



**T.C.
İSTANBUL ÜNİVERSİTESİ-CERRAHPAŞA
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ**



[DOKTORA TEZİ]

**[FARKLI VERİ TÜRLERİNİN BİRLEŞTİRİLMESİ ile EPİLEPSİ
HASTALIĞININ OTOMATİK TESPİTİ ve SINIFLANDIRILMASI]**

[Şengül BAYRAK]

**[DANIŞMAN
Doç. Dr. Eylem YÜCEL DEMİREL]**

**[II. DANIŞMAN
Doç. Dr. Hidayet TAKÇI]**

[Bilgisayar Mühendisliği Anabilim Dalı]

[Bilgisayar Mühendisliği Programı]

[İSTANBUL-2021]

Bu çalışma, [7.06.2021] tarihinde aşağıdaki jüri tarafından [Bilgisayar Mühendisliği Anabilim Dalı, Bilgisayar Mühendisliği Programı]nda [Doktora tezi] olarak kabul edilmiştir.

Tez Jürisi

[Doç. Dr.] [Eylem YÜCEL DEMİREL] (Danışman)
[İstanbul Üniversitesi – Cerrahpaşa
Mühendislik Fakültesi]

[Prof. Dr.] [Sabri ARIK]
[İstanbul Üniversitesi – Cerrahpaşa
Mühendislik Fakültesi]

[Prof. Dr.] [Ahmet SERTBAŞ]
[İstanbul Üniversitesi – Cerrahpaşa
Mühendislik Fakültesi]

[Doç. Dr.] [Neyir ÖZCAN SEMERCİ]
[Bursa Uludağ Üniversitesi
Mühendislik Fakültesi]

[Dr. Öğr. Üyesi] [Nilgün PALA AÇIKGÖZ]
[Biruni Üniversitesi
Tıp Fakültesi]



[20.04.2016 tarihli Resmi Gazete’de yayımlanan Lisansüstü Eğitim ve Öğretim Yönetmeliğinin 9/2 ve 22/2 maddeleri gereğince; Bu Lisansüstü teze, İstanbul Üniversitesi–Cerrahpaşa’nın abonesi olduğu intihal yazılım programı kullanılarak Lisansüstü Eğitim Enstitüsü’nün belirlemiş olduğu ölçütlere uygun rapor alınmıştır.]

ÖNSÖZ

Doktora eğitimimde bana inanarak yoluma ışık tutan ve bu tez çalışmasının gerçekleşmesindeki her aşamada yol göstericiliği, değerli bilgi ve kıymetli tecrübelerini benimle paylaşan ve bana farklı bakış açıları sunan tez danışmanım Sayın Doç. Dr. Eylem YÜCEL DEMİREL hocama sonsuz şükranlarımı arz ederim.

Yüksek lisans eğitimimden itibaren her zaman sonsuz desteğini gördüğüm ve derin bilgisi ile bu tezin oluşmasında ilham kaynağı olan ikinci tez danışmanım Sayın Doç. Dr. Hidayet TAKÇI hocama sonsuz şükranlarımı arz ederim.

Tez çalışmasının daha iyisi olabilmesi için kıymetli önerileri ve yorumlarıyla büyük katkı sağlamış olan Sayın Prof. Dr. Sabri ARIK hocama ve Sayın Dr. Öğr. Üyesi Nilgün PALA AÇIKGÖZ hocama minnettarım.

Tez çalışmam süresince tezime kıymetli katkı sağlayan Sayın Doç. Dr. Rüya ŞAMLI hocama teşekkürü borç bilirim.

Tezimde hasta bireylerin verilerinin toplanması sırasında sonsuz destek, yardımları ve sabrı için Avicenna Hastanesi Başhekim yardımcısı Sayın Dr. Lütfi HOCAOĞLU hocama teşekkürü borç bilirim.

Hayatımın her evresinde olduğu gibi bu tez çalışması sırasında da her zaman sonsuz sevgi ve her konuda desteklerini gördüğüm sevgili aileme, özellikle annem Kadife Bozyurt'a minnettarım.

Bu tez çalışmasını, hayat bulduğum biricik kıymetli kızım Ada'ya ithaf ediyorum.

Haziran 2021

Şengül BAYRAK

İÇİNDEKİLER

Sayfa No

ÖNSÖZ	iv
İÇİNDEKİLER.....	v
ŞEKİL LİSTESİ	ix
TABLO LİSTESİ	xiii
SİMGE VE KISALTMA LİSTESİ	xv
ÖZET	xviii
SUMMARY	xx
1. GİRİŞ.....	1
2. GENEL KISIMLAR	3
2.1. EPİLEPSİ HASTALIĞI.....	3
2.2. EPİLEPSİ HASTALIĞININ TANISINDA KULLANILAN TEKNİKLER	4
2.2.1. Klinik Muayene Bulguları	4
2.2.2. Laboratuvar Bulguları	5
2.2.3. Elektroensefalogram (EEG).....	5
2.2.4. Manyetik Rezonans Görüntüleme (MRG)	6
2.3. EPİLEPSİ HASTALIĞININ TANISINDA MAKİNE ÖĞRENMESİ YÖNTEMLERİ VE LİTERATÜR ÇALIŞMALARI	6
3. MALZEME VE YÖNTEM	12
3.1. VERİ ÜZERİNDE UYGULANAN İŞLEMLER	12
3.2. YÖNTEM.....	15
3.2.1. ID3 Yöntemi	15
3.2.2. Doğrusal Regresyon Yöntemi	16
3.2.3. Naive Bayes Yöntemi.....	17
3.2.4. DVM Yöntemi.....	18
3.2.5. k–NN Yöntemi	20
3.2.6. Yapay Sinir Ağları (YSA) Yöntemi	21

3.2.7. Uzun Kısa Süreli Bellek Ağları (Long Short Term Memory – LSTM) Yöntemi	26
3.3. MODEL DEĞERLENDİRME.....	29
3.3.1. Karışıklık Matrisi Oluşturma.....	29
3.3.2. İki Sınıflı Modelleme için Doğruluk Ölçütleri Belirleme	30
3.3.3. n -kat Çaprazlama Yöntemi ile Model Doğrulama.....	30
3.3.4. Alıcı İşlem Karakteristikleri (Receiver Operating Characteristic – ROC) Analizi.....	31
3.3.5. Kappa Katsayısı.....	31
3.4. ÇOK MODLU KARAR TABANLI VERİ FÜZYONU	32
3.4.1. Erken Füzyon.....	34
3.4.2. Geç Füzyon.....	35
3.4.3. Ara Füzyon	37
3.5. HASTA VERİLERİNİN TOPLANMASI	40
3.5.1. Seçilen Hasta Verilerinden İstatistiksel Özelliklerin Çıkarılması	41
3.5.1.1. Klinik Parametre Özellik Setinin Elde Edilmesi.....	42
3.5.1.2. Laboratuvar Bulguları Özellik Setinin Elde Edilmesi	43
3.5.1.3. EEG Özellik Setinin Elde Edilmesi.....	43
3.5.1.4. MRG Özellik Setinin Elde Edilmesi	49
3.6. GERÇEK EPİLEPSİ VERİ SETİNİN OLUŞTURULMASI	53
3.7. GEÇ FÜZYON TEKNİĞİNE ÖZELLİK SETLERİNİN UYGULANMASI	53
3.7.1. Algoritma 1'e Klinik Parametre Özellik Setinin Uygulanması.....	60
3.7.1.1. ID3 Yöntemine Klinik Parametre Özellik Setinin Uygulanması.....	61
3.7.1.2. Doğrusal Regresyon Yöntemine Klinik Parametre Özellik Setinin Uygulanması	64
3.7.1.3. Naive Bayes Yöntemine Klinik Parametre Özellik Setinin Uygulanması	66
3.7.1.4. DVM Yöntemine Klinik Parametre Özellik Setinin Uygulanması	67
3.7.1.5. k -NN Yöntemine Klinik Parametre Özellik Setinin Uygulanması.....	69
3.7.1.6. FFBBP Yöntemine Klinik Parametre Özellik Setinin Uygulanması	71
3.7.1.7. LSTM Yöntemine Klinik Parametre Özellik Setinin Uygulanması	73
3.7.2. Algoritma 1'e Laboratuvar Bulguları Özellik Setinin Uygulanması	78
3.7.3. Algoritma 1'e EEG Özellik Setinin Uygulanması	79
3.7.4. Algoritma 1'e MRG Özellik Setinin Uygulanması	80
3.7.5. Geç Füzyona Uygulanan Özellik Setlerinden Elde Edilen Sonuçların Çok Modlu Karar Tabanlı Veri Füzyonuna Uygulanması	82
4. BULGULAR	83

4.1. KLİNİK PARAMETRE ÖZELLİK SETİNE AİT SINIFLANDIRMA SONUÇLARI (MODEL 1)	84
4.1.1. Model için ID3 Yönteminin Sonuçları	84
4.1.2. Model 1 için Doğrusal Regresyon Yönteminin Sonuçları	85
4.1.3. Model 1 için Naive Bayes Yönteminin Sonuçları	88
4.1.4. Model 1 için DVM Yönteminin Sonuçları	88
4.1.5. Model 1 için k-NN Yönteminin Sonuçları	89
4.1.6. Model 1 için FFBP Yönteminin Sonuçları	90
4.2. LABORATUVAR BULGULARI ÖZELLİK SETİNE AİT SINIFLANDIRMA SONUÇLARI (MODEL 2)	94
4.2.1. Model 2 için ID3 Yönteminin Sonuçları	94
4.2.2. Model 2 için Doğrusal Regresyon Yönteminin Sonuçları	95
4.2.3. Model 2 için Naive Bayes Yönteminin Sonuçları	97
4.2.4. Model 2 için DVM Yönteminin Sonuçları	98
4.2.5. Model 2 için k-NN Yönteminin Sonuçları	98
4.2.6. Model 2 için FFBP Yönteminin Sonuçları	99
4.3. EEG ÖZELLİK SETİNE AİT SINIFLANDIRMA SONUÇLARI (MODEL 3)	103
4.3.1. Model 3 için ID3 Yönteminin Sonuçları	103
4.3.2. Model 3 için Doğrusal Regresyon Yönteminin Sonuçları	105
4.3.3. Model 3 için Naive Bayes Yönteminin Sonuçları	106
4.3.4. Model 3 için DVM Yönteminin Sonuçları	107
4.3.5. Model 3 için k-NN Yönteminin Sonuçları	107
4.3.6. Model 3 için FFBP Yönteminin Sonuçları	108
4.4. MRG ÖZELLİK SETİNE AİT SINIFLANDIRMA SONUÇLARI (MODEL 4)	111
4.4.1. Model 4 için ID3 Yönteminin Sonuçları	111
4.4.2. Model 4 için Doğrusal Regresyon Yönteminin Sonuçları	114
4.4.3. Model 4 için Naive Bayes Yönteminin Sonuçları	115
4.4.4. Model 4 için DVM Yönteminin Sonuçları	115
4.4.5. Model 4 için k-NN Yönteminin Sonuçları	116
4.4.6. Model 4 için FFBP Yönteminin Sonuçları	116
4.4.7. Model 4 için LSTM Yönteminin Sonuçları	125
4.5. ÇOK MODLU KARAR TABANLI VERİ FÜZYONUNUN SONUÇLARI (MODEL 5)	127
4.5.1. Model 1 için Analiz Sonuçları	127
4.5.2. Model 2 için Analiz Sonuçları	128

4.5.3. Model 3 için Analiz Sonuçları.....	130
4.5.4. Model 4 için Analiz Sonuçları.....	131
4.5.5. Model 5 için Analiz Sonuçları.....	132
4.6. EPİLEPSİ VERİ SETİNE AİT DENEYSEL SONUÇLAR	134
5. TARTIŞMA VE SONUÇ.....	136
KAYNAKLAR.....	139
EKLER.....	150
ÖZGEÇMİŞ	Hata! Yer işareti tanımlanmamış.



ŞEKİL LİSTESİ

Sayfa No

Şekil 2.1: Epileptik nöbet geçiren bireye ait örnek EEG	3
Şekil 2.2: Epilepsi hastalığının klinik olarak kesin tanısının konulma evreleri.....	4
Şekil 2.3: Epilepsi hastası bireye ait örnek MRG	6
Şekil 3.1: Veri işleme adımları.....	12
Şekil 3.2: Forbes'te yapılmış bir ankete göre veri hazırlamada geçen zaman	14
Şekil 3.3: Özellik mühendisliği uygulama adımları.....	14
Şekil 3.4: ID3 yöntemi akış diyagramı.....	16
Şekil 3.5: Doğrusal regresyon yöntemi akış diyagramı.....	17
Şekil 3.6: Naive Bayes yöntemi akış diyagramı.....	18
Şekil 3.7: İki boyutlu uzayda doğrusal ayrılabilen veri seti.....	18
Şekil 3.8: Doğrusal ayrılabilen sınıflar arasındaki muhtemel en büyük boşluk.....	19
Şekil 3.9: Doğrusal ayrılabilen sınıflar için hesaplanan destek vektörleri.....	20
Şekil 3.10: DVM yöntemi akış diyagramı.....	20
Şekil 3.11: k-NN yöntemi akış diyagramı.....	21
Şekil 3.12: İnsan beynine ait sinir hücresinin genel yapısı.....	21
Şekil 3.13: Hatanın geriye yayılma yöntemi genel blok diyagramı	23
Şekil 3.14: Gizli katmana ait ağırlık hata fonksiyonu.....	24
Şekil 3.15: Hatanın geriye yayılması yöntemi akış diyagramı.....	26
Şekil 3.16: LSTM ağında gradyan kaybı ve gradyan patlaması sorunu.....	26
Şekil 3.17: t zaman adımındaki LSTM ağına genel yapısı.....	27
Şekil 3.18: LSTM ağına genel akış diyagramı.....	29
Şekil 3.19: Erken füzyon	34
Şekil 3.20: Geç füzyon	35
Şekil 3.21: Ara füzyon	38
Şekil 3.22: Avicenna Hastane veri tabanı varlık-ilişki diyagramı.....	40
Şekil 3.23: Epilepsi veri seti için uygun hasta verisinin seçilme kriteri.....	40
Şekil 3.24: Epilepsi veri seti için istatistiksel özelliklerin çıkarılmasındaki işlem adımları...	42
Şekil 3.25: Standart 10-20 sistemi elektrot konumları – rutin EEG incelemesinde aktivasyon yöntemleri ve süreleri	43

Şekil 3.26: EEG sinyal verisinden anlamlı özelliklerin çıkarılmasında uygulanan işlem adımları.....	44
Şekil 3.27: EEG verisinin EEGLAB yazılımına aktarılması ve 10–20 kanal olarak tanımlanması.....	45
Şekil 3.28: EEG sinyaline yüksek geçirgen filtre uygulanması.....	46
Şekil 3.29: Kötü EEG kanallarının tanımlanması/reddedilmesi.....	46
Şekil 3.30: Kötü kanallardan arındırılmış EEG sinyalindeki kanallar için bağımsız bileşen analizi.....	48
Şekil 3.31: Metin tabanlı radyoloji raporlarının analizi için işlem adımları.....	50
Şekil 3.32: Zipf Yasası’na uygulanan MRG raporları için işlem adımları.....	51
Şekil 3.33: Zipf Yasası’na uygulanan MRG raporlarındaki sözcüklerin geçme sıklığı.....	52
Şekil 3.34: Gerçek epilepsi veri seti.....	53
Şekil 3.35: Sınıflandırıcı toplulukları oluşturmada dört seviye soru	53
Şekil 3.36: Hasta bir bireye nörolog tarafından epilepsi tanısı konulması adımları ile bu tez çalışmasındaki epilepsi sınıflandırma modelinin çalışma adımları.....	54
Şekil 3.37: Epilepsi hastalığının sınıflandırılmasında çok girdili modelleme.....	55
Şekil 3.38: Karar profillerinin (<i>DP</i>) birleştirilmesinin genel yapısı.....	57
Şekil 3.39: Epilepsi hastalığının sınıflandırılmasında çok modlu karar tabanlı veri füzyonu modeli.....	58
Şekil 3.40: Model 1.....	60
Şekil 3.41: Klinik parametre özellik seti için örnek ağaç kök düğümü.....	63
Şekil 3.42: Tablo 3.11’deki verilere göre Yaş – Sınıf (Epilepsi: 0, Sağlıklı: 1) ilişkisi.....	64
Şekil 3.43: Tablo 3.11’deki verilere göre Yaş –Sınıf değerleri için doğrusal model.....	66
Şekil 3.44: İleri beslemeli çok katmalı sinir ağı.....	71
Şekil 3.45: $t = 0$ için ileri beslemeli ağı eğitimi için hesaplama sonucu.....	74
Şekil 3.46: $t = 1$ için ileri beslemeli ağı eğitimi için hesaplama sonucu.....	75
Şekil 3.47: $t = 1$ için geri beslemeli ağı eğitimi için hesaplama sonucu.....	76
Şekil 3.48: $t = 0$ için geri beslemeli ağı eğitimi için hesaplama sonucu.....	77
Şekil 3.49: Model 2.....	78
Şekil 3.50: Model 3.....	79
Şekil 3.51: Model 4.....	81
Şekil 3.52: Geç füzyona uygulanan özellik setlerinden elde edilen sonuçların çok modlu karar tabanlı veri füzyonunun uygulanması.....	82
Şekil 4.1: Klinik muayene özellik seti.....	84
Şekil 4.2: ID3 yöntemine uygulanan klinik parametre özellik seti karar ağacı.....	85

Şekil 4.3: ID3 yöntemine uygulanan klinik parametre özellik seti için karışıklık matrisi.....	85
Şekil 4.4: Klinik parametre özellik setinin sınıflandırılmasında artık değerlerin normal olasılık dağılım grafiği.....	87
Şekil 4.5: Doğrusal regresyon yöntemine uygulanan klinik parametre özellik seti için karışıklık matrisi.....	88
Şekil 4.6: Naive Bayes yöntemine uygulanan klinik parametre özellik seti için karıştırma matrisi	88
Şekil 4.7: DVM yöntemine uygulanan klinik parametre özellik seti için karıştırma matrisi.....	89
Şekil 4.8: k-NN yöntemine uygulanan klinik parametre özellik seti için karıştırma matrisi.....	89
Şekil 4.9: Klinik parametre özellik seti için ÇKA yapısı.....	90
Şekil 4.10: FFBP yöntemine uygulanan klinik parametreler özellik seti için en iyi doğrulama performans grafiği.....	90
Şekil 4.11: FFBP yöntemine uygulanan klinik parametre özellik seti için karıştırma matrisi.....	91
Şekil 4.12: Laboratuvar bulguları özellik seti.....	94
Şekil 4.13: ID3 yöntemine uygulanan laboratuvar bulguları özellik seti karar ağacı.....	95
Şekil 4.14: ID3 yöntemine uygulanan laboratuvar bulguları özellik seti için karıştırma matrisi.....	95
Şekil 4.15: Laboratuvar bulguları özellik setinin sınıflandırılmasında artık değerlerin normal olasılık dağılım grafiği.....	97
Şekil 4.16: Doğrusal regresyon yöntemine uygulanan laboratuvar bulguları özellik seti için karışıklık matrisi.....	97
Şekil 4.17: Naive Bayes yöntemine uygulanan laboratuvar bulguları özellik seti için karışıklık matrisi.....	98
Şekil 4.18: DVM yöntemine uygulanan laboratuvar bulguları özellik seti için karıştırma matrisi.....	98
Şekil 4.19: k-NN yöntemine uygulanan laboratuvar bulguları özellik seti için karıştırma matrisi.....	99
Şekil 4.20: Laboratuvar bulguları özellik seti için ÇKA ağının yapısı.....	99
Şekil 4.21: FFBP yöntemine uygulanan laboratuvar bulguları özellik seti için en iyi doğrulama (validasyon) performans grafiği.....	99
Şekil 4.22: FFBP yöntemine uygulanan laboratuvar bulguları özellik seti için karıştırma matrisi.....	100
Şekil 4.23: EEG özellik seti.....	103
Şekil 4.24: ID3 yöntemine uygulanan EEG özellik seti karar ağacı.....	104
Şekil 4.25: ID3 yöntemine uygulanan EEG özellik seti için karışıklık matrisi.....	105

Şekil 4.26: EEG özellik setinin sınıflandırılmasında artık değerlerin normal olasılık dağılım grafiği.....	106
Şekil 4.27: Doğrusal regresyon yöntemine uygulanan EEG özellik seti için karıştırma matrisi.....	106
Şekil 4.28: Naive Bayes yöntemine uygulanan EEG özellik seti için karışıklık matrisi.....	107
Şekil 4.29: DVM yöntemine uygulanan EEG özellik seti için karışıklık matrisi.....	107
Şekil 4.30: k-NN yöntemine uygulanan EEG özellik seti için karışıklık matrisi.....	108
Şekil 4.31: EEG özellik seti için ÇKA ağının yapısı.....	108
Şekil 4.32: FFBP yöntemine uygulanan EEG özellik seti için en iyi doğrulama (validasyon) performans grafiği.....	108
Şekil 4.33: FFBP yöntemine uygulanan EEG özellik seti için karışıklık matrisi.....	109
Şekil 4.34: MRG özellik seti.....	111
Şekil 4.35: ID3 yöntemine uygulanan MRG özellik seti için karışıklık matrisi.....	114
Şekil 4.36: MRG özellik setinin sınıflandırılmasında artık değerlerin normal olasılık dağılım grafiği.....	114
Şekil 4.37: Doğrusal regresyon yöntemine uygulanan MRG özellik seti için karıştırma matrisi.....	115
Şekil 4.38: Naive Bayes yöntemine uygulanan MRG özellik seti için karışıklık matrisi.....	115
Şekil 4.39: DVM yöntemine uygulanan MRG özellik seti için karışıklık matrisi.....	116
Şekil 4.40: k-NN yöntemine uygulanan MRG özellik seti için karışıklık matrisi.....	116
Şekil 4.41: MRG özellik seti için ÇKA ağının yapısı.....	117
Şekil 4.42: FFBP yöntemine uygulanan MRG özellik seti için en iyi doğrulama (validasyon) performans grafiği.....	117
Şekil 4.43: FFBP yöntemine uygulanan MRG özellik seti için karışıklık matrisi.....	117
Şekil 4.44: LSTM ağına MRG özellik setinin uygulama adımları.....	125
Şekil 4.45: LSTM ağına MRG özellik setinin uygulanması.....	126
Şekil 4.46: LSTM yöntemine uygulanan MRG özellik seti için karışıklık matrisi.....	126

TABLO LİSTESİ

Sayfa No

Tablo 2.1: Uluslararası epilepsi nöbetlerinin özeti.....	3
Tablo 3.1: Karışıklık matrisi.....	29
Tablo 3.2: n -kat çaprazlama yöntemine göre doğrulama ($n = 10$).....	31
Tablo 3.3: Farklı topluluk metotlarının teorik analizleri.....	33
Tablo 3.4: Epilepsi veri setindeki hasta detayları.....	41
Tablo 3.5: Bireylere ait EEG sinyalinde elde edilen istatistiksel özellikler.....	49
Tablo 3.6: Radyolog tarafından yazılan MRG rapor verisi için ön işlem adımları.....	50
Tablo 3.7: Epilepsi_MRG_Zipf sözlüğünden ve Epilepsi_MRG_Zipf sözlüğünden elde edilen 9 farklı özellik setleri için oluşturulmuş doğrusal model sonuçları.....	52
Tablo 3.8: Gerçek epilepsi veri setinin özet açıklaması.....	56
Tablo 3.9: Klinik parametre özellik setinden seçilmiş örnek veri.....	62
Tablo 3.10: Tablo 3.9'daki yaş değişkeninin gruplandırılması.....	63
Tablo 3.11: Yaş – Sınıf bilgisi.....	64
Tablo 3.12: Klinik özellik setinden seçilmiş örnek veri.....	66
Tablo 3.13: Klinik özellik setinden seçilmiş örnek veri.....	69
Tablo 3.14: Gözlem değerlerinin $x = \{0, 2, 51\}$ noktasına uzaklıkları.....	70
Tablo 3.15: $k = 4$ için komşulukların belirlenmesi.....	70
Tablo 4.1: Özellik setlerinin optimum sınıflandırılmasında yöntemlere uygulanan ortak parametre değerleri.....	83
Tablo 4.2: ID3 yöntemine uygulanan klinik parametre özellik seti karar kuralları.....	85
Tablo 4.3: Klinik parametre özellik seti doğrusal regresyon modelleme sonuçları.....	86
Tablo 4.4: Klinik parametre özellik seti için giriş katmanından ara katmanına doğru ağırlık değerleri.....	91
Tablo 4.5: Klinik parametre özellik seti ara katmandan çıkış katmanına doğru ağırlık değerleri.....	92
Tablo 4.6: Klinik parametre özellik setinde bulunan özelliklerin katkısı.....	93
Tablo 4.7: ID3 yöntemine uygulanan laboratuvar bulguları özellik seti karar kuralları.....	95
Tablo 4.8: Laboratuvar bulguları özellik seti doğrusal regresyon modelleme sonuçları.....	96
Tablo 4.9: Laboratuvar bulguları özellik seti için giriş katmanından ara katmanına doğru ağırlık değerleri.....	100
Tablo 4.10: Laboratuvar bulguları özellik seti için ara katmandan çıkış katmanına doğru ağırlık	

değerleri.....	101
Tablo 4.11: Laboratuvar bulguları özellik setinde bulunan özelliklerin katkısı.....	102
Tablo 4.12: ID3 yöntemine uygulanan EEG özellik setine ait karar kuralları.....	104
Tablo 4.13: EEG özellik seti doğrusal regresyon modelleme sonuçları.....	105
Tablo 4.14: EEG özellik seti için giriş katmanından ara katmanına doğru ağırlık değerleri.....	109
Tablo 4.15: EEG özellik seti için ara katmandan çıkış katmanına ağırlık değerleri.....	110
Tablo 4.16: EEG özellik setinde bulunan özelliklerin katkısı.....	110
Tablo 4.17: ID3 yöntemine uygulanan MRG özellik setine ait karar kuralları.....	112
Tablo 4.18: MRG özellik seti için giriş katmanından ara katmanına doğru ağırlık değerleri..	118
Tablo 4.19: MRG özellik seti için ara katmandan çıkış katmanına doğru ağırlık değerleri...	123
Tablo 4.20: MRG özellik setinde bulunan özelliklerin katkısı.....	124
Tablo 4.21: Model 1 (Klinik Parametre özellik seti için uygulanan) modelleme sonuçları....	127
Tablo 4.22: Model 1 (Klinik parametre özellik seti modelleme) istatistiksel analizi.....	128
Tablo 4.23: Model 2 (Laboratuvar bulguları özellik seti modelleme) sonuçları.....	128
Tablo 4.24: Model 2 (Laboratuvar bulguları özellik seti modelleme) istatistiksel analizi....	129
Tablo 4.25: Model 3 (EEG özellik seti modelleme) sonuçları.....	129
Tablo 4.26: Model 3 (EEG özellik seti modelleme) istatistiksel sonuçları.....	130
Tablo 4.27: Model 4 (MRG özellik seti için uygulanan modelleme) sonuçları.....	130
Tablo 4.28: Model 4 (MRG özellik seti modelleme) istatistiksel sonuçları.....	131
Tablo 4.29: Özellik setleri için tahminleme sonuçları.....	133
Tablo 4.30: Çok modlu karar tabanlı veri füzyonu tekniği ile elde edilen tahminleme performansı.....	133
Tablo 4.31: Epilepsi veri seti için tekil sınıflandırma sonuçları.....	134
Tablo 4.32: Epilepsi veri seti için tekil sınıflandırma tahminleme sonuçları.....	134

SİMGE VE KISALTMA LİSTESİ

Simgeler	Açıklama
X_t	: Giriş vektörü
$cov_{a,b}$	: (a, b) gibi iki veri arasındaki değişimin kovaryans değeri
D	: Veri tabanı
x_i	: Veri tabanındaki bir kayıt bilgisi
Y	: Sınıf
H	: Entropi
H_0	: Optimal ayırma düzlemi
S	: Kazanım
E	: Hata değeri
p-değeri	: Hatanın maksimum kabul edilebilir düzeyi
F	: Doğrusal regresyon modelinin önemi
b	: Regresyon katsayıları
w	: Ağırlık değeri
$P(C_1 x_i)$	: x_i değerinin C_1 sınıfına ait olma olasılığı
k	: En yakın komşuluk
$d(i, j)$	: $(i$ ve $j)$ arasındaki uzaklık
d_k	: k değerinin hedef çıkış değeri
o_k	: k değerinin ağdan elde edilen asıl çıkış değeri
Δw_{kj}	: Ağırlıkların değişimi
μ	: Öğrenme katsayısı
$f(net_j)$	: Transfer fonksiyonu
α	: Momentum
N	: İterasyon sayısı
C_t	: Mevcut bellek hücresi (t anında)
H_t	: Mevcut hücre akışı (t anında)
f_t	: Unut kapısı (t anında)
I_t	: Giriş kapısı (t anında)
O_t	: Çıkış kapısı (t anında)
$Pr(a)$	: İki gözlem değeri için gözlemlenen uyumların toplam değere orantısı
$Pr(e)$	: Beklenen değerın şansa bağlı ortaya çıkma olasılığı

$x(n)$	: Giriş sinyali
$y(n)$	: Çıkış sinyali
$Kurt(R)$	: Kurtosis
DP	: Decision Profile (Karar Profili)
avg	: Average (Ortalama)

Kısaltmalar

Açıklama

ILAE	: International League Against Epilepsy (Uluslararası Epilepsi ile Savaş Derneği)
FIR	: Finite Impulse Response (Sonlu Dürtü Yanıtı)
MRG	: Manyetik Rezonans Görüntüleme
EEG	: Elektroansefalogram
WBC	: White Blood Cell (Beyaz Kan Hücresi)
RBC	: Red Blood Cell (Eritrosit)
HGB	: Hemoglobin
HCT	: Hematokrit
PLT	: Plates, Trombosit
MCV	: Mean Eritrosit Volume (Ortalama Eritrosit Hacmi)
MCH	: Mean Corpuscular Hemoglobin (Ortalama Korpüsküler Hemoglobin)
MCHC	: Mean Corpuscular Hemoglobin Concentration (Ortalama Korpüsküler Hemoglobin Konsantrasyonu)
TBA	: Temel Bileşen Analizi
DVM	: Destek Vektör Makinası
YSA	: Yapay Sinir Ağları
k-NN	: k Nearest Neighbor (k En Yakın Komşuluk)
FFBP	: Feed Forward Back Propagation (İleri Beslemeli Geri Yayılımlı)
ÇKA	: Çok Katmanlı Algılayıcı
CNN	: Convolutional Neural Network (Evrşimli Sinir Ağı)
PQN-NN	: Parallel Quasi-Newton Neural Network (Paralel Paralel Yarı Newton Sinir Ağı)
DS	: Dempster-Shafer
BBA	: Bağımsız Bileşen Analizi
RMSE	: Root Mean Squared Error (Kök Ortalama Kare Sapma)
LSTM	: Long Short-Term Memory (Uzun Kısa Süreli Hafıza)
SFFS	: Sequential Forward Feature Selection (Sıralı İleri Özellik Seçimi)

MPRAGE	: Magnetization Prepared Rapid Acquisition Gradient Echo (Hazırlanmış Hızlı Toplama Gradient Eko)
STFT	: Short Time Fourier Transform (Kısa Zamanlı Fourier Dönüşümü)
DWT	: Discrete Wavelet Transform (Ayrık Dalgacık Dönüşümü)
BOW	: Bag of Words (Sözcük Çantası)
ROC	: Receiver Operating Characteristic (Alıcı İşlem Karakteristikleri)



ÖZET

DOKTORA TEZİ

FARKLI VERİ TÜRLERİNİN BİRLEŞTİRİLMESİ İLE EPİLEPSİ HASTALIĞININ OTOMATİK TESPİTİ VE SINIFLANDIRILMASI

Şengül BAYRAK

İstanbul Üniversitesi-Cerrahpaşa

Lisansüstü Eğitim Enstitüsü

Bilgisayar Mühendisliği Anabilim Dalı

Danışman: Doç. Dr. Eylem YÜCEL DEMİREL

II. Danışman: Doç. Dr. Hidayet TAKCI

Epilepsi, beyin normal aktivitesinin, sinir hücrelerinde geçici olarak meydana gelen anormal elektriksel aktivite sonucu bozulması ile oluşan klinik bir durumdur. Dünyada yaklaşık 40 milyon epilepsi hastası, ülkemizde ise 700 bin civarında epilepsi hastası olduğu bilinmektedir. Bugüne kadar epilepsi hastalığının tam olarak bir sebebi açıklanamamıştır. Uluslararası Epilepsi ile Savaş Derneği (International League Against Epilepsy – ILAE) epileptik nöbet tiplerini sistematik bir şekilde sınıflandırmıştır. ILAE tarafından tanımlanmış nöbet tiplerine rağmen her hastanın geçirdiği nöbet kendisine özgü olduğundan farklılıklar gösterebilmektedir. Bu durum bazı hastalar için epilepsi tanısının konulmasını güçleştirmektedir. Bazı bireylerde başka bir bozukluğun yol açtığı belirtiler yanlışlıkla epilepsi tanısı olarak tanımlanabilmektedir. Bu yüzden, epilepsi hastalığının tanımlanmasında mühendislik ve tıp bilimleri arasında etkileşimli çalışmalar son yıllarda uygulama alanı bulmuştur. Epilepsi hastalığının tanısı sürecinde doktora yardımcı makine öğrenmesi yöntemleri ile geliştirilmiş çalışmaların artış göstermesi iki bilim arasındaki çalışmaların ilgi görmesine neden olmuştur. Özellikle mühendisliği yöntemleri kullanılarak geliştirilen matematiksel yöntemler yaklaşımları ile epilepsi hastalığının sınıflandırılması klinik yöntemlere güçlü bir alternatif oluşturmuş ve literatürde daha iyi sonuçların elde edildiği yeni çalışmalara ihtiyaç duyulmuştur. Bugüne kadar yapılmış çalışmalarda epilepsi hastalığının tanımlanmasında hastaya ait sadece elektroensefalogram (EEG) verisi ile önemli diğer verilerin; demografik özellikler, nörolojik muayene sırasında doktor epikriz rapor verisinden elde edilen klinik bulgular, laboratuvar bulgularından elde edilen kan değerleri, Manyetik Rezonans Görüntüleme (MRG) radyoloji raporlarından elde edilen verilerin bir arada kullanılarak yapıldığı kapsamlı bir çalışmaya

rastlanılmamıştır. Bu tez çalışmasında, epilepsi hastalığının tanımlanmasında 18 yaş üstü erişkin bireylere ait (i) demografik verilerinden hasta ile ilgili özellikler çıkarılmış, (ii) doktor epikriz rapor verisinden hasta şikayetleri çıkarılmış, (iii) laboratuvar bulgularından kan değerleri elde edilmiş, (iv) Sonlu Dürtü Yanıtı filtreleme ve Bağımsız Bileşen Analizi yöntemleri kullanılarak EEG verisine ait anlamlı özellikler çıkarılmış, (v) doğal dil işleme tekniklerinden ZipF yasası tekniği ile MRG rapor verisine ait anlamlı özellikler çıkarılmıştır. (i – v) adımları ile epilepsi hastalığının ayırıcı tanısında yararlı olabilecek tüm parametreler bir arada kullanılarak hastalıkla ilgili anlamlı özellikler çıkarılarak, hastalığa özgü ülkemizdeki epilepsi hastalarına ait emsal niteliğinde gerçek bir veri seti oluşturulmuştur. Ayrıca, hastalıkla ilgili anlamlı özelliklerin çıkarılması sayesinde hastalığın erken dönemde teşhisi için koruyucu tedaviye yansıtacak adımların atılması ile doktora yardımcı bir sistem oluşturulması hedeflenmiştir. Oluşturulan veri setindeki veriler kullanılarak epilepsi ve sağlıklı bireylere ait verilerin etkin olarak sınıflandırılması için çok modlu karar tabanlı veri füzyonu tekniği önerilmiştir. Çok modlu karar tabanlı veri füzyonu tekniği ile modelleme başarımı %93,75 doğruluk oranı olarak elde edilmiştir. Tekil modelleme tekniği ile sınıflandırılan epilepsi veri seti için ID3, DVM, ve FFBP yöntemlerinden sınıflandırmada en başarılı tahminleme doğruluğunu %83,33 olarak hesaplamıştır. Bu tezde önerilen çok modlu karar tabanlı veri füzyonu tekniği tekil modelleme tekniğinden %10,42 oranında daha başarılı olmuştur. Bu tez çalışmasında epilepsi tanısında optimum özellik seti oluşturularak sınıflandırma doğruluğu başarılı en iyi füzyon yöntemi araştırılmıştır.

Haziran 2021, 176 sayfa.

Anahtar kelimeler: Epilepsi, epilepsi sınıflandırma, medikal veri madenciliği, tahminsel modelleme, veri füzyonu, çok modlu karar tabanlı veri füzyonu

SUMMARY

[Ph.D. THESIS]

**| AUTOMATIC DETECTION AND CLASSIFICATION OF EPILEPSY
DISEASE BY COMBINING DIFFERENT DATA TYPES |**

[Şengül BAYRAK]

Istanbul University–Cerrahpasa

Institute of Graduate Studies

[Department of Computer Engineering]

Supervisor : [Assoc. Prof. Dr.] [Eylem YÜCEL DEMİREL]

[Co–Supervisor : Assoc. Prof. Dr. Hidayet TAKCI |

Epilepsy is a clinical condition that occurs when the normal activity of the brain is disrupted as a result of the abnormal electrical activity, which happens temporarily in the nerve cells. It is known that there are approximately 40 million epilepsy patients in the world and around 700 thousand epilepsy patients in our country. Until now, no exact cause of epilepsy disease has been explained. The International League Against Epilepsy (ILAE) systematically classified epileptic seizure types. Despite the types of seizures defined by the ILAE, each patient's seizure is unique, so they may differ. This situation makes it difficult to diagnose epilepsy for some patients. In some individuals, symptoms caused by another disorder may be mistakenly defined as epilepsy. Therefore, interactive studies between engineering and medical sciences have been applied in the definition of epilepsy disease in recent years. The increase in studies developed by physician assistant machine learning methods during the diagnosis of epilepsy has attracted attention between the two sciences. Classification epilepsy with the mathematical methods approaches have been developed by using feature engineering methods, has created a strong alternative to clinical methods, and new studies with better results are needed in the literature. In the studies conducted to date, only electroencephalogram (EEG) data of the patient and other important data which are demographic characteristics, clinical findings obtained from doctor epicrisis report data during the neurological examination, blood values obtained from

laboratory findings, Magnetic Resonance Imaging (MRI) radiology reports. In this thesis, for the diagnosis of epilepsy, adult individuals over the age of 18 (i) patient-related characteristics have been extracted from demographic data, (ii) patient complaints have been extracted from doctor epicrisis report data, (iii) blood values have been obtained from laboratory findings, (iv) the meaningful EEG features have been extracted by using the Finite Impulse Response filtering and Independent Component Analysis methods, (v) the significant features of the MRI report data have been extracted by the ZipF law technique which is one of the natural language processing techniques. By using all the parameters that may be useful in the differential diagnosis of epilepsy with steps (i – v) together, the significant characteristics related to the disease have been extracted, and a real data set has been created for the epilepsy patients in our country. In addition, it is aimed to create a system that helps doctors by taking the steps that will be reflected in preventive treatment for early diagnosis of the disease by removing the significant characteristics of the disease. Using the data in the data set created, it has been proposed multi-mode decision-based data fusion technique to effectively classify data belonging to epilepsy individuals and healthy individuals. Modeling performance has been achieved as 93.75% accuracy rate with the multimodal decision-based data fusion technique. With the singular classification technique applied to the epilepsy data set, the prediction success has been obtained with the ID3, DVM, and FFBP methods with an accuracy rate of 83.33%. According to the results of the thesis, it has been seen that the classification accuracy of the multimodal decision-based classification technique was 10.42% more successful than the classification accuracy of the singular classification technique. In this thesis, the best fusion method with successful classification accuracy has been investigated by generating the optimum feature set in the diagnosis of epilepsy.

June 2021, 176 pages.

Keywords: Epilepsy, epilepsy classification, medical data mining, predictive modeling, data fusion, multimodal decision-based data fusion

1. GİRİŞ

İnsanoğlunun uygarlığın ilk dönemlerinden bu yana baş etmek zorunda olduğu epilepsi hastalığı, hastaların günlük aktivitelerini kısıtlayan ve yaşamlarını tehdit eden önemli nörolojik bir rahatsızlıktır. Epilepsi, beyin normal aktivitesinin, sinir hücrelerinde geçici olarak meydana gelen anormal elektriksel aktivite sonucu bozulması ile oluşan klinik bir durumdur [1 – 3].

Bugüne kadar epilepsi hastalığının tam olarak bir sebebi bilimsel olarak açıklanamamıştır. Ancak doğuştan gelen bozukluklar (genlerin bulunduğu kromozomlarda meydana gelen hastalıklar, genetik etmenler, aileden gelen mutasyona uğramış genler, enzim eksikliği), gebelik sırasında ortaya çıkan bebeğin gelişimini etkileyecek mikrobik hastalıklar, gebelik sırasında annenin sigara/alkol kullanması, beyin zarlarında iltihap oluşması, beyinde meydana gelen tümörler, enfeksiyon, ateşli havale geçirilmesi, iskemik lezyon, tiroit hastalıkları, herhangi bir kaza sonucu beyin zedelenmesi, çeşitli nedenlerden dolayı travma geçirme, beslenme bozuklukları epilepsi hastalığına sebep olan etmenler olarak sıralanabilmektedir [2 – 6].

Dünyada yaklaşık 40 milyon, ülkemizde ise 700 bin civarında epilepsi hastası olduğu bilinmektedir [7]. 20. yy sonlarında ILAE epileptik nöbetleri sistematik bir şekilde sınıflayarak epilepsi hastalığı hakkında hekimlerin klinik pratiğine önemli katkı sağlamıştır. Hastalığın tanısı için nöbetin gözlenmesi ve bu gözlemlerin hekime aktarılması çok önemlidir.

ILAE tarafından tanımlanmış nöbet tiplerine rağmen her hastanın geçirdiği nöbet kendisine özgü olduğundan farklılıklar gösterebilmektedir. Bu durum bazı hastalar için epilepsi tanısının konulmasını güçleştirmektedir. Bazı kişilerde başka bir bozukluğun yol açtığı belirtiler yanlışlıkla epilepsi tanısı olarak tanımlanabilmektedir. Bu yüzden, epilepsi hastalığının tanımlanmasında mühendislik metotları ve tıp etkileşimi konularındaki çalışmalar son yıllarda uygulama alanı bulmuştur. Bunlardan biri olan makine öğrenmesi yöntemleri, epilepsi hastalığının teşhisi sürecinde doktora yardımcı çalışmaların son yıllarda uygulanmasıyla tıp alanında yoğun ilgi görmüştür. Geliştirilen matematiksel yöntem yaklaşımları ile epilepsi hastalığının sınıflandırılması klinik yöntemlere güçlü bir alternatif oluşturmuş ve literatürde daha iyi sonuçların elde edildiği yeni çalışmalara ihtiyaç duyulmuştur. Hastanın yaşam kalitesi açısından epilepsi hastalığının başlangıç safhasında teşhis edilebilmesi için geliştirilmiş

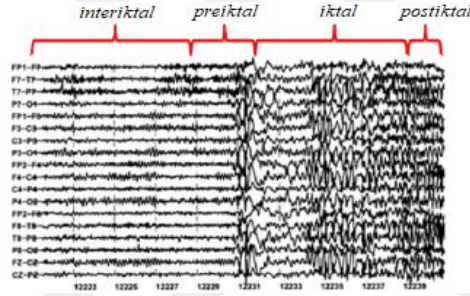
matematiksel yöntemler ile kestirilmesi hekimin doğru karar vermesinde de yardımcı önemli araçlar haline gelmiştir.

Bu tez çalışmasında, özgün epilepsi veri seti için Haliç Üniversitesi Etik Kurul kararıyla, Avicenna Hastanesi Nöroloji bölümünde, 30 Haziran 2006 ile 19 Mart 2020 tarihleri arasında takip edilen, klinik olarak kesin epilepsi tanısı konulmuş 97 birey ile vitamin eksikliği tanısı konulmuş 25 kontrol grubu sağlıklı bireye ait, demografik özellikler (yaş, cinsiyet), nörolojik muayene epikriz raporundan elde edilen özellikler (şuur kaybı, terleme, uyku bozukluğu, baş ağrısı, konuşmada zorlanma, görme bozukluğu, el-kollarda uyuşma, baş dönmesi, bilinç kaybı, unutkanlık, bayılma, geç algılama, dalma, ateş, kusma, çabuk yorulma, vücutta kasılma, çenede kilitlenme , ... gibi), laboratuvar bulgularından elde edilen kan test sonuçlarına (Beyaz Kan Hücresi – White Blood Cell – WBC, Kırmızı kan hücresi sayısı Eritrosit – RBC, Hemoglobin – HGB, Hematokrit – HCT, Plates, Trombosit sayısı – PLT, Ortalama Eritrosit Hacmi – Mean Eritrosit Volume – MCV, Ortalama eritrosit hemoglobini – Mean Corpuskuler Hemoglobin – MCH, Ortalama eritrosit hemoglobin konsantrasyonu – Mean Corpuskuler Hemoglobin Concentration – MCHC, ... gibi) ait veriler seçilerek, EEG sinyal verisine sinyal işleme yöntemleri; radyolog tarafından okunmuş beyin MRG rapor verisine doğal dil işleme teknikleri uygulanarak hastalıkla ilgili anlamlı istatistiksel özellikler çıkarılarak, gerçek epilepsi veri seti oluşturulmuştur. Epilepsi hastalığının ayırıcı tanısında yararlı olabilecek tüm parametreler bir arada kullanılarak hastalıkla ilgili anlamlı özellikler çıkarılmış ve hastalığa özgü ülkemizdeki hastalara ait emsal oluşturacak gerçek bir epilepsi veri seti oluşturulmuştur. Bu tez çalışması birçok çalışma için önemli bir kaynak oluşturacak, daha yüksek duyarlıklılı ve daha güvenilir sonuçların elde edilebileceği çalışmaların yapılabilmesi için özgün emsal niteliğinde oluşturulan veri seti literatüre kazandırılarak katkı sağlayacaktır. Ayrıca, hastalıkla ilgili anlamlı özelliklerin çıkarılması sayesinde hastalığın erken dönemde teşhisi için koruyucu tedaviye yansıyacak adımların atılması ile doktora yardımcı bir sistem olması hedeflenmektedir. Oluşturulmuş olan gerçek epilepsi veri seti ile epilepsi hastalığının otomatik tanımlanmasında ve sınıflandırılmasında çok modlu karar tabanlı veri füzyonu tekniği ile optimum modelleme yapılmıştır.

2. GENEL KISIMLAR

2.1. EPİLEPSİ HASTALIĞI

Halk arasında “Sara Hastalığı” olarak da bilinen epilepsi, epileptik nöbetler, bilinç kaybı, bayılma, titreme, kasılma, sabit bakma, nefes darlığı, nöbet esnasında; ağızdan köpük gelmesi, idrar kaçırma, istifra etme, nöbet sonrası; şaşkınlık, uyku hali, bilinç bulanıklığı, korku, endişe yaşanması ile kendisini göstermektedir [8, 9]. Epileptik nöbetin olduğu dönem iktal dönem, nöbetin bitiminden hemen sonraki dönem postiktal dönemdir. Nöbetler arası süre interiktal dönem olarak adlandırılmaktadır. Epileptik nöbet geçiren bir bireye ait örnek EEG, Şekil 2.1’de gösterilmiştir [7].



Şekil 2.1: Epileptik nöbet geçiren bireye ait örnek EEG.

Hastalığın tanısı için nöbetin gözlenmesi ve bu gözlemlerin doktora aktarılması çok önemlidir. Epileptik nöbetlerin sınıflamasında klinik ve EEG bulgularına dayanarak her bir nöbet ayrı sınıflandırılmaktadır [10]. Uluslararası epileptik nöbetler Tablo 2.1’de özetlenmiştir [7].

Tablo 2.1: Uluslararası epilepsi nöbetlerinin özeti.

I.	Parsiyel (fokal, lokal) nöbetler
A.	Basit parsiyel nöbetler (bilinç bozukluğu yok)
B.	Kompleks parsiyel nöbetler (bilinç bozukluğu var)
C.	Sekonder jeneraliz nöbetlere dönüşen parsiyel nöbetler
II.	Jeneralize nöbetler
A.	Absans (petit mal) nöbetler
B.	Miyoklinik nöbetler
C.	Tonik nöbetler
D.	Atonik nöbetler
E.	Klonik nöbetler
F.	Tonik – klonik (grand mal) nöbetler
III.	Sınıflandırılmayan epileptik nöbetler (yetersiz veri nedeniyle)

Tablo 2.1’e göre nöbetler (a) parsiyel nöbetler (beyinde nispeten küçük bir lokalizasyonda başlayan), (b) jeneralize nöbetler (bilateral simetrik ve lokal başlangıcı olmayan) olarak iki

geniş sınıfa ayrılmaktadır. Epilepsi tanısında hastaya ait klinik bulgular, laboratuvar bulguları, EEG ve MRG çok önemli veri kaynaklarıdır.

2.2. EPİLEPSİ HASTALIĞININ TANISINDA KULLANILAN TEKNİKLER

Klinik olarak epilepsi hastalığının kesin tanısı nörolog tarafından Şekil 2.2'deki evrelere göre konulmaktadır.



Şekil 2.2: Epilepsi hastalığının klinik olarak kesin tanısının konulma evreleri.

Şekil 2.2(a,b)'de klinik muayene sırasında hasta şikâyetlerine göre doktorun hasta için muayene notu oluşturma, Şekil 2.2(c)'de epileptik nöbetin metabolizmadan kaynaklandığının kesin olarak anlaşılabilmesi için laboratuvar ortamında kan testlerinin yapılması, Şekil 2.2(d)'de hastanın kafatası üzerinden belirlenmiş özel noktalara yerleştirilen elektrotlar aracılığı ile çekilen EEG verisi, Şekil 2.2(e)'de nöbetin epilepsi hastalığından kaynaklandığından emin olunduğunda bunun beyindeki yapısal bozukluk ve tümör gibi nedenlerden dolayı olup olmadığının MRG ile incelenmesi, Şekil 2.2 (f)'de Şekil 2.2 (a-e)'deki aşamalardan elde edilen sonuçlara göre nörolog tarafından hastaya klinik olarak epilepsi olup olmadığı teşhisinin konulduğunu göstermektedir [11].

2.2.1. Klinik Muayene Bulguları

Nörolojik muayene sırasında hekimin hastanın şikâyetleri (şuur kaybı, terleme, uyku bozukluğu, baş ağrısı, konuşmada zorlanma, görme bozukluğu, el-kollarda uyuşma, baş dönmesi, bilinç kaybı, unutkanlık, bayılma, geç algılama, dalma, ateş, kusma, çabuk yorulma, vücutta kasılma, çenede kilitlenme, ... gibi), yaş, cinsiyet ile ilgili bilgileri epikriz raporuna not etmesi ile elde edilmektedir.

2.2.2. Laboratuvar Bulguları

Epilepsi hastalığına tanı koymak ve takip etmek, hastanın hangi tipte nöbet geçirmekte olduğunu belirlemek, toksinler, enfeksiyonlar, ilaç veya alkol yoksunluğu, nöbetlere neden olabilen diyabet gibi altta yatan hastalıkları tanımlamak ve epilepsi hastalığını benzer belirti ve bulgularının bir bölümüne neden olabilen bayılma veya inme gibi rahatsızlıklardan ayırt etmek için laboratuvar bulgularından elde edilen kan test sonuçları (WBC, RBC, HGB, HCT, PLT, MCV, MCH, MCHC, ... gibi) kullanılmaktadır.

2.2.3. Elektroensefalogram (EEG)

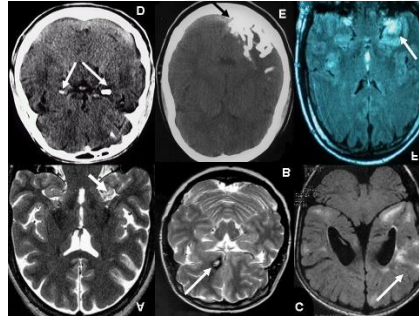
EEG, beynin elektriksel faaliyetlerinin kayıtlanması işlemi olup, uluslararası kabul edilen haritalandırma işlemine uyarak belirli standartlarda, dijital ortamda ya da kâğıt üzerine yazdırma ile yapılmaktadır. 1947 yılında Londra’da gerçekleştirilen ilk uluslararası EEG kongresinde, EEG çekiminde kullanılacak elektrotlar için standart bir yerleştirme metodunun gerekliliği üzerinde durulmuştur. 10–20 sistemi olarak tanımlanan yerleştirme düzeni kabul gören bir metot haline gelmiştir [12]. Beyin biyoelektliği beyin kabuğunda yer alan nöron ve uzantılardan oluşmaktadır. Beyin biyoelektliği kaydedilen EEG’deki tek tek nöronların biyoelektroniğini değil aynı zamanda çalışan bir grup nöronun elektliğinin de kaydedilmesidir. Birlikte çalışan nöronların toplam oluşturduğu elektrik ritmik bir düzeni göstermektedir. Beyin devamlı elektriksel faaliyet içerisinde olup, yaşamın belirli devrelerinde bu elektriksel faaliyeti, belirgin gelişkinlik düzeyleri göstermekte, yine günlük yaşamın belirli evrelerinde de (uyku ve uyanıklık gibi) belirli standartlarda elektriksel faaliyetini devam ettirmektedir [11].

Beyin biyoelektliği, ritmik özelliği olan ve eş zamanlı çalışma gösteren beyin dalgalarından oluşmaktadır. Beyni etkileyen hastalıklarda, başta epilepsi olmak üzere bu doğal ritimlerin yerini farklı, anormal ritimler ve elektrik dalgaları almaktadır. Beyin dalgaları frekans, amplitüt ve faz olmak üzere üç ana karakteri göstermektedir. Frekans; bir saniye içerisinde tekrarlayan dalga sayısını ifade etmektedir. Amplitüt; beyin dalgasının yüksekliği, voltajı o dalganın amplitütüdür. Faz; aynı beyin yarım küresinden alınan trasedeki dalgaların çıkış ve inişlerini temsil etmektedir. Beyin yarım küresinde normal işleyiş devam ediyorsa faz paralelliği göstermektedir. Beyin yapısını bozan hastalıklı bir alan bu paralelliği bozmaktadır ve hastalıklı alanın önündeki dalgalar alana doğru yani arkaya doğru yönelecektir. Arkasındaki bölgenin dalgaları ise alana doğru, yani öne doğru yönelecektir. Böylece, hastalıklı alanda

karşılaşacaktır. Bu durum faz karşılaşmasıdır ve hastalıklı alanın tanımlanmasında çok önemlidir [4 – 10]. Beyin dalgalarının amplitütü 0–200 mikro volt ve frekansları her birkaç saniyede 1 ile 50 ve daha fazlası olabilmektedir. Dalgaların özellikleri beyin aktivitesine ve durumuna göre değişmektedir. Çoğu zaman beyin dalgaları düzensizdir ve belirli bir ritim tanımlanamamaktadır. Tanımlanabilen bazı özel ritimler şöyledir. Alfa ritmi; saniyede 8–13 frekansında ve 40–50 mikro volt olan bu dalgalar uyanık ve dinlenme halinde görülmektedir. Beta ritmi; saniyede 14 frekansında ve üzerindeki dalgalar beta dalgalarıdır. Genellikle 10–20 mikro volt olan düşük amplitütlü dalgaları temsil etmektedir. Delta ritmi; saniyede 3,5 frekansın altında olan dalgalarıdır.

2.2.4. Manyetik Rezonans Görüntüleme (MRG)

Epilepsi tanısında, beynin yapısal ve fonksiyonel değerlendirmesi ve hastalığın yönetimi MRG bir diğer önemli veri kaynağıdır. Epileptik nöbetler ile nedensel olarak ilişkili olabilecek serebral lezyonların tespit edilmesinde rol oynamaktadır [8, 9]. Epilepsi hastası bireyine ait örnek MRG, Şekil 2.3'te gösterilmiştir [10].



Şekil 2.3: Epilepsi hastası bireye ait örnek MRG.

MRG çekimi yapılmış bir bireyin raporu radyolog tarafından yazılmaktadır. Radyoloji raporlarındaki insidans bulgularının kayıtları klinik görevlerde, uzmanlar tarafından önceden tanımlanmış bir dizi anahtar kelime kullanılarak oluşturulmaktadır.

2.3. EPİLEPSİ HASTALIĞININ TANISINDA MAKİNE ÖĞRENMESİ YÖNTEMLERİ VE LİTERATÜR ÇALIŞMALARI

ILAE tarafından tanımlanmış nöbet tiplerine rağmen her hastanın geçirdiği nöbet kendisine özgü olduğundan farklılıklar gösterebilmektedir. Bu durum bazı hastalar için epilepsi tanısının konulmasını güçleştirmektedir. Bazı kişilerde başka bir bozukluğun yol açtığı belirtiler

yanlılıkla epilepsi tanısı olarak tanımlanabilmektedir. Bu yüzden, epilepsi hastalığının tanımlanmasında mühendislik bilimindeki metotlar ve tıp bilimi konularındaki çalışmalar son yıllarda uygulama alanı bulmuşlardır. Bunlardan biri olan makine öğrenmesi yöntemleri ile epilepsi hastalığının teşhisi sürecinde doktora yardımcı çalışmaların son yıllarda uygulanmasıyla tıp alanında yoğun ilgi görmesine neden olmuştur. Geliştirilen matematiksel yöntem yaklaşımları ile epilepsi hastalığının sınıflandırılması klinik yöntemlere güçlü bir alternatif oluşturmuş ve literatürde daha iyi sonuçların elde edildiği yeni çalışmalara ihtiyaç duyulmuştur. Hastanın yaşam kalitesi açısından epilepsi hastalığının başlangıç safhasında teşhis edilebilmesi için oluşturulan matematiksel yöntemler ile kestirilmesi hekimin doğru karar vermesinde de yardımcı önemli araçlar haline gelmiştir. Son yıllarda, epilepsi hastalığının tanımlanmasında genellikle sadece EEG verisi ile epilepsi hastalığının sınıflandırılması yapılmıştır.

Bugüne kadar yapılan araştırmalarda [13 – 30] epilepsi ile ilgili çok sayıda çalışma olmasına rağmen epileptik nöbetlerin tespit edilmesi ve sınıflandırılmasında klinik muayene bulguları, laboratuvar bulgular, EEG ve MRG verisi ile oluşturulmuş olan veri seti ile yapılmış çalışmalar yetersizdir. Literatürde, epilepsi hastası bireylerin EEG sinyal verisinden dalgacık katsayısı, entropi, fraktal boyutlandırma, istatistiksel özellikler (ortalama, standart sapma ,..., gibi) çıkarılmış ve bu veriler kullanılarak çalışmalar yapılmıştır.

Bayrak ve diğ. [13], epilepsi hastalığının sınıflandırılmasında BONN EEG veri tabanı için uygulama geliştirmişlerdir. Çalışmalarında, Kaiser tabanlı Sonlu Dürtü Yanıtı ve Temel Bileşen Analizi (TBA) yöntemleri aracılığıyla ekstrakraniyal ve intrakraniyal EEG sinyal özelliklerini çıkarmışlardır. Alt uzay k-en yakın komşu yöntemine EEG sinyal özelliklerini uygulayarak %94,40 sınıflandırma doğruluğu elde etmişlerdir.

Bayrak ve diğ. [14], epilepsi hastalığının sınıflandırılmasında BONN EEG veri tabanı için uygulama geliştirmişlerdir. Çalışmalarında, Kaiser tabanlı Sonlu Dürtü Yanıtı ve TBA yöntemleri aracılığıyla ekstrakraniyal ve intrakraniyal EEG sinyal özelliklerini çıkarmışlardır. Levenberg–Marquardt, Bayesian Düzenlenmesi ve Ölçekli Konjugat Gradyan yöntemlerine çıkarılan özellikler uygulanarak modellenmiştir. Levenberg–Marquardt yöntemi ile %83,01 doğruluk oranında sınıflandırma başarımı elde etmişlerdir.

Abhiyana ve diğ. [15], 50 bireye ait fokal, 50 bireye ait fokal–olmayan bölgeden elde edilmiş EEG sinyalinin yaklaşık entropi, örnek entropi, yapay entropi, permütasyon entropi ve çok ölçekli entropi karakterizasyonunu t–test analizi p–değeri $< 0,05$ için hesaplamışlardır. Sıralı İleri Özellik Seçimi (Sequential Forward Feature Selection – SFFS) yöntemine göre seçilen en önemli özellikler Radyal Tabanlı Fonksiyon Destek Vektör Makinası yöntemlerine uygulanarak normal ve fokal sinyal sınıflandırması için %82,00 doğruluk oranı elde etmişlerdir.

Rajaguru ve diğ. [16], 20 epilepsi hastasına ait EEG verisinden elde edilen 25.600 örnek sinyale Hilbert Dönüşümü yöntemi uygulayarak özellik uzayını indirgemişlerdir. Elman Geri Yayımlı ve ÇKA yöntemlerine indirgenmiş özellik uzayı uygulanarak epileptik risk analizi sırasıyla %100 ve %97,91 sınıflandırma doğruluğu ile hesaplamışlardır.

Dohler ve diğ. [17], 144 epilepsi hastası bireye ait 3 boyutlu T1–ağırlıklı MRG veri setini Mıknatıslanma ÇKA, Hazırlanmış Hızlı Toplama Gradient Eko (Magnetization Prepared Rapid Acquisition Gradient Echo – MPRAGE) dizilimi ile 256×256 boyutlu veri matrislerine dönüştürerek hesaplamışlardır. Hücrel sinir ağlarına (Cellular Neural Networks – CNN) hesaplanan veri seti uygulanarak temporal lob epilepsisini %68 doğruluk oranı ile sınıflandırmışlardır.

Duque–Muñoz ve diğ. [18], 40 epilepsi hastası bireye ait EEG verisine Kısa Zamanlı Fourier Dönüşümü (Short Time Fourier Transform – STFT) uygulayarak sinyal verisine ait stokastik değişkenleri hesaplamışlardır. STFT yöntemi ile elde edilen durağan olmayan EEG davranışlarını normal, iktal ve interiktal aktivitelerini Destek Vektör Makinası (DVM) ve k–en yakın komşu yöntemleri ile sırasıyla %95,78 ve %96,58 doğruluk oranlarıyla sınıflandırmışlardır.

Gandhi ve diğ. [19], 300 normal, 300 epilepsi hastasına ait EEG verisine Ayrık Dalgacık Dönüşümü (Discrete Wavelet Transform – DWT) uygulayarak sinyale ait enerji, entropi, standart sapma, ortalama, basıklık, çarpıklık değerlerini hesaplayarak ayrılma ağacı elde etmişlerdir. Ağaçtaki her bir düğüme Harmoni Arama Yöntemi uygulayarak yeni özellik uzayı elde etmişlerdir. Olasılıksal Sinir Ağları yöntemine elde edilen yeni özellik uzayı uygulanarak normal ve epileptik hasta verilerini %100 doğruluk oranı ile sınıflandırmışlardır.

Dhiman ve diğ. [20], normal ve epilepsi hastası bireylere ait toplam 100 EEG segmentine ayrık dalgacık dönüşümü (Discrete Wavelet Transform – DWT) uygulayarak sinyale ait enerji,

entropi, standart sapma, ortalama, basıklık, çarpıklık değerlerini hesaplayarak optimal özellikleri genetik yöntemler ile seçerek hibrit yaklaşımda bulunmuşlardır. DVM yöntemine optimal özellik uzayını uygulayarak 10–kat çapraz doğrulama yöntemi ile %100 doğrulukla sınıflandırmışlardır.

Oğulata ve diğ. [21], 79 epilepsi hastasına ait EEG, yaş, cinsiyet, bilinç kaybı, nöbet zamanı bilgilerine dayanarak elde edilen veri setinden tonik klonik, absans, myoklinik nöbetleri olmak üzere epileptik nöbetleri sınıflandırmışlardır. ÇKA yöntemine uygulanan veri seti için epileptik nöbetleri sırasıyla %78,50, %80,00, %50,00, %91,60 olarak sınıflandırmışlardır.

Kamath ve diğ. [22], 3000 örnek pencereleme ile oluşan veri setinin zaman alanındaki hızlı frekans değişimlerini Teager Enerji ile hesaplayarak epileptik nöbetlerin teşhisinde yeni bir yaklaşımda bulunmuşlardır. Hasta ve sağlıklı bireylerin sınıflandırılmasında sinyale ait Z, O, N, F, S frekanslarından oluşan ZONF, S frekanslarını %97,00, NF,S frekanslarını %98,50, Z,S frekanslarını %100, ZNF,S frekanslarını %98,90 ile hesaplamışlardır.

Li ve diğ. [23], EEG sinyaline ait S, Z, N, F segmentlerine Zaman – Frekans Analizi ve Kalman filtresi metotlarını uygulayarak yüksek çözünürlükte spektral çıkarım yapmışlardır. Fourier dönüşümü ve Çok Ölçekli Radyal Tabanlı Fonksiyon metodolojilerini Parçacık Sürüşü Optimizasyonu ile birleştirerek sağlıklı ve epilepsi hastalarından zamandan bağımsız spektral çıkarımda bulunarak TBA ile özellik uzayını elde etmişlerdir. DVM yöntemine yeni özellik uzayı uygulanarak S–F & S–Z EEG segmentlerinin sınıflandırılmasında %100, S–N & S–NF EEG segmenti için $98,73 \pm 0,43$, S–Z EEG segmenti için %100, S–N EEG segmenti için %99,80, S–F EEG segmenti için %97,60, S ve N–F EEG segmenti için %98,73 doğruluk oranı elde etmişlerdir.

Übeyli [24], normal ve kısmi epilepsi tanımlamasını MIT–BIH heterojen veri tabanındaki elektrokardiyogram sinyalleri ile yapmıştır. DVM yöntemine db1 dalgacık katsayısı ile elde edilen özellikleri uygulayarak optimum sınıflandırma doğruluğunu %99,44 olarak elde edilmiştir.

Adam ve diğ. [25], EEG sinyal verisine Zirve Yakalama Yöntemi uygulayarak elde edilen tepe–sinyallerinden özellik uzayını çıkarmışlardır. Elde edilen özellik uzayı kural tabanlı sınıflandırıcı ile eğitilmiştir. Test işlemi için, filtrelenmiş EEG sinyalinden elde edilen özellikler eğitim aşamasında oluşturulan kural tabanlı ağaca uygulanarak optimum zirve noktasını

Geometrik Ortalama Metoduna göre belirlemişlerdir. Geometrik Ortalama Metoduna göre eğitim işleminin sonucunda %99,90, test işleminin sonucunda %98,58 doğruluk elde etmişlerdir.

Zainuddin ve diğ. [26], 4097 EEG sinyal noktasına K–Ortalamlar ve Yapay C–Ortalamlar yöntemlerini uygulayarak sinyaldeki ara düğümlerin karar verilmesini sağlamışlardır. Ara düğümlere tip–2 Yapay C–Ortalamlar ve Harmoni Arama yöntemlerini bir arada uygulayarak hibrit yöntem önermişlerdir. Çalışmalarının sonucunda epileptik nöbet ve sağlıklı bireylere ait sınıflandırma doğruluğunu $99,15 \pm 0.13$ elde etmişlerdir.

Sharif ve diğ. [27], 19 epilepsi hastasına ait EEG zaman serisine Poincaré metodu uygulamayı önermişlerdir. Bu metoda göre, her kanal için Genetik algoritma ve Shannon entropi hesabı ile 64 yapay kurallar tanımlayarak öznitelik uzayı elde etmişlerdir. Yapay kurallardan varyans hesabı yapmışlardır. Varyans hesabı ile iktal nöbet özellik setini elde etmişlerdir. Elde edilen özellik setine TBA yöntemini uygulayarak özellik seçimi yapmışlardır. Gauss RBF–SVM yöntemine elde edilen yeni özellik uzayını uygulayarak %91,80 duyarlılık elde etmişlerdir.

Wang ve diğ. [28], A, B normal (200 EEG segment), C, D interiktal (200 EEG segment), E iktal (100 EEG segment) epileptik nöbet verisi için kural tabanlı otomatik hastalık tespit sistemi geliştirmişlerdir. EEG segment verisine Durağan Sinyal Fourier Dönüşümü uygulayarak özellik çıkarımı yapmışlardır. Elde edilen özellikler ile C4.5, Rastgele Orman, DVM + C4.5 ve DVM + Rastgele Orman yöntemleri için ayrı ayrı kurallar çıkarmışlardır. DVM + Rastgele Orman yöntemi ile çıkarılan kurallar %98,96 ikili segment için, %96,00 üçlü segment için, %82,60 beşli segment için sınıflandırma doğruluğu elde etmişlerdir.

Martinez–del–Rincon ve diğ. [29], BONN, HRUCH, FREIBURG EEG sinyal veri setlerini normal ve epileptik olarak sınıflandırmışlardır. Özellik çıkarımında Sözcük Çantası (Bag of Words – BOW) modelinden esinlenerek sinyaldeki gürültü, sinyal zayıflaması ve artefakt varlığını sinyalde çok sık geçen veriyi hesaplayarak elde etmişlerdir. Verideki dalgacık katsayılarını hesaplayarak istatistiksel veriler elde etmişlerdir. A, E EEG segmentlerinin sınıflandırılmasında BOW + Polinomsal DVM yöntemi ile %99,84 sınıflandırma doğruluğu elde etmişlerdir.

Bayrak ve diğ. [30], bireylerin EEG sinyal verisini kullanarak epileptik nöbet tespiti için yeni bir klinik karar destek sistemi geliştirmişlerdir. Zaman–frekans alanından çıkarılan filtrelenmiş

özellikler hesaplanmıştır. Ayırık kosinüs dönüşümü, kesirli Fourier dönüşümü ve Hilbert dönüşümü yöntemleri ile elde edilen zaman–frekans özelliklerinin sınıflandırma doğruluklarını karşılaştırmışlardır. Momentumlu gradyan inişi, ölçeklendirilmiş eşlenik gradyan ve uyarlanabilir öğrenme oranı yöntemlerine 3 farklı özellik seti (DCT_EEG, FrFT_EEG ve HT_EEG) uygulanmıştır. Hilbert dönüşümü yöntemiyle elde edilen EEG sinyal özellik setine sınıflandırma doğruluğu ölçeklendirilmiş eşlenik gradyan yöntemi uygulandığında en başarılı sınıflandırma doğruluğunu % 87,38 olarak elde etmişlerdir.

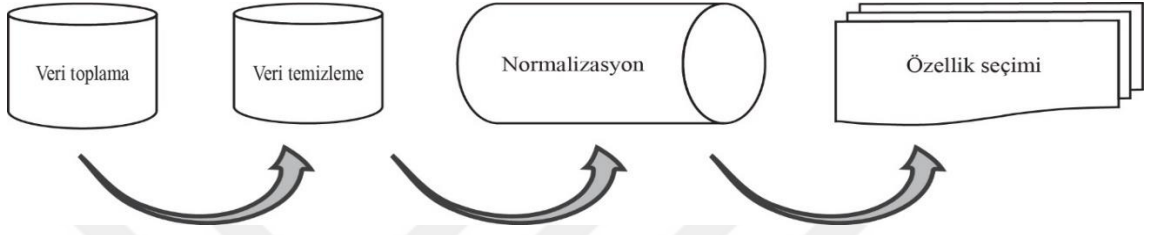
Yağmur Demirezen ve Sertbaş [31], ayırık kosinüs dönüşümü yönteminden elde edilen 1–D EEG verileri ile bu EEG verileri için ortalama, varyans, standart sapma, basıklık, çarpıklık gibi istatistiksel yöntemleri hesaplayarak veri seti elde etmişlerdir. Elde ettikleri veri setine PCA, LDA, İleri Seçim ve Geriye Doğru Seçim yöntemleri uygulamışlardır. Bağımsız Bileşen Analizi ile artefaktalardan arındırılan EEG veri setinin ÇKA yöntemi ile epilepsi ve sağlıklı sınıflamasından % 96,00 sınıflandırma başarımları elde etmişlerdir.

Literatürde son yıllara ait ulusal ve uluslararası yapılmış çalışmalarda, epilepsi hastalığının sınıflandırılmasında geliştirilmiş matematiksel yöntemlerden tespit edilen durumlar sırasıyla:

1. Sinyal işleme metotlarına EEG sinyal verisi uygulanarak sinyalden hastalıkla ilgili özellikler çıkarılmış,
2. Genellikle hazır EEG veri tabanlarındaki EEG sinyali ile uygulama geliştirilmiş,
3. Az sayıda çalışmada görüntü işleme tekniği MRG verisine uygulanarak lezyon varlığı araştırılmış ve geliştirilen matematiksel yöntemlerin genellikle tek bir modeli ele aldığı görülmüştür.

3. MALZEME VE YÖNTEM

Bir amaç için bir arada bulunan veri, SQL, Postgre SQL, Oracle gibi veri tabanlarında organize veri toplulukları halinde saklanırlar [32]. Veri tabanında saklanan verilerin anlamlı bir bütünlük oluşturabilmesi ve bu veriden çıkarımlarda bulunulabilmesi için veri işleme adımları Şekil 3.1’de gösterilmiştir.



Şekil 3.1: Veri işleme adımları.

3.1. VERİ ÜZERİNDE UYGULANAN İŞLEMLER

Yönteme uygulanacak verinin uygun hale getirilmesi için yeniden yapılandırılması gerekmektedir. Yeniden yapılandırma işlemleri veri toplama, veri temizleme, normalizasyon ve özellik seçimi işlemlerinden oluşmaktadır [33].

Veri işlemede ilk adım verinin toplanmasıdır. Belirli bir alana ait problemin çözümü için kaliteli verinin sistematik olarak bir araya toplanmasıdır. Böylece, toplanmış kaliteli verinin istatistiksel analizlerini yaparak ilgili alanda üretilen sorulara doğru ve güvenilir sonuçların alınmasını sağlamaktadır [34].

Veri toplama işlemi ile toplanan veride, verinin istenilen özelliklere sahip olmaması durumunda verinin temizlenmesi gerekmektedir. Veri temizliği için çeşitli teknikler vardır. Eksik değer içeren kayıtlar ise veri setinden çıkarılmaktadır ya da yerine genel bir sabit değer kullanılmaktadır. İlgili veri değişkeninin bulunduğu sütundaki sayısal değerlerin ortalaması alınarak eksik veri için bu değer atanmaktadır. Ya da ilgili veri değişkeninin sahip olduğu sayısal değerlerin tamamı yerine bir sınıfa ait örnek verilerin ortalama değeri hesaplanarak eksik değer yerine atanabilmektedir. Eksik veri için regresyon tahmini yapılabilmektedir [35].

Dönüşüm (Normalizasyon), veri tabanından seçilen verinin 0 – 1 gibi aralıklara indirgenmesi işlemidir. Normalizasyon işlemleri için min–maks normalizasyon, sıfır ortalama normalizasyon, ondalıklı normalizasyon yöntemleri vardır. Bu tez çalışmasında, veriyi doğrusal şekilde normalize eden min–maks normalizasyon yöntemi kullanılmıştır [36].

min bir verinin alabileceği en düşük değerdir, *maks* bir verinin alabileceği en büyük değerdir. Bir verinin 0 – 1 aralığında normalizasyonu Denklem 3.1’de verilmiştir.

$$a' = \frac{a - \min}{\max - \min} \quad (3.1)$$

Denklem 3.1’e göre a' verinin normalize edilmiş değeri, a ise verinin orijinal değeridir.

Özellik Seçimi (boyut indirgeme), verideki özellik sayısının azaltılmasıdır. Yöntemin yapısından dolayı ve elde edilecek sonuçların hassasiyetinden dolayı belirli özelliklerin birleştirilerek tek bir özellik olarak işlem görmesi gerekebilmektedir. Bu durumda dalgacık dönüşümü veya TBA yöntemleri birleştirme işlemi olarak kullanılmaktadır [37]. Bir özellik kümesinin varyans–kovaryans yapısını ve bu özelliklerin doğrusal birleşimleri aracılığıyla boyut indirgemesini sağlayan çok değişkenli istatistik yöntemi olan TBA yöntemi kullanılmıştır. Bu yöntemde l adet değişken doğrusal, ortogonal ve birbirinden bağımsız özellikler içeren k adet ($k \leq l$) değişkene dönüştürülmektedir [38]. TBA işlemi için (a, b) gibi iki veri arasındaki değişimin kovaryans değeri Denklem 3.2’deki gibi hesaplanmaktadır.

$$\text{cov}_{a,b} = \frac{\sum_{i=1}^n (a_i - \bar{a})(b_i - \bar{b})}{n-1} \quad (3.2)$$

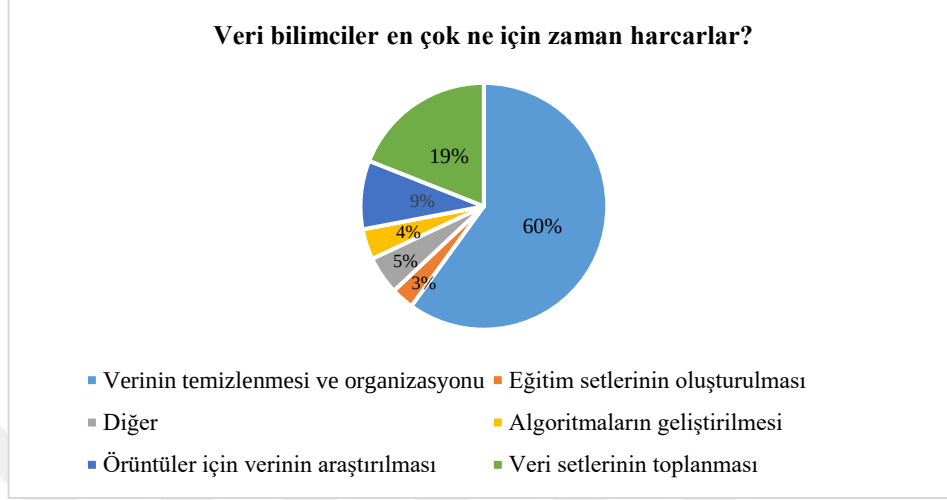
(x, y) için kovaryans matrisi Denklem 3.3’teki gibi hesaplanmaktadır.

$$k_{\text{cov}(a,b)} = \begin{bmatrix} \text{cov}_{a,a} & \text{cov}_{b,a} \\ \text{cov}_{a,b} & \text{cov}_{b,b} \end{bmatrix} \quad (3.3)$$

Matrisler aynı boyuttaysa çarpılarak öz değer ve öz vektörlerin hesabı yapılarak indirgeme özellik vektörü seçilip indirgeme çarpımı yapılmaktadır.

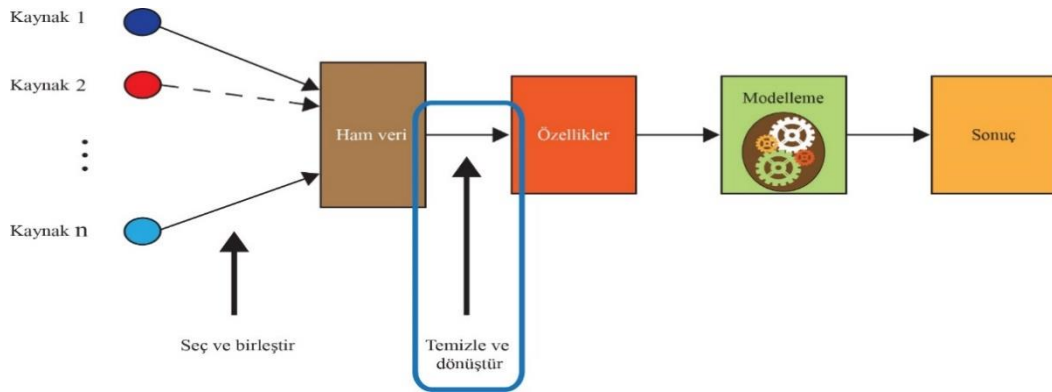
Özellik mühendisliği, öğrenme yönteminin tahmine dayalı gücünü artırmak için ham veriden yeni özellik setlerinin oluşturulmasıdır. Farklı kaynaklardan elde edilmiş verideki anlamlı özelliklerin seçilmesi sınıflandırma sonucunu etkilemektedir [39]. Veri hazırlama aşaması

zahmetli bir süreçtir. Bu konuda Forbes'te yapılmış bir ankete göre veri hazırlama işleminin veri bilimcilerin zamanlarının %79'unu aldığı Şekil 3.2'de gösterilmektedir [40].



Şekil 3.2: Forbes'te yapılmış bir ankete göre veri hazırlamada geçen zaman.

Makine öğrenimi yöntemi gereksinimleriyle uyumlu, uygun girdi veri setini hazırlamak ve makine öğrenimi modellerinin performansını iyileştirmek için özellik mühendisliği adımlarının uygulanması gerekmektedir. Imputasyon, aykırı değerlerin kullanımı, bölme, log dönüşümü, tek sıcak kodlama, gruplama işlemleri, özellik bölme, ölçekleme özellik mühendisliğinde kullanılan tekniklerdir. Özellik mühendisliği için uygulama adımları Şekil 3.3'te gösterilmiştir.



Şekil 3.3: Özellik mühendisliği uygulama adımları.

Şekil 3.3'e göre, farklı kaynaklardan elde edilen ham verinin temizlenmesi ve modelleme için uygun hale dönüştürülmesi gerekmektedir. Bu işlemlerden sonra modellemede kullanılacak özellikler seçilerek sınıflandırma işlemi yapılmaktadır [41].

3.2. YÖNTEM

Öğreticili öğrenme, eğitim ve test veri setleri ile oluşturulmaktadır. Eğitim veri setinde her bir kaydın özellikleri ve sınıf bilgisi vardır. Öğreticili öğrenmede, verinin hangi sınıfa ait olduğu bilgisi bulunan bir model oluşturulmaktadır. Eğitim veri setinde yer almayan veriler test veri seti olarak modellenmektedir ve modelin başarısı test veri setine göre değerlendirilmektedir. Öğreticili öğrenme işleminin matematiksel formülü Denklem 3.4'teki gibidir.

$$D = \{x_1, x_2, \dots, x_n\}; Y = \{y_1, y_2, \dots, y_m\} \quad (3.4)$$

Denklem 3.4'e göre, D bir veri tabanını ve x_i veri tabanındaki bir kayıt bilgisini, Y ise m adet sınıftan oluşan sınıf kümesini temsil etmektedir. Böylece, $f: D \rightarrow Y$ ve her bir x_i 'nin bir sınıfa ait olması gerekmektedir. Denklem 3.5'e göre her bir Y_j ayrı bir sınıftır ve her bir sınıf kendisine ait veri kayıtlarını içermektedir [42, 43].

$$Y_j = \{x_i | f(x_i) = Y_j, 1 \leq i \leq n, 1 \leq j \leq m \text{ ve } x_i \in D\} \quad (3.5)$$

3.2.1. ID3 Yöntemi

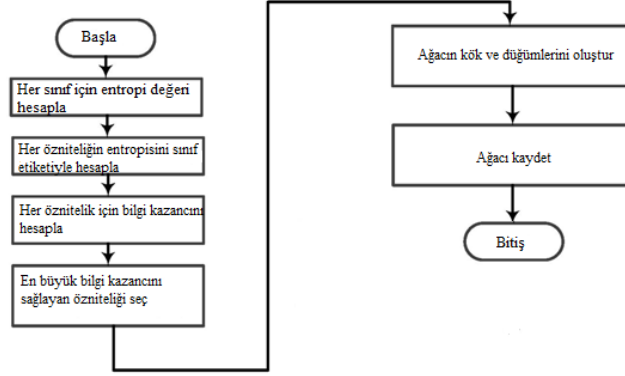
ID3 yöntemi, veri setindeki en ayırt edici özelliği bulmak için Denklem 3.6'daki entropi kavramını kullanmaktadır. Entropi beklentisizliğin maksimum değeridir ve $[0 - 1]$ arasında değer almaktadır [44, 45].

$$H(p_1, p_2, \dots, p_n) = -\sum (p_i \log(1/p_i)) \quad (3.6)$$

Denklem 3.6'ya göre, $\langle p_1, p_2, \dots, p_n \rangle$ olasılıkları ifade ederken, olasılıkların toplamı 1'dir. ID3 yöntemi, doğru sınıflandırma yapabilmek için veri setinin bölünme öncesindeki ve bölünme sonrasındaki fark değerini kullanmaktadır. Bu fark hangi alt bölüm için büyükse öncelikli düğüm ve dallanma hesaplanan bu büyük değere göre belirlenmektedir. Aradaki bu fark kazanımdır (S) ve Denklem 3.7'ye göre hesaplanmaktadır.

$$Kazanım(D; S) = H(D) - \sum_{i=1}^n P(D_i)H(D_i) \quad (3.7)$$

ID3 yönteminin genel akış diyagramı, Şekil 3.4'teki gibidir.



Şekil 3.4: ID3 yöntemi akış diyagramı.

Şekil 3.4’e göre, ID3 yönteminin işlem adımları ile karar ağacı ve karar kuralları elde edilmektedir.

3.2.2. Doğrusal Regresyon Yöntemi

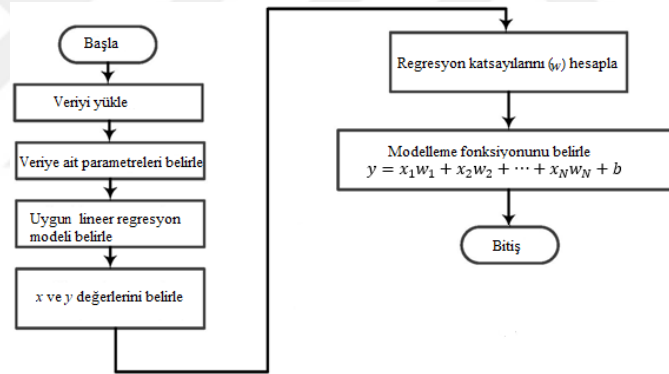
İki ya da daha çok bağımlı x değişkenlerinin birbiri ile doğrusal ilişkisinin fonksiyonel bilgisini y , tahminleyen istatistiksel bir yöntemdir [46 – 48]. Bu yöntem ile x değişkenlerinden birinin bile değeri bilindiğinde diğer x değişkenleri için tahminleme yapılabilmektedir. Değişkenler arasındaki ilişkinin eğiliminin hesaplanması için eğri uydurulmaktadır. Veriler arasındaki eğilimin uydurulacağı eğri için en küçük kareler metodu kullanılmaktadır. En küçük kareler metodu, her bir veri noktası için hatanın en küçük olduğu doğru ile ifade edilmektedir. Çıktı değeri ile hedef değeri arasındaki karesel farkların toplamı Denklem 3.8 ile hesaplanmaktadır. Denklem 3.8’e göre, E ile hata (maliyet) değeri, y_i ile çıkış değerleri, ve θ_0, θ_1 ile ağırlık değerleri hesaplanmaktadır.

$$E(\theta_0, \theta_1) = \frac{1}{2N} \sum_{i=1}^N (h_i(x) - y_i)^2 \quad (3.8)$$

Doğrusal regresyon yöntemiyle modellenen her bir problem için hata değerinin minimum olması beklenmektedir. Fonksiyonun gradyan azalma yöntemi ile lokal minimumuna doğru iteratif olarak yakınsama yapmaktadır. Denklem 3.9’da, fonksiyonun lokal minimumunu hesaplamak için birinci dereceden türevi hesaplanır ve gradyan azalma yöntemi kullanılmaktadır.

$$y = x_1 w_1 + x_2 w_2 + \dots + x_N w_N + b \quad (3.9)$$

Fonksiyonda hesaplanan yakınsama yönü gradyan doğrultusunda ters yönde bir doğru ile öğrenme katsayısı ile çarpılarak bir sonraki yeni ağırlık katsayısı hesaplanmaktadır. En uygun eğrinin uydurulması için ağırlıkların optimum değerleri elde edilene kadar eğitim işleminin devam ettirilmesi gerekmektedir. Doğrusal regresyon ile istatistiksel olarak anlamlı bir farkın olup olmadığına karar verilmesi durumunda, olası hata miktarı p-değeri ile hesaplanmaktadır [49]. p-değeri için hatanın maksimum kabul edilebilir düzeyi 0,05 olarak kabul görmüştür. Değişkenlerin regresyon analizi sonucunda bulunan p-değeri 0,05'in altında bir değer ise karşılaştırma sonucunda anlamlı farklılık bulunduğu anlamına gelmektedir. Aynı zamanda, doğrusal regresyon modelinin genel önemini test edebilmek için F değeri kullanılmaktadır. F değeri, karelerin ortalama regresyon toplamının, karelerin ortalama hata toplamına bölünmesiyle elde edilen oran ile hesaplanmaktadır. Değeri sıfırdan rastgele büyük bir sayıya kadar değişmektedir. Doğrusal regresyon yönteminin genel akış diyagramı Şekil 3.5'te verilmiştir.



Şekil 3.5: Doğrusal regresyon yöntemi akış diyagramı.

Şekil 3.5'e göre, doğrusal regresyon yönteminin işlem adımları ile değişkenler arası farklılık hesabı yapılmaktadır.

3.2.3. Naive Bayes Yöntemi

Sınıfı belli olmayan yeni bir verinin mevcut sınıflardan herhangi birine girme olasılığını hesaplayan istatistiksel bir yöntemdir. İki sınıflı bir sınıflandırma probleminde Bayes teoremi Denklem 3.10'daki gibi hesaplanmaktadır [43, 50].

$$P(C_1|x_i) = \frac{P(x_i|C_1)P(C_1)}{P(x_i|C_1)P(C_1)+P(x_i|C_2)P(C_2)} \quad (3.10)$$

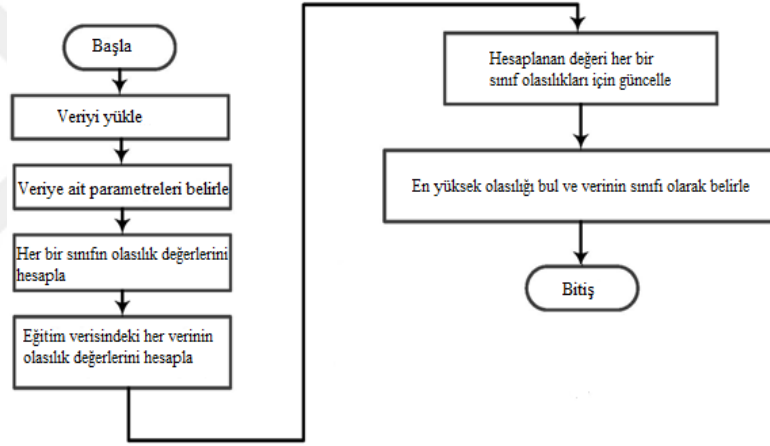
Denklem 3.10'a göre, C_1 ve C_2 iki ayrı sınıfı temsil etmektedir. $P(C_1|x_i)$, x_i değerinin C_1 sınıfına ait olma olasılığını hesaplamaktadır. $P(x_i)$, x_i değerinin veri tabanında bulunma sayısıdır. $P(C_1)$ ve $P(C_2)$ değerleri C_1 ve C_2 sınıflarının veri tabanında bulunma sayısıdır. m sınıflı bir problem için Bayes teoremi Denklem 3.11'deki gibidir.

$$P(x_i) = \sum_{j=1}^m P(x_i|C_j)P(C_j) \quad (3.11)$$

x_i 'nin C_1 sınıfına ait olma olasılığı Denklem 3.12'deki gibi hesaplanmaktadır [51].

$$P(C_1|x_i) = \frac{P(x_i|C_1)P(C_1)}{P(x_i)} \quad (3.12)$$

Naive Bayes yönteminin genel akış diyagramı, Şekil 3.6'daki gibidir.

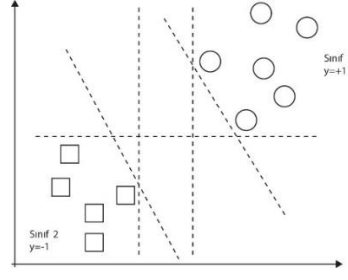


Şekil 3.6: Naive Bayes yöntemi akış diyagramı.

Şekil 3.6'ya göre, Naive Bayes yönteminin işlem adımları ile veri setindeki kayıtların ait oldukları sınıfla ilgili olasılık değerleri hesaplanmaktadır.

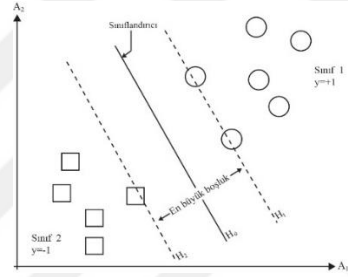
3.2.4. DVM Yöntemi

Verinin doğrusal veya doğrusal olmayan bir fonksiyon yardımıyla en uygun fonksiyonu tahmin etmesi esasına dayanarak sınıflandırılmasıdır. Doğrusal olmayan problemler için çekirdek fonksiyonları ile hesaplama yapılmaktadır. Çalışmamız, iki sınıflı sınıflandırma problemi olduğu için matematiksel işlemler 2-sınıflı sınıflandırmaya göre verilmiştir. n elemanlı D veri seti $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ şeklindedir ve $y \in \{+1, -1\}$ olarak kabul edilmektedir. Şekil 3.7'ye göre, verinin sınıflandırılması için hiper düzlemler arasındaki boşluğu en büyük olanı seçmek gerekmektedir [52, 53].



Şekil 3.7: İki boyutlu uzayda doğrusal ayrılabilen veri seti.

Şekil 3.8'e göre H_1 ve H_2 hiper düzlemlerinin orta değerini H_0 oluşturmaktadır. H_0 optimal ayırma hiper düzlemi olarak iki sınıfı birbirinden ayıran doğrusal hiper düzlemdir. H_0 , Denklem 3.13'teki gibi ifade edilebilmektedir [43].



Şekil 3.8: Doğrusal ayrılabilen sınıflar arasındaki muhtemel en büyük boşluk.

$$H_0 = \sum_{i=1}^n w_i x_i + b = 0 \quad (3.13)$$

Denklem 3.13'e göre, n öznitelik sayısını, $W = \{w_1, w_2, \dots, w_n\}$ ağırlık vektörünü göstermektedir. b sabit bir sayıyı temsil etmektedir. İki niteliğin var olduğu varsayılırsa $X = \{x_1, x_2\}$ değerleri için H_1 ve H_2 hiper düzlemleri Denklem 3.14 ve Denklem 3.15'teki gibi ifade edilmektedir.

$$H_1: W^T X + b = 1 \quad (3.14)$$

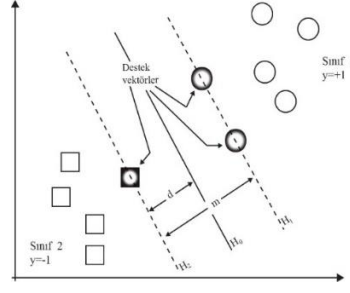
$$H_2: W^T X + b = -1 \quad (3.15)$$

Şekil 3.8'deki hiper düzlemin üst tarafındaki noktalar Denklem 3.16'ya göre hesaplanmaktadır.

$$W^T X + b > 0, y_1 = +1 \quad (3.16)$$

Şekil 3.9'daki hiper düzlemin alt tarafındaki noktalar Denklem 3.17'ye göre hesaplanmaktadır.

$$W^T X + b < 0, y_2 = -1 \quad (3.17)$$

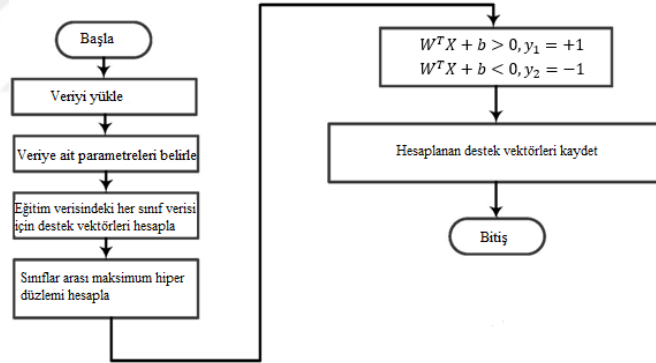


Şekil 3.9: Doğrusal ayrılabilen sınıflar için hesaplanan destek vektörleri.

H_1 ve H_2 hiper düzlemleri üzerinde yer alan gözlemler destek vektörleridir. z , hiper düzlem üzerindeki bir noktayı temsil etmektedir ve bir destek vektörüyle H_0 arasındaki uzaklık Denklem 3.18'deki gibi hesaplanmaktadır.

$$d = \frac{|wX_z' + b|}{\|w\|} \quad (3.18)$$

DVM yöntemi genel akış diyagramı, Şekil 3.10'daki gibidir.



Şekil 3.10: DVM yöntemi akış diyagramı.

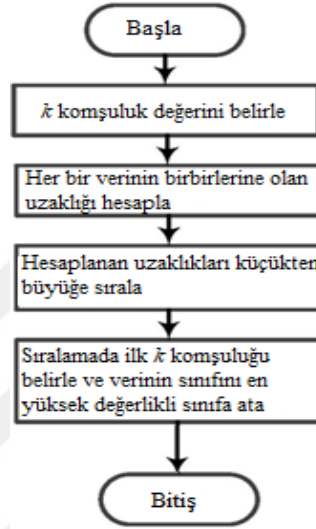
Şekil 3.10'a göre, DVM yönteminin işlem adımları ile sınıflar arası en uygun hiper düzlem hesaplanmaktadır.

3.2.5. k-NN Yöntemi

Sınıfları belli olan veri tabanından yararlanarak, yeni bir verinin hangi sınıfa ait olduğunu uzaklık hesabına göre bulan sınıflandırıcıdır. En küçük Öklid uzaklığına sahip k sayıdaki veri için seçim yapılabilmektedir. Öklid uzaklık formülü Denklem 3.19'a göre ifade edilmektedir [43].

$$d(i, j) = \sqrt{\sum_{k=1}^n (x_{ik} - x_{jk})^2} \quad (3.19)$$

Yöntemde k en yakın komşuluk parametresi önceden belirlenmektedir. Sınıfı bulunmak istenen verinin tüm verilere uzaklığı hesaplanmaktadır ve sıralanmaktadır. En küçük k tane veri seçilmektedir. En sık tekrarlanan kategori değeri, yeni verinin sınıfı olarak hesaplanmaktadır [54, 55]. k -NN yöntemi genel akış diyagramı, Şekil 3.11’de verildiği gibidir.

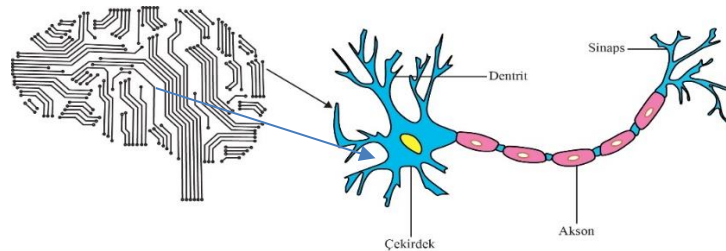


Şekil 3.11: k -NN yönteminin akış diyagramı.

Şekil 3.11’e göre, k -NN yönteminin işlem adımları ile veri setindeki kayıtların en yakın k komşuluğundaki değeri hesaplanmaktadır.

3.2.6. Yapay Sinir Ağları (YSA) Yöntemi

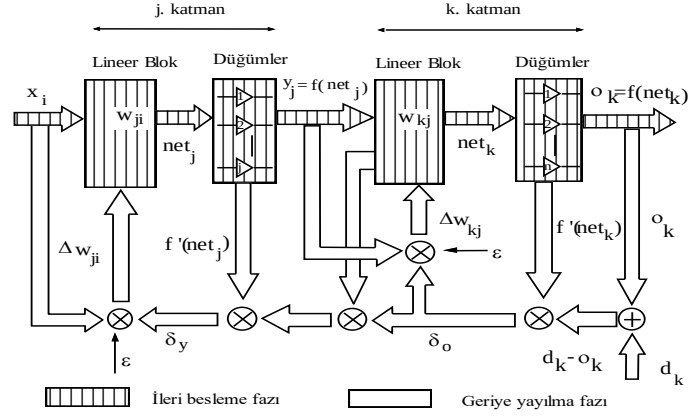
İnsan beyinde bilgi ve sinyaller nöronlar arasındaki akson bağlantılarıyla aktarılmaktadır [56]. YSA yöntemi, Şekil 3.12’deki gibi insan beyninin nörolojik yapısından esinlenilerek geliştirilmiştir. İnsan beynine ait nöronlar sinyal gönderimini nöronlar arasındaki bağlantı ile yapabilmektedir.



Şekil 3.12: İnsan beynine ait sinir hücresinin genel yapısı.

YSA gösterilen bir örnekten öğrenerek kendi kurallarını oluşturmaktadır. Öğrenme; giriş örneklerine veya bu girişlerin çıkışlarına bağlı olarak ağırlık bağlantı ağırlıklarını değiştiren veya ayarlayan öğrenme kuralı ile gerçekleştirilmektedir. Ağ, ağırlık bileşenindeki değerler değiştirilerek örneklere göre eğitilmektedir ve herhangi bir örnek ile yeniden karşılaştırıldığında uygun cevabı üretmektedir. Öğrenmede iyi bir model kullanıp, ağırlıkların bu modele göre değiştirilmesi esastır. Bu ağırlıkların değiştirilmesi belli bir hata oranına kadar tekrar edilmektedir. Son olarak ağırlık değerleri test etmede kullanılmak üzere kaydedilmektedir. Çok katmanlı algılayıcı, giriş ve çıkış katmanları arasında birden fazla katmanın kullanıldığı ileri yayımlı YSA sistemidir. Gizli katman olarak isimlendirilen bu ara katmanlarda, düğümleri giriş ve çıkış katmanlarına doğrudan bağlı olmayan işlem elemanları vardır.

Çok katmanlı ağlarda, veriler giriş katmanı tarafından kabul edilmektedirler. Ağ içinde yapılan işlemler sonucunda çıkış katmanında oluşan sonuç değer, istenen cevap ile karşılaştırılmaktadır. Bulunan cevap ile istenen cevap arasında herhangi bir ayrılık varsa, ağırlıklar bu farkı azaltacak şekilde yeniden düzenlenmektedir. Girişteki değer, ağırlıklar uygun noktaya ulaşana kadar değişmemektedir. Hesaplanan çıkışlar, istenilen cevaplarla karşılaştırılarak sonuçta gerekirse hatanın belirtilmesi gerekmektedir. Hata işareti gizli işlem elemanlarından çıkış katmanına olan ağırlıkları değiştirmekte kullanılmaktadır. Giriş katmanından gizli katmana gelen değer değiştirilip değiştirilmediğinin göz önünde bulundurulması gerekmektedir. Gizli birimlerden ne tür bir çıkış istendiği bilinemeyeceği için gizli birimlerin çıkışında hata işareti verilmesi kolay değildir. Bunun yerine her birimin çıkış biriminin hatalarına olan etkisi bilinmelidir. Bu ise hatalı birim için gizli birime bağlı olan çıkış birimlerinin hata işaretlerinin toplamı alınarak yapılmaktadır [43, 57]. Çok sayıda gizli katmana sahip sistemlerde, her sistemin hata işaretleri, bir önceki katmanın düzeltilmiş işlemlerinden çıkartılarak işlem tekrarlanmaktadır. Sonuç olarak, ağırlık düzeltme işlemi çıkış seviyesine bağlı ağırlıklardan başlar ve işlem ters yönde, giriş seviyesine varana kadar devam etmektedir. Sonuçta sistem hatalar yapar, ama bu ağdaki hatalardan bir şeyler öğrenip isteneni bulana kadar işleme devam etmektedir. Bu yönteme hatanın geriye yayılması yöntemi denilmektedir ve Şekil 3.13'te genel hatanın geriye yayılma eğitme yöntemi (Feed Forward Back Propagation – FFBP) genel blok diyagramı verilmiştir [58].



Şekil 3.13: Hatanın geriye yayılma yöntemi genel blok diyagramı.

Hatanın geriye yayılma yöntemi, çok katmanlı, ileri yayılımlı bir öğrenmeden elde edilen çıkışlar ile eldeki hedef çıkışlar arasındaki hataların karesinin ortalamasını minimum yapmak için geliştirilmiş iteratif gradyan bir yöntemdir ve eğitme işlemi için genelleştirilmiş delta kuralını kullanmaktadır. Şekil 3.13'te gösterilen şekilde yöntem, ana hatlarıyla şöyledir: Ağ mimarisi tanımlanır ve ağırlıklar bazı rastgele küçük sayılar ile başlatılarak, ağa ilk giriş sunulur. Burada n -boyutlu giriş örnekleri girildiğinde; $x_i = [x_1, x_2, \dots, x_n]^T$ benzer şekilde istenilen m -boyutlu çıkış örneklerini ise; $d_k = [d_1, d_2, \dots, d_m]^T$ belirtmektedir. x_i değerleri i katmanındaki nöronların çıkış değerleri ise, j katmanındaki bir nörona gelecek olan toplam giriş değeri Denklem 3.20'deki gibi hesaplanmaktadır.

$$net_j = \sum_{i=1}^n w_{ji} x_i \quad (3.20)$$

Gizli katmandaki j nöronunun transfer fonksiyonu çıkış değeri Denklem 3.21'deki gibi hesaplanmaktadır.

$$y_j = f_j(net_j) \quad (3.21)$$

Çıkış katmanındaki k nöronuna gelecek olan toplam giriş ise Denklem 3.22'deki gibi hesaplanmaktadır.

$$net_k = \sum_{j=1}^J w_{kj} y_j \quad (3.22)$$

Çıkış katmanındaki bir k nöronun lineer olmayan çıkışı Denklem 3.23'teki gibi hesaplanmaktadır.

$$o_k = f_k(net_k); k = 1, 2, \dots, m \quad (3.23)$$

Ağdan elde edilen çıkış değeri, gerçek çıkış değeri ile karşılaştırılmakta ve e_k hatası Denklem 3.24'teki gibi hesaplanmaktadır.

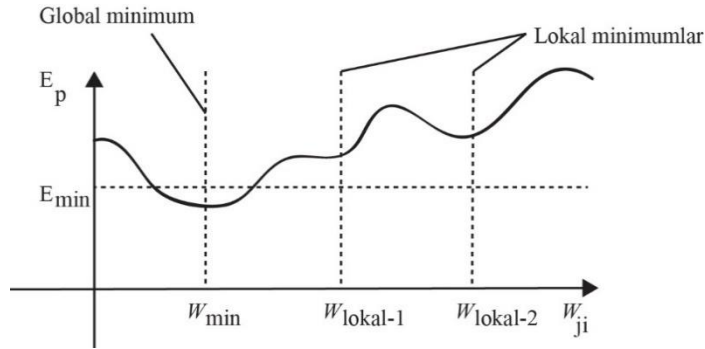
$$e_k = (d_k - o_k) \quad (3.24)$$

d_k ve o_k sırasıyla çıkış katmanındaki herhangi bir k nöronunun hedef değeri ve ağdan elde edilen asıl çıkış değerleridir. Her bir örnek veri için toplam karesel hata Denklem 3.25'teki gibi hesaplanmaktadır.

$$E = \frac{1}{2} \sum_k (d_k - o_k)^2 \quad (3.25)$$

Gizli katman için hata düzeyi Şekil 3.14'teki gibi lokal ve global minimum değerler olarak sınıflandırılmaktadır. Lokal minimumlar birden fazla olabilir. Global minimum, hatanın en küçük olduğu minimum değerdir. Öğrenme sırasında global minimuma ulaşmak amaçlanmaktadır ve ağırlıkların değişimi Denklem 3.26'daki gibi hesaplanmaktadır.

$$\Delta w_{kj} = -\varepsilon \frac{\partial E}{\partial w_{kj}} \quad (3.26)$$



Şekil 3.14: Gizli katmana ait ağırlık hata fonksiyonu.

Burada ε öğrenme oranı katsayısı adı verilen çoğunlukla 0,01 ile 0,9 arasında seçilen küçük değerde pozitif bir sayıdır. Amaç, uygun w seçimiyle, $E = \sum_p E_p$ toplam hatayı yeterince küçük yapabilmektir. Bu amacı gerçekleştirmek için, bir $p \in P$ örneği art arda ve rastgele biçimde seçilerek k . nöronda oluşan ve “delta” adı verilen hata işareti Denklem 3.27'deki gibi hesaplanmaktadır.

$$\Delta w_{kj} = \varepsilon \delta_0 y_j \quad (3.27)$$

Delta kuralı, E_p değerini düşürmek, ağırlığı $\delta_0 y_j$ değerine bağlı olarak değiştirmek demektir. Önceki gizli katmandaki tüm nöronlar ile çıkış katmanındaki k . nöron arasındaki bağ mukavemetleri, Denklem 3.27'deki Δw_{kj} miktarı ile ayarlanmaktadır. Eğer j nöronları gizli katmana ait nöronlar ise Denklem 3.28'deki gibi transfer (eşik) fonksiyonu olarak sigmoid fonksiyonu kullanılabilmektedir.

$$f(net_j) = y_j = \frac{1}{1+e^{-net_j}} \quad (3.28)$$

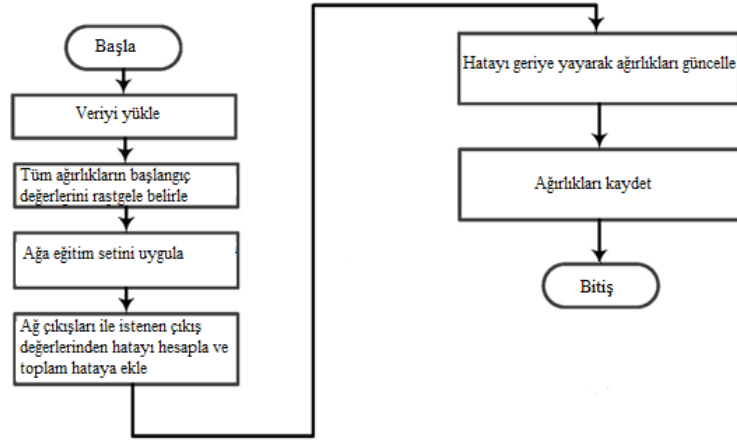
w_{ji} değerini öğrenme durumunda her eğitme örneklerinin seti için Δw_{ij} değerini hesaplamak gerekmektedir. Öğrenme oranı ε , hızlı öğrenmeyi sağlar; fakat dalgalanmalara sebep olabileceği için Rumelhart ve McClelland Denklem 3.29 ve Denklem 3.30'daki gibi bir tür momentum terimi dahil ederek değişim öngörmüşlerdir [59].

$$\Delta w_{ji}(N+1) = \varepsilon \delta_y x_i + \alpha \Delta w_{ji}(N) \quad (3.29)$$

$$\Delta w_{kj}(N+1) = \varepsilon \delta_0 y_j + \alpha \Delta w_{kj}(N) \quad (3.30)$$

Denklem 3.29 ve Denklem 3.30'a göre N öğrenme iterasyon sayısını göstermektedir. α değeri momentum terimi olan küçük değerli pozitif bir sayıdır. Momentum mevcut delta ağırlığı üzerinden önceki delta ağırlığının belirli bir kısmını beslemektedir. Böylece daha düşük öğrenme katsayısı ile daha hızlı öğrenme elde edilmesi sağlanmaktadır. Momentum katsayısı genellikle $0 < \alpha < 1$ aralığında değişen sabit bir sayıdır.

Hatanın geriye yayılması yöntemi, karesi alınmış hata fonksiyonunu minimize eden bir yöntem olup genelleştirilmiş delta kuralını eğitmek için kullanılmaktadır. Bu YSA modelinin eğitme işleminin akış diyagramı, Şekil 3.15'teki gibidir.

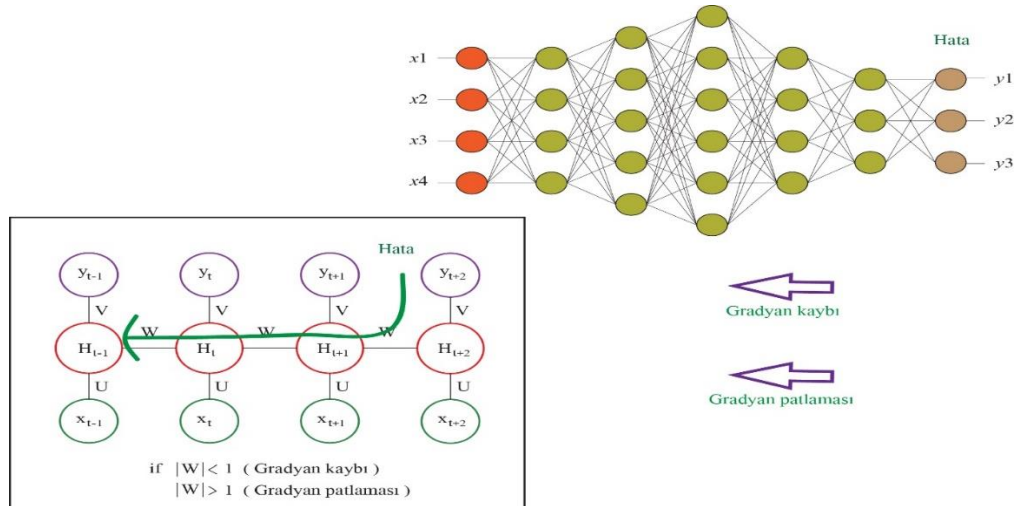


Şekil 3.15: Hatanın geriye yayılması yöntemi akış diyagramı.

Şekil 3.15'e göre, hatanın geriye yayılması yöntemi ile en küçük hata değeri bulunana kadar hesaplama devam etmektedir.

3.2.7. Uzun Kısa Süreli Bellek Ağları (Long Short Term Memory – LSTM) Yöntemi

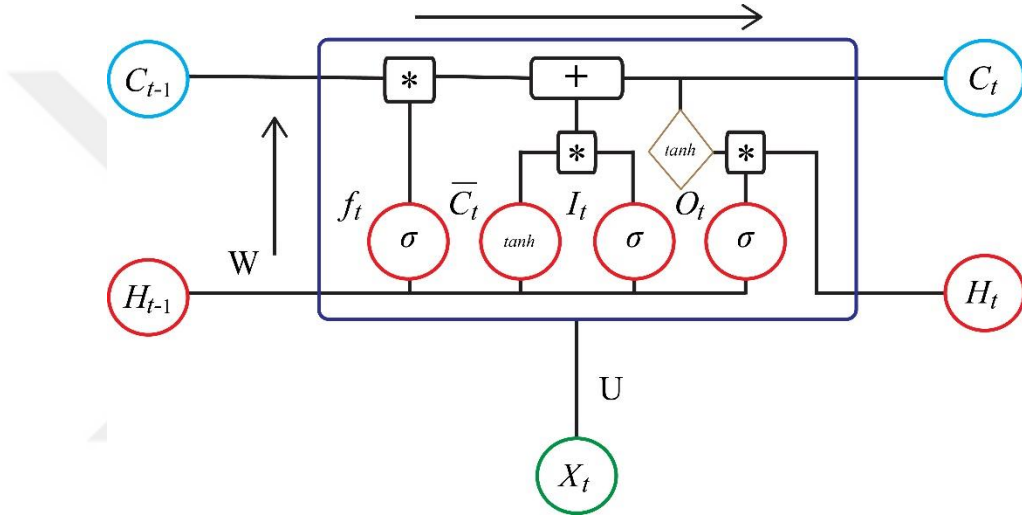
YSA yöntemi, ağdan hesaplanan hata değerinin gradyanlarını hesaplamaktadır, ardından ağırlıkları güncellemektedir. Geri yayımlı bir YSA ağı çıkış katmanından giriş katmanına kadar basit bir sinir ağına ağırlıkların güncellenmesinde sorun yaşanmayabilir, ancak derin bir sinir ağına Şekil 3.16'da gösterildiği gibi gradyan kaybı veya gradyan patlaması ismi verilen sorunlarla karşılaşılabilir [60].



Şekil 3.16: LSTM ağına gradyan kaybı ve gradyan patlaması sorunu.

Gradyan değerlerine geri dönüldüğünde, değerlerin üssel olarak küçülmesi ve bunun da gradyan kaybı sorununa neden olması veya üssel olarak daha büyük olması ve bunun da

patlayan gradyan sorununa neden olması mümkündür. Bu sorunu önlemek için düzeltilmiş doğrusal birimini bir aktivasyon fonksiyonu olarak uygulamak çözümlerden birisidir. LSTM ağında zaman adımları mevcuttur ancak her adım için LSTM ağında "bellek" olarak adlandırılan ekstra bilgiler mevcuttur [61]. Dolayısıyla LSTM hücresi sigmoid fonksiyonu içeren bir sinir ağında “unut kapısı (f)”, “giriş kapısı (I)” ve “çıkış kapısı (O)”, $\tanh$ fonksiyonlu bir sinir ağında “aday katmanı ($\bar{C}$)”, bir vektör olan “gizli durum (H)”, bir vektör olan “Bellek durumu (C)” bileşenlerinden oluşmaktadır. Şekil 3.17’de, t zaman adımındaki LSTM ağı diyagramı verilmiştir.



Şekil 3.17: t zaman adımındaki LSTM ağının genel yapısı.

Denklem 3.31 ile Denklem 3.36 arasındaki denklemlere göre, X_t giriş vektörünü, H_{t-1} önceki hücre çıkışı, C_{t-1} önceki hücre belleğini, H_t mevcut hücre çıkışı, C_t mevcut bellek hücresini temsil etmektedir.

$$f_t = \sigma(X_t * U_f + H_{t-1} * W_f) \quad (3.31)$$

$$\bar{C}_t = \tanh(X_t * U_c + H_{t-1} * W_c) \quad (3.32)$$

$$I_t = \sigma(X_t * U_i + H_{t-1} * W_i) \quad (3.33)$$

$$O_t = \sigma(X_t * U_o + H_{t-1} * W_o) \quad (3.34)$$

$$C_t = f_t * C_{t-1} + I_t * \bar{C}_t \quad (3.35)$$

$$H_t = O_t * \tanh(C_t) \quad (3.36)$$

t zaman adımındaki LSTM ağı diyagramına göre, aday katman dışındaki ağlar sigmoid aktivasyon fonksiyonuna sahip tek katmanlı sinir ağlarıdır. Aday katman $\tanh$ fonksiyonuna sahip tek katmanlı sinir ağıdır. Bu kapılar önce U giriş vektörü ve önceki gizli durum vektörünü (W) alır, ardından bunları birleştirir ve aktivasyon fonksiyonunu uygular. Son olarak bu geçitler vektörler üretir (Sigmoid için 0 ile 1 arasında, $\tanh$ için -1 ile 1 arasında). Böylece her zaman adımı için f, C', I, O vektörü hesaplanmaktadır. LSTM ağında zaman adımları vardır ve mevcut zaman adım değeri önceki zaman adım değerine bağlıdır, bu nedenle güncelleme yapmak için sonuna kadar geri gidilerek hesaplama yapılması gerekmektedir. LSTM, genellikle birkaç birim içeren “bloklar” içinde uygulanmaktadır. Bu tasarım, derin sinir ağlarında tipiktir ve paralel donanımlarla yapılan uygulamaları kolaylaştırmaktadır.

LSTM blokları, bilgi akışını kontrol eden üç veya dört “kapı” içermektedir. Bu kapılar, 0 ile 1 arasındaki bir değeri hesaplamak için mantıksal fonksiyonu kullanarak gerçekleştirilmektedir. Çoğaltma, bilginin belleğe girip çıkmak için kısmen izin vermek veya reddetmek için bu değerle uygulanmaktadır. Örneğin, bir “giriş” kapısı, yeni bir değer bellekte ne derece aktığını kontrol eder. Bir “unut” kapısı, bir değer bellekte kalma derecesini denetler. Bir “çıktı” kapısı, bloğun çıktı aktivasyonunu hesaplamak için bellekteki değerin ne kadar kullanıldığını kontrol eder. Örneğin; Ada odaya girer, Asya da odaya girer. Ada ...'ya "merhaba" dedi? Doğru tahmin yapabilmek için burada "Asya" yı C belleğine kaydeder. LSTM hücresi bilgileri ekleyebilir veya kaldırabilir. Örneğin; Ada ve Asya birlikte odaya girerler, daha sonra Ali odaya girer. Ada ... 'a "merhaba" dedi? Yapılan varsayım, hafızanın Asya'dan Ali'ye değişebileceğidir. Bu nedenle LSTM hücresi önceki bellek durumu C_{t-1} değerini alır ve unutma kapısı (f) ile Denklem 3.37'deki gibi çarpılmaktadır.

$$C_t = C_{t-1} * f_t \quad (3.37)$$

Hesaplanan unut kapısı değeri 0 ise, önceki bellek durumu tamamen unutulur, unut kapısı değeri 1 ise, önceki bellek durumu tamamen hücreye geçirilir.

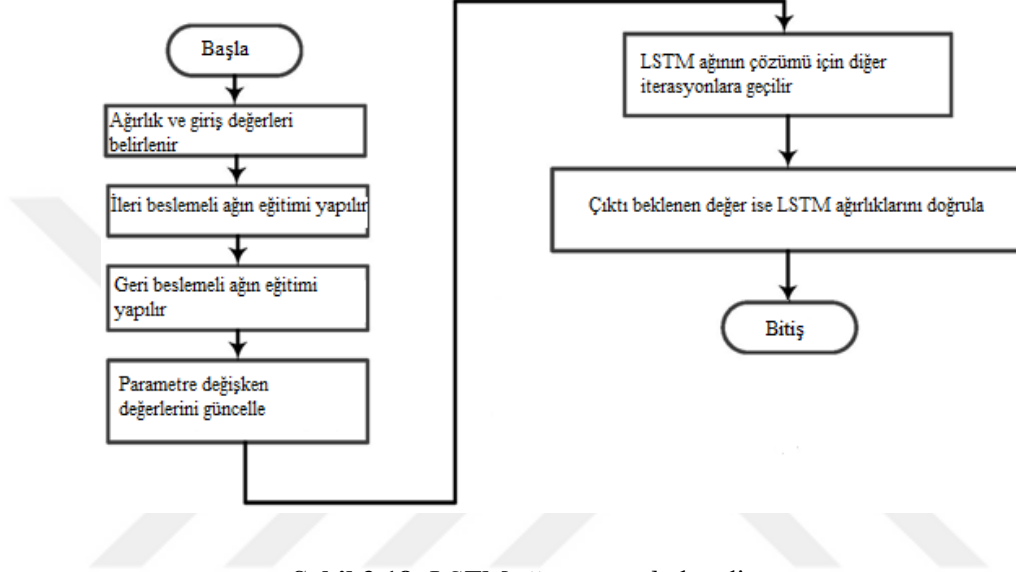
Mevcut t anında bellek durumu C_t ile yeni bellek durumu, giriş durumu ve C katman değeri Denklem 3.38'e göre hesaplanmaktadır.

$$C_t = C_t + (I_t * \bar{C}_t) \quad (3.38)$$

Denklemler 3.39'a göre, çıktı değeri için ağız durumunun C_t değerine dayalı olarak filtrelenmiş bir sürümü hesaplanarak gizli durum elde edilecektir.

$$H_t = \tanh(C_t) \quad (3.39)$$

C_t ve H_t değerleri bir sonraki aşamaya geçirilerek tekrar edilmektedir. LSTM ağının genel akış diyagramı, Şekil 3.18'deki gibidir.



Şekil 3.18: LSTM ağının genel akış diyagramı.

3.3. MODEL DEĞERLENDİRME

Veri setine en uygun sınıflandırma modelinin belirlenmesi için farklı modeller geliştirilerek modeller arası sınıflandırma doğruluklarının karşılaştırılmasıdır.

3.3.1. Karışıklık Matrisi Oluşturma

Sınıflandırma işlemlerinde gerçek verinin sınıfı ile öngörülen sınıf değerlerinin karşılaştırılmasıdır. İki sınıflı veri setine ait karışıklık matrisinde, Tablo 3.1'deki satırlar gerçek değerlere ilişkin sınıf etiketlerini, sütunlar öngörülen sınıf etiketlerini temsil etmektedir [62].

Tablo 3.1: Karışıklık matrisi.

	Tahmin Değeri: Epilepsi	Tahmin Değeri: Sağlıklı
Gerçek sınıf etiketi: Epilepsi	TP	FP
Gerçek sınıf etiketi: Sağlıklı	FN	TN

Tablo 3.1'e göre Epilepsi ve Sağlıklı olmak üzere iki sınıflı bir modelde;

TP tahmin değeri; Epilepsi sınıfı için başarılı öngörü sayısını ifade etmektedir (Doğru Pozitif).
 FP tahmin değeri; Epilepsi sınıfı için başarısız öngörü sayısını ifade etmektedir (Yanlış Pozitif).
 FN tahmin değeri; Sağlıklı sınıfı için başarısız öngörü sayısını ifade etmektedir (Yanlış Negatif).
 TN tahmin değeri; Sağlıklı sınıfı için başarılı öngörü sayısını ifade etmektedir (Doğru Negatif).

3.3.2. İki Sınıflı Modelleme için Doğruluk Ölçütleri Belirleme

İki sınıflı sınıflandırma modellerinde modelleme performansı için doğruluk (accuracy), duyarlılık (sensitivity), özgüllük (specificity) ve kesinlik (precision), F1 ölçüsü ölçütlerinden faydalanılmaktadır [63]. Bu ölçütler Denklem 3.40, Denklem 3.41, Denklem 3.42, Denklem 3.43 ve Denklem 3.44'e göre hesaplanmaktadır.

$$\text{Doğruluk} = \frac{TP+TN}{TP+TN+FP+FN} \quad (3.40)$$

$$\text{Duyarlılık} = \frac{TP}{TP+FN} \quad (3.41)$$

$$\text{Özgüllük} = \frac{TN}{TN+FP} \quad (3.42)$$

$$\text{Kesinlik} = \frac{TP}{TP+FP} \quad (3.43)$$

$$F1 = \frac{2TP}{2TP+FP+FN} \quad (3.44)$$

3.3.3. n -kat Çaprazlama Yöntemi ile Model Doğrulama

Veri setinin n adet eşit alt kümeye bölünüp $n - 1$ sayıda kümenin eğitim işlemi için, 1 tanesinin test işlemi için ayrılmasıdır. Tablo 3.2'de de gösterildiği gibi, bu işlem n kez tekrarlanır ve her defasında elde edilen doğruluk değeri ortalaması ilgili sınıflandırma modelinin doğruluk değeri olarak hesaplanmaktadır [64].

Tablo 3.2: n -kat çaprazlama yöntemine göre doğrulama ($n = 10$).

	1. adım	2. adım	3. adım	4. adım	5. adım	6. adım	7. adım	8. adım	9. adım	10. adım
1. adım	Test	Eğitim	Eğitim	Eğitim	Eğitim	Eğitim	Eğitim	Eğitim	Eğitim	Eğitim
2. adım	Eğitim	Test	Eğitim	Eğitim	Eğitim	Eğitim	Eğitim	Eğitim	Eğitim	Eğitim
3. adım	Eğitim	Eğitim	Test	Eğitim	Eğitim	Eğitim	Eğitim	Eğitim	Eğitim	Eğitim
4. adım	Eğitim	Eğitim	Eğitim	Test	Eğitim	Eğitim	Eğitim	Eğitim	Eğitim	Eğitim
5. adım	Eğitim	Eğitim	Eğitim	Eğitim	Test	Eğitim	Eğitim	Eğitim	Eğitim	Eğitim
6. adım	Eğitim	Eğitim	Eğitim	Eğitim	Eğitim	Test	Eğitim	Eğitim	Eğitim	Eğitim
7. adım	Eğitim	Eğitim	Eğitim	Eğitim	Eğitim	Eğitim	Test	Eğitim	Eğitim	Eğitim
8. adım	Eğitim	Eğitim	Eğitim	Eğitim	Eğitim	Eğitim	Eğitim	Test	Eğitim	Eğitim
9. adım	Eğitim	Eğitim	Eğitim	Eğitim	Eğitim	Eğitim	Eğitim	Eğitim	Test	Eğitim
10. adım	Eğitim	Eğitim	Eğitim	Eğitim	Eğitim	Eğitim	Eğitim	Eğitim	Eğitim	Test

3.3.4. Alıcı İşlem Karakteristikleri (Receiver Operating Characteristic – ROC) Analizi

Bir sınıflandırma modelinin hassasiyetini maksimum yapabilmek için Yanlış Negatif değerinin sıfıra indirgenmesi gerekmektedir. ROC analizi, sınıflandırma modelinin farklı Hassasiyet – Yanlış Pozitif Oranı değer çiftleri ile elde edilebilen toplam başarı eğrisidir. Eğri altında kalan alan sınıflandırma modelinin doğruluk oranını göstermektedir. AUC, TP ve FP eğrisinin altındaki alan Denklem 3.45'e göre hesaplanmaktadır [65].

$$AUC = \int_0^1 \left(\frac{TP}{TP+FN} \right) d \left(1 - \frac{TN}{TN+FP} \right) \quad (3.45)$$

3.3.5. Kappa Katsayısı

$Pr(a)$ ve $Pr(e)$ iki farklı gözlem arasındaki uyumu hesaplayan istatistiksel bir yöntemdir [66]. $Pr(a)$, iki gözlem değeri için gözlemlenen uyumların toplam değere orantısıdır. $Pr(e)$, beklenen değerin şansa bağlı ortaya çıkma olasılığıdır. $Kappa$ değeri hesabı Denklem 3.46'ya göre hesaplanmaktadır.

$$Kappa = \frac{Pr(a) - Pr(e)}{1 - Pr(e)} \quad (3.46)$$

$Kappa$ değeri -1 ile 1 arasında değer alabilmektedir. İki gözlem arasındaki uyum aşağıdaki gibi değerlendirilmektedir [66, 67].

Eğer $Kappa$ değeri < 0 ise iki gözlemci arasındaki uyum sadece tesadüfen bağlıdır,
Eğer $Kappa$ değeri $0.01 - 0.20$ ise iki gözlemcinin sonuçları önemsiz uyumludur.
Eğer $Kappa$ değeri $0.21 - 0.40$ ise iki gözlemcinin sonuçları yetersiz uyumludur.
Eğer $Kappa$ değeri $0.41 - 0.60$ ise iki gözlemcinin sonuçları orta düzeyde uyumludur.

Eğer Kappa değeri 0.61 – 0.80 ise iki gözlemcinin sonuçları iyi uyumludur.

Eğer Kappa değeri 0.81 – 1.00 ise iki gözlemcinin sonuçları çok iyi bir uyum düzeyine sahiptir.

3.4. ÇOK MODLU KARAR TABANLI VERİ FÜZYONU

Topluluk öğrenme yöntemleri, birçok makine öğrenmesi için en son teknoloji çözümlerdir [68]. Topluluk öğrenmesi çoklu makine öğrenmesi metotlarının kararlarını birleştiren şemsiye terimdir. Topluluk öğrenimi, belirli bir makine öğrenimi görevini çözmek için birden çok modelin üretilmesi ve birleşmesidir. Temel çalışma prensibi, çoklu makine öğrenmesi metotları ile önerilen her bir modelin tahminleme performansının ayrı ayrı birleştirilmesine dayanmaktadır [69]. Topluluk öğrenmesi yaklaşımı tarihte ilk defa İngiliz filozof ve istatistikçi olan Sir Francis Galton'un 1906 yılında İngiltere Plymouth'da bir gezisinde canlı hayvan pazarında "öküzün ağırlığı yarışması"na tanık olduğu sırada ortaya çıkmıştır. Panayırdaki bir öküz 800 katılımcıya gösterilmiş ve her katılımcıya bu öküzün kesildikten ve derisi yüzüldükten sonra kaç kilogram et vereceği sorulmuştur. Katılımcılar panayır meydanında sergilenen öküzün ağırlığına ilişkin tahminlerini bir kâğıda yazarak yarışmaya katılmışlardır. En yakın tahminde bulunan ise ödül kazanacaktır. Galton, yarışma sonrasında katılımcıların tahminlerini toplayıp, tüm tahminlerin ortalamasını hesaplamıştır. Sonuç olarak, yarışmada sergilenen öküzün gerçek ağırlığı 539 kilogram iken, yarışmaya katılan 800 kişinin tahminlerinin ortalaması 538,65 kilogram olmuştur. Sonuçta, uzman olmayanların çoğunluğunun oluşturduğu bu topluluk, öküzün gerçek ağırlığını 935 gramlık yanılğı ile bilebilmişlerdir. Galton, bu yarışma sayesinde topluluğun tahminlemede hata payının %1 olduğu sonucuna varmıştır. Galton bu tarz deneyleri defalarca tekrarlayarak konuyla alakasız olsalar bile kitlelerin de bir fikrinin, kanaatinin ve tahmininin olduğunu ve bu ölçünün doğruya işaret ettiğini ispatlamıştır [70]. Topluluk metodolojisinin sezgisel olarak açıklaması insan doğasından ve farklı görüşler toplama eğiliminden kaynaklanmaktadır. Ana fikir, “Öküzün Ağırlığı Yarışması”, James Surowiecki'nin “The Wisdom of Crowds (Kitlelerin Bilgeliği)” [71] kitabında yer aldığı gibi çok kişinin aklı biraraya geldiğinde neden az kişinin aklından daha üstündür? Kolektif akıl iş dünyasını, ekonomileri, toplumları ve ulusları nasıl biçimlendirir? sorularıyla referans vermiştir. Kitaba göre, “çok sayıdaki çok bilmeyen kişinin bilgisi”, “az sayıdaki çokbilmiş kişinin bilgisi” nden daha üstün olabildiği ifade edilmiştir. Farklı kaynaklardan gelen veriler doğası gereği farklı özellikler taşırlar. Tüm özellikleri bir araya getirerek tüm veri setini tek bir sınıflandırıcı ile sınıflandırmak pratik bir yol değildir. Bunun yerine özelliklerin farklı gruplar halinde ayrılarak sınıflandırılması ve sınıflandırıcıların

çıktılarının birleştirilmesi ile daha başarılı sınıflandırma sonucu hesaplanmaktadır. Topluluk öğrenmesi, toplulukların oluşturulmasını ve sınıflandırıcıların kaynaştırılmasını içermektedir. Sınıflandırma başarımı başarılı bir topluluk öğrenme modeli oluşturmada iki amaç ele alınmıştır [72]: (a) topluluktaki her bir sınıflandırıcının iyi başarımlı sonucu vermesi, (b) topluluktaki sınıflandırıcıların çok çeşitli olması gerekliliğidir. Toplulukların oluşturulması, bu sınıflandırıcıların paralel veya ardışık eğitiminde temel sınıflandırıcıların eğitimi ile elde edilebilmektedir. Farklı topluluk metodlarının teorik analizleri Tablo 3.3'te ifade edilmiştir.

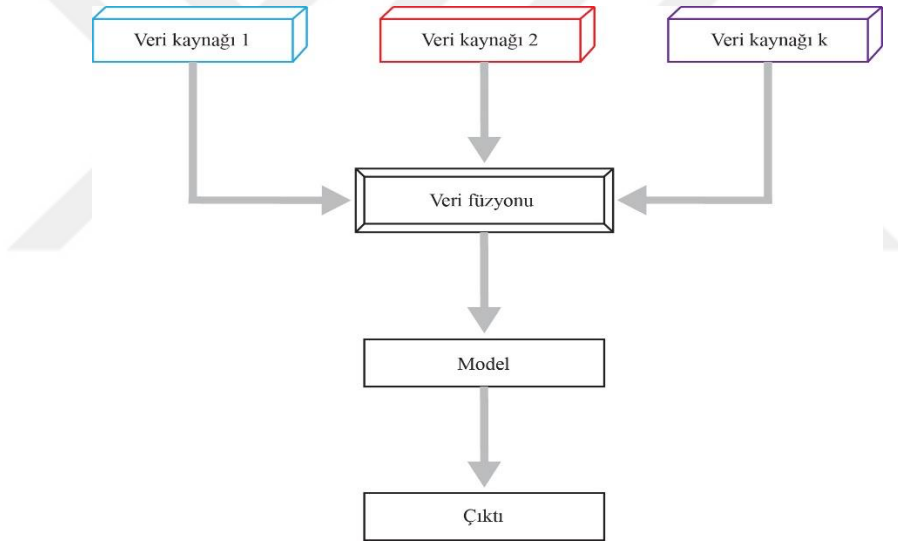
Tablo 3.3: Farklı topluluk metodlarının teorik analizleri.

Metot	Artıları	Eksileri
Çantalama (Bagging)	– varyans azaltma – farklı eğitim örnekleri aracılığıyla çeşitliliğin oluşturulması	– dengesiz verilerle uğraşmak için uygun değil
Arttırma (Boosting)	– önyargı azaltma – eğitimin sonraki aşamalarında yanlış sınıflandırılmış örneklerle odaklanarak çeşitlilik yaratma	– küçük boyutlu veri setleri ile uğraşmak
Rastgele alt uzay (Random Subspace)	– boyutluluk sorununun azaltılması – farklı özellik alt kümeleri aracılığıyla çeşitlilik yaratma	– Özellik alt uzayının boyutluluğunun nasıl tanımlanacağına dair net bir kılavuz yok
Sabit füzyon kuralları (Fixed rules of fusion)	– düşük hesaplama karmaşıklığı – bağımsız veya düşük korelasyona sahip sınıflandırıcıları küçük veriler üzerinde benzer performansla birleştirmek için uygundur	– yüksek korelasyona sahip sınıflandırıcıları farklı performansla birleştirmek için uygun değildir – her bir tekil sınıflandırıcının güvenilirliği dikkate alınmaz
Eğitilmiş füzyon kuralları (Trained rules of fusion)	– sabit kurallardan daha uygun – yüksek korelasyona sahip sınıflandırıcıları farklı performansla birleştirmek için	– yüksek hesaplama karmaşıklığı

Çok modlu karar tabanlı veri füzyonu ile ilgili literatürdeki çalışmalar araştırılmıştır. Lahat ve diğ. [73] çok modluluğu, çoklu sensörler tarafından gözlemlenen bir sistem olarak tanımlamıştır. Çok modluluğu kullanmanın amacı, tek tek sensörlerden önemli bilgileri çıkarmak ve karıştırmak ve bu karışık özelliği belirli bir sorunu çözmek için kullanmaktır. Böylece, beklenen çıktı tekil modellemelere göre daha zengin şekilde temsil edilecek ve daha başarılı sınıflandırma performansına sahip olacaktır. Çok modluluk veri analizi tıp, iş dünyası ve sürücüsüz araç teknolojisi ve oyun programlama gibi çeşitli çalışma alanlarına pratik bir çözüm olmuştur. Kamera, LIDAR, radar ve ultrasonik gibi yaygın uzaktan algılama cihazları sıklıkla kaynaştırılmaktadır [74, 76]. Çok modlu veri füzyonu için erken, geç ve ara füzyon [73, 75] olmak üzere üç farklı teknik vardır.

3.4.1. Erken Füzyon

Veri seviyesi füzyonu olarak da bilinen erken füzyon, Şekil 3.19’da gösterildiği gibi modelleme gerçekleştirilmeden önce birden çok verinin kaynaştırılmasının geleneksel bir yoludur. Bu yöntem giriş seviyesi füzyonu da denir. Erken füzyon tekniği için iki olası yaklaşım önerilmektedir [77]. İlk yaklaşım, iki sensör arasındaki korelasyonu kaldırarak veriyi birleştirmektir. İkinci yaklaşım, verileri daha düşük boyutlu ortak alanda kaynaştırmaktır. Erken füzyon ham verilere veya sensörlerden elde edilen önceden işlenmiş verilere uygulanabilmektedir. Veri özellikleri füzyondan önce verilerden çıkarılmaktadır. Veri kaynaklarının senkronizasyonu, bir veri kaynağı ayırıkken diğerleri sürekli olduğunda veri kaynaklarını tek bir özellik vektörüne dönüştürmek, erken veri kaynaşmasında önemli bir zorluktur.



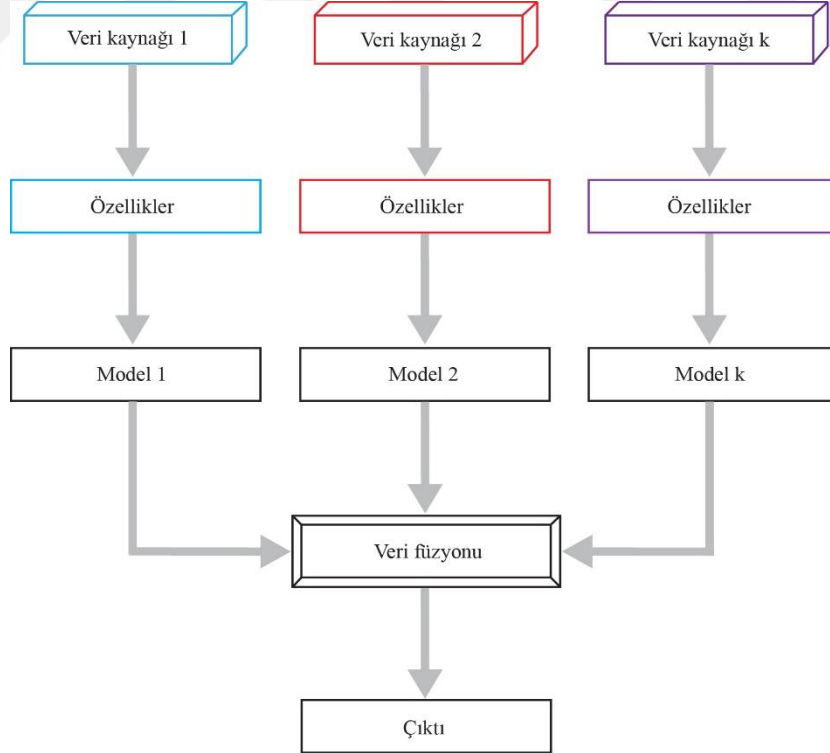
Şekil 3.19: Erken füzyon.

Erken veri kaynaşmasının arkasındaki varsayım, birden çok veri kaynağı arasındaki koşullu bağımsızlıktır. Sebe ve diğ. [78], “Bu varsayım her zaman doğru değildir, çünkü çoklu modeller, örneğin video ve derinlik ipuçları gibi yüksek derecede ilişkili özelliklere sahip olabilirler” demişlerdir. Owens ve diğ. [79] aynı zamanda farklı modellerin daha yüksek bir seviyede birbiriyle ilişkili bilgiler içerebileceğini belirtmektedir. Dolayısıyla, her bir modelin çıktısının birbirinden bağımsız olarak işlendiği varsayılabilir. Poria ve diğ. [80], çok modlu bir akıştaki özelliklerin birleştirilmesini içeren erken aşama veri füzyonunu uygulamışlardır. Erken evre veri füzyonunu kullanmanın iki dezavantajını vurgulamışlardır. Bunlardan ilki, füzyondan önce ortak bir zemin oluşturmak için yöntemlerden büyük miktarda

verinin çıkarılması gerekliliğidir. Veriler ortak matrislere sahip olduktan sonra, bir makine öğrenme yöntemi kullanılarak analiz edilmektedir. Bu yöntemin vurgulanan diğer dezavantajı, farklı yöntemlerin zamanda senkronize edilmesidir. Bu dezavantajı aşmanın yaygın bir yolu, veriyi ortak bir paydada toplamaktır. Martínez ve diğ. [81] eğitim, havuzlama ve evrişim füzyonunu önermektedir. Önerilen bu yöntemler, ardışık ayrık olayların sürekli verilerle kaynaşmasıyla elde edilmiştir.

3.4.2. Geç Füzyon

Karar füzyonu olarak da bilinen geç füzyon, Şekil 3.20'deki gibi veri kaynaklarını bağımsız olarak takip eder ve bunu bir karar verme aşamasında füzyon izlemektedir. Geç füzyon, topluluk sınıflandırıcılarının popülaritesinden esinlenmiştir [82]. Bu teknik, özellikle veri kaynakları örnekleme hızı, veri boyutu ve ölçü birimi bakımından birbirinden önemli ölçüde farklı olduğunda, erken füzyon yönteminden çok daha basittir. Geç füzyon genellikle daha iyi performans sağlamaktadır. Çünkü birden fazla modelden gelen hatalar bağımsız olarak ele alınır ki bu nedenle modellemeler arasındaki hata birbirlerinden ilişkisizdir.



Şekil 3.20: Geç füzyon.

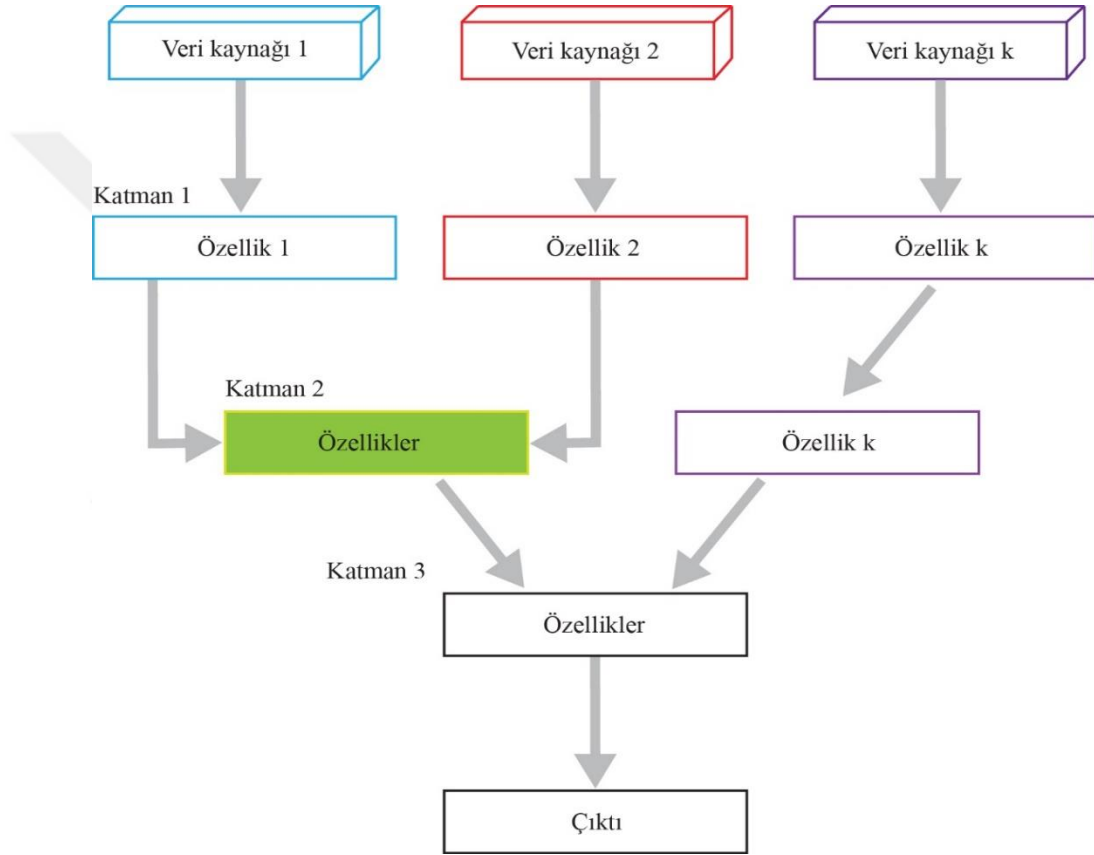
Ramachandram ve diğ. [83], geç füzyonun erken füzyondan daha iyi performans gösterdiğine dair kesin bir kanıt bulunmadığını savunmaktadır. Ancak, bazı araştırmacılar çok modlu veri problemlerini analiz etmek için geç füzyon tekniğini önermektedir [84 – 88]. Bağımsız olarak eğitilmiş modellerin her birinin nihayetinde nasıl birleştirileceğine karar vermenin en uygun yolunu belirlemek için farklı kurallar vardır. Bayes kuralları, maksimum–füzyon ve ortalama–füzyon, yaygın olarak geç füzyon tekniğine ait kurallardan bazılarıdır. Girdi veri akışları boyutluluk ve örnekleme oranı açısından önemli ölçüde değiştiğinde, geç füzyon kullanılması daha basit ve daha esnek bir yaklaşımdır. Senanayaka ve diğ. [84], vites kutusundaki arızaları tespit etmek için çok katmanlı bir karma öğrenme yöntemi önermişlerdir. Özellik düzeyinde veri füzyonu kullanılarak alan adı bilgisi özellikleri ÇKA sınıflandırıcısı ile görüntü veya örüntü için ise evrişimli sinir ağı (Convolutional Neral Network – CNN) sınıflandırıcısı ile sınıflandırılmıştır. Naive Bayes yöntemiyle iki yöntemin sonucunu birleştirilerek hata sınıflandırması yapmışlardır. MLP + CNN yöntemlerinin karar seviyesi füzyon aşamasında en başarılı doğruluk sonucunu hesaplamışlardır. Mi ve diğ. [85], sınıflandırıcıların kararlarını birleştirmek, bir örüntü tanıma sisteminin performansını arttırabileceğini öne sürmüşlerdir. Ultrason görüntülerinden katarakt operasyonlarında pratik gereksinimlerin tespiti için 15 farklı veri seti için yöntem önermişlerdir. BigDeli ve diğ. [86], birden fazla sensörden uzaktan algılama verilerinin birleşmesi, sınıflandırılmasını önermişlerdir. Bunun nedeni, ek kaynakların daha fazla bilgi sağlaması ve farklı bilgilerin birleştirilmesi, gözlemlenen alanın daha iyi anlaşılmasını sağlayacağını düşünmüş olmalarıdır. Veri füzyonu, LIDAR ve optik arazi örtüsü sınıflandırması için uzaktan algılama verileri daha fazla dikkat çekmiştir. Bu bildiride, birleşme füzyonu metodolojisi arazi örtüsü sınıflandırmasında hiperspektral ve LIDAR verileri kullanmışlardır. Önerilen füzyon yöntemi için bir DVM tabanlı sınıflandırıcı füzyon tekniği uygulanmıştır. Karar düzeyinde hiperspektral ve LIDAR verileri kullanmışlardır. İlk olarak, özellik alanları çıkarılmıştır. Ardından, DVM sınıflandırıcılara her özellik uygulanmıştır. Sınıflandırıcılardan sonuçlar alındıktan sonra, sınıflandırıcı füzyon yöntemi olarak Naive Bayes ve DVM sınıflandırıcılarının sonuçları iki veri kümesi halinde birleştirilmiştir. Deneysel sonuçlara göre önerilen veri füzyon yöntemi tek veri kümelerine kıyasla sınıflandırma doğruluğunu ve kappa katsayısını başarılı bir şekilde geliştirmiştir. Sonuçlar genel olarak hiperspektral ve LIDAR verileri üzerinde DVM yöntemiyle sınıflandırılmasında ayrı ayrı doğrulukları %88 ve %58 olarak elde edilirken karar füzyon metodolojisi ile sınıflandırma doğruluğu %91 olarak elde edilmiştir. Rappoport ve Shamir [87], RNA dizilimi, DNA

metilasyon, transkriptomik ve metilomik tekil omik veri setinin biyolojik ve tıp araştırmalarında tek başına kıymetsiz olduğunu son yıllarda büyük biyomedikal veri setlerinin önem kazandığını vurgulamışlardır. Çalışmalarında kümeleme yöntemlerine çoklu omik veri setlerini uygulamışlardır. Çoklu veri tiplerinin birden fazla kümeleme yöntemleri ile kümelemişlerdir. Kanser Genom Atlas veri setinden elde ettikleri veri seti ile 10 farklı kanser tipini yolak bazlı karşılaştırmışlardır. Kanser alt tipleri tanımlanırken, gen dizilimi ve DNA metilasyonunu birlikte kullanmışlardır. Çoklu omik kümeleme yaklaşımının verideki deneysel ve biyolojik gürültüyü azalttığını elde etmişlerdir. Toh ve diğ. [88], translasyonel tıpta makine öğrenmesi ile gelecek nesiller için hastaya daha hızlı tanı ve tedavi etmede düşük maliyetli ve ölçekte yüksek kalitede veri çıkarımı için yeni karar mekanizmalarının keşfedileceğini hedeflediklerini vurgulamışlardır.

3.4.3. Ara Füzyon

Hibrid füzyonu olarak da bilinen ara füzyon mimarisi, popüler derin sinir ağı temeline dayanmaktadır. Bu yöntem, model eğitiminin farklı aşamalarında veri kaynaşmasına izin veren en esnek yöntemdir. Sinir ağı tabanlı çok modlu veri füzyonu, sınıflandırma performansını büyük ölçüde arttırmıştır. Ara füzyon, giriş verilerini çoklu katmanlar aracılığıyla daha yüksek bir temsil seviyesine dönüştürmektedir. Her katman, giriş verilerinin ölçeğini, eğimini ve salınımlarını dönüştüren ve orijinal giriş verilerinin yeni bir temsilini veren doğrusal ve doğrusal olmayan işlevleri çalıştırmaktadır. Derin öğrenme çok modlu bağlamda ara füzyon, farklı model temsillerinin tek bir gizli katmana kaynaşmasıdır, böylece model, modellerin her birinin ortak bir temsilini öğrenmektedir. Farklı model özelliklerinin füzyonunun gerçekleştiği katmana füzyon katmanı veya paylaşılan temsil katmanı denilmektedir. Şekil 3.21'deki gibi farklı modeller aynı anda tek bir paylaşılan sunum katmanına birleştirilebilir veya bu, bir seferde bir veya birden fazla model kullanılarak kademeli olarak gerçekleştirilebilmektedir. Her ne kadar birden fazla yöntem özelliğini veya ağırlığı tek bir katmanda birleştirmek mümkün olsa da modelin aşırı sığmasına yol açabilir veya ağ her yöntem arasındaki ilişkiyi öğrenemeyebilir. Derin çok modlu füzyon performansını iyileştirmenin bir yolu, verinin boyutunu azaltmaktır. Yi ve diğ. [89] bir füzyon katmanı veya paylaşılan temsil katmanı oluşturduktan sonra ağı boyutsallığını azaltmak için oto-kodlayıcılar kullanmışlardır. Erken seviye füzyon ve geç füzyonun aksine, ara füzyon, farklı derinliklerde özellikleri birleştirmek için esneklik sunmuştur. Karpathy ve diğ. [90], bir araştırma makalesinde, eğitim video akışı

özelliklerinin birden fazla füzyon katmanında kademeli olarak kaynaştığı bir “yavaş kaynaştırma” ağı kullanmışlardır. Bu yaklaşım, büyük ölçekli bir video akışı sınıflandırma probleminde daha iyi performans göstermiştir. Neverova ve diğ. [91], iletişimsel jest tanımda son teknoloji ürünü öneren çalışmada yüksek derecede korelasyona sahip girdi modellerini (görsel girdi modelleri, sonra hareket girdi modelleri ve sonra ses girdi modelleri) önce ve sonra düşük korelasyonlu olanları kademeli olarak kaynaştıran kademeli bir füzyon yöntemini göstermektedir.



Şekil 3.21: Ara füzyon.

Liu ve Zhang [92], çalışmalarında yüz ifadesi tanımak için farklı veri setlerini kullanarak çok sınıflı görev tanımlamışlardır. Farklı öğrenme yöntemleri kullanarak farklı özellik setleri üzerinde eğitilmiş sınıflandırıcıların kaynaşmasını içeren bir topluluk öğrenmesi bakış açısı önermişlerdir. Topluluk öğrenme yaklaşımlarının, karar ağacı ve derin sinir ağları gibi popüler öğrenme yaklaşımlarına kıyasla önemli sınıflandırma doğruluğu sağladığı sonucuna varmışlardır. Çok seviyeli füzyon ile araştırma yapmışlardır ve %87 sınıflandırma doğruluğu elde etmişlerdir. Si ve diğ. [93], tamamen mekanize kömür madenciliği için bir endüstriyel uygulama örneğinde sistemin etkisini belirtmek için bir kesicinin kesme koşullarını verimli ve

doğru bir şekilde tanımlamışlardır. Çoklu sensör veri füzyonu tanımlama yöntemi olan paralel yarı Newton sinir ağı ve Dempster–Shafer teorisi ile sınıflandırma modeli geliştirmişlerdir. Titreşim sinyalleri ve 6 farklı kesme koşulunun akım sinyali tasarlanmış bir deney sisteminden toplanmış ve bazı özel durum özellikleri topluluk ampirik moduna dayalı çıkarılmıştır. Deneyde, üç sınıflandırıcı, ölçülen verilerin seçilmiş özellikleri ve DS teorisi, sınıflandırıcıya uygulanmıştır. Deneysel sonuçlar önerilen yöntemin rakip yöntemlerden daha yüksek algılama doğruluğu ve güvenilirliği ile çalışmıştır. 3 farklı sınıf olan kömür dikişinin farklı jeolojik koşulları için 6 farklı kesme özelliği için ortalama sınıflandırma doğruluğu hesaplanmıştır. Kuncheva ve diğ. [94], çoklu sınıflandırıcıların füzyon tekniğini, tekil sınıflandırıcı modellemeye göre sınıflandırma performanslarını artırmak amacıyla kullanmışlardır. Sui ve diğ. [95], şizofreni hastalarına ait FMRG, SMRG ve EEG verisini kullanarak çoklu model füzyonu ile modelleme yapmışlardır. Çalışmalarında kanonik korelasyon analizi, t–test ve DVM–RFE analizleri ile ayrı ayrı özellik seçimi yapmışlardır ve her bir yöntemden elde edilen optimum özellikleri 3–yollu füzyon ile birleştirmişlerdir. Marini ve diğ. [96], çalışmalarında gen, alan, yolak ve protein–protein etkileşim verisi ile gelecek jenerasyonlara epilepsi gen paneli sunmak için veri füzyonu çerçevesi sunmuştur. Çalışmalarında çoklu genom kaynakları kullanmışlardır. 1–4 yaşları arasındaki 85 epilepsi hastası ve 61 kontrol grubu için ROI modellemesi ile epileptik ve epileptik olmayan sınıflandırması yapmışlardır. Çalışmalarında 15 veri setini birleştirmişlerdir. Her bir veri seti eğitim ve test seti olarak ayrılmış ve test için Rastgele Orman Modeli uygulanmıştır. Çalışmalarını (a) veri toplama ve veri ön işleme (b) tek değişkenli analiz (c) çok değişkenli tahminleme modeli olmak üzere üçe ayırmışlardır. Tek değişkenli analiz ile çok değişkenli tahminleme ayrı ayrı yürütülmüştür. He ve ekibi [97], radyom, klinik tedaviler, tıbbi görüntüler ve patoloji gibi kaynaklardan gelen farklı tipteki verileri analiz ederek tümör derecelendirme, teşhis ve öngörü tahminlemesi yapmışlardır. Bununla birlikte, tahmin modellemesi için uygun bir sınıflandırıcı seçmenin yanı sıra çeşitli klinik bilgileri yönetmenin etkili bir yolunu keşfetmek, klinik bağlamda pratiği hala talep edilmektedir. Bu çalışmalarında, farklı sınıflandırıcıları çok kriterli karar verme çerçevesinde birleştirmek için sınıflandırıcı füzyon stratejisi önermişlerdir. Önerilen stratejiye dayanan hiyerarşik bir tahmin şeması, çoklu model özelliklerini ve çoklu sınıflandırıcıları güvenilir bir şekilde bağlamak için de araştırmışlardır. UCI veri setinden 10 farklı veri seti ile çalışmışlardır.

3.5. HASTA VERİLERİNİN TOPLANMASI

Haliç Üniversitesi Etik Kurul kararıyla bu tezde epilepsi hastalığı için oluşturulacak özgün veri seti için demografik özellikler, nörolojik muayene epikriz raporundan elde edilen klinik parametreler, laboratuvar bulguları, EEG sinyal verisi, MRG radyoloji rapor verisi, Avicenna Hastanesi Nöroloji bölümünde takip edilen, klinik olarak kesin epilepsi tanısı konulmuş hasta bireyler ve kontrol grubu sağlıklı bireylere ait verilerdir. Hasta verileri Avicenna Hastanesi'nde MS-SQL 2018 veri tabanından sorgulanmıştır. Hasta verilerinin SQL sorguları ile sorgulandığı Avicenna Hastane veri tabanı varlık-ilişki diyagramı Şekil 3.22'de gösterilmiştir.



Şekil 3.22: Avicenna Hastane veri tabanı varlık-ilişki diyagramı.

Bu tez çalışması için uygun hasta verisinin hastane veri tabanından sorgulanarak seçilebilmesi için hasta verilerinin Şekil 3.23'teki kaba koddaki ön koşulu sağlaması gerekmektedir.

```

IF (hasta.yaş >= 18){
  IF ((hasta.epikriz raporu == true) && (hasta.laboratuvar bulguları == true) && (hasta.EEG == true)
  && (hasta.epilepsi protokolünde çekilmiş MRG radyoloji raporu == true)) {
    IF ((hasta.tüm veriler ilgili muayene tarihinde yapılmışsa) && (hasta.nörolojik tanı ==
    (Epilepsi) || hasta.nörolojik tanı == (Sağlıklı)))
  }
}
İlgili hasta verisi oluşturulacak özgün veri seti için seçilir

```

Şekil 3.23: Epilepsi veri seti için uygun hasta verisinin seçilme kriteri.

Bu çalışmada oluşturulan epilepsi hastalığına özgün gerçek veri seti için Avicenna Hastane veri tabanından 30 Haziran 2006 ile 19 Mart 2020 tarihleri arasındaki hasta verileri sorgulanmıştır. Seçilen hasta bireyler ile oluşturulan epilepsi veri setindeki hasta verisinin detaylarını içeren bilgiler Tablo 3.4'te gösterilmiştir.

Tablo 3.4: Epilepsi veri setindeki hasta detayları.

Sınıf	Yaş	Kadın	Erkek
Epilepsi	18–88	59	38
Kontrol Grubu	18–65	12	13

3.5.1. Seçilen Hasta Verilerinden İstatistiksel Özelliklerin Çıkarılması

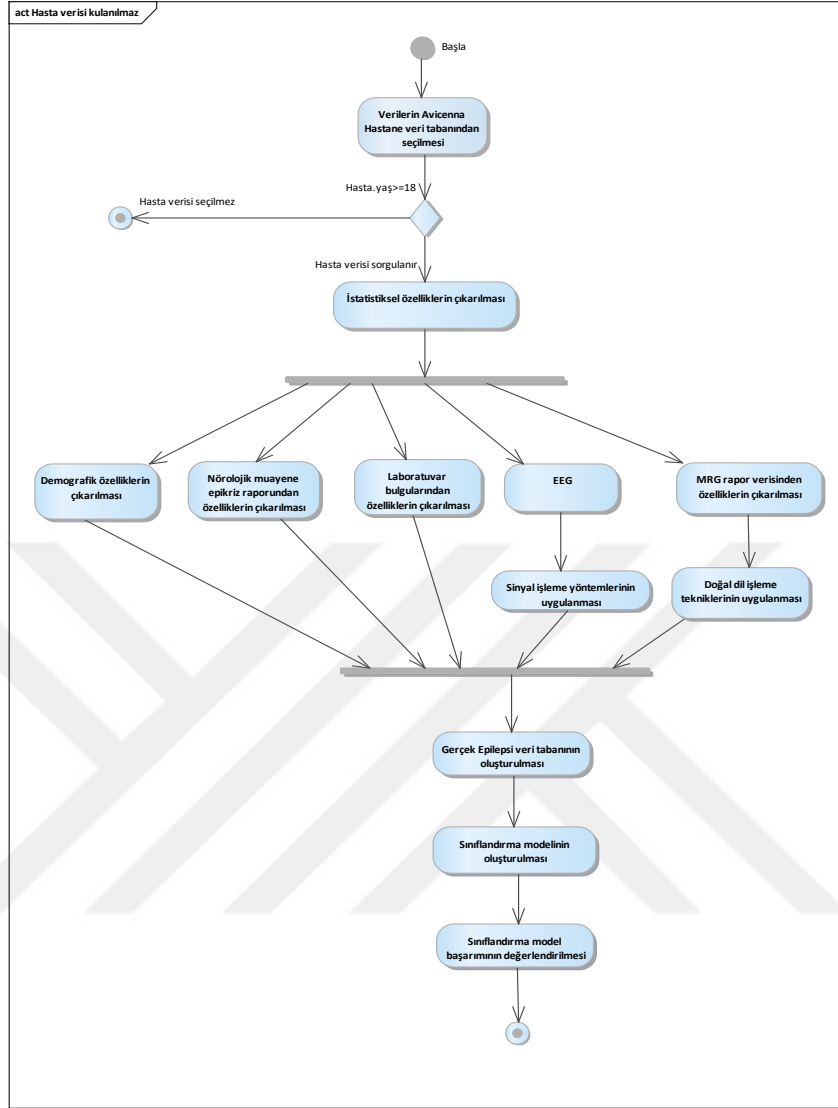
Şekil 3.23'teki gibi oluşturulması hedeflenen özgün epilepsi veri seti için Avicenna Hastanesi hasta veri tabanından yaşı 18'den büyük olan bireylere ait doktor epikriz raporundan elde edilen klinik parametreler, laboratuvar bulgularından elde edilen kan değerleri, EEG verisi ve MRG radyoloji rapor verisi olan epilepsi ve kontrol grubu sağlıklı bireylere ait verinin seçilmesi amaçlanmıştır. Seçilen hastanın farklı yıllarda kontrolleri varsa güncel olan en son yıla ait verisi kullanılmıştır. Epilepsi hastası ve kontrol grubu sağlıklı bireylere ait verilerden istatistiksel özellikler çıkarılmış ve özgün epilepsi veri seti oluşturulmuştur. İstatistiksel özelliklerin çıkarılmasında Şekil 3.24'teki aktivite diyagramındaki işlem adımları uygulanmıştır.

Şekil 3.24'e göre özgün veri tabanı için seçilmiş bireylere ait verilerden istatistiksel özellikler çıkarmak için işlem adımlarının ayrıntıları aşağıda verilmiştir.

Adım 1: Epilepsi hastası ve kontrol grubu bireylere ait demografik özellikler (yaş, cinsiyet) çıkarılmıştır. Daha sonra, epilepsi hastası ve kontrol grubu bireylere ait nörolojik muayene epikriz raporlarından hasta şikayetleri ile klinik parametreler çıkarılmıştır.

Adım 2: Epilepsi hastası ve kontrol grubu bireylere ait laboratuvar test sonuçlarındaki bulguları sorgulanarak çıkarılmıştır.

Adım 3: EEG sinyal verisinin hastane veri tabanından .e formatından .edf formatında dışarı aktarılması ve hastalıkla ilgili anlamlı ayırt edici özellikler sinyal işleme yöntemleriyle çıkarılmıştır.



Şekil 3.24: Epilepsi veri seti için istatistiksel özelliklerin çıkarılmasındaki işlem adımları.

Adım 4: Radyolog tarafından okunan MRG raporlarından hastalıkla ilgili epilepsi sözlüğü oluşturulması için doğal dil işleme tekniklerden ZipF yasası tekniği kullanılmış ve hastalıkla ilgili anlamlı özellikler çıkarılmıştır. İstatistiksel özelliklerin çıkarılmasıyla özgün epilepsi veri seti oluşturulmuştur.

3.5.1.1. Klinik Parametre Özellik Setinin Elde Edilmesi

Hastanın demografik özelliklerinden; doğum tarihi, yaş, cinsiyet bilgileri ile birlikte, hasta muayene raporları okunarak hastanın şikâyetleri (şuur kaybı, terleme, uyku bozukluğu, baş ağrısı, konuşmada zorlanma, görme bozukluğu, el-kollarda uyuşma, baş dönmesi, bilinç kaybı,

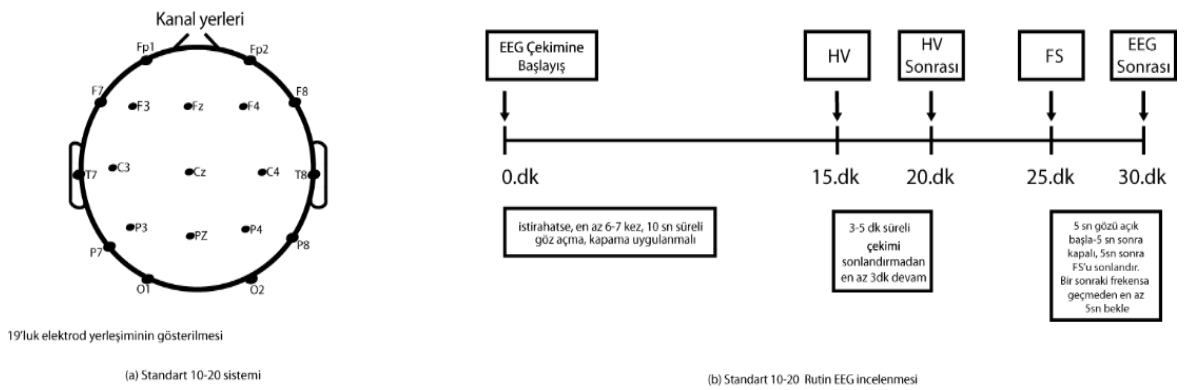
unutkanlık, bayılma, geç algılama, dalma, ateş, kusma, çabuk yorulma, vücutta kasılma, çenede kilitlenme , ... gibi) ile ilgili bilgileri doktor epikriz raporlarından sorgulanarak çıkarılmıştır.

3.5.1.2. Laboratuvar Bulguları Özellik Setinin Elde Edilmesi

Epilepsi hastalığına tanı koymak ve takip etmek, hastanın hangi tipte nöbet geçirmekte olduğunu belirlemek, toksinler, enfeksiyonlar, ilaç veya alkol yoksunluğu, nöbetlere neden olabilen diyabet gibi altta yatan hastalıkları tanımlamak ve epilepsi hastalığını benzer belirti ve bulguların bir bölümüne neden olabilen bayılma veya inme gibi rahatsızlıklardan ayırt etmek için laboratuvar bulgularından elde edilen kan test sonuçları (WBC, RBC, HGB, HCT, PLT, MCV, MCH, MCHC, ... gibi) sorgulanarak çıkarılmıştır.

3.5.1.3. EEG Özellik Setinin Elde Edilmesi

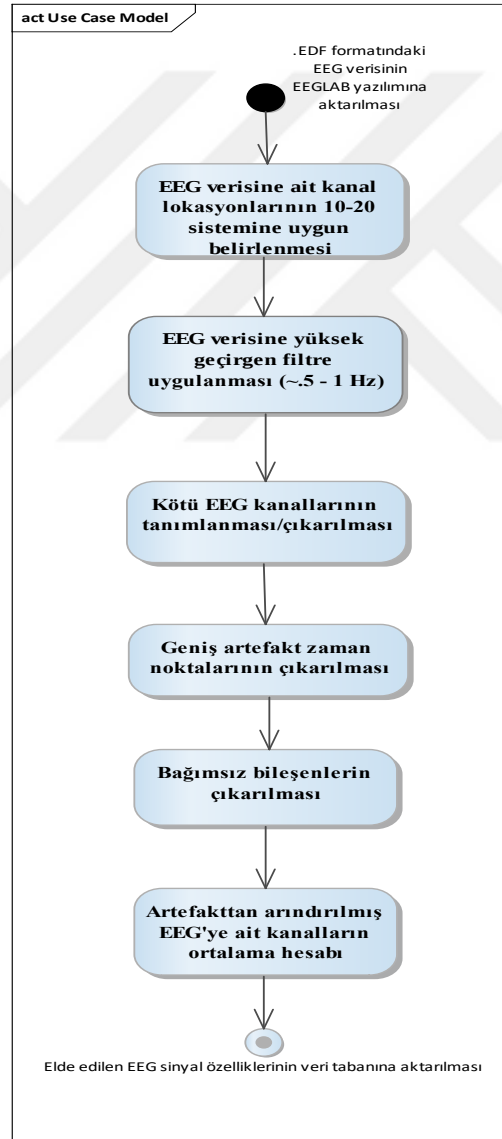
Bu tez çalışmasında analiz edilmiş 97'si epilepsi, 25'i kontrol grubu sağlıklı bireye ait EEG kayıtları Nicolet EEG [98] cihazı aracılığıyla çekilmiştir. Bu çalışmada, bireylere ait çekilmiş uyanıklık EEG verisi kullanılmıştır. EEG incelemeleri, 10–20 sistemine göre yerleştirilen elektrotlar ile 30 dakikalık kayıtlar olarak elde edilmiştir. Şekil 3.25(a)'daki daire şeklindeki siyah noktalarla belirtilen alanlar 10–20 elektrot sisteminde elektrotların yerleştirilme düzenini gösterirken Şekil 3.25(b)'de rutin EEG incelemesinde aktivasyon ve süresini göstermektedir. Hastalardan alınan EEG verileri Nicolet EEG cihazında elde edilen .e formatından Avrupa Veri Formatı (European Data Format – EDF)na dönüştürülerek çıkarılmıştır [99].



Şekil 3.25: Standart 10–20 sistemi elektrot konumları – rutin EEG incelemesinde aktivasyon yöntemleri ve süreleri.

EEG sinyali artefakt kirleticiler denilen hareket artefaktı (titreme, çığneme, emme), kas artefaktı, EKG artefaktı, ter artefaktı, göz kırpması, öksürme gibi istem dışı hareketlerle

kolaylıkla bozulabildiği için sinyaldeki bozulmaları gideren bilgisayar yazılımları zamanla ihtiyaç haline gelmiştir. Sinyal işleme, sinyalde bulunan anlamlı bilgilerin elde edilebilmesi için kullanılan yöntemleri içermektedir. Sinyal işleme ile bir sinyalin özellikleri zaman–frekans uzaylarında eşzamanlı olarak incelenmektedir. Bu teze ait oluşturulacak epilepsi veri setindeki hastalara ait EEG verisi sinyal işleme yöntemleriyle işlenerek hastalıkla ilgili çıkarılan anlamlı özellikleri epilepsi veri setine eklenmiştir. EEG verisinden anlamlı özelliklerin çıkarılmasında EEGLAB [100] yazılım aracı kullanılmıştır. EEG sinyal verisinden anlamlı özelliklerin çıkarılmasında uygulanan işlem adımları Şekil 3.26’da aktivite diyagramında gösterilmiştir.

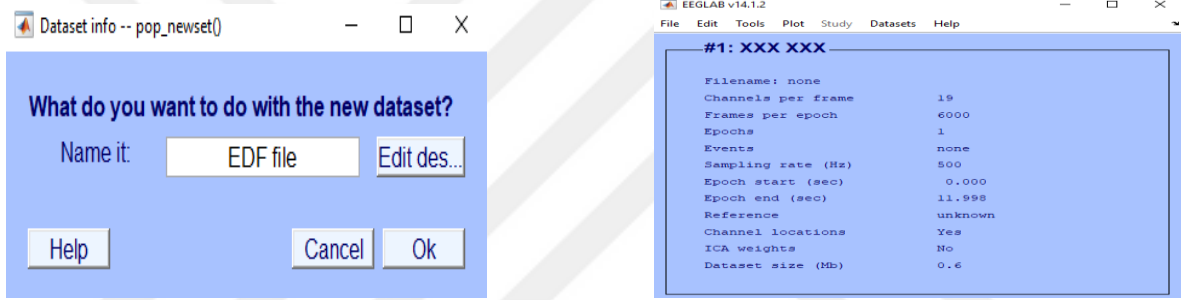


Şekil 3.26: EEG sinyal verisinden anlamlı özelliklerin çıkarılmasında uygulanan işlem adımları.

Şekil 3.26'ya göre hasta bireye ait örnek EEG sinyal verisi için anlamlı özelliklerin çıkarılmasında uygulanan işlem adımları aşağıda açıklanmıştır.

Adım 1: .edf formatındaki EEG verisinin EEGLAB yazılımına aktarılması aşamasında epilepsi ve kontrol grubu sağlıklı bireylere ait EEG verisi Şekil 3.27(a)'daki gibi .edf formatında yazılıma aktarılmıştır.

Adım 2: EEG verisine ait kanal lokasyonlarının 10–20 sistemine uygun belirlenmesi aşamasında epilepsi ve kontrol grubu sağlıklı bireylere ait EEG verisi 19 kanal olarak Şekil 3.27(b)'deki gibi tanımlanmıştır.



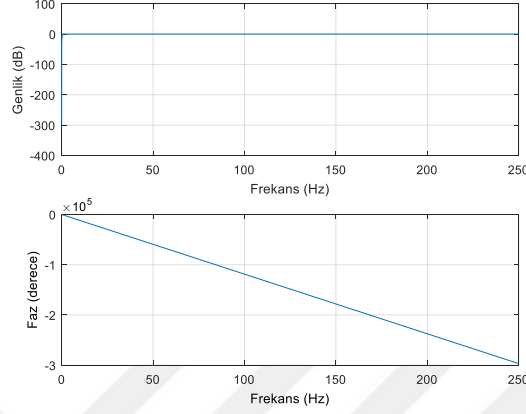
(a) .EDF formatındaki EEG verisinin yazılıma aktarılması (b) 19 kanallı EEG verisinin tanımlanması

Şekil 3.27: EEG verisinin EEGLAB yazılımına aktarılması ve 10–20 kanal olarak tanımlanması.

Adım 3: EEG verisine yüksek geçirgen filtre uygulanması ($\sim 0.5 - 1$ Hz) aşamasında EEG sinyalinden hastalıkla ilgili anlamlı özelliklerin çıkarılması için frekans alanı özellikleri sayısal filtre yöntemlerinden Sonlu Dürtü Yanıtı filtresi kullanılmıştır. Sonlu Dürtü Yanıtı filtresi, EEG sinyal verisindeki artefaktların temizlenmesinde yaygın olarak kullanılmaktadır [101, 102, 103]. Sonlu Dürtü Yanıtı; birim darbe cevapları sonlu sayıda terim içeren tamamen kararlı filtreler olduğundan EEG sinyalinin işlenmesinde kullanılmaktadır. Filtre çıkışı, sadece anlık giriş ve geçmiş çıkış değerlerine bağlıdır ve filtre çıkışından girişine geri besleme olmamaktadır. Veri setindeki her bir hastaya ait elde edilen EEG sinyaline ait giriş $x(n)$ ile çıkış ise $y(n)$ ile ifade edilen sayısal bir sistemin fark denklemi Denklem 3.47'deki gibi ifade edilmektedir.

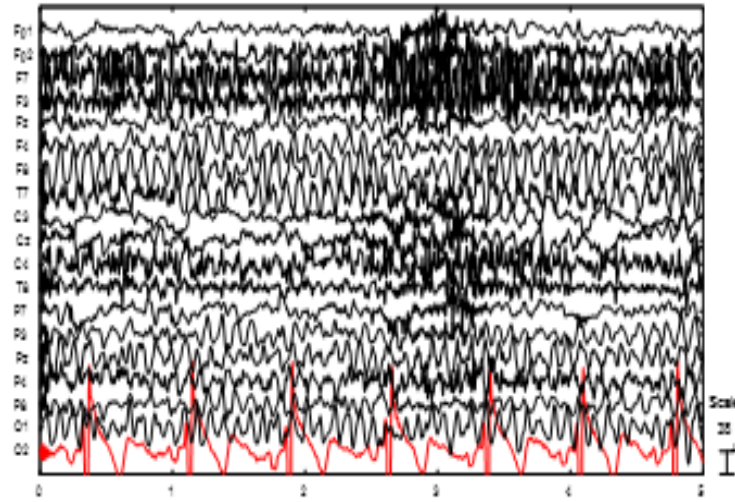
$$y(n) = b_0x(n) + b_1x(n-1) + \dots + b_mx(n-M) - a_1y(n-1) \dots - a_Ny(n-N) \quad (3.47)$$

EEG sinyalinde olası geniş amplitüdün çıkarılması amacıyla EEG'den bağımsız bileşenlerin hesaplanması için düşük geçirgenlikli band genişliği 0,5 ve yüksek geçirgenlikli band genişliği 40 olarak uygulanarak Şekil 3.28'deki grafik elde edilmiştir.



Şekil 3.28: EEG sinyaline yüksek geçirgen filtre uygulanması.

Adım 4: Kötü EEG kanallarının tanımlanması/reddedilmesi ve geniş artefakt zaman noktalarının reddedilmesi aşamalarında Kurtosis yöntemi [104] kullanılmıştır. Şekil 3.29'da Sonlu Dürtü Yanıtı filtresinden geçirilmiş bireye ait EEG sinyalindeki O2 kanalı geniş artefaktlı olduğundan veri setinden çıkarılmıştır ve O2 artefaktından arındırılan EEG verisi 18 kanal olarak elde edilmiştir.



Şekil 3.29: Kötü EEG kanallarının tanımlanması/reddedilmesi.

Adım 5: Filtreden geçirilerek kötü kanallardan arındırılan EEG sinyalinden bağımsız bileşenlerin hesaplanması aşamasında Şekil 3.30’da gösterildiği gibi kötü kanallardan arındırılmış kanal sayısı ile aynı sayıda bağımsız bileşenler hesaplanmıştır.

Bağımsız Bileşen Analizi (BBA), çok değişkenli verilerin bağımsız bileşenlerinin doğrusal birleşimi olarak hesaplanan istatistiksel yöntemdir. BBA sayesinde, 19 farklı kanaldan karışmış olarak gelen EEG sinyallerini birbirinden ayırt edebilmektedir. BBA çok değişkenli verilerin bir dizi bağımsız bileşenin doğrusal bileşeninden oluşmaktadır. Her bir n adet noktadan örneklenmiş p adet değişkenden oluşan veri kümesi W matrisi ile temsil edilmektedir.

W matrisi Denklem 3.48’deki gibi matris çarpımı ile gösterilmektedir. A karışım matrisini, Y bağımsız bileşenleri içeren kaynak matrisini göstermektedir.

$$W = AY \quad (3.48)$$

BBA hesaplamasında Y matrisi Denklem 3.49’a göre hesaplanmaktadır.

$$Y = A^{-1}Z \quad (3.49)$$

Denklem 3.47’ye göre BBA modelinin tanımlı olabilmesi için BBA’nın normalden farklı bir dağılım göstermesi gerekmektedir. Karışım sayısının da BBA sayısına eşit olduğu varsayılmaktadır [104], aksi durum normal-dışılık olarak bilinmektedir. Böyle bir durumda, bağımsız bileşenlerden birini kestirmek için Denklem 3.50 kullanılmaktadır.

$$R = W^T Z \quad (3.50)$$

Denklem 3.48’e göre belirlenecek matris olan W , A ’nın tersinin satır değerlerinden biri ise doğrusal dönüşüm bağımsız bileşenlerden birine eşittir. A matrisi bilinmediğinden W matrisini belirlemek için Denklem 3.51’deki rastlantı değeri olan kestirici ile hesaplanabilmektedir.

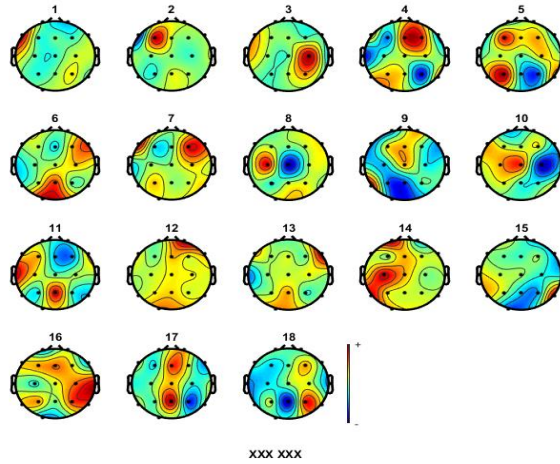
$$R = W^T Z = W^T AY = V^T Y \quad (3.51)$$

R , Y_i ’nin doğrusal birleşimidir. İki bağımsız rastlantı değişkeninin toplamı orjinal değişkenlerden normal dağılıma yaklaşıacağından $V^T Y$, Y_i ’nin herhangi birinden daha çok normal dağılıma uygun davranacaktır. Böylece A , V_i değerlerinden yalnızca biri sıfırdan farklı bir değer olacağı için W , $W^T Z$ değeri normal dışılığı maksimum yapan değer olarak

alınabilmektedir. Gerçekte W matrisleri n boyutlu uzayda normal dışılığın optimizasyonu $2n$ adet lokal maksimum değerli olarak hesaplanmaktadır. Birden çok bağımsız bileşenleri bulabilmek için lokal maksimum değerlerinin hepsi hesaplanmaktadır. Farklı BBA değerleri ilişkisiz olduğu için hesaplama işlemi karmaşık değildir. BBA hesaplanmasında normal–dışılığı kullanabilmek için R değerinin normal normal–dışılık ölçüsü kurtosis ile hesaplanabilmektedir [105]. Kurtosis normal–dışılığın klasik bir ölçütü basıklık (r) ya da 4. Dereceden birikinti değeri Denklem 3.52’ye göre hesaplanmaktadır.

$$Kurt(R) = E(R^4) - 3[E(R^2)]^2 \quad (3.52)$$

R değeri birim varyans olduğundan Denklem 3.52’de $E(R^4)-3$ değerine indirgenmektedir. $E(R^4)$, basıklığın 4. moment değerinin normalize edilmiş değeridir.



Şekil 3.30: Kötü kanallardan arındırılmış EEG sinyalindeki kanallar için bağımsız bileşen analizi.

Adım 6: Artefaktan arındırılmış EEG sinyalindeki her kanala ait ortalama değer hesaplanmıştır. Şekil 3.26’daki aşamalar veri seti için seçilen tüm bireylere ait EEG verisine uygulanmıştır ve sonuçta artefaktan arındırılmış EEG sinyali elde edilmiştir. Oluşturulan EEG özellik setindeki ilk 2 bireye ait EEG sinyalindeki kanallar için elde edilen istatistiksel özellikler Tablo 3.5’te gösterilmiştir. Tablo 3.5’e göre her bireye ait artefaktsız EEG verisindeki her kanalın temsil edilmesi için ilgili kanalın ortalama değeri hesaplanmıştır. Tablo 3.5’te kırmızı ile işaretli kısım EEG verisinden çıkarılan kötü kanaldır. İlgili çıkarılan kanalın veri setinde temsil edilmesi gerekmektedir. Bu yüzden kırmızı ile işaretlenen değer diğer bireylerin artefaktsız EEG sinyalindeki ilgili kanalının ortalaması hesaplanarak aykırı değer (outlier) [105,

106] olarak ifade edilmiştir. Aykırı değer, silinen kanal değeri için ilgili kanalın aritmetik ortalama değeridir. Tablo 3.5'e göre, EEG analiz hesaplamaları sırasında, epilepsi sınıfında 1. kayıttaki hastaya ait O2 kanalı kötü kanal olarak hesaplanmıştır (kırmızı renkle boyalı hücre) ve özellik setinden çıkarılmıştır. Son durumda diğer 97 epilepsi hastası bireye ait O2 kanalının hesaplanan aritmetik ortalama değeri, 1. kayıttaki hastanın O2 kanalına ait aykırı değeri (outlier) olmuştur.

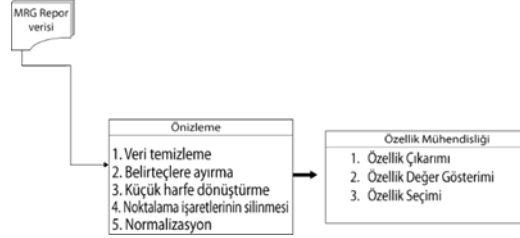
Tablo 3.5: Bireylere ait EEG sinyalinde elde edilen istatistiksel özellikler.

ID	Fp1	Fp2	F7	F3	Fz	F4	F8	T7	C3	Cz	C4	T8	P7	P3	Pz	P4	P8	O1	O2
1	0,023	0,177	0,384	0,241	0,099	0,384	0,094	0,485	0,390	0,315	0,228	0,098	0,470	0,150	0,379	0,861	0,163	0,237	0,804
2	0,049	0,057	0,477	0,071	0,177	0,110	0,209	0,065	0,222	0,603	0,389	0,409	0,535	0,658	0,131	0,342	0,055	0,671	0,971
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...

3.5.1.4. MRG Özellik Setinin Elde Edilmesi

MRG radyoloji raporlarının sürekli güncellenen, entegre ve paylaşılan veriler olan elektronik sağlık kaydı sistemlerine aktarılması, radyoloji araştırma ve uygulamalarında ilerleme potansiyeline yol açmıştır [107]. MRG radyoloji raporlarının çoğunluğu yapılandırılmamış ve serbest biçimli dile sahip olduğu için, bilgiyi elle çıkarmak zaman alıcıdır, genellikle yönetilemez bir görevdir ve insan hatasına eğilimlidir, bu yüzden özel uzmanlık gerektirmektedir. Doğal dil işleme yapılandırılmamış metin verilerinin analizinde yaygın olarak kullanılmaktadır [108]. Elektronik sağlık kayıt sistemlerinde metin tabanlı veri analizi için doğal dil işleme teknikleri; kural tabanlı ve istatistiksel makine öğrenmesi olmak üzere ikiye ayrılmaktadır. Kural tabanlı doğal dil işleme teknikleri, radyoloji raporlarındaki insidans bulgularının kayıtları gibi klinik görevlerde, uzmanlar tarafından önceden tanımlanmış bir dizi anahtar sözcükten faydalanarak hazırlanmaktadır. Makine öğrenmesine dayalı teknikler, sınıflandırmaya ulaşmak için önceden etiketlenmiş rapor içeriğinin sözcüksel ve klinik özelliklerini öğrenmektedir [109]. Son yıllarda, literatürde metin tabanlı klinik, radyolojik ve patolojik raporların sınıflandırılması ile ilgili akut akciğer hasarı, kıkırdak anormallikler tespiti, pulmoner nodüler bulgular, beyin tümörü varlığı, beyin kitleleri tipleri (metastaz, meningioma, gliomalar grade ii, gliomalar grade iii, glioblastomalar), onkolojik yanıt, akut iskemik inme tespiti, meme kanseri tespiti için çalışmalar yapılmıştır [110 – 118]. Literatürde, epilepsi

hastalığını beyin MRG radyoloji raporlarından sınıflandırmak için doğal dil işleme teknikleriyle yapılmış çalışmaların sayısı sınırlıdır. Tezde radyolog tarafından okunmuş beyin MRG raporları için ön işlem ve özellik mühendisliği teknikleri ile uygulanan işlem adımları Şekil 3.31’deki gibidir.



Şekil 3.31: Metin tabanlı radyoloji raporlarının analizi için işlem adımları.

Tezde MRG radyoloji rapor verisine uygulanan ön işlem adımlarının ayrıntıları Tablo 3.6’da gösterilmiştir.

Tablo 3.6: Radyolog tarafından yazılan MRG rapor verisi için ön işlem adımları.

Adım No	Önişleme adımları
1	HTML/XML kodlarının metin içerisinden temizlenmesi,
2	HTML/XML kodlarından temizlenen metnin belirteçlere ayrıştırılması,
3	Rapor metninin küçük harfe dönüştürülmesi
4	Türkçe diline ait bağlaç olan verilerin (örneğin, eğer, ve, için, gibi, vb.), sayıların, noktalama işaretlerinin, sembollerin silinmesi,
5	Normalizasyon

Bir metinde bulunan sözcüklerin geçme sıklığı hakkında bilgi veren Zipf Yasası’na ön işlem adımları uygulanarak elde edilmiş 97’si epilepsi, 25’i kontrol grubu sağlıklı bireylere ait MRG rapor verisinden istatistiksel özellikler çıkarılmıştır [119 – 122]. Dilbilimde bu yasaya göre herhangi bir yazılı metinde geçen sözcükler en sık kullanılanlardan en seyrek kullanılanlara doğru sıralandığında, elde edilen sıralama listesindeki her bir sözcüğün sıra numarası ile o sözcüğün metin içerisinde geçme sıklığı çarpımı her zaman sabit bir sayıyı vermektedir. Zipf Yasası’na göre, N sayıda sözcük bulunan bir metinde, K sıra numarasına sahip sözcüklerin dağılım fonksiyonu $f(K, s, N)$ Denklem 3.53’teki gibidir.

$$f(K, s, N) = \frac{\frac{1}{K^s}}{\sum_{n=1}^N (1/n^s)} \quad (3.53)$$

Denklem 3.53'e göre N sözcüklerin sayısını, K sözcüklerin sırasını, s sözcüklerin dağılımını karakterize eden üssel değeri temsil etmektedir. MRG raporunda bulunan sözcük sayısı N olursa ve klasik Zipf Yasası kullanılırsa s parametresinin üssel değeri 1 olmaktadır [123, 124]. En sık kullanılan sözcüğün dağılımı $f(K; s, N)$ Denklem 3.54'te ifade edilmektedir.

$$f(K; s, N) = \frac{1}{K^s H_{N,s}} ; (H_{N,s} \text{ N. harmonik sayısı}) \quad (3.54)$$

Denklem 3.53 ve Denklem 3.54'e göre, sıra numarası N olan bir sözcük için sıklığın $\frac{1}{N}$ olması gerekmektedir. Böylece en fazla sıklıkla kullanılan sözcük, ikinci sırada sıklıkla kullanılan sözcükten 2 misli daha fazla, üçüncü sıradaki sözcükten 3 misli daha fazla kullanılmaktadır. Bu tez çalışmasında Şekil 3.32'de, Zipf yasasına 97 epilepsi hastası ve 25 kontrol grubu sağlıklı bireylere ait MRG raporları uygulanarak istatistiksel analiz yapılmıştır.

Girdi: Ön işlem uygulanmış 122 bireye ait MRG raporları

Algoritma:

Adım 1: MRG raporunda bulunan sözcük sayısı N hesabı

Adım 2: K sıra numarasına sahip sözcüklerin dağılım fonksiyonu $f(K, s, N)$ hesabı

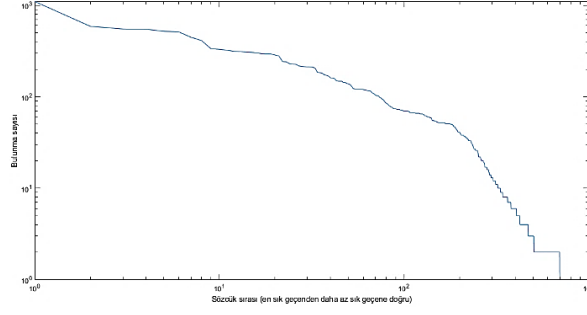
Adım 3: En sık kullanılan sözcüğün dağılımının hesabı

Adım 4: En sık kullanılanlardan en seyrek kullanılan sözcüğe doğru sözcük sıralaması

Çıktı: Epilepsi sözlüğü

Şekil 3.32: Zipf Yasası'na uygulanan MRG raporları için işlem adımları.

Zipf Yasası'na uygulanan MRG raporlarına ait en sık kullanılan sözcükten en seyrek kullanılan sözcüğe doğru yapılan sözcük sıralaması Şekil 3.33'te gösterilmiştir. Şekil 3.33'te görüldüğü gibi baştaki sıralamadaki sözcüklerin sıklığı diğerlerine göre çok daha fazla sıklıkta geçtiği gözlemlenmekte, diğer tüm sözcükler gittikçe azalan sayılarda gözlemlenmektedir. 122 bireye ait MRG rapor verisinde geçen sözcükler sıralandığında 789 sözcükten oluşan bir epilepsi sözlüğü elde edilmiştir ve bu sözlükte hazırlanan tüm metinlerde, en sık kullanılan sözcüğün 1101 defa geçtiği görülmüştür.



Şekil 3.33: Zipf Yasası'na uygulanan MRG raporlarındaki sözcüklerin geçme sıklığı.

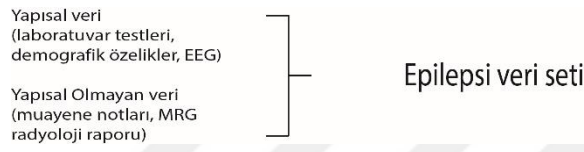
Tablo 3.7’de, Zipf Yasası’na, hasta bireylere ait MRG raporları uygulanarak 122×789 boyutlu Epilepsi_MRG_Zipf adı verilen epilepsi sözlüğü elde edilmiştir. Epilepsi_MRG_Zipf sözlük için doğrusal regresyon modeli [125, 126] geliştirilmiştir. Geliştirilen doğrusal regresyon modelinin başarımı kök ortalama kare sapmaya (Root Mean Squared Error – RMSE) yöntemine göre değerlendirilmiştir. Gerçek epilepsi veri tabanında yer alacak MRG rapor verisindeki en anlamlı sözcüklerin belirlenebilmesi için Epilepsi_MRG_Zipf sözlüğünden elde edilen 9 farklı sözlük için doğrusal regresyon modellemesi yapılmıştır. Tablo 3.7’ye göre, Epilepsi_MRG_Zipf sözlüğünden en seyrek sıklıkta geçen %36,37 sözcüğün çıkarılmasıyla elde edilen Epilepsi_MRG_Zipf_7 sözlüğü için geliştirilen doğrusal regresyon model başarımı en yüksektir. Bu yüzden, Epilepsi_MRG_Zipf_7 sözlüğü gerçek epilepsi veri setine eklenmiştir.

Tablo 3.7: Epilepsi_MRG_Zipf sözlüğü ve bu sözlükten elde edilen 9 farklı özellik seti için oluşturulmuş doğrusal model sonuçları.

Özellik seti adı	Özellik seti açıklama	Özellik seti boyutu	RMSE
Epilepsi_MRG_Zipf	122 bireye ait MRG rapor verisinden elde edilen sözlük	122×789	2,2806
Epilepsi_MRG_Zipf_1	Epilepsi_MRG_Zipf sözlüğünden en sık geçen %5 sözcük çıkarılmıştır	122×750	3,0153
Epilepsi_MRG_Zipf_2	Epilepsi_MRG_Zipf sözlüğünden en seyrek sıklıkta geçen %5 sözcük çıkarılmıştır	122×750	3,4524
Epilepsi_MRG_Zipf_3	Epilepsi_MRG_Zipf sözlüğünden en sık ve en seyrek geçen %5 sözcük çıkarılmıştır	122×711	3,6747
Epilepsi_MRG_Zipf_4	Epilepsi_MRG_Zipf sözlüğünden en sık geçen %10 sözcük çıkarılmıştır	122×709	3,0079
Epilepsi_MRG_Zipf_5	Epilepsi_MRG_Zipf sözlüğünden en seyrek sıklıkta geçen %10 sözcük çıkarılmıştır	122×709	3,0623
Epilepsi_MRG_Zipf_6	Epilepsi_MRG_Zipf sözlüğünden en sık ve en seyrek geçen %10 sözcük çıkarılmıştır	122×629	3,0305
Epilepsi_MRG_Zipf_7	Epilepsi_MRG_Zipf sözlüğünden en sık geçen %36,37 sözcük çıkarılmıştır	122×502	2,1741
Epilepsi_MRG_Zipf_8	Epilepsi_MRG_Zipf sözlüğünden en seyrek sıklıkta geçen %15 sözcük çıkarılmıştır	122×671	3,8609
Epilepsi_MRG_Zipf_9	Epilepsi_MRG_Zipf sözlüğünden en sık ve en seyrek geçen %15 sözcük çıkarılmıştır	122×554	3,0958

3.6. GERÇEK EPİLEPSİ VERİ SETİNİN OLUŞTURULMASI

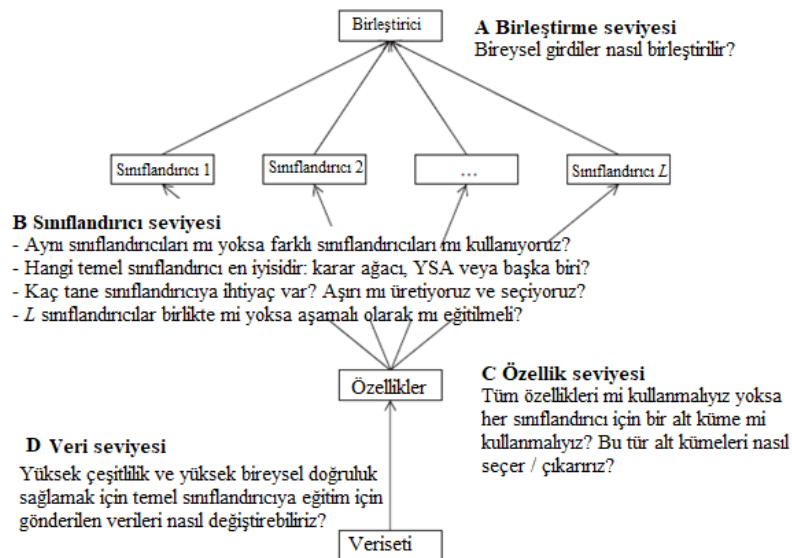
Şekil 3.34’te gösterilen gerçek epilepsi veri seti; demografik veriler ve klinik muayene sırasında epikriz raporlarından elde edilen klinik parametreler (122×48), laboratuvar bulgularından elde edilen kan değerleri (122×34), sinyal işleme tekniklerinden elde edilen EEG sinyal özellikleri (122×19) ve beyin MRG raporlarından elde edilen metinsel özelliklerden (122×502) oluşturulmuştur. Deneysel çalışmaların sonucunda gerçek epilepsi veri seti 122×603 boyutlu bir matris olarak elde edilmiştir.



Şekil 3.34: Gerçek epilepsi veri seti.

3.7. GEÇ FÜZYON TEKİNİĞİNE ÖZELLİK SETLERİNİN UYGULANMASI

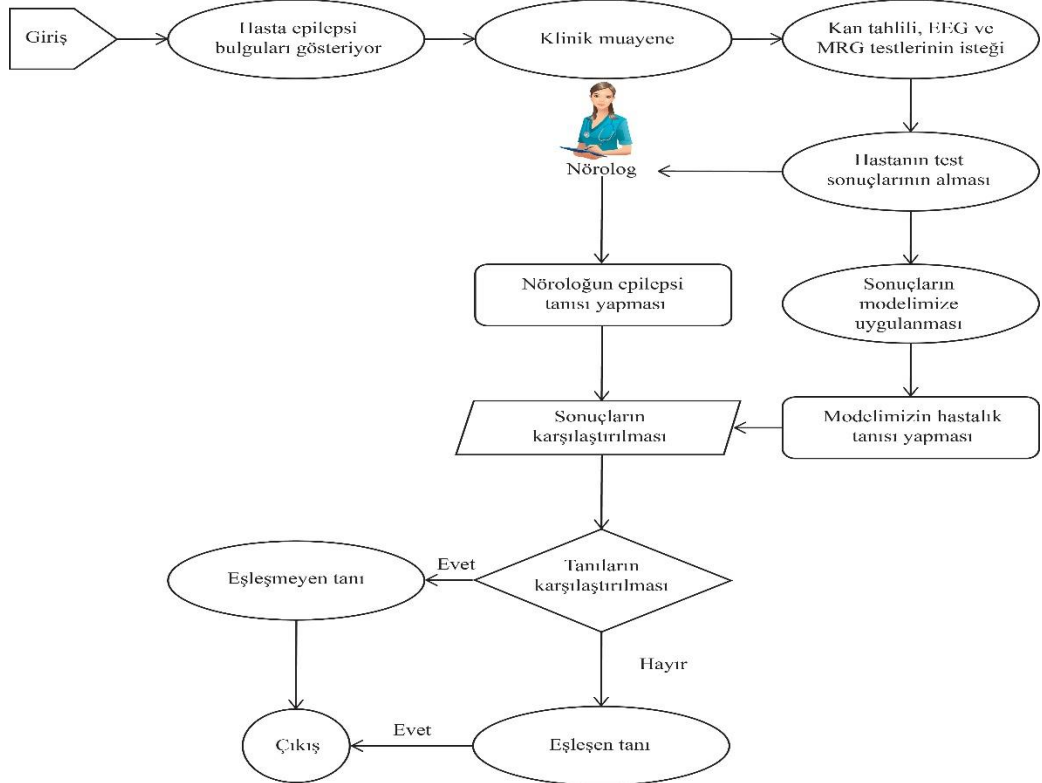
Farklı kaynaklardan gelen veriler doğası gereği farklı özellik taşırlar. Tüm özellikleri bir araya getirerek tüm veri setini tek bir sınıflandırıcı ile sınıflandırmak pratik bir yol değildir. Bunun yerine, Şekil 3.35’e göre, özelliklerin farklı gruplar halinde ayrılarak sınıflandırılması ve sınıflandırıcıların çıktılarını birleştirilmesi daha başarılı sınıflandırma sonucu üretmektedir [82].



Şekil 3.35: Sınıflandırıcı toplulukları oluşturmada dört seviye soru.

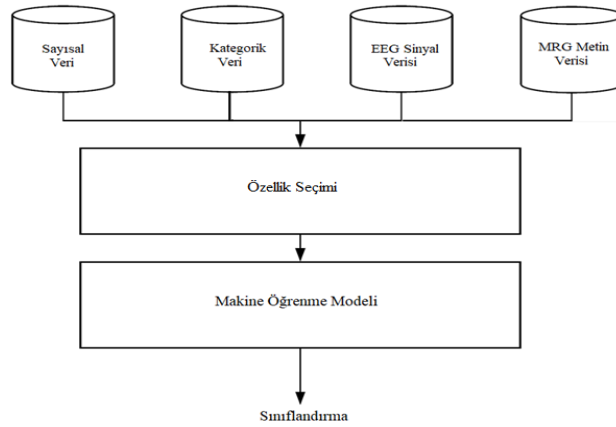
Epilepsi hastalığının tespit edilmesindeki hata oranını düşürmek için, görüntü ve sinyal verisinin birlikte kullanılması ile beyindeki bozukluklar hakkında daha fazla bilgi sağlanmaktadır. Ayrıca, bu tez için önerilen çok modlu yaklaşım sistemin özgüllüğünü artırmaktadır [82].

Bu tez çalışmasında oluşturulan epilepsi veri seti; klinik Parametreler, laboratuvar Bulguları, EEG, MRG olmak üzere 4 farklı özellik setinden elde edilmiştir. Literatürde de bu çalışmadaki gibi farklı özellik setlerinden oluşturulan veri setleri Özellik Çantalama gibi farklı modelleme yöntemleri ile modellenmiştir. Teorik ve ampirik sonuçlar göstermiştir ki; sınıflandırıcılar bağımsız yani bir sınıflandırıcı diğer sınıflandırıcı ile korele değil ise sınıflandırıcıların birleştirilmesi optimum olarak iyileştirme sağlamaktadır. Bir başka deyişle öznitelik bölümlene yöntemleri veri bölümlene yöntemlerinden daha iyi sınıflandırma başarımı gösterebilmektedir [126, 127]. Şekil 3.36’da bir hastanın klinik olarak nörolog tarafından epilepsi tanısı adımlarının, modelimizle birlikte nasıl çalıştığı gösterilmiştir.



Şekil 3.36: Hasta bir bireye nörolog tarafından epilepsi tanısı konulması adımları ile bu tez çalışmasındaki epilepsi sınıflandırma modelinin çalışma adımları.

Şekil 3.36'ya göre, nörolog tarafından epilepsi hastalığının tespiti de farklı kaynaklardan elde edilen özelliklerin birleştirilmesiyle yapılmaktadır. Bu tez çalışmasında da klinik parametreler, laboratuvar bulguları, EEG ve MRG ile elde edilen özellikler çok modlu sınıflandırma yaklaşımına göre ayrı ayrı modellenmiştir. Her bir modellemeden elde edilen sınıflandırma doğruluk oranları birleştirilerek epilepsi hastalığının sınıflandırılması yapılmıştır. Bu tez çalışmasında aralarında korelasyonu bulunmayan 4 farklı alt özellik seti oluşturulmuştur. Bu tez çalışmasında, tek bir işlev gören tek bir sınıflandırıcı oluşturmak yerine sınıflandırmada daha istikrarlı ve başarılı performans elde etme yeteneğini gösteren çok modlu sınıflandırma yaklaşımı önerilmiştir. Bu tez çalışmasında, iki sınıflı görev için farklı özellik setleri üzerinde eğitilmiş farklı sınıflandırıcılardan; ID3 karar ağacı, doğrusal regresyon, Naive Bayes, DVM, k-NN, FFBP ve derin öğrenme yaklaşımlarından LSTM yöntemlerinin sonuçları karar tabanlı tekniğine göre ile ortalama füzyon kuralına göre kaynaştırılmıştır. Şekil 3.37'ye göre, farklı kaynaklardan gelen veriyi işlemek için makine öğrenme yöntemleri ile çok girdili bir model önermektedir. Veri setinde sadece sayısal ve kategorik verileri değil, aynı zamanda sinyal, metin verileri de içermektedir. Geliştirdiğimiz sınıflandırma modelinin çoklu türdeki karışık veriyi çoklu giriş (yani aynı türde olmayan) olarak kabul eden ve bu girdiler hakkında tahmin hesaplayabilme yeteneğinin; sayısal, kategorik, sinyal, metin verisi gibi çoklu veri girişini kabul edebilen, karışık veri girişlerinden uçtan uca modelleyebilen, çoklu veri girişleri kullanarak modelin başarımının değerlendirebilen olması gerekmektedir. Makine öğreniminde, karışık veriler birden çok bağımsız veri türüne sahip olma kavramını ifade eder. Sayısal, kategorik ve sinyal, metin verisi dahil olmak üzere birden çok girişi kabul edebilen bir mimari tanımlanmıştır. Daha sonra bu karışık veriler üzerinde tek bir uçtan uca ağ eğitilmiştir. Epilepsi ve kontrol grubu sağlıklı bireylerin sınıflandırılması için optimum modelleme geliştirilmiştir.



Şekil 3.37: Epilepsi hastalığının sınıflandırılmasında çok girdili modelleme.

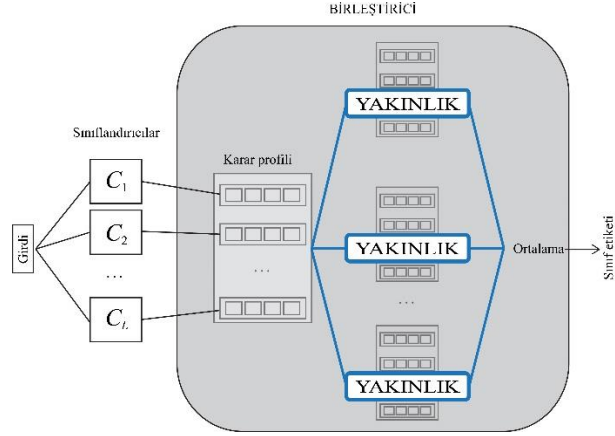
Şekil 3.37’de sayısal / sürekli veri; yaş ve kan değerlerini, kategorik veri; cinsiyet verisini, sinyal verisi; EEG’yi, metin verisi; MRG radyolojik rapor verisinden elde edilen özellikleri temsil etmektedir. Bu tez çalışmasında oluşturulan gerçek epilepsi veri setinin (122×603) özet açıklaması Tablo 3.8’de verilmiştir.

Tablo 3.8: Gerçek epilepsi veri setinin özet açıklaması.

Özellik Seti	Özellik Sayısı	Parametreler	Özelliklerin Açıklaması
Klinik Parametreler	48	Demografik özellikler (2)	Yaş; Cinsiyet
		Epikriz rapor sonuçları (42)	Şuur Kaybı; Terleme; Uyku Bozukluğu; Baş Ağrısı; Konuşmada Zorlanma; Jeneralize Tonik–Klonik Nöbet (senede); Görme Bozukluğu; El–Kollarda Uyuşma; Bas Dönmesi; Göz Kararması; Bilinç Kaybı; Halsizlik; Unutkanlık; Dil Dolandırma; Bayılma Hissi; Bayılma; Denge Bozukluğu; Nefes Daralması; Uyku Hali; Geç Algılama; Dalma(Donup Kalma); Ateş; Bulantı; Sağ El Uyuşması; Ağızda Uyuşma; Kafa Travması; Anlamsız Hareketler; Kusma; Kafa Karışıklığı; El Ayakta Atma; Tüm vücutta titreme; Tonik Kasılma; Uykuda Kasılma; Çenede Kilitlenme; Sağ Yüzde Uyuşma; Sol Kulak Ağrısı; Ellerde Uyuşma; Sağ Tarafda Güçsüzlük; Sol Bacakta Kasılma; Sol Kulakta İşitme Kaybı; Sol Göz Görme Kaybı; Sol Kolda Kasılma;
		Diğerleri (4)	Alkol Kullanımı; Sigara Kullanımı; Ayda Geçirdiği Atak Sayısı; Epilepsi soy geçmişi
Laboratuvar Bulguları	34	Kan Değerleri	NEUT%; LYMPH%; MONO%; EO%; BASO%; IG%; IG#; PDW; MPV; P–LCR; PCT; NEUT#; LYMPH#; MONO#; EO#; ASO#; WBC; RBC; HGB; HCT; PLT; MCV; MCH; MCHC; RDW–SD; RDW–CV; TSH (SENSİTİF); VİTAMİN B–12; ALT(ALANİN, AMİNOTR.,SGPT); AST (ASPARTAT AMİNOTR., SGOT); Kreatinin; CKD–EP (eGFR); Glukoz (Açlık Kan Şekeri); Sedimentasyon
EEG	19	Sinyal özellikleri	Fp1;Fp2;F7;F3;Fz; F4; F8; T7; C3; Cz; C4; T8; P7; P3; Pz; P4; P8; O1; O2
MRG	502	Metin özellikleri	patern; sirkülasyon; void; forme; parasellar; eksternal; kranioservikal; büyük; kemik; intrakranial; kortikomedulla; kontrast; kraniofasial; yer; osteofit; pansinuzit; parietooksipital; pontomezensefalik; post; prepostantral; rezorpsiyon; rosenmuller; sift; spleniumu; suboptimal; tail; tebral; thornwaldt; tip; torus; tubarius; yağlanma; yoğun; yöntem; yıl; çevresel; ...vb.*

*MRG raporlarından çıkarılan toplam özellik sayısı 502 olduğundan dolayı tüm özellikler Tablo 3.8’de verilememiştir.

Şekil 3.38’e göre her bir sınıflandırıcının karar profillerinin (DP) birleştirilmesinin genel yapısı gösterilmektedir.



Şekil 3.38: Karar profillerinin (DP) birleştirilmesinin genel yapısı.

Şekil 3.38'e göre, bu tez çalışması için geliştirilen esnek model ile farklı kaynaklardan elde edilen verileri işlemek için makine öğrenme yöntemlerini içeren çok girdili bir model tanımlanmıştır.

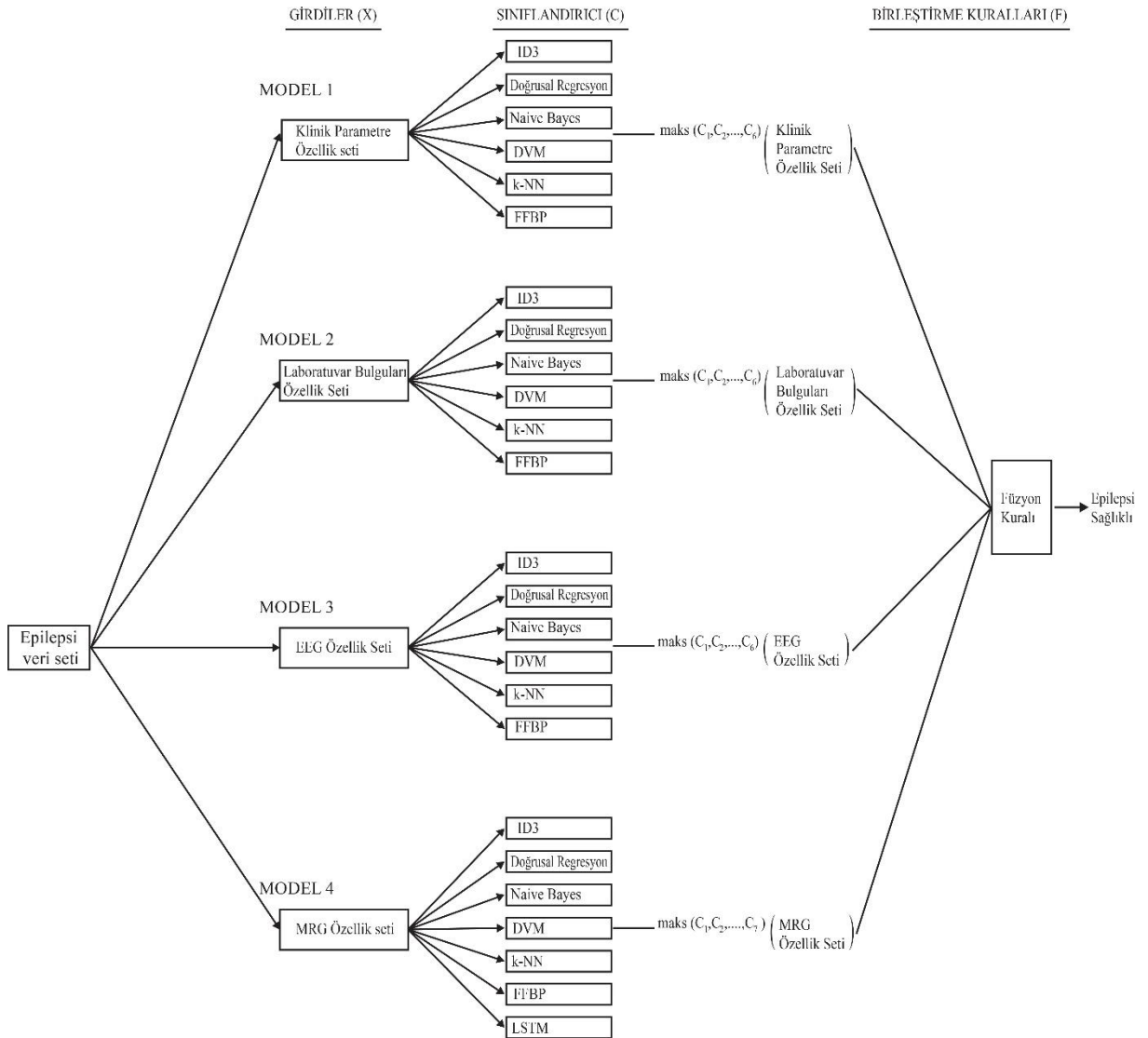
Girdi: Klinik parametre özellik seti, laboratuvar bulguları özellik seti, EEG özellik seti ve MRG özellik seti

Sınıflandırıcılar: C₁: ID3, C₂: Doğrusal regresyon, C₃: Naive Bayes, C₄: DVM, C₅: k-NN, C₆: FFBP, C₇: LSTM

Sınıf etiketi: Epilepsi veya Sağlıklı

Bu tez çalışmasında uygulanan adımlar:

- Topluluk Yöntemleriyle özellik bölümleme; bağımsız alt veri setleri elde etmek için eğitim setinden alt veri setleri üretilmiştir. Her bir alt özellik seti sınıflandırıcı ile eğitilir. Test aşaması eğitimin içerisinde 10-kat çaprazlama yöntemi kullanılmıştır.



Şekil 3.39: Epilepsi hastalığının sınıflandırılmasında çok modlu karar tabanlı veri füzyonu modeli.

Örnek test veri seti için önerilen çok modlu karar tabanlı veri füzyonu modeli kaba kodu Algoritma 1’de gösterilmiştir.

Algoritma 1: Örnek test veri seti için modelleme

- Adım 1 **Girdi:** Test veri seti $D = [x_1, \dots, x_n]^T$
 Adım 2 **Çıktı:** Sınıflandırma sonucu Epilepsi (0) veya Sağlıklı (1), $z \rightarrow (0, +1)$
 Adım 3

Sınıflandırıcılara test veri setindeki verinin uygulanması

```

foreach  $x$  in  $D$  do
    tahmin( $x$ ) = ID3( $x$ );
    if (tahmin( $x$ ) == gerçek( $x$ ))
         $ID3\_doğruluk$  ++;
    tahmin( $x$ ) = DoğrusalRegresyon( $x$ );
    if (tahmin( $x$ ) == gerçek( $x$ ))
         $DoğrusalRegresyon\_doğruluk$  ++;
    tahmin( $x$ ) = NaiveBayes( $x$ );
    if (tahmin( $x$ ) == gerçek( $x$ ))
         $NaiveBayes\_doğruluk$  ++;
    tahmin( $x$ ) = DVM( $x$ );
    if (tahmin( $x$ ) == gerçek( $x$ ))
         $DVM\_doğruluk$  ++;
    if (tahmin( $x$ ) == gerçek( $x$ ))
        tahmin( $x$ ) = k-NN( $x$ );
         $k - NN\_doğruluk$  ++;
    tahmin( $x$ ) = FFBP( $x$ );
    if (tahmin( $x$ ) == gerçek( $x$ ))
         $FFBP\_doğruluk$  ++;
    tahmin( $x$ ) = LSTM( $x$ );
    if (tahmin( $x$ ) == gerçek( $x$ ))
         $LSTM\_doğruluk$  ++;
end foreach
 $C_1$  =  $ID3\_doğruluk$ ;
 $C_2$  =  $DoğrusalRegresyon\_doğruluk$ ;
 $C_3$  =  $NaiveBayes\_doğruluk$ ;
 $C_4$  =  $DVM\_doğruluk$ ;
 $C_5$  =  $k - NN\_doğruluk$ ;
 $C_6$  =  $FFBP\_doğruluk$ ;
 $C_7$  =  $LSTM\_doğruluk$ ;

```

Adım 4 Maksimum Sınıflandırıcının Hesaplanması

Model[j] = maksimum ($C_1, C_2, C_3, C_4, C_5, C_6, C_7$)
 // j : Model numarası, k : Model sayısı, $j = 1 \dots k$

Algoritma 1'deki adımlar epilepsi veri setini oluşturan her bir özellik setine aşağıdaki sırayla uygulanmıştır. Algoritma 1'in klinik parametre özellik setine uygulanması ile Model 1, laboratuvar bulguları özellik setine uygulanması ile Model 2, EEG özellik setine uygulanması ile Model 3, MRG özellik setine uygulanması ile Model 4 oluşturulmuştur. Karar profili DP Algoritma 2'deki ortalama birleştirici ile belirlenmektedir.

Algoritma 2: Epilepsi veri seti için ortalama birleştirici yöntemi ile çok modlu karar tabanlı veri füzyonu modeli

Adım 1 Eğitim: Yok

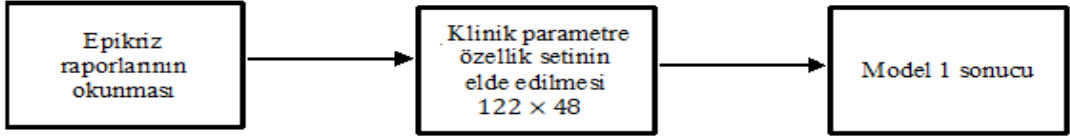
Adım 2 Çıktı: Sınıflandırma sonucu Epilepsi (0) veya Sağlıklı (1), $z \rightarrow (0, +1)$

$$DP = \text{avg} (\text{Model 1, Model 2, Model 3, Model 4})$$

Ortalama birleştirici yöntemi ile Algoritma 2’de DP değeri hesaplanmıştır. DP değeri, bu tez çalışmasında epilepsi veri setinin sınıflandırılması için önerilen çok modlu karar tabanlı veri füzyonu modeli için hesaplanan başarımlar sonucudur.

3.7.1. Algoritma 1’e Klinik Parametre Özellik Setinin Uygulanması

Algoritma 1’e Tablo 3.8’de ifade edilen klinik parametreler özellik seti uygulanmasıyla Model 1 oluşturulmuştur. Model 1 için geliştirilen modelleme Şekil 3.40’taki gibidir.



Şekil 3.40: Model 1.

Model 1: Klinik parametre özellik seti için modelleme

- Adım 1 **Girdi:** Test veri seti $D = [x_1, \dots, x_n]^T$
 Adım 2 **Çıktı:** Sınıflandırma sonucu Epilepsi (0) veya Sağlıklı (1), $z \rightarrow (0, +1)$
 Adım 3

Sınıflandırıcılara test veri setindeki verinin uygulanması

```

foreach  $x$  in  $D$  do
    tahmin( $x$ ) = ID3( $x$ );
    if (tahmin( $x$ ) == gerçek( $x$ ))
         $ID3\_doğruluk$  ++;
    tahmin( $x$ ) = DoğrusalRegresyon( $x$ );
    if (tahmin( $x$ ) == gerçek( $x$ ))
         $DoğrusalRegresyon\_doğruluk$  ++;
    tahmin( $x$ ) = NaiveBayes( $x$ );
    if (tahmin( $x$ ) == gerçek( $x$ ))
         $NaiveBayes\_doğruluk$  ++;
    tahmin( $x$ ) = DVM( $x$ );
    if (tahmin( $x$ ) == gerçek( $x$ ))
         $DVM\_doğruluk$  ++;
    if (tahmin( $x$ ) == gerçek( $x$ ))
        tahmin( $x$ ) = k-NN( $x$ );
         $k - NN\_doğruluk$  ++;
    tahmin( $x$ ) = FFBP( $x$ );
    if (tahmin( $x$ ) == gerçek( $x$ ))
         $FFBP\_doğruluk$  ++;
end foreach
 $C_1 = ID3\_doğruluk$ ;
 $C_2 = DoğrusalRegresyon\_doğruluk$ ;
 $C_3 = NaiveBayes\_doğruluk$ ;
 $C_4 = DVM\_doğruluk$ ;
 $C_5 = k - NN\_doğruluk$ ;
 $C_6 = FFBP\_doğruluk$ ;
  
```

- Adım 4 Maksimum Sınıflandırıcının Hesaplanması
 Model 1 = maksimum ($C_1, C_2, C_3, C_4, C_5, C_6$)
-

Model 1 ile, ID3, doğrusal regresyon, Naive Bayes, DVM, k-NN, FFBP yöntemlerine uygulanan klinik parametreler özellik setinin uygulama adımları matematiksel olarak açıklanmaktadır.

3.7.1.1. ID3 Yöntemine Klinik Parametre Özellik Setinin Uygulanması

Klinik parametre özellik setindeki epilepsi ve kontrol grubu sağlıklı bireylere ait örnekler Tablo 3.9'da verilmiştir. Tablo 3.9'daki veriler kullanılarak ID3 yöntemine göre karar ağacı oluşturulacaktır. Bunun için öncelikle kök düğüm hesaplanmıştır.

Tablo 3.9: Klinik parametre özellik setinden seçilmiş örnek veri.

Cinsiyet	Yaş	Sınıf
Kadın	24	Epilepsi
Erkek	22	Sağlıklı
Kadın	56	Epilepsi
Kadın	73	Epilepsi
Kadın	35	Sağlıklı
Kadın	54	Epilepsi
Kadın	40	Sağlıklı
Erkek	36	Epilepsi
Erkek	72	Epilepsi
Kadın	35	Sağlıklı
Erkek	28	Sağlıklı
Erkek	81	Epilepsi

Adım 1: Tablo 3.9'a göre, 7'si Epilepsi hastası, 5'i sağlıklı sınıfta bulunan toplam 12 bireye ait veri mevcuttur. Denklem 3.6 ile genel durum entropi değeri hesaplanmıştır.

$$H(p_1, \dots, p_{12}) = \frac{7}{12} \log\left(\frac{12}{7}\right) + \frac{5}{12} \log\left(\frac{12}{5}\right) = 0,242$$

Adım 2: Cinsiyet değişkeni için entropi değeri hesaplanmıştır.

$$H(Kadın) = \frac{4}{7} \log\left(\frac{7}{4}\right) + \frac{3}{7} \log\left(\frac{7}{3}\right) = 0,284$$

$$H(Erkek) = \frac{3}{5} \log\left(\frac{5}{3}\right) + \frac{2}{5} \log\left(\frac{5}{2}\right) = 0,289$$

Adım 3: Cinsiyet değişkeni için hesaplanan entropi değerlerinin ağırlıklı toplamı hesaplanmıştır.

$$Ağırlıklı toplam = \left(\frac{7}{12}\right) \times 0,284 + \left(\frac{5}{12}\right) \times 0,289 = 0,280$$

Adım 4: Cinsiyet değişkeni için kazanım Denklem 3.7 ile hesaplanmıştır.

$$Kazanım (Cinsiyet) = 0,242 - 0,280 = -0,038$$

Yaş değişkeni için entropi ve kazanım değerleri hesaplanmıştır. Bunun için Yaş değişkeninin gruplara ayrılması gerekmektedir.

Adım 1: Yaş değişkeni Tablo 3.9'daki verilere göre Tablo 3.10'daki gibi gruplandırılmıştır.

Tablo 3.10: Tablo 3.9'daki yaş değişkeninin gruplandırılması.

Yaş aralığı	Grup numarası
22 – 28	1. grup
29 – 40	2. grup
41 – 70	3. grup
71 – 81	4. grup

Tablo 3.10'daki her grup için entropi değeri hesaplanmaktadır.

$$H(1. grup) = \frac{1}{3} \log \left(\frac{3}{1} \right) + \frac{2}{3} \log \left(\frac{3}{2} \right) = 0,266$$

$$H(2. grup) = \frac{1}{4} \log \left(\frac{4}{1} \right) + \frac{3}{4} \log \left(\frac{4}{3} \right) = 0,243$$

$$H(3. grup) = \frac{2}{2} \log \left(\frac{2}{2} \right) = 0$$

$$H(4. grup) = \frac{3}{3} \log \left(\frac{3}{3} \right) = 0$$

Adım 3: Yaş değişkeni için hesaplanan entropi değerlerinin ağırlıklı toplamı hesaplanmaktadır.

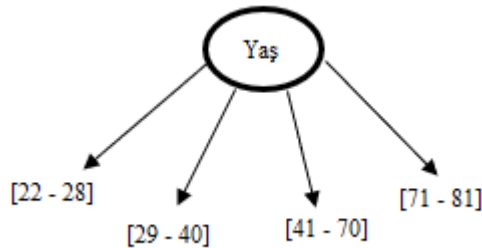
$$\text{Ağırlıklı toplam} = \left(\frac{3}{12} \right) \times 0,266 + \left(\frac{4}{12} \right) \times 0,243 + 0 + 0 = 0,147$$

Adım 4: Yaş değişkeni için kazanım Denklem 3.7 ile hesaplanmaktadır.

$$\text{Kazanım (Yaş)} = 0,242 - 0,147 = 0,095$$

$$\text{Kazanım (Cinsiyet)} = -0,038 \text{ ve } \text{Kazanım (Yaş)} = 0,095 \text{ olarak hesaplanmıştır.}$$

Bu kazanım değerlerinden en büyük olan *yaş* değişkenine aittir. Bu yüzden karar ağacının kök düğümü Şekil 3.41'deki gibi *yaş* değişkeni seçilecektir. Bu adımlar ile epilepsi veri setinden seçilmiş örnek veri için örnek bir karar ağacı hesaplaması yapılmıştır. Tüm bu adımlar klinik parametre özellik setine uygulanmıştır.

**Şekil 3.41:** Klinik parametre özellik seti için örnek ağaç kök düğümü.

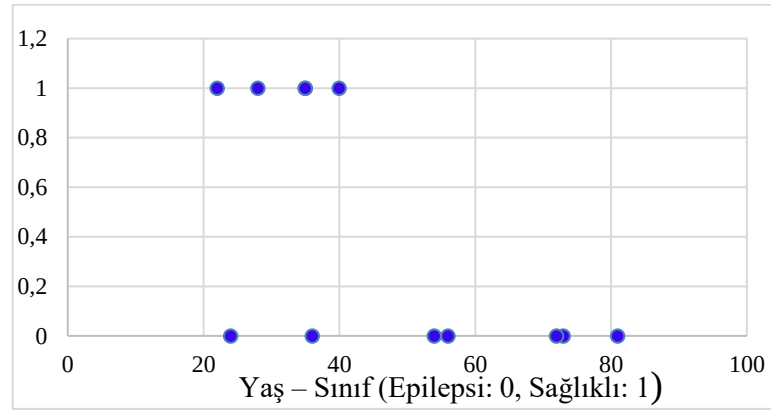
3.7.1.2. Doğrusal Regresyon Yöntemine Klinik Parametre Özellik Setinin Uygulanması

Klinik parametre özellik seti için Tablo 3.11’de yaş bilgisi ve sınıf bilgisi (Epilepsi: 0 ile etiketli ve Sağlıklı: 1 ile etiketli) arasındaki ilişki doğrusal regresyon yöntemiyle hesaplanmıştır.

Tablo 3.11: Yaş – Sınıf bilgisi.

Yaş (x)	Sınıf (y)
24	0
22	1
56	0
73	0
35	1
54	0
40	1
36	0
72	0
35	1
28	1
81	0

Adım 1: Tablo 3.11’deki verilere göre yaş – sınıf arasındaki ilişki Şekil 3.42’deki gibidir.



Şekil 3.42: Tablo 3.11’deki verilere göre Yaş – Sınıf (Epilepsi: 0, Sağlıklı: 1) ilişkisi.

Adım 2: Başlangıç parametreleri olan ağırlıklar θ_0 ve θ_1 ve öğrenme katsayısı belirlenir. Çıkış değerleri (y) Denklem 3.8’e göre hesaplanmıştır. Hesaplanan yeni y değerlerinin gerçek y değerlerinden farkının karesi alınıp toplanarak maliyet hata değeri hesaplanmıştır.

$$h_1(x_1) = \theta_0 + \theta_1 x_1 = 0$$

$$h_2(x_2) = \theta_0 + \theta_1 x_2 = 0$$

$$h_3(x_3) = \theta_0 + \theta_1 x_3 = 0$$

$$h_4(x_4) = \Theta_0 + \Theta_1 x_4 = 0$$

$$h_5(x_5) = \Theta_0 + \Theta_1 x_5 = 0$$

$$h_6(x_6) = \Theta_0 + \Theta_1 x_6 = 0$$

$$h_7(x_7) = \Theta_0 + \Theta_1 x_7 = 0$$

$$h_8(x_8) = \Theta_0 + \Theta_1 x_8 = 0$$

$$h_9(x_9) = \Theta_0 + \Theta_1 x_9 = 0$$

$$h_{10}(x_{10}) = \Theta_0 + \Theta_1 x_{10} = 0$$

$$h_{11}(x_{11}) = \Theta_0 + \Theta_1 x_{11} = 0$$

$$h_{12}(x_{12}) = \Theta_0 + \Theta_1 x_{12} = 0$$

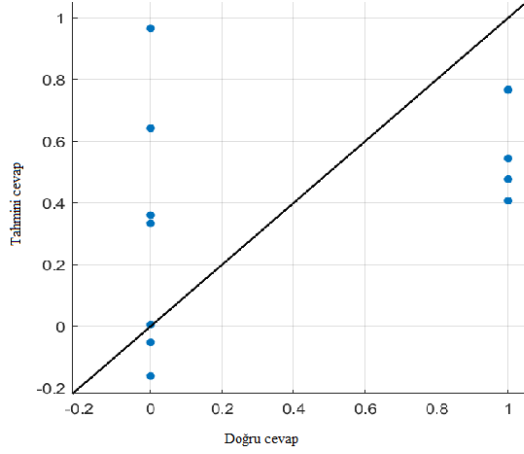
Toplam hata değeri Denklem 3.8'e göre $E(\Theta_0, \Theta_1) = \frac{1}{2N} \sum_{i=1}^N (h_i(x) - y_i)^2 = 1,7890$ olarak hesaplanmıştır.

Adım 3: Toplam hata değeri sıfırdan büyük bir değerdir. Bu fonksiyonu minimum yapacak x değerlerinin bulunabilmesi için Θ_0 ve Θ_1 için türev alınmaktadır.

$$\frac{dx}{d\Theta_0} E(\Theta_0, \Theta_1) = \frac{1}{N} \sum_{i=1}^N (h_i(x) - y_i)$$

$$\frac{dx}{d\Theta_1} E(\Theta_0, \Theta_1) = \frac{1}{N} \sum_{i=1}^N (h_i(x) - y_i) x_i$$

Adım 4: Türevin ters yönünde doğrultu üzerinde hareket edilerek minimum nokta öğrenme katsayısı ile (0,01) ile çarpılarak yeni ağırlıklar hesaplanmıştır. Bu ağırlıklara göre doğrusal doğru denklemi Şekil 3.43'teki gibi elde edilmiştir.



Şekil 3.43: Tablo 3.11’deki verilere göre Yaş –Sınıf değerleri için doğrusal model.

3.7.1.3. Naive Bayes Yöntemine Klinik Parametre Özellik Setinin Uygulanması

Tablo 3.12’de klinik parametreler özellik setine ait örnek veriler kullanılarak $x = \{\text{Erkek}, 51\}$ olarak verilen yeni verinin Epilepsi ve Sağlıklı sınıflarından hangisine dahil olduğu Bayes kuralı ile adım adım hesaplanmıştır.

Tablo 3.12: Klinik özellik setinden seçilmiş örnek veri.

Cinsiyet	Yaş	Sınıf
Kadın	24	Epilepsi
Erkek	22	Sağlıklı
Kadın	56	Epilepsi
Kadın	73	Epilepsi
Kadın	35	Sağlıklı
Kadın	54	Epilepsi
Kadın	40	Sağlıklı
Erkek	36	Epilepsi
Erkek	72	Epilepsi
Kadın	35	Sağlıklı
Erkek	28	Sağlıklı
Erkek	81	Epilepsi
Erkek	51	?

Adım 1: Her bir sınıf için $P(C_j)$ değerleri hesaplanmıştır.

$$P(\text{Epilepsi}) = P(C_1) = \frac{7}{12} = 0,583$$

$$P(\text{Sağlıklı}) = P(C_2) = \frac{5}{12} = 0,416$$

Adım 2: $P(x|C_j)$ değeri $x = \{Erkek, 51\}$ verisi için epilepsi ve sağlıklı sınıflarında olma olasılıkları hesaplanmıştır.

$$P(x|Epilepsi) = \frac{1}{7} = 0,142$$

$$P(x|Sağlıklı) = \frac{1}{5} = 0,2$$

Adım 3: $P_{x(i)}$ değeri $P_{x=\{Erkek,51\}}$ için Denklem 3.11 yardımıyla aşağıdaki gibi hesaplanmıştır.

$$P_{x(i)} = (0,583 \times 0,142) + (0,416 \times 0,2) = 0,165$$

Adım 4: $x = \{Erkek, 51\}$ verisinin Epilepsi sınıfında olma olasılığı Denklem 3.12 ile hesaplanmıştır.

$$P(Epilepsi|x) = \frac{0,583 \times 0,142}{0,165} = 0,501$$

$$P(Sağlıklı|x) = \frac{0,416 \times 0,2}{0,165} = 0,504$$

Bu adımlar ile $x = \{Erkek, 51\}$ verisinin olasılığı daha büyük olan Sağlıklı sınıfına ait olduğu hesaplanmıştır.

3.7.1.4. DVM Yöntemine Klinik Parametre Özellik Setinin Uygulanması

Klinik parametre özellik setine ait 4 farklı bireye ait aile soy geçmişinde epilepsi varlığı ve sigara kullanımı bilgisi kullanılarak DVM yöntemi ile örnek sınıflandırma hesaplaması yapılmıştır.

$$x_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, x_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, x_3 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, y = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, y \text{ bilgisi için Epilepsi sınıfı pozitif bölgede (+1)}$$

ve Sağlıklı sınıfı negatif bölgede (-1) temsil edilmektedir.

Adım 1: Denklem 3.13'teki fonksiyona göre hesaplama yapılmıştır.

$$L(w, b, a) = a_1 + a_2 + a_3 + a_4 - \frac{1}{2}(a_1 a_1 y_1 y_1 x_1^T x_1 + a_1 a_2 y_1 y_2 x_1^T x_2 + a_1 a_3 y_1 y_3 x_1^T x_3 + a_2 a_1 y_2 y_1 x_2^T x_1 + a_2 a_2 y_2 y_2 x_2^T x_2 + a_2 a_3 y_2 y_3 x_2^T x_3 + a_3 a_1 y_3 y_1 x_3^T x_1 + a_3 a_2 y_3 y_2 x_3^T x_2 + a_3 a_3 y_3 y_3 x_3^T x_3)$$

$x_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ değerlerine sahip olduğu için

$$L(w, b, a) = a_1 + a_2 + a_3 - \frac{1}{2}(a_1 a_1 y_1 y_1 x_1^T x_1 + a_1 a_3 y_1 y_3 x_1^T x_3 + a_3 a_1 y_3 y_1 x_3^T x_1 + a_3 a_3 y_3 y_3 x_3^T x_3) \text{ olarak elde edilmiştir.}$$

$$\begin{aligned} L(w, b, a) &= a_1 + a_2 + a_3 \\ &+ (a_1^2(1)^2[0 \ 1] \begin{bmatrix} 0 \\ 1 \end{bmatrix} + a_1 a_3(1)(-1)[0 \ 1] \begin{bmatrix} 1 \\ 1 \end{bmatrix} + a_3 a_1(-1)(1)[1 \ 1] \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ &+ a_3^2(-1)^2[1 \ 1] \begin{bmatrix} 1 \\ 1 \end{bmatrix}) \\ &= a_1 + a_2 + a_3 - \frac{1}{2}(a_1^2 - 2a_1 a_3 + 2a_3^2) \end{aligned}$$

Adım 2: $\sum_{i=1}^3 y_i a_i = 0$ olduğu için $a_1 + a_2 - a_3 = 0$ değeri $a_1 + a_2 = a_3$ hesaplanmıştır.

$L(a) = 2a_3 - \frac{1}{2}(a_1^2 - 2a_1 a_3 + 2a_3^2)$ ile a_1, a_3 değerlerini bulabilmek için fonksiyonun türevi alınarak sıfıra eşitlenmiştir.

$$\frac{\partial L}{\partial a_1} = -2a_1 + 2a_3 = 0 \text{ ile } a_1 = a_3 = 2 \text{ olarak hesaplanmıştır.}$$

Adım 4: $a = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix}$ değerleri elde edilmiştir ve w^* hesaplanmıştır.

$$w^* = \sum_{x_i \in \text{Destekvektör}} a_i y_i x_i = a_1 y_1 x_1 + a_3 y_3 x_3 = 2(1) \begin{bmatrix} 0 \\ 1 \end{bmatrix} + 2(-1) \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \end{bmatrix}$$

Adım 5: $b^* = \frac{1}{N_{\text{Destekvektör}}} \left(\sum_{s=1}^{N_{\text{Destekvektör}}} \left(\frac{1}{y_s} - x_s^T w^* \right) \right)$ ile

$$\begin{aligned} b^* &= \frac{1}{2} \left(\frac{1}{y_1} x_1^T \begin{bmatrix} -2 \\ 0 \end{bmatrix} + \frac{1}{y_3} - x_3^T \begin{bmatrix} -2 \\ 0 \end{bmatrix} \right) = \frac{1}{2} \left(\frac{1}{1} [0 \ 1] \begin{bmatrix} -2 \\ 0 \end{bmatrix} + \frac{1}{(-1)} - [1 \ 1] \begin{bmatrix} -2 \\ 0 \end{bmatrix} \right) \\ &= 1 \end{aligned}$$

Adım 5: Yeni gözlem değerleri için sınıflandırma $f(x) = \text{sign}(w^* x + b)$ genel formülü ile

$$f(x) = \text{sign}\left(\begin{bmatrix} -2 \\ 0 \end{bmatrix} [x_1 + x_2] + 1\right) = \text{sign}(-2x_1 + 1) \text{ olarak hesaplanmıştır.}$$

Adım 6: $x_4 = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$ gibi yeni bir veri için $\text{sign}(-2x_1 + 1) > 0$ olduğundan verinin pozitif bölgede yani Epilepsi sınıfına ait olduğu anlaşılmaktadır.

3.7.1.5. k-NN Yöntemine Klinik Parametre Özellik Setinin Uygulanması

Tablo 3.13'te klinik parametreözellik setine ait örnek veriler kullanılarak $x = \{0,2, 51\}$ olarak verilen yeni verinin epilepsi ve sağlıklı sınıflarından hangisi olduğunu, k-NN yöntemi adımları uygulanarak hesaplanmıştır.

Tablo 3.13: Klinik özellik setinden seçilmiş örnek veri.

Bayılma	Yaş	Sınıf
1	24	Epilepsi
0	22	Sağlıklı
0,48	56	Epilepsi
0,48	73	Epilepsi
0	35	Sağlıklı
1	54	Epilepsi
0	40	Sağlıklı
0,48	36	Epilepsi
0,48	72	Epilepsi
1	35	Sağlıklı
0	28	Sağlıklı
0,48	81	Epilepsi
0,2	51	?

Adım 1: $x = \{0,2, 51\}$ verisi için en yakın 4 komşu ($k = 4$) için hesaplama yapılmıştır.

Adım 2: $x = \{0,2, 51\}$ verisi için her her bir gözlem değerlerinin birbiri arasındaki Öklid uzaklıkları Denklem 3.19'a göre hesaplanmıştır.

$$d(i, j) = \sqrt{(1 - 0,2)^2 + (24 - 51)^2} = 27,014$$

$$d(i, j) = \sqrt{(0 - 0,2)^2 + (22 - 51)^2} = 29,000$$

$$d(i, j) = \sqrt{(0,48 - 0,2)^2 + (56 - 51)^2} = 5,007$$

$$d(i, j) = \sqrt{(0,48 - 0,2)^2 + (73 - 51)^2} = 22,001$$

$$d(i, j) = \sqrt{(0 - 0,2)^2 + (35 - 51)^2} = 16,001$$

$$d(i, j) = \sqrt{(1 - 0,2)^2 + (54 - 51)^2} = 3,104$$

$$d(i, j) = \sqrt{(0 - 0,2)^2 + (40 - 51)^2} = 11,001$$

$$d(i, j) = \sqrt{(0,48 - 0,2)^2 + (36 - 51)^2} = 15,002$$

$$d(i, j) = \sqrt{(0,48 - 0,2)^2 + (72 - 51)^2} = 21,001$$

$$d(i, j) = \sqrt{(1 - 0,2)^2 + (35 - 51)^2} = 16,019$$

$$d(i, j) = \sqrt{(0 - 0,2)^2 + (28 - 51)^2} = 23,000$$

$$d(i, j) = \sqrt{(0,48 - 0,2)^2 + (81 - 51)^2} = 30,001$$

Hesaplanan değerler Tablo 3.14'te gösterilmiştir.

Tablo 3.14: Gözlem değerlerinin $x = \{0,2, 51\}$ noktasına uzaklıkları.

Bayılma	Yaş	$d(i, j)$
1	24	27,014
0	22	29,000
0,48	56	5,007
0,48	73	22,001
0	35	16,001
1	54	3,104
0	40	11,001
0,48	36	15,002
0,48	72	21,001
1	35	16,019
0	28	23,000
0,48	81	30,001

Adım 3: Tablo 3.15'te hesaplanan değerler küçükten büyüğe sıralanmıştır.

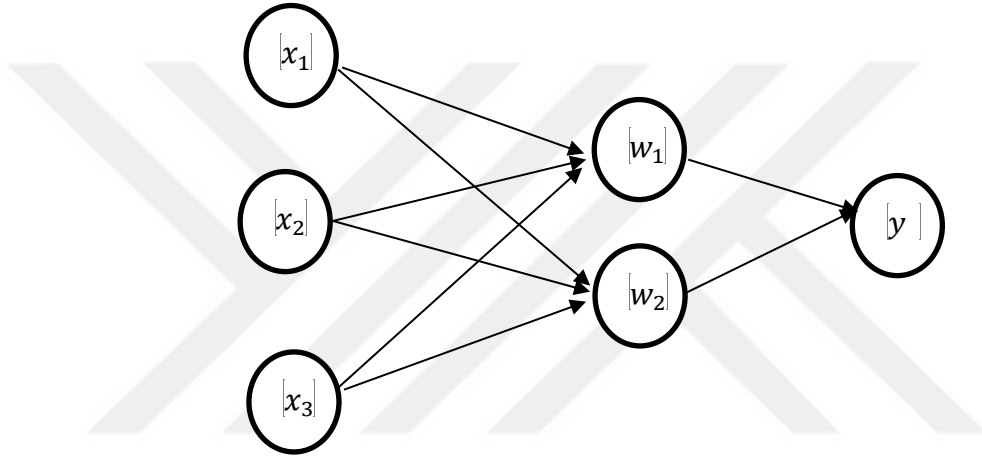
Tablo 3.15: $k = 4$ için komşulukların belirlenmesi.

Bayılma	Yaş	$d(i, j)$	Sıra	k komşunun sınıfı
1	24	27,014	10	
0	22	29,000	11	
0,48	56	5,007	2	Epilepsi
0,48	73	22,001	8	
0	35	16,001	5	
1	54	3,104	1	Epilepsi
0	40	11,001	3	Sağlıklı
0,48	36	15,002	4	Epilepsi
0,48	72	21,001	7	
1	35	16,019	6	
0	28	23,000	9	
0,48	81	30,001	12	

Adım 4: Tablo 3.15’te hesaplanan değerler küçükten büyüğe sıralandığında (0,2, 51) verisine en yakın gözlem değeri, $k = 4$ için en yakın komşuluk içinde 3 tane “Epilepsi”, 1 tane “Sağlıklı” sonucu vardır ve yeni verinin sınıfı “Epilepsi” olarak belirlenmiştir.

3.7.1.6. FFBP Yöntemine Klinik Parametre Özellik Setinin Uygulanması

Şekil 3.44’te klinik parametre özellik setine ait örnek veriler kullanılarak FFBP yöntemiyle çok katmanlı ileri beslemeli yapay sinir ağının hata ve ağırlıklarının uygulama adımları sırasıyla verilmiştir.



Şekil 3.44: İleri beslemeli çok katmanlı sinir ağı.

$$x_1 = 1; x_2 = -1; x_3 = 0$$

$$w_{11} = 0,4; w_{12} = 0,1; w_{21} = 0,3; w_{22} = 0,6; w_{31} = 0,1; w_{32} = 0,2; w_{33} = -0,1; w_{1y} = 0,1; w_{2y} = 0,2$$

$$\theta_1 = 0,1; \theta_2 = 0,1; \theta_y = 0,2$$

Öğrenme katsayısı $\mu = 0,01$ olsun ve ulaşılmak istenen $d = 1$ olsun.

Adım 1: $y_1 = (w_{11} \times x_1) + (w_{21} \times x_2) + (w_{31} \times x_3) + \theta_1 = (0,4 \times 1) + (0,3 \times (-1)) + (0,1 \times 0) + 0,1 = 0,2$

w_1 değeri, $w_1 = \frac{1}{1+e^{-y_1}} = \frac{1}{1+e^{-0,2}} = 0,549$ ve bu değer y düğümüne giriş değeri olarak kullanılmıştır. w_2 düğümüne giriş değeri hesaplanmış ve sonrasında ateşleme fonksiyonu hesaplanmıştır. Böylece y düğümü için ikinci giriş değeri hesaplanmış olacaktır.

$$\begin{aligned} y_2 &= (w_{12} \times x_1) + (w_{22} \times x_2) + (w_{32} \times x_3) + \theta_2 \\ &= (0,1 \times 1) + (0,6 \times (-1)) + (0,2 \times 0) + 0,1 = -0,4 \end{aligned}$$

$$w_2 \text{ düğümü değeri } w_2 = \frac{1}{1+e^{-y_2}} = \frac{1}{1+e^{-0,4}} = 0,401$$

$$y \text{ düğümünün girişi } y = (w_{1y} \times y_1) + (w_{2y} \times y_2) + \theta_y = (0,1 \times 0,2) + (0,2 \times -0,4) + 0,2 = 0,14$$

$$y \text{ düğümü çıkış değeri } y^{çıkış} = \frac{1}{1+e^{-y}} = \frac{1}{1+e^{-0,14}} = 0,534$$

Çıkış için ulaşılmak istenen değer 1 olduğundan y hata değeri hesaplanmıştır.

$$hata_y = y^{çıkış} \times (d - y^{çıkış})^2 = 0,534 \times (1 - 0,534)^2 = 0,115$$

$$hata_{y_1} = w_1 \times (d - w_1)^2 = 0,549 \times (1 - 0,549)^2 = 0,111$$

$$hata_{y_2} = w_2 \times (d - w_2)^2 = 0,401 \times (1 - 0,401)^2 = 0,143$$

Adım 2: Elde edilen hata değerleri ile yeni değerler hesaplanmıştır.

$$w_{11}^{yeni} = w_{11} + (\mu \times hata_{y_1} \times x_1) = 0,4 + (0,01 \times 0,111 \times 1) = 0,401$$

$$w_{12}^{yeni} = w_{12} + (\mu \times hata_{y_2} \times x_1) = 0,1 + (0,01 \times 0,143 \times 1) = 0,101$$

$$w_{21}^{yeni} = w_{21} + (\mu \times hata_{y_1} \times x_2) = 0,3 + (0,01 \times 0,111 \times (-1)) = 0,298$$

$$w_{22}^{yeni} = w_{22} + (\mu \times hata_{y_2} \times x_2) = 0,6 + (0,01 \times 0,143 \times (-1)) = 0,598$$

$$w_{31}^{yeni} = w_{31} + (\mu \times hata_{y_1} \times x_3) = 0,1 + (0,01 \times 0,111 \times (0)) = 0,1$$

$$w_{32}^{yeni} = w_{32} + (\mu \times hata_{y_2} \times x_3) = 0,2 + (0,01 \times 0,143 \times (0)) = 0,2$$

$$w_{1y}^{yeni} = w_{1y} + (\mu \times hata_y \times y_1) = 0,1 + (0,01 \times 0,115 \times 0,2) = 0,100$$

$$w_{2y}^{yeni} = w_{2y} + (\mu \times hata_y \times y_2) = 0,2 + (0,01 \times 0,115 \times (-0,4)) = 0,199$$

$$\theta_1^{yeni} = \theta_1 + (\mu \times hata_{y_1}) = 0,1 + (0,01 \times 0,111) = 0,101$$

$$\theta_2^{yeni} = \theta_2 + (\mu \times hata_{y_2}) = (0,1) + (0,01 \times 0,143) = 0,101$$

$$\theta_y^{yeni} = \theta_y + (\mu \times hata_y) = 0,2 + (0,01 \times 0,115) = 0,201$$

Adım 3: Klinik parametre özellik setindeki öğrenme işlemi her değer için ve verilen hata değerine ulaşıncaya kadar işlemler baştan sona dek yapılıp ve en son ağırlık değeri hesaplanmasıyla tamamlanmıştır.

3.7.1.7. LSTM Yöntemine Klinik Parametre Özellik Setinin Uygulanması

Klinik parametre özellik setine ait örnek veriler kullanılarak LSTM yönteminin uygulama adımları sırasıyla verilmiştir.

Adım 1: İç ağırlıklar tanımlanmıştır.

$$W_a = \begin{bmatrix} 0,4 \\ 0,2 \end{bmatrix}, U_a = [0,1], b_a = [0,3]$$

$$W_i = \begin{bmatrix} 0,1 \\ 0,3 \end{bmatrix}, U_i = [0,4], b_i = [0,8]$$

$$W_f = \begin{bmatrix} 0,5 \\ 0,6 \end{bmatrix}, U_f = [0,7], b_f = [0,1]$$

$$W_o = \begin{bmatrix} 0,8 \\ 0,4 \end{bmatrix}, U_o = [0,9], b_o = [0,2]$$

Adım 2: Sistemin giriş değerleri tanımlanmıştır.

$$x_0 = \begin{bmatrix} 2 \\ 4 \end{bmatrix} \text{ sınıf Epilepsi sınıfı (0 değeri ile etiketli)}$$

$$x_1 = \begin{bmatrix} 1 \\ 0,3 \end{bmatrix} \text{ sınıf Sağlıklı sınıfı (1 değeri ile etiketli)}$$

Adım 3: $t = 0$ için ileri beslemeli ağırlık eğitimi yapılmıştır.

$$\begin{aligned} a_0 &= \tanh(W_a^T \times x_0 + U_a \times out_{-1} + b_a) = \tanh\left(\begin{bmatrix} 0,4 & 0,2 \end{bmatrix} \times \begin{bmatrix} 2 \\ 4 \end{bmatrix} + [0,1] \times 0 + 0,3\right) \\ &= 0,956 \end{aligned}$$

$$\begin{aligned} i_0 &= \sigma(W_i^T \times x_0 + U_i \times out_{-1} + b_i) = \sigma\left(\begin{bmatrix} 0,1 & 0,3 \end{bmatrix} \times \begin{bmatrix} 2 \\ 4 \end{bmatrix} + [0,4] \times 0 + 0,8\right) \\ &= 4,272 \end{aligned}$$

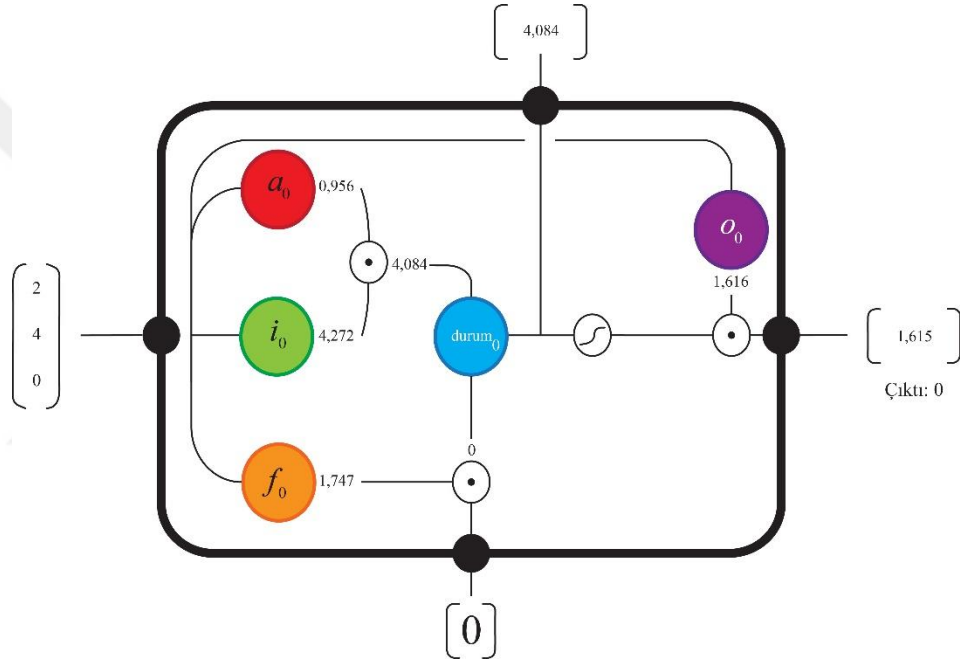
$$f_0 = \sigma(W_f^T \times x_0 + U_f \times out_{-1} + b_f) = \sigma([0,5 \quad 0,6] \times \begin{bmatrix} 2 \\ 4 \end{bmatrix} + [0,7] \times 0 + 0,1) = 1,747$$

$$o_0 = \sigma(W_o^T \times x_0 + U_o \times out_{-1} + b_o) = \sigma([0,8 \quad 0,4] \times \begin{bmatrix} 2 \\ 4 \end{bmatrix} + [0,9] \times 0 + 0,2) = 1,616$$

$$durum_0 = (a_0 \times i_0) + (f_0 \times durum_{-1}) = (0,956 \times 4,272) + (1,747 \times 0) = 4,084$$

$$out_0 = \tanh(durum_0) \times o_0 = \tanh(4,084) \times 1,616 = 1,615$$

$t = 0$ için ileri beslemeli ağı eğitimi işleminin sonucu Şekil 3.45'te gösterilmiştir.



Şekil 3.45: $t = 0$ için ileri beslemeli ağı eğitimi işlemi için hesaplama sonucu.

Adım 4: Durum ve çıktı değeri hesaplandıktan sonra bir sonraki zaman adımı $t = 1$ için ileri beslemeli ağı eğitimi yapılmıştır.

$$a_1 = \tanh(W_a^T \times x_1 + U_a \times out_0 + b_a) = \tanh([0,4 \quad 0,2] \times \begin{bmatrix} 1 \\ 0,3 \end{bmatrix} + [0,1] \times 1,615 + 0,3) = 0,726$$

$$i_1 = \sigma(W_i^T \times x_1 + U_i \times out_0 + b_i) = \sigma([0,1 \quad 0,3] \times \begin{bmatrix} 1 \\ 0,3 \end{bmatrix} + [0,4] \times 1,615 + 0,8) = 0,317$$

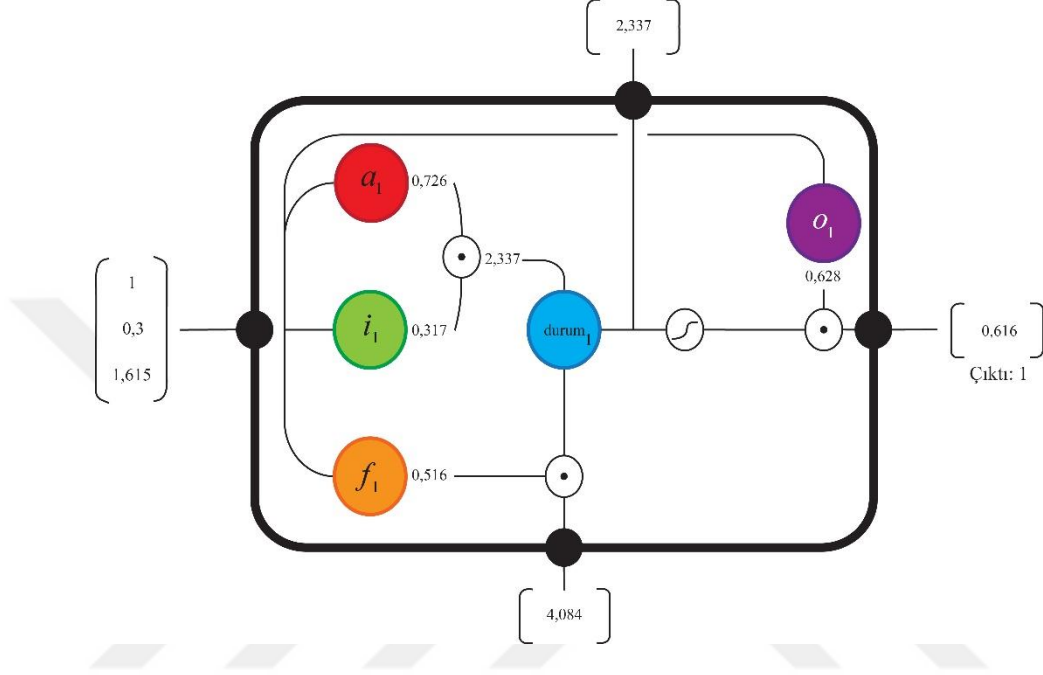
$$f_1 = \sigma(W_f^T \times x_1 + U_f \times out_0 + b_f) = \sigma([0,5 \quad 0,6] \times \begin{bmatrix} 1 \\ 0,3 \end{bmatrix} + [0,7] \times 1,615 + 0,1) = 0,516$$

$$o_1 = \sigma(W_o^T \times x_1 + U_o \times out_0 + b_o) = \sigma([0,8 \quad 0,4] \times \begin{bmatrix} 1 \\ 0,3 \end{bmatrix} + [0,9] \times 1,615 + 0,2) = 0,628$$

$$durum_1 = (a_1 \times i_1) + (f_1 \times durum_0) = (0,726 \times 0,317) + (0,516 \times 4,084) = 2,337$$

$$out_1 = \tanh(durum_1) \times o_1 = \tanh(2,337) \times 0,628 = 0,616$$

$t = 1$ için ileri beslemeli ağı eğitimi işleminin sonucu Şekil 3.46'da gösterilmiştir.



Şekil 3.46: $t = 1$ için ileri beslemeli ağı eğitimi işlemi için hesaplama sonucu.

Adım 5: $t = 1$ için geri beslemeli ağı eğitimi yapılır. Öncelikle beklenen çıktıdaki fark hesaplanmıştır.

$$\partial_x E(x, \hat{x}) = x - \hat{x} \text{ formülü ile } \Delta_1 = 0,616 - 1 = -0,384$$

Gelecekteki zaman adımları olmadığı için $\Delta out_1 = 0$ 'dır.

$$\delta out_1 = \Delta_1 + \Delta out_1 = -0,384 - 0 = -0,384$$

$$\begin{aligned} \delta durum_1 &= \delta out_1 \times o_1 \times (1 - \tanh^2(durum_1)) + \delta durum_2 \times f_2 \\ &= -0,384 \times 0,628 \times (1 - \tanh^2(2,337)) + 0 \times 0 = -0,009 \end{aligned}$$

$$\delta a_1 = \delta durum_1 \times i_1 \times (1 - a_1^2) = -0,009 \times 0,317 \times (1 - 0,527) = -0,001$$

$$\begin{aligned} \delta i_1 &= \delta durum_1 \times a_1 \times i_1 \times (1 - i_1) = -0,009 \times 0,726 \times 0,317 \times (1 - 0,317) \\ &= -0,001 \end{aligned}$$

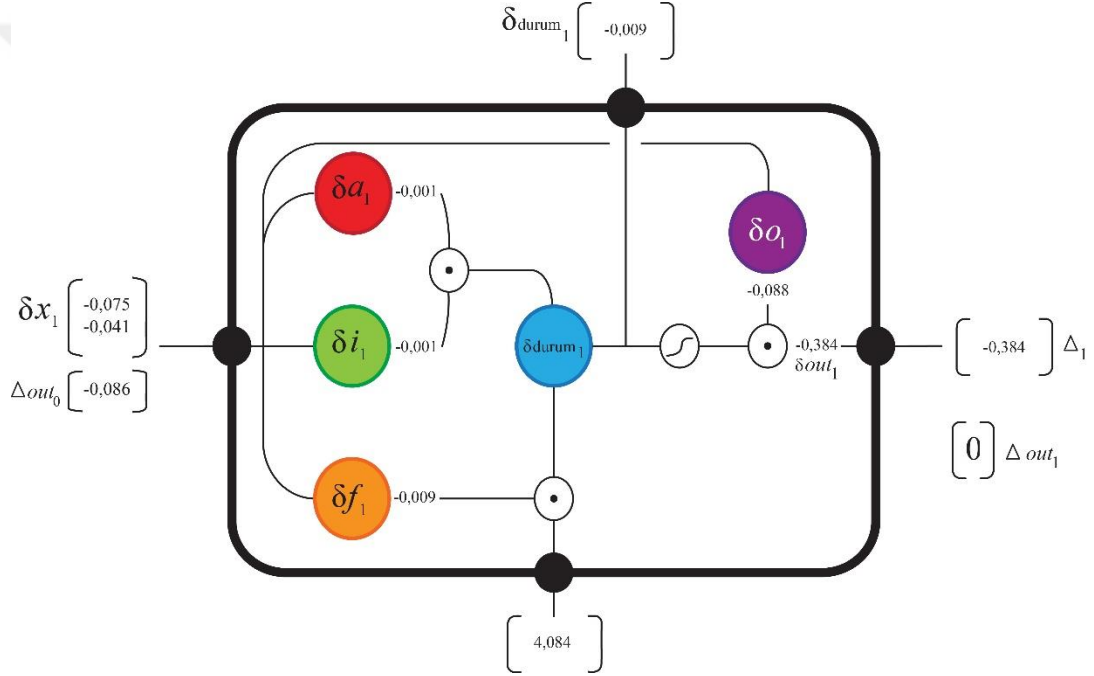
$$\begin{aligned} \delta f_1 &= \delta durum_1 \times durum_0 \times f_1 \times (1 - f_1) = -0,009 \times 4,084 \times 0,516 \times (1 - 0,516) \\ &= -0,009 \end{aligned}$$

$$\begin{aligned}\delta o_1 &= \delta out_1 \times \tanh(durum_1) \times o_1 \times (1 - o_1) \\ &= -0,384 \times \tanh(2,337) \times 0,628 \times (1 - 0,628) = -0,088\end{aligned}$$

$$\delta x_1 = W^T \times \delta gates_1 = \begin{bmatrix} 0,4 & 0,1 & 0,5 & 0,8 \\ 0,2 & 0,3 & 0,6 & 0,4 \end{bmatrix} \times \begin{bmatrix} -0,001 \\ -0,001 \\ -0,009 \\ -0,088 \end{bmatrix} = \begin{bmatrix} -0,075 \\ -0,041 \end{bmatrix}$$

$$\Delta out_0 = U^T \times \delta gates_1 = [0,1 \quad 0,4 \quad 0,7 \quad 0,9] \times \begin{bmatrix} -0,001 \\ -0,001 \\ -0,009 \\ -0,088 \end{bmatrix} = -0,086$$

$t = 1$ için geri beslemeli ağı eğitimi işleminin sonucu Şekil 3.47’de gösterilmiştir.



Şekil 3.47: $t = 1$ için geri beslemeli ağı eğitimi işlemi için hesaplama sonucu.

Adım 6: Δout değerini geri alıp hesaplama yapılmıştır. $t = 0$ geri beslemeli ağı eğitimi yapılmıştır.

$$\Delta_0 = \partial_x E = 1,615 - 0 = 1,615$$

$$\Delta out_0 = -0,086, t = 1 \text{ deniçin geri dönen değerdir,}$$

$$\delta out_0 = \Delta_0 + \Delta out_0 = 1,615 - 0,086 = 1,529$$

$$\begin{aligned}\delta durum_0 &= \delta out_0 \times o_0 \times (1 - \tanh^2(durum_0)) + \delta durum_1 \times f_1 \\ &= 1,529 \times 1,616 \times (1 - \tanh^2(4,084)) - 0,009 \times 0,516 = 0,0003\end{aligned}$$

$$\delta a_0 = \delta durum_0 \times i_0 \times (1 - a_0^2) = 0,0003 \times 4,272 \times (1 - 0,913) = 0,00008$$

$$\delta i_0 = \delta durum_0 \times a_0 \times i_0 \times (1 - i_0) = 0,0003 \times 0,956 \times 4,272 \times (1 - 4,272) = -0,003$$

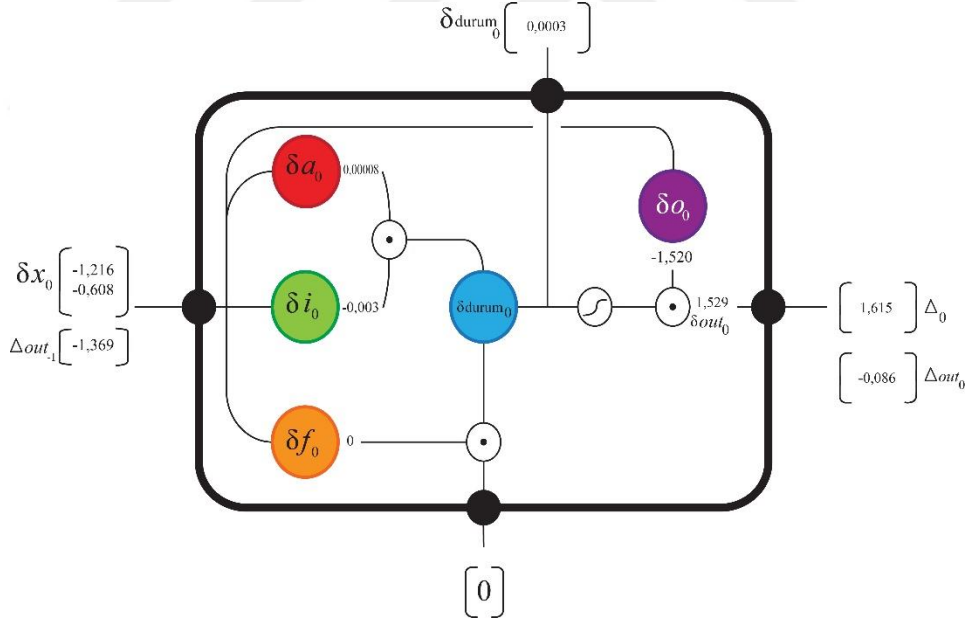
$$\delta f_0 = \delta durum_0 \times durum_{-1} \times f_0 \times (1 - f_0) = 0,0003 \times 0 \times 1,747 \times (1 - 1,747) = 0$$

$$\delta o_0 = \delta out_0 \times \tanh(durum_0) \times o_0 \times (1 - o_0) = 1,529 \times \tanh(4,084) \times 1,616 \times (1 - 1,616) = -1,520$$

$$\delta x_0 = W^T \times \delta gates_0 = \begin{bmatrix} 0,4 & 0,1 & 0,5 & 0,8 \\ 0,2 & 0,3 & 0,6 & 0,4 \end{bmatrix} \times \begin{bmatrix} 0,00008 \\ -0,003 \\ 0 \\ -1,520 \end{bmatrix} = \begin{bmatrix} -1,216 \\ -0,608 \end{bmatrix}$$

$$\Delta out_{-1} = U^T \times \delta gates_0 = [0,1 \quad 0,4 \quad 0,7 \quad 0,9] \times \begin{bmatrix} 0,00008 \\ -0,003 \\ 0 \\ -1,520 \end{bmatrix} = -1,369$$

$t = 0$ için geri beslemeli ağı eğitimi işleminin sonucu Şekil 3.48’de gösterilmiştir.



Şekil 3.48: $t = 0$ için geri beslemeli ağı eğitimi işlemi için hesaplama sonucu.

Adım 6: Geri yayılım eğitimi bitirdikten sonra, parametre değişkenlerinin güncellemesi gerekmektedir. Öğrenme oranının güncellemesi için Stokastik Gradyan İniş metodu kullanılmaktadır: $\mu = 0.1$ için ağırlıkların ne kadar değişeceği hesaplanmıştır.

$$\delta W = \sum_{t=0}^T \delta gates_t \otimes x_t = \begin{bmatrix} 0,00008 \\ -0,003 \\ 0 \\ -1,520 \end{bmatrix} \begin{bmatrix} 2 & 4 \end{bmatrix} + \begin{bmatrix} -0,001 \\ -0,001 \\ -0,009 \\ -0,088 \end{bmatrix} \begin{bmatrix} 1 & 0,3 \end{bmatrix} = \begin{bmatrix} -0,00084 & -0,00002 \\ -0,007 & -0,0123 \\ -0,009 & -0,002 \\ -3,128 & -6,106 \end{bmatrix}$$

$$\delta U = \sum_{t=0}^{T-1} \delta gates_{t+1} \otimes out_t = \begin{bmatrix} -0,001 \\ -0,001 \\ -0,009 \\ -0,088 \end{bmatrix} \begin{bmatrix} 1,615 \end{bmatrix} = \begin{bmatrix} -0,001 \\ -0,001 \\ -0,014 \\ -0,142 \end{bmatrix}$$

$$\delta b = \sum_{t=0}^T \delta gates_{t+1} = \begin{bmatrix} 0,00008 \\ -0,003 \\ 0 \\ -1,520 \end{bmatrix} + \begin{bmatrix} -0,001 \\ -0,001 \\ -0,009 \\ -0,088 \end{bmatrix} = \begin{bmatrix} -0,0009 \\ -0,004 \\ -0,009 \\ -1,608 \end{bmatrix}$$

Stokastik Gradyan İniş metodu ile güncellemek için $W^{yeni} = (W^{eski} - \mu) \times \delta W^{eski}$ formülü kullanılmaktadır [128].

$$W_a = \begin{bmatrix} 0,40084 \\ 0,20002 \end{bmatrix}, U_a = \begin{bmatrix} 0,101 \end{bmatrix}, b_a = \begin{bmatrix} 0,30009 \end{bmatrix}$$

$$W_i = \begin{bmatrix} 0,107 \\ 0,3123 \end{bmatrix}, U_i = \begin{bmatrix} 0,401 \end{bmatrix}, b_i = \begin{bmatrix} 0,8004 \end{bmatrix}$$

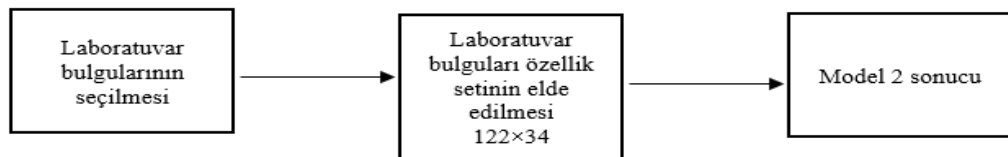
$$W_f = \begin{bmatrix} 0,509 \\ 0,602 \end{bmatrix}, U_f = \begin{bmatrix} 0,714 \end{bmatrix}, b_f = \begin{bmatrix} 0,1009 \end{bmatrix}$$

$$W_o = \begin{bmatrix} 0,828 \\ 0,406 \end{bmatrix}, U_o = \begin{bmatrix} 0,9142 \end{bmatrix}, b_o = \begin{bmatrix} 0,2608 \end{bmatrix}$$

Adım 7: Bir LSTM hücresinin çözümünün bir iterasyonu tamamlanmıştır.

3.7.2. Algoritma 1'e Laboratuvar Bulguları Özellik Setinin Uygulanması

Algoritma 1'e Tablo 3.8'de ifade edilen laboratuvar bulguları özellik seti uygulanmasıyla Model 2 oluşmuştur. Model 2 için geliştirilen modelleme Şekil 3.49'daki gibidir. Model 2 ile, ID3, doğrusal regresyon, Naive Bayes, DVM, k-NN ve FFBBP yöntemlerine uygulanan laboratuvar bulguları özellik setinden elde edilen en başarılı sınıflandırma doğruluk oranına sahip yöntemin sonucu hesaplanmıştır.



Şekil 3.49: Model 2.

Model 2: Laboratuvar bulguları özellik seti için modelleme

- Adım 1 **Girdi:** Test veri seti $D = [x_1, \dots, x_n]^T$
 Adım 2 **Çıktı:** Sınıflandırma sonucu Epilepsi (0) veya Sağlıklı (1), $z \rightarrow (0, +1)$
 Adım 3

Sınıflandırıcılara test veri setindeki verinin uygulanması

```

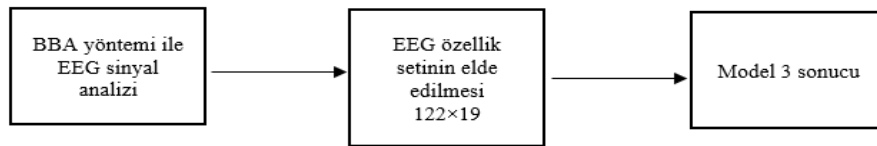
foreach  $x$  in  $D$  do
    tahmin( $x$ ) = ID3( $x$ );
    if (tahmin( $x$ ) == gerçek( $x$ ))
         $ID3\_doğruluk++$ ;
    tahmin( $x$ ) = DoğrusalRegresyon( $x$ );
    if (tahmin( $x$ ) == gerçek( $x$ ))
         $DoğrusalRegresyon\_doğruluk++$ ;
    tahmin( $x$ ) = NaiveBayes( $x$ );
    if (tahmin( $x$ ) == gerçek( $x$ ))
         $NaiveBayes\_doğruluk++$ ;
    tahmin( $x$ ) = DVM( $x$ );
    if (tahmin( $x$ ) == gerçek( $x$ ))
         $DVM\_doğruluk++$ ;
    if (tahmin( $x$ ) == gerçek( $x$ ))
        tahmin( $x$ ) = k-NN( $x$ );
         $k-NN\_doğruluk++$ ;
    tahmin( $x$ ) = FFBP( $x$ );
    if (tahmin( $x$ ) == gerçek( $x$ ))
         $FFBP\_doğruluk++$ ;
end foreach
 $C_1 = ID3\_doğruluk$ ;
 $C_2 = DoğrusalRegresyon\_doğruluk$ ;
 $C_3 = NaiveBayes\_doğruluk$ ;
 $C_4 = DVM\_doğruluk$ ;
 $C_5 = k-NN\_doğruluk$ ;
 $C_6 = FFBP\_doğruluk$ ;

```

- Adım 4 Maksimum Sınıflandırıcının Hesaplanması
 Model 2 = maksimum ($C_1, C_2, C_3, C_4, C_5, C_6$)

3.7.3. Algoritma 1'e EEG Özellik Setinin Uygulanması

Algoritma 1'e Tablo 3.8'de ifade edilen EEG özellik seti uygulanmasıyla Model 3 oluşmuştur. Model 3 için geliştirilen modelleme Şekil 3.50'deki gibidir.

**Şekil 3.50:** Model 3.

Model 3: EEG özellik seti için modelleme

- Adım 1 **Girdi:** Test veri seti $D = [x_1, \dots, x_n]^T$
 Adım 2 **Çıktı:** Sınıflandırma sonucu Epilepsi (0) veya Sağlıklı (1), $z \rightarrow (0, +1)$
 Adım 3

Sınıflandırıcılara test veri setindeki verinin uygulanması

```

foreach  $x$  in  $D$  do
    tahmin( $x$ ) = ID3( $x$ );
    if (tahmin( $x$ ) == gerçek( $x$ ))
        ID3_doğruluk ++;
    tahmin( $x$ ) = DoğrusalRegresyon( $x$ );
    if (tahmin( $x$ ) == gerçek( $x$ ))
        DoğrusalRegresyon_doğruluk ++;
    tahmin( $x$ ) = NaiveBayes( $x$ );
    if (tahmin( $x$ ) == gerçek( $x$ ))
        NaiveBayes_doğruluk ++;
    tahmin( $x$ ) = DVM( $x$ );
    if (tahmin( $x$ ) == gerçek( $x$ ))
        DVM_doğruluk ++;
    if (tahmin( $x$ ) == gerçek( $x$ ))
        tahmin( $x$ ) = k-NN( $x$ );
        k - NN_doğruluk ++;
    tahmin( $x$ ) = FFBP( $x$ );
    if (tahmin( $x$ ) == gerçek( $x$ ))
        FFBP_doğruluk ++;
end foreach
 $C_1$  = ID3_doğruluk;
 $C_2$  = DoğrusalRegresyon_doğruluk;
 $C_3$  = NaiveBayes_doğruluk;
 $C_4$  = DVM_doğruluk;
 $C_5$  = k - NN_doğruluk;
 $C_6$  = FFBP_doğruluk;

```

- Adım 4 Maksimum Sınıflandırıcının Hesaplanması
 Model 3 = maksimum ($C_1, C_2, C_3, C_4, C_5, C_6$)

Model 3 ile, ID3, doğrusal regresyon, Naive Bayes, DVM, k-NN ve FFBP yöntemlerine uygulanan EEG özellik setinden elde edilen en başarılı sınıflandırma doğruluk oranına sahip yöntemin sonucu hesaplanmıştır.

3.7.4. Algoritma 1'e MRG Özellik Setinin Uygulanması

Algoritma 1'e Tablo 3.8'de ifade edilen MRG özellik seti uygulanmasıyla Model 4 oluşmuştur. Model 4 için geliştirilen modelleme Şekil 3.51'deki gibidir.



Şekil 3.51: Model 4.

Model 4: MRG özellik seti için modelleme

- Adım 1 **Girdi:** Test veri seti $D = [x_1, \dots, x_n]^T$
 Adım 2 **Çıktı:** Sınıflandırma sonucu Epilepsi (0) veya Sağlıklı (1), $z \rightarrow (0, +1)$
 Adım 3

Sınıflandırıcılara test veri setindeki verinin uygulanması

```

foreach x in D do
  tahmin(x) = ID3(x);
  if (tahmin(x) == gerçek(x))
    ID3_doğruluk ++;
  tahmin(x) = DoğrusalRegresyon(x);
  if (tahmin(x) == gerçek(x))
    DoğrusalRegresyon_doğruluk ++;
  tahmin(x) = NaiveBayes(x);
  if (tahmin(x) == gerçek(x))
    NaiveBayes_doğruluk ++;
  tahmin(x) = DVM(x);
  if (tahmin(x) == gerçek(x))
    DVM_doğruluk ++;
  tahmin(x) = k-NN(x);
  if (tahmin(x) == gerçek(x))
    k-NN_doğruluk ++;
  tahmin(x) = FFBP(x);
  if (tahmin(x) == gerçek(x))
    FFBP_doğruluk ++;
  tahmin(x) = LSTM(x);
  if (tahmin(x) == gerçek(x))
    LSTM_doğruluk ++;
end foreach
C1 = ID3_doğruluk;
C2 = DoğrusalRegresyon_doğruluk;
C3 = NaiveBayes_doğruluk;
C4 = DVM_doğruluk;
C5 = k-NN_doğruluk;
C6 = FFBP_doğruluk;
C7 = LSTM_doğruluk;

```

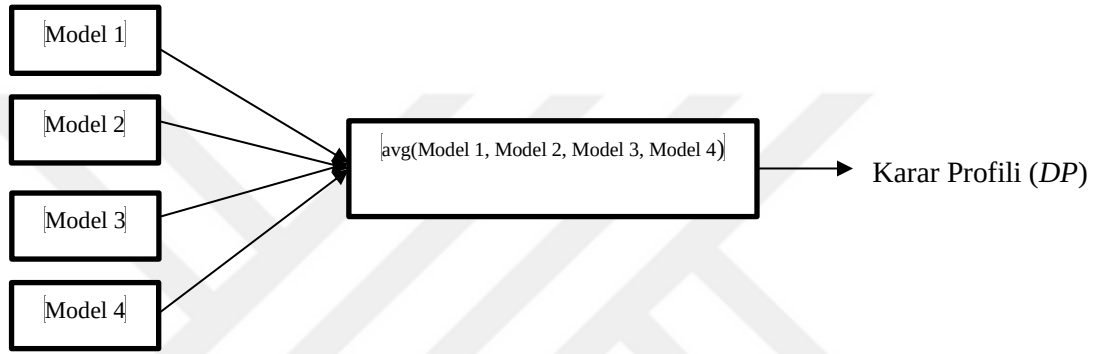
- Adım 4 Maksimum Sınıflandırıcının Hesaplanması

Model 4 = maksimum ($C_1, C_2, C_3, C_4, C_5, C_6, C_7$)

Model 4 ile ID3, doğrusal regresyon, Naive Bayes, DVM, k-NN, FFBP ve LSTM yöntemlerine uygulanan MRG özellik setinden elde edilen en başarılı sınıflandırma doğruluk oranına sahip yöntemin sonucu hesaplanmıştır.

3.7.5. Geç Füzyona Uygulanan Özellik Setlerinden Elde Edilen Sonuçların Çok Modlu Karar Tabanlı Veri Füzyonuna Uygulanması

Şekil 3.52’de, Algoritma 1’e uygulanan her bir özellik seti için sınıflandırmada maksimum doğruluk orana sahip sınıflandırıcı seçilmiştir.



Şekil 3.52: Geç füzyona uygulanan özellik setlerinden elde edilen sonuçların çok modlu karar tabanlı veri füzyonuna uygulanması.

Şekil 3.52’ye göre, bu tez çalışmasında önerilen çok modlu karar tabanlı veri füzyonu modelinin başarımlı sonucu, ortalama birleştirici ile hesaplanan değer olmuştur. |

4. BULGULAR

Gerçek hasta verileri ile oluşturulan epilepsi veri seti; klinik parametre özellik seti (122×48), laboratuvar bulguları özellik seti (122×34), EEG özellik seti (122×19), MRG özellik seti (122×502) için çok modlu karar tabanlı veri füzyonu modellemesi geliştirilerek epilepsi ve kontrol grubu sağlıklı bireyler sınıflandırılmıştır.

Çok modlu karar tabanlı veri füzyon modellemesinde epilepsi ve kontrol grubu sağlıklı bireylerin sınıflandırılması için ID3, doğrusal regresyon, Naive Bayes, lineer DVM, k-NN, FFBP yöntemlerine epilepsi veri setini oluşturan her bir özellik seti (klinik parametre özellik seti, laboratuvar bulguları özellik seti, EEG özellik seti ve MRG özellik seti) uygulanmıştır. Ayrıca, LSTM yöntemine MRG özellik seti uygulanarak modellenmiştir. Özellik setleri eğitim ve test veri seti olarak ayrılmıştır. Test veri seti, her özellik setinin 10–kat çaprazlama yöntemine göre bölünmesiyle elde edilmiştir. Özellik setlerinin optimum sınıflandırılmasında yöntemlere uygulanan ortak parametre değerleri Tablo 4.1’de verilmiştir. Bu değerler deneysel çalışmalar sırasında özellik setlerimiz için seçilen optimum değerlerdir.

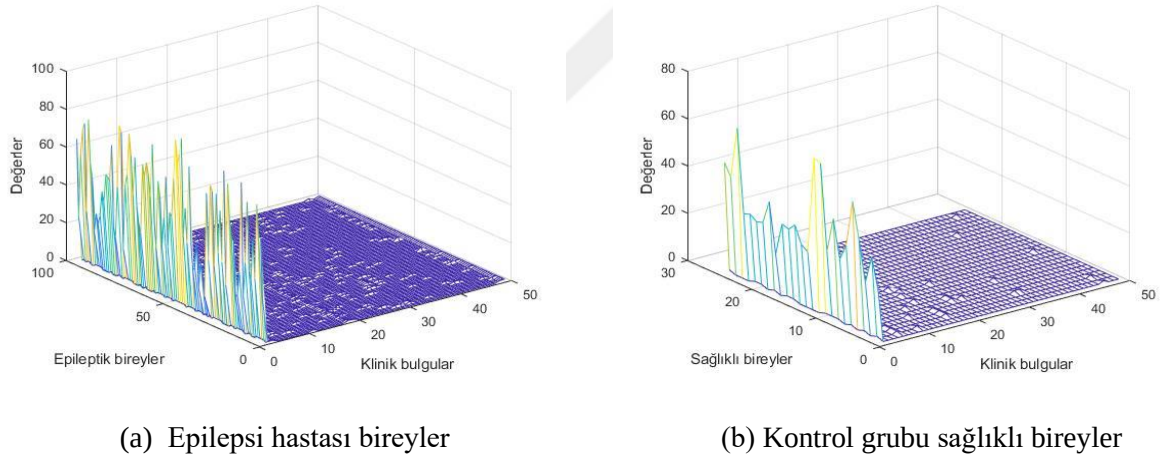
Tablo 4.1: Özellik setlerinin optimum sınıflandırılmasında yöntemlere uygulanan ortak parametre değerleri.

Sınıflandırıcılar	Model parametreleri ve değerleri	
ID3	maksimum bölünme sayısı	100
	bölünme kriteri	Entropi Kazancı
Naive Bayes	Tahminleyiciler için dağılım	Çekirdek
	fonksiyonu	
DVM	Çekirdek fonksiyonu	Lineer
	Sınıflandırma metodu	Bire–bir
k–NN	Uzaklık metriği	Öklid
	Komşu sayısı	10
FFBP	Aktivasyon fonksiyonu	tanh
	Çözümleyici	SCG
	Alfa	0,0001
	Öğrenme oranı	0,001
	Maksimum iterasyon	1000
	Ara katman nöron sayısı	10
LSTM	Eğitim fonksiyonu	adam
	Maksimum epok	100
	Başlangıç öğrenme katsayısı	0,001

Bu tez çalışmasında, ID3, doğrusal regresyon, Naive Bayes, lineer DVM, k-NN, FFBP yöntemlerine uygulanan klinik parametre özellik setinin modellenmesine Model 1; laboratuvar bulguları özellik setinin modellenmesine Model 2; EEG özellik setinin modellenmesine Model 3 ismi verilmiştir. ID3, doğrusal regresyon, Naive Bayes, lineer DVM, k-NN, FFBP ve LSTM yöntemlerine uygulanan MRG özellik setinin modellenmesine Model 4 ismi verilmiştir.

4.1. KLİNİK PARAMETRE ÖZELLİK SETİNE AİT SINIFLANDIRMA SONUÇLARI (MODEL 1)

Klinik parametre özellik seti (122×48) boyutlu bir matristir. Klinik muayene sırasında epilepsi hastası bireylerden elde edilen değerlerin özellik analizi Şekil 4.1(a)'da, kontrol grubu sağlıklı bireylerden elde edilen değerlerin özellik analizi Şekil 4.1(b)'deki grafiklerde gösterilmiştir.



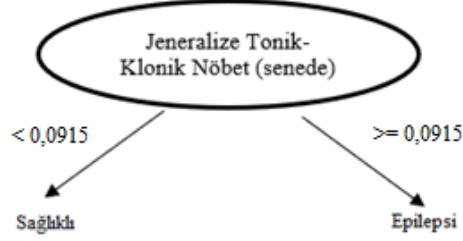
Şekil 4.1: Klinik muayene özellik seti.

ID3, doğrusal regresyon, Naive Bayes, lineer DVM, k-NN, FFBP yöntemlerine klinik muayene özellik seti uygulanarak epilepsi ve sağlıklı olarak sınıflandırma yapılmıştır. Bu yöntemlere ait sınıflandırma doğruluk oranlarından en başarılı olanı Model 1'in sonucu olmuştur.

4.1.1. Model için ID3 Yönteminin Sonuçları

ID3 yöntemine uygulanan klinik parametre özellik seti için elde edilen karar ağacı Şekil 4.2'de, karar kuralları ise Tablo 4.2'de gösterilmiştir. Eğitim işlemi sırasında ağacın dallarına ayırma işlemi öncesinde 200 değeri ile, ayırma işlemi sonrasında 50 değeri ile ve ağaç oluşturma

derinliği 10 değeri olarak düzenlenmiştir. Eğitim veri setinin sınıflandırılması sonucunda 3 yapraklı ve 2 düğümlü karar ağacı elde edilmiştir. Eğitim işlemi için sınıflandırma doğruluğu %100 olarak elde edilmiştir.



Şekil 4.2: ID3 yöntemine uygulanan klinik parametre özellik seti karar ağacı.

Tablo 4.2: ID3 yöntemine uygulanan klinik parametre özellik seti karar kuralları.

. If Jeneralize Tonik–Klonik Nöbet (senede) < 0,0915 then Sınıf = Sağlıklı (15 bireyde)
. If Jeneralize Tonik–Klonik Nöbet (senede) >= 0,0915 then Sınıf = Epilepsi (66 bireyde)

10–kat çaprazlama işlemi test başarımları için karışıklık matrisi Şekil 4.3’te gösterilmiştir. Şekil 4.3’e göre, test veri setinde ID3 yöntemi epilepsi hastası olan 10 bireye ait veriyi doğru sınıfa yani epilepsi sınıfı olarak tanımlamıştır. Test veri seti için ayrılmış olan 2 kontrol grubu sağlıklı bireye ait veri doğru sınıf olarak tanımlamıştır.

	Epilepsi	Sağlıklı
Epilepsi	10	0
Sağlıklı	0	2

Şekil 4.3: ID3 yöntemine uygulanan klinik parametre özellik seti için karışıklık matrisi.

Şekil 4.3’te, 10–kat çaprazlama işlemine göre epilepsi ve kontrol grubu sağlıklı bireylere ait klinik parametre özellik setinin sınıflandırılması için test başarımları %100 olarak elde edilmiştir.

4.1.2. Model 1 için Doğrusal Regresyon Yönteminin Sonuçları

Doğrusal regresyon yöntemine uygulanan klinik parametre özellik seti için 48 parametre ile lineer formül elde edilmiştir. y doğrusal regresyon denkleminde kestirimi yapılan katsayılar Tablo 4.3’te verilmiştir. Tablo 4.3’ e göre, gözlem sayısı 110, hata serbestlik derecesi 62, kök ortalama kare hata değeri 0,193, R–kare 0,872, düzeltilmiş R–kare 0,776, F istatistiği ile sabit

model 9,02, p-değeri 7e-15 olarak elde edilmiştir. Doğrusal regresyon yöntemi ile modellenen klinik parametreler özellik setine ait ortalama hata değeri, bağımsızlık derecesi, F değeri ve p-değeri sonuçları Tablo 4.3'te verilmiştir. Klinik parametre özellik setinin doğrusal regresyon analizinde, Tablo 4.3'e göre mavi renkle boyanan, x3: Jeneralize Tonik-Klonik Nobet(senede), x15: Unutkanlık, x22: Geç algılama, x23: Donup kalma, x24: Ateş, x44: Sol kolda kasılma parametrelerinin p-değerlerinin 0,05 değerinin altında olduğundan epilepsi ve kontrol grubu sağlıklı bireylerin sınıflandırılmasında anlamlı özellikler olduğu sonucu çıkarılmıştır.

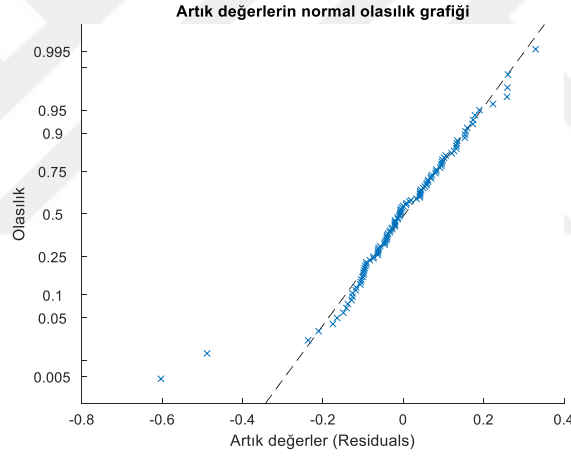
Tablo 4.3: Klinik parametre özellik seti doğrusal regresyon modelleme sonuçları.

Parametreler	Ortalama Hata	DF	F	p-Değeri
x1	0,137	1	3,681	0,059
x2	0,043	1	1,153	0,286
x3	0,531	1	14,209	0,0003
x4	0,002	1	0,075	0,784
x5	0,030	1	0,821	0,368
x6	0,064	1	1,735	0,192
x7	0,015	1	0,409	0,524
x8	0,083	1	2,218	0,141
x9	0,002	1	0,071	0,790
x10	0,0009	1	0,025	0,872
x11	0,029	1	0,798	0,375
x12	0,0001	1	0,003	0,951
x13	0,106	1	2,856	0,096
x14	0,047	1	1,282	0,261
x15	0,182	1	4,873	0,030
x16	0,031	1	0,845	0,361
x17	1,071	1	2,862	0,995
x18	1,369	1	36,578	9,232
x19	0,013	1	0,368	0,545
x20	0,006	1	0,173	0,678
x21	0,0009	1	0,024	0,876
x22	0,152	1	4,084	0,047
x23	0,513	1	13,711	0,0004
x24	0,330	1	8,838	0,004
x25	0,022	1	0,589	0,445
x26	0,0002	1	0,007	0,933
x27	0,067	1	1,794	0,185
x28	0,025	1	0,682	0,411
x29	0,029	1	0,775	0,382
x30	0,039	1	1,064	0,306
x31	0,101	1	2,698	0,105
x32	0,005	1	0,146	0,703
x33	0,012	1	0,329	0,568
x34	0,067	1	1,803	0,184
x35	0,011	1	0,299	0,585
x36	0,001	1	0,052	0,818
x37	0,066	1	1,788	0,186
x38	0,023	1	0,631	0,429

Tablo 4.3 (devam): Klinik parametre özellik seti doğrusal regresyon modelleme sonuçları.

Parametreler	Ortalama Hata	DF	F	p-Değeri
x39	0,008	1	0,228	0,634
x40	0,030	1	0,814	0,370
x41	0,005	1	0,148	0,701
x42	0	1	NaN	NaN
x43	0	1	NaN	NaN
x44	0,188	1	5,045	0,028
x45	0,024	1	0,659	0,419
x46	0,001	1	0,045	0,832
x47	0,0005	1	0,013	0,906
x48	0,050	1	1,336	0,252
Hata	2,321	62	0,037	

Doğrusal regresyon yöntemine uygulanan klinik parametre özellik setinin sınıflandırılmasında artık değerlerin normal olasılık dağılım grafiği Şekil 4.4'te verilmiştir. Eğitim işlemi için doğrusal regresyon yönteminin sınıflandırma doğruluğu %98,40 olarak elde edilmiştir.



Şekil 4.4: Klinik parametre özellik setinin sınıflandırılmasında artık değerlerin normal olasılık dağılım grafiği.

Şekil 4.5'e göre, 10–kat çaprazlama işlemine göre epilepsi ve sağlıklı kontrol grubu bireylere ait klinik parametre özellik setinin sınıflandırılmasında test başarıımı verilmiştir.

	Epilepsi	Sağlıklı
Epilepsi	10	0
Sağlıklı	0	2

Şekil 4.5: Doğrusal regresyon yöntemine uygulanan klinik parametre özellik seti için karışıklık matrisi.

Şekil 4.5'e göre, doğrusal regresyon yöntemine klinik parametre özellik setinin uygulanarak sınıflandırılmasında test veri setindeki epilepsi hastası ve kontrol grubu sağlıklı bireylere ait veri %100 doğrulukla sınıflandırılmıştır.

4.1.3. Model 1 için Naive Bayes Yönteminin Sonuçları

Naive Bayes yöntemine klinik parametre özellik seti uygulanmıştır. Çekirdek dağılımı kullanılarak Naive Bayes yöntemi ile eğitim setinin sınıflandırılmasında sınıflandırma doğruluk oranı %79,5 elde edilmiştir. Şekil 4.6'ya göre, 10-kat çaprazlama işlemine göre epilepsi ve kontrol grubu sağlıklı bireylere ait klinik parametre özellik setinin sınıflandırılmasındaki test başarımları %83,33 olarak elde edilmiştir.

	Epilepsi	Sağlıklı
Epilepsi	10	0
Sağlıklı	2	0

Şekil 4.6: Naive Bayes yöntemine uygulanan klinik parametre özellik seti için karışıklık matrisi.

Şekil 4.6'ya göre, Naive Bayes yöntemine klinik parametre özellik seti uygulanarak sınıflandırılmasında test veri setindeki epilepsi hastası bireylere ait veri doğru sınıflandırılırken, kontrol grubu sağlıklı bireylere ait veri yanlış sınıflandırılmıştır.

4.1.4. Model 1 için DVM Yönteminin Sonuçları

Lineer DVM yöntemine klinik parametre özellik seti uygulanarak sınıflandırılmıştır. Eğitim verisinin sınıflandırma doğruluğu %97,50 olarak elde edilmiştir. Test veri setinin

sınıflandırmadaki test başarımı Şekil 4.7’de karışıklık matrisinde gösterilmiştir. Test veri setinin sınıflandırma doğruluğu %83,33 olarak elde edilmiştir.

	Epilepsi	Sağlıklı
Epilepsi	10	0
Sağlıklı	2	0

Şekil 4.7: DVM yöntemine uygulanan klinik parametre özellik seti için karışıklık matrisi.

Şekil 4.7’ye göre, lineer DVM yöntemine klinik parametre özellik setinin uygulanarak sınıflandırılmasında test veri setindeki epilepsi hastası bireylere ait veri doğru sınıf olarak yani epilepsi sınıfı olarak tanımlanmıştır. Test veri seti için ayrılmış 2 kontrol grubu sağlıklı bireylere ait veri yanlış sınıf olarak yani epilepsi olarak tanımlanmıştır.

4.1.5. Model 1 için k–NN Yönteminin Sonuçları

k–NN yöntemine klinik parametre özellik seti uygulanmıştır. Öklid uzaklığına göre hesaplanan en yakın 10–komşuluk için eğitim veri setinin sınıflandırma doğruluğu %96,70 olarak elde edilmiştir. Test veri setinin test doğruluğu Şekil 4.8’de karışıklık matrisi ile verilmiştir ve test veri setinin sınıflandırma doğruluğu %66,66 olarak elde edilmiştir.

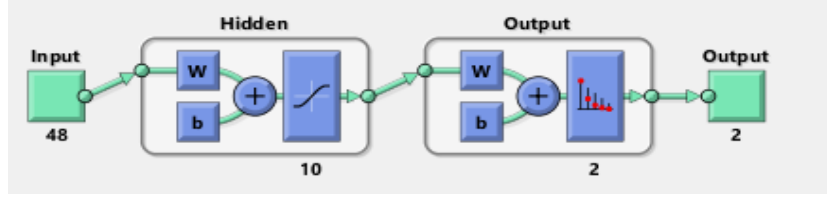
	Epilepsi	Sağlıklı
Epilepsi	8	1
Sağlıklı	3	0

Şekil 4.8: k–NN yöntemine uygulanan klinik parametre özellik seti için karışıklık matrisi.

Şekil 4.8’e göre, k–NN yöntemine klinik parametre özellik setinin uygulanarak sınıflandırılmasında test verisindeki epilepsi hastası bireylere ait 10 bireye ait verinin 8’i doğru sınıflandırılırken, kontrol grubu sağlıklı bireylere ait veri epilepsi sınıfı olarak hesaplanmıştır.

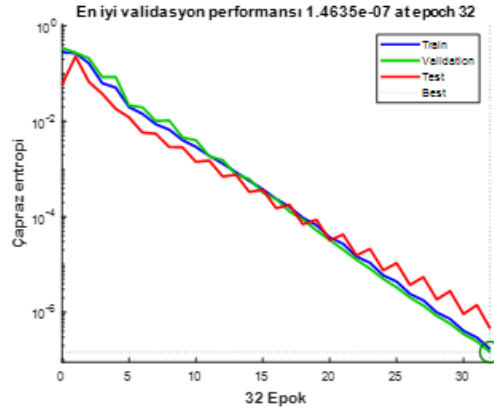
4.1.6. Model 1 için FFBP Yönteminin Sonuçları

Klinik parametre özellik setinin sınıflandırılması için $\mu = 0,001$ ile ara katman nöron sayısı 10 olarak düzenlenmiş ÇKA genel yapısı Şekil 4.9’da verilmiştir.



Şekil 4.9: Klinik parametre özellik seti için ÇKA yapısı.

FFBP yöntemine uygulanan klinik parametreler özellik seti için eğitim işleminden %98,40 sınıflandırma doğruluğu elde edilmiştir. Şekil 4.10’a göre en iyi doğrulama (validasyon) performansı 32 iterasyon ile sınıflandırma doğruluğu, %99,99 olarak elde edilmiştir.



Şekil 4.10: FFBP yöntemine uygulanan klinik parametreler özellik seti için en iyi doğrulama performansı grafiği.

FFBP yöntemine uygulanan klinik parametreler özellik setinin sınıflandırılmasında test veri setinin doğruluğu Şekil 4.11’de karışıklık matrisi ile verilmiştir ve test setinin sınıflandırma doğruluğu %100 olarak elde edilmiştir.

	Epilepsi	Sağlıklı
Epilepsi	11	0
Sağlıklı	0	1

Şekil 4.11: FFBP yöntemine uygulanan klinik parametre özellik seti için karışıklık matrisi.

Şekil 4.11'e göre, FFBP yöntemine klinik parametre özellik setinin uygulanarak sınıflandırılmasında test veri setindeki epilepsi hastası ve kontrol grubu sağlıklı bireylere ait veri doğru sınıf olarak hesaplanmıştır. FFBP yöntemi ile modellemede giriş katmanından (x) ara katmana doğru ağırlık değerleri (w) Tablo 4.4'te verilmiştir.

Tablo 4.4: Klinik parametre özellik seti için giriş katmanından ara katmanına doğru ağırlık değerleri.

Klinik Parametreler	Nöron "1"	Nöron "2"	Nöron "3"	Nöron "4"	Nöron "5"	Nöron "6"	Nöron "7"	Nöron "8"	Nöron "9"	Nöron "10"
Yaş	0,149296	0,149663	0,149193	0,150141	0,149829	0,149789	0,149054	0,148991	0,148948	0,149165
Cinsiyet	0,084678	0,084783	0,084685	0,084317	0,085154	0,084165	0,08413	0,084801	0,08475	0,084289
Jeneralize Tonik-Klonik Nöbet (senede)	0,303749	0,304262	0,304013	0,30487	0,304991	0,304132	0,303358	0,303431	0,3035	0,304154
Şuur Kaybı	0,015447	0,015168	0,015346	0,01489	0,015379	0,015157	0,015223	0,014679	0,015117	0,015099
Terleme	0,13149	0,13156	0,131783	0,13211	0,132476	0,132037	0,131526	0,131938	0,131107	0,131566
Uyku Bozukluğu	0,106517	0,107244	0,106647	0,106926	0,107232	0,106967	0,106453	0,106891	0,106953	0,107122
Baş Ağrısı	0,015641	0,015414	0,015848	0,0154	0,015849	0,01551	0,01522	0,015303	0,015586	0,015475
Konuşmada Zorlanma (Bozukluğu)	0,136117	0,135662	0,135761	0,1362566	0,1356762	0,1355478	0,135414	0,1361536	0,1358926	0,1360037
Görme Bozukluğu (Bulank görme)	0,0181195	0,0179946	0,0180205	0,017793	0,017704	0,0177694	0,0176693	0,0177963	0,0173502	0,0174622
El-Kollarda Uyuşma	0,0948473	0,0943695	0,0948252	0,09495	0,0950843	0,0948418	0,0943787	0,0946859	0,0947143	0,0946427
Baş Dönmesi	0,1358878	0,135173	0,1358071	0,1360316	0,1359989	0,1358334	0,1353947	0,1353688	0,1349994	0,1355413
Göz Karaması	0,0068484	0,0070968	0,0072368	0,0072976	0,0067203	0,0069027	0,007284	0,006982	0,0070603	0,0069679
Bilinç Kaybı	0,1805368	0,1797275	0,1799735	0,1805845	0,1802988	0,1802015	0,1799503	0,1800656	0,1800809	0,1806127
Halsizlik	0,0480901	0,0482366	0,0482838	0,0477854	0,0480634	0,0482056	0,0477829	0,0476675	0,0478599	0,0482661
Unutkanlık	0,2735729	0,2736763	0,2733296	0,2743416	0,2744618	0,2732547	0,2727827	0,2729718	0,2724686	0,2737009
Dil Dolandırma	0,1077061	0,108313	0,1076814	0,1085901	0,1079927	0,1078195	0,1080789	0,1076084	0,1074175	0,1076448
Bayılma Hissi	0,1430405	0,1429109	0,1432433	0,1430982	0,143075	0,1433121	0,1430197	0,1433435	0,1429898	0,1429659
Bayılma	0,4330447	0,4325528	0,4329076	0,4332365	0,4341379	0,4330653	0,4325053	0,4327699	0,4318072	0,4330869
Denge Bozukluğu	0,1704381	0,1700541	0,1701926	0,1697953	0,1702888	0,1700972	0,1699839	0,1696734	0,1699378	0,1702786
Nefes Daralması	0,0734472	0,0736784	0,0735464	0,0740749	0,0744679	0,0737922	0,0733024	0,0738189	0,0733181	0,0739345
Uyku Hali	0,0708429	0,0711263	0,0706396	0,0706875	0,071111	0,0711226	0,0706607	0,0712789	0,0709755	0,0705567
Geç Algılama	0,1020718	0,1021861	0,1023361	0,1020578	0,1020357	0,1017685	0,1021201	0,1018606	0,1016895	0,1016796
Dalma (Donup Kalma)	0,3113816	0,3113109	0,3121849	0,3122188	0,3129407	0,3116057	0,3110651	0,3119502	0,3109738	0,3122799
Ateş	0,0973166	0,0973887	0,0973261	0,0976671	0,0975759	0,0973268	0,0971777	0,0969788	0,0965056	0,0971813
Bulantı	0,2101145	0,210183	0,2098569	0,2106836	0,2107135	0,2099635	0,209501	0,2101406	0,2096838	0,2105739
Sağ El Uyuşması	0,0676431	0,0676797	0,0675495	0,0681751	0,0682868	0,0677636	0,0682193	0,067338	0,0678456	0,0678997
Ağızda Uyuşma	0,0890435	0,0897318	0,0896332	0,0898094	0,0893139	0,0899134	0,0892001	0,0890551	0,0890767	0,0899858
Kafa Travması	0,1348692	0,135033	0,135409	0,1356459	0,1355515	0,1349764	0,1349312	0,1347505	0,134892	0,1350607
Anlamsız Hareketler	0,1278954	0,1282783	0,1283644	0,1277476	0,1285639	0,1279199	0,1274334	0,1279707	0,1280625	0,1281685
Kusma	0,1372612	0,1368001	0,1371372	0,1373411	0,1378663	0,1368621	0,1366454	0,1372566	0,1370055	0,1370288
Kafa Karışıklığı	0,2284057	0,2287707	0,2283435	0,2293473	0,2292361	0,2289105	0,2281363	0,2285575	0,2284837	0,2286452
El Ayakta Atma	0,0275373	0,0271587	0,027503	0,0275475	0,0272485	0,0276237	0,0275441	0,0272026	0,0271965	0,0275038
Tonik Kasılma	0,0737404	0,073488	0,0734055	0,074156	0,0740852	0,0737926	0,0738131	0,0735062	0,0730932	0,0739495
Uykuda Kasılma	0,1086959	0,1090007	0,1088299	0,10966	0,109804	0,1092741	0,1086622	0,1089978	0,1086454	0,1094996

Tablo 4.4 (devam): Klinik parametre özellik seti için giriş katmanından ara katmanına doğru ağırlık değerleri.

Cenede Kilitlenme	0,1783716	0,178236	0,1787601	0,1787092	0,1787628	0,1787948	0,1777649	0,1778761	0,1777018	0,1788008
Sağ Yüzde Uyuşma	0,0668461	0,0673685	0,0670502	0,0671435	0,0672887	0,0668721	0,067127	0,0673545	0,0664654	0,0668803
Sol Kulak Ağrısı	0,0702956	0,0702446	0,070389	0,0705689	0,0698278	0,0701279	0,0698007	0,0703658	0,0696264	0,0697893
Ellerde Uyuşma	0,0604424	0,0598452	0,0603605	0,0603088	0,0607307	0,0599184	0,0605063	0,0604186	0,0598754	0,0598416
Sağ Tarafda Güçsüzlük	0,0865478	0,0873776	0,0874044	0,0867732	0,0868958	0,086969	0,0865623	0,0866852	0,0864612	0,0871647
Sol Bacakta Kasılma–Kramp	0,0258777	0,0265365	0,02625	0,0259859	0,0267538	0,0261451	0,0260175	0,0267197	0,026581	0,0260982
Sol Kulakta İşitme Kaybı	0,058407	0,05856	0,0582297	0,0584431	0,0587123	0,0585832	0,0578724	0,0586697	0,0585667	0,0586144
Sol Göz Görme Kaybı	0,0587639	0,0587028	0,0579811	0,0585485	0,0585018	0,0582829	0,0582247	0,0586417	0,0580095	0,0581159
Sol Kolda Kasılma	0,1409141	0,1406727	0,141242	0,1406359	0,141632	0,1405015	0,1404355	0,1405319	0,1402583	0,1411686
Alkol Kullanımı	0,0669124	0,0666549	0,0671477	0,0668387	0,0674075	0,0666808	0,0668142	0,0667607	0,0672197	0,0666211
Sigara Kullanımı	0,15692257	0,15686997	0,15684851	0,15725792	0,15747421	-0,1573750	-0,1570961	-0,1567557	-0,15705295	-0,15698780
Ayda Geçirdiği Atak Sayısı	0,1288107	0,12885767	0,12867347	0,1288746	0,12909356	0,12862965	0,12893256	0,12819701	0,12888956	0,12848481
Epilepsi Soy geçmiş	0,03342364	0,03359101	0,0339263	0,03328088	0,03404716	0,03404766	0,03332848	0,03338724	0,03402231	0,03390958
Eğilim (bias)	0,01600063	0,01568635	0,01642049	0,01701281	0,01699909	0,01621385	0,01542232	0,01547531	0,01514281	0,01646165

Ara katmandan (w) çıkış katmanına (y) ağırlık değerleri Tablo 4.5’te verilmiştir.

Tablo 4.5: Klinik parametre özellik seti için ara katmandan çıkış katmanına doğru ağırlık değerleri.

–	Epilepsi	Sağlıklı
Nöron "1"	0,56294618	-0,56309177
Nöron "2"	0,56322596	-0,56295898
Nöron "3"	0,56323704	-0,56347657
Nöron "4"	0,56540982	-0,56518875
Nöron "5"	0,56702935	-0,56642457
Nöron "6"	0,56356299	-0,56335315
Nöron "7"	0,56092308	-0,56139597
Nöron "8"	0,56231582	-0,56261983
Nöron "9"	0,56106270	-0,56053846
Nöron "10"	0,56404494	-0,56369691
Eğilim (bias)	-1,20431715	1,20393141

Klinik parametre özellik setinde bulunan özelliklerin katkısı Tablo 4.6’da gösterilmiştir. Tablo 4.6’ya göre, klinik parametre özellik setindeki: yaş, şuur kaybı, uyku bozukluğu, baş ağrısı, görme bozukluğu (bulanık görme), el kollarda uyuşma, göz kararması, halsizlik, bayılma hissi, bayılma, nefes daralması, uyku hali, ateş, sağ el uyuşması, el ayakta atma, tüm vücutta titreme, tonik kasılma, sağ yüzde uyuşma, sol kulak ağrısı, ellerde uyuşma, sağ tarafta güçsüzlük, sol bacakta kasılma–kramp, sol kulakta işitme kaybı, sol göz görme kaybı, sol kolda kasılma, alkol

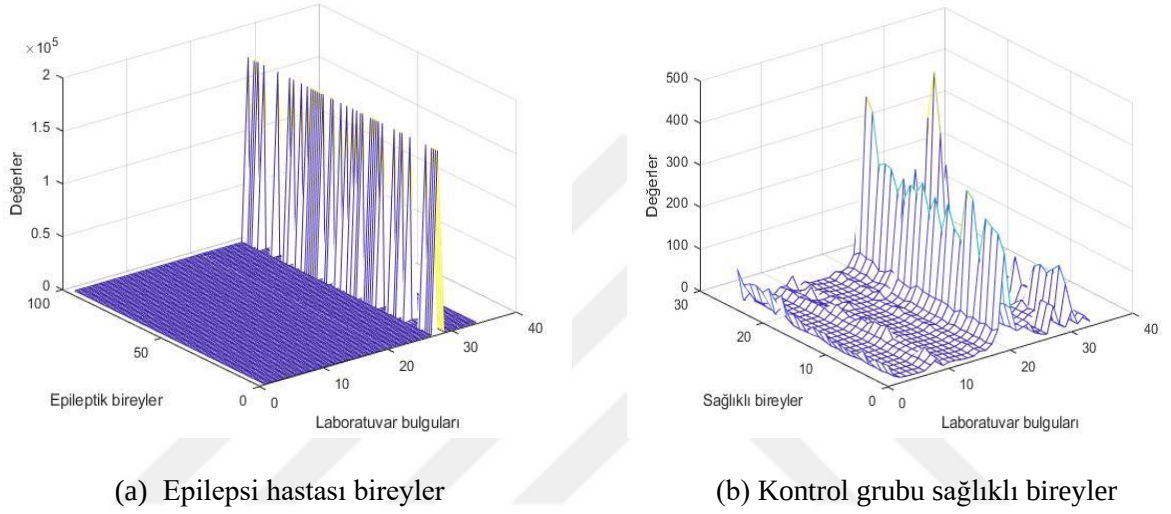
kullanımı, sigara kullanımı, ayda geçirdiği atak sayısı, epilepsi soy geçmişi özelliklerinin epilepsi hastalığının tespitinde ayırt edici özellikler olduğu görülmektedir.

Tablo 4.6: Klinik parametre özellik setinde bulunan özelliklerin katkısı.

Klinik parametreler	Hata oranı
Yaş	0,0000
Cinsiyet	0,0082
Jeneralize Tonik-KlonikNobet (senede)	0,0082
Şuur Kaybı	0,0000
Terleme	0,0082
Uyku Bozukluğu	0,0000
Baş Ağrısı	0,0000
Konuşmada Zorlanma (Bozukluğu)	0,0082
Görme Bozukluğu (Bulanık görme)	0,0000
El-Kollarda Uyuşma	0,0000
Baş Dönmesi	0,0082
Göz Kararması	0,0000
Bilinç Kaybı	0,0082
Halsizlik	0,0000
Unutkanlık	0,0082
Dil Dolandırma	0,0082
Bayılma Hissi	0,0000
Bayılma	0,0000
Denge Bozukluğu	0,0082
Nefes Daralması	0,0000
Uyku Hali	0,0000
Geç Algılama	0,0082
Dalma (Donup Kalma)	0,0082
Ateş	0,0000
Bulantı	0,0082
Sağ Elde Uyuşma	0,0000
Ağızda Uyuşma	0,0082
Kafa Travması	0,0082
Anlamsız Hareketler	0,0082
Kusma	0,0082
Kafa Karışıklığı	0,0082
El Ayakta Atma	0,0000
Tüm vücutta titreme	0,0000
Tonik Kasılma	0,0000
Uykuda Kasılma	0,0082
Çenede Kilitlenme	0,0082
Sağ Yüzde Uyuşma	0,0000
Sol Kulak Ağrısı	0,0000
Ellerde Uyuşma	0,0000
Sağ Tarafda Güçsüzlük	0,0000
Sol Bacakta Kasılma-Kramp	0,0000
Sol Kulakta İşitme Kaybı	0,0000
Sol Göz Görme Kaybı	0,0000
Sol Kolda Kasılma	0,0082
Alkol Kullanımı	0,0000
Sigara Kullanımı	0,0000
Ayda Geçirdiği Atak Sayısı	0,0082
Epilepsi Soy geçmişi	0,0000

4.2. LABORATUVAR BULGULARI ÖZELLİK SETİNE AİT SINIFLANDIRMA SONUÇLARI (MODEL 2)

Laboratuvar bulguları özellik seti (122×34) boyutlu bir matristir. Laboratuvar bulguları sırasında epilepsi hastası bireylerden elde edilen kan değerleri Şekil 4.12(a)'da, kontrol grubu sağlıklı bireylerden elde edilen kan değerlerine ait özellik analizi Şekil 4.12(b)'deki grafiklerde gösterilmiştir.

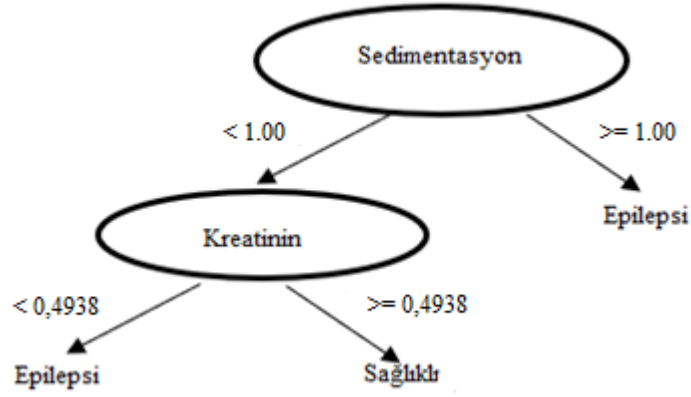


Şekil 4.12: Laboratuvar bulguları özellik seti.

Model 2, ID3, doğrusal regresyon, Naive Bayes, lineer DVM, k-NN, FFBP yöntemlerine laboratuvar bulguları özellik seti uygulanarak epilepsi ve sağlıklı olarak sınıflandırma yapılmıştır. Bu yöntemlere ait sınıflandırma doğruluk oranlarından en başarılı olanı Model 2'nin sonucu olmuştur.

4.2.1. Model 2 için ID3 Yönteminin Sonuçları

ID3 yöntemine uygulanan laboratuvar bulguları özellik seti için elde edilen karar ağacı Şekil 4.13'te, karar kuralları ise Tablo 4.7'de gösterilmiştir. Eğitim işlemi sırasında ağacın dallarına ayırma işlemi öncesinde 200 değeri ile, ayırma işlemi sonrasında 50 değeri ile ve ağaç oluşturma derinliği 10 değeri olarak düzenlenmiştir. Eğitim setinin sınıflandırılması sonucunda 5 yapraklı 3 düğümlü karar ağacı oluşmuştur. Eğitim işlemi sınıflandırma doğruluğu %98,40 olarak elde edilmiştir.



Şekil 4.13: ID3 yöntemine uygulanan laboratuvar bulguları özellik seti karar ağacı.

Tablo 4.7: ID3 yöntemine uygulanan laboratuvar bulguları özellik seti karar kuralları.

. If Sedimentasyon < 1,0000	
If Kreatinin < 0,4938 then Sınıf = Epilepsi (1 bireyde)	
else if Kreatinin >= 0,4938 then Sınıf = Sağlıklı (25 bireyde)	
else if Sedimentasyon >= 1,0000 then Sınıf = Epilepsi (96 bireyde)	

10–kat çaprazlama işlemi test başarımı için karışıklık matrisi Şekil 4.14’te gösterilmiştir. Şekil 4.14’e göre, test verisinde ID3 yöntemi epilepsi hastası ve kontrol grubu bireylere ait veriyi doğru sınıflandırmıştır.

	Epilepsi	Sağlıklı
Epilepsi	10	0
Sağlıklı	0	2

Şekil 4.14: ID3 yöntemine uygulanan laboratuvar bulguları özellik seti için karışıklık matrisi.

Şekil 4.15’e göre 10–kat çaprazlama işlemine göre epilepsi ve kontrol grubu sağlıklı bireylere ait laboratuvar bulguları özellik setinin sınıflandırılmasında test başarımı %100 olarak elde edilmiştir.

4.2.2. Model 2 için Doğrusal Regresyon Yönteminin Sonuçları

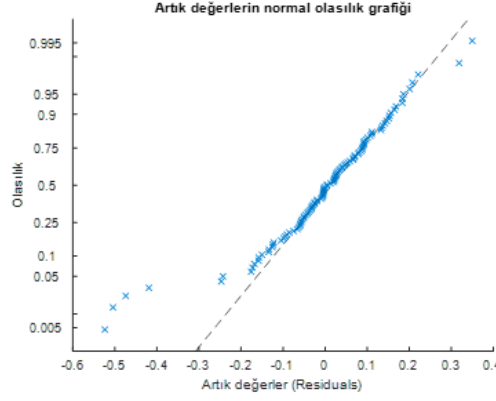
Doğrusal regresyon yöntemine uygulanan laboratuvar bulguları özellik seti için 34 parametre ile doğrusal formül elde edilmiştir. y doğrusal regresyon denkleminde kestirimi yapılan katsayılar Tablo 4.8’de verilmiştir. Tablo 4.8’e göre, gözlem sayısı 110, hata serbestlik derecesi 75, kök ortalama kare hata değeri 0,173, R–kare 0,877, düzeltilmiş R–kare 0,822, F istatistiği

ile sabit model 15,8, p-değeri 1,62e-22 olarak elde edilmiştir. Laboratuvar bulguları özellik setinin doğrusal regresyon analizinde, x3: PDW, x28: Vitamin B-12, x31: Kreatinin, x32: CKD-EP (eGFR) parametrelerinin p-değerleri 0,05 değerinin altında olduğundan epilepsi ve kontrol grubu sağlıklı bireylerin sınıflandırılmasında anlamlı özellikler olarak çıkarılmıştır.

Tablo 4.8: Laboratuvar bulguları özellik seti doğrusal regresyon modelleme sonuçları.

Parametreler	Ortalama Hata	DF	F	p-Değeri
x1	0,008	1	0,287	0,593
x2	0,0009	1	0,031	0,859
x3	0,007	1	0,263	0,609
x4	0,0028	1	0,283	0,595
x5	0,109	1	3,676	0,059
x6	0,043	1	1,463	0,230
x7	0,612	1	20,554	2,155
x8	0,409	1	13,759	0,000
x9	0,012	1	0,419	0,519
x10	0,589	1	19,802	2,936
x11	0,005	1	0,175	0,676
x12	0,004	1	0,141	0,708
x13	0,017	1	0,576	0,450
x14	0,0001	1	0,005	0,938
x15	0,004	1	0,162	0,687
x16	0,007	1	0,245	0,621
x17	0,005	1	0,198	0,657
x18	0,039	1	1,313	0,255
x19	0,049	1	1,660	0,201
x20	0,114	1	3,840	0,053
x21	0,077	1	2,609	0,110
x22	0,036	1	1,226	0,271
x23	0,100	1	3,386	0,069
x24	0,007	1	0,243	0,623
x25	0,004	1	0,163	0,686
x26	0,036	1	1,222	0,272
x27	0,002	1	0,077	0,780
x28	0,243	1	8,168	0,005
x29	0,000	1	0,003	0,953
x30	0,010	1	0,359	0,550
x31	0,375	1	12,605	0,000
x32	0,218	1	7,344	0,008
x33	0,013	1	0,440	0,509
x34	0,109	1	3,674	0,059
Hata	2,233	75	0,029	

Doğrusal regresyon yöntemine uygulanan laboratuvar bulgular özellik setinin sınıflandırılmasında artık değerlerin normal olasılık dağılım grafiği Şekil 4.15'te verilmiştir. Eğitim işlemi için doğrusal regresyon yönteminin sınıflandırma doğruluğu %95,10 olarak elde edilmiştir.



Şekil 4.15: Laboratuvar bulguları özellik setinin sınıflandırılmasında artık değerlerin normal olasılık dağılım grafiği.

Şekil 4.16'ya göre 10–kat çaprazlama işlemine göre epilepsi ve kontrol grubu sağlıklı bireylere ait laboratuvar bulguları özellik setinin sınıflandırılmasında test başarımı %100 olarak elde edilmiştir.

	Epilepsi	Sağlıklı
Epilepsi	10	0
Sağlıklı	0	2

Şekil 4.16: Doğrusal regresyon yöntemine uygulanan laboratuvar bulguları özellik seti için karışıklık matrisi.

Şekil 4.16'ya göre, doğrusal regresyon yöntemine laboratuvar bulguları özellik setinin uygulanarak sınıflandırılmasında test verisindeki epilepsi hastası ve kontrol grubu sağlıklı bireylere ait veri doğru sınıflandırılmıştır.

4.2.3. Model 2 için Naive Bayes Yönteminin Sonuçları

Naive Bayes yöntemine laboratuvar bulguları özellik seti uygulanmıştır. Çekirdek dağılım fonksiyonu kullanılarak Naive Bayes yöntemi ile eğitim veri setinin sınıflandırma doğruluk oranı %95,90 elde edilmiştir. Şekil 4.17'ye göre, 10–kat çaprazlama işlemine göre epilepsi ve kontrol grubu sağlıklı bireylere ait laboratuvar bulguları özellik setinin sınıflandırılmasında test başarımları oranı %83,33 olarak elde edilmiştir.

	Epilepsi	Sağlıklı
Epilepsi	10	0
Sağlıklı	2	0

Şekil 4.17: Naive Bayes yöntemine uygulanan laboratuvar bulguları özellik seti için karışıklık matrisi.

Şekil 4.17'ye göre, Naive Bayes yöntemine laboratuvar bulguları özellik setinin uygulanarak sınıflandırılmasında test veri setindeki epilepsi hastası bireylere ait veri doğru sınıflandırılırken, kontrol grubu sağlıklı bireylere ait veri yanlış sınıflandırılmıştır.

4.2.4. Model 2 için DVM Yönteminin Sonuçları

Lineer DVM yöntemine laboratuvar bulguları özellik seti uygulanarak sınıflandırılmıştır. Eğitim veri setinin sınıflandırma doğruluğu %93,40 olarak elde edilmiştir. Test veri setinin sınıflandırmadaki test başarımı Şekil 4.18'de karışıklık matrisinde gösterilmiştir ve test veri setinin sınıflandırma doğruluğu %100 olarak elde edilmiştir.

	Epilepsi	Sağlıklı
Epilepsi	10	0
Sağlıklı	0	2

Şekil 4.18: DVM yöntemine uygulanan laboratuvar bulguları özellik seti için karışıklık matrisi.

Şekil 4.18'e göre, lineer DVM yöntemine laboratuvar bulguları özellik setinin uygulanarak sınıflandırılmasında test veri setindeki epilepsi hastası ve kontrol grubu sağlıklı bireylere ait veri doğru sınıflandırılmıştır.

4.2.5. Model 2 için k-NN Yönteminin Sonuçları

k-NN yöntemine laboratuvar bulguları özellik seti uygulanmıştır. Öklid uzaklığına göre hesaplanan en yakın 10-komşuluk için eğitim veri setinin sınıflandırma doğruluğu %93,40 olarak elde edilmiştir. Test veri setinin test doğruluğu Şekil 4.19'da karışıklık matrisi ile verilmiştir ve test verisinin sınıflandırma doğruluğu %100 olarak elde edilmiştir.

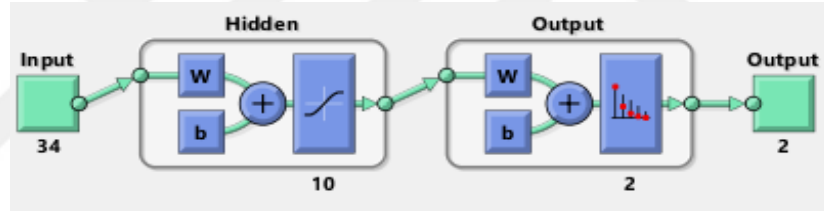
	Epilepsi	Sağlıklı
Epilepsi	10	0
Sağlıklı	0	2

Şekil 4.19: k-NN yöntemine uygulanan laboratuvar bulguları özellik seti için karışıklık matrisi.

Şekil 4.19'a göre, k-NN yöntemine laboratuvar bulguları özellik setinin uygulanarak sınıflandırılmasında test veri setindeki epilepsi hastası ve kontrol grubu sağlıklı bireylere ait veri doğru sınıflandırılmıştır.

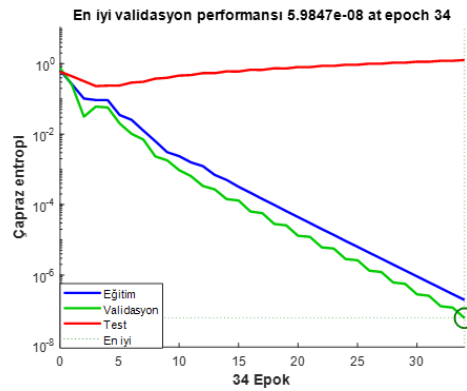
4.2.6. Model 2 için FFBP Yönteminin Sonuçları

Laboratuvar bulguları özellik setinin sınıflandırılması için $\mu = 0,001$ ile ara katman nöron sayısı 10 olarak düzenlenmiş ÇKA genel yapısı Şekil 4.20'de gösterilmiştir.



Şekil 4.20: Laboratuvar bulguları özellik seti için ÇKA yapısı.

FFBP yöntemine uygulanan laboratuvar bulguları özellik seti için eğitim işleminden %100 sınıflandırma doğruluğu elde edilmiştir. Şekil 4.21'e göre en iyi doğrulama (validasyon) performansı 34 iterasyon ile %99,99 sınıflandırma doğruluğu ile elde edilmiştir.



Şekil 4.21: FFBP yöntemine uygulanan laboratuvar bulguları özellik seti için en iyi doğrulama (validasyon) performans grafiği.

FFBP yöntemine uygulanan laboratuvar bulguları özellik setinin sınıflandırılmasında test veri setinin doğruluğu Şekil 4.22’de karışıklık matrisi ile verilmiştir ve test veri setinin sınıflandırma doğruluğunun %100 olduğu görülmektedir.

	Epilepsi	Sağlıklı
Epilepsi	10	0
Sağlıklı	0	2

Şekil 4.22: FFBP yöntemine uygulanan laboratuvar bulguları özellik seti için karışıklık matrisi.

Şekil 4.22’ye göre, FFBP yöntemine laboratuvar bulguları özellik setinin uygulanarak sınıflandırılmasında test veri setindeki epilepsi hastası ve kontrol grubu sağlıklı bireylere ait veri doğru sınıf olarak hesaplanmıştır. ÇKA modellemesinde giriş katmanından (x) ara katmanına doğru ağırlık değerleri (w) Tablo 4.9’da verilmiştir.

Tablo 4.9: Laboratuvar bulguları özellik seti için giriş katmanından ara katmanına doğru ağırlık değerleri.

Parametreler	Nöron "1"	Nöron "2"	Nöron "3"	Nöron "4"	Nöron "5"	Nöron "6"	Nöron "7"	Nöron "8"	Nöron "9"	Nöron "10"
NEUT%	0,01306432	0,01359224	0,01316065	0,01312973	0,01346018	0,01339206	0,01322306	0,01328054	0,01319398	0,01294434
LYMPH%	0,16918737	0,16871292	0,16916237	0,16901158	0,16926064	0,16935273	0,16855719	0,16917618	0,16925544	0,16909903
MONO%	0,06516363	0,06498847	0,06511117	0,06579236	0,06564901	0,06576378	0,06537638	0,06533036	0,06498025	0,06559161
EO%	-0,00405684	-0,00359874	-0,00358557	-0,00345100	-0,00401658	-0,00367631	-0,00367631	-0,00325199	-0,00333616	-0,00396074
BASO%	0,20502927	0,20504053	0,20441151	0,20464653	0,20470819	0,20503208	0,204555	0,20459234	0,20488037	0,2047353
IG%	0,25147778	0,25186744	0,25196398	0,25186032	0,25229378	0,2521117	0,25234232	0,25190743	0,25214084	0,2523165
IG#	0,39997591	0,40021666	0,40030191	0,39939002	0,39994214	0,40011413	0,40029549	0,39975553	0,39944907	0,40046796
PDW	-0,06817821	-0,06787098	-0,06790558	-0,06754164	-0,06827010	-0,06762759	-0,06830099	-0,06816242	-0,06810083	-0,06829782
MPV	-0,13349635	-0,13303982	-0,13337239	-0,13328679	-0,13349633	-0,13335318	-0,13356181	-0,13311188	-0,13377654	-0,13332175
P-LCR	0,50225879	0,50281679	0,5026405	0,50247065	0,50231746	0,50323204	0,50318213	0,50234622	0,50217459	0,50300503
PCT	-0,11052290	-0,11031315	-0,11012750	-0,11031285	-0,11044248	-0,11007299	-0,11001944	-0,11047830	-0,11017386	-0,11041960
NEUT#	-0,00545013	-0,00511220	-0,00498639	-0,00558761	-0,00477695	-0,00488631	-0,00513096	-0,00504814	-0,00554637	-0,00561083
LYMPH#	0,04536522	0,04550915	0,04582036	0,04556828	0,04555263	0,04538985	0,04603741	0,04542813	0,04586964	0,04542477
MONO#	-0,06393137	-0,06432620	-0,06436243	-0,06405770	-0,06445966	-0,06381336	-0,06380818	-0,06415356	-0,06422176	-0,06385163
EO#	-0,13053453	-0,13018608	-0,13061999	-0,13021245	-0,13060922	-0,13099331	-0,13071961	-0,13046468	-0,13080101	-0,13090806
BASO#	-0,11527307	-0,11542126	-0,11531670	-0,11551743	-0,11608843	-0,11589926	-0,11528993	-0,11549297	-0,11586556	-0,11564829
WBC	-0,12267033	-0,12266444	-0,12283594	-0,12211724	-0,12283032	-0,12266186	-0,12271996	-0,12277366	-0,12244388	-0,12273523
RBC	-0,14340505	-0,14306926	-0,14347485	-0,14325064	-0,14347331	-0,14359581	-0,14310315	-0,14307872	-0,14354088	-0,14327693
HGB	-0,13130728	-0,13135947	-0,13142691	-0,13150515	-0,13088555	-0,13095930	-0,13187808	-0,13082250	-0,13156012	-0,13158367
HCT	-0,13278345	-0,13212989	-0,13282291	-0,13203697	-0,13209323	-0,13191851	-0,13210142	-0,13206764	-0,13193449	-0,13271818
PLT	-0,08326957	-0,08340543	-0,08347211	-0,08390526	-0,08349606	-0,08336889	-0,08359697	-0,08369708	-0,08387425	-0,08353842
MCV	0,17686051	0,17693029	0,17734458	0,17716282	0,17664347	0,17654518	0,17741866	0,17729099	0,17679212	0,17742708
MCH	0,01884993	0,01864604	0,0188759	0,01883456	0,01898161	0,01890831	0,01835084	0,01841784	0,01871469	0,01905736

Tablo 4.9(devam): Laboratuvar bulguları özellik seti için giriş katmanından ara katmanına doğru ağırlık değerleri.

MCHC	-0,09455757	-0,09501949	-0,09475245	-0,09423687	-0,09484871	-0,09481585	-0,09434838	-0,09454686	-0,09443731	-0,09458484
RDW-SD	0,04745925	0,04727832	0,04695359	0,04747948	0,04740375	0,04712694	0,04751502	0,0477376	0,04720413	0,04689121
RDW-CV	0,04529358	0,04555549	0,04558026	0,045414	0,04582305	0,04557058	0,04541814	0,04531945	0,04559145	0,04595146
TSH	0,05651397	0,05590024	0,05652204	0,05642132	0,05607043	0,05593888	0,05649507	0,05643095	0,05622036	0,0567267
VİTAMİN B12	0,3374142	0,33785532	0,33773505	0,3374412	0,33779471	0,33791781	0,33831879	0,33721382	0,33786443	0,3382362
ALT	-0,03124464	-0,03111216	-0,03136166	-0,03116154	-0,03084063	-0,03119700	-0,03083914	-0,03098142	-0,03138492	-0,03156642
AST	0,27764429	0,27804908	0,27782957	0,27767799	0,27808482	0,27782646	0,27842548	0,27751446	0,27788166	0,27771694
Kreatinin	-0,41550424	-0,41520462	-0,41513248	-0,41465564	-0,41496210	-0,41580957	-0,41552423	-0,41506492	-0,41498871	-0,41535380
CKD-EP (eGFR)	0,40425108	0,40401804	0,40430368	0,4044864	0,40406522	0,40434831	0,40448709	0,40366318	0,40372871	0,40468462
Glukoz(Açlık Kan şekeri)	-0,07671042	-0,07637384	-0,07641555	-0,07641099	-0,07640904	-0,07648996	-0,07665993	-0,07635408	-0,07601579	-0,07609113
Sedimentasyon	0,36785464	0,36786668	0,36778834	0,36809254	0,36790108	0,36798824	0,36841606	0,36747617	0,36821781	0,36832133
Eğilim (bias)	-0,01193433	-0,01184061	-0,01202975	-0,01226960	-0,01166852	-0,01157944	-0,01132346	-0,01182799	-0,01219605	-0,01133058

Laboratuvar bulguları özellik seti için ara katmandan (w) çıkış katmanına (y) doğru ağırlık değerleri Tablo 4.10’da verilmiştir.

Tablo 4.10: Laboratuvar bulguları özellik seti için ara katmandan çıkış katmanına ağırlık değerleri.

–	Epilepsi	Sağlıklı
Nöron "1"	0,73686226	-0,73648992
Nöron "2"	0,73669204	-0,73715454
Nöron "3"	0,73717360	-0,73710024
Nöron "4"	0,73651550	-0,73640064
Nöron "5"	0,73687890	-0,73692020
Nöron "6"	0,73752093	-0,73781105
Nöron "7"	0,73851418	-0,73752251
Nöron "8"	0,73624231	-0,73573364
Nöron "9"	0,73636794	-0,73695380
Nöron "10"	0,73776921	-0,73834791
Eğilim (bias)	-2,16989249	2,16984790

Laboratuvar bulguları özellik setinde bulunan özelliklerin katkısı Tablo 4.11’de gösterilmiştir.

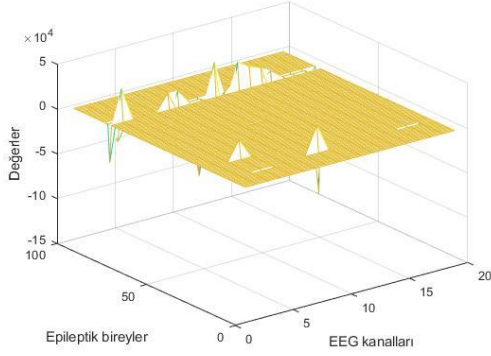
Tablo 4.11’e göre, laboratuvar bulguları özellik setindeki: NEUT%, EO%, IG%, PDW, MPV, PCT, NEUT#, LYMPH#, EO#, BASO#, RBC, HGB, HCT, MCV, MCH, RDW-SD, RDW-CV, ALT, Glukoz (Açlık kan şekeri) özelliklerinin epilepsi hastalığının tespitinde ayırt edici özellikler olduğu görülmektedir.

Tablo 4.11: Laboratuvar bulguları özellik setinde bulunan özelliklerin katkısı.

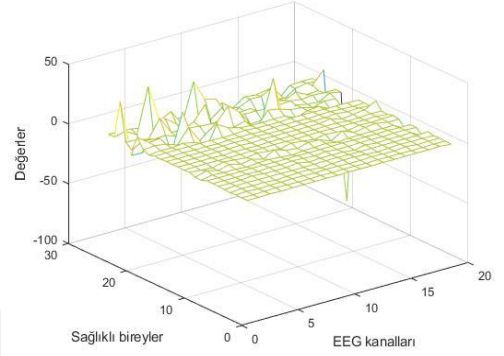
Laboratuvar bulguları parametreleri	Hata oranı
NEUT%	0,0000
LYMPH%	0,0164
MONO%	0,0082
EO%	0,0000
BASO%	0,0164
IG%	0,0000
IG#	0,0164
PDW	0,0000
MPV	0,0000
P-LCR	0,0082
PCT	0,0000
NEUT#	0,0000
LYMPH#	0,0000
MONO#	0,0082
EO#	0,0000
BASO#	0,0000
WBC	0,0082
RBC	0,0000
HGB	0,0000
HCT	0,0000
PLT	0,0082
MCV	0,0000
MCH	0,0000
MCHC	0,0082
RDW-SD	0,0000
RDW-CV	0,0000
TSH	0,0082
VİTAMİN B12	0,0082
ALT	0,0000
AST	0,0082
Kreatinin	0,0164
CKD-EP (eGFR)	0,0164
Glukoz(Açlık kan şekeri)	0,0000
Sedimentasyon	0,0164

4.3. EEG ÖZELLİK SETİNE AİT SINIFLANDIRMA SONUÇLARI (MODEL 3)

EEG özellik seti (122×19) boyutlu bir matristir. EEG çekimi sırasında epilepsi hastası bireylerden elde edilen sinyal değerler Şekil 4.23(a)'da, kontrol grubu sağlıklı bireylerden elde edilen sinyal özellik analizi Şekil 4.23(b)'deki grafiklerde gösterilmiştir.



(a) Epilepsi hastası bireyler



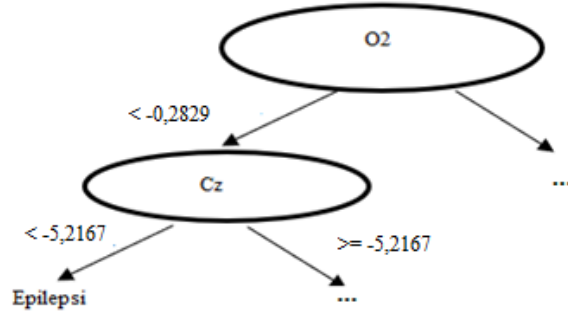
(b) Kontrol grubu sağlıklı bireyler

Şekil 4.23: EEG özellik seti.

Model 3, ID3, doğrusal regresyon, Naive Bayes, lineer DVM, k-NN, FFBP yöntemlerine EEG özellik seti uygulanarak epilepsi ve sağlıklı olarak sınıflandırılmıştır. Bu yöntemlere ait sınıflandırma doğruluk oranlarından en başarılı olanı Model 3'ün sonucu olmuştur.

4.3.1. Model 3 için ID3 Yönteminin Sonuçları

ID3 yöntemine uygulanan EEG özellik seti için elde edilen karar ağacı Şekil 4.24'te, karar kuralları ise Tablo 4.12'de gösterilmiştir. Eğitim işlemi sırasında ağacın dallarına ayırma işlemi öncesinde 200 değeri ile, ayırma işlemi sonrasında 50 değeri ile ve ağaç oluşturma derinliği 10 değeri olarak düzenlenmiştir. Eğitim setinin sınıflandırılması sonucunda 37 yapraklı 19 düğümlü karar ağacı oluşmuştur. Eğitim işlemi sınıflandırma doğruluğu %80,30 olarak elde edilmiştir.



Şekil 4.24: ID3 yöntemine uygulanan EEG özellik seti karar ağacı.

Tablo 4.12: ID3 yöntemine uygulanan EEG özellik setine ait karar kuralları.

If O2 < -0,2829
If Cz < -5,2167 then Sınıf= Epilepsi (4 bireyde)
else if Cz >= -5,2167
if F4 < -2,1356 then Sınıf= Sağlıklı (4 bireyde)
else if F4 >= -2,1356
If Fp1 < 0,5414
If F3 < -14,3434 then Sınıf= Epilepsi (3 bireyde)
else if F3 >= -14,3434
If O1 < 0,6592
If O2 < -0,6098
If F8 < 0,1575
If T8 < -0,1140 then Sınıf= Sağlıklı (2 bireyde)
else if T8 >= -0,1140
If P3 < 0,1459 then Sınıf= Sağlıklı (2 bireyde)
else if P3 >= 0,1459 then Sınıf= Epilepsi (4 bireyde)
else if F8 >= 0,1575 then Sınıf= Epilepsi (4 bireyde)
If O2 >= -0,6098 then Sınıf= Sağlıklı (12 bireyde)
else if O1 >= 0,6592 then Sınıf= Epilepsi (2 bireyde)
If Fp1 >= 0,5414 then Sınıf= Epilepsi (5 bireyde)
If O2 >= -0,2829
If P4 < -1,0503
If Fp2 < -8,9068 then Sınıf= Sağlıklı (1 bireyde)
else if Fp2 >= -8,9068
If C3 < 0,1524 then Sınıf= Epilepsi (3 bireyde)
else if C3 >= 0,1524 then Sınıf= Sağlıklı (1 bireyde)
else if P4 >= -1,0503
If Pz < -0,0488
If O2 < -0,0488 then Sınıf= Epilepsi (15 bireyde)
else if O2 >= -0,0488
If F3 < 0,0276 then Sınıf= Epilepsi (7 bireyde)
else if F3 >= 0,0276
If Fp1 < -0,0866
If T7 < -0,0765 then Sınıf= Sağlıklı (1 bireyde)
else if T7 >= -0,0765 then Sınıf= Epilepsi (3 bireyde)
else if Fp1 >= -0,0866 then Sınıf= Sağlıklı (2 bireyde)
If Pz >= -0,0488 then Sınıf= Epilepsi (47 bireyde)

10–kat çaprazlama işlemi test başarıımı için karışıklık matrisi Şekil 4.25’te gösterilmiştir. Şekil 4.25’e göre, test veri setinde ID3 yöntemi epilepsi hastası olan 10 bireyden 3’ünün verisini yanlış sınıfa yani sağlıklı sınıfı olarak tanımlamıştır. Kontrol grubu sağlıklı bireylere ait veriyi doğru sınıfa atamıştır.

	Epilepsi	Sağlıklı
Epilepsi	7	3
Sağlıklı	0	2

Şekil 4.25: ID3 yöntemine uygulanan EEG özellik seti için karışıklık matrisi.

Şekil 4.25'e göre, 10–kat çaprazlama işlemine göre epilepsi ve kontrol grubu sağlıklı bireylere ait EEG özellik setinin sınıflandırılmasında test başarımları %75,00 olarak elde edilmiştir.

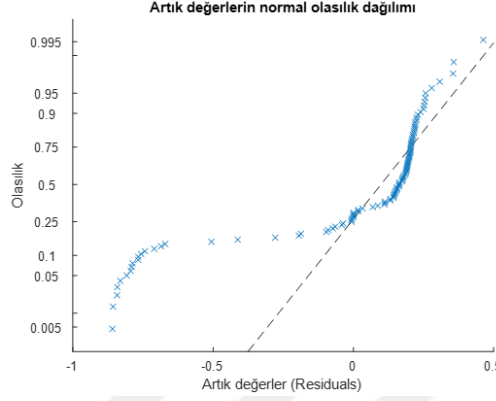
4.3.2. Model 3 için Doğrusal Regresyon Yönteminin Sonuçları

Doğrusal regresyon yöntemine uygulanan EEG özellik seti için 19 parametre ile doğrusal formül elde edilmiştir. y doğrusal regresyon denkleminde kestirimi yapılan katsayılar Tablo 4.13'te verilmiştir. Tablo 4.13'e göre gözlem sayısı 110, hata serbestlik derecesi 90, kök ortalama kare hata değeri 0,388, R–kare 0,255, düzeltilmiş R–kare 0,0978, F istatistiği ile sabit model 1,62, p –değeri 0,0677 olarak elde edilmiştir. Doğrusal regresyon yöntemi ile modellenen EEG özellik setine ait ortalama hata değeri, bağımsızlık derecesi, F değeri ve p –değeri sonuçları Tablo 4.13'te verilmiştir. EEG özellik setinin doğrusal regresyon analizinde, x_6 : F4, x_7 : F8, x_8 : T7, x_{14} : P3, x_{15} : Pz, x_{16} : P4 parametrelerinin p –değerleri 0,05 değerinin altında olduğundan epilepsi ve kontrol grubu sağlıklı bireylerin sınıflandırılmasında anlamlı özellikler olduğu sonucu çıkarılmıştır.

Tablo 4.13: EEG özellik seti doğrusal regresyon modelleme sonuçları.

EEG kanalları	Ortalama Hata	DF	F	p–Değeri
x1	0,178	1	0,287	0,278
x2	0,083	1	0,031	0,458
x3	0,226	1	0,263	0,223
x4	0,056	1	0,283	0,541
x5	0,146	1	3,676	0,326
x6	1,515	1	1,463	0,002
x7	2,190	1	20,554	0,000
x8	1,348	1	13,759	0,003
x9	0,009	1	0,419	0,801
x10	0,210	1	19,802	0,240
x11	0,040	1	0,175	0,606
x12	0,051	1	0,141	0,560
x13	0,024	1	0,576	0,688
x14	0,661	1	0,005	0,038
x15	0,517	1	0,162	0,002
x16	0,517	1	0,245	0,002
x17	0,055	1	0,198	0,545
x18	0,048	1	1,313	0,572
x19	0,136	1	1,660	0,344
Hata	13,552	90	0,150	

Doğrusal regresyon yöntemine uygulanan EEG özellik setinin sınıflandırılmasında artık değerlerin normal olasılık dağılım grafiği Şekil 4.26’da verilmiştir. Eğitim işlemi için doğrusal regresyon yönteminin sınıflandırma doğruluğu %71,30 olarak elde edilmiştir.



Şekil 4.26: EEG özellik setinin sınıflandırılmasında artık değerlerin normal olasılık dağılım grafiği.

Şekil 4.27’ye göre, 10–kat çaprazlama işlemine göre epilepsi ve kontrol grubu sağlıklı bireylere ait EEG özellik setinin sınıflandırılmasında test başarımları %91,66 olarak elde edilmiştir.

	Epilepsi	Sağlıklı
Epilepsi	9	1
Sağlıklı	0	2

Şekil 4.27: Doğrusal regresyon yöntemine uygulanan EEG özellik seti için karışıklık matrisi.

Şekil 4.27’ye göre, doğrusal regresyon yöntemine EEG özellik setinin uygulanarak sınıflandırılmasında test veri setindeki epilepsi hastası 10 bireye ait veriden sadece 1’ini yanlış sınıflandırmıştır. Kontrol grubu sağlıklı bireylere ait veri ise doğru sınıflandırılmıştır.

4.3.3. Model 3 için Naive Bayes Yönteminin Sonuçları

Naive Bayes yöntemine EEG özellik seti uygulanmıştır. Çekirdek dağılım fonksiyonu kullanılarak Naive Bayes yöntemi ile eğitim setinin sınıflandırma doğruluk oranı %70,50 olarak elde edilmiştir. Şekil 4.28’e göre, 10–kat çaprazlama işlemine göre epilepsi ve kontrol grubu sağlıklı bireylere ait EEG özellik setinin sınıflandırılmasında test başarımları %83,33 olarak elde edilmiştir.

	Epilepsi	Sağlıklı
Epilepsi	10	0
Sağlıklı	2	0

Şekil 4.28: Naive Bayes yöntemine uygulanan EEG özellik seti için karışıklık matrisi.

Şekil 4.28'e göre, Naive Bayes yöntemine EEG özellik setinin uygulanarak sınıflandırılmasında test veri setindeki epilepsi hastası bireylere ait veri doğru sınıflandırılırken, kontrol grubu sağlıklı bireylere ait veri yanlış sınıflandırılmıştır.

4.3.4. Model 3 için DVM Yönteminin Sonuçları

Lineer DVM yöntemine EEG özellik seti uygulanmıştır. Eğitim veri setinin sınıflandırma doğruluğu %75,40 olarak elde edilmiştir. Test veri setinin sınıflandırmadaki test başarımı Şekil 4.29'da karışıklık matrisinde gösterilmiştir ve test veri setinin sınıflandırma doğruluğu %83,33 olarak elde edilmiştir.

	Epilepsi	Sağlıklı
Epilepsi	10	0
Sağlıklı	2	0

Şekil 4.29: DVM yöntemine uygulanan EEG özellik seti için karışıklık matrisi.

Şekil 4.29'a göre, lineer DVM yöntemine EEG özellik setinin uygulanarak sınıflandırılmasında test veri setindeki epilepsi hastası bireylere ait veri doğru sınıflandırılırken, kontrol grubu sağlıklı bireylere ait veri yanlış sınıflandırılmıştır.

4.3.5. Model 3 için k-NN Yönteminin Sonuçları

k-NN yöntemine EEG özellik seti uygulanmıştır. Öklid uzaklığına göre hesaplanan en yakın 10-komşuluk için eğitim veri setinin sınıflandırma doğruluğu %78,70 olarak elde edilmiştir. Test veri setinin doğruluğu Şekil 4.30'da karışıklık matrisi ile verilmiştir ve test verisinin sınıflandırma doğruluğu %83,33 olarak elde edilmiştir.

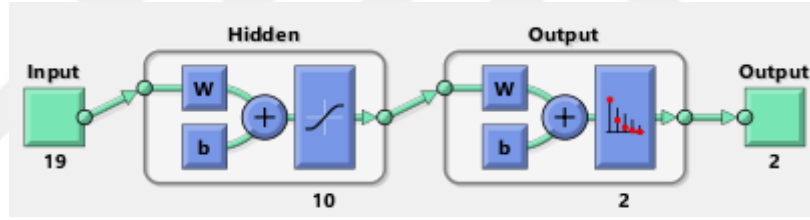
	Epilepsi	Sağlıklı
Epilepsi	10	0
Sağlıklı	2	0

Şekil 4.30: k-NN yöntemine uygulanan EEG özellik seti için karışıklık matrisi.

Şekil 4.30'a göre, k-NN yöntemine EEG özellik setinin uygulanarak sınıflandırılmasında test veri setindeki epilepsi hastası bireylere ait veri doğru sınıflandırılırken, kontrol grubu sağlıklı bireylere ait veri yanlış sınıflandırılmıştır.

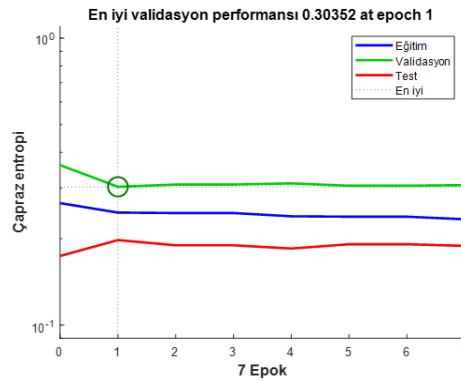
4.3.6. Model 3 için FFBP Yönteminin Sonuçları

EEG özellik setinin sınıflandırılması için $\mu = 0,001$ ile ara katman nöron sayısı 10 olarak düzenlenmiş ÇKA yapısı Şekil 4.31'de verilmiştir.



Şekil 4.31: EEG özellik seti için ÇKA yapısı.

FFBP yöntemine uygulanan EEG özellik seti için eğitim işleminden %79,50 sınıflandırma doğruluğu elde edilmiştir. Şekil 4.32'ye göre en iyi doğrulama (validasyon) performansı 7 iterasyon ile %76,60 sınıflandırma doğruluğu ile elde edilmiştir.



Şekil 4.32: FFBP yöntemine uygulanan EEG özellik seti için en iyi doğrulama (validasyon) performans grafiği.

FFBP yöntemine uygulanan EEG özellik setinin sınıflandırılmasında test veri setinin doğruluğu Şekil 4.33'te karışıklık matrisi ile verilmiştir ve test veri setinin sınıflandırma doğruluğu %83,33 olarak hesaplanmıştır.

	Epilepsi	Sağlıklı
Epilepsi	10	2
Sağlıklı	0	0

Şekil 4.33: FFBP yöntemine uygulanan EEG özellik seti için karışıklık matrisi.

Şekil 4.33'e göre, FFBP yöntemine EEG özellik setinin uygulanarak sınıflandırılmasında test verisindeki epilepsi hastası bireylere ait 12 bireye ait veriden 10'unu doğru sınıflandırabilmiştir. FFBP yöntemi ile modelleme giriş katmanından (x) ara katmanına (y) doğru ağırlık değerleri Tablo 4.14'te verilmiştir.

Tablo 4.14: EEG özellik seti için giriş katmanından ara katmanına doğru ağırlık değerleri.

EEG kanalları	Nöron "1"	Nöron "2"	Nöron "3"	Nöron "4"	Nöron "5"	Nöron "6"	Nöron "7"	Nöron "8"	Nöron "9"	Nöron "10"
Fp1	-0,02661008	-0,02687145	-0,02665708	-0,02622077	-0,02706109	-0,02615858	-0,02670525	-0,02708006	-0,02660848	-0,02671982
Fp2	-0,04210286	-0,04206953	-0,04240107	-0,04283255	-0,04275074	-0,04215717	-0,04261096	-0,04253728	-0,04194032	-0,04263558
F7	0,35040365	0,3504714	0,35049664	0,35061889	0,35122535	0,3507201	0,35080208	0,35081425	0,35071822	0,35039409
F3	-0,01211197	-0,01235153	-0,01284277	-0,01290779	-0,01294399	-0,01294945	-0,01286482	-0,01254184	-0,01269434	-0,01268416
Fz	0,01276898	0,01261807	0,01274042	0,01248833	0,01306103	0,01292156	0,01313924	0,0130638	0,01334775	0,01270112
F4	0,0206117	0,02053263	0,02045622	0,0211617	0,02070501	0,02065865	0,02060335	0,02048741	0,02053989	0,02054724
F8	-0,01799597	-0,01838724	-0,01845358	-0,01782945	-0,01871075	-0,01852721	-0,01787993	-0,01784594	-0,01856467	-0,01817998
T7	-0,02738698	-0,02715061	-0,02784009	-0,02797470	-0,02756612	-0,02758751	-0,02777369	-0,02724598	-0,02767345	-0,02796699
C3	0,12951877	0,12999008	0,13004793	0,1303656	0,13063028	0,12977556	0,13092414	0,12948765	0,12923199	0,12933059
Cz	0,02656138	0,02709527	0,02656153	0,02722438	0,02646322	0,02695553	0,02686156	0,02658339	0,02686497	0,02701461
C4	-0,02423449	-0,02377456	-0,02386299	-0,02391823	-0,02362399	-0,02394842	-0,02368287	-0,02349546	-0,02351909	-0,02345086
T8	0,0249953	0,02466309	0,02538356	0,02533851	0,02463892	0,02486557	0,02454398	0,02532614	0,02545083	0,02532865
P7	0,01305357	0,01355603	0,0132421	0,01294865	0,01276684	0,01312867	0,01359057	0,01287326	0,0126706	0,01322229
P3	-0,05744156	-0,05764898	-0,05799884	-0,05757826	-0,05794598	-0,05749361	-0,05784877	-0,05706246	-0,05769608	-0,05761334
Pz	-0,02404944	-0,02382860	-0,02383151	-0,02352723	-0,02431338	-0,02420404	-0,02425717	-0,02364892	-0,02339388	-0,02374403
P4	0,93222607	0,93515774	0,93480218	0,93552259	0,93751908	0,93348584	0,9376284	0,93426317	0,9325861	0,93220916
P8	-0,22427937	-0,22484203	-0,22462589	-0,22475807	-0,22531622	-0,22410979	-0,22524610	-0,22479937	-0,22417474	-0,22426543
O1	-0,10970149	-0,10991634	-0,10953033	-0,10973655	-0,10958626	-0,10995872	-0,11019178	-0,10940794	-0,10955743	-0,10918434
O2	-0,02673443	-0,02702272	-0,02743334	-0,02686793	-0,02745126	-0,02701365	-0,02736070	-0,02745311	-0,02732985	-0,02731757
Eğilim (bias)	0,040849	0,0412308	0,04084734	0,04113756	0,04133453	0,04083162	0,04138852	0,04134423	0,0405327	0,04055739

EEG özellik seti için ara katmandan (w) çıkış katmanına (y) doğru ağırlık değerleri Tablo 4.15'te verilmiştir.

Tablo 4.15: EEG özellik seti için ara katmandan çıkış katmanına ağırlık değerleri.

–	Epilepsi	Sağlıklı
Nöron "1"	0,48063863	–0,47965146
Nöron "2"	0,48312301	–0,48236729
Nöron "3"	0,48230650	–0,48278449
Nöron "4"	0,48288447	–0,48348562
Nöron "5"	0,48531331	–0,48509728
Nöron "6"	0,48147738	–0,48122822
Nöron "7"	0,48487967	–0,48534190
Nöron "8"	0,48175038	–0,48227576
Nöron "9"	0,48065631	–0,48056029
Nöron "10"	0,48022085	–0,48018005
Eğilim (bias)	–1,19673193	1,19660549

EEG özellik setinde bulunan özelliklerin öğrenmede ağırlık katkısı Tablo 4.16'da gösterilmiştir.

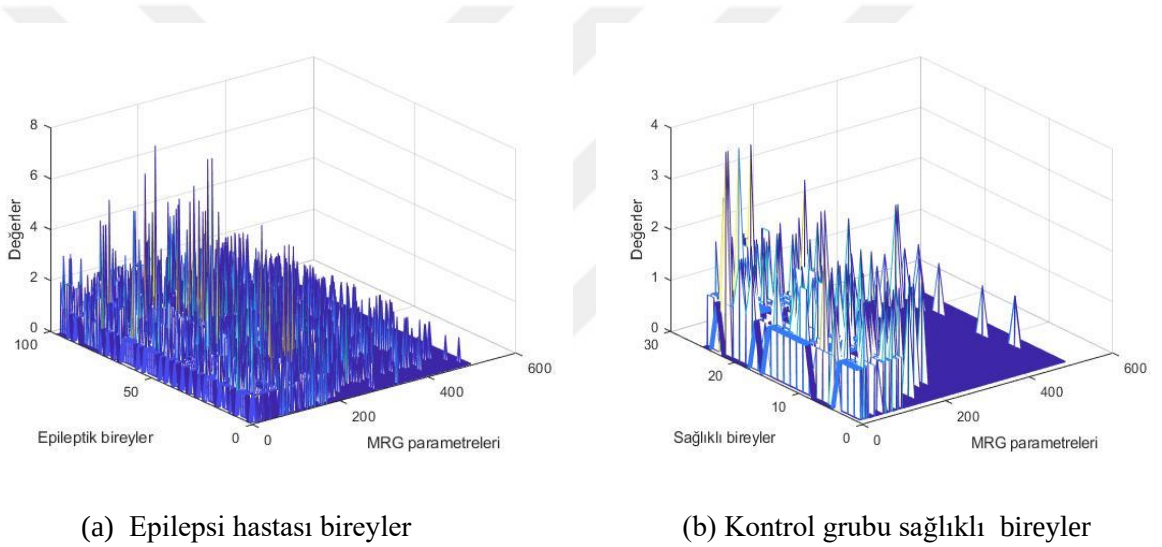
Tablo 4.16: EEG özellik setinde bulunan özelliklerin katkısı.

EEG kanalları	Hata Oranı
Fp1	0,1803
Fp2	0,1803
F7	0,1885
F3	0,1803
Fz	0,1803
F4	0,1803
F8	0,1803
T7	0,1803
C3	0,1803
Cz	0,1803
C4	0,1803
T8	0,1803
P7	0,1803
P3	0,1803
Pz	0,1803
P4	0,2131
P8	0,1803
O1	0,1803
O2	0,1803

Tablo 4.16'ya göre, EEG özellik setindeki tüm kanalların hastalığın tespitinde aynı hata oranında (0,1803) yani aynı etki oranına sahip olduğu görülmektedir. FFBP ağında tüm EEG kanallarının ayırt edici olduğu görülmektedir.

4.4. MRG ÖZELLİK SETİNE AİT SINIFLANDIRMA SONUÇLARI (MODEL 4)

MRG özellik seti (122×502) boyutlu bir matristir. Radyolog tarafından yorumlanan MRG raporlarından epilepsi hastası bireylerden elde edilen değerlerin özellik analizi Şekil 4.34 (a)'da, kontrol grubu sağlıklı bireylerden elde edilen değerlerin özellik analizi Şekil 4.34(b)'deki grafiklerde gösterilmiştir.



Şekil 4.34: MRG özellik seti.

Model 4'e göre, ID3, doğrusal regresyon, Naive Bayes, lineer DVM, k-NN, FFBP, LSTM yöntemlerine MRG özellik seti uygulanarak epilepsi ve sağlıklı olarak sınıflandırılmıştır. Bu yöntemlere ait sınıflandırma doğruluk oranlarından en başarılı olanı Model 4'ün sonucu olmuştur.

4.4.1. Model 4 için ID3 Yönteminin Sonuçları

ID3 yöntemine uygulanan MRG özellik seti için elde edilen karar ağacı karar kuralları Tablo 4.17'de gösterilmiştir. Eğitim işlemi sırasında ağacın dallarına ayırma işlemi öncesinde 200 değeri ile, ayırma işlemi sonrasında 50 değeri ile ve ağaç oluşturma derinliği 10 değeri olarak düzenlenmiştir. Eğitim setinin sınıflandırma sonucunda 71 yapraklı 36 düğümlü karar ağacı oluşmuştur. Eğitim işlemi setinin sınıflandırma doğruluğu %82,00 olarak elde edilmiştir.

Tablo 4.17 (devam): ID3 yöntemine uygulanan MRG özellik setine ait karar kuralları.

If ((mukozal < 1,5000) and (seviye < 1,0000) and (kemik < 0,5000) and (düzey < 1,5000) and (sekonder < 0,5000) and (basion < 1,5000) and (multipl < 0,5000) and (dural < 0,5000) and (hipoplazi < 1,0000) and (daralma < 0,5000) and (olgu < 1,0000) and (olgu >= 1,0000) then Sınıf = Epilepsi (1 bireyde)
If ((mukozal < 1,5000) and (seviye < 1,0000) and (kemik < 0,5000) and (düzey < 1,5000) and (sekonder < 0,5000) and (basion < 1,5000) and (multipl < 0,5000) and (dural < 0,5000) and (hipoplazi < 1,0000) and (daralma >= 0,5000) and (fissür < 0,5000) then Sınıf = Sağlıklı (6 bireyde)
If ((mukozal < 1,5000) and (seviye < 1,0000) and (kemik < 0,5000) and (düzey < 1,5000) and (sekonder < 0,5000) and (basion < 1,5000) and (multipl < 0,5000) and (dural < 0,5000) and (hipoplazi < 1,0000) and (daralma >= 0,5000) and (fissür >= 0,5000) and (atrofi < 1,5000) then Sınıf = Epilepsi (1 bireyde)
If ((mukozal < 1,5000) and (seviye < 1,0000) and (kemik < 0,5000) and (düzey < 1,5000) and (sekonder < 0,5000) and (basion < 1,5000) and (multipl < 0,5000) and (dural < 0,5000) and (hipoplazi < 1,0000) and (daralma >= 0,5000) and (fissür >= 0,5000) and (atrofi >= 1,5000) then Sınıf= Sağlıklı (1 bireyde)
If ((mukozal < 1,5000) and (seviye < 1,0000) and (kemik < 0,5000) and (düzey < 1,5000) and (sekonder < 0,5000) and (basion < 1,5000) and (multipl < 0,5000) and (dural < 0,5000) and (hipoplazi >= 1,0000) then Sınıf = Epilepsi (2 bireyde)
If ((mukozal < 1,5000) and (seviye < 1,0000) and (kemik < 0,5000) and (düzey < 1,5000) and (sekonder < 0,5000) and (basion < 1,5000) and (multipl < 0,5000) and dural >= 0,5000 and (sekel < 1,0000) and (paranazal < 1,5000) and (cilt < 1,0000) and (ekstraaksiyel < 1,0000) and (deviye < 1,5000) and (griak < 0,5000) then Sınıf = Epilepsi (5 bireyde)
If ((mukozal < 1,5000) and (seviye < 1,0000) and (kemik < 0,5000) and (düzey < 1,5000) and (sekonder < 0,5000) and (basion < 1,5000) and (multipl < 0,5000) and dural >= 0,5000 and (sekel < 1,0000) and (paranazal < 1,5000) and (cilt < 1,0000) and (ekstraaksiyel < 1,0000) and (deviye < 1,5000) and (griak >= 0,5000) then Sınıf = Epilepsi (81,82 % of 11 bireyde)
If ((mukozal < 1,5000) and (seviye < 1,0000) and (kemik < 0,5000) and (düzey < 1,5000) and (sekonder < 0,5000) and (basion < 1,5000) and (multipl < 0,5000) and dural >= 0,5000 and (sekel < 1,0000) and (paranazal < 1,5000) and (cilt < 1,0000) and (ekstraaksiyel < 1,0000) and (deviye >= 1,5000) then Sınıf = Epilepsi (83,33 % of 6 bireyde)
If ((mukozal < 1,5000) and (seviye < 1,0000) and (kemik < 0,5000) and (düzey < 1,5000) and (sekonder < 0,5000) and (basion < 1,5000) and (multipl < 0,5000) and dural >= 0,5000 and (sekel < 1,0000) and (paranazal < 1,5000) and (cilt < 1,0000) and (ekstraaksiyel >= 1,0000) and (cilt >= 1,0000) then Sınıf = Epilepsi (1 bireyde)
If ((mukozal < 1,5000) and (seviye < 1,0000) and (kemik < 0,5000) and (düzey < 1,5000) and (sekonder < 0,5000) and (basion < 1,5000) and (multipl < 0,5000) and dural >= 0,5000 and (sekel < 1,0000) and (paranazal >= 1,5000) then Sınıf = Epilepsi (2 bireyde)
If ((mukozal < 1,5000) and (seviye < 1,0000) and (kemik < 0,5000) and (düzey < 1,5000) and (sekonder < 0,5000) and (basion < 1,5000) and (multipl < 0,5000) and dural >= 0,5000 and (sekel >= 1,0000) then Sınıf = Epilepsi (4 bireyde)
If ((mukozal < 1,5000) and (seviye < 1,0000) and (kemik < 0,5000) and (düzey < 1,5000) and (sekonder < 0,5000) and (basion < 1,5000) and (multipl >= 0,5000) then Sınıf = Epilepsi (8 bireyde)
If ((mukozal < 1,5000) and (seviye < 1,0000) and (kemik < 0,5000) and (düzey < 1,5000) and (sekonder < 0,5000) and (basion >= 1,5000) then Sınıf = Epilepsi (3 bireyde)
If ((mukozal < 1,5000) and (seviye < 1,0000) and (kemik < 0,5000) and (düzey < 1,5000) and (sekonder >= 0,5000) then Sınıf = Epilepsi (6 bireyde)
If ((mukozal < 1,5000) and (seviye < 1,0000) and (kemik < 0,5000) and (düzey >= 1,5000) then Sınıf = Epilepsi (8 bireyde)
If ((mukozal < 1,5000) and (seviye < 1,0000) and (kemik >= 0,5000) then Sınıf = Epilepsi (9 bireyde)
If ((mukozal < 1,5000) and (seviye >= 1,0000) then Sınıf = Epilepsi (6 bireyde)
If ((mukozal < 1,5000) and (mukozal >= 1,5000) and (seri < 0,5000) then Sınıf = Sağlıklı (2 bireyde)
If ((mukozal < 1,5000) and (mukozal >= 1,5000) and (seri >= 0,5000) then Sınıf = Epilepsi (1 bireyde)

10–kat çaprazlama işlemi test başarımı için karışıklık matrisi Şekil 4.35’te gösterilmiştir. Şekil 4.35’e göre, test veri setinde ID3 yöntemi epilepsi hastası bireylere ait 12 veriden 7’sini ve kontrol grubu sağlıklı bireylere ait veriyi doğru sınıflandırmıştır.

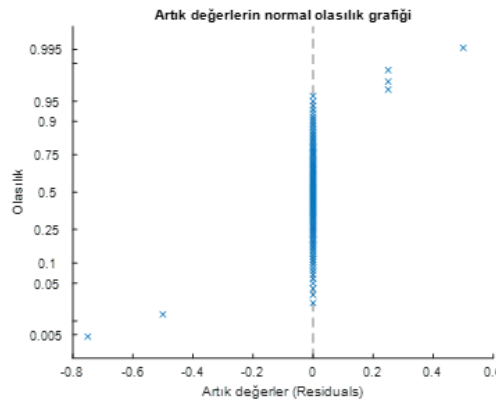
	Epilepsi	Sağlıklı
Epilepsi	7	3
Sağlıklı	0	2

Şekil 4.35: ID3 yöntemine uygulanan MRG özellik seti için karışıklık matrisi.

Şekil 4.35’e göre, 10–kat çaprazlama işlemine göre epilepsi ve kontrol grubu sağlıklı bireylere ait MRG özellik setinin sınıflandırılmasında test başarımı %75 olarak elde edilmiştir.

4.4.2. Model 4 için Doğrusal Regresyon Yönteminin Sonuçları

Doğrusal regresyon yöntemine uygulanan MRG özellik seti için 502 parametre ile lineer formül elde edilmiştir. y doğrusal regresyon denkleminde gözlem sayısı 110, hata serbestlik derecesi 6, kök ortalama kare hata değeri 0,546, R–kare 0,931, düzeltilmiş R–kare –0,248, F istatistiği ile sabit model 0,789, p –değeri 0,721 olarak elde edilmiştir. Doğrusal regresyona uygulanan MRG özellik setinin sınıflandırılmasında artık değerlerin normal olasılık dağılım grafiği Şekil 4.36’de verilmiştir. Eğitim işlemi için doğrusal regresyon yönteminin sınıflandırma doğruluğu %54,90 olarak elde edilmiştir.



Şekil 4.36: MRG özellik setinin sınıflandırılmasında artık değerlerin normal olasılık dağılım grafiği.

Şekil 4.37’ye göre 10–kat çaprazlama işlemine göre epilepsi ve kontrol grubu sağlıklı bireylere ait MRG özellik setinin sınıflandırılmasında test başarımı %41,66 olarak elde edilmiştir.

	Epilepsi	Sağlıklı
Epilepsi	4	6
Sağlıklı	1	1

Şekil 4.37: Doğrusal regresyon yöntemine uygulanan MRG özellik seti için karışıklık matrisi.

Şekil 4.37'ye göre, doğrusal regresyon yöntemine MRG özellik setinin uygulanarak sınıflandırılmasında test veri setindeki epilepsi hastası 10 bireye ait veriden 4'ü ve kontrol grubu sağlıklı bireylere ait 2 veriden 1 veri doğru sınıflandırılmıştır.

4.4.3. Model 4 için Naive Bayes Yönteminin Sonuçları

Naive Bayes yöntemine MRG özellik seti uygulanmıştır. Çekirdek dağılım fonksiyonu kullanılarak Naive Bayes yöntemi ile eğitim setinin sınıflandırma doğruluk oranı %79,50 elde edilmiştir. Şekil 4.38'e göre, 10-kat çaprazlama işlemine göre epilepsi ve kontrol grubu sağlıklı bireylere ait MRG özellik setinin sınıflandırılmasında test başarımları %83,33 olarak elde edilmiştir.

	Epilepsi	Sağlıklı
Epilepsi	10	0
Sağlıklı	2	0

Şekil 4.38: Naive Bayes yöntemine uygulanan MRG özellik seti için karışıklık matrisi.

Şekil 4.38'e göre, Naive Bayes yöntemine MRG özellik setinin uygulanarak sınıflandırılmasında test verisindeki epilepsi hastası bireylere ait veri doğru sınıflandırılırken, kontrol grubu sağlıklı bireylere ait veri yanlış sınıflandırılmıştır.

4.4.4. Model 4 için DVM Yönteminin Sonuçları

Lineer DVM yöntemine MRG özellik seti uygulanmıştır. Eğitim verisinin sınıflandırma doğruluğu %77,90 elde edilmiştir. Test veri setinin sınıflandırmadaki test başarımları Şekil 4.39'da karışıklık matrisinde gösterilmiştir ve test verisinin sınıflandırma doğruluğu %66,66 olarak elde edilmiştir.

	Epilepsi	Sağlıklı
Epilepsi	7	3
Sağlıklı	1	1

Şekil 4.39: DVM yöntemine uygulanan MRG özellik seti için karışıklık matrisi.

Şekil 4.39'a göre, lineer DVM yöntemine MRG özellik setinin uygulanarak sınıflandırılmasında test veri setindeki epilepsi hastası 10 bireyden 7'sini doğru sınıflandırırken, kontrol grubu sağlıklı bireylere ait 2 veriden 1'i doğru sınıflandırılmıştır.

4.4.5. Model 4 için k-NN Yönteminin Sonuçları

k-NN yöntemine MRG özellik seti uygulanmıştır. Öklid uzaklığına göre hesaplanan en yakın 10-komşuluk için eğitim veri setinin sınıflandırma doğruluğu %75,40 olarak elde edilmiştir. Test veri setinin test doğruluğu Şekil 4.40'de karışıklık matrisi ile verilmiştir ve test verisinin sınıflandırma doğruluğu %66,66 olarak elde edilmiştir.

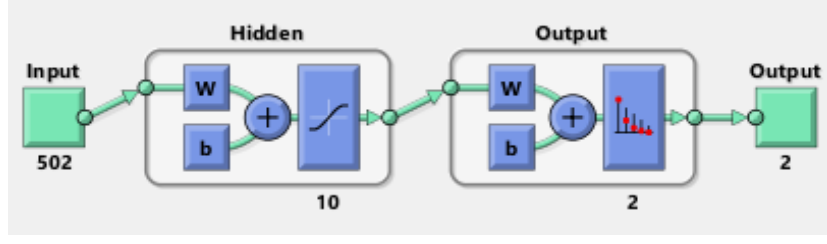
	Epilepsi	Sağlıklı
Epilepsi	7	4
Sağlıklı	0	1

Şekil 4.40: k-NN yöntemine uygulanan MRG özellik seti için karışıklık matrisi.

Şekil 4.40'e göre, k-NN yöntemine MRG özellik setinin uygulanarak sınıflandırılmasında test veri setindeki epilepsi hastası 11 bireyden 7'si ve kontrol grubu sağlıklı 1 bireye ait veri doğru sınıflandırılmıştır.

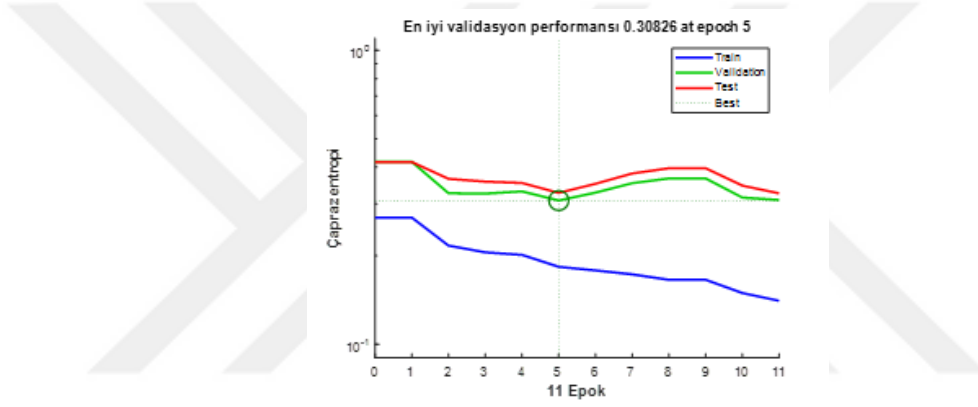
4.4.6. Model 4 için FFBP Yönteminin Sonuçları

MRG özellik setinin sınıflandırılması için $\mu = 0,001$ ile ara katman nöron sayısı 10 olarak düzenlenmiş ÇKA genel yapısı Şekil 4.41'de verilmiştir.



Şekil 4.41: MRG özellik seti için ÇKA yapısı.

FFBP yöntemine uygulanan MRG özellik seti için eğitim işleminden %79,50 sınıflandırma doğruluğu elde edilmiştir. Şekil 4.42'ye göre en iyi doğrulama (validasyon) performansı 11 iterasyon ile %69,20 sınıflandırma doğruluğu ile elde edilmiştir.



Şekil 4.42: FFBP yöntemine uygulanan MRG özellik seti için en iyi doğrulama (validasyon) performans grafiği.

FFBP yöntemine uygulanan MRG özellik setinin sınıflandırılmasında test veri setinin doğruluğu Şekil 4.43'te karışıklık matrisi ile verilmiştir ve test veri setinin sınıflandırma doğruluğunun %83,33 olduğu görülmektedir.

	Epilepsi	Sağlıklı
Epilepsi	10	2
Sağlıklı	0	0

Şekil 4.43: FFBP yöntemine uygulanan MRG özellik seti için karışıklık matrisi.

Şekil 4.43'e göre, FFBP yöntemine MRG özellik setinin uygulanarak sınıflandırılmasında test verisindeki epilepsi hastası bireylere ait 12 veriden 10 bireye ait veri doğru sınıflandırılmıştır.

FFBP modellemesinde giriş katmanından (x) ara katmanına doğru ağırlık değerleri (w) Tablo 4.18’de verilmiştir.

Tablo 4.18: MRG özellik seti için giriş katmanından ara katmanına doğru ağırlık değerleri.

Parametre	Nöron "1"	Nöron 2	Nöron "3"	Nöron "4"	Nöron "5"	Nöron "6"	Nöron "7"	Nöron "8"	Nöron "9"	Nöron "10"
patern	-0,01152227	-0,01090193	-0,01136191	-0,01199297	-0,01154735	-0,01140167	-0,01183292	-0,01121767	-0,01179860	-0,01143426
sirkülasyon	-0,01179193	-0,01167405	-0,01120139	-0,01168647	-0,01207518	-0,01169054	-0,01088894	-0,01152214	-0,01174210	-0,01169662
void	-0,01189164	-0,01165849	-0,01116341	-0,01155298	-0,01217622	-0,01146536	-0,01104606	-0,01107866	-0,01174121	-0,01134464
forme	-0,02022442	-0,02013569	-0,01987830	-0,02009556	-0,02054859	-0,02042804	-0,02022163	-0,02016322	-0,02009284	-0,01981561
parasellar	-0,03053387	-0,03022473	-0,03003432	-0,03072621	-0,03019060	-0,03005745	-0,02974014	-0,03027743	-0,03049964	-0,03030804
yerinde	-0,02016606	-0,02009361	-0,02066927	-0,02086692	-0,02060604	-0,02076737	-0,01983154	-0,02050710	-0,02012749	-0,01982020
eksternal	-0,00444059	-0,00327181	-0,00412368	-0,00399954	-0,00427520	-0,00366187	-0,00377758	-0,00415294	-0,00410245	-0,00374096
kranioservi kal	-0,00044345	-0,00001277	0,00048733	-0,00053401	-0,00010345	0,00045019	-0,00026137	-0,00046501	0,00018998	0,00029121
kemik	-0,00585363	-0,00601503	-0,00599137	-0,00648862	-0,00642128	-0,00652351	-0,00562640	-0,00584169	-0,00658178	-0,00578136
mesafe	-0,03681644	-0,03771163	-0,03748007	-0,03627453	-0,03629736	-0,03651658	-0,03744887	-0,03604672	-0,03697866	-0,03754055
intrakranial	0,24421022	0,24293987	0,24335153	0,24445671	0,24299511	0,24378062	0,24282386	0,24278045	0,24301555	0,24191001
kortikomed ulla	-0,02081244	-0,01981108	-0,02030972	-0,02043548	-0,02100325	-0,02044747	-0,01990936	-0,02093795	-0,02078195	-0,02026950
kontrast	0,0818574	0,08259264	0,08273745	0,08118255	0,08192437	0,08157647	0,08286303	0,08151038	0,08248123	0,08291164
kraniofasial	0,00173646	0,0014227	0,00179316	0,00104246	0,0009736	0,00153219	0,00203436	0,0010003	0,00115539	0,00198825
milimetrik	-0,14789394	-0,14767567	-0,14764814	-0,14779481	-0,14725344	-0,14826618	-0,14700254	-0,14670854	-0,14697558	-0,14740166
postkontrast	-0,00865458	-0,00904998	-0,00913430	-0,00802055	-0,00828156	-0,00806159	-0,00874709	-0,00875766	-0,00858386	-0,00843470
periferik	0,03151842	0,03133317	0,03167523	0,03124706	0,03210534	0,03161764	0,03110091	0,03124588	0,0311174	0,03158412
duralnoz	0,82355678	0,82552632	0,82638761	0,82093287	0,81791455	0,82207068	0,82386628	0,82001228	0,8197653	0,82269454
kas	0,01790743	0,01701218	0,01751011	0,0173788	0,01780275	0,01705625	0,0174206	0,01748233	0,01688531	0,01733438
hipetrofi	-0,16257760	-0,16487836	-0,16417731	-0,16218456	-0,16061938	-0,16305935	-0,16331989	-0,16269256	-0,16270346	-0,16309975
intrikiller	0,11863795	0,11972745	0,11991183	0,11873572	0,11890658	0,11870751	0,11959183	0,11869594	0,11870916	0,11936761
orbital	-0,19163450	-0,19014470	-0,19135466	-0,19171546	-0,19075668	-0,19130789	-0,18986730	-0,19005410	-0,19044238	-0,18972196
hücre	-0,04518232	-0,04427538	-0,04538519	-0,04448273	-0,04485644	-0,04461002	-0,04386183	-0,04420693	-0,04323703	-0,04363487
kalmılık	-0,03727328	-0,03791399	-0,03796724	-0,03734652	-0,03730439	-0,03785134	-0,03790781	-0,03711769	-0,03768252	-0,03773483
karotid	0,02182374	0,02162395	0,02098382	0,02155465	0,02188073	0,02113588	0,02164548	0,02121577	0,02137923	0,02115111
kave	0,02143763	0,02106619	0,02152678	0,02168598	0,02188907	0,02098939	0,02142072	0,02132508	0,02148237	0,02132162
kavernoz	-0,04496593	-0,04503242	-0,04504143	-0,04466188	-0,04517839	-0,04515579	-0,04479707	-0,04500794	-0,04438989	-0,04466945
kitle	0,04016055	0,03970463	0,03983212	0,03982063	0,03936487	0,03979491	0,03931927	0,03957466	0,03907975	0,03954592
klastrum	0,02176253	0,02165835	0,02169552	0,02178409	0,02143133	0,021786	0,02169959	0,02166898	0,0214844	0,02141848
koklea	0,02152795	0,02142035	0,02089569	0,02169133	0,02181999	0,02114536	0,02161654	0,02084125	0,02164634	0,0211699
meckel	0,02117964	0,02140297	0,02093891	0,02111958	0,02172289	0,02171351	0,02157003	0,02110279	0,02149784	0,02148864
nazofarenks	-0,04735603	-0,04838487	-0,04860273	-0,04730577	-0,04660658	-0,04741718	-0,04803602	-0,04737237	-0,04707355	-0,04827511
okuler	0,02134058	0,02138092	0,02099378	0,02128764	0,02190259	0,02171387	0,02110599	0,02128443	0,02076595	0,02156138
pedikul	0,02157736	0,0216432	0,0217523	0,02165022	0,02130314	0,02183265	0,02142124	0,02123392	0,021292	0,02128753
serbest	-0,03474466	-0,03471531	-0,03478078	-0,03552917	-0,03546712	-0,03456001	-0,03448478	-0,03531756	-0,03490984	-0,03503864
serebellopo ntis	0,02103551	0,02093958	0,02100065	0,02143363	0,02102825	0,02097633	0,0213078	0,0211791	0,02131803	0,02106546
bulb	0,01907616	0,0191874	0,0188276	0,01917809	0,01920396	0,01875468	0,01874142	0,01876517	0,01882458	0,01854913

Tablo 4.18(devam): MRG özellik seti için giriş katmanından ara katmanına doğru ağırlık değerleri.

Parametre	Nöron "1"	Nöron 2	Nöron "3"	Nöron "4"	Nöron "5"	Nöron "6"	Nöron "7"	Nöron "8"	Nöron "9"	Nöron "10"
kuadri	0,0283671	0,02826012	0,02877187	0,02897264	0,02837632	0,02823859	0,02831225	0,02843629	0,02861537	0,0280821
nigra	0,02868775	0,02827433	0,02836439	0,0288686	0,02880548	0,02895908	0,02863887	0,02798089	0,02842774	0,02812071
poligon	0,02445593	0,02401919	0,02461149	0,02423635	0,02495149	0,02465314	0,0238907	0,02449683	0,02395128	0,02392254
prepontis	0,02883771	0,02828819	0,0287305	0,02883961	0,02862364	0,02851265	0,02837534	0,02866112	0,02853493	0,02844103
retrobulbera lan	0,01917903	0,01894838	0,01903617	0,01910133	0,01940692	0,01945111	0,01879925	0,01916767	0,01918386	0,0194197
suprasellar	0,02823774	0,02825092	0,02882745	0,02855381	0,02825381	0,02862791	0,02866917	0,02818249	0,02804449	0,02829145
supraserebr al	0,0282496	0,02866885	0,02854775	0,02828276	0,02879746	0,02851862	0,02832076	0,02858159	0,02841628	0,02848024
willis	0,02461009	0,02430451	0,02491971	0,02416893	0,02444763	0,02420172	0,02463984	0,02417127	0,02452618	0,02475128
yan	0,02886386	0,02795564	0,02887655	0,02823877	0,02813292	0,02890338	0,02818176	0,02844402	0,02855425	0,02877611
diploea	0,01424229	0,01436834	0,01464815	0,01436208	0,01477432	0,0142954	0,01493449	0,01402499	0,01428606	0,0143569
geminalam bien	0,01627058	0,01666052	0,0165196	0,01705571	0,01691461	0,01705472	0,01633145	0,01608621	0,01598816	0,01609294
klivus	0,01455132	0,0142281	0,01493023	0,01505444	0,01483456	0,014847	0,01414274	0,01426513	0,01492046	0,01406114
optikum	0,03204367	0,03156258	0,03171368	0,03189607	0,03136469	0,03228986	0,03216341	0,03205708	0,0314845	0,03148897
tabula	0,01419065	0,01477599	0,01437246	0,01463473	0,01513602	0,01430832	0,01431226	0,01415555	0,01418479	0,01450089
adet	-0,06209960	-0,06252806	-0,06221705	-0,06183293	-0,06248951	-0,06203012	-0,06198858	-0,06218175	-0,06229332	-0,06190686
epilepsi	0,31436723	0,31331518	0,31441294	0,31372101	0,31234818	0,31394109	0,3125958	0,31280032	0,31163806	0,31257483
daralma	-0,31324831	-0,31496478	-0,31498744	-0,31132536	-0,30919498	-0,31249066	-0,31415454	-0,31222562	-0,31197728	-0,31316081
paranazal	0,02378206	0,0232154	0,02356205	0,02342394	0,02198081	0,02317233	0,02306466	0,02245527	0,02297138	0,02245548
protokol	0,34323577	0,34280301	0,34405622	0,3432251	0,34228893	0,3437321	0,34215158	0,34225597	0,34129421	0,34205007
pasaj	-0,30630409	-0,30844277	-0,30827835	-0,30556245	-0,30293218	-0,30658133	-0,30754381	-0,30615104	-0,30531334	-0,30664724
retansiyon	0,23960279	0,23587243	0,23737522	0,24010641	0,2386888	0,23843567	0,23634689	0,23662681	0,23762807	0,23573144
aksiyel	0,15022033	0,14748947	0,14859043	0,14988163	0,14910799	0,14920806	0,14705863	0,14835978	0,14814544	0,14694615
ethmoid	-0,21359372	-0,21050031	-0,21219595	-0,21276196	-0,21200039	-0,21175421	-0,20973420	-0,21011617	-0,20949898	-0,20940130
hafif	0,09956647	0,09947475	0,09979622	0,10001376	0,09993139	0,0999509	0,09906822	0,09948385	0,09926013	0,09933806
aerasyon	-0,06150176	-0,06215684	-0,06174961	-0,06210790	-0,06279164	-0,06214974	-0,06271471	-0,06213676	-0,06224229	-0,06262224
artım	0,22536291	0,22552871	0,22477922	0,22568067	0,22526941	0,2256964	0,22514864	0,22515127	0,22468751	0,22442046
infarkt	-0,01679375	-0,01661446	-0,01710532	-0,01657956	-0,01663531	-0,01690423	-0,01652103	-0,01718232	-0,01691737	-0,01647224
kanama	-0,01676105	-0,01647222	-0,01655368	-0,01718190	-0,01721403	-0,01682342	-0,01719707	-0,01716338	-0,01672420	-0,01686597
vertebral	0,0691286	0,06998017	0,0697013	0,06929339	0,06824033	0,06910818	0,06996365	0,06926361	0,07025	0,06935709
segment	0,07665782	0,07778596	0,07703018	0,07601089	0,07550347	0,0767502	0,07754354	0,07694182	0,07715667	0,07717792
madde	-0,06127756	-0,06026362	-0,06047155	-0,06118297	-0,06095581	-0,06128320	-0,06066141	-0,06099083	-0,06098410	-0,06049290
nukleus	-0,03463466	-0,03553279	-0,03450384	-0,03426958	-0,03476943	-0,03455187	-0,03491570	-0,03521071	-0,03502028	-0,03517493
derece	0,12177046	0,1213259	0,12219607	0,1221501	0,12155766	0,12201366	0,12133838	0,1218708	0,12206415	0,12121491
kalınlaşma	-0,23871665	-0,23765434	-0,23860527	-0,23895123	-0,23809941	-0,23831942	-0,23762465	-0,23800306	-0,23760243	-0,23719588
konka	-0,24890573	-0,25078003	-0,25050115	-0,24775793	-0,24660266	-0,24918269	-0,24981598	-0,24871669	-0,24879065	-0,24874811
nazi	0,00234609	0,00080495	0,00098579	0,00252262	0,00232436	0,00222883	0,00100753	0,001992	0,00141806	0,00233897
taban	0,02826482	0,02353536	0,02475324	0,02815329	0,027681	0,02697443	0,0239576	0,02473638	0,02589449	0,02393028
tutulum	0,08538059	0,08603694	0,08585506	0,08530391	0,08554569	0,08571937	0,08532515	0,08495901	0,0848999	0,08601945
atrofi	0,00826972	0,00762963	0,00729708	0,0082438	0,00884597	0,00801322	0,00820151	0,00773724	0,00804109	0,00826269
internial	-0,23825636	-0,23669115	-0,23691077	-0,23879484	-0,23720874	-0,23815223	-0,23612558	-0,23679983	-0,23736417	-0,23579639
kaudat	-0,03156350	-0,03193066	-0,03230507	-0,03126559	-0,03177147	-0,03153813	-0,03243954	-0,03168190	-0,03135155	-0,03159248

Tablo 4.18(devam): MRG özellik seti için giriş katmanından ara katmanına doğru ağırlık değerleri.

Parametre	Nöron "1"	Nöron 2	Nöron "3"	Nöron "4"	Nöron "5"	Nöron "6"	Nöron "7"	Nöron "8"	Nöron "9"	Nöron "10"
hipoplazi	0,07876666	0,08043514	0,07988107	0,0788379	0,07841256	0,07896035	0,07963743	0,07880686	0,07966121	0,0792474
orta	0,04531968	0,04486972	0,04533485	0,04532026	0,04501569	0,0447442	0,04454178	0,0453076	0,04464033	0,04514495
skleroz	-0,03778163	-0,03927582	-0,03894548	-0,03774083	-0,03816515	-0,03871294	-0,03910109	-0,03870825	-0,03871208	-0,03939113
düzey	-0,07352619	-0,07360654	-0,07419761	-0,07350588	-0,07352249	-0,07380012	-0,07336928	-0,07317934	-0,07328158	-0,07344476
griak	-0,06880102	-0,06742900	-0,06834519	-0,06830219	-0,06852102	-0,06814501	-0,06763887	-0,06761260	-0,06850592	-0,06764195
sfenoid	-0,39182550	-0,38797253	-0,39076556	-0,39022503	-0,38927573	-0,39043364	-0,38785521	-0,38697945	-0,38685398	-0,38626628
mastoid	-0,30465996	-0,30566982	-0,30571260	-0,30392359	-0,30253153	-0,30432009	-0,30514668	-0,30374877	-0,30310027	-0,30475791
sekonder	0,04487027	0,04471361	0,04466649	0,04473127	0,04543064	0,04524941	0,04505423	0,04522335	0,04467913	0,04476316
beyin	0,09197969	0,09289127	0,09357244	0,09099475	0,0910311	0,09230154	0,09238059	0,09133131	0,0920171	0,09201984
globus	-0,02623013	-0,02659007	-0,02624158	-0,02628838	-0,02663111	-0,02604695	-0,02678217	-0,02685013	-0,02649092	-0,02695844
mastikatuar	0,1696942	0,16747215	0,16831119	0,1692766	0,16881943	0,16865236	0,167328	0,16740037	0,16699522	0,16725784
pallidus	-0,02594550	-0,02697889	-0,02676758	-0,02621283	-0,02644633	-0,02609122	-0,02646931	-0,02705527	-0,02691436	-0,02621870
putamen	-0,02663041	-0,02659520	-0,02649448	-0,02649174	-0,02674451	-0,02619687	-0,02632622	-0,02639654	-0,02692518	-0,02634105
yaygın	0,18995953	0,19015633	0,19002971	0,1904531	0,1905537	0,19087847	0,19031793	0,19064411	0,18995627	0,19038242
subkortikal	0,23003939	0,23068448	0,23065238	0,2294366	0,22990578	0,22925943	0,22973166	0,22916985	0,22966828	0,22949101
deviye	-0,09794268	-0,10133862	-0,10011931	-0,09765802	-0,09667693	-0,09838664	-0,09971019	-0,09797334	-0,09897285	-0,09976388
venoz	-0,06746418	-0,06775237	-0,06801800	-0,06705327	-0,06719101	-0,06712378	-0,06737710	-0,06760862	-0,06761014	-0,06730451
iskemi	0,02912418	0,02858318	0,0288704	0,02878443	0,02903581	0,02901778	0,02897545	0,02851241	0,02881503	0,0287664
seri	-0,06938104	-0,07035517	-0,07047413	-0,06855518	-0,06834421	-0,06924557	-0,06988242	-0,06875327	-0,06867392	-0,06980058
nonspesifik	0,21432688	0,21471532	0,21431647	0,21414969	0,21406012	0,21380268	0,21439879	0,21355408	0,21393693	0,2135316
fseir	0,28658954	0,28578566	0,28620313	0,2868549	0,28594763	0,28654096	0,2853395	0,285265	0,28544909	0,2849972
stibul	0,13219138	0,13151646	0,13270173	0,13182192	0,13137749	0,13232925	0,1314351	0,13120461	0,13160641	0,13175485
değişiklik	-0,14504236	-0,14570340	-0,14541537	-0,14434498	-0,14413701	-0,14460363	-0,14518643	-0,14465121	-0,14451610	-0,14449916
intensitenor	-0,14149300	-0,14268856	-0,14276570	-0,14224603	-0,14129424	-0,14188017	-0,14256406	-0,14179696	-0,14186759	-0,14244126
mal										
majornoz	0,14312155	0,14224453	0,1430777	0,14229568	0,14260628	0,14278555	0,14212293	0,14167007	0,14227689	0,14212926
multipl	0,1573941	0,15811393	0,15838048	0,15771629	0,15818034	0,15753613	0,15759758	0,15797968	0,1574032	0,15748509
sekel	-0,03797065	-0,03771386	-0,03776526	-0,03739777	-0,03678806	-0,03717523	-0,03753178	-0,03688486	-0,03651518	-0,03731382
major	-0,11505198	-0,11461277	-0,11437086	-0,11476526	-0,11421856	-0,11387707	-0,11425238	-0,11375613	-0,11416523	-0,11454514
vestibul	-0,11002787	-0,10978002	-0,10912812	-0,10940917	-0,10879544	-0,10947710	-0,10895207	-0,10820479	-0,10919291	-0,10940932
ekstraaksiy	-0,09453118	-0,09514113	-0,09560395	-0,09398071	-0,09370810	-0,09414423	-0,09517302	-0,09429799	-0,09407027	-0,09494925
el										
empty	-0,04093803	-0,04196824	-0,04208935	-0,04080942	-0,04044274	-0,04097948	-0,04151352	-0,04126226	-0,04129557	-0,04136004
girus	-0,02889954	-0,02917001	-0,02966106	-0,02858238	-0,02832880	-0,02884831	-0,02867796	-0,02900739	-0,02841263	-0,02874757
gliotik	0,21273349	0,21312427	0,21370296	0,21180769	0,20990592	0,21212208	0,21247898	0,21167907	0,21084509	0,2126141
kemik	0,04309522	0,04361404	0,0431565	0,04364218	0,04339615	0,04345829	0,04343282	0,04331684	0,04311869	0,04366739
kesim	0,05503668	0,05499859	0,0554543	0,05510595	0,05463345	0,05558004	0,0546069	0,0551144	0,05482639	0,05528522
saptanma	0,0371539	0,03765944	0,03734623	0,03705529	0,03747196	0,03785666	0,03788302	0,03776336	0,03766168	0,038437
gadodiamid	-0,06628369	-0,06633799	-0,06600452	-0,06685576	-0,06652378	-0,06720389	-0,06587222	-0,06637513	-0,06554165	-0,06534046
iskemik	-0,00872708	-0,00836746	-0,00852479	-0,00852255	-0,00865515	-0,00898587	-0,00860961	-0,00846303	-0,00905488	-0,00867365
sebat	0,00401268	0,00440512	0,00420989	0,00419166	0,00462436	0,00430834	0,00459678	0,00429963	0,00467981	0,00458501
seviye	-0,02126179	-0,02217360	-0,02203455	-0,02093689	-0,02096378	-0,02159991	-0,02139316	-0,02171811	-0,02130694	-0,02151349
hadise	0,02113591	0,02032782	0,02038288	0,02112843	0,02157867	0,02069566	0,02110633	0,02116629	0,02086739	0,02129969

Tablo 4.18(devam): MRG özellik seti için giriş katmanından ara katmanına doğru ağırlık değerleri.

Parametre	Nöron "1"	Nöron 2	Nöron "3"	Nöron "4"	Nöron "5"	Nöron "6"	Nöron "7"	Nöron "8"	Nöron "9"	Nöron "10"
havalanma	-0,04917337	-0,05068065	-0,04992134	-0,04963347	-0,04899085	-0,04923941	-0,05059811	-0,05059120	-0,04946260	-0,05083520
izlenmedi	0,27294151	0,27407156	0,27408456	0,27209448	0,2716026	0,27249209	0,27328891	0,27170842	0,27196044	0,27374391
mukozal	-0,43723964	-0,43451753	-0,43592251	-0,43670904	-0,43567736	-0,43677179	-0,43351023	-0,43434736	-0,43359381	-0,43280801
bos	0,01964496	0,01916822	0,01903564	0,0191824	0,01888297	0,01950953	0,01974213	0,01888954	0,0190529	0,01956564
cm	0,25864209	0,25931544	0,25899601	0,25860224	0,25817285	0,25870616	0,25886781	0,25855362	0,25792729	0,25839152
fse-ir	-0,33869241	-0,33753043	-0,33861142	-0,33858044	-0,33764916	-0,33881704	-0,33685168	-0,33701360	-0,33620255	-0,33610097
girmeyen	-0,06774509	-0,06863216	-0,06838180	-0,06821243	-0,06766421	-0,06828852	-0,06869857	-0,06753414	-0,06784462	-0,06873721
koronal	-0,05014677	-0,04925524	-0,04960599	-0,04965489	-0,05014522	-0,04956830	-0,04951470	-0,04953772	-0,04947684	-0,04947848
oksipital	0,12826799	0,12854743	0,12848046	0,12858534	0,12835057	0,12839283	0,12789954	0,12733213	0,12743601	0,12766561
hat	0,03779715	0,03767911	0,03737622	0,0383364	0,03833426	0,03753803	0,03776605	0,03807641	0,03841129	0,03822982
inferior	-0,11890253	-0,11957056	-0,11909985	-0,11852037	-0,11760321	-0,11882706	-0,11906824	-0,11831460	-0,11773650	-0,11920274
operasyon	0,00858192	0,00881044	0,00870794	0,00949115	0,0090196	0,00858891	0,00893455	0,00944468	0,00891578	0,00945763
parsiyel	-0,14936923	-0,14990094	-0,15011458	-0,14868518	-0,14785859	-0,14869049	-0,14942801	-0,14888717	-0,14855589	-0,14853663
selluler	-0,24826926	-0,24908066	-0,24946584	-0,24817114	-0,24702799	-0,24828657	-0,24843490	-0,24831127	-0,24795022	-0,24788626
svo	-0,04604450	-0,04621965	-0,04650310	-0,04626081	-0,04546238	-0,04612553	-0,04633010	-0,04593396	-0,04560174	-0,04628027
alt	0,10347724	0,10426831	0,10441029	0,10417336	0,10374108	0,10379784	0,10325715	0,10370645	0,10359894	0,10342337
eski	-0,05351982	-0,05429147	-0,05404009	-0,05348332	-0,05282087	-0,05327891	-0,05369096	-0,05286069	-0,05273085	-0,05301616
hipointens	-0,00491047	-0,00500668	-0,00535313	-0,00495234	-0,00473797	-0,00466261	-0,00496816	-0,00454719	-0,00465541	-0,00482553
peritrigonal	0,00093006	0,00081333	0,00065546	0,00089077	0,00204563	0,00157842	0,0009115	0,00161978	0,0017906	0,00079402
radiatada	-0,10675570	-0,10572250	-0,10680867	-0,10645071	-0,10680781	-0,10651612	-0,10524877	-0,10648420	-0,10529978	-0,10606509
izointens	0,0014291	0,00191895	0,00120017	0,00138872	0,00222365	0,00113564	0,00143745	0,00164977	0,00166876	0,00120278
malazik	0,04356442	0,04366526	0,04388561	0,04392278	0,04389191	0,04400673	0,04387685	0,04386864	0,0439301	0,04336768
mikrovasku ler	0,09834324	0,0976164	0,09801091	0,09852247	0,09761681	0,09831066	0,09802002	0,0976827	0,09763633	0,09712497
olgu	0,01219152	0,01131948	0,01201829	0,01145458	0,01159215	0,01129138	0,01187933	0,01149249	0,01209578	0,01211188
patolojik	-0,05723212	-0,05516817	-0,05578032	-0,05828545	-0,05840549	-0,05754940	-0,05589014	-0,05720529	-0,05685310	-0,05562432
temporal	0,09519977	0,09445733	0,09488221	0,09474365	0,09433456	0,09499872	0,09450987	0,09453751	0,09486693	0,0950603
araknoid	-0,05855248	-0,05875863	-0,05820922	-0,05806160	-0,05811344	-0,05809684	-0,05851990	-0,05842937	-0,05814035	-0,05843961
hasar	-0,08552944	-0,08630804	-0,08625137	-0,08502381	-0,08497478	-0,08538680	-0,08515393	-0,08479746	-0,08422827	-0,08557074
lomber	0,001756	0,0015177	0,00195257	0,00213129	0,0023082	0,00248963	0,00193373	0,00230649	0,0027394	0,00261903
swi	0,07051995	0,06991995	0,07027693	0,07018469	0,07048055	0,07033047	0,06981334	0,0703141	0,0701984	0,06990343
t2a	0,04420729	0,04498838	0,04486625	0,04468533	0,04455799	0,04533208	0,04525526	0,04490061	0,0443198	0,04501214
tarzda	0,1947197	0,19533716	0,19572486	0,19507592	0,19423047	0,19475539	0,19488938	0,19395043	0,19415363	0,19474576
tonsil	0,01713777	0,01710802	0,01658445	0,01718309	0,01761743	0,0172579	0,01719145	0,01730998	0,0169696	0,01697343
alan	0,29061349	0,29270763	0,29201046	0,28849769	0,28876582	0,28893253	0,2908489	0,2885668	0,28886751	0,29113406
anfissür	-0,11957892	-0,12015321	-0,12033392	-0,11861809	-0,11739032	-0,11907050	-0,11986530	-0,11896420	-0,11834377	-0,11903101
bulging	-0,00580599	-0,00638356	-0,00571744	-0,00538516	-0,00579114	-0,00610581	-0,00559332	-0,00558716	-0,00567715	-0,00598381
dejenere	-0,00603140	-0,00628514	-0,00601573	-0,00644882	-0,00552435	-0,00648005	-0,00634981	-0,00577285	-0,00650199	-0,00662632
anjyopati	0,10198917	0,10158952	0,1020029	0,10229458	0,10215668	0,10181705	0,10174303	0,10156824	0,10165162	0,10111703
anamlı	-0,00900485	-0,00950003	-0,00941904	-0,00829940	-0,00847290	-0,00852611	-0,00934186	-0,00857207	-0,00869307	-0,00919992
hipokampal	-0,07571089	-0,07664294	-0,07623862	-0,07622409	-0,07541800	-0,07556121	-0,07571021	-0,07550418	-0,07554823	-0,07540355
iskemigloz is	-0,20821639	-0,20966819	-0,20947457	-0,20785744	-0,20648816	-0,20849177	-0,20834800	-0,20772749	-0,20721177	-0,20798591

Tablo 4.18(devam): MRG özellik seti için giriş katmanından ara katmanına doğru ağırlık değerleri.

Parametre	Nöron "1"	Nöron 2	Nöron "3"	Nöron "4"	Nöron "5"	Nöron "6"	Nöron "7"	Nöron "8"	Nöron "9"	Nöron "10"
lakun	0,02934259	0,02864323	0,02839691	0,02910653	0,0291015	0,02878967	0,02899184	0,02848667	0,02859688	0,02875815
lokalizasyon	-0,20161407	-0,20166761	-0,20184552	-0,20126799	-0,20078565	-0,20111853	-0,20141165	-0,20090594	-0,20053232	-0,20079620
porensefali	-0,03722942	-0,03712135	-0,03792098	-0,03736539	-0,03628113	-0,03742312	-0,03657042	-0,03680110	-0,03646098	-0,03695559
spesifik	-0,34225096	-0,34166015	-0,34234336	-0,34196198	-0,34188909	-0,34184969	-0,34115102	-0,34061679	-0,34038723	-0,34028685
spinal	-0,01114459	-0,01144144	-0,01158519	-0,01139052	-0,01103387	-0,01124755	-0,01135693	-0,01099772	-0,01146700	-0,01110079
basion	0,03825933	0,03812176	0,0378918	0,03809989	0,03892525	0,03837008	0,03837909	0,03799105	0,03809206	0,0381217
defekt	-0,00848202	-0,00905208	-0,00891446	-0,00822565	-0,00780673	-0,00817152	-0,00828898	-0,00786462	-0,00795989	-0,00833362
lob	0,17876789	0,17881698	0,17865253	0,17872369	0,17910984	0,1783354	0,17933196	0,17857228	0,17767988	0,17878428
menengiom	0,13062119	0,13063395	0,13076267	0,13027856	0,13016202	0,13042502	0,1298654	0,12971984	0,1298922	0,1300747
opistion	0,03857406	0,03823172	0,03853133	0,03883565	0,03827052	0,03807767	0,03846883	0,03888659	0,03838771	0,03834798
santral	-0,00654740	-0,00719490	-0,00668095	-0,00716644	-0,00639556	-0,00716159	-0,00700222	-0,00645325	-0,00672765	-0,00695681
art	0,07765263	0,07756591	0,0771648	0,07732541	0,07728199	0,07765377	0,07724964	0,07682402	0,07679337	0,07727588
bas	-0,02643500	-0,02590166	-0,02661349	-0,02575713	-0,02559435	-0,02614580	-0,02590151	-0,02590067	-0,02606673	-0,02575438
duyarlı	0,1376287	0,13763259	0,13784794	0,13707965	0,13697956	0,13794966	0,13727482	0,13741558	0,13696041	0,13737803
eklem	-0,02408654	-0,02429536	-0,02454327	-0,02448783	-0,02365231	-0,02432683	-0,02430659	-0,02369730	-0,02429089	-0,02389119
faset	-0,02466400	-0,02418326	-0,02438167	-0,02419233	-0,02414675	-0,02427329	-0,02397432	-0,02444806	-0,02428522	-0,02390744
frontalis	0,09326194	0,09359268	0,09332661	0,09322231	0,0931403	0,09323107	0,09338467	0,09338006	0,0929384	0,0929869
gland	0,12773567	0,12815621	0,12866815	0,12818613	0,12803172	0,12829749	0,12838035	0,12786026	0,12775008	0,12799988
hiperostozis	0,09297402	0,09318569	0,09383302	0,09354648	0,09291488	0,09334791	0,09275536	0,09320602	0,09292244	0,0931281
kirli	-0,05008988	-0,04997921	-0,05026266	-0,04978403	-0,04911300	-0,04980389	-0,05012149	-0,04929139	-0,04934644	-0,05004281
kllinik	-0,03913307	-0,03890105	-0,03956023	-0,03857576	-0,03879855	-0,03919805	-0,03852428	-0,03830635	-0,03891025	-0,03851514
kortikosubkortikal	-0,03214978	-0,03249823	-0,03270676	-0,03216379	-0,03194044	-0,03277207	-0,03214263	-0,03265820	-0,03204505	-0,03235120
mikrohemoraji	0,01742275	0,01711122	0,01723644	0,01760161	0,01807329	0,01786291	0,01767859	0,01799477	0,01758573	0,01728359
parietal	0,06540368	0,06493742	0,06476787	0,06478182	0,06462243	0,06494112	0,06496512	0,06447124	0,06464467	0,06457103
pneal	0,01043248	0,00986046	0,00970799	0,00965295	0,01064169	0,00995693	0,01038175	0,00975288	0,01017294	0,00970059
superior	0,02325266	0,02300942	0,02345514	0,02324867	0,02381157	0,02343882	0,02335231	0,02368927	0,02362477	0,02317185
uzak	-0,04030041	-0,04067000	-0,04026860	-0,04013215	-0,03960402	-0,04014556	-0,04018103	-0,03999088	-0,03901180	-0,03960221
yararlı	0,02276165	0,02253205	0,02240173	0,02261906	0,0229078	0,02280704	0,02263943	0,02279923	0,02322835	0,02270199
çevresi	0,06908116	0,06952155	0,06916014	0,06898984	0,06952106	0,06885902	0,06966963	0,06887498	0,0692934	0,0694303
deviasyon	0,1736409	0,17508719	0,17519009	0,17200906	0,17202192	0,17275746	0,1748595	0,17266147	0,17302257	0,17460938
doku	-0,04028296	-0,04079928	-0,04104825	-0,04053384	-0,04043651	-0,04050859	-0,04041517	-0,03983562	-0,04037342	-0,04069688
dotarem	0,27389289	0,27425841	0,27390982	0,27369378	0,27259186	0,27420394	0,27301204	0,27275928	0,27269998	0,2735134
ektazik	-0,03812720	-0,03878279	-0,03827297	-0,03812808	-0,03732350	-0,03819319	-0,03815857	-0,03810538	-0,03717968	-0,03752482
foramende	0,02487033	0,02452838	0,02443666	0,02477898	0,0247764	0,02509293	0,02464092	0,02459561	0,02500107	0,02456074
miktar	0,26351293	0,25981825	0,26186639	0,26374208	0,26253849	0,26294842	0,25938614	0,26061226	0,26076528	0,25982363
minimal	-0,16489769	-0,16390107	-0,16477872	-0,16456936	-0,16399981	-0,16483981	-0,16358393	-0,16366641	-0,16287138	-0,16292814
nuks	-0,00626807	-0,00702542	-0,00723660	-0,00645062	-0,00619926	-0,00613392	-0,00638516	-0,00614337	-0,00659205	-0,00699697
parietal	0,05884917	0,05814959	0,05859442	0,05882334	0,05832627	0,05851303	0,05850269	0,05858585	0,05796562	0,057804
cilt	-0,05643869	-0,05700339	-0,05653316	-0,05609325	-0,05571421	-0,05624203	-0,05677172	-0,05643091	-0,05608809	-0,05657327
dolikoektazik	-0,10822453	-0,10802965	-0,10776512	-0,10778204	-0,10760263	-0,10804918	-0,10792980	-0,10753965	-0,10751710	-0,10788177
fizyon	0,04787673	0,04708123	0,04769177	0,04769082	0,0476607	0,04745378	0,0474037	0,04707023	0,04733951	0,04765108

Tablo 4.18(devam): MRG özellik seti için giriş katmanından ara katmanına doğru ağırlık değerleri.

Parametre	Nöron "1"	Nöron 2	Nöron "3"	Nöron "4"	Nöron "5"	Nöron "6"	Nöron "7"	Nöron "8"	Nöron "9"	Nöron "10"
horn	0,16355688	0,16464126	0,164461	0,16315238	0,1634883	0,16392348	0,16376318	0,16368962	0,16289111	0,16320289
kanama	0,04695482	0,04676905	0,04668757	0,04711291	0,0468609	0,04639987	0,0463581	0,04678354	0,04695507	0,04693688
komponent	0,13371905	0,13419908	0,13322068	0,13423203	0,13376315	0,13367598	0,13388299	0,13360607	0,13362549	0,133723
kortiko	-0,03563031	-0,03552051	-0,03527832	-0,03551340	-0,03508643	-0,03572380	-0,03578413	-0,03499554	-0,03491252	-0,03545164
medulla	-0,01180094	-0,01211720	-0,01244672	-0,01152515	-0,01176534	-0,01131288	-0,01146347	-0,01196263	-0,01151756	-0,01175932
mezial	0,12473206	0,12519295	0,12481646	0,12487998	0,12392263	0,12524244	0,12499303	0,12422656	0,12374717	0,12504856
siderotik	0,18723352	0,1875993	0,18772621	0,18780863	0,18778415	0,18800483	0,18864866	0,18740937	0,18741379	0,18828702
t2- flair	0,1388338	0,13891761	0,13904027	0,13983812	0,13982663	0,1393287	0,1394848	0,13822039	0,13822277	0,1384527
chiari	-0,05099040	-0,05115192	-0,05136450	-0,05061737	-0,05042559	-0,05078281	-0,05111318	-0,05093151	-0,05063082	-0,05103144
disk	-0,00022976	-0,00030983	-0,00049732	-0,00065646	-0,00058069	-0,00012283	-0,00022601	0,00001085	-0,00035715	-0,00046329
düzgün	-0,00508763	-0,00503404	-0,00528681	-0,00457164	-0,00473059	-0,00465588	-0,00454628	-0,00498394	-0,00492164	-0,00517351
homojen	-0,02522062	-0,02556012	-0,02542404	-0,02423001	-0,02446335	-0,02474446	-0,02447896	-0,02483779	-0,02451634	-0,02499178
kraniotomi	0,02013318	0,02003311	0,02063918	0,02002094	0,0207121	0,02038099	0,02033382	0,02063734	0,02007147	0,02048732
lordoz	-0,00105661	-0,00081003	-0,00052304	-0,00060783	-0,00040603	-0,00095764	-0,00037638	-0,00039230	-0,00001760	-0,00050308
ossez	-0,02936628	-0,02946487	-0,03015076	-0,02907169	-0,02891609	-0,02925023	-0,02919938	-0,02964770	-0,02876879	-0,02941409
				...						
				...						
t1a	0,20114797	0,20083716	0,20087924	0,2002626	0,20081498	0,20042695	0,20081918	0,20076515	0,19946155	0,20083318
Eğilim (bias)	0,4347483	0,43506879	0,43079673	0,42795655	0,43190926	0,4320361	0,42991789	0,42835155	0,43072435	0,43255569

MRG özellik seti için, ara katmandan (w) çıkış katmanına (y) ağırlık değerleri Tablo 4,19’da verilmiştir.

Tablo 4.19: MRG özellik seti için ara katmandan çıkış katmanına ağırlık değerleri.

–	Epilepsi	Sağlıklı
Nöron "1"	1,23569037	-1,23565697
Nöron "2"	1,23684674	-1,23700034
Nöron "3"	1,23844768	-1,23841137
Nöron "4"	1,23301565	-1,23303497
Nöron "5"	1,22887322	-1,22904033
Nöron "6"	1,23421756	-1,23447377
Nöron "7"	1,23293551	-1,23326391
Nöron "8"	1,22951881	-1,22980679
Nöron "9"	1,22823230	-1,22768142
Nöron "10"	1,23182469	-1,23111254
Eğilim (bias)	-3,40425707	3,40422299

MRG özellik setinde bulunan özelliklerin FFBP ağında katkısını Tablo 4.20’de gösterilmiştir.

Tablo 4.20: MRG özellik setinde bulunan özelliklerin katkısı.

MRG özellik seti	Hata oranı	MRG özellik seti	Hata oranı	MRG özellik seti	Hata oranı
patern	0,0574	beyin	0,0656	cm	0,0738
sirkülasyon	0,0574	globus	0,0656	fse-r	0,0410
void	0,0574	mastikatuvar	0,0574	girmeyen	0,0574
forme	0,0574	pallidus	0,0656	konka	0,0656
parasellar	0,0574	putamen	0,0656	koronal	0,0574
yerinde	0,0574	yaygın	0,0656	oksipital	0,0656
eksternal	0,0574	subkortikal	0,0738	hat	0,0656
kranioservikal	0,0574	deviye	0,0574	inferior	0,0574
büyük	0,0574	venoz	0,0492	operasyon	0,0574
kemik	0,0574	iskemi	0,0656	parsiyel	0,0574
mesafe	0,0656	seri	0,0574	selluler	0,0574
intrakranial	0,0574	nonspesifik	0,0738	svo	0,0574
kortikomedulla	0,0574	fseir	0,0656	alt	0,0656
en	0,0574	stibul	0,0656	bulunan	0,0574
kontrast	0,0574	değişiklik	0,0574	eski	0,0574
kraniofasial	0,0574	gri	0,0574	hipointens	0,0574
yer	0,0574	intensitenormal	0,0574	radiatada	0,0574
milimetrik	0,0574	kayıp	0,0656	izointens	0,0574
postkontrast	0,0574	majornoz	0,0656	komu	0,0574
periferik	0,0574	multipl	0,0656	malazik	0,0656
duralnoz	0,0738	sekel	0,0574	mikrovaskuler	0,0656
kaplayan	0,0656	major	0,0492	olgu	0,0574
kas	0,0574	vestibul	0,0492	patolojik	0,0574
hipertrofi	0,0574	ekstraaksiyel	0,0574	temporal	0,0656
intra	0,0656	empty	0,0574	araknoid	0,0574
magna	0,0574	girus	0,0574	hasar	0,0574
mevcut	0,0656	gliotik	0,0656	lomber	0,0574
intrikler	0,0574	kemik	0,0656	swi	0,0656
k/TD>	0,0574	kesim	0,0656	t2a	0,0656
orbital	0,0574	saptan	0,0656	tarzda	0,0656
hücre	0,0656	gadodiamid	0,0656	tonsil	0,0574
kalınlık	0,0656	iskemik	0,0574	alanına	0,0820
karotid	0,0574	olabilecek	0,0574	anfissür	0,0574
kave	0,0574	sebat	0,0574	bulging	0,0574
kaverno	0,0656	seviye	0,0574	dejenere	0,0574
kitle	0,0574	hadise	0,0656	tutan	0,0656
klastrum	0,0574	havalanma	0,0492	yaklaşık	0,0574
koklea	0,0574	izlenmedi	0,0574	etmektedir	0,0656
meckel	0,0574	mukozal	0,0656	anjiopati	0,0656
nazofarenks	0,0656	bos	0,0574	anlamli	0,0574
okuler	0,0574	cm	0,0738	girmektedir	0,0656
pedikul	0,0574	fse-r	0,0410	hipokampal	0,0574
serbest	0,0574	girmeyen	0,0574	iskemiglioizis	0,0574
serebellopontis	0,0574	nazi	0,0574	lakun	0,0656
bulb	0,0574	taban	0,0656	lokalizasyonu	0,0574
kuadri	0,0574	tutuluma	0,0656	porensifali	0,0574
nigra	0,0574	atrofi	0,0574	spesifik	0,0656
poligon	0,0574	internial	0,0574	spinal	0,0574
prepons	0,0574	kaudat	0,0656	basion	0,0656
retrobulberalan	0,0574	hipoplazi	0,0656	defekti	0,0574
substansia	0,0574	sfenoid	0,0656	lob	0,0656
suprasellar	0,0574	mastoid	0,0492	menengiom	0,0656
supraserebral	0,0574	sekonder	0,0656	opistion	0,0656
willis	0,0574	beyin	0,0656	santral	0,0574
yan	0,0574	globus	0,0656	art	0,0738
diploea	0,0574	mastikatuvar	0,0574	bas	0,0574
geminalambien	0,0574	pallidus	0,0656	duyarlı	0,0656
klivus	0,0574	putamen	0,0656	eklem	0,0574
optikum	0,0574	yaygın	0,0656	faset	0,0574
tabula	0,0574	subkortikal	0,0738	frontalis	0,0656
adet	0,0574	deviye	0,0574	gland	0,0656
epilepsi	0,0820	venoz	0,0492	hiperostozis	0,0656
daralm	0,0656	iskemi	0,0656	kirli	0,0574
paranazal	0,0656	seri	0,0574	klinik	0,0574
protokol	0,0820	nonspesifik	0,0738	kortikosubkortikal	0,0574
pasaj	0,0574	fseir	0,0656	mikrohemoraji	0,0574
retansiyon	0,0656	stibul	0,0656	paryetal	0,0574

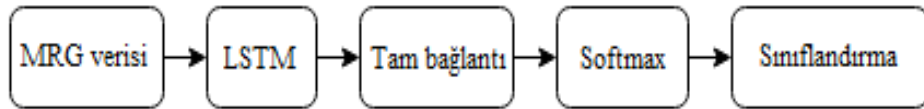
Tablo 4.20(devam): MRG özellik setinde bulunan özelliklerin katkısı.

MRG özellik seti	Hata oranı	MRG özellik seti	Hata oranı	MRG özellik seti	Hata oranı
aksiyel	0,0574	değişiklik	0,0574	hafif	0,0656
ethmoid	0,0656	gri	0,0574	pneal	0,0574
aerasyon	0,0574	intensitenormal	0,0574	superior	0,0574
artım	0,0656	kayıp	0,0656	uzak	0,0574
infarkt	0,0656	majornoz	0,0656	yararlı	0,0574
kanama	0,0656	multipl	0,0656	çevresi	0,0656
vertebral	0,0656	sekel	0,0574	deviasyon	0,0656
segment	0,0656	major	0,0492	doku	0,0574
madde	0,0574	vestibul	0,0492	dotarem	0,0574
nukleus	0,0656	ekstraaksiyel	0,0574	ektazik	0,0574
derece	0,0656	empty	0,0574	foramende	0,0574
kalınlaşma	0,0492	girus	0,0574	miktar	0,0656
nazi	0,0574	gliotik	0,0656	minimal	0,0574
taban	0,0656	kemik	0,0656	nuks	0,0574
tutululum	0,0656	kesim	0,0656	parietal	0,0656
atrofi	0,0574	peritrigonal	0,0574	cilt	0,0574
internal	0,0574	saptan	0,0656	dolikoektazik	0,0574
kaudat	0,0656	gadodiamid	0,0656	füzyon	0,0656
hipoplazi	0,0656	iskemik	0,0574	horn	0,0738
orta	0,0656	olabilecek	0,0574	kanama	0,0656
skleroz	0,0574	sebat	0,0574	komponent	0,0738
düzey	0,0574	seviye	0,0574	kortiko	0,0574
griak	0,0574	hadise	0,0656	medulla	0,0574
sfenoid	0,0656	havalanma	0,0492	mezial	0,0656
mastoid	0,0492	izlenmedi	0,0574	siderotik	0,0656
sekonder	0,0656	mukozal	0,0656	t2-flair	0,0738
		bos	0,0574	chiari	0,0574
				disk	0,0574
				düzgün	0,0574
				homojen	0,0574
				kraniotomi	0,0574
				lordoz	0,0574
				osseoz	0,0574
				t1a	0,0738
				...	...
				ak	0,0574

Tablo 4.20'ye göre, MRG özellik setindeki tüm özelliklerinin FFBP ağıının eğitimdeki katkısının hata oranı sıfıra yakın olması sebebiyle hastalığın tespitinde 502 özelliğin tamamının ayırt edici özellikler olduğu sonucuna varılmıştır.

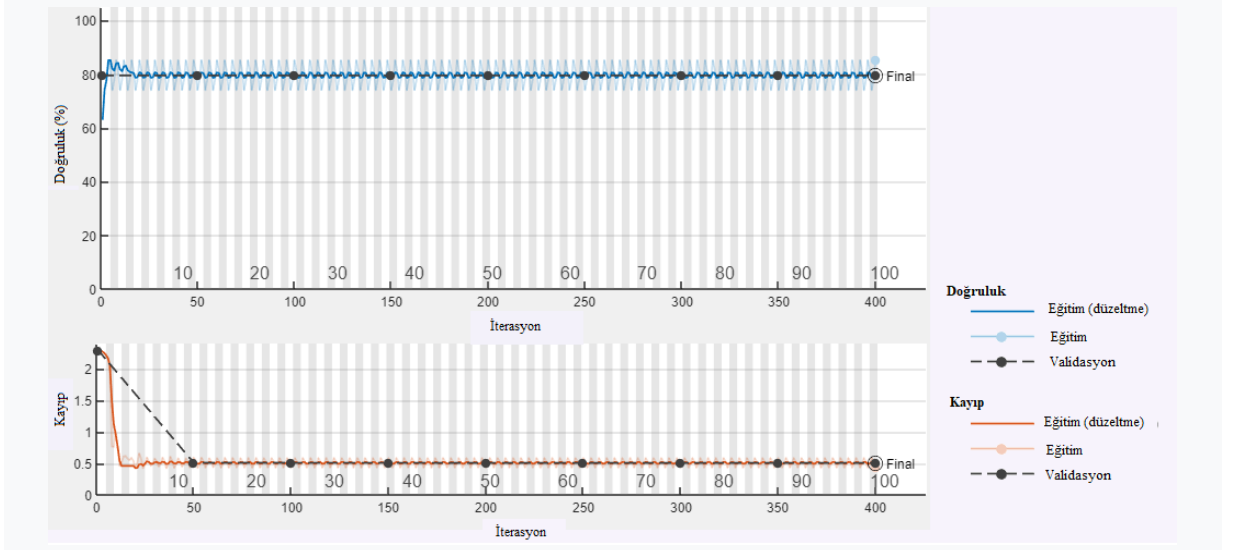
4.4.7. Model 4 için LSTM Yönteminin Sonuçları

LSTM derin öğrenme ağına uygulanan MRG özellik seti için uygulama adımları Şekil 4.44'te verilmiştir.



Şekil 4.44: LSTM ağına MRG özellik setinin uygulama adımları.

Şekil 4.44'e göre, MRG özellik seti LSTM ağına vektörel giriş olarak uygulanmıştır. Her bir giriş 1×502 olarak ağı girdi olarak uygulanmıştır. LSTM katmanı için gömme boyutu 100, gizli katman nöron sayısı 180, aktivasyon fonksiyonu olarak softmax fonksiyonu ve adam optimizasyon fonksiyonu olarak belirlenmiştir, Eğitim veri setini eşit şekilde bölerek eğitim yapabilmek için 27'lik mini gruplara bölünmüştür (mini_batch_size = 27). Maksimum epok sayısı 100, gradyan eşikini 1, öğrenme katsayısı başlangıç değeri 0,001 olarak belirlenen eğitim işleminden elde edilen ağı yapısı Şekil 4.45'te gösterilmiştir ve eğitim işleminden elde edilen sınıflandırma doğruluğu %83,30 olarak elde edilmiştir.



Şekil 4.45: LSTM ağına MRG özellik setinin uygulanması.

LSTM yöntemine uygulanan MRG özellik setinin sınıflandırılmasında test veri setinin doğruluğu Şekil 4.46'da karışıklık matrisi ile verilmiştir ve test veri setinin sınıflandırma doğruluğunun %83,33 olduğu görülmektedir.

	Epilepsi	Sağlıklı
Epilepsi	10	0
Sağlıklı	2	0

Şekil 4.46: LSTM yöntemine uygulanan MRG özellik seti için karışıklık matrisi.

Şekil 4.46'ya göre, LSTM yöntemine MRG özellik setinin uygulanarak sınıflandırılmasında test veri setindeki epilepsi hastası bireylere ait veri doğru sınıflandırılmıştır. Fakat kontrol grubu sağlıklı bireylere ait veri yanlış sınıflandırılmıştır.

4.5. ÇOK MODLU KARAR TABANLI VERİ FÜZYONUNUN SONUÇLARI (MODEL 5)

Model 5; Model 1, Model 2, Model 3 ve Model 4'ten elde edilen maksimum sınıflandırma doğruluk oranını hesaplayan yöntemlerin ortalama değeridir. Model 5, bu tez çalışmasında önerilen çok modlu karar tabanlı veri füzyonu ile modelleme başarımının sonucudur.

4.5.1. Model 1 için Analiz Sonuçları

Model 1, klinik parametre özellik seti için ID3, lineer regresyon, Naive Bayes, DVM, k-NN, FFBP yöntemleri ile modellenerek oluşturulmuştur. Eğitim işlemi için modellemelerden elde edilen sonuçlar Tablo 4.21'de verilmiştir.

Tablo 4.21: Model 1 (Klinik parametre özellik seti için uygulanan) modelleme sonuçları.

Yöntem	Doğruluk oranı (%)	Tahminleme hızı (obs/sn)	Eğitme zamanı (sn)
ID3	100	2300	1,34
Doğrusal regresyon	98,40	640	12,04
Naive Bayes	79,50	190	12,75
DVM	97,50	2600	1,09
k-NN	96,70	1100	2,51
FFBP	98,40	1200	0,01

Tablo 4.21'e göre, klinik parametre özellik seti için en başarılı modelleme %100 sınıflandırma doğruluğu ile ID3 yöntemi olmuştur. Klinik parametre özellik seti için en başarısız modelleme Naive Bayes yöntemi ile %79,50 sınıflandırma doğruluğu olarak elde edilmiştir. Klinik parametre özellik seti için yöntemler arası tahminleme hızı en başarılı yöntem Naive Bayes yöntemi olmuştur. ID3 yöntemi 2300 obs/sn ile en uzun tahminleme hızına sahiptir. Yöntemlerin eğitme zamanı (sn) küçükten büyüğe doğru FFBP (0,01 sn) < DVM (1,09 sn) < ID3 (1,34 sn) < k-NN (2,51 sn) < doğrusal regresyon (12,04 sn) < Naive Bayes (12,75 sn) olarak sıralanmaktadır.

Model 1 (Klinik parametre özellik seti için modelleme) test veri seti için tahminleme başarımı; doğruluk, duyarlılık, özgüllük, AUC, kesinlik, F1 ölçüsü ve kappa katsayısı için hesaplanmıştır ve hesaplanan istatistiksel analiz sonuçları Tablo 4.22'de verilmiştir.

Tablo 4.22: Model 1 (Klinik parametre özellik seti modelleme) istatistiksel analizi.

Yöntem	Doğruluk (%)	Duyarlılık (%)	Özgüllük (%)	AUC	Kesinlik (%)	F1	Kappa
ID3	100	100	100	1	100	1	1
Doğrusal regresyon	100	100	100	0,99	100	1	1
Naive Bayes	83,33	83,33	-	0,85	100	0,789	0
DVM	83,33	83,33	-	0,85	100	0,789	0
k-NN	66,66	96	95,45	0,99	88,89	1,33E-11	-0.1429
FFBP	100	100	100	1	100	1	1

Tablo 4.22'ye göre Model 1 için en başarılı tahminleme ID3, doğrusal regresyon ve FFBP yöntemleri ile elde edilmiştir. Bu yöntemlerin epilepsi ve sağlıklı sınıfları için doğruluk, duyarlılık, özgüllük ve kesinlik değerleri %100 olarak elde edilmiştir. Ayrıca bu üç yöntemin tahminleme başarımı %100 olarak hesaplanmıştır. F1 ölçüsü bu yöntemler için 1 olarak elde edilmiştir ve bu değer mükemmel hassasiyet ve geri çağırma değerinin elde edildiğini göstermektedir. Klinik parametre özellik seti için epilepsi sınıf ve sağlıklı sınıf arasındaki karşılaştırmalı uyuşmanın güvenilirliği ID3, doğrusal regresyon ve FFBP yöntemleri için Kappa katsayısı, 1 olarak elde edildiği için mükemmel uyumlu olduğu ispatlanmıştır.

4.5.2. Model 2 için Analiz Sonuçları

Model 2, laboratuvar bulguları özellik seti için ID3, lineer regresyon, Naive Bayes, DVM, k-NN, FFBP yöntemler ile modellenerek oluşturulmuştur. Eğitim işlemi için modellemelerden elde edilen sonuçlar Tablo 4.23'te verilmiştir.

Tablo 4.23: Model 2 (Laboratuvar bulguları özellik seti modelleme) sonuçları.

Yöntem	Doğruluk oranı (%)	Tahminleme hızı (obs/sn)	Eğitime zamanı (sn)
ID3	98,40	2800	0,93
Doğrusal regresyon	95,10	1100	1,73
Naive Bayes	95,90	510	4,23
DVM	93,40	1900	50,36
k-NN	93,40	1100	67,14
FFBP	100	800	1,00

Tablo 4.23'e göre, laboratuvar bulguları özellik seti için en başarılı modelleme %100 sınıflandırma doğruluğu ile FFBP yöntemi olmuştur. Laboratuvar bulguları özellik seti için en başarısız modellemeler DVM ve k-NN yöntemlerinden %93,40 sınıflandırma doğrulukları ile elde edilmiştir. Laboratuvar bulguları özellik seti için yöntemler arası tahminleme hızı en başarılı yöntem Naive Bayes yöntemi olmuştur. ID3 yöntemi 2800 obs/sn ile en uzun tahminleme hızına sahiptir. Yöntemlerin eğitime zamanı (sn) küçükten büyüğe doğru ID3 (0,93 sn) < FFBP (1 sn) < doğrusal regresyon (1,73 sn) < Naive Bayes (4,23 sn) < DVM (50,36 sn) < k-NN (67,14 sn) olarak sıralanmaktadır.

Model 2 (Laboratuvar bulguları özellik seti için modelleme) test veri seti için tahminleme başarımı; doğruluk, duyarlılık, özgüllük, AUC, kesinlik, F1 ölçüsü ve kappa katsayısı için hesaplanmıştır ve hesaplanan istatistiksel analiz sonuçları Tablo 4.24'te verilmiştir.

Tablo 4.24: Model 2 (Laboratuvar bulguları özellik seti modelleme) istatistiksel analizi.

Yöntem	Doğruluk(%)	Duyarlılık (%)	Özgüllük (%)	AUC	Kesinlik (%)	F1	Kappa
ID3	100	100	100	1	100	1	1
Doğrusal regresyon	100	100	100	1	100	1	1
Naive Bayes	83,33	83,33	-	0,85	100	0,789	0
DVM	100	100	100	1	100	1	1
k-NN	100	100	100	1	100	1	1
FFBP	100	100	100	1	100	1	1

Tablo 4.24'e göre Model 2 için en başarılı tahminleme ID3, doğrusal regresyon, DVM, k-NN ve FFBP yöntemleri ile elde edilmiştir. Bu yöntemlerin epilepsi ve sağlıklı sınıfları için doğruluk, duyarlılık, özgüllük ve kesinlik değerleri %100 olarak elde edilmiştir. Ayrıca bu beş yöntemin tahminleme başarımı %100 olarak hesaplanmıştır. F1 ölçüsü bu yöntemler için 1 olarak elde edilmiştir ve bu değer mükemmel hassasiyet ve geri çağırma değerinin elde edildiğini göstermektedir. Laboratuvar bulguları özellik seti için epilepsi sınıf ve sağlıklı sınıf arasındaki karşılaştırmalı uyuşmanın güvenilirliği ID3, doğrusal regresyon, DVM, k-NN ve FFBP yöntemleri için Kappa katsayısı, 1 olarak elde edildiği için mükemmel uyumlu olduğu ispatlanmıştır.

4.5.3. Model 3 için Analiz Sonuçları

Model 3, EEG özellik seti için ID3, doğrusal regresyon, Naive Bayes, DVM, k-NN, FFBP yöntemleri ile modellenerek oluşturulmuştur. Eğitim işlemi için modellemelerden elde edilen sonuçlar Tablo 4.25’te verilmiştir.

Tablo 4.25: Model 3 (EEG özellik seti modelleme) sonuçları.

Yöntem	Doğruluk oranı (%)	Tahminleme hızı (obs/sn)	Eğitime zamanı (sn)
ID3	80,30	3200	0,85
Doğrusal regresyon	71,30	980	1,70
Naive Bayes	70,50	770	2,67
DVM	75,40	3300	6,16
k-NN	78,70	2000	1,19
FFBP	79,50	2000	3,92

Tablo 4.25’e göre, EEG özellik seti için en başarılı modelleme ID3 yönteminden %80,30 sınıflandırma doğruluğu ile elde edilmiştir. EEG özellik seti için en başarısız modelleme Naive Bayes yönteminden %70,50 sınıflandırma doğruluğu ile elde edilmiştir. EEG özellik seti için algortimlar arası tahminleme hızı en başarılı yöntem Naive Bayes yöntemi olmuştur. DVM yöntemi 3300 obs/sn ile en uzun tahminleme hızına sahiptir. Yöntemlerin eğitime zamanı (sn) küçükten büyüğe doğru ID3 (0,85 sn) < k-NN (1,19 sn) < doğrusal regresyon (1,70 sn) < Naive Bayes (2,67 sn) < FFBP (3,92 sn) < DVM (6,16 sn) olarak sıralanmaktadır. Model 3 (EEG özellik seti için modelleme) test veri seti için tahminleme başarımı; doğruluk, duyarlılık, özgüllük, AUC, kesinlik, F1 ölçüsü ve kappa katsayısı için hesaplanmıştır ve hesaplanan istatistiksel analiz sonuçları Tablo 4.26’da verilmiştir.

Tablo 4.26: Model 3 (EEG özellik seti modelleme) istatistiksel sonuçları.

Yöntem	Doğruluk (%)	Duyarlılık (%)	Özgüllük (%)	AUC	Kesinlik (%)	F1	Kappa
ID3	75	100	40	0,77	70	0,695	0,437
Doğrusal regresyon	91,66	100	66,66	0,91	100	0,905	0,75
Naive Bayes	83,33	83,33	-	0,8	100	0,789	0
DVM	83,33	83,33	-	0,79	100	0,789	0
k-NN	83,33	83,33	-	0,82	100	0,789	0
FFBP	83,33	100	0	0,85	83,33	1,33E-15	0

Tablo 4.26'ya göre Model 3 için en başarılı tahminleme doğrusal regresyon yöntemi ile elde edilmiştir. Bu yöntemin epilepsi ve sağlıklı sınıfları için doğruluk oranı %91,66, duyarlılık oranı %100, özgüllük oranı %66,66 ve kesinlik değerleri %100 olarak elde edilmiştir. F1 ölçüsü bu yöntem için 0,905 olarak elde edilmiştir ve bu değer mükemmel hassasiyet ve geri çağırma değeri elde edildiğini göstermektedir. EEG özellik seti için epilepsi sınıf ve sağlıklı sınıf arasındaki karşılaştırmalı uyuşmanın güvenilirliği doğrusal regresyon yöntemi için Kappa katsayısı, 0,75 olarak elde edildiği için uyumlu olduğu ispatlanmıştır.

4.5.4. Model 4 için Analiz Sonuçları

Model 4, MRG özellik seti için ID3, doğrusal regresyon, Naive Bayes, DVM, k-NN, FFBBP ve LSTM yöntemleri ile modellenerek oluşturulmuştur. Eğitim işlemi için modellemelerden elde edilen sonuçlar Tablo 4.27'de verilmiştir.

Tablo 4.27: Model 4 (MRG özellik seti için uygulanan modelleme) sonuçları.

Yöntem	Doğruluk oranı (%)	Tahminleme hızı (obs/sn)	Eğitime zamanı (sn)
ID3	82,00	360	3,229
Doğrusal regresyon	54,90	130	18,447
Naive Bayes	79,50	29	75,015
DVM	77,90	310	21,84
k-NN	75,40	360	38,89
FFBBP	79,50	1200	1
LSTM	83,30	1500	28:53:00

Tablo 4.27'ye göre, MRG özellik seti için en başarılı modelleme LSTM yönteminden %83,30 sınıflandırma doğruluğu ile elde edilmiştir. MRG özellik seti için en başarısız modellemeler doğrusal regresyon yönteminden %54,90 sınıflandırma doğruluğu elde edilmiştir. MRG özellik seti için yöntemler arası tahminleme hızı en başarılı yöntem Naive Bayes yöntemi olmuştur. LSTM yöntemi 1500 obs/sn ile en uzun tahminleme hızına sahiptir. Yöntemlerin eğitime zamanı (sn) küçükten büyüğe doğru FFBBP (1 sn) < ID3 (3,229 sn) < doğrusal regresyon (18,447 sn) < DVM (21,84 sn) < k-NN (38,89 sn) < Naive Bayes (75,015) < LSTM (1733 sn) olarak sıralanmaktadır.

Model 4 (MRG özellik seti için modelleme) test veri seti için tahminleme başarımı; doğruluk, duyarlılık, özgüllük, AUC, kesinlik, F1 ölçüsü ve kappa katsayısı için hesaplanmıştır ve hesaplanan istatistiksel analiz sonuçları Tablo 4.28’de verilmiştir.

Tablo 4.28: Model 4 (MRG özellik seti modelleme) istatistiksel sonuçları.

Yöntem	Doğruluk (%)	Duyarlılık (%)	Özgüllük (%)	AUC	Kesinlik (%)	F1	Kappa
ID3	75,00	100	40	0,54	70	0,695	0,437
Doğrusal regresyon	41,66	80,00	14,2	0,53	40	0,293	-0,05
Naive Bayes	83,33	83,33	-	0,5	100	0,789	0
DVM	66,66	87,50	25	0,62	70	0,538	0,142
k-NN	66,66	100	20	0,62	63,6	0,5	0,225
FFBP	83,33	100	0	0,83	83,33	1,33E-15	0
LSTM	83,33	83,33	0	0,9	100	0,789	0

Tablo 4.28’e göre Model 4 için en başarılı tahminleme Naive Bayes, FFBP ve LSTM yöntemleri ile elde edilmiştir. Bu yöntemlerin epilepsi ve sağlıklı sınıfları için doğruluk oranları %83,33 olarak hesaplanmıştır. Naive Bayes ve LSTM yöntemlerinin duyarlılık oranları %83,33 ve FFBP yönteminin duyarlılık oranı %100 olarak hesaplanmıştır. Kesinlik değerleri Naive Bayes ve LSTM yöntemleri için %100 olarak elde edilmiştir. FFBP yönteminin kesinlik değeri 83,33 olarak elde edilmiştir. Ayrıca bu üç yöntemin tahminleme başarımı 0,5’in üzerinde hesaplanmıştır. F1 ölçüsü bu yöntemler için 1 değerine yakın değerler olarak elde edilmiştir ve bu değerler mükemmel hassasiyet ve geri çağırma değerinin elde edildiğini göstermektedir. MRG özellik seti için epilepsi sınıf ve sağlıklı sınıf arasındaki karşılaştırmalı uyuşmanın güvenilirliği ID3, doğrusal regresyon ve FFBP yöntemleri için Kappa katsayısı, 0 olarak elde edildiği için tesadüfen uyumlu olduğu sonucuna varılmıştır.

4.5.5. Model 5 için Analiz Sonuçları

Klinik parametre, laboratuvar bulguları, EEG, MRG özellik setlerinden oluşan epilepsi veri seti için elde edilen modelleme sonuçları Tablo 4.29’da verilmiştir. Her özellik seti için oluşturulan model için maksimum sınıflandırma doğruluğuna sahip yöntem seçilmiştir. Son olarak maksimum sınıflandırma doğruluklarının ortalaması hesaplanmıştır. Hesaplanan ortalama değer, bu tez çalışmasında önerilen çok modlu karar tabanlı veri füzyonu modellemesinin başarısıdır. Yöntemlere uygulanan özellik setlerinin epilepsi ve sağlıklı sınıflandırma doğruluk oranları Tablo 4.29’da verilmiştir.

Tablo 4.29: Özellik setleri için tahminleme sonuçları.

Özellik seti	Yöntem	Doğruluk oranı (%)
Klinik Parametre	ID3	100
	Doğrusal regresyon	100
	Naive Bayes	83,33
	DVM	83,33
	k-NN	66,66
	FFBP	100
Laboratuvar Bulguları	ID3	100
	Doğrusal regresyon	100
	Naive Bayes	83,33
	DVM	100
	k-NN	100
	FFBP	100
EEG	ID3	75,00
	Doğrusal regresyon	91,66
	Naive Bayes	83,33
	DVM	83,33
	k-NN	83,33
	FFBP	83,33
MRG	ID3	75,00
	Doğrusal regresyon	41,66
	Naive Bayes	83,33
	DVM	66,66
	k-NN	66,66
	FFBP	83,33
	LSTM	83,33

Bu tez çalışmasında önerilen çok modlu karar tabanlı veri füzyonu tekniği ile elde edilen sınıflandırma doğruluk oranı %93,75 elde edilerek Tablo 4.30’da gösterilmiştir.

Tablo 4.30: Çok modlu karar tabanlı veri füzyonu tekniği ile elde edilen tahminleme performansı.

Özellik seti	Özellik Çıkarımı	Yöntemler	Tahminleme performansı
Klinik parametre & Laboratuvar bulguları & EEG & MRG	BBA, ZipF	ID3, doğrusal regresyon, Naive Bayes, DVM, k-NN, FFBP, LSTM, 10-kat çaprazlama	Klinik parametre doğruluk oranı: %100 Laboratuvar bulguları doğruluk oranı: %100 EEG doğruluk oranı: %91,66 MRG doğruluk oranı: %83,33 Füzyon doğruluk oranı: %93,75

4.6. EPİLEPSİ VERİ SETİNE AİT DENEYSEL SONUÇLAR

Gerçek epilepsi veri seti (122×603) boyutlu bir matristir. Bu tez çalışmasında, farklı özellik setlerinin modellenerek kaynaştırılması ile elde edilen sınıflandırma başarımı, tekil modelleme ile elde edilen sınıflandırma başarımının bir karşılaştırılmıştır. Literatürdeki çalışmalarla doğru orantılı olarak, bu tez çalışmasında da epilepsi veri setindeki 4 farklı özellik setinin farklı yöntemlerle modellenmesiyle başarılı sınıflandırma doğruluğu elde edilmiştir. Bunun ispatının yapılabilmesi için epilepsi veri seti için tekil modelleme ID3, doğrusal regresyon, Naive Bayes, DVM, k-NN, FFBP yöntemleri ile uygulanmıştır. Tekil modelleme ile sınıflandırma sonuçları Tablo 4.31’de verilmiştir.

Tablo 4.31: Epilepsi veri seti için tekil sınıflandırma sonuçları.

Yöntem	Doğruluk oranı (%)	Tahminleme hızı (obs/sn)	Eğitime zamanı (sn)
ID3	86,70	300	2,89
Doğrusal regresyon	85,20	320	2,92
Naive Bayes	59,80	270	3,05
DVM	85,20	400	5,46
k-NN	84,40	360	2,45
FFBP	84,10	1	1

Tekil modelleme ile sınıflandırma sonuçları Tablo 4.32’de verilmiştir.

Tablo 4.32: Epilepsi veri seti için tekil sınıflandırma tahminleme sonuçları.

Yöntem	Doğruluk oranı (%)	Duyarlılık (%)	Özgüllük (%)	AUC	Kappa
ID3	83,33	96,03	100	0,92	0,84
Doğrusal regresyon	41,66	85,58	81,81	0,92	0,42
Naive Bayes	58,33	98,00	33,33	0,85	0,27
DVM	83,33	85,84	100	0,74	0,01
k-NN	75,00	84,21	87,50	0,59	0,56
FFBP	83,33	86,53	88,88	0,85	0,59

Bu tez çalışmasında önerilen çok modlu karar tabanlı veri füzyonu tekniği ile epilepsi sınıfı ve sağlıklı sınıfı %93,75 sınıflandırma doğruluğuna sahiptir. Tekil modelleme tekniği ile sınıflandırılan epilepsi veri seti için ID3, DVM,ve FFBP yöntemlerinden sınıflandırmada en başarılı tahminleme doğruluğunu %83,33 olarak hesaplamıştır. Bu tezde önerilen çok modlu

karar tabanlı veri füzyonu tekniğı tekil modelleme tekniğınden %10,42 oranında daha başarılı olmuştur. |



5. TARTIŞMA VE SONUÇ

Epilepsi hastalığı sebebi tam olarak açıklanamamış nörolojik bir rahatsızlıktır. Ülkemizde 700 bin civarında epilepsi hastası olduğu bilinmektedir. Epilepsi hastası her bireyin klinik laboratuvar bulguları, epileptik nöbeti ve beyin yapısındaki bozuklukları kendisine özgüdür. Bu durum bazı hastalar için epilepsi tanısının konulmasını güçleştirmektedir. Bazı bireylerde başka bir bozukluğun yol açtığı belirtiler yanlışlıkla epilepsi tanısı olarak tanımlanabilmektedir. Bu yüzden, epilepsi hastalığının tanımlanmasında mühendislik ve tıp bilimleri arasında etkileşimli çalışmalar son yıllarda uygulama alanı bulmuştur. Bunlardan biri olan epilepsi hastalığının tanısı sürecinde doktora yardımcı makine öğrenmesi yöntemleri ile geliştirilmiş çalışmaların artış göstermesi iki bilim arasındaki çalışmaların ilgi görmesine neden olmuştur. Özellik mühendisliği yöntemleri kullanılarak geliştirilen matematiksel yöntemler yaklaşımları ile epilepsi hastalığının sınıflandırılması klinik yöntemlere güçlü bir alternatif oluşturmuş ve literatürde daha iyi sonuçların elde edildiği yeni çalışmalara ihtiyaç duyulmuştur. Bugüne kadar yapılmış çalışmalarda epilepsi hastalığının tanımlanmasında hastaya ait sadece EEG verisi ile önemli diğer verilerin; demografik özellikler, nörolojik muayene sırasında doktor epikriz rapor verisinden elde edilen klinik bulgular, laboratuvar bulgularından elde edilen kan değerleri, MRG radyoloji raporlarından elde edilen verilerin bir arada kullanılarak yapıldığı kapsamlı bir çalışmaya rastlanılmamıştır.

Bu tez çalışmasında, epilepsi hastalığının tanımlanmasında 18 yaş üstü erişkin bireylere ait demografik verilerinden hasta ile ilgili özellikler çıkarılmış, doktor epikriz rapor verisinden hasta şikayetleri çıkarılmış, laboratuvar bulgularından kan değerleri elde edilmiş, EEG sinyal verisine ait anlamlı özellikler çıkarılmış, MRG rapor verisine ait anlamlı özellikler çıkarılmıştır. Böylece, epilepsi hastalığının ayırıcı tanısında yararlı olabilecek tüm parametreler bir arada kullanılarak hastalıkla ilgili anlamlı özellikler çıkarılarak, hastalığa özgü ülkemizdeki epilepsi hastalarına ait emsal niteliğinde gerçek bir veri seti oluşturulmuştur. Epilepsi kişiye özgü bir hastalık olduğundan her bir bireyin verisi bireye özgü değerlendirilerek epilepsi veri seti oluşturulmuştur. Ayrıca, hastalıkla ilgili anlamlı özelliklerin çıkarılması sayesinde hastalığın erken dönemde teşhisi için koruyucu tedaviye yansiyacak adımların atılması ile doktora yardımcı bir sistem oluşturulması hedeflenmiştir. Oluşturulan veri tabanındaki veriler kullanılarak epilepsi ve sağlıklı bireylere ait verilerin etkin olarak sınıflandırılması için çok modlu karar tabanlı veri füzyonu tekniği önerilmiştir.

Bu tez çalışmasında önerilen çok modlu karar tabanlı veri füzyonu tekniği öncelikle modeller arasında sınıflandırma doğruluğu, ROC analizi, eğitime zamanı, modelleme hızı ve istatistiksel özellikler için hesaplanmış değerlere göre karşılaştırılmıştır. Sonrasında, çok modlu karar tabanlı veri füzyonu tekniği ile tekil sınıflandırma tekniği arasındaki karşılaştırma sınıflandırma doğruluğu, ROC analizi, eğitime zamanı, modelleme hızı ve istatistiksel özellikler için hesaplanmış değerlere göre yapılmıştır. Son olarak, hastalığın erken dönemde teşhis edilebilmesi için ve doktora yardımcı bir sistem olabilmesi için epilepsi veri setinden çıkarılan ayırt edici özellikler verilmiştir.

Çok modlu karar tabanlı veri füzyonu modelleme başarımı ile %93,75 oranında sınıflandırma doğruluğu elde edilmiştir. Tekil modelleme tekniği ile sınıflandırılan epilepsi veri seti için ID3, DVM, ve FFBP yöntemlerinden sınıflandırmada en başarılı tahminleme doğruluğunu %83,33 olarak hesaplamıştır. Bu tezde önerilen çok modlu karar tabanlı veri füzyonu tekniği tekil modelleme tekniğinden %10,42 oranında daha başarılı olmuştur.

Çok modlu karar tabanlı veri füzyonu modelleme ile hastalıkla ilgili anlamlı özellikler çıkarılmıştır. Klinik parametre özellik setinin doğrusal regresyon analizinde, jeneralize tonik–klonik nöbet (senede), unutkanlık, geç algılama, donup kalma, ateş, sol kolda kasılma parametrelerinin p–değerlerinin 0,05 değerinin altında olduğundan epilepsi ve kontrol grubu sağlıklı bireylerin sınıflandırılmasında anlamlı özellikler olduğu sonucu çıkarılmıştır. FFBP yöntemine uygulanan klinik parametre özellik setindeki; yaş, şuur kaybı, uyku bozukluğu, baş ağrısı, el kollarda uyuşma, göz kararması, halsizlik, bayılma hissi, bayılma, nefes daralması, uyku hali, ateş, sağ el uyuşması, el ayakta atma, tüm vücutta titreme, sağ yüzde uyuşma, sol kulak ağrısı, ellerde uyuşma, sağ tarafta güçsüzlük, sol bacakta kasılma–kramp, sol kulakta işitme kaybı, sol göz görme kaybı, sol kolda kasılma, alkol kullanımı, sigara kullanımı, ayda geçirdiği atak sayısı, epilepsi soy geçmişi özelliklerinin epilepsi hastalığının tespitinde ayırt edici özellikler olduğu görülmektedir. Laboratuvar bulguları özellik setinin doğrusal regresyon analizinde, PDW, Vitamin B–12, Kreatinin, CKD–EP (eGFR) parametrelerinin p–değerleri 0,05 değerinin altında olduğundan epilepsi ve kontrol grubu sağlıklı bireylerin sınıflandırılmasında anlamlı özellikler olduğu sonucu çıkarılmıştır. FFBP yöntemine uygulanan laboratuvar bulguları özellik setindeki; NEUT%, EO%, IG%, PDW, MPV, PCT, NEUT#, LYMPH#, EO#, BASO#, RBC, HGB, HCT, MCV, MCH, RDW–SD, RDW–CV, ALT, Glukoz (Açlık kan şekeri) özelliklerinin epilepsi hastalığının tespitinde ayırt edici özellikler olduğu

görülmektedir. EEG özellik setinin doğrusal regresyon analizinde, F4, F8, T7, P3, Pz ve P4 kanallarının p-değerleri 0,05 değerinin altında olduğundan epilepsi ve kontrol grubu sağlıklı bireylerin sınıflandırılmasında anlamlı özellikler olduğu sonucu çıkarılmıştır. FFBP yöntemine uygulanan EEG özellik setindeki tüm kanalların hastalığın tespitinde aynı hata oranında (0,1803) yani aynı etki oranına sahip olduğu görülmektedir. MRG özellik setindeki tüm parametrelerin FFBP yöntemi ile elde edilen hata oranı sıfıra yakın olması sebebiyle hastalığın tespitinde 502 özelliğin tamamı ayırt edici özellikler olduğu görülmektedir. Bu sayede, bu çalışma hastalığın erken dönemde teşhisi için koruyucu tedaviye yansıyacak adımların atılması ile doktora yardımcı olacaktır. |



KAYNAKLAR

- [1]. Browne, T. R. and Holmes, G. L., 2008, *Handbook of epilepsy*, Jones & Bartlett Learning, USA.
- [2]. Shorvon, S., 2010, *Handbook of epilepsy treatment*, John Wiley & Sons, United States.
- [3]. Fisher, R.S., Acevedo, C., Arzimanoglou, A., Bogacz, A., Cross, J.H., Elger, C.E., Engel, J., 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. and Wiebe, S., 2014, ILAE official report: a practical clinical definition of epilepsy, *Epilepsia*, 55(4), 475–482.
- [4]. Acharya, U. R., Fujita, H., Sudarshan, V. K., Bhat, S. and Koh, J. E., 2015, Application of entropies for automated diagnosis of epilepsy using EEG signals: A review, *Knowledge-based systems*, 88, 85–96.
- [5]. Fisher, R. S., Cross, J. H., French, J. A., Higurashi, N., Hirsch, E., Jansen, F. E., Lagae, L., Moshé, S.L., Peltola, J., Roulet, Perez, E., Scheffer, I.E. and Zuberi, S.M., 2017, Operational classification of seizure types by the International League Against Epilepsy: Position Paper of the ILAE Commission for Classification and Terminology, *Epilepsia*, 58(4), 522–530.
- [6]. Shorvon, S., Perucca, E. and Engel Jr, J., 2015, *The treatment of epilepsy*, John Wiley & Sons, UK.
- [7]. www.itfnoroloji.org [Ziyaret tarihi: 14 Ekim 2018].
- [8]. Moghim, N. and Corne, D., 2011, Evaluating bio-inspired approaches for advance prediction of epileptic seizures, *In Nature and Biologically Inspired Computing (NaBIC), 2011 Third World Congress* 19 – 21 October 2011 Salamanca Spain, 540–545, IEEE.
- [9]. Fu, T., Kong, Q., Sheng, H. and Gao, L., 2016, Value of functionalized superparamagnetic iron oxide nanoparticles in the diagnosis and treatment of acute temporal lobe epilepsy on MRI, *Neural plasticity*, 2016, 1–12.
- [10]. Blümcke, I., Aronica, E., Miyata, H., Sarnat, H. B., Thom, M., Roessler, K., Rydenhag, B., Jehi, L., Wiebe, S. and Spreafico, R., 2016, International recommendation for a comprehensive neuropathologic workup of epilepsy surgery brain tissue: a consensus Task Force report from the ILAE Commission on Diagnostic Methods, *Epilepsia*, 57(3), 348–358.
- [11]. Bingöl, A. C., Çelik, M. and Gürtekin, Y., 2016, *Klinik nörofizyoloji laboratuvarları uygulama el kitabı*, 2nd ed., İstanbul: Türkiye Nörofizyoloji EEG–EMG Derneği Şubesi.
- [12]. Klem, G. H., Lüders, H. O., Jasper, H. H. And Elger, C., 1999, The ten-twenty electrode system of the International Federation. The International Federation of clinical

- neurophysiology, *Electroencephalography and clinical neurophysiology supplement*, 52, 3–6.
- [13]. Bayrak, Ş., Yücel, E. and Takçı, H., 2019, Classification of extracranial and intracranial EEG signals by using finite impulse response filter through ensemble learning, *27th signal processing and communications applications conference (SIU)*, 24 – 26 April 2019 Sivas, Turkey, IEEE, 1 – 4.
- [14]. Bayrak, Ş., Demirel, E. Y. and Şamlı, R., 2020, Classification of epileptic EEG signals based on finite impulse response filter and artificial neural networks training algorithms, *Uludağ University journal of the faculty of engineering*, 25(3), 1431–1444.
- [15]. Abhinaya, B. and Thanaraj, D. C. K. P., 2016, Feature extraction and selection of a combination of entropy features for real-time epilepsy detection, *International journal of engineering and computer science*, 5(4).
- [16]. Rajaguru, H. and Prabhakar, S. K., 2017, Hilbert transform with elman backpropagation and multilayer perceptrons for epilepsy classification, *Electronics, communication and aerospace technology (ICECA)*, 1, 571–576.
- [17]. Dohler, F., Chernihovskyi, A., Mormann, F., Elger, C. E. and Lehnertz, K., 2006, Detecting structural alterations in the brain using a cellular neural network based classification of magnetic resonance images, *10 th international workshop on cellular neural networks and their applications*, 28 – 30 August 2006 Istanbul, Turkey, IEEE, ISBN: 1-4244-0639-0, 1–4.
- [18]. Duque–Muñoz, L., Espinosa–Oviedo, J. J. and Castellanos–Dominguez, C. G., 2014, Identification and monitoring of brain activity based on stochastic relevance analysis of short-time EEG rhythms, *Biomedical engineering*, 13(1), 123.
- [19]. Gandhi, T. K., Chakraborty, P., Roy, G. G. and Panigrahi, B. K., 2012, Discrete harmony search based expert model for epileptic seizure detection in electroencephalography, *Expert systems with applications*, 39(4), 4055–4062.
- [20]. Dhiman, R. and Saini, J. S., 2014, Genetic algorithms tuned expert model for detection of epileptic seizures from EEG signatures, *Applied soft computing*, 19, 8–17.
- [21]. Oğulata, S. N., Şahin, C. and Erol, R., 2009, Neural network-based computer-aided diagnosis in classification of primary generalized epilepsy by EEG signals, *Journal of medical systems*, 33(2), 107–112.
- [22]. Kamath, C., 2013, A new approach to detect epileptic seizures in electroencephalograms using teager energy, *ISRN biomedical engineering*, 2013, 1–15.
- [23]. Li, Y., Wang, X. D., Luo, M. L., Li, K., Yang, X. F. and Guo, Q., 2018, Epileptic seizure classification of EEGs using time-frequency analysis based multiscale radial basis functions, *IEEE journal of biomedical and health informatics*, 22(2), 386–397.

- [24]. Übeyli, E. D., 2008, Support vector machines for detection of electrocardiographic changes in partial epileptic patients, *Engineering applications of artificial intelligence*, 21(8), 1196–1203.
- [25]. Adam, A., Shapiai, M. I., Tumari, M., Zaidi, M., Mohamad, M. S. and Mubin, M., 2014, Feature selection and classifier parameters estimation for EEG signals peak detection using particle swarm optimization, *The scientific world journal*, 1–13.
- [26]. Zainuddin, Z., Lai, K. H. and Ong, P., 2013, A hybrid algorithm for the initialization of wavelet neural networks: application in epileptic seizure classification, *International journal of applied physics and mathematics*, 3(5), 352.
- [27]. Sharif, B. and Jafari, A. H., 2017, Prediction of epileptic seizures from EEG using analysis of ictal rules on Poincaré plane, *Computer methods and programs in biomedicine*, 145, 11–22.
- [28]. Wang, G., Deng, Z. and Choi, K. S., 2017, Detection of epilepsy with Electroencephalogram using rule-based classifiers, *Neurocomputing*, 228, 283–290.
- [29]. Martinez-del-Rincon, J., Santofimia, M. J., del Toro, X., Barba, J., Romero, F., Navas, P. and Lopez, J. C., 2017, Non-linear classifiers applied to EEG analysis for epilepsy seizure detection, *Expert systems with applications*, 86, 99–112.
- [30]. Bayrak, Ş., Yücel, E., Takçı, H. and Şamlı, R., 2021, Classification of epileptic electroencephalograms using time-frequency and back propagation methods, *Computers, materials & continua*, DOI:10.32604/cmc.2021.015524.
- [31]. Yağmur, F. D., and Sertbaş, A., 2020, Automatic diagnosis of epilepsy from EEG signals using discrete cosine transform, *28th signal processing and communications applications conference (SIU)*, 5 – 7 October 2020 Gaziantep, Turkey, IEEE, 1 – 4.
- [32]. Silberschatz, A., Korth, H. F. And Sudarshan, S., 1997, *Database system concepts*, 4th ed., McGraw–Hill, New York.
- [33]. Mishra, P., Biancolillo, A., Roger, J. M., Marini, F. and Rutledge, D. N., 2020, New data preprocessing trends based on ensemble of multiple preprocessing techniques, *Trends in analytical chemistry*, 116045.
- [34]. Casley, D. J. and Kumar, K., 1988, *The collection, analysis, and use of monitoring and evaluation data*, Johns Hopkins University Press, Baltimore.
- [35]. Rahm, E. and Do, H. H., 2000, Data cleaning: Problems and current approaches, *IEEE data engineering bulletin*, 23(4), 3–13.
- [36]. Al Shalabi, L., Shaaban, Z. and Kasasbeh, B., 2006, Data mining: A preprocessing engine, *Journal of computer science*, 2(9), 735–739.

- [37]. Remeseiro, B. and Bolon-Canedo, V., 2019, A review of feature selection methods in medical applications, *Computers in biology and medicine*, 112, 103375.
- [38]. Wold, S., Esbensen, K. and Geladi, P., 1987, Principal component analysis, *Chemometrics and intelligent laboratory systems*, 2(1–3), 37–52.
- [39]. Mueller, J. P. and Massaron, L., 2018, *Artificial intelligence for dummies*, John Wiley & Sons, New Jersey.
- [40]. Source: <https://www.forbes.com/sites/gilpress/2016/03/23/data-preparation-most-time-consuming-least-enjoyable-data-science-task-survey-says/> [Ziyaret tarihi: 7 Ekim 2018].
- [41]. Turner, C. R., Fuggetta, A., Lavazza, L. and Wolf, A. L., 1999, A conceptual basis for feature engineering, *Journal of systems and software*, 49(1), 3–15.
- [42]. Hand, D. J. and Adams, N. M., 2014, Data mining, *Wiley statsref: statistics reference online*, 1–7.
- [43]. Bayrak, Ş., 2011, *Trigram özellik veri seti kullanılarak sınıflandırma yöntemleriyle dil tanıma*, Thesis (Master), Haliç University.
- [44]. Safavian, S. R. and Landgrebe, D., 1991, A survey of decision tree classifier methodology, *IEEE transactions on systems, man, and cybernetics*, 21(3), 660–674.
- [45]. Özkan, Y., 2008, *Veri madenciliği yöntemleri*, 1st ed., Papatya Yayıncılık, İstanbul, ISBN: 978-975-6797-82-2.
- [46]. Montgomery, D. C., Peck, E. A. and Vining, G. G., 2021, *Introduction to linear regression analysis*, 5th ed., John Wiley & Sons, New Jersey, ISBN: 978-1-118-62736-5.
- [47]. Zou, K. H., Tuncali, K. and Silverman, S. G., 2003, Correlation and simple linear regression, *Radiology*, 227(3), 617–628.
- [48]. Godfrey, K., 1985, Simple linear regression in medical research, *New England journal of medicine*, 313(26), 1629–1636.
- [49]. Elston, R. C., 1991, On Fisher's method of combining p-values, *Biometrical journal*, 33(3), 339–345.
- [50]. Al-Khurayji, R. and Sameh, A., 2017, An effective arabic text classification approach based on kernel naive bayes classifier, *International journal of artificial intelligence applications*, 8(6).
- [51]. Rish, I., 2001, An empirical study of the naive Bayes classifier, In *IJCAI 2001 workshop on empirical methods in artificial intelligence*, 3(22), 41–46.

- [52]. Noble, W. S., 2006, What is a support vector machine?, *Nature biotechnology*, 24(12), 1565–1567.
- [53]. Steinwart, I. and Christmann, A., 2008, *Support vector machines*, Springer Science & Business Media.
- [54]. Sarkar, M. and Leong, T. Y., 2000, Application of K–nearest neighbors algorithm on breast cancer diagnosis problem, *Proceedings of the AMIA Symposium*, 759–763.
- [55]. Xing, W. and Bei, Y., 2019, Medical health big data classification based on knn classification algorithm, *IEEE Access*, 8, 28808–28819.
- [56]. Van Gerven, M. and Bohte, S., 2017, Artificial neural networks as models of neural information processing, *Frontiers in computational neuroscience*, 11, 114.
- [57]. Kwon, K., Kim, D. and Park, H., 2017, A parallel MR imaging method using multilayer perceptron, *Medical physics*, 44(12), 6209–6224.
- [58]. Binfet, J. and Wilamowski, B. M., 2001, Microprocessor implementation of fuzzy systems and neural networks, *International Joint Conference on Neural Networks Proceedings*, 1, 234–239.
- [59]. McClelland, J. L., Rumelhart, D. E. and PDP Research Group, 1986, *Parallel distributed processing*, 2nd ed., MIT press, Cambridge.
- [60]. Graves, A., and Schmidhuber, J., 2005, Framewise phoneme classification with bidirectional LSTM and other neural network architectures, *Neural networks*, 18(5–6), 602–610.
- [61]. Schmidhuber, J. and Hochreiter, S., 1997, Long short–term memory, *Neural computation*, 9(8), 1735–1780.
- [62]. Luque, A., Carrasco, A., Martín, A. and de las Heras, A., 2019, The impact of class imbalance in classification performance metrics based on the binary confusion matrix, *Pattern recognition*, 91, 216–231.
- [63]. Visa, S., Ramsay, B., Ralescu, A. L. and Van Der Knaap, E., 2011, Confusion matrix–based feature selection, *MAICS*, 710, 120–127.
- [64]. Rodriguez, J. D., Perez, A. and Lozano, J. A., 2009, Sensitivity analysis of k–fold cross validation in prediction error estimation, *IEEE transactions on pattern analysis and machine intelligence*, 32(3), 569–575.
- [65]. Fawcett, T., 2006, An introduction to ROC analysis, *Pattern recognition letters*, 27(8), 861–874.

- [66]. Fleiss, J. L. and Cohen, J., 1973, The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability, *Educational and psychological measurement*, 33(3), 613–619.
- [67]. Banerjee, M., Capozzoli, M., McSweeney, L. and Sinha, D., 1999, Beyond kappa: A review of interrater agreement measures, *Canadian journal of statistics*, 27(1), 3–23.
- [68]. Dietterich, T. G., 2002, Ensemble learning, *The handbook of brain theory and neural networks*, 2, 110–125.
- [69]. Zhou, Z. H., 2009, Ensemble learning, *Encyclopedia of biometrics*, 1, 270–273.
- [70]. Galton, F., 1880, Statistics of mental imagery, *Mind*, 5(19), 301–318.
- [71]. Surowiecki, J., 2005, *The wisdom of crowds*, Anchor.
- [72]. Patel, T., 2008, *Ensemble of classifiers based data fusion of EEG and MRI for diagnosis of neurodegenerative disorders*, Thesis (PhD), Rowan University.
- [73]. Lahat, D., Adali, T. and Jutten, C., 2015, Multimodal data fusion: an overview of methods, challenges, and prospects, *Proceedings of the IEEE*, 103(9), 1449–1477.
- [74]. Khaleghi, B., Khamis, A., Karray, F. O. and Razavi, S. N., 2013, Multisensor data fusion: A review of the state-of-the-art, *Information fusion*, 14(1), 28–44.
- [75]. Uludağ, K., and Roebroek, A., 2014, General overview on the merits of multimodal neuroimaging data fusion, *Neuroimage*, 102, 3–10.
- [76]. LeCun, Y., Bengio, Y. and Hinton, G., 2015, Deep learning, *Nature*, 521(7553), 436–444.
- [77]. Tang, S. Y., Xiao, Y. N., Peng, Z. H. and Shen, H. B., 2020, Prediction modeling with data fusion and prevention strategy analysis for the COVID–19 outbreak, *Zhonghua liu xing bing xue za zhi= Zhonghua liuxingbingxue zazhi*, 41(4), 480–484.
- [78]. Sebe, N., Cohen, I., Garg, A. and Huang, T. S., 2005, *Machine learning in computer vision*, Springer Science & Business Media.
- [79]. Owens, A., Wu, J., McDermott, J. H., Freeman, W. T. and Torralba, A., 2016, Ambient sound provides supervision for visual learning, *In European conference on computer vision*, 801–816.
- [80]. Poria, S., Cambria, E. and Gelbukh, A., 2015, Deep convolutional neural network textual features and multiple kernel learning for utterance–level multimodal sentiment analysis, *Proceedings of the 2015 conference on empirical methods in natural language processing*, 19–21 September 2015 Lisbon, Association for Computational Linguistics, 2539–2544.

- [81]. Martínez, H. P. and Yannakakis, G. N., 2014, Deep multimodal fusion: Combining discrete events and continuous signals, *Proceedings of the 16th international conference on multimodal interaction*, 12–16 November 2014 Istanbul, New York, Association for Computing Machinery, ISBN: 978–1–4503–2885–2, 34–41.
- [82]. Kuncheva, L. I., 2014, *Combining pattern classifiers: methods and algorithms*, John Wiley & Sons.
- [83]. Ramachandram, D. and Taylor, G. W., 2017, Deep multimodal learning: A survey on recent advances and trends, *IEEE Signal processing magazine*, 34(6), 96–108.
- [84]. Senanayaka, J. S. L., Van Khang, H. and Robbersmyr, K. G., 2018, Multiple classifiers and data fusion for robust diagnosis of gearbox mixed faults, *IEEE Transactions on industrial informatics*, 15(8), 4569–4579.
- [85]. Mi, A., Wang, L. and Qi, J., 2016, A multiple classifier fusion algorithm using weighted decision templates, *Scientific programming*, 2016, 1–11.
- [86]. Bigdeli, B., Samadzadegan, F. and Reinartz, P., 2014, A decision fusion method based on multiple support vector machine system for fusion of hyperspectral and LIDAR data, *International journal of image and data fusion*, 5(3), 196–209.
- [87]. Rappoport, N. and Shamir, R., 2018, Multi-omic and multi-view clustering algorithms: review and cancer benchmark, *Nucleic acids research*, 46(20), 10546–10562.
- [88]. Toh, T. S., Dondelinger, F. and Wang, D., 2019, Looking beyond the hype: Applied AI and machine learning in translational medicine, *EBioMedicine*, 47, 607–615.
- [89]. Yi, D., Lei, Z. and Li, S. Z., 2015, Shared representation learning for heterogeneous face recognition, *In 2015 11th IEEE international conference and workshops on automatic face and gesture recognition*, 1, 1–7.
- [90]. Karpathy, A., Toderici, G., Shetty, S., Leung, T., Sukthankar, R. and Fei-Fei, L., 2014, Large-scale video classification with convolutional neural networks, *In proceedings of the IEEE conference on computer vision and pattern recognition*, 1725–1732.
- [91]. Neverova, N., Wolf, C., Taylor, G. and Nebout, F., 2015, Moddrop: adaptive multi-modal gesture recognition, *IEEE transactions on pattern analysis and machine intelligence*, 38(8), 1692–1706.
- [92]. Liu, H. and Zhang, L., 2019, Advancing ensemble learning performance through data transformation and classifiers fusion in granular computing context, *Expert systems with applications*, 131, 20–29.
- [93]. Si, L., Wang, Z., Liu, X., Tan, C., Xu, J. and Zheng, K., 2015, Multi-sensor data fusion identification for shearer cutting conditions based on parallel quasi-newton neural networks and the Dempster-Shafer theory, *Sensors*, 15(11), 28772–28795.

- [94]. Kuncheva, L. I., Bezdek, J. C. and Duin, R. P., 2001, Decision templates for multiple classifier fusion: an experimental comparison, *Pattern recognition*, 34(2), 299–314.
- [95]. Sui, J., Castro, E., He, H., Bridwell, D., Du, Y., Pearlson, G. D., Jiang, T. and Calhoun, V. D. , 2014, Combination of fMRI–SMRI–EEG data improves discrimination of schizophrenia patients by ensemble feature selection, *36th Annual international conference of the IEEE engineering in medicine and biology society*, 26–30 August 2014 Chicago, IEEE, 3889–3892.
- [96]. Marini, S., Limongelli, I., Rizzo, E., Malovini, A., Errichiello, E., Vetro, A., Da, T., Zuffari, O. and Bellazzi, R., 2016, A data fusion approach to enhance association study in epilepsy, *PloS one*, 11(12), e0164940.
- [97]. He, Q., Li, X., Kim, D. N., Jia, X., Gu, X., Zhen, X. and Zhou, L., 2020, Feasibility study of a multi-criteria decision-making based hierarchical model for multi-modality feature and multi-classifier fusion: Applications in medical prognosis prediction, *Information fusion*, 55, 207–219.
- [98]. Grant, A. C., Abdel-Baki, S. G., Omurtag, A., Sinert, R., Chari, G., Malhotra, S., Weedon, J., Fenton, A. A. and Zehtabchi, S., 2014, Diagnostic accuracy of microEEG: a miniature, wireless EEG device, *Epilepsy & Behavior*, 34, 81–85.
- [99]. Ihle, M., Feldwisch-Drentrup, H., Teixeira, C. A., Witon, A., Schelter, B., Timmer, J. and Schulze-Bonhage, A., 2012, EPILEPSIAE–A European epilepsy database, *Computer methods and programs in biomedicine*, 106(3), 127–138.
- [100]. Delorme, A. and Makeig, S., 2004, EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis, *Journal of neuroscience methods*, 134(1), 9–21.
- [101]. Diana, N. E., Kalsum, U., Sabiq, A., Jatmiko, W. and Mursanto, P., 2016, Comparing windowing methods on Finite Impulse Response (FIR) filter algorithm in Electroencephalography (EEG) data processing, *Journal of theoretical & applied information technology*, 88(3).
- [102]. Mmeremikwu, V. O. and Mbachu, C. B., 2018, EEG signal enhancement–FIR parzen window filter method, *American journal of engineering research (Ajer)*, 7(6), 51–55.
- [103]. Singh, V., Veer, K., Sharma, R. and Kumar, S., 2016, Comparative study of FIR and IIR filters for the removal of 50 Hz noise from EEG signal, *International journal of biomedical engineering and technology*, 22(3), 250–257.
- [104]. Mardia, K. V., 1970, Measures of multivariate skewness and kurtosis with applications, *Biometrika*, 57(3), 519–530.

- [105]. Barnett, V. and Lewis, T., 1994, *Outliers in statistical data*, 3rd ed., Wiley, England, ISBN: 978-0-471-93094-5.
- [106]. Bakar, Z. A., Mohamad, R., Ahmad, A. and Deris, M. M., 2006, A comparative study for outlier detection techniques in data mining, *2006 IEEE conference on cybernetics and intelligent systems*, 7–9 June 2006 Bangkok, IEEE, 1–6.
- [107]. Mujtaba, G., Shuib, L., Idris, N., Hoo, W. L., Raj, R. G., Khowaja, K., Shaikh, K. and Nweke, H.F., 2019, Clinical text classification research trends: Systematic literature review and open issues, *Expert systems with applications*, 116, 494–520.
- [108]. Lucini, F. R., Fogliatto, F. S., da Silveira, G. J., Neyeloff, J. L., Anzanello, M. J., Kuchenbecker, R. D. S. and Schaan, B. D., 2017, Text mining approach to predict hospital admissions using early medical records from the emergency department, *International journal of medical informatics*, 100, 1–8.
- [109]. Chen, P. H., Zafar, H., Galperin–Aizenberg, M. and Cook, T., 2018, Integrating natural language processing and machine learning algorithms to categorize oncologic response in radiology reports, *Journal of digital imaging*, 31(2), 178–184.
- [110]. Solti I., Cooke, C.R., Xia, F., Wurfel, M.M., 2009, Automated classification of radiology reports for acute lung injury: comparison of keyword and machine learning based natural language processing approaches, *Bioinformatics and biomedicine workshop*, 1–4 November 2009 Washington, IEEE, 314–319.
- [111]. Chen, L., R. Shah, T. Link, M. Bucknor, S. Majumdar and V. Pedoia, 2019, Text sentiment analysis for cartilage abnormalities detection from radiology reports, *Osteoarthritis and cartilage*, 27, S400–S401.
- [112]. Jianbo, Y., Zhu, H. and Tahmasebi, A., 2019, Classification of Pulmonary Nodular Findings based on Characterization of Change using Radiology Reports, *AMIA summits on translational science proceedings*, 285.
- [113]. Cheng, L.T.E, Zheng, J., Savova, G.K. and Erickson, B. J., 2010, Discerning tumor status from unstructured MRI reports—completeness of information in existing reports and utility of automated natural language processing, *Journal of digital imaging*, 23(2), 119–132.
- [114]. Zacharaki, E. I., Wang, S., Chawla, S., Yoo, D. S., Wolf, R., Melhem, E. R. and Davatzikos, C., 2009, Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme, *Magnetic resonance in medicine: an official journal of the international society for magnetic resonance in medicine*, 62(6), 1609–1618.
- [115]. Chen, Po–Hao, Zafar, H., Galperin–Aizenberg, M. and Cook, T., 2018, Integrating natural language processing and machine learning algorithms to categorize oncologic response in radiology reports, *Journal of digital imaging*, 31(2), 178–184.

- [116]. Chulho, K., Zhu, V., Obeid, J. and Lenert, L., 2019, Natural language processing and machine learning algorithm to identify brain MRI reports with acute ischemic stroke, *PloS one* 14(2), e0212778.
- [117]. Yala, A., Barzilay, R., Salama, L., Griffin, M., Sollender, G., Bardia, Lehman, C., Buckley, J.M., Coopey, S.B., Polubriaginof, F., J.E. Garber, Smith, B.L. , Gadd, M.A., Specht, M.C., Gudewicz, T.M., Guidi, A.J., Taghian, A. and Hughes, K.S., 2017, Using machine learning to parse breast pathology reports, *Breast cancer research and treatment*, 161(2), 203–211.
- [118]. Cummins, N., Amiriparian, S., Ottl, S., Gerczuk, M., Schmitt, M. and Schuller, B., 2018, Multimodal Bag-of-Words for cross domains sentiment analysis, *International conference on acoustics, speech and signal processing (ICASSP)*, 15 – 20 April 2018 Calgary, IEEE, 4954–4958.
- [119]. Zipf, G.K., 1935, *The Psychobiology of Language*, Taylor & Francis Group, London.
- [120]. Quiroz, J. C., Laranjo, L., Tufanaru, C., Kocaballi, A. B., Rezazadegan, D., Berkovsky, S. and Coiera, E., 2021, Empirical analysis of Zipf’s law, power law, and lognormal distributions in medical discharge reports, *International journal of medical informatics*, 145, 104324.
- [121]. Perotti, J. I. and Billoni, O. V., 2020, On the emergence of Zipf’s law in music, *Physica A: Statistical mechanics and its applications*, 549, 124309.
- [122]. Karami, A., Ghasemi, M., Sen, S., Moraes, M. F. and Shah, V., 2019, Exploring diseases and syndromes in neurology case reports from 1955 to 2017 with text mining, *Computers in biology and medicine*, 109, 322–332.
- [123]. Cao, X., Li, J., Wang, R., Wang, Y., Niu, Q. and Shi, J., 2019, Calibrating GloVe model on the principle of Zipf’s law, *Pattern recognition letters*, 125, 715–720.
- [124]. Chai, T. and Draxler, R. R., 2014, Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature, *Geoscientific model development*, 7(3), 1247–1250.
- [125]. Willmott, C. J. and Matsuura, K., 2005, Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance, *Climate research*, 30(1), 79–82.
- [126]. Zhou, L., Lai, K. K. and Yu, L., 2010, Least squares support vector machines ensemble models for credit scoring, *Expert systems with applications*, 37(1), 127–133.
- [127]. Bryll, R., Gutierrez–Osuna, R. and Quek, F., 2003, Attribute bagging: improving accuracy of classifier ensembles by using random feature subsets, *Pattern recognition*, 36(6), 1291–1302.

- [128]. Amari, S. I., 1993, Backpropagation and stochastic gradient descent method, *Neurocomputing*, 5(4–5), 185–196.



EKLER

EK 1.1. Girişimsel Olmayan Klinik Araştırmalar Etik Kurul Kararı

EK 1.2. Yöntemler için Geliştirilen MATLAB Kodlarından Örnekler

```

clear; clc;
%Verinin yüklenmesi
load('MRG.mat');

c = cvpartition(label,'KFold',10);
TrainIndex = c.training(1);
TestIndex = c.test(1);
TrainData = data(TrainIndex,:);
TrainLabel = label(TrainIndex);
TestData = data(TestIndex,:);
TestLabel = label(TestIndex);
clearvars c;

%%
%%ID3
ctree = fitctree(TrainData,TrainLabel);
view(ctree)
view(ctree,'mode','graph')
resuberror = resubLoss(ctree)
y_ID3 = predict(ctree,TestData)
cp1 = classperf(TestLabel, y)
sonuc = cp1.CorrectRate;
ID3_Model = fitctree(TrainData,TrainLabel,'OptimizeHyperparameters','auto')

%Karışıklık matrisi
figure
confusionchart(TestLabel, y_ID3);
conf_ID3 = confusionmat(TestLabel, y_ID3);

%ROC analizi
[tpr,fpr,thresholds] = plotroc(TestLabel, y_ID3)
clearvars cp result

%%
%%Lineer Regresyon
LRModel = fitlm(TrainData,TrainLabel)
table = anova(LRModel)
plotDiagnostics(LRModel)
plotResiduals(LRModel)
plotResiduals(LRModel,'probability')
plotSlice(LRModel)
y_LR = predict(LRModel, TestData);

%Test için karışıklık matrisi
figure
confusionchart(TestLabel, y_LR);
conf_LR = confusionmat(TestLabel, y_LR);

```

```

%ROC analizi
[tpr,fpr,thresholds] = plotroc(TestLabel,y_LR)
clearvars cp result

%%
%%Naive Bayes
NBModel = fitcnb(TrainData,TrainLabel,'Distribution','kernel');
y_NB = predict(NBModel, TestData);
cp2 = classperf(TestLabel, y_NB);
sonuc = cp2.CorrectRate;

%Karışıklık matrisi
figure
confusionchart(TestLabel, y_NB);
conf_NB = confusionmat(TestLabel, y_NB);

%ROC analizi
[tpr,fpr,thresholds] = plotroc(TestLabel,y_NB)

clearvars cp result

%%
%%Linear DVM
K = templateSVM('KernelFunction','Linear');
DVMModel = fitcecoc(TrainData, TrainLabel,'Learners', K);
y_DVM = predict(DVMModel, TestData);
cp3 = classperf(TestLabel, y_DVM);
sonuc = cp3.CorrectRate;

%Karışıklık matrisi
figure
confusionchart(TestLabel, y_DVM);
conf_dvm = confusionmat(TestLabel, y_DVM);

%ROC analizi
[tpr,fpr,thresholds] = plotroc(TestLabel,y_DVM)
clearvars cp result

%%
%%k-NN
kNNModel = fitcknn(TrainData,TrainLabel,'NumNeighbors',10);
y_knn = predict(kNNModel, TestData);
cp4 = classperf(TestLabel, y_knn);
sonuc = cp4.CorrectRate;

%Karışıklık matrisi
figure
confusionchart(TestLabel, y_knn);
conf_knn = confusionmat(TestLabel, y_knn);

%ROC analizi
[tpr,fpr,thresholds] = plotroc(TestLabel,y_knn)
clearvars cp result

%%

```

```

%%YSA
[~, loc] = ismember(TrainLabel, unique(TrainLabel));
y_one_hot = ind2vec(loc);

net = patternnet([10]);
net = train(net, TrainData, y_one_hot);
y_MLP = net(TestData);
cp5 = classperf(TestLabel, vec2ind(y_MLP));
sonuc = cp5.CorrectRate;

%%Karışıklık matrisi
figure
confusionchart(TestLabel, y_MLP);
conf_MLP = confusionmat(testLabel, y_MLP);

%%ROC analizi
[tp, fpr, thresholds] = plotroc(TestLabel, y_MLP)
clearvars cp result

%%
%%Verinin yüklenmesi
load('MRG.mat');
embeddingDimension = 100;
numHiddenUnits = 180;
numClasses = 2

layers = [ ...
    sequenceInputLayer(1)
    lstmLayer(numHiddenUnits, 'OutputMode', 'last')
    fullyConnectedLayer(numClasses)
    softmaxLayer
    classificationLayer]

%%Bir CPU üzerinde eğitim, bir GPU üzerinde eğitimden önemli ölçüde daha uzun sürebilir.

options = trainingOptions('adam', ...
    'MaxEpochs', 25, ...
    'GradientThreshold', 1, ...
    'InitialLearnRate', 0.001, ...
    'ValidationData', {XTrainAll, YTrainAll}, ...
    'Plots', 'training-progress', ...
    'Verbose', false);

%%LSTM ağının eğitilmesi
net = trainNetwork(XTrainAll, YTrainAll, layers, options);

%%Eğitilmiş LSTM ağını kullanarak eğitilmiş belgelerin sınıflandırılması.
YPredTrain = classify(net, XTrainAll);

%%Sınıflandırma doğruluğunun hesaplanması
accuracyTrain = sum(YPredTrain == YTrainAll)/numel(YPredTrain)

%%Karıştırma matrisi
figure
confusionchart(YTrainAll, YPredTrain);

```

```
%%Eğitilmiş LSTM ağına göre test metinlerinin sınıflandırılması  
YPredLSTM = classify(net,XTest);
```

```
%%Test metinlerinin sınıflandırma doğruluklarının hesaplanması  
accuracy = sum(YPredLSTM == YTest)/numel(YPredLSTM)
```

```
%%Test için karıştırma matrisi  
figure  
confusionchart(YTest, YPredLSTM);  
conf_LSTM = confusionmat(testLabel, YPredLSTM);
```

```
%ROC analizi  
[tpr,fpr,thresholds] = plotroc(testLabel,YPredLSTM)
```

|



