



İSKENDERUN TEKNİK

ÜNİVERSİTESİ

LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

**YÜKSEK
LİSANS
TEZİ**

**YAPAY ZEKA YÖNTEMLERİNİN
DOĞAL DİL İŞLEMEDE ETKİN
KULLANILMASI**

Yasemin ATAYOLU

**BİLGİSAYAR MÜHENDİSLİĞİ
ANABİLİM DALI**

KASIM 2024





**YAPAY ZEKA YÖNTEMLERİNİN DOĞAL DİL İŞLEMEDE ETKİN
KULLANILMASI**

YASEMİN ATAYOLU

**YÜKSEK LİSANS TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI**

**İSKENDERUN TEKNİK ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜ**

KASIM 2024

Yasemin ATAYOLU tarafından hazırlanan “YAPAY ZEKA YÖNTEMLERİNİN DOĞAL DİL İŞLEMEDE ETKİN KULLANILMASI” adlı tez çalışması aşağıdaki jüri tarafından OY BİRLİĞİ ile İskenderun Teknik Üniversitesi Bilgisayar Mühendisliği Anabilim Dalında YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

Danışman: Prof. Dr. Yakup KUTLU

Bilgisayar Mühendisliği Anabilim Dalı, İskenderun Teknik Üniversitesi

Bu tezin, kapsam ve kalite olarak Yüksek Lisans Tezi olduğunu

.....
.....

Başkan: Prof. Dr. Sertan ALKAN

Bilgisayar Mühendisliği Anabilim Dalı, İskenderun Teknik Üniversitesi

Bu tezin, kapsam ve kalite olarak Yüksek Lisans Tezi olduğunu

.....
.....

Üye: Dr. Öğr. Üyesi Mustafa YENİAD

Bilgisayar Mühendisliği Anabilim Dalı, Ankara Yıldırım Beyazıt Üniversitesi

Bu tezin, kapsam ve kalite olarak Yüksek Lisans Tezi olduğunu

.....
.....

Tez Savunma Tarihi: 20/11/2024

Jüri tarafından kabul edilen bu tezin Yüksek Lisans Tezi olması için gerekli şartları yerine getirdiğini onaylıyorum.

.....

Doç. Dr. Ersin BAHÇECİ
Lisansüstü Eğitim Enstitüsü Müdürü

ETİK BEYAN

İskenderun Teknik Üniversitesi Lisansüstü Eğitim Enstitüsü Tez Yazım Kurallarına uygun olarak hazırladığım bu tez çalışmada;

- Yükseköğretim Kuruluna gönderilen kopya ile tarafından Lisansüstü Eğitim Enstitüsü'ne verilen basılı ve/veya elektronik kopyaların birebir aynı olduğunu,
- Tez içinde sunduğum verileri, bilgileri ve dokümanları akademik ve etik kurallar çerçevesinde elde ettiğimi,
- Tüm bilgi, belge, değerlendirme ve sonuçları bilimsel etik ve ahlak kurallarına uygun olarak sunduğumu,
- Tez çalışmada yararlandığım eserlerin tümüne uygun atıfta bulunarak kaynak gösterdiğimi,
- Kullanılan verilerde herhangi bir değişiklik yapmadığımı,
- Bu tezde sunduğum çalışmanın özgün olduğunu, bildirir, aksi bir durumda aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi beyan ederim.

.../.../2024

YAPAY ZEKA YÖNTEMLERİNİN DOĞAL DİL İŞLEMEDE ETKİN KULLANILMASI

(Yüksek Lisans Tezi)

Yasemin ATAYOLU

İSKENDERUN TEKNİK ÜNİVERSİTESİ

LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

Kasım 2024

ÖZET

Dilin soyut ve dinamik doğası, dili yapay zekâ ile modellememizde belli sınırlılıklara yol açmaktadır. Bu sınırlılıklarını en aza indirmek için dili ve bağlamını çok boyutlu yapılar olarak ele almak ve modellemede veriye özgü yöntem stratejisi belirlemek gereklidir. Bu tez çalışmasında, yapay zekâların içerik üretme performansı, içeriklerin analizi ve içeriklerin sınıflandırılması ele alınmıştır. Yapay zekâların performansı eğitim bilimleri kapsamında ele alınmış ve içerik üretme kapasitelerinin etkili istem tasarımı ile nasıl geliştirilebileceği üzerine odaklanılmıştır. Eğitim bilimlerinde Bloom Taksonomisine uygun olarak soru tasarımı gerçekleştirilmiş ve taksonomiye uygun soru hazırlamada rehberlik eden referans noktalar özenle dikkate alınmıştır. İçeriklerin analizi için Sinema filmlerine ait içerikler film puanı ve türünün makine öğrenimi modelleriyle tahmini için kullanılmıştır. Bu bağlamda öznitelik seçimine ve girdi oluşturulmasına değinilmiş ve belirli özniteliklerle film puanı için regresyon modelleriyle tahmin yapılmıştır. Film özetlerini içeren tür sınıflandırılması ele alınmıştır. Günümüz önemli mental sağlık sorunları arasında olan anksiyete, depresyon ve intihar eğilimlerinin ilgili sosyal medya gönderileri ile belirlenmesi mümkün hale gelmiştir. Sosyal medya gönderilerinin sınıflandırılması için metin ön işleme tekniklerinin ve hiper parametre yöntemlerin başarıma etkileri ele alınmıştır. Diğer taraftan metin vektörel temsilleri için farklı yöntemler ele alınmış ve sınıflandırma performansı incelenmiştir. Sadece tekrarlayan ifadeleri kaldırarak uygulanan ön işlemde sonra yapılan Roberta modeli ile %87,2 başarımla elde edilmiştir. Bi-LSTM modeli eğitiminde ön eğitilmiş Roberta, word2vec ve TF-IDF vektörlerinin birleştirilmesiyle oluşturulmuş hibrit vektör sınıflandırma başarımlarını %13 kadar artırdığı görülmüştür. Yapay zekâ yöntemlerinin doğal dil işlemede etkin kullanılması için üç farklı problem farklı yöntem ve tekniklerinin değerlendirilmesi yapılmıştır. Yapılan uygulamalarda görüldüğü üzere gerek metin içeriklerinin çeşitliliği gerek ön işlemlerin sırası ve yapılması ve gerekse hiper parametre yöntemlerinin uygulanması başarımları ciddi şekilde etkilediği görülmüştür.

Anahtar Kelimeler : Metin sınıflandırma, Bloom Taksonomisi, Mental Sağlık, Sinema İçerikleri, Yapay Zekâ, Doğal Dil İşleme
Sayfa Adedi : 69
Danışman : Prof. Dr. Yakup KUTLU

EFFECTIVE USE OF ARTIFICIAL INTELLIGENCE METHODS IN NATURAL
LANGUAGE PROCESSING
(M. Sc. Thesis)

Yasemin ATAYOLU

ISKENDERUN TECHNICAL UNIVERSITY
INSTITUTE OF GRADUATE STUDIES

August 2024

ABSTRACT

The abstract and dynamic nature of language lead to certain limitations in modelling language with artificial intelligence. To minimize these limitations, it is essential to approach language and context as multidimensional structures and adopt data-specific methodological strategies for modelling. This thesis focuses on the content generation performance of artificial intelligence systems, the analysis of content, and content classification. The performance of artificial intelligence was evaluated within the scope of educational sciences, focusing on how their content generation capacities can be enhanced through effective prompt design. Questions were designed in accordance with Bloom's Taxonomy in educational sciences, and reference points guiding the preparation of taxonomy-compliant questions were carefully considered. For content analysis, the content of films was used to predict movie ratings and genres through machine learning models. Therefore, feature selection and input formation were addressed, and regression models were applied to predict movie ratings using selected features. Additionally, genre classification based on film summaries was explored. Contemporary mental health issues such as anxiety, depression, and suicidal tendencies have become identifiable through related social media posts. The impact of text preprocessing techniques and hyperparameter optimization methods on classification performance was examined. Furthermore, different methods for text vector representations were analyzed, and classification performance was evaluated. A Roberta model, applied after preprocessing by removing repetitive expressions, achieved 87.2% accuracy. It was observed that hybrid vector classification, created by combining pre-trained Roberta, word2vec, and TF-IDF vectors in training the Bi-LSTM model, increased classification accuracy by approximately 13%. Three different problems were evaluated using various methods and techniques for the effective use of artificial intelligence methods in natural language processing. The applications revealed that the diversity of textual content, the sequence and execution of preprocessing steps, and the implementation of hyperparameter methods significantly influence performance.

Key Words : Text Classification, Bloom's Taxonomy, Mental Health, Cinema
Content, Artificial Intelligence, Natural Language Processing
Page Number : 69
Supervisor : Prof. Dr. Yakup KUTLU

TEŐEKKÜR

Yüksek Lisans tez konusunun belirlenmesi, araştırılması ve yazımı sırasında sahip olduđu bilgi birikimi ve tecrübesi ile çalışmayı yönlendiren ve hiçbir yardımı esirgemeyen saygıdeđer danışman hocam Prof. Dr. Yakup KUTLU'ya sonsuz saygı ve teşekkürlerimi sunarım.

Bu tez çalışması ve yüksek lisans eğitimimin her aşamasında her türlü desteđi, fedakârlığı esirgemeyen anneme ve kardeşlerime, yanımda olan arkadaşlarıma, bu alanda gelişmemi sağlayan yüksek lisans derslerimin hocalarına, alanın erişilebilir kaynaklarını zenginleştirerek bilgi ve becerilerimi geliştirmeme yardımcı olan tüm emekçilere sonsuz teşekkürlerimi sunarım.

İÇİNDEKİLER

	Sayfa
ÖZET	iv
ABSTRACT.....	v
TEŞEKKÜR.....	vi
İÇİNDEKİLER	vii
ÇİZELGELERİN LİSTESİ.....	ix
ŞEKİLLERİN LİSTESİ.....	x
SİMGELER VE KISALTMALAR.....	xi
1. GİRİŞ.....	1
2. ÖNCEKİ ÇALIŞMALAR.....	7
3. MATERYAL VE YÖNTEM	12
3.1. Metin Önişleme Teknikleri	12
3.2. Yapay Zekâ Modelleri	13
3.3. Performans ölçümü	19
4. ARAŞTIRMA BULGULARI.....	21
4.1. Verilerin Elde Edilmesi	21
4.1.1. Bloom Taksonomisi soruları	21
4.1.2. Sinema filmi içerik verileri	22
4.1.3. Mental sağlık açısından sosyal medya gönderileri	26
4.2. Öznitelik elde edilmesi	29
4.3. Yapay Zekâ Modellerinin Performans Analizi.....	31
4.3.1. Bloom Taksonomisine göre soru üretimi ve sınıflandırması	31
4.3.2. Sinema filmi içeriklerinin analizi.....	32
4.3.3. Mental sağlık açısından sosyal medya gönderilerinin sınıflandırması	33

5. SONUÇ VE ÖNERİLER.....	45
KAYNAKLAR.....	47



ÇİZELGELERİN LİSTESİ

Çizelge	Sayfa
Çizelge 3.1. Ön işlem teknikleri ve kullanım açıklamaları.....	12
Çizelge 3.2. Büyük Dil Modelleri Karşılaştırması	18
Çizelge 4.1. Bloom taksonomisi soru dağılımı.....	21
Çizelge 4.2. Çevrim içi veri tabanları ve ilgili sinema filmleri içerikleri	23
Çizelge 4.3. Ait olduğu tür sayılarına göre film sayısının dağılımı	25
Çizelge 4.4. Film türü sınıflandırması için eğitim, doğrulama ve test sayıları.....	26
Çizelge 4.5. Mental sağlık gönderilerinin sınıflara göre veri istatistikleri	28
Çizelge 4.6. Üretilen yapay sorular ve uzman soruları sınıflandırmalarının F1 Skoru Sonuçları	32
Çizelge 4.7. Regresyon analizi sonuçları.....	33
Çizelge 4.8. Tür eşleştirmesi skorları	33
Çizelge 4.9. LLM ince ayar eğitimlerinde kullanılan parametreler.....	36
Çizelge 4.10. Büyük dil modellerinin sınıflandırma skorları	36
Çizelge 4.11. Farklı önerme tekniklerinin Roberta ince ayar sonuçları	39
Çizelge 4.12. Destek vektörleri sınıflandırma sonuçları	41
Çizelge 4.13. Bi-LSTM modellerinde kullanılan parametreler	43
Çizelge 4.14. Bi-LSTM sınıflandırma sonuçları.....	43

ŞEKİLLERİN LİSTESİ

Şekil	Sayfa
Şekil 1.1. Eliza diyalogu için script	2
Şekil 1.2. Örnek Eliza diyalogu	2
Şekil 3.1. Transformatör mimarisi	15
Şekil 4.1. Sinema filmleri IMDb puan dağılımı	24
Şekil 4.2. Sinema filmleri IMDb tür dağılımı.....	26
Şekil 4.3. Büyük dil modellerinin sınıflandırma performansı	37
Şekil 4.4. Metin ön işlem tekniklerinin sınıflara göre performans analizi.....	39
Şekil 4.5. Destek vektörlerinin farklı vektör alma yöntemleri ile sonuçları	41
Şekil 4.6. Bi-LSTM modelinin farklı vektör alma yöntemleri ile sonuçları.....	44

SİMGELER VE KISALTMALAR

Bu çalışmada kullanılmış simgeler ve kısaltmalar, açıklamaları ile birlikte aşağıda sunulmuştur.

Kısaltmalar

Açıklamalar

AI	Artificial Intelligence
ANN	Artificial Neural Network
ATN	Augmented Transition Networks
BP	Backpropagation
CNN	Convolutional Neural Network
DL	Deep Learning
FNN	Feedforward Neural Network
GPU	Graphics Processing Unit
HMM	Hidden Markov Models
LISP	List Processing
LSTM	Long Short-Term Memory
MIT	Massachusetts Institute of Technology
ReLU	Rectified Linear Unit
RL	Reinforcement Learning
RNN	Recurrent Neural Network
SL	Supervised Learning
SVM	Support Vector Machine
UL	Unsupervised Learning

1. GİRİŞ

Soru sormanın nedeni bilgi edinmek olsa da doğru cevapları alabilmek için soruları formüle etme ihtiyacı duyarız (Van Rooy, 2003; Jordania, 2006). Bu çeşitli makinelerin ortaya çıkmasından sonra makinelerin yapabildiklerine dahil etmek istediklerimizi nasıl gerçekleştirebileceğimiz konusunda da geçerliydi. Makinelere yöneltebileceğimiz soru yapısının “hesaplanabilirlik” temeline dayanması gerektiğini Turing ortaya koymuştur (1936). Alan Turing, zekanın bir ölçütü olarak daha sonra Turing testi olarak adlandırılacak yöntemi öneren ünlü makalesinde makinelere nasıl soru sormamız gerektiği ortaya koyarak makine zekasının temellerini atmıştır. “Soru-cevap yöntemi, dahil etmek istediğimiz insan çabasının neredeyse her alanını tanıtmak için uygun görünüyor.” (Turing, 1950:435)

McCulloch ve Pitts (1943), yapay sinir ağlarının teorik temellerini attıkları bir çalışma yayımlayarak biyolojik sinir hücrelerini temel alan matematiksel bir model önermiş ve bu hücrelerin belirli mantıksal işlemleri gerçekleştirebileceği öne sürmüşlerdir. Hebb (1949) yayımladığı kitabında, 'birlikte ateşlenen hücreler birlikte bağlanır' prensibiyle, öğrenmenin sinir hücreleri arasındaki eşzamanlı ve tekrar eden uyarılmalarla bağlantıların güçlenmesi yoluyla gerçekleştiğini öne sürmüştür. Cornell Havacılık Laboratuvarında Frank Rosenblatt tarafından geliştirilen ve girdiler ile çıktılar arasındaki ilişkiyi öğrenerek örüntü tanıma ve sınıflandırma yapmak amacıyla tasarlanmış Perceptron (Rosenblatt, 1957), biyolojik sinir ağlarından esinlenen ilk donanımsal yapay sinir ağıdır. Rosenblatt'ın perceptron teorisi, öğrenme ve genelleme süreçlerini, ölçülebilir fiziksel parametrelerle tahmin ederken, Hebb'in teorisi ise sinir hücrelerinin yakın zamanlı ateşlemesiyle bağlantıların güçlenmesine dayanıyordu (Rosenblatt, 1958). 1959 yılında Samuel makine öğrenimini, “makinelere açıkça programlanmadan öğrenmesini sağlayan algoritmalar” olarak tanımlamış ve dama oyunu gibi karmaşık oyunlarda makinelerin performansını artıracak otomatik öğrenme sistemleri üzerinde çalışmıştır (Samuel, 1959).

Chomsky'nin dil teorileri dilbilim ve doğal dil işleme (DDİ) alanını önemli ölçüde etkilemiş ve dilin gramer yapısı üzerine yoğunlaşan ilk nesil DDİ sistemlerinin temelini oluşturmuştur. Chomsky'nin sözdizimsel analiz (Chomsky, 1956;1957) ve biçimsel dilbilgisi (Chomsky, 1956; 1959) üzerine yaptığı çalışmalar dilin matematiksel ve hesaplamalı modellenmesinin araştırmalarının başlangıç noktası olmuştur (Joshi, 1991). Dil yapısının kural tabanlı bir

sistemle anlaşılabilirliğini öne süren Chomsky'nin çalışmaları, dilin hiyerarşik yapısını, matematiksel temsillerini, hesaplamalı dilbilimi ve dönüşümsel dilbilgisini açıklamaya yönelik katkılar sağlamıştır.

Doğal dil işleme insanların birbirlerine anlama ihtiyacından kaynaklı olarak diller arası makine çevirisi çalışmalarıyla başlamıştır. 1960'ların sonlarına kadar devam eden süreç, DDİ araştırmalarının başladığı ve makine çevirisine odaklanıldığı bir dönemdir. Bu dönemde dilin karmaşıklığı nedeniyle birçok problem tam olarak çözülememiş olsa da DDİ'nin temelini oluşturan ilkeler ve stratejiler geliştirilmiş ve dil işleme gereksinimleri belirlenmiştir (Jones, 1994).

ELIZA, 1964-1967 yılları arasında Joseph Weizenbaum tarafından geliştirilen doğal dil ile etkileşime giren ilk bilgisayar programıdır ve girdileri anlamadan sözdizimsel olarak dönüştüren bir çeviri işlemcisi olarak çalışmaktadır (Weizenbaum, 1966). ELIZA en iyi performansını Rogers'ın danışan odaklı terapi yöntemini (Rogers, 1951) uygulayan bir psikoterapist gibi etkileşimde bulunması gerektiğinde göstermiştir. Burada görev konuşanın girdilerine yönelik karar vermek değildir. Ancak verdiği yanıt konuşmacının ELIZA'nın kendisini anladığı yanılgısına düşmesine sebep olabilmektedir. Bu literatüre 'Eliza etkisi' olarak geçmiştir ve bu tanım insanların kendilerine yanıt veren programların olduğundan daha zeki algılamaları fenomeni için kullanılmıştır (Trapp ve Payr, 2003). Eliza'nın hazır olarak kullandığı konuşma kalıpları (script) Şekil 1.1.'de gösterilmiş ve Şekil 2.2.'de bu kalıpların kullanılarak nasıl bir diyalog kurulduğuna örnek verilmiştir.

((0 YOU ARE O (-SAD UNHAPPY *DEPRESSED* SICK) 0)
 (I AM SORRY TO HEAR YOU ARE 5) (DO YOU THINK COMING HERE WILL
 HELP YOU NOT TO BE 5) (I AM SURE ITS NOT PLEASANT TO BE 5) (CAN YOU
EXPLAIN WHAT MADE YOU 5))

Şekil 1.1. Eliza diyalogu için script (Weizenbaum, 1966:44)

Kullanıcı: "I am feeling *depressed*"
 ELIZA: "I am sorry to hear you are *depressed*. Can you explain what made you
depressed?"

Şekil 1.2 Örnek Eliza diyalogu

SHRDLU, 1968-1970 yılları arasında Terry Winograd tarafından geliştirilen, kullanıcıdan gelen doğal dil komutlarını anlayan ve bir sanal blok dünyasında işlemler gerçekleştiren bir sistemdir. Kısıtlı kelime dağarcığına sahip program, konuşmayı anlamadığında veya cevabı bilmediğinde bunu belirtiyordu. Komutun dilbilimsel analizi yapılması ve uygun aksiyon alması üzerine kurulmuştu (Winograd, 1972). ELIZA ve SHRDLU Massachusetts Teknoloji Enstitüsünde (MIT) geliştirilmiştir.

Woods tarafından Genişletilmiş Geçiş Ağları (ATN) dönüşümsel gramer işlevlerini ayrı bir dönüşüm bileşeni olmadan gerçekleştiren ve karmaşık dil yapılarını işleyebilen özyinelemeli modeller olarak geliştirilmiştir (Woods,1970). LUNAR, ay kayaçlarıyla ilgili soruları yanıtlamak için NASA tarafından geliştirilmiş bir sistemdi (Woods, 1972) ve ATN ayrıştırıcısını ve gramer formülasyonunu temel alıyordu (Woods, 1973). SHRDLU ve LUNAR, doğal dil işleme uygulamaları için kullanılabilen fonksiyonel bir programlama dili olan LIPS (McCarthy, 1960) ile programlanmıştır.

Quillian'ın kavramlar arasındaki genel ilişkileri yakalayarak insan bilgisinin bir dil temsilini oluşturma üzerine çalışmaları semantik/anlamsal ağların (semantic networks) temelini atmıştır (Quillian, 1966; Collins ve Quillian, 1969). Brachman semantik ağların bilgi temsilindeki rolünü, bu ilişkileri açık epistemolojik bağlantılarla nasıl hesaba katabileceğini ortaya koymuştur ve düğüm ve bağlantılar üzerine kurulu bir temsilin nasıl bir yorumlayıcı oluşturabileceğini göstermiştir (Brachman, 1979). Devam eden çalışmalarda kalıtım ağları (inheritance networks) geliştirilmiş ve dilbilgisi temsili, çıkarım mekanizmalarında miras hiyerarşileri ile bütünleşmiş bir yapı sunarak, dil işleme alanında daha derin ve anlamlı analizlerin yolu açılmıştır (Brachman ve Schmolze, 1989).

Yapay sinir ağları (ANN), biyolojik beyin sinir ağlarının işleyişinden esinlenen ve makinelerin öğrenimini sağlayan modellerdir. Rosenblatt ilk donanımsal yapay sinir ağı uygulaması olan Perceptron makinesini (Rosenblatt, 1957; 1958) ile yaptığı deneyleri 1962 tarihli kitabında anlatmıştır (Rosenblatt, 1962). Minsky ve Papert (1969) yayınladıkları kitapta (Wikipedia, 2024) Rosenblatt'ın Perceptron'unun yalnızca çok temel görevleri (lineer) yerine getirebileceğini örneğin, bir XOR fonksiyonunu öğrenmesinin imkânsız olduğunu matematiksel olarak göstermişler ve çok katmanlı (multilayer) sinir ağlarını eğitmenin zorluğuna da dikkat çekmişlerdir. 1967'de Amari, stokastik gradyan inişi ile eğitilen ve doğrusal olarak ayrıştırılamayan desenleri sınıflandırılabilen ilk çok katmanlı sinir ağını tanıttı (Amari, 1967) ve daha sonra beş katmanlı bu ağın bilgisayar deneylerini

öğrencisi Saito gerçekleştirdi (Schmidhuber, 2022). Little ve Shaw, beynin hafıza kapasitesinin, nöron ateşlemeleri, sinaptik modifikasyonlar ve sinaps sayısı ile nasıl ilişkilendiğini modelleyerek sinir ağlarının işleyişine dair yeni bir bakış açısı sunmuşlardır (Little ve Shaw, 1978).

Markov modeli, her olay geçişinin olasılığının yalnızca bir önceki olayın ulaşılan durumuna bağlı olduğu bir rastlantısal olaylar zincirini temsil eder (Pasma ve Rogers, 2020:74). Gizli Markov Modelleri (HMM), bir dizi durumu içeren ve her durumun belirli semboller ürettiği, durumlar arasında geçişlerin olasılıklarla belirlendiği modellerdir. (Rabiner, 1989). Temel fikir, HMM'nin sonsuz sayıda olası dizi üzerinde bir olasılık dağılımını tanımlayan sonlu bir model olmasıdır (Eddy, 1996). Konuşma tanıma, protein/DNA dizisi analizi, robot kontrolü ve metin verilerinden bilgi çıkarma gibi çeşitli alanlarda başarılı bir şekilde uygulanmışlardır (Zhai, 2003). Bu modeller, her durumu modelin iç yapısını temsil edecek şekilde tanımlar; geçiş olasılıkları bir durumdan diğerine geçişi, gözlem olasılıkları ise her durumun belirli bir gözlem (örneğin, bir dil modelinde kelime vb.) üretme olasılığını belirler. Bu yapı, her gözlemin, hangi durumdan geldiğine göre ne kadar olası olduğunu belirleyen olasılık dağılım fonksiyonları ile modellenir.

Maksimum Entropi Markov Modelleri (MEMM), Markov modelinin sıralı yapısını maksimum entropi ilkesi ile birleştirerek, etiketli veriye dayanarak her durumda en olası etiketin tahmin edilmesini sağlayan bir sıralı modelleme tekniğidir. MEMM, belirli bir gözlem ve önceki duruma bağlı olarak bir duruma ulaşma olasılığını temsil eden koşullu bir model kullanır, bu da HMM'nin bağımsızlık varsayımlarını esnetir (McCallum ve ark., 2000). Koşullu Rastgele Alanlar (CRF), dizi (sequence) etiketlemede, her etiketin yalnızca komşularına bağlı olduğu olasılıksal modellerdir. MEMM, dizi sonuna kadar gözlemlerle ilgilenen bağımsız tahminleri birleştirir; ancak bağımsız tahminlerin aşırıya kaçması, özellikle düşük frekanslı etiketler için "etiket yanlılığı" problemine yol açabilir. CRF'nin koşullu dağılımı, kenar ve köşe özellik fonksiyonları kullanılarak belirli parametrelerle ifade edilir, böylece etiket yanlılığı sorununu azaltır. HMM gibi çalışabilen CRF, diziye özel bir normalizasyon kullanarak gözlem dizilerinin daha basit hesaplanmasını sağlar. CRF, gözlem bağımlılıklarını daha zengin bir biçimde ifade edebilir, bu yüzden özellikle karmaşık yapıdaki bağımlılıklara sahip veriler üzerinde daha iyi performans gösterir. Özellikle doğal dil işleme görevlerinde, CRF, cümleler üzerindeki dilbilgisel analizlerde ve sözcük türü

etiketleme (POS) gibi görevlerde MEMM ve HMM'e göre daha yüksek doğruluk sağlamıştır (Lafferty ve ark., 2001).

Doğal dil işlemenin son yıllardaki gelişimden önce, soru cevaplama bir bilgi alma (IR) görevi olarak tanımlanıyordu (Kolomiyets ve Moens, 2011). Bu görev genel olarak veri tabanlarından sorgularla (query) bilgi çekmeye dayanmaktaydı. Doğal dil işlemenin gelişmesiyle birlikte soru cevaplama otomatik olarak yapılmasının önü açılmış oldu. Soru cevaplama sistemleri, bir bilgi tabanından veya belge gövdesinden ilgili bilgileri çıkararak doğal dilde sunulan soruları otomatik olarak cevaplamak üzere tasarlanmıştır (Alanazi ve ark., 2021).

J. L. Borges'in ifade ettiği gibi (1925), diller, her zaman onları aşan bir gerçekliğin sadeleştirilmiş biçimleridir. Bu perspektif, doğal dil işleme alanında, dilin soyut ve dinamik doğasının teknolojiyle modellenmesinin sınırlılıklarını anlamamıza ışık tutmaktadır. Yapay zekaların belli örüntü kalıplarıyla dili öğrenmesi bazı nüansların kaçırmasına yol açabilir. Bu bağlamda dili niceliksel temsillerden geri dönüştürmede başarılı yapay zekaların niteliksel bir analizde performansına bakılmak istenmiştir. Bu amaçla bu tez çalışmasında ilk olarak eğitim biliminden yararlanılmış ve Bloom Taksonomisine göre yapay sorular oluşturulmuş ve uzman sorularıyla kıyaslanmıştır.

Mevcut içerikler yeterli mi ve yapay zeka bu konuda yeterli bir şekilde kullanılabilir mi soruları akla gelmektedir. Bu amaçla sinema filmlerinin içerik analizi yapılmış ve filmlerin kısa tanıtım yazılarından türlerinin belirlenmesi çalışması yapılmıştır. Bunun için metin eşleştirme ve karşıt öğrenme yöntemleri ile 2D-CNN modeli eğitilmiştir. Ayrıca regresyon analizi ile IMDb film puanları tahminleri yapılmış ve makine öğrenimi yöntemleri karşılaştırılmıştır.

Düşünce, onu yönlendiren dilsel ifadeler aracılığıyla bilinçli bir şekilde ortaya çıkar. Dilsel ifadeler ve onların değerlendirmeleri de birer algı olduğundan, dilin varlığı, başka türlü çerçevelenemez olan düşünce üzerine düşünceler oluşturmayı mümkün kılar. Ayrıca, bir kelime, sabit bir anlam taşıyor gibi görünse de farklı bağlamlarda anlamı değişir ve farkları anlamak için dikkatli bir analiz gerekir (Jackendoff, 1996). Cümlelerdeki kelime sırası ve paragraflardaki cümle sırası, bilişsel işlev hakkında önemli bilgiler sağlayabilir ve dilsel karmaşıklığı ve bilişsel sağlığı değerlendirmek için daha gelişmiş doğal dil sözdizimsel analiz yöntemleri de kullanılabilir (Voletti ve ark. 2019). Tüm bunlar değerlendirildiğinde bu

tez çalışmasında son olarak çeşitli kaynaklardaki verilerden intihar eğilimi, depresyon, anksiyete ve semptomsuz/ normal olarak etiketlenmiş sosyal medya gönderilerini kullanarak bunların sınıflandırması üzerine deneysel çalışma gerçekleştirilmiştir.

Sosyal medyanın intihar davranışı üzerindeki etkilerine dair ortaya çıkan veriler, topluma yeni tehditler sunabileceği kadar, yardım ve önleme için de yeni fırsatlar yaratabileceğini göstermektedir (Luxton ve ark., 2012). Choudhury ve arkadaşları, ciddi depresyonu olan bireylerin sosyal medyadaki davranışlarının, bu semptomlara sahip olmayan bireylerden farklı olabileceği ortaya koymuştur (De Choudhury ve ark., 2013). Wang ve Bashir (2020) anksiyete semptomları gösteren bireylerle göstermeyenler arasında sosyal medya davranışlarında farklılıklar olduğunu ortaya koymuştur. Sosyal medya, bireylerin duygusal durumlarını ve günlük yaşamlarını yansıttığı için mental sağlık değerlendirmelerinde yeni bir kaynak olabilir (Wang ve Bashir, 2020). Sosyal medya verileri gerçek zamanlı olarak mevcuttur ve bu yüzden mental sağlık gözetimini ve riskinin tahminini kolaylaştırır (Gruebner ve ark., 2017).

Sonuç olarak yapay zekanın doğal dil işlemede etkili bir şekilde kullanılması için farklı uygulamaları içeren bir tez çalışması gerçekleştirilmiştir. Bu tez çalışması kapsamında;

1. Doğal dil işlemede son teknoloji yapay zekaların içerik üretme performansı incelenmiştir.
2. Sinema filmlerinin içerik analizi ile tür doğrulama ve film puanı tahmini yapılmıştır.
3. Çeşitli kaynaklardaki verilerden intihar eğilimi, depresyon, anksiyete ve semptomsuz/ normal olarak etiketlenmiş sosyal medya gönderilerini kullanarak bunların sınıflandırması üzerine odaklanılmıştır.
4. Büyük dil modellerinin sınıflandırma performansı değerlendirilmiştir ve metin ön işlem tekniklerinin etkisi incelenmiştir.
5. Farklı metin gömme yöntemlerinin etkisi makine öğrenimi yöntemlerinden SVM ve klasik derin öğrenme yöntemlerinden çift yönlü (ileri ve geri beslemeli) LSTM (BiLSTM) ile ölçülmüştür.

2. ÖNCEKİ ÇALIŞMALAR

İngilizce-Türkçe çeviri performanslarının karşılaştırıldığı çalışmada, Google Translate, Amazon Translate ve OpenAI GPT çevirileri değerlendirilmiş ve Türkçe-İngilizce-Türkçe döngüsü ile orijinal Türkçe ve çeviri Türkçe metinlerin benzerlikleri ölçülmüştür. Levenshtein mesafesi ve kosinüs benzerliği metrikleriyle elde edilen sonuçlarda, ortalama benzerlik OpenAI GPT ile 0.692, Amazon Translate ile 0.776 ve Google Translate ile 0.779 olarak bulunmuştur (Öncü ve Kutlu, 2024).

ChatGPT'nin otomatik olarak ürettiği çoktan seçmeli soruların (ÇSS) bilimsel ve klinik kabul düzeyini değerlendiren bir çalışmada, uzmanlar tarafından üretilen ÇSS ile karşılaştırmalı sonuçlar, ChatGPT tarafından üretilen ÇSS'lerin, konu uzmanları tarafından oluşturulan ÇSS'lerle benzer kalite düzeylerine sahip olduğunu göstermektedir (Kıyak, 2023). Bir başka çalışmada, büyük dil modellerinin (LLM'ler) kontrol edilebilir metin üretimiyle (CTG) eğitimdeki potansiyel etkileri değerlendirilmiştir. Elde edilen sonuçlar, CTG'nin eğitimde etkili bir araç olabileceğini gösterirken, aynı zamanda çalışmanın sınırlamalarına da dikkat çekmiştir (Elkins ve ark., 2023).

Makine öğrenimi yöntemleri kapsamında yapılan çalışmalara bakıldığında, Bloom'un Bilişsel Alan Taksonomisine dayalı soru sınıflandırma için farklı sayısal temsil ve sınıflandırıcılar kullanılan çalışmalar bulunmaktadır. Bu bağlamda, KNN (K-En Yakın Komşu), LR (Lojistik Regresyon) ve SVM (Destek Vektör Makineleri) sınıflandırıcıları ile gerçekleştirilen bir çalışmada, soruların Bloom taksonomisinin bilişsel alanına göre sınıflandırılması hedeflenmiştir (Mohammed ve Omar, 2020). TF-IDF (Terim Frekansı-Belge Frekansı Ters) ve Word2Vec gibi vektörleme yöntemleri, özellik mühendisliği teknikleri olarak kullanılmıştır. W2V-TFPOSIDF ve SVM yöntemleriyle bir veri kümesinde %83, diğer bir veri kümesinde ise %89 F1 skoru elde edilmiştir. Naive Bayes sınıflandırıcısının Bloom Taksonomisi Bilişsel Alanı'ndaki düzeylere göre sınav sorularını tahmin etme modeli olarak kullanılabilirliğinin ölçüldüğü çalışmada, veri kümesini 10 Katlı Çapraz Doğrulama ile farklı vektörleme yöntemlerinin performansını değerlendirilmiştir. 300 ara sınav ve final sınav sorusundan oluşan bu veri seti, TF-IDF yöntemi ile özellik çıkarımı sınıflandırma için kullanılmıştır. Yapılan değerlendirmeler neticesinde, N-Gram TF-IDF ile %85 Precision ve Words TF-IDF ile 82% Recall skoru elde ettiği belirtilmiştir (Aninditya ve ark., 2019). Derin öğrenme yöntemleri ile yapılan çalışmalara bakıldığında,

Bloom'un Taksonomisine dayalı soru sınıflandırma için yapılan çeşitli çalışmalar bulunmaktadır. Bu bağlamda, LSTM ve önceden eğitilmiş kelime gömme kullanımı incelendiği çalışmada Bloom'un taksonomisinin altı seviyesini temsil eden anahtar kelime/fiil sözlüğü kullanılmıştır (Shaikh ve ark., 2021). Yazılım mühendisliği dersine ait 844 sorunun Bloom Taksonomisine göre sınıflandırılması üzerine derin öğrenme tabanlı CNN ve LSTM kullanılan çalışmada, yapılan deneyler sonucunda önerilen CNN modelinin Bloom'un Taksonomisinin bilişsel boyutlarının tahmininde %80 doğruluk oranına ulaştığı görülmüştür. Bloom'un Taksonomisinin bilgi boyutları tahmininde ise LSTM modeli eğitim aşamasında daha başarılı olsa da test aşamasında CNN modelinin daha yüksek bir doğruluk oranına sahip olduğu belirlenmiştir (Laddha ve ark., 2021).

Bankacılık sektörüne yönelik bir makine okuma için soru-cevap (MRQA) sistemi geliştirilen çalışmada iki aşamalı ve her aşama için üç veri içeren bir eğitim süreci gerçekleştirilmiştir. Çalışma için eğitilen BERT modeliyle bir cümleden maskelenmiş kelimeleri ve bir sonraki cümleyi tahmin etme görevlerini gerçekleştirerek kelimenin bağlama göre anlamlı bir şekilde gömülmesini sağlamıştır. Açık alan ve kapalı alan soru cevaplama için ayrı veri setlerinde değerlendirilen model BERTurk ve mBERT modellerinden daha iyi sonuçlar vermiştir (Gemirter ve ark., 2021).

Türkçe metinlerden otomatik soru üretimi ve cevap çıkarımı için mT5 modelini kullanarak çoklu görevli bir yöntem öneren çalışmada otomatik soru cevaplama, soru üretimi ve cevap çıkarımı görevleri gerçekleştirilmiştir. Model, tek-görevli ayarlarda ayrı olarak eğitilmiş, çoklu-görevli ayarlarda ise bu görevlerin hepsini aynı anda gerçekleştirecek şekilde tasarlanmıştır. Performans değerlendirmesi, model boyutunu artırmanın hem küçük hem de büyük veri setlerinde önemli ölçüde iyileşme sağladığını, özellikle mT5-small'dan mT5-base'a geçişte belirgin bir artış gösterdiğini, mT5-large'a geçişin performans artırımını nispeten daha sınırlı olduğu ortaya koymuştur (Akyön ve ark., 2022).

IMDb verilerine dayalı çalışmalar, makine öğrenimi modellerini ve geçmiş verileri kullanarak film puanlarını doğru bir şekilde tahmin etmeyi amaçlamaktadır. Örneğin, aktörler, yönetmenler, türler, içerik derecelendirmeleri ve yapım şirketleri gibi geçmiş veriler kullanılarak yapılan bir çalışmada, 1D CNN modeli en iyi performansı sağlamış ve MSE 1.20, MAE ise 0.83 değerlerine ulaşmıştır (Abarja ve Wibowo, 2020). Başka bir çalışmada, XGBoost algoritması, 1.0 ile 10.0 arasında değerlendirilen filmler için 0.656 MAE elde ederek etkili bir sonuç göstermiştir (Zabaleta de Larrañaga, 2021). ALS (Alternating Least

Squares) modeliyle ortak filtreleme ve makine öğrenimi tekniklerini birleştiren hibrit bir öneri motoru, RMSE değerlerini 0.8959 ile 0.97613 arasında bir aralıkta elde etmiştir (Awan ve ark., 2021).

Bu alandaki film türü belirme çalışmalarında, GRU (Gated Recurrent Units) modeliyle %56 F-skoru elde edilmiştir (Hoang, 2018). Topolojik Veri Analizi (TDA) yöntemleriyle bu skor %15 artırılarak %71,88'e ulaşılmıştır (Doshi ve Zadrozny, 2018). Bi-LSTM (Çift Yönlü LSTM) modeli ise %67,68 F-skoru sağlamıştır (Ertuğrul ve Karagöz, 2018). 10 farklı türden yapay zeka ile oluşturulmuş film başlıkları ve özetleriyle yapılan bir çalışmada, bir ardışık sinir ağı modeli ile %80 doğruluk oranı elde edilmiştir (Tanvir ve ark., 2023).

Tohma ve ekibi, insan-robot etkileşimli Türkçe soru cevap sistemlerinde duygu analizi yapmak üzere hibrit model önerdikleri çalışmada, olumlu olumsuz duygu etiketlerine sahip down sendromlu çocukların eğitimi için kullanılan üç veri kümesi sınıflandırma performansı incelenmiştir. Çalışmada metin vektörleştirme için word2vec, TF-IDF ve Türkçe BERT kelime temsil yöntemleri kullanılmıştır. Metin vektörlerine, metinlerin olumlu ve olumsuz duygularını temsil eden kutupsallık puanları (polarity scores) eklenerek hibrit bir vektör formu elde edilmiştir. Polarite skorlarını eklemenin performansı artırdığı gözlenmiştir (Tohma ve ark., 2023).

Genel metin ön işleme yöntemleri çoğu veri kümesi için uygundur, ancak daha fazla anlamsal bilgi korumak için içerik odaklı yöntemler gereklidir ve gelişmiş metin ön işleme teknikleri, bu yöntemlere ek olarak değerlendirilir (Chai, 2023). Örneğin, durdurma kelimelerinin sınıflandırmaya etkisinin incelendiği bir çalışmada durdurma kelimelerini kaldırmak sınıflandırmayı iyileştirmiş ek olarak teknik durdurma kelimelerini de kaldırmak sınıflandırma başarımını daha da artırdığı görülmüştür (Sarica ve Luo, 2021). Farklı bir çalışmada üç veri kümesi üzerindeki farklı ön işlemenin etkisini ölçüldüğü deneysel çalışmalarda kelimelerin yalın halde (lemmatizasyon) ve küçük harfle kullanılması çoğu durumda yüksek performans gösterirken, noktalama işaretleri, URL'ler, kullanıcı atfı (@mentions) ve hastag (#) simgelerinin kaldırılması çoğu durumda düşük performans sergilediği gözlemlenmiştir (Naseem ve ark., 2021).

Villa-Pérez ve ekibi, İngilizce ve İspanyolca Twitter kullanıcılarının mental sağlık durumlarını otomatik olarak tespit etmek için makine öğrenimi ve derin öğrenme yöntemlerini kullanmıştır. Teşhis edilmiş 1.500 kullanıcı ve kontrol grubundan 1.700

kullanıcıya ait veri kümeleri oluşturulmuş, n-gram, q-gram, konu modelleme gibi metin özellikleri çıkarılmıştır. XGBoost ile İngilizce veri setinde ikili sınıflamada 0.819 ve çoklu sınıflamada 0.47, İspanyolca veri setinde ikili sınıflamada 0.824 ve çoklu sınıflamada 0.501 F1 skoru elde etmişlerdir (Villa-Pérez ve ark., 2023). Ansari ve ekibi mental sağlık verilerinin çalışmasında veri artırımı yöntemi denemişler ve ikili sınıflandırma çalışmalarında veri artırımının F1 skorunda %4 varan iyileştirmeler sağlamışlardır (Ansari ve ark., 2021). Shorten ve ekibi tarafından yapılan çalışmada metin verileri için geri çeviri artırımı, sembolik ve kural tabanlı artırma veri artırımı yöntemleri incelenmiştir (Shorten ve ark., 2021).

Depresyonun metinlerden tespitinde BERT tabanlı sınıflandırıcıların yorumlanabilirliğini ve genelleme kapasitesini artırmayı amaçlayan DECK (DEpression ChecKlist) adlı depresyona özgü davranışsal testler setini tanıtmaktan çalışma, iki Twitter tabanlı ve bir klinik veri setinde, 23 DECK testi kullanarak BERT, RoBERTa ve ALBERT modellerini değerlendirmiştir. Modellerin güçlü yönleri arasında birinci tekil şahıs zamirlerinin kullanımını tespit etme becerisi bulunurken; intihar düşüncesi ve somatik semptomlar gibi depresyon belirtilerindeki zayıflıkları öne çıkmıştır. Sonuçlar, DECK testlerinin eğitim sürecine dahil edilmesiyle modellerin dağılım dışı performanslarının önemli ölçüde arttığını (F1 skorunda %53,93'e kadar artış) göstermektedir. DECK testleri, depresyon tespiti modellerini geliştirmek için tamamlayıcı bir değerlendirme ve tanı aracı olarak potansiyel sunmuştur (Novikova ve Shkaruta, 2022). Başka bir çalışmada, Reddit platformundaki kullanıcıların yapılandırılmamış verileri analiz edilerek depresyon, anksiyete, bipolar bozukluk, ADHD ve PTSD gibi beş yaygın mental hastalık sınıflandırılmıştır. En iyi sonuçlar, önceden eğitilmiş RoBERTa transfer öğrenme modeli kullanılarak doğruluk ve F1 skoru 0.83 elde edilmiştir (Ameer ve ark., 2022).

Zhang ve arkadaşları (2022), sosyal medya metinlerinin duygusal içerik analizi yoluyla zihinsel sağlık bozukluklarının belirlenmesine katkıda bulunduğunu belirtmiştir. Şili Eğitim Sistemi ile ilgili yapılan bir çalışmada depresyon, anksiyete ve stresin pandemi sonrası çocuklar ve ergenlerde yaygın olarak görüldüğünü ortaya koymuş, risk faktörlerini belirlemede %60'ın üzerinde doğrulukla sonuçlar elde etmiştir (Martínez-Líbano ve Yeomans-Cabrera, 2023). An ve arkadaşları (2024), gençler arasında sosyal medya kullanımı ile depresyon arasındaki ilişkiyi incelemiş ve bu ilişkinin coğrafi konumlar, sosyal medyada

harcanan süre, ideolojik temsiller (örneğin, bir şeyleri kaçırma korkusu olarak bilinen FOMO) ve kişilik özellikleri tarafından etkilenebileceğini bulmuşlardır.

Jan ve ekibi (2020) metin sınıflandırmasında doğruluğu artırmak için Bi-LSTM modelini Word2vec CNN ve dikkat mekanizmasının birleştirilmişlerdir. İçerik tabanlı filtreleme, kullanıcıların geçmişteki tercihlerine veya etkileşimlerine dayalı olarak benzer özelliklere sahip öğeleri öneren bir sistem olup, film öneri sistemlerinde (Musa ve Zhihong, 2020; Sanwal ve Çalışkan, 2021; Nair ve Preethi, 2022) başlangıç filtreleme yöntemi olarak kullanılır. Yoğun temsiller (dense representations), TF-IDF ve BM25 gibi geleneksel yöntemlere kıyasla çift kodlayıcı bir çerçeve kullanarak metinleri daha etkili bir şekilde almaktadır (Karpukhin ve ark., 2020). Üretken modeller (LLMs), alınan pasaj sayısına paralel olarak belirgin iyileşmeler göstermektedir (Izacard ve Grave, 2020).

3.MATERYAL VE YÖNTEM

3.1. Ön İşlem Teknikleri

Doğal dil işleme süreçlerinde ön işlemler dilin temsillerini -bu tez çalışmasında metinler-analiz ve modelleme için uygun bir hale getirme aşamasıdır. Bu süreç, veriyi temizleyip yapılandırarak, makine öğrenimi ve derin öğrenme modellerinin daha etkili bir şekilde çalışmasını sağlar. Ön işleme, metindeki gereksiz karakterleri, anlam taşımayan yapıları ve dil modelini genelleme yapmasına engel unsurları ayıklayarak veri gürültüsünü azaltmayı hedefler. Bu aşamalar, modelin metindeki gerçek anlam ilişkilerini öğrenmesini kolaylaştırır. Durdurma kelimeleri (stop-words) tek başına herhangi bir anlam taşımayan veya metin için işlevi olmayan kelimelerdir (Çakır ve Güdamlasioğlu, 2016). Seçenekler uygulamaya bağlı olarak bu liste özelleştirilebilir. Çizelge 3.1.'de çeşitli metin ön işleme teknikleri ve modelin metni işlemesi açısından faydaları sunulmuştur.

Çizelge 3.1. Ön işlem teknikleri ve kullanım açıklamaları

Ön İşleme Tekniği	Açıklama
Durdurma Kelimelerinin (Stop-words) Çıkarılması	Metin için anlam taşımayan veya modelleme açısından işlevi olmayan kelimelerin kaldırılması işlemidir. Uygulamaya göre özelleştirilebilir.
Kök Bulma (Lemmatization)	Kelimenin anlamını koruyarak, farklı çekim ve biçimlerden tek bir kök formda temsil edilmesini sağlar. Kelimeleri temel köklerine indirger ve anlamlı bir yapı sunar.
Kökleştirme (Stemming)	Kelimelerin eklerinin kaldırılması işlemidir. Yalın formundan farklı bir kelime ortaya çıkabilir; bu nedenle kök bulmadan farklıdır.
URL'lerin Çıkarılması	Metindeki bağlantıları kaldırarak veri setini gereksiz bağlantılardan temizler, modelin odaklanmasını artırır.
HTML Etiketlerinin Silinmesi	Web içeriklerinde bulunan HTML kodlarının kaldırılması, analiz sürecinde anlam taşımayan yapıları temizleyerek daha odaklı bir veri sağlar.
Satır Sonları ve Sekme Karakterlerinin Silinmesi	Metni daha düzenli ve analiz edilebilir bir formata dönüştürmek için satır sonları ve sekme karakterleri kaldırılır.
Markdown Bağlantılarının Çıkarılması	Metindeki Markdown bağlantılarının silinmesiyle bağlantı parazitleri temizlenir ve metin daha anlaşılır hale gelir.
Kullanıcı Adları ve E-posta Adreslerinin Silinmesi	Özellikle sosyal medya verilerinde, kullanıcı adları ve e-posta gibi öğelerin çıkarılması, odaklanmayı artırır.
Noktalama İşaretlerinin Kaldırılması	Metindeki noktalama işaretlerinin kaldırılması, modelin anlamlı kelime gruplarına daha iyi odaklanmasına yardımcı olur.

Belirli Dil Karakterlerinin Tutulması	Sadece hedef dilin karakterlerinin bırakılması, veri tutarlılığını artırarak analiz sürecinde karakter çeşitliliğini sınırlandırır.
Küçük Harflere Dönüştürme	Harflerin büyük/küçük olmasından kaynaklanan farklılıkları önlemek için tüm metin küçük harfe dönüştürülür, analiz sürecini etkileyebilecek bu tür varyasyonlar giderilmiş olur.
Boşlukların Silinmesi	Fazla veya gereksiz boşlukların çıkarılması, metni daha düzenli ve parazitsiz bir formatta analiz edilmesine olanak tanır.
Tekrarlayan ifadeleri kaldırma	İkilemeler dışında cümle veya söz öbeği biçiminde tekrarlayan ve modelin öğrenmesinde anlamı olmayan tekrarlayan ifadeleri kaldırmak gereksiz işlemde tasarruf etmemizi sağlar.

3.2. Yapay Zekâ Modelleri

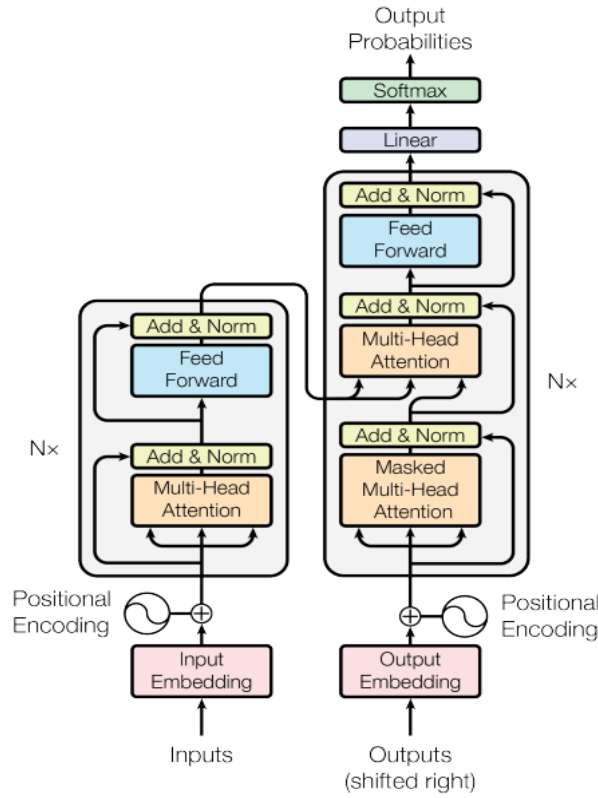
İleri beslemeli (feed-forward) sinir ağları, katmanlar halinde düzenlenmiş düğümlerden oluşur ve bu ağ yapısında veriler, giriş katmanından başlayarak bir yönlü olarak, gizli katmanlar üzerinden çıkış katmanına doğru ilerler. Geri yayılım (backpropagation, BP) algoritması hatanın çıkış katmanından gizli katmanlara ve oradan giriş katmanına doğru yayılmasını sağlar ve öğrenme gerçekleştirilir. Böylece, ağırlık ve eşik katsayıları, bu hata bilgisi kullanılarak ayarlanır ve ağ, istenen çıktıyı üretmeye yönelik optimize edilir (Svozil ve ark., 1997). Geri beslemeli nöral ağlar (Recurrent Neural Network, RNN), sıralı verileri modellemek için feed-forward sinir ağlarının basit bir uyarlamasıdır. Her adımda, giriş alır, gizli durumunu günceller ve tahmin yapar, böylece geçmiş bilgileri kullanarak daha doğru sonuçlar üretir (Sutskever ve ark., 2011). Konvolüsyonel sinir ağları (Convolutional Neural Network, CNN), evrişim işlemleri yoluyla verilerden özellikleri otomatik olarak çıkarmak üzere tasarlanmış ileri beslemeli sinir ağlarıdır (Li ve ark., 2021).

Fukushima (1980) biyolojik görme sistemlerinden esinlenerek nesnelerin farklı açılardan tanınmasını sağlayan, öz-organizasyon yeteneği sunan çok katmanlı hiyerarşik bir yapıya sahip Neocognitron modelini geliştirmiştir. Neocognitron, "derin" yapıda bir yapay sinir ağı olarak, Evrişimsel Sinir Ağlarını (Convolutional Neural Network, CNN) tanıtan ilk ağlardan biridir ve nörofizyolojik içgörülerini uygulayan ilk ağdır (Schmidhuber, 2015). 1982'de Hopfield, simetrik bağlantılarla çalışan, Hebbian öğrenme (Hebb, 1949) kuralıyla desenleri saklayan ve gürültülü verilerle doğru desenleri geri çağırabilen bir geri beslemeli (recurrent) ağ olan Hopfield ağını tanıttı (Hopfield, 1982). Seppo Linnainmaa, 1970 yılında otomatik türev hesaplama (automatic differentiation) tekniğini geliştirerek çok katmanlı

fonksiyonların türevini hesaplamaya yönelik adımları sundu ve böylece backpropagation algoritmasının matematiksel temelini oluşturdu (Linnainmaa, 1970; Schmidhuber, 2014). 1986'da Rumelhart ve ekibi, çok katmanlı perceptronların (MLP) eğitiminde backpropagation algoritmasının etkili bir yöntem olduğunu göstererek algoritmayı yaygınlaştırdı (Rumelhart ve ark., 1986). 1989'da LeCun ve arkadaşları, Neocognitron benzeri ağırlık paylaşımı yapan, evrimsel sinir ağlarına geri yayılım algoritmasını uygulamışlardır (LeCun ve ark., 1989).

Hochreiter ve Schmidhuber tarafından (1997) geliştirilen LSTM (Long Short-Term Memory) yöntemi, önceki RNN'lerin zorlandığı derin ve uzun süreli bağımlılık gerektiren görevlerde başarılı olmuş ve denetimsiz ön eğitime gerek kalmadan derin öğrenmenin temel sorunlarını aşabilmiştir. Ayrıca LSTM, dil görevlerinde önceki RNN ve HMM yapılarından daha iyi performans göstermiştir (Schmidhuber, 2015). Graves ve Schmidhuber, çift/iki yönlü LSTM (Bidirectional LSTM, BiLSTM) modelini tanıtarak, bunun özellikle konuşma tanıma gibi sıralı verilerde bağlam bilgisini iyileştirdiğini ve LSTM'e göre daha kısa eğitim süresinde daha iyi sonuçlar verdiğini göstermiştir (Graves ve Schmidhuber, 2005). Çift yönlü özelliği ile önceki/geçmiş bilgiye dayalı olarak çalışmasına ek olarak sonraki/gelecek bilgiye de erişebilen bir modeldir.

Denetimli öğrenme (supervised learning), etiketlenmiş veri kullanarak modelin doğru tahminler yapmayı öğrenmesini amaçlayan bir yöntemdir. Model, giriş-çıkış çiftlerinden öğrenir ve girdileri doğru etiketlere eşleştirmeyi öğrenir (Kotsiantis, 2007). Bu metin, denetimli ve öğretmensiz öğrenme kavramlarını karşılaştırarak özetlemektedir. Denetimli öğrenme, bir "öğretmen" veya denetleyicinin doğru yanıtlar veya hata ölçümleri sağladığı durumları içerir. Denetimsiz öğrenme (unsupervised learning), bir denetleyici veya öğretmen tarafından her gözlem için doğru yanıtlar veya hata geri bildirim olmadan yalnızca girdilerle çalışır ve bu girdilerin ortak yoğunluğunu doğrudan tahmin etmeye odaklanır. Denetimsiz öğrenme yöntemlerinde başarıyı ölçmek zordur ve sonuçlar sezgisel olarak değerlendirilir. Bu yöntemler arasında kümeleme, ana bileşen analizi ve karışım modelleme gibi teknikler yer alır (Hastie, 2009: 485-486). Transfer öğrenme, bir kaynaktan elde edilen bilgiyi, hedefteki tahmin fonksiyonunun öğrenilmesini iyileştirmek için kullanmayı amaçlar. Kaynak ve hedef, aynı özelliklere ya da aynı görevlere sahip olmayabilir, bu nedenle transfer öğrenme farklı özelliklere veya görevlere sahip iki alan arasında bilgi aktarımı sağlar (Pan ve Yang, 2009).



Şekil 3.1. Transformatör mimarisi (Vaswani ve ark., 2017:3)

Vaswani ve arkadaşları 2017’de “Attention is All You Need” başlıklı makalede, sıralı hesaplamaları ortadan kaldırarak giriş ve çıkış arasında tüm bağlamın dikkate alınmasıyla paralelleştirme imkânı sağlayan Transformer ağ mimarisini tanıttılar. Mimari encoder (kodlayıcı) ve decoder (kod çözücü) iki ana bileşenden oluşmaktadır. Ölçeklenmiş nokta çarpan (scaled dot-product) ve çok başlı (multi-head) bileşenlere sahip dikkat (attention) mekanizması, farklı konumlardaki bilgiye paralel olarak odaklanarak modelin daha güçlü ve detaylı bağlam ilişkileri öğrenmesini sağlar. Her katman, her pozisyona ayrı ayrı uygulanan bir feed-forward ağı, token'ları vektörlere dönüştürmek için gömmeler ve dizinin sırasını anlaması için konum kodlamaları (positional encodings) oluşur (Vaswani ve ark., 2017).

Aşağıdaki formüller bu makaleden alınmıştır (Vaswani ve ark., 2017:4,5);

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (1)$$

Temel dikkat (ölçeklendirilmiş nokta-çarpan dikkati) mekanizmasında (1):

- **Q**: Sorgu (query) vektörü
- **K**: Anahtar (key) vektörü
- **V**: Değer (value) vektörü
- d_k : Vektörlerin boyutu

Sorgu ve anahtar vektörlerinin birbirleriyle ne kadar ilişkili olduğunu anlamak için, bu vektörlerin nokta çarpımını alınır. Ancak bu değeri daha dengeli hale getirebilmek için, bu noktada d'ye bölme işlemi yaparız ve bu ilişkiyi softmax fonksiyonu ile normalize edilir ve elde edilen ağırlıkları, değer (V) vektörlerine uygulanır.

$$MultiHead(Q, K, V) = Concat(head_1, \dots, head_n)W^o \quad (2)$$

Burada;

$$head_i = Attention(QW_i^Q, KW_i^K, VW_i^V) \quad (3)$$

Çok başlı dikkat (2), (3);

- $head_i$: Farklı başlıklarla elde edilen dikkat sonuçları.
- Concat: Tüm başlıkların çıktıları birleştirilmesi.
- W^o : Birleştirilmiş çıktıyı modelin giriş boyutuna uygun hale getiren ağırlık matrisi.

Her başlık, sorguları, anahtarları ve değerleri farklı boyutlara indirger ve bu başlıklar paralel olarak çalışır. Her bir başlık ayrı bir dikkat fonksiyonu çalıştırır ve sonuçlar birleştirilir. Bu sayede model, farklı başlıklar üzerinden aynı anda farklı bilgi alt uzaylarına dikkat edebilir.

Büyük Dil Modelleri (Large Language Models, LLMs) ifadesinin, büyük (large) terimi, milyonlarca ya da milyarlarca parametre içeren ve karmaşık dil kalıplarını ve nüanslarını yakalamasına olanak tanıyan modelin ölçeğini vurgular. Dil modeli (Language Model) kullanımı, bu sistemlerin bir dil dizisinin olasılığını tahmin etmek üzere eğitildiğini ve bu sayede çeviri, özetleme ve soru cevaplama gibi görevler için etkili olduklarını gösterir. Büyük dil modelleri ön eğitilmiş modellerdir ve uzun dil bağlamlarını etkili bir şekilde ele almasını sağlayan transformers mimarisine dayanır. 2018 yılında GPT-1 (Generative Pre-trained Transformer- 1), üretken ön eğitilmiş dil modeli olarak tanıtılmıştır (Radford ve ark., 2018).

BERT (Bidirectional Encoder Representations from Transformers) iki yönlü dikkat mekanizmasını kullanarak dilin her iki yönünü aynı anda işleyerek bağlama duyarlı temsiller oluşturabilen bir dil modeli olarak tanıtılmıştır. Maskelenmiş dil modeli (MLM) ve sonraki cümle tahmini (NSP) kullanarak, dilin bağlamını daha iyi anlamasına olanak tanır (Devlin ve ark., 2018). RoBERTa (Robustly Optimized BERT pretraining Approach), BERT modelinin eğitiminde yapılan tasarım iyileştirmeleri ile geliştirilmiş bir dil modelidir. RoBERTa, daha uzun eğitim süresi, büyük batch'ler, daha fazla veri, sonraki cümle tahmini (NSP) hedefinin kaldırılması ve dinamik maskeleye gibi iyileştirmelerle BERT'in performansını artırmıştır (Liu ve ark., 2019).

XLNet, Auto-Regressive (AR) yaklaşımını benimseyerek dil modellemesi görevlerinde Auto-Encoding (AE) yönteminin avantajlarını birleştiren bir modeldir. Transformer-XL (uzun metinlerdeki bağımlılıkları öğrenebilen) mimarisi üzerine inşa edilen XLNet, permütasyon tabanlı bir dil modelleme yöntemi kullanarak maskelenmiş tokenler yerine kelimelerin sırasını karıştırır (Yang ve ark., 2019).

Raffel ve arkadaşları, metin tabanlı tüm dil işleme görevlerini bir metinden-metne dönüşüm problemi olarak ele alan birleşik bir çerçeve sunmuş ve bu çerçeveyi kullanarak transfer öğrenmede çeşitli dil işleme görevlerinin gerçekleştirilmesini sağlayan T5 (Text-to-Text Transfer Transformer) modelini tanıtmıştır (Raffel ve ark., 2020). Türkçe dahil 101 dilde eğitilen mT5, T5'in çok dilli bir versiyonu olarak geliştirilmiştir (Xue ve ark., 2020).

Lewis ve arkadaşları, metinleri gürültüye maruz bırakıp yeniden oluşturma yöntemini kullanarak, doğal dil üretimi, çeviri ve anlama görevleri için BART (Bidirectional and Auto-Regressive Transformers) modelini tanıtmıştır (Lewis ve ark., 2020). Türkçe dilinin içlerinde bulunmadığı 50 dilde eğitilen mBART, BART'ın çok dilli bir versiyonu olarak geliştirilmiştir (Tang ve ark., 2020). VBART, Türkçe için sıfırdan büyük bir veri kümesi üzerinde önceden eğitilmiş ilk sequence-to-sequence büyük dil modeli olup, BART mimarisine dayalı bir transformatör kodlayıcı-kod çözücü üzerine kuruludur ve soru üretimi ile soru yanıtlama görevlerinde ince ayarı yapılmıştır (Turker ve ark., 2024).

Kullanılan büyük dil modellerinin çeşitli özelliklerine göre karşılaştırılması Çizelge 3.2.'de verilmiştir.

Çizelge 3.2. Büyük Dil Modelleri Karşılaştırması

Özellik	BERT	RoBERTa	XLNet	T5
Mimari	kodlayıcı	kodlayıcı	kod çözücü	kodlayıcı- çözücü
Model Temeli	Bidirectional Transformer kullanır.	BERT'in temelini olarak geliştirilmiş bir dil modelidir.	Permütasyon tabanlı,Transformer-XL, iki akışlı dikkat.	Transformer tabanlı Text-to-Text
Ön Eğitim	Maskelenmiş Dil Modelleme (MLM) ve Sonraki Cümle Tahmini (NSP) kullanılmıştır.	Yalnızca Maskelenmiş Dil Modelleme (MLM) kullanılmıştır. NSP hedefi kaldırılmıştır.	AutoRegressive (AR) yaklaşımıyla sıralama bağımsızlığı	Span Corruption, Denoising (Bozuk metni yeniden oluşturma)
Ön eğitim Verisi	BOOKCORPUS, İngilizce Wikipedia	CC-NEWS, OPENWEBTEXT, STORIES	Wikipedia, BooksCorpus	Colossal Clean Crawled Corpus (C4)
Maskleme	Girdi token'larının %15'i rastgele maskelenir ve model bu maskeleri tahmin etmeye çalışır.	Dinamik maskelenme kullanır, her döngüde farklı maskeler oluşturulur.	Permütasyon tabanlı.	Girdi metnindeki rastgele %15 oranında maskelenmiş bölgelere özel <extra_id> token'ları eklenir
Tokenizasyon	WordPiece Tokenizer	Byte-Pair Encoding (BPE)	SentencePiece	SentencePiece

Destek vektör makineleri (Support Vector Machine, SVM, sınıflandırma ve regresyon görevi için çekirdek tabanlı bir makine öğrenimi modelidir (Vapnik, 1995). Örüntü tanıma uygulamalarında SVM gibi yapay sinir ağları dışındaki makine öğrenme yöntemleri öne çıkmıştır (Schölkopf ve ark., 1999).

Karar ağaçları (Decision Trees), verilerdeki koşulları test ederek tahminler yapan, her dalında bir koşul, her yaprak düğümünde ise bir sonuç üreten bir öngörüsül model türüdür (Rudin ve ark., 2022). Gradient Boosting Machine, işlev uzayında sayısal optimizasyon kullanan bir işlev tahmin yöntemidir (Friedman, 2001). Gradyan Artırmalı Karar Ağaçlarında (Gradient Boosting Decision Trees, GBDT), karar ağaçları ardışık olarak inşa edilir ve her ağaç, modele negatif gradyanın uydurulmasıyla oluşturulur. Bu süreçte, her bir zayıf sınıflayıcı (weak classifier), bir önceki ağacın hatalarını azaltacak şekilde modele eklenir. CatBoost kategorik veriyi doğrudan işleyebilen ve büyük veri kümeleri için verimli

çalıřan yöntemleri içerir. Özellikle ordered boosting gibi teknikler, aşırı uyumu azaltmaya katkı sağlar. Ancak, CatBoost'un performansı diđer GBDT algoritmalarına göre hiperparametre ayarlarına daha duyarlıdır ve bu ayarların doğru yapılması performansını büyük ölçüde etkiler (Hancock & Khoshgoftaar, 2020). XGBoost, karar ağaçlarını düzenleme (regularization) ile birleştirerek aşırı uyumu (overfitting) önlemeye yardımcı olur ve tahmin doğruluğunu artırır (Szczepek, 2022).

3.3. Performans Ölçümü

Regresyon modellerinde değerlendirme metrikleri: Ortalama Karesel Hata, Ortalama Mutlak Hata, Determinasyon Katsayısı, Karekök Ortalama Karesel Hata ölçüm değerleri ile kullanılmıştır.

Ortalama Karesel Hata (MSE), tahmin edilen değerler ile gerçek değerler arasındaki farkların karesinin ortalamasıdır. Bu metrik, hataların büyüklüğünü ölçmek için kullanılır ve hata miktarı büyüdükçe değeri de artar. Küçük bir MSE değeri, modelin daha iyi performans gösterdiği anlamına gelir.

Ortalama Mutlak Hata (MAE), tahmin edilen değerler ile gerçek değerler arasındaki farkların mutlak değerlerinin ortalamasıdır. Hata miktarını doğrudan ifade eder. MSE'den farklı olarak büyük hatalar daha az cezalandırılır, bu nedenle daha anlaşılır bir hata ölçüsüdür.

R^2 (Determinasyon Katsayısı), modelin veri içindeki deęişkenlięi ne kadar iyi açıkladığını gösterir. Bu metrik 0 ile 1 arasında deęer alır ve 1'e ne kadar yakınsa model o kadar başarılıdır. Örneęin, $R^2 = 0.90$ ise model verinin %90'ını açıklıyor demektir.

Karekök Ortalama Karesel Hata (RMSE), MSE'nin karekökünün alınmasıyla elde edilir. Hata birimi orijinal veri birimine geri döndüğü için yorumlanması daha kolaydır. Daha düşük bir RMSE değeri, daha iyi bir model performansını gösterir.

Sınıflandırma modellerinde değerlendirme metriklerinden ise doğruluk, duyarlılık, F1 skoru kullanılmıştır.

Precision (doęruluk), modelin "Pozitif" olarak tahmin ettięi sonuçların gerçekten ne kadar doęru olduęunu ölçer.

Recall (duyarlılık), modelin gerçek pozitifleri ne kadar iyi yakaladığını gösterir.

F1 skoru, doğruluk ve duyarlılık arasında bir denge kurarak modelin genel performansını tek bir değerle özetler.

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP})$$

$$\text{Recall (Sensitivity)} = \text{TP} / (\text{TP} + \text{FN})$$

$$\text{F1 score} = 2 * (\text{precision} * \text{recall}) / (\text{precision} + \text{recall})$$

TP = True Positives, TN = True Negatives, FP = False Positives, and FN = False Negatives.



4. ARAŞTIRMA BULGULARI

4.1. Verilerin Elde Edilmesi

Veriler üç deneysel çalışma için ayrı ayrı elde edilmiştir. Bloom taksonomisi için yapay soru üretilmiş ve uzman soruları kullanılmıştır. Sinema filmlerinin IMDb ve TMDB veri tabanlarında bulunan içerik bilgileri kullanılmıştır. Farklı platformlardan alınmış ve mental sağlık açısından tespit veya teşhisi yapılmış sosyal medya gönderileri sınıflandırılmıştır.

4.1.1. Bloom taksonomisi soruları

Bloom Taksonomisine göre yapay soru üretimi Microsoft Copilot, Google Gemini ve Open AI Chat GPT yapay zekâları ile elde edilmiştir. Sorular eğitim bilimlerinde ölçme değerlendirilmede kullanılan Yeniden Yapılandırılmış Bloom Taksonomisinin bilişsel alan seviyelerine uygun olarak oluşturulmuştur. Bloom'un yenilenmiş taksonomisine göre sınıflandırma görevi için fiil ve fiilimsilerin bilişsel süreç boyutunu tanımlayan öğeler olmasından (Stanny, 2016) dolayı cümlelerin bu öğelerinin bulunduğu kısımlar belirleyicidir. Ayrıca Stanny çalışmasında 30 farklı kaynaktan her basamağa ait fiilleri bir araya getirmiştir. Bu çalışmada her basamak için fiiller yönergede verilmek üzere seçilirken Stanny'nin çalışmasında derlediği ve 10 kaynaktan fazla kaynaktan ortak olarak geçen fiiller seçilmiş ve Türkçeye çevrilerek basamak fiilleri olarak kullanılmıştır. Seçilen belirli konularda her seviyede ve üç yapay zekada sorular oluşturulmuştur. Yenilenmiş Bloom Taksonomisine yönelik soru hazırlama ve soruların sınıflandırmasına yönelik çalışmalarda basamakların/boyutların tanımlaması Anderson ve Krathwohl (2001) çalışması referans alınarak yapılmaktadır. Çizelge 4.1.'de üretilen yapay soruların sayıları ve edinilen uzman sorularının sayıları verilmiştir.

Çizelge 4.1. Bloom taksonomisi soru dağılımı

Soru kaynağı	hatırlama	anlama	uygulama	analiz	değerlendirme	yaratma
Yapay soru	168	168	168	168	168	168
Uzman sorusu	25	32	26	28	25	26

Fiil ve fiil içeren cümlelerin bölümleri, sınıflandırma için bilişsel boyutun temel unsurlarıdır. Sorular, referans sorularından bağımsızdır. Cümlelerin alakasız bölümleri, sözcük türleri ve bağımlılık ilişkileriyle Türkçeye uyarlanmış Spacy kütüphanesi kullanılarak (Altınok, 2023) filtrelenmiştir. Uzman soruları yapay sorularla karşılaştırma için literatürde çeşitli kaynaklardan elde edilmiştir. Bu çalışmada, Bloom'un Yenilenmiş Taksonomisi kapsamında Chris Drew'un çalışması, Ayşe Gülçin Ertekin'in raporu ve Millî Eğitim Bakanlığı'nın çevre konuları kazanımları ve Mercan'ın çalışması uzman soruları olarak kullanılmıştır (Drew, 2023; Ertekin, 2017; MEB, 2018, Mercan, 2019).

Prompt tasarımı:

1. Yeniden Yapılandırılmış Bloom Taksonomisinin Bilişsel Süreç Basamağı [1-6]:
2. [Basamak adı], [Basamak tanımı]
3. Bu basamağa uygun sorular üretmeni istiyorum.
4. Öğretimde kullanılan dil gereği soruları emir kipiyle oluşturmalısın.
5. Soruların bilişsel süreç basamağını içermesini sağlamalısın.
6. Sorularda bu fiilleri kullanmalısın: [basamak fiilleri]
7. Soru kelimelerini kullanmamalısın, soruları verilen fiillerle tamamlamalısın.
8. Soruları tek bir fiil ile ifade etmelisin.
9. Her soruda aynı fiili tekrarlamamaya çalış.
10. Sorunun fiille anlam bütünlüğü oluşturmasına dikkat et.
11. Soruların [...] basamağında olmasına dikkat et. [...] basamağı [düşük/yüksek] bilişsel seviyeli basamaktır. [Basamak yetenekleri] ve bu kapasiteleri sorularla ortaya çıkarmayı hedefler.
12. Soru ve konuları tablo formatında tut. Sütunlar: Soru, konu
13. Talimatlara uyararak her konudan birer soru oluşturacaksın.
14. [Konular]

4.1.2. Sinema filmi içerik verileri

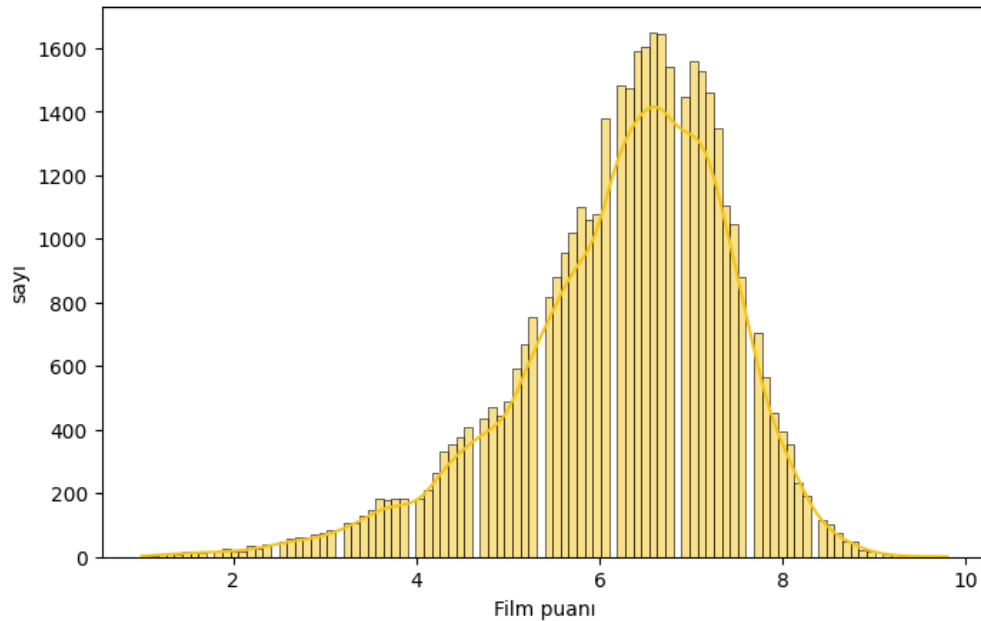
Son zamanlarda, film derecelendirmeleri ve tür seçimi, psikolojik yorumlar da dahil olmak üzere çeşitli alanlara dair içgörüler sunabilir (Puhazholi ve Francis, 2023). Sinema filmlerinin tür ve puan tahmini amacıyla The Movie Database (TMDB) ve IMDb veri tabanlarından elde edilen bilgiler kullanılarak kapsamlı bir analiz yapılmıştır. Bu çalışma kapsamında filmlerin kısa özetleri, tür tanımları, oylama bilgileri, film detayları ve ekip

bilgileri bir araya getirilmiştir. Çizelge 4.2. de kullanılan dosyalar ve dosyaların kullanılan içerikleri ve dövüş kulübü (fight club) filminin bilgileri örnek olarak verilmiştir.

Çizelge 4.2. Çevrim içi veri tabanları ve ilgili sinema filmleri içerikleri

Dosya	İçerik	Örnek
tmdb-movies-dataset-2023	Film özetleri	Fight Club film özeti: “A ticking-time-bomb insomniac and a slippery soap salesman channel primal male aggression into a shocking new form of therapy. Their concept catches on, with underground "fight clubs" forming in every town, until an eccentric gets in the way and ignites an out-of-control spiral toward oblivion.”
help.imdb.com	Tür tanımları	Drama türü tanımı: “Should contain numerous consecutive scenes of characters portrayed to effect a serious narrative throughout the title, usually involving conflicts and emotions. This can be exaggerated upon to produce melodrama. Subjective.”
title.ratings.tsv.gz	Oylama bilgileri	tconst/id: “tt0137523”, oy sayısı: “2233131”, oylama puanı: “8.8”
title.basics.tsv.gz	Yıl, süre, tür	yıl “1999”, süre (dk): “139”, tür: “Drama”
title.principals.tsv.gz	Oyuncular, yönetmen, yazar, besteci, yapımcı, editör ve görüntü yönetmeni	aktörler: “nm0001533, nm0000093, nm0340260, nm0001570”, yönetmen: “nm0000399”, yapımcılar “nm0149556, nm0068501, nm0513165”, yazarlar: “nm0657333, nm0880243”, aktris, görüntü yönetmeni, besteci, editör bilgileri bulunmuyor “0”

Veri tabanı ‘basics.tsv’ kullanılarak film türü "movie" olan başlıklar seçilmiş ve tconst, primaryTitle, startYear ve runtimeMinutes bilgileri kullanılmıştır. Veri tabanı ‘ratings.tsv’ tconst üzerinden eşleştirilmiş ve averageRating ve numVotes bilgileri kullanılmıştır. Veri tabanı ‘title.principals.tsv’ tconst üzerinden eşleştirilmiş ve nconst ve category sütunları kullanılmıştır. Bu kategoriler film ekibi bileşenlerinden oluşmaktadır ve bunlardan film sayısından az olan dolayısıyla çoğu film için anlam ifade etmeyen kategoriler çıkarılmıştır. Geriye kalan aktris, aktör, yönetmen, yazar, besteci, yapımcı, editör ve görüntü yönetmeni (actor, actress, producer, writer, editor, composer, director, cinematographer) bileşenleri öznitelik olarak kullanılmıştır. 1000’den az oy alan filmler daha anlamlı bir modelleme için filtrelenmiştir. Film türü "movie" olan ve 1000’den fazla oy alan filmler seçildiğinde 40,911 film elde edilmiştir. Film puanlarının film sayılarına göre dağılımı Şekil 4.1.’de gösterilmiştir. Örneğin “Esaretin Bedeli” filmi en yüksek IMDb puanına sahiptir ve 9,3 puandır. Puanlama 9,3’ten başlayarak 0’ a doğru ilerlemektedir.



Şekil 4.1. Sinema filmleri IMDb puan dağılımı

Filmlerin TMDB’deki kısa özetler ve IMDb’deki tür tanımları kullanılarak film türleri için doğrulama modeli yapılmıştır. Filmlere ait olan türler IMDb’den alınmıştır. Bir filme ait en fazla 3 en az 1 tür bulunmaktadır. Bu nedenle eğer bir filmin birden fazla türü varsa ayrı ayrı girdi olarak ele alınmıştır. 40,911 filminden nadir türler (verinin binde birinden az örnekte bulunan), TMDB’de özeti bulunmayan filmler veya seri halindeki filmlerin tekrarlayan

özetleri çıkarıldığında 40,376 film kalmıştır. Bu filmlere ait birden fazla tür olarak tanımlanan filmlerde her bir tür için bir kayıt tekrarı yapılarak veri seti ne dahil edilmiştir. Ayrıca bu durum tüm çoklu türe sahip filmler için uygulandığında veri seti boyutu 94,347 örnek şeklinde büyümüştür. Filmlerin ait olduğu tür sayıları Çizelge 4.3.'de verilmiştir. Toplamda 22 farklı tür bulunmaktadır. Türlerin film sayılarına göre dağılımı Şekil 4.2. de pasta dilim grafiğinde gösterilmiştir. Filmlere dahil olmayan türler negatif tür olarak alınarak veri çoğaltma yapılmıştır. Negatif örnekler IMDB'ye göre o filmin türü olmayan türlerden seçilmiştir. Filme ait kaç tür varsa o kadar ait olmayan türlerden seçilerek negatif örnekleri oluşturulmuştur. Bu şekilde 94,347 adet negatif örnek oluşturulmuştur. Ayrıca negatif örnekler oluşturulurken daha eşit bir dağılım yapılmış ve film türündeki dağılım daha dengeli hale getirilmiştir. Böylece modelin bazı türleri daha az görmesinden doğabilecek genelleme probleminin önüne geçilmiştir. Örneğin Kara Şövalye (2008) filmine ait IMDb'nin tür bilgisinin bulunduğu "title.principals.tsv.gz" veri tabanında aksiyon, suç ve drama türleri bulunmaktadır, bu film için 22 türden bu üç tür haricindeki 19 türden rastgele olarak tarih, gizem ve fantezi türleri ait olmayan negatif olarak atanmıştır.

Çizelge 4.3. Ait olduğu tür sayılarına göre film sayısının dağılımı

1 türe ait film sayısı	7240
2 türe ait film sayısı	12301
3 türe ait film sayısı	20835

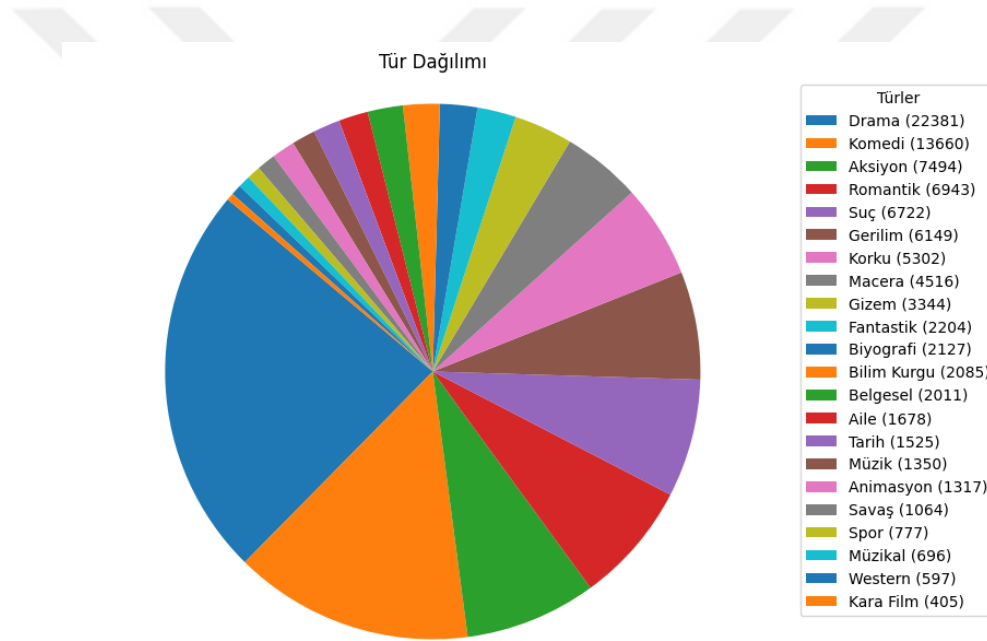
Veri çoğaltma birden fazla türe ait olan verilerin bölünmesi ile ait olan pozitif türlerde ve ait olmayan türler için pozitif sayısı kadar negatifleri oluşturularak veri çoğaltma yapılmıştır. Veri çoğaltma ile oluşturulan veri seti ile doğrulama modeli geliştirilmiştir. Ait olan ve olmayan türleri model tanımların verilmesi ile tahmin edecek şekilde eğitilmiştir. Örneğin "Kara Şövalye" filminin özetine karşılık "aksiyon" türünün tanımı verildiğinde film türü olduğunu veya "tarih" türünün tanımı verildiğinde filmin bu türde olmadığı modelin çıktısıdır.

Filmlerin tür doğrulaması için cümle gömme (Reimers ve Gurevych, 2020) yöntemini kullanan bir evrişimli sinir ağı (CNN) modeli geliştirilmiştir. Pozitif ve negatif örnekler için karşıt öğrenme (contrastive learning) kullanılmaktadır. Girdi olarak filmin özetini ve türün tanımı alan model türün o filme ait olup olmadığını doğrulamasını yapmaktadır. Verinin

%80'i eğitim, %10'u doğrulama ve %10'u test olarak ayrılmıştır ve Çizelge 4.4.'de örnek sayıları verilmiştir.

Çizelge 4.4. Film türü sınıflandırması için eğitim, doğrulama ve test sayıları

Bölünme	Ait olan tür sayısı	Ait olmayan tür sayısı
Eğitim verisi	75477	75477
Doğrulama verisi	9435	9435
Test verisi	9435	9435



Şekil 4.2. Sinema filmleri IMDb tür dağılımı

4.1.3. Mental sağlık açısından sosyal medya gönderileri

Sosyal medya verileri, kullanıcıların oluşturduğu çok miktardaki içerikten yararlanarak duygusal durumlar ve psikolojik koşullar hakkında içgörüler sağlamak için değerli bir kaynak oluşturur. Literatüre bakıldığında bu kaynaklar kullanılarak yapılmış birçok çalışma mevcuttur. Bu çalışmalar içinde gerek depresyon (Kabir ve ark., 2023), gerek anksiyete (Thakur, 2024), gerek intihar eğilimi (Ramírez-Cifuentes ve ark., 2020) teşhisleri yapılmaktadır. Bu çalışmada ise daha geniş ve kapsamlı bir yaklaşım benimsenerek benzer verilerin birleştirilmesiyle yeni bir veri seti oluşturularak yapay zeka modelleri ile performans

analizi yapılmıştır. Bu kapsamda, sosyal medya platformu çeşitliliği de sağlanmış ve her sınıf için en az iki farklı platformdan veri elde edilmiştir. Veri çeşitliliği sağlanan ve üç mental rahatsızlık ve normal durum olarak dört sınıf üzerinden sınıflandırma çalışması yapılmıştır. Twitter (X), Instagram ve Reddit sosyal medya platformlarından çeşitli veri setleri kullanılarak intihar eğilimi, depresyon, anksiyete ve normal durum analiz edilmiştir.

Instagram anksiyete veri seti: Bu veri seti, "Mpox Narrative on Instagram: A Labeled Multilingual Dataset of Instagram Posts on Mpox for Sentiment, Hate Speech, and Anxiety Analysis" başlıklı çalışmasından alınmış olup Instagram gönderilerinden oluşmaktadır (Thakur, 2024). Farklı dillerde yazılmış gönderiler içeren bu veri setinden, yalnızca İngilizce olan gönderiler filtrelenmiştir. Etiketlerin "No Stress/Anxiety Detected" veya "Stress/Anxiety Detected" olarak ayrıldığı veri seti gönderileri, anksiyete içerikli ve normal gönderiler olarak ilgili sınıflara dahil edilmiştir.

Twitter intihar eğilimi veri setleri: Burada kullanılan Twitter veri setleri kaggle veri bilimi platformundan edinilmiş iki veri seti ve github s.ürüm kontrol sistemi depolama alanı platformundan edinilmiş bir veri setinden oluşmaktadır. Veri setleri, intihar eğilimi gösteren ve göstermeyen gönderilerden oluşmaktadır. İlk veri setinde (Mahmud, 2023) "Potential Suicide post" etiketi taşıyan intihar eğilimli gönderileri ve "Not Suicide post" etiketi taşıyan normal gönderileri için kullanılmıştır. İkinci Twitter veri setinde (İnternet, 2024) "intention" etiketinin "1" olarak atandığı gönderiler intihar eğilimi sınıfı ve "0" olarak atandığı gönderiler normal sınıfı için kullanılmıştır. Üçüncü Twitter veri seti (Yadav, 2023) ön işlem yapılmamış, sadece Twitter platformuna ve intihar eğilimi etiketine sahip olan veri setidir. Bu veri seti sadece intihar eğilimi etiketi için kullanılmıştır.

Twitter depresyon Veri Seti: Bu birleşik Twitter veri seti Harvard Dataverse platformundan edinilmiştir (Helmy, 2024). Etiketlerin Depressed/0 ve Non-depressed/1 olarak ayrıldığı veri seti gönderileri depresyon ve normal gönderiler olarak ilgili sınıflara dahil edilmiştir.

Reddit Ruh Sağlığı Veri Setleri: Reddit platformundan alınan ve farklı dönemlere ait ruh sağlığı gönderilerini içeren bu veri setleri, "Natural Language Processing Reveals Vulnerable Mental Health Support Groups and Heightened Health Anxiety on Reddit During COVID-19" (Low ve ark., 2020) çalışmasından elde edilmiştir. Bu veri setleri, içinde intihar, anksiyete ve depresyon başlıklarının da bulunduğu çeşitli Reddit başlıklarının ilgili gönderilerini içerir. Veri seti farklı dönemlere göre (COVID-19 öncesi, 2018 ve 2019 ve

COVID-19 orta dönemi) gruplandırılmıştır ve ilgili etiketlerin her döneme ait örneklerinden belirli sayılarda alınarak normal sınıfı hariç her bir sınıf için (anksiyete, intihar eğilimi, depresyon) kullanılmıştır.

Veride daha eşit bir dağılım ve daha iyi bir genelleme sağlamak için sayıca fazla olan verilerden belirli sayılarda örnekler alınmıştır. Metin uzunluğu (kelime sayısı) kısaltmalarının (2-510) ardından kalan verilerin eşit dağılımına göre yapılmıştır. Instagram anksiyete verisi ve Twitter intihar eğilimi verileri 7000’şer civarı kalmış ve Twitter depresyon verisinden de 7000 seçilmiştir. Reddit anksiyete verilerinden normal olmayan üç sınıf için 14.000’ e tamamlayacak şekilde veriler eklenmiştir. Reddit verilerinden normal etiketli veri alınmamış ve Twitter intihar eğilimi veri setlerinde normal 6000’e yakın bir sayıda kaldığı için Instagram anksiyete ve Twitter depresyon verilerinden 6000’er daha alınarak 18.000 normal etiketli veri seçilmiştir (random state 42).

Dört ayrı veri setinden alınan normal gönderileri “Normal” etiketi ile toplam 18.000, iki veri setinden alınan anksiyete gönderileri “Anxiety” etiketiyle toplam 14.000, iki veri setinden alınan depresyon gönderileri “Depression” etiketiyle toplam 14.000 ve dört veri setinden alınan intihar eğilimi gönderileri “Suicidal” etiketi ile toplam 14.000 veri kullanılmıştır. Toplam veri 60.000 gönderiden oluşmaktadır. Mental sağlık verilerinin sınıflara göre veri istatistikleri Çizelge 4.5.’de verilmiştir. Veri sınıfların ve kaynakların eşit dağılımı ile %70’i (42.000) eğitim (train) verisi, %10’u (6.000) doğrulama (validation) verisi ve %20’si (12.000) test verisi olarak ayrılmıştır.

Çizelge 4.5. Mental sağlık gönderilerinin sınıflara göre veri istatistikleri

Sınıf	Normal	Anksiyete	Depresyon	İntihar e.
Gönderi sayısı	18000	14000	14000	14000
Kelime sayısı	50605	61849	33820	35414
Ort. kel. sayısı	27	120	80	112
Eğitim verisi	12600	9800	9800	9800
Doğrulama verisi	1800	1400	1400	1400
Test verisi	3600	1800	1800	1800

4.2. Öznitelik elde edilmesi

Geleneksel olarak metin benzerliği, metinlerin vektör temsilleri ile bu vektörler arası mesafe uzunluğunu hesaplanarak ölçülür. Bu yöntemin kullanıldığı en popüler metriklerden biri kosinüs benzerliğidir (Wang ve Dong, 2020). Bu tez çalışmasında öznitelik elde etmek için metinler sayısal temsillere dönüştürülmüştür. Sayısal temsiller çeşitli metin gömme yöntemleri kullanılarak elde edilmiştir.

Bloom Taksonomisine göre sınıflamada soruların vektörleri BERT modeli ile alınmıştır. Bert modelinin girdi metinlerini ayırdığı her token için çok başlı dikkat ve ardından gelen feed-forward ağının çıktılarının işlenmiş en son vektörlerinin ortalaması alınmıştır. El edilen vektör her token için 768 boyutunda son gizli katman vektörünün ortalaması olarak 768 boyutundadır.

Film puanı tahmininde öznitelikler filmin yılı, süresi, verilen oy sayısı ve kategorilerin değerlerinden oluşmaktadır. Kategorilerde aktris ve aktör bir filmde en fazla 10 adet bulunduğu için 10 öznitelik kaplar ve eğer bir filmin 10'dan az oyuncusu varsa padding (doldurma) uygulanarak "0" ile doldurulur. Aynı şekilde, yönetmen yazar, yapımcı, editör, besteci ve görüntü yönetmeni de en fazla 4 değerden oluştuğu için 4'er öznitelik kaplar ve eğer bir filmin bu kategorilerde değeri 4'ten az ise padding uygulanarak "0" ile doldurulur. Sonuç olarak her film için 48 boyutunda öznitelik vektörü elde edilmiş ve bu vektörler regresyon ile film puanı tahmininde kullanılmıştır.

Film tür tahmini için filmin özetleri ve tür tanımları, Sentence Transformer kullanılarak semantik anlamları yakalayan sayısal temsillere dönüştürülmüştür. Metin vektörleri Bert modelinin Sentence Transformer (bi-encoder) sinir ağıyla eğitilmiş "paraphrase-MiniLM-L6-v2" modeli (Reimers, 2019) kullanarak alınmıştır. Bu vektör 384 boyutundadır (normal Bert vektörü 768, bi encoder 768/2) ve tek cümle olmayan girdilerin de (verimizde bazı filmlerin özetleri birkaç cümle) tek girdi olarak alır ve 384'lük vektöre dönüştürür.

Mental sağlık sosyal medya gönderileri için metin gömme yöntemlerinden kelime bazında, kelimenin sıklığı bazında ve bağlamsal bazda sayısal temsil yöntemleri seçilmiştir.

Word2Vec kelime odaklı metin temsilleri elde edilen makine öğrenmesi temelli bir modelidir (Mikolov ve ark., 2013). Bu işlem, metinleri kelimelere ayırıp (tokenization) her kelimeyi yüksek boyutlu bir vektörle temsil eder. Model, kelimeler arasındaki semantik ilişkileri

öğrenir ve yakın anlamlı kelimeleri vektör uzayında birbirine yakın konumlandırır. Metinler, içerdikleri kelimelerin vektörlerinin ortalaması alınarak belge düzeyinde bir vektörle özetlenir. Bu temsiller, ham metni matematiksel bir forma dönüştürerek makine öğrenimi modelleri veya metin analizi görevlerinde kullanılabilir hale getirir. Metal sağlık gönderileri için Wor2Vec temsil yöntemlerinden biri olan CBOW (Continuous bag of words) kullanılmıştır. CBOW Word2vec, her kelimenin anlamsal anlamına dikkat edebildiği için daha iyi kelime gömme (word embedding) üretebilmektedir (Isnain ve ark., 2020). Oluşturulan vektör 768 boyutundadır ve en az 10 kez geçen kelimeler dikkate alınmış ve bir kelimenin bağlamını oluşturan pencere boyutu 5 kelime olarak ayarlanmıştır. CBOW algoritmasıyla kelime vektörlerini oluşturur ve bu vektörlerin ortalamasını alarak belge vektörlerini elde eder. Belge vektörü, belgedeki kelimelerin anlamsal özelliklerini özetleyen bir tekil vektör olarak temsil edilir.

TF-IDF (Term Frequency- Inverse Document Frequency), kelimenin belge içerisindeki frekansına dayalı metin temsilleri elde edilen istatistiksel bir modeldir (Ramos, 2003). TF-IDF yöntemiyle bir belge koleksiyonundaki kelimelerin ağırlıkları hesaplanarak bir matris oluşturulur ve sadece harflerden oluşan anlamlı kelimeler seçilerek boyut azaltımı yapılır. Bu süreçte, sık kullanılan işlevsiz kelimeler (stop-words) ve sayılar gibi ilgisiz özellikler filtrelenmiştir, böylece hem metin temsilleri optimize edilir hem de daha verimli ve genellenebilir modeller oluşturulur. Eğitim verisi üzerinde TF-IDF matrisi oluşturulmuştur. Bu matris filtrelenerek sadece alfabetik kelimeler, tüm belgelerde en az 291 kere geçen ve tüm belgelerin %15'inden az geçen kelimeler seçilmiştir. Sonuç olarak 768 boyutunda metin girdilerini temsil eden vektörler oluşturulmuştur.

Bağlamsal temsiller ön eğitilmiş transformatör tabanlı büyük dil modellerinden elde edilmiştir. Transformer modelleri (çalışmamızda BERT ve RoBERTa), metinlerdeki her bir kelime veya token için bağlama duyarlı (contextual) yüksek boyutlu vektörler üretir. Bu işlem sırasında metin, modele uygun şekilde tokenize edilmiş ve kısa metinler ise padding token'larıyla doldurulur. Modelin çıktısı olarak, her token için bağlamını dikkate alan bir vektör elde edilmiştir. Bu vektörler, kelimenin diğer kelimelerle ilişkisini temsil etmektedir. Daha sonra, padding token'larının etkisini kaldırmak için dikkat maskesi kullanılarak yalnızca gerçek token vektörleri seçilmiştir. Bu vektörler toplanır ve ortalaması alınarak metni temsil eden tek bir vektör oluşturulur. Sonuçta, her metin için bağlama duyarlı anlamını özetleyen 768 boyutlu bir vektör elde edilmiştir. Metinde bulunan her token için

son gizli katman vektörünün ortalaması o metnin vektörü olarak alınmıştır. Bu vektör çok başlı dikkat ve ardından gelen feed-forward ağın çıktılarının işlenmiş vektörlerinin ortalaması olan vektördür. Bert modeli ile metin temsilleri iki yönlü bağlamsal analiz yapması ile anlamsal sonuçlar verebilmektedir. Diğer kelime temsil yöntemlerine kıyasla daha başarılı sonuçlar elde ettiği çeşitli çalışmalar mevcuttur (Özçift ve ark.,2021; Savcı ve Das, 2023).

Ayrıca deneysel çalışmada bu vektörlerin birleşimi olan hibrit vektör elde edilmiştir. Hibrit vektörler Bert ve Roberta ile ayrı ayrı olarak Word2Vec ve TF- IDF vektörlerinin birleşimlerinden elde edilmiştir. Vektörlerin hibrit şekilde kullanımının tekil kullanımlara göre performansı analiz etmek için birleştirilmiş vektör elde edilmiştir.

4.3. Yapay Zekâ Modellerinin Performans Analizi

Deneysel çalışmada kullanılan yapay zekâ modelleri her görev için ayrı olarak belirlenmiştir. Farklı yöntemlerin etkisi için mental sağlık sosyal medya gönderileri üzerinde farklı yaklaşımlar test edilmiştir.

4.3.1. Bloom Taksonomisine göre soru üretimi ve sınıflandırması

Destek Vektörlerinde çeşitli çekirdek türleri, veri setlerini farklı şekillerde yüksek boyutlu uzaylara dönüştürmekte kullanılır (Cervantes ve ark., 2020). Bu tez çalışmasında doğrusal (linear) çekirdek kullanılmıştır. Karar fonksiyonu çoklu sınıflandırmada kullanılan ‘ovo’ (One-vs-One) her sınıf çifti için bir ikili sınıflandırma problemine bölmesi için seçilmiştir. Destek vektörleri yapay soruların %80 ile eğitilmiş ve %20’si test olarak ayrılmıştır. Çizelge 4.6.’de yapay sorularla eğitilen destek vektörleri modelinin yapay soruları ve uzman soruları sınıflandırmasının F1 skoru sonuçları verilmiştir.

Linear kernel: $K(x_i; x_j) = (x_i * x_j)$ (skaler çarpım)

Çizelge 4.6. Üretilen yapay sorular ve uzman soruları sınıflandırmalarının F1 Skoru Sonuçları

Bloom Taksonomisi	Copilot		Gemini		Chat GPT	
	Uzman S.	Yapay S.	Uzman S.	Yapay S.	Uzman S.	Yapay S.
Hatırlama	0.596	0.928	0.702*	0.603	0.596	0.824
Anlama	0.529	0.788	0.448	0.548	0.554	0.783
Uygulama	0.448	0.710	0.525	0.794	0.255	0.769
Analiz	0.421	0.841	0.346	0.722	0.327	0.829
Değerlendirme	0.303	0.928	0.517	0.743	0.698	0.939
Yaratma	0.745	0.928	0.526	0.928	0.646	0.909

4.3.2. Sinema filmi içeriklerinin analizi

40.911 film için birleşik IMDb veri tabanlarından öznitelikler kullanılarak dört regresyon modeli eğitilmiştir ve %10'luk bir kısım test için ayrılmıştır. Regresyon için veri işleme adımlarında “title.principals.tsv.gz” dosyasındaki film ekibi bileşenleri pivot matris ile filmlere göre ayrılmıştır. Daha sonra en fazla bileşen olan filmlere göre vektör boyutları belirlenmiş (aktör ve aktris için 10 değerleri için 4) ve film ekibinin ‘nconst’ bilgileri ya da boş olanları için ‘0’ ile doldurulmuştur. Tür bilgileri film puanları için girdilerden biri olarak kullanılmış ve her tür ve kombinasyonu için ayrı bir kod atanmıştır. Oy sayıları 0 ile 0.99 skalasında ölçeklendirilmiş ve süre sayısal değişkene çevrilmiştir.

Gradient Boosting, CatBoost, XGBoost ve Karar Ağacı (DT) modellerine ait performans metrikleri – MSE, MAE, R-kare ve RMSE Çizelge 4.7.’de özetlenmiştir. Modeller arasında ortak parametreler; 1.000 iterasyon, 0,1 öğrenme oranı ve 8 derinlik olup, bunlar doğruluk ve hesaplama verimliliği arasında denge sağlamak amacıyla optimize edilmiş değerlerdir. Bu modeller arasında Gradient Boosting algoritması, 0,711 MSE, 0,607 MAE, 0,485 R-kare ve 0,843 RMSE ile en iyi performansı göstermiştir.

Çizelge 4.7. Regresyon analizi sonuçları

Metrik	catboost	xgboost	Gradient Boosting	Karar Ağacı
MSE	0.765	0.756	0.711	1.034
MAE	0.629	0.621	0.607	0.749
R2	0.446	0.452	0.485	0.251
RMSE	0.874	0.869	0.843	1.017

Tür doğrulaması için 2 boyutlu Evrişimsel Sinir Ağları (Convolutional Neural Network, CNN) kullanılmıştır. Metin eşleştirme yaklaşımıyla pozitif ve negatif öneklerden öğrenme ve ait olan veya olmayan türleri tahmin etme yapılmıştır. IMDb'ye göre en fazla 3 türe ait 40,376 filmin o filme ait kaç tür varsa o kadar ait olmayan türlerden seçilerek negatif örnekleri oluşturulmuştur. Tür doğrulaması için veri işleme kısmında TMDB'den film özetleri ve IMDb'nin yardım sayfasında bulunan tür tanımları çekilmiştir. Daha sonra bu metinler cümle gömme yöntemi ile vektöre dönüştürülmüş ve modele bu sayısal temsiller verilmiştir. Tür doğrulaması skorları ait olan ve olmayan türler için Çizelge 4.8.'de eşleştirme skorları verilmiştir.

Çizelge 4.8. Tür doğrulaması skorları

Değer	Precision	Recall	F1-score	Test sayısı
Ait tür	0.84	0.86	0.85	9435
Ait olmayan tür	0.85	0.83	0.84	943

4.3.3. Mental sağlık açısından sosyal medya gönderilerinin sınıflandırması

LLM tabanlı modeller, önceden eğitilmiş dil modelleri (pre-trained) olmaları nedeniyle genelleme kapasiteleri yüksektir, genel özellikleri veri içinde kolay öğrenir ve başarıyı artırır. Bu tez çalışması için seçilen modeller açık kaynaklıdır ve kullanımları ücretsizdir. Sınıflandırma görevleri için düşük parametrelili LLM özellikle bellek yönetimi ve eğitim süresi açısından verimlidir ve büyük veri setlerinde bile başarılı sonuçlar verir.

Derin öğrenme modellerinde eğitim süreci parametrelerini ayarlarken modele ve göreve uygun olmasına dikkat edilmelidir. Batch boyutu, modelin bir defada işleminden geçirdiği ve güncelleme için kullandığı veri örneği sayısını ifade eder. Daha büyük batch boyutları daha hızlı ancak daha çok bellek gerektirirken, küçük batch boyutları daha fazla detay sağlar ve daha az bellek kullanır.

Model eğitimi için veri seti metin (text dataset) veya tensör (tensor dataset) yapısında kullanılabilir. Metin veri seti yapısı metin verileri için özel bir veri sınıfıdır, her öge için etiketleri döndürür. Tensör veri seti yapısı PyTorch'un hazır sınıfıdır, tensör verilerle doğrudan çalışır.

Epoch, modelin tüm eğitim veri setini bir kez işlediği bir döngüyü ifade eder. Her epoch, modelin öğrenme sürecinin devam etmesine olanak tanır ve modelin performansını artırmak için tekrar eden güncellemeler sağlar. Ancak çok fazla epoch kullanılması, aşırı öğrenmeye yol açabilir.

Öğrenme oranı (learning rate), modelin ağırlıklarının ne hızda güncelleneceğini belirleyen hiper parametredir, küçük bir değer eğitim sürecini yavaşlatabilir, büyük bir değer aşırı öğrenme gibi durumlara yol açabilir. Bu yüzden modele ve gerçekleştirilecek işleme bağlı olarak seçilmelidir.

Aşırı öğrenme (overfitting), bir makine öğrenimi modelinin eğitim verisine fazla uyum sağlaması sonucu, yeni verilere karşı düşük performans göstermesi durumudur. Dropout, bir nöronun devre dışı bırakılma olasılığını ifade eder. Eğitim sırasında, her ağırlık güncellemesinde bazı nöronlar rastgele kapanır ve ağırlık bu nöronları geçici olarak kullanmaması sağlanır.

Optimize edici algoritmalar derin öğrenme modellerini eğitmek ve modelin genelleme performansını artırmak için kullanılır. Standart Adam algoritmasının geliştirilmiş bir versiyonu olan AdamW'nin genelleme yeteneği, özellikle eğitim sürecinde daha düşük eğitim kaybı ve test hatalarına yol açar (Loshchilov, 2017)

Dinamik öğrenme oranı, derin öğrenme modellerinin eğitiminde stabilite ve verimlilik sağlar. Öğrenme oranı planlayıcıları (schedulers), her eğitim adımında oranı ayarlayarak modelin farklı eğitim aşamalarına uyumunu kolaylaştırır. Doğrusal planlayıcılar, başlangıçta düşük bir öğrenme oranı ile başlayıp, aşamalı olarak artan bir oran kullanarak modelin

kontrollü bir şekilde öğrenmesini destekler. Bu yaklaşım, döngü sayısı arttıkça modelin stabilitesini korumaya yardımcı olur.

Aktivasyon fonksiyonları modelin çıktılarını belirli bir biçimde dönüştüren ve modelin tahmin yeteneğini artıran matematiksel işlemlerdir. Çoklu sınıf görevlerinde Softmax aktivasyon fonksiyonu kullanılır ve her sınıfa ait tahminleri 1'e normalize eder. Sigmoid iki sınıflı tek çıktılı (binary) görevlerde kullanılır ve iki sınıf arasındaki olasılığı hesaplayarak modelin karmaşıklığını azaltır.

Hiper parametre optimizasyonu, hangi parametrelerin gerçekleşecek görev için en iyi performansı sergileyeceğini bulmaya yarayan bir yöntemdir. Makine öğrenmesi sistemlerinde temel görev, hiper parametreleri otomatik olarak ayarlayarak performansı en üst düzeye çıkarmaktır (Feurer ve Hutter, 2019). Optuna, dinamik arama alanı oluşturma ve verimli optimizasyon stratejileri ile hem araştırma hem de endüstri için kolayca uygulanabilen, güçlü bir hiperparametre optimizasyon aracıdır (Akiba ve ark., 2019). Optimizasyon denemesinde çeşitli hiper parametrelerin dinamik olarak önerilmesini sağlar ve budama stratejisi "başarısız olma ihtimali yüksek" olan denemelerin erken aşamalarda elenmesine (budanmasına) olanak tanır. Optuna ayrıca model optimizasyonuna da olanak tanır. Optuna, F1 skorundaki değişimleri izleyerek en iyi performansı sağlayan hiper parametre kombinasyonlarını bulur. Bu yöntemler, en iyi performansı veren denemeler üzerinde yoğunlaşarak zaman ve hesaplama kaynaklarının daha verimli kullanılmasını sağlar.

Çoklu sınıflandırma problemi için RoBERTa, BERT, XLNet ve T5 önceden eğitilmiş dil modelleri üzerinde ince ayar yapılmıştır. Sınıflandırma için decoder LLM performansını görmek amacıyla XLNet deneylere dahil edilmiştir. Bir encoder-decoder model olan T5 encoder modunda kullanılmıştır. Her model, sınıflandırma görevini 4 etiket üzerinden gerçekleştirmek üzere eğitilmiştir; "normal", "anxiety" (anksiyete), "depression" (depresyon) ve "suicidal" (intihara eğilimli).

Tüm modeller aynı eğitim parametreleri kullanılarak eğitilmiştir. Optuna ile yapılan çeşitli deneyler sonucu her model için optimum parametreler belirlenmiş ve model karşılaştırmalarında ortak olarak kullanılmıştır. Maksimum giriş uzunluğu Twitter gönderilerinin daha kısa olmasından dolayı 256 olarak belirlenmiştir. Batch boyutu,

RoBERTa, BERT ve XLNet için 32, T5 için 16 kullanılmıştır. Eğitimler GPU kullanılarak gerçekleştirilmiştir. Modellerin eğitimde kullanılan parametreler Çizelge 4.9.'de verilmiştir.

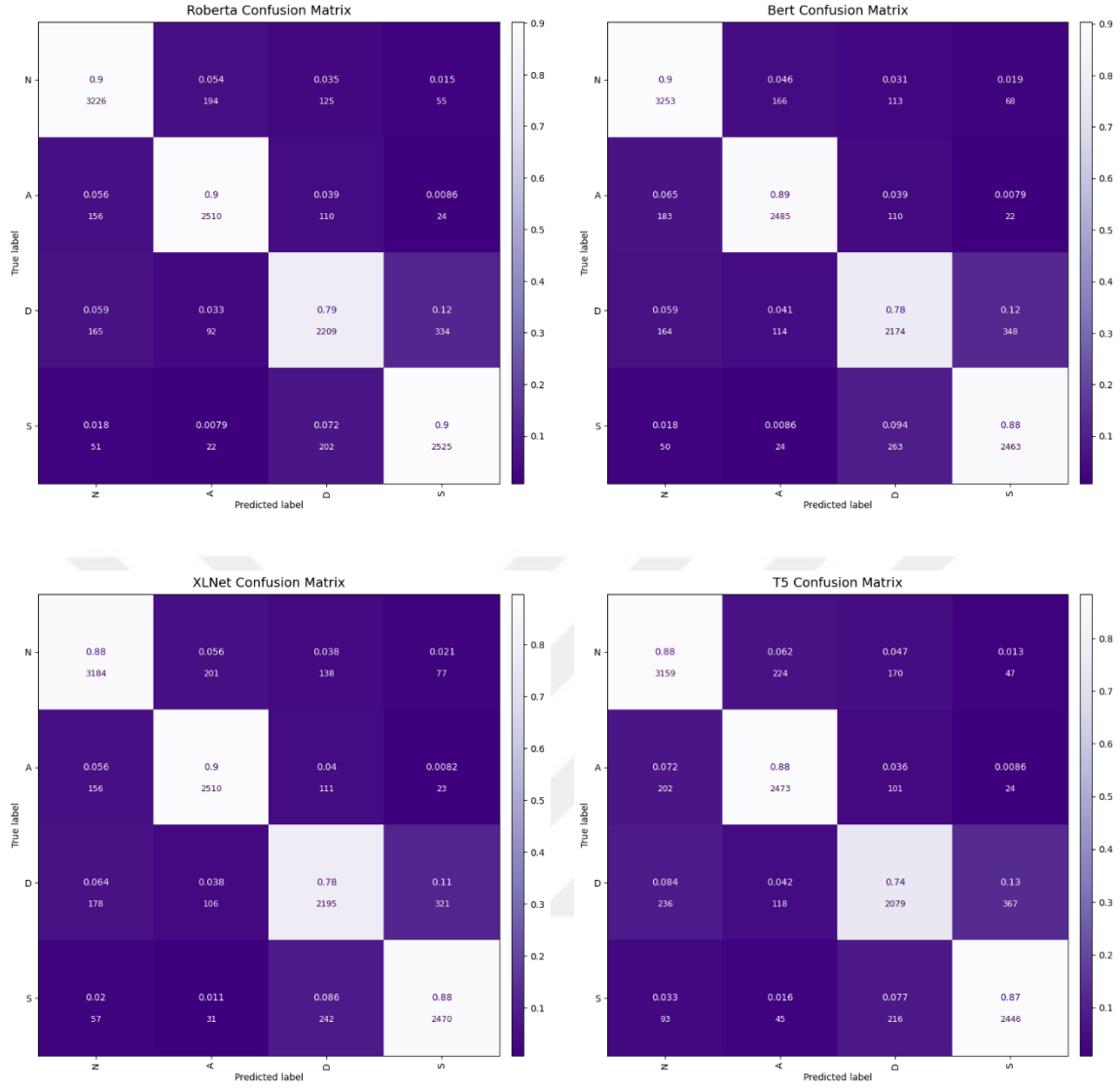
Çizelge 4.9. LLM ince ayar eğitimlerinde kullanılan parametreler

Parametre	Kullanılan
Veri seti Yapısı	Tensor Dataset
Epoch	3
Batch Boyutu	16/32
Maksimum giriş uzunluğu	256
dropout	0.1
Learning rate	3e -5
Scheduler	linear
Optimizasyon	AdamW
Aktivasyon Fonksiyonu	Softmax

Büyük dil modellerinin sınıflandırma sonuçları alınırken (Çizelge 4.10.) sadece tekrar eden ifadelerin kaldırılması ön işlemi uygulanmıştır. Roberta ince ayarı ile en yüksek başarımlar ve F1 skoru elde edilmiştir. Ayrıca Normal sınıfı için Roberta ve BERT, anksiyete sınıfı için Roberta ve XLNet, depresyon ve intihar eğilimi sınıfları için Roberta en yüksek sınıflandırma doğruluğunu vermiştir. Şekil 4.3.'de her model için sınıf bazlı doğruluk ve o sınıfa ait doğru tahmin sayıları karmaşıklık matrisleri ile verilmiştir.

Çizelge 4.10. Büyük dil modellerinin sınıflandırma skorları

Skor	RoBERTa	BERT	XLNet	T5
Acc	0.872*	0.864	0.863	0.846
F1	0.872*	0.864	0.862	0.845



Şekil 4.3. Büyük dil modellerinin sınıflandırma performansı

Çalışmamızda metin ön işleme teknikleri adım adım kullanılarak sınıflandırmaya etkilerine bakılmıştır. Genel olarak literatürde metin ön işleme adımlarının uygulamanın ve durdurma kelimelerinin kaldırılmasının sınıflamaya ve göreve olumlu etki yaptığı görülmektedir (Samad ve ark., 2020; Ladani ve Desai, 2020; Jefriyanto ve ark., 2023; Rahimi ve Homayounpour, 2023; HaCohen-Kerner ve ark., 2023; Fan ve ark., 2023; Jung ve ark., 2023). Ancak bu tez çalışmasında, mental sağlık açısından sosyal medya gönderilerinin sınıflandırılmasında harflerin küçük-büyük olmasının, farklı karakterlerin, bağlantıların, boşlukların, durdurma kelimelerinin ve sözcüklere getirilen eklerin anlamlı olduğu sonucuna varılmıştır. Çizelge 4.11.'de görüldüğü üzere her metin ön işleme adımı başarıyı biraz daha azaltmıştır. Çizelge 4.11. verilen karakter kısıtlaması, İngilizce karakterleri ve belli noktalama işaretlerini (.,;!?"'-) bırakarak diğer karakter türlerini filtreleme işlemidir.

Çizelgedeki “filtrelemeler” ifadesi URL, HTML etiketleri, bağlantılar ve bunlara ait metinlerin ve noktalama işaretlerinin kaldırılmış olduğu ön işlemi ifade etmektedir.

Çizelge 4.11. Farklı ön işleme tekniklerinin Roberta ince ayar sonuçları

Skor	tekrar eden ifadeleri kald.	ön işlemsiz	+ harf küçültme	+ karakter kısıtlaması	+ filtreleme ve temizlik	+ durdurma k. kaldır.	+ kökleştirme
Acc	0.872	0.872	0.871	0.869	0.862	0.828	0.810
F1	0.872	0.871	0.871	0.868	0.861	0.827	0.809

Tekrar eden ifadeleri kaldırma Şekil 4.3.’te uygulanan işlemdir ve sınıflandırma doğruluğu en yüksek ön işlem olmuştur. Tekrar eden ifadelerin kaldırılmasının etkisiyle modelin hali hazırda öğrendiği yapıların tekrar tekrar verilmesinin anlamsız olduğu hatta sınıflandırma doğruluğunu bir miktar düşürebileceği görülmüştür. Şekil 4.4.’te metin ön işlemlerinin adım adım ve sınıf sınıf Roberta modelin ince ayarı gerçekleştirilmesinde elde edilen karmaşıklık matrislerinde verilmiştir. Dikkat çeken bir nokta küçük harfe dönüştürmenin üç sınıfa etkisi nispeten olumsuzken depresyon sınıfına etkisi olumlu olmuş ve depresyon sınıfı doğruluğunu 0.80’ e çıkardığı görülmüştür.



Şekil 4.4. Metin ön işlem tekniklerinin sınıflara göre performans analizi

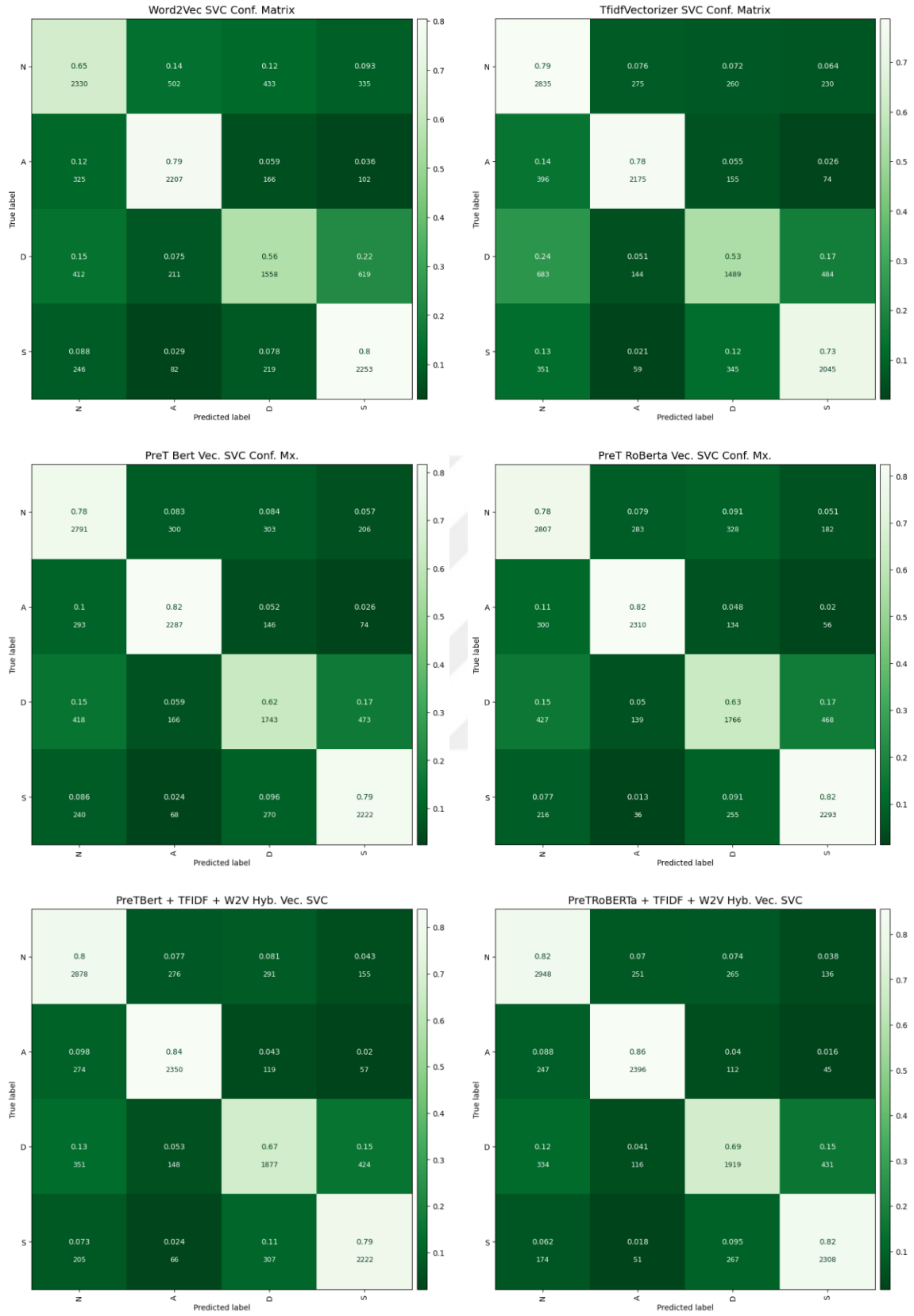
Mental sađlık gnderileri iin sınıflandırma alıřmamızda farklı metin gmme yntemleri ile destek vektrleri ve ift ynl uzun kısa vadeli bellek modelleri performansı analiz edilmiřtir.

Destek vektrleri iin dođrusal (linear) ekirdek ve ‘ovo’ karar fonksiyonu kullanılmıřtır. Word2vec, TF-IDF, Bert ve Roberta ile uygulanan metin gmme yntemleri ve bunların birleřtirilmesi ile elde edilen hibrit vektrlerin sınıflamaya etkisi ve destek vektrlerinin sınıflandırma performansı analiz edilmiřtir. izelge 4.12.’de destek vektrleri iin sınıflandırma skorları verilmiřtir. En yksek bařarım n eđitimli Roberta, word2vec ve TF-IDF vektrlerinin birleřtirilmesiyle oluřan hibrit vektr ile elde edilmiřtir.

izelge 4.12. Destek vektrleri sınıflandırma sonuları

Skor	Word2Vec	TF-IDF	Bert	Roberta	Bert + W.+ T.	Roberta + W. + T.
Acc	0.692	0.709	0.752	0.763	0.776	0.796
F1	0.695	0.712	0.753	0.764	0.777	0.797

Destek vektrleri ile yapılan sınıflandırma alıřmalarının her vektr iin karmařıklık matrisi Őekil 4.5.’te verilmiřtir. Hibrit vektr kombinasyonlarının tekil kullanımlara gre daha bařarılı sonular verdiđi grlmüřtir.



Şekil 4.5. Destek vektörlerinin farklı vektör alma yöntemleri ile sonuçları

Geleneksel makine öğrenimi yöntemlerinden uzun kısa vadeli bellek modelinin (LSTM) metnin hem öncesini hem sonrasını dikkate alan çift yönlü geri besleme dahil olan Bi-LSTM yöntemi ile farklı metin gömme yöntemlerinin sınıflandırma performansı test edilmiştir.

Eğitimde Tensor Dataset yapısı kullanılmış ve model, 10 epoch boyunca eğitilmiştir. Batch boyutu 16, 128 gizli katman, ve 0.1 dropout parametreleri belirlenmiştir. Model, AdamW optimizasyon algoritması ile $1e-3$ öğrenme oranı kullanılarak eğitilmiş ve çıktı aktivasyonu için Softmax fonksiyonu tercih edilmiştir. Çizelge 4.13.'de Bi-LSTM modeli için kullanılan eğitim parametreleri özetlenmiştir. Kullanılan parametrelerin açıklamaları bu bölümün (4.3.3.) bölümün başında verildiği gibidir. Hiper parametre yöntemi olarak grid search (ızgara arama) kullanılmış ve uygun parametreler seçilmiştir.

Çizelge 4.13. Bi-LSTM modellerinde kullanılan parametreler

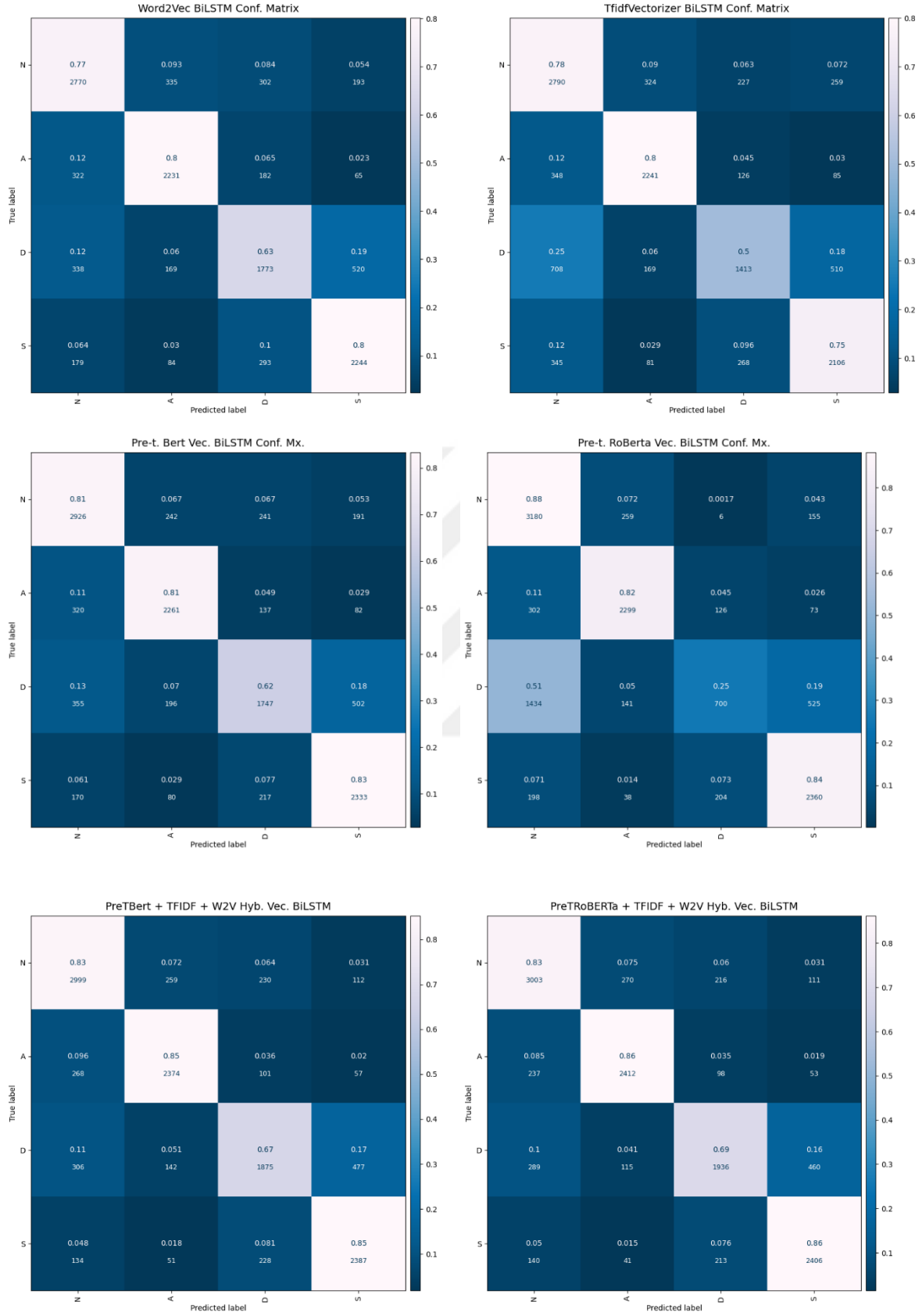
Parametre	Kullanılan
Veri seti Yapısı	Tensor Dataset
Epoch	10
Batch Boyutu	16
Gizli katman	128
dropout	0.1
Learning rate	$1e-3$
Scheduler	linear
Optimizasyon	AdamW
Aktivasyon Fonksiyonu	Softmax

Çizelge 4.14.'de farklı vektör temsil yöntemleriyle Bi-LSTM modeli kullanılarak elde edilen sınıflandırma sonuçları verilmiştir. Roberta + Word2Vec + TF-IDF kombinasyonu, en yüksek doğruluk %81,1 ve F1 skoru %81,3 ile en iyi performansı göstermiştir. Word2Vec ve TF-IDF gibi klasik yöntemler ise daha düşük başarı oranlarına sahiptir.

Çizelge 4.14. Bi LSTM sınıflandırma sonuçlar

Skor	Word2Vec	TF-IDF	Bert	Roberta	Bert + W.+ T.	Roberta + W. + T.
Acc	0.750	0.708	0.770	0.684	0.801	0.811
F1	0.751	0.712	0.772	0.711	0.802	0.813

Roberta vektörü Bert vektörüne göre sınıflandırmada daha düşük başarımlar gösterirken hibrit vektör kombinasyonunda Roberta ile olan birleşim en yüksek başarımları vermiştir. Destek vektörlerinde TF-IDF word2vec'den daha başarılı sonuç verirken, çift yönlü uzun kısa vadeli bellek ile modellemede word2vec TF-IDF'den daha başarılı sınıflandırma gerçekleştirmiştir. Her sınıfın ne kadar doğrulukla sınıflandırıldığı karmaşık matrisleriyle Şekil 4.6.'da her model için verilmiştir. Ön eğitilmiş Roberta modelinin tekil kullanımı normal sınıfı için en yüksek başarımları vermiştir. Diğer sınıflar için hibrit vektörler daha yüksek sınıflandırma sonucu vermiş ve destek vektörlerinde olduğu gibi Bi-LSTM ile sınıflandırmada da ön eğitilmiş Roberta, word2vec ve TF-IDF'in birleştirilmesi ile oluşturulmuş hibrit vektör en yüksek başarımlar ve F1 skoru sonucunu vermiştir.



Şekil 4.6. Bi-LSTM modelinin farklı vektör alma yöntemleri ile sonuçları

5. SONUÇ VE ÖNERİLER

Bu tez çalışmasında yapay zekâ modellerinin soru üretimi, sınıflandırma ve tahmin performanslarının analizi için deneysel çalışmalar yürütülmüştür. Bu çalışmalar kapsamında;

Yapay zekâ modellerinin içerik üretimi eğitim bilimleri alanında soru ve kazanım tasarımı için kullanılan Bloom taksonomisine göre incelenmiştir. Yapay soruların test verisi ve literatürden elde edilmiş uzman soruları ile başarımları ölçülmüş ve kontrol verisi olan uzman sorularının sınıflandırılmasında yapay test verilerinden daha düşük sonuçlar elde edilmiştir. Bu çalışma yapılırken nicel bir referans noktası olan fiillerden yararlanılmıştır. Ancak üretim daha nitel bir analiz gerektiren taksonomiye göre olması yapay zekâ modellerinin başarımları etkilemiştir. Uzman sorularıyla karşılaştırıldığında yapay zekâ üretimlerinin talimatlara uygun soru oluşturması taksonominin hedeflerini tam olarak karşılayamadığı gibi bir sonuç çıkmıştır. Bu da hedef gösterilen kategorilerde dile özgü nüanslar ve nicel olmayan beklentiler söz konusu olduğunda ve yapay zekâ modellerinin bu sınıflara ait büyük verilerle eğitilmediğinde zayıf kalabileceği görülmüştür. Yapay zekâların daha kolay anlayacağı ve daha az incelik gerektiren konular dışında ve eğitim bilimleri gibi kapsamlı alanlarda kullanımının dikkatle yapılması daha uygun olacaktır. Dilbilimci ve düşünür Chomsky, insan ve makine tarafından üretilen metinleri karşılaştırarak, 'dilin doğru açıklamaları karmaşıktır ve yalnızca büyük veri içinde yoğunlaşarak öğrenilemez' ve bu örüntülerin insan zekasına kıyasla yüzeysel ve şüpheli kalabileceğini çünkü rasyonellikle sınırlı olmadıklarını belirtmiştir (Chomsky ve ark., 2023).

Sinema filmlerinin içerik analizinde yıl, süre, kategori, yönetmen oyuncu gibi özelliklerin film puanlarına etkisi olduğu görülmüştür. Bu özneliklerin giriş olarak kullanıldığı regresyon analizinde Gradient Boosting algoritmasıyla 0.711 MSE ve 0.607 MAE gibi düşük kayıplarla başarılı bir film puanı tahmin modeli geliştirilmiştir. Ayrıca filmlerin kısa özetlerinden (bir veya birkaç cümlelik) ve filmler için IMDb'de tanımlanan türler ve ait olmayan türler ile tür doğrulama yapılmıştır. Tür doğrulama için geliştirilen 2 boyutlu CNN modelinde türlerin tanımları kullanılmış ve girdinin bir boyutu filmin özetinin sayısal temsili diğer boyutu türün tanımının sayısal temsili olarak verilmiştir. Sadece bu kısa özetler ve filmlerin ait olduğu ve olmadığı türlerin tanımlarıyla karşıt öğrenme yapılarak %84,45

başarım elde edilmiştir. Bu sınırlı girdiler varken modelin daha iyi öğrenmesini sağlamak amacıyla strateji geliştirmeye bir örnek teşkil etmektedir.

Aydın ve Irmak (2023) intihar eğilimi veri setinde ikili sınıflandırma yaparak %96,61 doğruluk elde etmişlerdir. Kaygı, depresyon ve stres tahminleri makine öğrenimi algoritmaları kullanılarak yapılan çalışmada depresyon için en yüksek F1 skoru Naive Bayes 0,836, stres için Rastgele Orman 0,711 ile elde edilmiş, anksiyete için ise hem Karar Ağacı hem de Naive Bayes ile en yüksek doğruluk değeri 0,733 sağlanmıştır (Priya ve ark., 2020). Cabral ve ekibi MM-EMOG ile sosyal medya gönderilerinden mental sağlık durumlarını sınıflandırmak için geliştirilmiş bir çoklu etiket duygu grafiği temsili sunmuşlardır ve Bert modeli ile %78'e kadar doğruluk elde etmişlerdir (Cabral ve ark., 2024). Ayrıca sağlık metinleri için özelleştirilmiş ClinicalBERT (Wang ve ark., 2023) ve mental sağlık metinleri için özelleştirilmiş MentalBERT (Ji ve ark., 2021) modelleri Bert modeline ince ayar yapılarak eğitilmiş göreve özel modellerdir.

Karmaşıklık matrislerinde sınıf bazlı sonuçlarımıza bakıldığında depresyon etiketli gönderilerin en çok intihar eğilimi gönderileri ile karıştığını görüyoruz. Ancak bu olası bir durumdur çünkü depresyon intiharla yakından ilişkilidir (Shneidman, 1993). Metin sayısal temsilleri ile özniteliklerinin çıkarılması için farklı yöntemler ele alınmış ve sınıflandırma performansı incelenmiştir. Ayrıca metin temsilleri olan vektörlerin hibrit kullanımının sınıflandırma performansına etkisi incelenmiştir. Birleştirilerek oluşturulmuş veri üzerinde çoklu sınıflandırmalarda daha az metin ön işleme adımının daha yüksek başarımla sonuçlandığı görülmüştür. Sadece tekrarlayan ifadeleri kaldırarak uygulanan ön işlemden sonra yapılan Roberta modeli ince ayarı ile %87,2 başarımla ve F1 skoru elde edilmiştir. Bi-LSTM modeli eğitiminde ön eğitilmiş Roberta, word2vec ve TF-IDF vektörlerinin birleştirilmesiyle oluşturulmuş hibrit vektör tekil kullanımlarına göre sınıflandırma başarımlarını %6 ile %13 arası artırmış ve 0.811'e çıkarmış F1 skorunu %6 ile %10 arası artırmış ve 0.813'e çıkarmıştır. Bağlamsal temsiller (LLM model vektörleri), sıralı yapıdaki veriyle çalışan BiLSTM'ler için kritik bir avantaj sağladığı görülmüştür. Hibrit yöntemler, farklı gömme yöntemlerinin güçlü yönlerini birleştirerek daha dengeli ve daha başarılı temsiller oluşturulmuştur. Metin ön işleme yöntemlerinin çoğaltılarak metnin eksilmesi doğruluk ve F1 skorlarını düşürmüştür.

Bu tez çalışmasında üç farklı problem farklı yöntem ve tekniklerinin değerlendirilmesi ile alınmış ve yapay zekâ yöntemlerinin doğal dil işlemede etkin kullanılması için incelenmiştir.

KAYNAKLAR

- Abarja, R. A., & Wibowo, A. (2020). Movie rating prediction using convolutional neural network based on historical values. *Int. J.*, 8, 2156-2164.
- Akiba, T., Sano, S., Yanase, T., Ohta, T., & Koyama, M. (2019, July). Optuna: A next-generation hyperparameter optimization framework. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining* (pp. 2623-2631).
- Akyön, F. Ç., Çavuşoğlu, A. D. E., Cengiz, C., Altınuç, S. O., & Temizel, A. (2022). Automated question generation and question answering from Turkish texts. *Turkish Journal of Electrical Engineering and Computer Sciences*, 30(5), 1931-1940.
- Alanazi, S. S., Elfadil, N., Jarajreh, M. ve Algarni, S. (2021). Question answering systems: a systematic literature review. *International Journal of Advanced Computer Science and Applications*, 12(3).
- Altınok, D. (2023, July). A diverse set of freely available linguistic resources for Turkish. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 13739-13750).
- Amari, S. (1967). A theory of adaptive pattern classifiers. *IEEE Transactions on Electronic Computers*, (3), 299-307.
- Ameer, I., Arif, M., Sidorov, G., Gómez-Adorno, H., & Gelbukh, A. (2022). Mental illness classification on social media texts using deep learning and transfer learning. *arXiv preprint arXiv:2207.01012*.
- An, T. M., Watts, S., & Tung, T. M. (2024). Investigating the Interaction Between Social Media Use and Depression among the Young. *Journal of Electrical Systems*, 20(4s), 2490-2503.
- Anderson, L. W., Krathwohl, D. R. (2001). A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives.
- Aninditya, A., Hasibuan, M. A., & Sutoyo, E. (2019, November). Text mining approach using TF-IDF and naive Bayes for classification of exam questions based on cognitive level of bloom's taxonomy. In *2019 IEEE International Conference on Internet of Things and Intelligence System (IoTaIS)* (pp. 112-117). IEEE.
- Ansari, G., Garg, M., & Saxena, C. (2021). Data augmentation for mental health classification on social media. *arXiv preprint arXiv:2112.10064*.
- Atayolu, Y., & Kutlu, Y. (2024). Similarity and Classification Analysis in Turkish Question Generation of AI tools According to Bloom's Taxonomy. *Tethys Env. Sci*, 1(2), 87-98.

- Awan, M. J., Khan, R. A., Nobanee, H., Yasin, A., Anwar, S. M., Naseem, U., & Singh, V. P. (2021). A recommendation engine for predicting movie ratings using a big data approach. *Electronics*, 10(10), 1215.
- Aydın, T., & Irmak, M. C. (2023, December). ELECTRA ve XLNET Modellerini Kullanarak X Verilerinden İntihara Meyilli İçerikleri Tespit Etme. In *SETSCI-Conference Proceedings* (Vol. 15, pp. 45-50). SETSCI-Conference Proceedings.
- Borges, J.L, (1925), *Inquisiciones.*, p. 65-66
- Brachman, R. J. (1979). On the epistemological status of semantic networks. In *Associative networks* (pp. 3-50). Academic Press.
- Brachman, R. J. ve Schmolze, J. G. (1989). An overview of the KL-ONE knowledge representation system. *Readings in artificial intelligence and databases*, 207-230.
- Cabral, R. C., Han, S. C., Poon, J., & Nenadic, G. (2024). MM-EMOG: Multi-Label Emotion Graph Representation for Mental Health Classification on Social Media. *Robotics*, 13(3), 53.
- Cervantes J., Garcia-Lamont F., Rodríguez-Mazahua L., Lopez A., A comprehensive survey on support vector machine classification: Applications, challenges and trends *Neurocomputing*, 408 (2020), pp. 189-215, 10.1016/j.neucom.2019.10.118
- Chai, C. P. (2023). Comparison of text preprocessing methods. *Natural Language Engineering*, 29(3), 509-553.
- Chomsky, N. (1956). Three models for the description of language. *IRE Transactions on information theory*, 2(3), 113-124.
- Chomsky, N. *Syntactic Structures*. The Hague: Mouton, 1957.
- Chomsky, N. (1959). On certain formal properties of grammars. *Information and control*, 2(2), 137-167.
- Chomsky, N., Roberts, I., & Watumull, J. (2023). Noam chomsky: The false promise of chatgpt. *The New York Times*, 8.
- Collins, A. M. ve Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior* 8, 240-247.
- Çakir, M. U. ve Güldamlasioğlu, S. (2016, June). Text mining analysis in Turkish language using big data tools. In *2016 IEEE 40th annual computer software and applications conference (COMPSAC)* (Vol. 1, pp. 614-618). IEEE.
- De Choudhury, M., Gamon, M., Counts, S., & Horvitz, E. (2013). Predicting depression via social media. In *Proceedings of the international AAAI conference on web and social media* (Vol. 7, No. 1, pp. 128-137).

- Devlin, J., Chang, M. W., Lee, K. ve Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.
- Doshi, P., & Zadrozny, W. (2018). Movie genre detection using topological data analysis. In *Statistical Language and Speech Processing: 6th International Conference, SLSP 2018, Mons, Belgium, October 15–16, 2018, Proceedings 6* (pp. 117-128). Springer International Publishing.
- Drew, C. (April 24, 2023). 119 Bloom's taxonomy examples. Helpful Professor. <https://helpfulprofessor.com/blooms-taxonomy-examples/>
- Eddy, S. R. (1996). Hidden markov models. *Current opinion in structural biology*, 6(3), 361-365.
- Elkins, S., Kochmar, E., Serban, I., & Cheung, J. C. (2023, June). How useful are educational questions generated by large language models?. In *International Conference on Artificial Intelligence in Education* (pp. 536-542). Cham: Springer Nature Switzerland.
- Ertekin, A.G. (2017). Öğrenme kazanımları ve sınav sorularının yenilenmiş BLOOM taksonomisine göre incelenmesi: İş ve meslek danışmanı ulusal yeterliliği örneği.
- Ertugrul, A. M., & Karagoz, P. (2018, January). Movie genre classification from plot summaries using bidirectional LSTM. In *2018 IEEE 12th International Conference on Semantic Computing (ICSC)* (pp. 248-251). IEEE.
- Fan, Y., Arora, C., & Treude, C. (2023, May). Stop words for processing software engineering documents: Do they matter?. In *2023 IEEE/ACM 2nd International Workshop on Natural Language-Based Software Engineering (NLBSE)* (pp. 40-47). IEEE.
- Feurer, M., & Hutter, F. (2019). Hyperparameter optimization. *Automated machine learning: Methods, systems, challenges*, 3-33.
- Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 1189-1232.
- Fukushima, K. (1980). Neocognitron: A self-organizing neural network for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, 36(4), 193–202.
- Gemirter, C. B., & Goularas, D. (2021). A Turkish question answering system based on deep learning neural networks. *Journal of Intelligent Systems: Theory and Applications*, 4(2), 65-75.
- Graves, A., & Schmidhuber, J. (2005, July). Framewise phoneme classification with bidirectional LSTM networks. In *Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005.* (Vol. 4, pp. 2047-2052). IEEE.

- Gruebner, O., Sykora, M., Lowe, S. R., Shankardass, K., Galea, S., & Subramanian, S. V. (2017). Big data opportunities for social behavioral and mental health research.
- HaCohen-Kerner, Y., Miller, D., & Yigal, Y. (2020). The influence of preprocessing on text classification using a bag-of-words representation. *PloS one*, 15(5), e0232525.
- Hancock, J. T., & Khoshgoftaar, T. M. (2020). CatBoost for big data: an interdisciplinary review. *Journal of big data*, 7(1), 94.
- Hastie, T. (2009). *The elements of statistical learning: data mining, inference, and prediction*.
- Hebb, D. O. *The organization of behavior*. New York: Wiley, 1949.
- Helmy, A., 2024, "Depression dataset for English tweets classified binary", <https://doi.org/10.7910/DVN/ISZCSA>, Harvard Dataverse, V1
- Hoang, Q. (2018). Predicting movie genres based on plot summaries. arXiv preprint arXiv:1801.04813.
- Hoperoft, J. E., & Ullman, J. D. (1979). *Introduction to automata theory, languages, and computation*. Addison-Wesley Publishing Company.
- Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the national academy of sciences*, 79(8), 2554-2558.
- IMDb. (n.d.). title.ratings.tsv. Retrieved September 21, 2023, 23:27 TRT, from <https://datasets.imdbws.com/title.ratings.tsv.gz>
- IMDb. (n.d.). title.basics.tsv. Retrieved September 21, 2023, 23:54 TRT, from <https://datasets.imdbws.com/title.basics.tsv.gz>
- IMDb. (n.d.). title.principals.tsv. Retrieved September 22, 2023, 00:13 TRT, from <https://datasets.imdbws.com/title.principals.tsv.gz>
- Isnain, A. R., Sihabuddin, A., & Suyanto, Y. (2020). Bidirectional long short term memory method and Word2vec extraction approach for hate speech detection. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 14(2), 169-178.
- Izacard, G. ve Grave, E. (2020). Leveraging passage retrieval with generative models for open domain question answering. arXiv preprint arXiv:2007.01282.
- Jackendoff, R. (1996). How language helps us think. *Pragmatics & Cognition*, 4(1), 1-34.
- Jang, B., Kim, M., Harerimana, G., Kang, S. U., & Kim, J. W. (2020). Bi-LSTM model to increase accuracy in text classification: Combining Word2vec CNN and attention mechanism. *Applied Sciences*, 10(17), 5841.

- Jefriyanto, J., Ainun, N., & Al Ardha, M. A. (2023). Application of Naïve Bayes Classification to Analyze Performance Using Stopwords. *Journal of Information System, Technology and Engineering*, 1(2), 49-53.
- Ji, S., Zhang, T., Ansari, L., Fu, J., Tiwari, P., & Cambria, E. (2021). Mentalbert: Publicly available pretrained language models for mental healthcare. *arXiv preprint arXiv:2110.15621*.
- Jones, K. S. (1994). Natural language processing: a historical review. *Current issues in computational linguistics: in honour of Don Walker*, 3-16.
- Jordania, J. (2006). Who asked the first question. The origins of human choral singing, intelligence, language and speech. Tblisi, Georgia: Logos.
- Joshi, A. K. (1991). Natural language processing. *Science*, 253(5025), 1242-1249.
- Jung, G., Shin, J., & Lee, S. (2023). Impact of preprocessing and word embedding on extreme multi-label patent classification tasks. *Applied Intelligence*, 53(4), 4047-4062.
- Kabir, M., Ahmed, T., Hasan, M. B., Laskar, M. T. R., Joarder, T. K., Mahmud, H., & Hasan, K. (2023). DEPTWEET: A typology for social media texts to detect depression severities. *Computers in Human Behavior*, 139, 107503.
- Karpukhin, V., Oğuz, B., Min, S., Lewis, P., Wu, L., Edunov, S., ... ve Yih, W. T. (2020). Dense passage retrieval for open-domain question answering. *arXiv preprint arXiv:2004.04906*.
- Kıyak, Y. S. (2023). A ChatGPT prompt for writing case-based multiple-choice questions. *Revista Española de Educación Médica*, 4(3).
- Kotsiantis, S. B., Zaharakis, I., & Pintelas, P. (2007). Supervised machine learning: A review of classification techniques. *Emerging artificial intelligence applications in computer engineering*, 160(1), 3-24.
- Kolomiyets, O. ve Moens, M. F. (2011). A survey on question answering technology from an information retrieval perspective. *Information Sciences*, 181(24), 5412-5434.
- Ladani, D. J., & Desai, N. P. (2020, March). Stopword identification and removal techniques on tc and ir applications: A survey. In *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)* (pp. 466-472). IEEE.
- Laddha, M. D., Lokare, V. T., Kiwelekar, A. W., & Netak, L. D. (2021). Classifications of the summative assessment for revised blooms taxonomy by using deep learning. *arXiv preprint arXiv:2104.08819*.
- Lafferty, J., McCallum, A., & Pereira, F. (2001, June). Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In *Icml* (Vol. 1, No. 2, p. 3).

- LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural computation*, 1(4), 541-551.
- Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., ... ve Zettlemoyer, L. (2019). Bart: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. *arXiv preprint arXiv:1910.13461*.
- Li, Z., Liu, F., Yang, W., Peng, S., & Zhou, J. (2021). A survey of convolutional neural networks: analysis, applications, and prospects. *IEEE transactions on neural networks and learning systems*, 33(12), 6999-7019.
- Linnainmaa, S. (1970). The representation of the cumulative rounding error of an algorithm as a Taylor expansion of the local rounding errors (Doctoral dissertation, Master's Thesis (in Finnish), Univ. Helsinki).
- Little, W. A., & Shaw, G. L. (1978). Analytic study of the memory storage capacity of a neural network. *Mathematical biosciences*, 39(3-4), 281-290.
- Liu, Y. (2019). Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*, 364.
- Loshchilov, I. (2017). Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*.
- Luxton, D. D., June, J. D., & Fairall, J. M. (2012). Social media and suicide: a public health perspective. *American journal of public health*, 102(S2), S195-S200.
- Martínez-Líbano, J., & Yeomans-Cabrera, M. M. (2024, October). Depression, anxiety, and stress in the Chilean Educational System: children and adolescents post-pandemic prevalence and variables. In *Frontiers in Education* (Vol. 9, p. 1407021). Frontiers Media SA.
- McCallum, A., Freitag, D., & Pereira, F. C. (2000, June). Maximum entropy Markov models for information extraction and segmentation. In *Icml* (Vol. 17, No. 2000, pp. 591-598).
- McCarthy, J. (1960). Recursive functions of symbolic expressions and their computation by machine, part I. *Communications of the ACM*, 3(4), 184-195.
- McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5, 115-133.
- MEB. (2018). Fen Bilimleri dersi öğretim programı (İlkokul ve Ortaokul 3., 4., 5., 6., 7. ve 8. sınıf). Ministry of National Education (Turkey), Ankara: Millî Eğitim Bakanlığı Yayınları.

- Mercan, S. I. (2019). Sosyal Bilgiler Öğretmen Adaylarının Yenilenmiş Bloom Taksonomisi Bilişsel Basamaklarına Göre Soru Sorma Becerilerinin İncelenmesi. *Third Sector Social Economic Review*, 54/1 (2019) 291-301
- Minsky, M., & Papert, S. (1969). *An introduction to computational geometry*. Cambridge tiass., HIT, 479(480), 104.
- Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*
- M. Nair, A., & Preethi, N. (2022). A pragmatic study on movie recommender systems using hybrid collaborative filtering. In *IoT and Analytics for Sensor Networks: Proceedings of ICWSNUCA 2021* (pp. 489-494). Springer Singapore.
- Mohammed, M., & Omar, N. (2020). Question classification based on Bloom's taxonomy cognitive domain using modified TF-IDF and word2vec. *PloS one*, 15(3), e0230442.
- Musa, J. M., & Zhihong, X. (2020, April). Item based collaborative filtering approach in movie recommendation system using different similarity measures. In *Proceedings of the 2020 6th International Conference on Computer and Technology Applications* (pp. 31-34).
- Naseem, U., Razzak, I., & Eklund, P. W. (2021). A survey of pre-processing techniques to improve short-text quality: a case study on hate speech detection on twitter. *Multimedia Tools and Applications*, 80, 35239-35266.
- Novikova, J., & Shkaruta, K. (2022). DECK: Behavioral tests to improve interpretability and generalizability of BERT models detecting depression from text. *arXiv preprint arXiv:2209.05286*.
- Öncü, M. H., & Kutlu, Y. (2024). Performance Comparison of Known AI Translation Tools. *Tethys Env. Sci*, 1(3), 117-126.
- Özçift, A., Akarsu, K., Yumuk, F., & Söylemez, C. (2021). Advancing natural language processing (NLP) applications of morphologically rich languages with bidirectional encoder representations from transformers (BERT): an empirical case study for Turkish. *Automatika: časopis za automatiku, mjerenje, elektroniku, računarstvo i komunikacije*, 62(2), 226-238.
- Pan, S. J., & Yang, Q. (2009). A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, 22(10), 1345-1359.
- Pasman, H. J., & Rogers, W. J. (2020). Logic based methods for dynamic risk assessment. In *Methods in Chemical Process Safety* (Vol. 4, pp. 61-122). Elsevier.
- Priya, A., Garg, S., & Tigga, N. P. (2020). Predicting anxiety, depression and stress in modern life using machine learning algorithms. *Procedia Computer Science*, 167, 1258-1267.

- Puhazholi, S., & Francis, F. S. (2023). Predicting Human Psychological Factors from Movie Ratings and Genres Using Genetic Algorithms. In 2023 International Conference on Network, Multimedia and Information Technology (NMITCON) (pp. 1-4). IEEE.
- Quillian, M. R. (1966). Semantic memory (pp. 227-270). Air Force Cambridge Research Laboratories, Office of Aerospace Research, United States Air Force.
- Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2), 257-286.
- Radford, A., Narasimhan, K., Salimans, T. ve Sutskever, I. (2018). Improving language understanding by generative pre-training.
- Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., ... ve Liu, P. J. (2020). Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of machine learning research*, 21(140), 1-67.
- Rahimi, Z., & Homayounpour, M. M. (2023). The impact of preprocessing on word embedding quality: A comparative study. *Language Resources and Evaluation*, 57(1), 257-291.
- Ramírez-Cifuentes, D., Freire, A., Baeza-Yates, R., Puntí, J., Medina-Bravo, P., Velazquez, D. A., ... & González, J. (2020). Detection of suicidal ideation on social media: multimodal, relational, and behavioral analysis. *Journal of medical internet research*, 22(7), e17758.
- Ramos, J. (2003, December). Using tf-idf to determine word relevance in document queries. In *Proceedings of the first instructional conference on machine learning* (Vol. 242, No. 1, pp. 29-48).
- Reimers, N. (2019). Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. arXiv preprint arXiv:1908.10084.
- Reimers, N., & Gurevych, I. (2020). Making monolingual sentence embeddings multilingual using knowledge distillation. arXiv preprint arXiv:2004.09813.
- Rogers, C. (1951). *Client Centered Therapy: Current Practice, Implications and Theory*. Houghton Mifflin, Boston.
- Rosenblatt, F. (1957). *The perceptron, a perceiving and recognizing automaton Project Para*. Cornell Aeronautical Laboratory.
- Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6), 386.
- Rosenblatt, F. (1962) *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms*. Spartan Books, Washington DC.

- Rudin, C., Chen, C., Chen, Z., Huang, H., Semenova, L., & Zhong, C. (2022). Interpretable machine learning: Fundamental principles and 10 grand challenges. *Statistic Surveys*, 16, 1-85.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *nature*, 323(6088), 533-536.
- Samad, M. D., Khounviengxay, N. D., & Witherow, M. A. (2020). Effect of text processing steps on twitter sentiment classification using word embedding. *arXiv preprint arXiv:2007.13027*.
- Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. *IBM Journal of research and development*, 3(3), 210-229.
- Sanwal, M., & ÇALIŞKAN, C. (2021). A hybrid movie recommender system and rating prediction model. *International Journal of Information Technology and Applied Sciences (IJITAS)*, 3(3), 161-168.
- Sarica, S., & Luo, J. (2021). Stopwords in technical language processing. *Plos one*, 16(8), e0254937.
- Savcı, P., & Das, B. (2023). Comparison of pre-trained language models in terms of carbon emissions, time and accuracy in multi-label text classification using AutoML. *Heliyon*, 9(5).
- Schmidhuber, J., & Hochreiter, S. (1997). Long short-term memory. *Neural Comput*, 9(8), 1735-1780.
- Schmidhuber, J. (2014). Who invented backpropagation. *More. [DL2]*.
- Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural networks*, 61, 85-117.
- Schmidhuber, J. (2022). Annotated history of modern ai and deep learning. *arXiv preprint arXiv:2212.11279*.
- Schölkopf, B., Burges, C. J., & Smola, A. J. (Eds.). (1999). *Advances in kernel methods: support vector learning*. MIT press.
- Shaikh, S., Daudpotta, S. M., & Imran, A. S. (2021). Bloom's learning outcomes' automatic classification using lstm and pretrained word embeddings. *IEEE Access*, 9, 117887-117909.
- Shneidman, E. S. (1993). Commentary: Suicide as psychache. *The Journal of nervous and mental disease*, 181(3), 145-147.
- Shorten, C., Khoshgoftaar, T. M. ve Furht, B. (2021). Text data augmentation for deep learning. *Journal of big Data*, 8(1), 101.

- Sipser, M. (1996). Introduction to the Theory of Computation. ACM Sigact News, 27(1), İkinci baskı 2006.
- Stanny, C. J. (2016). Reevaluating Bloom's Taxonomy: What measurable verbs can and cannot say about student learning. Education Sciences, 6(4), 37.
- Sutskever, I., Martens, J., & Hinton, G. E. (2011). Generating text with recurrent neural networks. In Proceedings of the 28th international conference on machine learning (ICML-11) (pp. 1017-1024).
- Svozil, D., Kvasnicka, V., & Pospichal, J. (1997). Introduction to multi-layer feed-forward neural networks. Chemometrics and intelligent laboratory systems, 39(1), 43-62.
- Szczepanek, R. (2022). Daily streamflow forecasting in mountainous catchment using XGBoost, LightGBM and CatBoost. Hydrology, 9(12), 226.
- Tang, Y., Tran, C., Li, X., Chen, P. J., Goyal, N., Chaudhary, V., ... ve Fan, A. (2020). Multilingual translation with extensible multilingual pretraining and finetuning. arXiv preprint arXiv:2008.00401.
- Tanvir, S., Abontee, A. P., Rayhan, M. A., Rahman, M. F., Sultana, T., & Ahmed, A. (2023). Movie Genre Prediction Using AI-Generated Data and Natural Language Processing. In 2023 26th International Conference on Computer and Information Technology (ICCIT) (pp. 1-4). IEEE.
- Thakur, N. (2024). MpoX Narrative on Instagram: A Labeled Multilingual Dataset of Instagram Posts on MpoX for Sentiment, Hate Speech, and Anxiety Analysis. arXiv preprint arXiv:2409.05292.
- Trappl, R. ve Payr, S. (2003). Emotions: From brain research to computer game development.
- Turing, A. M. (1936). On computable numbers, with an application to the Entscheidungsproblem. J. of Math, 58(345-363), 5.
- Turing, A. (1950). Computing machinery and intelligence. Mind, 59(236), pp.433–60.
- Turker, M., Ari, M. E. ve Han, A. (2024). Vbart: The turkish llm. arXiv preprint arXiv:2403.01308.
- Van Rooij, R. (2003). Questioning to resolve decision problems. Linguistics and Philosophy, 26, 727-763.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... ve Polosukhin, I. (2017). Attention is all you need. Advances in neural information processing systems, 30.
- Villa-Pérez, M. E., Trejo, L. A., Moin, M. B., & Stroulia, E. (2023). Extracting Mental Health Indicators From English and Spanish Social Media: A Machine Learning Approach. IEEE Access, 11, 128135-128152.

- Voleti, R., Liss, J. M., & Berisha, V. (2019). A review of automated speech and language features for assessment of cognitive and thought disorders. *IEEE journal of selected topics in signal processing*, 14(2), 282-298.
- Wang, T., & Bashir, M. (2020). Does social media behaviors reflect users' anxiety. A case study of twitter activities.
- Wang, J., & Dong, Y. (2020). Measurement of text similarity: a survey. *Information*, 11(9), 421.
- Wang, G., Liu, X., Ying, Z., Yang, G., Chen, Z., Liu, Z., ... & Chen, Y. (2023). Optimized glycemc control of type 2 diabetes with reinforcement learning: a proof-of-concept trial. *Nature Medicine*, 29(10), 2633-2642.
- Weizenbaum, J. (1966). ELIZA—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1), 36-45;44.
- Wikipedia contributors. (n.d.). Perceptrons (book). In Wikipedia. Retrieved August 3, 2024, from [https://en.wikipedia.org/wiki/Perceptrons_\(book\)](https://en.wikipedia.org/wiki/Perceptrons_(book))
- Winograd, T. (1972). Understanding natural language. *Cognitive Psychology*, 3(1), 1-191;8-15.
- Woods, W. A. (1970). "Transition Network Grammars for Natural Language Analysis." *Communications of the ACM*, 13(10), 591-606.
- Woods, W. (1972). The lunar sciences natural language information system. BBN report.
- Woods, W. A. (1973). An experimental parsing system for transition network grammars. *Natural language processing*, 111-154.
- Xue, L., Constant, N., Roberts, A., Kale, M., Al-Rfou, R., Siddhant, A., ... ve Raffel, C. (2020). mT5: A massively multilingual pre-trained text-to-text transformer. arXiv preprint arXiv:2010.11934.
- Yamada, I., Asai, A., Shindo, H., Takeda, H. ve Matsumoto, Y. (2020). LUKE: Deep contextualized entity representations with entity-aware self-attention. arXiv preprint arXiv:2010.01057.
- Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R. R. ve Le, Q. V. (2019). Xlnet: Generalized autoregressive pretraining for language understanding. *Advances in neural information processing systems*, 32.
- Zabaleta de Larrañaga, I. (2021). Using objective data from movies to predict other movies' approval rating through Machine Learning.
- Zhai, C. (2003). A brief note on the hidden Markov models (HMMs). Department of Computer Science, University of Illinois at Urbana-Champaign, IL, USA.

Zhang, T., Schoene, A. M., Ji, S., & Ananiadou, S. (2022). Natural language processing applied to mental illness detection: a narrative review. *NPJ digital medicine*, 5(1), 1-13.





teknoversite **AYRICALIĞINDASINIZ**

İSTE

