T.C.
MANİSA CELAL BAYAR ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ
ELEKTRİK ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI
ELEKTRİK ELEKTRONİK MÜHENDİSLİĞİ BİLİM DALI

PSİKOLOJİK UYARIMA DAYALI KAYGILI DURUMLARIN MAKİNE ÖĞRENMESİ İLE TESPİTİ

Damla ÇINAR

Danışman

Doç. Dr. Yücel KOÇYİĞİT



MANİSA–2024

DAMLA ÇINAR	**PSİKOLOJİİK UYARIMA DAYALI KAYGILI DURUMLARIN MAKİNE ÖĞRENMESİ İLE TEPİTİ**	**2024**

İÇİNDEKİLER

1. GİRİŞ 1
1.1 Tezin Konusu 2
1.2 Literatür Taraması 4
1.3 Tez Organizasyonu 7
2. GENEL BİLGİLER 8
2.1 Mental Bozukluklar 8
2.2 Makine Öğrenmesi 10
2.3 Elektroensefalografi 11
2.3.1 EEG Sinyalinin Öznitelik Çıkarma ve Filtreleme Yöntemleri 11
3 MATERYAL VE YÖNTEMLER 16
3.1 Proje Materyali 16
3.1.1 EMOTIV EPOC+ 17
3.1.2 EEG Sinyallerinin Ön işlemesi 18
3.1.3 Deney ve Veri Seti Etiketleme 19
3.1.4 EEG Sinyallerinin Sınıflandırılması 27
4 ARAŞTIRMA BULGULARI VE TARTIŞMA 31
4.1 İkili Sınıflandırma Öğrenme ve Test Sonuçlarının Analizi 32
4.1.1 İkili Sınıflandırma Yöntemi için K-En Yakın Komşu Algoritması Öğrenme ve Test Sonuçları Analizi 32
4.1.2 İkili Sınıflandırma Yöntemi için Yapay Sinir Ağları Algoritması Öğrenme ve Test Sonuçları Analizi 35
4.1.3 İkili Sınıflandırma Yöntemi için Destek Vektör Makinesi Algoritması Öğrenme ve Test Sonuçları Analizi 38
4.2 Dörtlü Sınıflandırma Öğrenme ve Test Sonuçlarının Analizi 43
4.2.1 Dörtlü Sınıflandırma Yöntemi için K-En Yakın Komşu Algoritması Öğrenme ve Test Sonuçları Analizi 44
4.2.2 Dörtlü Sınıflandırma Yöntemi için Yapay Sinir Ağları Algoritması Öğrenme ve Test Sonuçları Analizi 46
4.2.3 Dörtlü Sınıflandırma Yöntemi için Destek Vektör Makinesi Algoritması Öğrenme ve Test Sonuçları Analizi 48
5 SONUÇ VE ÖNERİLER 54
KAYNAKÇA 56

SİMGELER VE KISALTMALAR DİZİNİ

EEG	Elektroensefalografi
BBA	Beyin Bilgisayar Arayüzü
YZ	Yapay Zekâ
ML	Makine Öğrenmesi
ICD	Uluslararası Hastalık Sınıflandırması
DSM-V	Mental Bozuklukların Tanısal ve Sayımsal El Kitabı
EOG	Elektrookülogram
EMG	Elektromiyogram
EKG	Elektrokardiyogram
SDY	Sonlu Dürtü Yanıtı
KSA	Kör Sinyal Ayrıştırma
KKA	Kanonik Korelasyon Analizi
DVM	Destek Vektör Makinesi
YSA	Yapay Sinir Ağı
KNN	K-En Yakın Komşu Algoritması
OMS	Ortalama Mutlak Sapma
KOK	Kök Ortalama Kare
AÇK	Alıcı Çalışma Karakteristiği
DEHB	Dikkat Eksikliği ve Hiperaktivite Bozukluğu
FFBPNN	İleri Beslemeli Geri Yayılımlı Yapay Sinir Ağı
KOH	Kendini Organize Eden Haritalar
MDB	Majör Depresif Bozukluk
EMD	Empirik Mod Ayrıştırması
SAM	Kişisel Değerlendirme Mankeni

ŞEKİLLER DİZİNİ

Şekil 1 Elektroansefalogram (EEG) 1
Şekil 2 EMOTIV EPOC + ... 17
Şekil 3 Etiketleme Algoritması 20
Şekil 4 Katılımcı 23 için her bir kanala ait verilerin grafiği 21
Şekil 5 Katılımcı 23, Deney 1, Kanal 1 verilerinin grafiği 22
Şekil 6 Katılımcı 23 EEG Verileri Boyut İndirgeme, yeni boyutun grafiği 22
Şekil 7 Katılımcı 23, 5 saniyelik örnekleme grafiği 23
Şekil 8 Birleştirilmiş Veri Setinin Grafiği 23
Şekil 9 Destek Vektör Makinesi Küme Ayrışımı 28
Şekil 10 Yapay Sinir Ağları 30
Şekil 11 K-En Yakın Komşu Algoritması ile İkili Sınıflandırma için Öğrenme Grafiği 33
Şekil 12 K-En Yakın Komşu Algoritması ile İkili Sınıflandırma için Test Grafiği 34
Şekil 13 Yapay Sinir Ağları Algoritması ile İkili Sınıflandırma için Öğrenme Grafiği 36
Şekil 14 Yapay Sinir Ağları Algoritması İkili Sınıflandırma için Test Grafiği 37
Şekil 15 Destek Vektör Makinesi Algoritması ile İkili Sınıflandırma için Öğrenme Grafiği 39
Şekil 16 Destek Vektör Makinesi Algoritması ile İkili Sınıflandırma için Test Grafiği 40
Şekil 17 İkili Sınıflandırma için Tüm Modellerin Öğrenme Verileri F1 Skoru Karşılaştırma Grafiği 42
Şekil 18 İkili Sınıflandırma için Tüm Modellerin Test Verileri F1 Skoru Karşılaştırma Grafiği 43
Şekil 19 K-En Yakın Komşu Algoritması ile Dörtlü Sınıflandırma için Öğrenme Grafiği 44
Şekil 20 K-En Yakın Komşu Algoritması ile Dörtlü Sınıflandırma için Test Grafiği 45
Şekil 21 Yapay Sinir Ağları Algoritması ile Dörtlü Sınıflandırma için Öğrenme Grafiği 46
Şekil 22 Yapay Sinir Ağları Algoritması ile Dörtlü Sınıflandırma için Test Grafiği 47
Şekil 23 Destek Vektör Makinesi Algoritması ile Dörtlü Sınıflandırma için Öğrenme Grafiği 49
Şekil 24 Destek Vektör Makinesi Algoritması ile Dörtlü Sınıflandırma için Test Grafiği 50
Şekil 25 Dörtlü Sınıflandırma için Tüm Denemelerin Öğrenme Verileri F1 Skoru Karşılaştırma Grafiği 51
Şekil 26 Dörtlü Sınıflandırma için Tüm Denemelerin Test Verileri F1 Skoru Karşılaştırma Grafiği 52

TABLO DİZİNİ

Tablo 1 Hastalıkların DSM-V'e göre sınıflandırılması 8
Tablo 2 Bant Aralıkları 13
Tablo 3 K-En Yakın Komşu Algoritması ile İkili Sınıflandırma için Öğrenme Sonuçları..... 33
Tablo 4 K-En Yakın Komşu Algoritması ile İkili Sınıflandırma için Test Sonuçları 34
Tablo 5 Yapay Sinir Ağları Algoritması ile İkili Sınıflandırma için Öğrenme Sonuçları 36
Tablo 6 Yapay Sinir Ağları Algoritması ile İkili Sınıflandırma için Test Sonuçları . 37
Tablo 7 Destek Vektör Makinesi Algoritması ile İkili Sınıflandırma için Öğrenme Sonuçları..... 39
Tablo 8 Destek Vektör Makinesi Algoritması ile İkili Sınıflandırma için Test Sonuçları 40
Tablo 9 İkili Sınıflandırma için Tüm Modellerin Öğrenme Verileri F1 Skoru Karşılaştırması..... 41
Tablo 10 İkili Sınıflandırma için Tüm Modellerin Test Verileri F1 Skoru Karşılaştırması..... 42
Tablo 11 K-En Yakın Komşu Algoritması ile Dörtlü Sınıflandırma için Öğrenme Sonuçları..... 44
Tablo 12 K-En Yakın Komşu Algoritması ile Dörtlü Sınıflandırma için Test Sonuçları 45
Tablo 13 Yapay Sinir Ağları Algoritması ile Dörtlü Sınıflandırma için Öğrenme Sonuçları..... 46
Tablo 14 Yapay Sinir Ağları Algoritması ile Dörtlü Sınıflandırma için Test Sonuçları 47
Tablo 15 Destek Vektör Makinesi Algoritması ile Dörtlü Sınıflandırma için Öğrenme Sonuçları..... 49
Tablo 16 Destek Vektör Makinesi Algoritması ile Dörtlü Sınıflandırma için Test Sonuçları..... 50
Tablo 17 Dörtlü Sınıflandırma için Tüm Denemelerin Öğrenme Verileri F1 Skoru Karşılaştırması..... 51
Tablo 18 Dörtlü Sınıflandırma için Tüm Denemelerin Test Verileri F1 Skoru Karşılaştırması..... 51

TEŞEKKÜR

Çalışmamın her aşamasında bana hep destek olan, bilgi ve deneyimleri ile bana yol gösteren, bana olan inancını hiç kaybetmeyen sevgili danışman hocam Sayın Doç. Dr. Yücel KOÇYİĞİT'e, çalışmalarım sırasında manevi desteğini hep hissettiğim sevgili aileme ve arkadaşlarıma yürekten teşekkür ederim.

Damla Çınar
Manisa, 2024

ÖZET

Yüksek Lisans Tezi

Makine Öğrenmesi ile Psikolojik Uyarıma Dayalı Kaygılı Durumların Tespiti

Damla ÇINAR
Manisa Celal Bayar Üniversitesi
Lisansüstü Eğitim Enstitüsü
Elektrik Elektronik Anabilim Dalı

Danışman: Doç. Dr. Yücel KOÇYİĞİT

Psikolojik sağlık, bireylerin yaşam kalitesi üzerinde önemli bir etkiye sahiptir. Bu alanda, yapay zekâ ve makine öğrenmesi teknolojilerinin kullanımı, mental bozuklukların teşhisinde yenilikçi yaklaşımlar sunmaktadır. Bu teknolojiler, hastalıkların erken teşhisi ve kişiselleştirilmiş tedavi planlarının geliştirilmesinde geleneksel yöntemlere alternatifler sağlamaktadır. Günümüzde, semptomlara dayalı tanı yöntemleri yaygın olarak kullanılmakta, ancak biyobelirteçlerin kullanımı, teşhis süreçlerinde doğruluğu ve hızı artırmaktadır. Yapay zekâ ve makine öğrenmesi yöntemleri, biyobelirteçlerin analizinde ve yorumlanmasında önemli bir role sahip olup, psikiyatrik teşhislerin doğruluğunu artırarak geleneksel yöntemlerle ulaşılamayan derin içgörüler sunmaktadır.

-

Bu tez çalışması, psikolojik uyarımlara dayalı kaygı durumlarının tespitinde makine öğrenmesi algoritmalarının etkinliğini ve kullanılabilirliğini araştırmaktadır. Çalışma, EEG sinyallerinin kullanımına odaklanarak, bu sinyallerin analizi için gelişmiş algoritmalar ve yöntemlerin uygulanmasını içermektedir. Araştırmada, Intelligent Machines Lab tarafından hazırlanmış veri seti kullanılarak kaygı durumunun ortaya çıkışı incelenmiştir. Çalışmada, EEG sinyalleri üzerinde yapılan ikili ve dörtlü sınıflandırma görevlerinde K-En Yakın Komşu, Destek Vektör Makinesi, Yapay Sinir Ağı gibi algoritmaların performansları değerlendirilmiş, bu algoritmaların doğruluk, duyarlılık, özgüllük, kesinlik ve F1 skoru gibi çeşitli metrikler üzerinden analizleri yapılmıştır.

Anahtar Kelimeler: Yapay Zekâ, Makine Öğrenmesi, Mental Bozukluklar, Psikiyatrik Teşhis, EEG Sinyali Analizi, Biyobelirteçler,

2022, 70 sayfa

ABSTRACT

M.Sc. Thesis

Detection of Anxious Situations Based on Psychological Stimulation with Machine Learning

Damla ÇINAR

Manisa Celal Bayar University
Graduate School of Education
Department of Electrical and Electronics

Supervisor: Assoc. Dr. Yücel KOÇYİĞİT

Psychological health has a significant impact on individuals' quality of life. In this field, the use of artificial intelligence and machine learning technologies offers innovative approaches in the diagnosis of mental disorders. These technologies provide alternatives to traditional methods in the early diagnosis of diseases and the development of personalized treatment plans. Nowadays, diagnosis methods based on symptoms are commonly used, but the use of biomarkers increases the accuracy and speed in the diagnostic processes. Artificial intelligence and machine learning methods play a crucial role in the analysis and interpretation of biomarkers, enhancing the accuracy of psychiatric diagnoses and offering deep insights that are not attainable with traditional methods.

This thesis research investigates the effectiveness and applicability of machine learning algorithms in detecting anxiety states based on psychological stimuli. The study focuses on the use of EEG signals, applying advanced algorithms and methods for their analysis. In the research, a dataset prepared by the Intelligent Machines Lab was used to examine the emergence of anxiety states. In the study, the performance of algorithms such as K-Nearest Neighbor, Support Vector Machine, and Artificial Neural Network on binary and quaternary classification tasks on EEG signals were evaluated, analyzing various metrics like accuracy, sensitivity, specificity, precision, and F1 score.

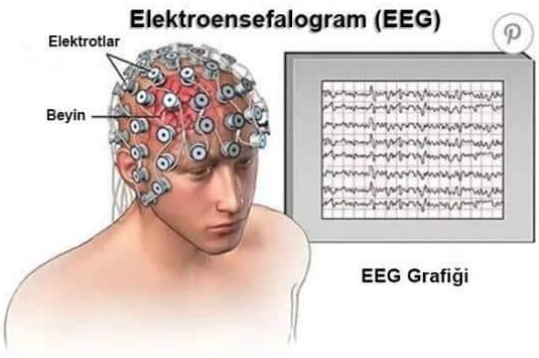
Keywords: Artificial Intelligence, Machine Learning, Mental Disorders, Psychiatric Diagnosis, EEG Signal Analysis, Biomarkers

2024, 70 pages

1. GİRİŞ

Makine öğrenmesi, verilerden öğrenme ve tahmin yapma yeteneğine sahip algoritmaların tasarımı ve uygulaması ile ilgilenen bir bilgisayar bilimi dalıdır. Makine öğrenmesi, yapay zekâ, veri madenciliği, istatistik, optimizasyon ve bilgisayar görüşü gibi alanlarla yakından ilişkilidir. Makine öğrenmesi, sınıflandırma, kümeleme, regresyon, boyut indirgeme, derin öğrenme, takviyeli öğrenme gibi çeşitli teknikleri içerir.

EEG (Elektroensefalografi), beyin hücrelerinin elektriksel aktivitesini ölçen ve kaydeden bir yöntemdir. EEG, beyin fonksiyonlarını, hastalıklarını, duygu durumlarını ve davranışlarını incelemek için kullanılan bir beyin-bilgisayar arayüzü (BBA) sistemidir. EEG, epilepsi, Alzheimer, Parkinson, uyku bozuklukları, depresyon, bağımlılık, travmatik beyin hasarı gibi birçok nörolojik ve mental bozukluğun tanı ve tedavisinde yardımcı olabilir. [1]



Şekil 1 Elektroansefalogram (EEG)

Makine öğrenmesi ile EEG analizi, EEG sinyallerinin işlenmesi, öznitelik çıkarımı, seçimi ve sınıflandırılması için makine öğrenmesi tekniklerinin kullanılmasıdır. Makine öğrenmesi ile EEG analizi, EEG sinyallerinin karmaşıklığını, gürültüsünü, değişkenliğini ve yüksek boyutluluğunu azaltmak, EEG sinyallerinden anlamlı bilgiler elde etmek ve EEG sinyallerini farklı kategorilere ayırmak için etkili bir yöntemdir.

1.1 Tezin Konusu

Günümüzde psikolojik sağlık, insan yaşamının kalitesini doğrudan etkileyen önemli bir faktördür ve bu alandaki yenilikler, toplum sağlığı açısından büyük önem taşımaktadır. Yapay zekâ (YZ) ve makine öğrenmesi (MÖ) teknolojilerinin psikolojik sağlık uygulamalarında kullanımı, son yıllarda bu alanda devrim niteliğinde gelişmelere yol açmıştır. Bu teknolojiler, mental bozuklukların teşhis ve tedavisinde kullanılan geleneksel yöntemlere alternatif sunarak, hastalıkların daha erken ve doğru bir şekilde teşhis edilmesini ve kişiselleştirilmiş tedavi planlarının geliştirilmesini sağlamaktadır. Özellikle büyük veri analizleri ve derin öğrenme algoritmaları, mental bozuklukların anlaşılması ve yönetilmesinde yeni ufuklar açmaktadır.

Mental bozuklukların tespitinde, doktorlar genellikle belirti tabanlı bir tanı yaklaşımı kullanırlar. Ancak, psikiyatri alanında, biyobelirteçlerin yerini klinik işaretler ve semptomlar almaktadır [2]. Bu biyobelirteçler, sağlıklı biyolojik süreçlerin veya tedaviye verilen farmakolojik tepkilerin fiziksel ölçüsü veya göstergesi olarak tanımlanabilir [3]. Yapay zekâ ve makine öğrenmesi yöntemleri, bu biyobelirteçlerin analizinde ve yorumlanmasında önemli bir rol oynayarak, psikiyatrik teşhislerin doğruluğunu ve hızını artırmaktadır. Bu teknolojiler, özellikle karmaşık ve çok boyutlu verilerin işlenmesinde etkili olup, geleneksel yöntemlerle elde edilemeyen içgörüler sunmaktadır.

Mental bozuklukların sınıflandırılması ve teşhisi uluslararası geçerliliği sahip iki ana kaynağa dayandırılır. Bu kaynaklar Amerika Psikiyatri Birliğinin Ruhsal Bozuklukların Tanısı ve Sayımsal El Kitabı [4] ve Dünya Sağlık Örgütünüm Hastalıkların Uluslararası Sınıflandırılması (ICD) [5] ismi verilen kaynaklardır. Mental bozuklukların teşhisinde, psikolojik değerlendirme ve DSM-V tanı kriterlerinin kullanıldığı, tedavi ve tanı takibinde kullanılan ortak psikometrik testler dışında objektif bir değerlendirme aracının bulunmadığı belirtilmiştir [6].

DSM-V gibi uluslararası geçerliliği olan kaynaklar mevcut olmasına rağmen, psikolojik değerlendirmeler ve tanı kriterleri genellikle subjektif ölçümlere dayanır ve objektif bir değerlendirme aracı sıklıkla eksiktir. Bu durum, biyolojik verilerle yapılan tanılara göre mental bozuklukların tanı güvenilirliğinin daha düşük olduğunu göstermektedir. Belirtilere dayalı tanı yaklaşımına alternatif olarak, model tabanlı yaklaşımlar, yani hesaplamalı psikiyatri, bu alanda nesnel ve ölçülebilir biyobelirteçlerle araştırmaların yapılmasını sağlayarak önemli gelişmelere yol açabilir.

Özellikle, kaygı duygu bozukluğu, günümüz toplumlarında en sık gözlemlenen mental bozukluklardan biridir. Bu bozukluk, insan hayatını ve verimini olumsuz etkileyebilir, depresyon ve intihara yol açabilir. Depresyonun 2030 yılından önce dünya çapında ikinci sakatlık ve ölüm nedeni olacağı tahmin edilmektedir [[7]– [8] – [9]].

Makine öğrenmesi yöntemlerinin mental bozuklukların tespitinde kullanılması, bu alanın modelleme ve tahmin çalışmaları için önemli bir potansiyel sunmaktadır. Yapay zekâ ve makine öğrenmesi, mental bozuklukların erken teşhisinde ve tedavi süreçlerinin optimize edilmesinde kritik bir role sahiptir. Bu teknolojilerin kullanımı, hasta bakımı kalitesini artırırken, sağlık hizmetlerinin maliyetini de azaltabilir.

Bu tez çalışmasının amacı, psikolojik uyarıma dayalı kaygılı durumların tespitinde makine öğrenmesi tekniklerinin kullanılabilirliğini ve etkinliğini araştırmaktır. Çalışmamız, EEG sinyallerinin kullanımına odaklanmakta olup, bu sinyallerin analizi için gelişmiş algoritmalar ve yöntemler uygulanmaktadır. Çalışmamızda, kaygı durumunun ortaya çıkışını incelemek için Intelligent Machines Lab tarafından hazırlanmış bir veri seti kullanılmaktadır [10].

Hazırlanmış tez çalışmasında, EEG sinyallerinin işlenmesi sürecinde filtreleme, gürültü giderme, segmentasyon, öznitelik çıkarma ve sınıflandırma gibi adımlar izlenmiştir. Amacımız, kaygı bozukluklarının daha doğru ve erken bir şekilde teşhis edilmesine olanak tanıyacak bir model geliştirmektir. Bu model, psikolojik durumların anlaşılmasında ve tanımlanmasında önemli bir rol oynayabilir ve kişiselleştirilmiş tedavi yaklaşımlarının geliştirilmesine katkı sağlamıştır.

1.2 Literatür Taraması

Başaran, çalışmasında, dikkat eksikliği ve hiperaktivite bozukluğu (DEHB) olan bireylerin EEG sinyallerini kullanarak, bu bozukluğun varlığını ve şiddetini kestirmeye yönelik bir yöntem geliştirmiştir. Bu amaçla, 30 DEHB'li ve 30 sağlıklı bireyden EEG sinyalleri toplanmış ve farklı frekans bantlarında güç spektrumu analizi yapılmıştır. Elde edilen öznitelikler, Geri Yayılmalı Yapay Sinir Ağı (FFBPNN), elman ağı ve Kendini Organize Eden Haritalar (KOH) gibi farklı yapay sinir ağları modellerine verilmiş ve bu modellerin DEHB'li bireyleri sağlıklı bireylerden ayırt etme ve DEHB'nin şiddetini belirleme performansları karşılaştırılmıştır. Sonuç olarak, FFBPNN modelinin %89 doğrulukla en iyi performansı gösterdiği, elman ağının %84 doğrulukla ikinci, KOH'un ise %70 doğrulukla üçüncü sırada yer aldığı bulunmuştur. Bu çalışma, DEHB'nin tanı ve tedavisinde EEG sinyallerinin ve yapay sinir ağlarının potansiyelini ortaya koymuştur [11].

Altun, çalışmasında makine öğrenme yöntemleri ile ham QEEG verisinden elde edilmiş olan biyobelirteçleri kullanarak, opioid kullanım bozukluğu sınıflandırılmasını amaçlamıştır [12]. QEEG, beyin dalgalarının frekans ve güç spektrumu analizi ile elde edilen bir yöntemdir. Opioid kullanım bozukluğu, opioid maddelerin aşırı ve bağımlılık yaratan şekilde kullanılması sonucu ortaya çıkan bir mental bozukluktur. Bu hastalık, beyin fonksiyonlarını, duygu durumunu ve davranışları etkileyebilir.

Çalışmada, 75 opioid kullanım bozukluğu olan ve 59 sağlıklı bireyden QEEG verileri toplanmıştır. Bu verilerden, teta ve beta frekans bantlarında, her elektrotun mutlak güç değeri hesaplanmıştır. Bu değerler, biyobelirteç olarak kullanılmıştır. Biyobelirteçler, biyolojik sistemlerin durumunu yansıtan ölçülebilir göstergelerdir. Biyobelirteçler, hastalıkların tanı, tedavi ve izleminde kullanılabilir.

Biyobelirteçler, yapay sinir ağları (YSA) gibi makine öğrenme yöntemleri ile işlenmiştir. Çalışmada, biyobelirteçlerin sınıflandırma performansını artırmak için, genetik algoritma kullanılmıştır. Genetik algoritma, doğal seçilim ve evrim ilkelerine dayanan bir optimizasyon yöntemidir.

Çalışmanın sonucunda, YSA modelinin, farklı frekans bantlarına göre, %94,73-%97,09 arasında değişen doğruluk değerleri ile, opioid kullanım bozukluğu olan ve olmayan bireyleri başarıyla sınıflandırdığı bulunmuştur. Bu çalışma, QEEG tabanlı biyobelirteçlerin ve makine öğrenme yöntemlerinin, opioid kullanım bozukluğu gibi mental bozuklukların tanı ve tedavisinde etkili araçlar olabileceğini göstermiştir.

Khodayari-Rostamabad, Reilly, vd., çalışmasında, majör depresif bozukluk (MDB) olan bireylerin, seçici serotonin geri alım inhibitörü (SSRI) adı verilen bir antidepresan ilaca nasıl yanıt vereceklerini, tedavi öncesi elektroensefalografi (EEG) verilerini kullanarak tahmin etmeyi amaçlamışlardır. Bu amaçla, 16 MDB'li ve 16 sağlıklı bireyden EEG verileri toplanmış ve farklı frekans bantlarında güç spektrumu analizi yapılmıştır. Elde edilen öznitelikler, kısmi en küçük kareler regresyonu (KPLRS) adı verilen bir makine öğrenmesi yöntemi ile işlenmiştir. KPLRS, girdi ve çıktı arasındaki ilişkiyi modelleyen bir regresyon yöntemidir. Regresyon, bir değişkenin diğer değişkenlere bağlı olarak nasıl değiştiğini belirlemek için kullanılan bir istatistiksel yöntemdir.

Çalışmada, KPLRS modelinin çıktısı, bireylerin SSRI ilacına verdikleri yanıtın tahmini olarak kullanılmıştır. Yanıt, Hamilton Depresyon Derecelendirme Ölçeği (HDRS) adı verilen bir klinik ölçek ile değerlendirilmiştir. HDRS, depresyonun şiddetini belirlemek için kullanılan bir ölçektir. Çalışmanın sonucunda, KPLRS modelinin, farklı frekans bantlarına göre, %85,7- %97,09 arasında değişen doğruluk değerleri ile, MDB'li bireylerin SSRI ilacına nasıl yanıt vereceklerini başarıyla tahmin ettiği bulunmuştur. Bu çalışma, EEG verilerinin ve makine öğrenmesi yöntemlerinin, depresyon gibi mental bozuklukların tedavisinde etkili araçlar olabileceğini göstermiştir [13].

Hosseinifard vd., dört EEG frekans bantlarıyla beraber lineer olmayan öznitelikleri de dahil ettikleri bir sınıflandırma çalışması yaparak depresyon hastaları ve sağlıklı bireyleri birbirinden ayırmışlardır [14]. Depresyon, beyin fonksiyonlarını, duygu durumunu ve davranışları etkileyebilir.

Çalışmada, 30 depresyon hastası ve 30 sağlıklı bireyden EEG verileri toplanmıştır. Bu verilerden, delta, teta, alfa ve beta frekans bantlarında, her elektrotun mutlak güç değeri, göreceli güç değeri, asimetri indeksi ve koherans değeri hesaplanmıştır. Bunlar, lineer özniteklikler olarak adlandırılmıştır. Ayrıca, her elektrotun Higuchi fraktal boyutu, Lyapunov üstel sayısı, entropi ve korelasyon boyutu gibi lineer olmayan öznitelikleri de hesaplanmıştır. Bu öznitelikler, EEG sinyallerinin karmaşıklığını, düzensizliğini ve kaotikliğini yansıtmaktadır.

Elde edilen öznitelikler, K-En Yakın Komşu (KNN), Doğrusal Ayırım Analiz (LDA) ve doğrusal lojistik regresyon (LJR) gibi farklı makine öğrenmesi yöntemleri ile işlenmiştir.

Çalışmada, öznitelik seçimi adı verilen bir teknik kullanılmıştır. Öznitelik seçimi, verilerden en anlamlı ve en az sayıda öznitelik seçmek için kullanılan bir makine öğrenmesi tekniğidir. Öznitelik seçimi, verilerin gürültüsünü, gereksizliğini ve kısmen eksikliğini azaltmak, verilerin anlaşılabilirliğini ve işlenebilirliğini artırmak, modelin performansını ve hızını iyileştirmek için önemlidir.

Çalışmanın sonucunda, KNN, LDA ve LJR modellerinin, farklı frekans bantlarına ve öznitelik seçimine göre, %73,3- %76,6 arasında değişen doğruluk değerleri ile, depresyon hastaları ve sağlıklı bireyleri başarıyla sınıflandırdığı bulunmuştur. Bu çalışma, EEG sinyallerinden elde edilen lineer ve lineer olmayan özniteliklerin ve makine öğrenmesi yöntemlerinin, depresyon gibi mental bozuklukların tanı ve tedavisinde etkili araçlar olabileceğini göstermiştir.

Mohammadi ve arkadaşları, depresyon hastalarını ve sağlıklı bireyleri EEG sinyallerine dayalı olarak ayırt etmek için bir makine öğrenmesi yöntemi geliştirmişlerdir. EEG, beyin hücrelerinin elektriksel aktivitesini ölçen bir yöntemdir.

Çalışmada, 30 depresyon hastası ve 30 sağlıklı bireyden EEG verileri toplanmıştır. Bu verilerden, alfa, beta, delta ve teta frekans bantlarında, her elektrotun mutlak güç değeri hesaplanmıştır. Bu değerler, EEG sinyallerinin öznitelikleri olarak kullanılmıştır.

Öznitelikler, C4.5 karar ağacı algoritması ile işlenmiştir. C4.5 karar ağacı algoritması, verileri ayrıştıran ve sınıflandıran bir makine öğrenmesi yöntemidir. Çalışmanın sonucunda, C4.5 karar ağacı algoritması, öznitelik seçimi yapılarak ve

farklı yöntemler kullanılarak, %80 doğruluk değeri ile, depresyon hastalarını ve sağlıklı bireyleri başarıyla ayırt ettiği bulunmuştur. Bu çalışma, EEG sinyallerinin ve makine öğrenmesi yöntemlerinin, depresyon gibi mental bozuklukların tanı ve tedavisinde etkili araçlar olabileceğini göstermiştir. [15].

Literatür taramasından görüldüğü üzere, kaygı durumunun hangi durumlarda ve hangi parametrelere bağlı olarak başladığı ile ilgili bir çalışmaya rastlanmamıştır.

1.3 Tez Organizasyonu

Hazırlanmış tez çalışması 5 bölümden oluşmaktadır. Tez çalışmasının ilk bölümü olan Giriş başlığı Tezin Konusu, Literatür Taraması, Tez Organizasyonu olarak üç alt başlıktan oluşmaktadır.

Hazırlanmış tez çalışmasında ikinci bölümde Genel Bilgiler başlığı altında Makine Öğrenmesi konusunun arka planı, EEG sinyallerinin yapısı ve nasıl ölçüldüğü hakkında bilgi verilmiştir. Devamında ise kullanılmış veri setinin özniteliklerinden, deney koşullarından, kayıt cihazı olan EMOTIV EPOC+ cihazının özelliklerinden bahsedilmiştir.

Hazırlanmış tez çalışmasında üçüncü bölümde Materyal ve Yöntemler başlığı altında EEG sinyallerine uygulanmış olan işlemlerden bahsedilmiştir. Deney sırasında toplanan EEG sinyallerine önceden yapılmış olan işlemlerden sonra, bu verilere ek olarak benim tarafımdan yapılan işlemler şunlardır: Verilerin boyutlarının değiştirilmesi, özniteliklerin çıkarılması ve sınıflandırma yöntemleri kullanılarak verilerin makine öğrenmesi teknikleriyle test edilmesi. Bu süreçte elde edilen başarı oranlarını belirlemek için yapılan matematiksel hesaplamalar detaylı bir şekilde açıklanmıştır.

Hazırlanmış tez çalışmasında dördüncü bölümde yapılan çalışmanın detaylıca anlatımı aktarılmıştır. Geliştirilmiş programdan ve çıktılarından, tez çalışması boyunca yapılan çalışmalardan bahsedilmiştir.

Hazırlanmış tez çalışmasında beşinci bölümde elde edilen sonuçlara ve EEG sinyallerinin en iyi sınıflandırmasının yapıldığı yöntemler önerilmiştir Tez çalışmasının geliştirilmesi için eklenebilecek adımlar önerilmiştir.

2. GENEL BİLGİLER

2.1 Mental Bozukluklar

Ruhsal rahatsızlık, kişinin duygusal, bilişsel ve davranışsal işlevlerinde bozulmaya neden olan, kişinin kendisine zarar vermesine sebep olabilecek veya başkaları için acı veya zarar yaratan, sosyal veya kültürel normlara uygun olmayan bir durumdur. Ruhsal rahatsızlık, beyin yapısı, kimyası, genetik, çevresel ve psikososyal faktörler gibi etkenler sayesinde meydana gelebilir. [16]

Tablo 1 Hastalıkların DSM-V'e göre sınıflandırılması [17]

DSM-V TANISAL ÖLÇÜTLER VE KODLAR			
İkiuçlu (Bipolar) ve İlişkili Bozukluklar	İkiuçlu (Bipolar) Bozukluk		
Nörogelişimsel Bozukluklar	Dikkat Eksikliği/Aşırı Hareketlilik Bozukluğu		
Depresyon Bozuklukları	Yeğin (Majör) Depresyon Bozukluğu		
Takıntı-Zorlantı Bozukluğu (Obsesif Kompulsif Bozukluk) ve İlişkili Bozukluklar	Takıntı-Zorlantı Bozukluğu (Obsesif Kompulsif Bozukluk)		
Madde ile İlişkili Bozukluklar ve Bağımlılık Bozuklukları	Madde ile İlişkili Bozukluklar	Opiyatla İlişkili Bozukluklar	Opiyat Kullanım Bozukluğu
Örselenme (Travma) ve Tetikleyici Etkenle (Stresörle) İlişkili Bozukluklar	Örselenme Sonrası Gerginlik (Travma Sonrası Stres) Bozukluğu		
Şizofreni Açılımı Kapsamında ve Psikozla Giden Diğer Bozukluklar	Şizofreni		

2.1.1 Mental Bozuklukların Sınıflandırılması

Mental bozukluklar DSM-V'ye göre Tablo 1'de belirtildiği gibi sınıflandırılmıştır. [17]

- İkiuçlu (Bipolar) ve İlişkili Bozukluklar
    - İkiuçlu (Bipolar) Bozukluk

Duygudurum dalgalanmaları ile karakterize bir durumdur. Kişi, aşırı mutluluk (mani) ve ağır depresyon dönemleri arasında gidip gelir.

- Nörogelişimsel Bozukluklar
    - Dikkat Eksikliği Bozukluğu

Dikkat süresinin kısa olması, kolayca dikkatin dağılması ve unutkanlık ile karakterize bir durumdur.

    - Aşırı Hareketlilik Bozukluğu

Kontrolsüz enerji seviyeleri, huzursuzluk ve dürtüsel davranışlar ile tanımlanır. Çocuklarda daha sık görülür.

- Depresyon Bozuklukları
    - Yeğin (Majör) Depresyon Bozukluğu

Depresyon ruh halinde uzun süreli bir düşüş, ilgi kaybı, umutsuzluk, suçluluk, değersizlik, uyku ve iştah bozuklukları gibi belirtilerle karakterize bir mental bozukluktur.

- Takıntı-Zorlantı Bozukluğu (Obsesif Kompulsif Bozukluk) ve İlişkili Bozukluklar
    - Takıntı-Zorlantı Bozukluğu (Obsesif Kompulsif Bozukluk)

Takıntılı düşünceler ve bunları gidermek için yapılan zorlayıcı davranışlarla karakterize bir durumdur.

- Madde ile İlişkili Bozukluklar ve Bağımlılık Bozuklukları
    - Madde ile İlişkili Bozukluklar

Alkol, uyuşturucu gibi maddelerin kullanımı sonucu psikolojik ve fizyolojik sorunlar yaşanmasıdır.

- Opiyat ile İlişkili Bozukluklar

Opiyatların (örneğim. morfin, heroin) kötüye kullanımı ve bağımlılığıdır.

- Opiyat Kullanım Bozukluğu

Opiyat bağımlılığı, tolerans gelişimi ve kullanımın olumsuz sonuçları ile tanımlanır.

- Örselenme (Travma) ve Tetikleyici Etkenle (Stresörle) İlişkili Bozukluklar
  - Örselenme Sonrası Gerginlik (Travma Sonrası Stres) Bozukluğu

Şiddetli travma sonrası gelişen, sürekli stres, korku ve kaygı hali.

- Şizofreni Açılımı Kapsamında ve Psikozla Giden Diğer Bozukluklar
  - Şizofreni

Gerçeklikten kopma, halüsinasyonlar, sanrılar ve düzensiz düşünme/davranış biçimleri ile karakterize bir psikotik bozukluktur.

2.2 Makine Öğrenmesi

Makine öğrenmesi yapay zekâ çalışmalarının bir alt kümesidir. Bilgisayarın geçmişte kaydedilmiş verilerden kendi kendine öğrenmesini ve öğrenmeye dayalı olarak bir sonuç çıkarmasını sağlayan bir teknolojidir. Makine öğrenmesi modelinde bir bilgisayar kararların alındığı bazı sınıfları ve ölçümleri içeren geçmiş deneyimlerden ve kalıplardan öğrenebilir ve insan ögesinin gerçek katılımı olmadan modelin performansına bağlı olarak sonuç sağlayabilir. Bu sebepten dolayı makine öğrenmesi genellikle istatistiksel ve olasılık matematiği, sınıflandırma teknikleri, regresyon teknikleri gibi çeşitli analitik yöntemleri uygular. Bu yöntemler ile bilgisayarın geçmiş verilerden öğrenmesine imkân tanınır ve öğrenme sonrasında büyük, karmaşık ve gürültülü verilerden elde edilebilecek sonuçlar tahmin edilir [18] [19]. Bu öğrenme yöntemleri denetimli ve denetimsiz öğrenme olmak üzere ikiye ayrılır. Denetimli öğrenme iki veya ikiden fazla veri kümesini önceden tanımlanmış

etiketlere göre eşlemeyi sağlayan bir fonksiyon geliştirmesidir. Denetimsiz öğrenme ise araştırmacı tarafından sağlanan herhangi bir etiketleme veya önişleme kullanılmadan verilerin nasıl düzenlendiğini belirlemeye çalışır [20].

2.3 Elektroensefalografi

Elektroensefalografi (EEG), beyin tarafından üretilen elektrik alanlarının yüksek zamansal ve uzaysal çözünürlüğüne sahip olan tek beyin görüntüleme yöntemidir [21]. EEG beyinden gelen elektrik sinyallerini kaydeder ve beyin aktivitesi hakkında değerli bilgiler elde etme fırsatı sağlar [22]. EEG sinyalleri, beyindeki nöronlar tarafından üretilen elektriksel aktivitenin izlenmesine olanak sağlar [23].

2.3.1 EEG Sinyalinin Öznitelik Çıkarma ve Filtreleme Yöntemleri

Öznitelik çıkarma işlemi, EEG verilerinden anlamlı bilgilerin elde edilmesi için kullanılan bir çok yöntemi ifade eder. EEG sinyalleri, beyin elektriksel aktivitesini yansıttığından, bu sinyallerden çıkarılan öznitelikler beyin fonksiyonları ve durumları hakkında önemli bilgiler sağlamaktadır.

2.3.1.1 EEG Sinyallerinin Zaman Alanı Öznitelikleri

2.3.1.1.1 Genlik

Sinyalin en yüksek ve en düşük değerleri arasındaki farktır. Genlik, sinyalin gücünü ve enerjisini gösterir. Sinyalin maksimum ve minimum değerlerinin farkı alınarak hesaplanır. [24]

$$A = x_{max} - x_{min} \tag{2.1}$$

2.3.1.1.2 Ortalama

Değerlerin toplamının örnek sayısına bölünmesi ile elde edilir. [24]

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \tag{2.2}$$

2.3.1.1.3 Standart Sapma

Varyans değerinin kareköküdür. Değerlerin yayılımının gösteriminde kullanılır. [24]

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2} \tag{2.3}$$

2.3.1.1.4 Hjört Parametreleri

Hjört parametreler, zaman serisinin karakteristiklerini tanımlamak için kullanılır. EEG sinyalleri gibi zaman serileri için üç temel Hjört parametresi vardır.

2.3.1.1.5 Aktivite

Aktivite, sinyalin gücünü veya varyansını ifade eder. Sinyalin genel dalgalanma yoğunluğunun bir ölçüsüdür.

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2 \tag{2.4}$$

2.3.1.1.6 Hareketlilik

Hareketlilik, sinyalin ortalama frekansının değişim hızını ölçer. Standart sapma ve frekans alanındaki güç spektrumunun bir ölçüsüdür.

$$\text{Hareketlilik} = \sqrt{\frac{\text{Sinyalin birinci türevinin varyansı}}{\text{Sinyalin varyansı}}} \tag{2.5}$$

$$\text{Sinyalin birinci türevinin varyansı} = \frac{1}{N-1}\sum_{i=1}^{N-1}(x_{i+1} - x_i)^2 \tag{2.6}$$

2.3.1.1.7 Karmaşıklık

Karmaşıklık, sinyalin dalga formunun ne kadar karmaşık olduğunu ölçer.

$$\text{Karmaşıklık} = \frac{\text{Sinyalin birinci türevinin hareketlilik değeri}}{\text{Sinyalin hareketlilik değeri}} \tag{2.7}$$

Hjört parametreleri EEG sinyallerinin analizinde sinyalin dinamik özniteliklerini anlamak için kullanılır. Bu parametreler beynin elektriksel aktivitesinin farklı durumlarını ayırt etmek için kullanılabilir.

2.3.1.2 EEG Sinyallerinin Frekans Alanı Öznitelikleri

2.3.1.2.1 Güç Spektral Yoğunluğu (PSD)

Sinyalin frekans bileşenlerinin gücünü gösterir ve genellikle Fourier dönüşümü veya Welch yöntemi gibi teknikler kullanılarak hesaplanır.

$$\mathrm{PSD(f)} = \lim_{T \to \infty} \frac{1}{T} |X(f)|^2 \quad (2.8)$$

PSD(f) değeri frekansta f olan güç spektral yoğunluğudur. T sinyalin süresidir. X(f) sinyalin Fourier dönüşümüdür ve $|X(f)|^2$ sinyalin o frekanstaki gücünü ifade eder.

2.3.1.2.2 Bant Güçleri

EEG sinyalleri belirli frekans bantlarına ayrılır. Her bir bant için toplam güç hesaplanır. Bantların frekans bilgileri Tablo 2'de belirtilmektedir. [25]

Tablo 2 Bant Aralıkları

Delta	0.5 – 4 Hz.
Teta	4 – 8 Hz.
Alpha	8 – 12 Hz.
Beta	12 – 30 Hz.
Gamma	30 Hz. ve üzeri

Bant gücü PSD'nin o banttaki frekans aralığında integralin alınması ile hesaplanır:

$$\text{Bant Gücü} = \int_{f_{low}}^{f_{high}} \mathrm{PSD(f)df} \quad (2.9)$$

Burada f_{low} ve f_{high} ilgili frekans bandının alt ve üst sınırlarıdır.

2.3.1.2.3 Spektral Entropi

Spektral entropi, sinyalin frekans dağılımının düzenliliğini veya düzensizliğini ölçer. Daha düzenli bir sinyal daha düşük spektral entropiye sahiptir. Entropi, genellikle güç spektrumunun normalleştirilmiş versiyonunun bir fonksiyonu olarak hesaplanır.

$$H = -\sum_{i=1}^{N} p(f_i) \log\big(p(f_i)\big) \quad (2.10)$$

$p(f_i)$ i'inci frekanstaki güç spektrumunun toplam güç spektrumuna oranıdır ve bu değerlerin toplamı 1 olmalıdır. N toplam frekans sayısını ifade eder. log fonksiyonu doğal algoritma olarak alınır.

Tanımlanmış öznitelikler EEG sinyallerinin analizinde ve beyin durumlarının sınıflandırılmasında kullanılabilir.

2.3.1.3 EEG Sinyallerinin Zaman-Frekans Öznitelikleri

2.3.1.3.1 Kısa Zamanlı Fourier Dönüşümü (STFT)

STFT sinyalin hem zaman hem de frekans bilgisini içeren bir temsil sunar. Sinyal kısa zaman dilimlerine ayrılarak her bir dilim üzerine Fourier dönüşümü uygular.

$$\mathrm{STFT}\{x(t)\}(f,\tau) = \int_{\infty}^{\infty} x(t)w(t-\tau)e^{-j2\pi ft}dt \qquad (2.11)$$

$x(t)$ zaman alanındaki sinyali temsil eder. $w(t-\tau)$ zaman dilimi etrafında yerel bir pencere fonksiyonunu temsil eder. τ zaman dilimi merkezi zamanını temsil eder. f frekansı temsil eder.

2.3.1.3.2 Wavelet Dönüşümü

Wavelet dönüşümü, sinyalin farklı frekanslardaki değişimlerini çözümlemek için kullanılır. Sinyal hem zaman, hem de ölçek açısından analiz edilir.

$$W_x(a,b) = \frac{1}{\sqrt{a}}\int_{\infty}^{\infty} x(t)\psi\left(\frac{t-b}{a}\right)dt \qquad (2.12)$$

a ölçek parametresini, b zaman kaymasını, $\psi(t)$ wavalet fonksiyonunu temsil etmektedir.

2.3.1.3.3 Hilbert-Huang Dönüşümü

Hilbert-Huang dönüşümü, özellikle doğrusal olmayan ve durağan olmayan veriler için uygundur. İlk olarak Empirik Mod Ayrıştırması (EMD) ile sinyaller,

verinin içerdiği farklı frekanslardaki dalgalanmaları temsil eden bir dizi yerel olarak durağan modlara ayrılır. Bu süreçte sinyal, farklı frekans bantlarına ayrıştırılarak analiz için daha yönetilebilir bileşenlere indirgenir. Ardından her mod Hilbert dönüşümüne tabi tutulur ve anlık frekans ve genlik bilgisini içeren bir zaman-frekans enerji spektrumu elde edilir. Bu yaklaşım, sinyalin iç yapısını ve anlık değişimlerini ortaya çıkararak zaman serisi analizinde derinlemesine bilgiler sunar.

$$\mathrm{HHT} = \mathrm{EMD}\{x(t)\} + \text{Hilbert Transform} \tag{2.13}$$

2.3.1.4 EEG Sinyallerinin Bağlantı ve Koherans Öznitelikleri

2.3.1.4.1 Fonksiyonel Bağlantı

Fonksiyonel bağlantı, farklı beyin bölgeleri arasındaki istatistiksel bağımlılıkları ölçer ve genellikle korelasyon veya kovaryans matrisleri ile temsil edilir.

2.3.1.4.2 Koherans

Koherans, iki sinyal arasındaki faz ve frekans uyumunu ölçer ve genellikle iki sinyalin çapraz spektral yoğunluğunun, her bir sinyalin güç spektral yoğunluklarının çarpımına bölümü ile hesaplanır.

$$\mathrm{Koherans}(f) = \frac{|P_{xy}(f)|^2}{P_{xx}(f)P_{yy}(f)} \tag{2.14}$$

$P_{xy}(f)$ iki sinyalin çapraz spektral yoğunluğunu temsil eder. $P_{xx}(f)P_{yy}(f)$ her bir sinyalin güç spektral yoğunluklarını temsil eder.

3 MATERYAL VE YÖNTEMLER

Materyal ve yöntemler bölümünde, hazırlanmış projede ilerlenmiş adımlar tanıtılmıştır.

3.1 Proje Materyali

Proje materyali olarak kullanılmış EEG sinyallerine ait veriler bir deney yapılarak toplanmıştır. İlgili veri seti https://ieee-dataport.org/open-access/dasps-database sitesinden alınmıştır [10].

Deney 23 sağlıklı katılımcı ile gerçekleştirilmiştir. Katılımcılar ortalama 30 yaşında, 13 kadın ve 10 erkekten oluşmaktadır. Deneyin amacı, deneye başlamadan önce katılımcılara detaylı olarak anlatılmıştır. Deney dikkat dağınıklığından kaçınmak ve katılımcının tam konsantrasyonunun sağlanması için izole bir ortamda gerçekleştirilmiştir.

Kaygı duygu durumunun ortaya çıkmasını incelemek üzere deney 23 sağlıklı kişi üzerinde 6 senaryonun her senaryo için öncelikle bir psikoterapist tarafından katılımcıya okunması, sonrasında katılımcının kendine hatırlatması şeklinde gerçekleştirilmiştir. Tüm süre boyunca deney ile EEG sinyalleri kaydedilmiş, süre dolduğunda her katılımcıdan Kendini Değerlendirme Anketi kullanılarak deney sırasında kendini nasıl hissettiğini derecelendirmesi istenmiştir.

Kayıt EMOTIV EPOC yazılımı kullanılarak gerçekleştirilmiştir. İlgili yazılım tüm kanalların veya sadece ihtiyaç duyulanların verilerini görüntülemeyi ve kaydetmesini sağlar. Üretilen ham veriler ‘.edf’ formatına sahiptir ve sonrasındaki işlemler için matlab betiği kullanılarak ‘.mat’ formatına dönüştürülebilir. [26]

3.1.1 EMOTIV EPOC+

EMOTIV EPOC+ EEG cihazı, beyin hücrelerinin elektriksel aktivitesini ölçen ve kaydeden bir cihazdır. EEG, elektroensefalografi adı verilen bir yöntemdir. EEG, beyin fonksiyonlarını, hastalıklarını, duygu durumlarını ve davranışlarını incelemek için kullanılan bir bilimsel yöntemdir.



Şekil 2 EMOTIV EPOC +

EMOTIV EPOC+ EEG cihazının bazı özellikleri şunlardır: [27]

- 14 kanallı, 128 veya 256 SPS örnekleme hızına, 2048 Hz dahili örnekleme hızına ve 14 bit çözünürlüğe sahiptir.
- AF3, AF4, F3, F4, FC5, FC6, F7, F8, T7, T8, P7, P8, O1, O2 adlı 14 elektrotu vardır. Bu elektrotlar, beyin bölgelerinin farklı fonksiyonlarını yansıtan EEG sinyallerini kaydetmek için kullanılır.
- Kablosuz olarak PC ve mobil cihazlara bağlanabilir. Şarj edilebilir ve 12 saate kadar pil ömrü vardır.
- 9 eksenli hareket sensörleri, baş hareketlerini algılayabilir.
- Ham EEG, zihinsel komutlar, performans metrikleri ve yüz ifadeleri algılamaları sunar.

EMOTIV EPOC+ EEG cihazı, ölçeklenebilir ve bağlamsal insan beyin araştırmaları ve gelişmiş beyin-bilgisayar arayüzü (BCI) uygulamaları için

tasarlanmıştır. Bu cihaz, profesyonel kalitede beyin verilerine hızlı ve kolay bir şekilde erişim sağlar.

3.1.2 EEG Sinyallerinin Ön işlemesi

Biyomedikal sinyal işleme sürecinde, gürültü ve artefaktların belirlenmesi ve temiz bir sinyalle öznitelik çıkarma aşamasına geçmek önemlidir. İyi bir sınıflandırma sonuçları elde etmek için aşağıdaki adımlar veri setine uygulanmıştır. Projede kullanılmış veriler '.edf' formatında kaydedilmiştir. Veri seti '.mat' formatında, Matlab yazılımında düzenlenmiştir. Veri setinin önişleme bölümü, çalışmanın yapıldığı veri setinden alınmıştır. [26]

3.1.2.1 Gürültü ve Artefaktların Belirlenmesi

Fizyolojik artefaktlar, beyinden farklı bir kaynaktan üretilir. Örneğin, elektrookülogram (EOG) artefaktları 4 Hz altındadır. Elektromiyogram (EMG) artefaktları 30 Hz'yi aşar. Elektrokardiyogram (EKG) artefaktları ise 1.2 Hz'dir. Ekstra fizyolojik artefaktlar ise insan vücuduyla ilgili değildir, çoğunlukla 50 Hz aralığındadır. Çevre veya elektroensefalografi (EEG) veri toplama parametreleri ile ilgili olabilir.

3.1.2.2 Gürültü Giderme Süreci

Sinyallerin temizlenmesi için EEGLab scripti kullanılmıştır. EEG sinyallerinin ilgili alt bandını keserek, temel çizgiyi çıkararak, Oküler (gözle ilgili) ve Masküler (kasla ilgili) artefaktlar çıkarılmıştır. Ham verilere 4-45 Hz Sonlu Dürtü Yanıtı (SDY) geçiş-bant filtresi uygulanmıştır. Otomatik artefakt çıkarma işlemi için EEGLAB araç kutusu kullanılmış ve EOG ve EMG artefaktları çıkarılmıştır. EMG ve EOG artefaktlarını çıkarmak için birkaç algoritma ilgili araç kutusunda bulunmaktadır. Gözlemlenen EEG verilerini en yüksek otokorelasyona sahip bileşenlere dönüştüren Kör Sinyal Ayrıştırma - Kanonik Korelasyon Analizi (KSA-KKA) yöntemi kullanılmıştır.

EEG ve EMG bantlarında güç tahmini için, analiz penceresi uzunluğunda eşit segment uzunluğu ile Hamming pencereli Welch periodogramı kullanılmıştır.

3.1.2.3 Kullanılan Yöntem ve Algoritmalar

Optimal Kör Kaynak Ayrıştırma için iWASOBI algoritması ve EOG artefaktlarını otomatik olarak düzeltmek için *eog fd* kriteri kullanıldı. Düşük fraktal boyutlu bileşenler, genellikle düşük frekanslı bileşenlerden oluşur ve bu da göz aktivitesi için uygun bir kriterdir.

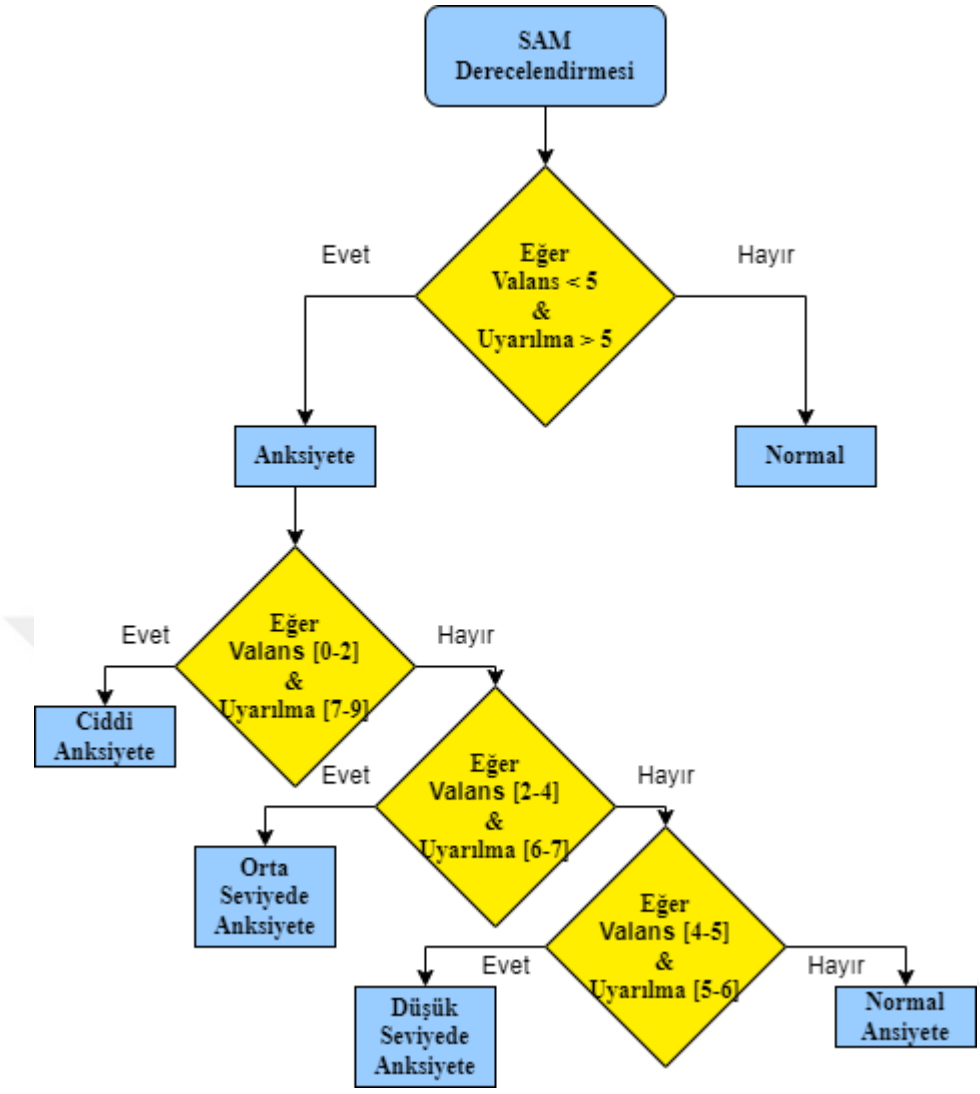
Bu süreçler, EEG verilerinden yüksek kaliteli ve temiz sinyaller elde etmek için kritik öneme sahiptir. Doğru sınıflandırma için temel oluşturur.

3.1.3 Deney ve Veri Seti Etiketleme

Deney her katılımcı için 6 dakika sürer ve 6 farklı duruma bölünmüştür. İlgi her denemenin ilk 30 saniyesindedir. Bu adımda 15 saniyelik Kişisel Değerlendirme Mankeni (SAM) bölümü çıkarılır. Sonuç olarak her bir katılımcı için 6 deneme ve 30 saniye elde edilir.

Katılımcılar, her 30 saniyelik uyarım sonrasında SAM anketini doldurarak, uyarım sırasında hissettikleri heyecanı ve duyguyu ifade eder. Deneyin sonunda, katılımcılardan en kaygı verici durumu belirtmeleri istenir. Bu bilgi, tüm denemeleri etiketlemek için kullanılmaktadır. Etiketleme algoritması Şekil 3 ile temsil edilmektedir.

Etiketleme sonucunda 156 Normal, 90 Şiddetli, 10 Orta ve 20 Hafif deneme elde edilir. Belirtilen valans değeri, bir deneyimin ya da uyarıcının hoşa gidici olup olmadığını ifade etmektedir. Genellikle bir skala üzerinde derecelendirilir; düşük valans (örneğin 0-2 arası) genellikle olumsuz veya hoşa gitmeyen hisleri ifade ederken, yüksek valans (örneğin 8-10 arası) hoşnutluk veya memnuniyeti ifade etmektedir. Orta aralık (örneğin 4-5) nötr duygularla ilişkilendirilir. Belirtilen uyarlama değeri, bir kullanıcının bireyde yarattığı fizyolojik uyanıklık derecesini belirtir. Düşük uyarılma sakinlik veya durgunluk hissini, yüksek uyarılma ise enerji, heyecan veya stres hissini temsil etmektedir. Sınıflandırma algoritmasında da belirtildiği üzere, valans düşük ve uyarılma yüksek ise, bu durum yüksek anksiyete seviyesine işaret etmektedir. Diyagram, farklı valans ve uyarılma kombinasyonlarına dayanarak katılımcıların anksiyete seviyelerini sınıflandırmak için kullanılmaktadır.



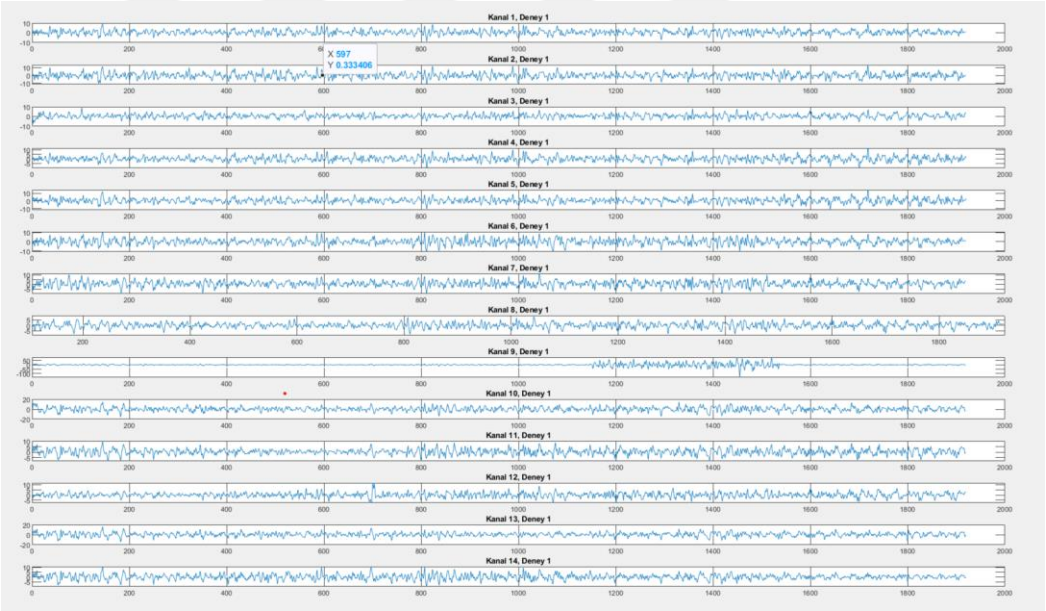
Şekil 3 Etiketleme Algoritması [10]

Katılımcı başına örnek sayısının arttırılması için, ana EEG sinyalinden 5 saniye denemesi için alt veri seti oluşturulmuştur. Bu süreçler, EEG verilerinden yüksek kaliteli ve temiz sinyaller elde etmek için kritik bir öneme sahiptir ve doğru sınıflandırma sonuçları için temel oluşturur.

3.1.3.1 Deney Seti EEG Sinyallerine Uygulanmış İşlemler

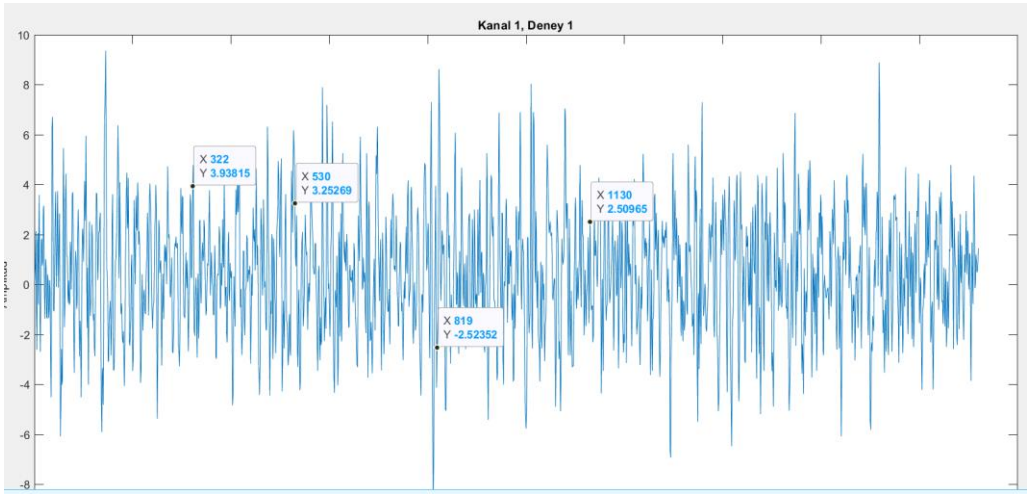
3.1.3.1.1 Boyut Azaltma

Veri setinden alınan örneklerde, her bir kullanıcı için oluşturulmuş EEG veri setinin boyutu 14x1920x12'dir. Veri setine uygulanacak istatistiksel işlemlerden dolayı, her bir kullanıcı için oluşturulmuş veri setlerine boyut indirgeme yapılması gerekmiştir. 14 rakamı ölçüm cihazı kanal sayısını, 1920 rakamı zaman noktasını, 12 rakamı ise valans ve uyarılma katsayıları için toplam ikili deneme sayısını temsil etmektedir. Şekil 4, 23. Katılımcı için her bir kanala ait verilerin grafiğini temsil etmektedir Şekil 5, 23. katılımcıya ait birinci deneyin birinci kanalını temsil etmektedir.



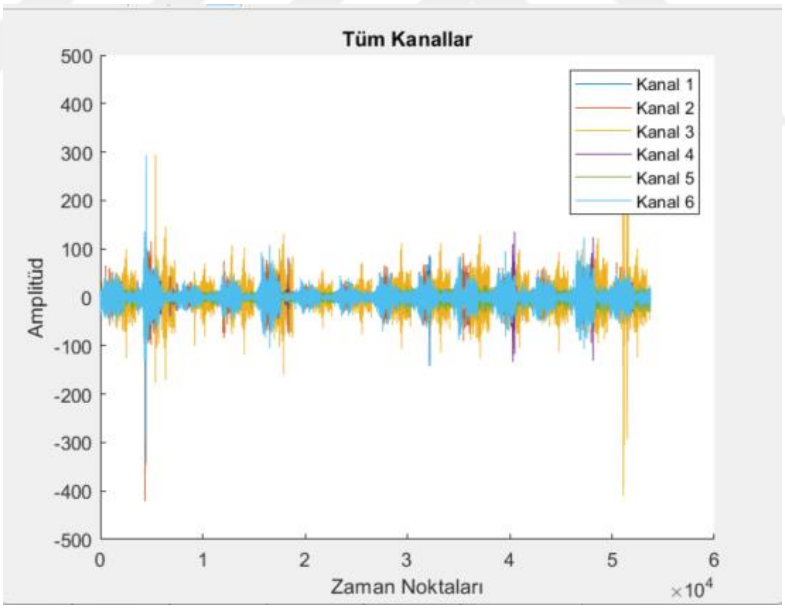
Şekil 4 Katılımcı 23 için her bir kanala ait verilerin grafiği

Veri seti içerisindeki her bir deneye ait 14 kanalın verileri, yatay olarak birleştirilir. İşlem tüm kanallar için gerçekleştirilir. Her saniyede kullanıcıdan 128 örnek alınmaktadır.



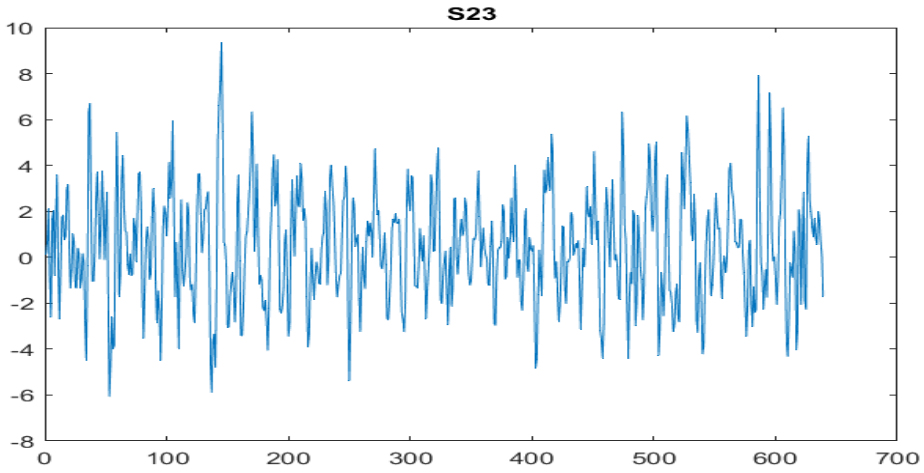
Şekil 5 Katılımcı 23, Deney 1, Kanal 1 verilerinin grafiği

Belirtilen işlem her bir kullanıcı için uygulandığında her bir kullanıcı için veri seti 6x53760 boyutlu matrise dönüştürülür. Şekil 6, 23. katılımcıya ait indirgenmiş verilerin grafiğini temsil etmektedir. İşlem tüm katılımcılar için gerçekleştirilmiştir.



Şekil 6 Katılımcı 23 EEG Verileri Boyut İndirgeme, yeni boyutun grafiği

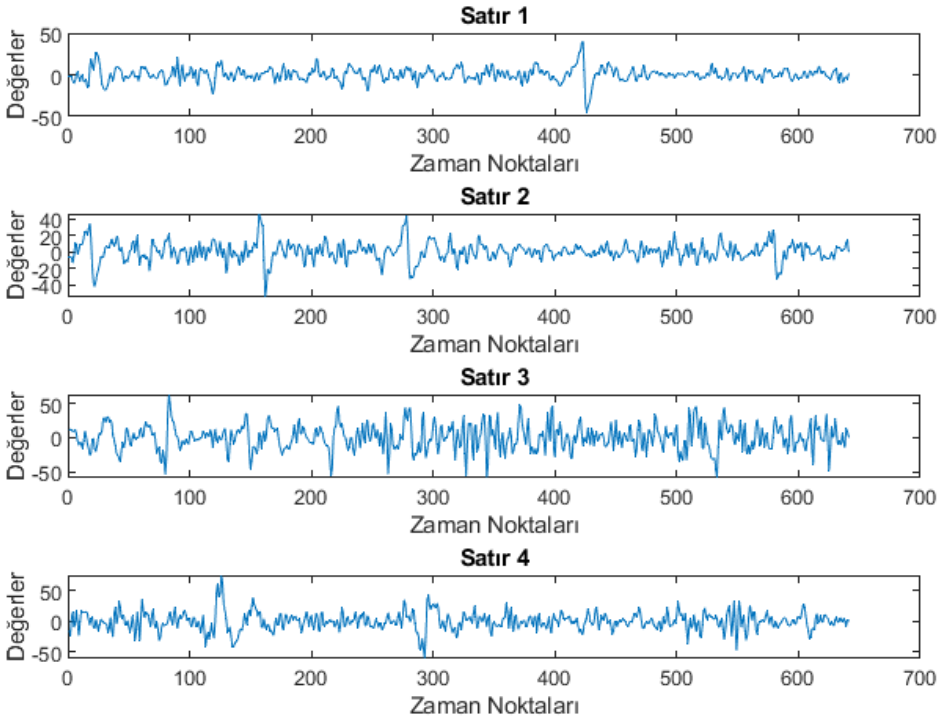
Veri setinde her deneme için 5'er saniyelik örnek seçimi yapılır. Hazırlanmış projede, her denemenin ilk 5 saniyesi seçilmiştir. Her bir kullanıcı için yeni bir matris oluşturulmuştur. Şekil 7, 5'er saniyelik örnek seçiminin grafiğini temsil etmektedir.



Şekil 7 Katılımcı 23, 5 saniyelik örnekleme grafiği

Her bir veri setinin son iki satırına ikili ve dörtlü sınıflandırmaya ait etiketler atanır ve veri seti kaydedilir.

Her bir veri seti birbiri ardına eklenerek, 5'er saniyelik kesitler için kombine edilmiş yeni bir veri seti oluşturulur. Şekil 8 kombinasyonu yapılmış veri setinin ilk dört satırına ait verilerinin grafiğini temsil etmektedir.



Şekil 8 Birleştirilmiş Veri Setinin Grafiği

Veri setindeki EEG sinyallerine uygulanmış istatistiksel öznitelik çıkarma işlemleri sırasıyla aşağıda belirtilmiştir.

3.1.3.1.2 Wavelet Dönüşümü

Düzenlenmiş veri seti için uygulanmış ilk işlem EEG Wavelet dönüşümüdür. Wavelet dönüşümü, EEG sinyallerinin zaman ve frekans özniteliklerini detaylıca analiz etmek, gürültünün azaltılması ve veriden anlamlı öznitelikler çıkarmak için kullanılmıştır.

3.1.3.1.3 Hilbert Dönüşümü

Kaygı durumları, beyin dalgalarının frekans bileşenlerinde değişikliklere neden olmaktadır. Hilbert dönüşümü, ilgili değişiklikleri zaman-frekans analizi yoluyla ortaya çıkarmak için kullanılmaktadır. Anlık genlik ve faz hesaplanarak Hilbert sınıflandırma modellerinde kullanılabilecek zengin öznitelik setleri üretilir.

3.1.3.1.4 Gürültü giderme

Hilbert dönüşümü sonucu hesaplanmış anlık genlik ve faz değerli veri seti için gürültü giderme işlemi yapılmıştır.

3.1.3.1.5 Hilbert – Huang Dönüşümü

Hilbert-Huang dönüşümü, zamanla değişen istatistiksel özniteliklere sahip ve doğrusal olmayan veri setlerini analiz etmek için geliştirilmiştir. Sinyalin zaman içinde nasıl değiştiğini anlamak, farklı frekans bileşenlerini ayırmak ve bu bileşenlerin beyin aktivitesiyle nasıl ilişkili olduğunu keşfetmek için kullanılmıştır.

13 adet istatistiksel öznitelikler ve hesaplamaları aşağıda belirtilmiştir.

3.1.3.1.6 Ortalama (Mean) ve Medyan

Ortalama değeri ve medyan değeri, sinyalin genel merkezi eğilimini temsil etmektedir. Medyan, sıralı bir veri setinde ortada bulunan değerdir. Veri seti sıralandıktan sonra, eğer N tek sayı ise ortadaki değer medyandır. N çift sayı ise, ortadaki iki değerin ortalaması medyandır.

$$\text{Ortalama} = \mu = \frac{1}{N}\sum_{i=1}^{N} x_i \qquad (3.1)$$

3.1.3.1.7 Dalga Formu Uzunluğu

Dalga formu uzunluğu, sinyalin şeklini ve kompleksitesini ölçer. Farklı beyin durumları dalga formlarında farklılıklar göstermektedir.

$$\text{Dalga Formu Uzunluğu} = \sum_{i=1}^{N-1} |x_{i+1} - x_i| \quad (3.2)$$

3.1.3.1.8 Ortalama Mutlak Sapma

Sinyalin ortalama değerinden ne kadar sapma gösterdiğini ölçer. Sinyalin kararlılığı veya oynaklığı ile ilgili bilgi vermektedir.

$$\text{OMS} = \frac{1}{N}\sum_{i=1}^{N} |x_i - \mu| \quad (3.3)$$

3.1.3.1.9 Enerji

Sinyalin genel gücünü veya enerjisini ölçmektedir. Yüksek enerji seviyeleri, yüksek beyin aktivitesini temsil edebilir.

$$\text{Enerji} = \sum_{i=1}^{N} x_i^2 \quad (3.4)$$

3.1.3.1.10 Standart Sapma

Sinyalin varyansının kareköküdür ve sinyalin ne kadar değişken olduğu hakkında bilgi verir.

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2} \quad (3.5)$$

3.1.3.1.11 Basıklık

Basıklık, veri dağılımının tepesinin ne kadar basık veya sivri olduğunu ifade eder.

$$\text{Basıklık} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i - \mu}{\sigma}\right)^4 - 3 \quad (3.6)$$

3.1.3.1.12 Çarpıklık

Çarpıklık, veri dağılımının simetrisini ölçer.

$$\text{Çarpıklık} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i - \mu}{\sigma}\right)^3 \quad (3.7)$$

3.1.3.1.13 Kök Ortalama Kare

Sinyalin genlik seviyesinin bir ölçüsüdür ve sinyalin gücünü ifade eder.

$$\text{KOK} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2} \quad (3.8)$$

3.1.3.1.14 Shannon Entropisi

Sinyalin bilgi içeriğinin ve düzensizliğinin bir ölçüsüdür. EEG sinyallerindeki entropi değişimleri ve farklı beyin durumlarıyla ilişkilendirilebilir.

$$H(X) = -\sum p(x_i)\log(p(x_i)) \quad (3.9)$$

3.1.3.1.15 Çapraz Entropi

İki olasılık dağılımı arasındaki benzerliği ölçmektedir. Sinyaller arasındaki ilişkilerin anlaşılmasında kullanılabilir. Anlık frekans değeri ve sınıf etiketinden tüketilmiş olasılık dağılımı arasında hesaplanmıştır.

$$H(p,q) = -\sum p(x_i)\log(q(x_i)) \quad (3.10)$$

3.1.3.1.16 Varyan

Veri noktalarının dağılımının ölçüsüdür.

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2 \quad (3.11)$$

3.1.3.1.17 Log Enerji Entropisi

Sinyalin enerji dağılımının logaritmik ölçüsüdür ve genellikle sinyalin karmaşıklığını ifade etmektedir. Denklemde belirtilen p_i i'nci zaman dilimindeki normalleştirilmiş enerji değeridir.

$$\text{Log Enerji Entropisi} = -\sum_{i=1}^{N} p_i \log p_i \tag{3.12}$$

Yukarıda belirtilen istatistiksel ve zaman-frekans öznitelikleri, EEG sinyallerinin analizi için hesaplanmıştır. 13 öznitelik seçilmiştir. EEG sinyalleri zamanla değişen istatistiksel özniteliklere sahiptir ve çok boyutludur. Her bir öznitelik, beynin elektriksel aktivitesinin farklı yönlerini yakalamaktadır. Her altı örneklemin işlenmesi sonucunda 13 farklı öznitelik hesaplanmıştır ve belirtilen işlem 6 kez tekrarlanarak 78 adet öznitelik elde edilmiştir. Hazırlanmış olan tez çalışmasında, ikili ve dörtlü sınıflandırma yapılarak sınıflandırma yöntemleri karşılaştırılmıştır.

3.1.4 EEG Sinyallerinin Sınıflandırılması

3.1.4.1 K-En Yakın Komşu Algoritması

KNN algoritması sınıflandırılacak verilerin, önceki veriler ile olan komşuluk ilişkisine göre sınıflandırmasını temel alır. KNN yeni noktaya en yakın noktayı bulmak için kullanılır. K adet bilinmeyen noktanın en yakın komşusunu temsil eder. KNN algoritmalarının en önemli avantajlarından biri, analitik olarak takip edilebilir olmalarıdır. Aynı zamanda yerel bilgilere uyarlanabilir ve tutarsız eğitim verilerine toleranslıdır. Dezavantajları olarak ise, büyük miktarda bellek kullanımına ihtiyaç duymaları ve verinin boyutu arttıkça, işlem yükünün artmasıdır. [28]

KNN algoritması çeşitli varyasyonlara sahiptir. Ağırlıklı K-En Yakın Komşular, Kosinüs K-En Yakın Komşular, İnce Ayarlı K-En Yakın Komşu algoritması bu varyasyonlardan birkaçıdır.

KNN algoritması yeni sınıflandırılacak olan örnek ile eğitim kümesindeki örneklerin uzaklıklarını ölçerek, en yakın k adet komşunun sınıf değerlerini kullanır ve sınıflandırma işlemini gerçekleştirir. Sınıflandırılacak veri ve eğitim verileri arasındaki mesafeyi ölçmek amacıyla genelde aşağıda belirtilmiş denklemlerde

sıralanmış olan Öklid (Denklem 3.15), Manhattan (Denklem 3.16) ve Minkowski (Denklem 3.17) uzaklık ölçüm formülleri kullanılır.

Aşağıdaki denklemlerde X sınıflandırılacak olan yeni veriyi, Y ise veri kümesindeki örneği temsil etmektedir. Veriler veri kümesi de öznitelik sayısı kadar eleman içeren vektör şeklinde tutulmaktadır. [28]

$$X = \{x_1 , x_2 , \dots . , x_n\} \tag{3.13}$$

$$Y = \{ y_1, y_2 , \dots . , y_n\} \tag{3.14}$$

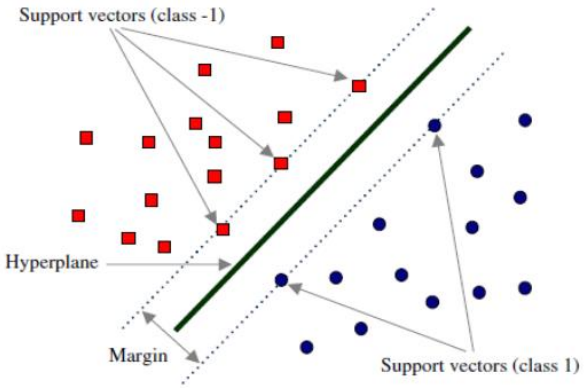
$$d(X,Y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \tag{3.15}$$

$$d(X,Y) = |x_i - y_i| \tag{3.16}$$

$$d(X,Y) = \sqrt{(\sum_{i=1}^{n}|x_i - y_i|^p)^{1/p}} \tag{3.17}$$

3.1.4.2 Destek Vektör Makinesi (DVM)

DVM sınıflandırma problemleri için kullanılması uygun olan bir makine öğrenmesi modelidir. DVM, yeni veri noktasının kolayca tespitini sağlayabilmek için en iyi karar sınırını oluşturmaktadır. En iyi karar sınırı hiper düzlem olarak adlandırılır. DVM hiper düzlemin kurulumuna yardımcı olan vektörleri seçer ve bu vektörleri 'Destek Vektör Makinesi' olarak isimlendirmektedir. [29]



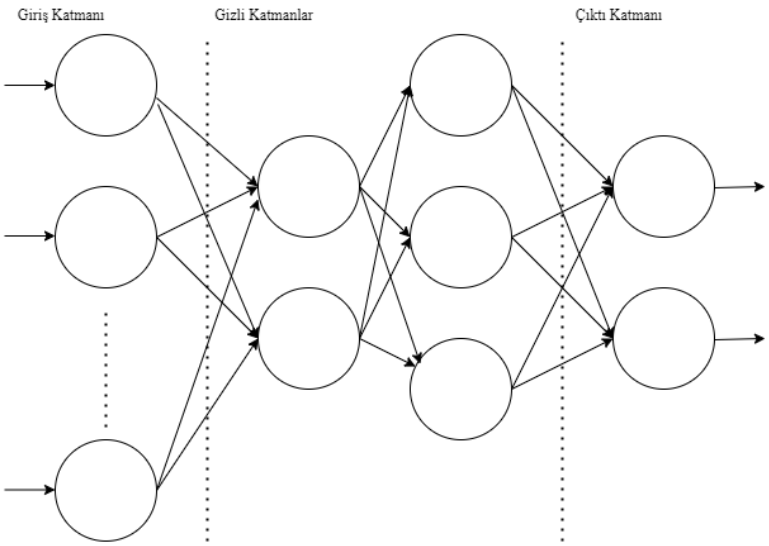
Şekil 9 Destek Vektör Makinesi Küme Ayrışımı *[29]*

Destek vektör makineleri çeşitli varyasyonlara sahiptir. Bunların arasında Gauss Destek Vektör Makinesi, DVM Çekirdek, Kübik DVM, Karesel DVM, Doğrusal DVM örnek gösterilebilir. Ayrıca, her varyasyon da kendi içinde sınıflandırılabilir. Destek Vektör Makineleri matematiksel olarak $K(x_i\ ,\ x_j) = \phi\ (\mathrm{x}) \cdot \phi(x_j)$ şeklinde ifade edilen bir kernel fonksiyonu yardımıyla doğrusal olmayan dönüşümler yapabilmektedir. Bu yöntem ile verilerin yüksek boyutta doğrusal olarak ayrımına olanak sağlar. Doğrusal olarak ayrılamayan iki sınıflı bir problemin çözümü ile ilgili matematiksel işlem aşağıdaki gibi yazılabilir. [30]

$$f(x) = sign(\sum_i a_i y_i \phi(\mathrm{x}) \cdot \phi(x_i) + b) \tag{3.18}$$

3.1.4.3 Yapay Sinir Ağları

Yapay sinir ağları, insan beyninin bilgi işleme şeklini taklit etmek için geliştirilmiş bir yapay zekâ modelidir. Yapay sinir ağları, basit yapay 'nöron' adı verilen birimlerden oluşmaktadır. Gerçek bir nöronun özelliklerini modeller ve genellikle bir aktivasyon fonksiyonu ile ilişkilendirilen matematiksel bir nörondur. Yapay nöronlari ağırlıklar ile birbirine bağlanır ve bu apırlıklar öğrenme süreci sırasında ayarlanır. Bir YSA, genellikle üç tür katmandan oluşmaktadır. Giriş katmanı, bir veya daha fazla gizli katman ve bir çıkış katmanı. Her katmandaki nöronlar, önceki katmandaki nöronların çıktılarını alır, ağırlıklarla çarpılır ve bir aktivasyon fonksiyonu tarafından işlenir.



Şekil 10 Yapay Sinir Ağları

Yapay Sinir Ağları çeşitli varyasyonlara sahiptir. Bunların arasında Geniş Yapay Sinir Ağı, İki Katmanlı Yapay Sinir Ağı, Üç Katmanlı Yapay Sinir Ağı örnek gösterilebilir.

4 ARAŞTIRMA BULGULARI VE TARTIŞMA

Yapılan çalışmada, ikili ve dörtlü sınıflı veri seti için makine öğrenme yöntemleriyle sınıflandırma sonuçları araştırılmıştır. İkili sınıflandırma sürecinde veri seti, "anksiyetik" veya "değil" olarak; dörtlü sınıflandırmada ise "ciddi anksiyete", "orta anksiyete", "düşük anksiyete" ve "anksiyete yok" olarak sınıflandırılmıştır.

Araştırmanın ilk bölümünde ikili sınıflandırmaya ait veriler sunulmuştur.

Veri seti %70 eğitim seti ve %30 test seti olarak ayrılmıştır. Ayrılan eğitim ve test setleri öğrenme ve test aşamaları için ayrı ayrı Matlab programının Classification Learner aracına yüklenerek K-En Yakın Komşu, Yapay Sinir Ağları, Destek Vektör Makinesi algoritmalarında eğitilmiştir. Aşağıda belirtilen tablolarda öğrenme ve test algoritmalarının analizinin sonuçları verilmiştir

Hesaplamalarda kullanılan TP değeri *Gerçek Pozitif* değerini temsil etmektedir. Bu terim gerçekte de pozitif olan durumları ifade etmektedir. Yani modelin, bir durumu doğru bir şekilde pozitif olarak sınıflandırdığı durum sayısını göstermektedir. TN değeri *Gerçek Negatif* değerini temsil etmektedir. Bu terim modelin negatif olarak tahmin ettiği ve gerçekte de negatif olan durumları belirtmektedir. Başka bir deyişle, modelin bir durumu doğru bir şekilde negatif olarak sınıflandırdığı durum sayısını ifade etmektedir. FP değeri *Yanlış Pozitif* değerini temsil etmektedir. Bu terim modelin pozitif olarak tahmin ettiği ancak gerçekte negatif olan durumları ifade etmektedir. Bu durum bir nevi 'yanlış alarm' olarak da adlandırılabilir. FN değeri *Yanlış Negatif* değerini temsil etmektedir. Bu terim modelin negatif olarak tahmin ettiği ancak gerçekte pozitif olan durumları ifade etmektedir. Belirtilen dört temel kavram, bir sınıflandırma modelinin performansını değerlendirmede kritik öneme sahiptir. TP, TN, FP, FN değerleri üzerinden hesaplanan *duyarlılık, özgüllük, kesinlik* ve *F1 skoru* gibi metrikler modelin ne kadar iyi performans gösterdiğini anlamamıza yardımcı olur.

Öğrenme algoritmalarının model performansını değerlendirebilmek için duyarlılık, özgüllük, kesinlik ve F1 skor değerleri hesaplanmıştır.

Duyarlılık değeri pozitif sınıfın ne kadar iyi tespit edilebildiğini göstermektedir.

$$\text{Duyarlılık} = \frac{\text{TP}}{\text{TP+FN}} \tag{4.1}$$

Özgüllük değeri, negatif sınıfın ne kadar iyi tespit edilebildiğini göstermektedir.

$$\text{Özgüllük} = \frac{\text{TN}}{\text{TN+FP}} \tag{4.2}$$

Kesinlik değeri, pozitif olarak tahmin edilen örneklerin ne kadarının gerçekten pozitif olduğunu göstermektedir.

$$\text{Kesinlik} = \frac{\text{TP}}{\text{TP+FP}} \tag{4.3}$$

F1 skoru değeri, duyarlılık ve kesinlik değerleri arasındaki dengeyi yakalamak için kullanılmaktadır.

$$\text{F1 skoru} = 2 \text{ x } \frac{\text{Duyarlılık x Kesinlik}}{\text{Duyarlılık+Kesinlik}} \tag{4.4}$$

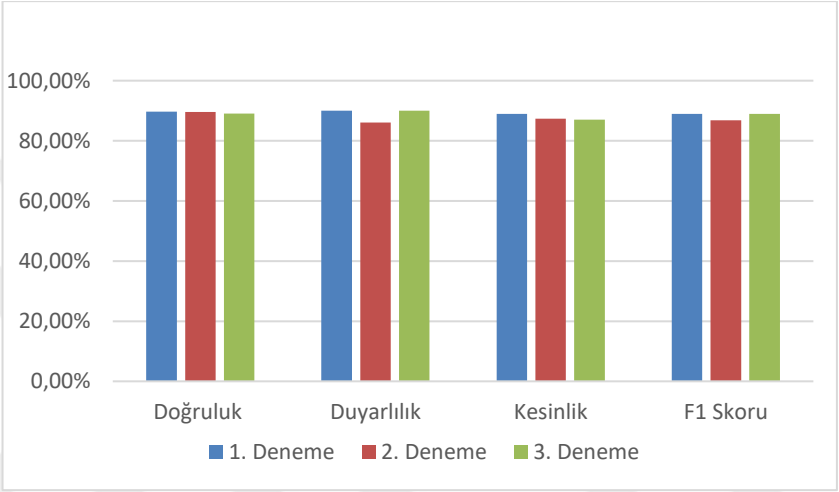
4.1 İkili Sınıflandırma Öğrenme ve Test Sonuçlarının Analizi

4.1.1 İkili Sınıflandırma Yöntemi için K-En Yakın Komşu Algoritması Öğrenme ve Test Sonuçları Analizi

K-En Yakın Komşu algoritması, sınıflandırma ve regresyon problemlerinde yaygın olarak kullanılan bir algoritmadır. Komşuluk sayısı ya da ‘k’ parametresi, KNN’nin karar verme sürecinde ne kadar komşusunu dikkate alacağını belirler. Belirli bir ‘k’ değeri, modelin genelleme ve aşırı öğrenme yeteneği üzerinde büyük bir etkiye sahiptir. Yüksek bir ‘k’ değeri genellikle daha düşük varyans ancak yüksek yanlılık ile sonuçlanırken düşük bir ‘k’ değeri yüksek varyans ve düşük yanlılık eğilimi göstermektedir.

Tablo 3 K-En Yakın Komşu Algoritması ile İkili Sınıflandırma için Öğrenme Sonuçları

Sıralama	Öğrenme Algoritması	Doğruluk (%)	Duyarlılık (%)	Özgüllük (%)	Kesinlik (%)	F1 Skoru (%)
1.Deneme	KNN	89.70	89.79	89.67	89.11	89.45
2.Deneme	KNN	89.60	86.14	87.83	87.44	86.79
3.Deneme	KNN	89.10	90.16	88.22	87.29	89.70

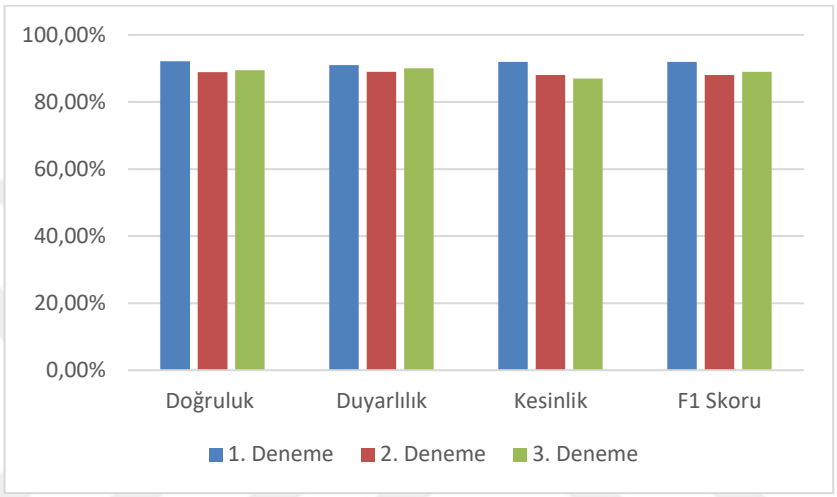


Şekil 11 K-En Yakın Komşu Algoritması ile İkili Sınıflandırma için Öğrenme Grafiği

KNN algoritması kullanılarak yapılan ikili sınıflandırma çalışmasında, farklı 'k' değerlerinin öğrenme sonuçları üzerindeki etkisi incelenmiştir. Tablo 3'te sunulan öğrenme sonuçları, algoritmanın farklı 'k' değerleri için performansını göstermektedir. Tablo 3'te, birinci denemede 'k' değeri 2 olarak seçilmiş ve birinci deneme %89.7 doğruluk oranı ile en yüksek performansı sergilemiştir. İkinci denemede 'k' değeri 10 olarak seçilmiş ve %89.6 doğruluk oranı hesaplanmıştır. Üçüncü denemede ise 'k' değeri 7 olarak seçilmiş ve %89.1 doğruluk oranı hesaplanmıştır. Her deneme için, komşuluk sayısı özelliği dışında tüm özellikler aynı tutulmuştur. İkinci ve üçüncü denemeler, benzer doğruluk oranlarına sahip olsalar dahi, duyarlılık ve kesinlik değerlerinde düşüş meydana gelmektedir. Bu durum 'k' değerinin artmasının modelin genelleme yeteneğini arttırdığını ancak aynı zamanda belirli sınıflar için duyarlılığı ve kesinliği azaltabileceğini göstermektedir. Şekil 11'de gösterilen öğrenme grafikleri, 'k' değerlerinin, 2,10 ve 7 olduğu durumlarda algoritmanın performansını görsel olarak sunmaktadır. Grafikler, farklı 'k'

Tablo 4 K-En Yakın Komşu Algoritması ile İkili Sınıflandırma için Test Sonuçları

Sıralama	Öğrenme Algoritması	Doğruluk (%)	Duyarlılık (%)	Özgüllük (%)	Kesinlik (%)	F1 Skoru (%)
1.Deneme	KNN	92.17	91.10	92.95	92.42	91.76
2.Deneme	KNN	88.9	89.01	88.89	87.73	88.36
3.Deneme	KNN	89.5	90.30	88.75	87.36	88.31



Şekil 12 K-En Yakın Komşu Algoritması ile İkili Sınıflandırma için Test Grafiği

değerlerinin algoritmanın doğruluk, duyarlılık, kesinlik ve F1 skoru ölçütler üzerindeki etkisini detaylı bir şekilde ortaya koymaktadır.

KNN algoritması kullanılarak gerçekleştirilen ikili sınıflandırma test sonuçları, algoritmanın farklı 'k' değerlerine göre nasıl performans gösterdiğini detaylı bir şekilde göstermektedir. Tablo 4'te sunulan test sonuçları, modelin doğruluk, duyarlılık, özgüllük, kesinlik ve F1 skoru ölçütleri üzerindeki performansını özetlemektedir.

'k' değerinin 2 seçildiği birinci denemenin %92.17 gibi oldukça yüksek bir doğruluk değeri verdiği Tablo 4'te görülmektedir. 'k' değerinin 10 ve 7 olarak belirlendiği ikinci ve üçüncü denemelerde ise, yüksek 'k' değerlerinin doğruluk oranlarında bir düşüşe neden olduğu gözlemlenmektedir. Elde edilen sonuçlar, 'k' değerinin model performansı üzerinde önemli bir etkisi olduğunu göstermektedir. Şekil 12'de görselleştirilen test grafikleri, 'k' değerinin 2, 10 ve 7 olduğu durumlar

için elde edilen sonuçları karşılaştırmalı olarak sunmaktadır. Grafikler, farklı ‘k’ değerlerinin modelin çeşitli performans ölçütleri üzerindeki etkisini anlaşılır bir şekilde göstermektedir.

Belirtilen bulgular, model duyarlılığı ve özgüllüğü arasındaki dengeyi koruma konusunda komşu sayısının kritik bir rol oynadığını göstermektedir. Tablo 4 ve Şekil 12 üzerinden yapılan analizler, en düşük komşu sayısına sahip modelin her iki tabloda da diğer modellerden daha iyi sonuç verdiği gözlemlenmiştir. Düşük ‘k’ değerlerinin belirli veri setleri ve problemler için daha iyi sonuçlar verebileceği çıkarımı yapılabilir. Ancak bu durum genel bir kural olarak kabul edilemez. Çünkü düşük ‘k’ değeri bazı durumlarda aşırı öğrenmeye yol açabilir. İdeal ‘k’ değerinin bulunabilmesi için, çapraz doğrulama gibi yöntemler kullanılarak modelin performansının doğrulanması önemlidir.

4.1.2 İkili Sınıflandırma Yöntemi için Yapay Sinir Ağları Algoritması Öğrenme ve Test Sonuçları Analizi

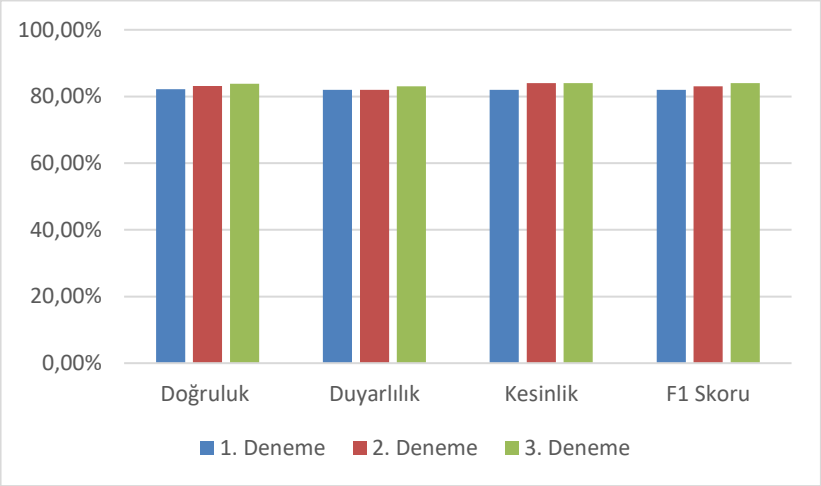
Yapay Sinir Ağlarının temel yapı taşlarından biri olan katman seçimi, model performansı üzerinde kritik bir rol oynamaktadır. Katman seçimi, modelin öğrenme kapasitesini, genelleme yeteneğini ve eğitim süresini doğrudan etkileyebilir. Bu nedenle, YSA tasarımında katman sayısı ve türünün doğru bir şekilde seçilmesi büyük bir önem taşımaktadır.

YSA, giriş, gizli ve çıkış katmanları olmak üzere üç temel katmandan oluşmaktadır. Giriş katmanı, seçilen öznitelikler ile modele veri sağlar. Gizli katmanlar, giriş verilerini işleyerek öznitelik çıkarımı yapmaktadır. Çıkış katmanı ise ikili sınıflandırma problemi için sonuçları üretir.

İkili sınıflandırmada için üç gizli katmanlı bir YSA kullanılırken, modelin doğru tahminler yapabilmesi için yeterli öğrenme kapasitesine sahip olması gerekmektedir. Bu kapasite, gizli katmanlardaki nöron sayısı ve katman sayısı ile doğrudan ilişkilidir. Ancak, çok fazla nöron veya katman kullanmak, modelin aşırı uyum sağlayıp genel veri setinde düşük performans göstermesine neden olabilir. Bu nedenle, ikili sınıflandırma probleminde katman seçimi yapılırken, modelin yeterince

Tablo 5 Yapay Sinir Ağları Algoritması ile İkili Sınıflandırma için Öğrenme Sonuçları

Sıralama	Öğrenme Algoritması	Doğruluk (%)	Duyarlılık (%)	Özgüllük (%)	Kesinlik (%)	F1 Skoru (%)
1.Deneme	YSA	82.2	81.91	82.45	81.54	81.73
2.Deneme	YSA	83.1	81.98	84.32	83.96	82.96
3.Deneme	YSA	83.8	82.99	84.63	84.11	83.55



Şekil 13 Yapay Sinir Ağları Algoritması ile İkili Sınıflandırma için Öğrenme Grafiği

genelleyebilmesi ve eğitim verilerine aşırı uyum sağlamaktan kaçınabilmesi için dikkatli bir dengenin kurulması gerekmektedir

Üç gizli katmanlı YSA tasarımında, katmanların ve nöronların sayısı önemli bir rol oynar. Katman ve nöron sayısının belirlenmesi genellikle deneme yanılma ile yapılmakta ve çapraz doğrulama teknikleri ile optimize edilmektedir.

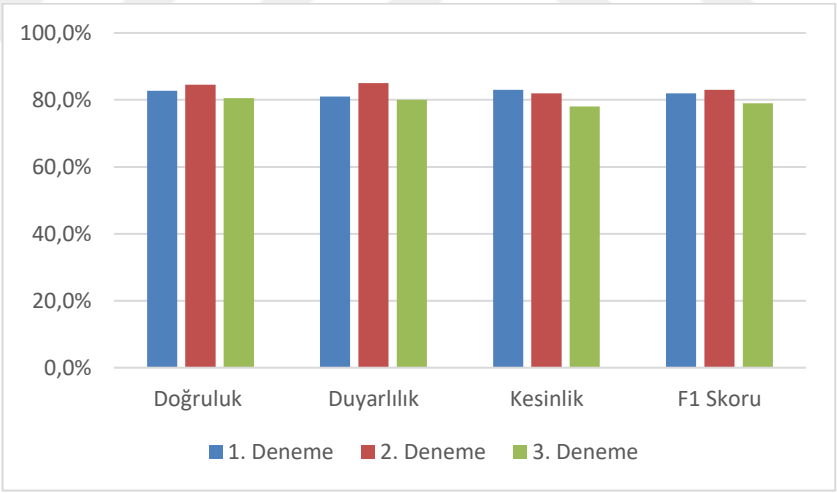
Tablo 5'te yer alan Yapay Sinir Ağları ile ikili sınıflandırma için Öğrenme sonuçları, üç farklı denemeyle modelin performansını özetler. Belirtilen denemelerde, üç farklı YSA denemesinin tamamında üç gizli katman seçilmiş, katman sayısı sabit tutulmuştur. Birinci denemede, her katman için 100 nöron kullanılmıştır.

İkinci ve üçüncü denemelerde, her katman için 1000 nöron kullanılmıştır.. Üçüncü modelin daha yüksek özgüllük ve duyarlılık göstermesi, kullanılan tüm özniteliklerin modelin ayrım gücünü arttırdığına işaret etmektedir. Bu durum, verinin zenginliğinin ve çeşitliliğinin, YSA modellerinin performansını olumlu yönde etkileyebileceğini gösterebilir.

Şekil 13'teki öğrenme grafiği, YSA'nın ikili sınıflandırma performansını görsel olarak sunmaktadır. Belirtilen grafik, farklı denemelerdeki model performansının karşılaştırılmasını sağlar ve öğrenme sürecindeki ilerlemeyi açıkca göstermektedir.

Tablo 6 Yapay Sinir Ağları Algoritması ile İkili Sınıflandırma için Test Sonuçları

Sıralama	**Öğrenme Algoritması**	**Doğruluk (%)**	**Duyarlılık (%)**	**Özgüllük (%)**	**Kesinlik (%)**	**F1 Skoru (%)**
1.Deneme	YSA	82.7	81.49	83.89	82.67	82.08
2.Deneme	YSA	84.5	85.28	83.76	81.59	83.39
3.Deneme	YSA	80.5	80.37	80.58	78.34	79.34



Şekil 14 Yapay Sinir Ağları Algoritması İkili Sınıflandırma için Test Grafiği

Yapay Sinir Ağları kullanılarak gerçekleştirilen ikili sınıflandırma test sonuçları, Tablo 6 ile sunulmuştur. Belirtilen tabloda, farklı YSA modellerinin doğruluk, duyarlılık, özgüllük, kesinlik ve F1 skoru performans metrikleri üzerinden değerlendirme yapılmıştır. İkinci deneme %84.5 doğruluk oranı ile dikkat çekici bir performans sergilerken, duyarlılık ve kesinlik açısından da güçlü sonuçlar elde

etmiştir. Bu, söz konusu modelin dengeli bir genelleme yeteneğine sahip olduğunu göstermektedir. Bu duruma karşın, birinci ve üçüncü modeller farklı metriklerde değişken performanslar sergilemiş, üçüncü model ise tüm metriklerde daha düşük sonuçlar vermiştir. Bu durum da aşırı uyum veya yetersiz genelleme yeteneğine işaret edebilir.

Şekil 14, YSA ile yapılan ikili sınıflandırma modellerinin test sonuçlarını grafik üzerinde karşılaştırmalı olarak sunar. Bu görselleştirme, her bir modelin farklı metriklerdeki performansını daha net bir şekilde ortaya koymaktadır.

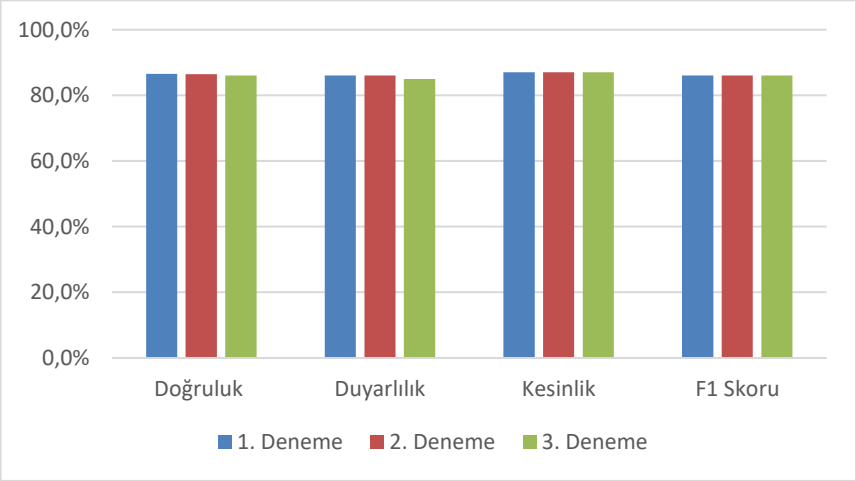
YSA'nın başarılı bir şekilde ikili sınıflandırma yapabilmesi için modelin doğru şekilde yapılandırılması gerekmektedir. Bu durum, yalnızca doğru nöron sayısını belirlemekle kalmaz, aynı zamanda modelin eğitim sürecinde kullanılacak özniteliklerinde dikkatlice seçilmesini gerektirir. Belirtilen yaklaşım, modelin hem eğitim hem de test verilerinde yüksek ve dengeli performans sergilelemesini sağlamak için kritik bir öneme sahiptir.

4.1.3 İkili Sınıflandırma Yöntemi için Destek Vektör Makinesi Algoritması Öğrenme ve Test Sonuçları Analizi

Destek Vektör Makinesi, ikili sınıflandırma problemlerinin çözümünde yaygın olarak kullanılan güçlü ve etkili bir makine öğrenmesi algoritmasıdır. Temel olarak, iki sınıfı en iyi şekilde ayırabilen hiper-düzlemi belirlemeyi amaçlamaktadır. Bu amaca ulaşmak için, sınıflar arasındaki marjı maksimize ederek, en net ayrımı sağlamaya çalışır.

Tablo 7 Destek Vektör Makinesi Algoritması ile İkili Sınıflandırma için Öğrenme Sonuçları

Sıralama	Öğrenme Algoritması	Doğruluk (%)	Duyarlılık (%)	Özgüllük (%)	Kesinlik (%)	F1 Skoru (%)
1.Deneme	DVM	86.5	85.59	87.50	87.14	86.36
2.Deneme	DVM	86.4	85.76	87.03	86.54	86.14
3.Deneme	DVM	86.0	84.99	86.91	86.54	85.76



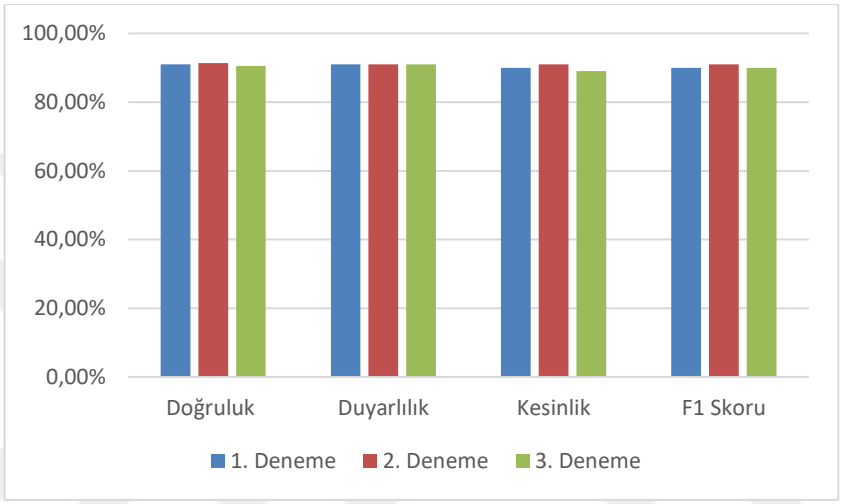
Şekil 15 Destek Vektör Makinesi Algoritması ile İkili Sınıflandırma için Öğrenme Grafiği

Destek Vektör Makinesi kullanılarak yapılan ikili sınıflandırma çalışmasının sonuçları Tablo 7'de sunulmuştur. Belirtilen tabloda, üç farklı DVM denemesinin doğruluk, duyarlılık, özgüllük, kesinlik ve F1 skoru performans metriklerine yer verilmiştir. Şekil 15'de ise, belirtilen öğrenme sonuçlarının grafiksel olarak gösterimi sağlanmış, böylece denemelerin performanslarını karşılaştırmak amaçlanmıştır.

Her üç deneme için de Gausyen kernel fonksiyonu kullanılmıştır. Gausyen kernel'in tercih edilmesinin nedeni, polinom veya kuadratik gibi diğer kernel yöntemlerine kuyasla daha yüksek doğruluk değerleri sunmasıdır. İkili sınıflandırma problemlerinde, modelin performansını maksimize etmek esastır. DVM algoritması ile yapılan denemelerde, Guasyen kernel kullanıldığında elde edilen sonuçlar, polinom veya kuadratik kernel kullanıldığında elde edilen sonuçlara göre daha üstün olmuştur. Polinom ve kuadratik kernel'ler veri setinin özniteliklerine ve dağılımına bağlı olarak bazı durumlarda yeterli ayrımı sağlayamayabilir.

Tablo 8 Destek Vektör Makinesi Algoritması ile İkili Sınıflandırma için Test Sonuçları

Sıralama	Öğrenme Algoritması	Doğruluk (%)	Duyarlılık (%)	Özgüllük (%)	Kesinlik (%)	F1 Skoru (%)
1.Deneme	DVM	90.70	90.84	90.52	89.53	90.18
2.Deneme	DVM	91.40	91.27	91.45	90.61	90.94
3.Deneme	DVM	90.50	90.81	90.23	89.17	89.98



Şekil 16 Destek Vektör Makinesi Algoritması ile İkili Sınıflandırma için Test Grafiği

Bu durum doğruluk değerinde düşüşe sebep olmuştur. Guasyen kernel'in, veri üzerinde yarattığı yüksek boyutlu uzay, modelin veri setindeki karmaşık yapıları daha iyi öğrenmesine ve böylece genel sınıflandırma performansını artırmasına olanak tanımaktadır.

Destek Vektör Makinesi kullanılarak gerçekleştirilen ikili sınıflandırma testlerinin sonuçları Tablo 8'de detaylandırılmıştır. Bu sonuçlar, DVM algoritmasının test verileri üzerinde nasıl bir performans sergilediğini ortaya koymaktadır.

Üç farklı DVM denemesinin test sonuçları, her bir modelin sınıflandırma probleminde benzer ve yüksek düzeyde başarı gösterdiğini göstermektedir. İkinci model %91.4 doğruluk oranıyla en yüksek performansı sergilerken, üçüncü model biraz daha düşük performansla öne çıkmaktadır.

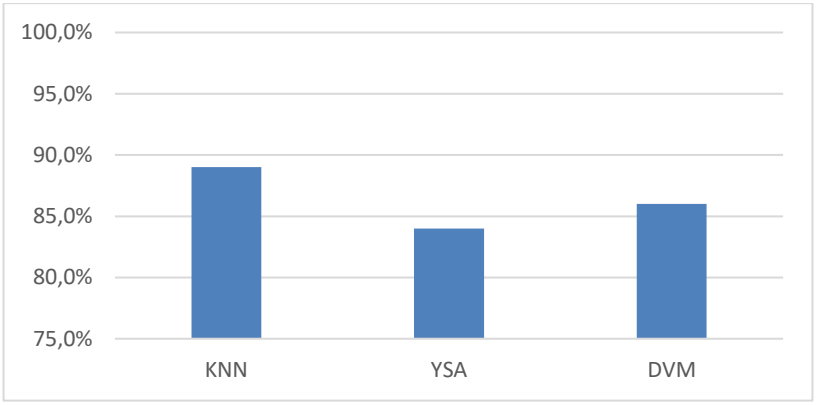
Şekil 16, DVM ile ikili sınıflandırma için elde edilen test sonuçlarının grafiksel sunumunu içerir. Belirtilen görsel, farklı modellerin performansını karşılaştırmalı olarak sunarak, algoritmanın test verileri üzerindeki etkinliğini vurgular.

Tablo 7 ve Tablo 8'in karşılaştırmalı analizi, ikinci modelin genel olarak diğerlerinden daha iyi bir performans gösterdiğini ortaya koymaktadır. Bu durum kullanlan öznitelik seçiminin, kernel fonksiyonunun ve parametre ayarlarının, DVM'nin performansını optimize etmek için önemini gösterir. Aynı zamanda, tüm modellerin performanslarının oldukça benzer sonuç vermesi, DVM algoritmasının kararlı yapısının bir göstergesi olarak değerlendirilebilir.

İkili Sınıflama öğrenme fazında algoritmaların en yüksek F1 skorları Tablo 9'da belirtilmiştir. Öğrenme aşamasında KNN en yüksek F1 skoruna sahip olup, bu metrik üzerinden diğer iki modele göre daha iyi performans sergilediğini göstermektedir. YSA'nın F1 skoru %84 ile en düşük performansı sergilerken, DVM'nin skoru %86 ile KNN'in ardından ikinci en yüksek skordur. Şekil 17'de F1 skoruna ait grafiği de bu sonuçları görselleştirmekte, KNN'nin diğer iki modele göre bariz bir üstünlük sağladığını vurgulamaktadır. Şekil 17, ikili sınıflandırma için elde edilen öğrenme sonuçlarının F1 skoru değerlerinin grafiksel sunumunu içerir. Belirtilen görsel, farklı modellerin F1 skoru performansını karşılaştırmalı olarak sunarak, algoritmanın test verileri üzerindeki etkinliğini vurgular.

Tablo 9 İkili Sınıflandırma için Tüm Modellerin Öğrenme Verileri F1 Skoru Karşılaştırması

Sıralama	**F1 Skoru (%)**
KNN	88.70
YSA	83.55
DVM	85.76



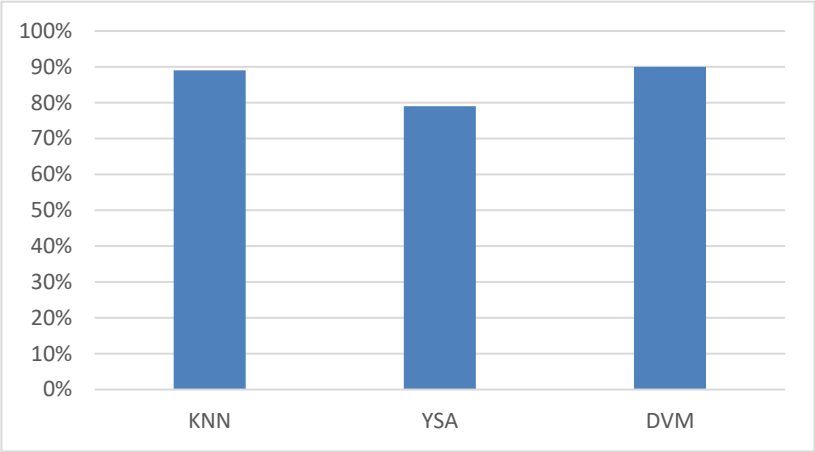
Şekil 17 İkili Sınıflandırma için Tüm Modellerin Öğrenme Verileri F1 Skoru Karşılaştırma Grafiği

Tablo 10'da belirtildiği üzere, test aşamasında, DVM en yüksek F1 skorunu elde etmiş, bu durum DVM'nin genelleyici kapasitesinin yüksek olduğunu işaret edebilir. KNN'nin performansı öğrenme ve test fazında sabit kalmıştır. Bu durum algoritmanın tutarlı bir performans sergilediğini gösterir. Ancak, YSA'nın F1 skoru test fazında önemli ölçüde düşmüş ve bu durum YSA modelinin öğrenme fazında aşırı öğrenme yapmış olabileceğine veya test verisine iyi genelleyemediğine işaret edebilir.

Şekil 18, ikili sınıflandırma için elde edilen test sonuçlarının F1 skoru değerlerinin grafiksel sunumunu içerir. Belirtilen görsel, farklı modellerin F1 skoru performansını karşılaştırmalı olarak sunarak, algoritmanın test verileri üzerindeki etkinliğini vurgular.

Tablo 10 İkili Sınıflandırma için Tüm Modellerin Test Verileri F1 Skoru Karşılaştırması

Sıralama	**F1 Skoru (%)**
KNN	88.31
YSA	79.34
DVM	89.98



Şekil 18 İkili Sınıflandırma için Tüm Modellerin Test Verileri F1 Skoru Karşılaştırma Grafiği

İkili sınıflandırma görevinde, öğrenme ve test aşamaları için seçilen modeller göz önünde bulundurulduğunda, DVM'nin test aşamasında en iyi performansı gösterdiği ve en yüksek genelleyici kabiliyete sahip olduğu söylenebilir. KNN'nin performansı her iki fazda da tutarlı kalmıştır. Bu durum algoritmanın güvenilir ve dengeli bir sınıflandırma yeteneğine sahip olduğunu göstermektedir. Öte yandan, YSA'nın öğrenme fazına kıyasla test fazında performansının düşmesi, modelin yapılandırmasının ve parametre ayarlarının gözden geçirilmesi gerektiğini öne sürmektedir.

4.2 Dörtlü Sınıflandırma Öğrenme ve Test Sonuçlarının Analizi

Araştırmanın ikinci bölümünde "ciddi anksiyete", "orta anksiyete", "düşük anksiyete" ve "anksiyete yok" tan oluşan dörtlü sınıflandırmaya ait veriler sunulmuştur.

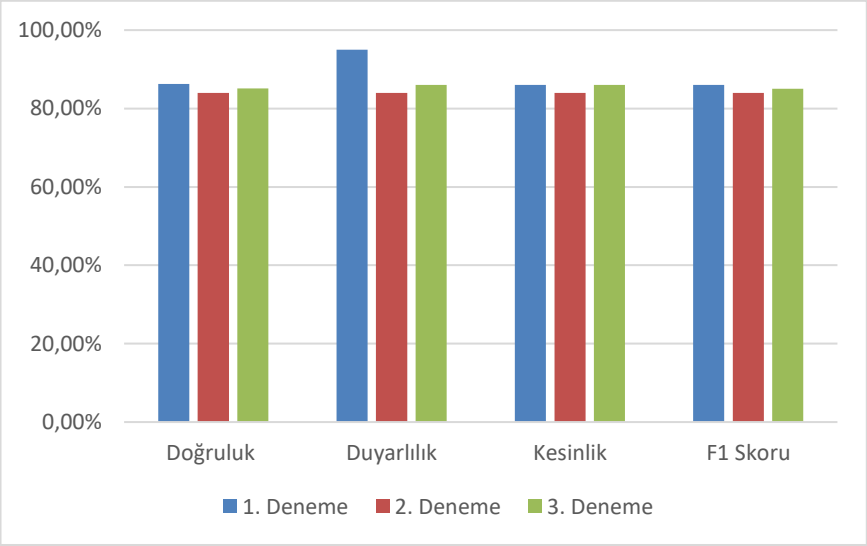
Veri seti %70 eğitim seti ve %30 test seti olarak ayrılmıştır. Ayrılan test setinin Matlab programının Classification Learner aracına yüklenmesi sonucu veri seti K-En Yakın Komşu, Yapay Sinir Ağları, Destek Vektör Makinesi algoritmalarında test edilmiştir.

4.2.1 Dörtlü Sınıflandırma Yöntemi için K-En Yakın Komşu Algoritması Öğrenme ve Test Sonuçları Analizi

Tablo 11'de, birinci algoritmada komşuluk sayısı 2, ikinci algoritmada komşu sayısı 10, üçüncü algoritmada komşuluk sayısı 7 olarak belirlenmiştir. Tablo 3'te de gözlemlendiği üzere, komşuluk sayısının artması doğruluk değerinde düşüşe sebep olmaktadır. Birinci model doğruluk ve duyarlılıkta en yüksek sonuçları vermektedir. Bu durum, pozitif sınıfın tespiti konusunda birinci algoritmanın güçlü bir performans sergilediğini ifade etmektedir. İkinci ve üçüncü modellerdeki özgüllük değerlerinin yüksek sonuç vermesi, modellerin negatif sınıfı iyi bir şekilde tespit edebildiğini göstermektedir. Ancak F1 skorundaki düşüş, genel performansın biraz daha düşük olduğunu gösterir.

Tablo 11 K-En Yakın Komşu Algoritması ile Dörtlü Sınıflandırma için Öğrenme Sonuçları

Sıralama	Öğrenme Algoritması	Doğruluk (%)	Duyarlılık (%)	Özgüllük (%)	Kesinlik (%)	F1 Skoru (%)
1.Deneme	KNN	86.3	95.4	86.3	86.3	86.3
2.Deneme	KNN	84.0	84.0	94.7	84.0	84.0
3.Deneme	KNN	85.1	85.7	95	85.7	85.1



Şekil 19 K-En Yakın Komşu Algoritması ile Dörtlü Sınıflandırma için Öğrenme Grafiği

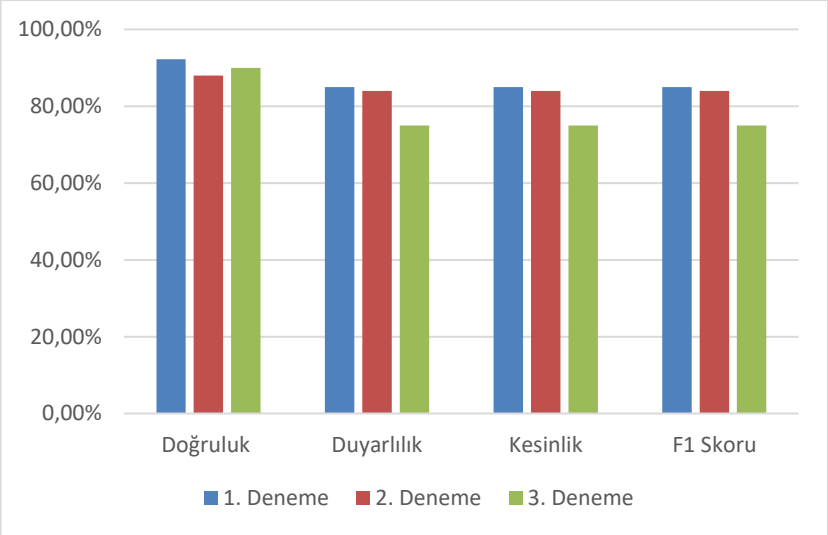
Şekil 19, KNN ile dörtlü sınıflandırma için elde edilen öğrenme sonuçlarının grafiksel sunumunu içerir. Belirtilen görsel, farklı modellerin performansını karşılaştırmalı olarak sunarak, algoritmanın öğrenme verileri üzerindeki etkinliğini vurgular. .

Tablo 12'de ilk modelin genel olarak test veri setinde en güçlü performansı sergilediği gözlemlenmektedir. Duyarlılık ve kesinliğin yüksek oranları, modelin gerçek pozitif durumları etkin bir şekilde tanıdığını ve yanlış pozitiflerin çok düşük olduğunu göstermektedir.

Şekil 20, KNN ile dörtlü sınıflandırma için elde edilen test sonuçlarının grafiksel sunumunu içerir. Belirtilen görsel, farklı modellerin performansını karşılaştırmalı olarak sunarak, algoritmanın test verileri üzerindeki etkinliğini vurgular.

Tablo 12 K-En Yakın Komşu Algoritması ile Dörtlü Sınıflandırma için Test Sonuçları

Sıralama	**Öğrenme Algoritması**	**Doğruluk (%)**	**Duyarlılık (%)**	**Özgüllük (%)**	**Kesinlik (%)**	**F1 Skoru (%)**
1.Deneme	KNN	92.20	85.0	95.00	85.00	85.00
2.Deneme	KNN	87.70	83.96	94.65	83.96	83.96
3.Deneme	KNN	89.50	75.24	91.75	75.24	75.24



Şekil 20 K-En Yakın Komşu Algoritması ile Dörtlü Sınıflandırma için Test Grafiği

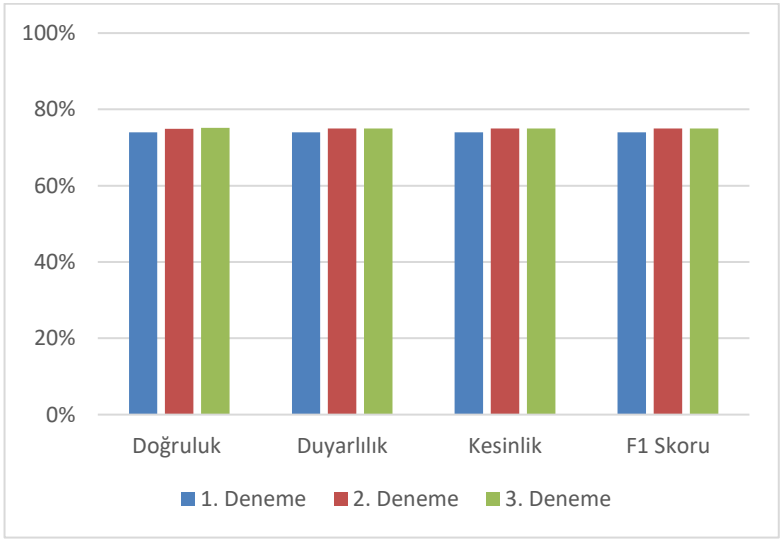
Tablo 11 ve Tablo 12 birlikte yorumlandığında her modelin dengeli bir performans sergilediği gözlemlenmiştir. Ancak test sonuçları, üçüncü model için, performans değeri için belirgin bir düşüş olduğunu gösterir. İlk model, hem öğrenme hem de test aşamalarında, diğer modellere kıyasla genel olarak daha iyi bir performans göstermektedir.

4.2.2 Dörtlü Sınıflandırma Yöntemi için Yapay Sinir Ağları Algoritması Öğrenme ve Test Sonuçları Analizi

Yapay Sinir Ağları Algoritması ile Dörtlü Sınıflandırma için Öğrenme Sonuçlarının belirtildiği Tablo 13 'te her üç modelin de doğruluk değerleri aynı çıkmıştır. Modellerin doğru sınıflandırma yapma konusunda orta düzeyde performans gösterdiğine işaret etmektedir. Genel olarak, bu modellerin öğrenme aşamasında gösterdiği performans, orta düzeyde bir başarı gösterdiğini ifade eder

Tablo 13 Yapay Sinir Ağları Algoritması ile Dörtlü Sınıflandırma için Öğrenme Sonuçları

Sıralama	Öğrenme Algoritması	Doğruluk (%)	Duyarlılık (%)	Özgüllük (%)	Kesinlik (%)	F1 Skoru (%)
1.Deneme	YSA	74.20	74.21	91.40	74.00	74.00
2.Deneme	YSA	74.90	74.87	91.62	74.87	74.87
3.Deneme	YSA	75.20	75.24	91.75	75.24	75.24



Şekil 21 Yapay Sinir Ağları Algoritması ile Dörtlü Sınıflandırma için Öğrenme Grafiği

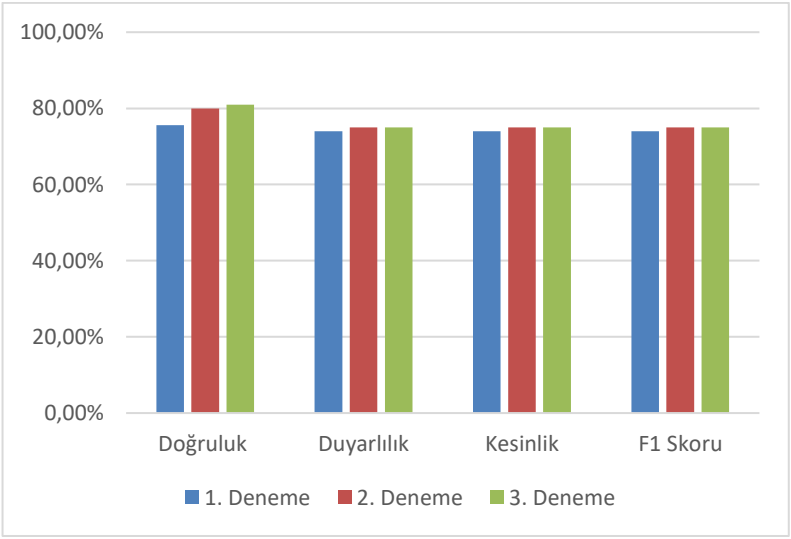
Şekil 21, YSA ile dörtlü sınıflandırma için elde edilen öğrenme sonuçlarının grafiksel sunumunu içerir. Belirtilen görsel, farklı denemelerin performansını karşılaştırmalı olarak sunarak, algoritmanın öğrenme verileri üzerindeki etkinliğini vurgular. .

Tablo 13 'te öğrenme sonuçları ile yapılan kıyaslamada, test sonuçlarında en yüksek doğruluğu üçüncü deneme göstermiştir. Kesinlik ve F1 skoru değerleri arasındaki yakınlık, pozitif sınıflandırılan durumların ve sınıflandırma dengesinin yakın olduğunu göstermektedir.

Şekil 22, YSA ile dörtlü sınıflandırma için elde edilen test sonuçlarının grafiksel sunumunu içerir. Belirtilen görsel, farklı denemelerin performansını karşılaştırmalı olarak sunarak, algoritmanın test verileri üzerindeki etkinliğini vurgular.

Tablo 14 Yapay Sinir Ağları Algoritması ile Dörtlü Sınıflandırma için Test Sonuçları

Sıralama	**Öğrenme Algoritması**	**Doğruluk (%)**	**Duyarlılık (%)**	**Özgüllük (%)**	**Kesinlik (%)**	**F1 Skoru (%)**
1.Deneme	YSA	75.6	74.21	91.4	74.21	74.21
2.Deneme	YSA	79.6	75.24	91.75	75.24	75.24
3.Deneme	YSA	80.5	74.87	91.62	74.87	74.87



Şekil 22 Yapay Sinir Ağları Algoritması ile Dörtlü Sınıflandırma için Test Grafiği

Tablo 13 ve Tablo 14 için yapılan karşılaştırmada tüm denemelerin benzer ve yüksek duyarlılık ve özgüllük değerleri ile dikkat çekmektedir.

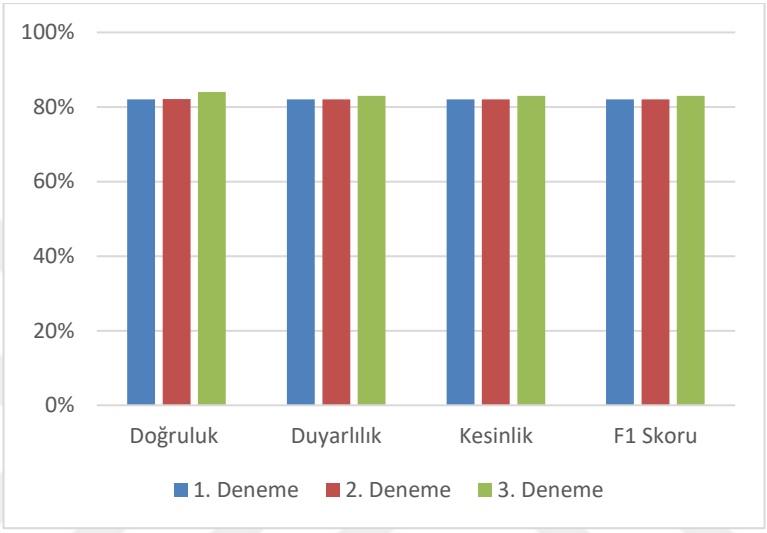
4.2.3 Dörtlü Sınıflandırma Yöntemi için Destek Vektör Makinesi Algoritması Öğrenme ve Test Sonuçları Analizi

Tablo 14'te, üçüncü denemenin hem doğruluk, hem de kesinlik değerleri karşılaştırıldığında diğer denemelere göre daha yüksek performans sergilediği gözlemlenmiştir. Yüksek özgüllük oranı, denemelerin negatif sınıfları doğru bir şekilde tespit ettiğini belirtmektedir. Bu durum yanlış pozitiflerin düşük olduğu anlamına gelmektedir.

Şekil 23, DVM ile dörtlü sınıflandırma için elde edilen öğrenme sonuçlarının grafiksel sunumunu içerir. Belirtilen görsel, farklı denemelerin performansını karşılaştırmalı olarak sunarak, algoritmanın öğrenme verileri üzerindeki etkinliğini vurgular.

Tablo 15 Destek Vektör Makinesi Algoritması ile Dörtlü Sınıflandırma için Öğrenme Sonuçları

Sıralama	**Öğrenme Algoritması**	**Doğruluk (%)**	**Duyarlılık (%)**	**Özgüllük (%)**	**Kesinlik (%)**	**F1 Skoru (%)**
1.Deneme	DVM	82.00	81.97	93.99	81.97	81.97
2.Deneme	DVM	82.10	82.11	94.04	82.11	82.11
3.Deneme	DVM	83.50	83.21	94.40	83.21	83.21



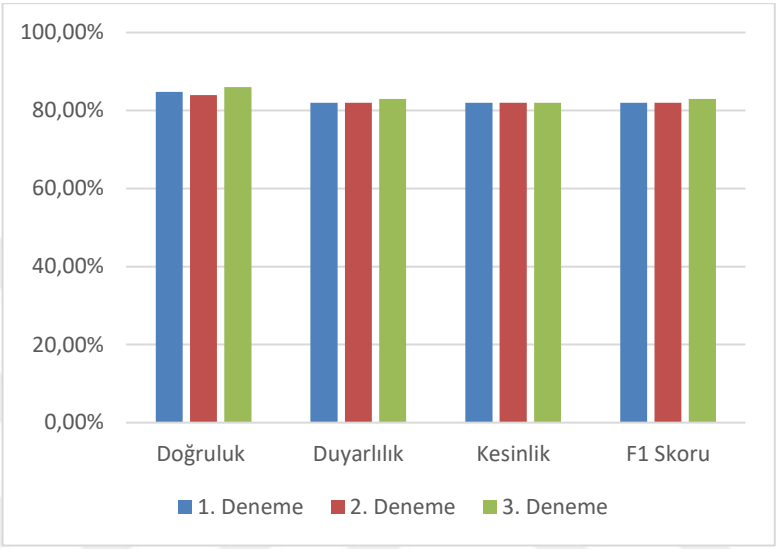
Şekil 23 Destek Vektör Makinesi Algoritması ile Dörtlü Sınıflandırma için Öğrenme Grafiği

Tablo 15'te belirtilen denemelerin genel olarak benzer bir performans sergilediği gözlemlenmiştir. Ancak üçüncü denemenin diğer denemelerle karşılaştırılmasında daha yüksek doğruluk ve F1 skoru değeri gösterdiği gözlemlenmiştir. Bu durum, denemenin test performansının diğer denemelere göre daha yüksek olduğunu gösterir.

Şekil 24, DVM ile dörtlü sınıflandırma için elde edilen test sonuçlarının grafiksel sunumunu içerir. Belirtilen görsel, farklı denemelerin performansını karşılaştırmalı olarak sunarak, algoritmanın test verileri üzerindeki etkinliğini vurgular.

Tablo 16 Destek Vektör Makinesi Algoritması ile Dörtlü Sınıflandırma için Test Sonuçları

Sıralama	Öğrenme Algoritması	Doğruluk (%)	Duyarlılık (%)	Özgüllük (%)	Kesinlik (%)	F1 Skoru (%)
1.Deneme	DVM	84.80	81.97	93.99	81.97	81.97
2.Deneme	DVM	84.30	82.11	94.04	82.11	82.11
3.Deneme	DVM	85.50	83.21	94.40	83.21	83.21



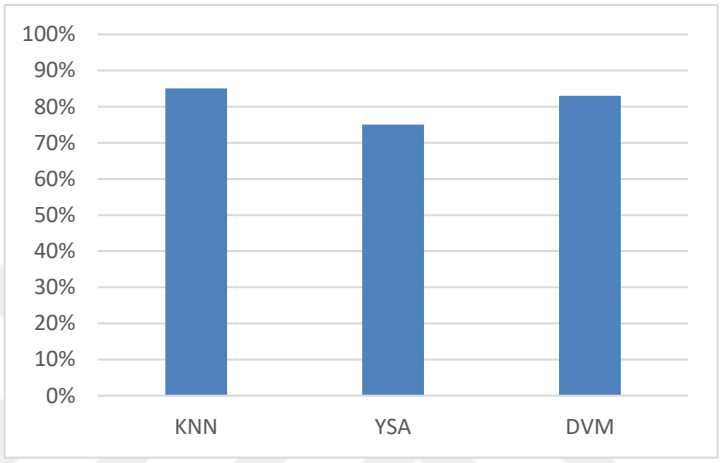
Şekil 24 Destek Vektör Makinesi Algoritması ile Dörtlü Sınıflandırma için Test Grafiği

Tablo 15 ve Tablo 16 incelendiğinde tüm denemeler duyarlılık ve özgüllük açısından tutarlı bir performans sergilemiş ve doğruluk oranları test aşamasında artış göstermiştir.

Tablo 17'de belirtildiği üzere KNN en yüksek F1 skoruna sahiptir. Bu durumu, DVM ve YSA takip etmektedir. Yüksek F1 skoru, denemenin hem duyarlılık, hemde kesinlik açısından iyi bir denge kurduğunu göstermektedir. Bu durum denemenin, gerçek pozitifleri, yanlış pozitiflerle karıştırmadan etkili bir şekilde tespit ettiğini ifade eder. KNN'nin öğrenme verisinde bu derece iyi uyum sağlaması, öğrenme sürecinde karmaşık desenleri yakalayabilme kapasitesinin yüksek olduğunu işaret etmektedir. Öğrenme verileri üzerinde DVM'nin sağlam bir performans gösterdiği, ancak KNN kadar etkili olmadığı gözlemlenmektedir.

Tablo 17 Dörtlü Sınıflandırma için Tüm Denemelerin Öğrenme Verileri F1 Skoru Karşılaştırması

Sıralama	F1 Skoru (%)
KNN	85.10
YSA	75.24
DVM	83.21

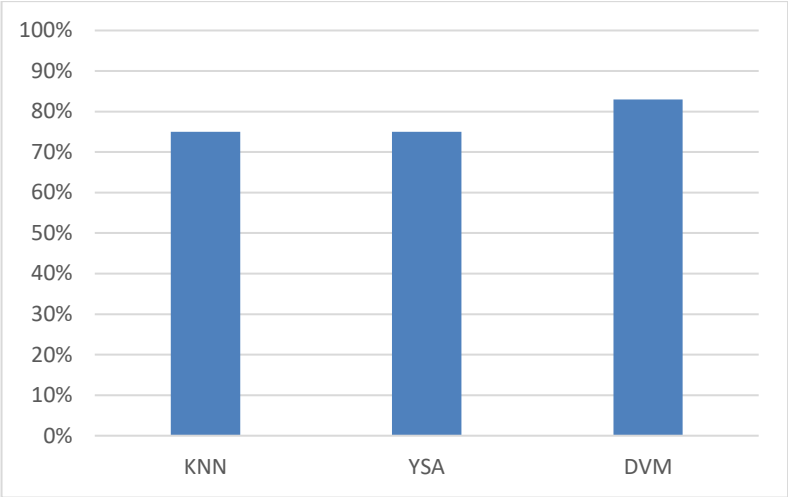


Şekil 25 Dörtlü Sınıflandırma için Tüm Denemelerin Öğrenme Verileri F1 Skoru Karşılaştırma Grafiği

Tablo 18'de belirtildiği üzere, test verisi üzerinde yapılan değerlendirmede, DVM en yüksek F1 skoru performansını sergilemiştir. Bu durum DVM'nin genelleme kabiliyetinin yüksek olduğunu göstermektedir. KNN'nin test F1 skoru, öğrenme aşamasına kıyasla düşüş göstermistir. Bu durum aşırı uyuma ve öğrenme verisine fazla özgü desenlerin öğrenilmesine işaret edebilir. YSA'nın F1 skoru öğrenme ve test aşamalarında tutarlı kalmış, ancak bu durum denemeler arasında en düşük skor olduğu için performansı artırma ihtiyacını göstermektedir.

Tablo 18 Dörtlü Sınıflandırma için Tüm Denemelerin Test Verileri F1 Skoru Karşılaştırması

Sıralama	F1 Skoru (%)
KNN	75.24%
YSA	74.87
DVM	83.21



Şekil 26 Dörtlü Sınıflandırma için Tüm Denemelerin Test Verileri F1 Skoru Karşılaştırma Grafiği

Her iki tablo ve grafik incelendiğinde, DVM'nin dörtlü sınıflandırma problemi için en sağlam model olduğu gözlemlenmektedir. Bu sonuç, DVM'nin karar sınırlarını belirlemede marjinal yaklaşımın, yüksek boyutlu veri uzayında etkinliğinin bir sonucu olabilir. YSA'nın hem öğrenme hem de test aşamalarında göreceli olarak düşük performans göstermesi, belki de model mimarisinin ve hiperparametre ayarlarının bu görev için daha iyi optimize edilmesi gerektiğine işaret etmektedir.

Bu çalışmada ele alınan KNN ve DVM, YSA modelleri, farklı öznitelikler göz önünde bulundurularak değerlendirilmiştir. İkili ve çoklu sınıflandırma testlerinde, KNN algoritmalarının yüksek doğruluk oranlarına ulaştığı gözlemlenmiştir. İkili sınıflandırmada %92.17'ye varan bir doğruluk, modelin keskin sınıflandırma kabiliyetine işaret etmektedir. Dörtlü sınıflandırmada ise %90.20 gibi etkileyici bir performans sergilenmiştir. Benzer şekilde, çok katmanlı yapay sinir ağları da %86'ya varan doğruluk oranları ile dikkate değer sonuçlar elde etmiştir. SVM algoritmalarının da %91.4 gibi yüksek bir doğruluk sergilediği ve özellikle çekirdek fonksiyonlar ve çok sınıflı kodlamada farklı yaklaşımların etkili olduğu görülmüştür.

Elde edilen yüksek doğruluk oranları, özellikle EEG veri setlerinin karmaşıklığı göz önüne alındığında, kullanılan makine öğrenimi algoritmalarının etkinliğini göstermektedir. Ancak, modelin karmaşıklığı arttıkça tahmin maliyeti ve eğitim süresi gibi kaynak tüketimi de artmaktadır.

Bu durum, gerçek zamanlı uygulamalar için önemli bir faktör olup, uygulamanın işlevselliği ile kaynak tüketimi arasında bir denge kurulması gerektiğini göstermektedir. Yapay sinir ağlarının daha uzun eğitim sürelerine ihtiyaç duyması, büyük veri setleri üzerinde çalışan ve hızlı tepki vermesi gereken sistemlerde bir engel teşkil edebilir. Model boyutu da önemli bir husustur ve daha büyük modeller, sınıflandırma performansını artırabilirken, aynı zamanda daha fazla depolama alanı ve işlem gücü gerektirebilir. Bu, özellikle gömülü sistemlerde ve mobil cihazlarda çalışacak rahatlama teknolojileri için önemli bir sınırlamadır.

Performans metriklerine gelince, duyarlılık ve özgüllük değerleri bize modellerin sınıflandırma hassasiyeti hakkında bilgi vermektedir. En yüksek özgüllük değerleri, modellerin yanlış pozitifleri başarıyla minimize edebildiğini ve gerçek negatif durumları doğru bir şekilde tanımlayabildiğini göstermektedir. Ancak, düşük duyarlılık değerleri bazı modellerin belirli sınıfları yanlış negatif olarak sınıflandırma eğiliminde olduğunu göstermektedir. Bu durum, özellikle yanlış negatiflerin yüksek risk taşıdığı tıbbi uygulamalar gibi alanlarda önemli olabilir. Son olarak, kesinlik ve F1 skorları modellerin genel performansı hakkında dengeli bir görüş sunmaktadır. Ancak, bu değerler aynı zamanda veri dengesizliği ve sınıflar arası farklılıklar gibi faktörlerden etkilenebilir. Bu nedenle, gerçek dünya verileri üzerinde yapılan daha kapsamlı testler, bu sonuçların genel geçerliliğini doğrulamak için gereklidir.

5 SONUÇ VE ÖNERİLER

Bu çalışmada, elektroensefalografi (EEG) verilerini kullanarak kişilerin anksiyete durum tespitini yapmak için K-En Yakın Komşu (KNN), Destek Vektör Makineleri ve Yapay Sinir Ağları modelleri geliştirilmiş ve karşılaştırılmıştır. Elde edilen sonuçlar, modelin doğruluğu, duyarlılık, özgüllük, kesinlik ve F1 skoru performans metrikleri üzerinden değerlendirilmiştir.

İkili sınıflandırma testlerinde KNN algoritmasının %92.17 gibi yüksek bir doğruluk oranına ulaştığı gözlemlenirken, dörtlü sınıflandırmada %92.2 'ye ulaşan bir doğruluk sergilenmiştir. Bunun yanı sıra, çok katmanlı yapay sinir ağları (YSA) modellerinin de benzer bir performans gösterdiği gözlenmiştir. Özellikle, ikili sınıflandırmada Tablo 5'de belirtilen üçüncü denemenin %83.8 doğruluk oranı elde etmesi YSA algoritmanın karmaşık veri setleri üzerindeki potansiyelini göstermektedir.

Performans metriklerine detaylı olarak bakıldığında, duyarlılık ve özgüllük oranlarının yanı sıra kesinlik ve F1 skorları da yüksek seviyelerde tutulmuştur. Bu metrikler, sınıflandırma modellerinin sadece genel doğruluğunu değil, aynı zamanda belirli sınıflara yönelik duyarlılığını ve dengeli bir performansını da göstermektedir. Özellikle, en yüksek özgüllük oranları, modellerin yanlış pozitifleri minimuma indirgediğini ve gerçek negatifleri başarılı bir şekilde tanımladığını belirtmektedir.

Algoritmaların performansları ikili ve dörtlü sınıflandırma arasında değişiklik gösterir. Genellikle dörtlü sınıflandırmada duyarlılık oranları, ikili sınıflandırmaya göre daha düşük çıktığı gözlemlenmiştir. Buradan, dörtlü sınıflandırmanın daha karmaşık bir görev olduğu ve algoritmaların daha fazla sınıfı ayırt etmesi gerektiğini göstermektedir.

İkili ve dörtlü sınıflandırma görevlerinde farklı algoritmaların güçlü ve zayıf yönleri bulunmaktadır. Güçlü ve zayıf yönlerin daha detaylı tespiti için detaylı öznitelik çıkarımı çalışması önerilmektedir.

EEG sinyal analizi gibi karmaşık veri setlerinde, algoritmaların performansını optimize etmek için veri ön işleme ve öznitelik çıkarımı, veri seti boyutu belirleme işlemlerinin daha ayrıntılı yapılması gerekmektedir. Sonraki çalışmalarda, veri seti içerisinden farklı bölgelerden, farklı zaman dilimlerinde veriler seçilip, öznitelik çıkarımı yapılarak çalışma iyileştirilebilir.

KAYNAKÇA

[1] Çağlıyan, U. K., & Çağlıyan, B. (2021). Epilepsi EEG Verilerinin Makine Öğrenmesi Teknikleriyle. Avrupa Bilim ve Teknoloji Dergisi, (23), 163-172

[2] Cook, I. A. H. A. M. K. A. F. H., & Cook, L. A. F. (2009). Quantitative EEG Analysis Methods and Clinical Applications. EEG Signals in Psychiatry: Biomarkers for Depression Management içinde (ss. 289-315). Artech House..

[3] Vawter, M. P., & Vawter, M. E. A. (2013). Encyclopedia of Systems Biology. Biomarkers, Clinical Relevance içinde (ss. 125-126). Springer New York.

[4] American Psychiatric Association. (2013). Diagnostic and Statistical Manual of Mental Disorders. American Psychiatric Publishing.

[5] World Health Organization. (2018). International Classification of Diseases 11th Revision: The global standard for diagnostic health information. World Health Organization.

[6] Aldemir, R. (2019). Dikkat eksikliği ve hiperaktivite bozukluğu olan çocuklarda ilaçla tedavi sürecinin EEG analizleriyle değerlendirilmesi. Erciyes Üniversitesi.

[7] Mathews, L. D., & Mathews, C. D. (2006). Projections of Global Mortality and Burden of Disease from 2002 to 2030. PlosMedicine, 3(11), e442.

[8] Ezzati, S. K. S. J. A., & Ezzati, M. (2013). Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: A systematic analysis for the Global Burden of Disease Study 2010. *The Lancet, 380*(9859), 2197-2223.

[9] World Health Organization. (2017). Depression and other common mental disorders.

[10] Baghdadi, A., Aribi, Y., Fourati, R., Halouani, N., Siarry, P., & Alimi, A. M. (2019). DASPS: A Database for Anxious States based on a Psychological Stimulation.

[11] Başaran, M. (2019). Dikkat eksikliği ve hiperaktivite bozukluğunun (DEHB) EEG sinyalleri kullanılarak yapay sinir ağları ile kestirimi. Dumlupınar Üniversitesi.

[12] Çapkan Altun, E. N. (2019). Opioid kullanım bozukluğu ve sağlıklı kontrol gruplarının sınıflandırılmasında QEEG tabanlı biyobelirteç ile makine öğrenme

yöntemleri kullanılarak retrospektif olarak sınıflandırılması [Doktora tezi, Üsküdar Üniversitesi].

[13] Khodayari-Rostamabad, A. R. J. P., Reilly, J. P., Hasey, G. M., de Bruin, H., & Maccrimmon, D. J. (2013). A machine learning approach using EEG data to predict response to SSRI treatment for major depressive disorder. *Clinical Neurophysiology, 124*(10), 1975–1985.

[14] Hosseinifard, B. M. M., Moradi, M. H., & Rostami, R. (2013). Classifying depression patients and normal subjects using machine learning techniques and nonlinear features from EEG signal. *Computer Methods and Programs in Biomedicine, 109*(3), 339–345.

[15] Mohammadi, M. A.-A., Al-Azab, F. R., Raahemi, B. R., Richards, G. J. N., Jaworska, S. D. d. l. S., Smith, D., ... Knott, V. (2015). Data mining EEG signals in depression for their diagnostic value. *BMC Medical Informatics and Decision Making, 15*(1), 108.

[16] Bolton, D. (2008). What is mental disorder? Oxford University Press.

[17] American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders: DSM-5.

[18] Tsang, W. (2003). Kernel methods in supervised and unsupervised learning.

[19] Pattanayak, S. (2017). Convolutional Neural Networks. Pro Deep Learning with TensorFlow.

[20] Nie, S. Z. M., Zheng, M., & Ji, Q. (2018). The Deep Regression Bayesian Network and Its Applications: Probabilistic Deep Learning for Computer Vision. *IEEE Signal Processing Magazine, 35*(1), 1.

[21] Lin, C.-T., Chuang, C.-H., Huang, C.-S., Tsai, S.-F., Lu, S.-W., Chen, Y.-H., & Ko, L.-W. (2014). Wireless and wearable EEG system for evaluating driver vigilance. *Biomedical Circuits and Systems, 8*(2), 165-176.

[22] Zhang, G., Davoodnia, V., Sepas-Moghaddam, A., Zhang, Y., & Etemad, A. (2020). Classification of Hand Movements From EEG Using a Deep Attention-Based LSTM Network. *IEEE Sensors Journal*, 20(6), 3113-3122

[23] Santos-Mayo, L., San-Jose-Revuelta, L. M., & Arribas, J. I. (2017). A Computer-Aided Diagnosis System With EEG Based on the P3b Wave During an Auditory Odd-Ball Task in Schizophrenia. *IEEE Transactions on Biomedical Engineering*, 64(2), 395-407

[24] Yazıcı, M. (2015). Classification of EEG signals using time domain features.

[25] Omar, S. Q., & Tepe, C. (2023). EEG İşareti tabanlı anksiyete sınıflandırması için dalgacık dönüşümü ile öznitelik çıkarma. Niğde Ömer Halisdemir Üniversitesi Mühendislik Bilimleri Dergisi, 12(3), 726-732.

[26] Baghdadi, A., Aribi, Y., Fourati, R., Halouani, N., Siarry, P., & Alimi, A. M. (2019). DASPS: A Database for Anxious States based on a Psychological Stimulation.

[27] Emotiv. (s.a.). Epoc+. https://www.emotiv.com/epoc/.

[28] Hacıbeyoğlu, M., Çelik, M., & Erdaş Çiçek, Ö. (2023). K En Yakın Komşu Algoritması ile Binalarda Enerji Verimliliği Tahmini. Necmettin Erbakan Üniversitesi Fen ve Mühendislik Bilimleri Dergisi, 5(2), 65-74

[29] Altınel, A. B. (2021). Türkçe Metinlerde Makine Öğrenmesi Algoritmalarının Duygu Analizi Problemi Üzerindeki Performansının Kıyaslanması. European Journal of Science and Technology, Special Issue 28, 1056-1061.

[30] Kavzoğlu, T., & Çölkesen, İ. (2010). Destek vektör makineleri ile uydu görüntülerinin sınıflandırılmasında kernel fonksiyonlarının etkilerinin incelenmesi. Harita Dergisi, 76(144), 73-82.

[31] Köroğlu, E. (2014). Amerikan Psikiyatri Birliği, Ruhsal Bozuklukların Tanısal ve Sayımsal Elkitabı (DSM-5) (5. baskı). Hekimler Yayın Birliği.