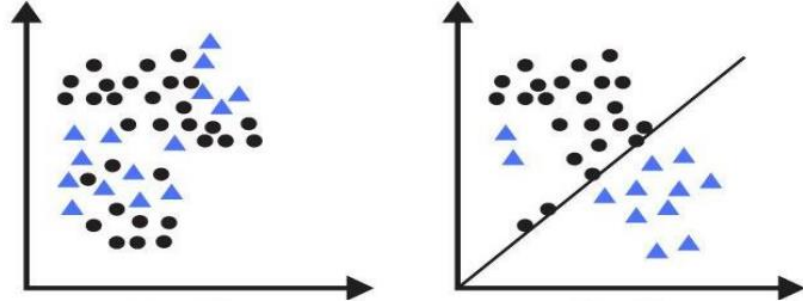
Öznitelik değişkenlerinin hem sürekli hem de kategorik olması, model parametrelerinin kolay bir şekilde yorumlanabilmesi, öznitelik değişkenlerinin normal dağıldığının varsayılmasına gerek olamaması gibi olumlu yanlarının yanında kategoriler arasında tam bir ayrım olduğunda çalışmaması, kategorileri ayırmak için eğri bir yüzey gerektiğinde düşük performans göstermesi gibi olumsuz yanları da bulunmaktadır (Gürsakar vd., 2022).

2.2.5. Diskriminant Analizi

Sınıflandırmada kullanılan bir başka yöntem de Diskriminant Analizi (Discriminant Analysis) dir. Genellikle makine öğrenmesi uygulamaları için ön işleme adımında boyut azaltma tekniği olarak kullanılan Diskriminant Analizi, denetimli öğrenme içinde en sık kullanılan boyut azaltma tekniğidir. Bu yöntem, genellikle değişkenleri iki veya daha fazla sınıfa dağıtmak, yani gruplar arasındaki farklılıkları modellemek için yaygın bir şekilde kullanılır (Murphy, 2022; Sakarkar vd., 2021).

Diskriminant analizi, iki veya daha fazla gruba ayrılmış nesneler üzerinde ölçülen çok değişkenli verilerle ilgilenir. Amaç, diğer nesneleri bu gruplardan birine veya diğerine sınıflandırmak veya bu gruplar arasındaki farkları tanımlamaktır. Her iki durumda da özellikle ilgi, hangi değişkenlerin lineer kombinasyonlarının gruplar arasındaki farkları en iyi şekilde vurguladığını belirlemektir (Fieller, 2016).

Diskriminant analizi, diskriminant fonksiyonları oluşturmak için verileri daha düşük boyutlu bir temsile yansıtan denetimli bir öğrenme algoritmasıdır. Diskriminant fonksiyonları, sınıf merkez noktalarının ayrılmasını en üst düzeye çıkarırken, bunlar boyunca her bir sınıfın varyansını en aza indiren orijinal (sürekli) değişkenlerin doğrusal kombinasyonlarıdır (Öngen Bilir, 2022). Hedef değişkenin ikiden fazla kategoriye ayrılmasında kullanılması tavsiye edilir. Amacı sınıflar arasındaki ayrımı en üst düzeye çıkaran yeni özellikler bulmaktır. İstatistik biliminde sınıflandırmanın genel adı diskriminant analizidir (Alpaydın, 2016). Diskriminant analizinin yapılabilmesi için sınıfların üyeliklerini gösteren hedef değişken ve gruplardaki bireylere ait özelliklerin bulunduğu öznitelik değişkenler olmalıdır. Sınıf üyeliğini gösteren hedef değişken kategorik, öznitelik değişkenler ise nicel değerler almalıdır (Gürsakaç, 2019).



Şekil 16: Diskriminant Analizi Öncesi ve Sonrası (Sakarkar vd., 2021)

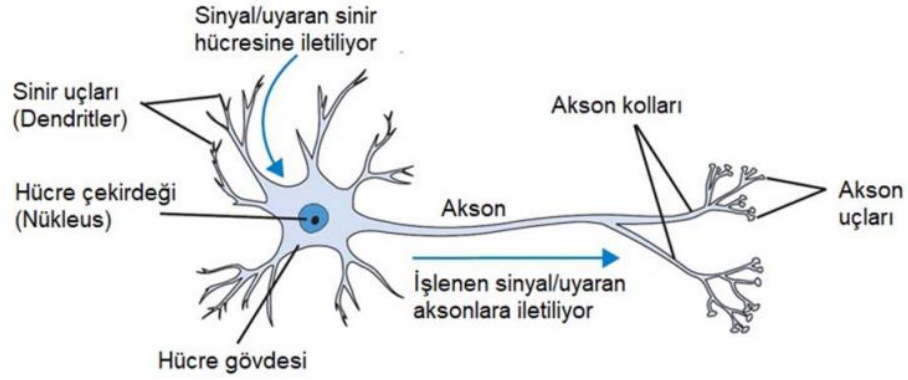
Boyut indirgemede kullanılabilmesi bu analizin olumlu tarafını oluştururken öznitelik değişkenlerinin sadece sürekli olduğu durumlarda kullanılabilmesi, verilerin normal dağıldığının ve kategorilerin eşit kovaryansa sahip olduğunun varsayılması gibi olumsuz tarafları da vardır.

2.2.6. Yapay Sinir Ağları

Yapay sinir ağları (YSA), beyin tarafından yapılan duyuşal işleminin ilk modellerinden esinlenip bir bilgisayarda model nöron ağı simüle edilerek oluşturulan gerçek nöronların işlemlerini taklit eden algoritmalar uygulayarak, ağı birçok türde sorunu çözmeyi “öğrenmesini” sağlayan bir makine öğrenmesi tekniğidir (Krogh, 2008; Rojas, 1996). Genellikle sinir ağı olarak adlandırılan yapay sinir ağı, biyolojik sinir ağlarından ilham alan matematiksel bir modeldir. Bir sinir ağı, birbirine bağlı bir yapay

nöron grubundan oluşur ve hesaplamaya bağlantılı bir yaklaşım kullanarak bilgileri işler (Bailer-Jones vd., 2001; Fausett, 1993; Gupta, 2013).

Biyolojik sinir ağlarından ilham alan YSA'lar, çok sayıda ara bağlantıya sahip çok sayıda basit işlemciden oluşan büyük ölçüde paralel bilgi işlem sistemleridir. YSA modelleri insan beyninin kullanıldığına inanılan bazı örgütsel kuralları kullanmaya çalışır. Düğümlerinde “yapay nöronlar” olan bir ağ ortaya koyulur. Bunlara yapay sinir ağları denir. Yapay nöronlar doğal nöronlardan ilham alınarak yapılmıştır. YSA'nın temel işlem ünitesi olan nöron, dış ortamdan veya komşu nöronlardan aldığı girdiği işleyerek sonraki nörona veya çıktı katmanına iletir (Özen, 2022).

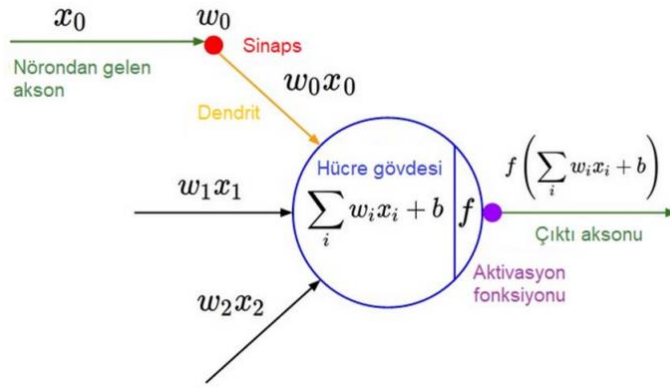


Şekil 17: Biyolojik Sinir Hücresinin Yapısı (Özen, 2022).

Biyolojik sinir hücrelerine benzeyen yapay sinir hücreleri oluşturulur. YSA larda kullanılan işlem çoklu girişli ve ağırlıklandırılmış bir toplama işlemidir. Bu işlemde ilk adım dış ortamdan gelen verilerin veya sinyallerin $(x_0, x_1, x_2, \dots, x_n)$ toplama işlemine giriş olarak kullanılmasıdır. Bu girdiler, genellikle bir veri noktasının veya olayın özelliklerini temsil ederler. Her bir giriş, bir ağırlık ile çarpılır. Bu ağırlıklar $(w_0, w_1, w_2, \dots, w_n)$, her bir girişin model tarafından ne kadar önemli olduğunu veya katkı sağladığını belirler. Ağırlıklar, modelin öğrenme süreci sırasında ayarlanabilir veya önceden belirlenmiş değerlere sahip olabilirler. Her bir giriş, karşılık gelen ağırlık ile çarpılır. Yani, “ $x_i * w_i$ ” işlemi yapılır ve her bir girişin ağırlıklı değeri elde edilir. Ağırlıklı girişlerin her biri toplanarak bir toplam değer elde edilir. Bu toplam, genellikle bir çıkış veya tahmin değeri olabilir. Toplama işlemi, girişlerin ve ağırlıkların doğru bir şekilde toplanmasıyla gerçekleştirilir. Matematiksel olarak ifade etmek gerekirse, toplama işlemi aşağıdaki gibi yazılabilir:

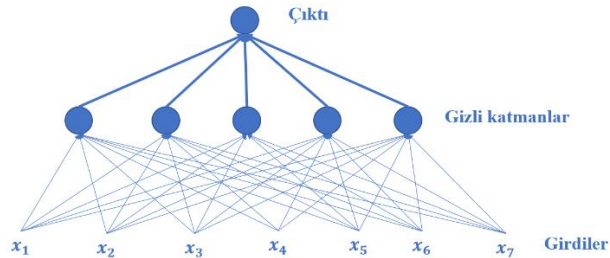
$$\Sigma = x_0w_0 + x_1w_1 + \dots + x_nw_n \quad (14)$$

Bu toplam değer aktivasyon fonksiyonu veya transfer fonksiyonu adı verilen bir fonksiyon ile değerlendirilir. Aktivasyon fonksiyonu veya transfer fonksiyonu, yapay sinir ağlarının veya benzeri hesaplama modellerinin çıktısını belirlemek için kullanılan önemli bir bileşendir. Bu fonksiyon, toplam değeri (ağırlıklı girişlerin toplamını) alır ve bu değeri bir sonraki nörona veya çıktı katmanına iletmek için değerlendirir. Aktivasyon fonksiyonu, ağın nonlinearity (doğrusal olmayanlık) eklemesine yardımcı olur ve modelin daha karmaşık ilişkileri öğrenmesini sağlar (Özen, 2022).

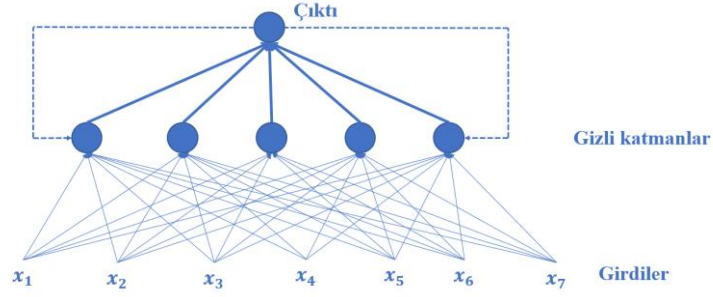


Şekil 18: Yapay Sinir Hücresinin Yapısı (Özen, 2022).

Yapay sinir hücrelerinden oluşan sisteme yapay sinir ağları denmektedir. Yapay sinir ağlarının birkaç farklı yapısı bulunmaktadır. Yapıda girdi katmanı, ara katman ve çıktı katmanı bulunur. Ağlarda sinyallerin akışı tek yönde ve geriye doğru bağlantılar yoksa buna ileri beslemeli YSA, sinyaller işlenmiş olan nörona tekrar geri dönebiliyorsa buna da geri beslemeli YSA denir (Khodayar & Regan, 2023; Venkateswaran & Ciaburro, 2017; Walczak, 2019).



Şekil 19: İleri Beslemeli Yapay Sinir Ağları Mimarisi



Şekil 20: Geri Beslemeli Yapay Sinir Ağları Mimarisi

Yapay sinir ağlarının birçok olumlu tarafları bulunmaktadır. Gerçek hayattaki karmaşık ilişkileri öğrenme yeteneğine sahiptir (Singh, 2021), bunun yanında büyük ölçekte gürültülü veriler ile çalışabilir (Lancashire vd., 2008; Singh, 2021) vbue bir kez eğitildikten sonra yeni veri setlerinde modelin ilişkileri ortaya çıkararak genelleştirme yapabilmektedir (Bailer-Jones vd., 2001). Ayrıca bilgiler bir veri tabanında değil tüm ağda depolanır, birkaç bilgi parçasının tek bir yerde kaybolmasının ağın çalışmasına engel olmaz. YSA lar eğitildikten sonra eksik bilgilerle dahi çıktı üretebilir, ağdaki bir veya daha fazla hücrenin bozulması çıktı üretimine engel olmaz, ağlar hataya dayanıklıdır, ağ yıpranmasının yavaştır ve aynı anda birden fazla işi gerçekleştirebilecek sayısal güce sahiptir (Gupta, 2013; Khodayar & Regan, 2023; Walczak, 2019).

Sayılan bu olumlu taraflarının yanında modelin nasıl ve neden belirli bir çıktı elde ettiğinin (kara kutu- black box) tam olarak bilinmemesi, ağ yapıları gereği paralel işlem gücüne sahip işlemcilerle ihtiyaç duyarak hesaplama maliyetinin yüksek olması, uygun ağ yapısının bulunması için deneme yanılma yoluna başvurulması, eğitim sürecinin tam olarak bilinmemesi (Mijwel, 2018) gibi olumsuz tarafları da vardır.

2.2.7. *Naive Bayes*

Naive Bayes (NB) algoritması, bir rassal değişken için olasılık dağılımı içinde koşullu olasılıklar ile ön olasılıklar arasındaki ilişkiyi gösteren Bayes teoremi prensiplerine dayalı olarak oluşturulmuş bir sınıflandırma algoritmasıdır. NB algoritması, anlaşılması kolay ve yüksek başarı oranlarına sahip bir makine öğrenmesi algoritmasıdır ve genellikle verinin özelliklerinin belirli bir sınıfa bağımlılığını koruyarak uygun sınıflandırmayı gerçekleştirir (Engin, 2022; Lantz, 2015; Orhan & Adem, 2012).

Bayes teoreminin formülü aşağıdaki gibidir:

$$P(A_i \setminus B) = \frac{P(A_i \setminus B) * P(A_i)}{\sum P(B \setminus A_i) * P(A_i)} = \frac{P(A_i \cap B)}{\sum P(B \setminus A_i) * P(A_i)} \quad (14)$$

Bayes teoreminde B olayının gerçekleşmesi koşuluna bağlı olarak herhangi bir A_i olayının gerçekleşme olasılığı her bir bileşik olasılığın, bileşik olasılıkların toplamına bölünmesi ile elde edilir. Ayrıca, Bayes formülünde paydada yer alan değere toplam olasılık, bu durumda elde edilen $P(A_i \setminus B)$ koşullu olasılığına da son olasılıklar adı verilir (Aytaç, 2012; Bayram Arlı, 2018).

Naive Bayes genellikle olayların olasılıklarının değerlendirilmesi için kullanılan ve yeni bilgiye göre bu olasılıkların güncellenmesini sağlayan bir yöntemdir. Naive Bayes algoritması, Bayes teoremini sınıflandırma problemlerine uygulamak için basit bir yöntem tanımlar. Bayesian yöntemleri kullanan tek makine öğrenmesi yöntemi değildir ancak en yaygın olanıdır. Özellikle metin sınıflandırması için kullanılan bir yöntemdir (Lantz, 2015).

Naive Bayes, veri setindeki tüm özelliklerin eşit derecede önemli ve bağımsız olduğunu varsayar. Birden çok özellik ve koşul altında bir durumun olasılığını hesaplamak için Bayes Teoremi, koşullu olasılıkların çarpımı yoluyla hesaplanabilir. Bu, özellikle sınıflandırma veya tahmin problemlerinde kullanılır. Hastalık teşhisi, müşteri davranış analizi, enerji verimliliği tahmini gibi alanlarda hastaların yaşam tarzı, müşterinin demografik bilgileri, bir evin yalıtım kalitesi gibi birçok özellik koşuluna bağlı olarak bir hesaplama yapmak gerekebilir. Birden fazla özellik koşulu altındaki bir verinin belli bir sınıfa ait olma olasılığını hesaplamak için, Bayes Teoremini bu özelliklerin her biri için ayrı ayrı kullanmak gerekecektir. Her bir özellik, bir koşul olarak düşünülür ve bu koşullu olasılıkların çarpımı, sonuçta elde edilen toplam olasılığı verir. (Engin, 2022).

Verinin belli bir sınıfa (S) ait olma olasılığı şu şekilde hesaplanabilir:

$$P(S \setminus A, B, C) = \frac{P(A \setminus S) * P(B \setminus S) * P(C \setminus S) * P(S)}{P(A) * P(B) * P(C)} \quad (15)$$

Burada $P(S \setminus A, B, C)$ verinin S sınıfına ait olma olasılığını, $P(A \setminus S)$, $P(B \setminus S)$ ve $P(C \setminus S)$ her bir özellik için sınıf S altındaki koşullu olasılıkları $P(S)$ sınıf S'nin koşulsuz

olasılığını $P(A)$, $P(B)$ ve $P(C)$ her bir özelliğin koşulsuz olasılıklarını temsil eder (Peng vd., 2004).

NB sınıflayıcı, verilen bir veri noktasının hangi sınıfa ait olduğunu belirlemek için kullanılır. Özellikle metin sınıflandırma gibi uygulamalarda, bir belgenin bir kategoriye ait olup olmadığını saptamak için kullanılır. Dolayısıyla, NB sınıflayıcının odak noktası sınıf tahminidir. NB sınıflayıcı, veri noktasının farklı sınıflara ait olma olasılıklarını hesaplar, ancak genellikle bu olasılıkların karşılaştırılmasıyla çalışır. Yani, sınıf tahminini belirlemek için en yüksek olasılığa sahip sınıf seçilir. Diğer sınıfların olasılıkları da hesaplanır, ancak en yüksek olasılık üzerinde durulur. NB sınıflayıcı, basit bir model yapısına sahiptir ve hesaplamaları hızlı yapar. Bu nedenle, özellikle büyük veri kümelerinde hızlı ve etkili bir sınıflandırma sağlar. Ancak, kesin olasılık değerlerine çok fazla vurgu yapmaz ve bu nedenle genellikle "naive" (basit) olarak adlandırılır (Engin, 2022; Orhan & Adem, 2012) .

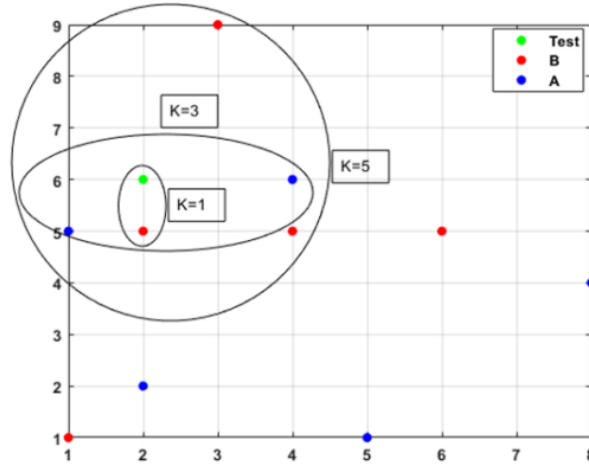
Örneklem dağılımının dengesiz olduğu veri kümeleri üzerinde çalışılabilen bir algoritma olması, veri sayısı az ya da çok olsa da sonuçlar için önemli bir kısıtlama olmaması, öğrenme ve test etme sürelerinin kısa olması, sınıfları hızlı ve kolay bir şekilde tahmin etmesi gibi olumlu taraflarının yanında test verisinde gözlemlenen bir değer eğitimi kümesinde hiç gözlemlenmemesi, yani sınıf frekansına sahip olmaması durumunda tahmin yapamaması, özellikler arasında bir tür bağımlılık olduğunda koşullu bağımsızlık varsayımının geçerli olmaması gibi olumsuz tarafları da vardır (Engin, 2022; Ng & Jordan, 2001).

2.2.8. *k-En Yakın Komşu*

k-en yakın komşu (k-Nearest Neighbor – k-NN) algoritması, örnek verilerin bulunduğu sınıfın ve en yakın komşunun, herhangi bir k değerine göre belirlendiği bir sınıflandırma yöntemidir. En eski, basit makine öğrenmesi yöntemlerinden biridir. Bu algoritmada karşılaşılan yeni bir örnek var olan örneklerle arasındaki benzerliğe göre sınıflandırılır (Bhatia, 2010; Joseph & Ramakrishnan, 2015).

k-NN algoritması, denetimli öğrenme kategorisine ait bir makine öğrenmesi yöntemidir. Bu algoritma, bir test örneği verildiğinde tahmin yapmak için eğitim örnekleri arasındaki benzerliği kullanır. k-NN algoritması bir veri noktasının sınıfını

belirlemek için seçilen bir özelliğin kendisine en yakın olan diğer özellikler arasındaki yakınlığı kullanır. k-NN algoritmasında kullanılan “k” parametresi, tahminlerin yapılması için kaç komşunun kullanılacağını belirtir. Genellikle k parametresi çoğunluğa dayalı tahminlerin yapılması ve sınıf etiketleri arasında eşit bir bölünme sağlanması için 1'den büyük ve tek sayı olarak tercih edilir (Bilgin & Eser, 2022b).



Şekil 21: K = 1-3-5 için Komşuluk Hesabı (Bilgin & Eser, 2022b).

k-NN algoritması, örneklerin n sayıda sayısal özelliklerle tanımlandığı eğitim verilerini kullanır. Tüm eğitim örnekleri, n boyutlu bir öznitelik uzayında bir noktayı temsil eder. Bilinmeyen bir örnek sınıflandırılmak istendiğinde, bu örneğe en yakın k komşuyu belirlemek için eğitim setindeki tüm örneklerin özellik uzayındaki mesafeleri hesaplanır. Bu k en yakın örnek, bilinmeyen örneğin sınıf etiketi belirlemek için kullanılır. Bu belirleme işlemi, k en yakın komşuların sınıf etiketlerinin çoğunluğu ile yapılır (Han vd., 2012).

k-NN algoritmasının performansı, örnekler arasındaki benzerliğin nasıl ölçüleceği ile yakından ilgilidir. Benzerlik, Öklid uzaklığı gibi farklı uzaklık ölçütleri kullanılarak hesaplanabilir. Bu, k-NN algoritmasında kritik bir faktördür. k-NN algoritmasının temel versiyonunda, sınıf etiketi çoğunluk oylamasına dayalı olarak belirlenir. Ancak, simetrik olmayan dağılıma sahip veri setlerinde, bazı sınıfların diğerlerine göre daha fazla örnek içermesi, yeni örneklerin sınıf etiketlerinin belirlenmesinde bu sınıfların daha belirleyici olmasına neden olabilir. Temel k-NN algoritmasının sınıf belirleme işlemi çoğunluk oylamasına dayalı olsa da simetrik olmayan dağılıma sahip veri setlerinde bazı sınıfların

daha baskın olması sebebiyle, farklı ağırlık değerleri atayan yöntemler de mevcuttur. Bu nedenle, k-NN algoritmasında uzaklık ölçütünün etki değerini farklı şekillerde ağırlıklandıran yöntemler geliştirilmiştir (Gärtner vd., 2004).

k-NN algoritması, sınıflandırma problemlerinde özellikle tercih edilen bir yöntemdir çünkü eğitim gerektirmez, uygulaması kolaydır ve analitik olarak takip edilebilir. Ayrıca, yerel bilgilere uyum sağlayabilir, paralel işlemcilerde çalışabilir ve gürültülü eğitim verilerine karşı dayanıklıdır. k-NN algoritması, sınıflandırma uygulamalarında önemli avantajlara sahip olsa da bazı dezavantajlar da mevcuttur. Özellikle, yüksek miktarda bellek alanı gerektirir, veri seti ve öznitelik boyutu arttıkça işlem yükü ve maliyet önemli ölçüde artar. Performansı, k komşu sayısı, uzaklık ölçütü ve öznitelik sayısı gibi parametrelere ve özelliklere bağlı olarak değişebilir (Bhatia, 2010).

ÜÇÜNCÜ BÖLÜM

UYUŞMAZLIK MAHKEMESİ KARARLARININ YAPAY ZEKÂ YÖNTEMLERİ İLE TAHMİNİ

Aristoteles, bütün anayasalarda üç unsurun bulunduğunu belirtir: Bu öğelerden birincisi, topluca görüşüp düşünmedir yani millet çapında önem taşıyan her şeyin tartışılmasıdır. İkincisi, yürütme, üçüncüsü ise yargı fonksiyonudur (Gözler, 2018). Locke, yasama organının, devlette egemenliğin temsilcisi olduğunu kabul etmekle birlikte, kanunların yapılmasında yürütmenin de yetki sahibi olmasını istemektedir. Bir devlette en büyük tehlike olan despotizm, yasama organıyla yürütme organının tek elde toplanmasından doğar. Bundan dolayı hukuk devletinde organların birbirinden ayrılması gerekir (Büyük, 2010).

Kuvvetler ayrılığı ilkesini ilk ortaya koyan Aristo olmakla birlikte ona son şeklini Montesquieu vermiştir. Locke'un fikirlerinden etkilenen Montesquieu, devletin üç erki olduğunu; yasama, yürütme ve yargıdan oluşan bu üç erkin kanun yapma işine katıldığını belirtmektedir (Büyük, 2010). Kuvvetler ayrılığı sistemi, bir devletin yönetiminde temel bir ilkedir ve kuvvetlerin farklı organlara ayrılmasını ve bu organların belirli yetki ve görevlere sahip olmasını ifade eder (Özbudun, 2008; Soysal, 1990; Teziç, 2003).

Devletimizde, yargı yetkisi, Türk milleti adına bağımsız ve tarafsız mahkemelerce kullanılır (Anayasa madde 9). Yargı biri organik ve şekli, diğeri maddi anlamda olmak üzere, iki noktadan hareket ederek tarif edilebilir. Organik anlamda yargı, yasama ve yürütme gibi devletin diğerk iki ana gücünden bağımsız bir şekilde faaliyet gösteren mahkemelerin ve yargı organlarının tüm faaliyetlerini ifade eder. Bu, yargının kendi içinde bağımsız bir kurum olduğu ve diğerk devlet organlarından ayrı bir şekilde çalıştığı anlamına gelir. Maddi anlamda yargı ise objektif hukukun somut bir olaya uygulanması anlamına gelir. Bu, mahkemelerin yasalara ve hukuki prensiplere dayalı olarak bir dava veya hukuki sorun hakkında karar verme işlemidir. Maddi anlamda yargı, mahkemelerin davalara, uyuşmazlıklara veya suçlara ilişkin kararlarını verme sürecini ifade eder (Kayacı, 2012; Kuru & Aydın, 2020).

Türkiye’de yargı sistemi alanında Fransız sistemini benimsemiştir ve ülkemizde “yargı ayrılığı sistemi” uygulanmaktadır. Yürütme organının yargısal denetimine ilişkin

kabul edilen iki temel sistemden biri olan “idari rejim” veya “yargı ayrılığı sistemi” yürütme organını denetleyecek mahkemelerin ayrı bir yargı kolu oluşturduğu bir sistemdir. Bu sistem, yürütme organının faaliyetlerinin yargı organları tarafından bağımsız bir şekilde denetlendiği ve yürütmenin hukuka uygunluğunun sağlandığı bir yapıyı ifade eder (Yeniay, 2020). Kesin hüküm verme yetkisine sahip ve başka bir mahkemeye tabi olmayan mahkemelere yüksek mahkeme, kararları aynı yüksek mahkemede temyiz edilen mahkemelerin oluşturduğu düzene de yargı kolu denilmektedir (Gözler, 2018). Yargısal denetim açısından ikinci olarak kabul edilen sistem, özel bir yargı kolunun olmadığı “yargı birliği sistemi” olarak adlandırılır. Bu sistem, “adli rejim” olarak da bilinir ve bu yaklaşıma sahip ülkelerde yargı organı tek bir yargı kolundan oluşur. Yürütme organının eylemleri ve özel hukuk anlaşmazlıkları, bu yargı kolundaki mahkemeler tarafından gözden geçirilir (Kaplan, 2017).

Türkiye’deki yargı kollarını tespit etmek için yüksek mahkemeleri belirlemek gerekir. Anayasamızın 146-160’ıncı maddelerinde yüksek mahkemeler, Anayasa Mahkemesi, Yargıtay, Danıştay ve UYM olarak sayılmıştır. O halde Türkiye’de yargı kolları, Anayasa Yargısı, adli yargı, idari yargı ve uyuşmazlık yargısı olarak ayrılmaktadır (Gözler, 2019).

Tablo 4: Türkiye’de Yargı Kolları (Gözler, 2019).

YARGI KOLU	ANAYASA YARGISI	ADLİ YARGI	İDARİ YARGI	UYUŞMAZLIK YARGISI
İlk Derece Mahkemesi	Anayasa Mahkemesi	Hukuk (Sulh- Asliye) Ceza (Asliye- Ağır)	İdare Mahkemeleri Vergi Mahkemeleri	Uyuşmazlık Mahkemesi
Ara Derece Mahkemesi		Bölge Adliye Mahkemeleri	Bölge İdare Mahkemeleri	
Üst Derece Mahkemesi		Yargıtay	Danıştay	

3.1. Uyuşmazlık Mahkemesinin Türk Yargı Sistemindeki Yeri ve Uyuşmazlık Mahkemesi Kararlarının Niteliği

Ülkemizde birden çok yargı kolu bulunduğu için bunların arasında uyuşmazlıklar çıktığında bu uyuşmazlıkları çözmesi gereken ve bu mahkemelerin üstünde olan bir mahkemenin varlığı gerekli hale gelmiştir (Gözübüyük & Tan, 2012). Bu gereklilikten UYM doğmuştur. Ülkemizde UYM ihtiyacı Tanzimat dönemiyle birlikte ortaya çıkmıştır. Daha önce, yalnızca İslam Hukuku uygulanırken, Gülhane Hattı Hümayunu ile birlikte Avrupa Hukuku ve idari yapıları da hukuk sistemimize dahil edilmiş ve böylece iki ayrı hukukun uygulandığı bir düzene geçilmiştir. Günümüzdeki adliye mahkemelerinin görevlerini yerine getirmek amacıyla “Divan-ı Ahkam-ı Adliye” ve kişi ile devlet arasındaki anlaşmazlıkların çözüme kavuşturulması için de “Şurayı Devlet” oluşturulmuştur (Anıl, 1992).

14.07.1945 tarihinde 4788 Sayılı Kanun’la UYM kurulmuş ve 1961 Anayasası ile UYM’ne anayasal bir hüviyet kazandırılmıştır. Yargı erkini Anayasa’nın kendilerine tanıdığı görev ve yetki alanlarıyla sınırlı olarak paylaştan, yüksek mahkemeler arasında yer alan UYM, Anayasa’nın 158. (1961 Anayasası 142.) maddesiyle görevlendirilmiş, *adli ve idari yargı mercileri arasındaki görev ve hüküm uyuşmazlıklarını kesin olarak çözmeye yetkili, bağımsız bir yüksek mahkemedir*. Bu tanıma 12.06.1979 tarihli 2247 Sayılı Uyuşmazlık Mahkemesinin Kuruluş ve İşleyişi Hakkında Kanun’un 1. Maddesinde de “*Uyuşmazlık Mahkemesi; Türkiye Cumhuriyeti Anayasası ile görevlendirilmiş, adli ve idari yargı mercileri arasındaki görev ve hüküm uyuşmazlıklarını kesin olarak çözmeye yetkili ve bu kanunla kurulup görev yapan bağımsız bir yüksek mahkemedir*” şeklinde yerini bulmuştur.

UYM, yargı ayrılığı ilkesi gereği ortaya çıkan görev anlaşmazlıklarını çözerek, kişilerin hak arama özgürlüğünü yerine getirmeyi amaçlayan ve hüküm uyuşmazlıklarını çözerek hukuki engelleri gideren özel yetkili bir yüksek mahkemedir. Bu nedenle, Yargıtay ve Danıştay gibi diğer yüksek mahkemelerin kararlarını iptal edip yerlerine hüküm verebilen tek yargı organıdır (Gözler, 2019).

UYM nin kuruluşu 12.06.1979 tarihli 2247 Sayılı Uyuşmazlık Mahkemesinin Kuruluş ve İşleyişi Hakkında Kanun’un 3. Maddesinde “*UYM bir Başkan ile altı asıl, altı yedek üyeden kurulur. UYM Başkanı, Anayasa Mahkemesince kendi üyeleri arasından*

seçilir. UYM'ne, Yargıtay Hukuk Genel Kurulu ile Danıştay Genel Kurulunca kendi daire başkan ve üyeleri arasından üçer asıl, üçer yedek üye seçilir. Toplantı yetersayısı, asıl üyelerin mazeretleri halinde yedek üyeler alınmak suretiyle sağlanır.” şeklinde anlatılmıştır.

UYM görüşmeleri ve kararları da 12.06.1979 tarihli 2247 Sayılı Uyuşmazlık Mahkemesinin Kuruluş ve İşleyişi Hakkında Kanun'un 26. Maddesinde “UYM'nde, üye tam sayısı hazır bulunmadıkça görüşmelere başlanamaz ve karar verilemez. Eksik üyenin yeri o üyenin geldiği yargı merciinin yedek üyesi ile, bu olmadığı takdirde başka bir yedek üye ile doldurulur. Mahkeme, her toplantı dönemi başında, yıllık çalışma planını yapar. UYM'nde incelemeler dosya üzerinde yapılır ve bu kanunda belli edilenler dışında çoğunlukla karar verilir.” şeklinde anlatılmıştır.

T.C.
UYUŞMAZLIK MAHKEMESİ

ESAS NO : 2018 / 530
KARAR NO : 2018 / 467
KARAR TR : 24.9.2018

ÖZET : Karayolunda meydana gelen trafik kazasında uğranılan maddi ve manevi zararın tazmini istemiyle açılan davanın, 2918 sayılı Yasanın 19.01.2011 tarihinde yürürlüğe giren 110. maddesi kapsamında ADLİ YARGI YERİNDE çözümülenmesi gerektiği hk.

K A R A R

Davacı lar : M. A., M. A.
Vekili : Av. H.E. A.
Davalı : Ulaştırma Denizcilik ve Haberleşme Bakanlığı-Karayolları Genel Müdürlüğü
Vekili : Av. M. Ş. H., Av.N. B.

O L A Y : Davacılar vekili dilekçesinde; karar verilmesi istemiyle 15.9.2017 tarihinde idari yargı yerinde dava açmıştır.

AKSARAY İDARE MAHKEMESİ; usulleri ile diğer hükümleri kapsadığı ve bu kanunun karayollarında uygulanacağı belirtilmiştir.

.....

Davacı vekili bu kez aynı istemle adli yargı yerinde dava açmıştır.

AKSARAY 4.ASLİYE HUKUK MAHKEMESİ:usul yönünden reddi gerektiği dosyanın 2247 sayılı Uyuşmazlık Mahkemesi'nin Kuruluşu ve İşleyişi Hakkında Kanun'un 19. maddesi gereğince yargı yerinin belirlenmesi için Uyuşmazlık Mahkemesi'ne gönderilmesi gerektiği kanaatine varılmış ve aşağıdaki şekilde hüküm kurulmuştur.

İNCELEME VE GEREKÇE:

Uyuşmazlık Mahkemesi'nin, Hicabi DURSUN'un Başkanlığında, Üyeler: Şükrü BOZER, Mehmet AKSU, Birol SONER, Süleyman Hilmi AYDIN, Aydemir TUNÇ ve Turgay Tuncay VARLI'nın katılımlarıyla yapılan 24.9.2018 günlü toplantısında:

I-İLK İNCELEME: ...

II-ESASIN İNCELENMESİ: ...dinlendikten sonra GEREĞİ GÖRÜŞÜLÜP DÜŞÜNÜLDÜ: ...görevsizlik kararının kaldırılması gerekmiştir.

S O N U Ç : Davanın çözümünde ADLİ YARGININ görevli olduğuna, bu nedenle Aksaray 4.Asliye Hukuk Mahkemesinin BAŞVURUSUNUN REDDİ ile aynı Mahkemece ayrıca verilen ... gün ve sayılı GÖREVSİZLİK KARARININ KALDIRILMASINA, gününde Üyenın KARŞI OYU ve OY ÇOKLUĞU İLE KESİN OLARAK karar verildi.

Başkan Hicabi DURSUN	Üye Şükrü BOZER	Üye Mehmet AKSU	Üye Birol SONER
----------------------------	-----------------------	-----------------------	-----------------------

Üye Süleyman Hilmi AYDIN	Üye Aydemir TUNÇ	Üye Turgay Tuncay VARLI
--------------------------------	------------------------	-------------------------------

KARŞI OY

3.11.2015 tarih ve 29521 sayılı Resmi Gazetede yayımlanan Uyuşmazlık Mahkemesi'nin 28.09.2015 gün ve E:2015/580 K:2015/592 sayılı kararında belirtilmiş olan düşünce doğrultusunda sayın çoğunluğun görüşüne katılamıyorum. 24.9.2018

ÜYE

Turgay Tuncay VARLI

Şekil 22 : UYM Karar Örneği

UYM, “hukuk” ve “ceza” bölümleriyle ayrılmakta olup, hukukla ilgili anlaşmazlıklar hukuk bölümünde, ceza ile ilgili anlaşmazlıklar ise ceza bölümünde ele alınarak çözümlenmektedir. Adli ve idari yargı mercileri tarafından, görevle ilgili olmaksızın kesin olarak verilmiş veya kesinleşmiş, aynı konuya ve sebebe ilişkin, taraflarından en az biri aynı olan ve kararlar arasındaki çelişki yüzünden hakkın yerine getirilmesi olanaksız bulunan hallerde hüküm uyuşmazlığının varlığı kabul edilir. Görev

uyuşmazlıkları ise olumlu ve olumsuz olmak üzere ikiye ayrılır. Olumlu görev uyuşmazlığı, aynı taraflar, sebep ve konuya sahip olan bir davada iki ayrı yargı düzenindeki mahkemelerin görevli olduklarına karar vermeleri ve davaya devam etmeleri durumu olarak tanımlanabilirken olumsuz görev uyuşmazlığı konusu, sebebi ve tarafları aynı olan bir davada, en az iki adli veya idari yargı yerinin kendilerini görevsiz ilan etmeleri ve verilen kararların kesin veya kesinleşmiş olması durumu şeklinde ifade edilmektedir (Dinç, 2016; Kuru & Aydın, 2020).

Bahsedilen bu uyuşmazlıkların yolları ve inceleme kurallarının anlatıldığı 2247 Sayılı Uyuşmazlık Mahkemesinin Kuruluş ve İşleyişi Hakkında Kanunda olumlu görev uyuşmazlığı çıkarma “*Görev uyuşmazlığı çıkarma; adli ve idari bir yargı merciinde açılmış olan davada ileri sürülen görev itirazının reddi üzerine ilgili Başsavcı veya Başkanunsözcüsü tarafından görev konusunun incelenmesinin Uyuşmazlık Mahkemesinden istenmesidir.*” şeklinde tanımlanmıştır. Aynı kanunun 14. maddesinde de olumsuz görev uyuşmazlığı “*Olumsuz görev uyuşmazlığının bulunduğu ileri sürülebilmesi için adli ve idari yargı mercilerinin tarafları, konusu ve sebebi aynı olan davada kendilerini görevsiz görmeleri ve bu yolda verdikleri kararların kesin veya kesinleşmiş olması gerekir.*” şeklinde ifade edilmiştir.

3.2. Uyuşmazlık Mahkemesi Kararlarının Yapay Zekâ Yöntemleri İle Tahmini

3.2.1. Veri Toplama Yöntemi

UYM almış olduğu kararları “<https://kararlar.uyusmazlik.gov.tr/>” adresinde yayınlamaktadır. Bu çalışma kapsamına sınıflandırma (classification) teknikleri için uygun olan hukuk bölümü kararlarından karar sonucu olarak adli yargı-idari yargı ayrımı yapan olumsuz görev uyuşmazlığı kararları dikkate alınmıştır.

Karar Sayısı	Esas Sayısı	Bölüm	Uygunsuzluk	Karar Sonucu
1985024	1985023	Hukuk Bölümü	Görev Uygunsuzluğu	Bazıyıcı Red - 1. Maddeye göre
2016024	2016021	Hukuk Bölümü	Görev Uygunsuzluğu	Adli - 18. Maddeye göre
2016028	2016012	Hukuk Bölümü	Görev Uygunsuzluğu	Adli - 18. Maddeye göre
2016094	2016093	Hukuk Bölümü	Görev Uygunsuzluğu	Adli - 18. Maddeye göre
208603	208606	Hukuk Bölümü	Görev Uygunsuzluğu	Adli - 14. Maddeye göre
2086029	208606	Hukuk Bölümü	Görev Uygunsuzluğu	Adli - 14. Maddeye göre
2017014	2017044	Hukuk Bölümü	Görev Uygunsuzluğu	Adli - 14. Maddeye göre
198011	198010	Hukuk Bölümü	Görev Uygunsuzluğu	Bazıyıcı Red - 18. Maddeye göre
2017045	2017004	Hukuk Bölümü	Görev Uygunsuzluğu	Adli - 14. Maddeye göre
2016092	2016153	Hukuk Bölümü	Görev Uygunsuzluğu	Adli - 18. Maddeye göre

Şekil 23: UYM Kararları Kamuya Açık Portal Sayfası

İnternet ortamında açık erişimli verilerin düzenli bir şekilde işlenebilmesi için kullanıcılar tarafından tek tek indirilmesi gerekmektedir. Veri sayısı çok fazla olduğu zaman insanlar tarafından bunun sorgusunun ve tek tek indirilme işlemlerinin yapılması, ardından dosyalar halinde kaydedilmesi ve dosyaların içeriğinin düzenli bir veri kümesi haline getirilmesi uzun bir zaman ve efor gerektirmektedir. Bu tür veri toplama ve işleme işlemlerinde web botları veya diğer adıyla web tarayıcıları otomasyon araçları kullanılarak insan işgücünden önemli ölçüde tasarruf edilir ve işlem hızı artırılır. Web botu, otomatikleştirilmiş bir yazılım programı veya robot olarak tanımlanır ve internet üzerinde otomatik olarak belirli görevleri gerçekleştiren bir bilgisayar programıdır. Web botları, belirli bir işlemi veya görevi tekrarlayarak, insan müdahalesi olmadan internet üzerinde veri toplama, analiz veya işlem yapma yeteneğine sahiptir (Suchacka vd., 2021).

Web bot yazılımları geliştirilirken çeşitli teknolojiler mevcuttur fakat bu teknolojiler içinde istemci sunucu mimarisinde en hızlı cevap veren teknolojilerinden bir tanesi Python temelli **Selenium** kütüphanesidir (Muthukadan, 2023). Selenium kütüphanesini kullanırken öncelikle ayarlanan bot tarafından site adresine ulaşım sağlanmıştır. Site adresine ulaşım sağlandıktan sonra ulaşılan site üzerinde bulunan seçim ve yazım kutucukları yani sorgu yapılacak koşulların seçilme işlemi yapılmıştır.

Şekil 24: UYM Kamuya Açık Platformunda Seçim ve Yazım Kutucukları

Bu çalışma kapsamında UYM kararlarında bölüm olarak hukuk bölümü, karar sonucu olarak adli ve idari, uyuşmazlık olarak da görev uyuşmazlığı seçilmiştir. Seçimler sonucunda kamuya açık platformda ekrana 5 koşullu sonuçlar çıkmıştır.

Karar Sayısı	Esas Sayısı	Bölüm	Uyuşmazlık	Karar Sonucu
2018/467	2018/530	Hukuk Bölümü	Görev Uyuşmazlığı	Adli - 19. Maddeye göre
2016/210	2016/163	Hukuk Bölümü	Görev Uyuşmazlığı	Adli - 19. Maddeye göre
2015/795	2015/780	Hukuk Bölümü	Görev Uyuşmazlığı	Adli - 14. Maddeye göre
2015/789	2015/774	Hukuk Bölümü	Görev Uyuşmazlığı	Adli - 14. Maddeye göre
2015/749	2015/370	Hukuk Bölümü	Görev Uyuşmazlığı	İdari - 14. Maddeye göre
2015/805	2015/790	Hukuk Bölümü	Görev Uyuşmazlığı	İdari - 19. Maddeye göre
2015/811	2015/796	Hukuk Bölümü	Görev Uyuşmazlığı	İdari - 14. Maddeye göre
2015/834	2015/823	Hukuk Bölümü	Görev Uyuşmazlığı	Adli - 14. Maddeye göre

Şekil 25: UYM Kamuya Açık Platformunda Koşullar

Bu koşulların web bot tarafından seçilip ardından sorgu yapılması sağlanmıştır. Bu yapılırken öncelikle ilgili web sitesinin arka planındaki HTML etiketlerine ulaşılmıştır. UYM kamuya açık platformunda kullanılan HTML etiketleri sabittir. Bu HTML etiketlerine sadece ilk erişimde ulaşılmış ve istenilen kriterler konulmuştur.

Kriterlere ait bilgiler verilen sıraya göre (karar sayısı, esas sayısı, bölüm, uyuşmazlık, karar sonucu) hepsinin arasına belirli bir süre (1 sn) konularak çekilmiştir. Bu çekilme esnasında araya belirli bir süre konularak seçilen davaların üst süte gelmesi engellenmiştir. Aksi takdirde web bot, bir kriteri çekmeden diğerini sorgulamaya çalışacak ve bir sorun teşkil edecektir.

Bu işlemlerden sonra web bot çalıştırılmış ve sorguların sonuçları gelmeye başlamıştır. Kayıtlar yapıldıktan sonra karar sayısı, esas sayısı, bölüm sayısı, uyuşmazlık ve karar sonucu şeklinde 5 tane ayrı veri gelmektedir. UYM kamuya açık platformunda çekilen her bir karar için oluşturulan PDF belgelerinin farklı bir formatta olduğu görülmüştür. Bu nedenle sadece bir kalıp oluşturamamış, her bir PDF belgesinin içinde farklı formatta olan içerikleri önileme yöntemleriyle düzenlenmiştir. Düzenlenen belgelerde bulunan her bir koşul değeri için ayrı bir sütun olacak şekilde bir EXCEL tablosu oluşturulmuştur. Bu çalışma kapsamında karar içeriği ile ilgilenildiği için PDF belgesinin içinde bulunan OLAY ile başlayan kısım ayrı bir sütun olarak alınmıştır. Buradaki karar sonucu sütunu hedef değişken, diğer sütunlar öznitelik değişkeni niteliğindedir. Bu çalışma kapsamında Olay sütunu öznitelik değişkeni olarak alınmıştır.

B	E	G	H	I	J
kararsayısı	Olay	esassayısı	Bolum	Uyusmazlik	Kararsonucu
2018/467	O L A Y	2018/530	Hukuk Bölümü	Görev Uyuşmazlığı	Adli
2016/210	O L A Y	2016/163	Hukuk Bölümü	Görev Uyuşmazlığı	Adli
2015/795	O L A Y	2015/780	Hukuk Bölümü	Görev Uyuşmazlığı	Adli
2015/789	O L A Y	2015/774	Hukuk Bölümü	Görev Uyuşmazlığı	Adli
2015/749	O L A Y	2015/370	Hukuk Bölümü	Görev Uyuşmazlığı	İdari
2015/805	O L A Y	2015/790	Hukuk Bölümü	Görev Uyuşmazlığı	İdari
2015/811	O L A Y	2015/796	Hukuk Bölümü	Görev Uyuşmazlığı	İdari
2015/834	O L A Y	2015/823	Hukuk Bölümü	Görev Uyuşmazlığı	Adli
2015/769	O L A Y	2015/751	Hukuk Bölümü	Görev Uyuşmazlığı	İdari
2015/791	O L A Y	2015/776	Hukuk Bölümü	Görev Uyuşmazlığı	Adli
2015/758	O L A Y	2015/677	Hukuk Bölümü	Görev Uyuşmazlığı	Adli
2015/775	O L A Y	2015/758	Hukuk Bölümü	Görev Uyuşmazlığı	İdari
2015/770	O L A Y	2015/752	Hukuk Bölümü	Görev Uyuşmazlığı	Adli
2015/826	O L A Y	2015/813	Hukuk Bölümü	Görev Uyuşmazlığı	Adli
2015/783	O L A Y	2015/766	Hukuk Bölümü	Görev Uyuşmazlığı	Adli
2015/831	O L A Y	2015/820	Hukuk Bölümü	Görev Uyuşmazlığı	Adli
2015/782	Olayda, da	2015/765	Hukuk Bölümü	Görev Uyuşmazlığı	Adli
2015/761	O L A Y	2015/704	Hukuk Bölümü	Görev Uyuşmazlığı	İdari
2015/760	O L A Y	2015/700	Hukuk Bölümü	Görev Uyuşmazlığı	Adli
2015/794	O L A Y	2015/779	Hukuk Bölümü	Görev Uyuşmazlığı	Adli
2015/796	O L A Y	2015/781	Hukuk Bölümü	Görev Uyuşmazlığı	Adli
2015/786	O L A Y	2015/770	Hukuk Bölümü	Görev Uyuşmazlığı	Adli
2015/785	O L A Y	2015/768	Hukuk Bölümü	Görev Uyuşmazlığı	Adli
2015/801	O L A Y	2015/786	Hukuk Bölümü	Görev Uyuşmazlığı	Adli
2015/828	O L A Y	2015/817	Hukuk Bölümü	Görev Uyuşmazlığı	Adli
2015/838	O L A Y	2015/830	Hukuk Bölümü	Görev Uyuşmazlığı	Adli
2015/787	O L A Y	2015/782	Hukuk Bölümü	Görev Uyuşmazlığı	Adli

Şekil 26: Çekilen Kararlarla İlgili Excel Tablosu

Bu çalışma kapsamında değerlendirilen 6020 adet belgenin okunup işlenmesi yaklaşık 12 saat sürmektedir. Fakat ilgili verilerin alındığı platformda veriler yapılandırılmamış halde bulunmaktadır. Bu nedenle tutarlı ve anlamlı sonuçlar elde edebilmek için öncelikle bu veriler üzerinde çeşitli veri ön işleme adımları uygulanmıştır. Bu ön işleme adımları temel olarak veri temizleme, veri bütünleştirme, veri değiştirme ve veri azaltma adımlarından oluşmaktadır. Farklı şekil formatının ne olduğunun bulunup bu formata göre ayrı bir ön işleme kodu yazılmıştır. Bu şekilde manuel olarak düzeltmelerle veri indirme işlemi yaklaşık 1 ay sürmüştür.

Elde edilen metinsel verilerin veri ön işleme adımlarından sonra doğal dil işleme süreçlerinde etkin bir şekilde kullanılabilmesi için farklı birkaç adım daha uygulanmıştır. Bu kapsamda metnin içerisinde bulunan durağan kelimelerin çıkarılması, metnin küçük harflere dönüştürülmesi ve eğer varsa içinde anlamsız kelime ve karakterlerden arındırılması, metnin kelime bazlı olarak parçalanması ve kelime köklerinin ayrıştırılması işlemleri yapılmıştır. Metin ön işleme adımları yapılırken Python kütüphanelerinden faydalandığı gibi literatürde daha etkin sonuçların elde edildiği görülen **KUSH** gibi metin ön işleme yöntemlerinden de yararlanılmıştır (Uçkan & Karıcı, 2020).

Tablo 5 : Ham Veri ve İşlenmiş Veri

Davacılar vekili dilekçesinde; 08.08.2013 günü sürücü M. A.'nın sevk ve idaresindeki 68 ... 677 plaka sayılı araç ile plakası tespit edilemeyen yabancı plakalı aracın çarpışması neticesinde, 68 ... 677 plakalı aracın şarampole yuvarlandığını, sürücü ile araçta yolcu olarak bulunan M. A.'nın ağır şekilde yaralandığını	dava vekil dilekçe gün sür m a.nın sevk idare plaka sayı araç plaka tespit etmek yabancı plaka ara çarpışmak netice plaka ara şarampol yuvarlamak sür araç yol olmak bulmak m a.nın ağır şekil yara
HAM VERİ	ÖN İŞLENMİŞ VERİ

Makine öğrenmesi algoritmalarının doğru ve anlamlı sonuçlar üretebilmesi için önceden tanımlanmış bir dizi veri kullanılarak eğitilmesi gerekmektedir. Ancak doğal dil işleme üzerinde çalışmanın en önemli problemlerinden biri makine öğrenmesi algoritmalarının doğrudan metin verileri üzerinde çalışamamasıdır. Bu problemin çözümü için verilerin matematiksel formda temsil edilmesi gerekmektedir. Verilerin bu formlara dönüştürülmesi için öznitelik çıkarma (Feature Extraction) ve kelime temsil (Word Embedding) yöntemleri kullanılmıştır. Öznitelik çıkarma yöntemlerinden kelime

çantası (Bag of Words) ve TF-IDF yöntemlerinden faydalanılmakla beraber kelime temsil yöntemlerinden Word2Vec ve Glove yöntemleri kullanılmıştır. Böylece veri kümesinde bulunan hedef ve öznitelik değişkenleri vektörel hale gelmiştir. Öznitelik değişkeni 4515x30616 tipinde bir matrise dönüşmüştür.

```
[12] %%Tunning CountVectorizer
      #vect_tunned = CountVectorizer(stop_words='english', ngram_range=(1,2), min_df=0.1, max_df=0.7, max_features=100)
      #vect_tunned.fit(x_train)
      x_train_dtm = vect.transform(x_train)
      x_test_dtm = vect.transform(x_test)
```

```
[13] x_train_dtm
```

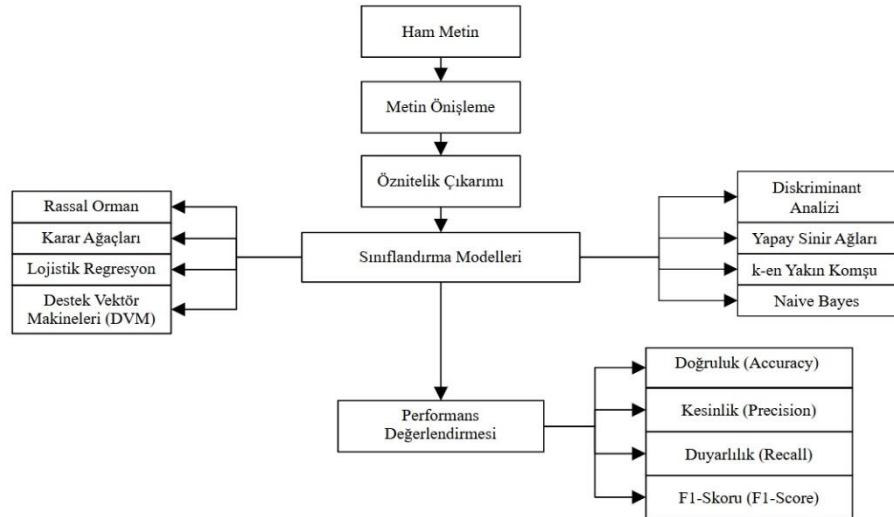
```
<4515x30616 sparse matrix of type '<class 'numpy.int64''
with 456953 stored elements in Compressed Sparse Row format>
```

```
y_train
```

```
2558  0
2085  0
1472  0
5625  0
1245  1
..
3884  0
5361  1
5398  0
5569  1
866   0
Name: target_encoded, Length: 4515, dtype: int64
```

Şekil 27 : Öznitelik ve Hedef Değişkenlerin Python Çıktısı

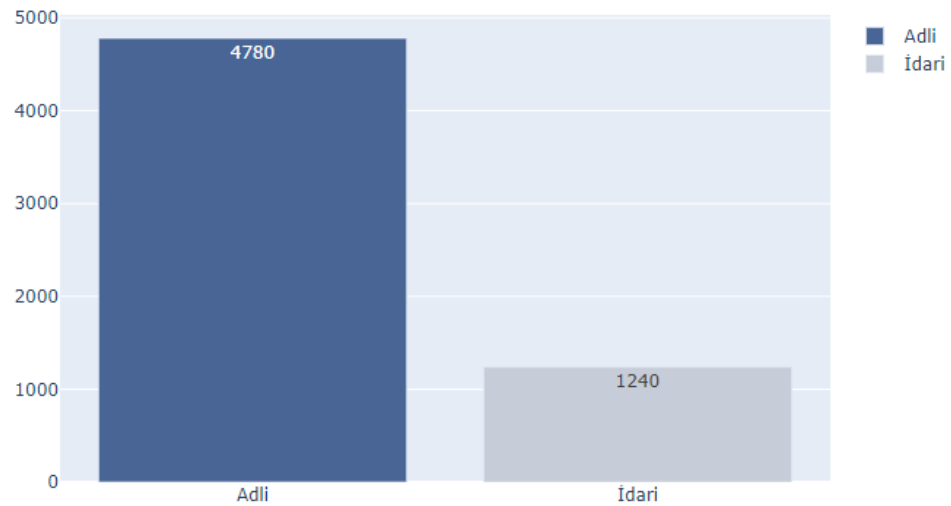
Belirtilen bu adımların tamamlanmasının ardından doğal dil işlemeye uygun makine öğrenmesi algoritmaları kullanılarak sınıflandırma işlemleri ve ardından da performans değerlendirmeleri yapılmıştır.



Şekil 28: Metin sınıflandırma akış şeması

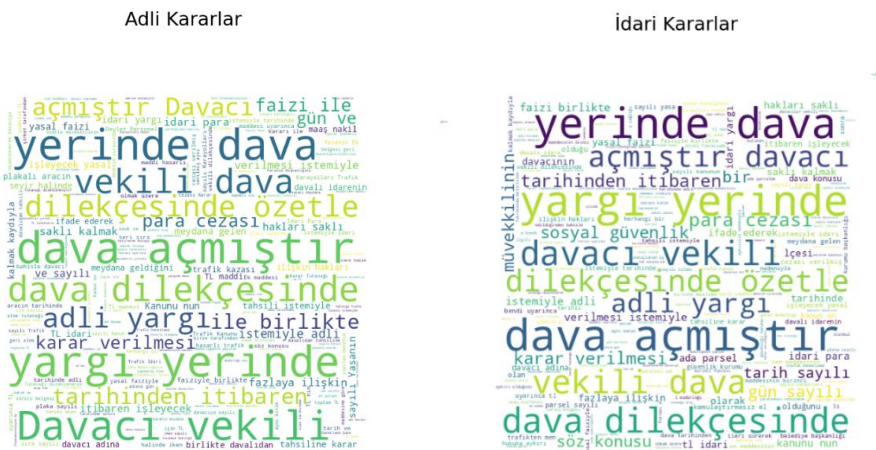
3.2.2. Yapay Zekâ Yöntemleri ile Tahmin

UYM'nin web ortamında bulunan olumsuz görev uyuşmazlığı “adli” ve “idari” kararlarından 6020 adedi analiz için alınmıştır. Bunların 4780 adedi adli, 1240 adedi idari karar niteliğindedir.



Şekil 29: Veri setinin dağılımı

Sonucu adli veya idari olan kararların en çok kullanıldığı kelimeler (kelime bulutları- word cloud) aşağıdaki şekil 30’da gösterildiği gibi bulunmuştur.

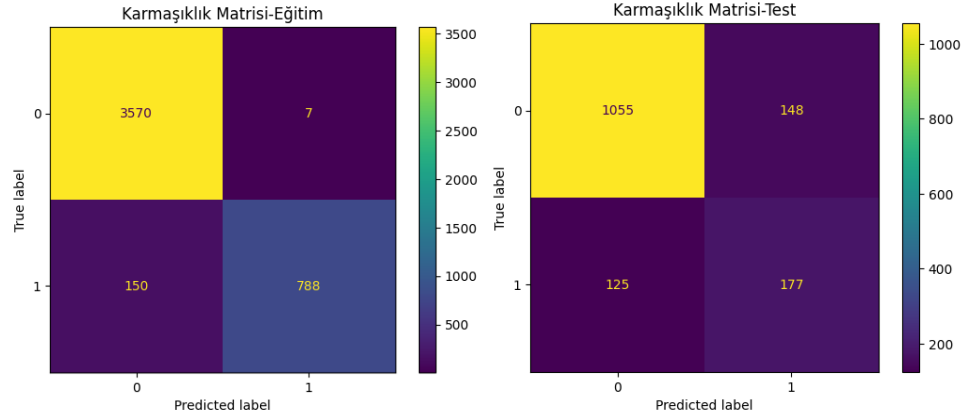


Şekil 30: Adli ve İdari Kararlardaki Kelime Bulutları

Bu kararların %75'i yani 4515 tanesi eğitim için ayrılmıştır. Eğitim verisi yardımıyla karar içerikleri ve sonuç arasındaki ilişkiyi öğrenen model oluşturulmuştur. Geriye kalan 1505 adet mahkeme kararı da oluşturulan modelin testi için kullanılmıştır.

Veri ön işleme adımından geçirilen kararlara sınıflandırma algoritmaları uygulanmıştır. Verilerin etiketlenilmesi için, adli kararlara 0, idari kararlara 1 değeri atanmıştır. Sınıflandırma algoritmalarının sonuçları aşağıda gösterilmiştir.

Karar Ağaçları Algoritması Sonuçları



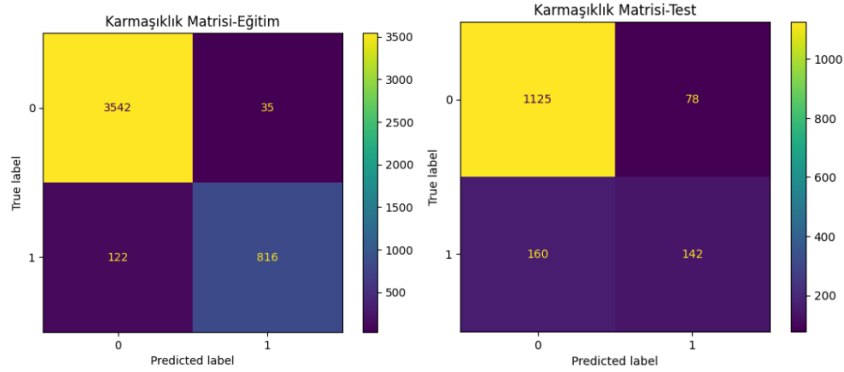
Şekil 31: Karar Ağaçları Algoritması Karmaşıklık Matrisi

Eğitim seti karmaşıklık matrisinde de görüldüğü gibi 3577 tane adli kararın 3570 tanesi, 938 tane idari kararın 788 tanesi doğru tahmin edilmiştir. Test seti karmaşıklık matrisinde de görüldüğü gibi 1203 tane adli kararın 1055 tanesi, 302 tane idari kararın 177 tanesi doğru tahmin edilmiştir.

Tablo 6: Karar Ağaçları Algoritması Değerlendirme Metrikleri

	Kesinlik (Precision)	Duyarlılık (Recall)	F1-Skoru (F1-Score)
Adli Kararlar	0.89	0.88	0.89
İdari Kararlar	0.54	0.59	0.56

Rassal Orman Algoritması Sonuçları



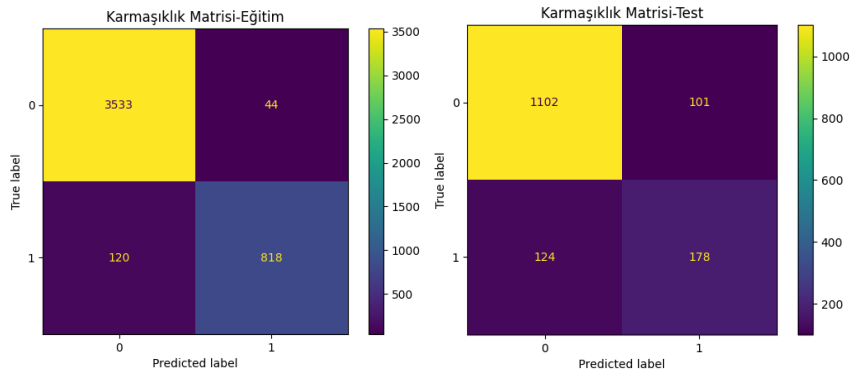
Şekil 32: Rassal Orman Algoritması karmaşıklık matrisi

Eğitim seti karmaşıklık matrisinde de görüldüğü gibi 3577 tane adli kararın 3542 tanesi, 938 tane idari kararın 816 tanesi doğru tahmin edilmiştir. Test seti karmaşıklık matrisinde de görüldüğü gibi 1203 tane adli kararın 1125 tanesi, 302 tane idari kararın 142 tanesi doğru tahmin edilmiştir.

Tablo 7: Rassal Orman Algoritması Değerlendirme Metrikleri

	Kesinlik (Precision)	Duyarlılık (Recall)	F1-Skoru (F1-Score)
Adli Kararlar	0.88	0.94	0.90
İdari Kararlar	0.65	0.47	0.54

Lojistik Regresyon Algoritması Sonuçları



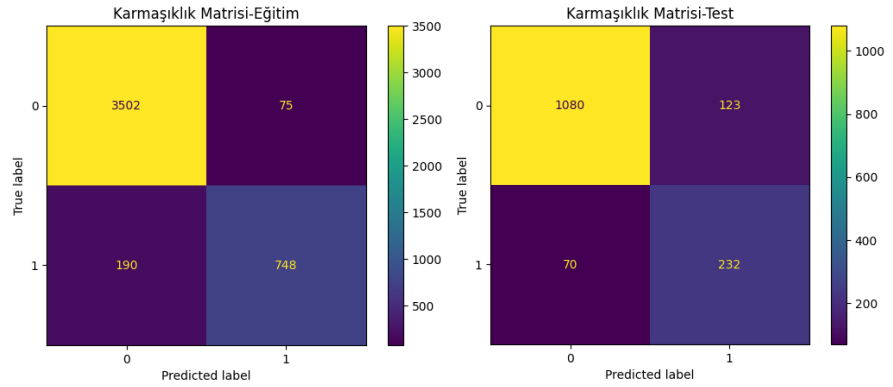
Şekil 33: Lojistik Regresyon Algoritması Karmaşıklık Matrisi

Eğitim seti karmaşıklık matrisinde de görüldüğü gibi 3577 tane adli kararın 3533 tanesi, 938 tane idari kararın 818 tanesi doğru tahmin edilmiştir. Test seti karmaşıklık matrisinde de görüldüğü gibi 1203 tane adli kararın 1102 tanesi, 302 tane idari kararın 178 tanesi doğru tahmin edilmiştir.

Tablo 8: Lojistik Regresyon Algoritması Değerlendirme Metrikleri

	Kesinlik (Precision)	Duyarlılık (Recall)	F1-Skoru (F1-Score)
Adli Kararlar	0.90	0.92	0.91
İdari Kararlar	0.64	0.59	0.61

Destek Vektör Makineleri Algoritması Sonuçları



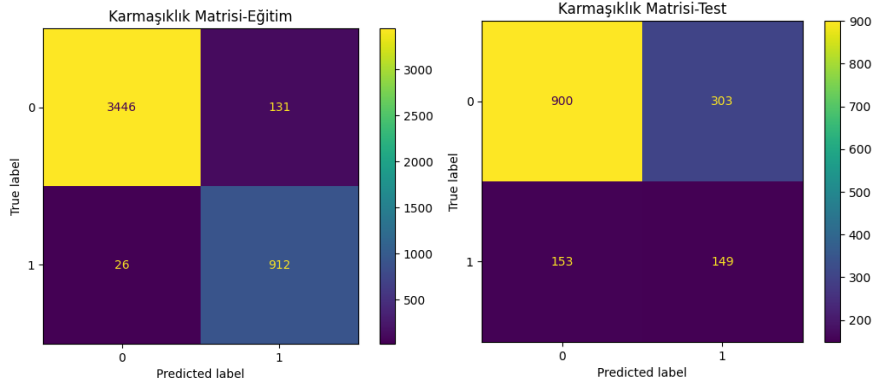
Şekil 34: Destek Vektör Makineleri Algoritması Karmaşıklık Matrisi

Eğitim seti karmaşıklık matrisinde de görüldüğü gibi 3577 tane adli kararın 3502 tanesi, 938 tane idari kararın 748 tanesi doğru tahmin edilmiştir. Test seti karmaşıklık matrisinde de görüldüğü gibi 1203 tane adli kararın 1080 tanesi, 302 tane idari kararın 232 tanesi doğru tahmin edilmiştir.

Tablo 9: Destek Vektör Makineleri Algoritması Değerlendirme Metrikleri

	Kesinlik (Precision)	Duyarlılık (Recall)	F1-Skoru (F1-Score)
Adli Kararlar	0.94	0.90	0.92
İdari Kararlar	0.65	0.77	0.71

Diskriminant Analizi Algoritması Sonuçları



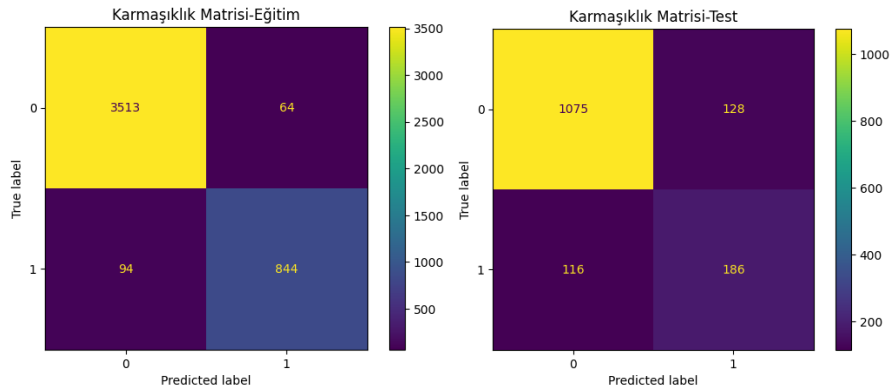
Şekil 35: Diskriminant Analizi Algoritması Karmaşıklık Matrisi

Eğitim seti karmaşıklık matrisinde de görüldüğü gibi 3577 tane adli kararın 3446 tanesi, 938 tane idari kararın 912 tanesi doğru tahmin edilmiştir. Test seti karmaşıklık matrisinde de görüldüğü gibi 1203 tane adli kararın 900 tanesi, 302 tane idari kararın 149 tanesi doğru tahmin edilmiştir.

Tablo 10: Diskriminant Analizi Algoritması Değerlendirme Metrikleri

	Kesinlik (Precision)	Duyarlılık (Recall)	F1-Skoru (F1-Score)
Adli Kararlar	0.85	0.75	0.80
İdari Kararlar	0.33	0.49	0.40

Yapay Sinir Ağları Algoritması Sonuçları



Şekil 36: Yapay Sinir Ağları Algoritması Karmaşıklık Matrisi

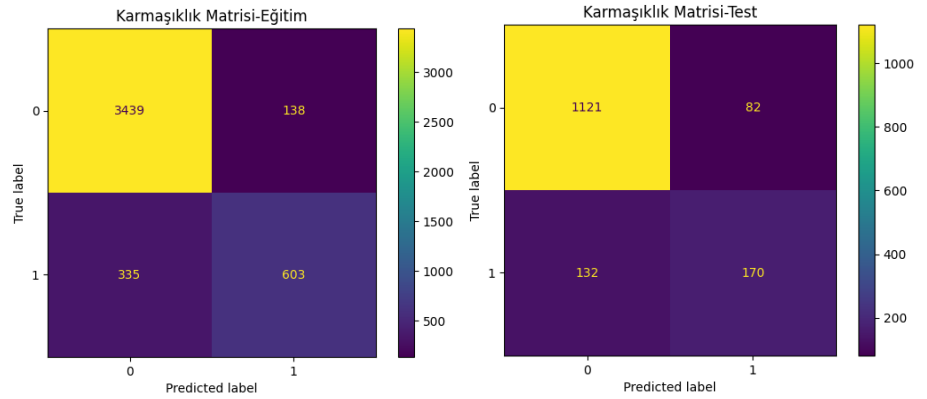
Eğitim seti karmaşıklık matrisinde de görüldüğü gibi 3577 tane adli kararın 3513 tanesi, 938 tane idari kararın 844 tanesi doğru tahmin edilmiştir. Test seti karmaşıklık matrisinde de görüldüğü gibi 1203 tane adli kararın 1075 tanesi, 302 tane idari kararın 186 tanesi doğru tahmin edilmiştir.

Tablo 11: Yapay Sinir Ağları Algoritması Değerlendirme Metrikleri

	Kesinlik (Precision)	Duyarlılık (Recall)	F1-Skoru (F1-Score)
Adli Kararlar	0.90	0.89	0.90
İdari Kararlar	0.59	0.62	0.60

k-en Yakın Komşu Algoritması Sonuçları

Eğitim seti karmaşıklık matrisinde de görüldüğü gibi 3577 tane adli kararın 3439 tanesi, 938 tane idari kararın 603 tanesi doğru tahmin edilmiştir. Test seti karmaşıklık matrisinde de görüldüğü gibi 1203 tane adli kararın 1121 tanesi, 302 tane idari kararın 170 tanesi doğru tahmin edilmiştir.

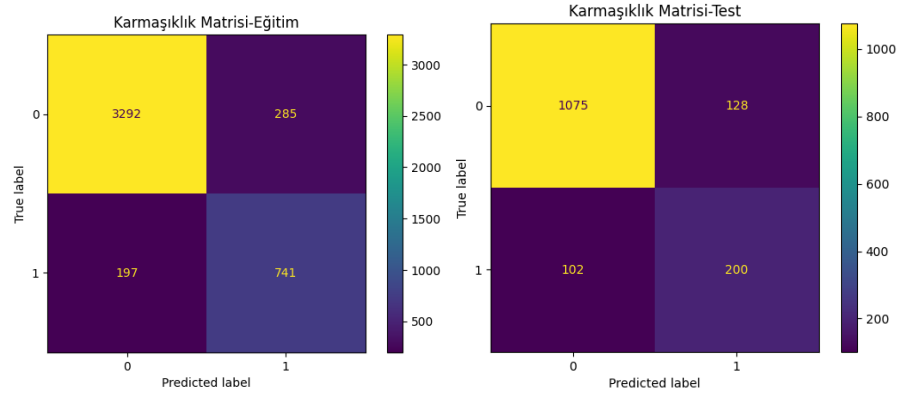


Şekil 37: k-en Yakın Komşu Algoritması Karmaşıklık Matrisi

Tablo 12: k-en Yakın Komşu Algoritması Değerlendirme Metrikleri

	Kesinlik (Precision)	Duyarlılık (Recall)	F1-Skoru (F1-Score)
Adli Kararlar	0.89	0.93	0.91
İdari Kararlar	0.67	0.56	0.61

Naive Bayes Algoritması Sonuçları



Şekil 38: Naive Bayes Algoritması Karmaşıklık Matrisi

Eğitim seti karmaşıklık matrisinde de görüldüğü gibi 3577 tane adli kararın 3292 tanesi, 938 tane idari kararın 741 tanesi doğru tahmin edilmiştir. Test seti karmaşıklık matrisinde de görüldüğü gibi 1203 tane adli kararın 1075 tanesi, 302 tane idari kararın 200 tanesi doğru tahmin edilmiştir.

Tablo 13: Naive Bayes Algoritması Değerlendirme Metrikleri

	Kesinlik (Precision)	Duyarlılık (Recall)	F1-Skoru (F1-Score)
Adli Kararlar	0.91	0.89	0.90
İdari Kararlar	0.61	0.66	0.63

Değerlendirme Metrikleri

Bütün algoritmaların değerlendirme metrikleri aşağıdaki Tablo 14’te verilmiştir. Tablo 14’te verilen değerler Python sonuç ekranında görülen “*macro avg*” değerleridir. Bu değerler, daha önceki tablolarda verilen adli ve idari kararlara ait değerlendirme metriklerinin aritmetik ortalaması alınarak elde edilen değerlerdir.

Tablo 14: Sınıflandırma Teknikleri Toplam Değerlendirme

	Doğruluk (Accuracy)	Kesinlik (Precision)	Duyarlılık (Recall)	F1-Skoru (F1-Score)
Karar Ağaçları	0.82	0.72	0.73	0.73
Rassal Orman	0.84	0.76	0.70	0.72
Lojistik Regresyon	0.85	0.77	0.75	0.76
Destek Vektör Makinelere	0.87	0.80	0.83	0.81
Diskriminant Analizi	0.70	0.59	0.62	0.60
Yapay Sinir Ağları	0.84	0.46	0.76	0.75
k-en Yakın Komşu	0.86	0.78	0.75	0.76
Naive Bayes	0.85	0.76	0.78	0.77

3.3. Tartışma ve Öneriler

Toplam değerlendirme sonuçlarına bakıldığında en iyi sonucu doğruluk değerinde %87, kesinlik değerinde %80, duyarlılık değerinde %83 ve F1 skorunda ise %81 ile DVM algoritması almıştır. Bütün algoritmalar da adli kararların tahmini idari kararların tahminine göre daha iyi sonuç vermiştir.

Tablo 15: Önerilen çalışma ile yapılan çalışmaların karşılaştırılması

	Mahkeme	Doğruluk Oranı
Aletras ve ark., (2016)	AİHM	%79
Liu & Chen, (2017)	AİHM	%79
Sulea, vd., (2017)	Fransız Yargıtayı	%96
Virtucio vd., (2018)	Filipinler Yüksek Mahkemesi	%59
Kowsrihawat vd. (2018)	Tayland Yüksek Mahkemesi	%67
Kaur & Bozic (2019)	AİHM	%82
Lage-Freitas vd. (2019)	Brezilya Mahkemeleri	%79 (F1)
Visentin ve ark., (2019)	AİHM	%86
Chalkidis vd. (2019)	AİHM	%82
Chalkidis vd. (2020)	AİHM	%83 (F1)
Quemy & Wrembel (2020)	AİHM	%96
Medvedeva ve ark., (2020)	AİHM	%75
Aydın, (2020)	Türk Yargıtayı	%81
Bertalan & Ruiz (2020)	São Paulo Adalet Mahkemesi	%98 (F1)
Strickson & De La Iglesia, (2020)	İngiltere Yüksek Mahkemesi	%69
Mumcuoğlu ve ark., (2021)	Türk Anayasa Mahkemesi	%67
Malik vd. (2021)	Hint Yüksek Mahkemesi	%77
Sert vd., (2022)	Türk Anayasa Mahkemesi	%90
Önerilen Çalışma	Türk Uyuşmazlık Mahkemesi	%87

Şu ana kadar Türk Uyuşmazlık Mahkemesi kararları ile çalışılmamıştır. Aynı algoritmaların kullanıldığı diğer çalışmalarla kıyaslandığında önerilen çalışma DVM algoritması %87 doğruluk değeri ile en doğru sonucu vermiştir. Yapılan diğer çalışmalar

farklı makine öğrenmesi algoritmalarından ziyade sinir ağıları kullanımına yöneliktir. Derin öğrenme algoritmaları ile hukuki metinler üzerinde yapay zekâ çalışmaları çok az sayıdadır. Bundan sonra yapılacak olan çalışmalarda derin öğrenme tabanlı algoritmalar ile hukuki metinlerin analizi yapılırsa performans değerlendirmelerinin daha iyi sonuçlar vereceği düşünülmektedir.

Önerilen çalışma için önce Entegre (Tümleşik) Geliştirme Ortamlarından (Integrated Development Environment-IDE) PyCharm ve Spyder kullanılmıştır. Ancak güncellemelerle ilgili yaşanan sıkıntılar kodların çalışmasını yavaşlatmış hatta bazen engellemiştir. Çözüm olarak Google Colab ortamında kodlar yazılmış ve çıktılar alınmıştır. Google serverlarının kullanımı sayesinde hız artmış ve güncellemelerle ilgili sıkıntılar ortadan kalkmıştır. Bundan sonra yapılacak çalışmalar için Google Colab ortamı önerilmektedir.

Bu çalışma yapay zekâ çalışma alanı anlamında dilden bağımsız bir çalışmadır. Hukuk alanında yapay zekâ çalışmalarında dil bağımlı çalışmalar da yapılmalıdır. Dil bağımlı çalışmalar, kelime düzeyinde morfolojik (biçim bilimsel) çalışmalar, söz dizimsel düzeyde cümle analizi çalışmaları, anlam düzeyindeki çalışmalar ve söylev (pragmatik) düzeyindeki çalışmalar olmak üzere dört ana odak etrafında şekillenmektedir. Günlük hayatta kullandığımız herhangi bir kelimenin hukuki metin ve/veya görülen dava üzerinde önemli etkileri olabilir. Bu nedenle dil bağımlı çalışmalar özellikle hukuk alanında yapılmalıdır.

Adalet sistemine yeni bir soluk getirecek yapay zekâ çalışmalarında diğer yapay zekâ çalışmalarında olduğu gibi toplumda ve özellikle çalışılan sektördeki aktörler tarafından yapay zekâ kavramına karşı endişe ve korkunun hâkim olduğu görülmektedir. Bu endişe ve korkuyu gidermek için yargı teşkilatı içindeki her birime yapay zekâ kavramına yönelik eğitici bilgilerin verilmesi gerekmektedir.

Bu çalışmada ve diğer hukuk alanındaki makine öğrenmesi çalışmalarında da görüldüğü gibi, bir öngörude bulunabilmek için mahkeme kararlarının hangileri olduğunun belirlenmesi ardından bu kararların sınıflandırılması gerekmektedir. Verilen hükümlerin çalışmaların amacına göre kodlanmasına ihtiyaç duyulmaktadır. Böyle türde kodlamalar yapmak gerçek yaşamdaki bütün mahkeme kararları için imkânsız gibi görülse de Çin Halk Cumhuriyeti gibi standartlaştırmayı bir devlet politikası haline getirmiş bir ülkede mümkün hale gelmiştir (Gültekin-Varkonyi, 2022). Çalışma

kapsamında UYM kararları web ortamında çekildikten sonra bu tez kapsamında geliştirilen yapay zekâ botu tarafından okunan mahkeme kararlarının standart bir form şeklinde olmaması nedeniyle veriler standart hale getirmiştir. Bu durum veri ön işleme adımında çok fazla vakit harcanmasına sebep olmuştur. Bu anlamda yukarıda açıklandığı sebeplerle ülkemizde mahkeme kararları ve hatta avukat dilekçelerine kadar yargı sisteminde kullanılan bütün belgelerin standart bir forma sokulması hukuk alanında yapay zekâ çalışmalarına çok büyük katkı sağlayacaktır.

Yapay zekâ uygulamalarının geliştirilmesi için öncelikle dijital verilere ihtiyaç duyulmaktadır. Bu nedenle hukuk alanında yapay zekâ çalışmaları yapabilmek için yargı sistemindeki tüm belgelerin dijital hale getirilmesi gerekmektedir. Standart hale gelmiş dijital hukuk verileri ile yapay zekâ çalışmalarında çok büyük aşamalar kaydedilecektir.

Hukuk alanında yapay zekâ çalışmalarına çok büyük katkı sağlayacak kavramlardan biri çevirim içi hukuk hizmetleridir.

Tania Sourdin, “Judges, Technology and Artificial Intelligence” adlı eserinde dört farklı çevrimiçi mahkeme modelinden bahsetmektedir (Sourdin, 2021). Bu modeller şunlardır:

1. Çevrimiçi Hakemlik: Taraflar, uyuşmazlıklarını bir hakem ya da panel tarafından çözülmek üzere sunarlar. Tarafların fiziksel olarak bir araya gelmesi gerekmez ve tüm süreç çevrimiçi olarak yürütülür.

2. Çevrimiçi Tahkim: Taraflar, uyuşmazlıklarını bir tahkimci veya panel tarafından çözülmek üzere sunarlar. Tahkim süreci, taraflar arasında bir anlaşmazlık olduğunda sözleşme tarafından önceden belirlenir.

3. Çevrimiçi Yargı: Taraflar, uyuşmazlıklarını çevrimiçi bir yargıç veya panel tarafından çözülmek üzere sunarlar. Bu tür bir mahkeme sistemi, yargıçların fiziksel olarak bir araya gelmesi gerektiği durumlarda kullanılabilir.

4. Hibrit Mahkemeler: Hibrit mahkemeler, çevrimiçi ve fiziksel mahkeme unsurlarını bir araya getirir. Bu tür mahkemeler, tarafların ihtiyaçlarına ve uyuşmazlık türüne göre değişkenlik gösterir.

Bu dört model, tarafların çevrimiçi ortamda uyuşmazlıklarını çözmelerine yardımcı olmak için farklı seçenekler sunar. Ancak her modelin kendine özgü avantajları ve dezavantajları vardır ve uyuşmazlığın türüne, tarafların ihtiyaçlarına ve yerel yargı sistemine uygun bir seçim yapılması önemlidir. Bu tür çevirim içi modellerin ülkemizde kullanılması ve yaygınlaşması ile hukuk alanında yapay zekâ çalışmalarına çok büyük katkı sağlayacaktır.

Çin Halk Cumhuriyeti'nde hayata geçirilen “akıllı mahkemeler” sistemi hem belgelerin dijitalleşmesi hem de standart hale getirilmesi süreci açısından ayrıca yapay zekâyâ sahip yargıçların kullanılması açısından incelenmesi gereken bir sistemdir.

Bu “akıllı mahkemeler” inşa edilirken dosyalama aşamasında akıllı rehberlik, duruşma sürecinde ses transkripsiyonu, deneme sürecinde otomatik belge oluşturma gibi akıllı yardım işlevleri ile “Akıllı Mahkemeler”in ön hazırlıklarının yapıldığı rapor edilmektedir (Cui, 2020). Yapay zekâ bu sistemlerin inşasındaki her adımda kullanılmaktadır.

İlk olarak, yapay zekâ belge işleme sürecine yardımcı olmak için kullanılmaktadır. Hukuki belgeler, mahkemelerde ses transkripsiyonu, karar üretimi gibi işlemler yapmak için yapay zekâ kullanılarak işlenmektedir. İddianame ve karar otomatik olarak oluşturulmaktadır. Hangzhou İnternet Mahkemesi gibi bazı mahkeme platformları, tarafların ilgili materyalleri girdiği zaman iddianameyi hızlıca oluşturmak için yapay zekâ kullanmaktadır. Mahkemenin verdiği karardan sonra, platform yazılı kararın tamamını veya bir kısmını otomatik olarak oluşturmaktadır (Cui, 2020).

İkincisi, yapay zekâ, duruşma transkripsiyonuna yardımcı olmak için ses dönüştürme sistemlerini kullanarak bir çeşit yapay zekâ kâtip gibi kullanılmaktadır. Böylece katiplerin yükü büyük ölçüde azalmaktadır.

Üçüncüsü, yapay zekâ, davaların işlenmesine yardımcı olmak için bir çeşit sanal yargıç gibi kullanılmaktadır. Büyük veri, makine öğrenmesi vb. teknolojiler temel alınarak, yapay zekâ destekli sistemler, birçok davayı öğrenerek, delil bilgilerini çıkarmayı ve doğru kararların öngörülmesini sağlamayı öğrenip yargıçların kararlarında referans sağlamaktadır. Yapay zekâ destekli sistemler, karar sürecini standartlaştırmakta kararın tutarlılığını artırmakta, insanların haksız, sahte veya yanlış şekilde suçlandığı

veya cezalandırıldığı durumların ortaya çıkmasını azaltmakta ve yargının halk nezdindeki güvenilirliğini artırmaktadır (Sourdin, 2021).

Dördüncüsü, yapay zekâ, adli hizmetlere yardımcı olmak için kullanılmaktadır. Mahkemeler, Xiamen Mahkemesi'nin “Legal System Canghai” adlı WeChat resmi hesabı robotu, Shenzhen Uluslararası Tahkim Mahkemesi'nin “Xiao 3i” adlı robotu vb. aracılığıyla halka hukuk danışmanlığı sağlamakta ayrıca Shenzhen Nanshan Adalet Bürosu'nun “Nan Xiao Fa” adlı varlık robotu gibi bazı varlık robotları ile halka, hukuk firmalarına ve mahkemelere akıllı müşteri hizmeti sağlayan Fa Gougou gibi robotlar bulunmaktadır (Sourdin, 2021).

Belki de bu sistemin kullanılmasında dikkat çeken en önemli husus Kıta-Avrupası hukuk sistemine sahip olan bir ülkenin yapay zekâ sistemleri kullanarak Anglo-Sakson hukuk sistemine benzer şekilde içtihadı dayalı bir sisteme geçmesidir (Sourdin, 2021).

Hukuki belgelerin dijitalleştirilmesi ve beraberinde yapay zekâ sistemlerinin yargı sistemi içinde kullanılması hiç şüphesiz güvenlik sorununu akla getirmektedir. Bunun için Çin Halk Cumhuriyeti'nde blok zincir (blockchain) denilen teknoloji kullanılmaktadır. Şanghai'daki kamu güvenlik organları, savcılıklar, mahkemeler ve adalet bürolarının iç ağları birbirine bağlı değildir ve aralarındaki veri iletimi “Bir merkez, Dört platform” temelinde, kanıt depolama ve toplama için beş özel blok zinciri bir sisteme eklenmiştir ve böylece veri değiştirme işlemleri önlenmektedir (Cui, 2020). Ülkemizde yapılacak yargı sistemindeki yapay zekâ çalışmalarında da güvenlik açısından blockchain vb teknolojilerinin kullanılması gerekmektedir.

SONUÇ

Bu çalışma kapsamında UYM kararları makine öğrenmesi yöntemleri ile tahmin edilmeye çalışılmıştır. Kullanılan denetimli makine öğrenmesi algoritmalarından Destek Vektör Makineleri algoritması %87 doğruluk oranı ile en yüksek sonucu vermiştir. Mahkeme kararlarının tahminine yönelik yapılan diğer çalışmalarla karşılaştırıldığında bu çalışma kapsamında önerilen uygulamanın performansının daha iyi olduğu görülmüştür.

Bu çalışma kapsamında yapılan sınıflandırma işlemleri aslında bir tahmin işlemidir. Yapılan tahmin işlemi ile sisteme girilecek yeni bir mahkeme kararının adli veya idari yargıdan hangisinin görev alanına gireceği öngörülmektedir. Hâkim (yargıç), davanın avukatları veya savcılar tarafından hazırlanan somut gerçeklere, delillere, iddialara, savunmalara, ifadelere, yasal dayanaklara, şüpheli veya sanıkların duruşma süresindeki fiziksel ve ruhsal durumlarına, diğer her türlü bilgi kaynağına göre davanın sonucuna yönelik tahminde bulunan kişidir. Bu çalışmada önerilen uygulama uyuşmazlık mahkemesi kararlarını tahmin etmektedir. Bu açıdan düşünüldüğünde bu çalışmada kullanılan uygulama, uyuşmazlık mahkemesi hâkimi gibi karar veren bir yapay zekâ sistemidir.

Bu çalışmada önerilen uygulama bir sanal yargıç uygulamasıdır. Bu tür bir uygulamanın olumlu ve olumsuz birçok yönü bulunmaktadır. İnsanlara nazaran tarafsız olması, objektif değerlendirmelerde bulunması, zamandan tasarruf sağlaması, çok büyük bir bellek ve öğrenme kapasitesine sahip olması gibi olumlu taraflarının yanında önyargılı verilere dayalı olarak önyargılı kararlar vermesi, insanların sanal yargıca duygusal bağlantılar veya merhamet beklentisiyle yaklaşması ve bunun karşılığında yapay zeka teknolojisinin bu tür faktörleri hesaba katmakta zorlanması, insan hâkim gibi bir takdir yetkisine sahip olmaması, istatistiksel hatalar yapma potansiyeline sahip olması gibi olumsuz tarafları da bulunmaktadır.

İnsanların hayatına önemli derecede etki edecek olan sanal yargıç türü uygulamalarda özellikle etik ve güvenlik konularına dikkat edilmesi gerekmektedir. Uluslararası kurumlar tarafından belirlenen prensiplere göre hareket edilmesi çok

önemlidir. Bu tür uygulamalar, temel haklara saygılı, ayrımcılık yapmayan, güvenli, şeffaf ve kullanıcı kontrolünde olmalıdır.

Yargılamada kullanılan yapay zekâ teknolojisinin, insanların temel haklarını ve özgürlüklerini koruyacak şekilde kullanılması gerekmektedir. Aynı zamanda yapay zekâ sistemleri, cinsiyet, ırk, etnik köken, din, cinsel yönelim, engellilik gibi herhangi bir ayrımcılık temelinde kullanılmamalıdır. Bununla beraber yapay zekâ sistemleri, yüksek kalite ve güvenlik standartlarına uygun olarak geliştirilmeli ve karar algoritması ve veri işleme süreçleri anlaşılır ve açıklanabilir olmalıdır. En önemli noktalardan biri de yapay zekanın kullanıcının kontrolünde olmasıdır. Kullanıcılar, yapay zekâ sistemlerinin nasıl çalıştığını anlamalı, etkileşimde bulunmalı ve son kararı vermekte kontrol sahibi olmalıdır.

Bu prensipler, yargıda yapay zekâ kullanımının etik ve hukuki bir çerçevede gerçekleştirilmesi gerektiğini vurgulamaktadır. Yapay zekanın etik kullanımı, insanların denetimi ve kontrolü altında olması, insan haklarına ve değerlerine saygılı olması, şeffaf ve hesap verebilir bir sistem oluşturulması, güvenilir ve doğru sonuçlar vermesi, manipülasyona karşı korunması, kişisel verilerin güvenliğinin sağlanması, ayrımcılık yapmaması, denetlenebilir ve hesap verebilir olması gibi prensipler, etik açıdan doğru bir yapay zekâ kullanımını desteklemektedir.

Ayrıca, yapay zekânın olumsuz etkilerinin hızla fark edilip raporlanması, kararlarına müdahale edilerek kontrolünün sağlanması, yapay zekânın toplumun gelişimine katkı sağlaması ve sürdürülebilir olması da önemli faktörlerdir. yapay zekanın toplumun iyiliği için kullanılması ve olumlu bir etki yaratması hedeflenmelidir.

Yapay zekâ, günümüzde birçok alanda önemli bir rol oynamaktadır. Bu alanlardan biri de hukuktur. Bireylerin yapay zekâ hakkında bilgi sahibi olmaları ve bu teknolojileri anlayabilmeleri önemlidir. Bireylerin yapay zekâ hakkında temel bir anlayışa sahip olmaları ve yapay zekâ teknolojilerini etkili bir şekilde kullanabilme yeteneğini geliştirmelerine yapay zekâ okur-yazarlığı denilmektedir. Hukuk alanında yapay zekâ uygulamalarının çoğalmasıyla birlikte hukuk sâjeleri bu teknolojilerden faydalanacaktır. Bu nedenle, hukuk alanında da yapay zekâ okur yazarlığının artırılmasına yönelik çalışmaların yapılması gerekmektedir.

Sanal yargıç gibi teknolojik uygulamaların ortaya çıkması sonrası bu tür gelişmeler için yasal mevzuatlar çıkarılmalıdır. Yeni teknolojik gelişmelerin hızla ilerlediği günümüzde, yasal mevzuatın bu gelişmelere uyum sağlaması ve bu teknolojilerin güvenli, etik ve verimli bir şekilde kullanılması büyük önem taşımaktadır. Teknolojik gelişmeler, toplumun güvenliği ve istikrarı üzerinde doğrudan etkiye sahip olabilir. Uygun yasal düzenlemeler olmadan, teknolojinin yanlış kullanımı veya kötü niyetli kişilerin suistimali toplumsal güveni tehdit edebilir. Yeni teknolojilerin yaygınlaşmasıyla kişisel verilerin toplanması ve işlenmesi artmıştır. Yasal mevzuat, bireylerin veri mahremiyetini ve özel hayatını korumak için gerekli düzenlemeleri içermelidir. Teknolojinin etik açıdan duyarlı konularda kullanımı, toplumsal değerlere uygunluğu gerektirir. Yasalar etik standartları belirlemeli ve teknolojiyi etik sınırlar içinde tutmalıdır.

Teknolojik gelişmeler, hukuki boşluklar yaratabilir ve var olan yasaların uygulanmasını zorlaştırabilir. Yeni teknolojilere özgü düzenlemeler, bu boşlukları doldurmak ve hukuki güvenliği sağlamak için gereklidir. Teknolojinin hızla ilerlemesi, yenilikçi fikirlerin önünü açar. Ancak bu inovasyonların güvenilir bir şekilde gelişmesi için yasal mevzuatın yol gösterici olması gerekir. Teknoloji, farklı ülkeler ve bölgeler arasında hızla yayılabilir. Uyumlu yasal mevzuat, uluslararası standartlara uyum sağlayarak iş birliği ve ticareti kolaylaştırır. Yeni teknolojiler genellikle bilinmeyen riskleri beraberinde getirir. Yasal düzenlemeler, bu riskleri belirlemek, yönetmek ve minimize etmek için bir çerçeve sunar.

Teknolojik gelişmelerin beraberinde getireceği riskler kamu düzenini ve güvenliği etkileyebilir. Sanal yargıç türü uygulamalar da ayrımcılık ve eşitsizlik gibi riskleri içinde barındırmaktadır. Yasalar, bu tür durumlar için hazırlıklı olunmasını ve toplumun güvenliğinin korunmasını sağlamalı, ayrımcılığı önlemek ve eşitlik ilkesini korumak amacıyla gerekli düzenlemeleri içermelidir. Bu anlamda sanal yargıç türü uygulamaların kullanımında dikkat edilecek hususlar yasal mevzuat şeklinde hazırlanmalıdır. Bu çalışma kapsamında UYM için uygulanma potansiyeline sahip sanal yargıç türü bir yazılım geliştirilmiştir. Bu türden yazılımların uygulamaya konulabilmesi için 6100 sayılı

Hukuk Muhakemeleri Kanunu (HMK) ve ilgili diğ er mevzuatlarda deęiřiklik yapılması gerekmektedir.

KAYNAKLAR

- Acı, Ç., & Çırak, A. (2019). Türkçe Haber Metinlerinin Konvolüsyonel Sinir Ağları ve Word2Vec Kullanılarak Sınıflandırılması. *Bilişim Teknolojileri Dergisi*, 219-228. <https://doi.org/10.17671/gazibtd.457917>
- Adalı, E. (2012). Doğal dil işleme. *Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi*, 5. <https://dergipark.org.tr/en/download/article-file/207209>
- Agrawal, R., Imielinski, T., & Swami, A. (1993). ACM SIGMOD international conference on Management of Data. *Mining association rules between sets of items in large databases*, 207-216.
- Akil, C. (2012). Türkiye Barolar Birliği Disiplin Kurulu Kararları Işığında Avukatın Görevini Özenle Yerine Getirme Yükümlülüğü Araştırma. *Hacettepe Hukuk Fakültesi Dergisi*, 2(1), 11-26.
- Akkurt, S. S. (2019). Yapay Zekânın Otonom Davranışlarından Kaynaklanan Hukukî Sorumluluk. *Uyuşmazlık Mahkemesi Dergisi*, 13, 39-59.
- Aletras, N., Tsarapatsanis, D., Preoțiuc-Pietro, D., & Lampos, V. (2016). Predicting judicial decisions of the European court of human rights: A natural language processing perspective. *PeerJ Computer Science*, 2016(10). <https://doi.org/10.7717/peerj-cs.93>
- Alpaydın, E. (2014). *Introduction to Machine Learning* (Third). The MIT Press.
- Alpaydın, E. (2016). *Machine learning : the new AI*. The MIT Press.
- Anıl, A. (1992). Uyuşmazlık Mahkemesi'nin Kuruluş ve İşleyişi. *Ankara Üniversitesi SBF Dergisi*, 47(01), 45-50.
- Arf, C. (1959). Halk Konferansları. *Makineler Düşünebilir mi ve Nasıl Düşünebilir*, 91-103.
- Arslan, H., Dadaş, İ. E., & Işık, Y. E. (2022). Classification of Support Tickets using Different Vectorization and Pre-Processing Methods. *Düzce Üniversitesi Bilim ve Teknoloji Dergisi*, 1433-1442. <https://doi.org/10.29130/dubited.1017422>
- Attewell, P., & Monaghan, D. (2015). *Data Mining for the Social Sciences* (1. bs). University of California Press.
- Awad, M., & Khanna, R. (2015). Support vector machines for classification. İçinde . *Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers*, (ss. 39-66).
- Ay, R. A. (2021). *Doğal Dil İşleme ve Kümeleme Yöntemleri İle Müşteri Yorumlarının İncelenmesi* [Yayımlanmış Yüksek Lisans Tezi]. KTO Karatay Üniversitesi Lisansüstü Eğitim Enstitüsü.

- Aydın, Ö. (2020). *Mobbing İçerikli Yargı Kararlarının Makine Öğrenmesi Algoritmaları ile Sınıflandırılması* [Yayımlanmış Yüksek Lisans Tezi]. Balıkesir Üniversitesi Fen Bilimleri Enstitüsü.
- Aytaç, M. (2012). *Matematiksel İstatistik* (4. bs). Ezgi Kitabevi.
- Babangida, N. M., Ul Mustafa, M. R., Wan Yusuf, K., Isa, M. H., & Baig, I. (2016). Evaluation of low degree polynomial kernel support vector machines for modelling Pore-water pressure responses. *MATEC Web of Conferences*, 59, 04003. <https://doi.org/10.1051/mateconf/20165904003>
- Bailer-Jones, C. A. L., Gupta, R., & Singh, H. P. (2001). An introduction to artificial neural networks. *Automated Data Analysis in Astronomy*, 1-18. arXiv preprint astro-ph/0102224
- Bak, B. (2018). Medeni Hukuk Açısından Yapay Zekânın Hukuki Statüsü ve Yapay Zekâ Kullanımından Doğan Hukuki Sorumluluk. *Türkiye Adalet Akademisi Dergisi*, 9(35), 211-232. <https://www.youtube.com/watch?v=3WhpAKJcggl>
- Balaban, M. E., & Kartal, E. (2015). *Veri Madenciliği ve Makine Öğrenmesi Temel Algoritmaları ve R Dili Uygulamaları* (1. bs). Çağlayan Kitabevi.
- Bartlett, P., Freund, Y., Lee, W. S., & Schapire, R. E. (1998). Boosting the margin: a new explanation for the effectiveness of voting methods. *The Annals of Statistics*, 26(5). <https://doi.org/10.1214/aos/1024691352>
- Baştanlar, Y., & Özuysal, M. (2014). Introduction to machine learning. İçinde M. Yousef & J. Allmer (Ed.), *miRNomics: MicroRNA biology and computational analysis* (C. 1107, ss. 105-128). Springer Science+Business Media. https://doi.org/10.1007/978-1-62703-748-8_7
- Bayındır, A. S. (2022). *Yapay Zekâ Teknolojilerinin Ortaya Koyduğu Buluşların Patentlenebilirliği* [Yayımlanmış Yüksek Lisans Tezi]. İhsan Doğramacı Bilkent Üniversitesi Ekonomi ve Sosyal Bilimler Enstitüsü.
- Bayram Arlı, N. (2018). *Betimsel İstatistik*. Ezgi Kitabevi.
- Bayram Arlı, N. (2021). Büyük Veri ve Veri Madenciliği. İçinde Z. Sabuncuoğlu, K. Kaymaz, & M. Paksoy (Ed.), *İşletme Yönetiminde Dijital Yaklaşımlar* (1. Baskı, ss. 293-319). Dora.
- Bayram Arlı, N., Barca, O., & Görentaş, M. B. (2022). Rassal Orman. İçinde N. Bayram Arlı, S. Gürsakal, & M. Engin (Ed.), *Denetimli Makine Öğrenmesi Algoritmaları R ve Python Uygulamaları* (1. bs, ss. 119-148). Nobel.
- Bayram Arlı, N., Gürsakal, S., & Önay Koçoğlu, F. (2022). Denetimli Makine Öğrenmesi Algoritmalarına Giriş. İçinde N. Bayram Arlı, S. Gürsakal, & M. Engin (Ed.), *Denetimli Makine Öğrenmesi Algoritmaları R ve Python Uygulamaları* (1. bs, ss. 1-8). Nobel. www.nobelkitap.com

- Berrar, D. (2018). Cross-validation. İçinde *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics* (C. 1-3, ss. 542-545). Elsevier. <https://doi.org/10.1016/B978-0-12-809633-8.20349-X>
- Bertalan, V. G., & Ruiz, E. E. S. (2020). Predicting Judicial Outcomes in the Brazilian Legal System Using Textual Features. *International Conference on the Computational Processing of Portuguese, DHandNLP@PROPOR 2020*, 22-32.
- Bhatia, N. (2010). Survey of Nearest Neighbor Techniques. *IJCSIS) International Journal of Computer Science and Information Security*, 8(2). <http://sites.google.com/site/ijcsis/>
- Bilgin, M. (2018). *Makine Öğrenmesi :Teorisi ve Algoritmaları* (A. Yılmaz, Ed.; 1. bs). Papatya Yayınları.
- Bilgin, M., & Eser, M. (2022a). Karar Ağaçları. İçinde N. Bayram Arlı, S. Gürsakal, & M. Engin (Ed.), *Denetimli Makine Öğrenmesi Algoritmaları R ve Python Uygulamaları* (1. bs, ss. 101-117). Nobel.
- Bilgin, M., & Eser, M. (2022b). K-En Yakın Komşu. İçinde N. Bayram Arlı, S. Gürsakal, & M. Engin (Ed.), *Denetimli Makine Öğrenmesi Algoritmaları R ve Python Uygulamaları* (1. bs, ss. 263-280). Nobel.
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45, 5-32.
- Büyük, C. (2010). Yasama Yürütme Yargı: Platon'dan Montesquieu'ya Devletin Fonksiyonlarına Genel Bir Bakış. *Dinbilimleri Akademik Araştırma Dergisi*, 10(1), 159-170.
- Castaneda, B., & Cockburn, J. C. (2005). Reduced Support Vector Machines Applied to Real-Time Face Tracking. *Proceedings. (ICASSP '05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.*, 673-676. <https://doi.org/10.1109/ICASSP.2005.1415494>
- Chalkidis, I., Androutsopoulos, I., & Aletras, N. (2019). Neural Legal Judgment Prediction in English. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 4317-4323. <https://doi.org/10.18653/v1/P19-1424>
- Chalkidis, I., Fergadiotis, M., Malakasiotis, P., Aletras, N., & Androutsopoulos, I. (2020). LEGAL-BERT: “preparing the muppets for court. *2020 conference on empirical methods in natural language processing: findings*, 2898-2904.
- Chowdhary, K. R. (2020). Natural Language Processing. İçinde *Fundamentals of Artificial Intelligence* (1. bs, ss. 603-649). Springer. https://doi.org/https://doi.org/10.1007/978-81-322-3972-7_19
- Condevaux, C. (2020). Neural Legal Outcome Prediction with Partial Least Squares Compression. *Stats*, 3(3), 396-411. <https://doi.org/10.3390/stats3030025>
- Cui, Y. (2020). *Artificial Intelligence and Judicial Modernization* (1. bs). Springer Singapore. <https://doi.org/https://doi.org/10.1007/978-981-32-9880-4>

- Cumhurbaşkanlığı Dijital Dönüşüm Ofisi Başkanlığı, & Sanayi ve Teknoloji Bakanlığı. (2021). *Ulusal Yapay Zekâ Stratejisi (UYZS) 2021-2025*. <https://cbddo.gov.tr/UYZS>
- Curran, N. M., Sun, J., & Hong, J.-W. (2020). Anthropomorphizing AlphaGo: a content analysis of the framing of Google DeepMind's AlphaGo in the Chinese and American press. *AI & SOCIETY*, 35(3), 727-735. <https://doi.org/10.1007/s00146-019-00908-9>
- Çamkerten, A. S. (2023). Hukuk Analitiği. *Necmettin Erbakan Üniversitesi Hukuk Fakültesi Dergisi, Necmettin Erbakan University*. <https://doi.org/10.51120/NEUHFD.2023.79>
- Dale, R. (2019). Law and Word Order: NLP in Legal Tech. *Natural Language Engineering*, 25(1), 211-217. <https://doi.org/10.1017/S1351324918000475>
- Dangeti, P. (2017). *Statistics for machine learning*. Packt Publishing.
- Dhar, V. (2016). The Future of Artificial Intelligence. *Big Data*, 4(1), 5-9. <https://doi.org/https://doi.org/10.1089/big.2016.29004.vda>
- Dinç, M. İ. (2016). “Olumsuz Görev Uyuşmazlığının Çözümü İçim Uyuşmazlık Mahkemesine Başvuru. *Danıştay Dergisi*, 60, 36-47.
- Dothan, S. (2020). A Guide to Quantitative Legal Research. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3719836>
- Durgesh, S., & Lekha, B. (2010). Data classification using support vector machine. *Journal of Theoretical and Applied Information Technology*, 12(1), 1-7.
- Dülger, M. V. (2021). Günümüz Yapay Zeka Teknolojisi ve 'Robot Yargıç/Avukat' Gerçeği (Artificial Intelligence Technology Today and the Reality of 'Robot Judge/Lawyer'). *hplusdergi*, 4-9. <https://doi.org/http://dx.doi.org/10.2139/ssrn.3792182>
- Engin, M. (2022). Naive Bayes. İçinde N. Bayram Arlı, S. Gürsakal, & M. Engin (Ed.), *Denetimli Makine Öğrenmesi Algoritmaları R ve Python Uygulamaları* (1. bs, ss. 281-294). Nobel.
- euRobotics. (2020). *About SPARC*. euRobotics. https://eu-robotics.net/divi_overlay/sparc/
- European Comission. (2021). *Functioning of the European ODR Platform: Statistical Report 2020*. <https://commission.europa.eu/system/files/2021-12/2021-report-final.pdf>
- European Commission. (2018). 237 *Final Communication From The Commission To The European Parliament, The European Council, The Council, The European Economic And Social Committee And The Committee Of The Regions Artificial Intelligence For Europe*. <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0237&from=EN>

- European Commission. (2018). *The European Commission's High-Level Expert Group On Artificial Intelligence. A Definition Of AI: Main Capabilities And Scientific Disciplines*.
https://ec.europa.eu/futurium/en/system/files/ged/ai_hleg_definition_of_ai_18_december_1.pdf
- European Consumer Consultative Group. (2018). *Policy Recommendations for a Safe and Secure Use of Artificial Intelligence, Automated Decision-Making, Robotics and Connected Devices in a Modern Consumer World*.
- European Economic and Social Committee. (2018). *Artificial intelligence: Anticipating It's Impact on Work to Ensure a fair Transition*. <https://www.eesc.europa.eu/en/our-work/opinions-information-reports/opinions/artificial-intelligence-anticipating-its-impact-jobs-ensure-fair-transition-own-initiative-opinion>
- European Parliament. (2017). *Civil Law Rules On Robotics*.
<https://oeil.secure.europarl.europa.eu/oeil/popups/printficheglobal.pdf?id=656479&l=en>
- Eyecioğlu Özmutlu, A. (2021). Doğal Dil İşleme. İçinde T. Talan & C. Aktürk (Ed.), *Bilgisayar Bilimlerinde Teorik Ve Uygulamalı Araştırmalar* (ss. 129-154). Efe Akademik Yayıncılık.
- Fan, J., Upadhye, S., & Worster, A. (2006). Understanding receiver operating characteristic (ROC) curves. *Canadian Journal of Emergency Medicine*, 8(1), 19-20.
- Faraz, A. (2016). A Comparison of Text Categorization Methods. *International Journal on Natural Language Computing*, 5(1), 31-44.
<https://doi.org/10.5121/ijnlc.2016.5103>
- Fausett, L. V. (1993). *Fundamentals of Neural Networks: Architectures, Algorithms And Applications* (1. bs). Pearson.
- Fieller, N. (2016). *Basics of Matrix Algebra for Statistics with R* (J. M. Chambers, T. Hothorn, D. T. Lang, & H. Wickham, Ed.). CRC Press Taylor & Francis Group.
- Fornaciari, T., & Poesio, M. (2013). Automatic deception detection in Italian court cases. *Artificial Intelligence and Law*, 21, 303-340.
<https://doi.org/https://doi.org/10.1007/s10506-013-9140-4>
- Gärtner, T., Lloyd, J. W., & Flach, P. A. (2004). Kernels and distances for structured data. *Machine Learning*, 57, 205-232.
<https://doi.org/https://doi.org/10.1023/B:MACH.0000039777.23772.30>
- Gedik, G. (2020). Robotlara karşı gerçek kişilerin korunması gerekliliği ve robot vergisi önerisi. *Marmara Üniversitesi Hukuk Fakültesi Hukuk Araştırmaları Dergisi*, 26(1), 24-48.

- Giwa, D. C., & Kodjovi, D. (2023). Artificial Intelligence and the Future of the Administration of Law and Justice in Nigeria. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4566440>
- Gözler, K. (2018). *Hukuka Giriş* (15. Baskı). Ekin Yayınevi.
- Gözler, K. (2019). *Türk Anayasa Hukuku Dersleri* (23. Baskı). Ekin Yayınevi.
- Gözübüyük, A. Ş., & Tan, T. (2012). *İdare Hukuku* (5. bs). Turhan Kitabevi.
- Gupta, N. (2013). Artificial neural network. *Network and Complex Systems*, 3(1), 24-28.
- Gültekin-Varkonyi, G. (2022). *Robot Yargıçlar*. Seçkin.
- Güner, C. (2020). Yapay Zekânın Verdiği Zarardan Doğan Sözleşme Dışı Sorumluluğa Uygulanacak Hukuk. *Uyuşmazlık Mahkemesi Dergisi*, 15, 229-272.
- Gürsakar, S. (2019). *Sosyal bilimlerde SPSS uygulamalı çok değişkenli istatistiksel analiz*. Dora.
- Gürsakar, S., Barca, O., & Tütüncü, T. E. (2022). Lojistik Regresyon Analizi. İçinde N. Bayram Arlı, S. Gürsakar, & M. Engin (Ed.), *Denetimli Makine Öğrenmesi Algoritmaları R ve Python Uygulamaları* (1. bs, ss. 53-83). Nobel.
- Hakim, H. A., Praja, C. B. E., & Ming-Hsi, S. (2023). AI in Law: Urgency of the Implementation of Artificial Intelligence on Law Enforcement in Indonesia. *Jurnal Hukum Novelty*, 14(1), 122-134.
- Han, J., Kamber, M., & Pei, J. (2012). *Data mining concepts and techniques* (Third). Morgan Kaufman. Han, J., Kamber, M., & Pei, J. (2012). Data mining concepts and techniques third edition. University of Illinois at Urbana-Champaign Micheline Kamber Jian Pei Simon Fraser University
- Haugeland, J. (1985). *Artificial Intelligence: The Very Idea*. MIT Press.
- Hesterberg, T. (2011). Bootstrap. *Wiley Interdisciplinary Reviews: Computational Statistics*, 3(6), 497-526. <https://doi.org/10.1002/wics.182>
- High-Level Expert Group on AI (AI HLEG). (2019). *Policy and Investment Recommendations for Trustworthy AI*. <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>
- Horowitz, J. L. (2001). The Bootstrap. İçinde J. Heckman & E. Leamer (Ed.), *Handbook of Econometrics* (C. 5, ss. 3159-3228). Elsevier.
- Houy, C., Niesen, T., Fettke, P., & Loos, P. (2013). Towards automated identification and analysis of argumentation structures in the decision corpus of the German Federal Constitutional Court. In *2013 7th IEEE International Conference on Digital Ecosystems and Technologies (DEST)*, 72-77. <https://doi.org/10.1109/DEST.2013.6611332>

- İçer, Z., & Buluz, B. (2019). Yapay Zekanın Ceza Muhakemesindeki Rolü ve Geleceği,. *9. Uluslararası Suç ve Ceza Film Festivali* .
- Ikonomakis, M., Kotsiantis, S., & Tampakas, V. (2005). Text Classification Using Machine Learning Techniques. *WSEAS transactions on computers*, 4(8), 966-974. <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=b0485fba23aabda526358f31cb5a382b66a08270>
- Jo, T. (2021). *Machine Learning Foundations Supervised, Unsupervised, and Advanced Learning*. Springer International Publishing.
- Joseph, F., & Ramakrishnan, N. (2015). Text categorization using improved K nearest neighbor algorithm. *Int J Trends Eng Technol*, 4, 65-68.
- Kadhim, A. I. (2019). Survey on supervised machine learning techniques for automatic text classification. *Artificial Intelligence Review*, 52(1), 273-292. <https://doi.org/10.1007/s10462-018-09677-1>
- Kaplan, G. (2017). *İdari Yargılama Hukuku*. Ekin Yayınevi.
- Kaplan, O. (2020, Kasım 20). Yapay Zeka Kavramına İdare Hukuku Açısından Bir Yaklaşım. *3. Uluslararası Sosyal Bilimlerde Kritik Tartışmalar Kongresi*.
- Kara Kılıçarslan, S. (2019). Yapay Zekanın Hukuki Statüsü ve Hukuki Kişiliği Üzerine Tartışmalar. *Yıldırım Beyazıt Hukuk Dergisi*, 2, 363-389. <https://doi.org/https://doi.org/10.33432/ybuhukuk.599224>
- Karahan Adalı, G. (2022). Denetimli Makine Öğrenmesi Aşamaları. İçinde N. Bayram Arlı, S. Gürsakal, & M. Engin (Ed.), *Denetimli Makine Öğrenmesi Algoritmaları R ve Python Uygulamaları* (1. Baskı, ss. 11-25). Nobel.
- Katz, M. D., Bommarito, M. J., & Blackman, J. (2017). A general approach for predicting the behavior of the Supreme Court of the United States. *PLoS ONE*, 12(4). <https://doi.org/10.1371/journal.pone.0174698>
- Kauffman, M. E., & Soares, M. N. (2020). AI in legal services: new trends in AI-enabled legal services. *Service Oriented Computing and Applications*, 14(4), 223-226. <https://doi.org/10.1007/s11761-020-00305-x>
- Kaur, A., & Bozic, B. (2019). Convolutional Neural Network-based Automatic Prediction of Judgments of the European Court of Human Rights. *Irish Conference on Artificial Intelligence and Cognitive Science.*, 458-469.
- Kayacı, M. T. (2012). *Türkiye’de Yargının İdaresi ve Hakimler ve Savcılar Yüksek Kurulu* [Yayımlanmış Yüksek Lisans Tezi]. Akdeniz Üniversitesi Sosyal Bilimler Enstitüsü.
- Kaynak Balta, B. (2020). Yapay Zeka Ürünlerinin Hukuki Niteliği Ve Fikri Eser Kavramı. *Ankara Hacı Bayram Veli Üniversitesi Hukuk Fakültesi Dergisi*, 24(3), 205-230. <https://doi.org/10.34246/ahbvuhfd.775498>

- Kelly, K. (2017). *Büyük Teknolojik Dönüşüm*. Optimist Kitap.
- Khodayar, M., & Regan, J. (2023). Deep Neural Networks in Power Systems: A Review. *Energies*, 16(12), 4773. <https://doi.org/10.3390/en16124773>
- Kılıç, B., & Öner, Y. (2021). *Yargıtay Kararlarının Suç Türlerine Göre Makine Öğrenmesi Yöntemleri İle Sınıflandırılması* (C. 4, Sayı 3). <https://dergipark.org.tr/en/download/article-file/2032425>
- Kim, H., Cho, H., & Ryu, D. (2022). Corporate Bankruptcy Prediction Using Machine Learning Methodologies with a Focus on Sequential Data. *Computational Economics*, 59(3), 1231-1249. <https://doi.org/10.1007/s10614-021-10126-5>
- Kolbjørnsrud, V., Amico, R., & Thomas, R. J. (2016). How Artificial Intelligence Will Redefine Management. *Harvard Business Review*, 2(1), 3-10.
- Kowsari, K., Meimandi, K. J., Heidarysafa, M., Mendu, S., Barnes, L., & Brown, D. (2019). Text classification algorithms: A survey. İçinde *Information (Switzerland)* (C. 10, Sayı 4). MDPI AG. <https://doi.org/10.3390/info10040150>
- Kowsrihawatt, K., Vateekul, P., & Boonkwan, P. (2018). Predicting Judicial Decisions of Criminal Cases from Thai Supreme Court Using Bi-directional GRU with Attention Mechanism. *2018 5th Asian Conference on Defense Technology (ACDT)*, 50-55. <https://doi.org/10.1109/ACDT.2018.8592948>
- Köroğlu, Y. (2017). Yapay Zeka'nın Teorik ve Pratik Sınırları. *Bogaziçi Üniversitesi Yayinevi*, 1-10.
- Krogh, A. (2008). What are artificial neural networks? *Nature Biotechnology*, 26, 195-197.
- Kuppala, J., Srinivas, K. K., Anudeep, P., Kumar, R. S., & Vardhini, P. A. H. (2022). Benefits of Artificial Intelligence in the Legal System and Law Enforcement. *2022 International Mobile and Embedded Technology Conference (MECON)*, 221-225. <https://doi.org/10.1109/MECON53876.2022.9752352>
- Kuru, B., & Aydın, B. (2020). *İstinaf Sistemine Göre Yazılmış Medeni Usul Hukuku Ders Kitabı* (4. Baskı). Yetkin Yayınları.
- Lage-Freitas, A., Allende-Cid, H., Santana, O., & de Oliveira-Lage, L. (2019). *Predicting Brazilian court decisions*.
- Lancashire, L. J., Lemetre, C., & Ball, G. R. (2008). An introduction to artificial neural networks in bioinformatics--application to complex microarray and mass spectrometry datasets in cancer studies. *Briefings in Bioinformatics*, 10(3), 315-329. <https://doi.org/10.1093/bib/bbp012>
- Lantz, B. (2015). *Machine Learning with R - Second Edition* (Second edition). Packt Publishing.

- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444. <https://doi.org/10.1038/nature14539>
- Lewis, T. (2014). *A Brief History of Artificial Intelligence*. LiveScience Retrieved. <https://www.livescience.com/49007-history-of-artificial-intelligence.html>
- Liaw, A., & Wiener, M. (2002). Classification and Regression by randomForest. *R News*, 2(3), 18-22.
- Liddy, E. D. (2001). Natural Language Processing. İçinde M. Decker (Ed.), *Encyclopedia of Library and Information Science* (2nd Edition). Syracuse University SURFACE.
- Liu, Z., & Chen, H. (2017). 2017 IEEE Symposium Series on Computational Intelligence (SSCI). *A predictive performance comparison of machine learning models for judicial cases*, 1-6. <https://doi.org/10.1109/SSCI.2017.8285436>
- Long, S., Tu, C., Liu, Z., & Sun, M. (2019). Automatic Judgment Prediction via Legal Reading Comprehension. In *Chinese Computational Linguistics: 18th China National Conference, CCL 2019*, 558-572. <https://doi.org/https://doi.org/10.48550/arXiv.1809.06537>
- Mahesh, B. (2020). Machine Learning Algorithms -A Review. *International Journal of Science and Research (IJSR)*, 9(1), 381-386. <https://doi.org/10.21275/ART20203995>
- Malik, V., Sanjay, R., Nigam, S. K., Ghosh, K., Guha, S. K., Bhattacharya, A., & Modi, A. (2021). *ILDC for CJPE: Indian Legal Documents Corpus for Court Judgment Prediction and Explanation*.
- Medvedeva, M., Vols, M., & Wieling, M. (2020). Using machine learning to predict decisions of the European Court of Human Rights. *Artificial Intelligence and Law*, 28(2), 237-266. <https://doi.org/10.1007/s10506-019-09255-y>
- Medvedeva, M., Wieling, M., & Vols, M. (2023). Rethinking the field of automatic prediction of court decisions. *Artificial Intelligence and Law*, 31(1), 195-212. <https://doi.org/10.1007/s10506-021-09306-3>
- Metsker, O., Trofimov, E., Sikorsky, S., & Kovalchuk, S. (2019). Text and data mining techniques in judgment open data analysis for administrative practice control. *Communications in Computer and Information Science*, 947, 169-180. https://doi.org/10.1007/978-3-030-13283-5_13
- Mijwel, M. M. (2018, Ocak 27). *Artificial Neural Networks Advantages and Disadvantages*. <https://www.linkedin.com/pulse/artificial-neural-networks-advantages-disadvantages-maad-m-mijwel>
- Möslein, F., & Bora Çınar, S. (2021). Yönetim Kurulu Toplantı Odasındaki Robotlar: Yapay Zeka ve Şirketler Hukuku. *İstanbul Hukuk Mecmuası*, 79(2), 699-728. <https://doi.org/10.26650/mecmua.2021.79.2.0004>

- Mugunthan, S. R., & Vijayakumar, T. (2021). Design of Improved Version of Sigmoidal Function with Biases for Classification Task in ELM Domain. *Journal of Soft Computing Paradigm*, 3(2), 70-82. <https://doi.org/10.36548/jscp.2021.2.002>
- Mumcuoğlu, E., Öztürk, C. E., Ozaktas, H. M., & Koç, A. (2021). Natural language processing in law: Prediction of outcomes in the higher courts of Turkey. *Information Processing and Management*, 58(5). <https://doi.org/10.1016/j.ipm.2021.102684>
- Murphy, K. P. (2022). *Probabilistic Machine Learning An Introduction* (C. Bishop, D. Heckerman, M. Jordan, & M. Kearns, Ed.). The MIT Press.
- Muthukadan, B. (2023). *Selenium with Python*. <https://selenium-python.readthedocs.io/>.
- Nabiyev, V. (2005). *Yapay Zeka* (6. Baskı). Seçkin Yayıncılık.
- Navlani, A. (2019). *Scikit-learn SVM Tutorial with Python (Support Vector Machines)*. <https://www.datacamp.com/tutorial/svm-classification-scikit-learn-python>.
- Newell, A., & Simon, H. (1956). The logic theory machine--A complex information processing system. *IEEE Transactions on Information Theory*, 2(3), 61-79. <https://doi.org/10.1109/TIT.1956.1056797>
- Ng, A., & Jordan, M. I. (2001). On Discriminative vs. Generative Classifiers: A comparison of logistic regression and naive Bayes. *Advances in Neural Information Processing Systems*, 21, 841-848.
- Nikolskaia, K., & Naumov, V. (2020). Artificial Intelligence in Law. *2020 International Multi-Conference on Industrial Engineering and Modern Technologies (FarEastCon)*, 1-4. <https://doi.org/10.1109/FarEastCon50210.2020.9271095>
- Oard, D. W. (1997). The State of the Art in Text Filtering. *User Modeling and User-Adapted Interaction*, 7, 141-178. <https://doi.org/https://doi.org/10.1023/A:1008287121180>
- Onan, A., & Korukoğlu, S. (2016). Metin Sınıflandırmada Öznitelik Seçim Yöntemlerinin Değerlendirilmesi. *Akademik Bilişim*.
- Orhan, U., & Adem, K. (2012). Naive Bayes Yönteminde Olasılık Çarpanlarının Etkileri The Effects of Probability Factors in Naive Bayes Method. *Elektrik- Elektronik ve Bilgisayar Mühendisliği Sempozyumu*, 722-724.
- O'Sullivan, C., & Beel, J. (2019). *Predicting the Outcome of Judicial Decisions made by the European Court of Human Rights*.
- Öngen Bilir, B. (2022). Doğrusal Diskriminant Analizi. İçinde N. Bayram Arlı, S. Gürsakal, & M. Engin (Ed.), *Denetimli Makine Öğrenmesi Algoritmaları R ve Python Uygulamaları* (1. bs, ss. 85-99). Nobel.
- Özbudun, E. (2008). *Türk Anayasa Hukuku*. Yetkin Yayınları.

- Özen, Z. (2022). Yapay Sinir Ağları. İçinde N. Bayram Arlı, S. Gürsakal, & M. Engin (Ed.), *Denetimli Makine Öğrenmesi Algoritmaları R ve Python Uygulamaları* (1. Baskı, ss. 215-262). Nobel.
- Özkan, N. (2021). Kripto Para Piyasasında Fiyat Değişimlerinin Makine Öğrenmesi Yöntemleriyle Tahmini. İçinde Y. A. Ünvan (Ed.), *Bankacılık ve Finans Çalışmaları* (ss. 92-115). Livre de Lyon.
- Özparlak, B. O. (2019). Otomatik Hukuk: Hukuk Sisteminde ve Mesleğinde Yapay Zekâ. *Hukuk Defterleri*, 20.
- Öztemel, E. (2020). Yapay zekâ ve insanlığın geleceği. İçinde M. Şeker, Y. Bulduklu, C. Korkut, & M. Doğrul (Ed.), *Bilişim Teknolojileri ve İletişim: Birey ve Toplum Güvenliği* (ss. 75-91). Türkiye Bilimler Akademisi.
- Patra, R. K., Konda, S., Rao, M. V., Balmur, K., & Madhukar, G. (2022). Predictive Model and Theory of Interaction. İçinde S. Kumar, R. Raja, S. Tiwari, & S. Rani (Ed.), *Cognitive Behavior and Human Computer Interaction Based on Machine Learning Algorithm* (ss. 179-211). Scrivener Publishing LLC.
- Peng, F., Schuurmans, D., & Wang, S. (2004). Augmenting Naive Bayes Classifiers with Statistical Language Models. *Information Retrieval*, 7, 317-345. <https://doi.org/https://doi.org/10.1023/B:INRT.0000011209.19643.e2>
- Pennington, J., Socher, R., & Manning, C. (2014). GloVe: Global Vectors for Word Representation. *2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1532-1543. <https://doi.org/10.3115/v1/D14-1162>
- Pirim, H. (2006). YAPAY ZEKA. *Journal of Yasar University*, 1(1), 81-93.
- Quemy, A., & Wrembel, R. (2020). On Integrating and Classifying Legal Text Documents. *International Conference on Database and Expert Systems Applications*, 385-399. https://doi.org/https://doi.org/10.1007/978-3-030-59003-1_25
- Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, 1, 81-106. <https://doi.org/https://doi.org/10.1007/BF00116251>
- Ramasubramanian, K., & Moolayil, J. (2019). *Applied Supervised Learning with R* (V. Mewada, A. Date, & S. Zagade, Ed.). Packt Publishing Ltd.
- Reiling, A. D. (Dory). (2020). Courts and Artificial Intelligence. *International Journal For Court Administration*, 11(2), 2-10. <https://doi.org/10.36745/ijca.343>
- RoboLaw. (2014). *RoboLaw: Project Overview*. RoboLaw. <http://www.robolaw.eu/projectdetails.htm>
- Rojas, R. (1996). *Neural Networks: A Systematic Introduction* (1. bs). Springer.
- Rokach, L., & Maimon, O. (2015). *Data Mining With Decision Trees Theory And Applications* (2. bs). World Scientific Publishing Co. Pte. Ltd.

- Sadiq, A. T., & Abdullah, S. M. (2012). Hybrid Intelligent Technique for Text Categorization. *2012 International Conference on Advanced Computer Science Applications and Technologies (ACSAT)*, 238-245. <https://doi.org/10.1109/ACSAT.2012.50>
- Sakarkar, G., Patil, G., & Dutta, P. (2021). *Machine Learning Algorithms Using Python Programming*. Nova Science Publishers, Inc.
- Say, C. (2021). *50 Soruda Yapay Zeka*. Bilim ve Gelecek Kitaplığı.
- Schröer, C., Kruse, F., & Gómez, J. M. (2021). A Systematic Literature Review on Applying CRISP-DM Process Model. *Procedia Computer Science*, 181, 526-534. <https://doi.org/10.1016/j.procs.2021.01.199>
- Sert, M. F., Yıldırım, E., & Haşlak, İ. (2022). Using Artificial Intelligence to Predict Decisions of the Turkish Constitutional Court. *Social Science Computer Review*, 40(6), 1416-1435. <https://doi.org/10.1177/08944393211010398>
- Shearer, C. (2000). The CRISP-DM Model: The New Blueprint for Data Mining. *Journal of Data Warehousing*, 5(4), 13-22.
- Shi, C., Sourdin, T., & Li, B. (2021). The Smart Court – A New Pathway to Justice in China? *International Journal for Court Administration*, 12(1). <https://doi.org/10.36745/ijca.367>
- Sil, R., Roy, A., Bhushan, B., & Mazumdar, A. K. (2019). Artificial Intelligence and Machine Learning based Legal Application: The State-of-the-Art and Future Research Trends. *2019 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)*, 57-62. <https://doi.org/10.1109/ICCCIS48478.2019.8974479>
- Silahtaroglu, G. (2013). *Veri madenciliği: Kavram ve Algoritmaları*. Papatya.
- Singh, G. (2021, Eylül 6). *Introduction to Artificial Neural Networks*. <https://www.analyticsvidhya.com/blog/2021/09/introduction-to-artificial-neural-networks/>.
- Sorhun, E. (2021). *Python ile Makine Öğrenmesi* (2. Baskı). Abaküs Kitap Yayın.
- Sourdin, T. (2021). *Judges, Technology and Artificial Intelligence The Artificial Judge*. Edward Elgar Publishing.
- Soysal, M. (1990). *100 Soruda Anayasanın Anlamı*. Gerçek Yayınevi.
- Sparkes, M. (2023). AI will advise a defendant in court. *New Scientist*, 257(3421), 8. [https://doi.org/10.1016/S0262-4079\(23\)00042-8](https://doi.org/10.1016/S0262-4079(23)00042-8)
- Strickson, B., & De La Iglesia, B. (2020). Legal Judgement Prediction for UK Courts. *Proceedings of the 2020 The 3rd International Conference on Information Science and System*, 204-209. <https://doi.org/10.1145/3388176.3388183>

- Sulea, O.-M., Zampieri, M., Malmasi, S., Vela, M., Dinu, L. P., & van Genabith, J. (2017). *Exploring the Use of Text Classification in the Legal Domain*.
- Sulea, O.-M., Zampieri, M., Vela, M., & van Genabith, J. (2017). Predicting the Law Area and Decisions of French Supreme Court Cases. *RANLP 2017 - Recent Advances in Natural Language Processing Meet Deep Learning*, 716-722. https://doi.org/10.26615/978-954-452-049-6_092
- Sundquist, C. P. (2020). Technology and the (Re) Construction of Law. *Journal of Legal Education*, 70, 402.
- Sümer, S. Y. (2021). CEZA YARGILAMASININ GELECEĞİ: ROBOT HAKİM. *Dokuz Eylül Üniversitesi Hukuk Fakültesi Dergisi*, 1543-1591. <https://doi.org/10.33717/deuhfd.998255>
- Szegedy, C., Wei Liu, Yangqing Jia, Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1-9. <https://doi.org/10.1109/CVPR.2015.7298594>
- Şeker, S. E. (2015). Doğal Dil İşleme (Natural Language Processing). *YBS Ansiklopedi*, 2(4), 14-31.
- Tan, A.-H. (1999). Text mining: The state of the art and the challenges. *the pakdd 1999 workshop on knowledge discovery from advanced databases*, 65-70.
- Tantuğ, A. C. (2016). Metin Sınıflandırma. *Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi*, 5(2).
- Tanveer, M., Rajani, T., Rastogi, R., Shao, Y. H., & Ganaie, M. A. (2021). Comprehensive Review On Twin Support Vector Machines. *Annals of Operations Research*, 1-46. <https://doi.org/10.1007/s10479-022-04575-w>
- Tegmark, M. (2017). *Life 3.0: Being human in the age of artificial intelligence*. Knopf Doubleday Publishing Group.
- Teziç, E. (2003). *Anayasa Hukuku* (8. bs). Beta.
- Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind, New Series*, 59, 433-460.
- Uçkan, T., & Karcı, A. (2020). Extractive multi-document text summarization based on graph independent sets. *Egyptian Informatics Journal*, 21(3), 145-157. <https://doi.org/10.1016/j.eij.2019.12.002>
- van Gelder, E. (2017, Aralık 10). *Digitalisering in buitengerechtelijke geschiloplossing*. KEIDUIDELIJK Alles over KEI. <http://www.keiduidelijk.nl/2017/12/10/digitalisering-in-buitengerechtelijke-geschiloplossing/>

- Vasdani, T. (2020). *Robot justice: China's use of Internet courts*. LexisNexis Canada Inc. <https://www.lexisnexis.ca/en-ca/ihc/2020-02/robot-justice-chinas-use-of-internet-courts.page>
- Venkateswaran, B., & Ciaburro, G. (2017). *Neural Networks with R*. Packt Publishing.
- Virtucio, M. B. L., Aborot, J. A., Abonita, J. K. C., Avinante, R. S., Copino, R. J. B., Neverida, M. P., Osiana, V. O., Peramo, E. C., Syjuco, J. G., & Tan, G. B. A. (2018). Predicting Decisions of the Philippine Supreme Court Using Natural Language Processing and Machine Learning. *2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC)*, 130-135. <https://doi.org/10.1109/COMPSAC.2018.10348>
- Visentin, A., Nardotto, A., & Osullivan, B. (2019). Predicting judicial decisions: A statistically rigorous approach and a new ensemble classifier. *Proceedings - International Conference on Tools with Artificial Intelligence, ICTAI, 2019-November*, 1820-1824. <https://doi.org/10.1109/ICTAI.2019.00275>
- Walczak, S. (2019). *Artificial Neural Networks* (ss. 40-53). <https://doi.org/10.4018/978-1-5225-7368-5.ch004>
- Wang, M., & Deng, W. (2021). Deep face recognition: A survey. *Neurocomputing*, 429, 215-244. <https://doi.org/10.1016/j.neucom.2020.10.081>
- Winston, P. H. (1992). *Artificial Intelligence* (Third edition). Addison-Wesley.
- Wong, T.-T. (2015). Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. *Pattern Recognition*, 48(9), 2839-2846. <https://doi.org/10.1016/j.patcog.2015.03.009>
- Yang, Z., Wang, P., Zhang, L., Shou, L., & Xu, W. (2019). A Recurrent Attention Network for Judgment Prediction. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11730 LNCS, 253-266. https://doi.org/10.1007/978-3-030-30490-4_21
- Yeniay, L. (2020). Adil Yargılanma Hakkı Bağlamında Yargı Ayrılığı Sistemi Üzerine Bir İnceleme. *Marmara Üniversitesi Hukuk Fakültesi Hukuk Araştırmaları Dergisi*, 26(1), 161-189.
- Yıldız, H. K., Gençtav, M., Usta, N., Diri, B., & Amasyalı, M. F. (2007). *Metin Sınıflandırmada Yeni Özellik Çıkarımı A New Feature Extraction Method for Text Classification*.
- Yılmaz, G. (2020). Yapay Zekânın Yargı Sistemlerinde Kullanılmasına İlişkin Avrupa Etik Şartı. *Marmara Üniversitesi Avrupa Topluluğu Enstitüsü Avrupa Araştırmaları Dergisi*, 28(1), 27-55. <https://doi.org/10.29228/mjes.8>
- Yılmaz, O. G. (2021). Yargı Uygulamasında Yapay Zekâ Kullanımı - Yapay Zekâ Hâkim Cübbesini Giyebilecek mi? *Adalet Dergisi*, 66, 379-415.

- Yin, Y., Yang, H., Zhao, Z., & Chen, S. (2019). A Judicial Sentencing Method Based on Fused Deep Neural Networks. *Artificial Neural Networks and Machine Learning – ICANN 2019: Text and Time Series*, 213-226.
- Zeytin, Z., & Gençay, E. (2019). Hukuk ve Yapay Zekâ: E-Kişi, Mali Sorumluluk ve Bir Hukuk Uygulaması. *Türk-Alman Üniversitesi Hukuk Fakültesi Dergisi*, 1(1), 39-70.
- Zhou, Z.-H. (2021). *Machine Learning*. Springer Singapore.
<https://doi.org/10.1007/978-981-15-1967-3>
- Zhu, X., & Goldberg, A. B. (2009). *Introduction to Semi-Supervised Learning* (C. 3). Springer. <https://doi.org/10.2200/S00196ED1V01Y200906AIM006>

EKLER

EK1: WEB BOT KODLARI

```
#constraints

BASE_URL = "https://kararlar.uyusmazlik.gov.tr/"

#veri çekme

import Uyumazlik.constraints as const

import os

from selenium import webdriver

from selenium.webdriver.common.keys import Keys

import pandas as pd

import lxml

import time

import re

import string


from selenium.webdriver.chrome.options import Options

class Uyumazlik(webdriver.Chrome):

    def __init__(self,driver_path=r"\\C:\\SeleniumDrivers"):

        self.driver_path=driver_path

        os.environ['PATH'] += self.driver_path

        super(Uyumazlik, self).__init__()

        self.implicitly_wait(15)
```

```

self.maximize_window()

self.text = ""

def __exit__(self, exc_type, exc_val, exc_tb):
    #self.quit()

    pass

def land_first_page(self):
    self.get(const.BASE_URL)

def get_deccisions_summary(self, decision):
    try:
        decision = "\"" + decision + "\""

        sub1 = "ÖZET"

        sub2 = "K A R A R"

        idx1 = decision.index(sub1)

        idx2 = decision.index(sub2)

        summary = ""

        for idx in range(idx1 + len(sub1) + 1, idx2):
            summary = summary + decision[idx]

    except:
        summary = ""

    return summary

```

```

def get_deccision_result(self,decision):

    try:

        decision = "\"" + decision + "\""

        sub1 = "S O N U Ç  :"

        sub2 = "          Başkan          Üye "

        idx1 = decision.index(sub1)

        idx2 = decision.index(sub2)

        result = "

        for idx in range(idx1 + len(sub1) + 1, idx2):

            result = result + decision[idx]

    except:

        result=""

    return result

def get_decision_action(self,decision):

    try:

        decision = "\"" + decision + "\""

        sub1 = " O L A Y      :"

        sub2 = "S O N U Ç  :"

        idx1 = decision.index(sub1)

        idx2 = decision.index(sub2)

        action = "

```

```

        for idx in range(idx1 + len(sub1) + 1, idx2):

            action = action + decision[idx]

    except:

        action=""

    return action

def find_string_index(self,htmldata):

    pass

def find_tags_with_text_proccesing(self,text):

    clear_blank=str(text).replace(" ","").lower()

    sub1 = "olay:"

    sub2 = "sonuç:"

    idx1 = clear_blank.index(sub1)

    idx2 = clear_blank.index(sub2)

    action = "

    for idx in range(idx1 + len(sub1) + 1, idx2):

        action = action + text[idx]

    print("olayson",action,"---")

def loadDecissions(self):

    pagecount = self.find_element_by_class_name("totalItemCount ").text

```

```

pg = int(pagecount.replace(" adet kayıt bulundu", ""))

items = self.find_elements_by_css_selector("div[data-rel='popover']")

count = 0

u_text=""

olay=""

sonuctext=""

for i in items:

    print(count,"dosyası",i.text)

    time.sleep(1)

    self.text += i.text + "%+%"

    count = count + 1

    if count % 5 == 1:

        link = i.find_element_by_tag_name("a").click()

        time.sleep(1)

        self.switch_to.window(self.window_handles[1])

        try:

            try:

                detay = self.find_element_by_class_name("WordSection1")

            except:

                detay = self.find_element_by_class_name("Section1")

        sonuc = detay.text

```

```

parcali = sonuc.split("\n")

for i in parcali:

    if (i[0:50].replace(" ", "").lower().find("olay") > -1 and
i[0:50].replace(" ",
                                                                    "").lower().find(
                                                                    ":")) > -1):

        if (i[0:50].replace(" ", "").lower().find(":") > 0):

            olay = i

self.close()

self.switch_to.window(self.window_handles[0])

summary_of_decision=""#self.get_deccisions_summary(sonuc)

action_of_decision=olay#self.get_decision_action(sonuc)

result_of_decision=sonuctext#self.get_deccision_result(sonuc)

self.text = self.text + sonuc + "%+%"

self.text = self.text + summary_of_decision + "%+%"

self.text = self.text + action_of_decision + "%+%"

self.text = self.text + result_of_decision + "%+%"

except:

    print(count,"numaralı dosyada hata alındı")

    self.close()

    self.switch_to.window(self.window_handles[0])

    continue

""""if count==4:

```



```

if str(i.text).lower().find("adli")>0:

    durum="0"

elif str(i.text).lower().find("idari")>0:

    durum="1"

else:

    durum="2"

self.text = self.text + durum + "%+%"

if count != 0 and count % 5 == 0:

    count = 0

    self.text = self.text + "%#%"

```

```

def text_excel(self, text):

    df_new = pd.DataFrame(

        columns=['kararsayisi', 'Karar', 'Ozet','Olay','Sonuc','esassayisi', 'Bolum',
'Uyusmazlik', 'Kararsonucu', 'pdf'])

```

```

for i in text.split("%#%"):

    temp = i.split("%+%"

    """for j in temp:

        print("-->",j)"""

    print("içişler ", temp)

    try:

        if len(temp) > 1:

```

```

        a_series = pd.Series(temp, index=df_new.columns)

        df_new = df_new.append(a_series, ignore_index=True)

    except:

        continue

    df_new.to_excel("Uyusmazlik_Kararlari_excel.xlsx",encoding="utf-8-sig")


def search_decissions(self):

    Bolum=self.find_element_by_css_selector('option[value="96b26fc4-ef8e-4a4f-a9cc-a3de89952aa1"]')

    Bolum.click()

    time.sleep(1)

    uyusmazlik_gorev =
self.find_element_by_css_selector('option[value="7b1e2cd3-8f09-418a-921c-bbe501e1740c"]')

    uyusmazlik_gorev.click()

    time.sleep(1)

    karar_sonucu=self.find_element_by_class_name("dropdown-toggle")

    karar_sonucu.click()

    time.sleep(1)


    karar_sonucu_adli=self.find_element_by_css_selector('input[value="bf7e4bf9-2181-46a4-abc2-51286202be96"]')

    karar_sonucu_adli.click()

    time.sleep(1)

```

```

karar_sonucu_idari=self.find_element_by_css_selector('input[value="fcd1fa44-f4d9-4d0f-970b-6769ece367a2"]')

        karar_sonucu_idari.click()

        time.sleep(1)

        karar_sonucu.click()

        search_button = self.find_element_by_css_selector('button[class="btn btn-primary"]')

        search_button.click()

        time.sleep(1)


    def getDecissions(self):

        """page_input =
self.find_element_by_css_selector('input[class="pageInput"]')

        page_input.clear()

        page_input.send_keys("226")

        page_input.send_keys(Keys.ENTER)

        time.sleep(2)"""


    while True:

        try:

            links = self.find_element_by_class_name('icon-chevron-right')

            data_max =

self.find_element_by_class_name('pageInput').get_attribute("data-max")

```

```

        current_page=
self.find_element_by_class_name('pageInput').get_attribute("value")

print(links.text,"--",data_max,"--",current_page)

if int(current_page) < 1:#int(data_max):

    links.click()

    time.sleep(2)

    self.loadDecissions()

else:

    links.click()

    break

except:

    continue

self.text_excel(self.text)

```

EK2: DOĞAL DİL İŞLEME KODLARI

```
import re

import string

import numpy as np

import random

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

from plotly import graph_objs as go

import plotly.express as px

import plotly.figure_factory as ff

from collections import Counter

from PIL import Image

from wordcloud import WordCloud, STOPWORDS, ImageColorGenerator

import nltk

from nltk.corpus import stopwords

from nltk.tokenize import word_tokenize

from tqdm import tqdm

import os

import nltk

import spacy

import random

from spacy.util import compounding
```

```

from spacy.util import minibatch

from collections import defaultdict

from collections import Counter

from sklearn.metrics import (

    precision_score,

    recall_score,

    f1_score,

    classification_report,

    accuracy_score

)

import plotly.io as pio

#pio.renderers.default = 'svg'

pio.renderers.default = 'svg'


primary_blue = "#496595"

primary_blue2 = "#85a1c1"

primary_blue3 = "#3f4d63"

primary_grey = "#c6ccd8"

primary_black = "#202022"

primary_bgcolor = "#f4f0ea"

primary_green = px.colors.qualitative.Plotly[2]

df = pd.read_excel("uk3.xlsx")

df_new=df[["Kararsonucu","Olay"]]

```

```

df = df.dropna(how="any", axis=1)#herhangi bir satır veya sutun boş ise kaldır

df_new.columns = ['target', 'message']

df_new.head()

#%%

df_new['message_len'] = df_new['message'].apply(lambda x: len(str(x).split(' ')))

df_new.head()

df_new=df_new.dropna(how="any")

import plotly.graph_objects as go

import plotly.io as pio

#pio.renderers.default = 'svg'

pio.renderers.default = 'svg'

balance_counts = df_new.groupby('target')['target'].agg('count').values

#figure out balancing situation about total count of target values separetly

fig = go.Figure()

fig.add_trace(go.Bar(

    x=['Adli'],

    y=[balance_counts[0]],

    name='Adli',

    text=[balance_counts[0]],

    textposition='auto',

    marker_color=primary_blue

))

fig.add_trace(go.Bar(

```

```

x=['İdari'],

y=[balance_counts[1]],

name='İdari',

text=[balance_counts[1]],

textposition='auto',

marker_color=primary_grey

))

fig.update_layout(

    title='<span style="font-size:32px; font-family:Times New Roman">Veri
setinin dağılımı</span>'

)

fig.show()

nltk.download('stopwords')

stop_words = stopwords.words('turkish')

def clean_text(text):

    text = str(text).lower()

    text = re.sub('[.*?\]', "", text)

    text = re.sub('https?://\S+|www\.\S+', "", text)

    text = re.sub('<.*?>+', "", text)

    text = re.sub('[%s]' % re.escape(string.punctuation), "", text)

    text = re.sub('\n', "", text)

    text = re.sub('\w*\d\w*', "", text)

    return text

```



```

stop_words = stopwords.words('turkish')

def remove_stopwords(text):

    text = ' '.join(word for word in text.split(' ') if word not in stop_words)

    return text

def preprocess_data(text):

    text = clean_text(text)

    text = ' '.join(word for word in text.split(' ') if word not in stop_words)

    return text

df_new["message_last"]=df_new["message"]

df_new['message'] = df_new['message'].apply(preprocess_data)

### Target encoding

from sklearn.preprocessing import LabelEncoder

le = LabelEncoder()

le.fit(df_new['target'])

df_new['target_encoded'] = le.transform(df_new['target'])

df.head()

file_id2 = '1Udmij_22ovv_pEDbwIybY0jd_UNySZPS'

downloaded = drive.CreateFile({'id': file_id2})

downloaded.GetContentFile('black-square.jpg')

### Tokens visualization

twitter_mask = np.array(Image.open('black-square.jpg'))

wc = WordCloud(

    background_color='white',

```

```

        max_words=200,

        mask=twitter_mask,

    )

    wc.generate(' '.join(str(text) for text in df_new.loc[df_new['target'] == 'Adli',
'message']))

    plt.figure(figsize=(18,10))

    plt.title('Adli Kararlar',

        fontdict={'size': 22, 'verticalalignment': 'bottom'})

    plt.imshow(wc)

    plt.axis("off")

    plt.show()

#visualization spam message

twitter_mask = np.array(Image.open('black-square.jpg'))

wc = WordCloud(

    background_color='white',

    max_words=200,

    mask=twitter_mask,

)

wc.generate(' '.join(str(text) for text in df_new.loc[df_new['target'] == 'İdari',
'message']))

plt.figure(figsize=(18,10))

plt.title('İdari Kararlar',

    fontdict={'size': 22, 'verticalalignment': 'bottom'})

plt.imshow(wc)

```

```

plt.axis("off")

plt.show()

### Vectorization

# 1-Split into train and test sets

from sklearn.model_selection import train_test_split

x = df_new['message']

y = df_new['target_encoded']

x_train, x_test, y_train, y_test = train_test_split(x, y, random_state=42)

print(len(x_train), len(y_train))

print(len(x_test), len(y_test))

#2-vektorization

from sklearn.feature_extraction.text import CountVectorizer

# instantiate the vectorizer

vect = CountVectorizer()

vect.fit(x_train)

### Tuning CountVectorizer

#vect_tunned = CountVectorizer(stop_words='english', ngram_range=(1,2),
min_df=0.1, max_df=0.7, max_features=100)

#vect_tunned.fit(x_train)

x_train_dtm = vect.transform(x_train)

x_test_dtm = vect.transform(x_test)

### TF-IDF

from sklearn.feature_extraction.text import TfidfTransformer

tfidf_transformer = TfidfTransformer()

```

```
tfidf_transformer.fit(x_train_dtm)
```

```
x_train_tfidf = tfidf_transformer.transform(x_train_dtm)
```

```
x_train_tfidf
```

EK3: MAKİNE ÖĞRENMESİ KODLARI

```
##### Modeling

import plotly.figure_factory as ff

x_axes = ['Adli', 'İdari']
y_axes = ['İdari', 'Adli']

def conf_matrix(z, x=x_axes, y=y_axes):

    z = np.flip(z, 0)

    # change each element of z to type string for annotations
    z_text = [[str(y) for y in x] for x in z]

    # set up figure

    fig = ff.create_annotated_heatmap(z, x=x, y=y, annotation_text=z_text,
    colorscale='Viridis')

    # add title

    fig.update_layout(title_text='<b>Confusion matrix</b>',

        xaxis = dict(title='Predicted value'),

        yaxis = dict(title='Real value')

    )

    # add colorbar

    fig['data'][0]['showscale'] = True

    return fig

#rassal orman

from sklearn.ensemble import RandomForestClassifier

from sklearn.inspection import permutation_importance

from sklearn.metrics import confusion_matrix, ConfusionMatrixDisplay
```

```

from sklearn.metrics import classification_report

rf=RandomForestClassifier(n_estimators=1000,
criterion='entropy',random_state=2)

rf.fit(x_train_dtm , y_train)

print("Accuracy on training set: {:.3f}".format(rf.score(x_train_dtm , y_train)))

y_tahmin_e = rf.predict(x_train_dtm)

cm = confusion_matrix(y_train,y_tahmin_e)

display = ConfusionMatrixDisplay(cm).plot()

plt.title('Karmaşıklık Matrisi-Eğitim')

plt.show()

y_tahmin_t = rf.predict(x_test_dtm)

cm = confusion_matrix(y_test,y_tahmin_t)

display = ConfusionMatrixDisplay(cm).plot()

plt.title('Karmaşıklık Matrisi-Test')

plt.show()

print("Accuracy on test set: {:.3f}".format(rf.score(x_test_dtm,y_test)))

cr= classification_report(y_test,y_tahmin_t)

print("Sınıflama Raporu")

print(cr)

#karar ağaçları

from sklearn.tree import DecisionTreeClassifier

dte = DecisionTreeClassifier(random_state=2, criterion='entropy')

dte.fit(x_train_dtm, y_train)

```

```

y_tahmin_e = dtc.predict(x_train_dtm)

cm = confusion_matrix(y_train,y_tahmin_e)

display = ConfusionMatrixDisplay(cm).plot()

plt.title('Karmaşıklık Matrisi-Eğitim')

plt.show()

y_tahmin_t = dtc.predict(x_test_dtm)

cm = confusion_matrix(y_test,y_tahmin_t)

display = ConfusionMatrixDisplay(cm).plot()

plt.title('Karmaşıklık Matrisi-Test')

plt.show()

print("Accuracy on test set: {:.3f}".format(dtc.score(x_test_dtm,y_test)))

cr= classification_report(y_test,y_tahmin_t)

print("Sınıflama Raporu")

print(cr)

```

#lojistik regresyon

```

from sklearn.linear_model import LogisticRegression as lr

from sklearn.metrics import confusion_matrix, ConfusionMatrixDisplay

from sklearn.metrics import classification_report

lrm= lr().fit(x_train_dtm, y_train)

y_tahmin_e = lrm.predict(x_train_dtm)

cm = confusion_matrix(y_train,y_tahmin_e)

display = ConfusionMatrixDisplay(cm).plot()

```

```

plt.title('Karmaşıklık Matrisi-Eğitim')

plt.show()

y_tahmin_t = lrm.predict(x_test_dtm)

cm = confusion_matrix(y_test,y_tahmin_t)

display = ConfusionMatrixDisplay(cm).plot()

plt.title('Karmaşıklık Matrisi-Test')

plt.show()

print("Accuracy on test set: {:.3f}".format(lrm.score(x_test_dtm,y_test)))

cr= classification_report(y_test,y_tahmin_t)

print("Sınıflama Raporu")

print(cr)

#diskriminant analizi

from sklearn. discriminant_analysis import LinearDiscriminantAnalysis

lda = LinearDiscriminantAnalysis()

x_train_dtm = np.squeeze(np.asarray(x_train_dtm))

lda.fit(x_train_dtm, y_train)

y_tahmin_e = lda.predict(x_train_dtm)

cm = confusion_matrix(y_train,y_tahmin_e)

display = ConfusionMatrixDisplay(cm).plot()

plt.title('Karmaşıklık Matrisi-Eğitim')

plt.show()

y_tahmin_t = lda.predict(x_test_dtm)

cm = confusion_matrix(y_test,y_tahmin_t)

```



```

display = ConfusionMatrixDisplay(cm).plot()

plt.title('Karmaşıklık Matrisi-Test')

plt.show()

print("Accuracy on test set: {:.3f}".format(lda.score(x_test_dtm,y_test)))

cr= classification_report(y_test,y_tahmin_t)

print("Sınıflama Raporu")

print(cr)

```

#destek vektör makineleri

```

import numpy as np

import pandas as pd

import seaborn as sns

from sklearn.preprocessing import MinMaxScaler

from sklearn.model_selection import train_test_split

from sklearn import svm

from sklearn.decomposition import PCA

import matplotlib.pyplot as plt

from sklearn.model_selection import GridSearchCV

dvm= svm.SVC(kernel='linear')

dvm.fit(x_train_tfidf, y_train)

y_tahmin_e = dvm.predict(x_train_tfidf)

cm = confusion_matrix(y_train,y_tahmin_e)

display = ConfusionMatrixDisplay(cm).plot()

plt.title('Karmaşıklık Matrisi-Eğitim')

```

```

plt.show()

y_tahmin_t = dvm.predict(x_test_dtm)

cm = confusion_matrix(y_test,y_tahmin_t)

display = ConfusionMatrixDisplay(cm).plot()

plt.title('Karmaşıklık Matrisi-Test')

plt.show()

print("Accuracy on test set: {:.3f}".format(dvm.score(x_test_dtm,y_test)))

cr= classification_report(y_test,y_tahmin_t)

print("Sınıflama Raporu")

print(cr)

#yapay sinir ağıları

from sklearn.neural_network import MLPClassifier

clf = MLPClassifier(solver='lbfgs', alpha=1e-5,

                    hidden_layer_sizes=(5, 2), random_state=1)

clf.fit(x_train_dtm, y_train)

y_tahmin_e = clf.predict(x_train_dtm)

cm = confusion_matrix(y_train,y_tahmin_e)

display = ConfusionMatrixDisplay(cm).plot()

plt.title('Karmaşıklık Matrisi-Eğitim')

plt.show()

y_tahmin_t = clf.predict(x_test_dtm)

cm = confusion_matrix(y_test,y_tahmin_t)

display = ConfusionMatrixDisplay(cm).plot()

```

```

plt.title('Karmaşıklık Matrisi-Test')

plt.show()

print("Accuracy on test set: {:.3f}".format(clf.score(x_test_dtm,y_test)))

cr= classification_report(y_test,y_tahmin_t)

print("Sınıflama Raporu")

print(cr)

# k-en yakın komşu

from sklearn.neighbors import KNeighborsClassifier

knn = KNeighborsClassifier(n_neighbors=5,metric='euclidean')

knn.fit(x_train_dtm, y_train)

y_tahmin_e = knn.predict(x_train_dtm)

cm = confusion_matrix(y_train,y_tahmin_e)

display = ConfusionMatrixDisplay(cm).plot()

plt.title('Karmaşıklık Matrisi-Eğitim')

plt.show()

y_tahmin_t = knn.predict(x_test_dtm)

cm = confusion_matrix(y_test,y_tahmin_t)

display = ConfusionMatrixDisplay(cm).plot()

plt.title('Karmaşıklık Matrisi-Test')

plt.show()

print("Accuracy on test set: {:.3f}".format(knn.score(x_test_dtm,y_test)))

cr= classification_report(y_test,y_tahmin_t)

print("Sınıflama Raporu")

```

```

print(cr)

#naive bayes

from sklearn.naive_bayes import MultinomialNB

nvb = MultinomialNB(force_alpha=True)

nvb.fit(x_train_dtm, y_train)

y_tahmin_e = nvb.predict(x_train_dtm)

cm = confusion_matrix(y_train,y_tahmin_e)

display = ConfusionMatrixDisplay(cm).plot()

plt.title('Karmaşıklık Matrisi-Eğitim')

plt.show()

y_tahmin_t = nvb.predict(x_test_dtm)

cm = confusion_matrix(y_test,y_tahmin_t)

display = ConfusionMatrixDisplay(cm).plot()

plt.title('Karmaşıklık Matrisi-Test')

plt.show()

print("Accuracy on test set: {:.3f}".format(nvb.score(x_test_dtm,y_test)))

cr= classification_report(y_test,y_tahmin_t)

print("Sınıflama Raporu")

print(cr)

```

ÖZGEÇMİŞ

Adı-Soyadı	Muhammed Burak		GÖRENTAŞ
Doğum Yeri ve Yılı			
Bildiği Yabancı	İngilizce		
Eğitim Durumu	Başlama- Bitirme		Kurum Adı
Lise	1992	1995	İzmit 19 Mayıs Lisesi
Lisans	1995	2000	Van Yüzüncü Yıl Üniversitesi
Lisans	2009	2018	Anadolu Üniversitesi
Lisans	2018	2022	Bursa Uludağ Üniversitesi
Yüksek Lisans	2013	2017	Van Yüzüncü Yıl Üniveristesi
Doktora	2018	2023	Bursa Uludağ Üniveristesi
Çalıştığı Kurum	Başlama - Ayrılma		Çalışılan Kurumun Adı
1.	2002	2004	Bayramlar Dersanesi
2.	2004	2012	Pianalitik Dersanesi
3.	2012		Van Yüzüncü Yıl Üniversitesi
Üye Olduğu Bilimsel ve Meslekî Kuruluşlar			
Katıldığı Proje ve Toplantılar			

Yayınlar:	Görentaş, M. B., Uçkan, T., Ayata, F., & Tekpınar, A., (2023). An Overview of the Relationship Between Artificial Intelligence and Law in Forensic Medicine. Euroasia Journal of Mathematics, Engineering, Natural & Medical Sciences International Indexed and Refereed , vol.10, no.28, 61-76. Görentaş, M. B., Uçkan, T., Bayram Arlı, N. (2023). <i>Uyuşmazlık Mahkemesi Kararlarının Makine Öğrenmesi Yöntemleri ile Sınıflandırılması</i>. Yüzüncü Yıl Üniversitesi Fen Bilimleri Enstitüsü Dergisi, [Yayıma Kabul Edilmiş Özgün Araştırma Makalesi]. DOI: 10.53433/yyufbed.1292275 Görentaş, M. B., & Uçkan, T., (2023). Ethics Of Artificial Intelligence In Law And Statistics . 5. ULUSLARARASI MARMARA BİLİMSEL ARAŞTIRMALAR VE İNOVASYON KONGRESİ (pp.199-210). İstanbul, Turkey Görentaş, M. B., Uçkan, T., Ayata, F., & Tekpınar, A., (2023). AN OVERVIEW OF THE RELATIONSHIP BETWEEN ARTIFICIAL INTELLIGENCE AND LAW IN FORENSIC MEDICINE. 7. ULUSLARARASI BİLİMSEL ARAŞTIRMALAR VE GÜNCEL GELİŞMELER KONGRESİ (pp.1234-1235). Ankara, Turkey Bayram Arlı, N., Barca, O., & Görentaş, M. B. (2022). Random Forest. Supervised Machine Learning Algorithms R and Python Applications (pp.119-149), Ankara: Nobel Yayınevi. Görentaş, M. B., & Saraçlı, S., (2019). The Evaluation of the Effect of the Earthquake on Socio-Economic DevelopmentLevel with the Cluster Analysis. International Conference on Data Science, Machine Learning and Statistics 2019 (DMS-2019) (pp.79). Van, Turkey Görentaş, M. B., Ayata, F., & Çelik, H. E., (2018). Eastern Region Socio-Economic Development Cluster Analysis of Provincial Level According to the Classification . ECSAC'18 EUROPEAN CONFERENCE ON SCIENCE, ART & CULTURE (pp.56). Antalya, Turkey Ayata, F., Biçek, E., Uludağ, F., & Görentaş, M. B., (2018). E-Learning Expectations, Readiness and Satisfaction Levels for Pedagogical Proficiency Students . ECSAC'18 European Conference on Science, Art & Culture (pp.70). Antalya, Turkey
Diğer:	
İletişim (e-posta):	
	Tarih: İmza: Adı-Soyadı: Muhammed Burak GÖRENTAŞ