



T.C.
AKSARAY ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
YÖNETİM BİLİŞİM SİSTEMLERİ ANABİLİM DALI

**KİTLESEL ÇEVİRİMİÇİ DERS PLATFORMLARINDA KURSLARA
YAPILAN YORUMLARIN METİN MADENCİLİĞİ KULLANILARAK
DUYGU ANALİZİ**

YÜKSEK LİSANS TEZİ

RAMAZAN DAŞGIN

DANIŞMAN
DOÇ. DR. KEMAL ADEM

AKSARAY 2023

T.C.
AKSARAY ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
YÖNETİM BİLİŞİM SİSTEMLERİ ANABİLİM DALI

**KİTLESEL ÇEVİRİMİÇİ DERS PLATFORMLARINDA KURSLARA
YAPILAN YORUMLARIN METİN MADENCİLİĞİ KULLANILARAK
DUYGU ANALİZİ**

YÜKSEK LİSANS TEZİ

RAMAZAN DAŞGIN

DANIŞMAN
DOÇ. DR. KEMAL ADEM

AKSARAY 2023

TELİF HAKKI VE TEZ FOTOKOPİ İZİN FORMU

Bu tezin tüm hakları saklıdır. Kaynak göstermek koşuluyla tezin teslim tarihinden itibaren(...) ay sonra tezden fotokopi çekilebilir.

YAZARIN

Adı : Ramazan

Soyadı : DAŞGIN

Bölümü : Yönetim Bilişim Sistemleri

İmza :

Teslim tarihi :

TEZİN

Türkçe Adı : Kitlel Çevrimiçi Ders Platformlarında Kurslara Yapılan Yorumların Metin Madenciliği Kullanılarak Duygu Analizi

İngilizce Adı : Sentiment Analysis of Comments on Courses on Massive Online Course Platforms Using Text Mining

ETİK İLKELERE UYGUNLUK BEYANI

Tez yazma sürecinde bilimsel ve etik ilkelere uyduğumu, yararlandığım tüm kaynakları kaynak gösterme ilkelerine uygun olarak kaynakçada belirttiğimi ve bu bölümler dışındaki tüm ifadelerin şahsıma ait olduğunu beyan ederim.

Yazar Adı Soyadı: Ramazan DAŞGIN

İmza:

T.C.
AKSARAY ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
JÜRİ ONAY SAYFASI

Ramazan DAŞGIN tarafından hazırlanan “Kitlesele Çevrimiçi Ders Platformlarında Kurslara Yapılan Yorumların Metin Madenciliği Kullanılarak Duygu Analizi” başlıklı tez çalışması aşağıdaki jüri tarafından oy birliği ile Aksaray Üniversitesi Yönetim Bilişim Sistemleri Anabilim Dalı’nda Yüksek Lisans Tezi olarak kabul edilmiştir.

İmza

Danışman 1: Doç. Dr. Kemal ADEM

Bilgisayar Mühendisliği, Sivas Bilim ve Teknoloji Üniversitesi

.....

Üye: Dr. Öğr. Üyesi Rahmi BAKİ

Yönetim Bilişim Sistemleri, Aksaray Üniversitesi

.....

Üye: Dr. Öğr. Üyesi Serhat KILIÇARSLAN

Yazılım Mühendisliği, Bandırma Onyediy Eylül Üniversitesi

.....

Tez Savunma Tarihi: 19 / 01 / 2023

Sosyal Bilimler Enstitüsü Yönetim Kurulu’nun tarih ve sayılı kararı ile onaylanmıştır.

Sosyal Bilimler Enstitüsü Müdürü

Doç. Dr. Sevilay USLU DİVANOĞLU

İmza

TEŞEKKÜR

Bu çalışmanın gerçekleşmesinde yardımlarını esirgemeyen ve yol gösteren danışman hocam Sayın Doç. Dr. Kemal ADEM'e, değerli fikirleri ile çalışmaya katkı sunan jüri üyelerine, metin madenciliği konusunda yardımlarından dolayı Bilgisayar Mühendisi Hüma ÜNLÜ'ye, her daim yanımda olan başta eşim Aslıhan DAŞGIN ve oğlum Mehmet Eren DAŞGIN olmak üzere destek ve yardımlarını esirgemeyen tüm aileme teşekkür ederim.

AKSARAY ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
KİTLESEL ÇEVİRİMİÇİ DERS PLATFORMLARINDA KURLARA
YAPILAN YORUMLARIN METİN MADENCİLİĞİ
KULLANILARAK DUYGU ANALİZİ

(Yüksek Lisans Tezi)

Ramazan DAŞGIN, Aksaray 2023

ÖZET

Çağımızda kolay ulaşılabilen, az maliyetli, zaman mekan kısıtı olmayan aktiviteler daha çok ilgi görmektedir. Son zamanlarda insanların sahip olduğu olanaklar ve teknolojinin gelişimi de düşünüldüğünde insanların ilgisinde yaşanan bu değişiklik eğitime de yansımıştır. İnsanlar artık istedikleri zaman, istedikleri yerden, kendi seçebilecekleri içeriklere ulaşmak istemektedir. Bu istekler sonunda Kitlesel Açık Çevrimiçi Ders (KAÇD) platformları ortaya çıkmaya başlamıştır. Bu platformlar üzerinde ücretli veya ücretsiz birçok konuda kurslar yer almaktadır. Birçok kişi bu kurslara kayıt olmadan önce, kurslara yapılan yorumlara ve kursa verilen puana göre değerlendirme yapılarak kayıt olmaktadır. Ancak tüm yorumları okuyarak bir kurs hakkında karar vermek kolay olmamaktadır. Bu çalışmada kullanıcıların yorumları okumasına gerek kalmadan kursları olumlu olumsuz olarak değerlendirmesi amacıyla KAÇD platformlarından birisi olan Udemy sitesinde bulunan kurslara yapılan yorumlar kullanılmıştır. Bu yorumlar üzerinden klasik makine öğrenmesi ve derin öğrenme yöntemleri kullanarak kurslar hakkında olumlu olumsuz değerlendirmeleri yapılmıştır. Klasik makine öğrenmelerinden BayesNet, J48 ve OneR algoritmaları ile en başarılı sonuç %91.576 doğruluk ile BayesNet algoritmasından elde edilmiştir. Veri setine Random, GloVe ve Word2Vec kelime gömmeleri uygulandıktan sonra, derin öğrenme modellerinden GRU ve CNN-LSTM hibrit mimarileri uygulanmış ve en başarılı sonuç GloVe kelime gömmesi kullanıldıktan sonra %95.67 doğruluk oranıyla GRU mimarisinden elde edilmiştir.

Bilim Kodu: 92432, 92416

Anahtar Kelimeler: KAÇD, Udemy, Metin Madenciliği, Derin Öğrenme, Makine Öğrenmesi, Duygu Analizi, Word2Vec, Glove

Sayfa Adedi: 67

Danışman: Doç. Dr. Kemal ADEM

AKSARAY UNIVERSITY
GRADUATE SCHOOL OF SOCIAL SCIENCES
SENTIMENT ANALYSIS OF COMMENTS ON COURSES
ON MASSIVE ONLINE COURSE PLATFORMS USING
TEXT MINING

(M.A. Thesis)

Ramazan DAŞGIN, Aksaray 2023

ABSTRACT

In our age, things that are easily accessible, less costly, and have no time and space restrictions attract more attention. Considering the possibilities that people have recently and the development of technology, this change in people's interest has also been reflected in education. People now want to access content that they can choose from wherever they want, whenever they want. As a result of these requests, Massive Open Online Course (MOOC) platforms began to emerge. There are many paid or free courses on these platforms. Before enrolling in these courses, many people register based on the comments made and the score given to the course. However, it is not easy to decide about a course by reading all the reviews. In this study, comments made on courses on Udemy, one of the MOOC platforms, were used in order to evaluate the courses positively and negatively without the need for users to read the comments. On these comments, positive and negative evaluations were made about the courses using classical machine learning and deep learning. With BayesNet, J48 and OneR algorithms from classical machine learning, the most successful result was obtained from BayesNet algorithm with an accuracy of 91.576%. After applying Random, GloVe and Word2Vec word embeddings to the dataset, hybrid architectures of GRU and CNN-LSTM from deep learning models were applied and the most successful result was obtained from GRU architecture with an accuracy of 95.67% after using GloVe word embedding.

Science Code: 92432, 92416

Key Words: MOOC, Udemy, Text Mining, Deep Learning, Machine Learning, Sentiment Analysis, Word2Vec, Glove

Number Pages: 67

Supervisor: Assoc. Prof. Dr. Kemal ADEM

İÇİNDEKİLER

TELİF HAKKI VE TEZ FOTOKOPİ İZİN FORMU	İ
ETİK İLKELERE UYGUNLUK BEYANI.....	İİ
JÜRİ ONAY SAYFASI.....	İİİ
TEŞEKKÜR.....	İV
ÖZET.....	V
ABSTRACT	VI
TABLolar LİSTESİ	İX
ŞEKİLLER LİSTESİ.....	X
SİMGELER VE KISALTMALAR LİSTESİ	Xİ
1. GİRİŞ	1
1.1. ARAŞTIRMANIN AMACI	3
1.2. ARAŞTIRMANIN KAPSAMI	4
1.3. ARAŞTIRMANIN ÖNEMİ	4
2. KURAMSAL TEMELLER	5
2.1. KİTLESEL AÇIK ÇEVİRİMİÇİ DERS (KAÇD)	5
2.1.1. <i>Kitlesel Açık Çevrimiçi Derslerinin (KAÇD) Tarihi</i>	7
2.1.2. <i>Kitlesel Açık Çevrimiçi Derslerinin (KAÇD) Türleri</i>	9
2.2. UDEMY	10
2.3. KAYNAK ARAŞTIRMASI.....	12
3. MATERYAL VE YÖNTEM	17
3.1. VERİ SETİNİN OLUŞTURULMASI	17
3.2. YAPAY ZEKA	17
3.3. METİN MADENCİLİĞİ.....	20
3.4. DUYGU ANALİZİ	22

3.5. TOKENLEŞTİRME.....	23
3.6. KELİME GÖMMELERİ	24
3.6.1. <i>Word2Vec</i>	25
3.6.2. <i>Glove</i>	26
3.7. KLASİK MAKİNE ÖĞRENMESİ YÖNTEMLERİ.....	27
3.7.1. <i>BayesNet</i>	27
3.7.2. <i>OneR (One Rule)</i>	28
3.7.3. <i>J48 (Karar Ağaçları Algoritması)</i>	29
3.8. DERİN ÖĞRENME YÖNTEMLERİ	29
3.8.1. <i>CNN (Convolution Neural Network- Evrişimsel Sinir Ağları)</i> ...	31
3.8.2. <i>LSTM (Long Short-Term Memory- Uzun Kısa Süreli Bellek)</i> ...	32
3.8.3. <i>GRU (Gated Recurrent Unit- Kapılı Tekrarlayan Birim)</i>	33
4. BULGULAR	34
4.1. KLASİK MAKİNE ÖĞRENMELERİYLE ELDE EDİLEN SONUÇLAR.....	36
4.2. DERİN ÖĞRENME YÖNTEMİYLE ELDE EDİLEN SONUÇLAR.....	37
5. SONUÇ VE ÖNERİLER	41
6. KAYNAKÇA	43
EKLER.....	52
AKSARAY ÜNİVERSİTESİ TEZ DEĞERLENDİRME FORMU	53

TABLÖLAR LİSTESİ

Tablo 1: KAÇD Avantaj ve Dezavantajları	6
Tablo 2: gKAÇD ve bKAÇD'lerin özellikleri.....	10
Tablo 3: Sayılarla Udey (Udey, 2022).....	12
Tablo 4: Olumlu Olumsuz Veri Örneği	17
Tablo 5: Klasik Makine Öğrenmesi Algoritmaları Karşılaştırması.....	36
Tablo 6: BayesNet Karşıtlık Matrisi.....	36
Tablo 7: OneR Karşıtlık Matrisi	36
Tablo 8: J48 Karşıtlık Matrisi.....	37
Tablo 9: GRU Modeli Sonuçları.....	37
Tablo 10: Random Kelime Gömmesi Uygulaması Sonrası CNN-LSTM Hibrit Model Sonucu	38
Tablo 11: GloVe Kelime Gömmesi Uygulaması Sonrası CNN-LSTM Hibrit Model Sonucu	38
Tablo 12: Word2Vec Kelime Gömmesi Uygulaması Sonrası CNN-LSTM Hibrit Modeli	39
Tablo 13: Derin Öğrenme Modellerinin Sonuçları	40

ŞEKİLLER LİSTESİ

Şekil 1: KAÇD’lerin tarihi (Hill, 2012).....	7
Şekil 2: KAÇD’lerin büyümesi grafiği (Shah, 2021).....	8
Şekil 3: Konuya Göre Ders Dağılımı (Shah, 2021).....	8
Şekil 4: Udeemy Marka Tanıma Verisi (Google Trends,2022)	10
Şekil 5: Kursların büyümesi (Shah, 2021).....	11
Şekil 6: Kategoriye Göre Kurs Yüzdeleri (Shah, 2021).....	11
Şekil 7: Yapay Zeka Tarihi Gelişimi (Berber, 2019)	18
Şekil 8: Yapay Zekâ, Makine Öğrenmesi ve Derin Öğrenmenin Kronolojisi (Kayaalp ve Süzen, 2018)	19
Şekil 9: Metin madenciliği Süreci (Yaman, 2022).....	20
Şekil 10: Metin Ön işleme Süreçleri (Gürbüz, 2022).....	21
Şekil 11: Duygu Analizi Yöntemleri (Yılmaz ve Orman, 2021).....	23
Şekil 12: Vektörler Arası Gösterim (Demirtaş, 2018)	24
Şekil 13: Word2Vec Mimarisi (Keskin, 2018)	25
Şekil 14: CBOW Mimarisi (Mikolov, Chen, Corrado ve Dean, 2013).	26
Şekil 15: Skip-Gram Mimarisi (Mikolov, Chen, Corrado ve Dean, 2013).....	26
Şekil 16: J48 Algoritması Ağaç Modeli (Çoban, 2011)	29
Şekil 17: Makine Öğrenmesi ve Derin Öğrenmenin Katmansal Karşılaştırılması (Aslan,2020).....	30
Şekil 18: CNN Katman Yapısı (Keskin, 2022).....	31
Şekil 19: LSTM Modeli Kapı Yapısı (Yıldız, 2020).....	32
Şekil 20: GRU Mimarisi (Canlı, 2022).....	33
Şekil 21: Çalışmada Uygulanan İşlem Basamakları	35

SİMGELER VE KISALTMALAR LİSTESİ

CNN	Convolutional Neural Network (Evrişimsel Sinir Ağları)
RNN	Reccurent Neural Network (Tekrarlayan Sinir Ağları)
GRU	Gated Recurrent Unit (Kapılı Tekrarlayan Birim)
LSTM	Long Short- Term Memory (Uzun- Kısa Süreli Hafıza)
KAÇD	Kitlesel Açık Çevrimiçi Dersler
MİT	Massachusetts Institute of Technology
bKAÇD	Bağlantıcı KAÇD
gKAÇD	Geleneksel KAÇD
CBOW	Sürekli Kelime Çantası
Skip-gram	Atlama Gramı
OneR	One Rule
GPU	Grafik İşleme Ünitesi

1. GİRİŞ

Son yıllarda yazılım teknolojilerindeki gelişmeler ve bu gelişmelere bağlı olarak internete olan erişimin kolaylaşması gibi sebeplerle eğitimin verilme şeklinin de değiştiği görülmektedir. Teknolojideki gelişmelerle beraber bilgisayar ve internete erişim konusunda yaşanan sıkıntıların günden güne azalmasıyla öğretimin mekanı ve zamanlaması konusunda teknolojin esneklik katması kaliteli bir eğitime erişebileceğimizin göstergesidir (Chan, Rajamohan, Gan ve Samsudin, 2021).

Özellikle son birkaç yıldır tüm dünyaya hakim olan COVID19 sonucu doğan pandemi sürecinde tüm bu değişimin kaçınılmaz olduğu ve çevrimiçi öğrenmenin ne kadar önemli olduğu açık bir şekilde görülmüştür (Telli ve Altun, 2021). 2019 yılında Çin’de ortaya çıkan, ilk zamanlar korkutucu ve çokta yıkıcı olarak görülmese de tüm dünyaya yayılmasıyla ülkeleri ekonomik ve kalkınma açısından etkileyen COVID19 için birçok alanda tedbirler alınması zorunlu hale gelmiştir (Battineni, Chintalapudi ve Amente, 2020). Karantina, sosyal mesafe, kişisel hijyen, izolasyon, evden çalışma gibi birçok tedbirin sonucunda tüm dünyada okullar ve eğitim kurumları da kapatılmak zorunda kalmıştır (Kansal, Gautam, Çintalapudi, Jain ve Battineni, 2021). UNESCO (Birleşmiş Milletler Eğitim Bilim ve Kültür Örgütü)’nun raporuna göre 22 ülkede toplam 290 milyon öğrencinin eğitim konusunda ciddi etkilendiği görülmektedir (Chawla, Mittal, Chawla ve Goyal, 2020). Tüm bu olaylar sonucunda geleneksel eğitimin değiştirilmesi veya yeniden değerlendirilmesi gerektiği görülmüş ve eğitim uzmanlarının ve hükümetlerin dijital eğitime odaklanmaya ve çevrimiçi eğitime yönelmeye başladıkları görülmüştür (Kansal vd., 2021). Yüz yüze eğitimin kesildiği bu dönemde çevrimiçi yöntemlere geçiş kaçınılmaz hale gelmiş ve herkes öğretim ihtiyacı için bir şekilde dijital yöntemlere başvurmaya gerek duymuştur. Özellikle de bu ihtiyaç üniversite öğrencileri ve kendini geliştirme çabası gösteren kişilerde daha fazla arayışa sebep olmuştur.

Öğretimde oluşan bu boşluk sonucunda daha önceden de kullanılan ilk olarak 1700’lü yıllarda Amerika’da mektup ile başlayıp teknoloji geliştikçe radyo-televizyon ve sonunda bilgisayar ve çoklu ortamlarla yaygınlaşan uzaktan eğitim tekrar gündeme gelmiştir (Şahin ve Durdu, 2021). Bu sürecin uzantısı olan ve geçmişi eski olmasa da kullanımı gün geçtikçe artan, ilk olarak üniversitelerin kendi öğrencilerine verdikleri derslerin yayınlanması şeklinde ortaya çıkan kitlesel açık çevrimiçi derslerin (KAÇD) bir eğitsel müdahale olarak yer almaya başladığı görülmüştür (Impey ve Formanek, 2021).

Genel olarak hedef kitlesi ücretsiz ders almak isteyen, bitirme sertifikası için az miktarda ücret ödeyen veya aldığı derslerin üniversitede geçerli ders kredisi yerine geçmesini istemeyen öğrenciler olan KAÇD'ler (Impey ve Formanek, 2021) bu dönemde fazlasıyla tercih edilir hale gelmiştir (Shah, 2021). 2008 yılında popüler olmaya başlayan ve dünya çapında yayılmaya devam eden KAÇD'ler, farklı alanlarda öğrenmeyi destekleyen, bilgilerin farklı yollarla ölçeklenebildiği, ücretsiz çevrimiçi kurslar olarak tanımlanmak ile birlikte birçok farklı tanımları da mevcuttur (İşgör Şimşek, 2015). Her ne kadar başlangıç olarak dünyaca ünlü üniversitelerin (MIT, Harvard, Stanford) gibi derslerini herkese açık ve ücretsiz sunduğu çevrimiçi bir ortam olarak ortaya çıksa da son zamanlarda farklılaşmaya başlamıştır (Demirci, 2014).

Günümüzde KAÇD'ler uzmanlık aranmaksızın herkes tarafından ders verilebilir olması, çok çeşitli içerikler barındırması sebebiyle birçok platformda karşımıza çıkmaya başlamıştır (Güneş, 2020). Bireylere zaman ve mekan bağımsız, her zaman her yerde öğrenme imkanı sağladığından sadece okullardaki öğrencilere değil tüm toplumun öğrenme sürecinde büyük rol almaya başlamıştır (Güneş, 2020). Öğrenenler çeşitli okul veya kurs gibi herhangi bir kuruma bağlı kalmadan öğrenmenin temelinde bulunan video, ödev, sınav gibi etkinliklere blog, forum gibi tartışma alanlarını da kullanarak çeşitli çevrimiçi süreçlerle öğrenimlerine devam edebilmişlerdir (KOP (2011)'den akt. Güneş, 2020). KAÇD'lerin kurslara gidecek zaman veya imkanı olmayan her yaştan, her kesimden kişiler için istediği kursları kimi zaman ücretsiz, kimi zaman uygun fiyatlı olarak alabilmesi de ayrıca tercih edilmesinde avantaj oluşturmaktadır (Güneş, 2020).

Kayıtlı kişi sayısının gün geçtikçe arttığı, dünyaca ünlü kurumların ücretsiz açık kurslarının yer aldığı Coursera, Edx, MiriadaX, FutureLearn ile beraber birde herkesin kurs yayımlayabildiği ve birçok alanda kurslar barındıran Udemy, Udacity, KhanAcademy gibi daha pek çok KAÇD platformu bulunmaktadır (Onan, 2020). Tüm bu sistemlerde ihtiyacı olan kişiler istedikleri alanda ve ulaşmak istedikleri amaçları doğrultusunda kurslara erişim sağlayarak ücretli-ücretsiz kayıtlarını olarak derslere erişim sağlamaktadırlar.

Herhangi bir kursa kayıt olmadan önce kursun ilgili olduğu alanla ilgili tüm kursları keşfetmek, potansiyelimize ve yeteneklerimize zarar vermemesi için bize uygun olanına karar vermek önemli bir süreçtir (Chan vd., 2021). Online alışverişlerde dahi bir ürün alırken öncelikle bazı niteliklerin incelendiği görülmektedir. Alınacak ürün incelenirken tüketici için en önemli şey o ürünü daha önceden kullanan kişilerin yorumu olduğu gözlenmiştir. Araştırmalar %80'den fazla müşterinin ürünü almadan önce yorumları okuduğunu, bir rapora

göre alışveriş yapanların %94'ü, başka bir rapora göre ise %99.9'unun yorumları okumadan ürünü almadığı belirtilmiştir. İş arama alanında ise iş arayanların %83'ü kararlarını desteklerken incelemeleri kullandıkları görülmüştür (Chan vd., 2021).

Çevrimiçi platformlarda da dersleri alırken eğitmenin niteliği, dersin içeriği gibi birçok açıdan yorumların ve derecelendirmelerin incelenmesi, farklı dersler arasında karşılaştırma yapılması ve bu karşılaştırmanın olabildiğince hızlı yapılabilmesi, özellikle zaman kaybı olmaması açısından büyük önem taşımaktadır (Chan vd., 2021).

Çevrimiçi ortamda kurslara yapılan planlamaların belli bir standardı olsa da yapılan yorumların belirli bir standardı bulunmamaktadır. Bu da yapılan yorumların değerlendirilebilirliğini zorlaştırmaktadır. İşin tam burası da yapılandırılmamış verilerden yararlı bilgiler elde edilerek bir sonuca ulaştırılabileceği işlemi gerçekleştiren metin madenciliği devreye girmektedir. Metin belgelerinin kataloglanması ihtiyacı sonucunda gelişmeye başlayan metin madenciliği, büyük metin verilerinden bilgi keşfederek bu verilerdeki ilişkiyi otomatik belirleme işlemi yapar. Metin madenciliği ile yapılandırılmamış bilgi işlenerek metinden daha anlamlı sayısal veriler çıkarılır ve çeşitli yöntemlerle metnin içerdiği bilgiler erişilebilir hale gelir (Alahyari vd., 2017). Bilgiler erişilebilir hale getirildikten sonra duygusal/ fikrîsel bir sonuca varılması için de metin madenciliğinin bir alt dalı olan duygu analizi devreye girer. Duygu analizinde temel görev, metinde yer alan görüş, cümle veya işletme unsurlarının olumlu, olumsuz, nötr gibi sınıflandırmanın yapılabilmesidir (Işık, 2019).

Bu çalışmada da kursların tercih edilmesinde başkalarının deneyimlerinden yararlanmak için kurslar hakkında olumlu olumsuz gibi sonuç elde edilebilecek, tüm öğrenci yorumlarını gözen geçirerek kursiyere zaman kazandıracak bir sistem üzerinde çalışılması hedeflenmiştir. Tüm yorumları bir kursiyerin inceleyerek karara varması zor olduğundan bu sorunun metin madenciliğinin alt dallarından olan duygu analizi yöntemiyle aşılması üzerinde çalışılmıştır.

1.1. Araştırmanın Amacı

Çalışan profesyonel kişilerden yeni lise mezunlarına kadar özellikle pandemi sonrası, birçok kişi çeşitli sebeplerle çevrimiçi dersleri veya kursları almaya başladı veya alıştı (Chan vd., 2021). Çeşitli kurslar veren siteler üzerinden herhangi bir kurs almadan önce hangi kursun istenildiği gibi bir içeriğe sahip olduğu, hangisinin anlatımının, kapsamının kişinin hedefine uygun olduğunu kursu alan kişinin araştırması gerekir. Çünkü uygun olmayan kursu seçmek, zaman ve güven kaybına sebep olabilir, potansiyel ve yeteneklere zarar verebilir. Genellikle

kurs hakkında yapılan derecelendirmeler, yorumlar kurs hakkında en iyi fikir aracıdır. Ancak tüm yorumları okuyarak kurs hakkında bir karara varmak zaman alıcı iş olmasına karşın son zamanlarda yaygınlaşan metin madenciliği ve derin öğrenme gibi yeni teknolojilerin sonucunda bu işlem kolaylaşmıştır.

İşte bu çalışmadaki temel amacımız; duygu analizi yöntemlerini kullanarak bir KAÇD platformu olan Udemy üzerindeki kurslara yapılan yorumlar ve derecelendirmelerden klasik makine öğrenmesi ve derin öğrenme yöntemlerini kullanarak olumlu, olumsuz gibi sonuçlar çıkarmaktır.

1.2. Araştırmanın Kapsamı

Çalışma kapsamında belirtilen amaçlar dahilinde kullanımı bulunan bir KAÇD web sitesi olan Udemy’de bulunan Kurslara yapılan derecelendirme ve yorumları kullanarak duygu analizi ve metin analizi yöntemleriyle derin öğrenme algoritmalarını da kullanarak bir sonuca ulaştırılması gerçekleştirilmektedir. Python dili kullanılarak yazılan bir program sonucunda Udemy’deki yer alan kurslardaki yorumlar çekilerek alakasız veriler temizlenmiş ve bu şekilde çalışma verilerinin son hali elde edilmiştir.

1.3. Araştırmanın Önemi

Çevrimiçi öğrenme platformlarındaki incelemeler, öğrencilerin farklı kurslar hakkında başkalarının deneyimlerinden bilgi edinerek önemli kararlar almaları yönünden değerli birer araçtır. Ancak bir kurs hakkında onlarca bazen yüzlerce yorumu okuyarak kursun olumlu olumsuz yönlerini belirlemek zaman alabilir. Bu sebeple bir kurs hakkında tüm öğrenci yorumlarını derecelendirmeler de dahil, gözden geçirip o kurs hakkında olumlu- olumsuz gibi bir yargıya hızlı bir şekilde ulaşmamızı sağlayacak bir sistem önermesi ve bu sistemin daha önceden Udemy üzerinden bir örneğinin bulunmaması açılarından çalışmamız önemlidir.

2. KURAMSAL TEMELLER

2.1. Kitlesele Açık Çevrimiçi Ders (KAÇD)

Son yıllarda gerek fiziksel kursları desteklemek gerekse bu kurslara gitme imkanı bulunmayan kişilere eğitim vermek amacıyla kullanılan e-öğrenme metodu olan uzaktan eğitime ek olarak, her yaş ve kesime hitap eden, okul ve kampüste olduğu gibi eğitimde bir sınır çizmeyen kitlesele yapılar ön plana çıkmaya başlamıştır (Güneş, 2020; Saadatmand, 2017). Zamana ve mekandan bağımsız, maddi durumu, konumu ve öğrenim durumu fark etmeden istediği alanda kendini geliştirmeye olanak tanıyan bu kitlesele yapılar kitlesele açık çevrimiçi dersler (KAÇD) olarak adlandırılır. KAÇD'lerin ortaya çıkarak popülerleştiği ilk senelerden bu zamana kadar birçok tanımı yapılmıştır.

Bu tanımlardan birkaçı şu şekildedir:

- İçeriğinin çevrimiçi olarak sunulduğu, bilgi edinmede kullanıcıya işbirlikçi ve sosyal bağlantılı olanaklar sunan aktif bir modeldir (İşgör Şimşek, 2015; Thompson, 2011).
- Dünyanın neresinde olursanız olun ücretsiz kaydolup ders alınan, içerikleri açık sunulduğu ve son zamanlarda ön plana çıkan uzaktan eğitim ortamıdır (Esposite (2012)'den akt. İşgör Şimşek, 2015).
- İsteyen herkesin internet üzerinden ulaşarak kayıt olabildiği, genel olarak üniversiteler tarafından sunulan derslerdir (Demirci (2013)'ten akt. Güneş, 2020).
- Yaşı, cinsiyeti, coğrafyası ne olursa olsun katılan herkese uzaktan eğitimle öğrenim içeriği sağlayan son yenilikçi yaklaşımdır (Onan, 2020).
- Çok sayıda öğrencinin ye aldığı gruplar için tasarlanan, internetin olduğu her yerden, herkes tarafından ulaşılabilen ve bir koşul olmaksızın giriş yapılabilen çevrimiçi derslerdir (Butaray (2015)'ten akt. Şahin, 2022).

KAÇD kavramın yer alan kelimeler de tanımlarda yer alan ifadelerle örtüşmektedir.

Kavramda yer alan kelimelerin anlamlarını Siemens (2013) çalışmasında:

- Kitlesele: Sistemde büyük sayıda katılımının yer alabileceği,
- Açık: Sistemde yer alan içeriklerin açık olarak kullanıcıya sunulabildiği,
- Çevrimiçi: Sistemin öğrenme içeriklerinin birçoğunu çevrimiçi olarak öğrencilere sunduğu,

- Ders: Sistemde başlama ve bitiş vaktinin belirli olduğunu, alınan derslerin sonradan kullanıcıların erişim hakkına sahip olduğunu, forum, blog gibi sosyal etkileşim imkanı sunduğunu, ifade etmiştir.

Günümüzdeki birçok teknolojiye de olduğu gibi yeni nesil imkanlara uygun öğrenme ortamları olan KAÇD’lerin çevrimiçi dünyayla iç içe olmasının getirdiği birtakım avantajları ve dezavantajlarının da olması kaçınılmazdır. Avantajları (Akpınar,2014) ve dezavantajları (İşgör Şimşek, 2015) Tablo 1’de yer almaktadır.

Tablo 1: KAÇD Avantaj ve Dezavantajları

KAÇD Avantajları	KAÇD Dezavantajları
Değişik içeriklere ulaşım imkanı olması	Ölçme ve değerlendirme, çoktan seçmeli sorular veya akran değerlendirmesi gibi yöntemlerle yapıldığından güvenilirliği düşük olabilir
Derslerin farklı uzman veya öğreticilerden alınabilmesi	Değerlendirme sürecinde kopya veya kaynak göstermeden alınması denetlenememektedir
Kırsal bölgeler de dahil birçok kitleye ulaşabilmesi	Öğrenenler sosyal ortamdan uzak olduğundan birçok öğrenci dersleri tamamlamamaktadır
Kişiye özel öğrenme imkanı ile en uygun gelişimi elde etmeye yardımcı olabilmesi	İçerikler çoklu ortam derslerinin tasarım ilkelerine uymadan düzenlenebilmektedir
Öğrenen kişi için eğitimin, dolayısıyla uzaktan eğitimin, ekonomik hale gelmesi	Derslerde tamamlanma oranları oldukça düşüktür
Yer alan içerikler ve etkinlikler ile uzman, alanın deneyimi, otoritesi ve dünya çapındaki imkanlara ulaşımın gerçekleşmesi	Eğitim konusunda belli sınırların ve hedeflerin olduğu bir süreçte kurs hakkında herhangi bir muhatapın olmaması dolayısıyla da öğrenen kişinin sorularını kurs esnasında sorarak anında cevap alabileceği bir etkileşimin bulunmaması

Tüm bu fırsatların yanında kısıtlı birçok yanının da olduğu görülen KAÇD’lerin bu konuda ciddi çalışmaları bulunan Hollands ve Tirthali (2014)’e göre 6 adet amacı bulunmaktadır. Bu amaçlar şu şekildedir:

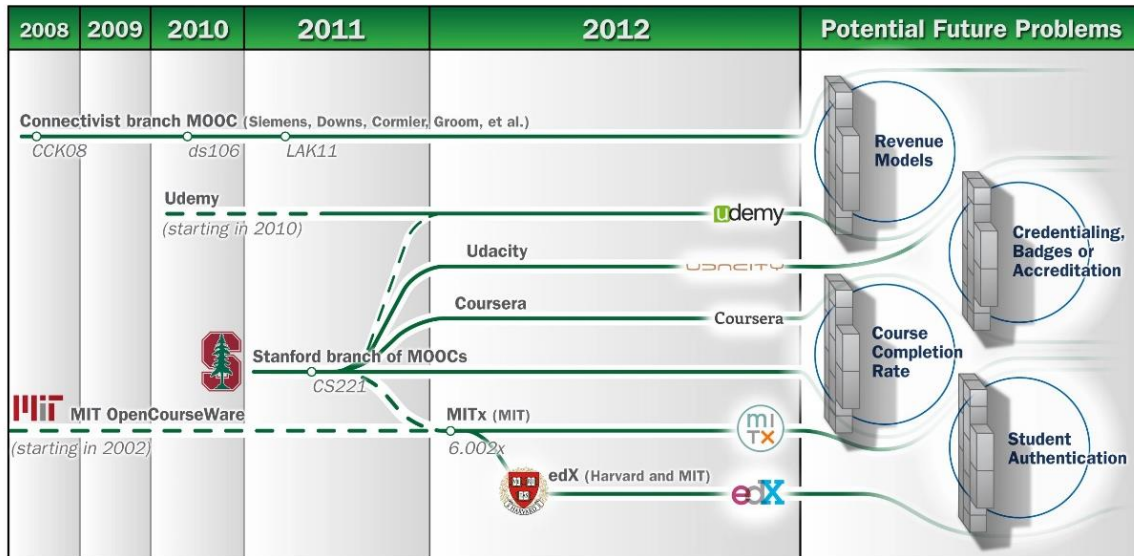
- Birçok kişiye ulaşarak kullanımı kolay hale getirmek,
- Dünyada öğrenen herkesin ilgisini çekerek marka haline gelmek,
- Birçok konuda oluşan giderleri olabildiğince aza indirip, geliri artış seviyesine getirerek ekonomiye katkı sağlamak,

- Öğrenen ve öğreticinin farklı kültürlerden olmasını sağlayarak eğitsel çıktıyı artırmak,
- Ders veren ve dersi alan açısından yenilikler meydana getirmek,
- Öğretme ve öğrenme üzerinde bazı pedagojik olan konularda gerçekleştirilemeyen araştırmaları yapabilmek.

2.1.1. Kitlese Açık Çevrimiçi Derslerinin (KAÇD) Tarihi

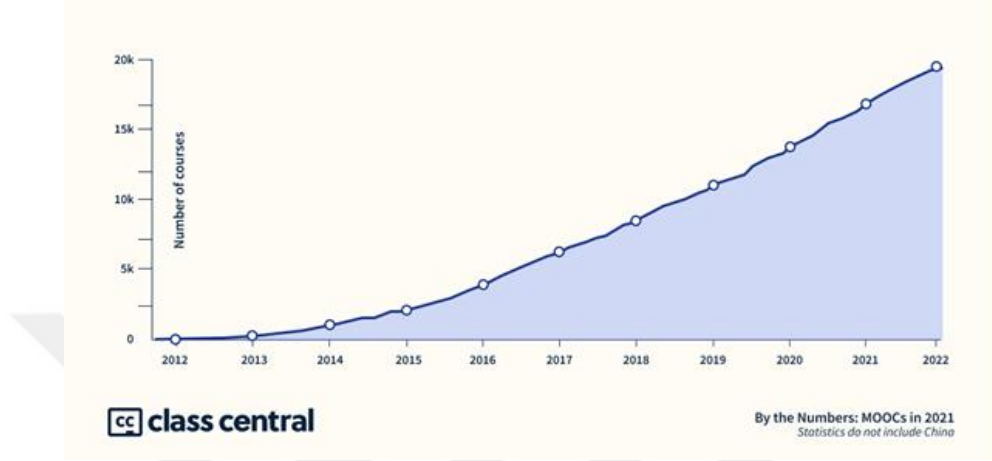
KAÇD'ler aslında, ilk olarak 1999 yılında tanıtılan ve 2002 yılında Massachusetts Institute of Technology (MIT) önderliğinde öğrenmeye hevesi olan ve ders materyallerine, video veya başka eğitimsel kaynaklara ihtiyaç hisseden tüm insanlarla internet aracılığıyla paylaşımlar gerçekleştiren "Açık Ders Kaynakları Girişimi"nin devamı olarak ortaya çıkmışlardır (Ata, 2018).

KAÇD terimi ise George Siemens ve Stephen Downes'in tasarladığı ve ABD'de sundukları on iki haftalık ve 24 öğrencinin kredili alabileceği bir çevrimiçi kurs olan "Connectivism and Connected Knowledge"ı tanımlamak için David Cormier ve Bryan Alexander adlı girişimciler ilk olarak 2008 yılında ortaya atılmıştır (Cormier ve Siemens (2010)'dan akt. Hollands ve Tirthali, 2014). 2008 yılında popülerliği artmaya başlamış ve dünyada çoğu üniversitenin internet ortamında yayılmasını sürdürmüştür (Kışla, Kumova Metin, Karaoğlu ve Demir, 2021). KAÇD'lerin tarihi Hill (2012) tarafından şekildeki gibi özetlenmiştir.



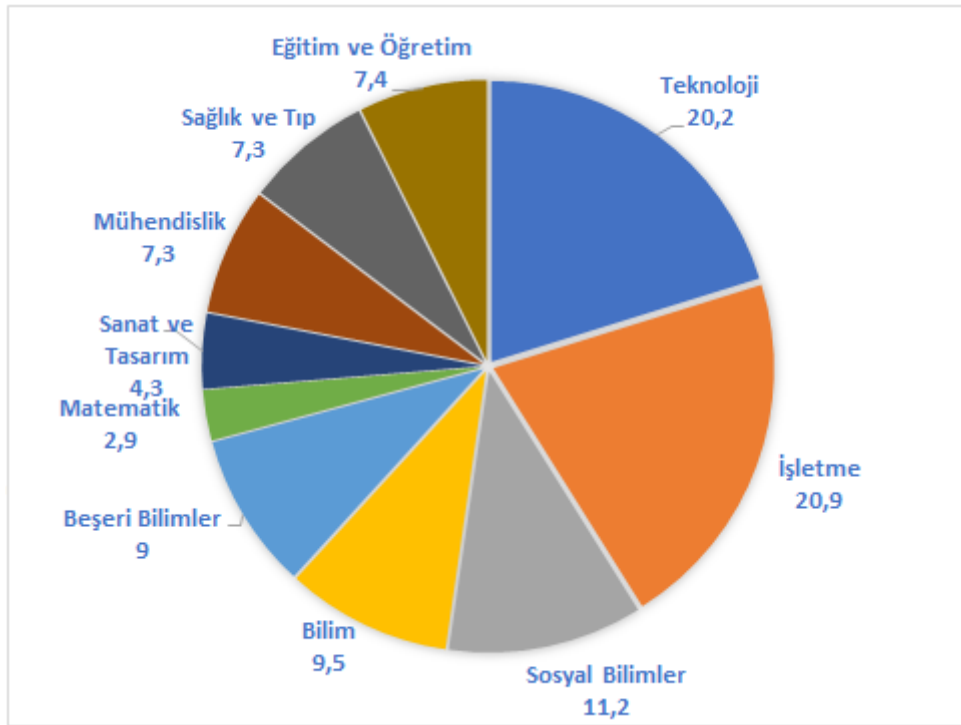
Şekil 1: KAÇD'lerin tarihi (Hill, 2012)

Başlarda MIT ve Harvard Üniversitelerinin iş birliği içinde oluşturdukları bu ortamların, daha sonradan tanınması Edx, Coursera, Udemy, Udacity, Khan Academy gibi birçok kursu barındıran KAÇD platformları ile gerçekleşmiştir (Kışla vd., 2021). KAÇD’lerin yıllar içindeki büyümesi Şekil 2’de yer almaktadır.



Şekil 2: KAÇD’lerin büyümesi grafiği (Shah, 2021)

Başlangıçta sadece üniversitede verilen derslerin yer aldığı bu platformlarda artık birçok konuda kurslar yer almaktadır (Şekil 3)



Şekil 3: Konuya Göre Ders Dağılımı (Shah, 2021)

2.1.2. Kitlese Açık Çevrimiçi Derslerinin (KAÇD) Türleri

KAÇD teriminin ilk kullanılmaya başlandığı zamanlardaki KAÇD'lere olan eğitim yaklaşımıyla KAÇD'lerin kullanılrlığının artmaya başlandığı yıllardaki eğitim yaklaşımları farklılaşmıştır. İşte bu eğitim yaklaşımlarını zamanla birbirinden ayırabilmek için bağlantıcı KAÇD (bKAÇD) ve geleneksel KAÇD (gKAÇD) terimlerinin türediğı görölmektedir (Selmanovic, Sayar ve Durdu, 2021).

bKAÇD: Downes ve arkadaşının bir üniversitenin sertifika programı için çevrimiçi olarak verecekleri derste, zorluklara rağmen, bağlantıcılık ilkelerini kullanmaya karar vermesiyle temelleri atılan bKAÇD'ler birinci nesil KAÇD olarak bilinmektedir (Hollands ve Tirthali, 2014). Bağlantıcılık ilkesinden devamla bKAÇD'lerde öğrenme blog, forum, Twitter, öğrenme toplulukları gibi ortamlar aracılığıyla gerçekleştirilirken içerik öğrenenler tarafından oluşturulur, dolayısıyla herkes kendi öğrenmesinin sorumluluğunu alır (İşgör ve Şimşek, 2015). Yine de eğitmenin ilk soruyu sorması veya haftalık sorular sorarak, çeşitli kaynaklar sunma hakkı vardır. Ancak gerçekte eğitmen katılımcıların faaliyetlerini gözden geçirip toparlayarak kolaylaştırıcı şekilde hareket eder. Kursun çıktıları ise katılımcıların oluşturduğu gönderiler, diyagramlar, resimler veya video gibi benzersiz ürünlerdir.

gKAÇD: 2011 yılında Andrew Ng ve arkadaşları tarafından Stanford Üniversitesinde halka açık olarak sunulan üç kursun eğitmen tarafından yönetilen gKAÇD'lerin başlangıcı olduğu düşünülmektedir(Selmanovic vd., 2021). Eğitimler, genelde 10 dakika veya daha kısa olan videolarla verilirken, ek okumalar, ödevler gibi ekstra etkinliklerle desteklenebilir. Öğrenme içeriklerine kayıt olarak erişim sağlanabilen gKAÇD'lerde geleneksel öğrenme modeli benimsenmiştir. Uzmanların hazırladığı içerikler, öğrenenin istenen zaman ve istediğı mekandan erişebileceğı web sitesi gibi bir alanda bulunurlar (Güneş, 2020). Coursera, Udemy, Udacity, edX, FutureLearn gibi halihazırda sıklıkla kullanılan platformlar gKAÇD sınıfındandır.

gKAÇD ve bKAÇD'lerin özellikleri Tablo2'de özet olarak yer almaktadır (Hollands ve Tirthali, 2014).

Tablo 2: gKAÇD ve bKAÇD'lerin özellikleri

Geleneksel KAÇD (gKAÇD)	Bağlantıcı KAÇD (bKAÇD)
• Önceden belirlenen, eğitmenin liderliğiyle yapılandırılan ve sıralanmış etkinlikler • Kısa ve içeriği veren video ve okumalar • Akran değerlendirmesi ve kısa sınavlarla verilen notlar • Tartışma forumlarına isteğe bağlı olarak katılma • Bir web sitesi üzerinden sağlayıcı aracılığıyla (Coursera, Udemy, Edx gibi)	• Öğrenci merkezli, sosyal ve teknik olarak öğrenme • “Atölye” gibi ir ortamda konuların oluşturulup/ araştırılması • Öğrencilerin oluşturduğu blog gönderisi, resim, video gibi ürünler • Tartışma formu, gruplar, sosyal ağlar önemli • Eğitimci, etkinlikleri gözden geçirme, özetleme, aktiviteye çevirme rolünde

2.2. Udemy

2010 yılında Eren Bali ve Oktay Çağlar tarafından Silikon Vadisinde kurulan tarayıcı tabanlı bir KAÇD sistemi (Bilge Şahin, 2022) olan Udemy dünyanın en önemli ve en fazla çevrimiçi kursa sahip olan platformudur. Platform tarayıcı üzerinden destek vermekle kalmayıp teknolojik alt yapılara uyum sağlayarak Android ve IOS işletim sistemlerine uygun mobil uygulamalar üzerinden de çalışmaktadır.

Udemy diğer KAÇD’lerden farklı olarak üniversiteler tarafından desteklenen kursları barındırmaz, verdikleri sertifikalar tanınmaz, kişinin bilgisini artırmasında yardımcı kar amaçlı bir sistemdir (Ramneet, Gupta ve Madhukar, 2021). Kurslarının neredeyse tamamı ücretli moddadır. Ücretli kursları Google’da en çok arananlarda olan Udemy, marka olarak tanınma açısından büyük artış göstermiştir (Şekil 4).

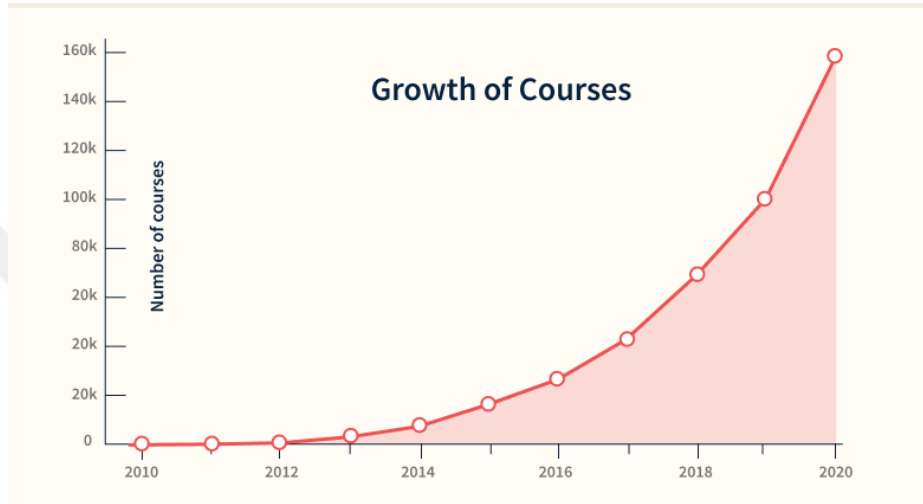


Şekil 4: Udemy Marka Tanıma Verisi (Google Trends,2022)

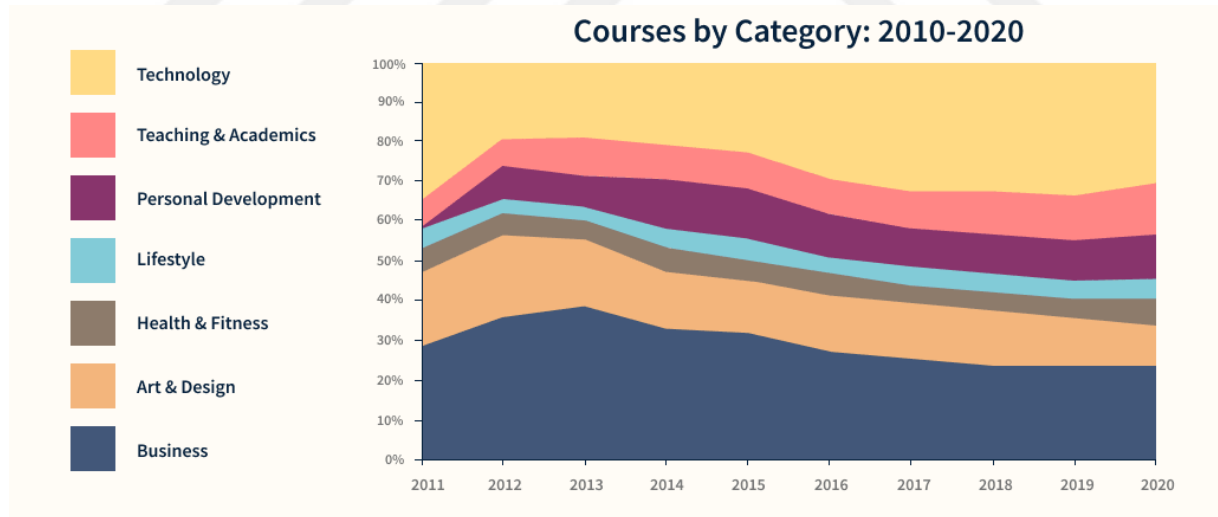
Bu platformda eğitimciler kurslarını ücretli veya ücretsiz olarak sunabilirler. Eğitimci olmak için kayıt olmak, belirli şartlara göre videoları hazırlayıp sisteme yüklemek yeterlidir.

Hazırlanan bu videolar denetlenerek onaylandıktan sonra kullanıcılara sunulur (Kışla vd., 2021).

Udemy’de çok çeşitli kategorilerde kurslar yer almaktadır ve Udemy ortaya çıktığından bu yana kurs sayılarında ciddi artışlar olup (Şekil 5), kurs kategorilerinin sistemde yer aldıkları yüzdelerinde de değişiklikler gözlenmektedir (Şekil 6).



Şekil 5: Kursların büyümesi (Shah, 2021)



Şekil 6: Kategoriye Göre Kurs Yüzdeleri (Shah, 2021)

Kursları bitiren kişilere bitirme sertifikasının sunulduğu Udemy’deki kursların güvenilirliği ve kalitesini kullanıcılar belirler (Kışla vd., 2021). Temmuz 2022 itibariyle Udemy’e ait öğrenen, eğitmen, kurs, kurs kaydı, dil ve kurumsal müşteri sayıları Tablo 3’te yer almaktadır.

Tablo 3: Sayılarla Udemy (Udemy, 2022)

Udemy	
Öğrenen	54 milyon
Eğitmen	71 bin
Kurs	204 bin
Kurs Kaydı	741 milyon
Dil	75
Kurumsal Müşteri	12500

Birçok dilde, binlerce kursu barındırmasında rağmen Udemy sisteminin hem olumlu hem de olumsuz yanları bulunmaktadır.

Olumsuz yönleri (İşgör Şimşek,2015; Şahin ve Durdu, 2021):

- Derslerin birçoğunun ücretli olması
- Kullanıcılara göre sistem arayüzünün karmaşık ve ticari olması
- Sistemin mobil ortamda dersleri filtrelemek için dil seçeneğinin ve şifre değiştirme seçeneğinin bulunmaması
- Kurslara ait videolarda alt yazı ve konuşma metni seçeneklerinin olmaması
- Tasarımda gerekli bazı fonksiyonların görünür olmamaları
- Eksik arayüz tasarımı

Olumlu yönleri (İşgör Şimşek,2015; Udemy,2022):

- Sistem arayüzünün Türkçe olması
- Kullanıcıların özellikle mobil ortamda ihtiyaç duyacağı düğmelerin dikkat çekici olması
- Kurs programında, kursun ders sayısı, süresi bilgilerinin yer alması
- Mobil cihazda kullanımında kursların cihaza indirilebilmesi
- Kurslar hakkında yapılan yorumların, tartışma ve duyuruların erişilebilir olması
- Farklı dillerde kurs alma imkanı
- Kurs veren olarak dünyanın her yerinde bulunan öğrenciye erişim imkanı
- Kursların farklı cihazlarda kullanılabilir olması

2.3. Kaynak Araştırması

Son zamanlarda ki teknolojik gelişmelerin neticesinde eğitimin şeklinin de değişmesi kaçınılmaz olmuştur. Yüz yüze eğitime ek olarak çevrimiçi eğitim de internete ulaşımın kolaylaşmasıyla tercih edilir hale gelmiştir. Bu tercihler Kitlese Açık Çevrim İçi Derslerin

meydana çıkarak, bu platformların çoğalmasını sağlamıştır. Çeşitli KAÇD platformlarında yer alan kurslarla ilgili hem metin madenciliği hem de bu platformların kullanılabilirliğiyle alakalı birçok çalışma yapılmıştır.

Öğrenci performansı ile ilgili tahmin modeli oluşturmak amacıyla veri madenciliği uygulaması kullanılan çalışmada Aydın (2007), önce sistemdeki veri kaynağını toparlayarak veri tabanı oluşturup, modelleme safhasında veri madenciliği uygulama adımlarını uygulayarak öğrenci bilgi sistemindeki ve e-öğrenme sistemindeki verileri hazırlamıştır. Veri madenciliği konusunda SPSS Clementine yazılımının kullanıldığı çalışmada, tahmin modellemesinde C5.0, Logistic Regresyon, Neural Net, CveRR, CHAID, QUEST algoritmaları kullanılmıştır. Modellere uygulanan geçerlilik testi sonunda en iyi tahmin modelinin %82.1'lik doğrulukla C5.0 algoritmasıyla ortaya çıkan karar ağacı modeli olduğu görülmüştür. 2000-2001 öğretim yılındaki anket verileri ve mezun öğrencilerinin kimliklerinin verilerinden oluşan veri kaynağının kullanıldığı kümeleme analizinde K-means algoritmasının kullanılmasıyla çalışmanın ikinci aşaması meydana gelmiştir. Mezun öğrencilerin, medeni hal, bilgisayar ve internet kullanımı, cinsiyet, mezun oldukları yaş ve mezuniyet gecikmesi özelliklerine göre gruplandırılarak yapılan kümeleme çalışması sonucunda, en kısa sürede mezun olan öğrencilerin bilgisayar ve internet olan aynı zamanda bekar olan erkek öğrenciler olurken, mezun olması en fazla gecikenlerin yaşı büyük olan, bilgisayar ve internet kullanımı sınırlı olan öğrenciler olduğu görülmüştür (Aydın, 2007).

İşgör Şimşek (2015), son yıllarda eğitimde sıklıkla kullanılan KAÇD'lerden olan Coursera, Udemy ve Udacity platformlarını ele alarak cinsiyet, yaş, teknolojik okur-yazarlık ve daha önceden çevrimiçi ders alma kategorilerine göre bu platformların mobil ortamdaki kullanılabilirliklerini değerlendirmek amacıyla nitel bir çalışma yapmıştır. Verileri toplamak için "Bilgisayar Sistemi Kullanılabilirlik Anket"nden yararlanarak hazırlanan 19 soruluk 7'li likert tipi ölçekle, 20-43 yaş aralığındaki 56 kişiye uygulamışlardır. Verilerin SPSS programı yardımıyla işlendiği anket, sistemin yararlılığı, bilgi kalitesi, arayüz kalitesi bölümlerinden oluşmakta olup, veriler değerlendirilirken Mauchly'nin varyans analizi ve tekrarlı ölçümler için ANOVA, düzeltme analizi olarak da Bonferroni kullanılmıştır. Kullanıcılara öncelikle mobil ortamda üç KAÇD için belirli görevler verildikten sonra bu anketler uygulanmış ve kullanıcıların demografik özellikleriyle bu anket arasında olan ilişki incelenmiştir. Yapılan araştırma sonucunda kullanılabilirlik açısından Coursera ve Udemy arasında fark olmayıp, Udacity ile aralarında anlamlı bir fark olduğu görülürken, kullanılabilirlikle demografik özelliklerden yaş faktörünün ilişkili olduğu görülmüştür. Aynı şekilde sistem yararlılığı,

arayüz kalitesi konusunda da düzeylerin Coursera ve Udemy için aynıyken, Udacity'den daha iyi kalite ve yararlılıkta olduğu görülmüştür (İşgör Şimşek, 2015).

Fakülte öğretim performansının değerlendirilmesinde öğrencilerin geri bildirimlerini kullanarak bakış açısına göre duygu analizinde derin öğrenme yaklaşımı kullanan ilk girişim olan çalışmalarında Sindhu vd. (2019), var olan metinden otomatik yön çıkarmak ve bunların karşılığı olan yönlendirmeyi bulan bir sistem önermişlerdir. Bu sistemde çift katmanlı LSTM modeliyle, denetimli bakış açısı temelli fikir madenciliği kullanılmıştır. Sukkur IBA Üniversitesi'nin son beş yıldaki öğrenci yorumlarını elle etiketleyip veri seti olarak kullandıkları çalışmada olumlu, olumsuz, nötr gibi yönelimler belirlenmiştir. Elde edilen verileri öncelikle metin ön işleme safhalarından geçirdikleri çalışmada paragrafları cümleye bölmede OpenNLP, cümleyi bölmek için nokta işareti, “ve” ve “ama” gibi bağlaçları kullanmışlardır. Veriler işlendikten sonra, alan uzmanlarının da yardımıyla 5000'den fazla yorumun olumluluk ve olumsuzluğuna göre yönleri belirlenmiş ve etiketlenmiştir. Sistemde gömme katmanının sayesinde görünüm çıkarmada %91, duygu polaritesini algılamada %93 gibi yüksek doğruluk değeri elde etmişlerdir (Sindhu vd., 2019).

Onan (2020), topluluk öğrenme ve derin öğrenmeyi kullanarak eğitim konusunda tahmin performansı yüksek olan, verimli çalışan bir duygu sınıflandırması sunmayı amaçladığı çalışmasında, 66000 KAÇD incelemesinin yer aldığı derlemi analiz etmiştir. Geleneksel denetimli, topluluk ve derin öğrenme yöntemlerinin tahmin performanslarını değerlendirdiği çalışmada aynı zamanda KAÇD değerlendirmeleri için duygu analizi konusunda metin temsil ve kelime yerleştirme şemaları da değerlendirilmiştir. Bu değerlendirmelerde makine, derin ve topluluk öğrenmesi yöntemleri kullanılmıştır. Çalışmada konvolüsyonel sinir ağı, tekrarlayan sinir ağı da dahil olmak üzere beş adet derin öğrenme mimarisi kullanılmış ve eğitimsel veri madenciliği konusunda duygu analizinde, topluluk ve derin öğrenme metotlarında daha başarılı olduğu görülmüştür. Çalışma sonunda %95.80 doğrulukla sınıflandırma performansı elde edilen GloVe sözcük yerleştirme Yöntemi ile kullanılan uzun kısa süreli bellek ağlarından yüksek performans elde edildiği görülmüştür (Onan,2020).

Chakravartky, Kameswari, Mydeen ve Seenivasan (2020), Coursera'daki makine öğrenmesi dersinden alınan yorumları veri seti olarak kullandıkları ve bu verilere fikir madenciliği uyguladıkları çalışmalarında, doğal dil işleme yöntemiyle görüş çıkarma, yapılandırılmamış metni işleme, kutup sınıflandırması adımlarını kullanmışlardır. Fikir madenciliği sonucunda elde edilen performans, KNN, Lojistik Regresyon gibi makine öğrenimi algoritmalarını kullanarak değerlendirmişlerdir. KNN algoritması ile özgüllük ve duyarlılık konularında daha

iyi sonuçlar elde edilirken, Lojistik Regresyon ile %94 gibi KNN'den daha fazla bir doğruluk değeri elde edildiği görülmektedir (Chakravartky, Kameswari, Mydeen ve Seenivasan, 2020)

Chan, Rajamohan, Gan ve Samsudin (2021), Coursera platformunu kullanarak yaptıkları çalışmalarında, kurs arayan kişilerin, istedikleri kurslarla ilgili fikir edinmeleri konusunda yardımcı olmak amacıyla yaptıkları çalışmalarında, metin temizleme, lemmatizasyon, metin madenciliği, duygu analizi ve görselleştirmeyi kapsayan bir metin analizini boru hattı önermişlerdir. Diğer önemli amacın birçok kursu bir şekilde karşılaştırarak kursa giriş yapılmasının kolaylaştırılmasını sağlamak olan çalışmada Python Jupyter Notebook'u kullanmışlardır. N-gram analizi ve kelime bulutu çalışmalarının kurslar hakkında olumlu olumsuz fikirler edinilmesi ve kurslar arası karşılaştırma yapılabilmesi açısından başarılı olan çalışmada olumsuz yorumların tespitinde eksik kaldığı görülmüştür. N-gram analizi sonucu doğrulukta 52/60, kullanılabilirlikte 55/60 puan aldıkları gözlemlenirken, model olumlu duyguları ayırt etmede %97'lik, olumsuz duyguları ayırt etmeden %46'lık bir ayırt etme oranı yakalamıştır (Chan vd., 2021).

Dina, Yunardi, Firdaus ve Airlangga (2021), çevrimiçi platformda kullanıcıların yaptıkları incelemeler ve KAÇD içeriğiyle ilgili gelecekte olabilecek tasarım eğilimlerini belirlemek amaçlı yaptıkları çalışmalarında verileri Coursera'dan elde etmişlerdir. Araştırmada amaçlarına ulaşmak için metin madenciliği kullandıktan sonra kullanıcı tercihlerinin belirlenmesi için özellik-duygu çiftleri (anahtar kelime çiftleri) çıkarılmıştır. Duygu analizi konusunda Destek Vektör Makinesinin kullanıldığı ve doğruluk, kesinlik, hatırlama ve F1 skoru değerlerinin %80'in üzerinde elde edildiği bu çalışmada, olumlu ve olumsuz özellik-duygu çiftlerinde kullanıcı tercihlerinin Coursera'da başarılı bir şekilde tanımlandığı anlaşılmaktadır (Dina ve ark., 2021).

Coursera platformundan toplanarak manuel olarak etiketledikleri öğrenci yorumlarından oluşan büyük bir veri kümesi kullanan Edalati, İmran, Kastrati ve Deudpota (2021), öğrencilerden gelen geri bildirimlerin analizi öğretmen ve öğrenme çıktılarının incelenmesi nedenleriyle fikir madenciliği kullanarak makine öğrenmesi modellerinin değerlendirilmesi amacıyla bu çalışmayı gerçekleştirmişlerdir. Öğrencilerin öğretimle ilgili tutumlarını, görüşlerini tahmin etmede derin öğrenme yöntemleri ile birlikte makine öğrenmesi algoritmalarından Rastgele Orman, Destek Vektör Makinesi ve Karar Ağacı algoritmalarını kullanmışlardır. Sonuçta Rastgele Orman algoritması ile en boy oranı tanımlamada %98.01, duyarlılık sınıflandırma görevlerinde %99.43 F1 skoru elde ederek başarılı sonuçlar elde etmişlerdir (Edalati, İmran, Kastrati ve Deudpota, 2021)

Sistemlerin etkili olarak kullanılmasının oldukça önemli olduğu çevrimiçi dünyada popüler olan KAÇD web sitelerinden Udemy, Coursera ve Udacity üzerinde çalışma yapan Şahin ve Durdu (2021), bu sitelerin kullanılabilirliklerini uzman temelli değerlendirme yönteminden bilişsel gezinti yöntemiyle değerlendirerek tespit edebildikleri problemlere ve sistemlerinin tasarımlarının geliştirilmesi ile kullanılabilirliklerinin artmasına yönelik öneri sunmuşlardır. Araştırmada sistem tasarımı, erişimi, eğitim kaynakları, kurs, profil ve yardım kategorileri altında fiiller gerçekleştirilirken sistemin oluşturduğu yanıtlara göre kullanım problemleri belirlendiği görülmektedir. Bu problemlerin, arayüzde bulunan gerekli özelliklerinin görünürlüğünün yetersizliği, tasarımdaki eksiklikler, bazı eksik özellikler olduğu görülen araştırmalarda en az kullanılabilirlik problemi düzeyinden en fazlaya sırasıyla Coursera, Udemy ve Udacity şeklinde olduğu görülmüştür (Şahin ve Durdu, 2021).

Kullanıcı deneyimlerinin değerlendirilmesi konusunda geleneksel yöntemlere ek olarak, Beyin Bilgisayar Arayüzü teknolojisi olan elektroansefalografi (EEG) yöntemi kullanmayı amaçladığı çalışmada Şahin (2022), Udemy platformunun kullanıcı deneyimi (KD) değerlendirmesi için 13 katılımcılı bir kullanıcı çalışması değerlendirmesi yapmıştır. Katılımcılar platformla etkileşime girmeden önce ve girdikten sonra KD değerlendirmesinde kullanılmak üzere veriler Etki Grid Anketi, KD anketinin kısa versiyonu (UEQ-S), Pozitif Etki ve Negatif Etki ölçeği (PANAS), Web Sitesi Analizi ve ölçüm envanteri (WAMMI) ile elde edilmiştir. Daha sonrasında EEG kullanımı için gerekli veriler ise Emotiv Epoc cihazı ile toplanmış ve Udemy platformunun kullanıcı deneyiminin olumlu olduğunu gösteren sonuçlar alınmıştır. Elde edilen tüm veriler sonuçta Geleneksel ve EEG yöntemlerindeki olarak karşılaştırıldığında birbirini destekler nitelikte olduğu görülmüştür (Şahin, 2022).

Literatürde yer alan KAÇD ile ilgili çalışmalardan da anlaşıldığı üzere KAÇD platformlarındaki kurslara yapılan yorumları kullanarak derin öğrenme yöntemleriyle olumlu olumsuz sınıflandırması yapan çalışmanın çok az olduğu ve özellikle Udemy platformuyla ilgili bu konuda hiç çalışma yapılmadığı görülmüştür. Bu çalışmada literatürdeki bu boşluğu doldurarak Udemy sitesindeki kurslara yapılan yorumların, klasik makine öğrenmesi alanında BayesNet, J48 ve OneR algoritmalarıyla, Random, GloVe ve Word2Vec kelime gömmelerinden sonra, derin öğrenme modellerinden, CNN, GRU ve LSTM mimarilerini kullanarak olumlu- olumsuz şeklinde sınıflandırması gerçekleştirilmiştir.

3. MATERİYAL VE YÖNTEM

3.1. Veri Setinin Oluşturulması

Bu çalışmada Udemy sitesinde yer alan kurslara yapılan yorumlar kullanılmıştır. Verilerin Udemy üzerinden alınması Udemy API kullanılarak yapılmıştır. Udemy’de yer alan kurslara verilen puanlara göre yorumlar olumlu olumsuz şeklinde sınıflandırılmıştır. Kurslara yapılan yorumlardan elde edilen verilerde, 1 ve 2.5 puan arası olumsuz, 4 ile 5 puan arası yorumlar ise olumlu kabul edilmiştir. 3 ve 3.5 puan olan yorumlar nötr olduğu için bu çalışmada dahil edilmemiştir. Verilerden anlamsız olabilecek 1 ile 3 harf arası olan kelimeler çıkarılmıştır. Anlamsız kelimelerin çıkarılması, yabancı dilde yazılmış yorumların ayıklanması gibi düzeltmelerden sonra 100000 olumlu ve 100000 olumsuz olmak üzere 200000 yorumdan oluşan veri seti elde edilmiştir.

Tablo 4’te veri setinde yer alan olumlu ve olumsuz örnekler gösterilmiştir.

Tablo 4: Olumlu Olumsuz Veri Örneği

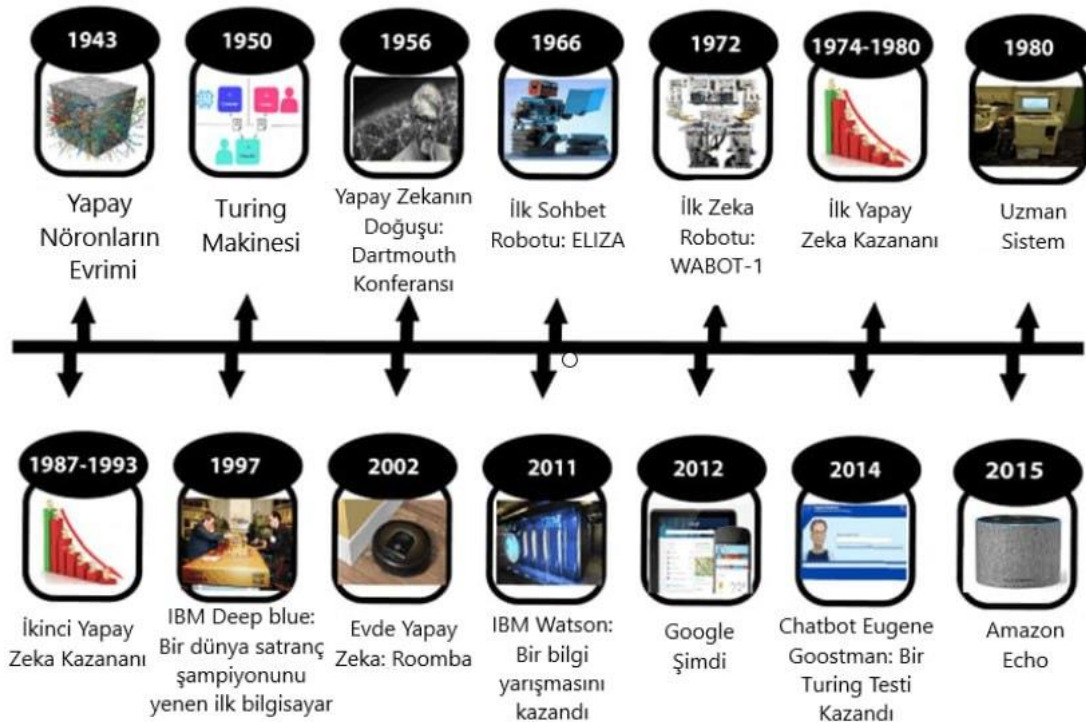
Olumlu (1)	şuanlık konular sıkmadan izleyebiliyorum güzel bi anlatım var
	şuana kadar satın aldığım en iyi eğitim oldu bu. tavsiye ederim
Olumsuz (0)	şöyle böyle. Şimdiye kadar öğretilen şey çok basit. çok temel
	Eğitmenin talimatları bu modülde güncel veya uygulanabilir değildir

3.2. Yapay Zeka

Teknolojik gelişmeler genel olarak bir ihtiyaç sonucunda ve herhangi bir canlıya ait olan bir özellikten etkilenecek ortaya çıkmıştır. Zeka kavramının da insana özgü bir olgu olduğu düşünüldüğünde son yıllarda birçok alanda ön plana çıkarak hayatımızı kolaylaştıran yapay zeka kavramı için de aynı şeyi söyleyebiliriz. Zekayı düşünmek, hareket etmek, karar vermek, algılamak ve olaylardan sonuç çıkaran yetenek olarak düşünersek, bu özelliklerin makineler tarafından bir insanmış gibi yapılabilir olması da yapay zeka olarak düşünülebilir (Mahmat, 2022). Birçok alanda kullanılan yapay zekanın aslında, uygulandığı alanın genişliği ve alana özgü tanımlamaların olması nedeniyle tam olarak ortak bir tanımla bulunmamaktadır (Akdemir, 2022). Fakat farklı boyutlarını ele alacak şekilde birden fazla tanımla da vardır. Bu tanımlarda da genel olarak insanca düşünmek ve davranmak, mantık çerçevesinde düşünmek ve davranmak olarak bahsedilip bu kavramların açıklaması yapılmıştır. Yapılan tanımlar incelendiğinde bir işi uygulayan bir insanmış gibi makinelerin yapabilir düzeyde olması genel ortak özellikler olduğu görülmektedir (Russel ve Norvig

(2016)'den akt., Akdemir, 2022).

Yapay zekanın temellerinin Eski Yunan Mitolojisi'ne dayandığı düşünülse de günümüz anlamında Alan Turing'in kendi adıyla anılan Turing testini oluşturarak yaptığı çalışmalara dayanmaktadır. 1950 yılında iki insan ve bir bilgisayar kullanılarak, sorulan sorulara verilen cevapların insana mı bilgisayara mı ait olduğunun ayırt edilmesi şeklinde uygulanan Turing testi sonucunda bilgisayarın büyük verileri işleyerek, konuşmayı yorumlayıp iletişim kurabileceği görülmüştür (Sahota ve Ashly (2019)'den akt., Akdemir, 2022). Bu testin uygulamasının ardından 1956 yılında yapay zekanın kurucu babalarından biri olarak kabul edilen John McCarthy, Dartmouth Kolej'de gerçekleştirilen, Otomata Teorisi, sinir ağları ve zeka çalışması konularının yer aldığı projede "Yapay Zeka" kavramını ilk kez kullanmıştır ve diğer araştırmacılar tarafından da bu uygun görülmüştür (Laplante (2017)'den akt., Akdemir, 2022).

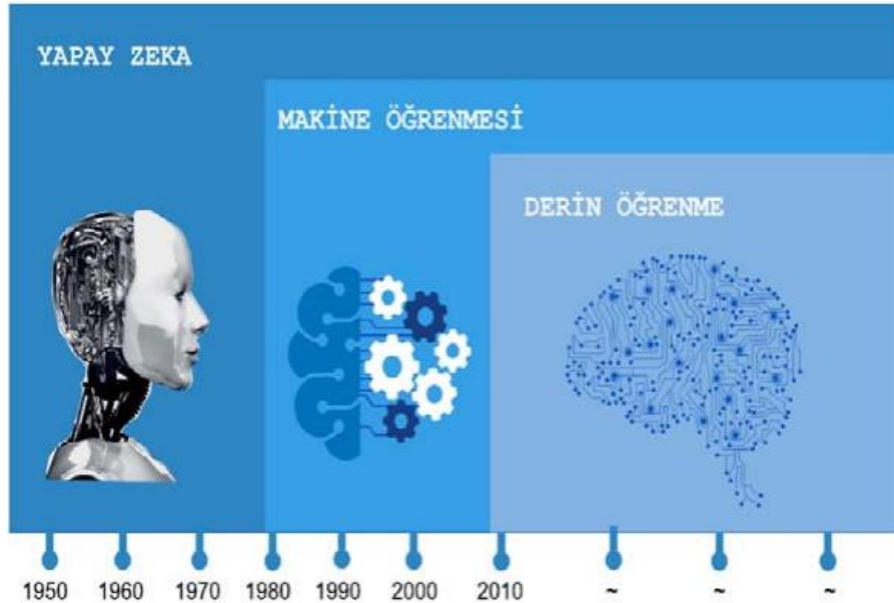


Şekil 7: Yapay Zeka Tarihi Gelişimi (Berber, 2019)

Yapay Zekanın tarihinin olduğu Şekil 7'de görüldüğü üzere, 1956 yılından sonra 1970 yılına kadar çalışmalar artarken, 1979-1980 yılları arası çeşitli sebeplerle azalan çalışmalar sonucu "Yapay Zeka Kışı" yaşanmıştır. 1980'de tekrar bilgisayar yazılımları ve uzman sistemlerin gelişmesiyle hız kazanan çalışmalar sonunda 1987-1993 yıllarında maliyet, cazibe kaybı,

uygulama alanının sınırlılığı gibi sebeplerle tekrar "Yapay Zeka Kışı" ortaya çıkmıştır. Bu yıllardan sonra teknolojik gelişmelerin de hız kazanmasıyla makine öğrenmesi algoritmalarının da gelişmesiyle tekrar ilgi çekici hale gelerek bir daha kış dönemine girilmemiştir.

1990'lı 2000'li yıllarda birçok Yapay Zeka uygulaması kullanıma sunulmuştur. 1997 yılında IBM'in Deep Blue bilgisayarı, konuşma tanıma yazılımının Windows işletim sistemine uygulanması, 2011'de Watson yazılımı, 2016 yılında Google tarafından geliştirilen AlphaGo yazılımı, makinelerin yapay zeka sayesinde başarılı olduğunu gösterirken, geliştirilen yazılımların tek bir amaç için yapıldığı ve zayıf yapıda olduğu görülmüştür (Akdemir, 2022). Yapay zekanın işte burada zayıf ve güçlü yapay zeka olarak ikiye ayrıldığı görülmektedir. Yalnızca konuşma tanıma, satranç oynama, İnternette arama yapma gibi sadece bir alana özgü görevleri yapabildiği için zayıf yapay zeka kavramı kullanılırken, insana ait işlerin neredeyse hepsini yapabilen ve hatalardan öğrenen güçlü yapay zeka kavramı kullanılmıştır (Lu vd. (2017)'den akt., Akdemir, 2022). Yapay zeka konusunda meydana gelen bu ayrım büyük verilerin artışıyla zamanla yerini makine öğrenmesine, makine öğrenmesi de yerini derin öğrenmeye bırakmıştır (Mahmat, 2022). Şekil 8'de görüldüğü gibi Yapay zeka makine öğrenmesini, makine öğrenmesi de derin öğrenmeyi kapsamaktadır.



Şekil 8: Yapay Zekâ, Makine Öğrenmesi ve Derin Öğrenmenin Kronolojisi (Kayaalp ve Süzen, 2018)

3.3. Metin Madenciliği

Son zamanlarda gelişen teknoloji ve artan veri yoğunluğu göz önüne alındığında bu verileri anlamlı hale getirmek büyük önem kazanmıştır. Ancak bazı metin verilerinin sonuç alınabilir şekilde olmasına karşın sonuç çıkarması zaman alan ve yapılandırılmamış halde bulunduğu göz önüne alınarak bu verilerin de işlenmesi için bir yöntem gerekmektedir. İşte bu yapılandırılmamış metin verilerini işleyip anlamlı bilgiler çıkarılmasını çeşitli veri madenciliği algoritmaları kullanarak gerçekleştirilmesi amacıyla metin madenciliği kullanılmaktadır (Işık, 2019). Kaynağı yapılandırılmamış, büyük verilerden oluşan ve bu verilerden önemli ve yararlı bilgiler elde edilmesi sürecine metin madenciliği denilmektedir (Yaman, 2022). Metin belgelerinin kataloglanması ihtiyacı sonucu ortaya çıkan bu yöntemle farklı yazılı kaynaklar içindeki bilgileri otomatik olarak ayrıştırarak yeni bilgilerin bilgisayar keşfiyle elde edildiği görülmüştür.



Şekil 9: Metin madenciliği Süreci (Yaman, 2022)

Metin madenciliği Şekil 9’da da görüldüğü üzere üzerinde çalıştığımız, dolayısıyla ihtiyacımız olan, yapılandırılmamış ya da yarı yapılandırılmış metinlerden oluşan verinin bir araya getirilmesi olan veri toplama aşaması ile başlamaktadır. Bu aşamada üzerinde çalışılması istenen veriler internet veya bilgisayar gibi ortamlardan elde edilen metinler, çevrimiçi veya çevrimdışı olabilen veri tabanlarından toplanabilen metinlerdir (Gürbüz, 2022). Elde edilen verilerin doğal dille yazılmış olduğu düşünüldüğünde bir sonraki aşama olan verilen ön işleme adımı kaçınılmaz olmaktadır. Veri ön işleme adımı ilginç ve nemli olan bilgilerin çıkarılması işlemini gerçekleştirmek amaç olduğundan, ortaya çıkacak sonucun da başarısında büyük yere sahiptir. Bu aşamada Şekil 10’da görüldüğü gibi noktalama işaretleri ve sayıları çıkarma, metni küçük harfe çevirme, tokenleştirme, metinde yer alan durak kelimeleri çıkarma, kök çözümleme, kelimelerin TF-IDF değerlerini bulma ve metin sınıflandırma aşamaları bulunmaktadır.



Şekil 10: Metin Ön işleme Süreçleri (Gürbüz, 2022)

Tüm bu aşamalardan geçirilerek bilgisayarın anlayabileceği hale getirilmiş olan metin verileri üzerinde uygun olan teknik kullanılarak analiz aşamasında hazır edilmektedir. Yapılandırılmamış metindeki değerli bilgilere ulaşmak için elde etme, yapılandırılmış olan veri tabanından bilgi çıkarmak için çıkarım, temel hatları ve genel anlamı bozmadan uzunluk azaltmak için özetleme, belgeleri sınıflara ayırmak için kategorizasyon, metinleri kümeleme, sınıflandırma ve analiz gibi işlemlerini yapmak için kümeleme gibi birçok farklı amaca hizmet eden teknikler bulunmaktadır (Yaman, 2022). Tüm bu sürecin sonunda metin analizi çeşitli yöntemlerle gerçekleştirilerek istenilen amaç doğrultusunda bir sonuca ulaşılmaktadır.

Yazılı iletişimi ve kaynakları kullanan ortamlarda gerçekleştirilmesi uzun zaman alan ve bazı görevlerin daha sistemli hale getirilmesinde avantaj sağlaması doğrultusunda metin madenciliği farklı alanlarla ilgilenmektedir. Yaman (2022) ve Gürbüz (2022)'ye göre göre bu alanlar ve açıklamaları şu şekildedir:

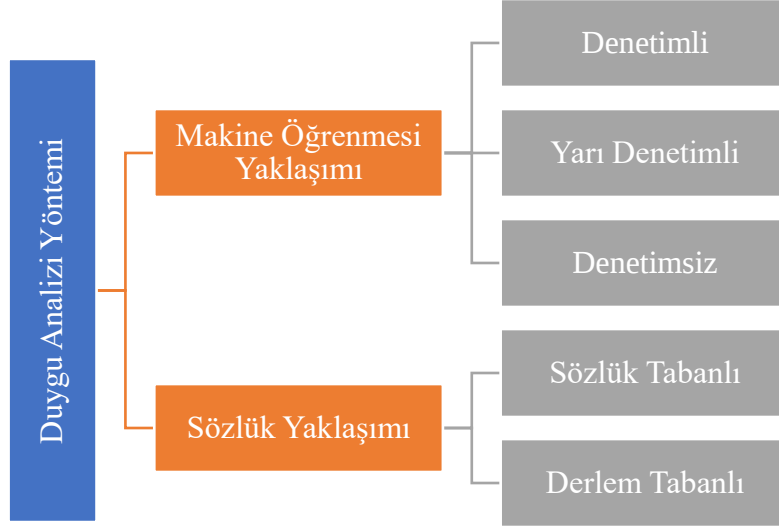
- Bilgi Erişimi (Information Retrieval): Arama motoru, anahtar kelime araması dahil olmak üzere ilgilenilen belge hakkında ön bilgilerin temin edilmesi ve saklanması
- Doğal Dil İşleme (Natural Language Process): Doğal dili çözümleyerek yorumlayan bir sistem tasarlanması ve uygulanması
- Adlandırılmış Varlıkların Tanınması (Named Entity Recognition): Metnin işlenmesi aşamasında istatistiki bazı özelliklerin çıkarılması

- Örüntüsü tanımlı varlıkların bulunması (Pattern Identified Entities): Metnin içerisinde yer alan özel bazı bilgilerin metin madenciliğine konu olarak kullanılması
- Eş Atıf (Coreference): Bir varlığa atıf yapan kelimelerin ve diğer başka terimlerin bulunması ve ayrılması
- İlişki, Kural, Olay Çıkarımları: Birçok farklı amaçla metnin içerisinde yer alan bazı bilgilerin çıkarılması
- Duygu Analizi (Sentimental Analysis): Metinde geçen duygusal ifadelerin çıkarılarak yazının olumlu/olumsuz olup olmamasına göre sınıflandırılması

3.4. Duygu Analizi

Duygu durum analizi, fikir madenciliği, duygu sınıflandırma, kanaat çıkarımı gibi birçok farklı isimlerle literatürde yer alan duygu analizi metin madenciliğinin önemli bir alanı olup, doğal dil işleme, veri madenciliği ve metin madenciliğini de bir araya getirdiği bilinmektedir (Işık, 2019). Amacı farklı kaynaklardaki kişiye ait bilgilerin tespiti, konu veya bir belge üzerinde duygusal eğilimleri belirlemek (Karakuş, 2022) olan duygu analizi, herhangi bir metin veya içerikte yer alan duygu ve düşünceleri gerekli bilgisayar algoritmaları kullanarak analiz eder ve sonuçta bu duyguları olumlu, olumsuz, nötr şeklinde bir sınıflandırmaya tabi tutar (Yılmaz ve Orman, 2021). Burada kullanılan duygu, düşünce veya görüşler kişilere ait olabilirken, kuruluşlara da ait olabilmektedir. Son zamanlarda artan sosyal medya hizmetleri, bloglar, forumlar, wikiler sayesinde duygu analizi konusunda çalışılan alanlar ve konular oldukça fazlalaşmıştır (Karakuş, 2021). Bu ortamlardaki veriler kullanarak tıp, eğitim, politika, medya, pazarlama, satış gibi birçok alanda duygu analizi sonuçları kullanılmaktadır (Aslan, 2020). Duygu analizi aynı zamanda duygu sınıflandırma, öznellik sınıflandırma, görüş özetleme, görüş alma, alay ve ironi, tür veya yazar tespiti, belirtilen kişilerin tespit edilmesi gibi birçok konuda da kullanılmaktadır (Karakuş, 2021).

Duygu analizi yapılırken yeteri kadar veri barındıran veri kümesine ihtiyaç vardır ve bu veri kümesi olumlu, olumsuz, nötr duygu şeklinde etiketlenmiş olması gerekmektedir. Daha sonra bu veriler gerekli yöntemler ve algoritmaları kullanarak ön işleme sürecine tabi tutulur. Bu ön işlem sonucunda veriler bilgisayarın daha çok anlayabileceği hali almış olurlar ve bu işlem sonunda sınıflandırma aşaması uygulanır. Yapılan sınıflandırma sonunda eğitilen verilerle test verileri arasında bir karşılaştırma yapılarak sistemin başarısı tespit edilir (Yılmaz ve Orman, 2021). Bu duygu analizi aşamalarında makine öğrenmesi ve sözlük tabanlı olmak üzere iki yaklaşım kullanılır (Şekil 11).



Şekil 11: Duygu Analizi Yöntemleri (Yılmaz ve Orman, 2021)

Şekil 11’de de görüldüğü üzere makine öğrenmesi tabanlı yaklaşım denetimli, yarı denetimli ve denetimsiz olarak üçe ayrılmaktadır. Denetimli yöntemlerden en yaygın kullanılanları Destek Vektör Makinesi, Naive Bayes, Maximum Entropy, OneR gibi tekniklerdir. Diğer iki yöntemde ise eğer yapacağımız çalışmada herhangi bir sınıflandırma yapmak için etiketli belge veya görüş kümesi elde etmek mümkün değilse kullanılmaktadır (Karakuş, 2021). Bizim yaptığımız çalışmada denetimli makine öğrenmesi yaklaşımıyla çalışılmıştır.

3.5. Tokenleştirme

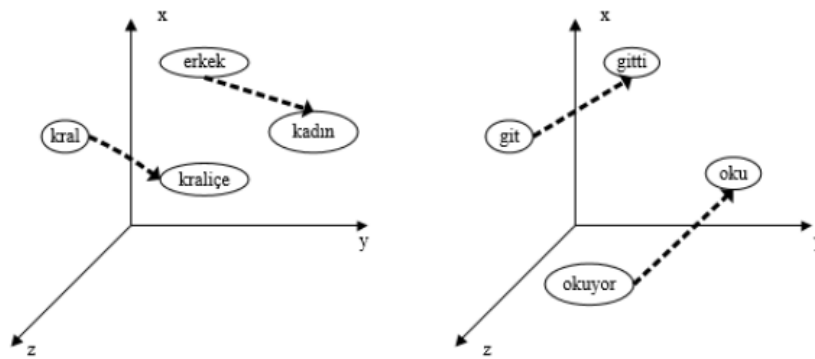
Çalışılan bir konuda elde edilen verileri öncesinde işlemek, çalışılan makine için, giriş boyutu azalacağından, işleri daha basit hale getirir (Barjenzi, 2021). Ele alınan metin verilerinin makinenin anlayacağı ve işleyeceği en küçük birim olan kelimelere ayırmadan önce bazı işlemlerden geçmesi gerekir. Metin ön işleme denilen bu aşamada, öncelikle metin cümlelere ayrılır. Cümlelere ayrılan metin daha kolay işlenebilmesi açısından küçük harflere çevrilerek var olan noktalama işaretleri ve sayılar cümlelerden çıkarılır (Gürbüz, 2022). Tüm bu aşamaların en önemlisi ve ana tekniği olan tokenizasyon işlemi gerçekleştirilerek ön işleme tamamlanmış olur.

Literatürde belirteçleme, tokenizasyon, tokenization olarak yer alan tokenleştirme, her bir veri kaydında bulunan metin dizilerini daha küçük birim olan kelimelere bölen sürecin adıdır ve işlenmemiş olan diziyi anlamlı belirteçlere böler (Dina ve ark., 2021 ve Gürbüz,2022). RNN, GRU, LSTM gibi popüler derin öğrenme mimarilerinin ham metni belirteçler düzeyinde işlemesinden dolayı önemli bir adımdır (Pai, 2022).

Tokenleştirme işlemi kelime, alt kelime ve karakter düzeyinde olmak üzere üçe ayrılır. Bunlardan kelime tokenizasyonu en yaygın kullanılanıdır. Metin parçalarını belirli bir sınırlayıcı kapsamında ayrı ayrı sözcüklere böler ve sınırlayıcıya bağlı kalarak farklı kelime belirteçleri oluşturur. Önceden eğitilmiş olan kelime gömmeleri Word2Vec ve GloVe, kelime tokenizasyonu içeriğinde gelir (Pai, 2022). Bu çalışmada Word2Vec ve GloVe kelime gömmesi yöntemlerini kullanacağımızdan kelime tokenizasyonu kullanılmıştır.

3.6. Kelime Gömmeleri

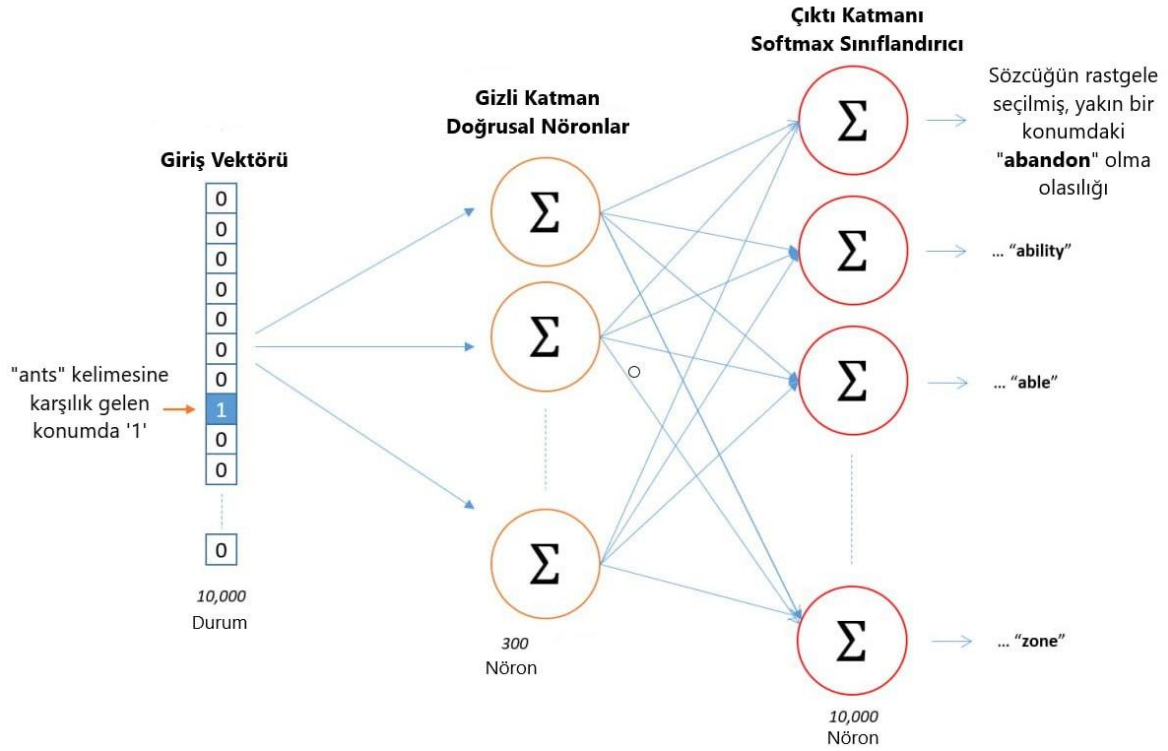
Makine öğrenmesi algoritmaları doğrudan metin verisiyle çalışmazlar. Metinler sayısal değerlere dönüştürülerek her kelime bir vektör ile temsil edilebilir hale getirilmelidir ki bu algoritmalar metne ait vektör değerini kullanabilsin (Omer, 2021). Metin ön işleme aşamalarından sonra bu işlemlerin gerçekleştirilmesi için kelime gömmeleri işin içine girer. Kelime gömmesi, aynı veya benzer anlama gelen birbirinden farklı kelimelerin sayısal vektör haline dönüştürülerek temsil edilmesini sağlayan tekniktir (Özdemir, 2021). Aynı zamanda sözcükler arasında olan anlamsal ve söz dizimsel benzerlikleri yakalar ve bu bağlamsal benzerliği koruyan vektör temsilleri oluşturur. Bu benzerlik vektör uzayında birbirine benzeyen kelimelerin vektörlerinin birbirine yakın noktalara eşlenmesiyle yani Öklid mesafesi ve kosinüs değerlerine göre hesaplanarak tanımlanır (Şekil 12). Bağlamsal olarak, büyük değerlik için yakın benzerlik gösterirken, değer küçük ise uzak benzerlik olarak ifade edilir (Demirtaş, 2018). Rastgele bir değerden başlatılan veriye göre vektör oluşturup kelime gömmesi yapılabilir, veri miktarı yeterli ise iyi sonuç verecektir. Bu çalışmada kelimelerin sayısal vektörlerini elde etmek için Random, Word2Vec ve GloVe yöntemleri kullanılmıştır.



Şekil 12: Vektörler Arası Gösterim (Demirtaş, 2018)

3.6.1. Word2Vec

Metin belgelerindeki kelime yerleřtirmelerinin öğrenilmesi ve kelimelerin vektörlere dönüřtürülmesinde kullanılan denetimsiz ve hesaplama yönünden verimli tahmin yapabilen bu yöntemi 2013 yılında Mikolov ve arkadaşları geliřtirmiřtir (Nergiz, 2021). Girdi, gizli ve çıktı katmanlarını içeren iki katmanlı sinir ağı yapısıyla, metni derin ağlar tarafından anlaşılan sayısal hale getirir (Çevik, 2019). Bu modelde bilginin anlamı da sözcüklerin anlamları da deęiřmez (Abdalrahman, 2020). Word2Vec'in mimari yapısı Şekil 13'te yer almaktadır.

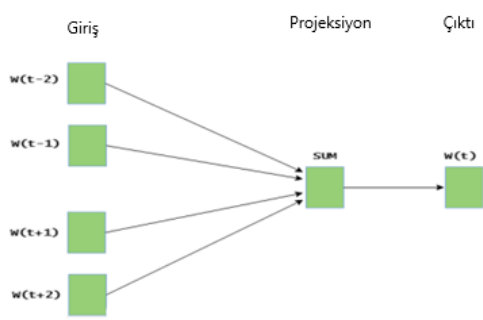


Şekil 13: Word2Vec Mimarisi (Keskin, 2018)

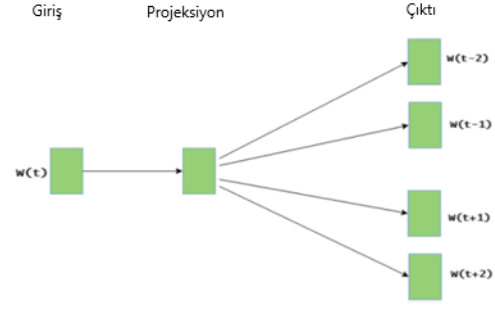
Büyük veriye sahip çalışmalarda verinin boyutunun azaltılması gerektięi için, kelimeler arasındaki kosinüs deęerinin hesaplayıp birbirine benzeyen sözcükler kümelenmekte ve bu yeni oluřan kümeler yeni boyuta sığdıęından verinin boyutunu küçültebilen Word2Vec yöntemi kullanılmaktadır. Sözcüklerin komřu sözcüklere göre eęitildięi yöntemde Skip-gram (atlama gramı) ve CBOW (sürekli kelime çantası) mimarileri kullanılarak eęitim gerçekteřtirilmektedir (Nergiz, 2021).

Word2Vec, hedefteki sözcüğün etrafında kaç sözcük olacaęını belirleyen pencere geniřlięi ve sözcüklerin vektörlerinin boyutunu tanımlayan embedding boyutu adlı parametrelere

göre kelime vektörleri oluşturur. Şekil 14'te mimarisi yer alan CBOW modeli, sözcüğe ait pencere boyutunda kendisinin çevrelendiği sözcüklere göre tahmin gerçekleştirilerek yerleştirmeyi öğrenir (Onan, 2020). Şekil 15'te mimarisi yer alan Skip-gram modelinde ise hedef sözcük kendisini çevreleyen sözcükleri tahmin ederek yerleştirme işlemini uygular (Abdalrahman, 2020).



Şekil 14: CBOW Mimarisi (Mikolov, Chen, Corrado ve Dean, 2013).



Şekil 15: Skip-Gram Mimarisi (Mikolov, Chen, Corrado ve Dean, 2013).

Skip-gram ve CBOW algoritmaları birbirlerinin tersi şeklinde çalışmaktadır. Skip-gram ortada bulunan kelimeyi alıp çevresinde bulunan kelimeleri tahmin etmekte, CBOW ise çevredeki kelimelere göre ortada bulunan kelimeyi tahmin etmektedir. Skip-gramda giriş orta kelime, çıkış çevredeki kelimelerdir. Pencere kavramı ise ortadaki kelimedenden itibaren kaç kelimenin alınacağını belirtmek için kullanılmaktadır. Eğer az sayıda verimiz varsa Skip-gram algoritması daha başarılı olabileceken, çok sayıda verinin varlığında hızlı sonuç için CBOW daha başarılı sonuçlar vermektedir (Koglak, 2020).

3.6.2. Glove

Sözcüklerin vektör temsillerini elde etmede kullanılan, denetimsiz tahmin modeli olan, küresel vektörler olarak da bilinen GloVe modelinde başlangıç noktası sözcüklerin birlikte görülme oranlarıdır. Word2Vec modelindeki yerel bağlam tabanlı öğrenime ek olarak küresel matris çarpanlarına ayırma yöntemi ile birleştirilmiştir (Onan, 2020). Vektör uzay modellerinde yer alan, genel matris vektörlerine dayanan gizli anlamsal analiz ile Word2Vec'de yer alan skip-gram analizi olmak üzere iki metodoloji bulunmaktadır. Bu metodolojilerden gizli anlamsal analizde istatistik verilerinde yüksek performans alınırken analogik görevlerde geri kaldığı, skip-gramda ise analogik görevlerde performans yüksek olurken istatistiksel konularda geri kaldığı görülmektedir (Tozyılmaz, 2019). Bu ilişkiden

yararlanarak Pennington, Socher ve Manning (2014), en az kareler yöntemini önermişlerdir ve modelde sözcüklerin beraber görülmesinin frekansları hesaplanırken aynı zamanda istatistiki yöntemleri de kullanabilen GloVe metodu ortaya çıkmıştır. Ele alınan eşit dağılım sonucu oluşan gürültülü verinin oluşumu probleminin en az kareler yöntemiyle önüne geçilmiştir.

Elimizdeki verilerin tek tek üzerinden geçerek, herhangi bir kelimenin verilerde kaç kez geçtiği hesaplanmakta ve optimizasyon buna göre gerçekleştirilmektedir. Word2Vec'te verilerin kontrolü sırasında kelime her karşısına çıktığında vektörler optimize edilirken, GloVe'de optimizasyon işlemi en sonda yapılmakta, bundan dolayı da eğitim daha hızlı gerçekleşmektedir (Koglak, 2020).

GloVe yönteminde eğitimde 1 milyon sözcük külliyyatından en fazla 42 milyon sözcük külliyyatına kadar farklı büyüklüklerde sözcük külliyyatı kullanılmıştır. Bu metin külliyyatlarında el alınan sözcük-sözcük beraber bulunma istatistiklerine göre eğitim gerçekleştirilmiş ve bu eğitime göre sözcük vektör uzaylarının lineer hali elde edilmiştir. Elde edilen sonuçlar analogik nedenselleme yöntemiyle test edilerek en iyi sonucun 300 boyutlu sözcük vektörlerinde alındığı görülmüştür ve Word2Vec yöntemine göre daha yüksek temsil skoru elde edilmiştir (Pennington ve ark., 2014)

3.7. Klasik Makine Öğrenmesi Yöntemleri

3.7.1. BayesNet

BayesNet, iki farklı duruma göre olasılık hesabına dayalı bir teorem olan Bayes teoremi mantığıyla çalışan, sınıflandırma problemlerine özel tasarlanmış bir bayes ağ çeşididir (Eraldemir, 2014). Kullanımının kolay olması ve sonuca hızlı ulaşılması açısından makine öğrenmesi ile yapılan durağan kod analizinde ilk seçilen algoritma olmuştur (Akgül, 2019).

Tahmin değişkenleri veya özellikleri aracılığıyla tanımlanmış olan örnek verilere etiket veya bir kategori atamayı amaçlayan BayesNet, denetimli sınıflandırma algoritmaları arasında yer almaktadır (Gürtürk, 2022). Çok terimi olan verilerin analizi sırasında istatistiksel sınıflandırma tekniğini kullanarak, buna göre olasılıkları hesaplar (Altınsoy, 2019).

BayesNet algoritması ağın yapısını öğrendikten sonra olasılık tablolarını öğrenme mantığıyla çalışmaktadır. Ağ yapısının iyi olduğunu belirlediğinde koşula bağlı olan bu olasılık tablosunu her değişkene özel tahmin etmektedir (Belen, 2018).

Birbirinden bağımsız olan durumların aynı zamanda oluşması ve iki durumun da ayrı ayrı oluşması durumlarına göre sonucun hesaplandığı bu algoritmanın formülü Eşitlik 1’de yer almaktadır (Eraldemir, 2014).

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)} \quad (1)$$

Eşitlik 1’de yer alan formülde yer alan ifadeler (Altınsoy, 2019):

$P(A|B)$: B durumu gerçekleştiğinde A durumunun gerçekleşme olasılığını, $P(A)$: A durumunun gerçekleşme olasılığını, $P(B|A)$: A durumu gerçekleştiğinde B durumunun gerçekleşme olasılığını, $P(B)$: B durumunun gerçekleşme olasılığını temsil etmektedir.

3.7.2. OneR (One Rule)

Hata temelli kural çıkarımına dayalı bu algoritma Holte (1993) tarafından önerilmiştir (Başer vd., 2020). Veride bulunan yordayıcı değişkenlerin her birine ait kurallar ürettikten sonra hatanın en az olduğu kuralı “bir kural” seçen basit bir yapıya sahip algoritmadır (Özden, 2018). Basit yapısına rağmen bu yöntem ile, birçok veri seti üzerinde başarılı sonuçlar elde edilmiştir (Kaynar vd., 2018).

Her değişkenin tek olarak kullanıldığında oluşabilecek hata değerini hesaplayıp değişkenleri sıralama amacıyla, veri setinde bulunan her değer için bir model oluşturmaktadır. Bu modeller sonunda elde edilen toplam hatalar küçükten büyüğe doğru sıralanarak, eşik değeri altında kalan modeli oluşturan değişken belirlenir ve veri setinden çıkarılır (Kaynar vd. 2018). Tahminlerin ayrı ayrı sınıflandırma üzerindeki etkisi sonunda ortaya çıkmış olan toplam hatanın küçük çıkması, tahminin ne kuvvetli olduğunu göstermektedir (Tavacı, 2011).

OneR algoritmasının çalışma mantığı şu şekildedir (Koca, 2022):

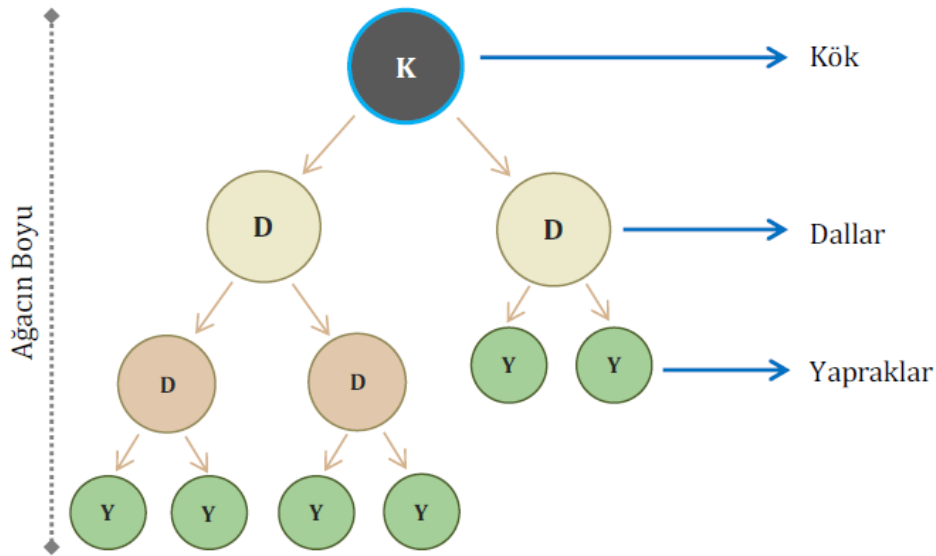
Her bir özellik için, aşağıdaki kural oluşturulur:

- Her bir sınıfın ne kadar görüldüğü sayılır
- Çok görülen sınıf bulunur
- Oluşturulan kural tanımlanır, bu özellik-değeri sınıflandırılır
- Kuralların hata oranı hesaplanır
- Kurallar en küçük hata oranıyla seçilir.

3.7.3. J48 (Karar Ağaçları Algoritması)

Weka veri madenciliği programında bulunan J48, C4.5 karar ağacı algoritması ve entropi değerlerini daha verimli hale getirmek amacıyla Ross Quinlan (1993) tarafından geliştirilmiş istatistiksel bir sınıflandırma algoritmasıdır (Gürtürk, 2022). J48’de anlaşılmasının kolaylığı ve karmaşık olmaması açısından karar verici olarak karar ağaçları tercih edilir. Ağaç karmaşıklığı işlemlerin doğruluğunda etkili olduğundan önemlidir (Akgül, 2019).

Karar ağaçları, kök, düğüm, dal ve yapraktan oluşan bir yapıya sahiptir (Özdemir, 2020). Bu yapıda bulunan düğüm sayısı, yaprak sayısı, ağacın derinliği, ağacın oluşturulmasındaki özellikler ağacın karmaşıklığını belirler. J48 algoritmasındaki amaç da bu karar ağacının boyutunun küçük tutulabilmesidir (Akgül, 2019). J48 algoritmasının ağaç yapısı Şekil 16’da yer almaktadır.



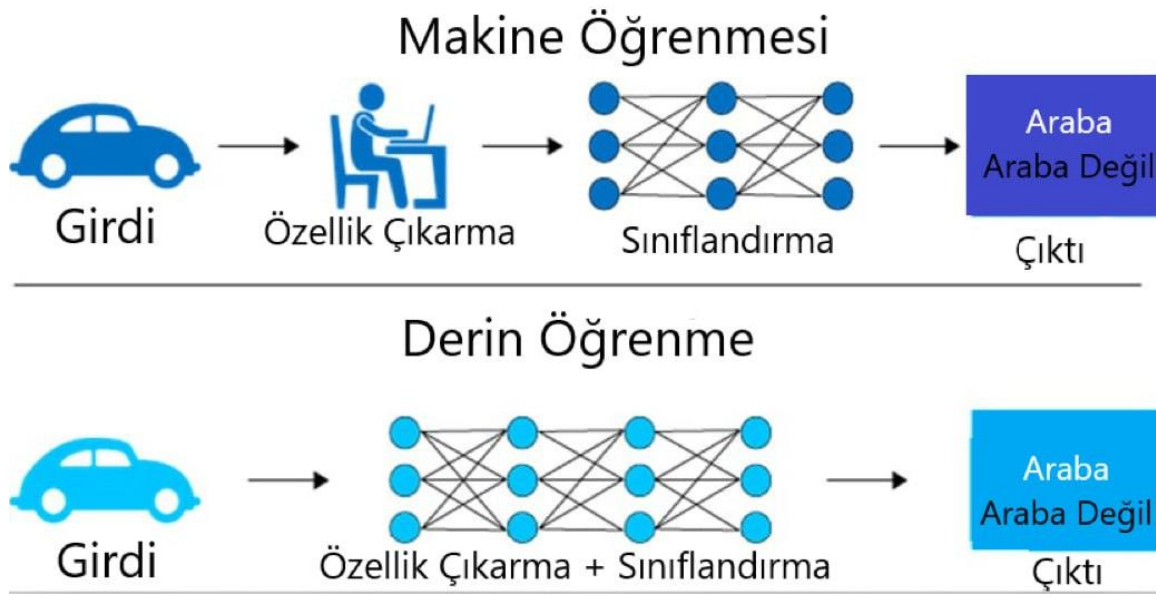
Şekil 16: J48 Algoritması Ağaç Modeli (Çoban, 2011)

J48 algoritmasında gerçekleştirilen sınıflandırmanın aşamaları, eldeki verilerin eğitim ve test gruplarına ayrılması, eğitim verilerindeki belirsizliklerin giderilip öğrenme aşamasına geçilmesi, karar ağacı modelinin oluşturulması, oluşan modelin test verileri ile karşılaştırılarak güvenilirliğinin test edilmesi ve elde edilen verilerin sınıflandırılması şeklindedir (Bolat, 2022).

3.8. Derin Öğrenme Yöntemleri

Makine öğrenmesinin alt dalı olarak bilinen derin öğrenme, insan beyninin ve beyindeki sinir hücrelerinin davranışı ve öğrenme şeklini örnek alarak tasarlanmıştır (Aslan, 2020).

Veri setleri kullanılarak eğitilen derin öğrenme modellerini kullanarak, mevcut veri setinin çıktılarıyla alakalı tahminleme işlemi gerçekleştirilebilir ve algoritmalar her çalıştığında değişiklikler yapılarak sonuç geliştirilmeye çalışılır ki bu bir insanın bir konuyla ilgili her düşündüğünde farklı fikirlere sahip olabilmesine benzer. Derin öğrenmede makine öğrenmesinden farklı olarak denetimli veya denetimsiz olarak yapılabilen özellik çıkarımı işlemi için ön işleme gerektirmez (Şekil 17).



Şekil 17: Makine Öğrenmesi ve Derin Öğrenmenin Katmansal Karşılaştırılması (Aslan,2020)

Derin öğrenmenin yapay sinir ağlarına benzer katmanlı bir yapısı vardır. Bu yapıda işlenmemiş girdilerden aşamalı bir şekilde üst düzey özelliklerin çıkarılmasında girdi ve çıktı katmanlarına ek olarak arada birden fazla gizli katman kullanılmaktadır (Mahmat, 2022). Her art arda gelen katman, kendisinden önceki katmanın çıktısını girdi olarak kullanır. Bu birden fazla katman yapısı sayesinde “derin” sıfatının derin öğrenme için kullanılmasına sebep olmuştur (Afatsun, 2020).

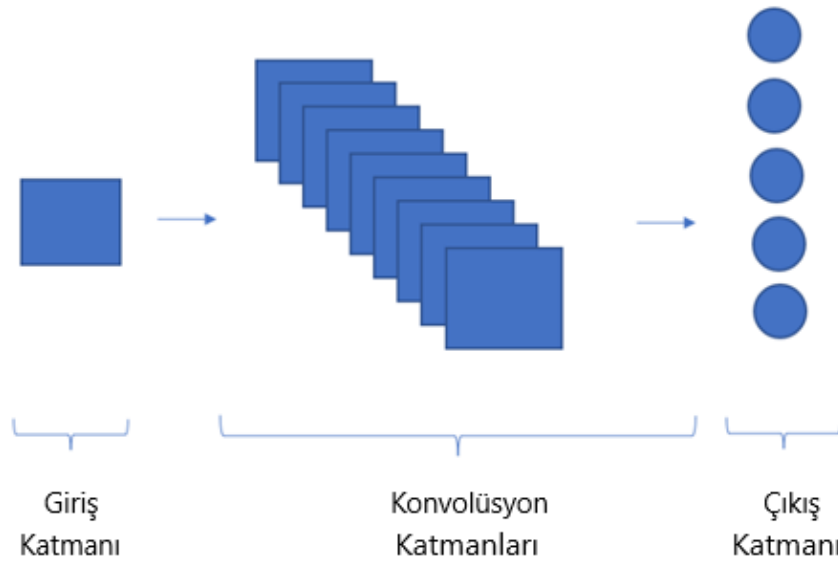
Teknolojik imkanların değişmesiyle doğan farklı ihtiyaçlara cevap vermek için bu katmansal yapılar farklılaştırılarak çözümler üretilmiştir. Bu çözümler derin öğrenme mimarilerinin farklılaşmasına sebep olmuştur. Literatür incelendiğinde bu mimarilerden en çok görüntü işleme, resim tanıma, nesne tespiti- takibi gibi alanlarda evrimsel sinir ağları (CNN), finansal uygulama, doğal dil işleme konularında uzun kısa süreli bellek (LSTM), yine doğal dil işleme, konuşma tanıma gibi alanlarda ise tekrarlayan yapay sinir ağları (RNN) kullanıldığı görülmektedir (Yılmaz ve Orman, 2021 ve Aslan, 2020).

3.8.1. CNN (Convolution Neural Network- Evrişimsel Sinir Ağları)

Derin öğrenme yöntemlerinin en yaygın olarak kullanılanlarından birisi olan CNN, 1989 yılında LeCun tarafından tanımlanmıştır (Seyidbayli, 2022). Ancak evrişimsel sinir ağlarının donanımsal ihtiyaçlarının yüksek olmasından dolayı başlarda gelişimi yavaşlamıştır. Donanımsal kapasitelerin gelişmesiyle 2010'lu yıllarda, CNN'lerin yeterli süre içinde çalışabilecek seviyelere gelinmiş ve kullanımı hızla yayılmaya başlamıştır (Yaşar, 2022).

Tek boyutlu olan zaman serilerinin ve iki boyutlu olan görsel verilerin işlenmesinde kullanılan bu yapay sinir ağı, son zamanlardaki donanımsal değişimlerin gerçekleşmesiyle üç boyutlu mimariler de gelişmeye başlamıştır (Tekinarslan, 2022). Dil içerikli olan sinyal ve diziler bir boyutlu, görüntü, ses gibi içerikler iki boyutlu, videolar, hacimli görüntüler üç boyutlu yapılarda kullanılır (LeCun ve ark., 2015).

Evrişim adındaki matematiksel işlemi kullanan CNN, birden fazla katman ve parametreden oluşmaktadır. Girdi verisinin içerisindeki ağırlıkları işlemektense öğrenebilir parametreler içeren filtreyi girdi üzerinde dolaştırma mantığıyla çalışır. Bu işlem sonunda modeldeki parametre sayısı azaldığından, girdi üzerinde yer alan komşu verilerin ilişkisi daha iyi yakalanmış olmaktadır (Özbilen, 2022). Bu ağlarda alınan giriş verileri katmanlar halinde işlenerek eğitim süreci gerçekleştirilir (Fındık, 2022). Yapısında yer katmanlar genel hatlarıyla Şekil 18'de görüldüğü gibi giriş katmanı, konvolüsyon (evrişimsel) katmanları ve çıkış katmanıdır.



Şekil 18: CNN Katman Yapısı (Keskin, 2022)

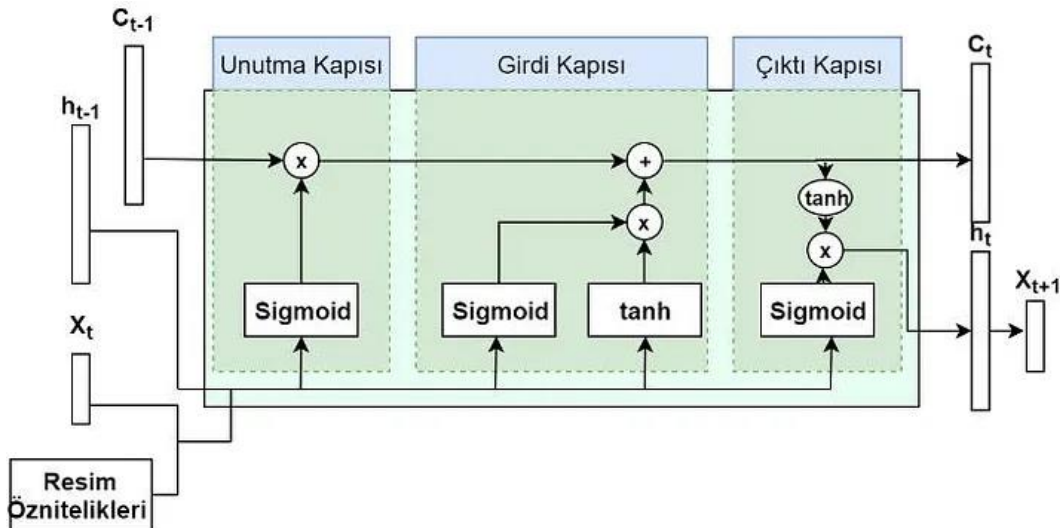
CNN modelinde, nesnelerin tanımlanmasında ön işlem yapılmadan nesnelerin arka plandan

ayrılması, geleneksel yöntemlere göre yüksek performanslı görüntü tanımlaması gerçekleştirmesi, zaman serisine dayanan işlemlerde tahminleri yüksek doğrulukla gerçekleştirmesi en önemli avantajlarından. Ancak bu ağların verimli çalışmasında güçlü Grafik İşleme Ünitesi (GPU) ve yüksek bellek kapasitelerine ihtiyaç olması açısından dezavantajları da mevcuttur (Fındık, 2022).

3.8.2. LSTM (Long Short-Term Memory- Uzun Kısa Süreli Bellek)

Derin öğrenme uygulamalarının zaman serilerinde kullanılması için Tekrarlayan Sinir Ağları (RNN) ortaya çıkmıştır. RNN de LSTM ve GRU da dahil olmak üzere farklı alt yöntemlere sahiptir. LSTM, bu yöntemler arasında en yaygın olarak kullanılan ve en çok bilinen modeldir (Çetintaş, 2022).

İnsan beyni, bir metin okuduğunda aradığı bilgiyi bulduktan sonra genelde metnin geri kalan kısmını unutması şeklinde çalışarak verileri işlemektedir. İnsan beyninin bu işleyişini taklit eden ve RNN'in gelişmiş versiyonu olan LSTM, Hochreiter ve Schmidhuber (1997) tarafından tasarlanmıştır (Eren, 2022). Bilgileri varsayılan olarak uzun süre saklayabilen LSTM, RNN'deki kaybolan gradyanlar problemini çözmek için kapı olarak adlandırılan, neyin hatırlanacağı neyin unutulacağına karar verilmesi için 3 yapı geliştirmiştir (Şekil 19). LSTM, bu kapılar aracılığıyla hazırlanan hücre yapılarına bilgilerin eklenmesi ya da kaldırılması işlevlerine sahiptir (Fındık, 2022).



Şekil 19: LSTM Modeli Kapı Yapısı (Yıldız, 2020)

Hücrelerin durumunu kontrol edip koruyan bu kapılar şu şekildedir (Keskin, 2022):

Unutma Kapısı: Bellekte var olan bilgilerin hangilerinin tutulup hangilerinin tutulmayacağını

aktivasyon fonksiyonları kullanarak karar verir.

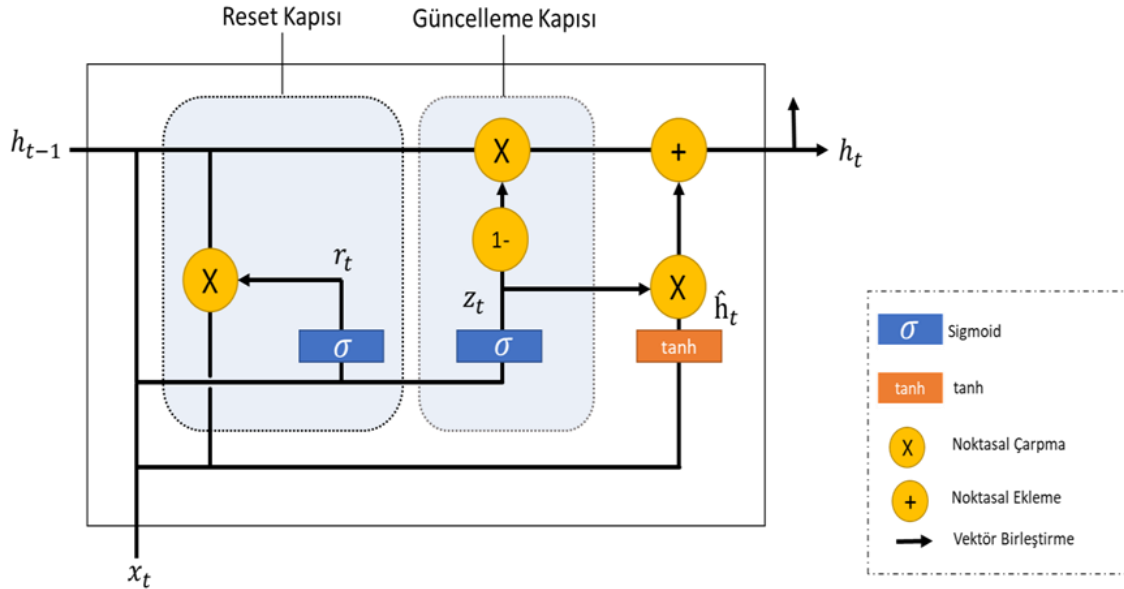
Girdi Kapısı: Verilerin ne olduğunu öğrenip, girdilerin hafıza hücrelerine akışını kontrol eder.

Çıktı Kapısı: Çıktıların ağı akışını belirleyerek tahmin yapar.

3.8.3. GRU (Gated Recurrent Unit- Kapılı Tekrarlayan Birim)

Tekrarlayan Sinir Ağları (RNN) mimarisinin bir çeşidi olan ve LSTM mimarisinin de birden fazla varyasyonundan en çok tercih edilen Kapılı Tekrarlayan Birim (GRU), Cho ve arkadaşları tarafından 2014 yılında ortaya konulmuştur. LSTM yapısı üzerindeki giriş ve unutma kapısını güncelleme kapısı olarak birleştirerek değişiklik yapmışlardır (Eren, 2022). GRU ve LSTM mimarileri birbirine benzediğinden başarılı sonuçlar üretebilmektedirler (Fındık, 2022).

GRU yapısının temelinde sıfırlama kapısı ve güncelleme kapısı yer almaktadır (Şekil 20). Model üzerinde var olan geçmiş bilgilerden ne kadarının unutulacağına karar vermede sıfırlama kapısı, modeldeki önceki zamanda var olan adımların hangilerinin ileriye dönük aktarılmasının belirlenmesinde güncelleme kapısı kullanılmaktadır (Ballı, 2022). GRU bu yapısıyla LSTM'ye göre daha basit, hızlı eğitilebilen, daha verimli yürütülebilen bir model olmaktadır (Gencer, 2022).



Şekil 20: GRU Mimarisi (Canlı, 2022)

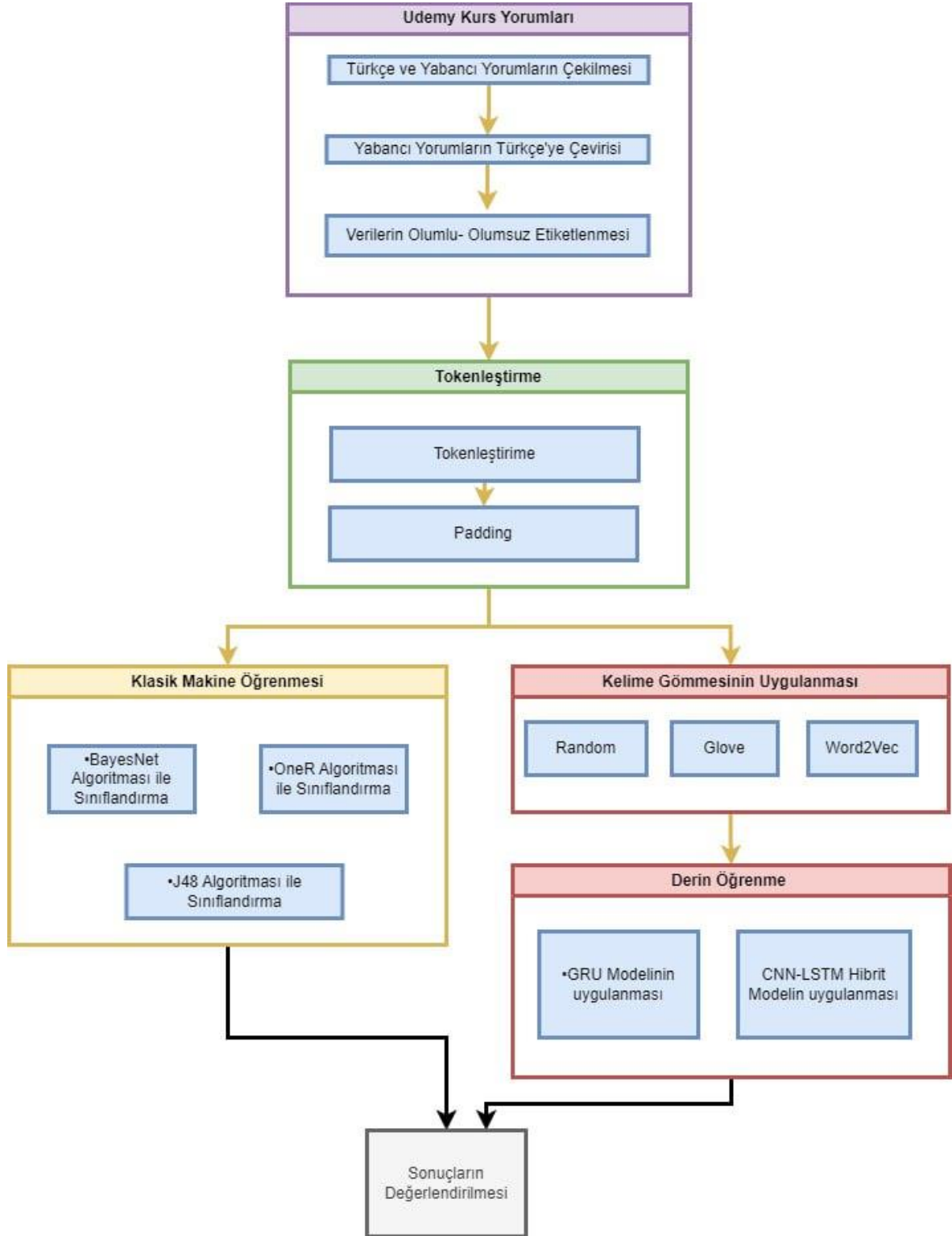
Yapısal olmayan metinlerin adlandırılmasında (Gencer, 2022), sermaye piyasası tahmini (Eren, 2022), rüzgar hızı tahminlemede (Fındık, 2022) kullanılan GRU, konuşma tanıma, insan genomu ve el yazısı analizi gibi birçok çalışmada da kullanılabilmektedir.

4. BULGULAR

Makine öğrenmesi kullanılacak çalışmalarda verilerin boyut ve çeşidine göre bilgisayar özelliklerine olan gereksinim değişmektedir. Derin öğrenme mimarilerinin çalışmasında ise veri setindeki metin sayısının fazla olmasından dolayı donanımsal olarak güçlü makinelere ihtiyaç duyulmaktadır. Bu çalışmada 8 Gb Ram, 2 Gb Nvidia Ekran Kartı, 128 Gb SSD donanımlarına sahip masaüstü bilgisayar kullanılmıştır.

Bu çalışmada Udemy sitesindeki kurslara yapılan yorumların olumlu ya da olumsuz olduğunu belirleyebilmek için derin öğrenme yöntemlerinin çalıştırılmasında Python dili kullanılarak, Jupyter Lab ve Spyder aracılığıyla kodların çalıştırılması gerçekleştirilmiştir. Klasik makine öğrenmesi yöntemleriyle bu işlemin gerçekleştirilmesi için Weka programı kullanılmıştır.

Bu çalışmada öncelikle Udemy sitesinde bulunan kurslara yapılan yorumlar, Udemy API ve Python programlama dili kullanılarak excel ve .csv dosyası halinde indirilmiştir. Bu yorumlardan Türkçe olmayan yorumların Türkçe'ye çevirisi yapılmıştır. Elde edilen veriler olumlu olumsuz olarak etiketlendikten sonra oluşan dosya .csv dosyası olarak kaydedilmiştir. Düzenlenen verilerin kelime gömmelerinin işleyebileceği hale getirilmesi için tokenleştirme işlemi uygulanmıştır. Tokenleştirme işleminden sonra elde edilen verilerle klasik makine öğrenmeleri uygulanmıştır. Yine aynı tokenleştirme sonucuyla öncelikle kelime gömmelerinden Random, GloVe ve Word2Vec yöntemleri uygulanıp, her birinden sonra ayrı ayrı GRU ve CNN-LSTM hibrit derin öğrenme modelleri uygulanmıştır (Şekil 21).



Şekil 21: Çalışmada Uygulanan İşlem Basamakları

4.1. Klasik Makine Öğrenmeleriyle Elde Edilen Sonuçlar

Bu çalışmada klasik makine öğrenmesi algoritmalarından BayesNet, OneR ve J48 kullanılarak, Udemy’de bulunan kursların yorumları olumlu olumsuz olarak sınıflandırma doğrulukları karşılaştırılmıştır. Weka programı kullanılarak, 10 fold ile çalıştırılan algoritmaların doğruluk ve F skorları Tablo 5’te yer almaktadır.

Tablo 5: Klasik Makine Öğrenmesi Algoritmaları Karşılaştırması

	Accuracy (Doğruluk)	F Skoru	Fold
BayesNet	91.596	0.916	10
OneR	85.1305	0.851	10
J48	82.9085	0.829	10

Tablo 5’te görüldüğü gibi BayesNet ile % 91.596 doğruluk ve %91.6 F skoru, OneR ile %85.1305 doğruluk ve %91.6 F skoru, J48 ile % 82.9085 doğruluk ve %82.9 F skoru değeri elde edilmiştir. Sonuçlardan da anlaşıldığı gibi klasik makine öğrenmesi algoritmalarından en yüksek doğruluk ve F skoru değeri BayesNet ile elde edilmiştir.

Bu çalışmada kullanılan BayesNet, OneR ve J48 algoritmalarının olumlu ve olumsuz yorumları sınıflandırmasıyla ilgili karşıtlık matrisleri Tablo 6, Tablo 7 ve Tablo 8’de yer almaktadır. *a* ile ifade edilen değer olumsuz, *b* ile ifade edilen değer ise olumlu yorumları ifade etmektedir.

Tablo 6: BayesNet Karşıtlık Matrisi

a	b	
93308	6692	a=0
10116	89884	b=1

Tablo 6’da yer alan BayesNet’ e ait karşıtlık matrisi değerlendirildiğinde, BayesNet algoritması olumsuz olan yorumlardan 93308 adetini doğru olarak sınıflandırırken, 6692 adetini yanlış sınıflandırdığı, olumlu yorumların ise, 89884 adetini doğru sınıflandırırken 10116 adetini yanlış sınıflandırdığı görülmektedir.

Tablo 7: OneR Karşıtlık Matrisi

a	b	
85784	14216	a=0
15523	84477	b=1

Tablo 7’de yer alan OneR’ye ait karşıtlık matrisi değerlendirildiğinde, OneR algoritmasının olumsuz olan yorumlardan 85784 adetini doğru olarak sınıflandırırken, 14216 adetini yanlış sınıflandırdığı, olumlu yorumların ise, 89884 adetini doğru sınıflandırırken 10116 adetini

yanlış sınıflandırdığı görülmektedir.

Tablo 8: J48 Karşıtlık Matrisi

a	b	
86979	13021	a=0
21162	78838	b=1

Tablo 8’de yer alan J48’e ait karşıtlık matrisi değerlendirildiğinde, J48 algoritmasının olumsuz olan yorumlardan 86979 adetini doğru olarak sınıflandırırken, 13021 adetini yanlış sınıflandırdığı, olumlu yorumların ise, 78838 adetini doğru sınıflandırırken 21162 adetini yanlış sınıflandırdığı görülmektedir.

Karşıtlık matrisleri ve doğruluk oranları birlikte değerlendirildiğinde 10 fold ile çalıştırılan klasik makine öğrenmesi algoritmalarından BayesNet algoritmasının daha başarılı olduğu görülmüştür.

4.2. Derin Öğrenme Yöntemiyle Elde Edilen Sonuçlar

Bu çalışmada Random, Glove ve Word2vec kelime gömmelerinin ardından, derin öğrenme modellerinden GRU ve CNN-LSTM hibrit modelleri kullanılarak, bu modellerin olumlu olumsuz yorumların tahminindeki başarıları değerlendirilmiştir. Kullanılan modeller her bir kelime gömmesi ayrı ayrı çalıştırdıktan sonra farklı derin öğrenme modelleri uygulanmıştır.

Embedding size değeri 300, learning rate (öğrenme oranı) 0.001 ve sigmoid fonksiyonu ile eğitilen GRU modeli sonucu elde edilen değerler Tablo 9’da yer almaktadır.

Tablo 9: GRU Modeli Sonuçları

Kelime Gömmesi	Doğruluk	Geri Çağırma	Kesinlik	Kayıp
Random	0.9499	0.9468	0.9458	0.2236
GloVe	0.9567	0.9582	0.9493	0.1193
Word2Vec	0.9441	0.9406	0.9396	0.1488

Tablo 9’ da Random, GloVe ve Word2Vec kelime gömmelerinin uygulandıktan sonra derin öğrenme modellerinden GRU sonucunda ortaya çıkan doğruluk, geri çağırma, kesinlik ve kayıp değerleri yer almaktadır. Random kelime gömmesinden sonra GRU modelinin eğitimi sonunda, %94.99 doğruluk, %94.68 geri çağırma, %94.58 kesinlik ve 0.2236 kayıp oranları elde edilmiştir. GloVe kelime gömmesinden sonra GRU ile gerçekleştirilen eğitim sonunda, %95.67 doğruluk, %95.82 geri çağırma, %94.93 kesinlik ve 0.1193 kayıp oranları elde edilmiştir. Word2Vec kelime gömmesi sonunda ise GRU ile, %94.91 doğruluk, %94.06 geri çağırma, %93.96 kesinlik ve 0.1488 kayıp oranları elde edilmiştir. GRU ile gerçekleştirilen

eğitimler karşılaştırıldığında en yüksek doğruluk, geri çağırma ve kesinlik değerinin veri setine GloVe kelime gömmesi uygulandıktan sonra modelin kullanımı sonunda elde edildiği görülmüştür.

CNN-LSTM hibrit modeli için, Random, Word2Vec ve GloVe kelime gömmeleri uygulandıktan sonra, evrişim katmanının 3, 4 ve 5 filtre boyutları, dropout katmanı için 0.1 ve 0.2 değerleri, pool size değerleri ise 1, 2, 3, 4 olarak değiştirilerek, doğruluk, kesinlik, geri çağırma ve kayıp değerleri elde edilmiştir. Elde edilen bu sonuçlardan her bir kelime gömmesi için CNN parametreleri filtre boyutu, dropout değeri ve pool size değerleri aynı belirlenen üç değer için elde edilen doğruluk oranları Tablo 10, Tablo 11 ve Tablo 12’de görülmektedir.

Tablo 10: Random Kelime Gömmesi Uygulaması Sonrası CNN-LSTM Hibrit Model Sonucu

CNN Parametreleri						
Filtre Boyutu	Drop out	Pool size	Doğruluk	Geri Çağırma	Kesinlik	Kayıp
3	0.2	1	0.9458	0.9580	0.9280	0.1515
3	0.2	4	0.9456	0.9408	0.9423	0.1437
4	0.1	2	0.9238	0.9099	0.9254	0.2278

Tablo10’da Random kelime gömmesi uygulandıktan sonra CNN-LSTM Hibrit model ile yapılan eğitim sonuçları yer almaktadır. Filtre boyutu 3 ve dropout değeri 0.2 iken pool size değeri 1 olarak ayarlandığında, %94.58 doğruluk, %95.80 geri çağırma, %92.80 kesinlik ve 0.1515 kayıp oranları elde edilmiştir. Filtre boyutu 3, dropout değeri 0.2 ve pool size değeri 4 olarak ayarlandığında ise %94.56 doğruluk, %94.08 geri çağırma, %94.23 kesinlik ve 0.1437 kayıp oranları elde edildiği görülmüştür. Filtre boyutu 4, dropout değeri 0.1 ve pool size değeri 2 olarak ayarlandığında %92.38 doğruluk, %90.99 geri çağırma, %92.54 kesinlik ve 0.2278 kayıp oranları elde edildiği görülmektedir. Random kelime gömmesi sonucunda en yüksek doğruluk oranının filtre boyutu 3, dropout değeri 0.2 ve pool size değeri 1 iken elde edildiği görülmektedir.

Tablo 11: GloVe Kelime Gömmesi Uygulaması Sonrası CNN-LSTM Hibrit Model Sonucu

CNN Parametreleri						
Filtre Boyutu	Drop out	Pool size	Doğruluk	Geri Çağırma	Kesinlik	Kayıp
3	0.2	1	0.9493	0.9464	0.9449	0.1415
3	0.2	4	0.9464	0.9289	0.9548	0.1506
4	0.1	2	0.9216	0.9001	0.9296	0.2303

Tablo11’de GloVe kelime gömmesi uygulandıktan sonra CNN-LSTM Hibrit model ile yapılan eğitim sonuçları yer almaktadır. Filtre boyutu 3 ve dropout değeri 0.2 iken pool size değeri 1 olarak ayarlandığında, %94.93 doğruluk, %94.64 geri çağırma, 94.49 kesinlik ve 0.1415 kayıp oranları elde edilmiştir. Filtre boyutu 3 ve dropout değeri 0.2 iken pool size değeri 4 olarak ayarlandığında ise %94.64 doğruluk, %92.89 geri çağırma, %95.48 kesinlik ve 0.1506 kayıp oranları elde edildiği görülmüştür. Filtre boyutu 4, droupout değeri 0.1 ve pool size değeri 2 olarak ayarlandığında %92.16 doğruluk, %90.01 geri çağırma, %92.96 kesinlik ve 0.2303 kayıp oranları elde edildiği görülmektedir. GloVe kelime gömmesi sonucunda en yüksek doğruluk oranının filtre boyutu 3, dropuout değeri 0.2 ve pool size değeri 1 iken elde edildiği, ancak değişen değerlere rağmen doğruluk oranları arasında farkın neredeyse hiç olmadığı görülmektedir.

Tablo 12: Word2Vec Kelime Gömmesi Uygulaması Sonrası CNN-LSTM Hibrit Modeli

CNN Parametreleri						
Filtre Boyutu	Drop out	Pool size	Doğruluk	Geri Çağırma	Kesinlik	Kayıp
3	0.2	1	0.9484	0.9454	0.9439	0.1415
3	0.2	4	0.8774	0.8264	0.9023	0.2637
4	0.1	2	0.9416	0.9466	0.9293	0.1912

Tablo12’de Word2Vec kelime gömmesi uygulandıktan sonra CNN-LSTM Hibrit model ile yapılan eğitim sonuçları yer almaktadır. Filtre boyutu 4, dropout değeri 0.1 ve pool size değeri 2 olarak ayarlandığında, %94.16 doğruluk, %94.66 geri çağırma, %92.93 kesinlik ve 0.1912 kayıp oranları elde edilirken, Filtre boyutu 3, dropout değeri 0.2 ve pool size değeri 4 olarak ayarlandığında ise %87.74 doğruluk, %82.64 geri çağırma, %90.23 kesinlik ve 0.2637 kayıp oranları elde edildiği görülmüştür. Filtre boyutu 3, droupout değeri 0.2 ve pool size değeri 1 olarak ayarlandığında %94.84 doğruluk, %94.54 geri çağırma, %94.39 kesinlik ve 0.1415 kayıp oranları elde edildiği görülmektedir. Word2Vec kelime gömmesi sonucunda en yüksek doğruluk oranının filtre boyutu 3, dropuout değeri 0.2 ve pool size değeri 1 iken elde edildiği görülmektedir.

Derin öğrenme modellerinden en iyi doğruluk oranları GRU ile kelime gömmelerinden GloVe kullanılarak, Random, GloVe ve Word2Vec kelime gömmeleri uygulandıktan sonra CNN-LSTM Hibrit modelinde farklı filtre boyutu, droupout ve pool size değerleri ile elde edilen sonuçlar değerlendirilmiştir (Tablo 13).

Tablo 13: Derin Öğrenme Modellerinin Sonuçları

	Doğruluk	Geri Çağırma	Kesinlik	Kayıp
GRU (GloVe)	0.9567	0.9582	0.9493	0.1193
CNN-LSTM (Random)	0.9458	0.9580	0.9280	0.1515
CNN-LSTM (GloVe)	0.9493	0.9464	0.9449	0.1415
CNN-LSTM (Word2Vec)	0.9484	0.9454	0.9439	0.1415

Tablo 13’te çalışmada kullanılan derin öğrenme modellerinden elde edilen doğruluk, kesinlik, geri çağırma ve kayıp sonuçları yer almaktadır. GloVe kelime gömmesi yapıldıktan sonra GRU derin öğrenme modeliyle %95.67 doğruluk, %95.82 geri çağırma, %94.93 kesinlik ve 0.1193 kayıp oranları elde edilmiştir. Random kelime gömmesi yapıldıktan sonra CNN-LSTM derin öğrenme modeliyle %94.58 doğruluk, %95.80 geri çağırma, %92.80 kesinlik ve 0.1515 kayıp oranları, GloVe kelime gömmesi yapıldıktan sonra CNN-LSTM derin öğrenme modeliyle %94.93 doğruluk, %94.64 geri çağırma, %94.49 kesinlik ve 0.1415 kayıp oranları, Word2Vec kelime gömmesi yapıldıktan sonra CNN-LSTM derin öğrenme modeliyle %94.84 doğruluk, %94.54 geri çağırma, %94.39 kesinlik ve 0.1415 kayıp oranları elde edildiği görülmektedir. Sonuçlar incelendiğinde en yüksek başarının GloVe kelime gömmesi sonunda GRU modeliyle elde edildiği görülmektedir.

5. SONUÇ VE ÖNERİLER

Son yıllarda teknolojinin gelişmesiyle birlikte eğitimin şekli de giderek değişmeye başlamıştır. Önceleri klasik olarak sınıflarda olan eğitim artık bir yere takılı kalmadan, istenilen her yerde, herkesin kendi şartlarına göre alabileceği, ulaşılabilmesi, istediği öğretimsel faaliyeti elde edebilmesi çok önemlidir. Tüm bu parasal duruma ve mekana bağlı olmadan eğitime ihtiyaç duyulmasının yanında teknolojinin de ilerlemesiyle farklı çözümler üretilmeye başlanmıştır. Bu çözümler herkese açık ve çoğu zaman parasız olan, herkesin istediği yerden istediği zaman ulaşabileceği Kitlesel Açık Çevrimiçi Dersler (KAÇD) adı altında toplanmıştır. Bu sistemler çevrimiçi siteler üzerinden eğiticiler tarafından hazırlanan kurslara, ilgili kursu almak isteyen kişinin kayıt olması şeklinde çalışmaktadır.

Coursera, FutureLearn, Khan Academy, Udemy gibi zamanla gelişmesine devam eden KAÇD'ler mevcuttur. Bu sistemlerde kurslar, bazıları az bir ücretle, bazıları ücretsiz bir şekilde yer almaktadır. Kullanıcılar kurslara kaydolarken, birçok çevrimiçi işlerimizde olduğu gibi, kursla ilgili yorumlara bakmaktadırlar. Kurslarla ilgili yorumlara bakarak değerlendirmenin gerçekleştirilmesi bazen çok zor olmaktadır. Bu çalışmada kullanıcıların kurs yorumlarını okumasına gerek kalmadan kursla ilgili olumlu-olumsuz olmak üzere bir yargıya varmaları amacıyla makine öğrenmesi ve derin öğrenme yöntemleri kullanılarak yorumlar sınıflandırılmıştır. Bu sınıflandırmanın gerçekleştirilmesi için gerekli olan veriler KAÇD platformu olan Udemy sitesindeki kurslara yapılan yorumlardan elde edilmiştir. 100000 olumlu, 100000 olumsuz yorum kullanılarak oluşturulan veri seti, klasik makine öğrenmesi algoritmalarından BayesNet, OneR ve J48 kullanılarak sınıflandırma yapılmış, doğruluk ve F skoru değerleri karşılaştırılmıştır. Derin öğrenme modelleri ile ise, Random, GloVe ve Word2Vec kelime gömmeleri kullanıldıktan sonra GRU ve CNN-LSTM Hibrit modelleri kullanılarak eğitim gerçekleştirilmiştir. Kullanılan bu modeller için ise doğruluk, kesinlik, geri çağırma ve kayıp değerleri karşılaştırılmıştır. Klasik makine öğrenmesi algoritmalarından %91.596 doğruluk ve %91.6 F skoru ile en başarılı sınıflandırma BayesNet ile gerçekleştirilmiştir. Derin öğrenme modellerinden GRU ile en başarılı sonuç, GloVe kelime gömmesi kullanıldıktan sonra %95.67 doğruluk, %95.82 geri çağırma, %94.93 kesinlik ve 0.1193 kayıp oranı ile elde edilmiştir. CNN-LSTM Hibrit modeli kullanılarak en iyi Random kelime gömmesinden sonra %94.58 doğruluk, %95.80 geri çağırma, %92.80 kesinlik ve 0.1515 kayıp değerleri, GloVe kelime gömmesinden sonra %94.93 doğruluk,

%94.64 geri çağırma, %94.49 kesinlik ve 0.1415 kayıp değerleri, Word2Vec kelime gömmesinden sonra %94.84 doğruluk, %94.54 geri çağırma, %94.39 kesinlik ve 0.1415 kayıp değerleri elde edilmiştir. CNN-LSTM modelleriyle en başarılı sonuç çok az bir farkla GloVe kelime gömmesi yapıldıktan sonra alındığı görülmüştür. GRU ve CNN-LSTM hibrit modellerinden en başarılı sonuç ise GloVe kelime gömmesi uygulandıktan sonra GRU kullanılmasıyla elde edilmiştir. Tüm sonuçlar klasik makine öğrenmesi ve derin öğrenme açısından incelendiğinde kullanılan algoritma ve modeller ile, Udemy’de bulunan kurs yorumlarını olumlu olumsuz ayırmada başarılı sonuçlar elde edildiğini göstermektedir.

Gelecekte yapılacak çalışmalarda, farklı kelime gömmeleri kullanılarak, farklı makine öğrenmesi algoritmaları ve derin öğrenme yöntemleriyle ve yeterli donanımına sahip bilgisayarlar kullanarak, yüksek başarıya sahip çalışmalar yapılması düşünülmektedir.

6. KAYNAKÇA

- Abdalrahman, G. A. (2020). Twitter Sentiment Analysis Using Deep Learning. (Yüksek Lisans Tezi), Fırat University Graduate School Of Natural And Applied Sciences Department Of Software Engineering, Elazığ.
- Afatsun, M. N. (2020). Derin Öğrenme Yöntemleri ile Türkçe Metinlerden Anlamlı Özet Çıkarma. (Yüksek Lisans Tezi), Ankara Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı, Ankara.
- Akdemir, D. M. (2022). E-Ticaret Sitelerinde Tüketicilerin Satın Alma ve Chatbotları Tekrar Kullanma Niyetini Etkileyen Faktörler Üzerine Bir Araştırma. (Yüksek Lisans Tezi), Dokuz Eylül Üniversitesi Sosyal Bilimler Enstitüsü Bilgi Yönetimi Anabilim Dalı, İzmir.
- Akgül, E. S. (2019). Makine Öğrenmesi ile Statik Kaynak Kod Analizi. (Yüksek Lisans Tezi), Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı, İstanbul.
- Altınsoy, F. (2019). Uzaktan Eğitim Öğrencilerinin Başarılarının Yapay Zeka Teknikleri ile Tahmini. (Yüksek Lisans Tezi), Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı, Isparta.
- Aslan, B. (2020). Derin Öğrenme ile Borsa Verileri Üzerinde Tahminleme Yapılması. (Yüksek Lisans Tezi), Ege Üniversitesi Bilgisayar Mühendisliği Anabilim Dalı, İzmir.
- Atık, A. ve Ata, A. (2018). Alternatif Dijital Eğitim Platformu Olarak Kitlel Çevrimiçi Açık Ders (Mooc) Uygulamaları. e-Journal of New World Sciences Academy, 13(4), 144-154.
- Aydın, S. (2007). Veri Madenciliği ve Anadolu Üniversitesi Uzaktan Eğitim Sisteminde Bir Uygulama. (Doktora Tezi), Anadolu Üniversitesi Sosyal Bilimler Enstitüsü, Eskişehir.
- Ballı, S. (2022). Derin Tekrarlayan Sinir Ağı Yöntemleri ile Erzurum Biyolojik Atıksu Arıtma Tesisinin Performans Tahmini. (Yüksek Lisans Tezi), Atatürk Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Ana Bilim Dalı, Erzurum
- Barjanzi, A. (2021). Investigating the Relationship Between Vocabulary Size Level,

Informal Exposure, Motivation, and the EFL Learners' Collocational Knowledge in Saudi Arabia. (Doktora Tezi), University of South Florida Department of Language, Literacy, ED.D., Exceptional Education, and Physical Education, ABD.

Başer, S. H., Hökelekli, O. ve Adem, K. (2020). Ortaöğretimde Öğrenim Gören Öğrenci Performanslarının Veri Madenciliği Yöntemleri ile Tahmin Edilmesi. Bilgisayar Bilimleri ve Teknolojileri Dergisi, 1 (1), 22-27. Retrieved from <https://dergipark.org.tr/en/pub/bibtet/issue/53797/702797>

Belen, N. (2018). Acil Tıp Kliniğine Göğüs Ağrısı Şikayeti ile Başvuran Hastalarda Otuz Günlük Kardiyak İlişkili Olayların Tanımlanmasında Klinik Karar Destek Sistemi (Clinical Decision Support System) Geliştirilmesi. (Uzmanlık Tezi), Sağlık Bilimleri Üniversitesi İzmir Bozyaka Eğitim ve Araştırma Hastanesi Acil Tıp Kliniği, İzmir

Berber, B. (2019). Yapay Zekanın Tarihi Gelişimi. 16.10.2022 tarihinde <https://teknofesor.com/yapay-zekanin-tarihi-gelisimi/> adresinden erişilmiştir.

Bilge Şahin, E. (2022). Bir Kitleleş Açık Çevrimiçi Ders Sisteminde Kullanıcı Deneyimi Değerlendirilmesinde Kullanılan Yöntemlerin Karşılaştırılması. (Yüksek Lisans Tezi), Kocaeli Üniversitesi Fen Bilimleri Enstitüsü, Kocaeli.

Bolat, H. (2022). Makine Öğrenme Yöntemleri ile Kahramanmaraş İlindeki Trafik Kazası Sonuçlarının Analizi ve Tahmin Edilmesi. (Yüksek Lisans Tezi), Munzur Üniversitesi Mühendislik Yönetimi Anabilim Dalı, Tunceli.

Canlı, H. (2022). IoT Tabanlı Akıllı Şehirlerde Derin Öğrenme ve Mobil Tabanlı Akıllı Park Sistemi Yaklaşımı. (Doktora Tezi), Düzce Üniversitesi Lisansüstü Eğitim Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı, Düzce.

Cetina, I. Goldbach, D. ve Manea, N. (2018). Udey: A Case Study In Online Education And Training. Revista Economica, Lucian Blaga University of Sibiu, Faculty of Economic Sciences,70(3), 46-54.

Chakravarthy, V.J. Kameswari, M. Devan Mydeen, H. ve Seenivasan, M. (2021). Opinion mining from student text review for choosing better online courses. IOP Conference Series: Materials Science and Engineering, 1070, 1-13

Chan, H. Y., Rajamohan, R., Gan, K.H. ve Samsudin, N. (2021, Eylül). Text Analytics on Course Reviews from Coursera Platform. 2021 IEEE International Conference on

Artificial Intelligence in Engineering and Technology (IICAIET), Kota Kinabalu, Malaysia.

Chaveesuk, S. Khalid, B. Bsoul-Kopowska, M. ve Rostańska, E. (2022). Comparative analysis of variables that influence behavioral intention to use MOOCs. PLoS ONE, 17(4), 1-26.

Çetintaş, A. O. (2022). Sivas İline Ait Meteorolojik Parametrelerin Arıma ve Lstm Yöntemleri Kullanılarak Modellenmesi ve Tahmini. (Yüksek Lisans Tezi), Ankara Üniversitesi Fen Bilimleri Enstitüsü Tarımsal Yapılar ve Sulama Anabilim Dalı, Ankara.

Çevik, F. (2019). Derin Öğrenme ve Kelime Gömme Modelleri Kullanarak Parkinson Hastalığının Analizi. (Yüksek Lisans Tezi), Doğu Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı, İstanbul.

Çoban, T. (2011). Makine Öğrenme Algoritmaları ile Web Sitelerinin Tıklamalarının Analizi. (Yüksek Lisans Tezi), Beykent Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı, İstanbul

Dehbi, A., Bakhouyi, A., Dehbi, R., ve Talea, M. (2022). MOOCs in Smart Education: Comparative Study by Applying AHP and COPRAS Method. International Journal of Emerging Technologies in Learning (iJET), 17(08), pp. 61–74.

Demirci, N. (2014). What is Massive Open Online Courses (MOOCs) and What is promising us for learning?: A Review-evaluative Article about MOOCs. Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi, 8(1), 231-256.

Demirtaş, M. A. (2018). Derin Öğrenme Yöntemleri ile Duygu Analizi ve Aktivasyon Fonksiyonlarının Karşılaştırılması. (Yüksek Lisans Tezi), Atatürk Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı, Erzurum.

Dina, N. Z., Yunardi, R. T., ve Firdaus, A. A. (2021). Utilizing Text Mining and Feature-Sentiment-Pairs to Support Data-Driven Design Automation Massive Open Online Course. International Journal of Emerging Technologies in Learning (iJET), 16(01), 134–151.

Edalati, M., Imran, A.S., Kastrati, Z., Daudpota, S.M. (2022). The Potential of Machine Learning Algorithms for Sentiment Classification of Students' Feedback on MOOC. In: Arai, K. (eds) Intelligent Systems and Applications. IntelliSys 2021. Lecture

Notes in Networks and Systems, vol 296. Springer, Cham.
https://doi.org/10.1007/978-3-030-82199-9_2

Eraldemir, S. G. (2014). EEG verileri kullanılarak metinsel okuma ve matematiksel işlemlerin analizi. (Yüksek Lisans Tezi), Mustafa Kemal Üniversitesi Fen Bilimleri Enstitüsü Elektrik-Elektronik Mühendisliği Ana Bilim Dalı, Hatay.

Eren, M. Ç. (2022). Yinelemeli Sinir Ağları ile Sermaye Piyasası Yön Tahmini Üzerine Bir Çalışma. (Yüksek Lisans), İstanbul Teknik Üniversitesi Lisansüstü Eğitim Enstitüsü Endüstri Mühendisliği Anabilim Dalı, İstanbul.

Fındık, M. M. (2022). Derin Öğrenme Tabanlı Hibrit Tahminleme Modeli Kullanarak Rüzgar Hızı Tahminlemesi. (Yüksek Lisans Tezi), Tokat Gaziosmanpaşa Üniversitesi Lisansüstü Eğitim Enstitüsü Elektrik-Elektronik Mühendisliği Anabilim Dalı, Tokat.

Gamage, D. Perera, I. ve Fernando, S. (2020). Exploring MOOC User Behaviors Beyond Platforms. International Journal of Emerging Technologies in Learning (IJET), 15(08), pp. 161–179.

Gencer, M. (2022). Yapısal Olmayan Metinler İçin Adlandırılmış Varlık Tanıma Algoritmaları ve Uygulamaları. (Yüksek Lisans Tezi), Dokuz Eylül Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Bilimleri Anabilim Dalı, İzmir.

Gomes, R.P. Ribeiro, V.G. ve Silveira, S.R. (2021). Educação não-formal online em design:uma análise conceitual dos cursos de Design Gráfico da plataforma Udemy. DAPesquisa - Publicação contínua, 16, 01-22

González Videgaray, M. Ruiz, R.R. (2020). Metodología de la Investigación y Proyectos: MOOC en Udemy con cerca de 15 000 alumnos. Alfa Omega Grupo Editor, (50-59). Ciudad de México.

Google Trends (2022). Udemy Zaman İçinde Gösterilen İlgi. 31.07.2022 tarihinde <https://trends.google.com/trends/explore?date=2012-12-24%202022-07-31veq=udemy> adresinden erişilmiştir.

Güneş, F. 2020, Kitleli Açık Çevrimiçi Ders Platformlarında Yayınlanan Siber Güvenlik İçeriklerinin Çoklu Ortam Tasarım İlkeleri Açısından İncelenmesi. (Yüksek Lisans Tezi), Gazi Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.

Gürbüz, I. (2022). Metin Madenciliği Yöntemi ile Araştırma Makalesi Sınıflandırması.

- (Yüksek Lisans Tezi), Gazi Üniversitesi Bilişim Sistemleri Ana Bilim Dalı, Ankara.
- Gürtürk, U. (2022). Türkiye'nin Siber Güvenlik Politikalarının Yazılım Mühendisliği Açısından Değerlendirilmesi ve Kritik Altyapıların Siber Saldırılarından Korunmasına Yönelik Olay Yönetim Sistemi Tasarım. (Yüksek Lisans Tezi), İstanbul Üniversitesi-Cerrahpaşa Lisansüstü Eğitim Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı, İstanbul.
- Hill, P. (2012). Four Barriers That MOOCs Must Overcome To Build a Sustainable Model. 10.09.2022 tarihinde <https://eliterate.us/four-barriers-that-moocs-must-overcome-to-become-sustainable-model/> adresinden erişilmiştir.
- Hochreiter, S. ve Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, vol. 9, no. 8, pp. 1735-1780, 15 Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.
- Hollands, F.M. ve Tirthali, D. (2014). MOOCs: Expectations and reality. Center for Benefit-Cost Studies of Education, Teachers College, Columbia University, NY.
- Holte, R.C. (1993) Very simple classification rules perform well on most commonly used data sets. *Machine Learning*, 11(1), 63-90.
- Impey, C. ve Formanek, M. (2021). MOOCS and 100 Days of COVID: Enrollment surges in massive open online astronomy classes during the coronavirus pandemic. *Social Sciences ve Humanities Open*, 4(1), 100177.
- Işık, N. (2019). Metin Madenciliği Yöntemleri ile E-Ticaret Markalarına Yönelik Sosyal Medya Yorumlarının Analizi. (Yüksek Lisans Tezi), Marmara Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul.
- İşgör Şimşek, E. 2015, Mobil Ortamlarda Kitlesele Açık Çevrimiçi Derslerin (Kaçd) Kullanılabilirliğinin Değerlendirilmesi. (Yüksek Lisans Tezi), Mimar Sinan Güzel Sanatlar Üniversitesi Fen Bilimleri Enstitüsü, İstanbul.
- Kansal, A.K. Gautam, J. Chintalapudi, N. Jain, S. ve Battineni, G. Google Trend Analysis and Paradigm Shift of Online Education Platforms during the COVID-19 Pandemic. *Infectious Disease Reports*, 13(2), 418-428.
- Karakuş, L. (2021). Eğitimde Metin Madenciliği: Türkçe Metinlerde Sözlük Tabanlı Duygu Analizi. (Yüksek Lisans Tezi), Akdeniz Üniversitesi Eğitim Bilimleri Enstitüsü Eğitim Bilimleri Ana Bilim Dalı, Antalya.

- Kayaalp, K. ve Süzen, A. (2018). Derin Öğrenme ve Türkiye'deki Uygulamaları. Iksad International Publishing House.
- Kaynar, O. Arslan, H. Görmez, Y. ve Işık, Y. E. (2018). Makine Öğrenmesi ve Öznelik Seçim Yöntemleriyle Saldırı Tespiti. Bilişim Teknolojileri Dergisi, 11 (2), 175-185. DOI: 10.17671/gazibtd.368583
- Keskin, M. (2018). Word2Vec, FastText, GloVe. 16.01.2023 tarihinde <https://medium.com/codable/word2vec-fasttext-glove-d4402fa8cce0> adresinden erişilmiştir.
- Keskin, C. (2022). Graf Tabanlı Yaklaşım ile Uzun Vadeli Trafik Akış Hızı Tahmini. (Yüksek Lisans Tezi), Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı, İstanbul.
- Kışla, T. Kumova Metin, S. Karaoğlu, B. ve Demir, E.K. (2021) Examining the Views of Engineering Students about Massive Open Online Courses. 2021 30th Annual Conference of the European Association for Education in Electrical and Information Engineering (EAEIE), Prague, Czech Republic. 2021
- Koca, B. (2022). Bilişim Teknolojileri Öğretmenlerinin Scratch Yazılımına İlişkin Özyeterlik İnançlarının Makine Öğrenmesi ve Derin Öğrenme Yöntemleri ile Sınıflandırılması. (Yüksek Lisans Tezi), Aksaray Üniversitesi Sosyal Bilimler Enstitüsü Yönetim Bilişim Sistemleri Anabilim Dalı, Aksaray.
- Koglak (2020). NLP: GloVe Nedir? 23.01.2023 tarihinde <https://koglakk.medium.com/nlp-glove-nedi%CC%87r-e21a1a0c57f5> adresinden erişilmiştir.
- Koglak (2020). NLP: Word2Vec Nedir? 23.01.2023 tarihinde <https://koglakk.medium.com/nlp-word2vec-nedi%CC%87r-594e093cd9ff> adresinden erişilmiştir.
- Labidi, A. Ouamani, F. ve Ben Saoud, N.B. (2021). An ontology based text approach for culture aware emotion mining: A Moodle plugin. 2021 International Conference of Women in Data Science at Taif University (WiDSTaif), Taif, Saudi Arabia.
- Lecun, Y., Bengio, Y., Hinton, G. (2015). Deep learning. Nature, 521(7553): 436–444.
- Mahmat, Z. (2022). Müşteri Şikayetlerinin Sınıflandırılması İçin Derin Öğrenme ve Makine Öğrenmesi Uygulamaları. (Yüksek Lisans Tezi), Atatürk Üniversitesi Fen Bilimleri Enstitüsü Endüstri Mühendisliği Ana Bilim Dalı, Erzurum.

- Mashudi, N.A. Mohd Izhar, M.A. ve Mohd Aris, S.A. (2022). Human-Computer Interaction in Mobile Learning: A Review. *International Journal of Advanced Computer Science and Applications (IJACSA)*, 13(3), 566-574.
- Mikolov, T., Chen, K., Corrado, G., Jeffrey, D. (2013). Efficient Estimation of Word Representations in Vector Space. *arXiv:1301.3781*.
<https://doi.org/10.48550/arXiv.1301.3781>
- Nergiz, G. (2021). Derin Öğrenme Teknikleri Kullanılarak Sosyal Ağlar Üzerinde Siber Zorbalık Tespiti. (Yüksek Lisans Tezi), Mersin Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı, Mersin.
- Omer, M. B. (2021). A Method For Predicting Title Of Given Text. (Yüksek Lisans Tezi), Istanbul Commerce University Graduate School Of Natural and Applied Sciences, İstanbul.
- Onan, A. (2020). Sentiment analysis on massive open online course evaluations: A text mining and deep learning approach. *Special Issue Article* 29(3), 572-589.
- Özbilen, M. L. (2022). Hisse Senedi Fiyat Tahmininde Otokodlayıcı ve Graf Evrişimli Ağının Uygulanması. (Yüksek Lisans Tezi), İstanbul Teknik Üniversitesi Lisansüstü Eğitim Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı, İstanbul
- Özdemir, E. (2020). Basketbol Maç Sonuçlarının Makine Öğrenmesi Yöntemleri ile Tahmin Edilmesi. (Yüksek Lisans Tezi), Muğla Sıtkı Koçman Üniversitesi Fen Bilimleri Enstitüsü Bilişim Sistemleri Mühendisliği Anabilim Dalı, Muğla.
- Özdemir, M. (2021). Derin Sinir Ağı Yöntemleri ile Wikipedia Toksik Yorum Sınıflandırması. (Yüksek Lisans Tezi), Giresun Üniversitesi Fen Bilimleri Enstitüsü İstatistik Anabilim Dalı, Giresun.
- Özden, C. (2018). The Effects Of Preprocessing Methods On Prediction Of Traffic Accident Severity. (Yüksek Lisans Tezi), Çukurova University Institute Of Natural And Applied Sciences Department Of Computer Engineering, Adana.
- Pennington, J., Socher, R., Manning, C. D. (2014). GloVe: Global Vectors for Word Representation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1532–1543, October 25-29, 2014, Doha, Qatar.
- Puspitarini, E.W. Maukar, A. Marisa, F. Haryanto, K.W. Pradana, T. (2021). Business

Models Canvas of MOOCs: an Investigation of Sustainable Practices for MOOC Universities. *Jurnal Ilmiah Bidang Teknologi Informasi dan Komunikasi*, 6(2), 133-137.

Qiu, R. (2020). *Udemy: Blended and e-Learning for Transforming Teaching and Learning*. In: Panth, B., Maclean, R. (eds) *Anticipating and Preparing for Emerging Skills and Jobs. Education in the Asia-Pacific Region: Issues, Concerns and Prospects*, vol 55. Springer, Singapore.

Ramneet, Gupta, D. Madhukar, M. (2020). *Bibliometric Analysis of MOOC using Bibliometrix Package of R*. 2020 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE), Bhubaneswar, India

Quinlan, J. R. (2014). *C4. 5: programs for machine learning*. Elsevier.

Selmanovic, D. Sayar, A. ve Durdu, P. O. (2021) *Cross cultural usability testing of MOOC platform*. 2021 5th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), Ankara, Turkey.

Seyidbayli, C. (2022). *Transfer Öğrenmesi Tabanlı Hibrit Evrişimsel Sinir Ağı Modelleri Kullanılarak Meme Kanseri Teşhisi*. Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü Mekatronik Mühendisliği Anabilim Dalı, İstanbul.

Shah, D. (2021). 157K Courses, 425M Enrollments: Breaking Down Udemy's Massive Catalog. 08.09.2022 tarihinde <https://www.classcentral.com/report/udemy-by-the-numbers/> adresinden erişilmiştir.

Shah, D. (2021). *By The Numbers: MOOCs in 2021*. 07.09.2022 tarihinde <https://www.classcentral.com/report/mooc-stats-2021/> adresinden erişilmiştir.

Sindhu, I. Muhammad Daudpota, S. Badar, K. Bakhtyar, M. Baber, J. ve Nurunnabi, M. (2019). *Aspect-Based Opinion Mining on Student's Feedback for Faculty Teaching Performance Evaluation*. *IEEE Access*, 7, 108729-108741

Şahin, E. ve Durdu, P. O. (2021). *Bilişsel Gezinti ile Kitlesel Açık Çevrimiçi Ders (KAÇD) Web Sitelerinin Kullanılabilirliğinin Değerlendirilmesi*, *Bilişim Teknolojileri Dergisi*, 14(4), 377-389.

Tavacı, H. (2011). *GSM şebekelerinde sahtekarlık yönetimi için veri madenciliği yöntemlerinin uygulanması*. (Yüksek Lisans Tezi), Bahçeşehir Üniversitesi Fen

Bilimleri Enstitüsü Bilgi Teknolojileri Ana Bilim Dalı, İstanbul.

- Tekinarslan, R. (2022). Çevrimiçi Öğrenme Ortamlarında Öğrenme Analitiği Verileri ve Makine Öğrenmesi Kullanarak Akademik Başarının Değerlendirilmesi. (Yüksek Lisans Tezi), Başkent Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı, Ankara.
- Telli, S.G. Altun, D. (2021). Coronavirus (Covid-19) Pandemisi Döneminde Çevrimiçi Öğrenme. Üniversite Araştırmaları Dergisi, 4(2), 90-107.
- Tozyılmaz, B. (2019). Abstractive Text Summarization On Wikihow Dataset Using Sentence Embeddings. (Yüksek Lisans Tezi), The Graduate School Of Natural And Applied Sciences Of Middle East Technical University, Ankara.
- Udemy (2022). About Udemy. 10.09.2022 tarihinde <https://about.udemy.com/> adresinden erişilmiştir.
- Yaman, U. C. (2022). Metin Madenciliği Teknikleri ile Türkçe Müşteri Yorumlarının Analizi. (Yüksek Lisans Tezi), Eskişehir Teknik Üniversitesi Endüstri Mühendisliği Anabilim Dalı, Eskişehir.
- Yaşar, H. (2022). Covid-19 Hastalığının Derin Öğrenme Yöntemleri Kullanılarak Tespiti. (Doktora Tezi), Konya Teknik Üniversitesi Lisansüstü Eğitim Enstitüsü Elektrik-Elektronik Mühendisliği Anabilim Dalı, Konya.
- Yıldız, E. (2020). LSTM ve Metin Sınıflandırması. 14.01.2023 tarihinde <https://medium.com/@iamemreyildiz/lstm-ve-metin-s%C4%B1n%C4%B1fland%C4%B1rmas%C4%B1-c1a48eb53e01> adresinden erişilmiştir.
- Yılmaz, M. C. ve Orman, Z. (2021). LSTM Derin Öğrenme Yaklaşımı ile Covid-19 Pandemi Sürecinde Twitter Verilerinden Duygu Analizi. Acta Infologica , 5 (2) , 359-372.
- Yuan, L. ve Powell, S. (2013). MOOCs and Open Education: Implications for Higher Education. <http://publications.cetis.ac.uk/2013/667> sayfasından erişilmiştir.

EKLER



Aksaray Üniversitesi Tez Değerlendirme Formu		
Öğrencinin Adı Soyadı:		EVET
Kapak		
1	Tez Başlığı tutanaktaki başlıkla aynı mı?	
2	Kapaktaki ay ve yıl savunmaya girilen tarihle tutarlı mı?	
3	Kapak format kılavuzdaki kapak formatına uygun mu?	
4	Kapakta yazılan tüm yazılar doğru olarak verilmiş mi?	
İçindekiler		
5	Sayfa numaraları tam verilmiş mi?	
6	Şekil, Çizelge vb. listeleri verilmiş mi? Sıralaması doğru mu?	
7	Özet, Abstract, Giriş, Sonuçlar vb. bölümler var mı?	
8	Yazım hataları kontrolü yapıldı mı?	
Giriş		
9	Hazırlanan tezin önemini anlatıyor mu?	
10	İkinci ve Üçüncü dereceden başlık içermemeli kuralına uyuldu mu?	
Özet/Abstract		
11	Kılavuza uygun mu?	
12	Ay ve yıl savunmaya girilen tarihle tutarlı mı?	
13	Özet; tek sayfa, tek aralık, tek paragraf kuralına uygun olarak yazıldı mı?	
14	Bilim kodu, sayfa adedi, anahtar kelimeler ve tez danışmanı yazıldı mı?	
Kaynakça		
15	Kaynakların tamamına metin içinde atıf yapıldı mı?	
16	Kaynak formatı Kılavuzdaki kaynak formatına uygun olarak hazırlanmış mı?	
17	Atıf formatı kılavuzdaki atıf formatına uygun mu?	
Atıf Yöntemi APA 6 ☐ CMS ☐ İSNAD ☐		
Genel Değerlendirme		
18	Etik Beyan açıklaması okundu, uyuldu ve imzalandı mı?	
19	Kabul/Onay sayfası kılavuzdaki formata uygun olarak düzenlenmiş mi?	
20	Kabul /Onay sayfasında belirtilen oy birliği/oy çokluğu seçeneklerinden uygun olanı savunmayla tutarlı olacak şekilde belirlenmiş mi?	
21	Sayfa kenar boşluklar ve sayfa numaraları kılavuzdaki formatına uygun mu?	
22	Paragraf boşlukları ve metin satır aralığı kılavuza uygun olacak şekilde düzenlenmiş mi?	
23	Başlık yazımları kılavuzdaki başlık formatlarına uygun mu?	
24	Yazı tipi ve boyutu kılavuzdaki yazı tipi ve boyutu formatına uygun mu?	
25	Şekil, Çizelge vb. açıklama ve numaralandırmaları kılavuzdaki formata uygun olarak yazılmış mı?	

Bu tezin tarafımdan “Tez yazım kuralları” okunarak dikkatlice hazırlanmış olduğunu ve doğabilecek her türlü olumsuzluktan sorumlu olacağımı kabul ederim.

Öğrencinin imzası