



**KTO KARATAY ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
ELEKTRİK – ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI
ELEKTRİK VE BİLGİSAYAR MÜHENDİSLİĞİ TEZLİ YÜKSEK LİSANS
PROGRAMI**

**MAKİNE ÖĞRENMESİ ALGORİTMALARI İLE EPİLEPSİLİ ÇOCUKLARIN
KLİNİK VE LABORATUVAR BULGULARININ İNCELENMESİ**

Osman ÇAKSEN

Yüksek Lisans Tezi

**KONYA
Ocak 2022**

MAKİNE ÖĞRENMEİ ALGORİTMALARI İLE EPİLEPSİLİ ÇOCUKLARIN
KLİNİK VE LABORATUVAR BULGULARININ İNCELENMESİ

Osman ÇAKSEN

KTO Karatay Üniversitesi
Lisansüstü Eğitim Enstitüsü
Elektrik – Elektronik Mühendisliği Anabilim Dalı
Elektrik Ve Bilgisayar Mühendisliği Tezli Yüksek Lisans Programı

Yüksek Lisans Tezi

Tez Danışmanı: Dr. Öğr. Üyesi Ali ÖZTÜRK

Konya
Ocak 2022

BİLDİRİM

Enstitü tarafından onaylanan Yüksek Lisans tezimin tamamını veya herhangi bir kısmını basılı veya dijital biçimde arşivleme ve aşağıda belirtilen koşullar dahilinde erişime açma iznini KTO Karatay Üniversitesine verdiğimi bildiririm. Bu izinle, Üniversiteye verilen kullanım hakları dışındaki tüm fikri mülkiyet haklarım bende kalacak ve gelecekteki çalışmalar (makale, kitap, lisans, patent vb.) için tezimin tamamının veya bir bölümünün kullanım hakları yalnızca bana ait olacaktır.

Tezimin bütünüyle kendi çalışmam olduğunu, başkalarının haklarını ihlal etmediğimi ve tezimin tek yetkili sahibi olduğumu beyan ve taahhüt ederim. Telif hakkı bulunan ve sahiplerinden yazılı izinle kullanılması zorunlu olan kaynakları, yazılı izin alarak kullandığımı ve istenildiğinde izinlerin suretlerini Üniversiteye teslim etmeyi taahhüt ederim.

Yükseköğretim Kurulu tarafından yayımlanan “Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge” kapsamında, tezim, aşağıda belirtilen koşullar haricince, YÖK Ulusal Tez Merkezi ve KTO Karatay Üniversitesi Açık Erişim Sisteminde erişime açılır.

☐ Enstitü / Fakülte Yönetim Kurulu kararı ile tezimin erişime açılması mezuniyet tarihimden itibaren 2 yıl ertelenmiştir.¹

☐ Enstitü / Fakülte Yönetim Kurulunun gerekçeli kararı ile tezimin erişime açılması mezuniyet tarihimden itibaren ... ay ertelenmiştir.²

☐ Tezimle ilgili gizlilik kararı verilmiştir.³⁴

26 Ocak 2022

Osman ÇAKSEN

¹ MADDE 6(1) Lisansüstü tezle ilgili patent başvurusu yapılması veya patent alma sürecinin devam etmesi durumunda, tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulu iki yıl süre ile tezin erişime açılmasının ertelenmesine karar verebilir.

² MADDE 6(2) Yeni teknik, materyal ve metotların kullanıldığı, henüz makaleye dönüşmemiş veya patent gibi yöntemlerle korunmamış ve internetten paylaşılması durumunda 3. şahıslara veya kurumlara haksız kazanç imkanı oluşturabilecek bilgi ve bulguları içeren tezler hakkında tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulunun gerekçeli kararı ile altı ayı aşmamak üzere tezin erişime açılması engellenebilir.

³ MADDE 7(1) Ulusal çıkarları veya güvenliği ilgilendiren, emniyet, istihbarat, savunma ve güvenlik, sağlık vb. konulara ilişkin lisansüstü tezlerle ilgili gizlilik kararı, tezin yapıldığı kurum tarafından verilir. Kurum ve kuruluşlarla yapılan iş birliği protokolü çerçevesinde hazırlanan lisansüstü tezlere ilişkin gizlilik kararı ise, ilgili kurum ve kuruluşun önerisi ile enstitü veya fakültenin uygun görüşü üzerine üniversite yönetim kurulu tarafından verilir. Gizlilik kararı verilen tezler Yükseköğretim Kuruluna bildirilir.

⁴ MADDE 7(2) Gizlilik kararı verilen tezler gizlilik süresince enstitü veya fakülte tarafından gizlilik kuralları çerçevesinde muhafaza edilir, gizlilik kararının kaldırılması halinde Tez Otomasyon Sistemine yüklenir.

ETİK BEYAN

KTO Karatay Üniversitesi Lisansüstü Eğitim Enstitüsü Tez/Proje Hazırlama ve Yazım Kurallarına uygun olarak Dr. Öğr. Üyesi Ali ÖZTÜRK danışmanlığında tarafımdan üretilen bu tez çalışmasında; sunduğum tüm veri, enformasyon, bilgi ve belgeleri bilimsel etik kuralları çerçevesinde elde ettiğimi, tüm değerlendirme, analiz, bulgu ve sonuçları bilimsel usullere uygun olarak sunduğumu, tez çalışmasında yararlandığım kaynakların tümüne bilimsel normlara uygun biçimde atıfta bulunarak kaynak gösterdiğimi, tezimin/projemin kaynak gösterilen durumlar dışında özgün olduğunu bildirir, aksi bir durumda aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi beyan ederim.

26 Ocak 2022

Osman ÇAKSEN

TEŞEKKÜR

Çalışmalarım boyunca kıymetli tecrübelerinden faydalandığım, yardım ve katkılarıyla beni yönlendiren, sabırla beni destekleyen akademik danışmanım Dr. Öğr. Üyesi Ali ÖZTÜRK'e teşekkürlerimi ve saygılarımı sunarım. Karşılaştığım her zorlukta yanımda olan, hiçbir fedakarlıktan kaçınmayan en büyük şansım babam Prof. Dr. Hüseyin ÇAKSEN'e, annem Zehra ÇAKSEN'e, kardeşlerime ve bu süreçte desteklerini esirgemeyen değerli arkadaşım Arş. Gör. Usame DURAK'a teşekkürlerimi sunarım.

26 Ocak 2022

Osman ÇAKSEN



ÖZET

Osman ÇAKSEN

Makine Öğrenmesi Algoritmalarıyla Epilepsili Çocukların Klinik Ve Laboratuvar

Bulgularının İncelenmesi

Yüksek Lisans Tezi

Konya, 2022

Dirençli epilepsi hastalarının tanımlanması kolay değildir ve uzun deneme yanılma süreçleri gerekmektedir. Bu çalışmada dirençli epilepsi tanısını desteklemek için hasta öykülerinden oluşturulan verilere makine öğrenmesi yöntemleri uygulanmıştır.

Bu amaçla, 500 epilepsili çocuk hastanın öykülerinden oluşturulan veri seti kullanılarak karar ağaçları, destek vektör makineleri, rastgele orman, yapay sinir ağları, lojistik regresyon, k-en yakın komşu ve Gaussian Naive Bayes algoritmaları ile sınıflandırma işlemi gerçekleştirilmiştir. Algoritmaların performansını belirlemek için doğruluk, özgüllük, hassasiyet ve f-skor istatistiksel ölçütleri kullanılmıştır. Elde edilen sonuçlara göre rastgele orman algoritması 0.85 doğruluk değeri, 0.86 özgüllük değeri, 0.98 hassasiyet değeri ve 0.90 f-skor ile en iyi performansa sahiptir.

İkincil olarak, bağımsız değişken olan dirençli epilepsiye etkisini görebilmek için, “sınıflandırıcı ile özellik değerlendirici” özellik seçimi yöntemi ile elde edilen özellikler (kullanılan ilaç sayısı, akraba evliliği, ailede febril konvülsiyon öyküsü, ailede epilepsi olma durumu, ilk elektroensefalografi ve mental retardasyon) ile yeni bir veri seti oluşturulmuştur.

Aynı algoritmalar yeni veri kümesine uygulanmış ve performanslarını belirlemek için doğruluk, özgüllük, hassasiyet ve f-skor ölçümleri kullanılmıştır. Elde edilen sonuçlara göre en yüksek doğruluk ölçütü 0.85 ile destek vektör makinelerine, en yüksek özgüllük değeri 0.98 ile destek vektör makineleri ve rastgele orman algoritmalarına, en yüksek hassasiyet 0.87 ile Gaussian Naive Bayes ve yapay sinir ağları algoritmalarına, en yüksek f-skor 0.92 ile destek vektör makineleri algoritmasına aittir.

Anahtar Kelimeler:

Epilepsi, retrospektif, febril konvülsiyon, EEG, mental retardasyon

ABSTRACT

Osman ÇAKSEN

Examination Of Clinical And Laboratory Findings Of Children With Epilepsy Using Machine Learning Algorithms

Master's Thesis

Konya, 2022

Identification of patients with resistant epilepsy is not easy and requires long trial and error processes. In this study, machine learning methods were applied to data generated from patient histories to support the diagnosis of resistant epilepsy.

For this purpose, the data set created from the stories of 500 children with epilepsy, decision trees, support vector machines, random forest, artificial neural networks, logistic regression, k nearest neighbor and Gaussian Naive Bayes algorithms were used for classification. Statistical measures of accuracy, specificity, sensitivity and f-score were used to determine the performance of the algorithms. According to the results obtained, the random forest algorithm has the best performance with 0.85 accuracy value, 0.86 specificity value, 0.98 sensitivity value and 0.90 f-score.

Secondly, in order to see its effect on resistant epilepsy which is the independent variable, a new data set was created with features (number of drugs used, consanguineous marriage, family history of febrile convulsions, epilepsy in the family, first electroencephalography and mental retardation) obtained by the classifier and feature evaluator feature selection method.

The same algorithms were applied to this new data set and measures of accuracy, specificity, sensitivity and f-score were used to determine their performances. According to the results obtained, the highest accuracy criterion belongs to support vector machines with 0.85, the highest specificity value belongs to support vector machines and random forest with 0.98, the highest sensitivity belongs to Gaussian Naive Bayes and artificial neural networks with 0.87, and the highest f-score belongs to support vector machines with 0.92.

Keywords:

Epilepsy, retrospective, febrile convulsion, EEG, mental retardation

İÇİNDEKİLER

KABUL VE ONAY	i
BİLDİRİM.....	ii
ETİK BEYAN	iii
TEŞEKKÜR	iv
ÖZET	v
ABSTRACT	vi
İÇİNDEKİLER.....	vii
TABLolar DİZİNİ.....	ix
ŞEKİLLER DİZİNİ	x
KISALTMALAR DİZİNİ	xi
1. GİRİŞ.....	1
2. LİTERATÜR ÖZETİ	3
3. EPİLEPSİ	9
3.1. Epidemiyoloji.....	9
3.2. Etiyoloji	9
3.3. Epilepsinin Sınıflaması	11
3.4. Tanı	13
3.5. Ayırıcı Tanı.....	13
3.6. Tedavi	17
4. YAPAY ZEKÂ.....	26
4.1. Makine Öğrenmesi.....	27
4.2. Sağlıkta Yapay Zekâ	34
5. ÇALIŞMADA KULLANILAN ALGORİTMALAR	38
5.1. Gaussian Naive Bayes (GNB)	38
5.1.1. Naive Bayesin Yapısı.....	39
5.2. K En Yakın Komşu (KYK)	40
5.3. Karar Ağaçları (KA)	43
5.3.1. KA Avantaj ve Dezavantajları	45
5.4. Destek Vektör Makinaları (DVM).....	47
5.4.1. Doğrusal destek vektör regresyon.....	50
5.4.2. Doğrusal olmayan destek vektör regresyon	51
5.5. Lojistik Regresyon.....	51
5.6. Yapay Sinir Ağları (YSA)	53
5.6.1. Yapay Sinir Ağları Yapısı.....	55
	vii

5.7. Rasgele Orman (RO)	57
5.8. Özellik Seçimi Yöntemleri	60
5.8.1. Fisher Skor	61
5.8.2. Relief	61
5.8.3. Ki-Kare Testi	61
5.8.4. Sınıflandırıcı ile Özellik Değerlendirici (Classifier Attribute Eval):	61
5.8.5. Korelasyon ile Özellik Değerlendirici (Correlation Attribute Eval):	62
5.8.6. Filtreleme ile Özellik Değerlendirici (Filtered Attribute Eval):	62
5.8.7. Kazanım Oranı ile Özellik Değerlendirici (Gain-Ratio Attribute Eval):	62
5.8.8. Bilgi Kazanım ile Özellik Değerlendirici (Info-Gain Attribute Eval):	62
5.8.9. OneR Özellik Değerlendirici (OneR Attribute Eval):	62
5.8.10. Önem ile Özellik Değerlendirici (Significance Attribute Eval):	63
5.8.11. SVM ile Özellik Değerlendirici (SVM Attribute Eval):	63
5.8.12. Simetrik Belirsizlik ile Özellik Değerlendirici (Symmetrical Uncert Attribute Eval):	63
6. PERFORMANS ÖLÇÜTLERİ	64
7. DENEYSEL SONUÇLAR	66
8. SONUÇ VE ÖNERİLER	76
KAYNAKLAR	80
ÖZGEÇMİŞ	88
ETİK KURUL	89

TABLolar DİZİNİ

Tablo 1. Epileptik nöbetlerin sınıflandırılması.....	12
Tablo 2. Epileptik nöbeti taklit eden sık durumlar	14
Tablo 3. Epilepsi sendromlarının başlangıç ve remisyon yaşları, prognoz ve tedavi seçenekleri	18
Tablo 4. Çocuklarda antiepileptik ilaçların nöbet tipine göre kullanımları ve dozları...	22
Tablo 5. Çalışmada kullanılan değişkenler.....	66
Tablo 6. Minkowski değerleri	68
Tablo 7. Euclidean değerleri.....	68
Tablo 8. Destek vektör makineleri değerleri	68
Tablo 9. Karar ağacı sınıflandırıcısı değerleri	68
Tablo 10. Rastgele orman gini parametresi değerleri.....	69
Tablo 11. Rastgele orman entropy parametresi değerleri.....	69
Tablo 12. Lojistik regresyon değerleri	70
Tablo 13. Doğruluk, özgüllük, hassasiyet ve f-skoru oranları.....	71
Tablo 14. Özellik seçimi için oluşturulan veri setine uygulanan algoritmaların doğruluk, özgüllük, hassasiyet ve f-skoru oranları	74

ŞEKİLLER DİZİNİ

Şekil 1. Epilepsilerin etiyolojik çerçevesi ile birlikte sınıflandırması.....	10
Şekil 2. Makine öğrenmesinin yapay zekâ-derin öğrenme ile ilişkisi ve makine öğrenmesinde yaygın olarak kullanılan öğrenme türleri	27
Şekil 3. Makine öğrenmesinin türleri ve alt grupları.....	28
Şekil 4. Kişiselleştirilmiş tedaviyi teşvik etmek için omik verilerini ve sağlık hizmetleri analitiğini entegre etmeye yönelik bir çerçeve	35
Şekil 5. Makine öğrenmesi algoritmaları ile çalışılan sağlık alanları.....	36
Şekil 6. Sağlıkta makine öğrenmesi çalışma alanları	37
Şekil 7. Basit bayes ağının birleşik olasılık dağılımı	38
Şekil 8. K-en yakın sınıflandırma örneğine ait grafik	42
Şekil 9. Karar ağacı Düğümleri	44
Şekil 10. Destek vektör makineleri veri ayrımı	48
Şekil 11. Doğrusal olarak ayrılabilen verilerin sınıflandırılması	49
Şekil 12. Doğrusal olarak ayrılamayan verilerin sınıflandırılması.....	50
Şekil 13. S şeklindeki cevap fonksiyonu grafiği	53
Şekil 14. Nöronlardan oluşan, ilişkili nöron bağlantıları içeren ve matematiksel bir model olan yapay sinir ağları modeli	54
Şekil 15. Yapay sinir ağları katmanları	55
Şekil 16. Rastgele orman sınıflandırma örneğine ait grafik	59
Şekil 17. Doğruluk oranları kıyaslanması	70
Şekil 18. Doğruluk oranları kıyaslanması	72
Şekil 19. Tüm değerlerin kıyaslamalı gösterimi.....	73
Şekil 20. Özellik seçimindeki verilere uygulanan algoritmaların kıyaslanması	75
Şekil 21. Özellik seçimindeki verilere uygulanan algoritmaların kıyaslanması	77
Şekil 22. Orijinal veri kümesi üzerinde algoritmaların kıyaslanması	77

KISALTMALAR DİZİNİ

Kısaltma	Açıklama
DVM	Destek Vektör Makineleri
EEG	Elektroensefalografi
KA	Karar Ağaçları
RO	Rastgele Orman
UESD	Uluslararası Epilepsi ile Savaş Derneği
YSA	Yapay Sinir Ağları



1. GİRİŞ

Epilepsi, herhangi bir ani merkezi sinir sistemi hasarının neden olmadığı ve nöbetlerin tekrarlamalarının nörobiyolojik, bilişsel, psikolojik ve sosyal sonuçları ile nöbet oluşturmaya kalıcı bir yatkınlık ile karakterize edilen kronik bir beyin hastalığıdır (Beghi, 2020). Epilepsinin genel prognozu, nöbet serbestliği ile ölçüldüğünde hastaların çoğunda olumludur. Çalışmalar tutarlı bir şekilde, vakaların yaklaşık yarısının uzun süreli nöbet remisyonu sağlama eğiliminde olduğunu göstermiştir. Bununla birlikte, epilepsinin uzun vadeli prognozu hakkında daha yeni raporlar, erken ve geç remisyon, tekrarlayan-düzelen bir seyir ve hatta kötüleşen bir seyir (remisyonun ardından nüks ve aralıksız nöbetler ile karakterize edilen) dahil olmak üzere farklı prognostik paternler tanımlamıştır (Beghi, 2020). Hastaların çoğunluğu mevcut antiepileptik ilaçlara (AEİ) olumlu yanıt verirken, dirençli epilepsi tipini geliştiren diğerleri ilaç tedavisine karşı direnç göstermekte ve ameliyat gibi ileri tedavilere ihtiyaç duymaktadır. Bu tür hastaların tanımlanması kolay değildir ve AEİ'lar uzun deneme yanılma süreçleri gerektirmektedir (Delen, Davazdahemami, Eryarsoy, Tomak ve Valluru, 2020).

Yapay zekâ, yeni bakım sağlama stratejileri, bilinçli karar verme ve hasta katılımını kolaylaştırma yoluyla sağlık hizmetlerini geliştirme potansiyeline sahip bir matematik mühendisliği alanıdır (Tack, 2019). Makine öğrenimi, karar vermeyi otomatikleştirmek ve hasta verilerine dayalı tahminler yapmak için kullanılabilen bir dar yapay zekâ biçimidir (Tack, 2019). Makine öğrenimi algoritmaları, tıp, e-posta filtreleme, konuşma tanıma ve bilgisayarla görme gibi ihtiyaç duyulan görevleri yerine getirmek için geleneksel algoritmalar geliştirmenin zor veya olanaksız olduğu çok çeşitli uygulamalarda kullanılmaktadır (Hu, Niu, Carrasco, Lennox ve Arvin, 2020). Veri madenciliği, makine öğreniminden farklıdır (Terzi, Küçüksille, Ergin ve İlker, 2011). Makine öğrenimi yaklaşımları üç grupta incelenir: Denetimli öğrenme: Bilgisayara, bir "öğretmen" tarafından verilen örnek girdiler ve istenen çıktılar sunulur ve amaç, girdileri çıktılara eşleyen genel bir kuralı öğrenmektir. Denetimsiz öğrenme: Öğrenme algoritmasına hiçbir etiket verilmez, girdisinde yapıyı bulmak için kendi başına bırakılır. Takviyeli öğrenme: Bir bilgisayar programı, belirli bir amacı (araç sürmek veya bir rakibe karşı oyun oynamak gibi) gerçekleştirmesi gereken dinamik bir ortamla etkileşime girer

(Bishop, 2006). Makine öğrenimi, açık talimatlar yerine verileri yorumlayarak performansı artırabilen algoritmalar geliřtirmek için istatistiksel ve bilgisayar bilimi ilkelerinden yararlanır. Görüntü tanıma, dil işleme ve veri madenciliğinde yaygın kullanımın yanı sıra, makine öğrenimi teknikleri, otomatik görüntüleme analizinden hastalık tahminine kadar tıbbi uygulamalarda artan ilgi görmektedir (Abbasi ve Goldenholz, 2019).

Bu çalışmada, Ocak 2011-2020 yılları arasında Necmettin Erbakan Üniversitesi Meram Tıp Fakóltesi Çocuk Nörolojisi Polikliniğine başvuran 1 ay-18 yaş arasındaki 500 epilepsili hastanın (220 kız ve 280 erkek) dosya kayıtları retrospektif olarak taranmıştır. Hastaların yaşı, semptomları, altta yatan hastalıklar, özgeçmiş ve soygeçmişleri, nörolojik muayene ve elektroensefalografi (EEG) bulguları, uygulanan tedaviler ve tedaviye cevapları ve beyin manyetik rezonans görüntüleme (MRG) sonuçları değerlendirilmiştir. Kaydedilen hastaların verileri makine öğrenmesi teknikleri (kural çıkarma, sınıflandırma ve kümeleme yöntemleri) kullanılarak analiz edilmiş ve sınıflandırılmıştır.

Amaç, hekimlerinin klinik mantığını taklit etmek için epilepsili çocukların klinik ve laboratuvar verilerinden dirençli epilepsi ile ilgili özellikleri çıkarmak için makine öğrenimini tekniklerinden bir sistem tasarlayarak, dirençli epilepsili hastaların tanısında yapay zekâ uygulamalarının etkinliğini belirlemektir.

2. LİTERATÜR ÖZETİ

Epilepsi, farklı etiyolojilere, fenotipik ve genotipik özelliklere sahip heterojen bir durumdur. Klinik ve araştırma yönleri, epidemiyolojiden moleküllere, klinik deneyler ve sonuçlar, gen ve ilaç keşfi, görüntüleme, EEG, patoloji, epilepsi cerrahisi, dijital teknolojiler ve daha pek çok şey arasında değişmektedir. Epilepsi verileri terabaytlar ve petabaytlar cinsinden toplanarak mevcut yeteneklerin sınırlarını zorlar. Diğer hastalıklara öncülük eden modern bilgi işlem gücü ve makine ve derin öğrenmedeki ilerlemeler, epilepsi için de heyecan verici olanaklar sunar. Bununla birlikte, bu tür verileri elde etmek, standartlaştırmak, iyileştirmek ve kullanılabilir hale getirmek için dikkatlice tasarlanmış yaklaşımlar olmadan, bir başarısızlık riski vardır. Bu nedenle, yakın paydaş girdileri ile ilgili ontolojilerin dikkatli bir şekilde oluşturulması, epilepsi verilerinin genellenmesi gibi daha iddialı büyük veri girişimleri için gerekli olan iskeleyi sağlar (Lhatoo, Bernasconi, Blumcke, Braun, Buchhalter, Denaxas, Galanopoulou, Josephson, Kobow, Lowenstein, Ryvlin, Schulze-Bonhage, Sahoo, Thom, Thurman, Worrell, Zhang ve Wiebe, 2020). Artan hesaplama yetenekleri, etkili makine öğrenimi algoritmalarının kullanılabilirliği ve daha büyük veri kümelerinin birikimi ile klinisyenler ve araştırmacılar, bu tekniklere aşinalıktan ve epilepsideki uygulamalarında hâlihazırda kaydedilen önemli ilerlemeden giderek daha fazla yararlanacaklardır (Abbasi ve Goldenholz, 2019).

An, Malhotra, Dilley, Han-Burgess, Valdez, Robertson, Clark, Westover ve Sun (2018) 1,376,756 epilepsili hastanın verilerini analiz etmiştir. Amaç, ilk AEİ reçete ederken, tıbbi ve eczane talep bilgilerin hangi epilepsi hastalarının AEİ'lara dirençli hale geleceğini tahmin edip edemeyeceğini belirlemektir. Bunlardan 582,258'i tüm dâhil etme kriterlerini karşılamış; 49,916 hasta AEİ dirençlidir. En az dört farklı AEİ için başvuruda bulunan bir hasta AEİ dirençli olarak tanımlanmıştır. Grubu karakterize etmek için demografik özellikleri, komorbiditeleri, ilaçları, işlemler, epilepsi durumunu ve ödeyen durumunu yansıtan 1,270 aday öngörücü "özellikler" oluşturulmuştur. Eğitim veri setinde (528,640 hasta), tahmini özellikleri seçmek için ANOVA F-değeri testleri gerçekleştirilmiş ve lojistik regresyon, destek vektör makineleri (DVM) ve rastgele orman (RO) dâhil olmak üzere çeşitli tahmin algoritmaları eğitilmiştir. Karşılaştırma modeli olarak yalnızca yaş

ve cinsiyet içeren bir model kullanılmıştır. Sonuçta, kıyaslama modeli için uzatılmış bir test setinde (53,618 hasta), en iyi model 0,664 (0,658, 0,671) ile karşılaştırıldığında, (%95 güvenlik aralığı) 0,753 (0,747, 0,759) alıcı çalışma özelliği eğrisi altında elde edilmiştir. Dahası, ilaç direnci için tahmin edilen olasılıklar, gözlemlenen frekanslarla yakından eşleşmiştir. İki AEİ yetersizliği yönünden bekleme kıyaslandığında, model ortalama 2,25 yıl önce ilaç direncini tahmin etmiştir (An vd., 2018).

Delen, Davazdahemami, Eryarsoy, Tomak ve Valluru (2020) doktorları tarafından yapılan epilepsi ile ilgili ilk tanı ile birlikte komorbiditelerini ve demografik bilgilerini kullanarak ilaca dirençli hastaları tanımlamak için veri dengelemeye yönelik yeni bir yaklaşımla donatılmış birkaç tahmine dayalı analitik makine öğrenme yöntemini kullanmışlardır. Amaç, elektronik tıbbi kayıtlardan elde edilen klinik ve demografik veriler kullanılarak risk altındaki hastaları tahmin etmeyi amaçlamaktır. Epilepsili 34,000'den fazla hastanın geçmiş tıbbi kayıtları kullanarak, pek çok tıbbi deneme yanılma kursundan geçmeden önce, refrakter epilepsi tipi (yani hastaların ilaç tedavisine dirençli) olan hastaları belirlemek için bir makine öğrenimi modeli elde edilmiştir. İki toplu sınıflandırma algoritması olan RO ve eğim artırma, genel olarak basit karar ağacı (KA) modelinden daha doğru sonuçlar sağlamış ve eğim artırma, üç eğitim veri setinin hepsine ilişkin en iyi sonuçları vermiştir. Önerilen model, refrakter epilepsi geliştiren hastaların yüzde 75'ini komorbiditelerini, demografik bilgilerini ve doktorları tarafından konulan ilk tanıları kullanarak doğru bir şekilde sınıflandırabilmiştir. Elde edilen umut verici sonuçlar, tıbbi kararları kolaylaştırmada makine öğrenme tekniklerinin potansiyel kullanımını vurgulamıştır ve tıp uzmanları için bir klinik karar destek sistemi geliştirmek için önerilen yaklaşımı genişletme olasılığını önermiştir (Delen vd., 2020).

Bennett, Kanber, Hoskote, Cardoso, Ourselin, Duncan and Winston (2019) tek taraflı temporal lob epilepsili 104 hastada çok modellenli MRG kullanarak nöbet başlangıcının bilinen tarafını tahmin etmek için çağdaş makine öğrenme yöntemlerini kullanmışlardır (82 MRG pozitif, 22 MRG negatif). Çalışmada, hem MRG pozitif hem de MRG negatif vakalarda nöbet tarafını öngören görüntü özelliklerini vurgulamak için "Önem Haritaları" adlı bir görselleştirme yaklaşımı geliştirilmiştir. Nöbet tarafı MRG pozitifte 0,981 (%95 güven aralığı = 0,974-0,989) ve MRG negatifte 0,842 (%95 güven aralığı = 0,736-0,949)

alıcı çalışma karakteristik eğrisi altındaki bir alanla tahmin edilebilmiştir. Hipokampustan kaynaklanan bilinen görüntü özellikleri, MRG pozitif olgularda nöbet lateralitesinin önde gelen belirleyicileri iken, MRG negatif olgularda yaygın temporal lob anormallikleri ortaya çıkmıştır. Sonuçta, görsel okumada fark edilmeyen gizli anormallikler, MRG negatif temporal lob epilepside, sadece hipokampus yerine tüm temporal lobu kapsayan bir uzaysal model ile tespit edilmiştir. Bu, MRG negatif temporal lob epilepsinin hafif fakat yaygın temporal lob anormallikleriyle ilişkili olabileceğini düşündürmektedir. Bu anormallikler, aşikâr lezyon yoksa yakından inceleme ve devralma sonrası işlemleri belirtirler (Bennett vd., 2019).

Ülkemizde EEG sinyalleri, epilepsi ve makine öğrenme yöntemleriyle ilgili çeşitli çalışmalar yapılmıştır. Yalçın (2012) sezgisel algoritma öğrenmeli yapay sinir ağları (YSA) ile epilepsi hastalığının teşhisini araştırmıştır. Çalışmada, yedi tane parçacık sürü optimizasyon tabanlı yapay sinir ağı (PSO) modeli içerisinde sınıflandırma için en uygun olanlar PSO3 ve PSO7'dir (Yalçın, 2012).

Saleem (2017) DVM ve dalgacık dönüşümü kullanarak EEG'de epilepsi nöbeti tanımayı araştırmıştır. Çalışmada, beyindeki nöronal aktivite üzerinde, gürültüyü filtrelemek için kullanılan maksimum çakışma ayırık dalgacık analizi ve sinyallerin öznitelik çıkarma ve DVM'nin EEG sinyallerinin sınıflandırılması için kullandığı en yaygın genel makine öğrenme tekniklerine dayanan iki brain-computer interface sistemi sunulmuştur. Diğer sistem ise EEG sinyallerinden epilepsi nöbetlerini belirleyebilen maksimal overlap ayırık dalgacık analizi, birincil bileşen analizi ve Destek Vektör Makinesi üzerine kuruludur. Bu çalışma normal kişi ile anormal "epileptik" hastalarını EEG kullanarak araştırmakta ve verilerin boyutsallığının azaltılması olarak temel bileşen analizinin önemi açıklanmaktadır (Saleem, 2017).

Yıldırım (2018) farklı zaman ölçekli EEG işaretlerinden epilepsinin tespitini araştırmıştır. Çalışmada, k-en yakın komşu (KYN), DVM, aşırı öğrenme makineleri (AÖM) sınıflandırıcısı açısından değerlendirme yapıldığında KYN ve AÖM sınıflandırıcılarının başarımlarının birbirine yakın olduğu ve tüm özellik çıkarma yöntemlerinde genelde en iyi sonuçları DVM sınıflandırıcısının verdiği görülmüştür (Yıldırım, 2018).

İmak (2019) epileptik EEG sinyallerinin makine öğrenmesi yöntemleri ile sınıflandırmasını yapmıştır. Çalışmada, farklı yöntemlerle elde edilen öznelilikler AÖM, DVM, KYK ve KA ile sınıflandırılmıştır. Kullanılan öznelilik çıkarma yöntemleri ve AÖM ile tanı kararının çok kısa sürede, iyi başarımla vermesi sağlanmıştır. Genel olarak en iyi sonuçlar AÖM ile kaydedilmiştir (İmak, 2019).

Kaya (2019) EEG beyin aktivitesi görüntüleme ve makine öğrenimi veri analizi yöntemlerini kullanarak bireylerin zihinsel dikkat durumlarını (odaklanmış, dalgın ve uykulu) izlemeye yönelik pasif bir beyin-bilgisayar ara yüzü geliştirmiştir. Çalışmada 1-5 Hz ve 10-15 Hz frekans bantlarında meydana gelen frontal ve parietal loblarda EEG aktivitesindeki değişikliklerin, bireylerin dikkat durumundaki değişikliklerle ilişkili olduğu saptanmıştır. DVM tabanlı zihinsel durum algılayıcı sürekli performans görevi versiyonunu kullanarak deneysel ortamda bu tür değişiklikleri, bireylerin dikkat durumunu %96.70 (en iyi) ve %91.72 (ortalama) doğrulukla belirlemiştir (Kaya, 2019).

Şentürk (2019) EEG sinyallerini epilepsinin nöbet öncesi tespitine yönelik analiz ve modellemesini araştırmıştır. Çalışmanın ilk aşamasında varyasyonel kip ayrışım yöntemi kullanılarak EEG sinyalleri analiz edildiğinde epilepsi ortaya çıkmadan önce bazı pencerelerde epilepsi durumuna benzer durum görüldüğü ve epilepsi öncesi epileptik aktivitenin tahmin edilebileceği ön görülmüştür. Çalışmanın ikinci adımında EEG sinyallerinden epileptik aktiviteye ait dalga formu matematiksel olarak modellenmiştir. Yapay arı kolonisi algoritması, parçacık sürü optimizasyon, genetik algoritma ve diferansiyel algoritması algoritmaları kullanılarak modele ait parametreler belirlenmiştir. Yapılan genel analizler ve elde edilen sonuçlar ile modelleme aşamasında parçacık sürü optimizasyon ve genetik algoritma algoritmalarının performansının daha iyi olduğu tespit edilmiştir. Sonuçta EEG sinyallerinden epilepsi öncesi epilepsi olabileceği durumu hakkında bilgi veren bir analiz yöntemi sunulmuştur. Ayrıca bireylerin epileptik dalga formunu belirten matematiksel model ve bu modeli ortaya koyan algoritmalar belirlenmiştir (Şentürk, 2019).

Zhunis (2019) EEG işaretlerinde epilepsi nöbet tahmini ve tespitini araştırmıştır. DVM, KA, KYK sınıflandırıcılarının başarımları analize edilmiştir. BONN veri tabanı kullanılarak EEG verilerinde sağlıklı, nöbet öncesi ve nöbet anı tespit edilmiştir. Yapılan

deneylerde epilepsi tespiti ve tahmininde en iyi başarıyı ampirik kip ayrışımı uygulanılmasının ve sınıflandırıcı olarak da DVM, KA durumunda elde edildiği görülmüştür (Zhunis, 2019).

Türk (2019) EEG işaretlerinden epilepsi türlerinin sınıflandırılmasında skalogram tabanlı derin öğrenme yaklaşımını incelemiştir. Çalışmada, nöbet öncesi EEG kayıtlarında %90.44, normal EEG kayıtlarında %91.29, parsiyel epileptik nöbetler EEG kayıtlarında %89.63 ve jeneralize epileptik nöbetler (nöbet anı) EEG kayıtlarında ise %96.50 oranında doğruluk başarıyı elde edilmiştir (Türk, 2019).

Bacaksız (2020) uyarlamalı yankılaşım teorisi ile epilepsi verilerini öbeklemiştir ve bir epilepsi modeli oluşturmuştur. Çalışmada temel ART yapısı incelenmiş ve bu yapıya ilişkin diferansiyel denklemlerden faydalanılarak küçük boyutlu bir problemin çözümünde neler olduğu durum uzayındaki davranışlar elde edilerek açıklanmıştır. Ayrıca epilepsiye ilişkin bir model olan Jansen modeli incelenerek bu model ile elde edilen neticeler sunulmuştur (Bacaksız, 2020).

Dal (2020) epileptik/epileptik olmayan nöbetleri birbirinden ayıracak elektromiyografi, elektrokardiyografi, ivmeölçer gibi çeşitli sensörler yardımıyla epileptik aktiviteyi izleyerek ölçülen sinyallerin analizini yapmıştır. Analizler sırasında üç farklı EEG kanalı, elektrokardiyografi ve elektromiyografiden oluşan veri grupları ile kriz tespiti, kriz türü tespiti ve tüm verilerin ayrıştırılması biçimde gruplandırılmıştır. Bu veri gruplarının hepsinde özellik seçmeden %100 sınıflandırma doğruluğu elde edilmiştir. Elektromiyografi ve elektrokardiyografi verilerinin kriz ve kriz türü tespitinde yeterli oldukları gözlenmiştir (Dal, 2020).

Yeşilkaya (2020) EEG sinyalleri ve makine öğrenme yöntemlerini kullanarak duygusal durum kestirimini araştırmıştır. Çalışmada, çok değişkenli görgül kip ayrışım gibi ileri sinyal analiz yöntemi ve derin öğrenme yaklaşımı kullanılarak duygu durum kestiriminde başarılı sonuçlar elde edildiğini ve gelecek insan-makine etkileşim sistemlerinde kullanılabileceğini bildirmiştir (Yeşilkaya, 2020).

Bu tez çalışmasında, yukarıdaki çalışmalardan farklı olarak epilepsili çocukların klinik ve laboratuvar verilerinden dirençli epilepsi ile ilgili özellikleri çıkarmak için makine öğrenimini tekniklerinden bir sistem tasarlayarak, dirençli epilepsili hastaların tanısında yapay zekâ uygulamalarının etkinliği araştırılmıştır.



3. EPİLEPSİ

Aralarında 24 saatten uzun süre geçmiş, iki veya fazla tetiklenmemiş nöbet, epilepsi olarak tanımlanır. Bu tanımlama, Uluslararası Epilepsi ile Savaş Derneği (UESD) tarafından 2014 yılında yayınlanan rapor ile genişletilmiş; 10 yıl içinde tekrarlama olasılığı yüksek olan (>%60) tek tetiklenmemiş (ya da refleks) nöbet ve özgül bir epilepsi sendromu tanısı alan durumlar da epilepsi tanımına katılmıştır. Aynı raporda, iyileşme kriterleri de tartışılmıştır, iyileşme terimi yerine çözümlenme teriminin daha uygun olacağı bildirilmiştir. Özgül bir yaş döneminde görülen belirli epilepsi sendromları için o yaş geçtikten sonra; diğer epilepsiler için ise en az 5 yıl AEİ kullanmaksızın 10 yıl nöbetsiz olma durumunda, epilepsinin çözümlendiği kabul edilmektedir (Fisher, Acevedo, Arzimanoglou, Bogacz, Cross, Elger, Engel, Forsgren, French, Glynn, Hesdorffer, Lee, Mathern, Moshé, Perucca, Scheffer, Tomson, Watanabe ve Wiebe, 2014).

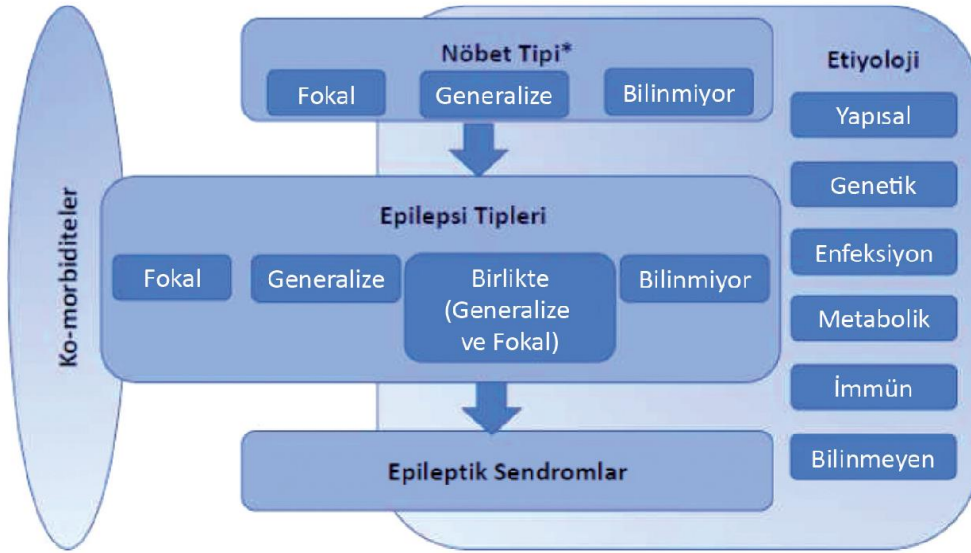
İlaça dirençli epilepsi, tolere edilen, uygun şekilde seçilen ve kullanılan iki AEİ (monoterapi olarak veya kombinasyon halinde) sürekli nöbetsizliği elde etmede başarısız olması olarak tanımlanmaktadır (Kwan, Arzimanoglou, Berg, Brodie, Allen Hauser, Mathern, Moshé, Perucca, Wiebe ve French, 2010).

3.1. Epidemiyoloji

Çocuklarda yapılan çeşitli çalışmalarında epilepsi prevalansı 4-13.5/1000 arasında saptanmıştır (Aydın, Ergor, Ergor ve Dirik, 2002; Aygün, 2010; Canpolat, Kumandas, Poyrazoglu, Gumus, Elmali, ve Per, 2014; Hüseyinoğlu, Ozben, Arhan Palanci, ve Gunes, 2012; Serdaroğlu, Ozkan, Aydın, Gücüyener, Tezcan, ve Aycan, 2004). Bu çalışmalardaki farklı sonuçlar genetik sebepler ve coğrafik bölge farklılıklarıyla ilişkili olabilir.

3.2. Etiyoloji

Şekil 1'de UESD tarafından 2017 yılında yapılan etyolojik çerçevesi ile birlikte epilepsilerin sınıflandırması görülmektedir (Scheffer vd., 2017).



Şekil 1. Epilepsilerin etiyolojik çerçevesi ile birlikte sınıflandırması
(Kaynak: Scheffer vd., 2017)

Yapısal epilepsi, önceden semptomatik epilepsi olarak adlandırılmaktadır. Genetik olabilen veya olmayan altta yatan bir yapısal beyin bozukluğunun neden olduğu epilepsi sendromu anlamına gelmektedir. Bu grupta, eski inme veya hipoksik-iskemik hasar gibi etiyolojiler ve tuberoskleroza (aynı zamanda genetik bir etiyolojiye sahiptir) sekonder epilepsi örnek gösterilebilir (Mikati ve Tchapyjnikov, 2020a).

Genetik epilepsi, önceden idiyopatik epilepsi olarak da isimlendirilmiştir. Epilepsi sendromunun, epilepsi dışında yapısal veya metabolik beyin bozukluğuna neden olmayan bilinen veya tahmin edilen genetik defekt(ler)in direkt sonucu olduğuna işaret etmektedir. Örneğin çocukluk çağı absans epilepsisi ve Dravet sendromu (Mikati ve Tchapyjnikov, 2020a). Epilepsi ile ilişkili hastalıkların birçoğunda iyon kanalı bozuklukları saptanmıştır (Wei, Yan, Su, He, Lin, Wang, Shi, Yi ve Liao, 2017). Genetik alanındaki çalışmalar ve gelişmelerle birlikte yeni genler keşfedilmektedir. Epilepsi ile ilişkili 1000'in üzerinde gen tanımlanmıştır (Wang, Lin, Liu, Xu, Shi, Yi, He ve Liao, 2017).

Enfeksiyöz epilepsi, bakteriyel menenjit veya herpes simpleks virüs ensefaliti gibi akut enfeksiyonlardan ziyade tüberküloz ve insan immün yetmezlik virüsü gibi kronik enfeksiyöz durumlara sekonder olarak gelişen epilepsileri tanımlar (Mikati ve

Tchapyjnikov, 2020a). Metabolik hastalıkların yaygın ensefalopati yapan formları çoğunlukla epileptik nöbetlere ve epilepsiye yol açmaktadır (Ünal, Haliloğlu ve Coşkun, 2012).

İmmün aracılı epilepsi, immün aracılı santral sinir sistemi inflamasyonuna sekonder ortaya çıkan epilepsileri tanımlamaktadır. Bu grup bozukluklar özel dikkat gerektirir, çünkü steroid ve intravenöz immünglobulin gibi immünoterapiler birinci basamak tedaviler olabilmektedir. Anti-N-metil-D-aspartik asit reseptör ensefaliti ve anti-lösin-zengin glioma-inaktif limbik ensefaliti gibi otoimmün ensefalitler, immün aracılı epilepsi örnekleridir (Mikati ve Tchapyjnikov, 2020a).

Eskiden kullanılan kriptojenik epilepsi, epilepsiye neden olan ve nörolojik işlevi etkileyen, altta yatan bir beyin bozukluğunun olduğu tahmin edilen, ancak altta yatan bozukluğun bilinmediği epilepsi sendromu anlamına gelmektedir. Bu bozukluk artık bilinmeyen epilepsi olarak adlandırılmakta olup, altta yatan epilepsi nedeninin hala belirlenmediğini işaret etmektedir (Mikati ve Tchapyjnikov, 2020a).

3.3. Epilepsinin Sınıflaması

Epileptik nöbetlerin klinik ve elektrografik sınıflaması tüm dünyada genel kabul görmüş ve ortak bir dil oluşmasını sağlamıştır. Şekil 1'de epilepsilerin sınıflaması görülmektedir. Epilepsiler, nöbetin başlangıcına göre dört gruba ayrılır: fokal başlangıçlı, jeneralize başlangıçlı, bilinmeyen başlangıçlı ve sınıflandırılmayan. Fokal nöbetlerde (önceden parsiyel olarak bilinen) ilk klinik ve EEG değişiklikleri beyin bir yarım küresindeki sınırlı alandaki nöronların aktivasyonunu gösterir. Fokal nöbetler, motor ve motor olmayan ayrıca bilincin bozulduğu (kompleks parsiyel nöbet) ve bilincin bozulmadığı (basit parsiyel nöbet) biçiminde gruplandırılır (Mikati ve Tchapyjnikov, 2020a). Jeneralize nöbetlerde, ilk klinik ve EEG değişiklikleri, her iki hemisferin tümünün senkronize etkilendiğini gösterir. Bilinmeyen başlangıçlı nöbetler; nöbetin fokal veya jeneralize olduğunu belirlemek için yeterli klinik bilginin olmadığı nöbetlerdir. Sınıflandırılmayan nöbet; klinik özellikleri alışılmışın dışında olan ve yeterli incelemeye rağmen durumu tespit edilemeyen nöbetlerdir. Tablo 1'de epileptik nöbetlerin sınıflandırılması görülmektedir (Mikati ve Tchapyjnikov, 2020a).

Tablo 1. Epileptik nöbetlerin sınıflandırılması

NÖBET TİPLERİ
Fokal Başlangıçlı Nöbetler
Motor başlangıç
Tonik
Klonik
Atonik
Myoklonik
Hyperkinetik
Epileptik spazmlar
Otomatizmalar
Motor olmayan başlangıç
Duraklama
Duyusal
Kognitif
Emosyonel
Otonom
Farkındalık tanımlayıcısı
Farkındalık korunmuş
Farkındalık bozulmuş
Jeneralizasyon tanımlayıcısı
Fokal başlayan bilateral tonik-klonik (daha önce sekonder jeneralize nöbet olarak bilinen)
Jeneralize Başlangıçlı Nöbetler
Motor
Tonik-klonik
Tonik
Klonik
Atonik
Myoklonik
Myoklonik-atonik
Myoklonic-tonik-klonik
Epileptik spazmlar
Motor olmayan (Absans)
Tipik
Atipik
Myoklonik
Göz kapağı miyoklonisi
Bilinmeyen Başlangıçlı Nöbetler
Motor
Tonik-klonik
Epileptik spazmlar
Motor olmayan
Duraklama
Sınıflandırılmayan

2010 UESD nöbet sınıflamasında temel olarak, jeneralize, fokal ve bilinmeyen olarak sınıflanmıştır. Jeneralize nöbetler; tonik-klonik, absans, miyoklonik, klonik, tonik ve atonik olarak gruplandırılmıştır. Auralar ise otonomik, motor, diskognitif olmak üzere genelleştirilmiştir (Berg, Berkovic, Brodie, Buchhalter, Cross, van Emde Boas, Engel, French, Glauser, Mathern, Moshé, Nordli, Plouin ve Scheffer, 2010). Son olarak UESD 2017 yılında büyük ölçüde 1981 yılında rapor edilen mevcut sınıflamaya dayanan ancak önemli değişiklikler içeren bir sınıflama yapmıştır (Fisher, Cross, D'Souza, French, Haut, Higurashi, Hirsch, Jansen, Lagae, Moshé, Peltola, Roulet Perez, Scheffer, Schulze-Bonhage, Somerville, Sperling, Yacubian ve Zuberi, 2017). Daha önce jeneralize nöbet kategorisinde olan bazı nöbetler yeni nöbet sınıflamasında fokal nöbet tipleri olarak değiştirilmiştir. Bir hemisferde sınırlı nöronal ağlardan kaynaklanan nöbetler için “parsiyel” yerine “fokal” terimi UESD’in 2010 önerilerindeki gibi kullanılırken, şuur yerine farkındalık kelimesi kullanılmıştır (Fisher vd., 2017). Bunlara ek olarak; fokal nöbetler için bilinç değişiklikleri temel alınarak yapılan basit ve kompleks ayrımının yapılmaması ve bir kişide iki taraflı motor bulgular olsa bile; eğer fokal bulgu ve işaretler varsa fokal nöbet tanısının önce değerlendirilmesi gerektiği savunulmuştur. Jeneralize nöbetlerin spesifik bir odak olmadan iki taraflı kortikal ve subkortikal ağlar aracılığı ile çok hızlı bir şekilde yayılarak ortaya çıktığı, tüm kortekse yayılımına gerek olmadığı ileri sürülmüştür (Fisher vd., 2017, Scheffer, Berkovic, Capovilla, Connolly, French, Guilhoto, Hirsch, Jain, Mathern, Moshé, Nordli, Perucca, Tomson, Wiebe, Zhang ve Zuberi, 2017) (Şekil 1).

3.4. Tanı

Epilepsi tanısı klinik verilere ve bulgulara dayanarak konmaktadır (Kaya Özçora, Canpolat ve Kumandaş, 2020).

3.5. Ayırıcı Tanı

Yanlış epilepsi tanısının %5-40 kadar yüksek olduğu tahmin edilmektedir. Genellikle

epileptik olmayan paroksizmal bozuklukları epilepsiden ayırmak için gereken tek şey, kapsamlı bir klinik muayeneye ek olarak dikkatli ve ayrıntılı bir öyküdür, ancak bazen bir EEG veya daha ileri testler gerekli olabilir. Cep telefonlarında video kaydının hazır olması, paha biçilmez bilgiler sağlayabilir (Mikati ve Obeid, 2020). Tablo 2'de epileptik nöbeti taklit eden sık durumlar verilmiştir. Bir çocuğa epilepsi tanısı konmadan önce tablo 2'de görülen epileptik olmayan paroksizmal bozukluklarla ayırıcı tanısının yapılması gerekir (Fine ve Wirrell, 2020).

Tablo 2. **Epileptik nöbeti taklit eden sık durumlar**

Epileptik nöbeti taklit eden durumlar	Etkilenen yaş	Klinik ipuçları
Benign uyku miyoklonusu	Yenidoğan ve erken infantil	• ≥ 1 uzuvda ya da yüzde miyoklonus, $< 3-5$ sn kısa süreli ve değişken duraksamalarla oluşur. • Yalnızca uykuda ortaya çıkar ve uyandığında ortadan kalkar. • Sağlıklı bebeklerdir.
Jitteriness	Yenidoğan	• ≥ 1 uzuvları etkiler, genellikle olay jitteriness den köken alır. • Genellikle anatomik olmayan modelde yayılır (ör. Sol bacak sağ kola). • Bebek sarılmadığında, uyarıldığında, şaşırdığında veya ağladığında artar, ancak bebek sarıldığında veya etkilenen uzuv nazikçe tutulduğunda baskılar.
Bebeklik döneminin benign miyoklonusu	Bebeklik	• Farkındalık değişmeden her biri 5 saniyeden az süren ≥ 1 uzvun kısa sarsılması. • Hem uyanıklıkta hem de uykuda ortaya çıkar. • Sağlıklı bebek.
Ürperme atakları	Geç bebeklik	• Farkındalıkta değişiklik olmadan titreme gibi omuz sallanarak kısa süreli sertleşme. • Genellikle heyecan veya hayal kırıklığı ile tetiklenir. • Aksi takdirde normal bebek.
Nefes tutma nöbetleri - siyanotik veya soluk	Bebeklik, erken çocukluk	• Acı, ağlama ve korku ile tetiklenir. • Çocuk genellikle ağlar (soluk nefes tutarken ağlama olmayabilir), ekspirasyonun sonunda nefesini tutar ve sonra kısa süreli tonik hale gelir. • İlişkili renk değişikliği - siyanotik veya soluk.

Sandifer sendromu	Bebeklik, erken çocukluk	• Sırt kavisli, uzuvların distonik duruşu ve başın döndürülmesi / eğilmesi. • Beslenerek ve yatarak kıskırtılabilir. • Oturarak hafifletilebilir. • Genellikle nörolojik olarak anormal çocuklarda görülür. • Gastroözofageal reflü nedeniyle olur.
Stereotipiler	Bebeklik-çocukluk	• Basit (vücut sallama, baş vurma gibi) veya karmaşık (parmak hareketleri veya bilek ekzasyonu/ekstansiyonu gibi) davranışlar. • Bunlar dokunsal olarak kesilebilir ve bazen sözlü uyarma ile kesilebilir. • Sağlıklı bireylerde ortaya çıkabilir ancak daha çok otizmli veya zihinsel engelli kişilerde görülür.
Hiperekpleksi	Bebeklikten ergenliğe kadar ancak yaşla birlikte daha az şiddetli hale gelir	• Bebeklik: bebekler hipertondiktir ancak spastik değildir; aşırı irkilme, gürültü veya dokunma, uzuvların ekstansiyonu ve boyun retraksiyonu ile görülür; bu bazen apne ve siyanoz ile ilişkilendirilebilir. • Parmak ucuyla glabellaya ardışık vurulduğunda alışma olmaksızın sürekli irkilme görülür.
Kendini uyarıcı davranış	Erken çocukluk	• Ritmik kalça ekzasyonu ve bacak çaprazlama ile adduksiyon, sıklıkla uzak bir ifade ile birlikte. • Kesilebilir, ancak kesilirse çocuk huzursuz olabilir.
İyi huylu paroksizmal vertigo	Erken çocukluk	• Anksiyete başlangıcı, dengesizlik hissi; çocuk genellikle ebeveyni kavrır. • Nistagmus ile ilişkili olabilir.
Döngüsel kusma	Çocukluk	• Saatler sürebilen ve arasına serpiştirilebilen ve haftalar ila aylar süren semptomsuz dönemlerde tekrarlayan kusmanın paroksizmal olayları.
Hayal kurmak	Çocukluk	• Dik dik bakmak, okul ödevi gibi sessiz aktivitelerle meşgul olduğunda ortaya çıkma olasılığı daha yüksektir. • Dokunsal stimülasyon ile kesintiye uğrayabilir.
Parasomniler	Çocukluk ve nadiren ergenlik	• Gece terörü, uyurgezerlik ve kafa karışıklığı uyarılma, derin REM dışı uykudan kaynaklanan davranışlardır; genellikle uykuya daldıktan sonraki ilk birkaç saat içinde; tipik olarak 3-5 dakikadan uzun sürer ve aralıklı olarak ortaya çıkarlar. • Bunlar geceden ayırt edilmelidir. kısa olan frontal lob nöbetleri (tipik olarak <2 dk.), Çok sık (gece başına birden fazla) ve gece boyunca ortaya çıkıyor.
Öfke nöbetleri/öfke saldırıları	Çocukluktan ergenliğe	• Öfke nöbetleri öncelikle küçük çocuklarda görülür ve bir uyarıcıya yanıt olarak nispeten kısa süreli davranışsal kontrolsüzlükleri içerir; bilinç bozulmamıştır.

Tikler	Çocukluk ve ergenlik	• Öfke reaksiyonları ağırlıklı olarak yaşlılarda görülür çocuklar ve gençler ve küçük uyaranlarla tetiklenirken karakteristik olarak orantısızdır; yarım saat veya daha uzun sürebilen bu dönemlerde hastalar genellikle agresiftir. • Genellikle günde birkaç kez meydana gelen istemsiz, ani, hızlı, tekrarlayan, ritmik olmayan, basit veya karmaşık hareketler veya seslendirmeler. • Sadece birkaç saniye için de olsa bunlar kesintiye uğrayabilir ve bastırılabilir. • Uyku sırasında tikler azalır.
Uykuda periyodik bacak hareketleri	Çocukluk ve ergenlik	• Ayak parmakları, ayak bilekleri, dizler ve kalçalarda tekrarlayan basmakalıp fleksiyon. • Uyanma ile çözülür.
Vazovagal senkop	Çocukluk ve ergenlik	Tipik olarak uzun süre ayakta durmakla tetiklenir, dehidrasyon, duruş değişikliği, sıcak ortam veya duygusal üzüntü (örneği kan alımı). • Daha önce baş dönmesi, bulanık görme, kulaklarda çınlama, solukluk, terleme, karın rahatsızlığı. • Kısa süreli olabilen ton kaybı miyoklonik jerkler veya tonik duruş. • Farkındalığa hızlı dönüş, ancak daha sonra kısa bir süre sersemlemiş kalabilir.
Postüral ortostatik taşikardi sendromu	Gençlik	• Epizodik sersemlik dönemleri, göğüs ağrısı, bulanık görme, karın ağrısı. • Ayakta durmakla gelir ve oturarak ya da uzanmakla gider.
Panik ataklar	Gençlik	• Nefes darlığı, boğulma hissi, çarpıntı, göğüs ağrısı, parestezi baş dönmesi, terleme, titreme ve bayılma hissi ile birlikte ani yaklaşan kıyamet hissi ile sadece dakikalar süren kısa bölümler • Hasta çok korkmuş ama farkındadır • Postiktal uykululuk / kafa karışıklığı yok.
Narkolepsi / katapleksi	Ergenlik, ara sıra çocukluk	• Gündüz aşırı uyku hali, katapleksi (güçlü duyguya yanıt olarak ton), hipnagogik halüsinasyonlar ve uyku felci.
Hemiplejik migren	Ergenlik bazen geç çocukluk	• Odak zayıflığı, konuşma bozukluğu, görsel semptomlar ve parestezi daha önce başlar. • Tipik migren benzeri baş ağrısı. • Genellikle aile öyküsü olumludur.

Psikojenik epileptik olmayan nöbetler	Ergenlik, ara sıra çocukluk	• 2 ana semiyoloji: 1) motor fenomeni olmayan tepkisiz dönemler veya 2) tuhaf, düzensiz sarsıntı ve itme ile motor fenomeni. • Genellikle uzamış > 15-60 dk. • Genellikle minimal postictal. • Başından itibaren sık ve tepkisiz.
Kardiyak senkop - uzun QT	Herhangi bir yaş	• Solukluk, atoni ile birlikte ani bilinç kaybı, veya tonik duruş. • Genellikle korku, egzersiz, sürpriz ile tetiklenme. • Ailede senkop geçmişi olabilir.

(Kaynak: Fine ve Wirrell, 2020)

3.6. Tedavi

Epilepside, her hasta için birçok seçenek olduğundan, hangi ilacın kullanılacağını seçimi her zaman kişiselleştirilmiş ve randomize kontrollü çalışmalardan elde edilen karşılaştırmalı etkililik verilerine dayanan bir karar olmalıdır. Pek çok AEİ pediyatrik popülasyonda araştırılmamış olsa da, bu ilaçların çocuklarda endikasyon dışı kullanımı yaygındır. Bazı AEİ'ların nöbet şiddetlenmelerine neden olabildiği dikkate alınmalıdır. Örneğin, Dravet sendromunda lamotrijin ile miyoklonik nöbetler artabilir ve absans nöbetleri karbamazepin (CBZ) ve tiagabin (TGB) ile alevlenebilir (Mikati ve Tchapyjnikov, 2020b). Tablo 3'de epilepsi sendromlarının başlangıç ve remisyon yaşları, prognoz ve tedavi seçenekleri görülmektedir (Mikati ve Tchapyjnikov, 2020b).

Tablo 3. Epilepsi sendromlarının başlangıç ve remisyon yaşları, prognoz ve tedavi seçenekleri

Spesifik sendromlar	Başlangıç yaşı	Remisyon yaşı	Prognoz	Monoterapi veya ek tedavi*	Muhtemel Ek tedavi **	Cerrahi
Neonatal						
Benign neonatal nöbetler	Neonatal	Neonatal	İyi	LEV, TPM, PB	-	Hayır
Erken miyoklonik ensefalopati ve Ohtahara sendromu	Neonatal, süt çocukluğu	Kötü, Ohtahara sendromu West sendromuna dönüşüyor.	Kötü	PB, steroid, VGB	BZD, ZNS, TPM, LEV, ketojenik diyet	Hayır
Benign ailesel neonatal nöbetler	Neonatal, erken süt çocukluğu	Neonatal, Erken süt çocukluğu	İyi	LEV, TPM, PB	-	Hayır
İnfant						
Benign infantile nöbetler (ailesel olmayan)	İnfant	İnfant	İyi	LEV, TPM, PB	-	Hayır
Benign infantile ailesel nöbetler	İnfant	İnfant	İyi	LEV, TPM, OXC, CBZ, PB	-	Hayır
Süt çocukluğunun migratuvar fokal nöbetlerle seyreden epilepsisi	İnfant	Remisyon olmaz	Kötü	LEV, PB, OXC, CBZ, PHT, TPM, QND	BZD, bromidler, LAC, VPA, ZNS	Hayır
West sendromu	İnfant	Değişken	Değişken	ACTH, steroidler, VGB	BZD, FBM, IVIG, TPM, ZNS, ketojenik diyet	Lezyonektomi ± kortikal rezeksiyon
Dravet sendromu (ağır infantil miyoklonik epilepsi)	İnfant	Remisyon olmaz	Şiddetli	CLB, stiripentol, VPA (sadece 2 yaşından sonra)	BZD, TPM, LEV, ZNS, Ketojenik diyet	Hayır

Benign infantile miyoklonik epilepsi Çocukluk	3 ay-3 yıl	3-5 yıl	Değişken	LEV, TPM, BZD	VPA, ZNS	Hayır
Sentrotemporal dikenli benign çocukluk çağı epilepsisi	3-13 yıl	16 yıl	İyi	OXC, CBZ, LEV, VPA	LAC, PER	Hayır
Erken ve geç başlangıçlı idiyopatik oksipital epilepsi	2-8 yıl; 6-17 yıl	12 yıl veya daha erken; 18 yıl	İyi	OXC, CBZ, LEV, VPA	LAC, PER	Hayır
Otozomal dominant noktürn frontal lob epilepsisi	Çocukluk çağı		Değişken	OXC, CBZ, LEV	CLB, PB, PHT, LAC, PER, GBP, TPM	Hayır
Ailesel lateral temporal lob epilepsisi	Çocukluk çağı-adölesan dönem		Değişken	OXC, CBZ, LEV	CLB, PB, PHT, GBP, TPM, VPA, LAC, PER	Hayır (bazı olgular hariç)
Febril nöbetli jeneralize epilepsi artı	Çocukluk çağı-adölesan dönem		Değişken	ESM, LTG, LEV, VPA (nöbet tipine göre)	CLB, TPM, PER	Hayır
Hipokampal skleroz ile ilişkili mezial temporal lob epilepsisi	Okul çağı veya daha erken	Uzun süreli	Değişken	OXC, CBZ, LEV	CLB, GBP, LAC, PB, PER, PHT, ZNS, TPM, VPA	Temporal rezeksiyon
Rasmussen sendromu	6-12 yıl	İlerleyici	Kötü	LEV, OXC, CBZ, plazmaferez, immunglobulinler	LAC, PB, PER, PHT, TPM	Fonksiyonel hemisferektomi
Hemikonvulziyon-hemipleji sendromu	1-5 yıl	Kronik	Şiddetli	OXC, CBZ, LEV	CLB, GBP, LAC, PB, PER, PHT, ZNS, TPM, VPA	Fonksiyonel hemisferektomi
Miyoklonik astatik nöbetli epilepsi	3-5 yıl	Değişken	Değişken	ESM, TPM, VPA, LEV, ZNS	BZD, Ketojenik diyet, LTG, PER, steroidler	Hayır

Çocukluk çağı absans epilepsisi	5-6 yıl	10-12 yıl	İyi	ESM, LTG, VPA	Acetazolamid, CZP, ketojenik diyet, ZNS	Hayır
Miyoklonik absans epilepsi	1-12 yıl	Değişken	İhtiyatlı	ESM, VPA, CZP	ZNS, LTG	Hayır
Lennox-Gastaut sendromu	3-10 yıl	Remisyon olmaz	Şiddetli	CLB, LTG, RFD, TPM, VPA	BZD, FBM, IVIG, PER, steroid, ZNS, ketojenik diyet	Kallozotomi
Landau-Kleffner sendromu	3-6 yıl	8-12 yıl	İhtiyatlı	Noktürnal DZP, steroid, VPA, LEV	CLB, ESM, IVIG, LTG, ketojenik diyet	Çoklu subpial transeksiyon, nadiren lezyonektomi
Yavaş uykuda sürekli diken dalga ile giden epilepsi	4-7 yıl	8-12 yıl	İhtiyatlı	Noktürnal DZP, steroidler, VPA, LEV	CLB, ESM, IVIG, LTG, Ketojenik diyet	Hayır
Diğer görsel sensitif epilepsiler	2-5 yıl	Belirsiz	Değişken	VPA	BZD, LEV, LTG, ZNS	Hayır
Febril nöbetler	3-5 yıl	3-6 yıl	İyi	BZD (sadece ateşli dönemlerde)	-	Hayır
Jüvenil						
Jüvenil absans epilepsi	10-12 yıl	Genellikle yaşam boyu	İyi	ESM, LTG, VPA	Çocukluk çağı absans epilepsisi ile aynı	Hayır
Jüvenil miyoklonik epilepsi	12-18 yıl	Genellikle yaşam boyu	İyi	LEV, TPM, VPA	BZD, LTG, PB, PER, PRM, ZNS	Hayır
Jeneralize tonik-klonik nöbetli epilepsi	12-18 yıl	Genellikle yaşam boyu	İyi	LEV, LTG, TPM, VPA	BZD, CBZ, PER, ZNS	Hayır
İdiyopatik fotosensitif oksipital lob epilepsisi	10-12 yıl	Belirsiz	Değişken	VPA, LEV	BZD, LTG, ZNS	Hayır

Progresif miyoklonik epilepsiler (Unverricht-Lundborg, Lafora, seroid lipofusinos, vb.)	Geç süt çocukluk döneminden adölesan dönemi	İlerleyici	Kötü	TPM, VPA, ZNS, LEV	BZD, PB, CLB, PER, ketojenik diyet	Hayır
Değişken başlangıç yaşı	Değişken	Uzun süreli	Değişken	LEV, OXC, CBZ, TPM, VPA	PHT, PB, CLB, GBP, LAC, PER, ZNS	Lezyonektomi ± kortikal rezeksiyon
Bölge ve nedeni ile tanımlanan mezial temporal lob epilepsisi	Değişken	Uzun süreli	Değişken	LEV, OXC, CBZ, TPM, VPA	CLB, GBP, LAC, PB, PER, PHT, ZNS	Temporal rezeksiyon
Spesifik nedenlerle tanımlanan mezial temporal lob epilepsisi	Değişken	Uzun süreli	İhtiyatlı	OXC, CBZ, LEV, TPM, VPA	CLB, LEV, PB, PHT, ZNS, GBP	Bazı kortikal alanlarda lezyonektomi
Startle (irkilme) epilepsi	Değişken	Uzun süreli				Hayır
Refleks nöbetler	Değişken	Mevcut değil		LEV, VPA	LTG, ZNS	
İlaç veya diğer kimyasallarla indüklenen nöbetler	Değişken	Mevcut değil		Sebebin ortadan kaldırılması	-	Hayır
Erken posttravmatik nöbetler	Değişken	Mevcut değil		LEV, PHT	-	Hayır

Tablo 4'de çocuklarda antiepileptik ilaçların nöbet tipine göre kullanımları ve dozları görülmektedir (Rubenstein ve Levy, 2019).

Tablo 4. Çocuklarda antiepileptik ilaçların nöbet tipine göre kullanımları ve dozları

AEİ	Pediyatrik doz	Etki mekanizması	Nöbet tipi
Karbamazepin	10–35 mg/kg/gün	Voltaj kapılı sodyum kanallarını hedefleyen sodyum kanal blokeri	Fokal nöbetler
Klobazam	<2 yaş: 0.5–1 mg/kg/gün >2 yaş: 10–20 mg/gün	GABA-A reseptörlerinin aktivasyonu	Lennox-Gastaut sendromu ilişkili nöbetlerde antiepileptik ilaç eklenir
Etosüksimid	15–40 mg/kg/ gün	Düşük voltajlı kalsiyum kanallarını hedef alır.	Absans nöbetler
Eslikarbazepin	400–1,200 mg/gün ağırlığa göre	Voltaj kapılı sodyum kanallarını hedefleyen sodyum kanal blokeri	Fokal nöbetler
Felbamat	15–45 mg/kg/gün	Voltaj kapılı sodyum kanalları, yüksek voltajlı kalsiyum kanalları ve GABA-A reseptörleri üzerinde hareket ettiği düşünülüyor.	Fokal nöbetler, Lennox-Gastaut sendromu ilişkili atakları azaltmada
Lakozamid	11–29 kg: 6–12 mg/kg/gün 30–49 kg: 4–8 mg/kg/gün	Voltaj kapılı sodyum kanallarını hedefleyen sodyum kanal engelleyici	Fokal nöbetler
Lamotrijin	Monoterapi: 4.5–7.5 mg/kg/gün Ek olarak: 5–15 mg/kg/gün	Voltaj kapılı sodyum kanallarını hedefler ve muhtemelen yüksek voltajlı kalsiyum kanallarına da etki eder.	Fokal ve jeneralize nöbetler
Levetirasetam	20–60 mg/kg/gün	Sinaptik vezikül proteini 2A'yı hedef alır, kesin mekanizması bilinmemektedir.	Çoğu nöbet türü
Okskarbazepin	30–40 mg/kg/gün	Voltaj kapılı sodyum kanallarını hedefleyen sodyum kanal engelleyici	Fokal nöbetler

Fenitoin	4–8 mg/kg/gün Adolesanlar için 600 mg/güne kadar çıkılabilir.	Voltaj kapılı sodyum kanallarını hedefleyen sodyum kanal engelleyici	Fokal nöbetler ve jeneralize nöbetler
Fenobarbital	Bebekler: 5–6 mg/kg/gün Çocuklar <5: 6–8 mg/kg/gün	GABA-A reseptörlerinin aktivasyonu.	Çoğu nöbet türü
Primidon	10–25 mg/kg/gün	GABA-A reseptörlerinin aktivasyonu.	Çoğu nöbet türü
Rufinamid	45 mg/kg/gün	Voltaj kapılı sodyum kanallarını hedefleyen sodyum kanal engelleyici	Fokal nöbetler, Lennox-Gastaut sendromu ilişkili atakları azaltmada
Tiagabin	<12 yaş: 4–8 mg/kg/gün >12 yaş: en fazla 32 mg/gün	Sinaptik GABA geri alımını bloke ederek GABA konsantrasyonunu artırır.	Fokal nöbetler
Topiramate	5–9 mg/kg/gün	Voltaj kapılı sodyum kanalları, yüksek voltaj aktiviteli kalsiyum kanalları ve GABA-A reseptörleri üzerinde hareket ettiği düşünülür.	Çoğu nöbet türü
Valproat	30–60 mg/kg/gün	Sodyum ve kalsiyum kanallarını hedeflemesi düşünülür.	Tüm nöbet türleri
Vigabitran	Bebeklerde: 50– 150 mg/gün. Büyük çocuklarda: 500– 3,000 mg/gün	Mitokondriyal enzim olan GABA-transaminazın inhibisyonuyla GABA konsantrasyonunu artırır.	İnfantil spazmlar, fokal nöbetler
Zonisamid	5–8 mg/kg/gün	Voltaj kapılı sodyum kanallarını hedefler ve muhtemelen düşük voltajlı kalsiyum kanallarına etki eder.	GABA-A =Gama- aminobutirikasid A.

(Kaynak: Rubenstein ve Levy, 2019)

AEİ tedavisinin öncelikli amacı ilaçların yan etkisi olmaksızın nöbet kontrolünü sağlamaktır. Parsiyel başlangıçlı nöbetlerin tedavisinde etkinliği bilinen en iyi ilaçlar

arasında CBZ, okskarbazepin (OXC), levatirasetam (LEV), valproik asit (VPA), lamotrijin (LTG) iken fenitoin (PHT), zonisamid (ZNS) ve topiramat (TPM) diğer seçenekler arasında sayılabilir. Ayrıca eslikarbazepin (ESL), lakosamide (LCM), fenobarbital (PB), pregabalin (PGB), TGB ve vigabatrin (VGB)de kullanılabilir. Jeneralize tonik klonik nöbetlerde VPA ilk tercih iken LTG de düşünülebilir. Bunun yanı sıra VPA, LEV, TPM, ZNS, LTG, felbamat (FBM) ve perampanel (PER) fokal ve/veya jeneralize epilepsilerin tedavisinde kullanılan geniş spektrumlu AEİ'lerdir. Primer jeneralize epilepsi sendromlarında rufinamid (RFM) ve etosüksimid (ESM) kullanılıyorken parsiyel epilepside tipik olarak kullanılmazlar. Absans nöbetlerde VPA ve ESM ilk tercih olarak kullanılabilir. LEV ise miyoklonik nöbetlerde daha etkilidir. Kompleks parsiyel nöbetlerin tedavisinde ise CBZ etkiliyken jeneralize tonik klonik nöbetlere ilerleyen parsiyel nöbetlerde CBZ ve VPA'nın etkinliği eşit olarak gözlenmiştir. Ayrıca gabapentin (GBP), LTG, TPM, OXC, LEV ZNS ve TGB dirençli parsiyel epilepsili hastalarda ve TPM dirençli jeneralize tonik klonik nöbetlerde kullanılabilir. Tonik ve atonik nöbetlerde ilk tercih VPA iken infantil spazm tedavisinde ise ilk tercih adrenokortikotropik hormon veya VGB'dir. Bununla birlikte Dravet sendromunda VPA veya TPM önerilirken, Lennox-Gastaut sendromunda ilk tercih olarak VPA düşünülmektedir (İpek, Kömür ve Okuyaz, 2020).

AEİ başlarken düşük dozda, monoterapi ile başlanmalıdır. Gerekli olması halinde doz kademeli olarak artırılabilir. Nöbetlerin tekrarladığı olgularda ilaç dozu maksimum farmakolojik doza veya tolere edilebilecek doza kadar artırılabilir. Günlük aktiviteyi engellemeyecek şekilde iki doz halinde verilebilen uzun etkili formlar tercih edilebilir. Aynı zamanda hasta ve ilaç uyumuna dikkat edilmelidir. Yanıt alınamayan durumlarda başka uygun ilaca geçilmelidir. Mevcut kullanılan ilaçlara tekrar yanıt alınamaması halinde ilaç düzeyleri başta olmak üzere ilaç etkileşimleri, nöbeti tetikleyici diğer faktörlerin araştırılması gerekebilir. Ayrıca gerekli durumda tanı yeniden gözden geçirilip nonpileptik paroksizmal olaylar, hareket bozuklukları ya da ilerleyici özellik gösteren nörolojik hastalıklar yönünden araştırılmalıdır (İpek, Kömür ve Okuyaz, 2020). Acil olmayan durumlarda veya yükleme gerekli olmadığında, seçilen antiepileptik tedavinin idame dozu başlanır. OXC, CBZ, TPM ve PER gibi bazı ilaçlarda, sedasyon ve benzeri yan etkilerinin tolere edilebilmesi için daha düşük dozlarla başlanır ve kademeli olarak

idame dozuna yükseltilir. ZNS, PB, PHT veya VPA gibi diğer bazı ilaçlarda idame dozundan başlamak genellikle tolere edilir. LEV ve GBP gibi bazı ilaçlarda ise, her iki yaklaşım da kullanılabilir. Hastalar olası yan etkiler hakkında bilgilendirilmeli ve bunlar takipte kontrol edilmelidirler (Mikati ve Tchapyjnikov, 2020b).

Vitamin duyarlı epilepsiler ayrıca özel dikkat gerektirir çünkü erken teşhis edilirlse ve hassas tedavi uygulanırsa, nöbet kontrolü ve nörogelişimsel sonuçlar üzerinde önemli bir etkiye sahip olunabilir. Örnek olarak antikitin eksikliği ile ilişkili epilepsiler için piridoksin, biyotinidaz eksikliği için biotin kullanımı, serebral folat eksikliği için folat kullanımı, biyotin-tiamin duyarlı bazal gangliyon hastalığı (epilepsi eşlik eder, beyindeki tiamin transportunda defekt vardır) gösterilebilir (Mikati ve Tchapyjnikov, 2020b).



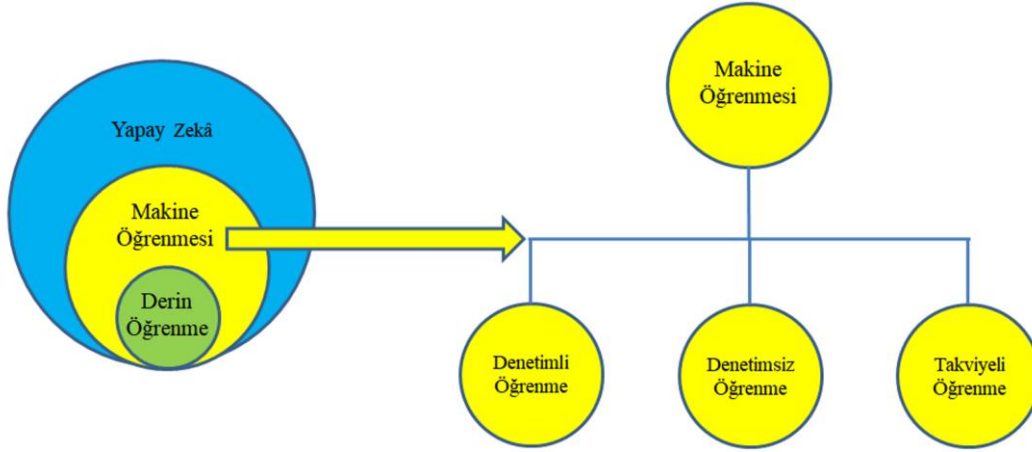
4. YAPAY ZEKÂ

Yapay zekâ, minimum insan müdahalesi ile akıllı davranışı modellemek için bir bilgisayarın kullanımını ima eden genel bir terimdir. Yapay zekânın genel olarak robotların icadıyla başladığı kabul edilir. Terim, zorla çalıştırma olarak kullanılan biyosentetik makineler anlamına gelen Çekçe robota kelimesinden türemiştir. Akıllı makineler yapma bilimi ve mühendisliği olarak tanımlanan yapay zekâ, resmi olarak 1956'da doğmuştur. Bu terim, robotik, tıbbi teşhis, tıbbi istatistik ve bugünün omikleri dahil olmak üzere insan biyolojisi gibi tıpta geniş bir yelpazedeki ögelere uygulanabilir (Hamet ve Tremblay, 2017).

Tıpta yapay zekânın iki ana dalı vardır: sanal ve fiziksel. Sanal branş, derin öğrenme bilgi yönetiminden, elektronik sağlık kayıtları da dahil olmak üzere sağlık yönetim sistemlerinin kontrolüne ve hekimlerin tedavi kararlarında aktif rehberliğine kadar bilişim yaklaşımlarını içerir. Fiziksel dal en iyi şekilde yaşlı hastaya veya ilgili cerraha yardımcı olmak için kullanılan robotlarla temsil edilir. Bu dalda ayrıca, benzersiz bir yeni ilaç dağıtım sistemi olan hedeflenen nanorobotlar da yer almaktadır. Bu uygulamaların toplumsal ve etik karmaşıklıkları, daha fazla yansıma, tıbbi yararlarının, ekonomik değerlerinin kanıtlanması ve daha geniş uygulamaları için disiplinler arası stratejilerin geliştirilmesini gerektirir (Hamet ve Tremblay, 2017).

Yapay zekâ, tıbbi uygulamaları dönüştürmeye hazırdır. Yapay zekâ, hassas tıp, nüfus sağlığı ve doğal dil işleme dahil olmak üzere sağlık ve tıp uygulamalarının çeşitli alanlarında incelenmiştir. Yapay zekânın bilgisayarla görme olarak bilinen görsel görevlere uygulanması, tıp camiasında önemli bir ilgi oluşturmuştur. Bu nedenle, yapay zekânın radyoloji, patoloji, oftalmoloji ve dermatoloji gibi görsel odaklı uzmanlıklarla ilgili olduğuna inanılmaktadır. Yapay zekânın gelişiminin arkasındaki yakıt, büyük dijital veri kümelerinin mevcudiyetidir; derin öğrenme algoritmaları, bir görüntüdeki lezyonu tanımlamak gibi belirli bir görevi gerçekleştirmek üzere kendilerini eğitmek için bu veri kümelerini kullanmaktadır (Kulkarni, Seneviratne, Baig ve Khan, 2020).

Şekil 2'de makine öğrenmesinin yapay zekâ-derin öğrenme ile ilişkisi ve makine öğrenmesinde yaygın olarak kullanılan öğrenme türleri gösterilmiştir.



Şekil 2. Makine öğrenmesinin yapay zekâ-derin öğrenme ile ilişkisi ve makine öğrenmesinde yaygın olarak kullanılan öğrenme türleri

4.1. Makine Öğrenmesi

Makine öğrenmesi, bilgisayarların verilerden nasıl öğrendiğine odaklanan bilimsel disiplindir. Verimli hesaplama algoritmalarına vurgu yaparak, verilerden ve bilgisayar bilimlerinden ilişkileri öğrenmeyi amaçlayan istatistiklerin kesiştiği noktada ortaya çıkar. Matematik ve bilgisayar bilimi arasındaki bu evlilik, milyarlarca veya trilyonlarca veri noktası içerebilen devasa veri kümelerinden istatistiksel modeller oluşturmanın benzersiz hesaplama zorluklarından kaynaklanmaktadır (Deo, 2015). KYK algoritması, Naive Bayes (NB) sınıflandırıcı, KA, lojistik regresyon analizi, k-ortalamlar algoritması, DVM'leri ve YSA makine öğrenmesi yöntemlerinden bazılarıdır. Bunların bir kısmı tahmin ve kestirim, bir kısmı kümeleme ve bir kısmı da sınıflandırma yapabilme kabiliyetine sahiptir (Atalay ve Çelik, 2017). Şekil 3'te makine öğrenmesinin türleri ve alt grupları görülmektedir (Atul, 2020).



Şekil 3. Makine öğrenmesinin türleri ve alt grupları

(Kaynak: Atul, 2020)

Makine öğrenmesi, üç türe ayrılmıştır:

- Gözetimli öğrenme
- Gözetimsiz öğrenme
- Takviyeli öğrenme (Atul, 2020).

Gözetimli öğrenme, öğrenmenin bir öğretmen tarafından yönlendirildiği öğrenmedir. Öğretmen olarak hareket eden bir veri seti vardır ve rolü modeli veya makineyi eğitmektir. Model eğitildikten sonra, kendisine yeni veri verildiğinde bir tahmin yapmaktadır veya karar vermektedir (Atul, 2020).

Gözetimsiz öğrenmede, model gözlem yoluyla öğrenir ve verilerdeki yapıları bulmaktadır. Modele bir veri kümesi verildiğinde, içinde kümeler oluşturarak veri kümesindeki kalıpları ve ilişkileri otomatik olarak bulmaktadır. Yapamayacağı şey, kümeye etiket eklemek, sanki bunu bir grup elma veya mango diyemez, ancak tüm elmaları mangolardan ayıracaktır. Modele elma, muz ve mango resimleri sunulursa bazı modellere ve ilişkilere dayanarak kümeler oluşturulur ve veri kümesi bu kümelere bölünür. Şimdi modele yeni bir veri beslenirse, onu oluşturulan kümelere ekler (Atul, 2020).

Takviyeli öğrenmede ise, bir ajanın çevre ile etkileşime girme ve en iyi sonucun ne olduğunu bulma yeteneğidir. Vuruş ve yargılama yöntemini takip eder. Temsilci, doğru veya yanlış cevap için bir puanla ödüllendirilir veya cezalandırılır ve kazanılan olumlu ödül puanları temelinde model kendisini eğitir. Ve yeniden eğitildikten sonra, kendisine sunulan yeni verileri tahmin etmeye hazır hale gelir (Atul, 2020).

Regresyon analizi, iki ya da daha çok nicel değişken arasındaki ilişkiyi ölçmek için kullanılan bir analiz yöntemidir (Gujarati ve Porter, 2018). Doğrusal diskriminant analizi ya da doğrusal ayırma analizi öznel özelliklerin bir doğrusal birleşimini bularak veriyi sınıflara ayırmaya yarayan bir metottur (Doğan, Orman, Örkücü ve Örkücü, 2019).

Polinom regresyonu, bağımsız değişken x ve bağımlı değişken y arasındaki ilişkinin x 'te n 'inci derece polinomu olarak modellendiği bir regresyon analizi biçimidir. Polinom regresyon, x 'in değeri ile $E(y | x)$ olarak gösterilen karşılık gelen y 'nin koşullu ortalaması arasındaki doğrusal olmayan bir ilişkiye uyar. Polinom regresyon, verilere doğrusal olmayan bir modele uysa da, istatistiksel bir tahmin problemi olarak doğrusaldır, yani regresyon fonksiyonu $E(y | x)$, verilerden tahmin edilen bilinmeyen parametrelerde doğrusaldır. Bu sebeple, polinom regresyon, çoklu doğrusal regresyonun özel bir biçimi olarak kabul edilir (Chang, Hsieh, Chang, Ringgaard ve Lin, 2010).

KA, ağaç benzeri bir karar modelini ve bunların olası sonuçlarını, şans olaylarının sonuçları, kaynak maliyetleri ve fayda dâhil olmak üzere kullanan bir karar destek aracıdır. Sadece şartlı kontrol ifadeleri içeren bir algoritmayı görüntülemenin bir yoludur. KA, bir hedefe ulaşma olasılığı en yüksek olan ancak aynı zamanda makine öğreniminde popüler bir araçtır (Kamieński, Jakubczyk ve Szufel, 2017).

Rastgele karar ormanları veya RO'lar sınıflandırma, regresyon ve diğer görevler için, eğitim aşamasında çok sayıda KA oluşturarak problemin tipine göre sınıf (sınıflandırma) veya sayı (regresyon) tahmini yapan bir toplu öğrenme metodudur (Ho, 1998). Sınıflandırma, örüntü tanımanın bir örneğidir. İstatistikte sınıflandırma, kategori üyeliği bilinen gözlemleri (veya örnekleri) içeren bir eğitim veri setine dayalı olarak yeni bir gözlemin hangi kategorilere (alt popülasyonlara) ait olduğunu belirleme problemidir. Örneğin, belirli bir e-postayı "istenmeyen posta" veya "istenmeyen posta olmayan" sınıfına atamak gibi. Belirli bir hastaya, hastanın gözlenen özelliklerine (cinsiyet, kan basıncı, belirli semptomların varlığı veya yokluğu vb.) göre bir tanı atamaktır. Makine öğrenimi terminolojisinde, sınıflandırma, denetimli öğrenmenin bir örneği olarak kabul edilir, yani doğru tanımlanmış gözlemlerden oluşan bir eğitim setinin mevcut olduğu öğrenmedir. Karşılık gelen denetimsiz işlem, kümeleme olarak bilinir ve bazı içsel benzerlik veya mesafe ölçülerine göre verilerin kategoriler halinde gruplandırılmasını içerir (Gentile, 2020).

KYK sınıflandırması, sadeliği ve verimliliği nedeniyle en iyi 10 veri madenciliği algoritmasından biri olarak kabul edilmiştir. Böylece KYK yöntemi, sınıflandırma, regresyon ve kayıp değer atama gibi veri madenciliği uygulamalarında başarıyla geliştirilmiştir. Standart bir KYK yönteminin ana fikri, çoğunluk kuralına göre bir test veri noktasının etiketini tahmin etmektir, yani test veri noktasının etiketi, özellik alanında en benzer k eğitim veri noktasının ana sınıfı ile tahmin edilir. KYK yöntemi, basit uygulaması ve üstün performansı nedeniyle veri madenciliği ve makine öğrenimi uygulamalarında yaygın olarak kullanılmaktadır (Zhang, Li, Zong, Zhu ve Wang, 2017).

Lojistik regresyon, sağlık bilimlerinde kullanılan en popüler çok değişkenli yöntemdir. Lojistik model veya logit model olarak da adlandırılan lojistik regresyon, birden fazla bağımsız değişken ile kategorik bir bağımlı değişken arasındaki ilişkiyi analiz eder ve verileri bir lojistik eğriye uydurarak bir olayın meydana gelme olasılığını tahmin eder. İki lojistik regresyon modeli vardır, ikili lojistik regresyon ve çok terimli lojistik regresyon. İkili lojistik regresyon, tipik olarak, bağımlı değişken ikili olduğunda ve bağımsız değişkenler sürekli veya kategorik olduğunda kullanılır. Bağımlı değişken ikili

olmadığında ve ikiden fazla kategoriden oluştuğunda, çok terimli bir lojistik regresyon kullanılabilir (Park, 2013).

Bayes teoremi, çıkarımsal istatistikler ve birçok gelişmiş makine öğrenimi modeli için temel öneme sahiptir. Bayesci akıl yürütme, yeni kanıtlar ışığında hipotezlerin olasılığını güncellemeye yönelik mantıklı bir yaklaşımdır ve bu nedenle bilimde haklı olarak çok önemli bir rol oynar. Bayes analizi, sık kullanılan istatistiksel yaklaşımların geliştirilmediği soruları yanıtlamamıza olanak tanır. Aslında, bir hipoteze olasılık atama fikri, frekans paradigmasının bir parçası değildir. Basit ama güçlü bir makine öğrenme algoritmasının temelini oluşturan Bayes teoremi, NB sınıflandırıcısıdır (Berrar, 2019). NB algoritması basit bir olasılık sınıflandırmasıdır. Bu algoritma belirli bir veri setindeki değerlerin frekansını ve kombinasyonunu hesaplayarak bir dizi olasılık hesaplar. Verilerdeki belirli özniteliklerin olasılığı, bir olasılık dizisinde üyeler olarak görünür ve eğitim veri setinden sınıftaki her bir öznitelik değerinin frekansı hesaplanarak elde edilir. Eğitim veri seti, sınıflandırma algoritmalarını eğitmek için kullanılan bir alt kümedir. Eğitim süreci, bilinmeyen değerleri tahmin etmek için bilinen değerleri kullanır (Wibawa, Kurniawan, Murti, Adiperkasa, Putra, Kurniawan ve Nugraha, 2019).

DVM, istatistiksel öğrenme teorisindeki matematiksel temeli nedeniyle makine öğrenimi problemlerine ilkeli bir yaklaşım sunar. DVM, çözümünü eğitim girdisinin bir alt kümesi olarak oluşturur. DVM, sınıflandırma, regresyon, yenilik algılama görevleri ve özellik indirgeme için yaygın olarak kullanılmıştır (Awad ve Khanna, 2015).

Kümeleme yöntemleri genel olarak iki gruba ayrılabilir: hiyerarşik ve bölümsel yöntemler. Bölmeli kümeleme yöntemleri, bir dizi veri noktasını aynı anda alt kümelere böler. Kullanılacak kümeleme algoritmasının seçiminin, mesafe hesaplamasının seçiminden daha önemli olduğu belirtilmelidir (Thrun, 2018).

Tekil değer ayrışımı, bilim ve teknolojinin çeşitli alanlarında kullanılan bir matris çarpanlara ayırma yöntemidir. Ayrıca, makine öğrenimi, veri madenciliği ve teorik bilgisayar bilimindeki son büyük gelişmeler nedeniyle, Tekil değer ayrışımının gittikçe daha önemli olduğu bulunmuştur. Sadece güçlü bir araç ve teori değil, aynı zamanda bir sanattır. Tekil değer ayrışımı, matrisleri veri biliminin bir "dili" haline getirir (Tekil Değer Ayrışımı, 2020-2021).

Temel bileşen analizi, gözlemlerin birbiriyle ilişkili birkaç nicel bağımlı değişken tarafından tanımlandığı bir veri tablosunu analiz eden çok değişkenli bir tekniktir. Amacı, temel bileşenler adı verilen bir dizi yeni ortogonal değişken olarak temsil etmek için istatistiksel verilerden önemli bilgileri çıkarmak ve gözlemler ile değişkenler arasındaki benzerlik modelini nokta haritalarında noktalar olarak göstermektir. Matematiksel olarak, Temel bileşen analizi, pozitif yarı tanımlı matrislerin öz ayrışmasına ve dikdörtgen matrislerin tekil değer ayrıştırmasına bağlıdır. Özvektörler ve özdeğerler tarafından belirlenir. Özvektörler ve özdeğerler, kare matrislerle ilişkili sayılar ve vektörlerdir. Birlikte, korelasyon, kovaryans veya çapraz çarpım matrisleri gibi bu matrisin yapısını analiz eden bir matrisin öz ayrışmasını sağlarlar. Temel bileşen analizini gerçekleştirmek pratikte oldukça basittir. Bir veri kümesi $m \times n$ matrisi olarak düzenlenir; burada m , ölçüm türlerinin sayısı ve n , deneme sayısıdır. Her ölçüm türü veya satır x_i için ortalamanın çıkarılmasıdır. Tekil değer ayrıştırması veya ortak varyansın özvektörlerini hesaplanır. Günümüz hayatında bilerek veya bilmeyerek çok değişkenli veri analizi ve görüntü sıkıştırmanın alternatif olarak kullanıldığı Temel bileşen analizinin birçok ilginç uygulamasının olduğu tespit edilmiştir (Mishra, Sarkar, Taraphder, Datta, Swain, Saikhom, Panda ve Laishram, 2017).

K-ortalamlar yöntemi, bölmeli kümeleme yönteminin belirgin ve iyi bilinen bir örneğidir. Aşağıdaki gibi ilerler: Küme sayısı seçildikten sonra, girdi alanında centroidler adı verilen küme merkezlerinin rastgele başlatılması gerçekleştirilir. Ardından, her bir merkeze en yakın veri noktaları o merkeze atanır. Veri noktalarının eşlenmesinden sonra, merkezler, atanan noktalardan ilgili merkez noktalarına olan mesafeler en aza indirilecek şekilde hareket ettirilir. Bu işlem tekrar tekrar gerçekleştirilir. Özetle, k-ortalama merkezleri, bireysel veri noktalarından ziyade ortalama noktalardır (Thrun, 2018).

İlişkilendirme kuralı madenciliği, finans, tıp, internet ve diğer alanlarda yaygın olarak kullanılan veri madenciliği alanındaki önemli araştırma yöntemlerinden biridir. Orijinal birliktelik kuralları madenciliği sepet analizi problemi için ileri sürülmüştür, amacı farklı emtia işlem veri tabanları arasındaki ilişkiyi bulmak, müşteri satın alma modellerinin genel kurallarını elde etmektir. Bu kurallar, tüccarlara satın alma, envanter ve raf tasarımını makul bir şekilde düzenleme konusunda rehberlik edebilir (Zhan).

İlişkilendirme kuralı madenciliği süreci iki aşamadan oluşmaktadır. İlk aşama, öncelikle veri toplama ekibinden tüm yüksek frekanslı bulmalıdır. İkinci aşama, yüksek frekanslı ekip tarafından birliktelik kurallarında ekibin güvenini karşılamasının hesaplanmasıdır. (Zhan, Zhu, Zhang, Wang, Wang and Liu, 2018).

Apriori algoritması, veri madenciliğinde popüler ve klasik bir algoritmadır. Birçok algoritma birliktelik kuralı madenciliği kapsamına girer ancak Apriori algoritması tipik algoritmalarından biridir. Agarwal tarafından 1993 yılında tanıtıldı; öge kümeleri arasında ilişki bulmaya yardımcı olan güçlü bir algoritmadır. Apriori algoritmasının temel bir özelliği, "sık bir öge setlerinin her alt kümesi hala sık öge kümesidir ve sık olmayan bir öge kümesinin her üst kümesi bir sık öge kümesi değildir". Bu özellik, apriori algoritmasında sık kullanılan tüm öge kümelerini keşfetmek için kullanılır (Rajeswari, 2015).

Sık kalıp büyümesinde, sık kalıp sık örüntü anlamına gelir (Han, Pei ve Yin, 2000). İlk geçişte, algoritma işlemlerin veri kümesindeki öğelerin (öznitelik-değer çiftleri) oluşumlarını sayar ve bu sayıları bir "başlık tablosunda" saklar. İkinci geçişte, işlemleri bir dijital ağaca ekleyerek sık kalıp-ağaç yapısını oluşturur. Ağacın hızlı bir şekilde işlenebilmesi için, her işlemdeki öğeler, eklenmeden önce veri kümesindeki sıklıklarının azalan sırasına göre sıralanmalıdır. Her işlemdeki minimum destek gereksinimini karşılamayan öğeler atılır. Çoğu işlem en sık öğeleri paylaşıyorsa, sık kalıp ağacı ağaç köküne yakın yüksek sıkıştırma sağlar. Ana veri kümesinin bu sıkıştırılmış sürümünün yinelemeli işlenmesi, aday öğeler oluşturmak ve bunları tüm veritabanında test etmek (uygunluk algoritmasında olduğu gibi) yerine, sık sık öge kümelerini doğrudan büyütür. Büyüme, başlık tablosunun altından, yani en küçük desteğe sahip öğeden, o öğeyle biten tüm sıralı işlemleri bularak başlar. Bu eşyayı ara. Orijinal sık kalıp ağacının üzerine yansıtıldığı yeni bir koşullu ağaç oluşturulur. Öngörülen ağaçtaki tüm düğümlerin destekleri, her düğümün çocuk sayımlarının toplamını alarak yeniden sayılır. Minimum desteği karşılamayan düğümler (ve dolayısıyla alt ağaçlar) budanır. Yinelemeli büyüme, koşullu hiçbir öge minimum destek eşiğini karşılamadığında sona erer. Kökten sonuca doğru ortaya çıkan yollar sık öge kümeleri olacaktır. Bu adımdan sonra, işleme orijinal sık kalıp ağacının bir sonraki en az desteklenen başlık ögesiyle devam eder. Özyinelemeli

süreç tamamlandığında, tüm sık kullanılan öge setleri bulunacak ve ilişkilendirme kuralı oluşturma başlayacaktır (Witten, Frank ve Hal, 2011).

4.2. Sağlıkta Yapay Zekâ

Yapay zekâ kullanan yaşam bilimleri araştırmacıları, her zamankinden daha hızlı yenilik yapma baskısı altındadır. Büyük, çok düzeyli ve entegre veri kümeleri, yeni içgörülerin kilidini açma ve atılımları hızlandırma vaadini sunar. Her zamankinden daha fazla veri mevcut olmasına rağmen, yalnızca bir kısmı küratörlüğünü yapar, entegre eder ve analiz eder. Yapay zekâ, bilgisayarların verilerden nasıl öğrendiğine ve insan düşünce süreçlerini nasıl taklit ettiğine odaklanır. Yapay zekâ, öğrenme kapasitesini artırır ve sağlık hizmetlerinin geleceğini dönüştüren ölçeklerde karar destek sistemi sağlamaktadır (Noorbakhsh-Sabet, Zand, Zhang ve Abedi, 2019).

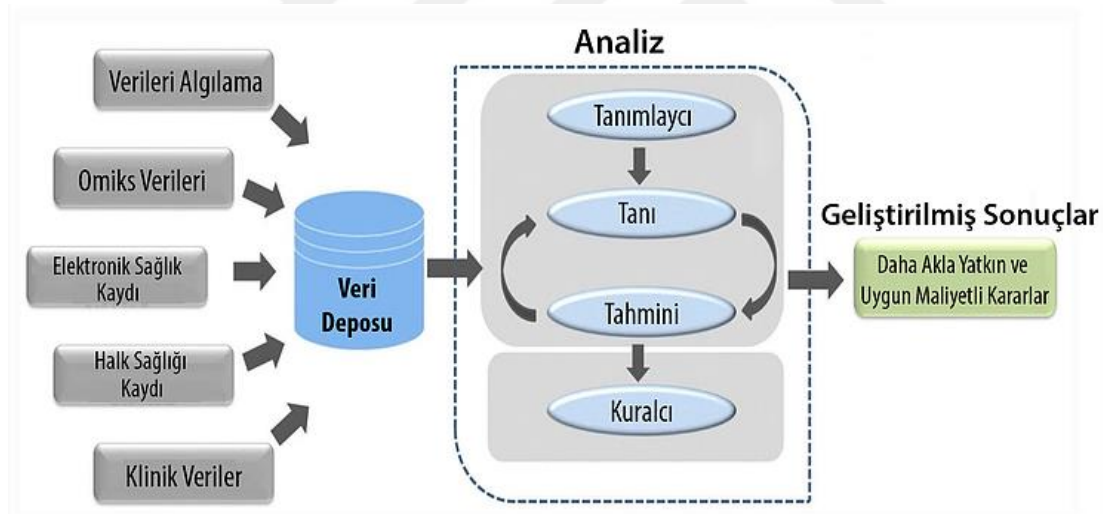
Birçok düşük ve orta gelirli ülkede bilgi teknolojisi altyapısı ve mobil bilgi işlem gücündeki eşzamanlı gelişmeler, yapay zekanın küresel sağlık alanına özgü zorlukların üstesinden gelmeye ve sağlıkla ilgili sürdürülebilir kalkınma hedeflerinin başarısını hızlandırmaya yardımcı olabileceğine dair umutları arttırmıştır. Yapay zekâ güdümlü sağlık müdahaleleri ve yeni teknolojiler hakkında etik ve kanıta dayalı kararlar almak için geleneksel olarak kullanılan araçların, yöntemlerin ve korumaların yapay zekâyâ uygulanıp uygulanamayacağı hakkında bir dizi temel soru gündeme gelmiştir. Yapay zekânın konuşlandırılması, tüberküloz ve sıtma dahil olmak üzere öncelikle bulaşıcı hastalıklara odaklanan müdahalelerle, düşük ve orta gelirli ülkelerde ortak olan çok çeşitli sağlık sorunları için şimdiden başlamıştır. Yapay zekâ türleri değişir, ancak çoğu bir tür makine öğrenimi veya sinyal işleme kullanmaktadır. Çoğu zaman sinyal işleme olmak üzere diğer yaklaşımlarla birlikte makine öğrenimi gibi, çeşitli makine öğrenimi yöntemleri sıklıkla birlikte kullanılır. (Schwalbe and Wahl, 2020)

Yapay zekâ güdümlü sağlık müdahaleleri, küresel sağlık araştırmacılarıyla ilgili dört kategoriye ayrılır: (1) teşhis, (2) hasta morbidite veya mortalite risk değerlendirmesi, (3) hastalık salgını tahmini ve gözetimi ve (4) sağlık politikası ve planlaması. Bununla birlikte, küresel sağlıkta yapay zekâ güdümlü müdahale araştırmalarının çoğu, yaygın kullanım veya geniş ölçekte dağıtım için gereken etik, düzenleyici veya pratik hususları

açıklamamaktadır. Alanın gelişmeye devam etmesine rağmen, yapay zekâ güdümlü sağlık müdahaleleri, düşük ve orta gelirli ülkelerde daha iyi sağlık sonuçlarına yol açabilir. Bu müdahaleleri geliştirmenin ve dağıtmanın bazı zorlukları bu ortamlara özgü olmasa da, küresel sağlık topluluğunun geliştirme, test etme ve kullanım için kılavuzlar oluşturmak için hızlı bir şekilde çalışması ve eşitlikçi ve etik kullanımı kolaylaştırmak için kullanıcı odaklı bir araştırma gündemi geliştirmesi gerekecektir. (Schwalbe and Wahl, 2020).

Yapay zekâ sağlık hizmetlerinde hızla gelişmektedir ve sağlık kuruluşlarının şu anda karşılaştığı en acil sorunlardan bazılarını çözmek için çeşitli yapay zekâ uygulamaları geliştirilmiştir. Sağlık liderlerinin yapay zekâ teknolojilerinin durumunu ve bu tür teknolojilerin sağlık hizmetlerinin verimliliğini, güvenliğini ve erişimini iyileştirmek ve değere dayalı bakım sağlamak için nasıl kullanılabileceğini anlamaları çok önemlidir. (Chen & Decary, 2020)

Şekil 4'de kişiselleştirilmiş tedaviyi teşvik etmek için omik verilerini ve sağlık hizmetleri analitiğini entegre etmeye yönelik bir çerçeve görülmektedir.



Şekil 4. Kişiselleştirilmiş tedaviyi teşvik etmek için omik verilerini ve sağlık hizmetleri analitiğini entegre etmeye yönelik bir çerçeve

(Kaynak: Dash 2019)

Türkiye’de makine öğrenmesi algoritmaları kullanılarak sağlık alanında yapılan 68 çalışma incelenmiştir. Çalışmaların en fazla dolaşım ve solunum sistemleri hastalıkları alanında yapıldığı görülmektedir (Şekil 5) (Adar ve Kılıç Delice, 2019a).

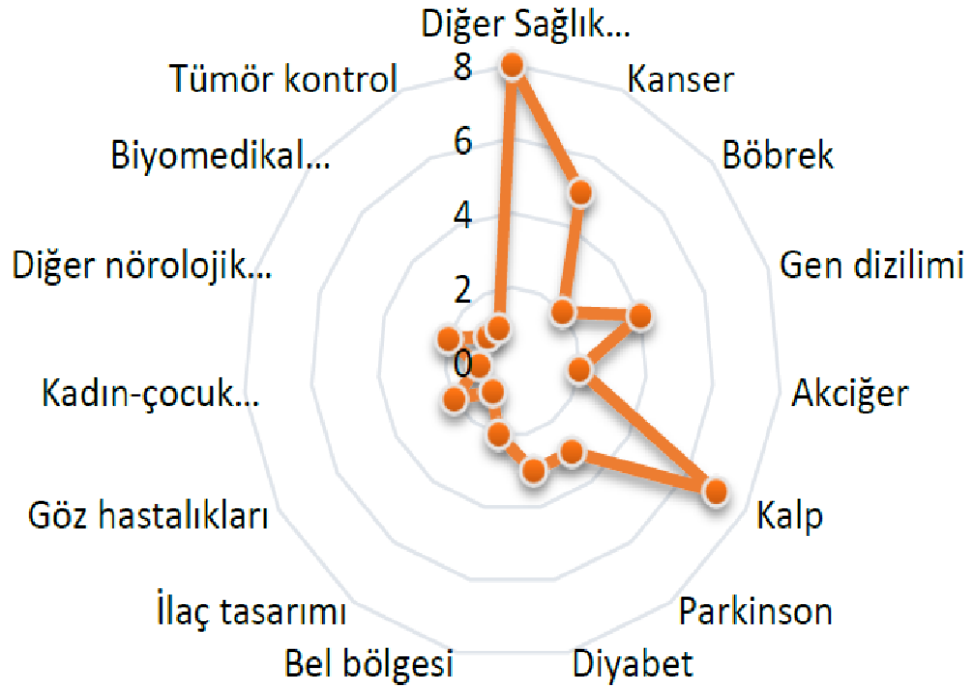
Ülkemizde 2019 yılında yapılan bir çalışmada Türkiye’de makine öğrenmesi ile ilgili sağlık alanında 44 tez çalışması yapıldığı rapor edilmiştir. Şekil 6’da görüldüğü gibi kalp rahatsızlıkları, gen dizilimi ve kanser konuları ile ilgili daha fazla çalışma olduğu görülmektedir (Adar ve Kılıç Delice, 2019b).



Şekil 5. Makine öğrenmesi algoritmaları ile çalışılan sağlık alanları

(Kaynak: Adar ve Kılıç Delice, 2019a)

Sağlık alanında yapılan çalışmalar



Şekil 6. Sağlıkta makine öğrenmesi çalışma alanları

Kaynak: (Adar ve Kılıç Delice, 2019b)

5. ÇALIŞMADA KULLANILAN ALGORİTMALAR

5.1. Gaussian Naive Bayes (GNB)

Bayes Teoremi, olasılık kuramı içinde incelenen önemli bir kuramdır. Bu teorem, bir rassal değişken için olasılık dağılımı içinde koşullu olasılıklar ile uç (marjinal) olasılıklar arasındaki ilişkiyi gösterir. Bayes Teoremi, stokastik bir süreç sırasında ortaya çıkan bir rastgele olay A ile bir diğer rastgele olay B için koşullu olasılıkları ve uç olasılıkları arasındaki ilişkisi denklem 1’de gösterilmiştir (Papoulis, 1984; Domingos ve Pazzani, 1997, Budak, 2014; Budak vd., 2016).

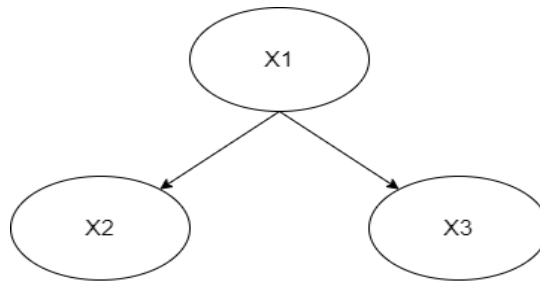
$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (1)$$

$P(A|B)$: A olayı gerçekleştiği durumda B olayının meydana gelme olasılığıdır.

$P(B|A)$: B olayı gerçekleştiği durumda A olayının meydana gelme olasılığıdır.

$P(A)$ ve $P(B)$: A ve B olaylarının önsel olasılıklarıdır.

Burada önsel olasılık Bayes teoremine öznellik katar. Diğer bir ifadeyle $P(A)$ henüz elde veri toplanmadan A olayı hakkında sahip olunan bilgidir. Diğer taraftan $P(B|A)$ ardıl olasılıktır veri toplandıktan sonra, A olayının gerçekleşmiş olduğu durumlarda B olayının gerçekleşme ihtimali hakkında bilgi verir. Şekil 7’de basit bir Bayes ağının birleşik olasılık dağılımı verilmiştir. Denklemi ise Denklem 2’de verilmiştir.



Şekil 7. Basit bayes ağının birleşik olasılık dağılımı

$$P(X_1, X_2, X_3) = P(X_2|X_1). P(X_3|X_1). P(X_1) \quad (2)$$

Şekil 7’de X_1 ’den X_2 ’ye bir kenarın var olduğu görülmektedir. Bu durumda X_1 , X_2 ’nin ebeveyni olur ve X_1 ’in X_2 üzerinde doğrudan bir etkisi vardır. X_2 ’nin bir ebeveyni olduğundan, kendisinin yerel olasılık dağılımı da koşullu olur. $i=1, 2, 3, \dots, n$ olmak üzere her i ’ye ait ebeveynler kümesi ebeveyn(X_i) olarak gösterilirse Bayes ağları için genelleştirilebilecek bileşik olasılık dağılımı ifadesi denklem 3’deki hali alır (Özkan, 2008).

$$P(X_1 = x_1, \dots, X_n = x_n) = \prod_{i=1}^n P\left(X_i = x_i \mid \text{ebeveyn}(X_i)\right) \quad (3) \quad (3)$$

Örnek vermek istersek, bir kişinin kadın mı erkek mi olduğuna dair bir veri seti içinde kayıt esnasındaki kişinin uzunluğu, kilosu ve ayak numarası gibi veriler de vardır. Bir insan böyle bir analizi kendi beyinde yaparak tahmin yöntemleri ile kişinin kadın mı erkek mi olduğuna karar verebilir. Dolayısıyla öğretilen sisteme kişinin uzunluğu 1.80, ağırlığı 95 ve ayak numarası 42 şeklinde bir bilgi verildiğinde eğitilmiş sistem daha önce gerçekleşmiş istatistiklerden faydalanarak kişinin kadın mı erkek mi olduğu ihtimalini hesaplar ve tahminini bildirir (Han ve ark., 2012).

5.1.1. Naive Bayesin Yapısı

Naive Bayes sınıflandırıcısının temeli Bayes teoremine dayanır. Bayes sınıflandırma algoritması, tembel (*lazy*) bir öğrenme algoritmasıdır. Aynı zamanda dengesiz veri kümelerinde de çalışabilir. Algoritmanın çalışma şekli şöyledir: Bir eleman için her durumun olasılığı hesaplanır ve olasılık değeri en yüksek olana göre sınıflandırılır. Az bir eğitim verisiyle çok yüksek başarı performanslarına ulaşılabilir. Test kümesindeki bir kelimenin eğitim kümesinde gözlemlenemeyen bir değeri varsa olasılık değeri olarak 0 verir. Bu tahmin yapılamayacağı anlamına gelir. Bu durum genellikle Sıfır Frekans (*Zero Frequency*) adıyla bilinir. Bu durumu gidermek için düzeltme teknikleri kullanılabilir. En basit düzeltme tekniklerinden biri *Laplace* tahmini olarak bilinir.

Kullanım alanlarına örnek olarak gerçek zamanlı tahmin, çok sınıflı tahmin, metin sınıflandırması, spam filtreleme, duyarlılık analizi, tıbbi tanı koyma ve öneri sistemleri verilebilir. (Hatipoğlu, 2018)

Naive Bayes sınıflandırıcılar gözetimli öğrenme ortamında çok verimli bir şekilde eğitilebilir. Naive Bayes'in bir avantajı, sadece sınıflandırma için gerekli parametreleri tahmin etmek için az sayıda eğitim verisi gerektirmesidir (Han vd., 2012).

Gaussian NB Sınıflandırıcı, Bayes Teoremi'nden faydalanılarak oluşturulmuş kolaylıkla anlaşılabilir ve uygulanabilir bir yöntemi denklem 4'de gösterilmiştir.

$$P(e|T) = \frac{P(T|e)P(e)}{P(T)} \quad (4)$$

$P(e)$ = e olayının önsel olasılığı

$P(T)$ = T eğitim verisinin önsel olasılığı

$P(T|e)$ = e olayı verildiğinde T'nin koşullu olasılığı

$P(e|T)$ = T eğitim verisi verildiğinde e'nin koşullu olasılığı

Bu yöntemle bir örneğin hedef niteliğin sınıf değerlerine ait olma olasılıkları bulunabilmektedir (Han ve Kamber, 2006). Gaussian NB, diğer algoritmalarla karşılaştırmak için temel seviye sınıflandırıcı olarak önerilen iyi bilinen bir istatistiksel öğrenme algoritmasıdır. NB, belirli bir sınıf için girdilerin birbirinden bağımsız olduğunu varsayarak sınıf koşulu olasılıkları tarafsız tahmin eder. NB, sınıfları koşullu marjinal yoğunlukları bulmak için ayırıcı sınıflar problemini azaltır, bu da belirli bir örneğin olası hedef sınıflardan biri olma olasılığını temsil eder (Molina, P'erez de la Blanca ve Taylor, 1994).

5.2. K En Yakın Komşu (KYK)

KYK, sınıfları belirli olan bir örnek kümedeki gözlem değerlerinden yararlanarak, örneğe katılacak yeni bir gözlemin hangi sınıfa ait olduğunu tespit etmek amacıyla kullanılan bilinen eski basit sınıflandırma algoritmalarından birisidir. (Özkan, 2008). Büyük ölçekli bir veri kümesinden, bir sorgu örneğine en çok benzer bir veri alt kümesi bulmayı

amaçlayan KYK algoritmaları; boyut indirgeme, model sınıflama ve görüntü alımı gibi geniş kapsamlı uygulamalarda temel bir bileşen olarak kullanılmaktadır (Jégou, Douze ve Schmid, 2011; Zhang, Ji ve Zhang, 2016).

Algoritmanın temeli, örnek veri ile k adet komşu arasındaki mesafenin ölçülmesine dayanmaktadır. Aradaki mesafe hangi mesafe ile daha az ise örnek o sınıfa dahil edilmektedir. Bu algoritma (KYK) sınıflandırılmak istenen yeni bireyin daha önceki bireylerde k tanesine yakınlığına bakıldığı bir algoritmadır (Cover ve Hart, 1967). Sınıflandırma esnasında test örnekleri eğitim örnekleri ile birbirleriyle karşılaştırılır. Yüksek k değerlerine aşırı uyma eğiliminde olmasından dolayı dikkat edilmesi gerekmektedir (Hastie vd., 2009). En yakın komşu algoritması yöntemini uygulayabilmek için belirlenmesi gereken uzaklık ölçüm yöntemi olarak genellikle “Öklid Uzaklığı” ya da “Kosinüs Benzerliği” ölçülerinden biri kullanılır (Kesgin, 2007).

Algoritmada sınıflandırılmak üzere alınan bir bilinmeyene en yakın k adet komşu belirlenir ve bu komşuların ait oldukları sınıflara bakarak, sınıflandırılmak istenen bilinmeyen bu sınıflardan en yakın mesafede olduğu birine ya da birden fazlasına ataması yapılacağı belirlenir. En sık kullanılan k değerleri veri kümesinin büyüklüğüne göre 3, 5 ve 7’dir (Silahtaroglu, 2008). Sınıflandırma işlemi yapılacak eleman, bir uzaklık fonksiyonuyla ölçülen en yakın komşular arasında en yakın olan sınıfa dağıtılır (Li, Zhang, Zhang, Pang, Lam, Hui ve Zhang, 2012). Kullanılan uzaklık ölçüsü aracılığıyla gözlem değerleri arasındaki mesafesinin hesaplanmasının ardından, sınıflamaya esas olacak en yakın k adedi dikkate alınarak değerlendirme yapmak suretiyle yargıya ulaşılır (Silahtaroglu, 2008). KYK algoritmasının formülü denklem 5’te gösterilmiştir.

$$x(x, y) = \sqrt{\sum_{j=1}^k w_j (x - z_j)^2} \quad (5)$$

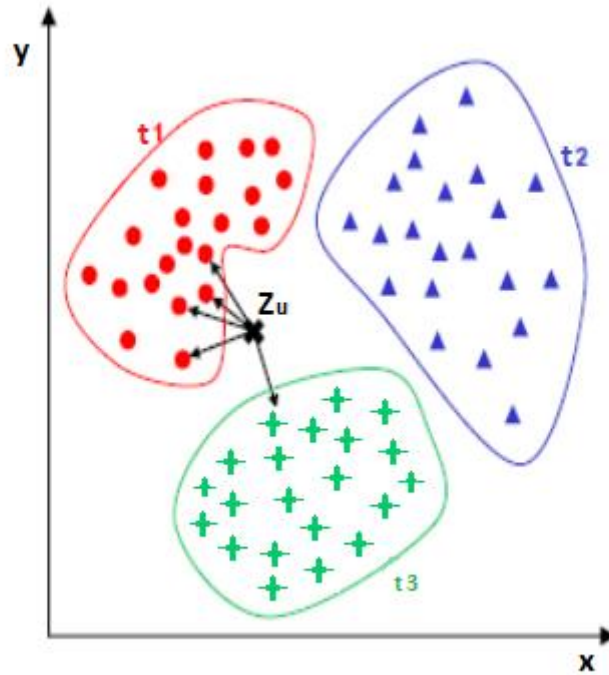
Yöntemde, sınıflandırma yapılacak verilerin öğrenme kümesindeki normal davranış verilerine benzerlikleri hesaplanarak; en yakın olduğu düşünülen k verinin ortalamasıyla, belirlenen eşik değere göre sınıflara atamaları yapılır. Yani sınıflandırmada kullanılan bu algoritmaya göre sınıflandırma sırasında çıkarılan özelliklerden, sınıflandırılmak istenen

Yeni bireyin önceki bireylerden k tanesine yakınlığına bakılmaktadır. Önemli olan, her bir sınıfın özelliklerinin önceden net bir şekilde belirlenmiş olmasıdır.

Yöntemin performansını;

- K en yakın komşu sayısı,
- Eşik değeri,
- Benzerlik ölçütü etkilemektedir (Çalışkan & Soğukpınar, 2008).

t_1 , t_2 ve t_3 olmak üzere üç adet sınıf olduğu varsayalım ve bilinmeyen bir u z örneği sınıflandırılmaya çalışalım. $k=6$ değeri için 6 en yakın komşu incelendiğinde 5 adet komşunun t_1 sınıfında olduğu, 1 adet komşunun ise t_3 sınıfında olduğu görülmektedir. Baskın olan taraf t_1 olduğu için u z örneği t_1 sınıfına dahil olduğuna ait şekil 8’de gösterilmektedir.



Şekil 8. K-en yakın sınıflandırma örneğine ait grafik

KYK’da bahsedilen avantajları yanı sıra bazı kısıtlamalara da sahiptir. Algoritmanın sınıflandırma başarısı özellikle K parametresinin seçimine bağlıdır. K parametresinin çok küçük seçilmesi durumunda algoritma gürültü içeren veriye karşı daha hassas

çalışacaktır. Tersı durumda ise, başka sınıflara ait olan veriler komşuluğa dahil olacaktır. Diğer önemli bir nokta ise uygun uzaklık ölçütünün seçilmesi ile ilgilidir. Uzaklık ölçütü uygulamadan önce farklı birime sahip veriler var ise standardize edilmesi gereklidir. Bağımsız değişkenlerin sürekli olması halinde uygun olan Euclidean uzaklık ölçütünün, değişken sayısının çok fazla olduğu yüksek boyutlu verilerde iyi sınıflama yapamadığı bilinmektedir. Ayrıca, yüksek boyutlu verilerde KYK yüksek bellek kapasitesine ihtiyaç duyar. Bu nedenle daha az sayıda ve ilgili değişkenler ile sınıflandırılma yapılması önerilir (Lantz 2019; Steinbach & Tan, 2009).

5.3. Karar Ağaçları (KA)

KA, eğitim ve testinin hızlı olması, sonuçlarının daha kolay yorumlanabilmesi ve etkin olması sebebiyle sınıflandırmada sıklıkla kullanılan yöntemlerden biridir (Witten, Frank ve Hall, 2011). KA kolay yorumlanabilir olması, diğer sistemlerle kolay entegre edilebilmesi, yapısı dolayısıyla anlaşılabilir kurallar ortaya koyması ve güvenilir olması sebebiyle sınıflama ve regresyon modellerinden en yaygın kullanım alanına sahip bir yöntemlerden biridir (Güner, 2015).

KA oluşturulurken kullanılan algoritmaya göre ağacın yapısı değişebilir. Değişik ağaç yapıları farklı sınıflandırma sonuçları verebilir (Dunham, 2003). İstatistiksel olarak anlamlı grupları bulan KA, sınıfları kolay ve anlaşılabilir ağaç diyagramları halinde ifade eden bir makine öğrenme yöntemidir (Doğan ve Özdamar, 2003).

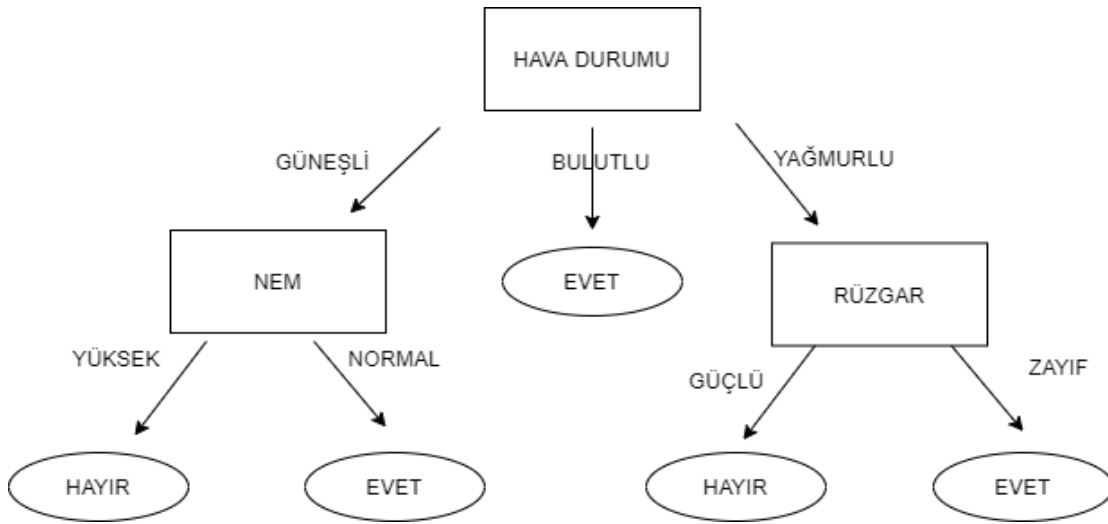
Gözetimli bir öğrenme yöntemi olan KA'lar girdi ve çıktı kümelerinden oluşmaktadır. KA, sınıf çıktılarının durumu ile girdi değişkenleri arasındaki yapıyı keşfeder (Tan, Steinbach ve Kumar, 2005). KA yapılarında aynı sınıfa ait olan veriler yapraklarda bulunur (Huang vd., 2003). Bir KA kök, dal, yaprak ve bunlar arasındaki düğümlerden oluşan bir sisteme sahiptir. Bir dal üzerinde bulunan yaprakta, olası bütün çıktı sınıflarına karşılık gelen sonuç değerleri bulunmaktadır (Kantardzic, 2011). KA yapısında veri miktarının fazlalığı sonucu etkilemez. KA'nın yapısı için oluşturulan teknikler hesaplama yöntemi olarak ucuz ve hızlıdır (Tan, Steinbach ve Kumar, 2006). KA yapısında çok fazla sınıf bulunması durumunda, ağacın yapısı genişlemekte, bunun sonucunda düğüm sayısı

artarken düğümlerdeki sınıf bilgisi azalmaktadır. Bu durumda da sistemin güvenilirlik oranı düşmektedir (Seidman, 2001).

Karar ağaçları parametrik olmayan bir veri analizi yöntemidir. Parametrik yöntemlerin sahip olması gereken homojen varyans, normal dağılım, gözlemlerin birbirinden bağımsız olması en azından aralıklı ölçek olması varsayımlarının her zaman sağlanması zor olduğu için ilgili şart ve/veya varsayımlarının sağlanamadığı durumlarda parametrik olmayan veri analiz yöntemleri kullanılmaktadır. Yani karar ağaçları parametrik yöntemlerin sahip olması gereken varsayımlara dayanmamaktadır. Bu da başka bir tercih nedeni olmaktadır (Saitoğlu, 2015).

KA genelden özele ve aşağıya doğru eğitilmiş veriden üretilmektedir. Karar ağacının ilk aşaması kök düğümdür. Ağaç örneklemdaki tüm verileri içeren bir kök düğümlerle başlamaktadır. Kök düğüm bağımlı değişkeni göstermektedir (Oğuzlar, 2004).

KA diğer bileşenleri dallar ve yapraklardır. Bu yapısı ile ağaç görünümüne benzemektedir. Kök düğümlerle başlayan karar ağaçları aşağı doğru inildikçe çok sayıdaki veri kümesini küçük gruplara bölerek dallara ayırmaktadır. Bu kökten başlayıp dallara doğru büyüyen ağaç yapısında ki her boğuma düğüm, oluşan bu ağaçlarda homojen olmayan düğümlere çocuk düğüm, homojen düğümlere ise terminal düğüm denir (Pehlivan, 2006). KA düğümlerine ait örnek bir şekil 9’da gösterilmiştir.



Şekil 9. Karar ağacı Düğümleri

5.3.1. KA Avantaj ve Dezavantajları

Sınıflandırmada kullanılan KA'nın avantajları;

KA; kolay anlaşılabilme ve yorumlanabilme özelliği nedeniyle kullanışlıdır.

- KA, büyük boyutlardaki veri kümelerine kolaylıkla uygulanabilmektedir.
- Veri madenciliği modeli arasında en sık kullanılan yöntem olduğu için bu alan da yayımlanmış fazla sayıda doküman ve aynı zamanda bilgisayar programı bulunmaktadır (Diler, 2016).
- Hem sürekli hem kesikli değişkenler için uygun olup, ayrıca karmaşık veri setlerine kolayca uygulanabilmektedirler (Oğuzlar, 2004).

KA'nın sağladığı avantajlar yanında dezavantajları da vardır. Bu dezavantajlar ise aşağıdaki gibi belirtilmiştir.

- En uygun ağacı bulmak bazen zor olabilmektedir. Faydalı karar ağacını oluşturmaya çalışılırken birden fazla karmaşık yapıda ve büyük ağaçlar üretilmektedir.
- Girdi miktarının az olduğu hallerde ağaç gerekli bilgiyi elde edememektedir.
- Birtakım algoritmalar sadece kategorik haldeki verilerle çalışmaktadır. Bu yüzden sürekli değişkenlerin sınıflamasında başarılı olmayabilmektedir.
- Zamana olan gereksinim ve ağaç karmaşıklıkları öğrenme kümesinin değişken adedine ve hacmine bağlı olarak değişiklik göstermektedir (Diler, 2016).
- Sürekli değişken değerlerini tahmin etmekte fazla başarılı olmamaktadır,
- Sınıf sayısı fazla ve öğrenme kümesi örnekleri sayısı az olduğu durumlarda model oluşturma fazla başarılı olamamaktadır.
- Yer ve zaman karmaşası değişken sayısına, öğrenme kümesi örnekleri sayısına ve oluşan ağacın yapısına bağlı olmaktadır.

- Ağaç oluşturma ve ağaç budama karmaşıklığı çok olabilmektedir (Köktürk, 2012).

Bilgi kazancı ve bilgi kazanç oranının kullanımı Quinlan (1993) tarafından ortaya atılmıştır. Bu yöntemle göre KA'nda hangi özelliğe göre dallanmanın yapılacağını belirlemek üzere entropi kurallarını içeren bilgi teorisi kullanılmıştır. Entropi bir sistemdeki düzensizliğin ya da belirsizliğin ölçüsüdür. Tek değişkenli karar ağaçlarında ID3 algoritması bilgi kazancı yaklaşımını kullanmaktadır. Bu algoritmanın geliştirilmiş hali olan C4.5 algoritması bölünme bilgisi kavramı ile bilgi kazancından yararlanarak hesaplanan kazanç oranı yaklaşımını kullanmaktadır (Quinlan, 1993). Bir veri setinin $C, C_1, C_2, \dots, C_n$ şeklinde birkaç sınıftan oluştuğu ve T 'nin sınıf değerlerini gösterdiği düşünülürse, bir sınıfa ait olasılık

$P_i = (C_i / |T|)$ olur ve sınıflara ait entropi denklem 6'da gösterilmiştir.

$$Entropi(T) = \sum_{i=1}^n p_i \log_2(p_i) \quad (6)$$

şeklinde hesaplanır. Veri setindeki B özneliğine göre T sınıf değerleri $T_1, T_2, \dots, T_n$ şeklinde alt kümelere ayrıldığı göz önüne alınsın. B öznitelik değerleri kullanılarak T sınıf değerlerinin bölünmesi sonucunda elde edilecek kazanç denklem 7 de gösterilmiştir.

$$Kazanç(B, T) = Entropi(T) - \sum_{i=1}^n \frac{T_i}{T} Entropi(T_i) \quad (7)$$

T kümesi için B özneliğinin değerini belirlenmesinde bölünme bilgisi denklem 8'de gösterilmiştir.

$$Bölünme Bilgisi(B) = - \sum_{i=1}^k \frac{T_i}{T} \log_2\left(\frac{T_i}{T}\right) \quad (8)$$

Bu durumda kazanç oranı bilgisi denklem 9'da gösterilmiştir.

$$Kazanç\ Oranı(B) = - \sum_{i=1}^k \frac{Kazanç(B, T)}{Bölünme\ Bilgisi(B)} \quad (9)$$

Bu eşitlik, sınıflandırma işleminde kullanılacak ayırma ile elde edilen bilgi oranını verir. Bu ölçüt kullanılarak, ağacın her bir düğümünde kazanç oranı maksimum olacak şekilde T eğitim kümesi tekrarlı bir şekilde ayrılır. İşleme her bir yaprak düğümü sadece bir sınıfa ait gözlem değerleri içerene kadar tekrar edilir (Kavzoğlu & Çölkesen, 2010).

5.4. Destek Vektör Makinaları (DVM)

DVM, istatistiksel teoriler üzerine dayanan güçlü bir makine öğrenme algoritmasıdır. Geleneksel yapay zekâ algoritmalarında öğrenme için çok sayıda eğitim verisi gerekmektedir. Bunun ile birlikte birçok sınıflandırma algoritmasında, düşük yakınsama oranı, yerel minimuma takılma ve ezberleme problemleriyle karşılaşmaktadır (Lu, Wang, Leung, Lo, Yuen, Xu ve Fan, 2002). Bu algoritma ikinci dereceden optimizasyon problemini çözerek, yüksek boyutlu değişken uzayda doğrusal olmayan karar sınırlarını tanımlama yeteneğine sahiptir (Hsu, Chang ve Lin, 2010).

Temel DVM teorisi, iki sınıftan puanlar içeren doğrusal olmayan şekilde ayrılabilir bir veri kümesi için, sınıfları bölen sonsuz sayıda çizginin olduğunu belirtir. İki sınıfı (yani karar sınırını) en iyi şekilde ayıran bir çizginin seçimi, destek vektörleri olarak bilinen sadece eğitim numunelerinin bir alt kümesi kullanılarak gerçekleştirilir. Destek vektörleri arasındaki maksimum sınır mesafesi optimal karar sınırını temsil etmek için alınır. Sınıfların doğrusal olarak ayrıştırılamadığı problemler için DVM, Kernel işlevini kullanarak giriş değişkenlerinin örtülü bir dönüşümünü kullanır. Kernel işlevleri, DVM'nin doğrusal bir düzlem kullanarak doğrusal olmayan ayrılabilir destek vektörlerini ayırmasına izin verir (Yu, Porwal, Holden ve Dentith, 2012). Çoğu uygulama için performansı optimize etmek için uygun bir çekirdek işlevinin ve çekirdek genişliğinin seçilmesi gerekir (Hsu vd., 2010). Örnek uzayında x girdi vektörleri, w üst düzlemin

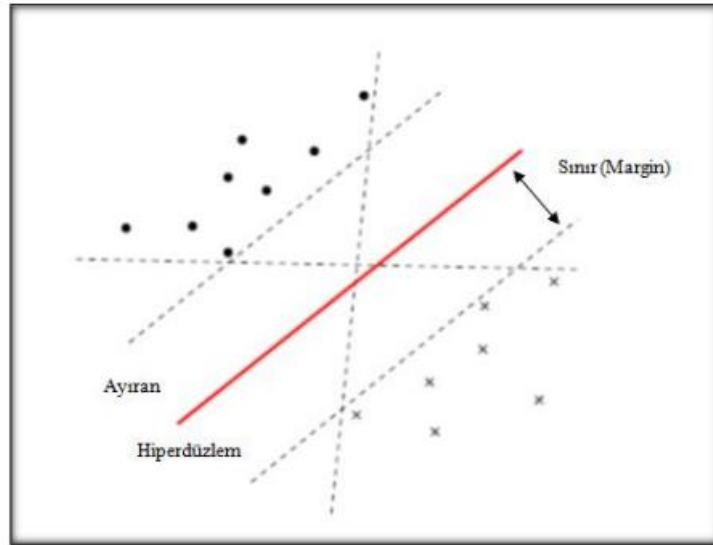
yönünü belirleyen ağırlık vektörü ve w_0 üst düzlemin koordinat sistemindeki yerini belirleyen değişken olmak üzere denklem 10'da gösterilmiştir.

$$g(x) = w^T x + w_0 \quad (10)$$

ve bu üst uzayda, ϕ radyal işlevi olmak üzere, ayırıcı işlevi denklem 11'de gösterilmiştir.

$$g_i(x) = \sum_{j=1}^n w_j \phi_{ij}(x) \quad (11)$$

Verilerin en iyi şekilde ayrıldığını teknik olarak hesaplayan yapı geometrik sınırlardır. Hiper düzlemin uzaklığı veri setindeki en yakın noktaları ayıracak genel ifadeyle verileri en iyi şekilde sınıflandıracak en geniş sınır anlamına gelir. Bir veri setini ayırmak için hiper düzlemde kesikli çizgiler olarak gösterilebilir ancak en iyi genellemeyi yapan yani en iyi sınır (margin) olan şekil 10'da gösterildiği gibi kırmızı çizgidir (Menon Krishna, 2009).

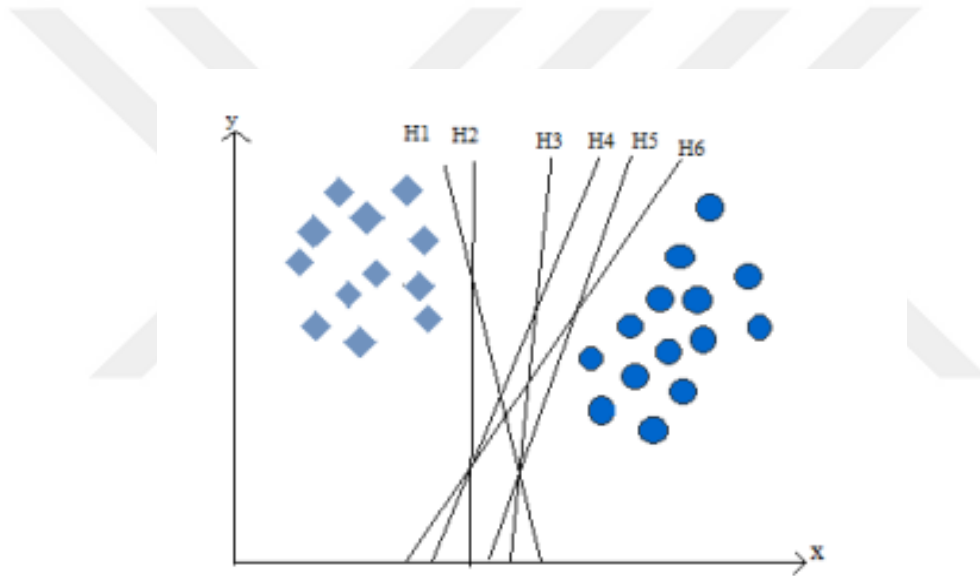


Şekil 10. Destek vektör makineleri veri ayrımı

DVM'de veri setindeki verilerin hiper düzlemin doğru tarafında, doğru sınıfta olmaları yeterli değildir aynı zamanda iyi bir genelleştirme için hiper düzlemden belirli bir uzaklıkta bulunması da gerekmektedir. En uygun hiper düzlemin belirlenmesi sınıflandırmanın başarısını doğrudan etkilemektedir (Han vd., 2012).

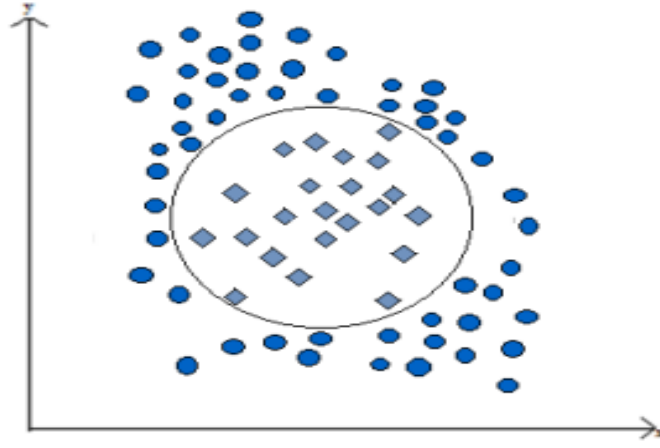
Destek vektör makineleri, sınıflandırma problemlerinin çözümü için verileri marjini en büyükleyecek en uygun düzlem veya hiper düzlem ile ayırmayı amaçlayan bir sınıflandırıcıdır. Marjin, iki farklı sınıfa ait birbirine en yakın veri noktaları arasındaki uzaklık olarak tanımlanmaktadır. Bu noktalar destek vektörleri olarak adlandırılmaktadır. Çok boyutlu uzayda sınıflandırma hatasını en küçükleyen en büyük marjinli hiper düzlemin belirlenmesi amaçlanmaktadır (Han vd., 2012).

Doğrusal olarak ayrılabilen veriler bir düzlem ile ayrılırken doğrusal olarak ayrılamayan veriler doğrusal olmayan haritalama yöntemi ile yüksek boyutlu bir uzaya aktarılmakta, burada hiper düzlem ile sınıflara ayrıldıktan sonra veri noktalarının girdi uzayına iz düşümleri alınmaktadır (Akşehirli vd., 2013). Şekil 11’de doğrusal olarak ayrılabilen verilerin sınıflandırılmasına yönelik örnek görsele yer verilmiştir.



Şekil 11. Doğrusal olarak ayrılabilen verilerin sınıflandırılması

Şekil 12’de doğrusal olarak ayrılamayan verilerin doğrusal olmayan haritalama yöntemi ile sınıflandırıldığı ve veri noktalarının girdi uzayına iz düşümlerinin alındığı bir görsel yer almaktadır.



Şekil 12. Doğrusal olarak ayrılamayan verilerin sınıflandırılması

Doğrusal olarak ayrılabilen veriler bir düzlem ile ayrılırken doğrusal olarak ayrılamayan veriler doğrusal olmayan haritalama yöntemi ile yüksek boyutlu bir uzaya aktarılmakta, burada hiper düzlem ile sınıflara ayrıldıktan sonra veri noktalarının girdi uzayına iz düşümleri alınmaktadır (Akşehirli vd., 2013).

5.4.1. Doğrusal destek vektör regresyon

ε -insensitive destek vektör regresyonda amaç tüm eğitim verisi için gerçek yanıt değişkeni y_i 'lerden maksimum ε sapmaya sahip olan olası en düz $f(\mathbf{x})$ fonksiyonunu bulmaktır. Diğer bir deyişle hatalar ε 'dan küçük olduğu sürece önemsenmeyecek, ancak bu değerden daha büyük bir sapma kabul edilmeyecektir (Burges 1998; Vapnik 1998).

İlk olarak doğrusal durum ele alınsın. f fonksiyonu aşağıda denklem 11'de ifade edilmektedir (Vapnik 1995).

$$f(\mathbf{x}) = \mathbf{w} \cdot \mathbf{x} + b \quad \mathbf{w} \in \mathbb{R}^D, b \in \mathbb{R} \quad (11)$$

Fonksiyonun düz olması $\mathbf{w}$ 'nin küçük olması anlamına gelmektedir. Bunu sağlamanın bir yolu $\|\mathbf{w}\|^2$ normunu minimize etmektir. Konveks optimizasyon problemi denklem 12 'de gösterilmiştir (Cortes & Vapnik 1995).

$$\min \frac{1}{2} \|\mathbf{w}\|^2 \quad y_i - (\mathbf{w} \cdot \mathbf{x}_i + b) \leq \varepsilon \quad (\mathbf{w} \cdot \mathbf{x}_i + b) - y_i \leq \varepsilon \quad (12)$$

Yukarıdaki ifadelerden görüldüğü üzere $\mathbf{w}$ ağırlıkları $\mathbf{x}_i$ örneklerinin doğrusal kombinasyonu şeklinde yazılabilmektedir. Böylece destek vektörler ile temsil edilen fonksiyonun kompleksliği girdi örneklerinin boyutundan bağımsız olacak ve sadece destek vektörlerin sayısına bağlı olacaktır. Şu noktaya da dikkat çekmek gerekir ki model $\mathbf{x}_i$ örneklerinin çarpımları şeklinde ifade edilebilmektedir.

5.4.2. Doğrusal olmayan destek vektör regresyon

Destek vektör regresyonun doğrusal olmayan duruma genelleştirilmesi için $\mathbf{x}_i$ örnekleri $\Phi: \mathbb{R}^D \mapsto \mathcal{H}$ fonksiyonu yardımı ile özellik uzayına haritalanmaktadır. Destek vektör algoritmaları sadece $\mathbf{x}_i$ örneklerinin iç çarpımlarına bağlı olduğundan Φ fonksiyonunu bilmemize gerek kalmadan $k(\mathbf{x}_i, \mathbf{x}_j) = \Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_j)$ fonksiyonunu hesaplamak yeterli olacaktır.

Destek vektör makineleri matematiksel olarak $k(\mathbf{x}_i, \mathbf{x}_j) = \Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_j)$ şeklinde ifade edilen bir kernel fonksiyonu yardımıyla doğrusal olmayan dönüşümler yapılabilmekte ve bu şekilde verilerin yüksek boyutta doğrusal olarak ayırılmasına imkân sağlamaktadır. Sonuç olarak, kernel fonksiyonu kullanarak doğrusal olarak ayıramayan iki sınıflı bir problemin çözümü ile ilgili karar kuralı denklem 13’de gösterilmiştir (Osuna vd., 1997).

$$f(x) = \text{sign} + \sum_i (a_i \phi(x) \cdot \phi(x_i) + b) \quad (13)$$

5.5. Lojistik Regresyon

Lineer regresyon iki değişken arasındaki ilişkiyi ölçmek için kullanılan makine öğrenmesi yöntemlerinden biridir (Seber ve Lee, 2012). Bu yöntemde iki değişken arasındaki bağlantıya en yakın doğrusal sonuç elde edilir. Örnek verilerinin mümkün olan en çoğunu kapsayacak şekilde geçecek bir doğru denklemi elde edilir. Bu denklem ile gelecekteki veriler öngörülür. Öznitelik sayısı birden fazla olursa çoklu doğrusal regresyon olarak isimlendirilir. Lojistik regresyon ise, çok değişkenli verileri

sınıflandırmak için kullanılır (Wright, 1995).

Lojistik Regresyon Analizinin temel amacı diğer regresyon yöntemlerinde olduğu gibi bağımsız değişkenler ile bağımlı değişken arasındaki nedensellik ilişkisini incelemektir. Lojistik regresyon ise, çok değişkenli verileri sınıflandırmak için kullanılır (Wright, 1995). Lojistik regresyon yönteminde bağımlı değişkenin sürekli olması gibi bir varsayım yoktur, özellikle bağımlı değişkenin iki veya daha çok kalitatif değer aldığı durumlarda kullanılır (Ulupınar, 2007). Lojistik regresyon, s, bağımsız x değişkeninin $-\infty$ ile $+\infty$ arasında değerler alabilen doğrusal işlevi olmak üzere denklem 14’de gösterilmiştir.

$$f(s) = \frac{e^s}{1 + e^s} = \frac{e^s}{1 + e^{-s}} \quad (14)$$

Lojistik regresyon analizi normal doğrusal regresyon analizlerinde farklı bir yöntemle yorumlanır. Lojistik regresyon yönteminde bağımlı değişken nitel(kategorik) olmak zorunda iken, bağımsız değişken(ler) her türlü ölçüm biçimine sahip olabilir. Yani lojistik regresyon analizinde bağımsız değişkenlerin ölçüm biçimi hakkında kısıtlayıcı durum yoktur (Sümbüloğlu vd., 2007).

Lojistik regresyon analizi, sınıflama ve atama işlemi yapmaya yardımcı olan bir regresyon yöntemidir. Normal dağılım varsayımı, süreklilik varsayımı önkoşulu yoktur. Bağımlı değişken üzerinde açıklayıcı değişkenlerin etkileri olasılık olarak elde edilerek, risk faktörlerinin olasılık olarak belirlenmesi sağlanır (Özdamar, 2002; Lemeshow ve Hosmer, 2000).

$$P = \frac{e^{0+X_1+\dots_k X_k}}{1 + e^{0+X_1+\dots_k X_k}} \quad (15)$$

Gösterilen Denklem 15’te (Özdamar, 2002);

P: İncelenen olayın gözlenme olasılığını,

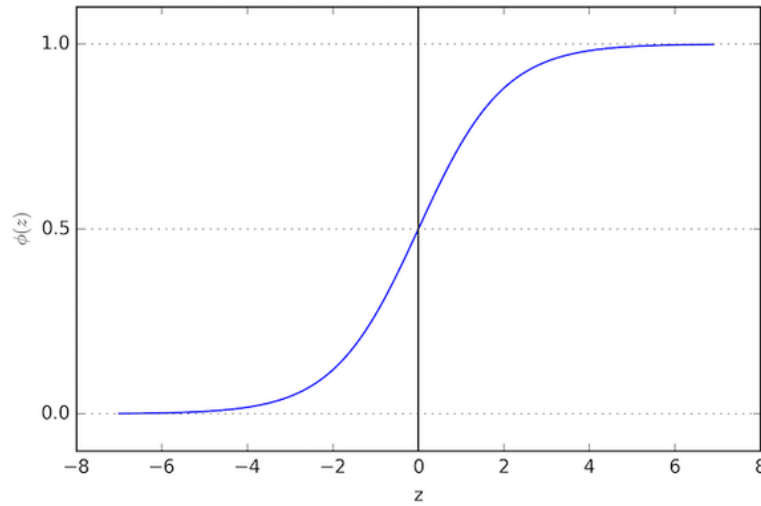
β_0 : Bağımsız değişkenler sıfır değerini aldığı anda bağımlı değişkenin değerini başka bir ifadeyle sabiti, $\beta_1 \beta_2 \dots \beta_k$: Bağımsız değişkenlerin regresyon katsayılarını,

$X_1 X_2 \dots X_k$: Bağımsız değişkenleri,

k: Bağımsız değişken sayısını,

e: 2.71 sayısını göstermektedir.

Lojistik modelde, bağımlı değişken kesikli değer alacağı için, bağımlı değişkeni açıklayan değişken (x) ile olan ilişkisi şekilde açıkça görülemez. Bu nedenle, yanıt değişkeni yerine $f(z)$ (π) olasılık değeri ile x arasında çizim yapılır. Bu çizimde $f(z)$ fonksiyonu sürekli olup 0'dan 1'e doğru sigmoid bir şekilde artarak devam eder ve şekil 13'te gösterilmiştir. Bu teoremden eğri biçimi S şeklinde olduğundan bu adı almıştır (Wright, 1995).

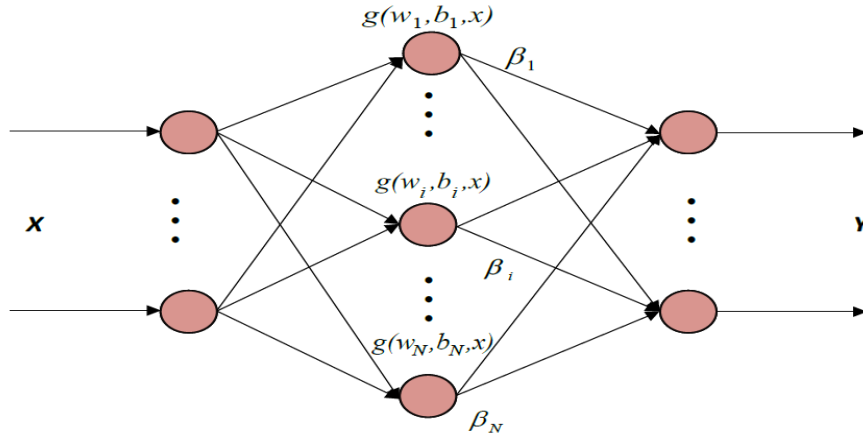


Şekil 13. S şeklindeki cevap fonksiyonu grafiği

5.6. Yapay Sinir Ağları (YSA)

YSA, insan sinir sistemini taklit ederek öğrenmeyi hedefleyen denetimli bir makine öğrenmesi yöntemidir. YSA modeli, bir katmanda bulunan nöronların takip eden katmandaki nöronlara bağlanması ile oluşturulur. En sık kullanılan YSA modeli olan çok katmanlı algılayıcı sinir ağı modeli; girdi katmanı, gizli katman ve çıktı katmanı olmak

üzere üç farklı katmandan meydana gelmektedir. Girdi katmanı verilerin okunduğu katmandır. Her bir nöron bir özelliği temsil ettiği için özellik sayısı kadar nöron içermektedir. Çıktı katmanı ise sınıfların belirlendiği katmandır. Bu katman oluşturulan modele göre tek bir nöron içerebileceği gibi sınıf çeşidi sayısı kadar nöron da içerebilmektedir (Er ve Çiğ, 2020). Nöronlardan oluşan, ilişkili nöron bağlantıları içeren ve matematiksel bir model olan YSA modeli şekil 14’de gösterilmiştir.



Şekil 14. Nöronlardan oluşan, ilişkili nöron bağlantıları içeren ve matematiksel bir model olan yapay sinir ağı modeli

Kaynak: (Er ve Çiğ, 2020)

Ağ oluşumunda kullanılan nöronlar arasında oluşturulan bağlantılar ağırlık denen sayısal değerler ile ilişkilendirilir. Her ağırlık, ağda transfer edilen ve sinyal ile çarpılan belirli bir değere sahiptir (Pappu ve Gummadi, 2017; Yang, Le Blond, Aggarwal, Wang ve Li, 2017). YSA’lar, problem hakkında bilgi sahibi olmaksızın girdi ve çıktı verileri arasındaki doğrusal olmayan ilişkileri tanımak için eğitilebilirler. Eğitimden sonra YSA’nın çalışma süresi son derece hızlıdır, çünkü yalnızca birkaç basit, birbirine bağlı işlem birimi içermektedir. YSA’ların model tanıma, genelleme ve enterpolasyon yapabilme özellikleri vardır. Bu nedenle eğitilen ağa bilinmeyen bir girdi uygulandığında uygun bir çıktı üretebilirler. Deterministik hata ve sistem modelleri yerine yalnızca stimüle edilmiş eğitim verileri gerektirirler (Niebur ve Germond, 1991).

5.6.1. Yapay Sinir Ağları Yapısı

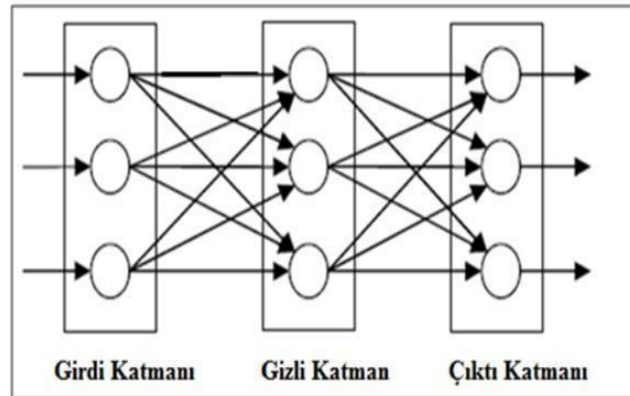
YSA girdi katmanı, gizli katman ve çıktı katmanı olmak üzere 3 katmandan oluşmaktadır (Nabiyev, 2016).

Girdi katmanı YSA'nın öğrenme sürecini gerçekleştirmesi için ağı sokulan girdilerin sisteme ilk giriş yaptığı katmandır. Yapay sinir ağı için iş akışının başlangıcıdır. Giriş katmanında hesaplama yapılmamakta sadece bilgiler gizli katmanlara aktarılmaktadır.

Gizli katman giriş katmanı ile çıktı katmanı arasında yer alan katmandır. Gizli katmanın, YSA'nın çevresi ile doğrudan bir bağlantısı yoktur. Gizli katmanda giriş katmanından alınan bilgiler işlenerek çıktı katmanına aktarılır. Gizli katmanlarda ileri ve geri yönlü işlemler gerçekleşebilir. Bir YSA'nın gizli katman sayısı ve her bir gizli katmanın sahip olduğu nöronların sayısı farklı durumlar için değişebilir. Gizli katmanların sayısı yanlış belirlendiği takdirde hesaplama süresinin artması, hesaplama hatalarının gerçekleşmesi gibi çeşitli olumsuz sonuçlar meydana gelebilmektedir.

Çıktı katmanı YSA'ya giriş yapan girdilerden sonuçlarının üretildiği katmandır. Çıktı katmanı hesaplamalardan ve ağıdan dış dünyaya doğru bilgi aktarımından sorumludur.

YSA'da bulunan tüm katmanlar Şekil 15'te gösterilmiştir.



Şekil 15. Yapay sinir ağları katmanları

Yapay nöronlarda ise ister çok katmanlı olsun ister basit yapıda olsun aşağıdaki bileşenler bulunmaktadır (Örkcü H. H. (2009) , Dede, G. (2008).

Girdi: Dışarıdan ağa girilen değerler ($x_1, x_2, \dots, x_m$) olarak ifade edilmektedir. Girdiler (x), m elemanlı sütun vektörü olarak denklem 16'da gösterilmiştir.

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_m \end{bmatrix} \quad (16)$$

Ağırlık: Dışarıdan ağa girilen değerlerin önemini ve hücredeki etkisini gösteren sayılar ağırlık ($w_1, w_2, \dots, w_m$) olarak ifade edilmektedir. Ağırlıklar, m elemanlı satır vektörü olarak düzenlenmiş hali denklem 17'de gösterilmiştir.

$$W = [w_1, w_2, \dots, w_m] \quad (17)$$

Toplama Fonksiyonu: Toplama fonksiyonu bileşeninde, yapay sinir hücresine giren tüm girdiler, ilgili ağırlıkları ile çarpılır. Bulunan tüm değerler toplanır ve böylece net girdi hesaplanır. Net girdi değerine eşik değeri (θ) eklenerek NET değer elde edilir. Toplam fonksiyonu sonucunda elde edilen NET değer aktivasyon fonksiyonuna iletilir.

Aktivasyon Fonksiyonu: Toplam fonksiyonundan iletilen NET değeri işleyerek yapay sinir hücresinin çıktısını oluşturan fonksiyondur.

Çıktı: Aktivasyon fonksiyonu sonucunda üretilen değerlere çıktı denir. Çıktı fonksiyonun belirlenmesi işlemleri denklem 18' gösterilmiştir.

$$NET = \sum_{i=1}^m w_i x_i + 0 \quad (18)$$

Yapay Sinir Ağlarının Avantajları ve Dezavantajları

YSA'nın avantajlarından bazıları aşağıda verilmiştir Vural B. B. (2007):

- Çeşitli örnekler üzerinden genelleme yapabilir.
- Algılamaya yönelik olaylarda Uzman Sistemlerle birlikte kullanılarak oldukça başarılı sonuçlar ortaya koymaktadır.
- YSA'lar yeni koşullara uyum sağlar ve buna ilişkin durumları öğrenebilmesi oldukça kolaydır. Her ne kadar bilgiler eksik de olsa kendi kendini eğitebilmesi sayesinde eksik

bilgiler ile performanslı bir şekilde yapay sinir ağıları çalışabilir.

YSA'nın dezavantajlarından bazıları aşağıda verilmiştir (Nabiyev, 2016):

- YSA'nın ağ yapısı genel olarak deneme yanılma yöntemiyle belirlendiğinden dolayı yerel çözümlere takılabilmektedir.
- Öğrenme katsayısı hücre sayılarının belirlenmesinde kesin bir kural olarak tanımlanmadığı için bazı durumlarda nöron sayısı tanıma işlemini ve eğitim hızını etkilemektedir.
- YSA'nın eğitilmesinde örnekler önem taşırlar ancak bir örnek seçim kuralı bulunmamaktadır.

Bilgiler ağa gömülü oldukları için yorumlanması ve çözümü elde etmek ve ağın takibini yapmak oldukça zordur.

5.7. Rasgele Orman (RO)

RO 2001 yılında Leo Breiman tarafından geliştirilmiştir. RO, torbalama yöntemi ve Tim Kam Ho tarafından önerilen rastgele alt uzay yöntemlerinin birleşiminden oluşmaktadır. Rastgele alt uzay yönteminde en uygun dallara bölünmeyi sağlayacak değişken tüm değişkenler arasından rastgele seçilmiş az sayıda değişken tarafından belirlenir (Uzbaşı, 2017). RO, ağaç tipi sınıflandırıcılar topluluğudur, ormandaki her ağaç bağımsız olarak örneklenen bir rasgele vektörün değerlerine ve aynı dağılıma bağlıdır (Breiman, 1999). RO karar ağaçları için tanımlanmış olmasına rağmen, bu yaklaşım tüm sınıflandırıcılar için geçerlidir. RO yönteminin önemli bir avantajı, çok sayıda girdi değişkenlerini ele almasıdır (Skurichina ve Duin, 2002).

RO yöntemi hem kategorik hem sürekli hem de her ikisinin yer aldığı veri setlerinde; aynı zamanda büyük veya küçük boyutlu veri setlerinde rahatlıkla kullanılabilir. Yöntemin dezavantajı olarak, sınıflandırma ağacı yönteminin aksine çıktı olarak bir ağaç vermemesidir (Akman vd., 2011). Bu şekilde rastgele tahminci seçmenin avantajı, topluluktaki ağaçlar arasında daha az korelasyon elde edildiği için oluşan modelin doğruluğu daha yüksektir (Suchetana vd., 2017). Bu yöntemde de sınıflandırma ve regresyon ağaçlarında olduğu gibi bölünme kriteri olarak daha önce gini indeksi

kullanılmaktadır. Gini indeksinin azalması istenen bir durumdur çünkü saflığın arttığına işaretler ve bu indeksin nihai olarak sıfıra eşit olması demek maksimum saflık demektir (Watts vd., 2011).

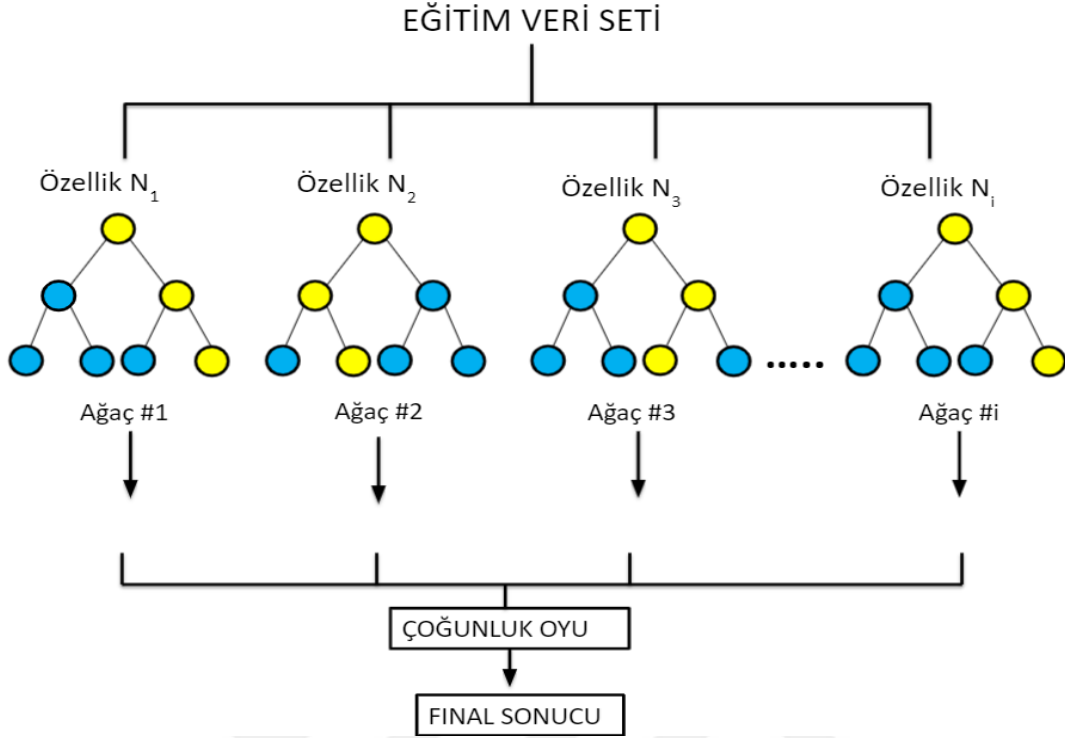
RO algoritması adımları aşağıdaki gibidir (Akman, Genç, ve Ankaralı, 2011):

- Orijinal veri setinden N adet örneklem seçilir. Seçilen örneklem belli bir kısmı oluşturulacak ağaçlar için öğrenme verisi kullanılır
- Budanmamış sınıflama ağacı aşağıdaki adımları takip ederek oluşturulur.
 - i) Her düğümde bütün öznelikler içerisinde en iyi özneliği seçmek yerine rastgele m tane öznelik seçilir ve bunların içerisinde en iyi dallara ayıracak (en çok bilgi kazancı sağlayacak) olanı belirlenir.
 - ii) Belirlenen tahmin değişkeni için en iyi dallanma kriteri entropi veya Gini indeksi ile hesaplanır ve hesaplanan değere göre veri setini her düğümde iki alt dala ayrılır.
 - iii) Yukarıdaki verilen işlemleri aşağıya doğru yaprak düğüm elde edilinceye kadar her düğümde tekrar edilir. Breiman (2001) tarafından varsayılan m değeri sınıflama ağaçları kurulurken ise $p^{1/2}$ olarak önerilmiştir. Burada p değeri toplam öznelik sayısını ifade etmektedir.
- N tane ağacın ayrı ayrı yapmış olduğu tahminleri bir araya getirerek yeni bir tahminde bulunur;

Sınıflama ağaçları için en çok oyu alan sınıf (majority voting) final sonucu olarak seçilir, yani her gözlem için ayrı ayrı verilen kararlar incelenir ve en çok hangi sınıf olarak tahmin edilmişse gözlemin sınıf kararı belirlenir. Örnek RO yapısı şekil 16'da gösterilmiştir.

Veri setinden hata oranını hesaplamak için ilk olarak her karar ağacı oluşturulurken, bootstrap örnekleme, ağaç oluşturulacak veri (in bag observation) ve ağaç oluşturmak için kullanılmayan veri (OOB verisi) olmak üzere ikiye ayrılır. OOB verisiyle ağacı test edilir ve hata oranı tahmini yapılır. Ardından bireysel ağaçların yaptığı OOB tahminlerini bir araya getirilir. Bu tahminlerden ormanın hata oranı kestirimi hesaplanır. Bu çalışmada

RO sınıflandırıcısının seçilmesinde, sınıflandırıcının kendi hata oranını hesaplayabiliyor olması önemli etkenlerden biridir.



Şekil 16. Rastgele orman sınıflandırma örneğine ait grafik

RO algoritmasını başlatmak için kullanıcı tarafından 2 parametre tanımlanmalıdır. Bu parametreler, en iyi bölünmeyi belirlemek için her bir düğümde kullanılan değişkenlerin sayısı (m) ve geliştirilecek ağaçların sayısı N 'dir. İlk olarak eğitim veri setinin 2/3'ünden önyükleme örnekleri oluşturulur. Out- of-bag (OOB) verisi olarak da adlandırılan, eğitim veri setinin 1/3'lük geri kalan kısmı hataları test etmek için kullanılır. Sonra her bir önyüklemeli örnekten budama olmadan ağaç geliştirilir. Her bir düğümde m değişkenleri tüm değişkenler arasından rastgele olarak seçilir ve bu değişkenler arasından en iyi dal belirlenir. Yeterli öngörü gücü ile yeterli miktarda düşük korelasyon sağlayan değişken sayısının seçimi son derece önemlidir (Horning 2010). Breiman'e (2002) göre toplam M adet değişken sayısının kare köküne eşit alınan m değişken sayısı genel olarak optimum sonuca en yakın sonucu verir. RO ağaç üretmek için CART (Classification and Regression Tree) algoritmasını kullanır (Beriman, 2001). Her bir düğüm de dallar CART algoritmasının kriterine (örn. GINI indeksi) göre oluşturulur. GINI indeksi sınıf homojenliğini ölçer ve denklemi 19' da gösterilmiştir.

$$\sum_{j \neq i} (f(C_i, T) / |T|) (f(C_j, T) / |T|) \quad (19)$$

Burada T eğitim veri setini, C_i rastgele seçilen bir pikselin ait olduğu sınıfı C_i ve $f(C_i, T) / |T|$ seçilen örneğin C_i sınıfına ait olma olasılığını göstermektedir (Pal, 2005). GINI indeksi büyüdükçe sınıf heterojenliği artarken, GINI indeksi azaldıkça sınıf homojenliği artar. Bir alt düğümün GINI indeksi bir üst düğümün GINI indeksinden daha az olduğunda o dal başarılıdır. GINI indeksi sıfıra ulaştınca yani her bir yaprak düğümde bir sınıf kaldığında ağaç dallanma işlemi sonlanır (Watts vd. 2011). Tüm N adet ağaç üretildiğinde N tane ağaçtan elde edilen tahmin sonuçlarına dayanılarak aday pikselin sınıfı belirlenir (Liaw & Wiener 2002).

5.8. Özellik Seçimi Yöntemleri

Özellik seçimi (*feature selection*), orijinal veri setini temsil edebilecek en iyi alt kümenin seçimi olarak tanımlanmaktadır. Özellik seçimi (diğer adıyla nitelik seçimi veya değişken seçimi), kullanılan algoritmaya göre özellikleri değerlendirerek veri setindeki n adet özellik arasından en iyi k adet özelliği seçme işlemidir (Forman, 2003).

Özellik seçimi, ilgilenilen problem için en faydalı ve en önemli özellikleri seçerek veri kümesindeki özellik sayısının azaltılmasını amaçlamaktadır. Özellik sayısının azaltılması analiz sürecinde uygulamacıya birçok avantaj sağlamaktadır.

Özellik seçme işleminin avantajları (Ladha ve Depha, 2011):

- Özellik kümesinin boyutunu düşürür ve algoritma hızını artırır,
- İlgili olmayan ve gürültülü veriyi ortadan kaldırır,
- Veri kalitesini geliştirir,
- Veri kümesini daha basit bir şekilde tanımlanabilir, görselleştirilebilir ve anlaşılabilir hale getirir,
- Veri kümesini oluşturmak için gerekli olan veri toplama işleminde kaynak tasarrufu sağlar,
- Veri depolamak için gerekli olan hafıza miktarını azaltır,
- Elde edilen modelin başarısını artırır.

5.8.1. Fisher Skor

İki sınıfı birbirinden ayırmaya yardımcı olacak özellikler bulunmaya çalışan bu yöntem ile özellik seçim işlemi, özelliklerin hesaplanan skorlara göre büyükten küçüğe doğru sıralanmasının ardından en üst sıradan başlanarak istenilen sayıda özelliğin seçilmesi şeklinde yapılmaktadır. Fisher skorunun yüksek olması, ilgili özelliğe ilişkin iki sınıf arasındaki ortalama farkın büyük olduğunu ayrıca ilgili sınıflardaki değerinde küçük sapmalarının olduğunu ifade etmektedir. Bu nedenle, iki sınıfı birbirinden en iyi ayıracak özellikleri seçmek için Fisher skoru yüksek olan özellikler tercih edilmektedir (Bolón vd., 2014).

5.8.2. Relief

Relief istatistiksel modelinin geliştirilmiş versiyonudur. Relief metodu, veri setinden bir örnek ele alarak ilgili örneğin, kendi sınıflarındaki diğer örneklerle yakınlığını ve farklı sınıflarla olan uzaklığına bağlı bir model oluşturarak öznelik seçme işlemini gerçekleştirmektedir (Bolón vd., 2014).

5.8.3. Ki-Kare Testi

Gözlenen ve beklenen frekanslar arasındaki farkın anlamlı olup olmadığı temeline dayanan Ki-Kare testi sık kullanılan özellik seçim yöntemlerinden biridir. Yöntem, özellikler (X) ile Y arasında ilişki olup olmadığını test edilmektedir. Yapılan test sonucunda, Y ile ilişkisi olmadığı tespit edilen özellikler veri setinden çıkartılır. Belirli bir sayıda özellik seçimi yapılmak istenildiğinde, özelliklerin hesaplanan χ^2 değerine göre büyükten küçüğe doğru sıralanmasının ardından en üst sıradan başlanarak istenilen sayıda özelliğin seçilmesi şeklinde işlem yapılır. Ki-kare değerinin hesaplanmasında kullanılan genel eşitlikler aşağıda verilmektedir (Kavzoğlu,2014).

5.8.4. Sınıflandırıcı ile Özellik Değerlendirici (Classifier Attribute Eval):

Bir özelliğin değerini, kullanıcı tarafından belirlenmiş bir sınıflandırma algoritması yardımıyla hesaplar.

5.8.5. Korelasyon ile Özellik Değerlendirici (*Correlation Attribute Eval*):

Bir özelliğin değerini, özellik ve sınıf arasındaki korelasyonu ölçerek hesaplar.

5.8.6. Filtreleme ile Özellik Değerlendirici (*Filtered Attribute Eval*):

Keyfi olarak seçilen ve özelliklerin sırasını değiştirmeyen bir filtreleme yöntemi ile düzenlenmiş veri için keyfi olarak seçilen bir özellik değerlendirici kullanır

5.8.7. Kazanım Oranı ile Özellik Değerlendirici (*Gain-Ratio Attribute Eval*):

Bir özelliğin değerini, özelliğin sınıfa göre kazanım oranını ölçerek hesaplar. Kazanım oranı simetrik olmayan bir ölçüt olup, bilgi kazancının çok çeşitli değerlere sahip özellikleri seçme eğiliminin önüne geçmek için kullanılmaktadır. Kazanım Oranı 0-1 aralığında değer almaktadır. Oran 1'e eşit olduğunda X bilgisinin tamamen Y bilgisini tahmin edebildiğini, 0'a eşit olduğunda ise Y ile X arasında hiçbir ilişki olmadığını gösterir. Kazanç Oranı denklem 20'deki gibi hesaplanır (Novakovic vd.,2011).

$$Kazanç\ Oranı = \frac{Bilgi\ Kazancı}{H(X)} \quad (20)$$

5.8.8. Bilgi Kazanım ile Özellik Değerlendirici (*Info-Gain Attribute Eval*):

Bir özelliğin değerini sınıfa göre bilgi kazanım oranını ölçerek hesaplar. Bilgi kazancı simetrik bir ölçüt olup ölçütte X gözlemlendikten sonra Y hakkında kazanılmış bilgi ile Y gözlemlendikten sonra X hakkında kazanılmış bilgi birbirine eşittir. Yöntemin zayıf yanı, daha fazla bilgiye sahip olmasa bile çok çeşitli değerlere sahip özellikler lehine önyargılı şekilde sonuç vermesidir (Tripoliti ve Manis, 2012).

5.8.9. OneR Özellik Değerlendirici (*OneR Attribute Eval*):

Bir özelliğin değerini OneR sınıflandırma algoritmasını kullanarak hesaplar. OneR sınıflandırma algoritması eğitim verisi üzerinde ve hata oranına göre özellikleri sıralar.

Tüm nümerik değerli özellikleri sürekli kabul eder ve onları birkaç ayırık aralığa böler (Tripoliti ve Manis, 2012).

5.8.10. Önem ile Özellik Değerlendirici (*Significance Attribute Eval*):

Olasılıksal önemi, özellik-sınıflar ve sınıflar-özellik ilişkisi şeklinde iki yönlü bir fonksiyon olarak hesaplayarak bir özelliğin değerini bulur.

5.8.11. SVM ile Özellik Değerlendirici (*SVM Attribute Eval*):

Bir özelliğin değerini, destek vektör makinesi (*Support Vector Machine; SVM*) sınıflandırma algoritmasını kullanarak bulur. Özellikler, SVM tarafından atanan ağırlığın karesi alınarak sıraya konulur.

5.8.12. Simetrik Belirsizlik ile Özellik Değerlendirici (*Symmetrical Uncert Attribute Eval*):

Bir özelliğin değerini özelliğin sınıfa göre simetrik belirsizliğini ölçerek hesaplar. Bu ölçüye göre eğer bir özellik bir sınıf ile yüksek korelasyona sahip fakat veri setindeki diğer özellikler ile düşük korelasyona sahipse iyi bir özelliktir (Tripoliti ve Manis, 2012).

6. PERFORMANS ÖLÇÜTLERİ

K-katlamalı çapraz doğrulama yöntemi modelin değerlendirilmesi için önemli olup, modelin geliştirilme aşamasında aşırı öğrenmeyi ve eksik öğrenmeyi tespit etmekte (Singh, Tiwari ve Tentu 2018) ve modelin test edilme aşamasında en iyi modeli oluşturmayı hedeflemektedir (Arlot ve Celisse, 2010). Aşırı öğrenme, modelin eğitim kümesindeki örüntüler yerine gözlemleri öğrenmesi durumunda ortaya çıkmaktadır. Bu durumda oluşturulan model eğitim aşamasında kullanılan veri kümesini öğrenir, ancak yeni gelen gözlemler için başarılı bir tahmin yapamaz. Genelde aşırı öğrenme modelleri eğitim aşamasını küçük hata oranı ile tamamlarken, test aşamasında büyük bir hata oranı ile tahmin etmektedir (Arlot ve Celisse, 2010). Eksik öğrenme ise, modelin gözlemlerdeki örüntüyü eksik bir şekilde öğrenmesi durumunda ortaya çıkmaktadır. Sınıflandırma modellerinde aşırı öğrenmeyi ve eksik öğrenmeyi önlemek için k-katlamalı çapraz doğrulama yöntemi kullanılmaktadır. K-katlamalı çapraz doğrulama yönteminde, ilk olarak eğitim sürecinde kullanılacak eğitim kümesi karıştırılır ve eşit büyüklükteki k adet alt kümelere bölünür. Bu işlemler k-kez tekrarlanarak her iterasyonda sıradaki alt küme eğitim veri kümesinden çıkarılır ve test kümesi olarak kullanılır. Tüm parçalar için değerlendirme süreci tamamlandığında, çapraz doğrulama modeli tüm veriler için bir performans ölçütü ve sonuçlar üretmektedir (Wiens, Dale, Boyce, ve Kershaw, 2008). Sınıflandırma modelinin başarısı doğru sınıfa atanan örnek sayısı ve yanlış sınıfa atanan örnek sayısı karşılaştırılarak belirlenmektedir. Model başarısını değerlendirmek için kullanılan temel kavramlar doğruluk, özgüllük, hassasiyet ve f-score ölçütüdür. Bu ölçütleri hesaplayabilmek için karışıklık matrisinden (confusion matrix) yararlanılır. Karışıklık matrisinde satırlar modelin öngördüğü sınıf sayılarını, sütunlar ise test kümesindeki gerçek sınıf sayılarını ifade etmektedir.

Önerilen yöntemin başarımlar kriterleri doğruluk, özgüllük, hassasiyet ve F-skoru oranlarına dayalı olarak yapılmıştır. Sınıflandırma problemlerindeki değerlendirme ölçütleri, karışıklık matrisi adı verilen her sınıf için doğru ve yanlış olarak sınıflandırılmış örnek sayılarının bulunduğu bir matris kullanılarak yapılır. Yanlış pozitifler, yanlış negatifler, doğru pozitifler ve doğru negatifler kavramları aşağıda belirtildiği gibi tanımlanır ve denklem 21’de gösterilmiştir.

- Yanlış Pozitifler: Negatif sınıftan olan, pozitif olarak tahmin edilen örnekler.
- Yanlış Negatifler: Gerçek sınıfı pozitif olan negatif olarak tahmin edilen örnekler.
- Doğru Pozitifler: Pozitif sınıfa ait doğru tahmin edilen örnekler.
- Doğru Negatifler: Negatif sınıfa ait olarak doğru tahmin edilen örnekler.

$$\text{Doğruluk} = \frac{\text{Doğru Negatif} + \text{Doğru Pozitif}}{\text{Doğru Negatif} + \text{Doğru Pozitif} + \text{Yanlış Negatif} + \text{Yanlış Pozitif}} \quad (21)$$

Hassasiyet ölçümü, ikili problemlerde her bir sınıf için sınıflandırıcının etkinliğini değerlendirir. Gerçek pozitif oran olarak bilinen hassasiyet, pozitif sınıfa ait tahmin edilen verilerin gerçek pozitif verilerine oranıdır. Hassasiyet ölçümü denklem 22’de verilmiştir.

$$\text{Hassasiyet} = \frac{\text{Doğru Pozitif}}{\text{Doğru Pozitif} + \text{Yanlış Negatif}} \quad (22)$$

Özgüllük, pozitif bir tahminin doğru olma olasılığını tahmin eden bir ölçüdür. Özgüllük ölçümü denklem 23’de verilmiştir.

$$\text{Özgüllük} = \frac{\text{Doğru Negatif}}{\text{Doğru Negatif} + \text{Yanlış Pozitif}} \quad (23)$$

F-skoru, pozitif kestirim oranı ve hassasiyet ölçülerinin uyumlu bir ortalaması olup denklem 24’de gösterildiği gibi hesaplanır.

$$F - \text{skor} = \frac{\text{Doğru Pozitif} + \text{Doğru Pozitif}}{\text{Doğru Pozitif} + \text{Doğru Pozitif} + \text{Yanlış Negatif} + \text{Yanlış Pozitif}} \quad (24)$$

7. DENEYSEL SONUÇLAR

Bu çalışmada 2020 yılında Necmettin Erbakan Üniversitesi Meram Tıp Fakültesi Çocuk Nörolojisi Polikliniğinde izlenen epilepsili hastaların verileri toplanmıştır. Veriler üzerinde makine öğrenmesi algoritmalarından olan NB, YSA, KA, KYK, lojistik regresyon, RO ve DVM kullanılarak sınıflandırma işlemi gerçekleştirilmiştir. Deneyler, Intel i7 2.6 GHz işlemci, 16 GB RAM özelliklerine sahip Windows 10 64 bit işletim sistemi kurulu bilgisayarda yapılmıştır.

Veri kümesindeki verilerin yaklaşık olarak yüzde dördü kayıp veri içermektedir. Bu kayıp veriler, veri ön işleme tekniklerinden ortalama veri ile doldurma yöntemi ile veri kümesi güncellenmiştir. Veri setindeki özelliklerde ayrıştırma işlemi yaparak kategorik hale dönüştürülmüştür. Çalışmada kullanılan veri setinin değişkenleri tablo 5'te gösterilmiştir.

Tablo 5. Çalışmada kullanılan değişkenler

Değişken adı	Açıklama ve değerler	Türü
Boy (cm)	15-186	(0, 1)
Yaş (yıl)	0-18	(0, 1)
Ağırlık (kg)	0-92	(0, 1)
İlk nöbet yaşı	0-16	(0, 1)
Baş çevresi (cm)	30-72	(0, 1)
Mental retardasyon	(0) yok (1) var	(0, 1)
İzlem süresi (yıl)	0.1-15	(0, 1)
Hastalıkta kullanılan ilaç sayısı	1-6	(0, 1)
Özgeçmiş	(0) anormal (1) normal	(0, 1)
Cinsiyet	(0) kız (1) erkek	(0, 1)
Akraba evliliği	(0) yok (1) var	(0, 1)
Prematurite	(0) yok (1) var	(0, 1)
Kardeş sayısı (hasta dahil)	1-8	(0, 1)
Doğumda hipoksi öyküsü	(0) yok (1) var	(0, 1)
Ek hastalık	(0) yok (1) var	(0, 1)
İlk EEG	(0) anormal (1) normal	(0, 1)
Ailede epilepsi öyküsü	(0) yok (1) var	(0, 1)
Ailede febril konvulsiyon öyküsü	(0) yok (1) var	(0, 1)
Ateşli nöbet	(0) yok (1) var	(0, 1)
Epilepsi tipi	(0) fokal (1) jeneralize	(0, 1)
Fizik muayene	(0) anormal (1) normal	(0, 1)
Manyetik rezonans görüntüleme	(0) anormal (1) normal	(0, 1)
Motor retardasyon	(0) yok (1) var	(0, 1)
Dirençli epilepsi	(0) yok (1) var	(0, 1)

Çalışmada 500 epilepsili hastanın 84'ünde dirençli epilepsi bulunmaktadır. Hastalar ilaç kullanan ve kullanmayan olarak iki kısma ayrılmıştır. İlaç kullanan 449 epilepsi hastasının 82'sinde (%18), ilaç kullanmayan 47 hastanın ikisinde (%4) dirençli epilepsi bulunmaktadır.

Hastaların 132'sinde (%26) anne-baba arasında akraba evliliği bulunmaktadır. Akraba evliliği olan bu hastaların 24'ünde (%4) dirençli epilepsi tanısı konulmuştur.

Hastaların 42'inde (%8) febril konvülsiyon öyküsü bulunmaktadır. Febril konvülsiyon öyküsü olan bu hastaların 4'ünde (%9) dirençli epilepsi bulunmaktadır.

Ailede epilepsi öyküsü olan 165 (%33) hastanın 31'inde (%18) dirençli epilepsi bulunmaktadır.

Hastaların 317'inde (%63) ilk EEG anormaldir. İlk EEG'si anormal olan bu hastaların 64'ünde (%20) dirençli epilepsi bulunmaktadır.

Toplam 500 epilepsili hastanın 357'sinde (%71) mental retardasyon bulunmaktadır. Mental retardasyonlu hastaların 45'ine (%13) dirençli epilepsi tanısı konulmuştur.

Veri setindeki Boy, Yaş, Ağırlık, İlk nöbet yaşı, Baş çevresi, Mental retardasyon, İzlem süresi özelliklerini kategorikleştirmek için küçükten büyüğe doğru sıralanıp ortanca değerler aşağısındaki değerler için 0, yukarısındaki değerler için 1 girilmiştir.

Veri seti ayrımında rastgelelik durumundan oluşan overfitting probleminin önüne geçebilmek için her bir algoritmada ve her bir parametresinde gerçek sonuçları görebilmek için K katmanlı çapraz doğrulama yöntemi kullanılmıştır ve k değeri 10 olarak seçilmiştir. Elde edilen k değerlerinin ortalaması alınmış ve standart sapma oranları hesaplanmıştır.

KYK algoritmasında Manhattan, Jaccard, Hamming, Cosine, Minkowski ve Euclidean parametreleri içerisinde Minkowski ve Euclidean parametreleri kullanılmış olup tablo 6'da ve tablo 7'de doğruluk oranları gösterilmiştir. Komşu değerleri k 5'e kadar test edilmiştir ve her bir k değerindeki minkowski ve euclidean değerleri eşit ve en iyi değer $k=2$ bulunmuştur.

Tablo 6. **Minkowski değerleri**

k değeri	Minkowski
k=1	0.726
k=2	0.824
k=3	0.79
k=4	0.826
k=5	0.806

Tablo 7. **Euclidean değerleri**

k değeri	Euclidean
k=1	0.726
k=2	0.824
k=3	0.79
k=4	0.826
k=5	0.806

DVM’de ise poly ve sigmoid parametreleri denenmiştir ve en iyi değer sigmoid parametresinde alınmıştır. Elde edilen doğruluk oranları tablo 8’de gösterilmiştir.

Tablo 8. **Destek vektör makineleri değerleri**

Destek vektör makineleri	Doğruluk Oranı
Poly	0.832
Sigmoid	0.834

KA algoritması için ise Gini ve entropy parametreleri denenmiştir ve en iyi değer Gini parametresinde alınmıştır. Elde edilen doğruluk oranları tablo 9’da gösterilmiştir.

Tablo 9. **Karar ağacı sınıflandırıcısı değerleri**

Karar ağacı sınıflandırıcısı	Doğruluk Oranı
Gini	0.75
Entropy	0.718

RO algoritmasında ise ağaç sayısı 20’den 120’ye kadar denenmiş olup her birinde ayrı ayrı entropy ve Gini parametreleri uygulanmıştır. En iyi değer entropy parametresinin 100 ağaç sayısında gözlemlenmektedir. Elde edilen sonuçların doğruluk oranları tablo 10 ve 11’de gösterilmiştir.

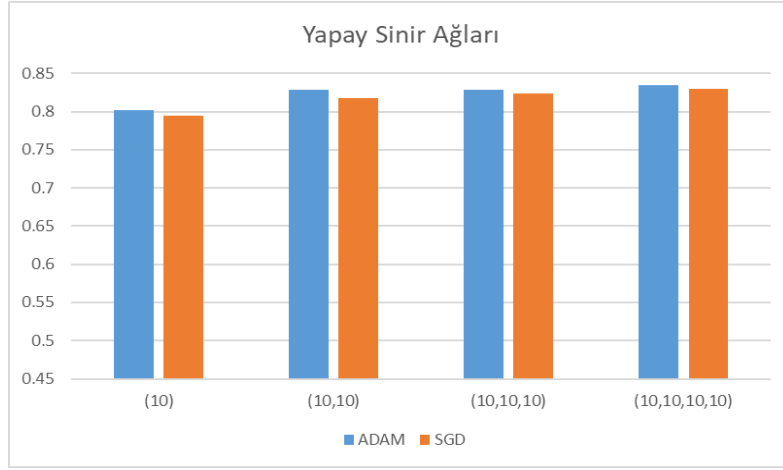
Tablo 10. Rastgele orman gini parametresi değerleri

k değeri	Gini
k=20	0.84
k=40	0.834
k=60	0.838
k=80	0.834
k=100	0.836
k=120	0.838

Tablo 11. Rastgele orman entropy parametresi değerleri

k değeri	Entropy
k=20	0.838
k=40	0.836
k=60	0.842
k=80	0.844
k=100	0.846
k=120	0.844

YSA’ında ise gizli katman sayısına ve solver parametresine göre incelenmiştir. Gizli katman sayısı (10), (10-10), (10-10-10), (10-10-10-10) solver parametresinde ise ADAM ile SGD parametreleri uygulanmıştır. En iyi sonuç 0.832 doğruluk oranıyla adam parametresinde ve gizli katman sayısı (10-10-10-10) ‘da gözlemlenmektedir. Elde edilen diğer sonuçların doğruluk oranları Şekil 17’de gösterilmiştir.



Şekil 17. Doğruluk oranları kıyaslanması

Lojistik regresyonda ise multiclass parametresini ovr ve auto olarak iki farklı şekilde incelenmiştir. Doğruluk oranlarının birebir aynı çıktığı tablo 12’de gösterilmiştir.

Tablo 12. Lojistik regresyon değerleri

Multiclass	Doğruluk Oranı
Ovr	0.808
Auto	0.808

NB algoritmasında olasılığı 0 olan durumlarda Laplace doğrulama yöntemi kullanılmıştır. KYK algoritmasında k komşu değeri 4 seçilmiştir. DVM algoritmasında kullanılan Kernel fonksiyonlarından sigmoid kullanılmıştır. KA’da Gini ve entropi kullanılabilmesini mümkün kılan criterion yönteminde entropi kullanılmıştır. YSA için solver adam yöntemi kullanılmıştır. RO algoritmasında ise criterion yönteminde entropi parametresi kullanılmıştır.

Bağımsız değişkene etki eden bağımlı değişkeni bulabilmek için sınıflandırıcı ile özellik değerlendirici yöntemi kullanılmıştır. Kullanılan bu yöntemle, karmaşık semptomları olan dirençli epilepsi tanısı için başarılı bir tıbbi karar destek mekanizması haline geçebileceği düşünülmektedir. Bu yapay zekâ yönteminin kullanımı, daha düşük tanı maliyeti, daha kısa tanı süresi yöntemlere gerek olmaması gibi avantajlara sahiptir. İncelenen veri setinde ise kullanılan ilaç sayısı, akraba evliliği, ailede febril konvülsiyon

öyküsü, ailede epilepsi olanlarda, ilk EEG, mental retardasyon parametreleri en çok etkilenen parametreler arasındadır.

Veri setinde yapılan istatistiklerin sonucunda kullandığımız algoritmalar arasında en iyi değer RO algoritmasında ağaç sayısı 100' de ve Entropy parametresinde KYK algoritmasında Minkowski parametresinde k=4 değerinde, DVM algoritmasında Sigmoid parametresinde, KA'larında Gini parametresinde, YSA'nda solver adam yönteminde elde edilmiştir.

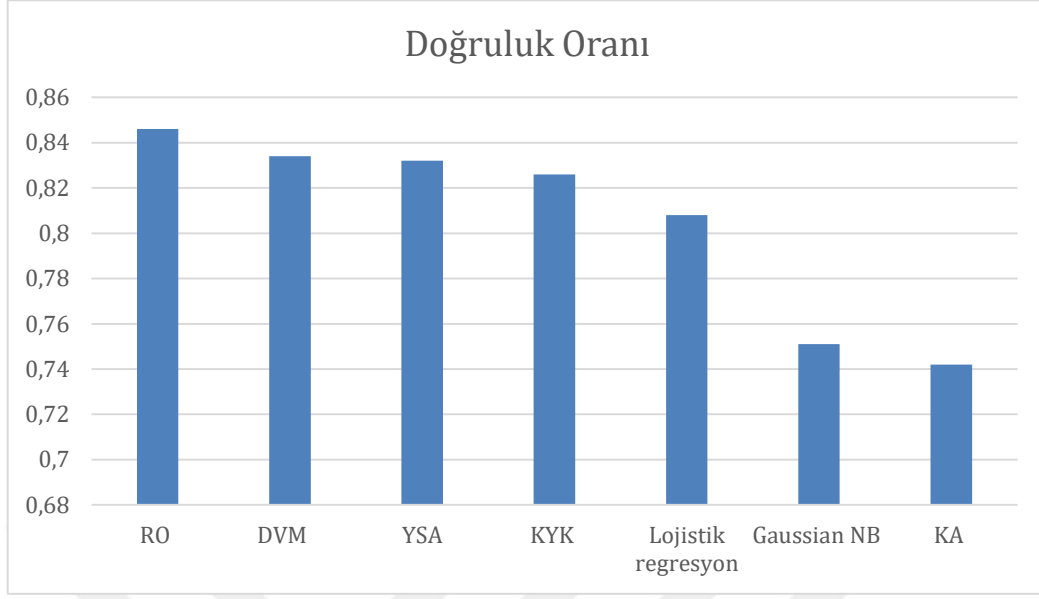
Deney sonuçlarının detaylı bir analizinin yapılabilmesi için kullanılan algoritmaların başarımları istatistiksel olarak karşılaştırılmıştır. Necmettin Erbakan Üniversitesi Meram Tıp Fakültesinde takip edilen epilepsili çocuklardan aldığımız veri kümesi için en başarılı sınıflandırma algoritmasını belirlemek için algoritmalar kıyaslanmıştır. Algoritmaları başarıma göre kıyaslarken veri kümesinde başarımları en yüksek olan algoritmadan en düşük olan algoritmaya kadar sırasıyla sahip olduğu değerleri verilmiştir. Çalışmamızda kullanılan algoritmaların doğruluk oranları kıyaslanması şekil 18'de görülmektedir.

Kullandığımız her bir algoritmanın kendi içerisinde farklı parametreler uygulanmış verinin eğitim ve test için farklı oranlarda bölünerek test verileri üzerinde doğruluk, özgüllük, hassasiyet ve f-skoru oranları hesaplanmış olup tablo 13'de gösterilmiştir.

Tablo 13. Doğruluk, özgüllük, hassasiyet ve f-skoru oranları

Model	Doğruluk	Özgüllük	Hassasiyet	F-skoru
Lojistik regresyon	0.808 ± 0.05	0.8123	0.9471	0.8346
KYK	0.826 ± 0.04	0.834	0.9221	0.8375
DVM	0.834 ± 0.01	0.834	0.95	0.8566
Gaussian NB	0.751 ± 0.09	0.8333	0.8318	0.7905
KA	0.742 ± 0.07	0.8423	0.8310	0.8326
YSA	0.832 ± 0.02	0.8285	0.9098	0.8498
RO	0.846 ± 0.02	0.861	0.98	0.8965

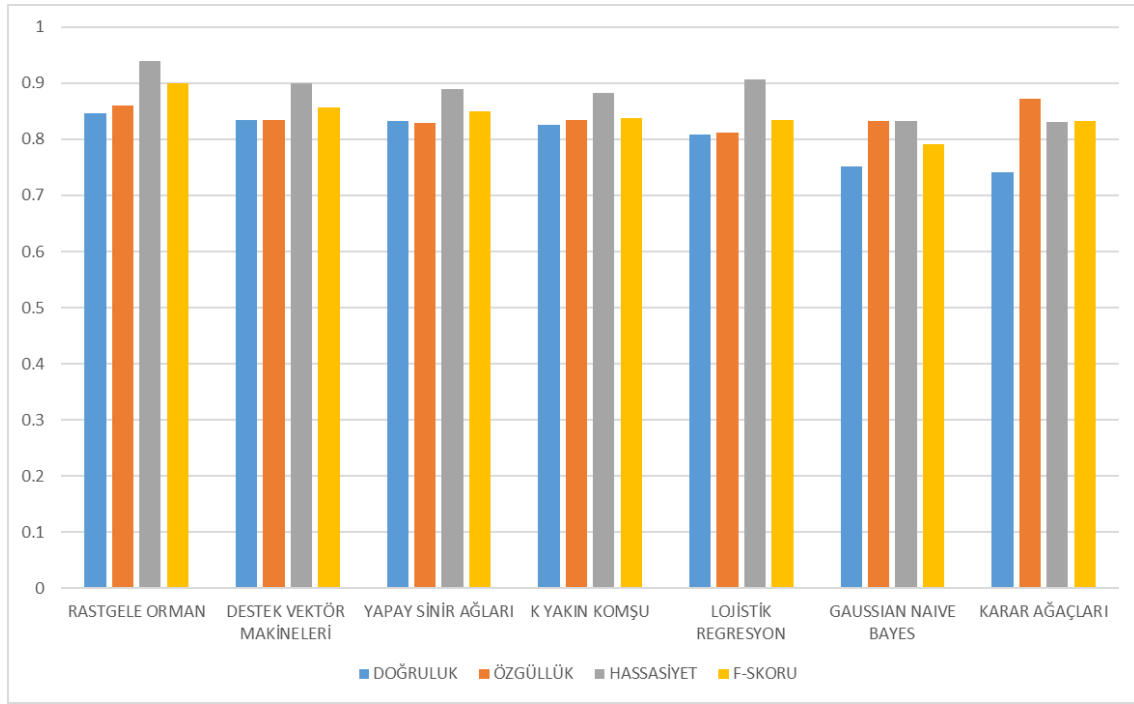
KYK, k-en yakın komşuluk; DVM, destek vektör makineleri; NB, naive Bayes; KA, karar ağacı; YSA, yapay sinir ağları; RO, rastgele orman.



Şekil 18. Doğruluk oranları kıyaslanması

Şekil 18’de görüldüğü gibi kullandığımız algoritmaları kendi aralarında doğruluk oranına göre karşılaştırsak eğer RO algoritması diğer algoritmalarından daha yüksek değerde olduğu ve sonrasında sırasıyla DVM, YSA, KYK, lojistik regresyon, Gaussian NB ve KA algoritmaları gelmektedir.

Çalışmamızda kullanılan algoritmaların doğruluk, özgüllük, hassasiyet ve f-skor değerleri şekil 19’da görülmektedir.



Şekil 19. Tüm değerlerin kıyaslamalı gösterimi

İstatiksel şekilde belirtilen Necmettin Erbakan Üniversitesi Meram Tıp Fakültesi takipli epilepsili çocuk hastaların demografik hasta öyküleri verileri, RO algoritması doğruluk derecesiyle en iyi sonucu ürettiği ve diğer algoritmalara göre tercih edilebileceği söylenebilir.

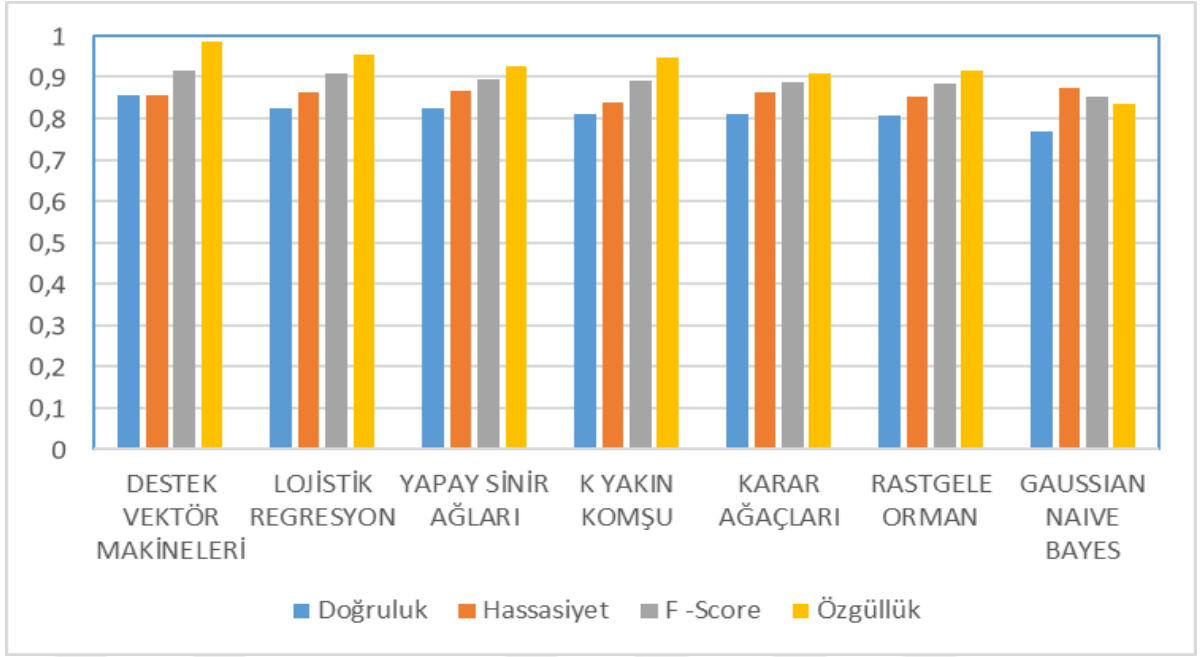
Bağımsız değişken olan dirençli epilepsiye etkisini görebilmek için, sınıflandırıcı özellik değerlendirilmesi özellik seçimi yönteminde öne çıkan (kullanılan ilaç sayısı, akraba evliliği, ailede febril konvülsiyon öyküsü, ailede epilepsi olanlarda, ilk EEG, mental retardasyon) özelliklerle yeni bir veri seti oluşturulmuştur. Bu veri setine uygulanan algoritmalar ve elde edilen oranları tablo 14'te gösterilmiştir.

Tablo 14. Özellik seçimi için oluşturulan veri setine uygulanan algoritmaların doğruluk, özgüllük, hassasiyet ve f-skoru oranları

Model	Doğruluk	Özgüllük	Hassasiyet	F-skoru
Lojistik regresyon	0.8242 $\pm$ 0.04	0.9549	0.8639	0.9071
KYK	0.8181 $\pm$ 0.05	0.9474	0.8400	0.8905
DVM	0.8545 $\pm$ 0.02	0.9850	0.8562	0.9161
Gaussian NB	0.7697 $\pm$ 0.07	0.8346	0.8740	0.8538
KA	0.8121 $\pm$ 0.05	0.9098	0.8643	0.8864
YSA	0.8260 $\pm$ 0.02	0.9248	0.8662	0.8945
RO	0.8061 $\pm$ 0.03	0.9173	0.8531	0.8841

KYK, k-en yakın komşuluk; DVM, destek vektör makineleri; NB, naive Bayes; KA, karar ağacı; YSA, yapay sinir ağları; RO, rastgele orman.

Elde edilen sonuçların doğruluk oranlarına göre kıyaslarsak DVM algoritması diğer algoritmalarından daha iyi sonuçlar vermiştir. Sonrasında sırasıyla YSA, lojistik regresyon, KYK, KA, RO ve Gaussian NB algoritmaları gelmektedir. Doğruluk oranlarıyla beraber modelin diğer performans ölçütleri şekil 20’de gösterilmiştir.



Şekil 20. Özellik seçimindeki verilere uygulanan algoritmaların kıyaslanması

Şekil 20’de görüldüğü gibi özellik seçimi için oluşturulan veri setine uyguladığımız tüm algoritmaların doğruluk, hassasiyet, özgüllük, F skor ölçütlerini incelediğimizde doğruluk oranında en iyi değer DVM algoritmasında, en kötü değer Gaussian NB algoritmasında, özgüllük ölçütüne göre incelediğimizde en iyi değer DVM algoritmasında en kötü değer Gaussian NB algoritmasında, hassasiyet ölçütüne göre incelediğimizde en iyi değer Gaussian NB algoritmasında en kötü değer KYK algoritmasında, f-skoru ölçütüne göre incelediğimizde en iyi değer DVM algoritmasında, en kötü değer Gaussian NB algoritmasında, olduğu gözlemlenmektedir. Bu duruma göre DVM algoritması hassasiyet ölçütü hariç diğer tüm ölçütlerde en iyi Gaussian NB algoritması hassasiyet ölçütü hariç diğer tüm ölçütlerde en kötü değer olarak gözlemlenmiştir.

8. SONUÇ VE ÖNERİLER

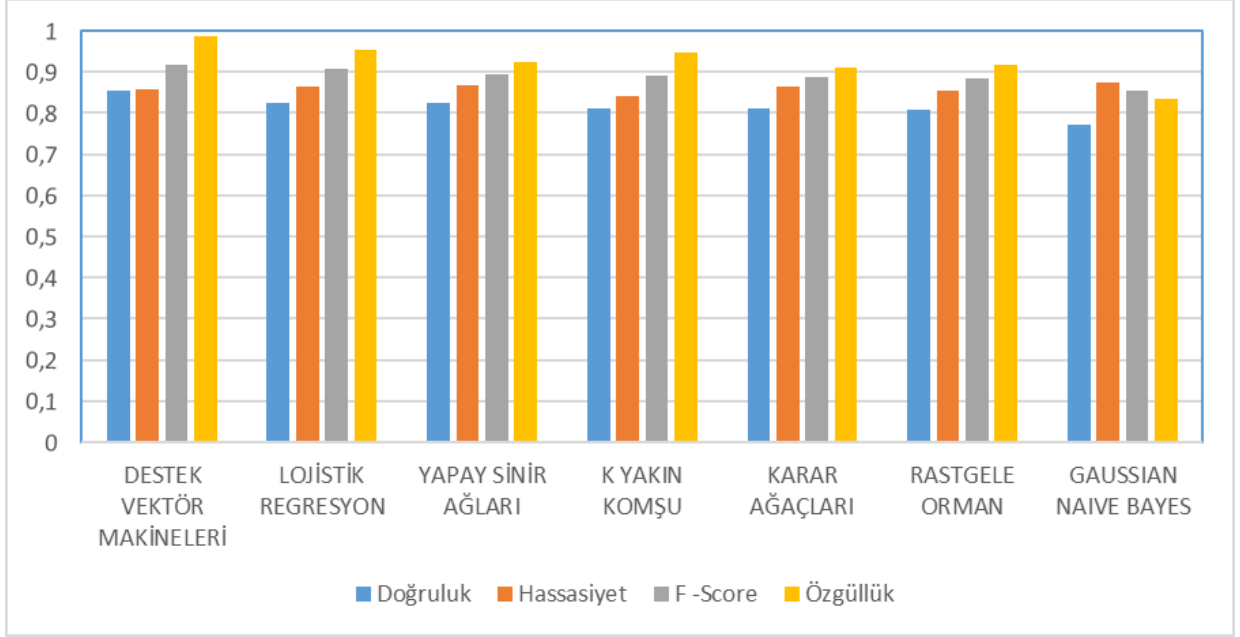
Hastanelerde doktorlar tarafından yapılan teşhis ve tedavilerde daha doğru sonuçlar elde etmek, doktorun teşhisine yardımcı olmak ve insan kaynaklı hataları engellemek amacıyla makine öğrenmesi tabanlı karar destek sistemleri kullanılmaktadır. Ayrıca çok fazla sayıda hastanın bilgilerini barındıran veri kümelerinin analizi, istatistiki bilgilerin çıkarılması ve bilimsel çalışmalarda kullanılabilmesi için hızlı bir şekilde yorumlanması gerekmektedir.

Bu çalışmada, Ocak 2011-2020 yılları arasında Necmettin Erbakan Üniversitesi Meram Tıp Fakültesi Çocuk Nörolojisi Polikliniğine başvuran 1 ay-18 yaş arasındaki 500 epilepsili hastanın dosya kayıtları retrospektif olarak taranmıştır. Hastaların cinsiyeti, yaşı, boyu, kilosu, ilk nöbet yaşı, baş çevreleri, mental retardasyonları, izlem süreleri, kullandıkları ilaç sayıları, hastalığa dair özgeçmişleri, hastanın ailesinde akraba evliliği var mı, hasta prematür mi, hastanın kendisi dahil kaç kardeşi olduğu, hastanın doğumunda hipoksi öyküsü bulunuyor mu, herhangi bir ek hastalığı var mı, ilk EEG'si normal mi, hastanın ailesinde epilepsi tanısı konulan birey var mı, ailede febril konvülsiyonlu birey var mı, hasta ateşli nöbet geçirdi mi, hastasının epilepsi tipi nasıl, fizik muayene kontrolü sonucu nasıl, MRG'si ve motor retardasyonun sonuçları değerlendirilmiştir. Epilepsili çocuk hastasının değerlendirilen sonuçlarına göre uzman kişi tarafından dirençli epilepsidir veya dirençli epilepsi değildir şeklinde tanısı konulmuştur. Tez çalışmasında kaydedilen hastaların verileri makine öğrenmesi teknikleri kullanılarak sınıflandırılmış ve doktorun kararına yakın sonuçlar bulunmaya çalışılmıştır.

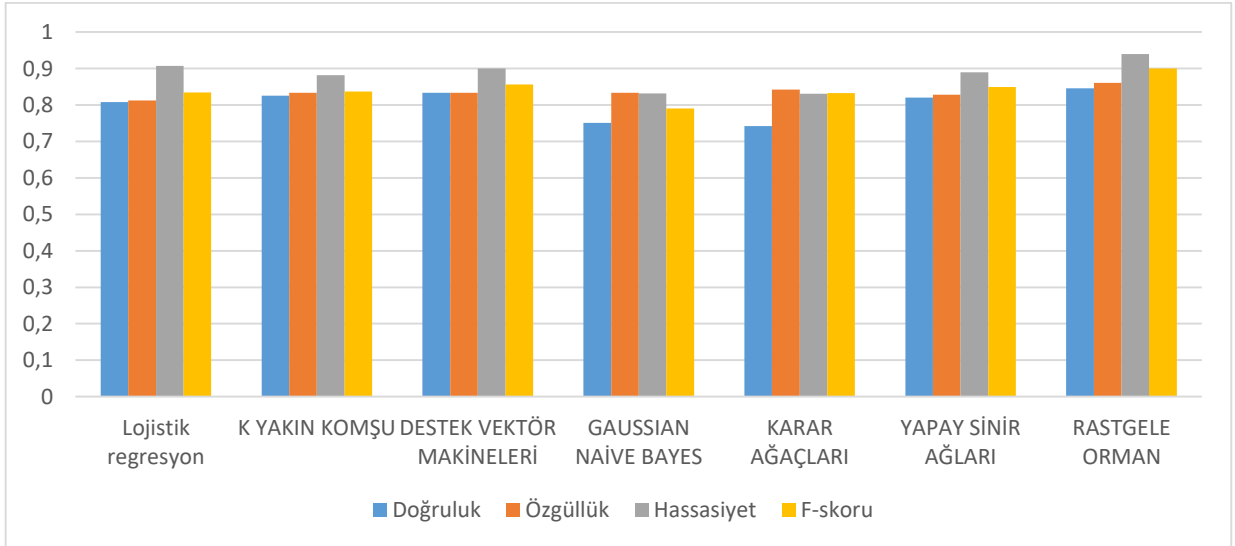
Necmettin Erbakan Üniversitesi Meram Tıp Fakültesi takipli epilepsili çocuk hastaların demografik hasta öyküleri verileri eksik ve fazla veriler temizlenmiş olup sonrasında analiz edilmiştir. Analiz sonucunda daha gerçekçi veriler elde etmek ve aşırı öğrenmenin önüne geçmek için k katlamalı çapraz doğrulama yöntemi kullanılmıştır.

Özellik seçimi aşamasında, filtreler, denetimli ve öznitelik klasörü altında bulunan ilgili ve oldukça kullanışlı performans ölçütlerini belirlemek için Weka'da öznitelik seçme özelliği kullanılmıştır (Frank vd., 2016). Bu tezde, en iyi performansı elde etmek için sınıflandırıcı özellik değerlendirmesi öznitelik seçim seçeneği kullanılmıştır. Bize

sağlamış olduğu belirlenmiş bir sınıflandırma algoritması seçme olanağı ile RO algoritması seçilmiş olup yöntem sonucundaki öne çıkan özellikler olan kullanılan ilaç sayısı, akraba evliliği, ailede febril konvülsiyon öyküsü, ailede epilepsi olanlarda, ilk EEG ve mental retardasyondur. Bu özellikler ayrı veri seti şeklinde toplanıp doğruluk, özgüllük, hassasiyet ve f-skoru oranları hesaplanmıştır ve elde edilen değerler şekil 21’de gösterilmiştir. Orijinal veri setinden elde edilen değerler ise şekil 22’de gösterilmiştir.



Şekil 21. Özellik seçimindeki verilere uygulanan algoritmaların kıyaslanması



Şekil 22. Orijinal veri kümesi üzerinde algoritmaların kıyaslanması

Şekil 21 ve 22’de görüldüğü gibi algoritmaların sonuçları birbiriyle karşılaştırılmak istenirse;

Gaussian NB algoritmasında; özellik seçiminden oluşturulan veri setindeki tüm oranlar orijinal veri setindeki orana göre daha yüksek olduğu,

KA algoritmasında; özellik seçiminden oluşturulan veri setindeki tüm oranlar orijinal veri setindeki orana göre daha yüksek olduğu,

Lojistik regresyon algoritmasında; özellik seçiminden oluşturulan veri setindeki hassasiyet oranı hariç diğer oranların orijinal veri setindeki orana göre daha yüksek olduğu,

YSA’da algoritmasında; özellik seçiminden oluşturulan veri setindeki hassasiyet oranı hariç diğer oranların orijinal veri setindeki orana göre daha yüksek olduğu,

KYK algoritmasında; özellik seçiminden oluşturulan veri setindeki doğruluk ve hassasiyet oranı hariç diğer oranların orijinal veri setindeki orana göre daha yüksek olduğu,

RO algoritmasında; özellik seçiminden oluşturulan veri setindeki yalnızca özgüllük oranı orijinal veri setindeki orana göre daha yüksek olduğu,

DVM algoritmasında; özellik seçiminden oluşturulan veri setindeki hassasiyet oranı hariç diğer oranların orijinal veri setindeki orana göre daha yüksek olduğu gözlemlenmiştir.

Buna göre özellik seçiminden alınan özelliklerin doğruluk, özgüllük, hassasiyet ve f-skoru oranlarının orijinal veri setinden daha yüksek olduğu gözlemlenmiştir.

Yapılan çalışmada elde edilen sonuçlara göre doğruluk ölçütü açısından RO algoritması %85 ile en yüksek performansa sahip algoritma olarak gözlemlenmektedir. RO algoritmasının en yakın takipçisi DVM algoritması doğruluk ölçütüne göre %84 ile ikinci sırada gelmektedir. Daha sonra sırasıyla YSA, KYK, Lojistik regresyon, Gaussian NB ve KA algoritmaları gelmektedir.

Tez çalışmasında, algoritmalar orijinal veri üzerinde ve özellik seçiminden sonra elde edilen veri kümesi üzerinde karşılaştırılmıştır. Her iki veri kümesi ile elde edilen doğruluk, özgüllük, hassasiyet ve f-skor oranları üzerinden algoritmalar karşılaştırılmıştır. İkinci veri kümesinden elde edilen en iyi doğruluk başarımlarının ilkinde göre daha yüksek olduğu gözlemlenmiştir. Kliniklerde dirençli epilepsi teşhisi için çok sayıda semptom ve laboratuvar bulgusuna ihtiyaç duyulmasına rağmen, DVM modelinin aynı teşhisi yüksek doğrulukla elde edebilmesi için özelliklerden sadece kullanılan ilaç sayısı, akraba evliliği, ailede febril konvülsiyon öyküsü, ailede epilepsi olma durumu, ilk EEG ve mental retardasyon gereklidir. Bu makine öğrenmesi yönteminin kullanımıyla, daha düşük tanı maliyeti ve daha kısa tanı süresi ile benzer sonuç elde edilecektir.

Bu çalışma ileride yapılacak olan çalışmalar için bir başlangıçtır. Bu olaya benzer olacak biçimde, daha büyük verilerde algoritmaların hızları ve hafızalarının tüketimi ile algoritmaların gerçekleştirdikleri performanslar arasındaki karşılaştırmalar da kıyaslanabilir. Daha farklı algoritmalar kullanılarak sınıflandırma başarımları ve çalışma süreleri göre kıyaslanabilir. Böylelikle de sınıflandırma başarımlarının yüksek veya çalışma sürelerinin az olduğu yöntemler tespit edilebilir.

Bu çalışmada dirençli epilepsinin otomatik teşhisi için önerilen yöntemlerin uzman bir doktor eşliğinde kullanılması düşünülebilir. Bu durum hem hasta olan kişiler hem sağlık sektörü çalışanları açısından düşük maliyet, daha çok fayda ve yüksek hız sağlayacaktır.

KAYNAKLAR

- Abbasi, B., & Goldenholz, D. M. (2019). Machine learning applications in epilepsy. *Epilepsia*, 60(10), 2037-2047.
- Adar, T., & Kılıç Delice, E. (2019b). Türkiye’de makine öğrenmesi ile ilgili yapılan tez çalışmalarına yönelik bir literatür taraması. *UEMK Proceedings Book* içinde (ss. 957-988). Gaziantep, Turkey: Gaziantep University.
- Adar, T., & Kılıç, Delice, E. (2019a). Sağlık alanında makine öğrenme algoritmalarının kullanılmasına yönelik bir literatür taraması. *UEMK Proceedings Book* içinde (ss. 928-956). Gaziantep, Turkey: Gaziantep University.
- Altman, N. S. (1991). An introduction to kernel and nearest-neighbor nonparametric regression. *The American Statistician*, 46, 175–185.
- An, S., Malhotra, K., Dilley, C., Han-Burgess, E., Valdez, J. N., Robertson, J., Clark, C., Westover, M. B., & Sun, J. (2018). Predicting drug-resistant epilepsy - A machine learning approach based on administrative claims data. *Epilepsy & Behavior: E&B*, 89, 118-125.
- Atalay, M., & Çelik, E. (2017). Büyük veri analizinde yapay zekâ ve makine öğrenmesi uygulamaları. *Mehmet Akif Ersoy Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 9(22), 155-172.
- Atul, H. (2020). *What is machine learning? Machine learning for beginners*. Erişim adresi: <https://www.edureka.co/blog/what-is-machine-learning/>. Erişim tarihi: 16 Ağustos 2021.
- Awad, M., & Khanna, R. (2015). *Efficient Learning Machines*. Apress, Open.
- Aydin, A., Ergor, A., Ergor, G., & Dirik, E. (2002). The prevalence of epilepsy amongst school children in Izmir, Turkey. *Seizure*, 11(6), 392-396.
- Aygün, B. (2010). *Erzurum il merkezinde 7-16 yaş okul çağı çocuklarında epilepsi prevalansının araştırılması*. Uzmanlık tezi, Atatürk Üniversitesi Tıp Fakültesi, Çocuk Sağlığı ve Hastalıkları Anabilim Dalı, Erzurum.
- Bacaksız, A. F. (2020). *Uyarlamalı yankılaşım teorisi ile epilepsi verilerinin öbeklenmesi ve bir epilepsi modeli*. Yüksek lisans tezi, İstanbul Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.
- Beghi, E. (2020). The epidemiology of epilepsy. *Neuroepidemiology*, 54(2), 185-191.
- Bennett, O. F., Kanber, B., Hoskote, C., Cardoso, M. J., Ourselin, S., Duncan, J. S., & Winston, G. P. (2019). Learning to see the invisible: A data-driven approach to finding the underlying patterns of abnormality in visually normal brain magnetic resonance images in patients with temporal lobe epilepsy. *Epilepsia*, 60(12), 2499-2507.
- Berg, A. T., Berkovic, S. F., Brodie, M. J., Buchhalter, J., Cross, J. H., van Emde Boas, W., Engel, J., French, J., Glauser, T. A., Mathern, G. W., Moshé, S. L., Nordli, D., Plouin, P., & Scheffer, I. E. (2010). Revised terminology and concepts for organization of

seizures and epilepsies: report of the ILAE Commission on Classification and Terminology, 2005-2009. *Epilepsia*, 51(4), 676-685.

Berrar, D. (2018). Bayes' theorem and naive bayes classifier. Reference Module in Life Sciences. İçinde. Amsterdam, Elsevier.

Bishop, C. M. (2006), Pattern recognition and machine learning. New York: Springer-Verlag.

Bolón-Canedo, V., Sánchez-Marroño, N., Alonso-Betanzos, A., Benítez, J. M., & Herrera, F. (2014). A review of microarray datasets and applied feature selection methods. *Information Sciences*, 282, 111-135.

Budak, C. (2014). *Medikal görüntülerdeki darbe gürültüsünün makine öğrenme teknikleri ile giderilmesi*. Doktora Tezi, Fırat Üniversitesi, Fen Bilimleri Enstitüsü. Elazığ.

Budak, C., Turk, M., & Toprak, A. (2016). Removal of impulse noise in digital images with naive Bayes classifier method. *Turkish Journal of Electrical Engineering & Computer Sciences*, 24(4), 2717-2729.

Canpolat, M., Kumandas, S., Poyrazoglu, H. G., Gumus, H., Elmali, F., & Per, H. (2014). Prevalence and risk factors of epilepsy among school children in Kayseri City Center, an urban area in Central Anatolia, Turkey. *Seizure*, 23(9), 708-716.

Chang, Y.-W., Hsieh, C.-J., Chang, K.-W., Ringgaard M., & Lin, C.-J. (2010). Training and testing low-degree polynomial data mappings via linear SVM. *Journal of Machine Learning Research*, 11, 1471-1490.

Chen, M., & Decary, M. (2020). Artificial intelligence in healthcare: An essential guide for health leaders. *Healthcare Management Forum*, 33(1), 10-18.

Cutler, A., Cutler, D. R., and Stevens, J. R. (2012). Random forests. C. Zhang & Y. Ma (Ed.), *Ensemble machine learning* içinde. Boston, MA: Springer.

Çalışkan, S. K., & Soğukpınar, İ. (2008). KxKNN: K-means ve k en yakın komşu yöntemleri ile ağlarda nüfuz tespiti. 2. Ağ ve Bilgi Güvenliği Sempozyumu içinde (ss. 120-124). Girne, Kıbrıs.

Dal, S. (2020). *Nörolojik krizlerden epilepsi nöbetinin kestirimi*. Yüksek lisans tezi, Batman Üniversitesi, Fen Bilimleri Enstitüsü, Batman.

Dash, S., Shakyawar, S.K., Sharma, M., & Kaushik, S. (2019). Big data in healthcare: management, analysis and future prospects. *Journal of Big Data* 6, 54.

Dede, G. (2008). *Yapay sinir ağları ile konuşma tanıma*. Yüksek lisans tezi, Ankara Üniversitesi, Fen Bilimleri Enstitüsü, Ankara.

Delen, D., Davazdahemami, B., Eryarsoy, E., Tomak, L., & Valluru, A. (2020). Using predictive analytics to identify drug-resistant epilepsy patients. *Health Informatics Journal*, 26(1), 449-460.

Deo, R. C. (2015). Machine learning in medicine. *Circulation*, 132(20), 1920-1930.

Diler, S. (2016). *Veri madenciliği süreçleri ve karar ağaçları algoritmaları ile bir uygulama*. Yüksek Lisans Tezi, Yüzüncü Yıl Üniversitesi, Fen Bilimleri Enstitüsü, Van.

- Doğan, M. I., Orman, A., Örkücü, M., & Örkücü, H. H. (2019). A new approach based on regression analysis and mathematical programming to multi-group classification problems. *Journal of the Faculty of Engineering and Architecture of Gazi University*, 34(4), 1939-1955.
- Domingos, P., & Pazzani, M. (1997). Beyond independence: Conditions for the optimality of the simple Bayesian classifier. *Machine Learning*, 29, 103-130.
- Er, M. B., & Çiğ, H. (2020). Türk müziği uyaranları kullanılarak insan duygularının makine öğrenmesi yöntemi ile tanınması. *Gazi Üniversitesi Fen Bilimleri Dergisi Part C: Tasarım ve Teknoloji*, 8(2), 458-474.
- Fine, A., & Wirrell, E. C. (2020). Seizures in children. *Pediatrics in Review*, 41(7), 321-347.
- Fisher, R. S., Acevedo, C., Arzimanoglou, A., Bogacz, A., Cross, J. H., Elger, C. E., Engel, J., Jr, Forsgren, L., French, J. A., Glynn, M., Hesdorffer, D. C., Lee, B. I., Mathern, G. W., Moshé, S. L., Perucca, E., Scheffer, I. E., Tomson, T., Watanabe, M., & Wiebe, S. (2014). ILAE official report: a practical clinical definition of epilepsy. *Epilepsia*, 55(4), 475-482.
- Fisher, R. S., Cross, J. H., D'Souza, C., French, J. A., Haut, S. R., Higurashi, N., Hirsch, E., Jansen, F. E., Lagae, L., Moshé, S. L., Peltola, J., Roulet Perez, E., Scheffer, I. E., Schulze-Bonhage, A., Somerville, E., Sperling, M., Yacubian, E. M., & Zuberi, S. M. (2017). Instruction manual for the ILAE 2017 operational classification of seizure types. *Epilepsia*, 58(4), 531-542.
- Forman, G. (2003). An extensive empirical study of feature selection metrics for text classification. *Journal of Machine Learning Research*, 3, 1289-1305.
- Gallo, C. (2015) Artificial neural networks tutorial. Encyclopedia of Information Science and Technology (3. Bs). İçinde. Hershey, IGI Global.
- Gentile, V. (2020). Statistical classification. From Codemotion. Erişim adresi: <https://www.codemotion.com/magazine/Glossary/statistical-classification/>. Erişim tarihi: 16 Ağustos 2021.
- Gujarati, D. N., & Porter, D. C. (çev. Şenesen, Ü., & Şenesen, G. G.). (2018.) Temel Ekonometri 5. Basımdan Çeviri (Ekonomik Baskı), Literatür Yayınları.
- Hamet, P., & Tremblay, J. (2017). Artificial intelligence in medicine. *Metabolism*, 69S, S36-S40.
- Han, J., & Kamber, M. (2006). Data mining: concepts and techniques (2. Bs). San Francisco, Morgan Kaufmann Publishers.
- Han, J., Kamber, M., & Pei, J. (2011). Data mining: concepts and techniques. (3. Bs). Amsterdam, Elsevier.
- Han, J., Pei, J., & Yin, Y. (2000). Mining frequent patterns without candidate generation. *ACM SIGMOD Record*, 29(2), 1-12.

- Hatipoğlu, E. (2018). Machine learning – Classification – Naive Bayes – Part 11. Erişim adresi: https://medium.com/@ekrem.hatipoglu/machine-learning-classification-_naive-bayes-part-11-4a10cd3452b4. Erişim tarihi: 7 Kasım 2021.
- Ho, T. K. (1998). The random subspace method for constructing decision forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20(8), 832-844.
- Horning, N. (2010). Random Forests : An algorithm for image classification and generation of continuous fields data sets. *International Conference on Geoinformatics for Spatial Infrastructure Development in Earth and Allied Sciences (GISIDEAS)* içinde (ss. 1-6). Hanoi, Vietnam.
- Hsu, C. W., Chang, C. C., & Lin, C. J. (2010). *A practical guide to support vector classification*. Department of computer science. National Taiwan University, Taipei 106, Taiwan.
- Hu, J., Niu, H., Carrasco, J., Lennox, B., & Arvin, F. (2020). Voronoi-based multi-robot autonomous exploration in unknown environments via deep reinforcement learning. *IEEE Transactions on Vehicular Technology*, 69(12), 14413-14423.
- Huseyinoglu, N., Ozben, S., Arhan, E., Palanci, Y., & Gunes, N. (2012). Prevalence and risk factors of epilepsy among school children in eastern Turkey. *Pediatric Neurology*, 47(1), 13-18.
- İmak, A. (2019). *Epileptik EEG sinyallerinin makine öğrenmesi yöntemleri ile sınıflandırılması*. Yüksek lisans tezi, Fırat Üniversitesi, Fen Bilimleri Enstitüsü, Elazığ.
- İpek, R., Kömür, M., & Okuyaz, Ç. (2020). Antiepileptik ilaçlar. S. Kumandaş vd. (Ed.), Çocuk nörolojisi, çocukluk çağı epilepsileri (1. Bs) içinde. Ankara, Türkiye Klinikleri.
- Jégou, H., Douze, M., & Schmid, C. (2011). Product quantization for nearest neighbor search. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(1), 117-128.
- Kamiński, B., Jakubczyk, M., & Szufel, P. (2017). A framework for sensitivity analysis of decision trees. *Central European Journal of Operations Research*, 26, 135-159.
- Kaya Özçora, G. D., Canpolat, M., & Kumandaş, S. (2020). Epilepsi tanı ve tedavisi. S. Kumandaş vd. (Ed.), Çocuk nörolojisi, çocukluk çağı epilepsileri (1. Bs) içinde. Ankara, Türkiye Klinikleri.
- Kaya, M. (2019). *Makine öğrenmesi yöntemleri kullanılarak EEG verisinden zihinsel dikkat durumu tespiti*. Yüksek lisans tezi, Mersin Üniversitesi, Fen Bilimleri Enstitüsü, Mersin.
- Köktürk, F. (2012). *K-en yakın komşuluk, yapay sinir ağları ve karar ağaçları yöntemlerinden sınıflandırma başarılarının karşılaştırılması*. Doktora Tezi, Bülent Ecevit Üniversitesi, Sağlık Bilimleri Enstitüsü, Zonguldak.
- Kulkarni, S., Seneviratne, N., Baig, M. S., & Khan, A. (2020). Artificial intelligence in medicine: Where are we now?. *Academic Radiology*, 27(1), 62–70.
- Kwan, P., Arzimanoglou, A., Berg, A. T., Brodie, M. J., Allen Hauser, W., Mathern, G., Moshé, S. L., Perucca, E., Wiebe, S., & French, J. (2010). Definition of drug resistant

epilepsy: consensus proposal by the ad hoc Task Force of the ILAE Commission on Therapeutic Strategies. *Epilepsia*, 51(6), 1069-1077.

Ladha, L., & Deepa, T. (2011). Feature selection methods and algorithms. *International Journal on Computer Science and Engineering*, 3(5), 1787-1797.

Lantz, B. (2019). *Machine learning with R: Expert techniques for predictive modeling*. (3rd Ed). Birmingham, Packt Publishing Ltd.

Lhatoo, S. D., Bernasconi, N., Blumcke, I., Braun, K., Buchhalter, J., Denaxas, S., Galanopoulou, A., Josephson, C., Kobow, K., Lowenstein, D., Ryvlin, P., Schulze-Bonhage, A., Sahoo, S. S., Thom, M., Thurman, D., Worrell, G., Zhang, G. Q., & Wiebe, S. (2020). Big data in epilepsy: Clinical and research considerations. Report from the Epilepsy Big Data Task Force of the International League Against Epilepsy. *Epilepsia*, 61(9), 1869-1883.

Li, C., Zhang, S., Zhang, H., Pang, L., Lam, K., Hui, C., & Zhang, S. (2012). Using the K-nearest neighbor algorithm for the classification of lymph node metastasis in gastric cancer. *Computational and Mathematical Methods in Medicine*, 2012, 876545.

Liaw, A., & Wiener, M. (2002). Classification and regression by random forest. *R News*, 2(3), 18-22.

Lu, W., Wang, W., Leung, A. Y. T., Lo, S. M., Yuen, R. K. K., Xu, Z., & Fan, H. (2002). Air pollutant parameter forecasting using support vector machines. *Proceedings of the 2002 International Joint Conference on Neural Networks. IJCNN'02 (Cat. No.02CH37290)* içinde (ss. 630–635). Honolulu, HI, USA.

Menon A. K. (2009). *Large-scale support vector machines: Algorithms and theory*. Research Exam, University of California, San Diego.

Mikati, M. A., & Obeid, M. M. (2020). Conditions that mimic seizures. R. M. Kliegman vd. (Ed.), *Nelson textbook of pediatrics* (21. Bs) içinde. Philadelphia, Elsevier.

Mikati, M. A., & Tchapyjnikov, D. (2020a). Seizures in childhood. R. M. Kliegman vd. (Ed.), *Nelson textbook of pediatrics* (21. Bs) içinde. Philadelphia, Elsevier.

Mikati, M. A., & Tchapyjnikov, D. (2020b). Treatment of seizures and epilepsy. R. M. Kliegman vd. (Ed.), *Nelson textbook of pediatrics* (21. Bs) içinde. Philadelphia, Elsevier.

Mishra, S. P., Sarkar, T., Taraphder, S., Datta, S., Swain, D. P., Saikhom, R., Panda, S., & Laishram, M. (2017). Multivariate statistical data analysis-principal component analysis (PCA). *International Journal of Livestock Research*, 7(5), 60–78.

Nabiyev, V. V. (2016). *Yapay zekâ*. Ankara, Seçkin Yayıncılık.

Noorbakhsh-Sabet, N., Zand, R., Zhang, Y., & Abedi, V. (2019). Artificial intelligence transforms the future of health care. *The American Journal of Medicine*, 132(7), 795–801.

Novakavic, J., & Strbac, P., & Bulatovic, D. (2011). Toward Optimal Feature Selection Using Ranking Methods and Classification Algorithms, *Yugoslav Journal of Operations Research*, 21(1), 119-135.

- Oğuzlar, A. (2004). CART analizi ile hanehalkı işgücü anketi sonuçlarının özetlenmesi. *Atatürk Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 18(3-4), 79-90.
- Okkan, U., Dalkılıç, H. Y. (2012). Radyal tabanlı yapay sinir ağları ile kemer barajı aylık akımlarının modellenmesi. *İMO Teknik Dergi*, 5957-5966.
- Örkcü H. H. (2009). *Ayırma analizine matematiksel programlama ve yapay sinir ağları yaklaşımları*. Doktora tezi, Gazi Üniversitesi, Fen Bilimleri Enstitüsü, Ankara.
- Özdamar, K. (2002). Paket programlar ile istatistiksel veri analizi (2. Bs). Eskişehir, Kaan Kitabevi.
- Özkan, Y. (2008). Veri madenciliği yöntemleri (1. Bs). İstanbul, Papatya Yayıncılık.
- Pal M. (2005), Random forest classifier for remote sensing classification. *International Journal of Remote Sensing*, 26(1), 217-222.
- Papoulis, A. (1984). Probability, random variables, and stochastic processes. New York, McGraw-Hill.
- Park, H. A. (2013). An introduction to logistic regression: from basic concepts to interpretation with particular attention to nursing domain. *Journal of Korean Academy of Nursing*, 43(2), 154–164.
- Quinlan, J. R. (1993). C4.5: Programs for machine learning. San Mateo, CA, Morgan Kaufmann.
- Rajeswari, K. (2015). Improved apriori algorithm – A comparative study using different objective measures. *(IJCSIT) International Journal of Computer Science and Information Technologies*, 6 (3), 3185-3191.
- Rubenstein, S., & Levy, A. (2019). Seizures in childhood: aetiology, diagnosis, treatment, and what the future may hold. *EMJ Neurology*, 7(1), 62-70.
- Saitoğlu, Y. S. (2015). *Sınıflama ve regresyon ağaçları*. Doktora Tezi, Marmara Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul
- Saleem A. M. (2017). *Epilepsy seizure detection in EEG signals using wavelet transforms and support vector machines*. Master Thesis. Firat University. The graduate school of natural and applied science, Elazığ.
- Scheffer, I. E., Berkovic, S., Capovilla, G., Connolly, M. B., French, J., Guilhoto, L., Hirsch, E., Jain, S., Mathern, G. W., Moshé, S. L., Nordli, D. R., Perucca, E., Tomson, T., Wiebe, S., Zhang, Y. H., & Zuberi, S. M. (2017). ILAE classification of the epilepsies: Position paper of the ILAE Commission for Classification and Terminology. *Epilepsia*, 58(4), 512-521.
- Schwalbe, N., & Wahl, B. (2020). Artificial intelligence and the future of global health. *Lancet*, 395(10236), 1579–1586.
- Seber, G. A., & Lee, A. J. (2012). Linear regression analysis (2. Bs). Hoboken, John Wiley & Sons.

- Serdaroğlu, A., Ozkan, S., Aydin, K., Gücüyener, K., Tezcan, S., & Aycan, S. (2004). Prevalence of epilepsy in Turkish children between the ages of 0 and 16 years. *Journal of Child Neurology*, 19(4), 271–274.
- Steinbach M, & Tan PN. (2009). kNN: k-Nearest Neighbors. X. Wu, V. Kumar. (Ed.), The top ten algorithms in data mining, içinde. Boca Raton, CRC Press.
- Sümbüloğlu, K. & Akdağ, B. (2007). Regresyon Yöntemleri ve Korelasyon Analizi, Hatiboğlu Basım ve Yayın, Ankara.
- Şentürk, T. (2019). *EEG sinyallerinin epilepsinin nöbet öncesi tespiti için analiz ve modellenmesi*. Yüksek lisans tezi, Erciyes Üniversitesi, Fen Bilimleri Enstitüsü, Kayseri.
- Tack, C. (2019). Artificial intelligence and machine learning applications in musculoskeletal physiotherapy. *Musculoskeletal Science & Practice*, 39, 164–169.
- Tekil Değer Ayırımı. Açıklandı. ICHI.PRO. (2020-2021). Erişim adresi: <https://ichi.pro/tr/tekil-deger-ayrisimi-aciklandi-256667345978106>. Erişim tarihi: 16 Ağustos 2021.
- Terzi, Ö., Küçüksille, E. U., Ergin G., & İlker, A. (2011). Veri madenciliği süreci kullanılarak güneş ışıması tahmini. *SDU International Technologic Science* 3, 29-37.
- Thrun, M. C. (2018). Projection-Based Clustering through Self-Organization and Swarm Intelligence. Marburg: Springer.
- Tripoliti, E., Manis, G. (2011). Feature selection in HRV analysis of young and elderly subjects. Á. Jobbágy (Ed.), 5th European conference of the international federation for medical and biological engineering. IFMBE proceedings, vol 37 içinde. Berlin, Heidelberg, Springer.
- Türk, Ö. (2019). *EEG işaretlerinden epilepsi türlerinin sınıflandırılmasında skalogram tabanlı derin öğrenme yaklaşımı*. Doktora tezi, Dicle Üniversitesi, Fen Bilimleri Enstitüsü, Diyarbakır.
- Ugoni, A., & Walker, B. F. (1995). The Chi square test: an introduction. *COMSIG Review*, 4(3), 61–64.
- Ünal, Ö., Haliloğlu, G., & Coşkun, T. (2012). Epileptik ensefalopatiler ve doğuştan metabolizma hastalıkları. *Çocuk Sağlığı ve Hastalıkları Dergisi*, 55(3), 141-188.
- Wang, J., Lin, Z. J., Liu, L., Xu, H. Q., Shi, Y. W., Yi, Y. H., He, N., & Liao, W. P. (2017). Epilepsy-associated genes. *Seizure*, 44, 11-20.
- Watts J. D., & Lawrence R. L. (2008). Merging random forest classification with an object-oriented approach for analysis of agricultural lands. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII. Part B7. Beijing.
- Wei, F., Yan, L. M., Su, T., He, N., Lin, Z. J., Wang, J., Shi, Y. W., Yi, Y. H., & Liao, W. P. (2017). Ion channel genes and epilepsy: functional alteration, pathogenic potential, and mechanism of epilepsy. *Neuroscience Bulletin*, 33(4), 455-477.

- Wibawa, A. P., Kurniawan, A. C., Murti, D. M. P., Adiperkasa, R. P., Putra, S. M., Kurniawan, S. A., & Nugraha, Y. R. (2019). Naïve Bayes classifier for journal quartile classification. *International Journal of Recent Contributions from Engineering, Science & IT (iJES)*, 7(2), 91-99.
- Witten, I. H., Frank E., & Hal M. A. (2011). Data mining. Practical machine learning tools and techniques (3. Bs). Amsterdam: Elsevier.
- Witten, I., Frank, E., Hall, M. A., & Pal, C. (2016). Data mining: Practical machine learning tools and techniques (4. Bs). İçinde. Amsterdam, Morgan Kaufmann Publishers Inc.
- Wright, R. E. (1995). Logistic regression. L. G. Grimm vd. (Ed.), Reading and understanding multivariate statistics içinde. American Psychological Association.
- Yalçın, N. (2012). *Sezgisel algoritma öğrenmeli yapay sinir ağları ile epilepsi hastalığının teşhisi*. Yüksek lisans tezi, Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, Konya.
- Yeşilkaya, B. (2020). *EEG sinyalleri ve makine öğrenme yöntemlerini kullanılarak duygusal durum kestirimi*. Yüksek lisans tezi, İzmir Katip Çelebi Üniversitesi, Fen Bilimleri Enstitüsü, İzmir.
- Yıldırım, M. (2018). *Farklı zaman ölçekli EEG işaretlerinden epilepsinin tespiti*. Yüksek lisans tezi, Dicle Üniversitesi, Fen Bilimleri Enstitüsü, Diyarbakır.
- Zhan, F., Zhu, X., Zhang, L., Wang, X., Wang, L., & Liu, C. (2019). Summary of association rules ESMA 2018. IOP Conference Series: Earth and Environmental Science 252, 032219.
- Zhang, S., Li, X., Zong, M., Zhu, X., & Wang, R. (2017). Learning k for kNN classification. *ACM Transactions on Intelligent Systems and Technology*, 8(3), 1–19.
- Zhang, Y., Ji, X., & Zhang, S. (2016). An approach to EEG-based emotion recognition using combined feature extraction method. *Neuroscience Letters*, 633, 152-157.
- Zhunis, A. (2019). *EEG işaretlerinde epilepsi nöbet tahmini ve tespiti*. Yüksek lisans tezi, Kocaeli Üniversitesi, Fen Bilimleri Enstitüsü, Kocaeli.

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı : Osman ÇAKSEN

EĞİTİM DURUMU

Lisans Öğrenimi : 2019, KTO Karatay Üniversitesi, Mühendislik Fakültesi,
Bilgisayar Mühendisliği

Yüksek Lisans Öğrenimi : 2022, KTO Karatay Üniversitesi, Elektrik- Elektronik
Mühendisliği Ana bilim Dalı, Elektrik ve Bilgisayar
Mühendisliği Tezli Yüksek Lisans Programı

Bildiği Yabancı Diller : İngilizce

Bilimsel Faaliyetleri : -

İŞ DENEYİMİ

Stajlar : 2017, Stajyer, GMG Bilişim 2018, Stajyer, Atiker Yazılım

Projeler : 2022, Fundus Kamerada CNN Kullanarak Papil Ödem
Tespiti, Necmettin Erbakan Üniversitesi

Çalıştığı Kurumlar : 2020, Bilgi İşlem Müdürü, Empero Endüstriyel
Mutfak Ekipmanları

Tarih: 26 Ocak 2022

ETİK KURUL

Evrak Tarih ve Sayısı: 28.04.2021-7863



T.C.
KTO KARATAY ÜNİVERSİTESİ REKTÖRLÜĞÜ
Tıp Fakültesi Dekanlığı
İlaç ve Tıbbi Cihaz Dışı Araştırmalar Etik Kurul Başkanlığı

Sayı : E-41901325-050.99-7863
Konu : Dr. Öğr. Üyesi Ali Öztürk (Etik
Kurul Kararı Hk.)

28.04.2021

Sayın Dr. Öğr. Üyesi Ali ÖZTÜRK
Öğretim Üyesi

27.04.2021 tarihli İlaç ve Tıbbi Cihaz Dışı Araştırmalar Etik Kurulu toplantısında başvurunuz değerlendirilmiş olup ilgili karar ekte sunulmuştur.

Bilgilerinizi rica ederim.

Prof. Dr. Taner ZİYLAN
İlaç ve Tıbbi Cihaz Dışı Araştırmalar Etik
Kurul Başkanı

Ek:Dr. Öğr. Üyesi Ali ÖZTÜRK (Etik Kurul Kararı Hk.)_1 (1 sayfa)

Bu belge, güvenli elektronik imza ile imzalanmıştır.

Belge Doğrulama Kodu :BSLUSD3TF

Belge Takip Adresi : <https://turkiye.gov.tr/kto-karatay-universitesi-oby>

Adres:Akabe Mahallesi Alaeddin Kapı Caddesi No:130 Karatay / Konya

Telefon:444 1251 Faks:0332 202 00 44

e-Posta:tipozelkalem@ktratay.edu.tr Web:www.karatay.edu.tr

Kep Adresi:ktokaratayuniversitesi@hs01.kep.tr

Bilgi için: Başra İRİŞ KAVUNER

Unvanı: Sekreter

Tel No: 444 1251 - 7279



T.C.
KTO KARATAY ÜNİVERSİTESİ TIP FAKÜLTESİ
İLAÇ VE TIBBİ CİHAZ DIŞI ARAŞTIRMALAR ETİK KURUL KARARI

Toplantı Sayısı: 4

Toplantı Tarihi: 27.04.2021

Karar Sayısı: 2021/007: Dr. Öğr. Üyesi Ali ÖZTÜRK'ün, “Makine Öğrenmesi Tekniğiyle Epilepsili Çocukların Klinik ve Laboratuvar Bulgularının İncelenmesi” başlıklı araştırma projesi çalışması ile ilgili 17.03.2021 tarihli dilekçesi ve ekleri görüşüldü.

Görüşme sonucunda araştırma projesi çalışmasının Dr. Öğr. Üyesi Ali ÖZTÜRK'ün, sorumluluğunda yürütülmesinin uygun olduğuna oy birliği ile karar verildi.

Not: Çalışma ile ilgili gerekli izin ve yasal sorumluluk araştırmacılara aittir.

Sorumlu Araştırmacı: Dr. Öğr. Üyesi Ali ÖZTÜRK

Yardımcı Araştırmacı: Prof. Dr. Hüseyin ÇAKSEN, Prof. Dr. Saim AÇIKGÖZOĞLU ve Osman ÇAKSEN

