YAPAY ZEKÂ TEKNİKLERİNİN HUKUK ALANINDA UYGULANMASI

Mustafa Emirkan OKURSOY

T.C.
BURSA ULUDAĞ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YAPAY ZEKA TEKNİKLERİNİN HUKUK ALANINDA UYGULANMASI

Mustafa Emirkan OKURSOY
0009-0007-9166-2550

Prof. Dr. Tülin İNKAYA
(Danışman)

Dr. Öğr. Üyesi Emine Gökçe KARABEL
(2. Danışman)

YÜKSEK LİSANS TEZİ
ENDÜSTRİ MÜHENDİSLİĞİ ANABİLİM DALI

BURSA – 2025
Her Hakkı Saklıdır

TEZ ONAYI

Mustafa Emirkan OKURSOY tarafından hazırlanan "YAPAY ZEKÂ TEKNİKLERİNİN HUKUK ALANINDA UYGULANMASI" adlı tez çalışması aşağıdaki jüri tarafından oy birliği ile Bursa Uludağ Üniversitesi Fen Bilimleri Enstitüsü Endüstri Mühendisliği Anabilim Dalı'nda **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Danışman: Prof. Dr. Tülin İNKAYA

Başkan	**:**	Prof. Dr. Tülin İNKAYA 0000-0002-6260-0162 Bursa Uludağ Üniversitesi, Mühendislik Fakültesi, Endüstri Mühendisliği Anabilim Dalı	İmza
Üye	**:**	Doç. Dr. Burcu Çağlar GENÇOSMAN 0000-0003-0159-8529 Bursa Uludağ Üniversitesi, Mühendislik Fakültesi, Yöneylem Araştırması Anabilim Dalı	İmza
Üye	**:**	Doç. Dr. Mustafa Özgür CİNGİZ 0000-0003-4469-1440 Bursa Teknik Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, Bilgisayar Bilimleri Anabilim Dalı	İmza

Yukarıdaki sonucu onaylarım

Prof. Dr. Ali KARA
Enstitü Müdürü
../../….

B.U.Ü. Fen Bilimleri Enstitüsü tez yazım kurallarına uygun olarak hazırladığım bu tez çalışmasında;

- tez içindeki bütün bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi,
- görsel, işitsel ve yazılı tüm bilgi ve sonuçları bilimsel ahlak kurallarına uygun olarak sunduğumu,
- başkalarının eserlerinden yararlanılması durumunda ilgili eserlere bilimsel normlara uygun olarak atıfta bulunduğumu,
- atıfta bulunduğum eserlerin tümünü kaynak olarak gösterdiğimi,
- kullanılan verilerde herhangi bir tahrifat yapmadığımı,
- ve bu tezin herhangi bir bölümünü bu üniversite veya başka bir üniversitede başka bir tez çalışması olarak sunmadığımı

beyan ederim.

30/06/2025

MUSTAFA EMİRKAN OKURSOY

TEZ YAYINLANMA
FİKRİ MÜLKİYET HAKLARI BEYANI

Enstitü tarafından onaylanan lisansüstü tezin/raporun tamamını veya herhangi bir kısmını, basılı (kâğıt) ve elektronik formatta arşivleme ve aşağıda verilen koşullarla kullanıma açma izni Bursa Uludağ Üniversitesi'ne aittir. Bu izinle Üniversiteye verilen kullanım hakları dışındaki tüm fikri mülkiyet hakları ile tezin tamamının ya da bir bölümünün gelecekteki çalışmalarda (makale, kitap, lisans ve patent vb.) kullanım hakları tarafımıza ait olacaktır. Tezde yer alan telif hakkı bulunan ve sahiplerinden yazılı izin alınarak kullanılması zorunlu metinlerin yazılı izin alınarak kullandığını ve istenildiğinde suretlerini Üniversiteye teslim etmeyi taahhüt ederiz.

Yükseköğretim Kurulu tarafından yayınlanan "**Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge**" kapsamında, yönerge tarafından belirtilen kısıtlamalar olmadığı takdirde tezin YÖK Ulusal Tez Merkezi / B.U.Ü. Kütüphanesi Açık Erişim Sistemi ve üye olunan diğer veri tabanlarının (Proquest veri tabanı gibi) erişimine açılması uygundur.

Danışman Adı-Soyadı 30.06.2025 Prof. Dr. Tülin İNKAYA	Öğrencinin Adı-Soyadı 30.06.2025 Mustafa Emirkan OKURSOY
İmza Bu bölüme kişinin kendi el yazısı ile okudum anladım yazmalı ve imzalanmalıdır.	İmza Bu bölüme kişinin kendi el yazısı ile okudum anladım yazmalı ve imzalanmalıdır.

BURSA ULUDAĞ ÜNİVERSİTESİ
LİSANSÜSTÜ TEZ TANITIMI
ÖĞRENCİ VE DANIŞMAN FORMU

FR 3.4.6_27

ÖĞRENCİ
Adı SOYADI: Mustafa Emirkan Okursoy
ÜNVANI : Araş. Gör.
FEN BİLİMLERİ ENSTİTÜSÜ
ENDÜSTRİ MÜHENDİSLİĞİ ABD
E-POSTA : emirkanokursoy@uludag.edu.tr
PROGRAMI: YÜKSEK LİSANS
ORCID : 0009-0007-9166-2550
TÜBİTAK ID : TBTK-0126-5914
WOS RESEARCHER ID : KVB-9021-2024
Google Scholar ID : 4fdPr6oAAAAJ

DANIŞMAN
Adı SOYADI : Tülin İNKAYA
ÜNVANI : Prof. Dr.
FEN BİLİMLERİ ENSTİTÜSÜ
ENDÜSTRİ MÜHENDİSLİĞİ ABD
E-POSTA : tinkaya@uludag.edu.tr
YÖKSİS ARAŞTIRMACI ID : 156681
ORCID : 0000-0002-6260-0162
TÜBİTAK ID : TBTK-0005-6960
WOS RESEARCHER ID : AAH-2155-2021
SCOPUS AUTHOR ID : 24490728300
Google Scholar ID : Xxq-RHYAAAAJ

BURSA ULUDAĞ ÜNİVERSİTESİ YÜKSEK LİSANS EĞİTİMİ BOYUNCA BİLİMSEL ÇALIŞMALARI VE FAALİYETLERİ*

1. Okursoy, M. E., & İnkaya, T. (2025). Yüksek Mahkeme kararlarının sınıflandırılması için veri dengeleme ve açıklanabilir yapay zekâ tabanlı karar destek sistemi. Veri Bilimi, 8(1), 11–23. https://dergipark.org.tr/tr/pub/veri/issue/92570/1557010
2. Okursoy, M. E., & İnkaya, T. (2024, September 2–6). *Yapay zekâ ile hukuk alanında bir karar destek sistemi*. In M. Gök (Ed.), *Proceedings of the 6th International Conference on Data Science and Applications (ICONDATA'24)* (p. 6). https://icondata.org
3. Ocak 2023 – Mart 2025 tarihleri arasında TÜBİTAK 22AG001 numaralı "Elektrikli Taşıtlar için Batarya Teknolojileri Araştırma ve Geliştirme Platformu (BATEG)" projesinde bursiyer olarak görev alınmıştır.
4. Mart 2025 tarihinden itibaren TÜBİTAK 2210/C Öncelikli Alanlara Yönelik Yurt İçi Yüksek Lisans Burs Programında bursiyer olarak görev alınmıştır.

BİRLEŞMİŞ MİLLETLER SÜRDÜRÜLEBİLİR KALKINMA HEDEFLERİ

☐		
☐		
☐		
☐		
☐		
☐		
☐		
☐		
☒		Yapay zekâ, Doğal dil işleme, Makine öğrenmesi
☐		
☐		
☐		
☐		
☐		
☐		
☒		Karar analizi, adalete erişim, karar sınıflandırma
☐		

ÖZET

Yüksek Lisans/Doktora Tezi

YAPAY ZEKÂ TEKNİKLERİNİN HUKUK ALANINDA UYGULANMASI

Mustafa Emirkan OKURSOY

Bursa Uludağ Üniversitesi
Fen Bilimleri Enstitüsü
Endüstri Mühendisliği Anabilim Dalı

Danışman: Prof. Dr. Tülin İNKAYA

Yapay zekâ ve doğal dil işleme teknikleri, hukuk alanında karar destek, belge analizi ve metin sınıflandırma gibi görevlerde artan bir gereksinimi karşılamakta; bu sayede yüksek hacimli yargı belgelerinin incelenmesi için gereken zaman ve emek azaltılabilmektedir. Bu tezde, Türkiye Cumhuriyeti Yüksek Mahkemelerine (Yargıtay ve Danıştay) ait 6.425 karar incelenmiş ve bu kararlar "onama" ve "bozma" olarak sınıflandırılmıştır. Veri setinde gözlemlenen sınıf dengesizliği, Sentetik Azınlık Sınıfı Üst Örnekleme Tekniği (Synthetic Minority Over-Sampling Technique, SMOTE) Rastgele Alt Örnekleme, bu ikisinin hibrit şekli ve büyük dil modeli ChatGPT (GPT-3.5-turbo) aracılığıyla üretilen sentetik onama kararları olmak üzere dört farklı veri dengeleme yöntemiyle giderilmiştir. Öznitelik mühendisliği kapsamında, metinlerde "doktrin/öğreti", "karşı oy/muhalefet" ve "direnme" ifadelerinin varlığına dayalı öznitelikler tanımlanmış ve metin temsillerine dâhil edilmiştir. Metinler, Terim Frekansı–Ters Belge Frekansı (TF-IDF) ve Word2Vec vektörleri ile temsil edilmiştir. Sınıflandırma işlemi Karar Ağacı, Rassal Orman, K-En Yakın Komşu, Destek Vektör Makineleri, XGBoost, LightGBM, Yapay Sinir Ağı ve Uzun Kısa Süreli Bellek (LSTM) gibi çeşitli makine öğrenmesi ve derin öğrenme algoritmalarıyla gerçekleştirilmiştir. Model performansları F-ölçütü, Geometrik Ortalama ve Dengeli Doğruluk İndeksi gibi sınıf dengesine duyarlı metriklerle değerlendirilmiştir. Elde edilen sonuçlar Friedman ve Nemenyi testleri ile istatistiksel olarak analiz edilmiştir. Özellikle büyük dil modeli ile yapılan üretken veri dengeleme yaklaşımının, geleneksel yöntemlere kıyasla daha istikrarlı ve başarılı sonuçlar verdiği gözlemlenmiştir. Bu çalışmada üretken yapay zekâ destekli veri dengeleme stratejisi, bağlamsal öznitelik mühendisliği ve farklı metin temsillerinin çoklu sınıflandırma algoritmalarıyla sistematik karşılaştırması yoluyla, Türkçe hukuk metinlerinin sınıflandırılmasına yönelik yöntemsel bir çerçeve sunulmaktadır. Ayrıca büyük dil modellerinin azınlık sınıfı tahmininin performansını artırabileceği gösterilmiştir.

Anahtar Kelimeler: Doğal Dil İşleme, Yargı Kararlarının Sınıflandırılması, Veri Dengeleme, Makine Öğrenmesi, Büyük Dil Modelleri
2025, vi + 127 sayfa.

ABSTRACT

MSc Thesis

APPLICATION OF ARTIFICIAL INTELLIGENCE TECHNIQUES IN THE FIELD OF LAW

Mustafa Emirkan OKURSOY

Bursa Uludağ University
Graduate School of Natural and Applied Sciences
Department of Industrial Engineering

Supervisor: Prof. Dr. Tülin İNKAYA

Artificial intelligence and natural language processing techniques address an increasing need in the legal domain for tasks such as decision support, document analysis, and text classification; thus, the time and effort required to examine high-volume judicial documents can be reduced. In this thesis, 6,425 decisions from the high courts of the Republic of Turkey (Court of Cassation and Council of State) were examined and classified as either "approval" or "reversal". The class imbalance observed in the dataset was addressed using four different data balancing methods: Synthetic Minority Over-sampling Technique (SMOTE), Random Undersampling, hybrid form of these two and synthetic approval decisions generated via the large language model ChatGPT (GPT-3.5-turbo). As part of feature engineering, three binary features based on the presence of the expressions “doctrine/teaching,” “dissenting opinion/opposition,” and “resistance” in the texts were defined and included in the text representations. The texts were represented using Term Frequency–Inverse Document Frequency (TF-IDF) and Word2Vec vectors. The classification task was carried out using various machine learning and deep learning algorithms such as Decision Tree, Random Forest, K-Nearest Neighbors, Support Vector Machines, XGBoost, LightGBM, Artificial Neural Network, and Long Short-Term Memory (LSTM). Model performances were evaluated using class imbalance-sensitive metrics such as F-score, Geometric Mean, and Index of Balanced Accuracy (IBA). The obtained results were statistically analyzed using the Friedman and Nemenyi tests. It was observed that the generative data balancing approach using the large language model yielded more stable and successful results compared to traditional methods. This study presents a methodological framework for the classification of Turkish legal texts through a systematic comparison of generative AI-supported data balancing strategies, contextual feature engineering, and different text representations with multiple classification algorithms. Moreover, it has been shown that large language models can enhance the performance of minority class prediction.

Key words: Natural Language Processing, Judicial Decision Classification, Data Balancing, Machine Learning, Large Language Models
2025, vi + 127 pages.

TEŞEKKÜR

Tez çalışmam süresince bilgi, deneyim ve rehberlikleriyle bana büyük destek veren çok değerli danışman hocam Prof. Dr. Tülin İnkaya'ya ve kıymetli katkılarıyla sürece yön veren Dr. Öğr. Üyesi Emine Gökçe Karabel'e en içten teşekkürlerimi sunarım. Ayrıca, tezimin yürütülmesinde sağladıkları maddi katkılardan dolayı TÜBİTAK 1004 Mükemmeliyet Merkezi Destek Programı kapsamında desteklenen 22AG001 numaralı "Elektrikli Taşıtlar İçin Batarya Teknolojileri Araştırma ve Geliştirme Platformu" projesine ve 2210-C Öncelikli Alanlara Yönelik Yurt İçi Yüksek Lisans Burs Programı (2024/2) kapsamında sunulan burs desteği için Türkiye Bilimsel ve Teknolojik Araştırma Kurumu'na (TÜBİTAK) şükranlarımı sunarım.

Mustafa Emirkan OKURSOY

İÇİNDEKİLER

Sayfa

ÖZET viii
ABSTRACT ix
TEŞEKKÜR x
SİMGELER ve KISALTMALAR DİZİNİ xiii
ŞEKİLLER DİZİNİ xv
ÇİZELGELER DİZİNİ xvi
1. GİRİŞ 1
2. KURAMSAL TEMELLER ve KAYNAK ARAŞTIRMASI 3
2.1. Hukuk Alanında Doğal Dil İşleme Uygulamaları 4
2.2. Diğer Disiplinlerdeki Doğal Dil İşleme Uygulamaları 12
2.3. Veri Dengeleme 14
3. MATERYAL ve YÖNTEM 20
3.1. Materyal 20
3.2. Yöntem 23
3.2.1. Veri Ön İşleme ve Öznitelik Mühendisliği 23
3.2.2. Terim Frekansı – Ters Doküman Frekansı 23
3.2.3. Word2Vec 25
3.2.4. Veri Dengeleme Yöntemleri 28
3.2.5. Rastgele Alt Örnekleme 28
3.2.6. Üst Örnekleme 29
3.2.7. Sentetik Azınlık Sınıfı Üst Örnekleme Tekniği (SMOTE) 29
3.2.8. Büyük Dil Modelleri ile Veri Dengeleme 31
3.2.9. Makine Öğrenmesi ve Derin Öğrenme Modelleri 35
3.2.10. K-En Yakın Komşuluk Yöntemi 35
3.2.11. Karar Ağacı 36
3.2.12. Rassal Orman 38
3.2.13. Destek Vektör Makineleri 39
3.2.14. XGBoost 41
3.2.15. LightGBM 43
3.2.16. Yapay Sinir Ağları 44
3.2.17. Uzun-Kısa Süreli Bellek Ağı 46
3.2.18. Önerilen Yaklaşım 49
3.2.19. Performans Metrikleri 51
4. BULGULAR VE TARTIŞMA 54
4.1. Deneysel Koşullar 54
4.2. Hiperparametre Optimizasyonu 56
4.3. Model Performanslarının Karşılaştırılması 57
4.4. İstatistiksel Analiz 60
5. SONUÇ 74
KAYNAKLAR 76
EKLER 83
EK 1. KNN Sınıflandırma Modeli Sonuçları 84
EK 2. DT Sınıflandırma Modeli Sonuçları 86

EK 3. RF Sınıflandırma Modeli Sonuçları 88
EK 4. SVM Sınıflandırma Modeli Sonuçları 90
EK 5. XGBOOST Sınıflandırma Modeli Sonuçları 92
EK 6. LightGBM Sınıflandırma Modeli Sonuçları 94
EK 7. ANN Sınıflandırma Modeli Sonuçları 96
EK 8. LSTM Sınıflandırma Modeli Sonuçları 98
EK 9. Nemenyi Kritik Fark Diyagramları 100
Özgeçmiş 127

SİMGELER ve KISALTMALAR DİZİNİ

Kısaltmalar	Açıklama
ABA	Amerikan Barolar Birliği (American Bar Association)
ACM	Bilgisayar Makineleri Derneği (Association for Computing Machinery)
AG	Anonim Şirket (Aktiengesellschaft)
AI	Yapay Zekâ (Artificial Intelligence)
AİHM	Avrupa İnsan Hakları Mahkemesi (European Court of Human Rights)
ANN	Yapay Sinir Ağı (Artificial Neural Network)
API	Uygulama Programlama Arayüzü (Application Programming Interface)
AUC	ROC Eğrisi Altındaki Alan (Area Under the ROC Curve)
BC	İkili Sınıflandırma (Binary Classification)
BERT	Çift Yönlü Dönüştürücülerden Kodlayıcı Temsiller (Bidirectional Encoder Representations from Transformers)
BERTIMBAL	BERT Tabanlı Dengesiz Öğrenme Yaklaşımı (BERT-based Imbalanced Learning Approach)
BIRCH	Hiyerarşilerle Dengeli Yinelemeli Azaltma ve Kümeleme (Balanced Iterative Reducing and Clustering using Hierarchies)
BiLSTM	Çift Yönlü Uzun-Kısa Süreli Bellek (Bidirectional Long Short-Term Memory)
C4	C4.5 Karar Ağacı Algoritması (C4.5 Decision Tree Algorithm)
CAEN	Bağlam Duyarlı Açıklama Ağı (Context-Aware Explanation Network)
CART	Sınıflandırma ve Regresyon Ağacı (Classification and Regression Tree)
CBOW	Sürekli Kelime Torbası (Continuous Bag-of-Words)
CD	İstatistiksel Kritik Fark (Critical Difference)
CNN	Evrişimli Sinir Ağı (Convolutional Neural Network)
DF	Belge Frekansı (Document Frequency)
DistilBERT	Damıtılmış Çift Yönlü Kodlayıcı Temsiller (Distilled Bidirectional Encoder Representations from Transformers)
FCD	Odaksal Kortikal Displazi (Focal Cortical Dysplasia)
FN	Yanlış Negatif (False Negative)
FP	Yanlış Pozitif (False Positive)
G-Mean	Geometrik Ortalama (Geometric Mean)
GCN	Graf Evrişimli Sinir Ağı (Graph Convolutional Network)
GRU	Kapılı Tekrarlayan Birim (Gated Recurrent Unit)
IBA	Dengeli Doğruluk İndeksi (Index of Balanced Accuracy)
IDF	Ters Belge Frekansı (Inverse Document Frequency)
KNN	K-En Yakın Komşu (K-Nearest Neighbor)
LA-TAG	Metin Öznitelikli Grafiklerde Büyük Dil Modeli Tabanlı Zenginleştirme (Large Language Model-based Augmentation on Text-Attributed Graphs)
LDA	Gizil Dirichlet Ayrıştırması (Latent Dirichlet Allocation)
LLM	Büyük Dil Model(ler)i (Large Language Model(s))
LSI	Gizil Anlam İndeksleme (Latent Semantic Indexing)
LSTM	Uzun-Kısa Süreli Bellek (Long Short-Term Memory)
LightGBM	Hafif Gradyan Artırma Makinesi (Light Gradient Boosting Machine)
MEDIQA	Tıbbi Soru-Cevap Veri Kümesi (Medical Question Answering Dataset)

MEL	İspanyolca Hukuk Modeli (Modelo de Español Legal)
MIT	Massachusetts Teknoloji Enstitüsü (Massachusetts Institute of Technology)
ML	Makine Öğrenmesi (Machine Learning)
MTS	Mesial Temporal Skleroz (Mesial Temporal Sclerosis)
MT-DNN	Çok Görevli Derin Sinir Ağı (Multi-Task Deep Neural Network)
NAACL	Kuzey Amerika Hesaplamalı Dilbilim Derneği Şubesi (North American Chapter of the Association for Computational Linguistics)
NIST	Ulusal Standartlar ve Teknoloji Enstitüsü (National Institute of Standards and Technology)
NLP	Doğal Dil İşleme (Natural Language Processing)
NLTK	Doğal Dil Araç Takımı (Natural Language Toolkit)
NLU	Doğal Dil Anlama (Natural Language Understanding)
OECD	Ekonomik İşbirliği ve Kalkınma Örgütü (Organisation for Economic Co-operation and Development)
PCA	Temel Bileşenler Analizi (Principal Component Analysis)
PTSD	Travma Sonrası Stres Bozukluğu (Post-Traumatic Stress Disorder)
REGBR	Brezilya Düzenleyici Derlemi (RegBR – Brazilian Regulatory Corpus)
RNN	Tekrarlayan Sinir Ağı (Recurrent Neural Network)
ROC	Alıcı Çalışma Karakteristiği (Receiver Operating Characteristic)
RUS	Rastgele Alt Örnekleme (Random Undersampling)
SHAP	Shapley Katkısal Açıklamalar (Shapley Additive exPlanations)
SIGKDD	Bilgi Keşfi ve Veri Madenciliği Özel İlgi Grubu (Special Interest Group on Knowledge Discovery & Data Mining)
SMOTE	Sentetik Azınlık Sınıfı Üst Örnekleme Tekniği (Synthetic Minority Over-Sampling Technique)
SVM	Destek Vektör Makineleri (Support Vector Machines)
SetFit	Cümle Dönüştürücü İnce Ayarı (Sentence Transformer Fine-Tuning)
TF	Terim Frekansı (Term Frequency)
TF-IDF	Terim Frekansı – Ters Belge Frekansı (Term Frequency – Inverse Document Frequency)
TN	Gerçek Negatif (True Negative)
TNR	Özgüllük (True Negative Rate)
TP	Gerçek Pozitif (True Positive)
TPR	Duyarlılık (True Positive Rate)
UYAP	Ulusal Yargı Ağı Projesi (National Judiciary Informatics System)
XGBoost	Aşırı Gradyan Artırma (Extreme Gradient Boosting)
XLM	Çapraz Dilli Dil Modeli (Cross-lingual Language Model)
YZ	Yapay Zekâ (Artificial Intelligence)

ŞEKİLLER DİZİNİ

	Sayfa
Şekil 3.1. Veri setinden rastgele seçilen bir mahkeme kararının ekran görüntüsü	21
Şekil 3.2. Onama ve Bozma kararlarının sayısal dağılımı	22
Şekil 3.3. Karar metinlerinin kelime sayılarına göre dağılımı	22
Şekil 3.4. SMOTE algoritmasının sözde kodu	30
Şekil 3.5. LLM ile Sentetik Mahkeme Kararları Üretimi Akış Şeması	34
Şekil 3.6. ChatGPT ile üretilen bir sentetik onama kararının ekran görüntüsü	34
Şekil 3.7 Önerilen yaklaşımın akış şeması	50
Şekil 4.1. XGBoost Modelinin Birinci Aşamada F-Ölçütüne Göre Nemenyi Diyagramı	66
Şekil 4.2. LSTM Modelinin Birinci Aşamada Geometrik Ortalamaya Göre Nemenyi Diyagramı	67
Şekil 4.3. LSTM Modelinin Birinci Aşamada IBA'ya Göre Nemenyi Diyagramı	68

ÇİZELGELER DİZİNİ

	Sayfa
Çizelge 3.1. Karmaşıklık Matrisi	51
Çizelge 4.1. Sınıflandırma modellerinin en iyi hiperparametre değerleri	57
Çizelge 4.2. TF-IDF ve Word2Vec için F-ölçütü, G-Ort ve IBA özet istatistikleri (Ortalama, Min, Maks, Standart Sapma)	58
Çizelge 4.3. Sınıflandırma Modellerinin En Yüksek F-ölçütü Sonuçları	58
Çizelge 4.4. Sınıflandırma Modellerinin En Yüksek Geometrik Ortalama Sonuçları	59
Çizelge 4.5. Sınıflandırma Modellerinin En Yüksek IBA Sonuçları	59
Çizelge 4.6. Birinci aşamada F-ölçütü ve Geometrik Ortalama için Friedman Test İstatistikleri	63
Çizelge 4.7. Birinci Aşamada F-Ölçütüne Göre En Başarılı Beş Kombinasyon	64
Çizelge 4.8. Birinci Aşamada Geometrik Ortalama Ölçütüne Göre En Başarılı Beş Kombinasyon	65
Çizelge 4.9. Birinci Aşamada IBA Ölçütüne Göre En Başarılı Beş Kombinasyon	65
Çizelge 4.10. İkinci aşamada F-ölçütü, Geometrik Ortalama ve IBA için Friedman Test İstatistikleri	70
Çizelge 4.11. İkinci Aşamada F-Ölçütüne Göre Ortalama Sıralamalar	71
Çizelge 4.12. İkinci Aşamada Geometrik Ortalamaya Göre Ortalama Sıralamalar	71
Çizelge 4.13. İkinci Aşamada IBA Ölçütüne Göre Ortalama Sıralamalar	72

1. GİRİŞ

Yapay zekâ (YZ) ve doğal dil işleme (Natural Language Processing, NLP) teknikleri, dijitalleşme süreciyle birlikte hukuk alanında yaygınlık kazanmaktadır. Karar destek sistemleri, içerik sınıflandırma, belge özetleme ve mevzuat analizi gibi NLP uygulamaları, hukuki süreçlerin hızını artırmakta ve karar vericilere tutarlı, veriye dayalı analiz olanağı sunmaktadır. Hukuki belgelerin anlaşılması, yorumlanması ve sınıflandırılması gibi karmaşık görevlerde yapay zekâ tabanlı çözümler etkili bir alternatif sunmaktadır.

Türkiye özelinde Ulusal Yargı Ağı Projesi (UYAP), e-tebligat ve e-duruşma gibi uygulamalarla yargının dijital dönüşüm süreci son yıllarda önemli ilerlemeler kaydetmiş, yapay zekâ tabanlı çözümlerin kullanımı için de elverişli bir altyapı oluşmuştur. Ancak hukuki metinlerin kendine özgü karmaşık yapısı ve Türkçe'nin dilsel özellikleri, doğal dil işleme (NLP) uygulamalarının önünde ciddi engeller teşkil etmektedir. Bu durum, özellikle yüksek mahkeme kararlarının makine öğrenmesi yöntemleriyle sınıflandırılmasında ortaya çıkan zorlukları daha da belirgin hale getirmektedir. Hukuk alanındaki sınıflandırma çalışmalarında sıkça karşılaşılan sınıf dengesizliği problemi, modelin azınlıkta kalan karar türlerini doğru bir şekilde öğrenebilmesini ve genelleştirebilmesini güçleştirmektedir. Dolayısıyla, veri dağılımındaki bu dengesizliğin üstesinden gelebilmek için üretken yapay zekânın da dâhil olduğu yenilikçi ve etkili çözümlerin geliştirilmesi önem arz etmektedir.

Üretken yapay zekâ özellikle veri yetersizliği ve sınıf dengesizliği gibi kritik problemlere yönelik olarak sağladığı sentetik veri üretimi kapasitesiyle makine öğrenmesi uygulamalarında önem kazanmıştır. Bu teknolojinin etkin kullanımı, sınıflandırma ve tahmin performansını artırırken modelin genelleştirme kabiliyetini de iyileştirmektedir.

Yapılan tez çalışması, Türkçe yüksek mahkeme kararlarının onama ve bozma olarak sınıflandırılması problemine üretken yapay zekânın avantajlarını yenilikçi bir yaklaşımla entegre etmektedir. Bu bağlamda tez çalışması, GPT-3.5-turbo büyük dil modelinden yararlanılarak azınlık sınıfa ait sentetik örneklerin oluşturulması ve bu örneklerin geleneksel yeniden örnekleme yöntemleriyle karşılaştırmalı olarak değerlendirilmesi

bakımından özgündür. Yapılan çalışmada literatürde sıklıkla kullanılan SMOTE ve rastgele alt örnekleme yöntemlerine alternatif bir veri dengeleme stratejisi geliştirilmiştir. Buna ek olarak, hukuk metinlerinde sıklıkla yer alan ancak daha önce bağlamsal öznitelik olarak incelenmemiş "doktrin", "karşı oy" ve "direnme" gibi ifadelerin varlığının, karar sınıflandırmasına etkisi de ilk kez sistematik bir biçimde incelenmiş, böylelikle hukuk diline özgü özelliklerin sınıflandırma performansına etkisine dair literatüre özgün bir katkı sağlanmıştır. Son olarak, kullanılan metin temsil yöntemleri terim frekansı–ters belge frekansı (Term Frequency–Inverse Document Frequency, TF-IDF) ve kelimeleri vektörlere dönüştürme yöntemi (Word to Vector, Word2Vec) ve sınıflandırma algoritmaları (klasik makine öğrenmesi yöntemleri ve LSTM ağları) arasındaki etkileşim kapsamlı bir deneysel tasarımla değerlendirilmiş, istatistikel analizlerle literatürdeki yöntemsel karşılaştırmalara yeni bir bakış açısı kazandırılmıştır.

Bu tezin diğer bölümleri şu şekilde düzenlenmiştir. Bölüm 2'de, yapay zekâ ve doğal dil işlemenin hukuk ve diğer disiplinlerdeki güncel uygulamaları, sınıf dengesizliği problemi ve ilgili yöntemlerin kuramsal temelleri literatür taraması çerçevesinde ele alınmıştır. Bölüm 3'te, çalışmada yararlanılan veri seti, veri önişleme ve öznitelik mühendisliği adımları, TF-IDF ve Word2Vec tabanlı metin temsilleri, geleneksel ve büyük dil modeli tabanlı veri dengeleme stratejileri, klasik ve derin öğrenme temelli sınıflandırma algoritmaları ile performans ölçütleri sunulmaktadır. Bölüm 4'te, tasarlanan deneysel çalışmalar sonucunda elde edilen bulgular, hiperparametre optimizasyonu ve model karşılaştırmaları çerçevesinde, Friedman–Nemenyi gibi istatistiksel analizler ışığında tartışılmıştır. Son olarak, Bölüm 5'te araştırmanın temel sonuçları özetlenmiş, bulguların hukuki yapay zekâ uygulamalarına yönelik pratik yansımaları değerlendirilmiş ve gelecekte yapılabilecek çalışmalar için öneriler sunulmuştur.

2. KURAMSAL TEMELLER ve KAYNAK ARAŞTIRMASI

Doğal dil işleme teknikleri, bilgisayarların insan dilini tanıma, anlama ve üretme yeteneğini amaçlayan bir yapay zekâ dalı olarak 1950'lerdeki kural tabanlı sözdizim modellerinden doğmuş; 1980'lerde istatistiksel yaklaşımların, 1990-2000'lerde makine öğrenimli yöntemlerin ve 2010'larda derin sinir ağlarının sahneye çıkmasıyla kesintisiz bir evrim geçirmiştir. Günümüzde büyük ve önceden eğitilmiş modeller sayesinde, dil modelleme, duygu analizi, içerik sınıflandırma, özetleme ve makine çevirisi gibi görevlerde ölçeğin getirdiği doğruluk artışları elde edilirken, uygulama alanları da sanal asistanlardan sağlık bilişimine, finansal kestirimlerden hukuki belge analizine kadar genişlemiştir (Pajila vd., 2023).

Son yıllarda doğal dil işleme teknikleri, yalnızca dil verisinin analizine dayalı uygulamalarla sınırlı kalmayıp, yüksek güvenlik ve uzmanlık gerektiren sektörlerde dönüştürücü bir rol üstlenmeye başlamıştır. Chen vd. (2024) tarafından gerçekleştirilen kapsamlı bir derleme çalışmasında, büyük dil modelleri (Large Language Models, LLM) ve NLP yaklaşımlarının sağlık, finans ve hukuk gibi yüksek riskli ve birbirinden farklı alanlarda artan etkisi ayrıntılı biçimde ele alınmıştır. Söz konusu alanlarda, tanı destek, risk analizi ve mevzuat yorumlama gibi kritik görevlerde bu teknolojilerin dönüştürücü etkiler yarattığı belirtilmiş; ayrıca veri gizliliği ve açıklanabilirlik gibi yeni araştırma konularına dikkat çekilmiştir. Özellikle disiplinler arası anlayışla geliştirilen sistemlerin, yöntemsel yeniliklerin yanı sıra çok paydaşlı iş birliklerine de ihtiyaç duyduğu vurgulanmıştır.

Bu bölümde öncelikle hukuk alanındaki doğal dil işleme uygulamalarına odaklanılmıştır. Türkçe dilinde ve Türk hukukuna yönelik araştırmaların sayıca sınırlı olduğu gözlemlendiğinden, literatür taraması Türk hukukuna yönelik yapılan çalışmalardan başlanarak verilmiştir. Ardından, uluslararası bağlamda hukuk metinlerine üzerinde doğal dil işleme yöntemlerini uygulayan incelenmiş ve yöntemsel yenilikler karşılaştırmalı olarak değerlendirilmiştir. Çalışmada kullanılan veri kümesinin dengesiz bir dağılıma sahip olması nedeniyle, izleyen kısımda sınıf dengesizliğini gidermeye yönelik stratejileri ele alan araştırmalara yer verilmiştir. Son olarak doğal dil işlemenin hukuk dışındaki

diğer metin işleme alanlarındaki güncel uygulamalarını irdeleyen çalışmalara sunulmuştur.

2.1. Hukuk Alanında Doğal Dil İşleme Uygulamaları

Yapay zekâ ve özellikle doğal dil işleme yöntemleri, son on yılda hukukun neredeyse tüm alt disiplinlerinde araştırma ve uygulama gündeminin merkezine yerleşmiştir (Ariai ve Demartini, 2025). Hukuk metinlerinin uzunluğu, kendine özgü terimleri yoğun içeren yapısı, özetleme, madde eşleştirme, hüküm tahmini ve metin arama gibi görevlerde makine öğrenmesine dayalı yaklaşımların kullanılmasını önemli hale getirmiş, böylece hukuki doğal dil işleme (Legal NLP) kavramını ortaya çıkarmıştır. Güncel derleme çalışmaları, hukuki doğal dil işlemenin kapsamını yargı kararlarının sınıflandırılması, hukuksal varlıkların tanınması, argüman madenciliği, soru-cevap ve çok dilli mevzuat eşleştirmesi şeklinde ifade etmekte; metin uzunluğu, kavramsal yoğunluk ve açık kaynaklardaki veri eksikliğini başlıca zorluklar olarak belirtmektedir (Ariai ve Demartini, 2025).

2020'lerin ortalarında büyük dil modelleri, doğal dil işleme ekosisteminin ana itici gücü konumuna gelmiş; hukuk bazlı ön eğitim ve yönergeler ile ince ayar gibi yöntemler sayesinde hüküm tahminlerinin doğruluğunun anlamlı düzeyde artış gösterdiği bildirilmiştir. Yapılan son çalışmalarda büyük dil modellerinin halüsinasyon ve hata zincirlemesi riskine işaret edilerek, nedensel akıl yürütme ve daha güvenilir bilgi işleme yöntemlerine geçilmesinin gerekliliği vurgulanmıştır (Lai vd. 2024).

Yapay zekâ tabanlı araçlar hukuk mesleğinde özellikle belge inceleme, sözleşme analizi ve içtihat taraması gibi alanlarda giderek daha fazla kullanılmaktadır. Reuters'ın haberine göre, bu araçların entegrasyonu hızla artarken, Amerikan Barolar Birliğinin (ABA) 2023 Teknoloji Raporu, Amerika Birleşik Devletleri'nde avukatların yaklaşık %10'unun halihazırda yapay zekâ tabanlı araçlar kullandığını, %20'sinin ise kullanmayı düşündüğünü göstermektedir. Ne var ki bu alandaki mevzuat eksikliği, kişisel verilerin korunması kaygısı ve kurumsal sorumluluğun paylaşılmasına dair riskler yapay zekâ uygulamalarının hukuk sektörü tarafından içselleştirilmesine engel olmaktadır (American Bar Association, 2023; Thomson Reuters, 2024).

Literatürde bu alanda yapılmış Türk hukuk sistemine ilişkin Türkçe çalışmalar sınırlıdır. Aşağıda sırasıyla sunulan çalışmalarda, ele aldıkları hukuki alt problem, kullandıkları yöntemsel yaklaşım ve rapor ettikleri temel bulgular bakımından kısaca özetlenerek, bu alandaki çalışmaların geldiği aşama ve yapılan tez çalışmasında izlenen yöntemlere ışık tutulmuştur.

Arslan vd. (2024) tarafından geliştirilen LAWNAV adlı çevrimiçi yönlendirme sistemi, Türkçe serbest metinleri analiz ederek kullanıcıları uygun hukuk bürolarına yönlendirme işlevi sunmaktadır. Çalışmanın literatüre katkısı, Türkçe doğal dil işleme tekniklerinin hukuk danışmanlığı alanındaki uygulamasını pratik bir sistem üzerinden göstermesi ve yerel kullanım senaryolarına odaklanarak dijital hukuki erişimi artırmasıdır. Ayrıca, yapay zekâ tabanlı hukuk teknolojilerinde basit ancak etkili yöntemlerin işlevselliğini göstermesi yönüyle literatürdeki bir boşluğu doldurmaktadır.

Turan vd. (2023) tarafından yürütülen çalışmada, Türkiye Anayasa Mahkemesi'nin bireysel başvuru kararları "İhlal" ve "İhlal Yok" şeklinde sınıflandırılmıştır. Karar metinlerinin "İnceleme ve Gerekçe" bölümleri bağımsız, "Hüküm" bölümü ise bağımlı değişken olarak tanımlanmıştır. Metinler ön işlemeye tabi tutularak Terim frekansı–ters döküman frekansı (Term Frequency–Inverse Document Frequency, TF-IDF) yöntemi ile sayısallaştırılmış; ardından k-en yakın komşu (k-Nearest Neighbors, KNN), destek vektör makineleri (Support Vector Machines, SVM), karar ağacı, rassal orman (Random Forest) ve aşırı gradyan artırma (Extreme Gradient Boosting, XGBoost) algoritmalarıyla sınıflandırılmıştır. Çalışmada Shapley katkısal açıklamalar (Shapley Additive Explanations, SHAP) yöntemiyle yüksek doğruluğa sahip ancak yorumlanabilirliği sınırlı olan kara kutu modelleri şeffaflaştırılmıştır. Bu sayede model çıktıları hukukçular tarafından daha kolay yorumlanabilir hâle gelmiştir (Turan vd., 2023).

Görentaş vd. (2023) tarafından yapılan bir çalışmada Uyuşmazlık Mahkemesi'nin olumsuz görev uyuşmazlığına ilişkin kararlarını makine öğrenmesi yöntemleriyle adli ve idari olarak sınıflandırmaya tabi tutulmuştur. Karar metinleri, sırasıyla sözcükleme (tokenizasyon), durak sözcüklerin temizlenmesi ve kök bulma işlemlerinden geçirilmiş; ardından TF-IDF ile vektörleştirilmiştir. TF-IDF vektör temsilleri üzerinde lojistik regresyon, destek vektör makineleri, karar ağaçları ve rassal orman algoritmaları

uygulanmış ve en yüksek başarı destek vektör makinesi modeli ile elde edilmiştir. Geçmişte yapılmış çalışmalara kıyasla daha yüksek sınıflandırma doğruluğuna ulaşılarak yapay zekâ tekniklerinin yargı uygulamalarına etkin şekilde entegre edilebileceği gösterilmiş ve ileride sanal yargıç gibi uygulamalara yönelik yapılacak çalışmalara teorik ve teknik bir temel hazırlanmaya çalışılmıştır (Görentaş vd., 2023).

Görentaş ve Uçkan (2023) tarafından gerçekleştirilen bir çalışmada Türkiye'deki yargı verisi üzerinde kümeleme uygulaması yapılmıştır. Çalışma kapsamında Uyuşmazlık Mahkemesi'nin olumsuz görev uyuşmazlığına ilişkin kararları üzerinde TF-IDF ile sayısallaştırma ve temel bileşenler analizi (Principal Component Analysis, PCA) ile iki boyuta indirgeme yapılarak görselleştirme uygulanmıştır. Kümeleme işleminde Hiyerarşilerle Dengeli Yinelemeli Azaltma ve Kümeleme (Balanced Iterative Reducing and Clustering using Hierarchies, BIRCH) yöntemi, büyük ve dağınık karar kümelerini dengeli ve ayrık gruplara ayırmada en yüksek başarıyı sağlamıştır. Böylece benzer içerikli davaların otomatik gruplanmasının hukukçulara emsal karar tarama, karar eğilimi analizi ve dava stratejisi geliştirme süreçlerinde önemli zaman tasarrufu sağlayabileceği sonucuna ulaşılmıştır (Görentaş ve Uçkan, 2023).

Karabel ve Aydemir (2023), Türk medeni yargı sisteminin dijital dönüşümünü üç ana başlıkta inceleyerek literatüre pratik ve yöntemsel bir çerçeve kazandırmıştır. İlk olarak, Ulusal Yargı Ağı Bilişim Sistemi (UYAP), elektronik tebligat (e-tebligat) ve elektronik duruşma (e-duruşma) uygulamalarının yargılama sürelerini kısaltarak adalete erişimi kolaylaştırdığı gösterilmiştir. İkinci olarak, Kanada'daki Sivil Uyuşmazlık Çözüm Mahkemesi (Civil Resolution Tribunal) ile İngiltere'deki Çevrimiçi Medeni Para Alacağı Talebi Sistemi (Online Civil Money Claim) ve Çevrimiçi Tazminat Talepleri (Damages Claims Online) pilot projeleri karşılaştırmalı olarak analiz edilmiş; küçük meblağlı uyuşmazlıklar için çevrimiçi ve zorunlu usulün Türk hukukuna uyarlanabilirliği tartışılmıştır. Üçüncü olarak, yapay zekâ (Artificial Intelligence, AI) tabanlı sistemlerin hâkime veya taraflara destek amaçlı kullanımının olası faydaları, algoritmik önyargı ve şeffaflık eksikliği gibi risklerle birlikte değerlendirilmiş ve robot hâkim modelinin mevcut koşullarda uygun olmadığı sonucuna ulaşılmıştır. Çalışma ile UYAP sisteminin AI destekli ön inceleme için yeniden yapılandırılması, e-duruşma altyapısının yaygınlaştırılması, küçük değerdeki davalar için çevrimiçi usulün yasalaştırılması ve

geliştirilecek AI sistemlerinin bağımsız denetime tabi tutulması yönünde somut reform önerileri sunularak, yargı sisteminde yapay zekâ kullanımı ve dijitalleşme ile ilgili literatüre katkıda bulunulmuştur (Karabel ve Aydemir, 2023).

Akça vd. (2022) tarafından yapılan bir çalışmada Yargıtay kararlarından oluşan dokuz sınıflı veri setiyle Türkçe hukuk metni sınıflandırmasında kaynak eksikliğini gidermiş ve çeşitli yöntemleri kapsamlı biçimde karşılaştırılmıştır. Değerlendirilen modeller arasında saf Bayes sınıflandırıcısı (Naive Bayes), lojistik regresyon, radyal temel fonksiyonlu destek vektör makinesi (Radial Basis Function Support Vector Machine, RBF-SVM), FastText tabanlı çift yönlü uzun-kısa süreli bellek ağı (FastText-based Bidirectional Long Short-Term Memory, FastText-BiLSTM), ayrıca hem genel hem de hukuk özelinde ince ayarlanmış çift yönlü dönüştürücü temsilleri (Bidirectional Encoder Representations from Transformers, BERT) ve seyreltilmiş BERT (DistilBERT) yer almıştır. Testlerde hukuk alanına özgü ince ayarlı BERT modeli, görece az yinelemeyle dahi en yüksek doğruluğu sağlayarak alan-özelinde dönüştürücü uyarlamasının üstünlüğünü göstermiştir. Böylece çalışma, Türkçe hukuk metni sınıflandırması literatürüne hem yeni bir veri kaynağı sunmakta hem de dönüştürücü-tabanlı yaklaşımların etkinliğini deneysel olarak kanıtlayan yöntemsel bir çerçeve kazandırmaktadır (Akça vd., 2022).

Mumcuoğlu vd. (2021) tarafından, Türkiye'deki yüksek mahkeme karar sonuçları doğal dil işleme (NLP) ve makine öğrenmesi teknikleriyle tahmin edilerek literatüre üç yönlü katkı sağlanmıştır. İlk olarak, Anayasa Mahkemesi ile bölge adliye ve idare mahkemelerinin hukuk, ceza, idari ve vergi dairelerinden derlenen çok katmanlı ve kamuya açık bir yargı metni koleksiyonu oluşturularak araştırmacıların kullanımına sunulmuştur. İkinci olarak, klasik sınıflandırma algoritmaları ile LSTM, GRU ve BiLSTM gibi derin öğrenme modelleri, dikkat (attention) mekanizması eklenmiş biçimde sistematik olarak karşılaştırılmıştır. Bu karşılaştırmalar sonucunda, uzun metinlerde derin öğrenme modellerinin tutarlı bir biçimde üstün performans gösterdiği ortaya konmuştur. Son olarak, mahkeme türleri arasındaki öngörülebilirlik düzeyindeki farklılıklar metin uzunluğu ve dava konularının yapısal özellikleriyle ilişkilendirilmiştir. Böylece, makine öğrenmesine dayalı öngörü modellerinin etik, şeffaflık ve uygulamaya dönük etkileri tartışmaya açılmıştır. (Mumcuoğlu vd., 2021).

Bu bölümün devamında farklı ülkelerdeki hukuk sistemlerini ele alan çalışmalara değinilmiştir. Hukuk alanında doğal dil işleme hakkındaki farklı ülkelerde yapılmış çalışmalar, çok dilli mevzuat analizinden yüksek mahkeme kararlarının tahminine, sözleşme analitiğinden yasa dili çözümlemesine kadar geniş bir yelpazede yapay zekâ ve derin öğrenme temelli yöntemlerin başarıyla uygulanabildiğini göstermektedir.

Siino vd. (2025) tarafından 2017-2023 döneminde yayımlanan 61 makaleyi incelenerek büyük dil modellerinin (Large Language Models, LLM) hukuki doğal dil işleme alanındaki kullanımı sistematik şekilde haritalandırılmıştır. Yayın hacmindeki belirgin artışı ve belge özetleme, sözleşme analizi ile yargı sonucu tahmini gibi görev çeşitliliğine değinilmiş; alana özgü önceden eğitilmiş modellerin geleneksel yöntemlere göre tutarlı üstünlük sağladığı göstererek model uyarlamasının önemini vurgulanmıştır. Ayrıca çok dilli veri kümeleri ile etik ve şeffaflık boşluklarına dikkat çekilerek geleceğe yönelik somut araştırma konuları önerilmiş ve böylece hukuki NLP çalışmaları tematik, metodolojik ve coğrafi boyutlarda sentezlenerek parçalı bilgi birikimini bütünleştiren, araştırmacılar için referans olabilecek bir çerçeve sunulmuştur (Siino vd., 2025).

Betancur Sánchez vd. (2025), İspanyolca hukuk metinlerindeki dilsel özgüllüğün genel amaçlı modellerle yeterince yakalanamadığı boşluğu doldurmak için çapraz dilli RoBERTa'nın büyük sürümü (Cross-lingual Language Model RoBERTa-large, XLM-RoBERTa-large) üzerinde alana özgü ön eğitim gerçekleştirerek İspanyolca Hukuk Modeli (Modelo de Español Legal, MEL) adlı yeni bir dil modeli geliştirmiştir. Resmî mevzuat, parlamento tutanakları ve yargı kararlarından derlenen veriyle eğitilen MEL modeliyle az miktarda kaynak bulunan bir dilde alana özgü ön eğitimin etkinliği gösterilerek İspanyolca yerel model eksikliği giderilmeye çalışılmış, farklı dillerde hukuk alanına özgü model uyarlamasının önemi de ifade edilmiştir (Betancur Sánchez vd., 2025).

Lam vd. (2025), Kanada iş hukukunda öznel bir yargısal çıktı olan makul ihbar süresinin öngörülmesi için derin öğrenme temelli bir çerçeve geliştirmiştir. Çalışmada, elle tanımlanan özelliklere dayalı modeller ile dönüştürücü (transformer) tabanlı ve Kanada istihdam hukukuna özel olarak uyarlanmış dil modellerinin performansları karşılaştırılmış; Bardal faktörleriyle tutarlı ısı haritaları üreterek model şeffaflığı

sağlanmıştır. Ön eğitimden geçmiş hukuki dil modellerinin ilk kez Kanada iş hukuku verisine uygulanmasıyla, önyargısız karar destek sistemlerinin manuel etiketleme gerektirmeden özet metinler üzerinden dahi yüksek isabetle makul ihbar süresi tahmini yapabileceği gösterilmiştir. Bu yönleriyle çalışma, hem alana uyarlanmış dönüştürücü (transformer) modellerinin üstünlüğünü ortaya koymakta, hem de Kanada iş hukuku için referans olabilecek şeffaf ve yeniden üretilebilir bir karar destek çerçevesi sunarak literatüre katkı sağlamaktadır (Lam vd., 2025).

Wajeed vd. (2025) tarafından yapılan çalışmada, metin sınıflandırmasında açıklanabilir makine öğrenimi yöntemleri kapsamlı biçimde karşılaştırılmış; performans-yorumlanabilirlik ikilemine yenilikçi bir çözüm olarak bağlama duyarlı açıklama ağları (Context-Aware Explanation Networks, CAEN) geliştirilmiştir. Özellikle hukuki belgelerin sınıflandırılması gibi yüksek açıklanabilirlik gerektiren alanlara yönelik uygulamalarda öne çıkan CAEN'in bağlam-duyarlı dikkat katmanı ile kullanıcı profiline göre özelleştirilen açıklama üreticisi, veri kümelerinde hem BERT-tabanlı modellerden hem de açıklanabilirlik yöntemlerinden daha yüksek doğruluk, sadakat ve kullanıcı memnuniyeti sağlamış; böylece bağlam bilgisini nedensel doğrulukla birleştirip kullanıcı odaklı şeffaflık ilkesini hayata geçiren ilk derin öğrenme mimarisi olarak önemli bir boşluğu doldurmuştur (Wajeed vd., 2025)

Jesus vd. (2023), Brezilya'da kâr amacı gütmeyen (üçüncü sektör) kuruluşlarını düzenleyen normatif metinleri saptamak için BERT tabanlı hukuki doğal dil işleme (Bidirectional Encoder Representations from Transformers-based Legal Natural Language Processing, BERTIMBAL) yaklaşımını geliştirmiş ve ulusal mevzuat veri tabanı üzerinde uygulamıştır. Model yaklaşık 6 000 dengesiz etiketli cümleyle ince ayarlandıktan sonra, anahtar sözcük aramalarının tespit edemediği çok sayıda yasa ve kararnamedeki üçüncü sektör hükümlerini yüksek doğrulukla belirlemiştir. Dönüştürücü (transformer) tabanlı modellerin sınırlı eğitim verisiyle bile geniş düzenleyici metin koleksiyonlarında ölçeklenebilir ve güvenilir biçimde kullanılabileceğini gösterilerek, mevzuat madenciliğinde yapay zekânın sistematik haritalama kapasitesi ortaya konulmuştur (Jesus vd., 2023).

Almuzaini vd. (2023), Suudi Arabistan aile hukuku kararlarında gerekçe belirlemek ve ilgili içtihat kaynaklarını sunmak amacıyla geliştirdikleri TaSbeeb adlı entegre karar destek sistemiyle, hukuki doğal dil işleme alanına yöntemsel katkılar sunmuştur. Çalışma, geniş ölçekli ve zenginleştirilmiş Arapça hukuk metinleri üzerinde derin öğrenme modellerinin uygulanabilirliğini göstermekle birlikte, karar gerekçelerinin otomatik saptanması için iki kademeli etiketleme ve yığma tabanlı sınıflandırma stratejilerini kullanarak yenilikçi bir yaklaşım ortaya koymuştur. Ayrıca Kur'ân, hadis ve fıkıh içtihatlarını içeren bilgi tabanını yargı kararlarıyla eşleştiren mimarisi, gerekçe tespiti ile ilgili içtihat kaynağının eş zamanlı sunulmasını sağlamış; böylece Arapça hukuk metinleri için etiketli büyük veri ve uzmanlaşmış dil modelleriyle desteklenen derin öğrenme uygulamalarına referans olabilecek bir çerçeve oluşturmuştur (Almuzaini vd., 2023).

Bhattacharya vd. (2023) tarafından yargı kararlarındaki cümleleri retorik rollerine ayırmak amacıyla geliştirilen DeepRhole adlı derin öğrenme tabanlı çerçeve hukuk metni yapısal çözümlemesine katkılar sunmuştur. Hindistan ve Birleşik Krallık Yüksek Mahkemelerine ait kararları içeren, olay, içtihat, gerekçe gibi retorik etiketlere sahip kapsamlı bir eğitim verisi oluşturarak kaynak eksikliğini azaltılmıştır. Derin öğrenme mimarileriyle bu bileşenleri yüksek doğrulukla otomatik olarak ayıran bir model geliştirilerek metin madenciliğine dayalı hukuk analitiği için örnek bir çalışma yapılmıştır. Model çıktısı benzer içerikli kararların karşılaştırmalı incelemesine uygun biçimde sunularak yargı süreçlerinin analizi ve karar şeffaflığının artırılması için veri bazlı bir katkı sağlanmıştır (Bhattacharya vd., 2023).

Chen vd. (2022) tarafından gerçekleştirilen bir çalışmada, yüksek terim çeşitliliğine sahip ABD Yüksek Mahkemesi kararları kullanılarak alan kavramı temelli öznitelik çıkarımıyla zenginleştirilmiş rassal orman modeli ile dönüştürücü tabanlı derin öğrenme ağları ve TF-IDF temelli yaklaşımlar karşılaştırılmıştır. Analizlerde, rassal orman modelinin doğruluk ve hesaplama maliyeti bakımından hem derin öğrenme mimarilerini hem de TF-IDF tabanlı yöntemleri geride bıraktığı sonucuna ulaşılmıştır. Böylece, hukuk metinlerinin sınıflandırılmasında konuya özgü kavramsal yapıların geleneksel makine öğrenmesine entegre edilmesinin, doğruluk kadar verimlilik açısından da etkili bir alternatif sunabileceği ortaya konulmuştur (Chen vd., 2022).

Michel vd. (2022) tarafından Avusturya Gelir Vergisi Kanunu üzerinde yürütülen çalışmada, yasa metinlerindeki cümleler yükümlülük, izin, yasak ve tanım kategorilerine ayrıştırılmış; böylece normatif ifadelerin makine öğrenmesiyle tespitine yönelik yöntemsel bir temel ortaya konmuştur. Sınıflandırma aşamasında FastText sözcük yerleştirmeleri (FastText embeddings, FastText) ile evrişimli sinir ağı (Convolutional Neural Network, CNN) mimarisi kullanılmış; bu sayede dolaylı biçimde ifade edilen normatif yapıların tutarlı ve doğru şekilde ayrıştırılabildiği gösterilmiştir. Gerçekleştirilen çalışmalarla manuel kural etiketleme yükünün azaltılabileceği ve mevzuatın sistematik şekilde analiz edilebileceği gösterilerek, kural çıkarım sistemlerinin otomasyonu ile dijital mevzuat incelemeleri için literatüre katkı sağlanmıştır (Michel vd., 2022).

Ash ve Marian (2020) tarafından yürütülen bir araştırmada, uluslararası vergi hukukuna ilişkin binlerce ikili vergi anlaşması doğal dil işleme (Natural Language Processing, NLP) yöntemleriyle incelenmiş; metinlerde zaman içinde gelişen ortak dil örüntüleri ile bu sürecin ardındaki kurumsal etkiler belirlenmiştir. Yapılan analizlerde Ekonomik İşbirliği ve Kalkınma Teşkilatı (Organisation for Economic Co-operation and Development, OECD) kılavuzlarının kısa ve orta vadede dilsel yakınsamayı yönlendirdiği, Birleşmiş Milletler'in (United Nations, BM) ise uzun vadede kaynak ülke merkezli yaklaşımın taşıyıcısı olduğu ortaya konmuştur. Bulgular sayesinde uluslararası vergi anlaşmalarında dilsel uyumun ölçülebilir biçimde haritalandığı, kurumsal aktörlerin metin standardizasyonundaki göreli etkilerinin nicel olarak ayrıştırıldığı ve artan dilsel benzerliğe karşın yorum farklılıklarının uygulamada ayrışmalara yol açabileceği gösterilerek hukuk dilinin standardizasyon sürecine yönelik yeni bir kavramsal çerçeve sunulmuştur (Ash ve Marian, 2020).

Aletras vd. (2016) tarafından yürütülen çalışmada, Avrupa İnsan Hakları Mahkemesi (European Court of Human Rights, ECHR) kararlarının yalnızca dava metinlerine dayanarak öngörülebilirliği incelenmiştir. ECHR'nin 3., 6. ve 8. maddelerine ilişkin dengeli bir karar kümesi oluşturulmuş ve metinler yapısal bölümlerine ayrıştırılarak değerlendirilmiştir. Sınıflandırma sürecinde makine öğrenmesi ve doğal dil işleme yöntemleri kullanılmış; olayların anlatıldığı bölümlerin karar tahmininde en yüksek belirleyiciliğe sahip olduğu gösterilmiştir. Böylece ECHR kararlarının öngörülmesinde ML ve NLP tekniklerinin uygulanabilir olduğu, dava metinlerinin yapısal olarak

bölümlendirilmesinin tahmin doğruluğunu artırdığı ve kararların büyük ölçüde dosyadaki olgusal içeriğe dayandığı ortaya konmuştur. Bulguların karar süreçlerinde somut olgulara verilen önemi vurgulayan hukuki realizm yaklaşımıyla örtüştüğü de belirtilmiştir (Aletras vd., 2016).

Bu tez çalışması kapsamında yürütülen araştırmada, Okursoy vd. (2025) tarafından Yargıtay ve Danıştay'a ait dengesiz bir Türkçe yüksek mahkeme kararları kümesi üzerinde veri dengeleme ve açıklanabilir yapay zekâ temelli bir karar destek sistemi geliştirilmiştir. Kapsamlı ön-işleme adımlarının ardından TF-IDF temsilleri kullanılmış, SMOTE ve rastgele alt örnekleme stratejileri tekil ve hibrit biçimlerde uygulanmış, klasik ve derin öğrenme tabanlı çoklu modeller denenmiş ve model çıktıları açıklanabilir kılınmaya çalışılmıştır. Böylece, Türk hukuku bağlamında azınlık sınıfı karar tahminine yönelik geniş ölçekli bir veri seti ile sistematik bir dengeleme-açıklanabilirlik çerçevesi literatüre kazandırılmıştır.

2.2. Diğer Disiplinlerdeki Doğal Dil İşleme Uygulamaları

Bu bölümde birbirinden farklı alanlarda yürütülmüş güncel doğal dil işleme çalışmalarına yer verilmiştir. Hukuk dışındaki farklı alanlarda yapılan çalışmalar incelenerek, model tasarımı, özellik mühendisliği ve değerlendirme metrikleri açısından ortak ve farklı noktaların belirlenmesi amaçlanmıştır. Bu doğrultuda, tez kapsamında doğal dil işleme tekniklerinin uygulandığı yerli ve yabancı akademik çalışmalar değerlendirilmiştir.

Ong vd. (2025) tarafından gerçekleştirilen bir çalışmada, doğal dil işleme (Natural Language Processing, NLP) ve açıklanabilir yapay zekâ (Explainable Artificial Intelligence, XAI) tekniklerinin kurumsal sürdürülebilirlik analizinde karşılaşılan öznellik sorunlarını gidermedeki rolü ele alınmıştır. Çalışmada, sürdürülebilirlik raporlarında yer alan verilerin eksiklik, belirsizlik, güvensizlik ve aşırı karmaşıklık gibi sorunlar barındırdığı, buna karşın insan analistlerin sınırlı kaynak ve öznellik nedeniyle değerlendirme sürecinde yetersiz kalabildiği vurgulanmıştır. Bu bağlamda, açıklanabilir doğal dil işleme (XNLP) yöntemleri aracılığıyla konu çıkarımı, etki sınıflandırması, otomatik derecelendirme ve etkileşimli diyalog sistemleri gibi görevlerin daha güvenilir, şeffaf ve tekrarlanabilir bir biçimde gerçekleştirilebileceği gösterilmiştir.

Açıklanabilirlik, yorumlanabilirlik ve sadakat (faithfulness) gibi XAI bileşenlerinin entegrasyonunun ise model güvenilirliğini artırdığı ifade edilmiştir (Ong vd., 2025).

Gorovaia ve Makrominas (2025) tarafından gerçekleştirilen bu çalışmada, çevresel ihlallerde bulunan şirketlerin kurumsal sosyal sorumluluk (KSS) raporları, çevreye duyarlı bir imaj oluşturmak amacıyla manipülatif biçimde kullanılmaktadır. Çalışma kapsamında, doğal dil işleme (Natural Language Processing, NLP) teknikleriyle Amerika Birleşik Devletleri'ndeki halka açık şirketlerin 2008–2022 yılları arasındaki KSS raporları analiz edilmiştir. Analizlerde, çevresel suçlara karışan firmaların daha uzun, daha olumlu içerikli ve daha sık yayımlanan raporlar hazırladığı; bu raporların çevresel vurgusunun yoğun ancak okunabilirliğinin düşük olduğu ortaya konmuştur. Bu bulgular, yeşil aklama (greenwashing) davranışının dikkat saptırma, seçici açıklama ve sembolik ayrışma gibi stratejilerle gerçekleştirildiğine işaret etmektedir. Çalışmada ayrıca çevresel sorumluluk düzeyini nesnel şekilde ölçmek için para cezası alıp almama durumu ikili bir gösterge olarak kullanılmıştır. (Gorovaia ve Makrominas 2025).

García-Méndez vd. (2022) tarafından gerçekleştirilen bu çalışma, finans haberlerinde yer alan ifadelerin geçmişe mi yoksa geleceğe mi yönelik olduğunu söylem düzeyinde ayırt edebilen bir sınıflandırma sistemi geliştirerek literatüre özgün bir katkı sunmaktadır. Bağımlılık çözümlemesi, yan tümce bölütleme ve varlık tanıma gibi doğal dil işleme tekniklerinin makine öğrenmesiyle bütünleştirildiği modelde, hem sözcüksel n-gram öznitelikleri hem de sözdizimsel ve pragmatik bağlamdan elde edilen zamansal göstergeler kullanılmıştır. Özellikle fiil çekimlerinin metin içindeki konumsal dağılımı ve varlık–eylem ilişkileri, zaman yöneliminin belirlenmesinde belirleyici olmuştur. Çalışma, finansal haber analizinde geleceğe yönelik içeriklerin güvenilir biçimde tespit edilebileceğini ortaya koyarak, doğal dil işleme tabanlı karar destek sistemleri için işlevsel bir temel sunmaktadır.

Ziems ve Wu (2021), yazılım kaynak kodlarındaki güvenlik açıklarını doğal dil işleme kapsamında ele alarak problemi, çok sınıflı metin sınıflandırma olarak yeniden formüle etmişlerdir. NIST tarafından sağlanan geniş kapsamlı C/C++ veri kümesini kullanarak (123 farklı güvenlik açığı sınıfı ve hatasız durumları içeren toplam 124 sınıf), LSTM, Bi-LSTM ve BERT tabanlı modelleri karşılaştırmış; özellikle dönüştürücü (transformer)

tabanlı BERT modelinin bağlamsal temsillerini çift yönlü bellek yapılarıyla birleştirmenin sınıflandırma performansını önemli ölçüde artırdığını ortaya koymuşlardır. Bu çalışma, kodlardaki güvenlik açıklarını NLP yöntemleriyle tanımlamanın etkinliğini ve zorluğunu göstererek, literatüre dönüştürücü tabanlı mimarilerin kod analizi ve güvenlik açığı saptamada yüksek performans sağlayabileceğine dair somut ve yenilikçi bir katkı sunmuştur.

Ahmad vd. (2020) kapsamında, ABD gazilerine ilişkin travma sonrası stres bozukluğu (Post-Traumatic Stress Disorder, PTSD) ile ilgili karar metinlerinden oluşturulan veri kümesindeki cümleler, çift yönlü uzun-kısa süreli bellek (Bidirectional Long Short-Term Memory, BiLSTM) mimarisine dayanan derin öğrenme modeliyle söylemsel işlevlerine göre otomatik olarak sınıflandırılmıştır. Her cümle, olay, gerekçe ve bağlam dâhil altı retorik kategoriye ayrılmış; modelin söylem sınıflandırmasında yüksek doğruluk sağladığı belirlenmiştir. Böylece hukuk metinlerinin yapısal analizinde derin öğrenmenin etkinliği ortaya konmuş, gerekçe ile bağlam ayrımının otomatik tespiti için yöntemsel bir temel sunulmuş ve söz konusu yaklaşımın farklı hukuk alanlarına genişletilerek yargı süreçlerindeki iş yükünü azaltma potansiyeli vurgulanmıştır (Ahmad vd., 2020).

Liu vd. (2019), dönüştürücü tabanlı dil modeli olan BERT'in önceden eğitilmiş temsillerini çoklu görevli derin sinir ağı yapısıyla birleştirerek, doğal dil anlama (NLU) alanında çok görevli bir öğrenme mimarisi (MT-DNN) geliştirmiştir. Önerilen yaklaşım, tek ve ikili cümle sınıflandırması, metin benzerliği ve alaka sıralaması gibi farklı görevleri ortak temsiller üzerinden ele alarak, görevler arası bilgi paylaşımını ve aktarım öğrenmesini mümkün kılmıştır. Model, BERT tabanlı paylaşımlı katmanlar üzerine inşa edilmiş ve görevler arası ortak öğrenmenin ardından her görev için ayrı ayrı ince ayar yapılmıştır. Bu çalışma, çoklu görevli öğrenmenin az veriyle gerçekleştirilen senaryolarda dahi genelleme kapasitesini artırarak, tek görevli modellere kıyasla daha güçlü dil temsilleri sunduğunu ortaya koymaktadır.

2.3. Veri Dengeleme

Sınıf dengesizliği, bir sınıflandırma probleminde sınıflar arasında ciddi oranda örnek sayısı farkı bulunması durumudur. Bu dengesizlik, makine öğrenmesi algoritmalarının

çoğunluk sınıfa ait örneklere aşırı uyum göstermesine ve karar sınırlarının bu sınıf lehine kaymasına yol açabilmektedir (He ve Garcia, 2009). Bu bağlamda, örnek sayısı fazla olan sınıf çoğunluk sınıf, az sayıda gözleme sahip olan ise azınlık sınıf olarak tanımlanır. Dengesiz veri setlerinde genel doğruluk (accuracy) gibi küresel performans ölçütleri yüksek çıkabilse de, azınlık sınıfına yönelik tahmin hatalarının ciddi seviyelere ulaşması uygulama açısından sorun teşkil etmektedir. Özellikle tıp, finans ve hukuk gibi hata maliyetinin yüksek olduğu alanlarda bu sorun daha kritik hale gelmekte olup, bu durum kapsamlı literatürde çeşitli çalışmalarla ortaya konmuştur (Rout vd., 2018).

Dengesiz veri sorununa yönelik çözüm yaklaşımları, literatürde üç ana kategoride sınıflandırılmıştır: veri düzeyindeki örnekleme yöntemleri (data-level sampling methods), algoritma düzeyinde maliyet duyarlı öğrenme teknikleri (algorithm-level cost-sensitive learning) ve topluluk temelli ya da melez stratejiler (ensemble-based or hybrid strategies) (He ve Garcia, 2009; Rout vd., 2018).

Veri düzeyindeki yöntemler sınıflar arası dengesizliği doğrudan örnekleme yoluyla dengelemeyi hedeflemektedir. Bu kapsamda, rastgele üst örnekleme (random oversampling) ve rastgele alt örnekleme (random undersampling) en temel uygulamalar arasında yer almaktadır (He ve Garcia, 2009). Ancak, yinelenen örneklerden kaynaklanabilecek aşırı uyum (overfitting) riskine karşı, daha karmaşık teknikler geliştirilmiştir. Bu bağlamda, Sentetik Azınlık Sınıfı Aşırı Örnekleme Tekniği (Synthetic Minority Over-sampling Technique, SMOTE), azınlık sınıfına ait örneklerin yakın komşuları arasında yeni örnekler oluşturarak karar sınırlarını daha iyi genelleştirmeyi hedeflemektedir (He ve Garcia, 2009). SMOTE'un türevleri olan Sınır-Çizgisine Dayalı SMOTE (Borderline-SMOTE, B-SMOTE) ve Uyarlanabilir Sentetik Örnekleme Yöntemi (Adaptive Synthetic Sampling, ADASYN) gibi yöntemler, özellikle karar sınırına yakın karmaşık bölgelerde etkili sonuçlar vermektedir. Bununla birlikte, bilgilendirilmiş eksiltme yöntemleri (informed undersampling methods) arasında yer alan Yakınlık Azaltma Yöntemi (NearMiss), Tomek Bağlantıları (Tomek Links) ve Tek Taraflı Seçim (One-Sided Selection, OSS) gibi algoritmalar, çoğunluk sınıfından silinecek örneklerin veri yapısına göre seçilmesi esasına dayalı olup bu yolla bilgi kaybı azaltılmaya çalışılmaktadır.

Algoritma düzeyinde, maliyet duyarlı öğrenme (cost-sensitive learning) yaklaşımları öne çıkmaktadır. Bu tip yöntemlerde, sınıflar arası hata maliyetleri farklılaştırılmakta; azınlık sınıfına ait yanlış sınıflandırmaların etkisi modele daha yüksek ağırlıklarla yansıtılmaktadır. Söz konusu stratejiler, karar ağaçlarında eşik kaydırma, düğüm bölme ve budama gibi karar mekanizmalarına; yapay sinir ağlarında ise hata fonksiyonlarına ve öğrenme oranlarına maliyet faktörleri entegre edilerek uygulanmaktadır (He ve Garcia, 2009).

Veri ve algoritma düzeyli yaklaşımların birlikte kullanılmasıyla geliştirilen topluluk öğrenme yöntemleri (ensemble learning methods), sınıf dengesizliği ile mücadelede etkili ve esnek çözümler sunmaktadır. Bu grupta öne çıkan tekniklerden biri olan SMOTE ile Güçlendirilmiş Sınıflandırma (SMOTE-Boost), SMOTE ile elde edilen örneklerin ardışık olarak yeniden örneklenmesi yoluyla çoğunluk sınıfına karşı duyarlılığı artırmaktadır. Benzer şekilde, Kolay Topluluk Yöntemi (EasyEnsemble), çoğunluk sınıfından farklı örnek alt kümeleri alarak her biri için zayıf sınıflayıcılar eğitmekte ve nihai kararı bu sınıflayıcıların birleştirilmesiyle vermektedir. Öte yandan, Komşuluk Dengeli Topluluk Yöntemi (Neighbourhood-Balanced Bagging), hem azınlık hem de çoğunluk sınıfının yerel yapısını dikkate alarak örnekleme yaparak dengeleme sağlamaktadır (Rout vd., 2018). Literatürde bu tür melez yöntemlerin, bilgi kaybını azalttığı ve genel genelleme hatasını düşürdüğü belirtilmiştir.

Doğal dil işleme alanındaki sınıflandırma çalışmalarında sınıf dengesizliği özellikle azınlık sınıfının yeterince temsil edilmediği durumlarda, öğrenme algoritmalarının yanlı tahminlerde bulunmasına ve performans düşüşüne yol açmaktadır. Bu nedenle dengesiz veri setleri, hem hukuk alanında hem de diğer alanlarda gerçekleştirilen makine öğrenmesi uygulamalarında önemli bir problem olarak karşımıza çıkmaktadır. Literatürde, sınıf dengesizliği problemini çözmek amacıyla çeşitli veri dengeleme stratejileri önerilmiştir. Aşağıda doğal dil işleme kapsamında veri dengeleme konusunda yapılmış güncel çalışmalar aktarılmıştır. Bu çalışmaların incelenmesi, tez kapsamında hukuk kararlarının sınıflandırılması için önerilen veri dengeleme yöntemlerinin belirlenmesinde yol göstericidir.

Wang vd. (2025) tarafından geliştirilen LA-TAG (Large Language Model-based Augmentation on Text-Attributed Graphs) adlı çerçeve, metin öznitelikli grafiklerdeki sınıf dengesizliği sorununa yönelik yenilikçi bir çözüm sunarak literatüre katkı sağlamaktadır. Büyük dil modelleri kullanılarak azınlık sınıfına ait düğümler için sentetik metinler üretilmiş, bu metinler semantik olarak uygun biçimde grafik yapısına entegre edilmiş ve son aşamada güncellenmiş grafik, Grafik Evrişimsel Ağ (Graph Convolutional Network, GCN) ile eğitilerek sınıflandırma gerçekleştirilmiştir. Çalışma, metin düzeyinde gerçekleştirilen bu özel artırma stratejisiyle klasik yöntemleri geride bırakmış; özellikle azınlık sınıflardaki performans farkını azaltarak hem F-ölçütü hem de genel doğrulukta iyileşme sağlamıştır. Bu yönüyle LA-TAG, metinsel düğüm sınıflandırmasında LLM destekli veri artırımı ve bağlantı tahmini kombinasyonunun etkili biçimde kullanılabileceğini gösteren bir yaklaşım sunmaktadır.

Thapa vd. (2025) tarafından gerçekleştirilen bir çalışmada, Devanagari alfabeli sınırlı sayıda kaynak bulunan diller için dil tanımlama, nefret söylemi tespiti ve hedef sınıflandırmasına yönelik üç alt görevden oluşan bir yarışma düzenlenmiştir. Özellikle sınıf dengesizliği sorununun etkisini azaltmak amacıyla çeşitli veri dengeleme stratejileri uygulanmıştır. Bu bağlamda, bazı katılımcılar azınlık sınıfları temsil eden örneklerin sayısını artırmak için veri artırımı (data augmentation) tekniklerini kullanmış; bazıları ise çoğunluk sınıflardan örnekleri azaltarak rastgele alt örnekleme (random undersampling) yöntemini uygulamıştır. Ek olarak, odaklanma kaybı (focal loss) fonksiyonu, sınıf ağırlıklandırması (class weighting), dikkat mekanizmaları (attention mechanisms) ve düşük güven puanlı örneklerin büyük dil modelleri (large language models) aracılığıyla yeniden sınıflandırılma (reclassification) kullanılmıştır. Bu yaklaşımların, özellikle az temsil edilen topluluk hedefli nefret söylemlerinin daha doğru tespitinde etkili olduğu görülmüştür (Thapa vd., 2025).

Beliveau vd. (2024) tarafından yapılan bir çalışma, sınıf dengesizliği ve örnek azlığının belirgin olduğu Danca beyin MRG raporlarında üç önemli patolojik bulgunun (FCD, MTS, HA) tespiti için çağdaş doğal dil işleme yöntemlerinin etkinliğini karşılaştırmalı olarak inceleyerek literatüre katkı sağlamaktadır. Danca diline özel ve çok dilli BERT modelleri, cümle gömme temelli SetFit yaklaşımı ve büyük dil modelleri değerlendirilmiş; bu modeller, sınıf dengesini gözeten özel ön işleme ve ince ayar

stratejileriyle test edilmiştir. Özellikle DanskBERT, azınlık sınıfların başarılı biçimde tanımlanmasında öne çıkmış; böylece düşük oranlı klinik bulguların tespitinde BERT tabanlı modellerin güvenilirliğini ortaya koymuştur. Çalışma, dengesiz tıbbi metinlerde dil ve görev uyumlu modellerin etkili sonuçlar verebileceğini göstererek, az sayıda kaynak bulunan diller için NLP uygulamalarına yönelik özgün bir yöntemsel çerçeve sunmaktadır.

Al Katat vd. (2024) tarafından gerçekleştirilen sistematik derleme çalışması, Arapça duygu analizi (Arabic Sentiment Analysis, ASA) alanındaki güncel yöntemleri inceleyerek sınıf dengesizliği sorununa yönelik yaklaşımları kapsamlı biçimde değerlendirmektedir. İncelenen çalışmaların büyük bölümünde verinin dengesiz yapısının model performansı üzerinde olumsuz etkiler yarattığı vurgulanmakta; bu nedenle çeşitli veri dengeleme tekniklerine başvurulduğu gözlemlenmektedir. Özellikle sentetik örnekleme yöntemlerinden SMOTE ve onun nominal değişkenleri dikkate alan bir türevi olan Nominal Kategorik Özniteliklerle Sentetik Sınıfı Azınlık Aşırı Örnekleme (Synthetic Minority Over-sampling Technique for Nominal and Continuous features, SMOTENC), azınlık sınıfının temsil gücünü artırmak amacıyla tercih edilmiştir. Öte yandan, bazı çalışmalar veriyi yapay olarak artırmak yerine, orijinal dengesiz veri kümesi üzerinde çalışarak performans kıyaslamaları yapmış ve dengesiz verilerle daha yüksek doğruluk oranlarının elde edilebildiğini bildirmiştir. Ayrıca, farklı örnekleme stratejilerinin birlikte değerlendirildiği karşılaştırmalı analizlerde, örnekleme yönteminin seçiminde hem veri kümesinin yapısal özelliklerinin hem de uygulanan model türünün belirleyici olduğu sonucuna ulaşılmıştır. Bu yönüyle çalışma, ASA literatüründe veri dengesizliği sorununu farklı boyutlarıyla ele alarak hem algoritmik hem de ön işleme düzeyinde uygulanabilir stratejiler sunması bakımından literatüre katkı sağlamaktadır (Al Katat vd., 2024).

Hu vd. (2024) tarafından yürütülen çalışmada, afet temalı tweet'lerin ikili duygu sınıflandırmasında karşılaşılan veri dengesizliği problemi çeşitli yöntemlerle giderilmeye çalışılmıştır. Özellikle eksik anahtar kelime ve konum bilgilerinin, ad öbeği tanıma (Named Entity Recognition) teknikleri aracılığıyla otomatik olarak doldurulması sağlanmış; böylece azınlık sınıfa ait örneklerin temsili güçlendirilmiştir. Sınıflandırma aşamasında ise farklı öğrenicilerin çıktılarını birleştirerek genelleme kabiliyetini artıran

topluluk yöntemlerden yararlanılmıştır. Bu kapsamda, her bir modelin farklı veri altkümeleriyle eğitildiği torbalama yöntemi (bagging), çoklu model çıktılarının çoğunluk kuralına göre birleştirildiği oylama yöntemi (voting) ve bir üst düzey öğreniciyle farklı modellerin çıktılarının bir araya getirildiği yığma yöntemi (stacking) uygulanmıştır. Bu stratejiler aracılığıyla hem çoğunluk sınıfının baskın etkisi azaltılmış hem de azınlık sınıfın daha doğru tahmin edilmesi sağlanmıştır. Böylece sınıf dengesizliğinden kaynaklanan öğrenme sorunlarının önüne geçilerek daha dengeli ve güvenilir bir sınıflandırma başarısı elde edilmiştir (Hu vd., 2024).

Fudickar vd. (2024) tarafından yürütülen bir çalışmada, bel ağrısı olan hastaların uygun tedavi birimlerine yönlendirilmesine yönelik makine öğrenimi temelli bir karar destek sisteminin doğruluğu artırılmaya çalışılmıştır. Özellikle veri setindeki sınıf dengesizliği sorunu dikkat çekmiş; rehabilitasyon kategorisi ağırlıklı iken, nöroşirürji gibi kategoriler azınlıkta kalmıştır. Bu dengesizlik, Sentetik Azınlık Üst Örnekleme Tekniği (Synthetic Minority Oversampling Technique, SMOTE) uygulanarak giderilmeye çalışılmıştır. SMOTE ile her bir sınıfın eğitim verilerinde eşit sayıda örnekle temsil edilmesi sağlanmıştır. Ayrıca, elektronik sağlık kayıtlarından doğal dil işleme ile elde edilen sevk nedenleri modele dahil edilerek veri zenginleştirilmiş; bu ek nitelikler, özellikle anesteziyoloji ve rehabilitasyon sevkleri, modelin performansını anlamlı şekilde artırmıştır. Bu yönüyle çalışma veri dengesizliğinin üstesinden gelinmesinde hem örnekleme temelli dengeleme yöntemlerinin hem de nitel veri zenginleştirmenin etkili bir biçimde kullanılabileceğini göstermektedir (Fudickar vd., 2024).

3. MATERYAL ve YÖNTEM

Bu bölümde, çalışmada kullanılan veri seti, öznitelik çıkarım süreçleri, uygulanan veri dengeleme ve makine öğrenmesi yöntemleri ile performans değerlendirme kriterlerine ilişkin bilgilere yer verilmiştir.

3.1. Materyal

Bu çalışmada kullanılan mahkeme kararı metinleri, Türkiye Cumhuriyeti Adalet Bakanlığı tarafından işletilen Ulusal Yargı Ağı Projesi (UYAP) üzerinden temin edilmiştir. UYAP, Türkiye'de adli ve idari yargı birimleri ile Adalet Bakanlığı teşkilatını tekil, bütünleşik ve elektronik imza destekli bir platformda buluşturan merkezi bilişim altyapısıdır.

UYAP'ın içerik odaklı alt modüllerinden biri olan Mevzuat ve İçtihat Programı, hâkim-savcılar, avukatlar ve araştırmacılar için kanun, yönetmelik, kararname ve yüksek yargı içtihatlarını çevrimdışı dahi erişilebilir kılan kapsamlı bir bilgi bankasıdır. UYAP Mevzuat ve İçtihat Programı, yapılan tez çalışması kapsamında yürütülen hukuk alanındaki doğal dil işleme çalışmalarında kullanılan metinlerin güvenilirliği ve doğruluğu açısından tercih edilen birincil kaynaktır.

UYAP Mevzuat ve İçtihat Programı'ndan alınan ham veri, 1975-2016 yılları arasında verilmiş toplam 6425 yüksek mahkeme kararından oluşmaktadır. Bu kararlar; Yargıtay Ceza ve Hukuk Daireleri, Ceza Genel Kurulu, Hukuk Genel Kurulu ile Danıştay Daireleri ve İdari Dava Daireleri Kurulunu kapsamaktadır. Şekil 3.1'de veri seti içerisinde yer alan örnek bir mahkeme kararının ekran görüntüsü verilmiştir.

YARGITAY

2. HUKUK DAİRESİ

Esas Numarası: 2010/5750

Karar Numarası: 2011/7526

Karar Tarihi: 03.05.2011

Taraflar arasındaki davanın yapılan muhakemesi sonunda mahalli mahkemece verilen hüküm temyiz edilmekle evrak okunup, gereği görüşülüp düşünüldü:

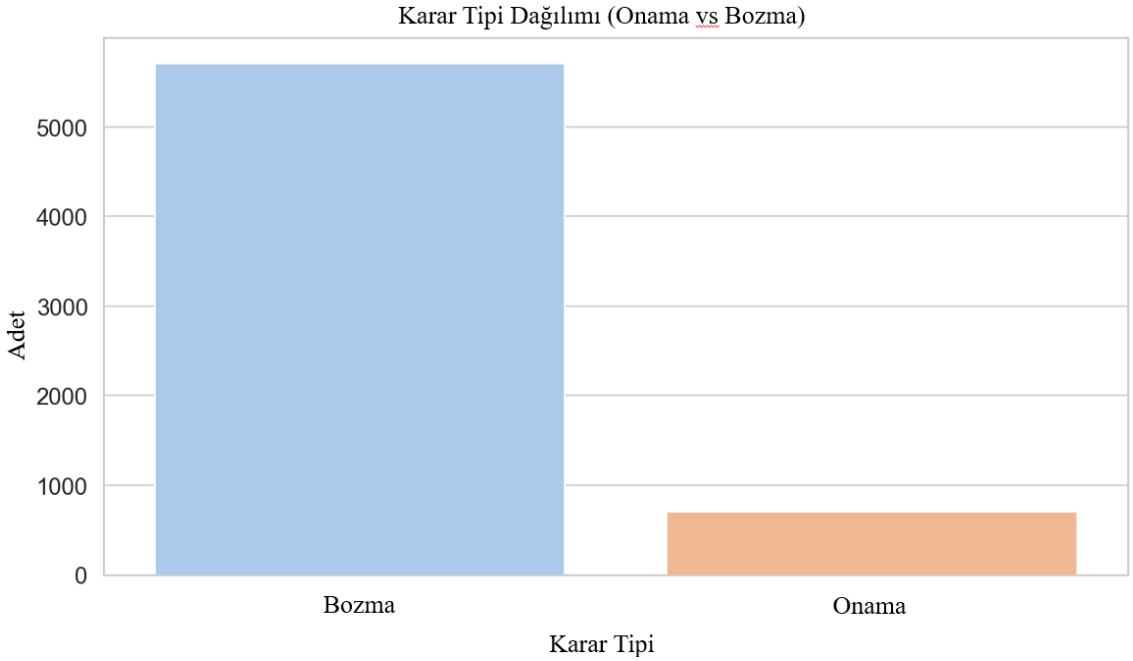
KARAR : Dava, evliliğin tespitine ve tesciline yöneliktir. Görev kamu düzenine ilişkindir. Mahkemece yargılamanın her aşamasında kendiliğinden dikkate alınması gerekir. 4787 sayılı Aile Mahkemelerinin Kuruluş, Görev ve Yargılama Usullerine Dair Kanun'un 4. maddesi, 4721 sayılı Türk Medeni Kanunu'nun ikinci kitabından üçüncü kısım hariç olmak üzere (TMK m. 118-395) kaynaklanan bütün davaların Aile Mahkemesinde bakılacağını hükme bağlamıştır. Şu halde görevsizlik kararı verilmesi gerekirken, yargılamaya devamla yazılı şekilde hüküm tesisi doğru görülmemiştir.

SONUÇ : Temyiz edilen hükmün yukarıda açıklanan nedenlerle (BOZULMASINA), bozma sebebine göre diğer temyiz isteklerinin şimdilik incelenmesine yer olmadığına, temyiz peşin harcının istek halinde yatırana geri verilmesine, işbu kararın tebliğinden itibaren 15 gün içinde karar düzeltme yolu açık olmak üzere, 03.05.2011 gününde oybirliğiyle karar verildi.

Şekil 3.1. Veri setinden rastgele seçilen bir mahkeme kararının ekran görüntüsü

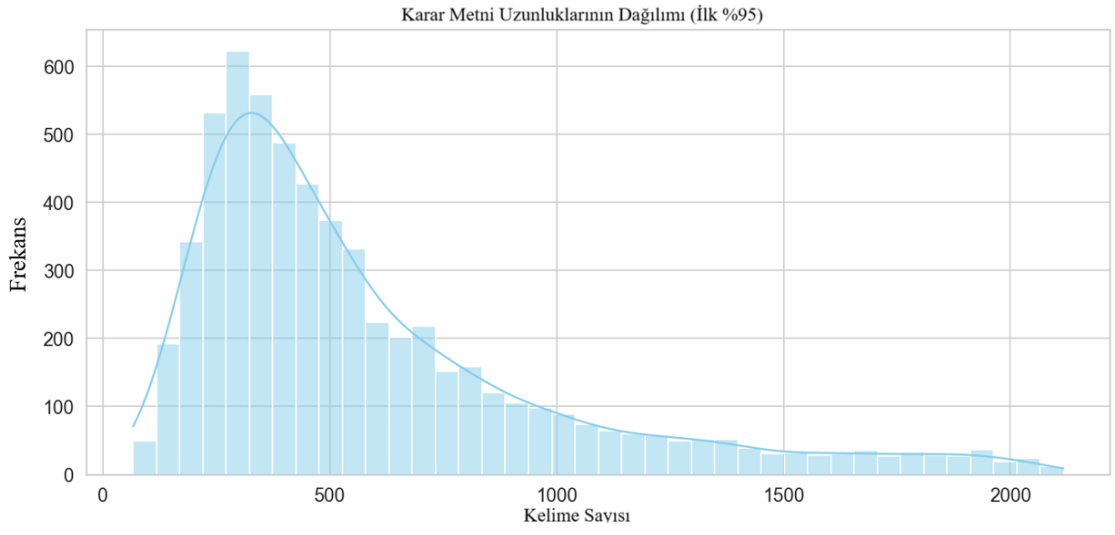
Yüksek Mahkeme kararlarının gerekçe bölümü, mahkemenin hangi olgulara ve hangi yasa hükümlerine dayanarak sonuca ulaştığını adım adım açıklar; dosyadaki olayları özetler, ilgili kanun maddelerini gösterir ve bu maddelerin uyuşmazlığa nasıl uygulandığını gerekçelendirir. Hüküm (sonuç) bölümü ise bu gerekçeye dayanarak verilen bağlayıcı kararı tek cümleyle ifade eder (ör. “bozulmasına” veya “onamasına”), ayrıca yargılama giderleri, harç iadesi, karar düzeltme-temyiz süreleri ve oy birliği/karşı oy bilgisi gibi infazı doğrudan ilgilendiren emir ve talimatları içerir.

Veri setinde 708’i onama (%11), 5 704’ü bozma (%89) hükmü içeren karar bulunmakta olup çalışma kapsamına girmeyen beş adet Anayasa Mahkemesi kararı ise analiz sürecinden önce çıkarılmıştır. Veri setindeki kararların dağılımı Şekil 2’de gösterilmiştir.



Şekil 3.2. Onama ve Bozma kararlarının sayısal dağılımı

Şekil 3.3'te karar metinlerinin kelime sayılarına göre dağılımı gösterilmiş, veri setindeki uç (aykırı) değerlerin etkisini azaltmak amacıyla karar metni uzunluklarının en küçükten en büyüğe sıralandığında ilk %95'lik kısmını oluşturan gözlemler dikkate alınmıştır. Mahkeme kararlarındaki ortalama kelime sayısı yaklaşık 600'dür.



Şekil 3.3. Karar metinlerinin kelime sayılarına göre dağılımı

3.2. Yöntem

Bu bölümde kullanılan veri ön işleme ve sınıflandırma yöntemleri ile önerilen yaklaşım açıklanmıştır.

3.2.1. Veri Ön İşleme ve Öznitelik Mühendisliği

Mahkeme kararlarının sınıflandırılmasında metinlerin makine öğrenmesi algoritmalarıyla uyumlu biçimde temsil edilebilmesi amacıyla bir veri ön işleme süreci uygulanmıştır.

Vektörizasyon işleminde, farklı yaklaşımlar temel alınarak öznitelik matrisleri oluşturulmuştur. Bu kapsamda, her terimin belge içerisindeki ağırlığını istatistiksel olarak ölçen TF-IDF yöntemi uygulanmıştır. Buna ek olarak, metinlerin anlamsal boyutunu yakalamak amacıyla Word2Vec yöntemi kullanılmıştır. Bu yöntemde, sözcüklerin benzer bağlamlarda birlikte bulunma olasılıklarına dayalı olarak öğrenilen yoğun vektör temsilleri aracılığıyla, her belge sürekli bir vektör uzayında temsil edilmiştir.

Sınıf dağılımındaki dengesizliklerin giderilmesi amacıyla çeşitli veri dengeleme stratejilerine başvurulmuştur. Bu kapsamda, sentetik üst örnekleme (SMOTE, rastgele alt örnekleme ve bu iki yöntemin birlikte kullanıldığı hibrit dengeleme teknikleri) uygulanmıştır. Ayrıca, metin yapısına uygun biçimde anlamsal bütünlüğü koruyarak örnek üretme yeteneğine sahip büyük dil modelleri aracılığıyla sentetik örnekleme yapılmıştır.

Son olarak, öznitelik uzayı uzman görüşü doğrultusunda içerik temelli göstergelerle zenginleştirilmiştir. Bu doğrultuda, belirli kavramların metinlerde geçip geçmediğini tespit etmeye yönelik ikili nitelikler tanımlanmış; böylece her metin hem istatistiksel hem anlamsal hem de içerik odaklı açılardan temsili hâle getirilmiştir. Bu bölümde kullanılan vektörizasyon ve veri dengeleme yöntemleri açıklanmıştır.

3.2.2. Terim Frekansı – Ters Doküman Frekansı

TF-IDF, vektör uzayı modelinde bir terimin belge için önemini ölçen istatistiksel bir ağırlıklandırma yöntemidir. TF-IDF, hem terimin belge içindeki sıklığını (Terim Frekansı - TF), hem de tüm belge koleksiyonundaki nadirliğini (Ters Doküman Frekansı - IDF)

dikkate alır. TF-IDF'in bu çift bileşenli yapısı, kelimelerin istatistiksel ayırt ediciliğini yansıtan bir temsil biçimi olarak tanımlanmıştır. Bir kelimenin belge içinde sık, ancak belge koleksiyonunda nadir olması, kelimenin içeriğe dair güçlü bir ayırt edicilik taşıdığını gösterir (Zhang vd., 2011; Kim ve Gil, 2019).

Bu nedenle TF-IDF, belge sınıflandırma, bilgi erişimi ve anahtar kelime çıkarımı gibi metin madenciliği uygulamalarında öznitelik mühendisliğinin temel yöntemlerinden biridir. Yani TF değeri arttıkça, söz konusu terimin o belge için taşıdığı bağlamsal önem artmaktadır. TF'in formülasyonu Denklem 3.1 ile verilmiştir.

$$TF_{i,j} = \frac{n_{i,j}}{\sum_k n_{k,j}} \tag{3.1}$$

Burada $n_{i,j}$, t_i teriminin d_j belgesinde kaç kez geçtiğini, $\sum_k n_{k,j}$ ise d_j belgesindeki tüm terimlerin toplam frekansını göstermektedir. Belge Frekansı (DF) ise bir terimin kaç farklı belgede geçtiğinin bir ölçüsüdür. Bu değer yükseldikçe bir terimin bulunduğu belge sayısının arttığı anlamına gelir, yani o terimin ayırt edicilik gücünün düşük olduğu sonucu çıkar. Bundan ötürü ters belge frekansı (IDF) formüle edilerek nadir rastlanan ve ayırt ediciliği yüksek terimlere daha yüksek ağırlık değerleri atanması sağlanmıştır. Denklem 3.2'de IDF formüle edilmiştir.

$$IDF_{i,j} = \ log \frac{|D|}{|d_j \in D : t_i \in d_j|} \tag{3.2}$$

Bu formülde $|D|$, veri setindeki toplam belge sayısını; t_i, i'inci terimi, d_j j'inci belgeyi temsil etmektedir. $\left|d_j \in D : t_i \in d_j\right|$ ifadesi ise terim t_i'nin yer aldığı belge sayısını göstermektedir (Kim ve Gil, 2019).

TF-IDF'in kolay uygulanabilir, yorumlanabilir ve çok boyutlu verilerde dahi makul doğruluk sağlayabilmesi sayesinde bilgi erişimden belge sınıflandırmaya kadar pek çok metin madenciliği görevinde hâlen temel bir yöntem olarak kullanılmaktadır (Kim ve Gil, 2019).

Ancak TF-IDF'in kelime sırasını, çok kelimeli kalıpları ve bağlamı dikkate almayan yapısı, derin semantik ilişkileri yakalama konusunda sınırlıdır. Bundan dolayı güncel çalışmalarda çoğu kez konu modelleme veya anlamsal gömme yöntemleriyle birlikte hibrit şekilde kullanılmaktadır.

Sonuç olarak TF-IDF, derin öğrenme temelli yaklaşımların yükselişine rağmen; hesaplama maliyeti düşük, yorumlanabilir ve basamaklı modelleme kurgularına kolayca entegre edilebilir olması nedeniyle, metin temsili literatüründe önemli bir yere sahiptir (Zhang vd., 2011; Kim ve Gil, 2019).

3.2.3. Word2Vec

Word2Vec, kelimeleri sürekli vektörler (dense vectors) şeklinde temsil eden ve bu temsiller arasında anlamsal ve sözdizimsel ilişkileri yakalamayı amaçlayan bir kelime gömme (word embedding) yöntemidir. Sürekli vektörler, her kelimeyi sabit uzunlukta, genellikle düşük boyutlu ve tüm bileşenleri sıfır olmayan reel sayılardan oluşan bir vektörle ifade eder. Word2Vec, kelimelerin kullanıldıkları bağlamlara göre benzer anlamlar taşıyabileceği düşüncesine dayanır; yani benzer cümle ya da çevrelerde geçen kelimelerin anlam açısından da birbirine yakın olacağı varsayılmaktadır (Mikolov vd., 2013a). Model, her kelimeye $\mathbb{R}^d$ uzayında yoğun (dense) bir uzayında bir vektör atayarak, bu vektörleri olasılıksal bir dil modeli kapsamında optimize eder. Bu süreçte amaç, bir kelimenin belirli bir bağlamda gözlemlenme olasılığını en çoklayan (maximum-likelihood) bir parametre tahmini gerçekleştirmektir (Mikolov vd., 2013a; Mikolov vd., 2013b).

Word2Vec'in temel mimarilerinden biri olan atlamalı ikili sözcük modeli (Skip-Gram), her hedef kelimeyi kullanarak, belirli bir pencere içinde yer alan bağlam sözcüklerini tahmin etmeyi amaçlar. Modelin optimize ettiği ortalama log-olasılık fonksiyonu Denklem 3.3'te tanımlanmıştır:

$$\mathcal{L} = \frac{1}{T}\sum_{t=1}^{T}\sum_{\substack{-c\le j\le c \\ j\neq 0}} log\, p\left(w_{t+j} \mid w_t\right) \tag{3.3}$$

Burada $\mathcal{L}$ modelin öğrenmeye çalıştığı amaç fonksiyonudur (loss function). Buradaki ifade, tüm eğitim kümesi boyunca elde edilen ortalama log-olasılık değeridir ve maksimize edilmeye çalışılır. w_t, cümledeki t'inci sıradaki kelime; model için hedef (merkez) kelimedir. T, eğitim kümesindeki toplam kelime sayısı olup ortalama alınması için kullanılır. Model, her bir pozisyon t için hedef kelime w_t'yi ele alarak işlem yapar. c, bağlam penceresinin yarıçapıdır. Örneğin $c = 2$ ise, hedef kelimenin sağında ve solunda iki kelimeye kadar olan komşular dikkate alınır. $-c \leq j \leq c$ ve $j \neq 0$ hedef kelimenin çevresindeki (bağlamdaki) kelimeleri belirlemek için kullanılır. $j = 0$ atlanır çünkü w_t kendi kendine bağlam oluşturmaz. w_{t+j}, w_t'nin bağlamında yer alan komşu kelimelerdir. Buradaki j, bağlam konumunu temsil eder. $log\, p(w_{t+j} \mid w_t)$ merkez kelime (w_t)'nin verildiği durumda, bağlamdaki kelime w_{t+j}'nin doğru tahmin edilme log-olasılığıdır. Modelin çıktısı budur ve tüm bağlam kelimeleri için toplanır.

Word2Vec'in Skip-Gram modeli içinde, verilen bir hedef kelime bağlamında, belirli bir çıktı kelimesinin görülme olasılığı (softmax) fonksiyonu Denklem 3.4'teki formül ile hesaplanır:

$$p(w_O \mid w_I) = \frac{exp\left(\left(v'_{w_O}\right)^{\top} \boldsymbol{v}_{w_I}\right)}{\sum_{w=1}^{|V|} exp\left((\boldsymbol{v}'_{\boldsymbol{w}})^{\top} \boldsymbol{v}_{w_I}\right)} \tag{3.4}$$

Burada $|V|$ sözlük boyutu, w_I ve w_O sırasıyla girdi ve çıktı kelimeleridir. $\left(v'_{w_O}\right)^{\top}$ bağlam kelimesinin transpozu alınan vektörü, v_{w_I} girdi kelimesinin vektörü, $(\boldsymbol{v}'_{\boldsymbol{w}})^{\top}$ Word2Vec modelinde bir bağlam kelimesinin çıktıya özgü vektörünün transpozudur. Bu formül sayesinde, eğitim sürecinde model semantik olarak benzer veya ilişkili kelimelerin, vektör uzayında birbirine yakın yerlerde temsil edilmesi gerektiğini öğrenir. Örneğin: "queen" ve "king" gibi sözcükler benzer bağlamlara sahip olduklarından, vektörleri de yakınlaşacaktır.

Softmax fonksiyonunun her adımdaki yüksek hesaplama yükü nedeniyle, negatif örnekleme (negative sampling) adı verilen daha verimli bir yöntem önerilmiştir. Negatif örnekleme yaklaşımında sadece gerçek bağlam kelimesi ve rastgele seçilen sınırlı sayıda

negatif kelime (o adımda gözlenen bağlamda yer almayan, yani hedef kelimeyle ilişkisi yanlış kabul edilen sözcük) dikkate alınarak model güncellenmektedir. Negatif Örnekleme (Negative Sampling) için kullanılan yaklaşık öğrenme formülü Denklem 3.5'te gösterilmiştir.

$$log\,\sigma\left({v'_{w_O}}^{\top} v_{w_I}\right) + \sum_{i=1}^{k} E_{w_i \sim P_n(w)}\left[log\,\sigma\left(-{v'_{w_i}}^{\top} v_{w_I}\right)\right] \tag{3.5}$$

Burada $\sigma(.)$ sigmoid fonksiyonunu, E fonksiyonun beklenen değerini, k her pozitif örnek için seçilen negatif örnek sayısını, $P_n(w)$ negatif örneklerin seçildiği gürültü dağılımını, v'_{w_i} rastgele negatif örneklerin çıktı vektörlerini göstermektedir.

Word2Vec mimarisinde iki temel yaklaşım bulunmaktadır: Skip-Gram ve Continuous Bag-of-Words (CBOW). Skip-Gram modeli, belirli bir merkez kelimeyi kullanarak bu kelimenin etrafındaki bağlam kelimelerini tahmin etmeyi amaçlar. Özellikle düşük frekanslı ve anlamsal olarak zengin kelimelerin temsillerini öğrenmede bu yapı etkilidir. Öte yandan, CBOW modeli ise bu süreci tersine çevirerek, çevredeki bağlam kelimelerini kullanarak merkez kelimeyi tahmin eder. Bağlam vektörlerinin ortalaması alınarak hedef kelimenin olasılığı hesaplanır. CBOW, sık karşılaşılan kelimeler için daha istikrarlı temsiller üretir ve Skip-Gram'e kıyasla daha hızlı eğitilir. Her iki model de bağlam ile kelime arasındaki ilişkileri istatistiksel olarak modellemeyi hedefler ve negatif örnekleme gibi tekniklerle ölçeklenebilir hale getirilir (Mikolov vd., 2013a; Mikolov vd., 2013b).

Word2Vec öğrenilen vektörlerin anlamsal ve sözdizimsel düzenliliklerini kodlayabilmektedir. Örnek bir analoji Denklem 3.6'da ifade edilmiştir:

$$\overrightarrow{\text{king}} - \overrightarrow{\text{man}} + \overrightarrow{\text{woman}} \approx \overrightarrow{\text{queen}} \tag{3.6}$$

Bu tür lineer ilişkiler, modelin kelimeler arası anlam farklarını ve cinsiyet, ülke-başkent gibi ilişkileri vektör düzleminde başarıyla yansıttığını göstermektedir (Mikolov vd., 2013c).

Word2Vec ayrıca çok sözcüklü ifadelerin (multi-word expressions) tek bir vektörle temsil edilmesine de olanak sağlar. "word2phrase" aracı kullanılarak örneğin “New York” ya da “San Francisco” gibi kalıplaşmış yapılar tek bir belirteç olarak modellenebilir ve bu yapıların tekil vektör temsilleri elde edilebilir (Mikolov vd., 2013c).

Sonuç olarak, Word2Vec modeli, kelimeler arasındaki anlamsal ve sözdizimsel ilişkileri yoğun vektör temsilleri üzerinden başarılı bir şekilde öğrenerek, doğal dil işleme alanında birçok görev için güçlü ve etkili bir temel oluşturmaktadır.

3.2.4. Veri Dengeleme Yöntemleri

Sınıflandırma problemlerinde, öğrenme algoritmalarının azınlık sınıfa ait örnekleri yeterince temsil edememesi nedeniyle performans düşüklüğü meydana gelebilmektedir. Bu sorunun giderilmesine yönelik olarak literatürde çeşitli veri dengeleme yöntemleri geliştirilmiştir. Bu bölümde, çalışmada kullanılan veri dengeleme yaklaşımlarına yer verilmiştir.

3.2.5. Rastgele Alt Örnekleme

Rastgele alt örnekleme (Random Undersampling, RUS), dengesiz veri kümelerinde çoğunluk sınıfına ait gözlemlerin rastgele elenerek sınıf dağılımının dengelenmesini amaçlayan bir örnekleme tekniğidir. Rastgele alt örnekleme, özellikle doğal dağılımın sınıflandırıcı öğrenimi için elverişli olmadığı senaryolarda, veri kümesi hacmini küçülterek hesaplama yükünü azaltması bakımından literatürde sıkça önerilmektedir (Chawla, 2005).

RUS’un temel avantajı, model eğitimi süresini önemli ölçüde kısaltması ve bellek gereksinimini düşürmesidir. Bununla birlikte, rastgele silme işlemi karar sınırına yakın veya temsil gücü yüksek örnekleri ortadan kaldırabileceği için sınıflandırıcının genel hatları eksik öğrenmesine yol açma riski taşımaktadır. RUS yönteminde, çoğunluk sınıfına ait örneklerin ne oranda eleneceğini belirleyen örnekleme oranı (sampling ratio), yöntemsel bir hiperparametre olarak yapılandırılabilir. Bu parametrenin değeri, sınıflar arası dengeyi sağlamakla birlikte, bilgi kaybı riskini de belirlediğinden, modelin genel genelleme performansını doğrudan etkilemektedir (Mishra, 2017).

Büyük veri bağlamında yapılan ampirik çalışmalar, azınlık sınıf oranının %0,1 ile %1 arasında olduğu son derece dengesiz senaryolarda dahi, çoğunluk sınıfının %50'ye kadar rastgele azaltılmasıyla elde edilen modellerin, tam veri kümesiyle eğitilen modellere benzer eğri altındaki alan (Area under the curve, AUC) değerleri sunduğunu göstermiştir (Hasanin ve Khoshgoftaar, 2018). Dolayısıyla bilgi kaybı belirli sınırlar içinde tutulduğu sürece, RUS büyük hacimli veri setlerinde pratik bir dengeleme aracı olarak öne çıkmaktadır.

3.2.6. Üst Örnekleme

Üst örnekleme (oversampling), sınıf dağılımında dengesizlik içeren veri kümelerinde, azınlık sınıfına ait örneklerin temsiliyetini artırmak amacıyla sıklıkla başvurulan bir yeniden örnekleme stratejisi olarak literatürde yer bulmuştur. Bu yöntemde, azınlık sınıfına ait örneklerin rastgele seçilerek tekrar edilmesi yoluyla veri kümesine yeni gözlemler eklenmektedir (He ve Garcia, 2009). Böylece, sınıf dağılımındaki oransal dengesizlik giderilmeye çalışılmakta ve öğrenme algoritmalarının azınlık sınıfına ilişkin kavramları daha sağlıklı biçimde öğrenebilmesi hedeflenmektedir. Ancak, aynı örneklerin tekrar edilmesi sonucunda modelin bu kopya gözlemlere aşırı uyum sağlaması (overfitting) riski oluşabilmektedir. Bu durum, eğitim verisinde yüksek doğruluk sergilenmesine karşın, test verisinde genelleme başarısının düşmesine neden olabilmektedir. Dolayısıyla, üst örnekleme yöntemlerinin uygulanmasında özellikle aşırı uyuma yol açabilecek durumlara karşı SMOTE gibi daha gelişmiş sentetik örnekleme tekniklerinin tercih edilmesi önerilmektedir (He ve Garcia, 2009).

3.2.7. Sentetik Azınlık Sınıfı Üst Örnekleme Tekniği (SMOTE)

Bir üst örnekleme yöntemi olarak Chawla vd. (2002) tarafından önerilen Sentetik Azınlık Sınıfı Üst Örnekleme Tekniği (Synthetic Minority Over-sampling Technique, SMOTE) yöntemi, azınlık sınıfından alınan örnekler ile bunların k-en yakın komşuları arasındaki doğrusal interpolasyon yoluyla sentetik örnekler üreterek sınıf dengesini sağlamayı amaçlamaktadır. Böylece, eğitim verisi içinde azınlık sınıfın temsil gücü artırılırken, veri uzayındaki genel dağılım yapısı korunmuş olur (Chawla vd., 2002).

SMOTE algoritması, her bir azınlık sınıfı örneği için önce k-en yakın komşuyu belirler, ardından bu komşular arasından rastgele seçilenler ile örnekler arasında aralık değeri × fark formülüyle yeni veriler oluşturur. Algoritmanın sözde kodu Şekil 3.4'te verilmiştir:

Algorithm 1: SMOTE(T, N, k) – Sentetik Azınlık Örnekleme Tekniği

Girdi: T : Azınlık sınıfına ait toplam örnek sayısı
$N\%$: Üretilecek sentetik örnek yüzdesi
k : Her örnek için dikkate alınacak en yakın komşu sayısı

Çıktı : $\frac{N}{100} \times T$ adet sentetik azınlık örneği

if $N < 100$ **then** Azınlık örneklerini rastgele karıştır;
// Alt örnekleme gerekli
$T \leftarrow \frac{N}{100} \times T$;
// Yeni gerçek örnek sayısı
$N \leftarrow 100$;
// Oranı 100'e yuvarla

$N \leftarrow \text{int}(\frac{N}{100})$;
// Yüzdeyi tam kat hâline getir
$numattrs \leftarrow$ öznitelik sayısı;
$Sample[][] \leftarrow$ azınlık veri kümesi;
$newindex \leftarrow 0$;
// Sentetik örnek sayacı
for $i \leftarrow 1$ **to** T **do**
  k en yakın komşuyu bul, dizine kaydet: $nnarray$;
  for $n \leftarrow 1$ **to** N **do**
    $nn \leftarrow$ rastgele $[1, k]$ arası tamsayı;
    for $attr \leftarrow 1$ **to** $numattrs$ **do**
      $dif \leftarrow Sample[nnarray[nn]][attr] - Sample[i][attr]$;
      $gap \leftarrow \text{rand}(0, 1)$;
      $Synthetic[newindex][attr] \leftarrow Sample[i][attr] + gap \times dif$;
    end
    $newindex \leftarrow newindex + 1$;
  end
end

Şekil 3.4. SMOTE algoritmasının sözde kodu

SMOTE'nin sözde kodunda kullanılan değişkenler şöyle tanımlanmıştır: T azınlık sınıfına ait orijinal örnek (gözlem) sayısını ifade eder; $\%N$sentetik örnek üretim oranını gösterir (örneğin $\%200$ seçildiğinde her gerçek örnek için iki sentetik örnek üretilir); k her azınlık örneği için dikkate alınan en yakın komşu sayısını belirtir; $\left(N = \frac{N\backslash\%}{100}\right)$ ifadesi, yüzdelik değerin 100'lük dilimlere bölünmüş tamsayı katsayısını temsil ederken; $numattrs$ veri kümesindeki öznitelik (feature) sayısını tanımlamaktadır. $Sample_{i,j}$ i'inci azınlık örneğinin j'inci öznitelik değerini göstermektedir. $nnarray$, her azınlık örneğinin k en yakın komşusunun indislerini barındıran dizidir. difdeğişkeni, seçilen komşu ile hedef örnek arasındaki öznitelik farkını $Sample_{nn} - Sample_i$ ifade eder;

$[0,1]$ aralığında üretilen rastgele sürekli bir değer olup sentetik noktanın konumunu belirler; $Synthetic_{s,j}$ oluşturulan s'inci sentetik örneğin j'inci öznitelik değerini tutar; $newindex$ ise üretilen sentetik örnek sayacını izleyerek her üretimde bir artar (Chawla vd., 2002).

Teorik olarak SMOTE'nin en önemli katkısı, basit çoğaltma yöntemlerinin aksine sentetik örnekleri veri uzayında karar sınırlarını genişletecEk Şekilde üretmesidir. Böylece karar ağaçları veya destek vektör makineleri gibi modeller, azınlık sınıfı için daha genel ve kapsayıcı karar bölgeleri öğrenir; bu da aşırı öğrenme (overfitting) riskini azaltır ve yanlış negatif oranını düşürür (Chawla vd., 2002).

3.2.8. Büyük Dil Modelleri ile Veri Dengeleme

Büyük dil modelleri (Large Language Models, LLM), milyarlarca parametreye sahip derin öğrenme tabanlı yapılardır ve çok büyük metin koleksiyonları üzerinde eğitilerek dilin yapısal ve anlamsal örüntülerini öğrenebilmektedir. Büyük dil modelleri bağlama duyarlı biçimde tutarlı metin üretimi, soru yanıtlama, özetleme, çeviri ve sınıflandırma gibi çok çeşitli doğal dil işleme görevlerinde etkin biçimde kullanılmaktadır (Brown vd., 2020). Günümüzde hem akademik araştırmalarda hem de endüstriyel uygulamalarda sıklıkla başvurulan modeller arasında Üretken Önceden Eğitilmiş Dönüştürücü (Generative Pre-trained Transformer, GPT), Yollar Tabanlı Dil Modeli (Pathways Language Model, PaLM), Meta'nın Büyük Dil Modeli (Large Language Model Meta AI, LLaMA) ve Özelleştirilebilir Dil Modeli (Customizable Language Model, Claude) gibi örnekler yer almaktadır. Söz konusu modeller, ince ayar (fine-tuning), yönerge temelli öğrenme (instruction tuning) ve iç-bağlamda öğrenme (in-context learning) gibi gelişmiş stratejilerle özelleştirilebilmekte ve kullanıcıya özgü çıktılar üretebilmektedir (Ding vd., 2024).

Makine öğrenmesi uygulamalarında sıklıkla karşılaşılan sınıf dengesizliği problemi, azınlık sınıfına ait örneklerin yetersizliği nedeniyle sınıflandırma modellerinin çoğunluk sınıfı lehine öğrenme eğiliminde olması şeklinde tanımlanmaktadır. Bu durum, özellikle hukuk, tıp ve siber güvenlik gibi yüksek hata maliyetine sahip alanlarda, azınlık sınıfına ait kritik bilgilerin göz ardı edilmesine neden olmaktadır (He & Garcia, 2009). Sınıf

dengesizliğini gidermek amacıyla geleneksel son dönemde büyük dil modelleri kullanılarak gerçekleştirilen sentetik örnek üretimi, azınlık sınıfının temsilini güçlendirmeye yönelik yenilikçi bir yaklaşım olarak öne çıkmaktadır. LLM tabanlı veri artırımı, yalnızca örnek sayısını artırmakla kalmamakta; aynı zamanda azınlık sınıfa ait anlamsal çeşitliliği genişleterek modelin karar sınırlarını daha isabetli biçimde öğrenmesine katkı sağlamaktadır. Ding vd. (2024) tarafından gerçekleştirilen kapsamlı bir derlemede, büyük dil modelleri aracılığıyla üretilen sentetik verilerin, yönerge ince ayarı, bağlam içi örnekleme ve hizalama süreçleriyle bütünleştirilerek sınıf dengesizliği sorununa anlamlı çözümler sunduğu ortaya konmuştur.

Sağlık ve kamu politikası gibi özel uygulama alanlarına yönelik çalışmalarda da benzer iyileşmeler sağlandığı bildirilmiştir. Sharma ve Feldman (2023), doktor-hasta diyaloglarını konu başlığına göre sınıflandıran bir görevde GPT-3.5-turbo kullanarak sentetik diyalog-etiket çiftleri üretmiş; az örnekli etiketlerde doğruluk oranını %30'a varan oranda artırmıştır. Öte yandan, büyük dil modeli ChatGPT'nin farklı granülerlik düzeylerinde (örneğin, küresel ve yerel) sentetik veriler üretebildiği ampirik olarak gösterilmiştir (Busker vd., 2025). Bu sayede, politika analizi gibi veri erişiminin kısıtlı olduğu alanlarda, gerçek veriye ihtiyaç duyulmadan temsili veri setlerinin oluşturulabileceği ortaya konmuştur. Valizadeh-Aslani vd. (2024), önyargılı sınıf dağılımı içeren metin sınıflandırma görevleri için iki aşamalı bir ince ayar stratejisi geliştirmiş; ilk aşamada büyük dil modeli ile her sınıf için eşit sayıda örnek türetilmiş, ikinci aşamada ise standart ince ayar uygulanmıştır. Bu yaklaşım sayesinde kuyruk (tail) sınıflarındaki performans kaybı azaltılmış ve geleneksel yeniden örnekleme tekniklerine kıyasla anlamlı iyileştirmeler elde edilmiştir.

Yapılan bu çalışmalar incelendiğinde LLM ile sentetik veri üretimi, geleneksel veri dengeleme yöntemlerine bir alternatiftir. Yapılan tez çalışmasında OpenAI tarafından geliştirilen ChatGPT'nin GPT-3.5-turbo modeli klasik veri dengeleme yöntemlerine bir alternatif olarak kullanılmıştır.

Bu yaklaşımda temel amaç, azınlık sınıfa ait örneklerin sayısını artırmak suretiyle sınıf dağılımındaki dengesizliği gidermek ve sınıflandırma modellerinin azınlık sınıfa ilişkin öğrenme başarımını artırmaktır. ChatGPT gibi büyük dil modelleri, sahip oldukları

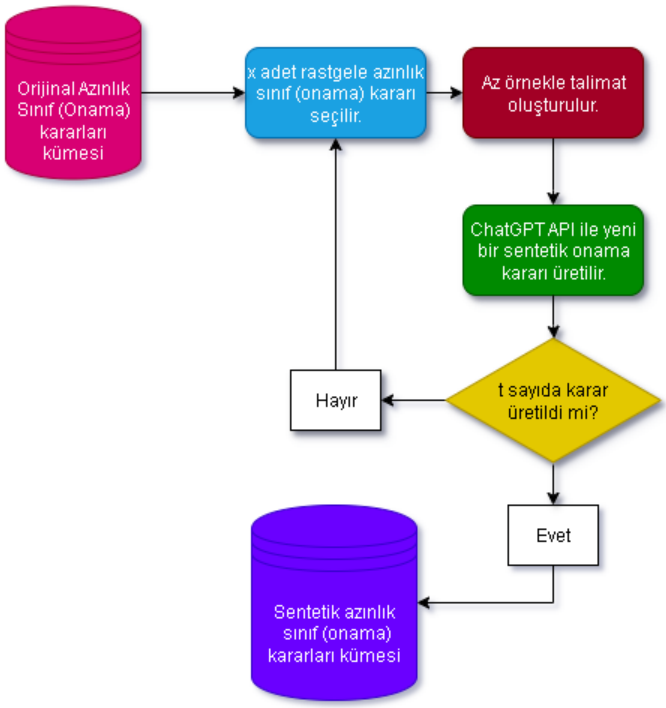
bağlamsal dil modelleme yetenekleri sayesinde, hedef sınıfa ait örüntüleri anlamlı biçimde taklit edebilen ve dilbilgisel olarak tutarlı, içerik açısından geçerli sentetik metinler üretebilen sistemlerdir. Bu çerçevede, veri dengeleme amacıyla gerçekleştirilen üretim süreci, dikkatle yapılandırılmış istemler (prompt'lar) aracılığıyla modelin yönlendirilmesi esasına dayanmaktadır.

Algoritmanın ilk adımı, üretilecek metinlerin bağlamsal özelliklerini öğrenebilmesi için modelin örnek alacağı temsili bir azınlık sınıfı örneklem kümesinin tanımlanmasıdır. Bu örneklem, hedef sınıfa ait karar metinlerinden rastgele ya da bilinçli bir örnekleme stratejisiyle seçilen temsil edici metin parçalarını içermektedir. İkinci adımda, seçilen bu örnekler modele giriş olarak verilecek bir istem (prompt) biçiminde yapılandırılır. Bu istemde genellikle birkaç örnek karar metni numaralandırılmış biçimde verilmekte ve modelden benzer yapıda yeni bir karar metni üretmesi beklenmektedir. Böylece, az örnekle öğrenme (few-shot learning) ilkesi çerçevesinde modelin üretken kapasitesi belirli bir şablon doğrultusunda yönlendirilmektedir.

Üçüncü adımda, metin üretimi süreci, parametrik olarak denetlenen bir model çağrısıyla yürütülür. Üretilecek metinlerin çeşitliliği ve gerçekçiliği; sıcaklık (temperature), çekirdek örnekleme oranı (top-p), tekrar cezası (frequency penalty) ve varlık cezası (presence penalty) gibi hiperparametreler aracılığıyla kontrol altına alınır. Parametre ayarlaması sayesinde modelin hem tekrarlayıcı içerikler üretmesi önlenmekte, hem de önceki örneklerden bağımsız ancak biçimsel olarak benzer yapılar oluşturması sağlanmaktadır. Her bir üretim döngüsünde, modele farklı örneklerin sunulması yoluyla çeşitli ve özgün sentetik karar metinleri elde edilir. Bu yapı hem anlamsal çeşitliliği artırmakta hem de yapay örneklerin birbirine fazla benzemesini engellemektedir.

Son adımda, üretilen metinler azınlık sınıfına ait olarak etiketlenir ve mevcut veri kümesiyle birleştirilir. Böylece sınıflar arası dağılım daha dengeli hâle getirilir ve makine öğrenmesi algoritmaları her sınıf için daha temsili öğrenme gerçekleştirebilir. ChatGPT ile yürütülen bu üretim süreci, yalnızca istatistiksel olarak dengeli bir veri seti oluşturmakla kalmamakta, aynı zamanda yüksek nitelikli ve bağlamsal bütünlüğe sahip örneklerin dâhil edilmesiyle modelin genelleme yeteneğini de artırmaktadır. Algoritmanın akış şeması Şekil 3.5'de gösterilmiştir. Şekil 3.5'de t adet rastgele

mahkeme kararının her seferinde x adet azınlık sınıfa ait karar seçilerek büyük dil modeline öğretilmesinin akış şeması gösterilmiştir.



Şekil 3.5. LLM ile Sentetik Mahkeme Kararları Üretimi Akış Şeması

Şekil 3.6'da ise ChatGPT tarafından üretilmiş örnek bir sentetik kararın ekran görüntüsü verilmiştir.

Yargıtay 1. Ceza Dairesi 2015/18-23 E., 2015/142 K.

KASTEN YARALAMA SUÇU

5237 SAYILI TÜRK CEZA KANUNU [Madde 86]

"İçtihat Metni"

Davalı hakkında kasten yaralama suçundan verilen kararın usul ve esasa ilişkin yönlerinin incelenmesi neticesinde, mahkeme kararının yerinde olduğu sonucuna varılmıştır. Bu nedenle, yerel mahkemenin verdiği karar onanmış ve temyiz talebi reddedilmiştir.

Şekil 3.6. ChatGPT ile üretilen bir sentetik onama kararının ekran görüntüsü

Üretilen metin, daire, esas, karar numaralarının belirtilmesi, suçun tanımlanması, ilgili kanun maddesinin açıkça gösterilmesi ve "onama/temyiz talebinin reddi" hükmünün net biçimde sunulması gibi temel öğeleriyle Yargıtay kararlarının resmî şablonuna uymaktadır. "Usul ve esasa ilişkin yönlerin incelenmesi" kalıbı ile kullanılan resmî dil; gerçek karar metinlerinde rastlanan terminolojiyi ve anlatım tonunu inandırıcı şekilde yansıtmaktadır.

3.2.9. Makine Öğrenmesi ve Derin Öğrenme Modelleri

Sınıflandırma aşamasında K-En Yakın Komşuluk, Karar Ağacı, Rassal Orman, Destek Vektör Makineleri, XGBoost, LightGBM ve Yapay Sinir Ağları gibi klasik algoritmaların yanı sıra, uzun ve karmaşık cümle yapılarındaki ardışık bağıntıları yakalamak üzere Uzun Kısa Süreli Bellek (Long Short Term Memory, LSTM) tekniği kullanılmıştır.

3.2.10. K-En Yakın Komşuluk Yöntemi

K-En Yakın Komşuluk (K-Nearest Neighbor, KNN) sınıflandırma algoritması, etiketli eğitim veri setindeki her bir yeni örneğin, en yakın komşu olarak belirlenen k adet eğitim örneğinin çoğunluk sınıfına atanması ilkesine dayanır. Parametreye dayalı varsayım olmayan (non-parametrik) bir sınıflandırıcıdır. KNN algoritması günümüzde basitliği ve etkinliği nedeniyle pek çok uygulama alanında tercih edilmiştir. KNN algoritmasının eğitim aşamasında yalnızca veri depolama işlemi gerçekleştirilmiş olup, model üzerinde ekstra bir genelleme adımı uygulanmamıştır; sınıflandırma, yalnızca test örneğinin eğitim kümesindeki örneklere olan uzaklıklarına bakılarak yapılmaktadır.

Öklid uzaklığı (Euclidean distance), KNN algoritmasında en sık tercih edilen mesafe ölçütüdür. İki n boyutlu vektör arasındaki uzaklık aşağıdaki formülle hesaplanmaktadır:

$$d(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \tag{3.7}$$

Burada $x = (x_1, x_2, \dots, x_n)$, test örneğini, $y = (y_1, y_2, \dots, y_n)$, eğitim kümesindeki bir örneği, n özellik (boyut) sayısını, $d(x, y)$ ise bu iki örnek arasındaki Öklid uzaklığını temsil eder. KNN algoritmasının sınıflandırma başarımı, komşu sayısını belirleyen k parametresinin uygun şekilde seçilmesine doğrudan bağlıdır (Fix ve Hodges, 1951; Taunk vd., 2019)

Küçük k değerleri, modelin eğitim verisindeki gürültüye aşırı duyarlılık göstermesine yol açarken, büyük k değerleri genelleme yetisini artırmakta; ancak bu durum, karar sınırlarının belirginliğini azaltarak sınıflar arasındaki ayrımı güçleştirebilmektedir. Ayrıca, her test örneği için tüm eğitim örnekleriyle mesafe hesaplaması yapılması gerektiğinden, algoritmanın hesaplama maliyeti oldukça yüksektir (Taunk vd., 2019).

3.2.11. Karar Ağacı

Karar ağacı (Decision Tree, DT) algoritması, makine öğrenmesi alanında hem sınıflandırma hem de regresyon problemleri için yaygın biçimde kullanılan yöntemlerden biridir. Karar ağaçlarının temel ilkeleri, 1960'lı yıllardan itibaren veri odaklı sınıflandırma yöntemleriyle çeşitli biçimlerde ortaya konmuş ve uygulanmıştır (Morgan ve Sonquist, 1963). Bu yöntemler, 1984 yılında yayımlanan Sınıflandırma ve Regresyon Ağaçları (Classification and Regression Trees, CART) başlıklı monografiyle birlikte sistematik bir yapıya kavuşturulmuştur (Breiman vd., 1984). Devamında, doğruluk ve genelleme kapasitesini artırmak amacıyla çeşitli iyileştirmelere gidilmiş; özellikle C4.5 gibi algoritmaların önerilmesiyle yönteme önemli katkılar sağlanmıştır (Quinlan, 1993). Zamanla karar ağaçları, karmaşık karar süreçlerini hiyerarşik alt kararlara bölerek yorumlanabilir "if-then" kuralları üretme yetenekleri sayesinde hem akademik çalışmalarda hem de çeşitli uygulamalarda yaygın biçimde kullanılmaya başlanmıştır.

Algoritma düzeyinde, ağaç inşası Gini indeksi, bilgi kazancı ve kazanç oranı gibi ölçütlerle yinelenen veri ayrımları yoluyla gerçekleştirilmekte; aşırı uyumun (overfitting) önüne geçmek için ön / son budama stratejileri uygulanmaktadır.

Algoritma, her adımdaki en iyi ayırıcı özniteliği seçip kök düğüm olarak yerleştirir, ardından veri kümesini bu özelliğin değerlerine göre dallara ayırarak alt kümeler meydana getirir. Her bölünme işleminde kümenin saflaşması, yani yalnızca aynı sınıftan

örnekler bulunan kümelerin oluşturulması hedeflenir. Bu yinelemeli süreç, alt kümeler tümüyle homojen hâle gelince ya da başka bölünme olmayana kadar devam eder; son aşamada yaprak düğümler oluşturulur ve ilgili sınıf etiketleri atanır. Öznitelik seçiminde yaygın olarak kullanılan safsızlık ölçütleri entropi ile gini indeksidir. Entropi veri setindeki düzensizliği, gini indeksi ise heterojenliği ölçer. Entropi formülünde bu oranların logaritmalarıyla çarpımları toplanarak belirsizlik ölçülürken, Gini safsızlığında bu oranların kareleri toplanarak saf kümeden sapma hesaplanır. Her iki ölçüt de veri kümesindeki sınıf çeşitliliğini değerlendirerek homojen alt kümeler oluşturmayı hedefler (Fürnkranz, 2011). Her iki ölçütün matematiksel formülleri sırasıyla Denklem 3.8 ve Denklem 3.9'da gösterilmiştir.

$$Entropi(S) = -\sum_{i=1}^{c} \frac{|S_i|}{|S|} log\left(\frac{|S_i|}{|S|}\right) \tag{3.8}$$

$$Gini(S) = 1 - \sum_{i=1}^{c} \left(\frac{|S_i|}{|S|}\right)^2 \tag{3.9}$$

Burada S, üzerinde işlem yapılan örnek kümesini; $|S|$ bu kümedeki toplam örnek sayısını; c, sınıf sayısını; S_i, i. sınıfa ait örneklerin yer aldığı alt küme yerine geçen i. düğümü ve $|S_i|$ bu düğümdeki örnek sayısını ifade eder. $|S_i|$'nin $|S|$'e oranı ise i. sınıfa ait örneklerin tüm küme içindeki payını temsil eder.

Entropi ve Gini safsızlık ölçütlerinin yanı sıra, öznitelik seçimi aşamasında veri kümesini ne ölçüde daha homojen alt kümelere ayırdığını nicel olarak belirlemek için Bilgi Kazancı (Information Gain, IG) ölçütü de yaygın biçimde kullanılmaktadır. Bilgi kazancının formülü Denklem 3.10'da gösterilmiştir:

$$IG(S, A) = Entropi(S) - \sum_{i=1}^{m} \frac{|S_\mathrm{i}|}{|S|} \cdot Entropi(S_\mathrm{i}) \tag{3.10}$$

Burada $IG(S, A)$, öznitelik A kullanılarak küme S'nin bölünmesiyle elde edilen bilgi kazancını gösterir. $Entropi(S)$ ve $Entropi(S_\mathrm{i})$, sırasıyla tam kümenin ve A'nın i. değerine sahip örneklerden oluşan alt küme S_i'nin (toplam m farklı değer) entropisidir. $|S|$ veri kümesinin tamamındaki, $|S_\mathrm{i}|$ ise her alt kümedeki örnek sayısını ifade eder. Yani

$\frac{|S_i|}{|S|}$ bölme işlemi ilgili alt kümenin toplam küme içindeki payını temsil eder; bu paylarla ağırlıklandırılıp toplandığında özniteliğin bölme sonrası beklenen belirsizliği elde edilir. Başlangıç belirsizliği $Entropi(S)$'den bu değer çıkarılarak, özniteliğin ne kadar kazanç sağladığı hesaplanır: sonuç pozitifse, A veri kümesini daha homojen alt parçalara ayırmış demektir (Suthaharan, 2016).

Karar ağaçları parametrik olmayan yapıları sayesinde sayısal ve kategorik öznitelikleri birlikte işleyebilmekte, eksik verilerle başa çıkabilmekte ve ürettiği kuralların alan uzmanlarınca kolaylıkla doğrulanabilmesi nedeniyle sağlık bilişimi, hukuk analitiği ve finans gibi alanlarda yaygın olarak kullanılmaktadır.

3.2.12. Rassal Orman

Rassal Orman (Random Forest, RF) yöntemi, karar ağaçlarının rastgele örneklenmiş veri altkümeleri (bootstrap) ve rastgele seçilmiş öznitelik altkümeleri üzerinde eğitilerek oluşturduğu çok sayıda bağımsız ağaçtan meydana gelen bir topluluk (ensemble) sınıflandırıcısıdır. Her ağaç, test örneği için bir oy kullanır; nihai sınıf etiketi çoğunluk oyu ile belirlenir. Söz konusu yapı, tekil ağaçlardaki yüksek varyansı düşürürken önyargıyı sınırlı tutarak güçlü bir genelleme yeteneği ortaya koyar (Liu vd., 2012).

Rassal orman yaklaşımında, her biri aynı dağılımdan bağımsız olarak seçilen rastgele tahmincilerden ($\{h(x, \Theta_k), k = 1, \ldots\}$) oluşmaktadır. Burada x giriş vektörünü, $h(x, \Theta_k)$ k'ıncı karar ağacı tarafından üretilen çıktıyı simgelemektedir. Model K sayısı kadar çalıştırılarak bir sınıflandırıcı dizisi üretilmesinin ardından nihai sınıf etiketi, her bir aday Y için indikatör fonksiyonu değerlerinin toplamına bakılarak tayin edilir. Denklem 3.11'de tüm ağaçların ortak kararı formüle edilmiştir:

$$H(x) = arg\ \max_{Y} \sum_{i=1}^{K} I(h_i(x) = Y) \tag{3.11}$$

Burada $h_i(x)$ rassal ormandaki i'inci ağacın tahminidir. $I(\cdot)$ indikatör fonksiyonu, Y ise çıktı değişkenini ifade etmektedir.

RF'nin başlıca üstünlükleri literatürde yüksek sınıflandırma doğruluğu, gürültüye karşı tolerans ve aşırı uyuma (overfitting) direnç olarak rapor edilmiştir (Parmar vd., 2019). Model, sürekli ve kategorik değişkenleri eşzamanlı işleyebilmekte; her ağaçtaki bölme kriterlerine dayanarak değişken önem skorları üreterek yorumlanabilirliği artırmaktadır. Ayrıca, yöntemin parametre ayarlarına karşı duyarlılığı görece düşüktür; genellikle yeterli sayıda ağaç kullanılması ve öznitelik sayısının karekköküne dayalı seçim kuralının uygulanması, tatmin edici bir sınıflandırma performansı sağlamaktadır (Liu vd., 2012).

Geliştirilmesinden bu yana RF; biyoenformatik, tıbbi tanı, uzaktan algılama, web sayfası sınıflandırması ve müşteri kaybı öngörüsü gibi çok çeşitli alanlarda, uyarlanabilir artırma (Adaptive Boosting, AdaBoost) ve SVM gibi güçlü yöntemlerle rekabet eden, hatta çoğu durumda onları geride bırakan sonuçlar vermiştir (Parmar vd., 2019). Böylece RF, hem teorik temelleri güçlü hem de pratikte esnek bir topluluk öğrenme (ensemble learning) yöntemi olarak makine öğrenmesine ilişkin deneysel çalışmalarda tercih edilebilecek bir araç konumuna gelmiştir.

3.2.13. Destek Vektör Makineleri

Destek Vektör Makineleri doğrusal ve doğrusal olmayan ayırma sınırlarını en geniş aralık (margin) ilkesine dayanarak öğrenen, denetimli bir makine öğrenmesi yaklaşımıdır. Özellikle yüksek boyutlu veride sağlam genelleme yeteneği sunması, onu metin madenciliğinden biyoinformatiğe kadar pek çok alanda başvurulan bir yöntem hâline getirmiştir (Cervantes vd., 2020). Son yıllarda gerçekleştirilen kapsamlı incelemeler, SVM'nin ikili sınıflandırma dışında çoklu sınıf, büyük veri ve dengesiz veri kümeleri gibi güncel zorluklara uyarlanması üzerinde yoğunlaştığını göstermektedir (Cervantes vd., 2020).

Kuramsal temelde SVM, eğitim örneklerini en az hata ile ayıran optimal hiperdüzlemi belirlemek üzere ikinci dereceden kısıtlı optimizasyon problemini çözer; hiperdüzleme en yakın örnekler "destek vektörleri" olarak adlandırılır ve karar sınırını bütünüyle tanımlar (Hearst, 1998). Cezalandırma parametresi ve gevşeklik değişkenleri aracılığıyla tanımlanan "yumuşak aralık" formülasyonu, doğrusal olarak tam ayrılmayan gerçek-

dünya verilerinde hatayı denetleyerek modelin yapısını esnekleştirir (Cervantes vd., 2020).

Doğrusal olmayan ayrım problemlerinde ise çekirdek hilesi (kernel trick) kullanılarak örnekler daha yüksek boyutlu bir öznitelik uzayına gömülür; burada doğrusal bir hiper-düzlem aslında girdiler uzayında eğrisel bir sınır üretir (Suthaharan, 2016). Yalnızca çekirdek fonksiyonu hesaplanarak bu uygulama gerçekleştirildiğinden, çok boyutlu dönüştürmelerin doğrudan hesaplanması gereğini ortadan kaldırır (Hearst, 1998).

Destek Vektör Makinelerinde doğrusal olmayan sınıflandırma problemleri için kullanılan temel tekniklerden biri, çekirdek fonksiyonları (kernel functions, KF) aracılığıyla veri noktalarının daha yüksek boyutlu öznitelik uzaylarına (feature spaces, FS) gömülmesidir. Bu yöntem, giriş uzayında (input space) doğrusal olmayan bir problemi, dönüştürülmüş uzayda doğrusal olarak ayırabilir kılmak üzere tasarlanmıştır. Çekirdek fonksiyonları, söz konusu dönüştürmenin açık biçimde hesaplanmasına gerek kalmaksızın iç çarpım (inner product) benzerliklerini tanımlayarak, öğrenme algoritmalarının yalnızca çekirdek matrisine (kernel matrix) dayanmasını mümkün kılar.

Literatürde sıklıkla kullanılan çekirdek fonksiyonları arasında doğrusal çekirdek (linear kernel), polinom çekirdek (polynomial kernel), Gauss tipi radyal tabanlı fonksiyon (radial basis function, RBF kernel) ve sigmoid çekirdek (sigmoid kernel) yer almaktadır. Bu çekirdeklerin her biri, farklı yapısal varsayımlara ve karar sınırı biçimlerine karşılık gelmektedir. Örneğin RBF çekirdeği, sınırlı etki alanına sahip lokal karar sınırları üretirken; polinom çekirdeği daha küresel eğrileri yansıtır. Bu nedenle çekirdek seçimi, hem model kapasitesi hem de genelleme performansı üzerinde doğrudan etkilidir. Schölkopf ve Smola, 2002).

SVM'nin temel üstünlükleri, iyi tanımlı konveks optimizasyon sayesinde global minimum garanti etmesi, az sayıda destek vektörü sayesinde bellek gereksinimini sınırlı tutması ve yüksek boyutlu uzaylarda dahi aşırı uyumu (overfitting) sınırlandıran Yapısal Risk Minimizasyonu ilkesine dayanmasıdır (Cervantes et al., 2020). Bununla birlikte, çoklu sınıf genişletmeleri için ikili parçalara ayırma stratejilerinin karmaşıklığı, büyük veri kümelerinde eğitim süresinin uzunluğu ve sınıf dağılımı dengesiz veri setlerinde sınır

yüzeyinin çoğunluk sınıfına kayma eğilimi önemli sınırlılıklar olarak rapor edilmiştir (Cervantes vd., 2020).

Uygulamada SVM; metin kategorilendirme, yüz tanıma, protein katlanma tahmini, kredi kartı sahteciliği tespiti gibi heterojen problemlerde yüksek doğruluk sağlamış; özellikle öznitelik uzayı boyutunun örnek sayısından büyük olduğu durumlarda rakip yöntemlere kıyasla üstün performans göstermiştir (Cervantes vd., 2020).

3.2.14. XGBoost

Aşırı Gradyan Artırma (Extreme Gradient Boosting, XGBoost), 2016 yılında Tianqi Chen tarafından geliştirilmiş ve Carlos Guestrin ile birlikte yaygınlaştırılmış, açık kaynaklı bir gradyan artırma kütüphanesidir. XGBoost, Gradyan Artırma Karar Ağaçları (Gradient Boosting Decision Tree, GBDT) çerçevesini ikinci mertebe türev bilgisi, $\ell 2$-düzenlileştirme ($\ell 2$ regularization) ve yaprak başına sabit ceza (γ) ile genişleterek aşırı öğrenmeyi (overfitting) baskılayan düzenlileştirilmiş bir hedef fonksiyonu tanımlar (Chen ve Guestrin, 2016; XGBoost Geliştiricileri, 2025). Bu formülasyon, her yinelemede eklenen zayıf öğrenicinin yalnızca artımsal kaybını değil, aynı zamanda model karmaşıklığını da optimize eder. Böylece XGBoost, $\ell 1/\ell 2$ düzenlileştirmesi olmayan klasik GBDT'ye kıyasla daha iyi genelleşme sergilemekte ve hiperparametre aralığını geniş bir problem kümesi için duyarlı kılmaktadır (XGBoost Geliştiricileri, 2025).

Model, Newton benzeri ikinci mertebeden genişletmeleri, veri ağırlıklarıyla (Hessian) ölçeklenmiş gradyan istatistiklerini ve yaprak düzeyinde kapalı form çözümleri kullanarak verimli bir ayrık optimizasyon problemine indirgemektedir. Bu ayrık optimizasyon sürecinde, aday bölme noktaları, örnek ağırlıklarını dikkate alan ağırlıklı yüzdelik kestirimi (weighted quantile sketch) algoritması aracılığıyla yaklaşık yüzdelik değerlere göre seçilmektedir. Söz konusu yöntem, özellikle büyük ve seyrek veri kümelerinde örnek ağırlıklarını gözeterek bellek kullanımı ve hesaplama karmaşıklığını doğrusal düzeye indirme avantajı sağlamaktadır. (Chen ve Guestrin, 2016).

Her bir ağacın herhangi bir yaprağı için en uygun ağırlık, Denklem 3.12'de sunulan formülle hesaplanır.

$$w_j^* = -\frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i + \lambda} \tag{3.12}$$

Burada w_j^* ifadesi, j'inci yaprağa ait optimal skor değerini temsil etmektedir. I_j, söz konusu yaprağa düşen gözlem indekslerini, g_i ve h_i sırasıyla i'inci gözleme ilişkin kayıp fonksiyonunun birinci ve ikinci türevini ifade etmekte olup, bu türevler modelin optimizasyon sürecinde yön ve eğim bilgisini belirlemede kullanılır. Regülasyon parametresi olarak formüle dâhil edilen λ ise, modelin aşırı öğrenmesini önlemek amacıyla $\ell 2$ normuna dayalı cezalandırmayı sağlamaktadır.

XGBoost algoritmasında, her bir yaprak düğümüne atanacak optimal ağırlık değeri, modelin ikinci mertebe optimizasyon yaklaşımı doğrultusunda belirlenmektedir. Bu yapı sayesinde XGBoost, her yaprak için gradyan ve eğrilik bilgilerini dikkate alarak kapalı formda hesaplanan ağırlıklarla modelin genelleme performansını artırmayı hedeflemektedir (Chen ve Guestrin, 2016).

Ek olarak, XGBoost; öğrenme oranı küçültmesi (learning rate shrinkage – η), satır ve sütun alt örneklemesi (row and column subsampling), eksik değerler için varsayılan yön stratejisi (default direction strategy for missing values) ve Grafik İşlem Birimi/disk dışı eğitim (GPU/out-of-core training) gibi uygulama düzeyi iyileştirmeler içerir. Bu sayede hem tek makinede hem de dağıtık ortamlarda, milyarlarca örnek içeren veri kümeleri üzerinde saniyeler mertebesinde tekrar süreleri sağlanmaktadır (Chen ve Guestrin, 2016).

Düzenlileştirme terimleri ile desteklenen kapalı formül, yaprak ağırlıklarının doğrudan hesaplanmasına olanak tanıyarak ağaç büyütme sürecini yığın belleği dostu ve hesaplama açısından doğrusal (lineer) tutar. Bu öznitelikler sınıf dengesizliği barındıran metin sınıflandırma problemleri de dâhil olmak üzere çok sayıda makine öğrenmesi görevinde XGBoost'u hem teorik olarak sağlam hem de pratikte verimli olabilecek bir seçenek hâline getirmektedir (Chen ve Guestrin, 2016).

3.2.15. LightGBM

LightGBM, Microsoft araştırmacıları tarafından geliştirilen, histogram tabanlı ve yaprak odaklı (leaf-wise) bir gradyan artırmalı karar ağacı (Gradient boosting decision tree, GBDT) algoritmasıdır. Gradyana dayalı tek taraflı örnekleme (Gradient-based one-side sampling, GOSS) ve özel öznitelik paketleme (Exclusive Feature Bundling, EFB) gibi yenilikçi teknikler sayesinde, yüksek boyutlu ve büyük hacimli veri kümelerinde XGBoost gibi popüler çerçevelere kıyasla daha hızlı ve daha az bellek kullanan bir alternatif sunmaktadır (Ke vd., 2017).

LightGBM'in resmî dokümantasyonunda, model karmaşıklığı, aşırı öğrenme (overfitting) kontrolü ve eğitim süresi üzerinde doğrudan etkili üç temel hiperparametre tanımlanmıştır: yaprak sayısı (num_leaves), yapraktaki minimum gözlem sayısı (min_data_in_leaf) ve öznitelik alt örnekleme oranı (feature_fraction). Ayrıca, hem merkezi işlem birimi (CPU) hem de grafik işlem birimi (GPU) mimarilerinde eğitim sürecini hızlandırmaya yönelik çok sayıda yapılandırma seçeneği sunulmaktadır.

LightGBM'de bir karar ağacındaki aday bölünmelerin kalitesi, düğüm içi toplam gradyan varyansındaki azalmaya dayalı bilgi kazancı ile değerlendirilmektedir. Bu bilgi kazancı, aşağıda Denklem 3.13'de gösterilen varyans temelli formülle hesaplanır:

$$V_{j|O}(d) = \frac{1}{n_O}\left(\frac{\left(\sum_{\{x_i \in O : x_{ij} \le d\}} g_i\right)^2}{n_{jl|O}\,(d)} + \frac{\left(\sum_{\{x_i \in O : x_{ij} > d\}} g_i\right)^2}{n_{jr|O}\,(d)}\right) \tag{3.13}$$

Burada O mevcut düğümdeki gözlem kümesini, g_i ilgili gözlemin negatif gradyanını, n_O toplam örnek sayısını, $n_{jl|O}\,(d)$ ve $n_{jr|O}\,(d)$ sırasıyla sol ve sağ alt düğümlerdeki örnek sayılarını temsil etmektedir. Diğer aday kesme noktaları için $V_{j|O}(d)$ hesaplanarak en yüksek kazançlı bölünme seçilir; böylece model, gradyan varyansını en büyük ölçüde azaltacak ayrımı gerçekleştirir (Ke vd., 2017).

Özetle LightGBM histogram binleme yöntemiyle sürekli öznitelikleri daha az sayıda aralığa indirger ve yaprak odaklı büyüme stratejisiyle, her adımda kaybı en çok azaltan

dalı genişleterek klasik derinlik odaklı GBDT algoritmalarına kıyasla daha verimli bir eğitim süreci sunar. GOSS ve EFB gibi yöntemler sayesinde, büyük boyutlu veri kümelerinde hem işlem süresi kısalmakta hem de doğruluk seviyesi korunmaktadır.

Modelin ifade gücü ve genelleme yeteneği, çeşitli hiperparametrelerle dengelenmektedir. Yaprak sayısını belirleyen parametre, modelin karar sınırlarını şekillendirirken; her yaprakta bulunması gereken asgarî örnek sayısını tanımlayan ayar, aşırı öğrenmeyi önlemeye yardımcı olur. Eğitim sırasında rastgele seçilecek öznitelik oranını belirleyen parametre ise hem öğrenme süresini kısaltır hem de değişkenler arası korelasyondan kaynaklanan sapmaları azaltır (Microsoft Corporation, 2025).

Sonuç olarak LightGBM; düşük bellek tüketimi, hızlı eğitim süresi ve güçlü genelleme kabiliyeti ile özellikle dengesiz ve karmaşık yapıya sahip veri kümeleri için etkili bir sınıflandırma algoritması olarak öne çıkmaktadır.

3.2.16. Yapay Sinir Ağları

Yapay Sinir Ağları (Artificial Neural Networks – ANN), biyolojik sinir sisteminden esinlenilerek geliştirilmiş, çok katmanlı ve veri odaklı hesaplama modelleridir. Bu ağlar, bilgi işleme görevini üstlenen yapay nöronlardan oluşur ve sinyallerin katmanlar arasında taşınması, ağırlıklı bağlantılar aracılığıyla gerçekleştirilir. Her bir yapay nöron, kendisine gelen sayısal girdileri belirli ağırlıklarla çarpar, bu değerleri toplar ve ardından bir aktivasyon fonksiyonu uygulayarak bir çıktı üretir. Nöronlar katmanlar hâlinde düzenlenmiş olup, bu katmanlar arasında ileriye dönük bilgi akışı sağlanır. Öğrenme süreci boyunca bağlantı ağırlıkları, modelin çıktılarındaki hataları en aza indirecEk Şekilde sürekli olarak güncellenir (Abraham, 2005).

Yapay sinir ağları, özellikle 1980'li yıllarda çok katmanlı geri yayılım (backpropagation) algoritmasının geliştirilmesiyle önemli bir ilerleme kaydetmiş ve bu tarihten itibaren yaygın biçimde kullanılmaya başlanmıştır (Abraham, 2005; Yang ve Wang, 2020). Katmanlı mimarisi sayesinde ağ, girdilerden başlayarak her bir katmanda veri temsillerini giderek daha soyut ve yüksek düzeyli biçimlere dönüştürür. Bu hiyerarşik temsil yeteneği, karmaşık örüntülerin öğrenilmesini ve yüksek doğrulukla modellenmesini mümkün kılmaktadır. Bu nedenle ANN, günümüzde bilgisayarla görme, doğal dil işleme,

ses tanıma ve sinirbilimsel veri analizi gibi birçok disiplinde çağdaş makine öğreniminin temel araçlarından biri olarak konumlanmıştır.

Bu bağlamda, tek bir yapay nöronun temel işlevi, giriş sinyallerini ağırlıkları ile çarparak toplamak ve bu toplama bir aktivasyon fonksiyonu uygulayarak bir çıktı üretmektir. Temel işlem Denklem 3.14'de formüle edilmiştir:

$$O = f(\text{net}) = f\left(\sum_{j=1}^{n} w_j x_j\right) \tag{3.14}$$

Denklem 3.14, bir yapay nöronun hesap birimini tanımlamaktadır. Burada x_j nörona gelen j girdi değerini, w_j, j. girdiye karşılık gelen öğrenilebilir sinaptik ağırlığı ifade eder. Toplam girdi sayısı n ile gösterilmiştir. $\sum_{j=1}^{n} w_j x_j$, tüm girdilerin kendi ağırlıklarıyla çarpılıp toplanmasıyla elde edilen "net" uyarımı tanımlar. $f(\cdot)$ genellikle sigmoid, ReLU gibi doğrusal olmayan ve türevlenebilir bir aktivasyon fonksiyonunu ifade eder. O aktivasyon fonksiyonunun uygulanması sonucu hesaplanan nöron çıktısını temsil eder. Bu noktada net ifadesiyle nöron gövdesine iletilen birleşik uyarım düzeyi gösterilmiştir. Söz konusu toplam ise Denklem 3.15'de gösterilen toplam çarpımı eşitliği ile tanımlanır:

$$\text{net} = w^{\top} x = w_1 x_1 + w_2 x_2 + \cdots + w_n x_n \tag{3.15}$$

Burada x_j girdileri, w_j öğrenme sürecinde güncellenen bağlantı ağırlıklarını temsil eder (Yang ve Wang, 2020).

Yapay sinir ağlarında öğrenme, temel olarak ağın çıktısıyla beklenen (hedef) çıktılar arasındaki farkın minimize edilmesi süreci olarak tanımlanır. Bu fark genellikle bir kayıp (loss) fonksiyonu ile sayısallaştırılır ve ağın ağırlıkları, bu kaybı azaltacak şekilde güncellenir. Öğrenme süreci, denetimli öğrenme çerçevesinde, örnek-girdi çiftleri kullanılarak gerçekleştirilir. Her eğitim örneği için ağın mevcut ağırlıklarıyla bir ileri geçiş (forward pass) gerçekleştirilir ve elde edilen çıktı ile hedef değer arasındaki hata hesaplanır. Çok katmanlı yapılarda sadece çıkış katmanında değil, aynı zamanda gizli katmanlardaki ağırlıkların da güncellenmesi gerektiğinden, ağ boyunca hata geriye doğru

yayılır. Bu algoritma geri yayılım (backpropagation) olarak adlandırılır (Abraham, 2005). Geri yayılım algoritması, zincir kuralına dayalı olarak ağın katmanlarında yer alan tüm ağırlıklar için gradyanların hesaplanmasını sağlar. Öncelikle hata fonksiyonunun, çıkış katmanındaki nöronların çıktılarıyla olan kısmi türevleri alınır; ardından bu hata, her bir katman boyunca geriye doğru iletilerek, önce çıkış katmanına bağlı gizli katmanların, daha sonra girişe yakın katmanların ağırlıklarına ilişkin gradyanlar elde edilir. Bu gradyanlar, ağırlık güncellemelerinde kullanılır. Ağırlık güncellemesi Denklem 3.16'da formüle edilmiştir:

$$\Delta w_{ij} = -\eta \frac{\partial E}{\partial w_{ij}} \tag{3.16}$$

Burada Δw_{ij} ağırlık güncellemesini, η öğrenme oranını, E ise toplam hata fonksiyonunu temsil eder. Ağırlıklar, i yönünde güncellenerek hata azaltılmaya çalışılır. Böylece her eğitim örneği ağın parametrelerini daha doğru tahminler yapacak şekilde düzenler (Yang ve Wang, 2020).

Yapay sinir ağından çoğu uygulama ve çalışmada faydalanılmaktadır. Yapay sinir ağları sınıflandırma, regresyon, kümeleme ve örüntü tanıma gibi pek çok görevde etkili çözümler sunarak hem bilimsel araştırmalarda hem de endüstriyel sistemlerde önemli bir yer edinmiştir.

3.2.17. Uzun-Kısa Süreli Bellek Ağı

Derin öğrenme (deep learning), çok katmanlı yapay sinir ağları aracılığıyla veriden yüksek düzeyde soyut temsil öğrenmeyi amaçlayan bir makine öğrenmesi yaklaşımıdır. Bu çerçevede, Yinelenen Sinir Ağları (Recurrent Neural Networks, RNN), zaman serisi ya da sıralı verilerdeki bağlamsal ilişkileri modellemek amacıyla geliştirilen özel bir yapay sinir ağı türüdür. RNN, önceki zaman adımlarındaki bilgiyi döngüsel bağlantılar aracılığıyla ağ belleğinde tutarak, her bir girdi adımını önceki girdilerle birlikte değerlendirir (Graves, 2012). Bununla birlikte, klasik RNN'in gradyanların zaman içinde üstel biçimde sönümlenmesi veya büyümesi gibi sorunlar, uzun bağımlılıkların öğrenilmesini zorlaştırmaktadır. Bu sorunu aşmak amacıyla geliştirilen uzun kısa süreli

bellek (Long Short-Term Memory, LSTM) ağı, yinelenen sinir ağlarının kronik kaybolan / patlayan gradyan sorununu aşmak için geliştirilen özel bir hücre mimarisidir. Klasik RNN, hata sinyali zaman içinde üstel biçimde zayıflayarak çok uzun dizilerde öğrenmeyi olanaksız hâle getirirken, LSTM hücresi hücre durumu (c_t) adı verilen kalıcı bir bellek ve üç denetleyici kapı (giriş, unutma, çıkış) üzerinden bu sinyali korumayı başarır (Hochreiter ve Schmidhuber, 1997). LSTM, her biri kendine özgü giriş (i_t), unutma (f_t) ve çıkış (o_t) kapılarıyla denetlenen bir hücre durumunu (c_t) korur; böylece bilginin, gerekli olduğu sürece bozulmadan saklanması ve istenildiğinde okunması sağlanır (Graves, 2012). Bu yapı konuşma tanımadan seri finansal tahminlere kadar pek çok alanda RNN'in performansını belirgin biçimde iyileştirmiştir (Van Houdt vd., 2020).

Denklem 3.17- Denklem 3.22 arasında gösterilen eşitlikler LSTM bloğunun ileri geçişini özetler (zaman indeksi t soldan sağa artar). Tüm vektörel çarpımlar eleman seviyesinde Hadamard çarpımı ($\odot$) ile gösterilmiştir; $\sigma(\cdot)$ lojistik sigmoid, $g(\cdot)$ ve $h(\cdot)$ ise genellikle $tan\,h(\cdot)$ fonksiyonlarıdır:

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i), \tag{3.17}$$

$$f_t = \sigma\big(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f\big), \tag{3.18}$$

$$\widetilde{c_t} = g(W_{xc}x_t + W_{hc}h_{t-1} + b_c), \tag{3.19}$$

$$c_t = f_t \odot c_{t-1} \; + \; i_t \odot \widetilde{c_t}, \tag{3.20}$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_o), \tag{3.21}$$

$$h_t = o_t \odot h(c_t) \tag{3.22}$$

Burada t zaman adımını (dizideki örnek sırasını), $x_t \in R^{d_x}$ t'inci adımda modele giren d_x boyutlu girdi vektörünü, $h_{t-1} \in R^{d_h}$ bir önceki adım gizli durumunu (blok çıktısını), c_{t-1}, $c_t \in R^{d_h}$ sırasıyla önceki ve güncel hücre durumlarını (uzun süreli bellek

vektörleri), $i_t, f_t, o_t \in [0,1]^{d_h}$ sırasıyla giriş, unutma ve çıkış kapısı aktivasyonlarını, $\widetilde{c_t}$ aday hücre içeriğini; yeni bilginin lineer olmayan dönüşümünü ifade eder. Burada d_h, gizli katmandaki birim sayısını ifade eder ve tüm kapı, hücre durumu ve gizli durum vektörlerinin ortak boyutudur. $W_{..}$ terimi, kapılar ve hücre durumu hesaplamalarında kullanılan ağırlık matrislerini temsil eder. Burada ilk alt indis, bilginin kaynağını (x: giriş vektörü, h: önceki gizli durum, c: mevcut hücre durumu), ikinci alt indis ise hedef bileşeni (i: giriş kapısı, f: unutma kapısı, c: aday hücre içeriği, o: çıkış kapısı) ifade eder.

Bu parametrelerin tümü geri yayılım ile eğitilip, hücre durumundaki gradyan $\frac{\partial E}{\partial c_t}$ doğrudan c_{t-1} üzerinden iletilerek LSTM mimarisine özgü olarak tanımlanan Sabit Hata Döngüsü (Constant Error Carousel, CEC) mekanizması sayesinde zaman içinde bozulmadan korunur (Hochreiter ve Schmidhuber, 1997). Unutma kapısının (forget gate) aktivasyonu değeri bire yakın olduğunda, hücre durumu önceki adımlardaki bilgisini korumaya devam eder ve bu sayede hata gradyanı zaman içinde sönümlenmeden aktarılabilir. Bu özellik gradyanların zamanla üstel olarak azalmasına yol açan kaybolan gradyan (vanishing gradient) problemine karşı etkili bir çözümdür (Graves, 2012).

LSTM mimarisi, zamana bağlı verilerdeki kısa ve uzun vadeli bağımlılıkları içsel olarak modelleyebilmesi sayesinde, geleneksel klasik makine öğrenmesi yaklaşımlarına güçlü bir alternatif sunmaktadır. K-en Yakın Komşu, Destek Vektör Makineleri veya Rassal Orman gibi yöntemler doğrusal olmayan ayırımlar ve sınıflandırma görevlerinde başarılı olsa da, girdileri sabit uzunlukta vektörlere indirgemeyi ve zaman eksenindeki bağımlılıkları dolaylı öznitelik mühendisliğiyle yakalamayı zorunlu kılar. Buna karşılık LSTM, hücre durumunda (c_t) saklanan iç belleği ve denetleyici kapıları aracılığıyla geçmiş bilgiler üzerinde "seçici unutma" ve "yeniden çağırma" işlemleri yaparak sıralı bağlamı doğrudan öğrenir; bu sayede dil modelleme, konuşma tanıma ve finansal zaman serisi kestirimi gibi doğrusal olmayan, yüksek derecede otokorelasyonlu problemlerde üstün performans gösterir (Hochreiter & Schmidhuber, 1997; Graves, 2012).

Ayrıca, LSTM'ler uçtan uca diferansiyellenebilir derin ağlar olduklarından, temsil öğrenimi (representation learning) sürecini veriyle birlikte otomatik optimize eder. Böylece klasik modellerde gerekli olabilen karmaşık öznitelik çıkarımı adımları tipik

olarak ortadan kalkar. Özellikle değişken uzunluktaki dizilerin işlenmesinde, her bir örneği eş boyutlu vektöre sabitleme gereksinimi bulunmadığından, bilgi kaybı en aza iner ve model girdi sırasını doğal hâliyle kullanır (Van Houdt et al., 2020, s. 118). LSTM kapı yapıları sayesinde kaybolan gradyan problemine karşı dayanıklıdır; geriye yayılım sırasında hata sinyali hücre durumunda lineer olarak aktarılır ve uzun dizilerde bile öğrenme kararlılığı korunur (Graves, 2012). Bu nedenle LSTM, klasik makine öğrenmesi yöntemlerine kıyasla zaman içindeki ilişkileri daha iyi yakalayabilir. Klasik yöntemler genellikle sadece tek bir zaman adımındaki bilgiye veya sınırlı bir geçmişe bakabilirken, LSTM uzun süreli geçmiş bilgileri de dikkate alabilir. Bu özelliği sayesinde LSTM dil, konuşma veya finansal veri gibi zamanla değişen verilerin analizinde daha başarılı sonuçlar vermekte ve bu tür veri setlerinde sıkça tercih edilmektedir.

3.2.18. Önerilen Yaklaşım

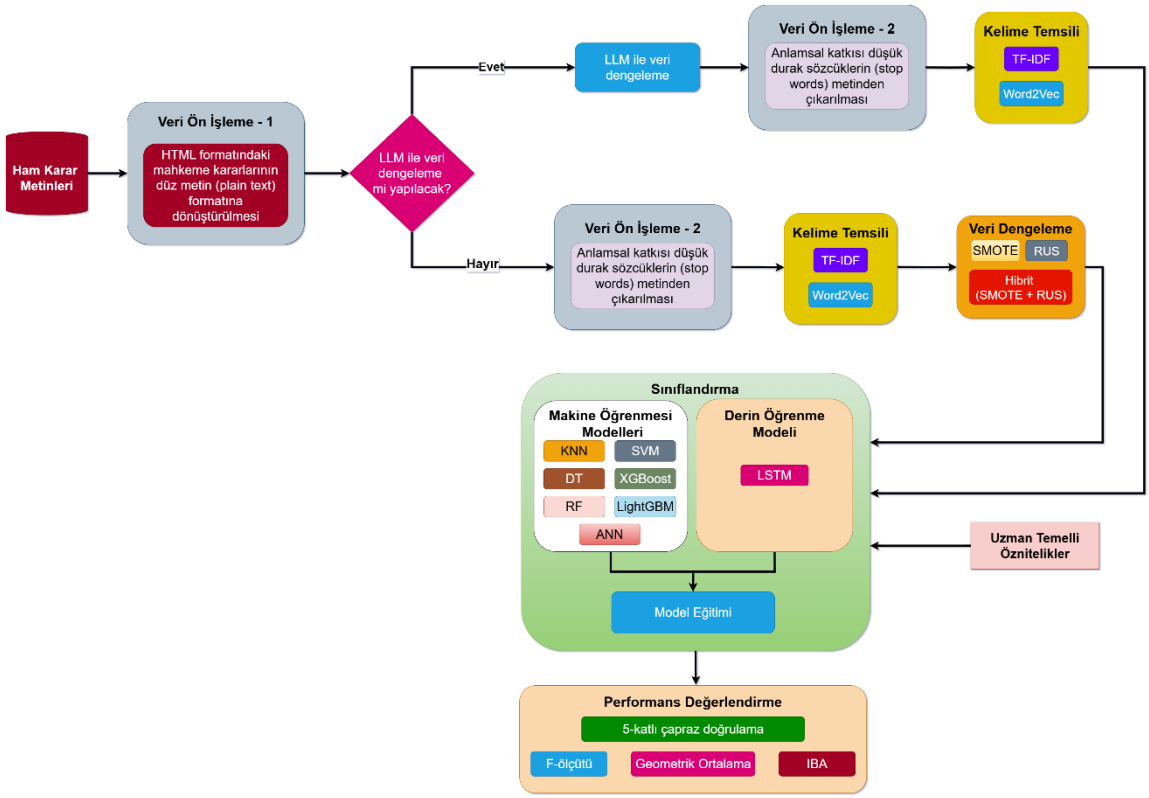
Bu çalışmada önerilen yaklaşımda, yüksek mahkeme kararlarının otomatik sınıflandırılması amacıyla TF-IDF ve Word2Vec tabanlı metin temsilleri ile hem derin öğrenme (LSTM) hem de klasik makine öğrenmesi algoritmaları kullanılmış; veri setindeki sınıf dengesizliği, LLM ile üretilen sentetik örnekler ve geleneksel yöntemler ile giderilmiştir.

İlk aşamada, ham karar metinleri toplanarak analiz için uygun hâle getirilmiştir. Bu kapsamda, mahkeme kararlarının HTML formatındaki metinleri düz metin biçimine dönüştürülmüş ve anlamsal katkısı düşük olan önemsiz sözcükler (stop words) çıkarılarak temel temizlik işlemleri uygulanmıştır.

Sonrasında, sınıflar arası dengesizliğin giderilmesi amacıyla hem LLM ile veri dengeleme hem de klasik veri dengeleme yöntemleri ayrı ayrı uygulanmıştır. LLM ile veri dengelemede, GPT-3.5-turbo büyük dil modeli kullanılarak azınlık sınıfa ait sentetik örnekler üretilmiş; ardından stop words temizlenmiş ve metinler TF-IDF ile Word2Vec yöntemleriyle sayısal olarak temsil edilmiştir. Klasik veri dengeleme yöntemlerinde ise, önemsiz sözcüklerin çıkarılmasının ardından metinler TF-IDF ve Word2Vec ile vektörleştirilmiş ve ardından SMOTE, Rastgele Alt Örnekleme (RUS) ve Hibrit (SMOTE + RUS) teknikleriyle dengeleme işlemi gerçekleştirilmiştir.

Devamında, uzman görüşüyle belirlenen üç ikili öznitelik (“doktrin/öğreti”, “karşı oy/muhalefet” ve “direnme”) veri setine eklenmiştir. Elde edilen vektör temsilleri ve ek öznitelikler, klasik makine öğrenmesi modelleri (KNN, SVM, DT, RF, XGBoost, LightGBM, ANN) ile derin öğrenme modeli (LSTM) kullanılarak sınıflandırılmıştır. Modellerin eğitimi, 5-katlı çapraz doğrulama yöntemiyle gerçekleştirilmiştir.

Son aşamada ise, model performansları; F-ölçütü, Geometrik Ortalama ve Dengelenmiş Doğruluk İndeksi (IBA) gibi sınıf dengesine duyarlı metriklerle değerlendirilmiştir. Tüm bu adımların genel akışı Şekil 3.7’de özetlenmiştir.



Şekil 3.7. Önerilen yaklaşımın akış şeması

3.2.19. Performans Metrikleri

Model performansını değerlendirmek için Çizelge 3.1'deki iki sınıflı karmaşıklık matrisi kullanılmıştır.

Çizelge 3.1. Karmaşıklık Matrisi

	Tahmin Edilen Pozitif	**Tahmin Edilen Negatif**
Gerçek Pozitif	Gerçek Pozitif (TP)	Yanlış Negatif (FN)
Gerçek Negatif	Yanlış Pozitif (FP)	Gerçek Negatif (TN)

Karmaşıklık matrisi (confusion matrix), bir sınıflandırma modelinin performansını değerlendirmek amacıyla sıklıkla başvurulan temel bir tablodur. Karmaşıklık matrisi dört temel bileşenden oluşur: Gerçek Pozitif, modelin pozitif olarak tahmin ettiği ve gerçekte de pozitif olan örnekleri; Yanlış Negatif, negatif tahmin edilen ancak gerçekte pozitif olan örnekleri; Yanlış Pozitif, pozitif tahmin edilmesine rağmen gerçekte negatif olan örnekleri; ve Gerçek Negatif ise hem modelin hem de gerçek etiketin negatif olduğu örnekleri ifade eder. Bu dört değer aracılığıyla, modelin başarımını çeşitli yönleriyle değerlendirmeye olanak tanıyan birçok performans metriği türetilmektedir.

Bu metriklerden en yaygın olanı doğruluk (accuracy) olup, doğru sınıflandırılan tüm örneklerin (TP + TN) toplam örnek sayısına oranı olarak tanımlanır. Ancak doğruluk, özellikle sınıf dağılımının dengesiz olduğu veri kümelerinde, çoğunluk sınıfına eğilimli modellerin başarımını olduğundan yüksek gösterme riski taşır (He ve Garcia, 2009). Bu sebeple, bu çalışma kapsamında doğruluk tek başına yeterli bir değerlendirme ölçütü olarak kullanılmamış; bunun yerine, karmaşıklık matrisinden türetilen ve sınıflar arası ayrımı daha hassas biçimde yansıtan diğer performans metriklerine de yer verilmiştir. Denklem 3.23'te doğruluk metriğinin matematiksel ifadesi sunulmuştur.

$$Doğruluk = \frac{TP + TN}{TP + TN + FP + FN} \tag{3.23}$$

Kesinlik (Precision), modelin pozitif olarak sınıflandırdığı örneklerin ne kadarının gerçekten pozitif olduğunu gösterir (He ve Garcia, 2009). Denklem 3.24'te kesinliğin formülü verilmiştir:

$$Kesinlik = \frac{TP}{TP + FP} \tag{3.24}$$

Duyarlılık (Recall / TPrate) gerçek pozitif örneklerin ne kadarının doğru şekilde tespit edildiğini ifade eder. Özellikle azınlık sınıfın yakalanma oranını ölçmek için önemlidir (He ve Garcia, 2009). Duyarlılık ölçütünün formülü Denklem 3.25'te gösterilmiştir:

$$Duyarlılık\ (TPrate) = \frac{TP}{TP + FN} \tag{3.25}$$

Özgüllük (Specificity / TNrate) gerçek negatif örneklerin ne kadarının doğru şekilde tanındığını ifade eder. Duyarlılık ile değerlendirilerek sınıflar arası denge hakkında bilgi verir (He ve Garcia, 2009). Denklem 3.26'da özgüllük ölçütünün formülü gösterilmiştir:

$$Özgüllük\ (TNrate) = \frac{TN}{TN + FP} \tag{3.26}$$

F-Ölçütü, kesinlik ve duyarlılığın harmonik ortalamasıdır. Dengeli bir değerlendirme sağlar ve özellikle dengesiz veri kümelerinde tercih edilir (He ve Garcia, 2009). Denklem 3.27'de F-ölçütünün formülü verilmiştir:

$$F - ölçütü = 2 \cdot \frac{Kesinlik \cdot Duyarlılık}{Kesinlik + Duyarlılık} \tag{3.27}$$

Geometrik Ortalama, duyarlılık ve özgüllüğün geometrik ortalamasını alarak, hem pozitif hem de negatif sınıfın başarısını dengeli biçimde ölçer (He ve Garcia, 2009). Denklem 3.28'de geometrik ortalama formüle edilmiştir:

$$Geometrik\ Ortalama = \sqrt{TPrate \cdot TNrate} \tag{3.28}$$

Dengeli Doğruluk İndeksi (Index of Balanced Accuracy, IBA) Geometrik ortalama değeri ile duyarlılık ve özgüllük farkını dikkate alan ağırlıklı bir ölçüttür. α parametresi ile modelin azınlık sınıfa ne ölçüde duyarlı olması gerektiği ayarlanabilir (He ve Garcia, 2009). Bu çalışmada α = 0,1 olarak belirlenmiştir.

$$IBA = \left(1 + \alpha \cdot (TPrate - TNrate)\right) \cdot TPrate \cdot TNrate \tag{3.29}$$

Bu bağlamda, sınıflandırma modellerinin etkinliğini gösterebilmek için F-ölçütünde ≥ 0,75, geometrik ortalamada ≥ 0,75 ve IBA'da (α = 0,1) ≥ 0,60 değerlerinin yakalanması hedeflenmiştir.

4. BULGULAR VE TARTIŞMA

Bu bölümde, gerçekleştirilen deneysel çalışmalara ilişkin sayısal bulgular sistematik biçimde sunulmuş ve elde edilen sonuçların anlamlılığı tartışılmıştır. Öncelikle her model–veri dengeleme bileşimi için beş tekrar ve beş katlı çapraz-doğrulama sonunda hesaplanan performans metrikleri verilmiştir. Ardından, bu metrikler arasındaki farklılıkların istatistiksel önemini belirlemek üzere Friedman sıralı varyans analizi uygulanmış; anlamlı bulunan karşılaştırmalar, Nemenyi post-hoc testi ile detaylandırılmıştır.

4.1. Deneysel Koşullar

Türkçe yüksek mahkeme kararlarına ilişkin veri kümesi, UYAP Mevzuat ve İçtihat Bilgi Sistemi'nden elde edilmiştir. Karar metinleri ilk olarak Python ortamında BeautifulSoup kütüphanesi aracılığıyla yalın metne dönüştürülmüş; ardından her belgenin sonunda yer alan "SONUÇ:" ya da "HÜKÜM:" başlıklı paragraf ayrıştırılmıştır. Bu paragrafta geçen "ONAMA", "BOZMA" veya "DÜZELTİLEREK ONAMA" anahtar sözcükleri temel alınarak her karar otomatik biçimde ilgili sınıfa etiketlenmiştir. Böylelikle onama-bozma ayrımı, karar metninin hüküm kısmına dayalı bir kural üzerinden gerçekleştirilmiştir.

Metinlerin geri kalanında olası veri sızıntısını önlemek amacıyla, gerekçe bölümlerinin son cümleleri incelenmiş; "-nması gerekir" kalıbıyla biten ve "boz" ya da "ona" köklerini içeren satırlar metinden çıkarılmıştır. Sonraki önişleme adımlarında metinler küçük harfe dönüştürülmüş; yalnızca rakamlardan oluşan belirteçler ile Türkçe durak sözcükler (stop words) metinden çıkarılmıştır. Ardından Zemberek kütüphanesi kullanılarak kök bulma işlemi uygulanmış, böylece çekim ekleri sıyrılmış ve sözcük varyantları tek bir temsile indirgenmiştir (Akın ve Akın, 2007).

Öznitelik uzayı, alınan uzman görüşü doğrultusunda tanımlanan üç ikili değişkenle zenginleştirilmiştir. "Doktrin" veya "öğreti" ifadelerinin varlığı contains_doktrin_ogreti, "karşı oy/karşıoy" ya da "muhalefet" terimlerinin varlığı contains_karsioy_muhalefet ve "direnme" sözcüğünün varlığı contains_direnme özniteliği olarak kodlanmıştır. Bu üç değişken, karar metinlerinin orijinal sırası korunarak oluşturulan bir dosyaya kaydedilmiştir.

Metin temsili iki farklı yöntemle gerçekleştirilmiştir. İlk yöntemde, TF-IDF matrisi oluşturulmuştur. İkinci yöntemde ise Köksal (2018) tarafından eğitilmiş 400 boyutlu Türkçe Word2Vec modeli yüklenmiş; her karar metni, içerdiği kelime vektörlerinin aritmetik ortalamasıyla yoğun bir vektör uzayında temsil edilmiştir.

Sınıf dengesizliğini gidermek için LLM ile veri dengeleme yapılmıştır. GPT-3.5-turbo modeli kullanılarak sentetik üst örnekleme gerçekleştirilmiştir. Az örnekle talimatlandırma (few-shot prompting*)* kurgusunda rastgele seçilen üç gerçek onama kararı modele örnek olarak verilmiş, modelin üretim davranışını kontrol eden temperature*,* top_p*,* frequency_penalty ve presence_penalty parametrelerinin en uygun değerlerini belirlemek amacıyla ızgara taraması yapılmıştır. En iyi sonuçlar temperature = 1.0, top_p = 0.8, frequency_penalty = 2.0 ve presence_penalty = −1.2 değerleriyle elde edilmiş; bu ayarlarla üretilen 708 sentetik karar metni azınlık sınıfa eklenerek CHATGPT %100 senaryosu oluşturulmuştur. LLM aracılığıyla sentetik veri üretiminin yanı sıra geleneksel veri dengeleme yöntemleri de ayrıca uygulanmıştır. Çalışmada uygulanan geleneksel veri dengeleme yöntemleri şunlardır:

- Azınlık sınıfına ait örnek sayısını artırmak amacıyla, SMOTE algoritması kullanılmıştır. Bu stratejide sırasıyla %100, %200, %300, %400 ve %600 oranlarında azınlık sınıfından sentetik örnekler üretilmiştir.

- Rastgele alt örnekleme (random undersampling) yöntemi uygulanarak azınlık sınıfının çoğunluk sınıfına oranı sırasıyla 1:1, 1:1.5 ve 1:2 olacak şekilde çoğunluk sınıfından rastgele örnekler çıkartılmıştır.

- Hem SMOTE hem rastgele alt örnekleme (random undersampling) bir arada kullanıldığı bu yaklaşımda, öncelikle azınlık sınıfın çoğunluk sınıfına oranı 1:2 oranında olacak şekilde çoğunluk sınıfından örnekler çıkarılmış, ardından azınlık sınıfı için %100 ve %200 oranlarında SMOTE uygulanarak azınlık sınıftaki örnek sayısı sentetik şekilde artırılmıştır.

Bu çalışmada deneysel analizler, Python programlama dili kullanılarak gerçekleştirilmiş; model geliştirme ve değerlendirme süreçlerinde scikit-learn, imbalanced-learn, xgboost, lightgbm, keras, tensorflow, gensim ve pandas kütüphanelerden yararlanılmıştır. Ham

karar metinlerinden oluşturulan veri kümesi, sınıf dağılımı korunacak biçimde %80 eğitim ve %20 test alt kümelerine ayrılmıştır. Eğitim verisi üzerinde değerlendirme sürecinde, tabakalı 5-katlı çapraz doğrulama (stratified 5-fold cross-validation) yöntemi uygulanmış ve bu işlem beş bağımsız tekrar ile desteklenmiştir. Böylece her model-konfigürasyon kombinasyonu için toplamda 25 farklı eğitim-test döngüsü gerçekleştirilmiş; elde edilen metriklerin aritmetik ortalaması alınarak genel model performansları hesaplanmıştır. Test veri kümesi ise model eğitimi, hiperparametre optimizasyonu veya veri dengeleme süreçlerine dahil edilmeksizin, yalnızca nihai performans değerlendirmelerinde kullanılmıştır.

4.2. Hiperparametre Optimizasyonu

Bu çalışmada kullanılan tüm sınıflandırma modelleri için hiperparametre optimizasyonu gerçekleştirilmiştir. Hiperparametre optimizasyonu, sınıflandırma modellerinin performansını en üst düzeye çıkarmak ve modelin öğrenme kapasitesini veri kümesine en uygun biçimde ayarlamak amacıyla gerçekleştirilmiştir. Her modelin yapısal olarak sahip olduğu hiperparametreler doğrudan modelin sınırlayıcı gücünü ve genelleme yeteneğini etkiler. Bu nedenle rastgele veya varsayılan değerlerle eğitilen bir modelin yerine, sistematik biçimde en uygun parametre kombinasyonlarıyla eğitilmiş bir modelin daha yüksek doğruluk, daha dengeli sınıf başarımı ve daha güvenilir öngörüler üreteceği öngörülmektedir. Bu sebeplerle çalışmada kullanılan sınıflandırma algoritmaları KNN, DT, RF, SVM, XGBoost, LightGBM, ANN ve LSTM için hiperparametre optimizasyonu ayrı ayrı gerçekleştirilmiştir. Optimizasyon süreci, klasik makine öğrenmesi modellerinde scikit-learn kütüphanesindeki GridSearchCV sınıfı kullanılarak tabakalı 5-katlı çapraz doğrulama yöntemiyle yürütülmüştür. Değerlendirme ölçütü olarak F-ölçütü tercih edilmiştir. LSTM modeli için ise hiperparametre taraması, özel olarak yapılandırılmış bir Python döngüsü aracılığıyla gerçekleştirilmiş; her kombinasyon için 5 katlı çapraz doğrulama sonucunda ortalama F-ölçütü değeri hesaplanarak en uygun yapılandırma belirlenmiştir. Uygulamalar hem TF-IDF hem de Word2Vec temsilleri üzerinden yapılmış; her bir modelin çapraz doğrulama sürecinde en yüksek performansı gösterdiği hiperparametre kombinasyonları belirlenmiştir. Elde edilen en iyi hiperparametre değerleri Çizelge 4.1'de gösterilmiştir.

Çizelge 4.1. Sınıflandırma modellerinin en iyi hiperparametre değerleri

Model	Temsil	Hiperparametre (Aralık)	Hiperparametrelerin Seçilen Değerleri
KNN	TF-IDF	n_neighbors ∈ {1, 3, 5, 7}	n_neighbors = 7
KNN	Word2Vec	n_neighbors ∈ {1, 3, 5, 7}	n_neighbors = 3
DT	TF-IDF / Word2Vec	max_depth ∈ {3, 5, 7}	max_depth = 5
RF	TF-IDF / Word2Vec	n_estimators ∈ {50, 100} max_depth ∈ {3, 5, 7}	n_estimators = 50 max_depth = 7
SVM	TF-IDF	C ∈ {0.1, 1, 10, 100, 1000} kernel ∈ {'linear', 'rbf'}	C = 10 kernel = 'rbf'
SVM	Word2Vec	C ∈ {0.1, 1, 10, 100, 1000} kernel ∈ {'linear', 'rbf'}	C = 1000 kernel = 'rbf'
XGBoost	TF-IDF / Word2Vec	n_estimators ∈ {50, 100} max_depth ∈ {3, 5, 7}	n_estimators = 50 max_depth = 7
LightGBM	TF-IDF / Word2Vec	n_estimators ∈ {50, 100} max_depth ∈ {3, 5, 7}	n_estimators = 100 max_depth = 7
ANN	TF-IDF / Word2Vec	hidden_layer_sizes ∈ {(50,), (100,)} activation ∈ {'tanh', 'relu'} solver ∈ {'sgd', 'adam'} learning_rate_init ∈ {0.01, 0.1} batch_size ∈ {16, 32} max_iter ∈ {20, 30}	hidden_layer_sizes = (100,) activation = 'relu' solver = 'adam' learning_rate_init = 0.01 batch_size = 32 max_iter = 30
LSTM	TF-IDF	activation ∈ {'tanh', 'relu'} optimizer ∈ {'sgd', 'adam'}	activation = 'relu' optimizer = 'adam'

4.3. Model Performanslarının Karşılaştırılması

Bu bölümde, sınıflandırma deneylerinin çıktıları senaryo kavramı çerçevesinde karşılaştırılmıştır. Bu çalışmada senaryo kavramı, dört bileşenden oluşan bir konfigürasyonu ifade etmektedir: seçilen sınıflandırma algoritması, kullanılan metin temsil yöntemi (TF-IDF veya Word2Vec), uygulanan veri dengeleme stratejisi (baz durum, SMOTE, rastgele alt örnekleme, hibrit yaklaşım ya da GPT-3.5-turbo temelli %100 üst örnekleme) ve uzman temelli özniteliğin modele dâhil edilme durumu (Var/Yok). Her senaryo, beş tekrarlı tabakalı beş katlı çapraz doğrulama tasarımına tabi tutulmuştur. Sınıflandırma modellerinin tüm sonuçları Ek-1'de verilmiştir. F-ölçütü, Geometrik Ortalama ve IBA ölçütlerine ait aritmetik ortalama, minimum, maksimum ve

standart sapma istatistikleri ve her sınıflandırma modelinin en yüksek F-ölçütü, geometrik ortalama ve IBA değerlerini veren senaryoları Çizelge 4.3-4.5'te gösterilmiştir.

Çizelge 4.2. TF-IDF ve Word2Vec için F-ölçütü, G-Ort ve IBA özet istatistikleri (Ortalama, Min, Maks, Standart Sapma)

Gösterim	Ölçüt	Ortalama	Min.	Maks.	Std. Sapma
TF-IDF	F-ölçütü	0,519	0,058	0,789	0,172
TF-IDF	Geometrik Ortalama	0,679	0,169	0,822	0,137
TF-IDF	IBA	0,466	0,027	0,657	0,155
Word2Vec	F-ölçütü	0,448	0,009	0,766	0,180
Word2Vec	Geometrik Ortalama	0,638	0,049	0,839	0,125
Word2Vec	IBA	0,410	0,004	0,689	0,137

Çizelge 4.3. Sınıflandırma Modellerinin En Yüksek F-ölçütü Sonuçları

Model	Temsil	Veri Dengeleme	Ek Öznitelikler	Doğruluk	TPR	TNR	F-ölçütü	Geo. Ort.	IBA
KNN	TF-IDF	CHATGPT (%100)	Yok	0,895	0,555	0,979	0,678	0,737	0,521
DT	TF-IDF	CHATGPT (%100)	Yok	0,898	0,593	0,974	0,698	0,760	0,555
RF	TF-IDF	CHATGPT (%100)	Yok	0,898	0,511	0,995	0,666	0,712	0,483
SVM	TF-IDF	CHATGPT (%100)	Yok	0,919	0,657	0,984	0,762	0,804	0,625
XGBoost	**TF-IDF**	**CHATGPT (%100)**	**Var**	**0,927**	**0,685**	**0,987**	**0,789**	**0,822**	**0,656**
LightGBM	TF-IDF	CHATGPT (%100)	Var	0,927	0,672	0,991	0,785	0,816	0,644
ANN	TF-IDF	CHATGPT (%100)	Yok	0,905	0,706	0,955	0,748	0,821	0,657
LSTM	Word2Vec	CHATGPT (%100)	Var	0,916	0,693	0,971	0,766	0,820	0,654

Çizelge 4.4. Sınıflandırma Modellerinin En Yüksek Geometrik Ortalama Sonuçları

Model	Gösterim	Veri Dengeleme	Ek Öznitelikler	Doğruluk	TPR	TNR	F-ölçütü	Geo. Ort.	IBA
KNN	TF-IDF	RUS (1:1,5)	Yok	0,773	0,737	0,777	0,418	0,757	0,571
DT	TF-IDF	CHATGPT (%100)	Yok	0,898	0,593	0,974	0,698	0,760	0,555
RF	TF-IDF	Hibrit (RUS 1:2 + SMOTE 200%)	Var	0,743	0,790	0,737	0,404	0,763	0,585
SVM	TF-IDF	CHATGPT (%100)	Var	0,916	0,661	0,979	0,757	0,804	0,627
XGBoost	TF-IDF	CHATGPT (%100)	Var	0,927	0,685	0,987	0,789	0,822	0,656
LightGBM	TF-IDF	CHATGPT (%100)	Yok	0,927	0,674	0,989	0,785	0,817	0,646
ANN	TF-IDF	CHATGPT (%100)	Yok	0,905	0,706	0,955	0,748	0,821	0,657
LSTM	**Word2Vec**	**CHATGPT (%100)**	**Var**	**0,909**	**0,740**	**0,951**	**0,764**	**0,839**	**0,689**

Çizelge 4.5. Sınıflandırma Modellerinin En Yüksek IBA Sonuçları

Model	Gösterim	Veri Dengeleme	Ek Öznitelikler	Doğruluk	TP-Oranı	TN-Oranı	F-ölçütü	G-Ort	IBA
KNN	TF-IDF	RUS (1:1,5)	Yok	0,773	0,737	0,777	0,418	0,757	0,571
DT	TF-IDF	CHATGPT (%100)	Yok	0,898	0,593	0,974	0,698	0,760	0,555
RF	TF-IDF	Hibrit (RUS 1:2 + SMOTE 200%)	Var	0,743	0,790	0,737	0,404	0,763	0,585
SVM	TF-IDF	CHATGPT (%100)	Var	0,916	0,661	0,979	0,757	0,804	0,627
XGBoost	TF-IDF	CHATGPT (%100)	Var	0,927	0,685	0,987	0,789	0,822	0,656
LightGBM	TF-IDF	CHATGPT (%100)	Yok	0,927	0,674	0,989	0,785	0,817	0,646
ANN	TF-IDF	CHATGPT (%100)	Var	0,901	0,710	0,949	0,740	0,820	0,657
LSTM	**Word2Vec**	**CHATGPT (%100)**	**Var**	**0,909**	**0,740**	**0,951**	**0,764**	**0,839**	**0,689**

Veri kümesindeki dengesiz sınıf dağılımı dikkate alındığında, klasik makine öğrenmesi algoritmaları ile derin öğrenme tabanlı yaklaşımların sınıflandırma performansları arasında belirgin farklar olduğu Çizelge 4.2 ve Çizelge 4.3'te gözlemlenmiştir. Özellikle karar ağaçlarına dayalı yöntemler, çoğunluk sınıfını doğru tanımada üstün başarı gösterirken; LSTM modeli, azınlık sınıfını daha etkin biçimde tespit edebilme yetenekleriyle öne çıkmaktadır. XGBoost, çoğunluk sınıfına ait örneklerde hatalı pozitifleri en aza indirerek özgüllüğü (TNR = 0,99) ve kesinliği (Precision = 0,93) yüksek tutmuş, böylece Çizelge 4.3'te raporlanan 0,789'luk en iyi F-ölçütü değerine ulaşmıştır.

XGBoost–TF-IDF konfigürasyonu yanlış-pozitifleri baskılayarak kesinliği 0,93 seviyesine çıkarabilmiş olsa da, özgüllük ile duyarlılık arasındaki dengesizlik (TPR = 0,69) nedeniyle İBA değeri 0,656'da sınırlı kalmıştır (bkz. Çizelge 4.5). LSTM ise kelimeler arasındaki bağlam bilgisini kullanarak azınlık sınıfını (TPR = 0,74) yakalama oranını artırmış; Çizelge 4.4 ve Çizelge 4.5'te gösterildiği üzere Geometrik Ortalama ve IBA ölçütlerinde öne geçmiştir. LSTM-Word2Vec, özgüllüğü %95 düzeyinde korumuş ve dengesizlik odaklı performans ölçütlerinde zirveye yerleşmiştir; F-ölçütte (0,764) nispeten geride kalmasının nedeni kesinliğin 0,79 seviyesinde kalmasıdır. Beş tekrarın ortalamaları incelendiğinde Çizelge 4.2, TF-IDF temsilinin Word2Vec'e kıyasla F-ölçütte yaklaşık 0,08; Geometrik Ortalama'da 0,04 ve IBA'da 0,06 puan daha yüksek sonuç verdiğini göstermektedir. Bu fark özellikle XGBoost ve LightGBM gibi ağaç tabanlı modellerde büyürken, Word2Vec ancak LSTM'de TF-IDF düzeyine yaklaşabilmiştir. Dolayısıyla hukuki metinlerde sık geçen kalıp ifadeleri doğrudan sayabilen TF-IDF, yanlış-pozitifleri azaltarak yüksek kesinlik ve dengeli doğruluk üretirken; Word2Vec'in anlamsal vektörleri LSTM dışında ayrım gücünü tam olarak yansıtamamıştır. Azınlık (onama) hatalarının kritik olduğu yargı kararlarının sınıflandırılmasında, Çizelge 4.5'e göre LSTM-Word2Vec konfigürasyonu daha dengeli bir alternatif sunmakta; yüksek TPR sayesinde azınlık sınıfını güvenilir biçimde tespit etmektedir. Buna karşılık, yüksek özgüllük ve kesinlik için XGBoost ile TF-IDF tercih edilebilir.

4.4. İstatistiksel Analiz

Bu bölümde, kelime temsili, veri dengeleme yöntemi ve sınıflandırıcıdan oluşan farklı kombinasyonlar arasındaki sınıflandırma performansı farklarının istatistiksel olarak anlamlı olup olmadığı incelenmiştir. Ortalama performans değerleri rastlantısal dalgalanmalara duyarlı olabileceğinden, değerlendirme sıralama temelli istatistiksel yöntemlerle gerçekleştirilmiştir. Bu çalışmada, birden fazla modelin çapraz doğrulama tekrarları altında göreli performans sıralamalarını karşılaştırmak amacıyla parametrik varsayımlara ihtiyaç duymayan Friedman testi ve sonrasında Nemenyi post-hoc testi tercih edilmiştir.

Her bir kombinasyon üç temel bileşenden oluşmaktadır:

- Kelime temsili: TF-IDF vektörleri ya da önceden eğitilmiş Word2Vec vektörleri,
- Veri dengeleme yöntemi: ChatGPT tabanlı sentetik üst örnekleme, Rastgele Alt Örnekleme (RUS), SMOTE (%100 veya %200), Hibrit (RUS 1:2 + SMOTE %100) ya da Hibrit (RUS 1:2 + SMOTE %200),
- Sınıflandırıcı: KNN, DT, RF, SVM, XGBoost, LightGBM, ANN ve LSTM.

Karşılaştırma süreci, istatistiksel testlerin güvenilirliğini artırmak ve çoklu karşılaştırmalardan kaynaklanan hata olasılığını azaltmak amacıyla iki aşamalı bir yapıda gerçekleştirilmiştir.

Birinci aşamada, her sınıflandırıcının kendi içinde, kelime temsili ve dengeleme stratejilerine dayalı kombinasyonları karşılaştırılmış; F-ölçütü, Geometrik Ortalama ve IBA performans metriklerine göre en iyi kombinasyonlar belirlenmiştir.

İkinci aşamada, her sınıflandırıcı için belirlenen en iyi kombinasyonlar bir araya getirilerek genel bir karşılaştırma yapılmış; böylece yalnızca performans açısından öne çıkan yapılandırmalar istatistiksel olarak yeniden test edilmiştir.

Her iki aşamadaki karşılaştırmalar sıralama temelli Friedman testi ile yürütülmüştür. Bu test, parametrik varyans analizinde gerekli olan normal dağılım varsayımına ihtiyaç duymaksızın sıralama temelli farkların test edilmesine olanak tanır (Friedman, 1937). Test istatistiği Denklem 4.1'de verilmiştir:

$$\chi_F^2 = \frac{12}{kb(k+1)} \sum_{j=1}^{k} R_j^2 - 3b(k+1) \qquad (4.1)$$

Burada χ_F^2 Friedman test istatistiğidir ve ki-kare dağılımı ile karşılaştırılır. k, karşılaştırılan grup (kombinasyon) sayısını ifade eder. b blok sayısını belirtir; bu çalışmada her bir tekrar (örneğin çapraz doğrulama tekrarı) bir blok olarak kabul edilmiştir. R_j , j'inci kombinasyonun tüm bloklar üzerindeki toplam sıralama değeridir; yani her blokta j'inci kombinasyonun aldığı sıra değerlerinin toplamıdır. $\sum_{j=1}^{k} R_j^2$ tüm

kombinasyonların (1'den k'ya kadar) toplam sıralamalarının karelerinin toplamıdır. $\frac{12}{kb(k+1)}$ katsayısı test istatistiğinin ölçeklenmesini ve uygun dağılımla karşılaştırılabilmesini sağlar. $3b(k+1)$ düzeltme terimidir; formülün dağılımın beklentisiyle uyumlu hale getirilmesini sağlar. Friedman analizi, kombinasyonların ortalama sıralamalarının rastgele mi, yoksa sistematik farklılıklar içerip içermediğini sınar (Friedman, 1937). Uygulanan sıralama temelli Friedman testi, kombinasyonların çapraz doğrulama tekrarları (bloklar) içerisindeki göreli sıralamalarının yalnızca rastlantısal mı yoksa sistematik olarak farklılaşıp farklılaşmadığını sınamak amacıyla kullanılmıştır. Testin sıfır hipotezi (H_0): "Tüm kombinasyonların ortalama sıralamaları eşittir." biçimindedir. Buna karşılık, alternatif hipotez (H_1): "En az bir kombinasyonun ortalama sıralaması diğerlerinden anlamlı düzeyde farklıdır." şeklinde tanımlanmıştır.

H_0 reddedildiğinde, performans farklılıklarının hangi kombinasyon çiftleri arasında istatistiksel olarak anlamlı olduğunu belirlemek amacıyla Nemenyi çoklu karşılaştırma testi uygulanmıştır. Bu post-hoc analiz, kombinasyonların ortalama sıralamaları arasındaki ikili farkların istatistiksel anlamlılığını değerlendirerek farklılaşmanın yönünü ve kapsamını ortaya koymayı amaçlamaktadır. Nemenyi testi, özellikle parametrik varsayımlardan bağımsız olarak çoklu karşılaştırmalar yapabilmek amacıyla geliştirilmiştir. Nemenyi testi, ortalama sıralamalar arasındaki farkı kritik fark (Critical Difference, CD) değeriyle karşılaştırılmaktadır. CD formülü Denklem 4.2'de verilmiştir:

$$CD = q_\alpha \cdot \sqrt{\frac{k(k+1)}{6b}} \tag{4.2}$$

Burada CD, Kritik Fark (Critical Difference), ortalama sıralamalar arasındaki farkın istatistiksel olarak anlamlı sayılması için aşılması gereken eşiği belirtir. k karşılaştırılan model sayısıdır. q_α nemenyi post-hoc testi kapsamında kullanılan bir kritik katsayıdır ve istatistiksel olarak güven düzeyine (α) karşılık gelen aralık dağılımından alınır. Friedman testi sonrası yapılan Nemenyi çoklu karşılaştırma analizlerinde, iki kombinasyonun ortalama sıralamaları arasındaki farkın anlamlı olup olmadığını belirlemek için Denklem

4.2 kullanılır. Fark CD'den büyükse, kombinasyonlar arasında istatistiksel olarak anlamlı fark vardır (Demšar, 2006; Nemenyi, 1963).

Demšar (2006), Friedman ve Nemenyi testlerinin makine öğrenmesindeki karşılaştırmalı deneylerde nasıl uygulanması gerektiğini sistematik olarak açıklamış ve bu testlerin sonuçlarının CD diyagramlarıyla görsel olarak ifade edilmesini önermiştir. Bu yaklaşımın temel avantajı, çok sayıda yöntem karşılaştırmasının hem istatistiksel olarak hem de sezgisel olarak anlamlandırılmasını sağlamasıdır (Demšar, 2006).

Bu doğrultuda, bu çalışmada da her bir performans metriği (F-ölçütü, Geometrik Ortalama ve IBA) için kombinasyonların ortalama sıralamaları görselleştirilmiş; ardından, kritik fark (CD) değeri esas alınarak istatistiksel olarak anlamlı fark göstermeyen kombinasyonlar çizgilerle aynı grupta gösterilmiştir.

İki aşamalı yapı, tek adımda yapılacak çok sayıda ikili karşılaştırmanın yol açacağı yanlış-pozitif hata oranını azaltmak için tercih edilmiştir. İlk aşamada her sınıflandırıcı için tüm kombinasyonlar analiz edilmiş; anlamlılık bulunduğunda Nemenyi testi ile ikili farklar değerlendirilmiştir. Metriklere bağlı olarak farklı kombinasyonlar öne çıktığından toplam dokuz yöntem ikinci aşamaya taşınmıştır. İkinci aşamada bu dokuz yöntem tek bir analiz kümesinde karşılaştırılmış; Friedman testi ile anlamlılık incelenmiş ve gerekli ise Nemenyi testi ile detaylandırılmıştır. Çizelge 4.6'da Birinci aşamada F-ölçütü, Geometrik Ortalama ve IBA için Friedman Test İstatistikleri gösterilmiştir.

Çizelge 4.6. Birinci aşamada F-ölçütü ve Geometrik Ortalama için Friedman Test İstatistikleri

Ölçüt	**k**	**χ^2**	**sd**	**p**
F-ölçütü	160	789,283	159	<0,001
Geometrik Ortalama	160	788,845	159	<0,001
IBA	160	788,860	159	<0,001

Friedman testi, her model-ölçüt kombinasyonunda kombinasyonların ortalama sıralamalarının eşit olduğu yönündeki sıfır hipotezini test etmektedir. Çizelge 4.6'da raporlanan χ^2 değerleri ile bunlara karşılık gelen $p<0,001$ düzeyleri, 159 serbestlik

derecesi altında kritik değerin oldukça üzerinde yer almaktadır. Bu sonuç, incelenen kombinasyonların başarı sıralamalarının yalnızca rastlantısal değişimlerle açıklanamayacağını ve aralarında istatistiksel olarak anlamlı farklılıklar bulunduğunu göstermektedir.

Friedman testi sonucunda kombinasyonlar arasında istatistiksel olarak anlamlı farklar saptandıktan sonra, bu farkların hangi kombinasyon çiftleri arasında ortaya çıktığını belirlemek amacıyla sıralama temelli bir çoklu karşılaştırma yöntemi olan Nemenyi post-hoc testi uygulanmıştır. Bu analizde, kombinasyonların ortalama sıralamaları arasındaki farklar, Kritik Fark (Critical Difference, CD) değeri ile karşılaştırılmış ve yalnızca bu değeri aşan farklılıklar istatistiksel olarak anlamlı kabul edilmiştir. Aşama 1 kapsamında elde edilen grafiklerde CD değeri 13,26 olarak hesaplanmış olup yalnızca belirgin biçimde öne çıkan kombinasyonların istatistiksel üstünlük göstermesine olanak tanımıştır. Buna karşılık, bazı kombinasyonlar benzer ortalama sıralamalara sahip olduklarından aynı grupta yer almış ve istatistiksel fark göstermemiştir. Üretilen tüm Nemenyi grafiklerine Ek 9'da yer verilmiştir. Aşama 1'e ilişkin F-ölçütü, Geometrik Ortalama ve IBA açısından en başarılı beş kombinasyon sırasıyla Çizelge 4.7, Çizelge 4.8 ve Çizelge 4.9'da özetlenmiştir.

Çizelge 4.7. Birinci Aşamada F-Ölçütüne Göre En Başarılı Beş Kombinasyon

Sıra	**Kombinasyon**	**Ortalama Sıralama Değeri**
1	TF-IDF + ChatGPT + XGBoost	1,4
2	TF-IDF + ChatGPT + LightGBM	1,6
3	Word2Vec + ChatGPT + LSTM	3,4
4	TF-IDF + ChatGPT + SVM	4,0
5	TF-IDF + ChatGPT + LSTM	4,6

Çizelge 4.8. Birinci Aşamada Geometrik Ortalama Ölçütüne Göre En Başarılı Beş Kombinasyon

Sıra	Kombinasyon	Ortalama Sıralama Değeri
1	Word2Vec + ChatGPT + LSTM	1,0
2	TF-IDF + ChatGPT + XGBoost	2,6
3	TF-IDF + ChatGPT + LightGBM	3,2
4	TF-IDF + ChatGPT + ANN	3,4
5	TF-IDF + ChatGPT + LSTM	4,8

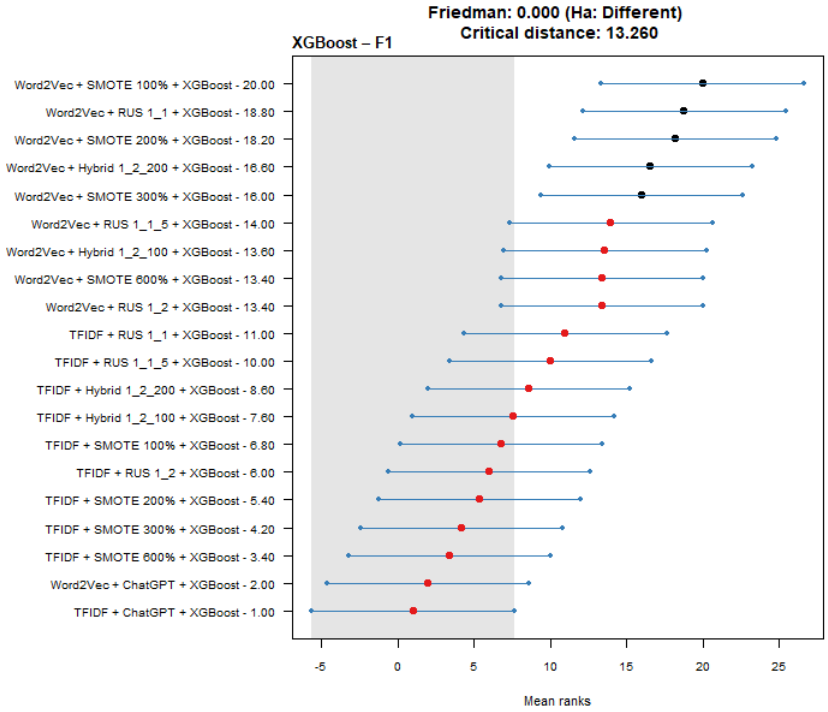
Çizelge 4.9. Birinci Aşamada IBA Ölçütüne Göre En Başarılı Beş Kombinasyon

Sıra	Kombinasyon	Ortalama Sıralama Değeri
1	Word2Vec + ChatGPT + LSTM	1
2	TF-IDF + ChatGPT + XGBoost	2,6
3	TF-IDF + ChatGPT + ANN	3,4
4	TF-IDF + ChatGPT + LightGBM	3,4
5	TF-IDF + ChatGPT + LSTM	4,8

Üç ölçütte de görülen sıra değerleri, ChatGPT tabanlı dengelemenin sistematik bir üstünlük sağladığını açık biçimde ortaya koymaktadır. F-ölçütünde TF IDF temsiliyle XGBoost ve LightGBM (1,4 ve 1,6) ilk iki sırayı alırken, Geometrik Ortalama ve IBA'da Word2Vec + ChatGPT + LSTM'in 1,0 ortalama sıra ile tek başına birinci olması, bu kombinasyonun özellikle sınıf dengesizliği ve hata dağılımına duyarlı metriklerde daha istikrarlı bir performans verdiğini göstermektedir. TF IDF + ChatGPT + XGBoost'un her üç ölçütte üst sıralarda yer alması, TF IDF temsiliyle XGBoost ve LightGBM modellerinin F ölçütü bakımından avantajlı olduğunu; buna karşın Word2Vec + ChatGPT + LSTM'in özellikle IBA ve Geometrik Ortalama gibi dengesiz veri koşullarını daha iyi yansıtan ölçütlerde öne çıktığını işaret etmektedir. İlk beş içinde yer alan diğer

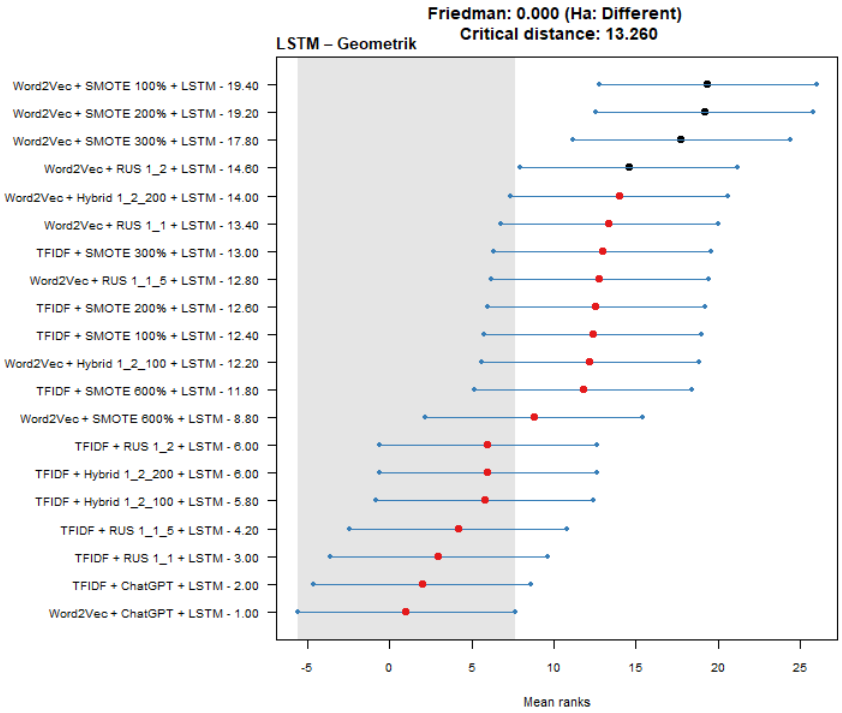
kombinasyonların (örneğin TF IDF + ChatGPT + ANN/LightGBM/LSTM) de CD bandı içinde birbirine yakın sonuçlar verdiği görülmekte; bu durum, ChatGPT tabanlı senaryoların genel olarak üst performans bandını oluşturduğunu, ancak temsil ve model seçimine bağlı olarak ölçüt bazında küçük yer değişikliklerinin ortaya çıktığını göstermektedir.

Şekil 4.1-Şekil 4.3'deki Nemenyi diyagramları, Çizelge 4.7–4.9'da birinci sırada yer alan kombinasyonların (F ölçütü için TF IDF + ChatGPT + XGBoost; Geometrik Ortalama ve IBA için Word2Vec + ChatGPT + LSTM) yalnızca tüm senaryolarla değil, aynı modelin alternatif dengeleme ve temsil yapılandırmalarıyla da kritik uzaklık (CD) eşiği bakımından nasıl ayrıştığını görselleştirmektedir.



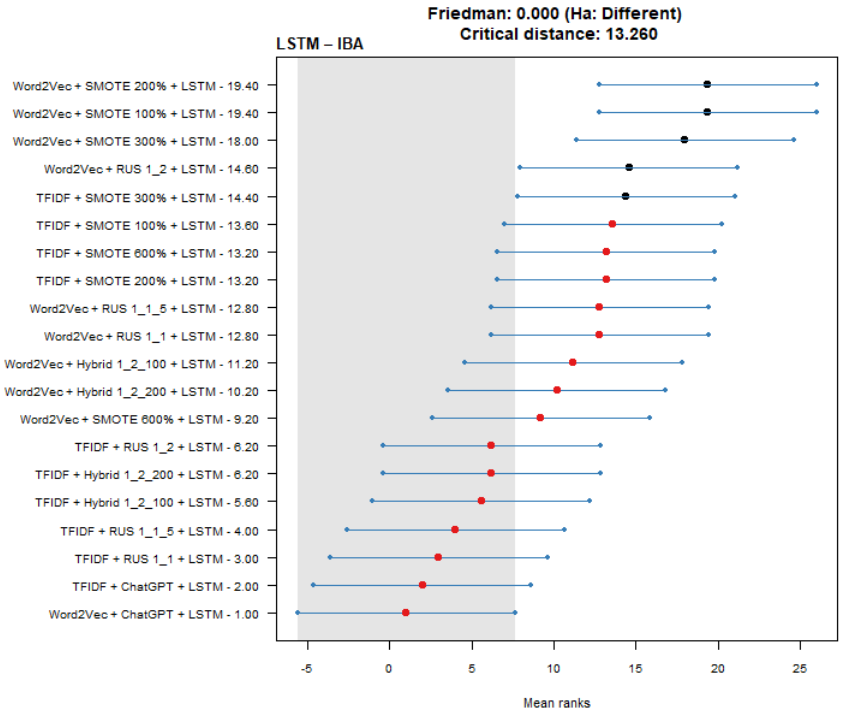
Şekil 4.1. XGBoost Modelinin Birinci Aşamada F-Ölçütüne Göre Nemenyi Diyagramı

Şekil 4.1'de yer alan Nemenyi diyagramında TF-IDF + ChatGPT + XGBoost kombinasyonunun diğer LSTM modelleri arasında en düşük ortalama sırayla öne çıktığını, ancak CD=13.260 bandı içinde kalan kombinasyonlarla istatistiksel olarak ayrışmazken, Word2Vec + SMOTE %100, %200, RUS (1:1) ve Hibrit (RUS 1:2 + SMOTE %200) yapılandırmalarından anlamlı biçimde daha iyi olduğunu göstermektedir.



Şekil 4.2. LSTM Modelinin Birinci Aşamada Geometrik Ortalamaya Göre Nemenyi Diyagramı

Şekil 4.2'deki Nemenyi diyagramında, Word2Vec + ChatGPT + LSTM kombinasyonunun en düşük ortalama sırayla sol uçta konumlandığını; CD=13.260 bandı içinde kalan diğer kırmızı LSTM yapılandırmalarıyla istatistiksel olarak ayrışmazken, Word2Vec temsilleriyle %100, %200 ve %300 oranında SMOTE ve RUS 1:2 uygulandığı durumlardan istatistiksel olarak anlamlı düzeyde farklıdır.



Şekil 4.3. LSTM Modelinin Birinci Aşamada IBA'ya Göre Nemenyi Diyagramı

Diyagram, Word2Vec + ChatGPT + LSTM kombinasyonunun en düşük ortalama sırayla sol uçta yer aldığını; CD=13.260 bandı içinde kalan diğer kırmızı LSTM varyantlarıyla anlamlı fark göstermediğini, ancak sağ taraftaki Word2Vec + SMOTE %100, %200, %300, RUS (1:2) yapılandırmalarından istatistiksel olarak belirgin biçimde üstün olduğunu ortaya koymaktadır.

Birinci aşamanın genel bulguları, kombine edilen üç bileşenin (temsili, dengeleme, sınıflandırıcı) performansı üzerinde belirgin ve ölçüt-bağımlı etkiler bulunduğunu göstermektedir. Özellikle ChatGPT tabanlı sentetik örnekleme, çoğu model ve temsil kombinasyonunda üst sıralarda yer almış ancak CD bandı içinde kalan alternatif kombinasyonlar ile farkın her zaman istatistiksel olarak anlamlı olmadığı görülmüştür. Buna karşılık, klasik dengeleme stratejileri (RUS/SMOTE/Hibrit) genellikle orta-alt sıralarda kümelenmiş ve çoğunlukla birbirleriyle ayrışmamıştır.

Birinci aşama analizlerinde, her sınıflandırıcı için F-ölçütü veya Geometrik Ortalama açısından en iyi performans gösteren kombinasyonlar belirlenmiş ve bu kombinasyonlar ikinci aşama karşılaştırmalarına dahil edilmiştir. Bu kapsamda, toplam dokuz kombinasyon değerlendirmeye alınmıştır:

- TF-IDF + ChatGPT + ANN
- TF-IDF + ChatGPT + DT
- TF-IDF + ChatGPT + KNN
- TF-IDF + ChatGPT + SVM
- TF-IDF + ChatGPT + RF
- TF-IDF + ChatGPT + LightGBM
- TF-IDF + ChatGPT + XGBoost
- TF-IDF + Hibrit (RUS 1:2 + SMOTE %200) + RF
- Word2Vec + ChatGPT + LSTM

İkinci aşamada, bu kombinasyonların tümü aynı analiz kümesi içerisinde ele alınarak karşılaştırılmıştır. Kombinasyonların her performans ölçütü açısından ortalama sıralamaları hesaplanmış; Friedman testi ile gruplar arasında istatistiksel olarak anlamlı farklılık bulunup bulunmadığı test edilmiştir. Anlamlı fark olması halinde hangi kombinasyon çiftleri arasında fark bulunduğunu belirlemek amacıyla Nemenyi post-hoc testi uygulanmıştır. Bu aşama, yalnızca ön seçimden geçen yapılandırmaların karşılaştırılmasına olanak sağlayarak güvenilirlik düzeyini yükseltmeyi ve çoklu karşılaştırmalardan kaynaklanan istatistiksel hata olasılığını azaltmayı hedeflemektedir.

Friedman testi için elde edilen χ^2 değerleri (F-ölçütü: 39,253; geometrik ortalama: 38,720) sekiz serbestlik derecesiyle kritik değerin ($\chi^2_{(8,0.05)} = 15,51$) üzerindedir; dolayısıyla sıfır hipotezi reddedilmiş ve kombinasyonların sıralama dağılımlarında istatistiksel olarak anlamlı farklılıklar bulunduğu sonucuna ulaşılmıştır. İlgili test istatistikleri Çizelge 4.9'da sunulmuştur.

Çizelge 4.10. İkinci aşamada F-ölçütü, Geometrik Ortalama ve IBA için Friedman Test İstatistikleri

Ölçüt	k	χ^2	sd	p
F-ölçütü	9	39,253	8	<0,001
Geometrik Ortalama	9	38,720	8	<0,001
IBA	9	39,040	8	<0,001

İkinci aşamada, Friedman testi ile kombinasyonlar arasında anlamlı farklar saptanmasının ardından, hangi yapılandırmalar arasında bu farkların oluştuğunu belirlemek amacıyla Nemenyi post-hoc testi uygulanmıştır. Ortalama sıralamalar arasındaki farkların istatistiksel anlamlılığı, bu aşama için hesaplanan kritik fark (CD = 5,372) değeri temelinde değerlendirilmiştir. Bu kapsamda yalnızca ortalama sıralamaları arasındaki fark 5,372'yi aşan kombinasyonlar istatistiksel olarak birbirinden farklı kabul edilmiştir.

F-ölçütü Geometrik Ortalama ve IBA ölçütleri açısından kombinasyonların ortalama sıralamaları Çizelge 4.11, Çizelge 4.12 ve Çizelge 4.13'de sunulmuştur. F-ölçütü açısından TF-IDF + ChatGPT + XGBoost kombinasyonsu 1,4 ortalama sıra değeri ile en iyi performansı sergilemiştir. Geometrik Ortalama ölçütü açısından ise Word2Vec + ChatGPT + LSTM, 1,0 ortalama sıra değeri ile ilk sırada yer almıştır. İkinci aşamaya ilişkin Nemenyi grafikleri Ek-9'da verilmiştir.

Çizelge 4.11. İkinci Aşamada F-Ölçütüne Göre Ortalama Sıralamalar

Sıra	Kombinasyon	Ortalama Sıralama Değeri
1	TF-IDF + ChatGPT + XGBoost	1,4
2	TF-IDF + ChatGPT + LightGBM	1,6
3	Word2Vec + ChatGPT + LSTM	3,2
4	TF-IDF + ChatGPT + SVM	3,8
5	TF-IDF + ChatGPT + ANN	5
6	TF-IDF + ChatGPT + DT	6
7	TF-IDF + ChatGPT + KNN	7,2
8	TF-IDF + ChatGPT + RF	7,8
9	TF-IDF + Hibrit (RUS 1:2+SMOTE %200) + RF	9

Çizelge 4.12. İkinci Aşamada Geometrik Ortalamaya Göre Ortalama Sıralamalar

Sıra	Kombinasyon	Ortalama Sıralama Değeri
1	Word2Vec + ChatGPT + LSTM	1
2	TF-IDF + ChatGPT + XGBoost	2,4
3	TF-IDF + ChatGPT + LightGBM	3,2
4	TF-IDF + ChatGPT + ANN	3,4
5	TF-IDF + ChatGPT + SVM	5
6	TF-IDF + Hibrit (RUS 1:2+SMOTE %200) + RF	6,4
7	TF-IDF + ChatGPT + DT	6,6
8	TF-IDF + ChatGPT + KNN	8
9	TF-IDF + ChatGPT + RF	9

Çizelge 4.13. İkinci Aşamada IBA Ölçütüne Göre Ortalama Sıralamalar

Sıra	Kombinasyon	Ortalama Sıralama Değeri
1	Word2Vec + ChatGPT + LSTM	1
2	TF-IDF + ChatGPT + XGBoost	2,4
3	TF-IDF + ChatGPT + LightGBM	3,2
4	TF-IDF + ChatGPT + ANN	3,4
5	TF-IDF + ChatGPT + SVM	5
6	TF-IDF + Hibrit (RUS 1:2+SMOTE %200) + RF	6
7	TFIDF + ChatGPT + DT	7
8	TF-IDF + ChatGPT + KNN	8
9	TF-IDF + ChatGPT + RF	9

Birlikte ele alındığında, Friedman testi kombinasyonların performans dağılımlarında anlamlı çeşitlilik olduğunu ortaya koymuş; Nemenyi testi ise bu farkın yönünü ve büyüklüğünü netleştirerek hangi kombinasyonların ölçüt bazında istatistiksel olarak anlamlı farka sahip olduğunu belirlemiştir. Yapılan istatistiksel analizler sonucunda, farklı kelime temsili, veri dengeleme stratejisi ve sınıflandırıcı bileşimlerinden oluşan kombinasyonlar arasında istatistiksel olarak anlamlı performans farkları olduğu ortaya konmuştur.

İki aşamalı analiz, kombinasyonların yalnızca mutlak performansları değil, istatistiksel olarak güvenilir biçimde ayrışma düzeylerini de ortaya koymuştur. Özellikle F-ölçütü açısından TF-IDF + ChatGPT + XGBoost yapılandırması en düşük ortalama sıralama ile en başarılı kombinasyon olarak öne çıkarken, Geometrik Ortalama ölçütü bakımından Word2Vec + ChatGPT + LSTM yapılandırması istatistiksel olarak en iyi performansa sahip kombinasyon olarak belirlenmiştir. IBA sonuçları da bu tabloyu desteklemektedir: Word2Vec + ChatGPT + LSTM en düşük ortalama sırayı elde etmiş, onu TF-IDF + ChatGPT + XGBoost ve TF-IDF + ChatGPT + LightGBM izlemiştir. Bu sonuçlar, sınıf dengesizliği içeren veri yapılarında, açıklık ve genelleyici başarı açısından TF-IDF

temsilli XGBoost modeli ile Word2Vec temsilli LSTM modelinin tercih edilebilirliğini güçlendirmektedir. Dolayısıyla, uygulanacak sınıflandırma görevinin amacına ve öncelikli metrik türüne bağlı olarak bu iki yapılandırmadan biri öne çıkarılabilir: Sınıflandırma başarımının genel dengesine (precision-recall dengesi) odaklanıldığında TF-IDF + ChatGPT + XGBoost; azınlık sınıfın duyarlılığına öncelik verildiğinde ise Word2Vec + ChatGPT + LSTM yapılandırması kullanılabilir. Sonuç olarak, elde edilen bulgular doğrultusunda en uygun model seçiminin, probleme özgü öncelikli başarı ölçütlerine göre yapılandırılmış kombinasyonlar arasından, istatistiksel olarak üstünlük gösteren alternatifler dikkate alınarak yapılması uygun olacaktır.

İkinci aşamada yeniden değerlendirilen dokuz kombinasyon üzerinde yürütülen Friedman testi, ortalama sıralamalar arasında istatistiksel olarak anlamlı farklar bulunduğunu ortaya koymuştur. Ardından gerçekleştirilen Nemenyi post-hoc analizi (CD = 5,372), ChatGPT tabanlı dengeleme stratejilerinin, özellikle belirli model ve temsil yapılandırmaları altında, diğer yöntemlerden istatistiksel olarak anlamlı düzeyde ayrıldığını göstermiştir. TF-IDF + ChatGPT + XGBoost bileşimi, F-ölçütü bakımından diğer tüm kombinasyonlardan anlamlı şekilde ayrışırken; Word2Vec + ChatGPT + LSTM kombinasyonu, Geometrik Ortalama metriği açısından benzer bir farklılık ortaya koymuştur. ChatGPT ile dengeleme yapılmayan yapılandırmalardan yalnızca biri (TF-IDF + Hibrit 1:2 + SMOTE %200 + RF), hiçbir metrikte kritik fark eşiğini aşamamış ve ChatGPT destekli yöntemlerin oluşturduğu gruba dahil olabilmiştir. Benzer biçimde, IBA metriğinde de Word2Vec + ChatGPT + LSTM en düşük ortalama sırayı elde etmiş, TF-IDF + ChatGPT + XGBoost ve LightGBM tarafından takip edilmiştir. Metin sınıflandırma problemlerinde kullanılan dengeleme stratejileri bağlamında, ChatGPT ile üretilen sentetik örneklerin hem LightGBM ve XGBoost ile hem de LSTM birlikte uygulandığında performans düzeyinde katkılar sağlayabildiğini göstermektedir. Ayrıca TF-IDF yönteminin genel doğruluk açısından, Word2Vec yönteminin ise sınıf dengesine duyarlı ölçütler bakımından avantaj sağladığı gözlemlenmiştir. Uzman görüşü temel alınarak modele dahil edilen üç ek özniteliğin performans ölçütlerine katkısı çok sınırlı (%1-2 seviyesinde) kalmıştır.

5. SONUÇ

Hukuki kararların sistematik olarak sınıflandırılması, hem karar vericilere destek sağlayacak öngörücü araçların geliştirilmesi hem de yargı süreçlerinin şeffaflık ve hesap verebilirlik ilkeleri çerçevesinde iyileştirilmesi açısından büyük önem taşımaktadır. Bu tez çalışması, Türkçe yüksek mahkeme kararlarının sınıflandırılmasına yönelik literatürdeki mevcut boşlukları dikkate alarak literatüre katkı sunmayı amaçlamıştır. Bu tez çalışmasında, Türkçe yüksek mahkeme kararlarının onama ve bozma olarak sınıflandırılmasına yönelik kapsamlı bir doğal dil işleme ve makine öğrenmesi yaklaşımı geliştirilmiştir. Literatürde sıkça karşılaşılan sınıf dengesizliği problemine yönelik olarak bu çalışmada, klasik yeniden örnekleme yöntemlerine ek olarak üretken yapay zekâ tabanlı sentetik veri üretiminin etkinliği araştırılmıştır. Özellikle büyük dil modeli GPT-3.5-turbo kullanılarak oluşturulan sentetik örneklerin, geleneksel yöntemlere kıyasla sınıflandırma performansını anlamlı ölçüde artırdığı belirlenmiştir. Deneylerde, derin öğrenme mimarilerinden LSTM modelinin Word2Vec temsilleriyle üstün performans sergilediği, gradyan artırma tabanlı modeller olan XGBoost ve LightGBM'nin ise özellikle TF-IDF temsilleriyle daha başarılı sonuçlar sunduğu gözlenmiştir. Bu çalışma, söz konusu bağlamsal özniteliklerin hukuki karar sınıflandırmasındaki önemini sistematik biçimde ortaya koyarak literatüre katkı sağlamaktadır.

Sonuç olarak, tez çalışması literatüre üç yönlü katkı sağlamaktadır. Birincisi, Türkçe yargı kararlarında sınıf dengesizliği problemine yönelik olarak üretken yapay zekâ destekli dengeleme stratejilerinin uygulanabilirliğini ve başarısını ortaya koymaktadır. İkincisi, farklı metin temsil yöntemleri ile çeşitli makine öğrenmesi modelleri arasındaki etkileşimleri kapsamlı bir deneysel tasarımla karşılaştırmalı olarak analiz etmektedir. Üçüncüsü ise, hukuk diline özgü uzman temelli özniteliklerin model başarısına katkısını istatistiksel olarak göstermekte ve bu sayede alan yazınına içerik-temelli bir açıklama zemini kazandırmaktadır.

Gelecek çalışmalarda, sınıflandırma görevine hukuki bağlamda neden-sonuç çıkarımı gibi ileri düzey görevlerin de entegre edilmesi, açıklanabilirlik odaklı yöntemler kullanılabilir ve BERT gibi önceden eğitilmiş hukuk-dil modelleriyle karşılaştırmalı analizler gerçekleştirilebilir. Ayrıca, karar metinlerinin yalnızca sonuç bölümleriyle

değil, gerekçe yapılarının da analizine dayalı çok katmanlı modellemeler, yargı sürecinin bütüncül biçimde anlaşılmasına katkı sunabilir.

KAYNAKLAR

Abraham, A. (2005). Artificial neural networks. In P. H. Sydenham & R. Thorn (Eds.), *Handbook of measuring system design* (pp. 901–908). John Wiley & Sons.

Adalet Bakanlığı Bilgi İşlem Dairesi Başkanlığı. (2015). *UYAP Mevzuat Programı yenilenen yazılımı ve tasarımı ile yayında*. https://uyap.gov.tr/amp/UYAP-Mevzuat-Programi-Yenilenen-Yazilimi-ve-Tasarimi-ile-Yayinda (Erişim tarihi: 25 Nisan 2025)

Adalet Bakanlığı. (2020). *UYAP Bilişim Sistemi genel bilgi*. https://uyap.gov.tr/Genel-Bilgi (Erişim tarihi: 25 Nisan 2025)

Ahmad, S., Harris, D., & Sahibzada, I. (2020). Understanding legal documents: Classification of rhetorical role of sentences using deep learning and natural language processing. In *2020 IEEE International Conference on Smart Computing (ICSC)* (pp. 464–467). IEEE. https://doi.org/10.1109/ICSC.2020.00089

Akça, O., Bayrak, G., Issifu, A. M., & Ganiz, M. C. (2022). Traditional machine learning and deep learning-based text classification for Turkish law documents using transformers and domain adaptation. In *2022 International Conference on Innovations in Intelligent Systems and Applications (INISTA)* (pp. 1–6). IEEE. https://doi.org/10.1109/INISTA55318.2022.9894051

Akın, A. A., & Akın, M. D. (2007). Zemberek, an open source NLP framework for Turkic languages. *Structure, 10*, 1–5.

Al Katat, S., Zaki, C., Hazimeh, H., El Bitar, I., Angarita, R., & Trojman, L. (2024). Natural language processing for Arabic sentiment analysis: A systematic literature review. *IEEE Transactions on Big Data, 10*(5), 576–594. https://doi.org/10.1109/TBDATA.2024.3366083

Aletras, N., Tsarapatsanis, D., Preoţiuc-Pietro, D., & Lampos, V. (2016). Predicting judicial decisions of the European Court of Human Rights: A natural language processing perspective. *PeerJ Computer Science, 2*, e93. https://doi.org/10.7717/peerj-cs.93

Almuzaini, H. A., & Azmi, A. M. (2023). TaSbeeb: A judicial decision support system based on deep learning framework. *Journal of King Saud University – Computer and Information Sciences, 35*(8), 101695. https://doi.org/10.1016/j.jksuci.2023.101695

Ariai, F., & Demartini, G. (2025). *Natural language processing for the legal domain: A survey of tasks, datasets, models, and challenges*. arXiv. https://arxiv.org/abs/2410.21306

Arslan, A., Talaş, U., & Çubukçu, B. (2024). *LAWNAV: Yapay zekâ destekli hukuk danışmanı yönlendirme uygulaması*. ResearchGate. https://www.researchgate.net/publication/384500931_LAWNAV_YAPAY_ZEKA_DESTEKLİ_HUKUK_DANIŞMANI_YÖNLENDIRME_UYGULAMASI

Ash, E., & Marian, O. Y. (2020). The making of international tax law: Empirical evidence from tax treaties text. *Florida Tax Review, 24*(1). UC Irvine School of Law Research Paper No. 2019-02.

Barros de Jesus, M., Goes, A. da S., Santiago, L. de G., Xavier, M. A. P., & Bevilacqua, S. (2023). Artificial intelligence on legal language processing: Using deep learning to find the regulatory law framework for the third sector. *Revista do Serviço Público, 74*(2), 439–461. https://doi.org/10.21874/rsp.v74i2.8091

Beliveau, V., Kaas, H., Prener, M., Ladefoged, C. N., Elliott, D., Knudsen, G. M., Pinborg, L. H., & Ganz, M. (2024). *Classification of radiological text in small and imbalanced datasets in a non-English language*. arXiv. https://arxiv.org/abs/2409.20147

Betancur Sánchez, D., Aldama García, N., Barbero Jiménez, Á., Guerrero Nieto, M., Marsà Morales, P., Serrano Salas, N., García Hernán, C., Haya Coll, P., Montiel Ponsoda, E., & Calleja Ibáñez, P. (2025). *MEL: Legal Spanish language model*. arXiv. https://arxiv.org/abs/2501.16011

Bhattacharya, P., Paul, S., Ghosh, K., Ghosh, S., & Wyner, A. (2021). DeepRhole: Deep learning for rhetorical role labeling of sentences in legal case documents. *Artificial Intelligence and Law, 31*, 1–38. https://doi.org/10.1007/s10506-021-09304-5

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. Wadsworth.

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, I., & Amodei, D. (2020). *Language models are few-shot learners*. *Advances in Neural Information Processing Systems, 33*, 1877–1901. https://proceedings.neurips.cc/paper/2020/hash/1457c0d6bfcb4967418bfb8ac142f64a-Abstract.html

Busker, T., Choenni, S., & Bargh, M. S. (2025). Exploiting GPT for synthetic data generation: An empirical study. *Government Information Quarterly, 42*, 101988. https://doi.org/10.1016/j.giq.2024.101988

Cervantes, J., Garcia-Lamont, F., Rodríguez-Mazahua, L., & Lopez, A. (2020). A comprehensive survey on support vector machine classification: Applications, challenges and trends. *Neurocomputing, 408*, 189–215. https://doi.org/10.1016/j.neucom.2019.10.118

Chawla, N. V. (2005). Data mining for imbalanced datasets: An overview. In O. Maimon & L. Rokach (Eds.), *Data mining and knowledge discovery handbook*. Springer. https://doi.org/10.1007/0-387-25465-X_40

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research, 16*, 321–357. https://doi.org/10.1613/jair.953

Chen, H., Wu, L., Chen, J., Lu, W., & Ding, J. (2022). A comparative study of automated legal text classification using random forests and deep learning. *Information Processing & Management, 59*(2), 102798. https://doi.org/10.1016/j.ipm.2021.102798

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). ACM. https://doi.org/10.1145/2939672.2939785

Chen, Z. Z., Ma, J., Zhang, X., Hao, N., Yan, A., Nourbakhsh, A., Yang, X., McAuley, J., Petzold, L., & Wang, W. Y. (2024). *A survey on large language models for critical societal domains: Finance, healthcare, and law*. arXiv. https://arxiv.org/abs/2405.01769

Demsar, J. (2006). Statistical comparisons of classifiers over multiple data sets. *Journal of Machine Learning Research, 7*, 1–30.

Ding, B., Qin, C., Zhao, R., Luo, T., Li, X., Chen, G., Xia, W., Hu, J., Luu, A. T., & Joty, S. (2024). *Data augmentation using large language models: Data perspectives, learning paradigms and challenges*. arXiv. https://arxiv.org/abs/2403.02990

Fix, E., & Hodges, J. L. (1951). *Discriminatory analysis, nonparametric discrimination: Consistency properties* (Technical Report No. 4). USAF School of Aviation Medicine.

Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association, 32*(200), 675–701.

Fudickar, S., Bantel, C., Spieker, J., Töpfer, H., Stegeman, P., Schiphorst Preuper, H. R., Reneman, M. F., Wolff, A. P., & Soer, R. (2024). Natural language processing of referral letters for machine learning–based triaging of patients with low back pain to the most appropriate intervention: Retrospective study. *Journal of Medical Internet Research, 26*, e46857. https://doi.org/10.2196/46857

Fürnkranz, J. (2011). Decision tree. In C. Sammut & G. I. Webb (Eds.), *Encyclopedia of machine learning* (pp. 263–267). Springer. https://doi.org/10.1007/978-0-387-30164-8_196

García, V., Mollineda, R. A., & Sánchez, J. S. (2009). Index of balanced accuracy: A performance measure for skewed class distributions. In H. Araujo, A. M. Mendonça, A. J. Pinho, & M. I. Torres (Eds.), *Pattern recognition and image analysis. IbPRIA 2009. Lecture notes in computer science* (Vol. 5524, pp. 441–448). Springer. https://doi.org/10.1007/978-3-642-02172-5_57

García-Méndez, S., de Arriba-Pérez, F., Barros-Vila, A., & González-Castaño, F. J. (2022). Detection of temporality at discourse level on financial news by combining natural language processing and machine learning. *Expert Systems with Applications, 197*, 116648. https://doi.org/10.1016/j.eswa.2022.116648

Gorovaia, N., & Makrominas, M. (2025). Identifying greenwashing in corporate-social responsibility reports using natural-language processing. *European Financial Management, 31*(1), 427–462. https://doi.org/10.1111/eufm.12509

Görentaş, M. B., & Uçkan, T. (2023). Makine öğrenmesi yöntemleri kullanılarak mahkeme kararlarının kümelenmesi. *Computer Science, 8*(2), 148–158. https://doi.org/10.53070/bbd.1318518

Görentaş, M. B., Uçkan, T., & Bayram Arlı, N. (2023). Uyuşmazlık Mahkemesi kararlarının makine öğrenmesi yöntemleri ile sınıflandırılması. *Yüzüncü Yıl Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 28*(3), 947–961. https://doi.org/10.53433/yyufbed.1292275

Graves, A. (2012). Long short-term memory. In *Supervised sequence labelling with recurrent neural networks* (pp. 37–45). *Studies in Computational Intelligence* (Vol. 385). Springer. https://doi.org/10.1007/978-3-642-24797-2_4

Hasanin, T., & Khoshgoftaar, T. (2018). The effects of random undersampling with simulated class imbalance for big data. In *2018 IEEE International Conference on Information Reuse and Integration (IRI)* (pp. 70–79). IEEE. https://doi.org/10.1109/IRI.2018.00018

He, H., & Garcia, E. A. (2009). Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering, 21*(9), 1263–1284. https://doi.org/10.1109/TKDE.2008.239

Hearst, M. A., Dumais, S. T., Osuna, E., Platt, J., & Schölkopf, B. (1998). Support vector machines. *IEEE Intelligent Systems and Their Applications, 13*(4), 18–28. https://doi.org/10.1109/5254.708428

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation, 9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Hu, Z., Dychka, I., Potapova, K., & Meliukh, V. (2024). Augmenting sentiment analysis prediction in binary text classification through advanced natural language processing models and classifiers. *International Journal of Information Technology and Computer Science, 16*, 16–31.

Karabel, E. G., & Aydemir, D. (2023). Medeni usul hukukunda yargılamanın hızlandırılması ve adalete erişim hakkı bakımından çevrimiçi yargılama ve yapay zekânın kullanımı. *Marmara Üniversitesi Hukuk Fakültesi Hukuk Araştırmaları Dergisi, 29*(1), 530–573. https://doi.org/10.33433/maruhad.1263609

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017). LightGBM: A highly efficient gradient boosting decision tree. In *Advances in Neural Information Processing Systems* (Vol. 30). Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf

Kim, S. W., & Gil, J. M. (2019). Research paper classification systems based on TF-IDF and LDA schemes. *Human-Centric Computing and Information Sciences, 9*, 30. https://doi.org/10.1186/s13673-019-0192-7

Kite-Jackson, D. W. (2024). *2023 Artificial Intelligence (AI) TechReport*. American Bar Association. https://www.americanbar.org/groups/law_practice/resources/tech-report/2023/2023-artificial-intelligence-ai-techreport/ (Erişim tarihi: 26 Nisan 2025)

Kotsiantis, S., Kanellopoulos, D., & Pintelas, P. (2005). Handling imbalanced datasets: A review. *GESTS International Transactions on Computer Science and Engineering, 30*, 25–36.

Köksal, A. (2018). *Pre-trained Word2Vec model for Turkish* [GitHub repository]. GitHub. https://github.com/akoksal/Turkish-Word2Vec (Erişim tarihi: 7 Mayıs 2025)

Lai, J., Gan, W., Wu, J., Qi, Z., & Yu, P. S. (2024). Large language models in law: A survey. *AI Open, 5*, 181–196. https://doi.org/10.1016/j.aiopen.2024.09.002

Lam, J., Chen, Y., Zulkernine, F., & Dahan, S. (2025). Legal text analytics for reasonable notice period prediction. *Journal of Computational and Cognitive Engineering*. https://doi.org/10.47852/bonviewJCCE52024104

Liu, X., He, P., Chen, W., & Gao, J. (2019). Multi-task deep neural networks for natural language understanding. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (pp. 4487–4496). Association for Computational Linguistics. https://doi.org/10.18653/v1/P19-1441

Liu, Y., Wang, Y., & Zhang, J. (2012). New machine learning algorithm: Random forest. In B. Liu, M. Ma, & J. Chang (Eds.), *ICICA 2012: Lecture notes in computer science* (Vol. 7473, pp. 246–252). Springer-Verlag. https://doi.org/10.1007/978-3-642-34062-8_32

Michel, M., Djurica, D., & Mendling, J. (2022). Identification of decision rules from legislative documents using machine learning and natural language processing. In *Proceedings of the 55th Hawaii International Conference on System Sciences* (pp. 6247–6256). University of Hawai‘i at Mānoa. https://doi.org/10.24251/HICSS.2022.757

Microsoft Corporation. (2025). *LightGBM documentation*. https://lightgbm.readthedocs.io/ (Erişim tarihi: 16 Mayıs 2025)

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013a). *Efficient estimation of word representations in vector space* [arXiv preprint]. arXiv. https://arxiv.org/abs/1301.3781

Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013b). Distributed representations of words and phrases and their compositionality. In *Advances in Neural Information Processing Systems, 26* (pp. 3111–3119). Curran Associates, Inc. https://papers.nips.cc/paper_files/paper/2013/hash/9aa42b31882ec039965f3c4923ce901b-Abstract.html

Mikolov, T., Yih, W.-T., & Zweig, G. (2013c). Linguistic regularities in continuous space word representations. In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)* (pp. 746–751). Association for Computational Linguistics. https://aclanthology.org/N13-1090

Mishra, S. (2017). Handling imbalanced data: SMOTE vs. random undersampling. *International Research Journal of Engineering and Technology, 4*(8), 317–320.

Morgan, J. N., & Sonquist, J. A. (1963). Problems in the analysis of survey data, and a proposal. *Journal of the American Statistical Association*, *58*(302), 415–434. https://doi.org/10.2307/2283276

Mumcuoğlu, E., Öztürk, C. E., Ozaktas, H. M., & Koç, A. (2021). Natural language processing in law: Prediction of outcomes in the higher courts of Turkey. *Information Processing & Management, 58*(5), 102684. https://doi.org/10.1016/j.ipm.2021.102684

Nemenyi, P. B. (1963). *Distribution-free multiple comparisons* (Doctoral dissertation). Princeton University.

Okursoy, M. E., & İnkaya, T. (2025). Yüksek Mahkeme kararlarının sınıflandırılması için veri dengeleme ve açıklanabilir yapay zekâ tabanlı karar destek sistemi. *Veri Bilimi,* 8(1), 11–23. https://dergipark.org.tr/tr/pub/veri/issue/92570/1557010

Ong, K., Mao, R., Satapathy, R., Shirota Filho, R., Cambria, E., Sulaeman, J., & Mengaldo, G. (2025). Explainable natural language processing for corporate sustainability analysis. *Information Fusion, 115*, Article 102726. https://doi.org/10.1016/j.inffus.2024.102726

Pajila, P. J. B., Sudha, K., Kalai Selvi, D. M., Kumar, V. N., Gayathri, S., & Subramanian, R. S. (2023). A survey on natural language processing and its applications. In *Proceedings of the 4th International Conference on Electronics and Sustainable Communication Systems (ICESC)* (pp. 996–1001). IEEE. https://doi.org/10.1109/ICESC57686.2023.10193469

Parmar, A., Katariya, R., & Patel, V. (2019). A review on random forest: An ensemble classifier. In J. Hemanth et al. (Eds.), *ICICI 2018: Lecture notes in data engineering*

and communications technologies (Vol. 26, pp. 758–763). Springer Nature Switzerland AG. https://doi.org/10.1007/978-3-030-03146-6_86

Quinlan, J. R. (1993). *C4.5: Programs for machine learning*. Morgan Kaufmann.

Rout, N., Mishra, D., & Mallick, M. K. (2018). Handling imbalanced data: A survey. In M. Reddy, K. Viswanath, & S. K. M. (Eds.), *International proceedings on advances in soft computing, intelligent systems and applications*. *Advances in Intelligent Systems and Computing* (Vol. 628, pp. 431–441). Springer. https://doi.org/10.1007/978-981-10-5272-9_39

Schölkopf, B., & Smola, A. J. (2001). *Learning with kernels: Support vector machines, regularization, optimization, and beyond*. The MIT Press. https://doi.org/10.7551/mitpress/4175.001.0001

Sharma, A., & Feldman, D. I. (2023). Team Cadence at MEDIQA-Sum 2023: Using ChatGPT as a data augmentation tool for classifying clinical dialogue. In *CLEF 2023: Conference and Labs of the Evaluation Forum* (September 18–21, 2023, Thessaloniki, Greece). *CEUR Workshop Proceedings*. http://ceur-ws.org

Siino, M., Falco, M., Croce, D., & Rosso, P. (2025). Exploring LLMs applications in law: A literature review on current legal NLP approaches. *IEEE Access, 13*, 18253–18276. https://doi.org/10.1109/ACCESS.2025.3533217

Suthaharan, S. (2016). Decision tree learning. In *Machine learning models and algorithms for big data classification* (pp. 209–235). *Integrated Series in Information Systems* (Vol. 36). Springer. https://doi.org/10.1007/978-1-4899-7641-3_10

Suthaharan, S. (2016). *Machine learning models and algorithms for big data classification: Thinking with examples for effective learning* (Vol. 36). Springer Science+Business Media. https://doi.org/10.1007/978-1-4899-7641-3

Szeto, A., & Das, A. N. (2024). Classification of Notices to Airmen using natural language processing. In *AIAA SCITECH 2024 Forum* (p. 2585). https://doi.org/10.2514/6.2024-2585

Taunk, K., De, S., Verma, S., & Swetapadma, A. (2019). A brief review of nearest neighbor algorithm for learning and classification. In *2019 International Conference on Intelligent Computing and Control Systems (ICCS)* (pp. 1255–1260). IEEE. https://doi.org/10.1109/ICCS45141.2019.9065747

Thapa, S., Rauniyar, K., Jafri, F. A., Adhikari, S., Sarveswaran, K., Bal, B. K., Veeramani, H., & Naseem, U. (2025). Natural language understanding of Devanagari script languages: Language identification, hate speech and its target detection. In K. Sarveswaran, A. Vaidya, B. K. Bal, S. Shams, & S. Thapa (Eds.), *Proceedings of the First Workshop on Challenges in Processing South Asian Languages (CHiPSAL 2025)* (pp. 71–82). International Committee on Computational Linguistics. https://aclanthology.org/2025.chipsal-1.7/

Thomson Reuters. (2025). *Balancing innovation and caution: How lawyers should integrate AI into legal practice*. https://www.reuters.com/legal/legalindustry/balancing-innovation-caution-how-lawyers-should-integrate-ai-into-legal-practice-2025-04-23/ (Erişim tarihi: 26 Nisan 2025)

Turan, T., Küçüksille, E. U., & Kemaloğlu, N. (2023). Prediction of Turkish Constitutional Court decisions with explainable artificial intelligence. *Bilge*

International Journal of Science and Technology Research, 7. https://doi.org/10.30516/bilgesci.1317525

ValizadehAslani, T., Shi, Y., Wang, J., Ren, P., Zhang, Y., Hu, M., Zhao, L., & Liang, H. (2024). Two-stage fine-tuning with ChatGPT data augmentation for learning class-imbalanced data. *Neurocomputing, 592*, 127801. https://doi.org/10.1016/j.neucom.2024.127801

Van Houdt, G., Mosquera, C., & Nápoles, G. (2020). A review on the long short-term memory model. *Artificial Intelligence Review, 53*(8), 5929–5955. https://doi.org/10.1007/s10462-020-09838-1

Wajeed, M. A., Quadry, K. M., Mallik, M. A., & Khanna, K. R. (2025). Explainable machine learning for text classification: A novel approach to transparency and interpretability. *Journal of Theoretical and Applied Information Technology, 103*(3).

Wang, L., Wang, Y., Ni, B., Zhao, Y., Wang, H., Ma, Y., & Derr, T. (2025). *SaVe-TAG: Semantic-aware vicinal risk minimization for long-tailed text-attributed graphs*. arXiv. https://arxiv.org/abs/2410.16882

Wang, S., Dai, Y., Shen, J., & Xuan, J. (2021). Research on expansion and classification of imbalanced data based on SMOTE algorithm. *Scientific Reports, 11*(1), 24039. https://doi.org/10.1038/s41598-021-03430-5

XGBoost Developers. (2025). *XGBoost parameters* (Version 3.0.1) [Çevrimiçi dokümantasyon]. https://xgboost.readthedocs.io/en/stable/parameter.html

Yang, G. R., & Wang, X. J. (2020). Artificial neural networks for neuroscientists: A primer. *Neuron, 107*(6), 1048–1070. https://doi.org/10.1016/j.neuron.2020.09.005

Zhang, W., Yoshida, T., & Tang, X. (2011). A comparative study of TF*IDF, LSI and multi-words for text classification. *Expert Systems with Applications, 38*(3), 2758–2765. https://doi.org/10.1016/j.eswa.2010.08.066

Ziems, N., & Wu, S. (2021). Security vulnerability detection using deep learning natural language processing. In *IEEE INFOCOM 2021 – IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)* (pp. 1–6). IEEE. https://doi.org/10.1109/INFOCOMWKSHPS51825.2021.9484500

EKLER

EK 1 KNN Sınıflandırma Modeli Sonuçları

EK 2 DT Sınıflandırma Modeli Sonuçları

EK 3 RF Sınıflandırma Modeli Sonuçları

EK 4 SVM Sınıflandırma Modeli Sonuçları

EK 5 XGBOOST Sınıflandırma Modeli Sonuçları

EK 6 LightGBM Sınıflandırma Modeli Sonuçları

EK 7 ANN Sınıflandırma Modeli Sonuçları

EK 8 LSTM Sınıflandırma Modeli Sonuçları

EK 9 Nemenyi Kritik Fark Diyagramları

EK 1. KNN Sınıflandırma Modeli Sonuçları

VDY	KTY	EÖ	A	P	R	F1	GM	IBA	TPR	TNR
CHATGPT (%100)	TF-IDF	Var	0,891	0,834	0,562	0,671	0,739	0,524	0,562	0,972
CHATGPT (%100)	TF-IDF	Yok	0,894	0,879	0,541	0,670	0,729	0,508	0,541	0,981
CHATGPT (%100)	Word2Vec	Var	0,877	0,777	0,532	0,632	0,716	0,490	0,532	0,962
CHATGPT (%100)	Word2Vec	Yok	0,869	0,739	0,529	0,616	0,710	0,483	0,529	0,954
Hibrit (RUS 1:2, SMOTE 100%)	TF-IDF	Var	0,186	0,115	0,941	0,203	0,226	0,068	0,941	0,092
Hibrit (RUS 1:2, SMOTE 100%)	TF-IDF	Yok	0,171	0,117	0,995	0,209	0,260	0,074	0,995	0,068
Hibrit (RUS 1:2, SMOTE 100%)	Word2Vec	Var	0,305	0,133	0,954	0,233	0,462	0,230	0,954	0,225
Hibrit (RUS 1:2, SMOTE 100%)	Word2Vec	Yok	0,304	0,132	0,954	0,233	0,461	0,229	0,954	0,224
Hibrit (RUS 1:2, SMOTE 200%)	TF-IDF	Var	0,155	0,113	0,968	0,202	0,199	0,051	0,968	0,054
Hibrit (RUS 1:2, SMOTE 200%)	TF-IDF	Yok	0,160	0,116	0,996	0,208	0,235	0,061	0,996	0,056
Hibrit (RUS 1:2, SMOTE 200%)	Word2Vec	Var	0,253	0,126	0,971	0,223	0,398	0,172	0,971	0,164
Hibrit (RUS 1:2, SMOTE 200%)	Word2Vec	Yok	0,254	0,126	0,970	0,223	0,399	0,172	0,970	0,165
Dengeleme Yok	TF-IDF	Var	0,898	0,843	0,093	0,167	0,303	0,085	0,093	0,998
Dengeleme Yok	TF-IDF	Yok	0,909	0,730	0,286	0,409	0,530	0,263	0,286	0,987
Dengeleme Yok	Word2Vec	Var	0,892	0,555	0,244	0,336	0,485	0,220	0,244	0,972
Dengeleme Yok	Word2Vec	Yok	0,894	0,551	0,248	0,340	0,491	0,224	0,248	0,974
RUS (1:1)	TF-IDF	Var	0,201	0,118	0,959	0,210	0,277	0,096	0,959	0,107
RUS (1:1)	TF-IDF	Yok	0,681	0,232	0,807	0,359	0,732	0,544	0,807	0,666
RUS (1:1)	Word2Vec	Var	0,670	0,208	0,710	0,322	0,687	0,474	0,710	0,665
RUS (1:1)	Word2Vec	Yok	0,663	0,206	0,717	0,320	0,686	0,473	0,717	0,657
RUS (1:1,5)	TF-IDF	Var	0,397	0,224	0,800	0,221	0,421	0,201	0,800	0,347
RUS (1:1,5)	TF-IDF	Yok	0,776	0,293	0,718	0,415	0,749	0,558	0,718	0,783
RUS (1:1,5)	Word2Vec	Var	0,756	0,250	0,601	0,352	0,682	0,458	0,601	0,775

EK 1. KNN Sınıflandırma Modeli Sonuçları

VDY	KTY	EÖ	A	P	R	F1	GM	IBA	TPR	TNR
RUS (1:1,5)	Word2Vec	Yok	0,750	0,244	0,601	0,347	0,679	0,454	0,601	0,769
RUS (1:2)	TF-IDF	Var	0,568	0,322	0,636	0,237	0,504	0,275	0,636	0,560
RUS (1:2)	TF-IDF	Yok	0,829	0,351	0,632	0,450	0,733	0,527	0,632	0,853
RUS (1:2)	Word2Vec	Var	0,805	0,290	0,526	0,373	0,664	0,427	0,526	0,839
RUS (1:2)	Word2Vec	Yok	0,799	0,282	0,523	0,366	0,660	0,423	0,523	0,834
SMOTE 100%	TF-IDF	Var	0,366	0,124	0,760	0,210	0,461	0,225	0,760	0,318
SMOTE 100%	TF-IDF	Yok	0,325	0,136	0,954	0,238	0,484	0,251	0,954	0,246
SMOTE 100%	Word2Vec	Var	0,565	0,183	0,849	0,301	0,670	0,464	0,849	0,529
SMOTE 100%	Word2Vec	Yok	0,563	0,182	0,843	0,299	0,667	0,460	0,843	0,529
SMOTE 200%	TF-IDF	Var	0,272	0,120	0,885	0,211	0,396	0,170	0,885	0,196
SMOTE 200%	TF-IDF	Yok	0,290	0,131	0,967	0,231	0,446	0,214	0,967	0,206
SMOTE 200%	Word2Vec	Var	0,484	0,164	0,898	0,278	0,623	0,406	0,898	0,432
SMOTE 200%	Word2Vec	Yok	0,480	0,163	0,896	0,276	0,619	0,402	0,896	0,429
SMOTE 300%	TF-IDF	Var	0,228	0,120	0,948	0,213	0,353	0,137	0,948	0,138
SMOTE 300%	TF-IDF	Yok	0,275	0,129	0,972	0,228	0,427	0,197	0,972	0,188
SMOTE 300%	Word2Vec	Var	0,442	0,156	0,922	0,267	0,593	0,371	0,922	0,383
SMOTE 300%	Word2Vec	Yok	0,437	0,155	0,921	0,266	0,589	0,366	0,921	0,377
SMOTE 600%	TF-IDF	Var	0,201	0,120	0,986	0,214	0,319	0,111	0,986	0,104
SMOTE 600%	TF-IDF	Yok	0,257	0,128	0,981	0,226	0,404	0,177	0,981	0,167
SMOTE 600%	Word2Vec	Var	0,381	0,146	0,946	0,252	0,541	0,312	0,946	0,311
SMOTE 600%	Word2Vec	Yok	0,379	0,145	0,947	0,252	0,540	0,310	0,947	0,308

VDY: Veri dengeleme yöntemi, KTY: Kelime temsili yöntemi, EÖ: Ek öznitelikler, A: Doğruluk (Accuracy), P: Kesinlik (Precision), R: Duyarlılık (Recall), F1: F-ölçütü, GM: Geometrik ortalama (G-Mean), IBA: Index of Balanced Accuracy, TNR: True Negative Rate, TPR: True Positive Rate

EK 2. DT Sınıflandırma Modeli Sonuçları

VDY	KTY	EÖ	A	P	R	F1	GM	IBA	TPR	TNR
CHATGPT (%100)	TF-IDF	Var	0,897	0,860	0,577	0,689	0,750	0,541	0,577	0,976
CHATGPT (%100)	TF-IDF	Yok	0,897	0,860	0,577	0,689	0,750	0,541	0,577	0,976
CHATGPT (%100)	Word2Vec	Var	0,882	0,854	0,491	0,623	0,693	0,457	0,491	0,979
CHATGPT (%100)	Word2Vec	Yok	0,883	0,881	0,478	0,619	0,685	0,446	0,478	0,984
Hibrit (RUS 1:2, SMOTE 100%)	TF-IDF	Var	0,735	0,251	0,696	0,368	0,716	0,512	0,696	0,740
Hibrit (RUS 1:2, SMOTE 100%)	TF-IDF	Yok	0,738	0,252	0,681	0,366	0,711	0,503	0,681	0,745
Hibrit (RUS 1:2, SMOTE 100%)	Word2Vec	Var	0,569	0,153	0,638	0,246	0,597	0,360	0,638	0,561
Hibrit (RUS 1:2, SMOTE 100%)	Word2Vec	Yok	0,579	0,155	0,626	0,248	0,598	0,361	0,626	0,574
Hibrit (RUS 1:2, SMOTE 200%)	TF-IDF	Var	0,715	0,241	0,724	0,361	0,718	0,516	0,724	0,714
Hibrit (RUS 1:2, SMOTE 200%)	TF-IDF	Yok	0,712	0,237	0,720	0,356	0,714	0,511	0,720	0,711
Hibrit (RUS 1:2, SMOTE 200%)	Word2Vec	Var	0,523	0,148	0,693	0,243	0,589	0,354	0,693	0,502
Hibrit (RUS 1:2, SMOTE 200%)	Word2Vec	Yok	0,520	0,149	0,711	0,246	0,593	0,360	0,711	0,496
Dengeleme Yok	TF-IDF	Var	0,898	0,587	0,272	0,368	0,513	0,247	0,272	0,976
Dengeleme Yok	TF-IDF	Yok	0,898	0,582	0,273	0,368	0,514	0,247	0,273	0,975
Dengeleme Yok	Word2Vec	Var	0,873	0,290	0,104	0,152	0,314	0,092	0,104	0,968
Dengeleme Yok	Word2Vec	Yok	0,873	0,297	0,105	0,154	0,318	0,093	0,105	0,969
RUS (1:1)	TF-IDF	Var	0,709	0,234	0,699	0,349	0,704	0,495	0,699	0,710
RUS (1:1)	TF-IDF	Yok	0,709	0,234	0,697	0,349	0,703	0,494	0,697	0,711
RUS (1:1)	Word2Vec	Var	0,598	0,159	0,611	0,252	0,602	0,363	0,611	0,596
RUS (1:1)	Word2Vec	Yok	0,598	0,159	0,614	0,252	0,603	0,365	0,614	0,596
RUS (1:1,5)	TF-IDF	Var	0,780	0,283	0,628	0,388	0,707	0,492	0,628	0,798
RUS (1:1,5)	TF-IDF	Yok	0,778	0,282	0,631	0,387	0,707	0,493	0,631	0,796
RUS (1:1,5)	Word2Vec	Var	0,685	0,169	0,471	0,248	0,577	0,326	0,471	0,712

EK 2. DT Sınıflandırma Modeli Sonuçları

VDY	KTY	EÖ	A	P	R	F1	GM	IBA	TPR	TNR
RUS (1:1,5)	Word2Vec	Yok	0,686	0,170	0,470	0,249	0,577	0,326	0,470	0,713
RUS (1:2)	TF-IDF	Var	0,839	0,355	0,496	0,407	0,658	0,419	0,496	0,882
RUS (1:2)	TF-IDF	Yok	0,843	0,362	0,487	0,409	0,655	0,413	0,487	0,887
RUS (1:2)	Word2Vec	Var	0,750	0,195	0,406	0,263	0,565	0,308	0,406	0,793
RUS (1:2)	Word2Vec	Yok	0,746	0,191	0,401	0,258	0,560	0,303	0,401	0,788
SMOTE 100%	TF-IDF	Var	0,884	0,474	0,328	0,383	0,557	0,292	0,328	0,953
SMOTE 100%	TF-IDF	Yok	0,881	0,456	0,322	0,373	0,551	0,287	0,322	0,951
SMOTE 100%	Word2Vec	Var	0,828	0,238	0,253	0,241	0,471	0,212	0,253	0,900
SMOTE 100%	Word2Vec	Yok	0,808	0,221	0,290	0,249	0,501	0,238	0,290	0,873
SMOTE 200%	TF-IDF	Var	0,848	0,358	0,449	0,394	0,632	0,384	0,449	0,897
SMOTE 200%	TF-IDF	Yok	0,844	0,352	0,456	0,392	0,635	0,388	0,456	0,892
SMOTE 200%	Word2Vec	Var	0,763	0,205	0,391	0,267	0,561	0,303	0,391	0,809
SMOTE 200%	Word2Vec	Yok	0,768	0,211	0,395	0,274	0,566	0,308	0,395	0,814
SMOTE 300%	TF-IDF	Var	0,825	0,325	0,530	0,401	0,674	0,441	0,530	0,862
SMOTE 300%	TF-IDF	Yok	0,827	0,330	0,540	0,408	0,681	0,451	0,540	0,863
SMOTE 300%	Word2Vec	Var	0,717	0,189	0,473	0,270	0,594	0,344	0,473	0,748
SMOTE 300%	Word2Vec	Yok	0,710	0,186	0,477	0,267	0,593	0,343	0,477	0,739
SMOTE 600%	TF-IDF	Var	0,793	0,287	0,579	0,382	0,688	0,462	0,579	0,819
SMOTE 600%	TF-IDF	Yok	0,795	0,295	0,598	0,394	0,700	0,479	0,598	0,820
SMOTE 600%	Word2Vec	Var	0,632	0,166	0,577	0,257	0,606	0,366	0,577	0,639
SMOTE 600%	Word2Vec	Yok	0,636	0,169	0,584	0,262	0,612	0,373	0,584	0,642

VDY: Veri dengeleme yöntemi, KTY: Kelime temsili yöntemi, EÖ: Ek öznitelikler, A: Doğruluk (Accuracy), P: Kesinlik (Precision), R: Duyarlılık (Recall), F1: F-ölçütü, GM: Geometrik ortalama (G-Mean), IBA: Index of Balanced Accuracy, TNR: True Negative Rate, TPR: True Positive Rate

EK 3. RF Sınıflandırma Modeli Sonuçları

VDY	KTY	EÖ	A	P	R	F1	GM	IBA	TPR	TNR
CHATGPT (%100)	TF-IDF	Var	0,898	0,961	0,508	0,664	0,711	0,481	0,508	0,995
CHATGPT (%100)	TF-IDF	Yok	0,898	0,962	0,506	0,663	0,710	0,479	0,506	0,995
CHATGPT (%100)	Word2Vec	Var	0,895	0,966	0,491	0,650	0,699	0,464	0,491	0,996
CHATGPT (%100)	Word2Vec	Yok	0,895	0,965	0,489	0,648	0,697	0,462	0,489	0,995
Hibrit (RUS 1:2, SMOTE 100%)	TF-IDF	Var	0,811	0,323	0,645	0,430	0,732	0,527	0,645	0,832
Hibrit (RUS 1:2, SMOTE 100%)	TF-IDF	Yok	0,808	0,317	0,641	0,424	0,728	0,521	0,641	0,828
Hibrit (RUS 1:2, SMOTE 100%)	Word2Vec	Var	0,671	0,206	0,695	0,318	0,680	0,465	0,695	0,668
Hibrit (RUS 1:2, SMOTE 100%)	Word2Vec	Yok	0,674	0,208	0,696	0,320	0,683	0,468	0,696	0,671
Hibrit (RUS 1:2, SMOTE 200%)	TF-IDF	Var	0,744	0,270	0,771	0,400	0,756	0,573	0,771	0,741
Hibrit (RUS 1:2, SMOTE 200%)	TF-IDF	Yok	0,744	0,269	0,768	0,399	0,754	0,570	0,768	0,741
Hibrit (RUS 1:2, SMOTE 200%)	Word2Vec	Var	0,508	0,161	0,822	0,270	0,621	0,399	0,822	0,469
Hibrit (RUS 1:2, SMOTE 200%)	Word2Vec	Yok	0,505	0,161	0,827	0,270	0,620	0,398	0,827	0,465
Dengeleme Yok	TF-IDF	Var	0,893	1,000	0,030	0,058	0,169	0,027	0,030	1,000
Dengeleme Yok	TF-IDF	Yok	0,893	0,980	0,030	0,058	0,169	0,027	0,030	1,000
Dengeleme Yok	Word2Vec	Var	0,887	0,537	0,005	0,009	0,049	0,004	0,005	0,996
Dengeleme Yok	Word2Vec	Yok	0,890	0,560	0,007	0,013	0,066	0,006	0,007	1,000
RUS (1:1)	TF-IDF	Var	0,790	0,296	0,650	0,406	0,723	0,516	0,650	0,807
RUS (1:1)	TF-IDF	Yok	0,791	0,298	0,654	0,409	0,726	0,520	0,654	0,808
RUS (1:1)	Word2Vec	Var	0,692	0,209	0,644	0,316	0,670	0,447	0,644	0,697
RUS (1:1)	Word2Vec	Yok	0,688	0,207	0,645	0,313	0,668	0,445	0,645	0,693
RUS (1:1,5)	TF-IDF	Var	0,854	0,373	0,470	0,415	0,650	0,405	0,470	0,902
RUS (1:1,5)	TF-IDF	Yok	0,852	0,366	0,467	0,410	0,648	0,402	0,467	0,899
RUS (1:1,5)	Word2Vec	Var	0,831	0,301	0,395	0,341	0,590	0,332	0,395	0,885

EK 3. RF Sınıflandırma Modeli Sonuçları

VDY	KTY	EÖ	A	P	R	F1	GM	IBA	TPR	TNR
RUS (1:1,5)	Word2Vec	Yok	0,830	0,301	0,401	0,343	0,594	0,337	0,401	0,883
RUS (1:2)	TF-IDF	Var	0,871	0,411	0,375	0,391	0,591	0,331	0,375	0,933
RUS (1:2)	TF-IDF	Yok	0,871	0,411	0,373	0,390	0,589	0,329	0,373	0,933
RUS (1:2)	Word2Vec	Var	0,874	0,399	0,260	0,313	0,496	0,230	0,260	0,950
RUS (1:2)	Word2Vec	Yok	0,872	0,389	0,259	0,309	0,495	0,229	0,259	0,948
SMOTE 100%	TF-IDF	Var	0,898	0,662	0,155	0,250	0,391	0,141	0,155	0,990
SMOTE 100%	TF-IDF	Yok	0,899	0,679	0,163	0,261	0,400	0,148	0,163	0,990
SMOTE 100%	Word2Vec	Var	0,893	0,708	0,058	0,106	0,237	0,052	0,058	0,997
SMOTE 100%	Word2Vec	Yok	0,893	0,726	0,058	0,107	0,238	0,053	0,058	0,997
SMOTE 200%	TF-IDF	Var	0,897	0,575	0,247	0,344	0,490	0,224	0,247	0,977
SMOTE 200%	TF-IDF	Yok	0,897	0,573	0,251	0,348	0,494	0,227	0,251	0,977
SMOTE 200%	Word2Vec	Var	0,889	0,498	0,154	0,234	0,388	0,139	0,154	0,980
SMOTE 200%	Word2Vec	Yok	0,890	0,515	0,151	0,232	0,384	0,136	0,151	0,982
SMOTE 300%	TF-IDF	Var	0,890	0,508	0,322	0,393	0,555	0,290	0,322	0,961
SMOTE 300%	TF-IDF	Yok	0,890	0,508	0,326	0,396	0,558	0,293	0,326	0,960
SMOTE 300%	Word2Vec	Var	0,872	0,388	0,268	0,316	0,503	0,237	0,268	0,947
SMOTE 300%	Word2Vec	Yok	0,871	0,380	0,260	0,308	0,496	0,230	0,260	0,947
SMOTE 600%	TF-IDF	Var	0,861	0,396	0,493	0,439	0,668	0,429	0,493	0,907
SMOTE 600%	TF-IDF	Yok	0,860	0,394	0,501	0,441	0,673	0,435	0,501	0,904
SMOTE 600%	Word2Vec	Var	0,770	0,251	0,547	0,344	0,660	0,425	0,547	0,797
SMOTE 600%	Word2Vec	Yok	0,774	0,257	0,550	0,350	0,664	0,430	0,550	0,802

VDY: Veri dengeleme yöntemi, KTY: Kelime temsili yöntemi, EÖ: Ek öznitelikler, A: Doğruluk (Accuracy), P: Kesinlik (Precision), R: Duyarlılık (Recall), F1: F-ölçütü, GM: Geometrik ortalama (G-Mean), IBA: Index of Balanced Accuracy, TNR: True Negative Rate, TPR: True Positive Rate

EK 4. SVM Sınıflandırma Modeli Sonuçları

VDY	KTY	EÖ	A	P	R	F1	GM	IBA	TPR	TNR
CHATGPT (%100)	TF-IDF	Var	0,916	0,892	0,657	0,756	0,802	0,623	0,657	0,980
CHATGPT (%100)	TF-IDF	Yok	0,916	0,894	0,656	0,756	0,802	0,623	0,656	0,981
CHATGPT (%100)	Word2Vec	Var	0,877	0,705	0,655	0,679	0,781	0,594	0,655	0,932
CHATGPT (%100)	Word2Vec	Yok	0,875	0,703	0,645	0,672	0,775	0,584	0,645	0,932
Hibrit (RUS 1:2, SMOTE 100%)	TF-IDF	Var	0,884	0,465	0,329	0,384	0,559	0,294	0,329	0,953
Hibrit (RUS 1:2, SMOTE 100%)	TF-IDF	Yok	0,883	0,476	0,571	0,518	0,725	0,507	0,571	0,921
Hibrit (RUS 1:2, SMOTE 100%)	Word2Vec	Var	0,831	0,340	0,557	0,422	0,693	0,467	0,557	0,866
Hibrit (RUS 1:2, SMOTE 100%)	Word2Vec	Yok	0,830	0,337	0,556	0,419	0,692	0,465	0,556	0,864
Hibrit (RUS 1:2, SMOTE 200%)	TF-IDF	Var	0,880	0,443	0,323	0,373	0,553	0,288	0,323	0,949
Hibrit (RUS 1:2, SMOTE 200%)	TF-IDF	Yok	0,883	0,476	0,566	0,516	0,722	0,503	0,566	0,922
Hibrit (RUS 1:2, SMOTE 200%)	Word2Vec	Var	0,832	0,341	0,558	0,423	0,695	0,469	0,558	0,866
Hibrit (RUS 1:2, SMOTE 200%)	Word2Vec	Yok	0,830	0,337	0,556	0,419	0,692	0,465	0,556	0,864
Dengeleme Yok	TF-IDF	Var	0,907	0,846	0,197	0,318	0,442	0,180	0,197	0,996
Dengeleme Yok	TF-IDF	Yok	0,918	0,861	0,313	0,458	0,557	0,290	0,313	0,994
Dengeleme Yok	Word2Vec	Var	0,904	0,676	0,308	0,421	0,548	0,282	0,308	0,978
Dengeleme Yok	Word2Vec	Yok	0,907	0,676	0,308	0,421	0,548	0,282	0,308	0,981
RUS (1:1)	TF-IDF	Var	0,596	0,194	0,843	0,316	0,690	0,489	0,843	0,565
RUS (1:1)	TF-IDF	Yok	0,793	0,321	0,775	0,453	0,785	0,616	0,775	0,796
RUS (1:1)	Word2Vec	Var	0,723	0,243	0,715	0,363	0,719	0,517	0,715	0,724
RUS (1:1)	Word2Vec	Yok	0,721	0,243	0,716	0,362	0,719	0,517	0,716	0,722
RUS (1:1,5)	TF-IDF	Var	0,748	0,265	0,709	0,385	0,730	0,531	0,709	0,753
RUS (1:1,5)	TF-IDF	Yok	0,853	0,400	0,659	0,497	0,760	0,565	0,659	0,877
RUS (1:1,5)	Word2Vec	Var	0,796	0,298	0,623	0,403	0,713	0,500	0,623	0,817

EK 4. SVM Sınıflandırma Modeli Sonuçları

VDY	KTY	EÖ	A	P	R	F1	GM	IBA	TPR	TNR
RUS (1:1,5)	Word2Vec	Yok	0,793	0,294	0,618	0,398	0,709	0,494	0,618	0,815
RUS (1:2)	TF-IDF	Var	0,876	0,461	0,442	0,442	0,638	0,390	0,442	0,930
RUS (1:2)	TF-IDF	Yok	0,881	0,472	0,581	0,520	0,730	0,516	0,581	0,919
RUS (1:2)	Word2Vec	Var	0,831	0,339	0,558	0,421	0,694	0,467	0,558	0,865
RUS (1:2)	Word2Vec	Yok	0,830	0,336	0,556	0,419	0,693	0,466	0,556	0,863
SMOTE 100%	TF-IDF	Var	0,907	0,821	0,200	0,321	0,445	0,183	0,200	0,995
SMOTE 100%	TF-IDF	Yok	0,918	0,862	0,309	0,454	0,553	0,286	0,309	0,994
SMOTE 100%	Word2Vec	Var	0,906	0,653	0,310	0,419	0,550	0,284	0,310	0,979
SMOTE 100%	Word2Vec	Yok	0,906	0,661	0,312	0,422	0,551	0,285	0,312	0,980
SMOTE 200%	TF-IDF	Var	0,906	0,803	0,199	0,318	0,444	0,182	0,199	0,994
SMOTE 200%	TF-IDF	Yok	0,918	0,865	0,306	0,452	0,551	0,284	0,306	0,994
SMOTE 200%	Word2Vec	Var	0,906	0,661	0,313	0,423	0,552	0,286	0,313	0,980
SMOTE 200%	Word2Vec	Yok	0,906	0,661	0,307	0,418	0,547	0,281	0,307	0,980
SMOTE 300%	TF-IDF	Var	0,905	0,782	0,198	0,315	0,442	0,181	0,198	0,993
SMOTE 300%	TF-IDF	Yok	0,918	0,862	0,304	0,449	0,549	0,282	0,304	0,994
SMOTE 300%	Word2Vec	Var	0,906	0,661	0,310	0,421	0,550	0,284	0,310	0,980
SMOTE 300%	Word2Vec	Yok	0,906	0,663	0,310	0,420	0,549	0,283	0,310	0,980
SMOTE 600%	TF-IDF	Var	0,903	0,735	0,197	0,310	0,441	0,180	0,197	0,991
SMOTE 600%	TF-IDF	Yok	0,918	0,864	0,301	0,445	0,546	0,278	0,301	0,994
SMOTE 600%	Word2Vec	Var	0,906	0,662	0,305	0,416	0,546	0,279	0,305	0,981
SMOTE 600%	Word2Vec	Yok	0,907	0,672	0,307	0,420	0,548	0,281	0,307	0,981

VDY: Veri dengeleme yöntemi, KTY: Kelime temsili yöntemi, EÖ: Ek öznitelikler, A: Doğruluk (Accuracy), P: Kesinlik (Precision), R: Duyarlılık (Recall), F1: F-ölçütü, GM: Geometrik ortalama (G-Mean), IBA: Index of Balanced Accuracy, TNR: True Negative Rate, TPR: True Positive Rate

EK 5. XGBOOST Sınıflandırma Modeli Sonuçları

VDY	KTY	EÖ	A	P	R	F1	GM	IBA	TPR	TNR
CHATGPT (%100)	TF-IDF	Var	0,926	0,932	0,677	0,784	0,818	0,648	0,677	0,988
CHATGPT (%100)	TF-IDF	Yok	0,925	0,929	0,677	0,783	0,817	0,648	0,677	0,987
CHATGPT (%100)	Word2Vec	Var	0,909	0,925	0,592	0,721	0,764	0,561	0,592	0,988
CHATGPT (%100)	Word2Vec	Yok	0,909	0,927	0,586	0,718	0,761	0,556	0,586	0,988
Hibrit (RUS 1:2, SMOTE 100%)	TF-IDF	Var	0,855	0,410	0,698	0,516	0,781	0,600	0,698	0,875
Hibrit (RUS 1:2, SMOTE 100%)	TF-IDF	Yok	0,855	0,409	0,697	0,515	0,781	0,599	0,697	0,875
Hibrit (RUS 1:2, SMOTE 100%)	Word2Vec	Var	0,785	0,270	0,555	0,363	0,671	0,440	0,555	0,813
Hibrit (RUS 1:2, SMOTE 100%)	Word2Vec	Yok	0,782	0,270	0,567	0,365	0,677	0,448	0,567	0,809
Hibrit (RUS 1:2, SMOTE 200%)	TF-IDF	Var	0,848	0,396	0,710	0,508	0,783	0,604	0,710	0,865
Hibrit (RUS 1:2, SMOTE 200%)	TF-IDF	Yok	0,852	0,404	0,716	0,516	0,788	0,612	0,716	0,869
Hibrit (RUS 1:2, SMOTE 200%)	Word2Vec	Var	0,745	0,242	0,610	0,346	0,681	0,458	0,610	0,762
Hibrit (RUS 1:2, SMOTE 200%)	Word2Vec	Yok	0,745	0,244	0,622	0,351	0,688	0,467	0,622	0,761
Dengeleme Yok	TF-IDF	Var	0,919	0,796	0,357	0,491	0,593	0,331	0,357	0,989
Dengeleme Yok	TF-IDF	Yok	0,919	0,793	0,357	0,490	0,593	0,331	0,357	0,988
Dengeleme Yok	Word2Vec	Var	0,897	0,771	0,133	0,227	0,362	0,121	0,133	0,992
Dengeleme Yok	Word2Vec	Yok	0,900	0,763	0,133	0,226	0,362	0,121	0,133	0,995
RUS (1:1)	TF-IDF	Var	0,785	0,312	0,789	0,447	0,786	0,619	0,789	0,784
RUS (1:1)	TF-IDF	Yok	0,784	0,312	0,793	0,448	0,787	0,621	0,793	0,783
RUS (1:1)	Word2Vec	Var	0,688	0,214	0,680	0,325	0,684	0,468	0,680	0,690
RUS (1:1)	Word2Vec	Yok	0,689	0,214	0,681	0,326	0,685	0,469	0,681	0,689
RUS (1:1,5)	TF-IDF	Var	0,843	0,387	0,711	0,500	0,781	0,601	0,711	0,860
RUS (1:1,5)	TF-IDF	Yok	0,842	0,385	0,710	0,498	0,780	0,600	0,710	0,859
RUS (1:1,5)	Word2Vec	Var	0,793	0,272	0,524	0,358	0,657	0,420	0,524	0,826

EK 5. XGBOOST Sınıflandırma Modeli Sonuçları

VDY	KTY	EÖ	A	P	R	F1	GM	IBA	TPR	TNR
RUS (1:1,5)	Word2Vec	Yok	0,797	0,279	0,534	0,367	0,665	0,430	0,534	0,829
RUS (1:2)	TF-IDF	Var	0,870	0,442	0,646	0,524	0,761	0,565	0,646	0,898
RUS (1:2)	TF-IDF	Yok	0,872	0,446	0,645	0,527	0,761	0,566	0,645	0,900
RUS (1:2)	Word2Vec	Var	0,842	0,332	0,420	0,370	0,612	0,358	0,420	0,895
RUS (1:2)	Word2Vec	Yok	0,840	0,327	0,417	0,366	0,609	0,355	0,417	0,893
SMOTE 100%	TF-IDF	Var	0,920	0,751	0,409	0,528	0,633	0,379	0,409	0,983
SMOTE 100%	TF-IDF	Yok	0,917	0,732	0,396	0,512	0,622	0,366	0,396	0,982
SMOTE 100%	Word2Vec	Var	0,898	0,611	0,214	0,316	0,458	0,194	0,214	0,983
SMOTE 100%	Word2Vec	Yok	0,898	0,619	0,210	0,312	0,453	0,190	0,210	0,984
SMOTE 200%	TF-IDF	Var	0,918	0,718	0,421	0,530	0,641	0,389	0,421	0,979
SMOTE 200%	TF-IDF	Yok	0,917	0,716	0,418	0,526	0,638	0,386	0,418	0,980
SMOTE 200%	Word2Vec	Var	0,894	0,542	0,250	0,341	0,493	0,226	0,250	0,973
SMOTE 200%	Word2Vec	Yok	0,894	0,547	0,235	0,328	0,478	0,213	0,235	0,976
SMOTE 300%	TF-IDF	Var	0,917	0,697	0,438	0,537	0,653	0,405	0,438	0,976
SMOTE 300%	TF-IDF	Yok	0,917	0,700	0,437	0,537	0,653	0,404	0,437	0,977
SMOTE 300%	Word2Vec	Var	0,891	0,513	0,278	0,359	0,517	0,250	0,278	0,967
SMOTE 300%	Word2Vec	Yok	0,890	0,506	0,270	0,351	0,510	0,243	0,270	0,967
SMOTE 600%	TF-IDF	Var	0,915	0,672	0,451	0,539	0,662	0,416	0,451	0,972
SMOTE 600%	TF-IDF	Yok	0,915	0,676	0,453	0,541	0,663	0,418	0,453	0,973
SMOTE 600%	Word2Vec	Var	0,877	0,428	0,323	0,366	0,551	0,287	0,323	0,946
SMOTE 600%	Word2Vec	Yok	0,876	0,422	0,323	0,365	0,551	0,286	0,323	0,945

VDY: Veri dengeleme yöntemi, KTY: Kelime temsili yöntemi, EÖ: Ek öznitelikler, A: Doğruluk (Accuracy), P: Kesinlik (Precision), R: Duyarlılık (Recall), F1: F-ölçütü, GM: Geometrik ortalama (G-Mean), IBA: Index of Balanced Accuracy, TNR: True Negative Rate, TPR: True Positive Rate

EK 6. LightGBM Sınıflandırma Modeli Sonuçları

VDY	KTY	EÖ	A	P	R	F1	GM	IBA	TPR	TNR
CHATGPT (%100)	TF-IDF	Var	0,926	0,941	0,671	0,783	0,815	0,643	0,671	0,990
CHATGPT (%100)	TF-IDF	Yok	0,926	0,941	0,671	0,783	0,815	0,643	0,671	0,989
CHATGPT (%100)	Word2Vec	Var	0,908	0,938	0,578	0,715	0,756	0,549	0,578	0,990
CHATGPT (%100)	Word2Vec	Yok	0,907	0,933	0,571	0,708	0,752	0,542	0,571	0,990
Hibrit (RUS 1:2, SMOTE 100%)	TF-IDF	Var	0,857	0,413	0,698	0,519	0,782	0,601	0,698	0,877
Hibrit (RUS 1:2, SMOTE 100%)	TF-IDF	Yok	0,855	0,409	0,693	0,513	0,778	0,595	0,693	0,875
Hibrit (RUS 1:2, SMOTE 100%)	Word2Vec	Var	0,790	0,278	0,564	0,373	0,679	0,450	0,564	0,818
Hibrit (RUS 1:2, SMOTE 100%)	Word2Vec	Yok	0,789	0,277	0,563	0,371	0,678	0,448	0,563	0,817
Hibrit (RUS 1:2, SMOTE 200%)	TF-IDF	Var	0,849	0,397	0,711	0,509	0,784	0,606	0,711	0,866
Hibrit (RUS 1:2, SMOTE 200%)	TF-IDF	Yok	0,849	0,398	0,715	0,511	0,786	0,609	0,715	0,865
Hibrit (RUS 1:2, SMOTE 200%)	Word2Vec	Var	0,755	0,251	0,614	0,356	0,688	0,467	0,614	0,773
Hibrit (RUS 1:2, SMOTE 200%)	Word2Vec	Yok	0,754	0,251	0,618	0,357	0,690	0,469	0,618	0,771
Dengeleme Yok	TF-IDF	Var	0,918	0,835	0,326	0,468	0,568	0,302	0,326	0,992
Dengeleme Yok	TF-IDF	Yok	0,919	0,830	0,331	0,472	0,572	0,307	0,331	0,992
Dengeleme Yok	Word2Vec	Var	0,897	0,826	0,117	0,204	0,339	0,107	0,117	0,993
Dengeleme Yok	Word2Vec	Yok	0,900	0,812	0,116	0,202	0,338	0,106	0,116	0,997
RUS (1:1)	TF-IDF	Var	0,790	0,320	0,797	0,457	0,793	0,630	0,797	0,790
RUS (1:1)	TF-IDF	Yok	0,789	0,318	0,788	0,452	0,788	0,621	0,788	0,790
RUS (1:1)	Word2Vec	Var	0,705	0,223	0,674	0,335	0,691	0,476	0,674	0,709
RUS (1:1)	Word2Vec	Yok	0,704	0,223	0,679	0,336	0,692	0,478	0,679	0,707
RUS (1:1,5)	TF-IDF	Var	0,847	0,394	0,713	0,507	0,784	0,607	0,713	0,864
RUS (1:1,5)	TF-IDF	Yok	0,847	0,393	0,708	0,505	0,782	0,602	0,708	0,864
RUS (1:1,5)	Word2Vec	Var	0,806	0,292	0,526	0,375	0,665	0,429	0,526	0,841

EK 6. LightGBM Sınıflandırma Modeli Sonuçları

VDY	KTY	EÖ	A	P	R	F1	GM	IBA	TPR	TNR
RUS (1:1,5)	Word2Vec	Yok	0,804	0,289	0,528	0,373	0,665	0,429	0,528	0,839
RUS (1:2)	TF-IDF	Var	0,876	0,458	0,637	0,532	0,759	0,562	0,637	0,906
RUS (1:2)	TF-IDF	Yok	0,876	0,459	0,642	0,534	0,762	0,566	0,642	0,905
RUS (1:2)	Word2Vec	Var	0,850	0,353	0,427	0,385	0,620	0,367	0,427	0,902
RUS (1:2)	Word2Vec	Yok	0,851	0,356	0,422	0,385	0,617	0,363	0,422	0,905
SMOTE 100%	TF-IDF	Var	0,919	0,750	0,403	0,523	0,628	0,373	0,403	0,983
SMOTE 100%	TF-IDF	Yok	0,918	0,742	0,391	0,511	0,619	0,362	0,391	0,983
SMOTE 100%	Word2Vec	Var	0,899	0,631	0,198	0,300	0,440	0,180	0,198	0,986
SMOTE 100%	Word2Vec	Yok	0,899	0,642	0,195	0,298	0,437	0,177	0,195	0,986
SMOTE 200%	TF-IDF	Var	0,919	0,727	0,422	0,533	0,642	0,391	0,422	0,980
SMOTE 200%	TF-IDF	Yok	0,917	0,715	0,413	0,522	0,635	0,382	0,413	0,979
SMOTE 200%	Word2Vec	Var	0,893	0,537	0,236	0,327	0,478	0,213	0,236	0,975
SMOTE 200%	Word2Vec	Yok	0,894	0,542	0,237	0,328	0,478	0,214	0,237	0,975
SMOTE 300%	TF-IDF	Var	0,916	0,697	0,431	0,531	0,648	0,398	0,431	0,977
SMOTE 300%	TF-IDF	Yok	0,917	0,701	0,436	0,536	0,652	0,403	0,436	0,977
SMOTE 300%	Word2Vec	Var	0,888	0,490	0,282	0,357	0,520	0,253	0,282	0,963
SMOTE 300%	Word2Vec	Yok	0,887	0,477	0,270	0,344	0,509	0,242	0,270	0,963
SMOTE 600%	TF-IDF	Var	0,915	0,664	0,463	0,545	0,670	0,427	0,463	0,971
SMOTE 600%	TF-IDF	Yok	0,914	0,659	0,461	0,541	0,668	0,425	0,461	0,970
SMOTE 600%	Word2Vec	Var	0,872	0,407	0,347	0,374	0,569	0,306	0,347	0,937
SMOTE 600%	Word2Vec	Yok	0,872	0,407	0,349	0,375	0,571	0,308	0,349	0,937

VDY: Veri dengeleme yöntemi, KTY: Kelime temsili yöntemi, EÖ: Ek öznitelikler, A: Doğruluk (Accuracy), P: Kesinlik (Precision), R: Duyarlılık (Recall), F1: F-ölçütü, GM: Geometrik ortalama (G-Mean), IBA: Index of Balanced Accuracy, TNR: True Negative Rate, TPR: True Positive Rate

EK 7. ANN Sınıflandırma Modeli Sonuçları

VDY	KTY	EÖ	A	P	R	F1	GM	IBA	TPR	TNR
CHATGPT (%100)	TF-IDF	Var	0,899	0,775	0,699	0,734	0,814	0,646	0,699	0,948
CHATGPT (%100)	TF-IDF	Yok	0,898	0,772	0,699	0,732	0,814	0,646	0,699	0,948
CHATGPT (%100)	Word2Vec	Var	0,886	0,826	0,610	0,685	0,760	0,559	0,610	0,954
CHATGPT (%100)	Word2Vec	Yok	0,901	0,889	0,577	0,697	0,751	0,543	0,577	0,981
Hibrit (RUS 1:2, SMOTE 100%)	TF-IDF	Var	0,785	0,282	0,601	0,382	0,696	0,475	0,601	0,808
Hibrit (RUS 1:2, SMOTE 100%)	TF-IDF	Yok	0,789	0,289	0,615	0,392	0,705	0,488	0,615	0,810
Hibrit (RUS 1:2, SMOTE 100%)	Word2Vec	Var	0,762	0,256	0,603	0,359	0,686	0,463	0,603	0,781
Hibrit (RUS 1:2, SMOTE 100%)	Word2Vec	Yok	0,760	0,253	0,598	0,355	0,683	0,458	0,598	0,780
Hibrit (RUS 1:2, SMOTE 200%)	TF-IDF	Var	0,805	0,305	0,588	0,400	0,698	0,476	0,588	0,832
Hibrit (RUS 1:2, SMOTE 200%)	TF-IDF	Yok	0,795	0,292	0,590	0,389	0,695	0,473	0,590	0,821
Hibrit (RUS 1:2, SMOTE 200%)	Word2Vec	Var	0,746	0,242	0,610	0,346	0,682	0,458	0,610	0,763
Hibrit (RUS 1:2, SMOTE 200%)	Word2Vec	Yok	0,751	0,248	0,612	0,352	0,686	0,463	0,612	0,769
Dengeleme Yok	TF-IDF	Var	0,868	0,416	0,457	0,433	0,647	0,400	0,457	0,918
Dengeleme Yok	TF-IDF	Yok	0,869	0,419	0,461	0,436	0,649	0,404	0,461	0,920
Dengeleme Yok	Word2Vec	Var	0,866	0,391	0,302	0,339	0,530	0,266	0,302	0,937
Dengeleme Yok	Word2Vec	Yok	0,870	0,392	0,310	0,345	0,539	0,273	0,310	0,940
RUS (1:1)	TF-IDF	Var	0,685	0,223	0,741	0,342	0,708	0,505	0,741	0,678
RUS (1:1)	TF-IDF	Yok	0,685	0,222	0,738	0,341	0,707	0,503	0,738	0,678
RUS (1:1)	Word2Vec	Var	0,692	0,219	0,697	0,333	0,694	0,482	0,697	0,692
RUS (1:1)	Word2Vec	Yok	0,689	0,216	0,690	0,329	0,689	0,475	0,690	0,689
RUS (1:1,5)	TF-IDF	Var	0,723	0,242	0,707	0,361	0,715	0,511	0,707	0,725
RUS (1:1,5)	TF-IDF	Yok	0,723	0,245	0,715	0,364	0,719	0,517	0,715	0,724

EK 7. ANN Sınıflandırma Modeli Sonuçları

VDY	KTY	EÖ	A	P	R	F1	GM	IBA	TPR	TNR
RUS (1:1,5)	Word2Vec	Var	0,753	0,251	0,623	0,358	0,692	0,472	0,623	0,769
RUS (1:1,5)	Word2Vec	Yok	0,748	0,248	0,626	0,355	0,691	0,472	0,626	0,763
RUS (1:2)	TF-IDF	Var	0,748	0,258	0,678	0,373	0,716	0,509	0,678	0,757
RUS (1:2)	TF-IDF	Yok	0,752	0,259	0,663	0,371	0,710	0,500	0,663	0,763
RUS (1:2)	Word2Vec	Var	0,782	0,270	0,566	0,365	0,676	0,446	0,566	0,809
RUS (1:2)	Word2Vec	Yok	0,780	0,269	0,577	0,367	0,681	0,454	0,577	0,805
SMOTE 100%	TF-IDF	Var	0,878	0,450	0,431	0,438	0,633	0,382	0,431	0,934
SMOTE 100%	TF-IDF	Yok	0,876	0,440	0,426	0,430	0,628	0,377	0,426	0,931
SMOTE 100%	Word2Vec	Var	0,859	0,363	0,365	0,363	0,579	0,317	0,365	0,920
SMOTE 100%	Word2Vec	Yok	0,860	0,366	0,364	0,364	0,579	0,317	0,364	0,921
SMOTE 200%	TF-IDF	Var	0,881	0,459	0,407	0,429	0,617	0,362	0,407	0,940
SMOTE 200%	TF-IDF	Yok	0,881	0,459	0,418	0,435	0,625	0,371	0,418	0,938
SMOTE 200%	Word2Vec	Var	0,857	0,364	0,392	0,376	0,598	0,340	0,392	0,914
SMOTE 200%	Word2Vec	Yok	0,856	0,355	0,374	0,363	0,584	0,324	0,374	0,915
SMOTE 300%	TF-IDF	Var	0,886	0,485	0,394	0,432	0,609	0,352	0,394	0,947
SMOTE 300%	TF-IDF	Yok	0,883	0,474	0,408	0,435	0,619	0,364	0,408	0,942
SMOTE 300%	Word2Vec	Var	0,850	0,347	0,401	0,372	0,603	0,345	0,401	0,906
SMOTE 300%	Word2Vec	Yok	0,849	0,344	0,398	0,368	0,600	0,342	0,398	0,905
SMOTE 600%	TF-IDF	Var	0,888	0,493	0,384	0,430	0,603	0,345	0,384	0,950
SMOTE 600%	TF-IDF	Yok	0,887	0,490	0,386	0,429	0,604	0,346	0,386	0,949
SMOTE 600%	Word2Vec	Var	0,845	0,340	0,426	0,378	0,617	0,364	0,426	0,897
SMOTE 600%	Word2Vec	Yok	0,845	0,339	0,423	0,376	0,616	0,362	0,423	0,897

VDY: Veri dengeleme yöntemi, KTY: Kelime temsili yöntemi, EÖ: Ek öznitelikler, A: Doğruluk (Accuracy), P: Kesinlik (Precision), R: Duyarlılık (Recall), F1: F-ölçütü, GM: Geometrik ortalama (G-Mean), IBA: Index of Balanced Accuracy, TNR: True Negative Rate, TPR: True Positive Rate

EK 8. LSTM Sınıflandırma Modeli Sonuçları

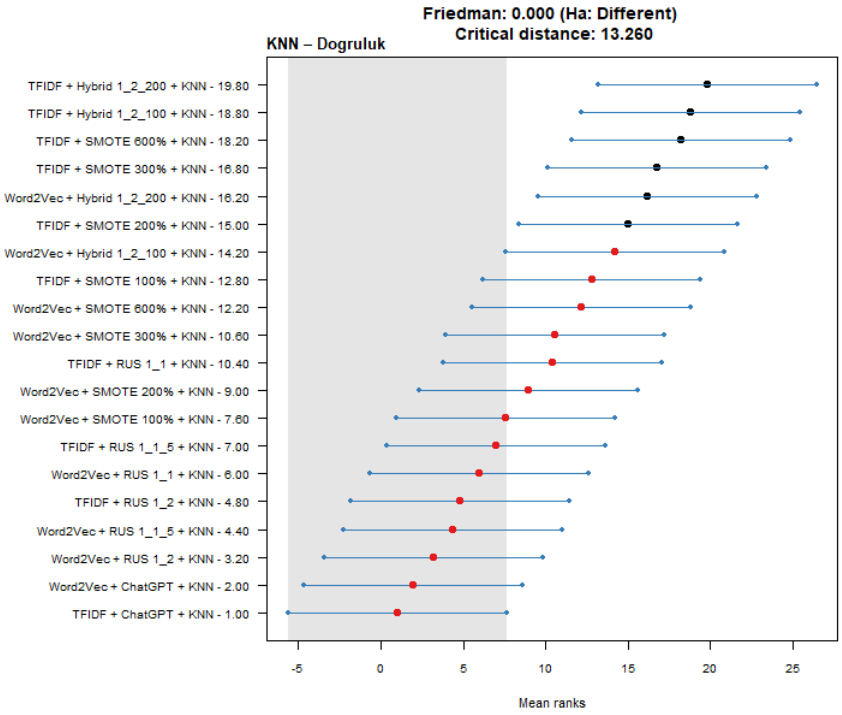
VDY	KTY	EÖ	A	P	R	F1	GM	IBA	TPR	TNR
CHATGPT (%100)	TF-IDF	Var	0,913	0,859	0,676	0,755	0,810	0,638	0,676	0,972
CHATGPT (%100)	TF-IDF	Yok	0,907	0,854	0,671	0,751	0,804	0,632	0,671	0,964
CHATGPT (%100)	Word2Vec	Var	0,912	0,824	0,711	0,762	0,826	0,667	0,711	0,961
CHATGPT (%100)	Word2Vec	Yok	0,909	0,812	0,707	0,755	0,823	0,661	0,707	0,959
Hibrit (RUS 1:2, SMOTE 100%)	TF-IDF	Var	0,858	0,410	0,611	0,488	0,736	0,527	0,611	0,888
Hibrit (RUS 1:2, SMOTE 100%)	TF-IDF	Yok	0,857	0,406	0,608	0,484	0,733	0,524	0,608	0,888
Hibrit (RUS 1:2, SMOTE 100%)	Word2Vec	Var	0,741	0,247	0,598	0,341	0,668	0,440	0,598	0,759
Hibrit (RUS 1:2, SMOTE 100%)	Word2Vec	Yok	0,761	0,253	0,558	0,343	0,658	0,424	0,558	0,786
Hibrit (RUS 1:2, SMOTE 200%)	TF-IDF	Var	0,861	0,417	0,592	0,485	0,726	0,513	0,592	0,894
Hibrit (RUS 1:2, SMOTE 200%)	TF-IDF	Yok	0,857	0,407	0,610	0,486	0,735	0,527	0,610	0,888
Hibrit (RUS 1:2, SMOTE 200%)	Word2Vec	Var	0,639	0,191	0,690	0,298	0,658	0,436	0,690	0,633
Hibrit (RUS 1:2, SMOTE 200%)	Word2Vec	Yok	0,649	0,191	0,668	0,296	0,655	0,430	0,668	0,647
Dengeleme Yok	TF-IDF	Var	0,901	0,569	0,450	0,500	0,655	0,409	0,450	0,957
Dengeleme Yok	TF-IDF	Yok	0,900	0,562	0,444	0,494	0,651	0,403	0,444	0,956
Dengeleme Yok	Word2Vec	Var	0,902	0,590	0,398	0,471	0,618	0,361	0,398	0,964
Dengeleme Yok	Word2Vec	Yok	0,900	0,578	0,406	0,471	0,623	0,368	0,406	0,961
RUS (1:1)	TF-IDF	Var	0,802	0,327	0,720	0,447	0,764	0,578	0,720	0,813
RUS (1:1)	TF-IDF	Yok	0,811	0,331	0,678	0,443	0,747	0,551	0,678	0,827
RUS (1:1)	Word2Vec	Var	0,781	0,277	0,538	0,353	0,654	0,419	0,538	0,811
RUS (1:1)	Word2Vec	Yok	0,788	0,287	0,533	0,359	0,654	0,418	0,533	0,819
RUS (1:1,5)	TF-IDF	Var	0,841	0,375	0,646	0,472	0,746	0,545	0,646	0,865
RUS (1:1,5)	TF-IDF	Yok	0,847	0,388	0,639	0,480	0,746	0,544	0,639	0,872
RUS (1:1,5)	Word2Vec	Var	0,820	0,316	0,516	0,389	0,664	0,426	0,516	0,858

EK 8. LSTM Sınıflandırma Modeli Sonuçları

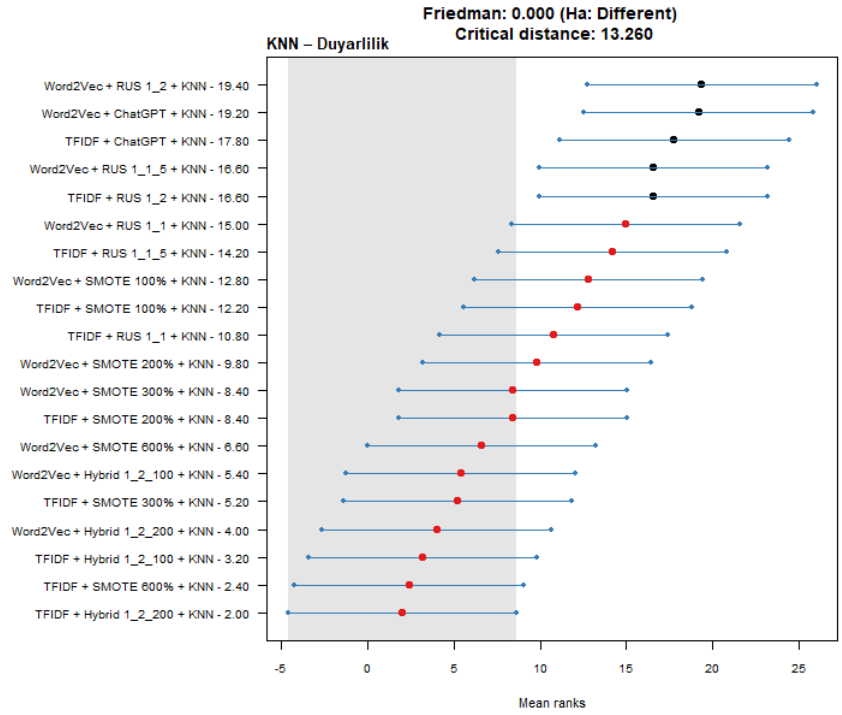
VDY	KTY	EÖ	A	P	R	F1	GM	IBA	TPR	TNR
RUS (1:1,5)	Word2Vec	Yok	0,822	0,327	0,503	0,387	0,655	0,415	0,503	0,862
RUS (1:2)	TF-IDF	Var	0,860	0,414	0,599	0,488	0,731	0,519	0,599	0,893
RUS (1:2)	TF-IDF	Yok	0,863	0,419	0,591	0,488	0,727	0,514	0,591	0,897
RUS (1:2)	Word2Vec	Var	0,848	0,376	0,483	0,415	0,655	0,413	0,483	0,894
RUS (1:2)	Word2Vec	Yok	0,841	0,353	0,492	0,405	0,657	0,416	0,492	0,884
SMOTE 100%	TF-IDF	Var	0,904	0,592	0,462	0,515	0,665	0,421	0,462	0,959
SMOTE 100%	TF-IDF	Yok	0,904	0,589	0,453	0,508	0,658	0,412	0,453	0,960
SMOTE 100%	Word2Vec	Var	0,892	0,528	0,429	0,468	0,637	0,386	0,429	0,950
SMOTE 100%	Word2Vec	Yok	0,897	0,554	0,412	0,467	0,626	0,372	0,412	0,957
SMOTE 200%	TF-IDF	Var	0,907	0,608	0,453	0,516	0,659	0,414	0,453	0,963
SMOTE 200%	TF-IDF	Yok	0,901	0,571	0,464	0,508	0,665	0,421	0,464	0,955
SMOTE 200%	Word2Vec	Var	0,893	0,533	0,430	0,469	0,637	0,387	0,430	0,950
SMOTE 200%	Word2Vec	Yok	0,892	0,525	0,412	0,456	0,625	0,371	0,412	0,951
SMOTE 300%	TF-IDF	Var	0,904	0,596	0,457	0,513	0,661	0,416	0,457	0,960
SMOTE 300%	TF-IDF	Yok	0,903	0,586	0,453	0,507	0,658	0,412	0,453	0,959
SMOTE 300%	Word2Vec	Var	0,881	0,470	0,443	0,453	0,643	0,394	0,443	0,936
SMOTE 300%	Word2Vec	Yok	0,880	0,470	0,441	0,448	0,640	0,391	0,441	0,934
SMOTE 600%	TF-IDF	Var	0,904	0,593	0,455	0,511	0,659	0,415	0,455	0,960
SMOTE 600%	TF-IDF	Yok	0,902	0,577	0,457	0,506	0,660	0,416	0,457	0,957
SMOTE 600%	Word2Vec	Var	0,839	0,353	0,535	0,423	0,683	0,452	0,535	0,876
SMOTE 600%	Word2Vec	Yok	0,848	0,378	0,503	0,424	0,667	0,430	0,503	0,891

VDY: Veri dengeleme yöntemi, KTY: Kelime temsili yöntemi, EÖ: Ek öznitelikler, A: Doğruluk (Accuracy), P: Kesinlik (Precision), R: Duyarlılık (Recall), F1: F-ölçütü, GM: Geometrik ortalama (G-Mean), IBA: Index of Balanced Accuracy, TNR: True Negative Rate, TPR: True Positive Rate

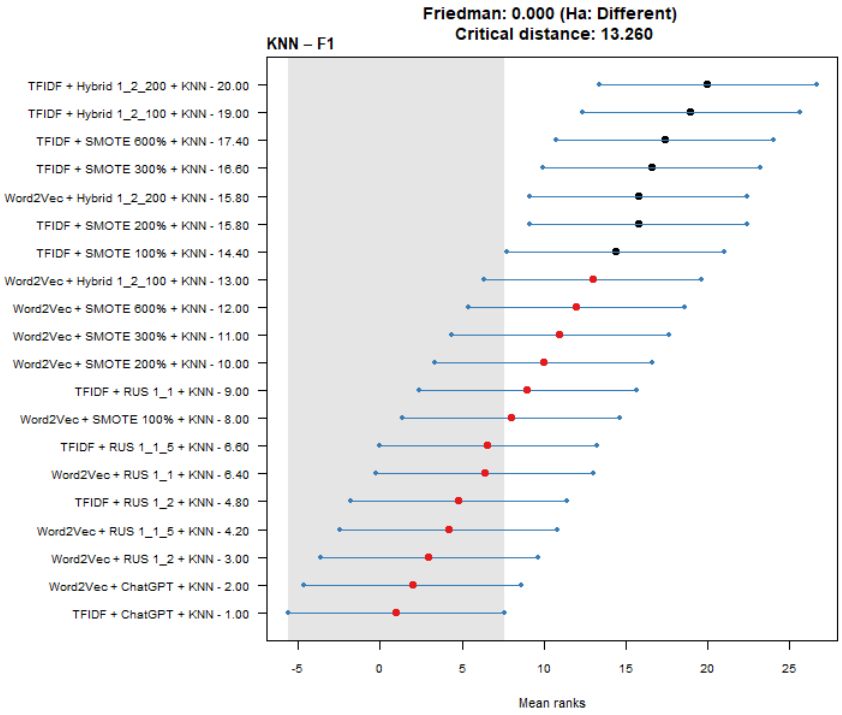
EK 9. Nemenyi Kritik Fark Diyagramları



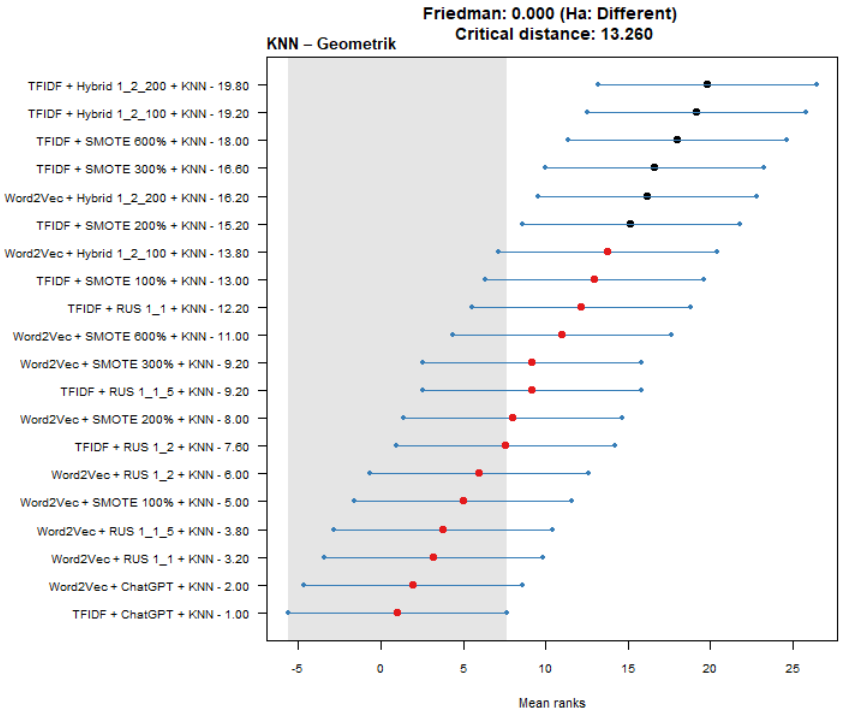
Ek Şekil 9.1. KNN – Doğruluk (Aşama 1)



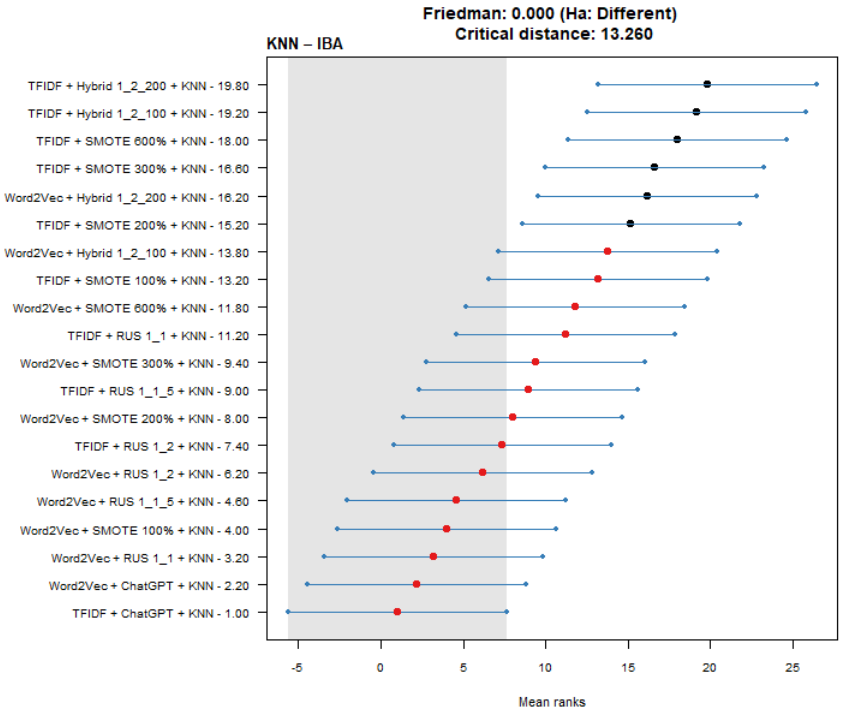
Ek Şekil 9.2. KNN – Duyarlılık (Aşama 1)



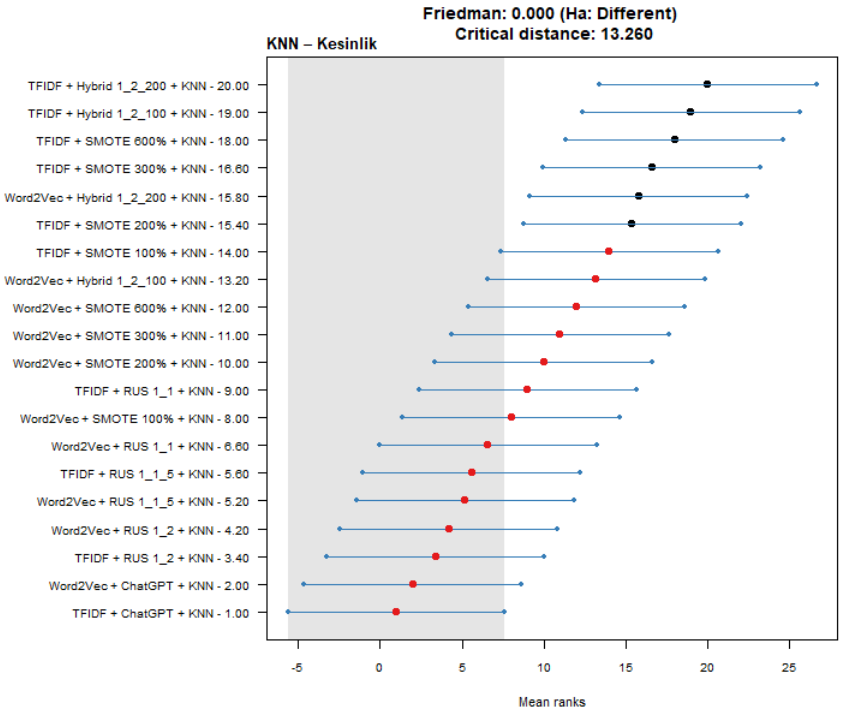
Ek Şekil 9.3. KNN – F-ölçütü (Aşama 1)



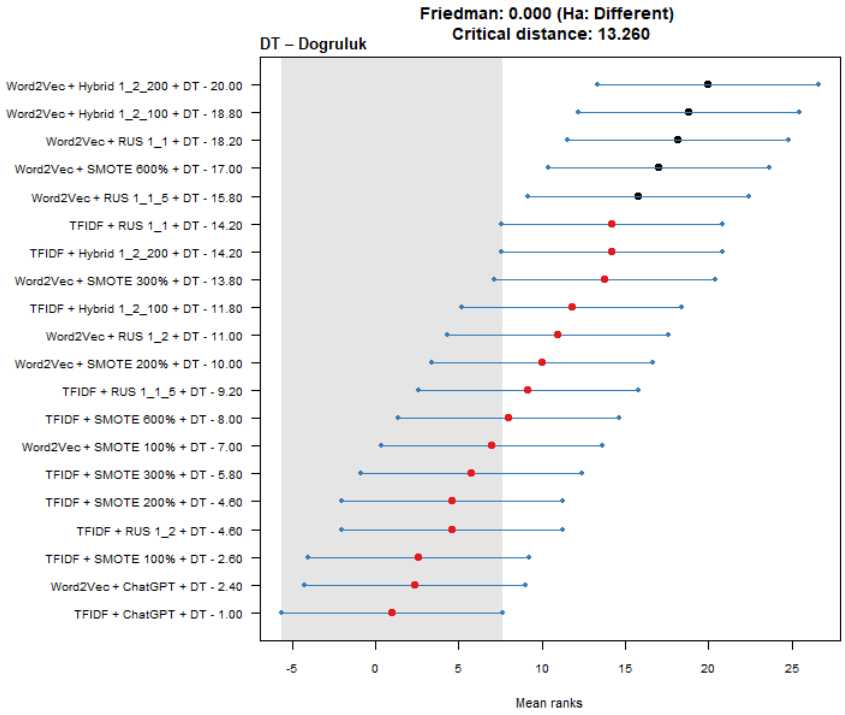
Ek Şekil 9.4. KNN – Geometrik Ortalama (Aşama 1)



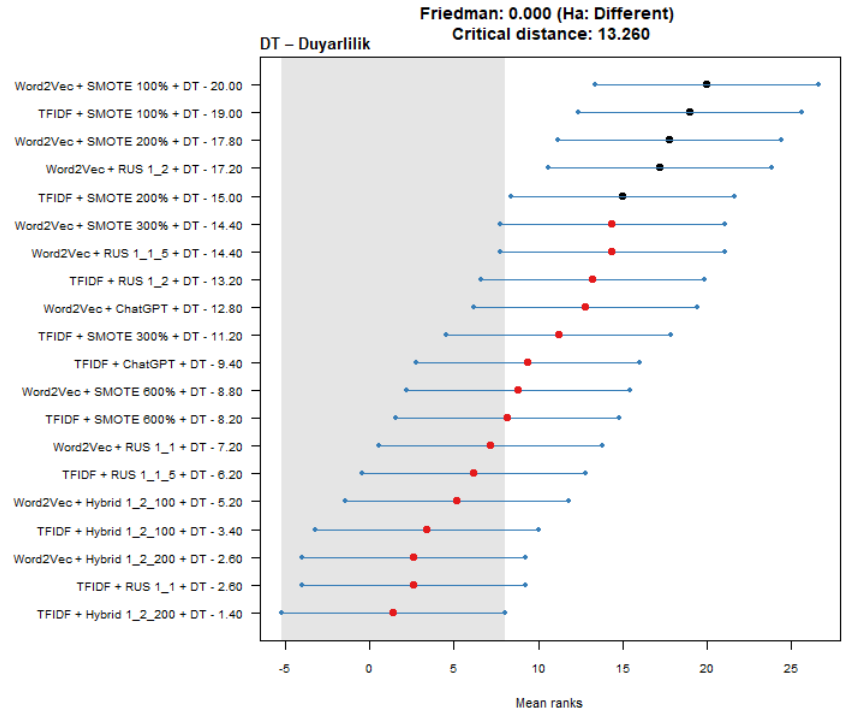
Ek Şekil 9.5. KNN – IBA (Aşama 1)



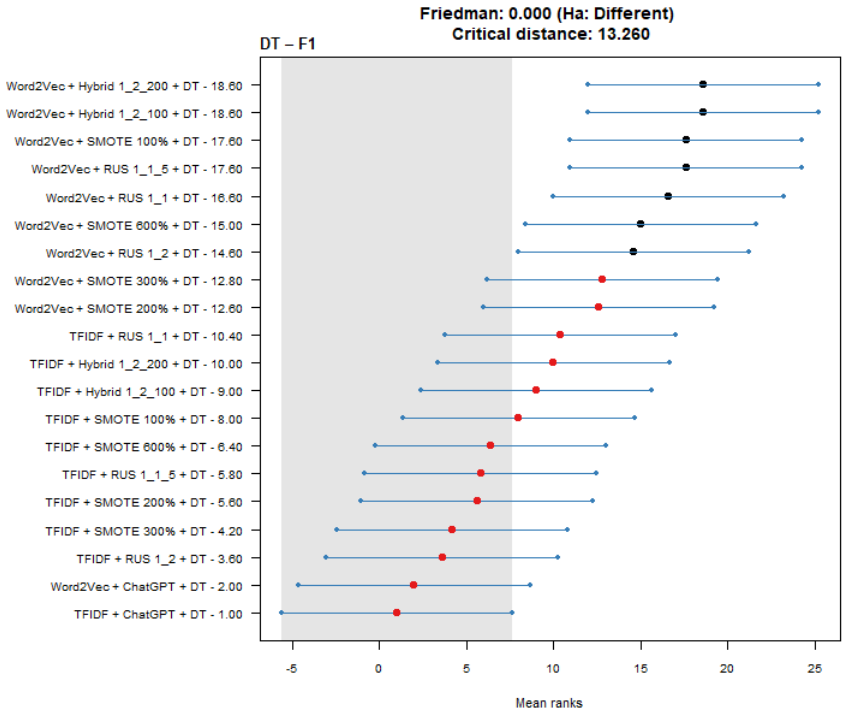
Ek Şekil 9.6. KNN – Kesinlik (Aşama 1)



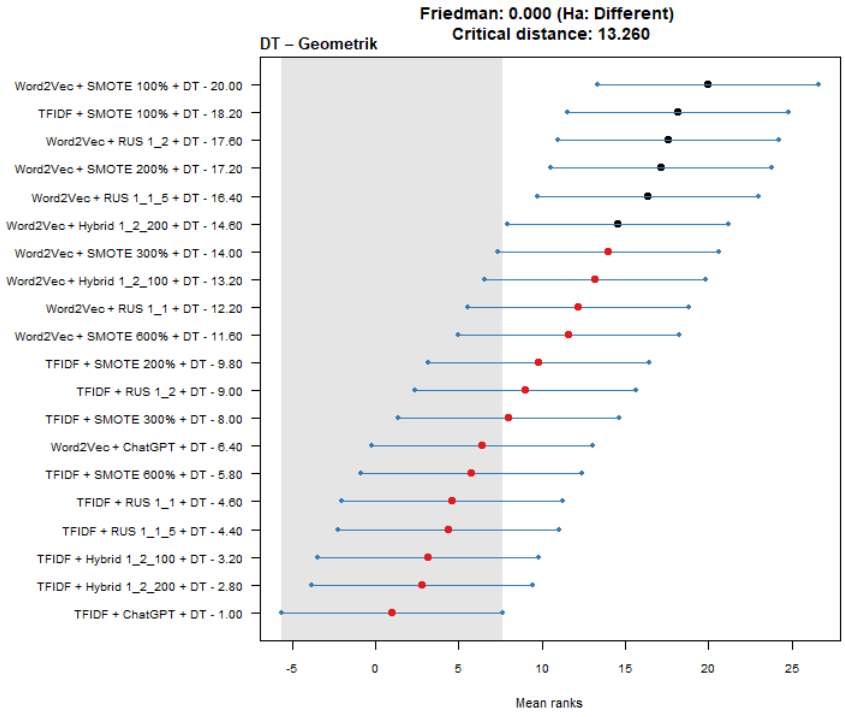
Ek Şekil 9.7. DT – Doğruluk (Aşama 1)



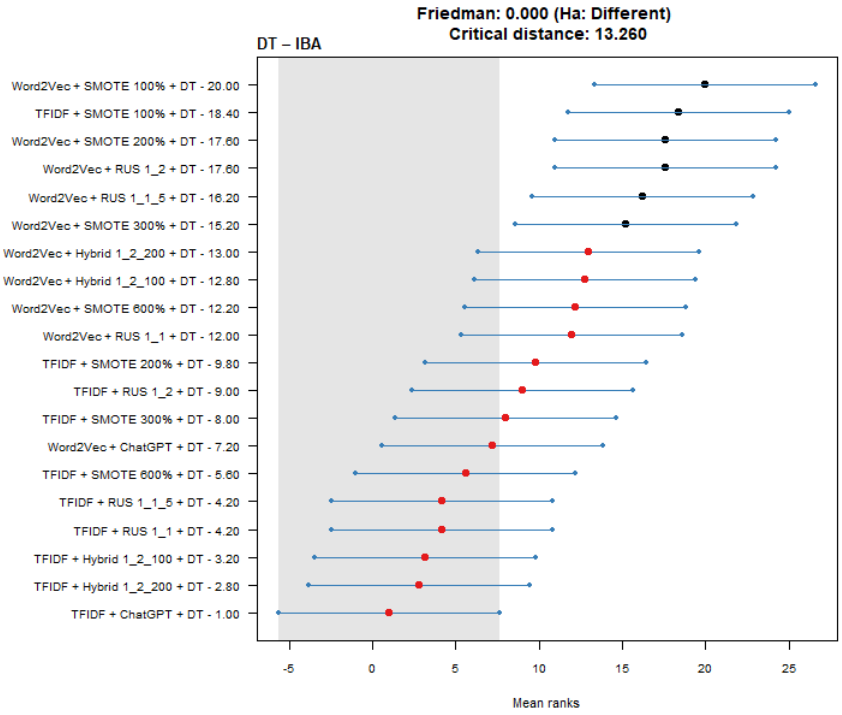
Ek Şekil 9.8. DT – Duyarlılık (Aşama 1)



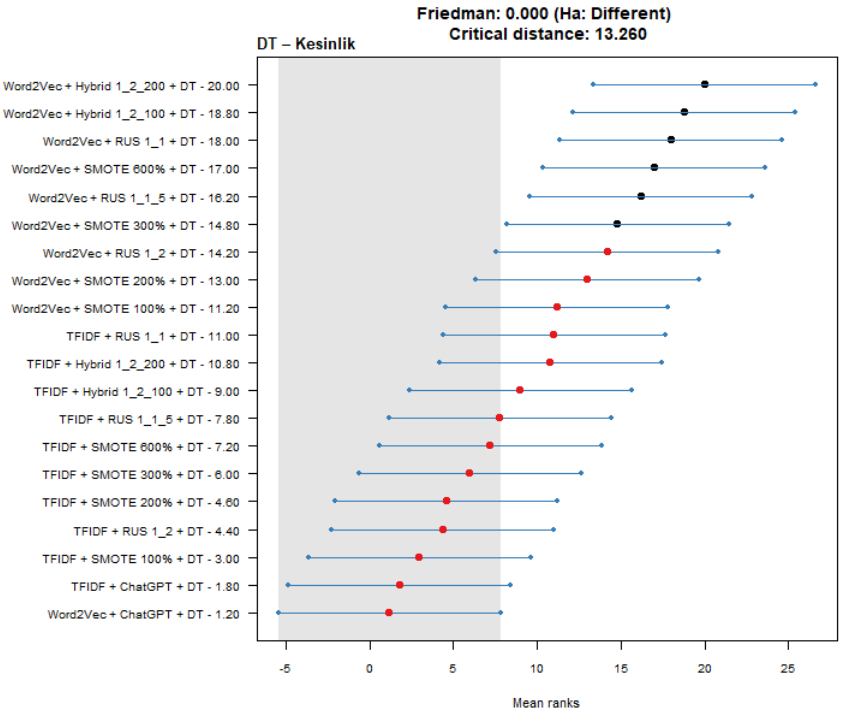
Ek Şekil 9.9. DT – F-ölçütü (Aşama 1)



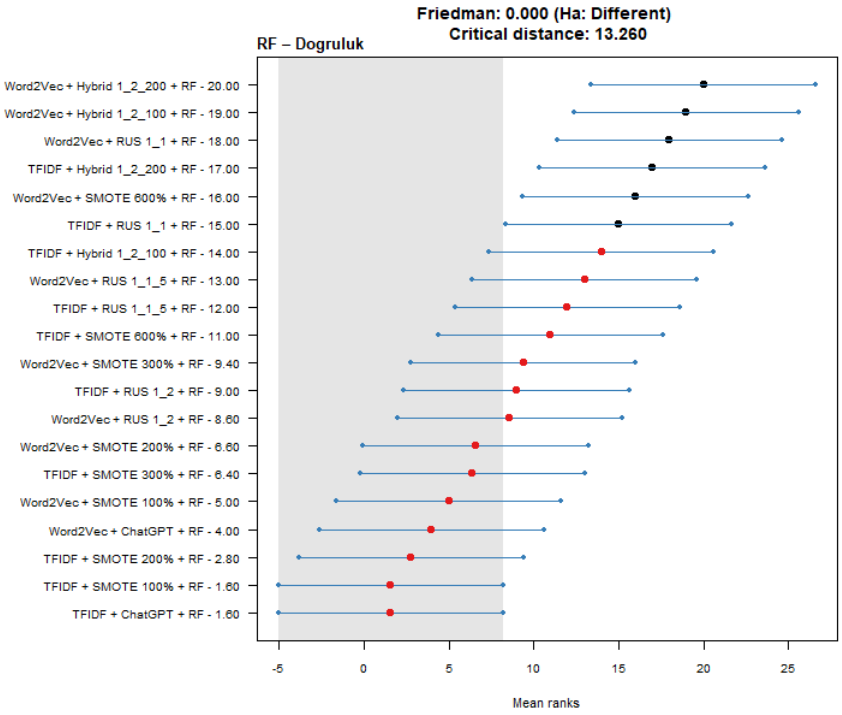
Ek Şekil 9.10. DT – Geometrik Ortalama (Aşama 1)



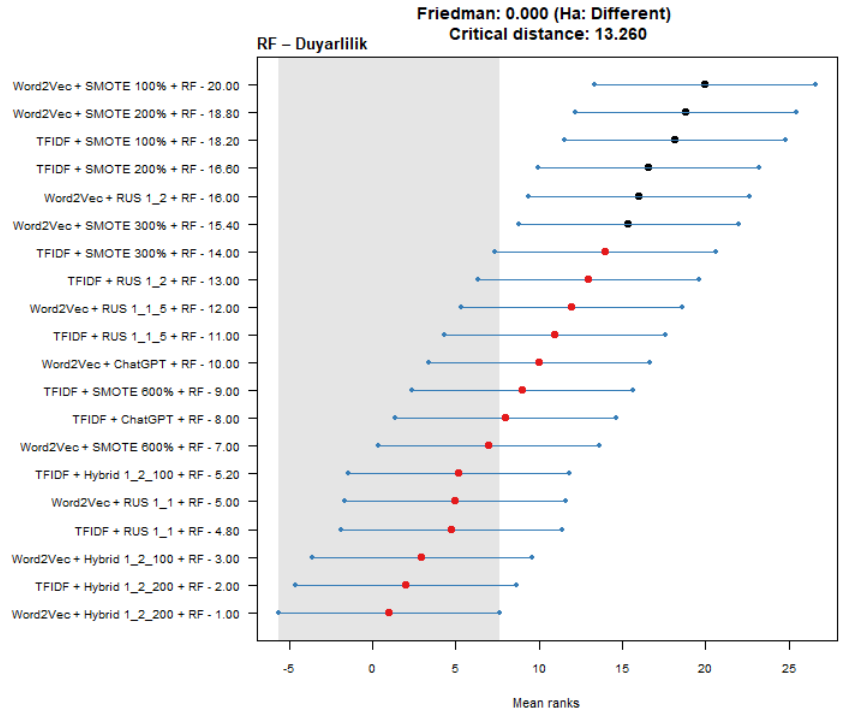
Ek Şekil 9.11. DT – IBA (Aşama 1)



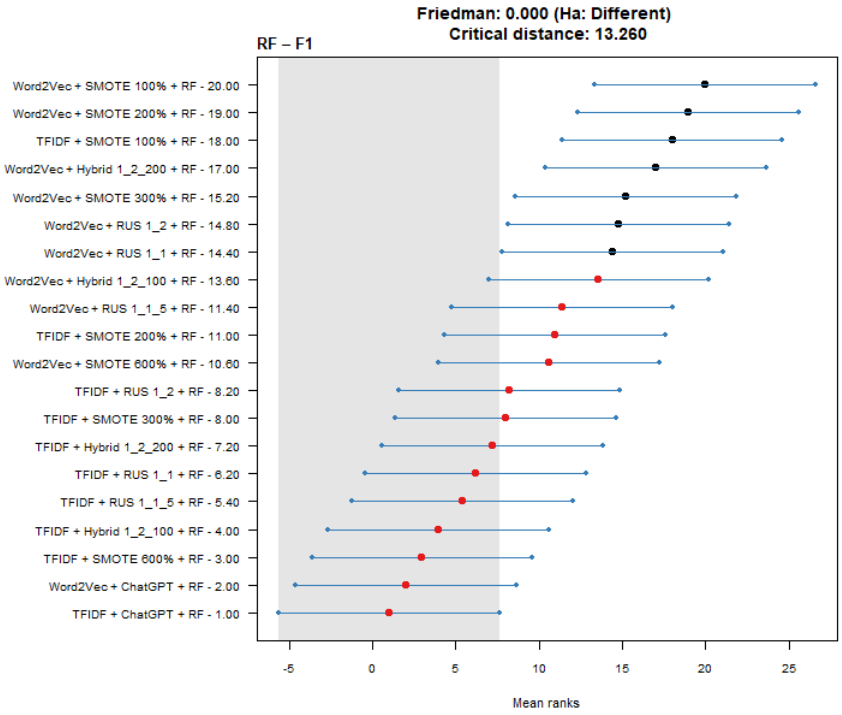
Ek Şekil 9.12. DT – Kesinlik (Aşama 1)



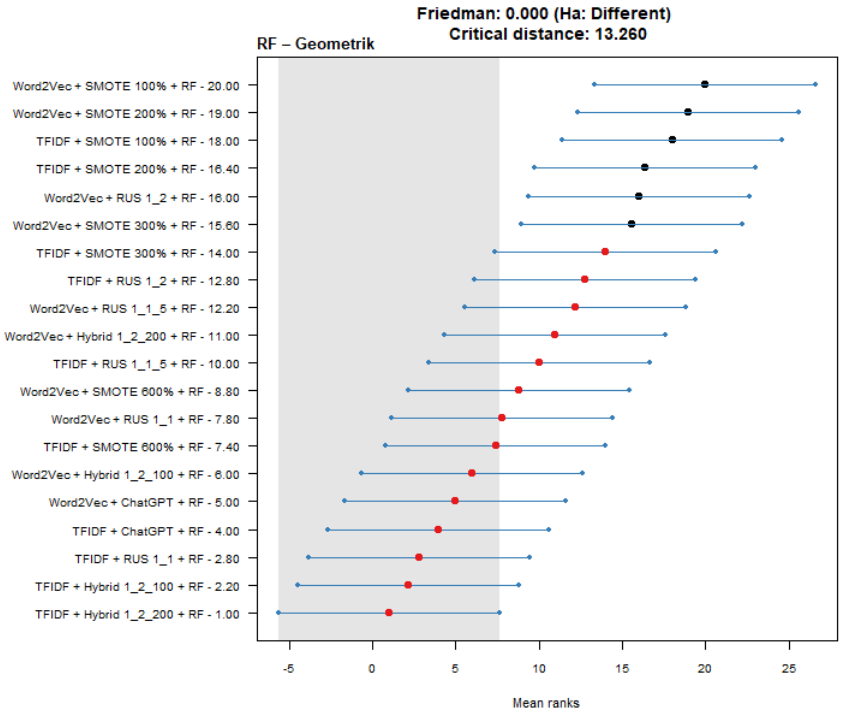
Ek Şekil 9.13. RF – Doğruluk (Aşama 1)



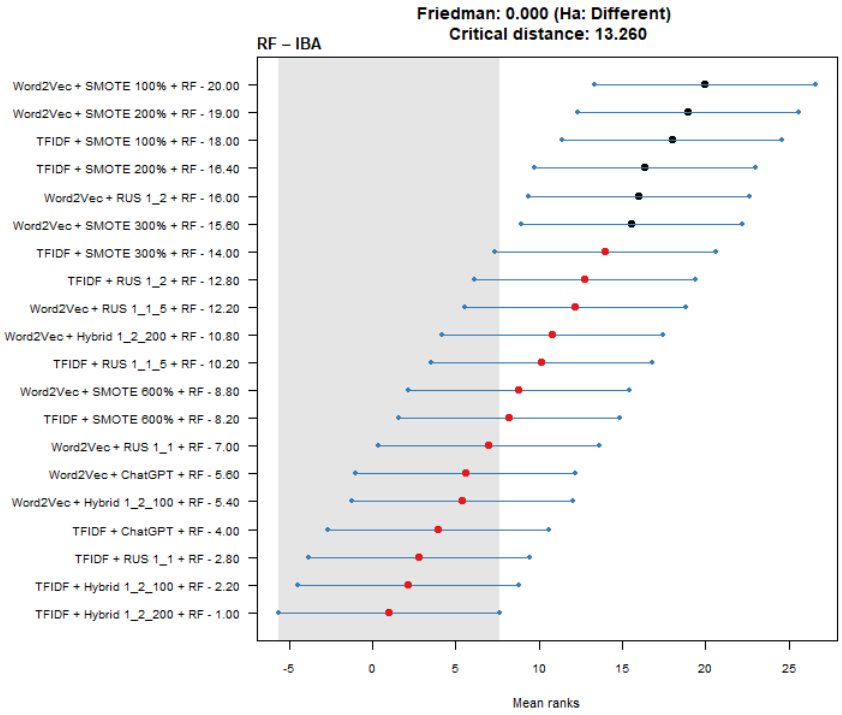
Ek Şekil 9.14. RF – Duyarlılık (Aşama 1)



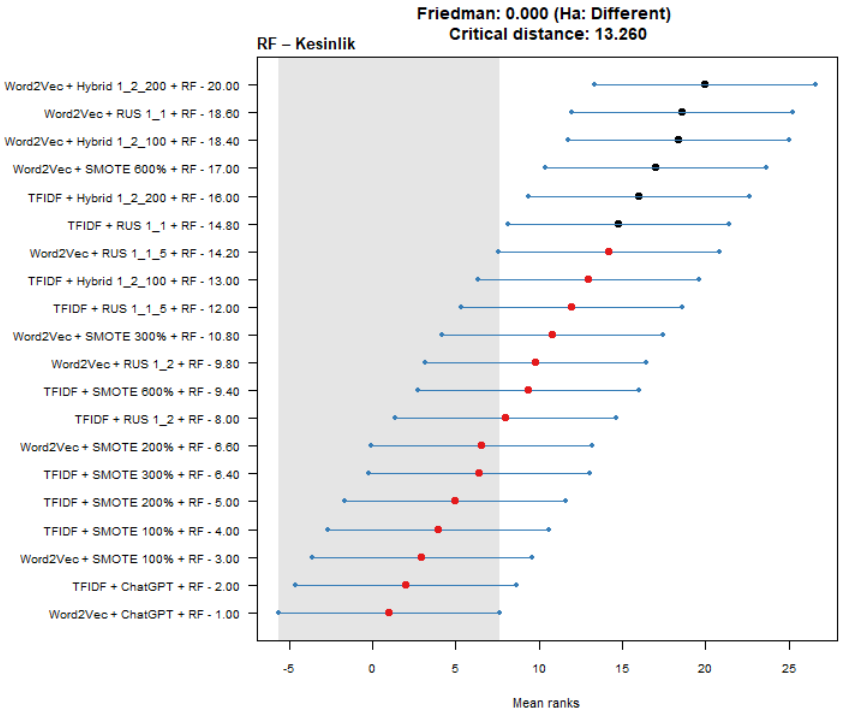
Ek Şekil 9.15. RF – F-ölçütü (Aşama 1)



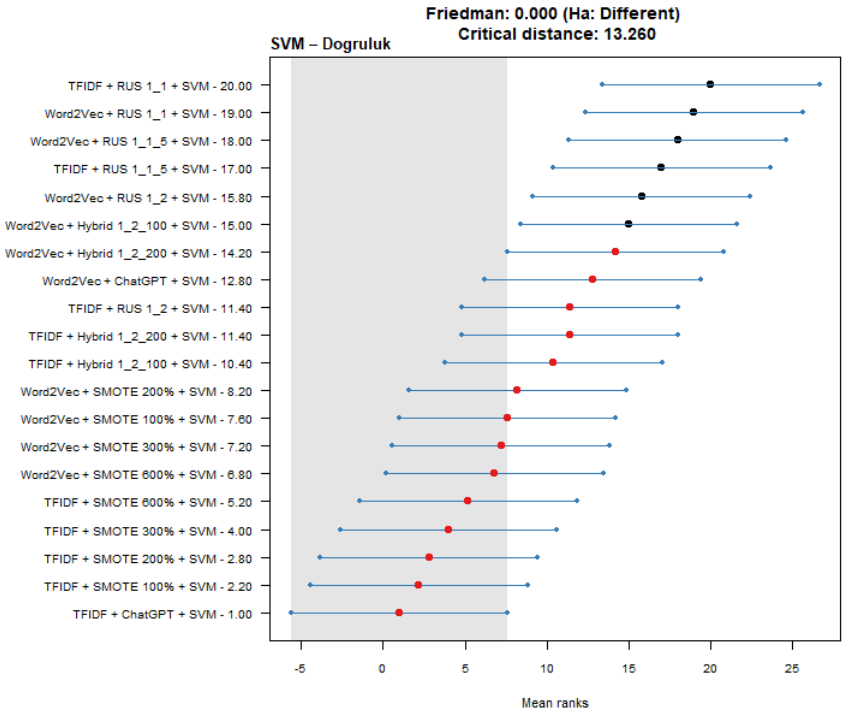
Ek Şekil 9.16. RF – Geometrik Ortalama (Aşama 1)



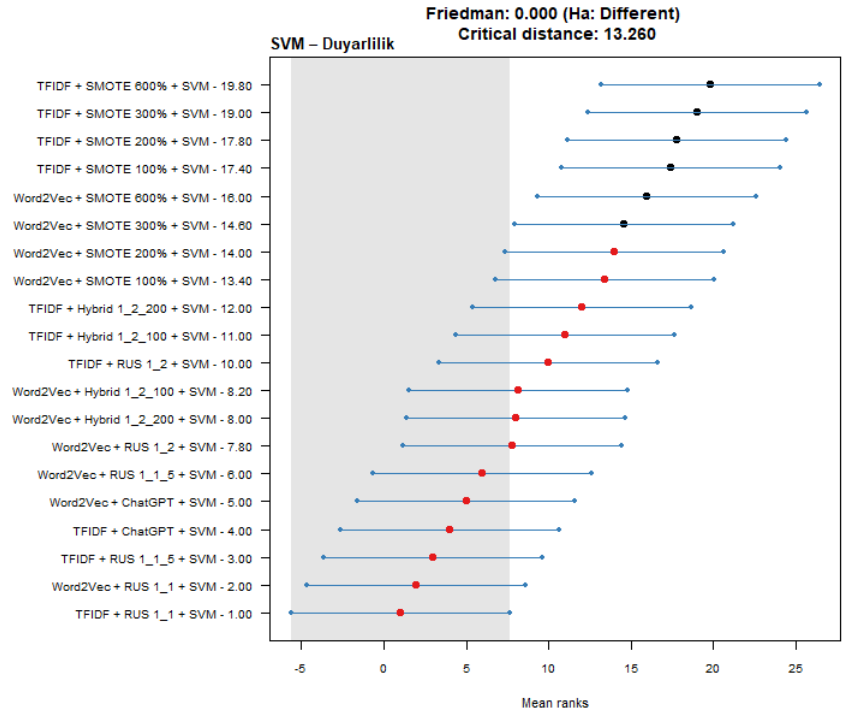
Ek Şekil 9.17. RF – IBA (Aşama 1)



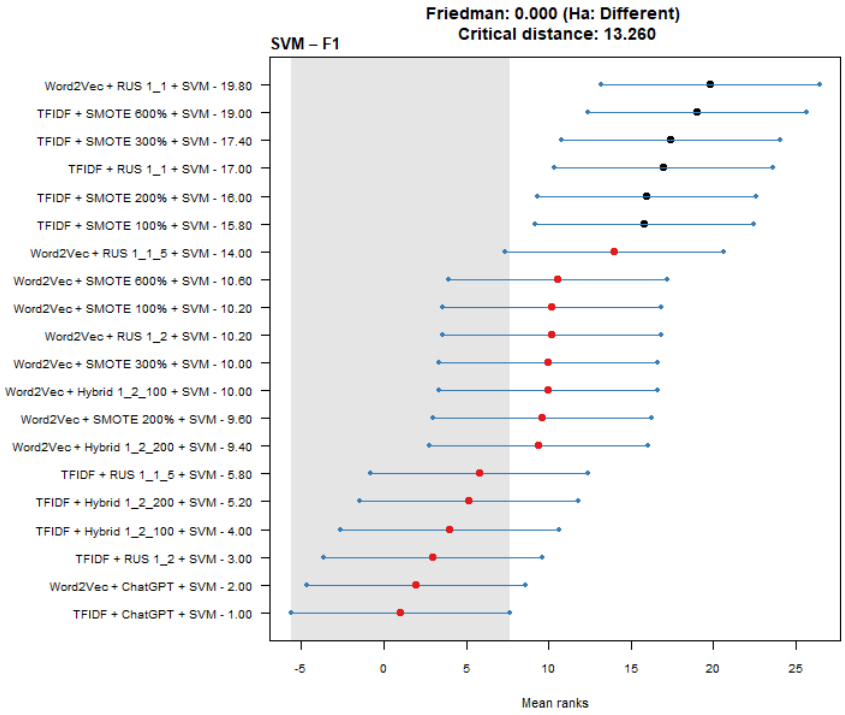
Ek Şekil 9.18. RF – Kesinlik (Aşama 1)



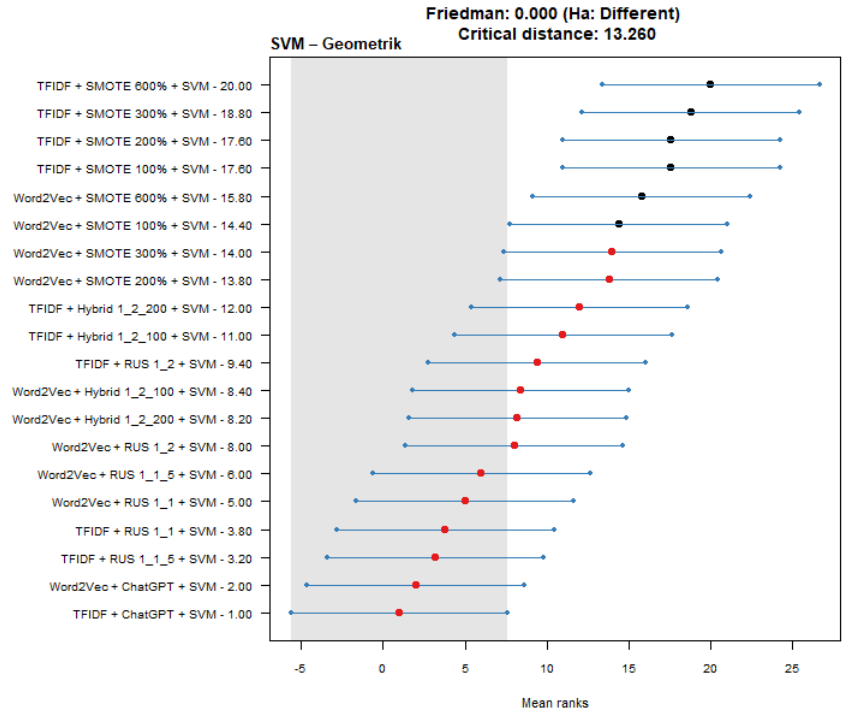
Ek Şekil 9.19. SVM – Doğruluk (Aşama 1)



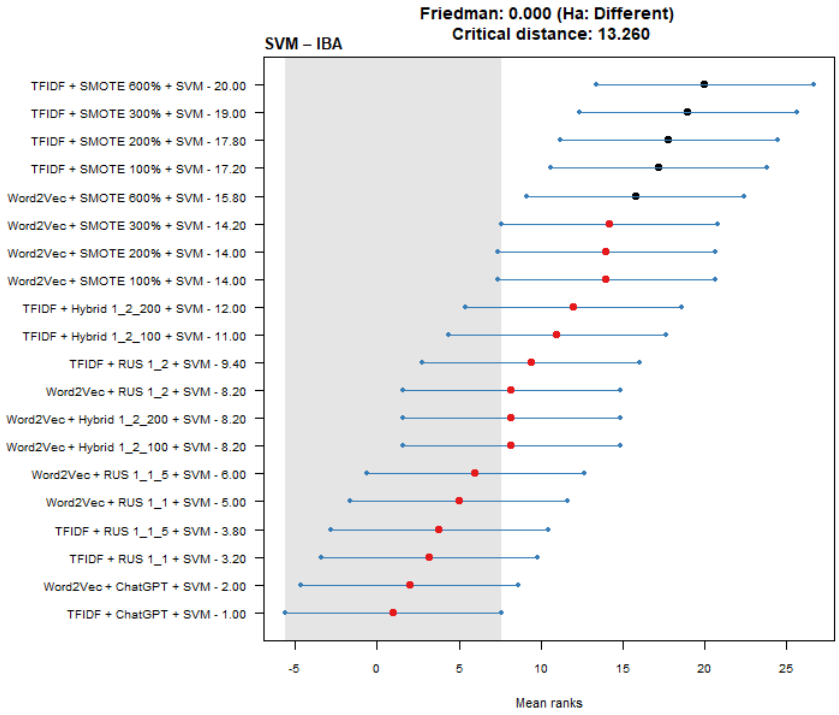
Ek Şekil 9.20. SVM – Duyarlılık (Aşama 1)



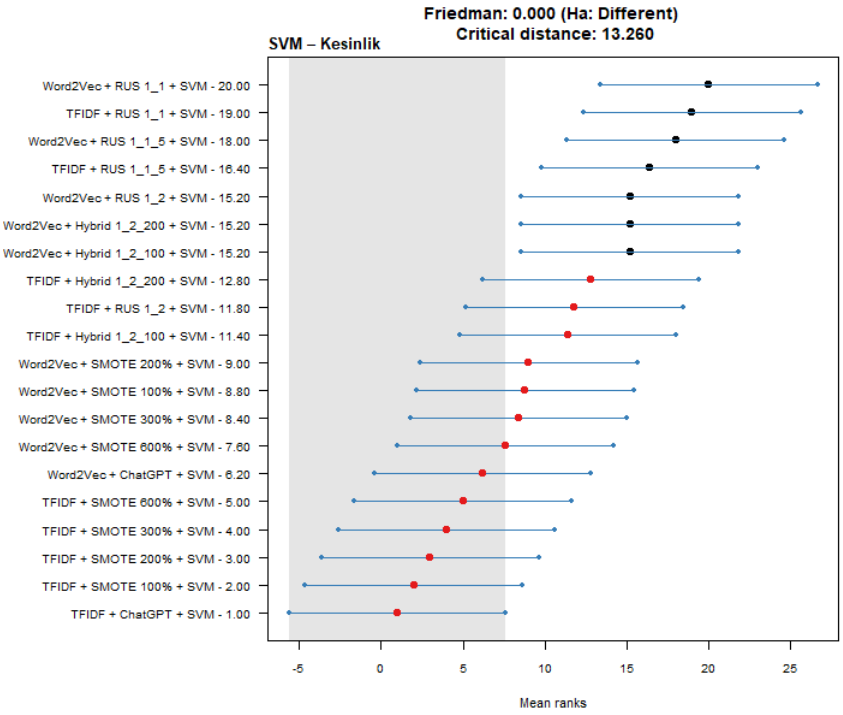
Ek Şekil 9.21. SVM – F-ölçütü (Aşama 1)



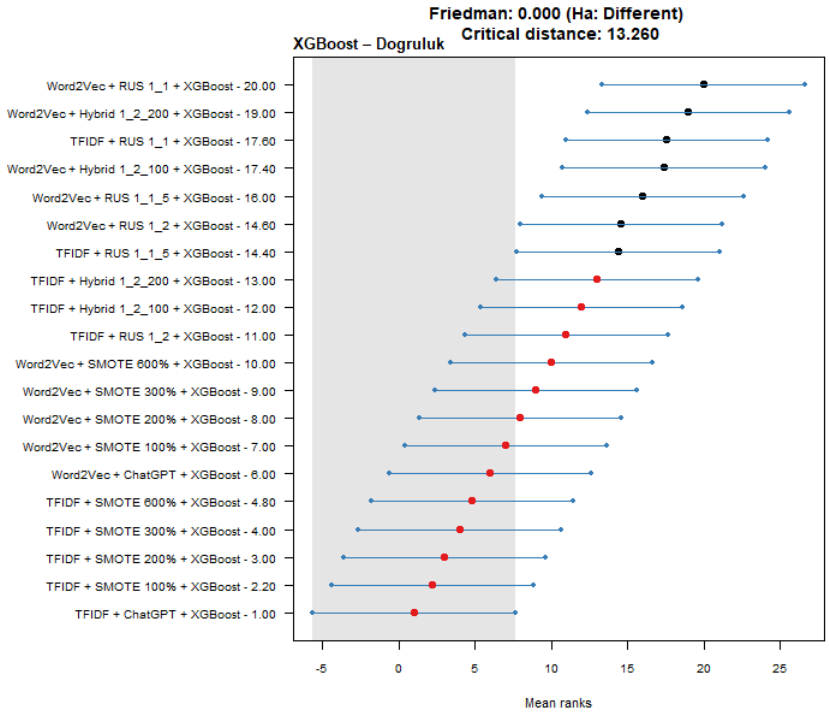
Ek Şekil 9.22. SVM – Geometrik Ortalama (Aşama 1)



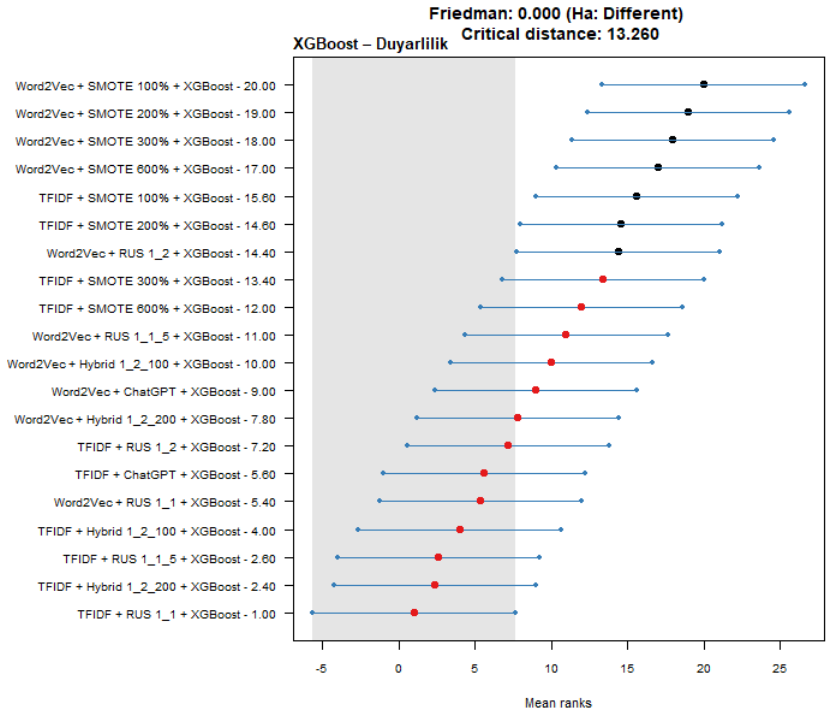
Ek Şekil 9.23. SVM – IBA (Aşama 1)



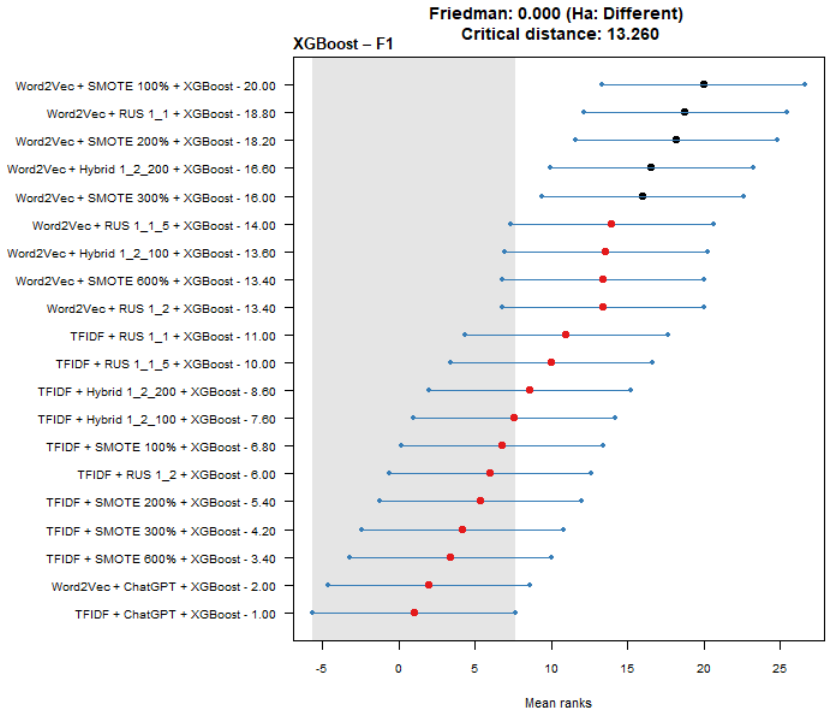
Ek Şekil 9.24. SVM – Kesinlik (Aşama 1)



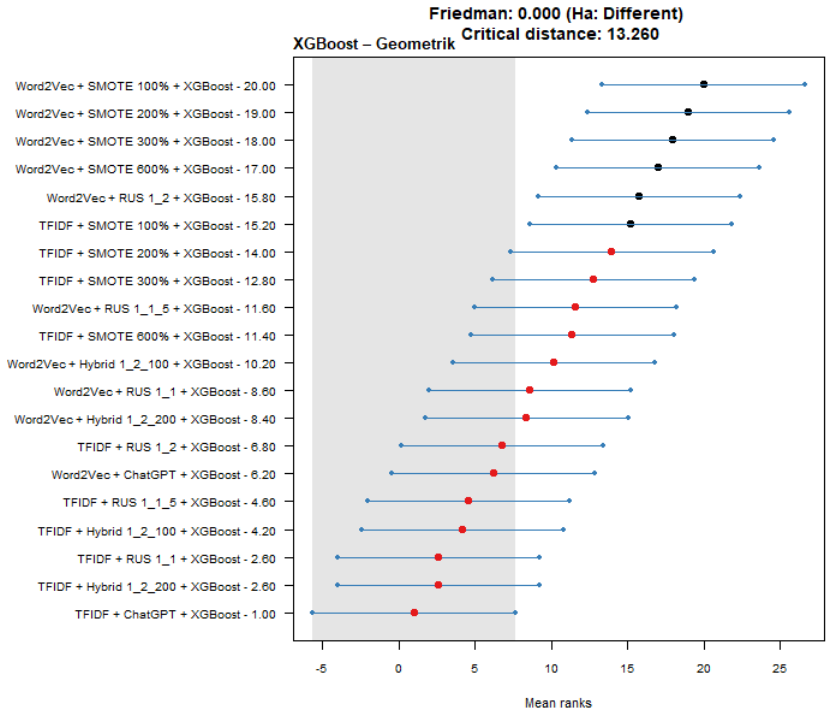
Ek Şekil 9.25. XGBoost – Doğruluk (Aşama 1)



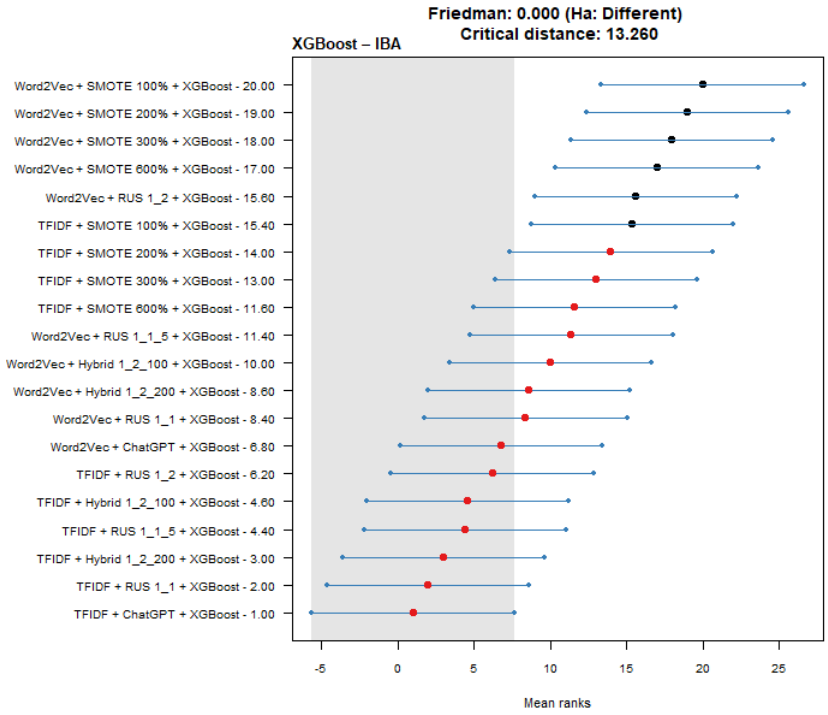
Ek Şekil 9.26. XGBoost – Duyarlılık (Aşama 1)



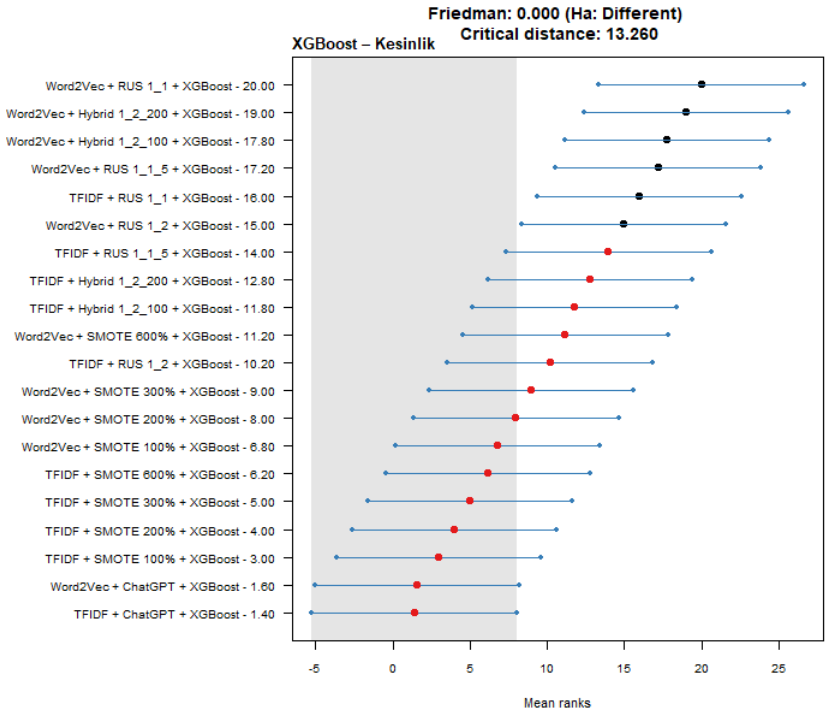
Ek Şekil 9.27. XGBoost – F-ölçütü (Aşama 1)



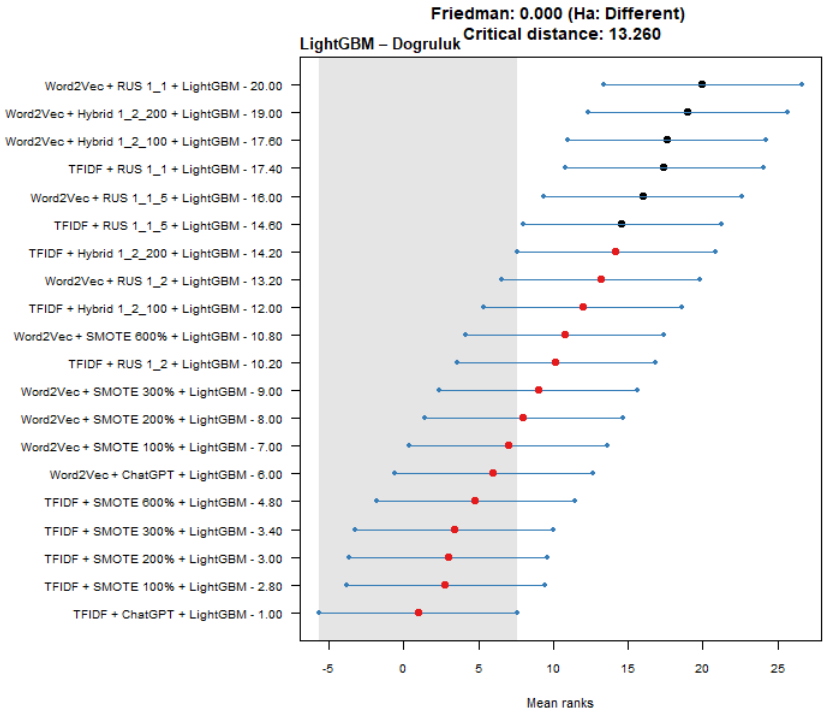
Ek Şekil 9.28. XGBoost – Geometrik Ortalama (Aşama 1)



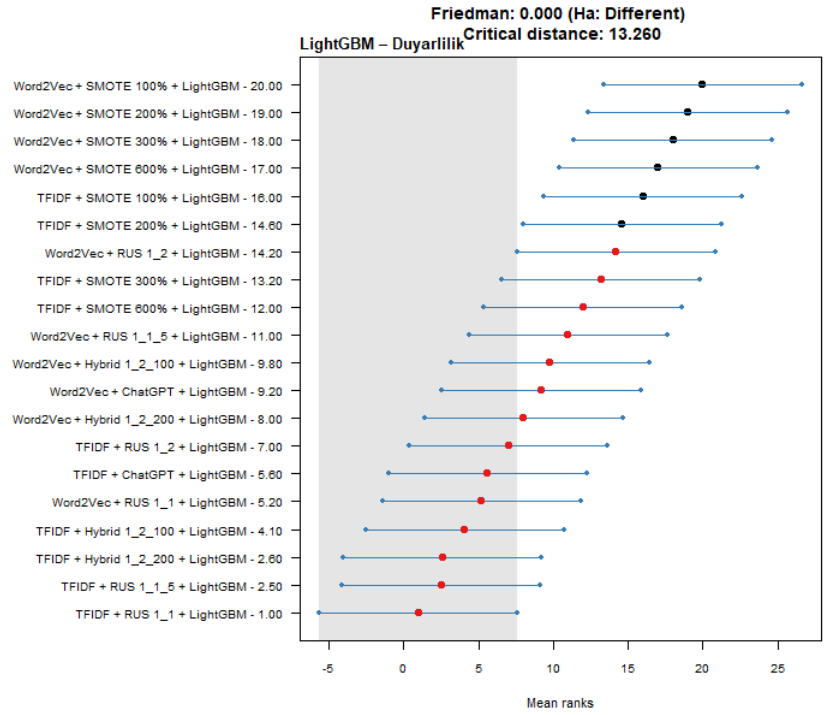
Ek Şekil 9.29. XGBoost – IBA (Aşama 1)



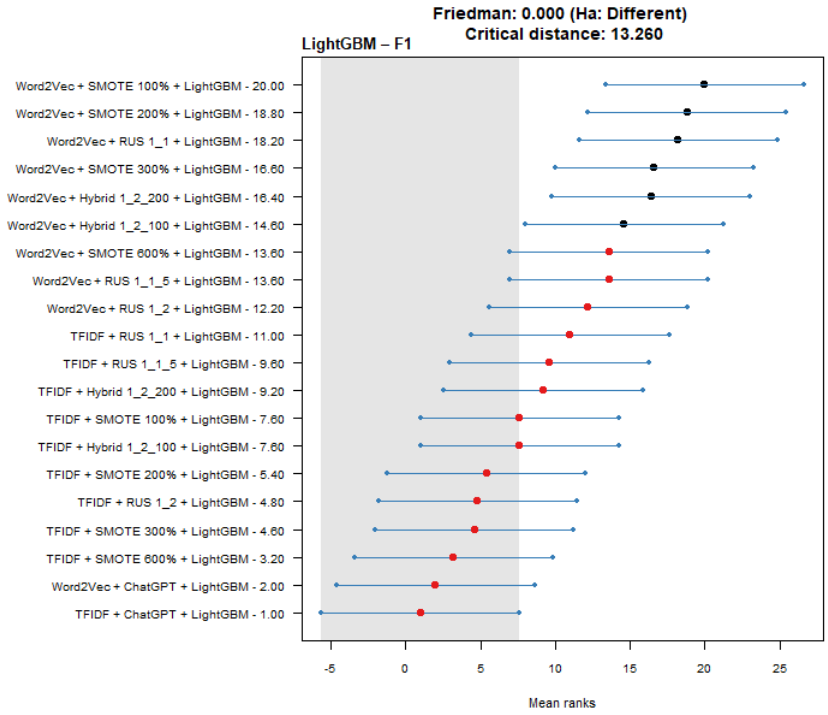
Ek Şekil 9.30. XGBoost – Kesinlik (Aşama 1)



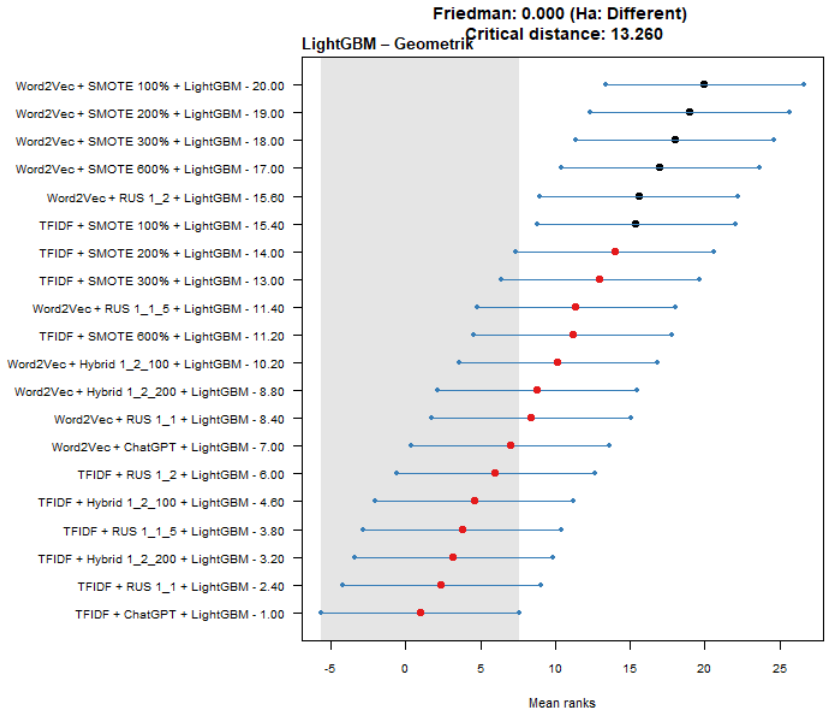
Ek Şekil 9.31. LightGBM – Doğruluk (Aşama 1)



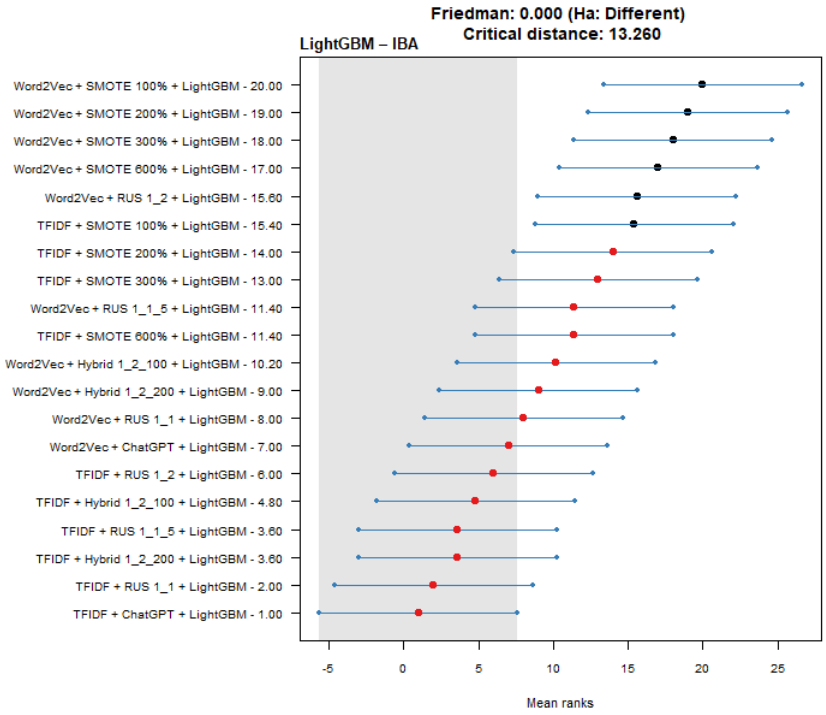
Ek Şekil 9.32. LightGBM – Duyarlılık (Aşama 1)



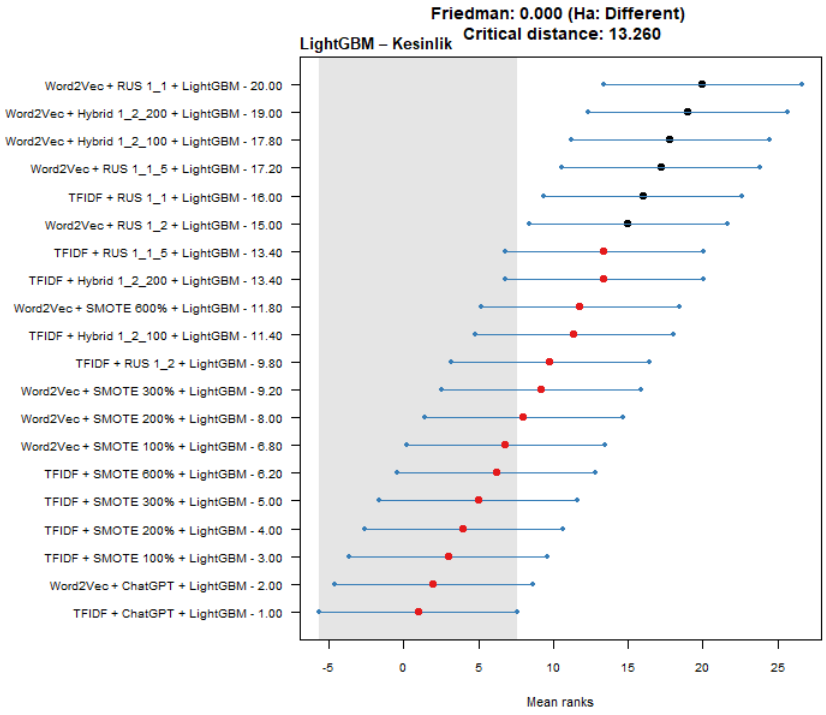
Ek Şekil 9.33. LightGBM – F-ölçütü (Aşama 1)



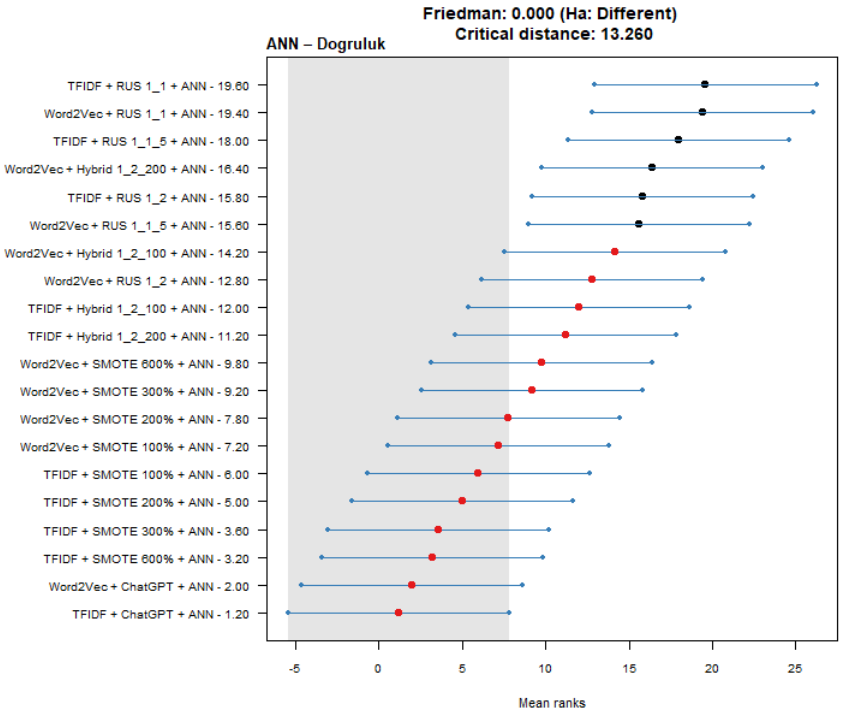
Ek Şekil 9.34. LightGBM – Geometrik Ortalama (Aşama 1)



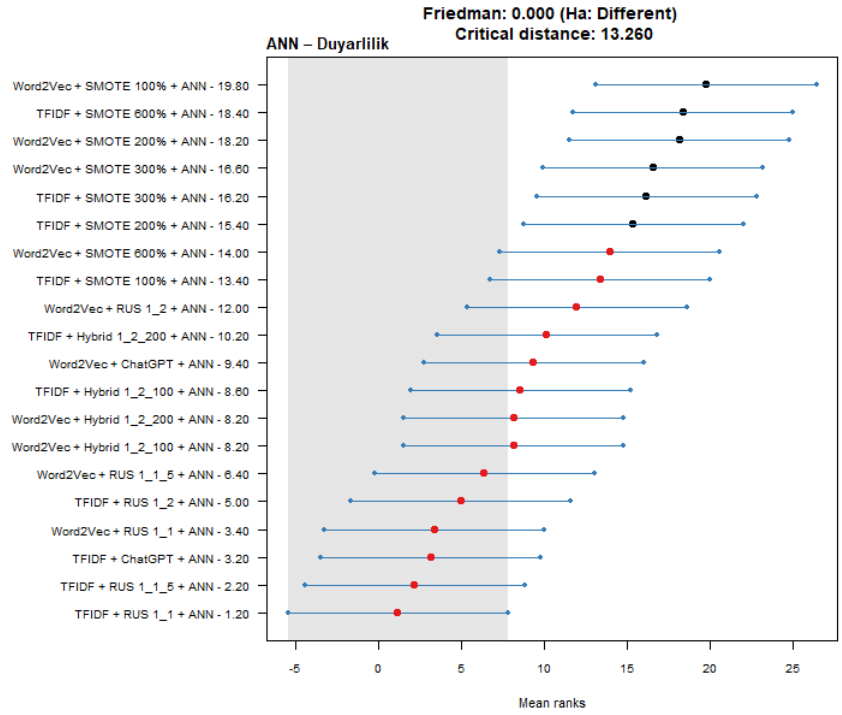
Ek Şekil 9.35. LightGBM – IBA (Aşama 1)



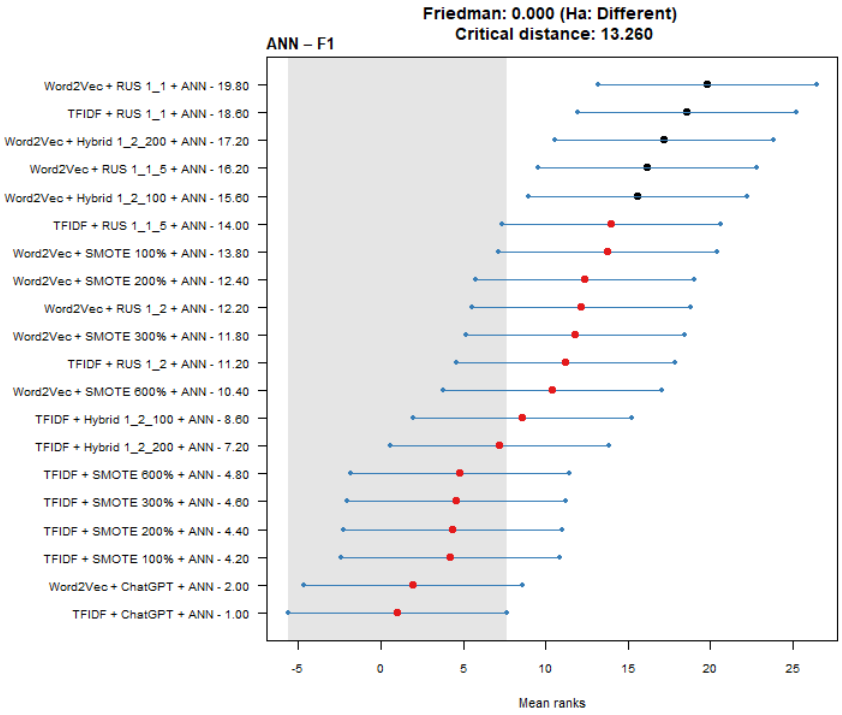
Ek Şekil 9.36. LightGBM – Kesinlik (Aşama 1)



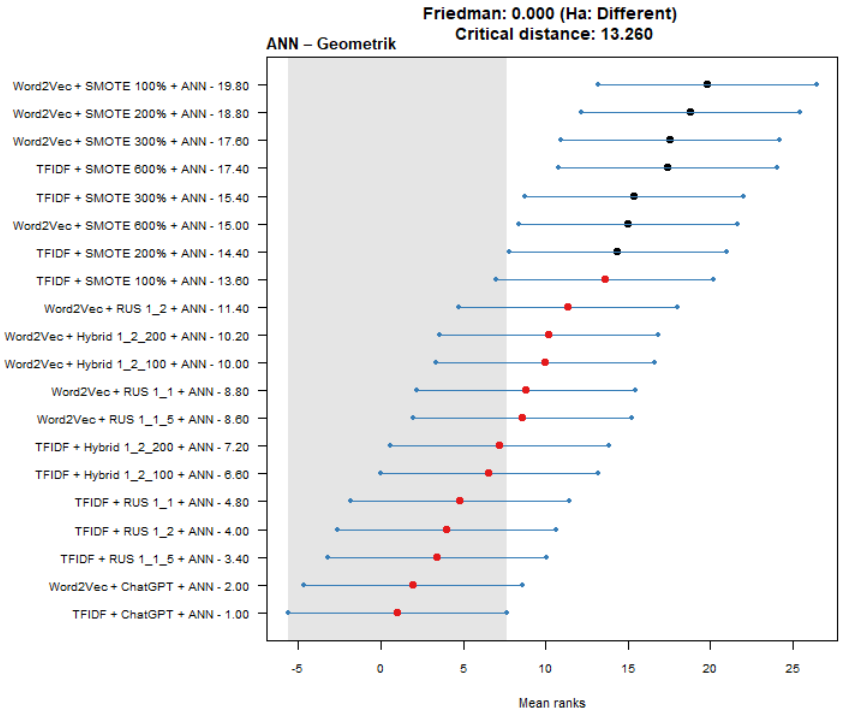
Ek Şekil 9.37. ANN – Doğruluk (Aşama 1)



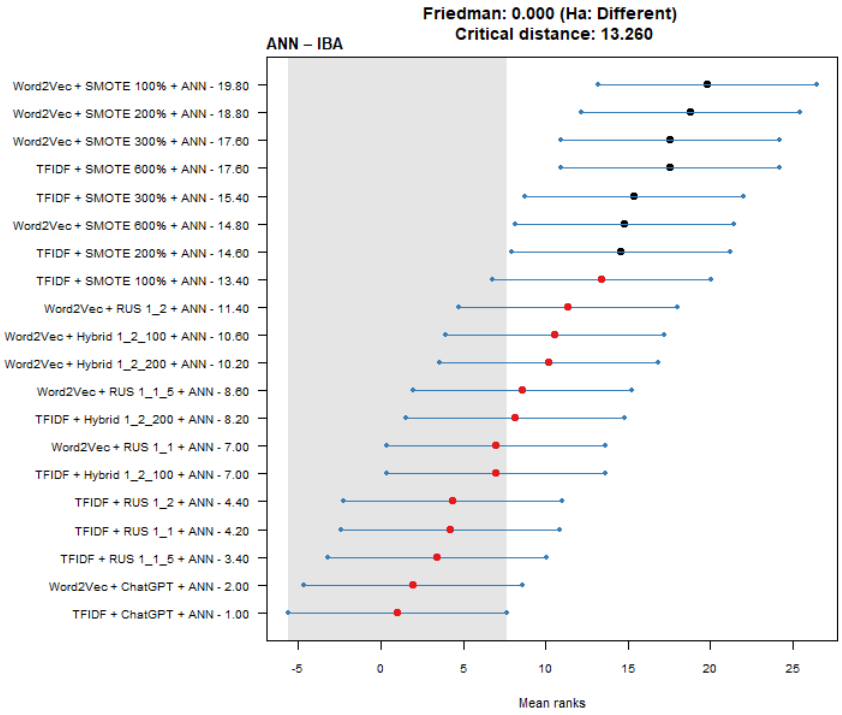
Ek Şekil 9.38. ANN – Duyarlılık (Aşama 1)



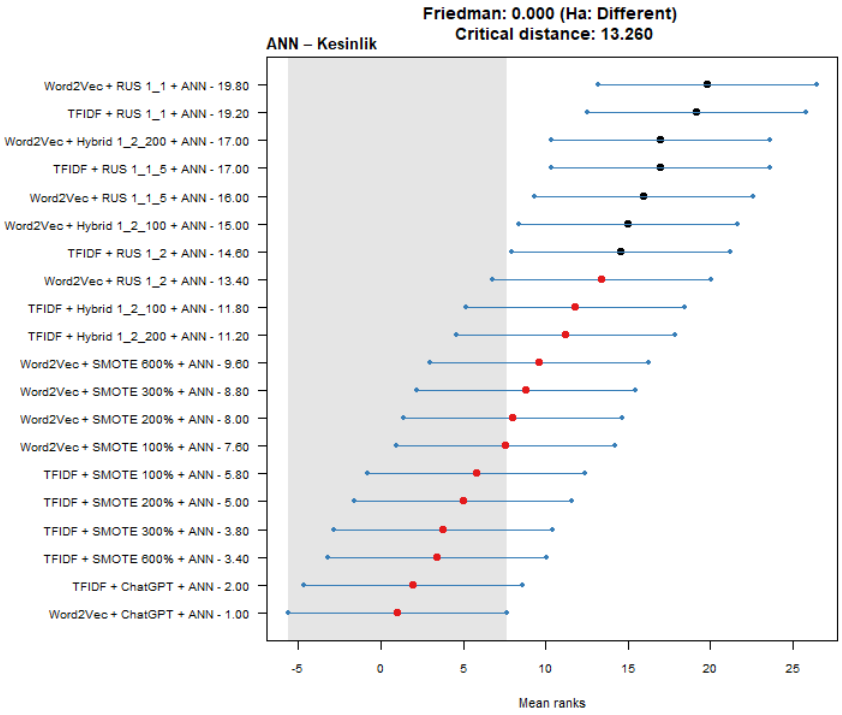
Ek Şekil 9.39. ANN – F-ölçütü (Aşama 1)



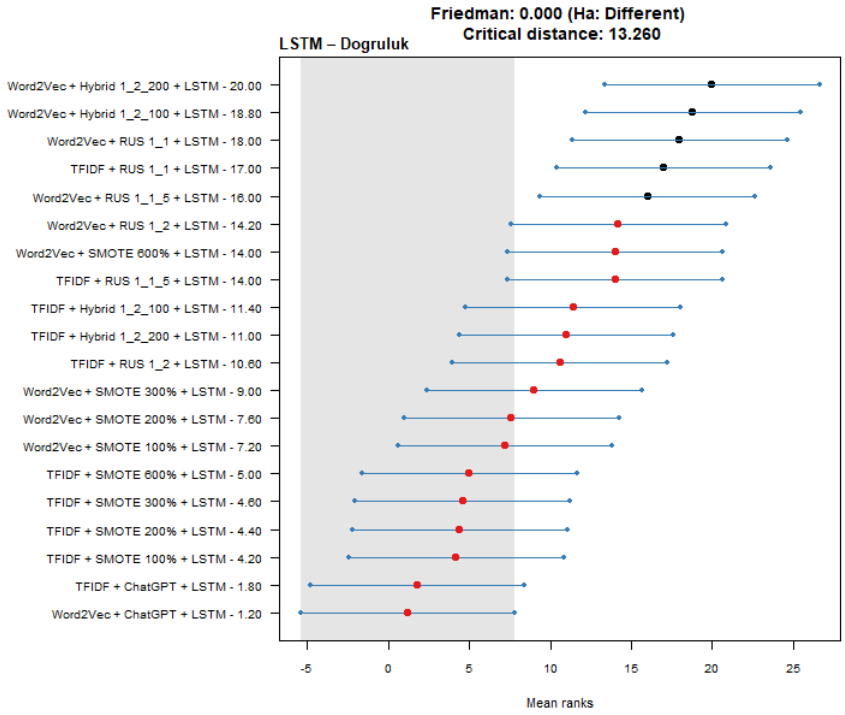
Ek Şekil 9.40. ANN – Geometrik Ortalama (Aşama 1)



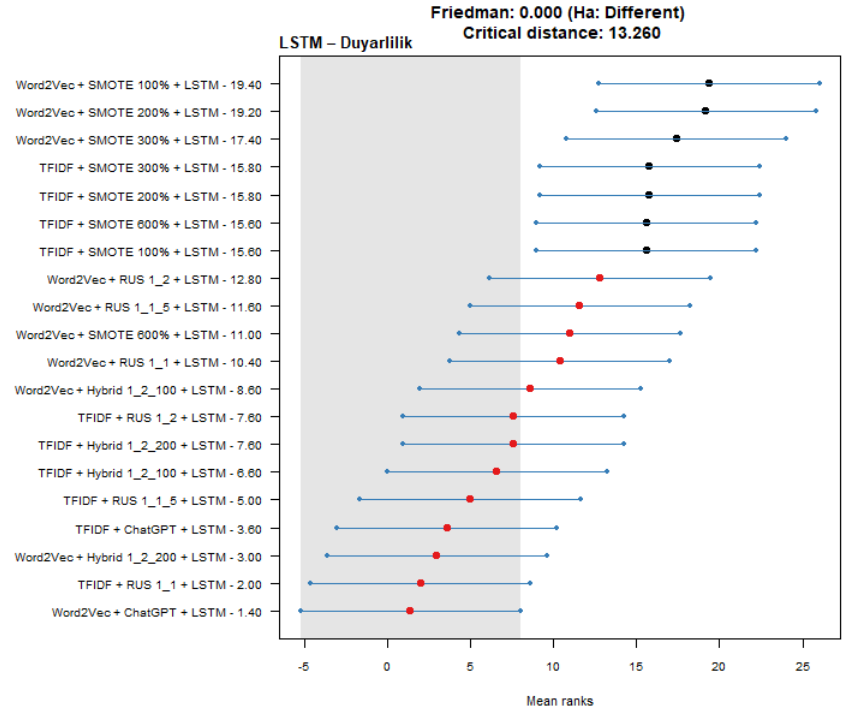
Ek Şekil 9.41. ANN – IBA (Aşama 1)



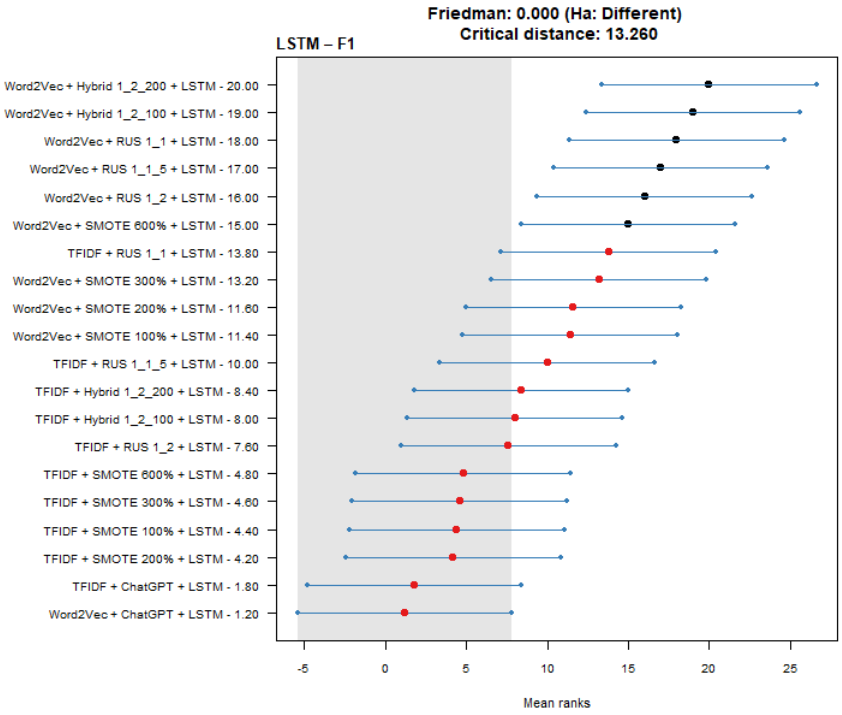
Ek Şekil 9.42. ANN – Kesinlik (Aşama 1)



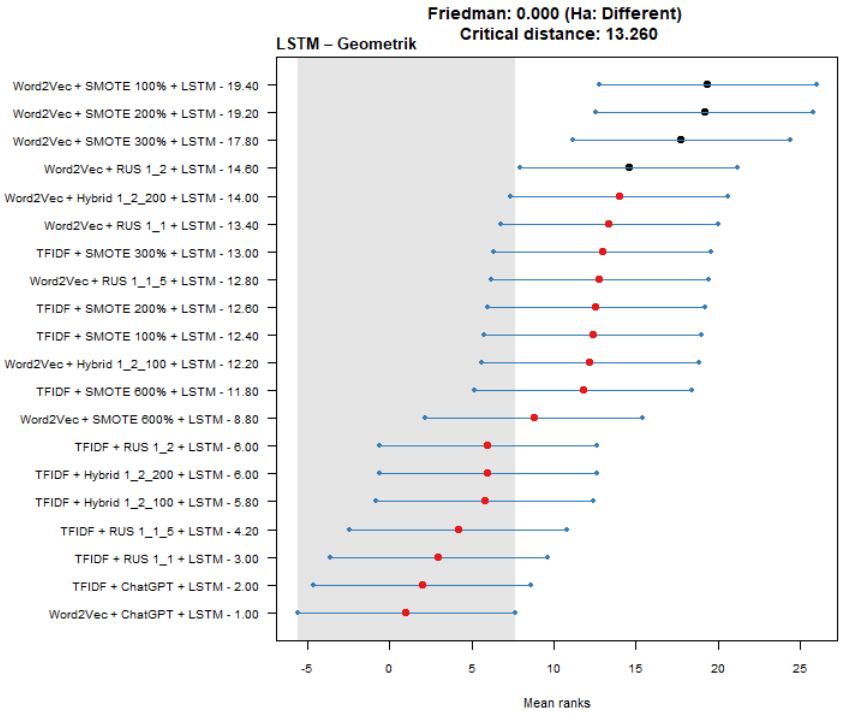
Ek Şekil 9.43. LSTM – Doğruluk (Aşama 1)



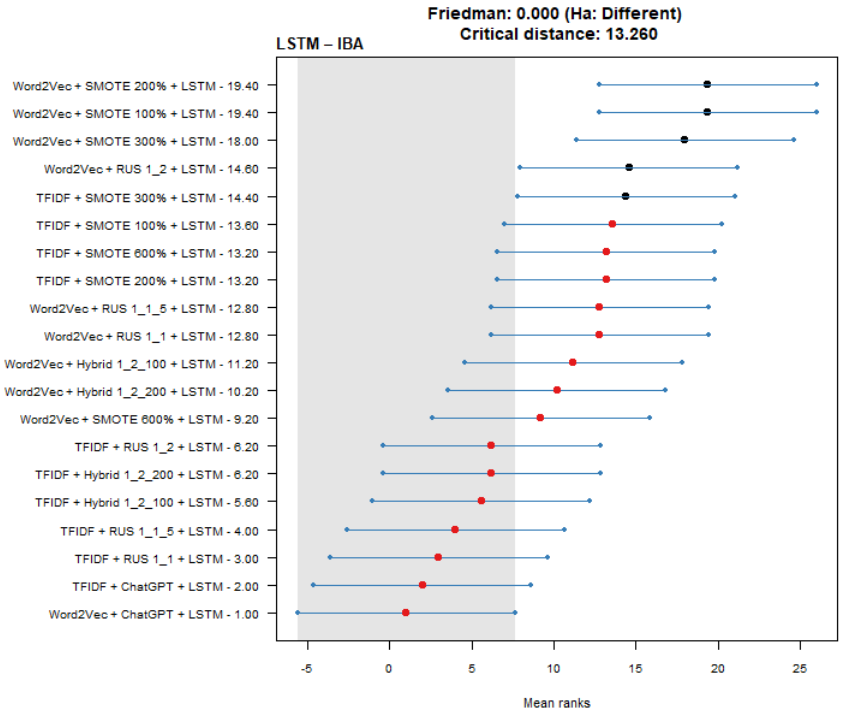
Ek Şekil 9.44. LSTM – Duyarlılık (Aşama 1)



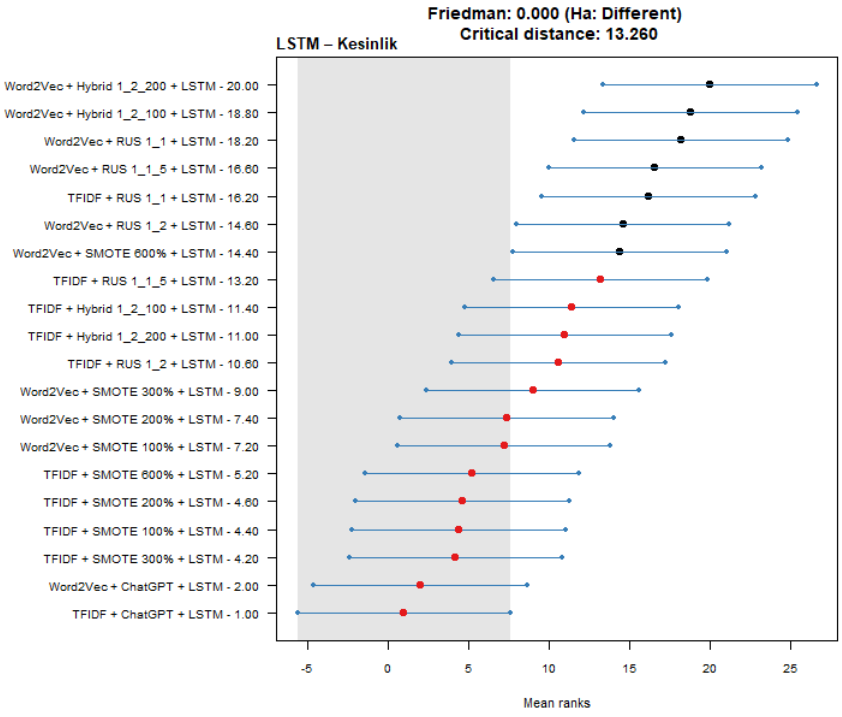
Ek Şekil 9.45. LSTM – F-ölçütü (Aşama 1)



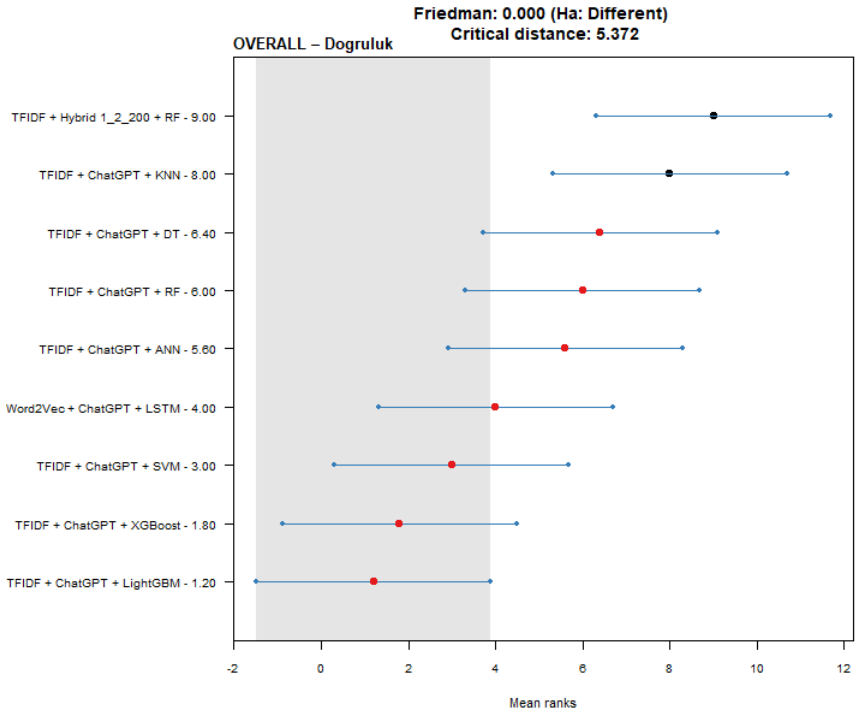
Ek Şekil 9.46. LSTM – Geometrik Ortalama (Aşama 1)



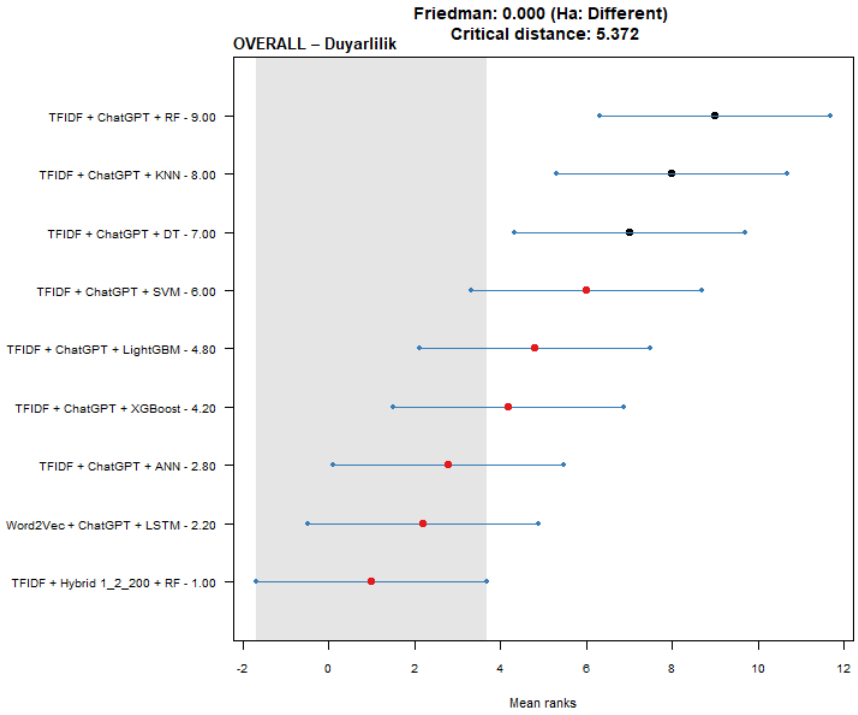
Ek Şekil 9.47. LSTM – IBA (Aşama 1)



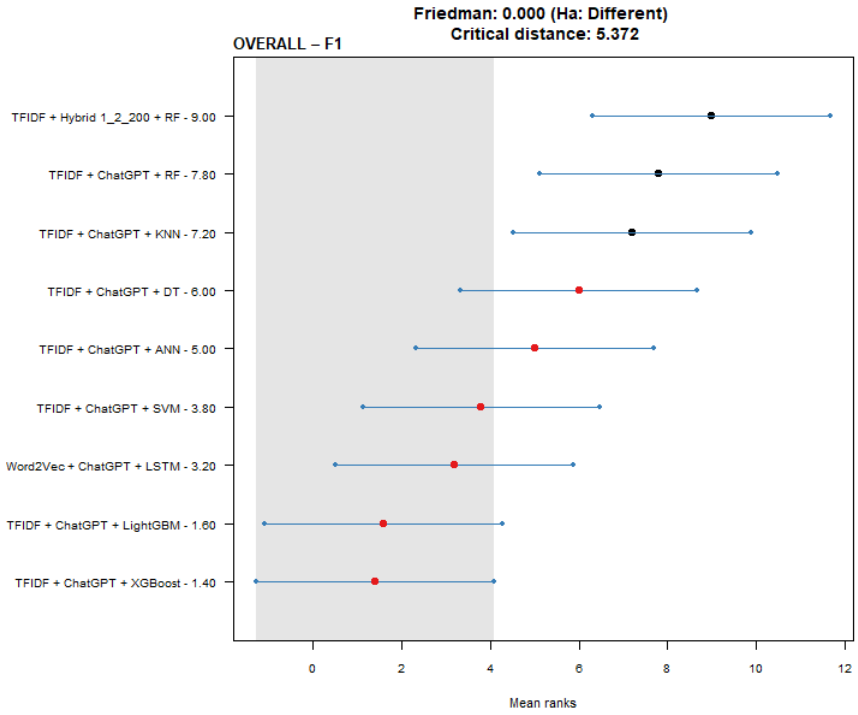
Ek Şekil 9.48. LSTM – Kesinlik (Aşama 1)



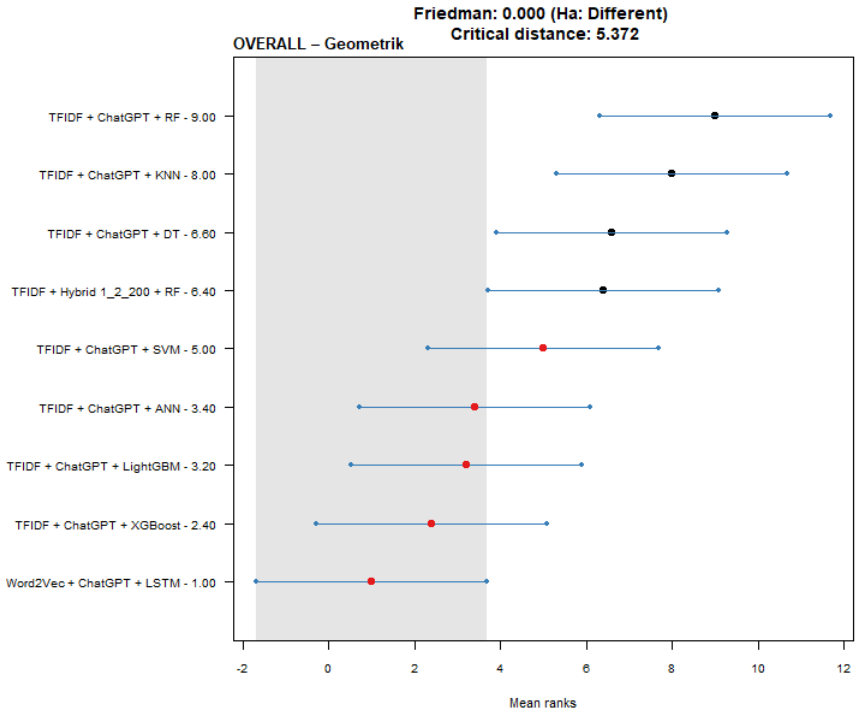
Ek Şekil 9.49. Doğruluk (Aşama 2 – Tüm Modeller)



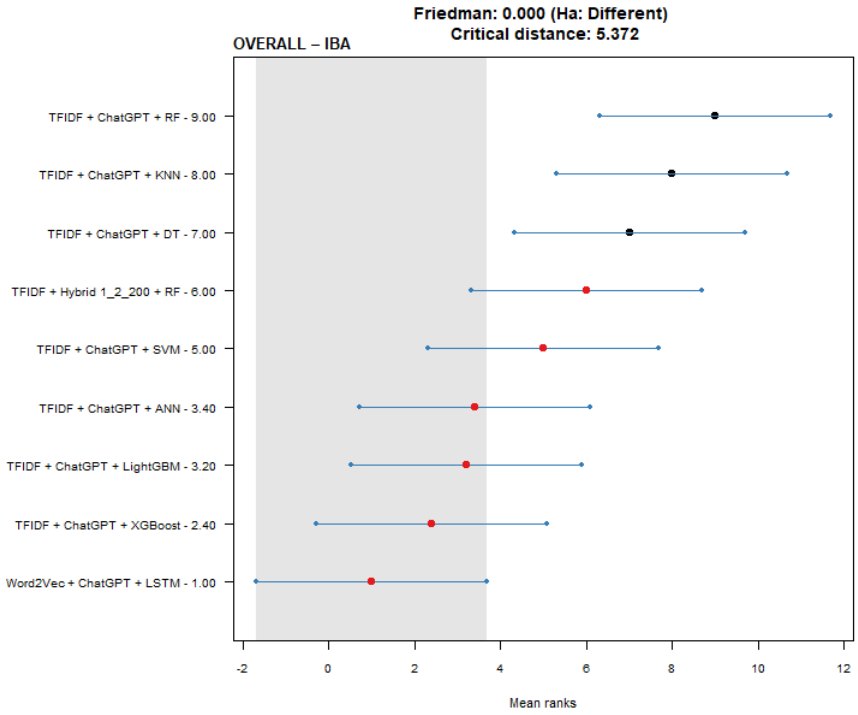
Ek Şekil 9.50. Duyarlılık (Aşama 2 – Tüm Modeller)



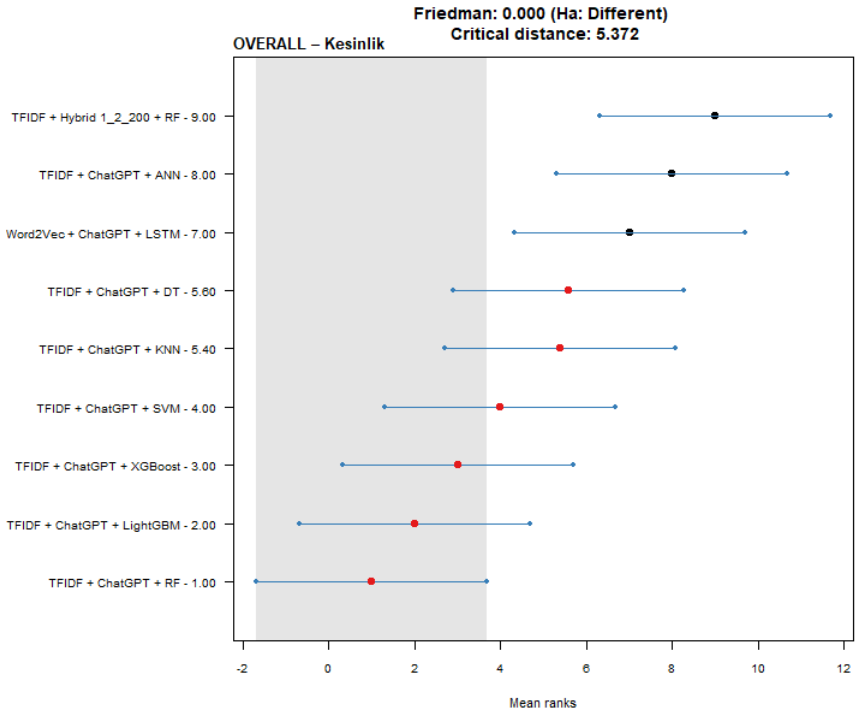
Ek Şekil 9.51. F-ölçütü (Aşama 2 – Tüm Modeller)



Ek Şekil 9.52. Geometrik Ortalama (Aşama 2 – Tüm Modeller)



Ek Şekil 9.53. IBA (Aşama 2 – Tüm Modeller)



Ek Şekil 9.54. Kesinlik (Aşama 2 – Tüm Modeller)

Özgeçmiş

Adı Soyadı : Mustafa Emirkan Okursoy

Doğum Yeri ve Tarihi : Nilüfer / 06.04.1999

Yabancı Dil : İngilizce

Eğitim Durumu

Lise : Özel Bursa Tan Anadolu Lisesi

Lisans : İstanbul Teknik Üniversitesi İşletme Mühendisliği

Yüksek Lisans : Bursa Uludağ Üniversitesi Endüstri Mühendisliği

Çalıştığı Kurum/Kurumlar : Bursa Uludağ Üniversitesi Mühendislik Fakültesi, Endüstri Mühendisliği Bölümü

İletişim (e-posta) : emirkanokursoy@uludag.edu.tr

Yayınları : Okursoy, M. E., & İnkaya, T. (2025). Yüksek Mahkeme kararlarının sınıflandırılması için veri dengeleme ve açıklanabilir yapay zekâ tabanlı karar destek sistemi. Veri Bilimi, 8(1), 11–23. https://dergipark.org.tr/tr/pub/veri/issue/92570/1557010