

**BEYİN GÖRÜNTÜLEME TEKNİKLERİNİN
ALZHEIMER HASTALIĞI ERKEN TANI
TAHMİNİNDE KULLANILMASI**

Yüksek Lisans Tezi

Savaş OKYAY

Eskişehir, 2016

**BEYİN GÖRÜNTÜLEME TEKNİKLERİNİN ALZHEIMER HASTALIĞI
ERKEN TANI TAHMİNİNDE KULLANILMASI**

Savaş OKYAY

YÜKSEK LİSANS TEZİ

Bilgisayar Mühendisliği Anabilim Dalı

Danışman: Yrd. Doç. Dr. Nihat ADAR

Eskişehir

Anadolu Üniversitesi

Fen Bilimleri Enstitüsü

Haziran, 2016

JÜRİ VE ENSTİTÜ ONAYI

Savaş OKYAY'ın “Beyin Görüntüleme Tekniklerinin Alzheimer Hastalığı Erken Tanı Tahmininde Kullanılması” başlıklı tezi 09/06/2016 tarihinde aşağıdaki jüri tarafından değerlendirilerek “Anadolu Üniversitesi Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliği”nin ilgili maddeleri uyarınca, **Bilgisayar Mühendisliği** Anabilim dalında Yüksek Lisans tezi olarak kabul edilmiştir.

	<u>Unvanı-Adı Soyadı</u>	<u>İmza</u>
Üye (Tez Danışmanı)	: Yrd. Doç. Dr. Nihat ADAR
Üye	: Yrd. Doç. Dr. Muzaffer DOĞAN
Üye	: Yrd. Doç. Dr. Kemal ÖZKAN

.....

Enstitü Müdürü

ÖZET

BEYİN GÖRÜNTÜLEME TEKNİKLERİNİN ALZHEIMER HASTALIĞI ERKEN TANI TAHMİNİNDE KULLANILMASI

Savaş OKYAY

Bilgisayar Mühendisliği Anabilim Dalı

Anadolu Üniversitesi, Fen Bilimleri Enstitüsü, Haziran, 2016

Danışman: Yrd. Doç. Dr. Nihat ADAR

Demans tipine göre hastaların beyinleri karakteristik farklılıklar gösterir. Beyin korteks kalınlığı, bazı bölgelerin hacmi veya yüzey alanı gibi ölçülendirmeleri, hastalık tiplerinin belirlenmesinde etkilidir. Tıbbi görüntüleme standartlarına uygun beyin görüntüleri manyetik rezonans görüntüleme cihazlarından elde edilebilir. Görüntülerin başlık kısmı, hastalara, hastalıklara, ayrıca görüntüleme çalışmalarına olmak üzere farklı tür birçok teknik olan veya olmayan bilgi barındırır. Beyin görsel dosyaları ve başlıkları üzerinden görüntü işleme tekniklerinin de yardımı ile hasta beyinlerinin fiziksel özellikleri sayısal olarak çıkarılabilir. Elde edilen sayısal verinin sınıflandırma algoritmaları içerisinde düzlemde vektörler şeklinde ifade edilmesi sonrasında numuneler sınıflandırılabilir. Bu çalışmada, 19 adet Alzheimer hastalığı, 19 adet frontotemporal demans ve 25 adet vasküler demans olmak üzere üç farklı hastalıktaki 63 numunenin manyetik rezonans görüntüleri kullanılmaktadır. Dilimlenmiş beyin görüntü setleri Freesurfer beyin analiz yazılım aracı ile birlikte işlenmektedir. Programın başarılı analizi sonrasında üretilen istatistik bilgileri içerisinde farklı öznitelik grupları oluşturulmaktadır. Öznitelik matrisleri, sarmalama yaklaşımı öznitelik seçim yöntemi uygulanacak şekilde genetik algoritmaya gönderilmektedir. Bulunan değerli öznitelikler, hastalığı belirlemede etkili olan uzuvlar, tartışılmaktadır. Farklı sınıflandırma yaklaşımları ve genetik algoritma parametreleri uygulanarak %95.2'yi gören doğruluk sonuçları ve hata matrisleri elde edilmektedir.

Anahtar Sözcükler: Genetik algoritma, Öznitelik seçimi, Freesurfer, Sınıflandırma, Demans.

ABSTRACT

THE USE OF BRAIN IMAGING TECHNIQUES FOR PREDICTING EARLY DIAGNOSIS OF ALZHEIMER'S DISEASE

Savaş OKYAY

Department of Computer Engineering

Anadolu University, Graduate School of Sciences, June, 2016

Supervisor: Assist. Prof. Dr. Nihat ADAR

Brains of dementia patients show characteristic differences according to the type of disease. Brain measurements such as cerebral cortical thickness, volumes or surface areas of some specific regions are effective in determining the type of disease. Brain images that are appropriate to the medical imaging standards can be obtained from magnetic resonance imaging devices. The headers of the images contain many different types of technical or non-technical information about patients, diseases, also imaging studies. Through brain imaging files and headers, the physical characteristics of patient brains can be extracted with the help of image processing techniques. After the numerical data is expressed as vectors in the plane during the classification, samples can be classified. In this study, magnetic resonance scans of 63 samples, having three disease types: 19 Alzheimer's disease, 19 frontotemporal dementia, and 25 vascular dementia, are used. Sliced brain image sets are processed with Freesurfer brain analyzing software tool. Different feature groups are created via generated statistical information after successful analysis of the program. Feature matrices are sent to the genetic algorithm that is used as a wrapper method for feature selection. Significant features, the brain regions that are effective in identifying the disease, are discussed. Applying different classification algorithms, and also genetic algorithm parameters, accuracy results up to %95.2 and confusion matrices are achieved.

Keywords: Genetic algorithm, Feature selection, Freesurfer, Classification, Dementia.

TEŞEKKÜR

Öncelikle bu tez çalışması olmak üzere birlikte yer aldığımız tüm akademik çalışmalarda, ayrıca idari çalışmalarda yapmış olduğu nokta atışı öneriler ile çalışma verimini yükselten, herhangi başka bir konuda dara düştüğümde desteğini esirgemeyen, kendime örnek almaya çalıştığım, engin bilgi birikimine sahip Sayın Nihat Adar'a en içten teşekkürlerimi sunarım.

Lisans ve yüksek lisans eğitimim boyunca yapmış olduğu öğretim ve tavsiyeler için Sayın Kemal Özkan'a bütün kalbimle teşekkür ederim. Eğitim-öğretim hayatım boyunca üzerimde emeği olan diğer bütün hocalarıma da teşekkürlerimi sunarım.

Çalışmalarına verdiği öneriler ile seviye atlattıran derin bilgi birikimli arkadaşım Sayın Cihan Uyanık'a en içten teşekkürlerimi sunarım.

Bu çalışmada kullanılan veri setinin oluşmasındaki katkılarından ötürü Sayın Demet Belgin Özbabalık Adapınar ve Sayın Baki Adapınar'a teşekkürlerimi sunarım.

Son olarak, beni bugünlere ulaştıran sevgili aileme minnettarım.

Savaş Okyay

Haziran 2016

09/06/2016

ETİK İLKE VE KURALLARA UYGUNLUK BEYANNAMESİ

Bu tezin bana ait, özgün bir çalışma olduğunu; çalışmamın hazırlık, veri toplama, analiz ve bilgilerin sunumu olmak üzere tüm aşamalardan bilimsel etik ilke ve kurallara uygun davrandığımı; bu çalışma kapsamında elde edilemeyen tüm veri ve bilgiler için kaynak gösterdiğimi ve bu kaynaklara kaynakçada yer verdiğimi; bu çalışmanın Anadolu Üniversitesi tarafından kullanılan “bilimsel intihal tespit programı”yla tarandığımı ve hiçbir şekilde “intihal içermediğini” beyan ederim. Herhangi bir zamanda, çalışmamla ilgili yaptığım bu beyana aykırı bir durumun saptanması durumunda, ortaya çıkacak tüm ahlaki ve hukuki sonuçlara razı olduğumu bildiririm.

.....

(İmza)

Savaş OKYAY

İÇİNDEKİLER

	<u>Sayfa</u>
BAŞLIK SAYFASI	i
JÜRİ VE ENSTİTÜ ONAYI.....	ii
ÖZET	iii
ABSTRACT.....	iv
TEŞEKKÜR	v
ETİK İLKE VE KURALLARA UYGUNLUK BEYANNAMESİ.....	vi
İÇİNDEKİLER	vii
TABLolar DİZİNİ	ix
ŞEKİLLER DİZİNİ.....	xi
GÖRSELLER DİZİNİ	xii
SİMGELER ve KISALTMALAR DİZİNİ.....	xiii
1. GİRİŞ	1
1.1. Literatür Taraması	3
1.2. Tez Organizasyonu	10
2. VERİ SETİ.....	12
3. YÖNTEM.....	14
3.1. Dosyaların Analizi.....	15
3.2. Öznitelik Çıkarma.....	16
3.3. Normalizasyon.....	22
3.4. Sınıflandırıcılar	23
3.4.1. Naïve Bayes sınıflandırıcı.....	23
3.4.2. Destek vektör makineleri	24
3.5. Öznitelik Seçimi.....	26
3.5.1. Genetik algoritma	27
3.5.2. Genetik algoritma parametreleri	29
3.5.2.1. Kromozom tipi.....	30
3.5.2.2. Fitness belirleme ve ölçeklendirme	31
3.5.2.3. Popülasyon	32
3.5.2.3.1. Başlangıç popülasyonu	32
3.5.2.3.2. Popülasyon kurgulama.....	33
3.5.2.3.3. Popülasyon büyüklüğü	33

3.5.2.4. Yeni nesil oluřturma.....	34
3.5.2.4.1. Seim	34
3.5.2.4.2. aprazlama	35
3.5.2.4.3. Mutasyon	37
3.5.2.5. Durdurma lutleri	37
3.6. apraz Doęrulama ve Hata Matrisleri	38
4. DENEYSEL ALIŐMALAR.....	40
4.1. 3-Sınıflı Deneysel alıŐmalar ve Sonuları.....	41
4.2. 2-Sınıflı Deneysel alıŐmalar ve Sonuları.....	43
4.2.1. AH v dięerleri sınıflandırması deneysel alıŐmaları ve sonuları.....	44
4.2.2. FTD v dięerleri sınıflandırması deneysel alıŐmaları ve sonuları.....	46
4.2.3. VaD v dięerleri sınıflandırması deneysel alıŐmaları ve sonuları.....	48
4.3. Deneysel Sonular ve Deęerlendirme.....	50
5. SONULAR VE YORUMLAR.....	53
KAYNAKA.....	55
ZGEMİŐ	59

TABLULAR DİZİNİ

Sayfa

Tablo 2.1. Kullanılan Görüntü Setlerinin Teknik Özellikleri	12
Tablo 2.2. Muayene Özellikleri	12
Tablo 2.3. Tıbbi Görüntü Cihazına Göre Hastalık Dağılım Tablosu.....	13
Tablo 2.4. Cinsiyete Göre Hastalık Dağılım Tablosu	13
Tablo 3.1. Freesurfer Beyin Yapılandırma Aşamaları.....	17
Tablo 3.2. Öznitelik Grupları	21
Tablo 3.3. 4×4 Boyutlarında Örnek Bir Matrisin Ham Değerleri.....	22
Tablo 3.4. 4×4 Boyutlarında Örnek Normalizasyon.....	22
Tablo 3.5. Hata Matrisi Örneği	38
Tablo 4.1. 3-Sınıflı Genetik Algoritma Birinci Aşama Parametreleri	41
Tablo 4.2. Naïve Bayes ile 3-Sınıflı Genetik Algoritma Birinci Aşama Sonuçları	42
Tablo 4.3. 3-Sınıflı Genetik Algoritma İkinci Aşama Parametreleri	42
Tablo 4.4. Naïve Bayes ile 3-Sınıflı Genetik Algoritma İkinci Aşama Sonuçları.....	43
Tablo 4.5. 2-Sınıflı Genetik Algoritma Birinci Aşama Parametreleri	43
Tablo 4.6. 2-Sınıflı Genetik Algoritma İkinci Aşama Parametreleri	44
Tablo 4.7. Naïve Bayes ile AH v Diğerleri Sınıflandırması Genetik Algoritma Birinci Aşama Sonuçları	44
Tablo 4.8. Naïve Bayes ile AH v Diğerleri Sınıflandırması Genetik Algoritma İkinci Aşama Sonuçları	45

Tablo 4.9. SVM ile AH v Diğerleri Sınıflandırması Genetik Algoritma Birinci Aşama Sonuçları.....	45
Tablo 4.10. SVM ile AH v Diğerleri Sınıflandırması Genetik Algoritma İkinci Aşama Sonuçları.....	46
Tablo 4.11. Naïve Bayes ile FTD v Diğerleri Sınıflandırması Genetik Algoritma Birinci Aşama Sonuçları	46
Tablo 4.12. Naïve Bayes ile FTD v Diğerleri Sınıflandırması Genetik Algoritma İkinci Aşama Sonuçları	47
Tablo 4.13. SVM ile FTD v Diğerleri Sınıflandırması Genetik Algoritma Birinci Aşama Sonuçları.....	47
Tablo 4.14. SVM ile FTD v Diğerleri Sınıflandırması Genetik Algoritma İkinci Aşama Sonuçları.....	48
Tablo 4.15. Naïve Bayes ile VaD v Diğerleri Sınıflandırması Genetik Algoritma Birinci Aşama Sonuçları	48
Tablo 4.16. Naïve Bayes ile VaD v Diğerleri Sınıflandırması Genetik Algoritma İkinci Aşama Sonuçları	49
Tablo 4.17. SVM ile VaD v Diğerleri Sınıflandırması Genetik Algoritma Birinci Aşama Sonuçları.....	49
Tablo 4.18. SVM ile VaD v Diğerleri Sınıflandırması Genetik Algoritma İkinci Aşama Sonuçları.....	49

ŞEKİLLER DİZİNİ

	<u>Sayfa</u>
Şekil 3.1. Lineer SVM Modeli Örneği	25
Şekil 3.2. Kernel Fonksiyonlar ile Lineerleştirilmiş SVM Modeli Örneği	25
Şekil 3.3. Kromozom Modeli	27
Şekil 3.4. Popülasyon Modeli.....	28
Şekil 3.5. Genetik Algoritma Tabanlı Sarmalama Öznelik Seçimi.....	29
Şekil 3.6. ‘Double Vector’ Tipinde Kromozom Örneği.....	30
Şekil 3.7. ‘Bit String’ Tipinde Kromozom Örneği.....	30
Şekil 3.8. Kromozom Aracılığı ile Sınıflandırıcıya Gönderilen Alt Öznelik Seti Örneği	31
Şekil 3.9. ‘Single Point’ Çaprazlama Örneği	36
Şekil 3.10. ‘Scattered’ Çaprazlama Örneği	36
Şekil 4.1. Deneysel Çalışma Aşamaları.....	41

GÖRSELLER DİZİNİ

Sayfa

Görsel 3.1. Yöntem Adımları.....	14
Görsel 3.2. Bir Numunenin 256×256 Çözünürlüğünde T1 Sekansı Gri-Seviye Kesiti .	16
Görsel 3.3. Bir Numunenin Koronal-Aksiyel-Sagittal Kesitleri.....	18
Görsel 3.4. Bir Numunenin Freesurfer ile Analizi Sonrası Üç Boyutlu Modeli.....	19
Görsel 3.5. Bir Numunenin aseg.stats Dosyasına Ait Kısmi Ölçülendirme Örneği	19
Görsel 3.6. Bir Numunenin lh.aparc.stats Dosyasına Ait Kısmi Ölçülendirme Örneği.	20
Görsel 3.7. “Sağ Yüzey Alanları” Öznitelik Grubuna Ait Kısmi Değerler	21

SİMGELER ve KISALTMALAR DİZİNİ

ADNI	: Alzheimer's disease neuroimaging initiative
AH	: Alzheimer hastalığı
DICOM	: Digital imaging and communications in medicine
FTD	: Frontotemporal demans
GA	: Genetik algoritma
GSS	: Genetik algoritma tabanlı sarmalama öznitelik seçimi
HBB	: Hafif bilişsel bozulma
MMSE	: Mini mental durum değerlendirme
MR	: Manyetik rezonans
MRG	: Manyetik rezonans görüntüleme
NB	: Naïve Bayes
NIfTI	: Neuroimaging informatics technology initiative
PACS	: Picture archiving communication systems
SBY	: Subjektif bellek yakınması
SVM	: Support vector machines
VaD	: Vasküler demans

1. GİRİŞ

Yaş ilerledikçe karşılaşılma oranı ciddi şekilde artan demans hastalıkları, nöropsikiyatrik bozukluk olarak tanımlanmaktadır. Bu tür hastalıklar, bilişsel bozukluklar ve zihin işlevlerinin yitilmesi, kişiliklerin değişkenlik göstermesi ile ilgilidir [1]. Demans hastalıklarının en sık karşılaşılanlarından bazıları Alzheimer hastalığı (AH), vasküler demans (VaD) ve frontotemporal demans (FTD) olarak sıralanabilir. 2003 yılında yapılan bir araştırmanın sonuçlarına göre, demanslı yaşlı sayısının 2030 yılında 63 milyonu, 2050 yılında ise 114 milyonu görebileceği tahmin edilmektedir [2].

Alois Alzheimer, sonralarda kendi adı ile anılacak, nedeni tam olarak bilinmeyen ancak 21. kromozomun uzun kolunda protein birikimi sonucu beyinde plakların oluşumu ile dikkat çeken hastalığı 1907 yılında tanımlamıştır [3]. AH, en sık karşılaşılan demans nedeni olarak vurgulanmaktadır [4]. Çeşitli nöropsikiyatrik ve davranışsal bozuklukların arttığı, bilişsel işlevlerin azaldığı hastalık ile yaşam kalitesi azalmakta, hatta ölüm ile de sonuçlanabilmektedir [5]. Tüm demans vakalarının yaklaşık olarak yarısını oluşturan, çoğu zaman giderek fazlalaşan unutkanlık şüphesi ile tespit edilen AH'nin karşılaşılma riskinin, diğer yaşla ilerleyen hastalıklara göre daha kesin bir şekilde yaşa bağlı olarak arttığı vurgulanmaktadır. 65 yaş üzerindeki nüfusta, Alzheimer hastalığı ile %3-11 arasındaki bir oranda, 85 yaş üzerinde ise %20-47 arasındaki bir oranda karşılaşılmaktadır [6]. 2005 yılında yapılan bir çalışmada, hastalığın dünya genelinde yaklaşık 24 milyon insanı etkilediği raporlanmıştır ve yaşam süresi ile etkileşimli olarak 2040 yılında 80 milyon kişiyi etkileyebileceği öngörülmektedir. Aynı araştırmanın sonuçlarına göre Türkiye'de ise 250-300 bin AH'ye sahip vatandaşımız olduğu tahmin edilmektedir [7].

Medikal çalışmalarda multi-enfarkt demans ismi ile de karşılaşılabilecek VaD, en sık rastlanma sıralamasında ikinci sırada bulunur; 50 yaşın üzerindeki erkeklerde karşılaşılma oranı daha fazladır. En sık rastlanma sıralamasında AH'yi takip eden VaD'nin belirtileri ve tanı ölçütleri ne kadar AH'ye benzese de, klinik tanı konulabilmesi için laboratuvar verisinin varlığı önemlidir [3]. Patolojik değişikliklerin yanı sıra AH'ye benzer nöropatolojik dejeneratif değişiklikler de görülebilir; başka bir ifade ile VaD'nin AH'ye klinik açıklamalar ile ne derece benzediği yapılan çalışmalarda gösterilebilmektedir [8].

Oldukça sık görülen bir diğer demans tipi ise, daha çok yürütücü işlevlerin bozulmasından kaynaklanan kognitif etkilenmenin karşılaştığı FTD'dir; sosyal ve kişisel ilişkilerdeki bariz değişiklikler ile fark edilebilir [9]. Belirgin davranış problemlerini tetikleyen, alışlagelmişin dışında bulguları olan, kişilik değişimi ve sosyal davranımında bozukluğun hastalık boyunca ön planda olduğu FTD diğer demanslarla karşılaştırıldığında, hiçbir diğer demans tipinde olmadığı kadar, %40-60 arasında bir oranda, aileseldir ve görülme sıklığının %3'ü geçmediği belirtilmektedir. Bu nedenle bu hastalık gündeme çok fazla gelememektedir. Bu demans tipini diğer demans tiplerinden ayırmak nispeten kolay olarak belirtilmektedir ancak diğer psikiyatrik bozukluklar da işin içine girdiğinde ayırım zorlaşabilmektedir [10].

Medikal testler sırasında, bunama şikâyeti ile gelen hastalar başta olmak üzere demans hastalıklarından birine sahip olması muhtemel hastaların beyin görüntüleri üzerinden kortikal hacim, kalınlık, yüzey alanı gibi beyin belirgin özellikleri gözlemlenerek klinik tanı konulmaya çalışılır.

Manyetik rezonans görüntüleme (MRG), vücut anatomisine ışık tutan görüntüleme tekniklerinden biridir. Radyo dalgaları ve mıknatıs kullanımı yardımı ile ilgili uzvun görüntüsünün oluşmasını sağlayan, radyasyon içermeyen bir tekniktir. Bir manyetik alan içerisinde sabit duran bir hastaya gönderilen atomlar içerisinde protonların mıknatıs üzerinden sapma etkisi ile ortaya çıkan enerjinin sinyale dönüştürülmesi üzerine kuruludur. T1 ve T2 sekansları, MRG'nin metodolojisini oluşturur. T2 ağırlıklı sekansta patolojik sinyal intensite farklılıkları saptanmaya çalışılırken T1 ağırlıklı sekansta yumuşak doku kontrastı hatasız elde edilebilir ve anatomik değerlendirme için zemin hazırlanmış olur [11].

MRG cihazlarından çıktı olarak *Neuroimaging Informatics Technology Initiative* (NIFTI) vb. girişimler kapsamında medikal görüntüleme standartlarına uygun görüntü dosyaları elde edilebilir. *nii* (NIFTI veri tipi) veya *dicom* (Digital Imaging and Communications in Medicine) dosya uzantılı medikal görüntü dosyaları beyni iki boyutlu dilimlenmiş resimler üzerinden üç boyutlu şekilde ifade etmeye çalışır. Bahsedilen dilimlenmiş beyin görüntüleri farklı eksenlere göre izdüşümsel iki boyutlu olarak elde edilebilir. Her bir görüntü dosyasının içerisinde, ilgili beyin bölümünün bir eksene göre iz düşüm görseli ve bu görsele ait görüntüleme çalışmasının nerede, ne zaman gerçekleştiği gibi teknik olmayan bilgilerinin haricinde görsel dosyanın teknik bilgileri de bulunur. Söz konusu görüntüleme çalışmaları çoğu zaman gri seviye olarak işlenir.

Görüntü işleme tekniklerinin kullanılması ile görsel dosyalar üzerinden arzu edilen ölçülemelere ulaşılabilir. Girdi-çıkı analizi yapılırken farklı ölçülemelerin farklı bir sonucu tetiklediği varsayıldığında, dikkat edilmesi gereken özelliklerin doğru tespit edilmesi büyük önem taşımaktadır. Her bir girdi örneği için, aranacak özellik görüntünün farklı yerlerinde olsa bile zorlanmadan bulunabilir olmalıdır. Görüntü dosyalarından elde edilen tanımlayıcı sayısal veri bloklarının bir düzlemde noktalar şeklinde ifade edilmesinden sonra, noktaların arasından çekilecek sınır çizgileri ile örnekleri kümelere ayırmak mümkündür. Teoride, benzer özellikleri taşıyan örneklerin aynı kümelere girmesi sağlanarak sınıflandırma yapılmış olur. Demans hastalıklarını içeren veri seti ile yapılacak sınıflandırma çalışmalarında klinik tanılarının doğru tahmin edilebilmesi için, medikal görsel dosyalar üzerinde beyin bölümlerine yönelik önem taşıyan noktalar belirlenmelidir.

Sınıflandırma işlemleri yapılırken girdi olarak gösterilen özelliklerden bazılarının diğerlerine göre çok daha kıymetli nitelendirilmesi mümkündür. Bazı özelliklerin ise yapılması istenilen sınıflandırmada etkisi olmayabilir. Bu doğrultuda, girdi olarak tüm öznelik seti yerine kıymetli özneliklerin seçilmesi, sınıflandırma tahmininin daha doğru olması adına önemlidir.

Bu çalışmanın ilk evrelerinde, AH, FTD ve VaD olmak üzere üç demans tipini içeren bir veri setine ait elde edilen tüm özneliklerinin birlikte kullanılması ile sınıflandırma çalışmaları yapılmıştır. Karşılaşılan doğruluk sonuçlarının %30 civarında bulunması ve tatmin edici olmayışı üzerine alt öznelik setlerinin tespit edilmesi ve doğruluk sonuçlarının yükseltilmesi, çalışmanın motivasyonunu oluşturmuştur. Bu doğrultuda bu çalışmada, değerli öznelik setlerinin tespit edilmesi amacı ile tüm öznelikler içerisinden diğerlerine göre daha kıymetli olabilecek özneliklerin, sarmalama metotlu öznelik seçim yaklaşımının farklı sınıflandırma algoritmaları üzerinden ne şekilde sonuçlandığı anlatılmaktadır.

1.1. Literatür Taraması

Medikal ve bilgisayar bilimleri arasındaki disiplinler arası çalışmalar kapsamında beyin görüntüleme alanında yapılan çalışmaların her geçen gün daha geniş bir alana yayıldığı söylenebilir. Özellikle son yıllarda bu alanda birçok önemli çalışma yapıldığı görülmektedir. Araştırmacıların ücretsiz erişimine sunulan standart veri setlerinin

kullanıldığı çalışmalar olduğu gibi kendi veri setleri üzerinden elde edilen bulgular ile literatüre katkıda bulunan çalışmaların sayısı da bir hayli fazladır. Son yıllarda uluslararası alanda medikal görüntülerin bir havuzda toplanarak kapsamlı veri setlerinin oluşturulması amacına hizmet eden iki önemli girişim dikkat çekmektedir. ABD merkezli Alzheimer Disease Neuroimaging Initiative (ADNI) ve Avrupa Birliği AddNeuroMed programı, AH gibi önde gelen demans hastalıkları için veri toplamak ve doğrulamak için yapılandırılmıştır. İki program da sırası ile Kuzey Amerika ve Avrupa'dan toplanılan MR görüntülerini araştırmacıların erişimine sunmaktadır. Detay olarak ise, tüm programlardan elde edilecek bulguların kendi içerisinde karşılaştırılabilirliğinin sağlanabilmesi amacı ile AddNeuroMed programının ADNI ile uyumlu olacak şekilde tasarlandığı belirtilmektedir [12]. AddNeuroMed veri seti, University of Kuopio, Finlandiya, University of Perugia, İtalya, Aristotle University of Thessaloniki, Yunanistan, King's College London, Birleşik Krallık, University of Lodz, Polonya ve University of Toulouse, Fransa olmak üzere 6 farklı üniversitenin işbirliği ile ortaya çıkmıştır [13]. ADNI veri seti ise Kuzey Amerika'da bulunan 55 merkezi katılımcı sayesinde toplanmıştır; ayrıca, bazı hastalar için bir hastaya ait farklı zaman dilimlerinde elde edilen beyin görüntülerini de içerisinde barındırmaktadır [14]. Her iki veri seti de AH, hafif bilişsel bozulma (HBB) gibi başlıca hastalıklara sahip hastaların ayrıca sağlıklı kontrol bireylerinin MR görüntülerini, klinik tanıları gibi çalışmalar sırasında kullanılacak veriyi içermektedir.

Beyin görüntüleme, modelleme ve demans hastalıklarının sınıflandırılması gibi bilgisayar bilimleri çalışmalarında beynin sanal olarak inşa edilmesi, modellemelerden öznitelik çıkarılması gibi işlemler için bazı yazılım araçları da kullanılabilir. Literatürde yardımcı yazılım araçları ile elde edilen özniteliklerin sınıflandırma işlemlerinde kullanılması üzerine kurulu çalışmaların sayısı oldukça fazladır.

2013 yılında AddNeuroMed topluluğu içerisinde 116 AH hastası, 119 HBB hastası ve 110 kontrol bireyleri olmak üzere toplam 345 numunenin bulunduğu veri seti ile yapılan bir deneysel çalışmada karar ağaçları, yapay sinir ağları, destek vektör makinaları (SVM) gibi farklı tekniklerin belirtilen veri seti üzerindeki sınıflandırma performansı incelenmiştir. FreeSurfer v4.5.0 beyin analiz yazılım aracı, görüntü işleme ve öznitelik çıkarımı için kullanılmıştır. Analiz sonrasında elde edilen 34 kortikal kalınlık ve 23 hacimsel veri olmak üzere toplam 57 ölçü, çok değişkenli analizlerin girdisi olarak belirlenmiştir. Çoğu numunede 0 olarak hesaplanan bazı özniteliklerin analizler sırasında

kullanılacak öznitelikler listesinden çıkarıldığı belirtilmiştir. Sağ ve sol beyin bölümleri için ayrı ayrı hesap edilmiş beyin ölçülendirmelerinin ortalamaları alınarak öznitelik sayısının azaltıldığı, bu sayede verinin yorumlanabilirliğinin daha kolay hale geldiği ve tahmin sonuçlarının önemli bir şekilde değişme uğramayacağı ifade edilmiştir. Hacimsel ölçülendirmeler hastanın kafatası içi hacmine göre normalize edilmiş olup kalınlık ölçülendirmeleri ham haliyle işleme alınmıştır. 10 katlamalı çapraz doğrulama test tekniği kullanılarak farklı sınıflandırma algoritmaları ve parametreler ile doğruluk sonuçları hesaplanmıştır. AH v kontrol bireyleri sınıflandırmasında farklı teknikler üzerinden bireyin yaşı, eğitim seviyesi gibi çeşitli ilave parametrelerin de eklenmesi ile doğruluk sonuçları %81.4 ve %88.1 arasında değişen değerler ile raporlanmıştır. SVM sınıflandırma modeli gerçekleştirirken önem sıralamasında üst sıralarda gözlemlenen özellikler *hippocampus* (tek başına en önemli), *amygdala*, *entorhinal cortex*, *parahippocampal gyrus*, *superior temporal gyrus*, *inferior lateral ventricle*, *orta temporal gyrus*, *insula* olarak belirtilmiştir. Tüm teknikler ile yapılan analizler sonucunda en değerli öznitelikler ise *hippocampus*, *amygdala*, *entorhinal cortex*, *inferior lateral ventricles*, *cerebrospinal uid*, *inferior*, *superior*, *middle temporal gyri* ve *temporal pole* olarak listelenmiştir [13].

2011 yılında gerçekleştirilen başka bir çalışmada ADNI veri tabanına kayıtlı, bireylerin T1 sekanslı MR görüntüleri kullanılmıştır. Kontrol bireyleri, ayrıca AH ve HBB hastalıklarına sahip yaşları 55 ve 90 arasında değişen 509 numunenin beyin taramaları Freesurfer yazılım aracına girdi olarak gösterilerek her bir numuneye ait kortikal öznitelikler çıkarılmıştır. Ayrıntılı bir şekilde veri setindeki bazı numunelerin Freesurfer ile modelleme sürecinden başarılı çıkamadığı da not edilmiştir. Aynı çalışma setindeki bireylerin çeşitli beyin bölgeleri üzerinden farklı sınıflandırma algoritma sonuçlarının karşılaştırılması amaçlanmıştır. 81 kontrol bireyi, 106 HBB hastası ve 69 AH hastası eğitim ve optimizasyon için ayrılmış; kalan 81 kontrol bireyi, 104 HBB hastası ve 68 AH hastası metotların performansının belirlenmesi aşamasında kullanılmıştır; başka bir ifade ile veri seti neredeyse yarı yarıya eğitim ve test amaçlı bölünmüştür. Öznitelikler çıkarıldıktan sonra normalize edilmiştir. Bununla birlikte, farklı gruplandırmalar yapılarak, bir başka deyiş ile farklı yaklaşımlar kullanılarak test aşamaları gerçekleştirilmiştir. Voksel esaslı (ele alınan vokseldeki gri madde, beyaz madde ve *cerebrospinal uid*), verteks esaslı (kortikal yüzeydeki öznitelikler) veya istenilen bölge esaslı (sadece *hippocampus* içerir) gruplandırmalar ile yaklaşımlar

belirlenmiştir. Bu yaklaşımların başarısı %84'ün üzerinde olarak raporlanmıştır. AH v kontrol bireyleri, HBB v kontrol bireyleri ve iki HBB seviyesinin farkını ölçmeye yönelik toplamda üç farklı 2-sınıf temelli sınıflandırma testleri yapılmıştır. AH v kontrol bireyleri sınıflandırmasının doğruluk sonuçlarının AH v HBB sınıflandırmasının doğruluk sonuçlarından yüksek olduğu belirtilmiştir. Hastalığın seviyesi, cinsiyet, yaş, genlerin yapısı, eğitim seviyesi, MR görüntülerinin kalitesi gibi farklı türlerden birçok etkenin başarı sonucunu etkileyebileceği yorumlanmıştır [15].

Freesurfer v4.3 beyin analiz yazılım aracı kullanılarak 2011 yılında ADNI veri tabanına kayıtlı numuneler üzerinden özniteliklerin çıkarıldığı başka bir çalışmada, lojistik regresyon, SVM, radyal tabanlı fonksiyon, C4.5 ağacı gibi çeşitli teknikler üzerinden bulgular kıyaslanmıştır. 813 numune yazılım aracına girdi olarak alınmış ancak 239 tanesi süreci başarı ile tamamlayamamıştır. Cinsiyet dengesini sağlayabilmek için 50 adet erkek HBB numunesinin verisi de rastgele elimine edildikten sonra toplamda 524 AH ve HBB hastasının, ayrıca kontrol bireylerinin T1 MR görüntüleri analiz edilmiştir. Freesurfer analizi sonrası her bir çıktının tümleyici bilgi taşıyabileceği düşünülerek tüm çıktılar bir sonraki adıma geçirilmiştir. Böyle olunca bir numune için 328 sayısal değer bir vektör şeklinde ele alınmıştır. Hacimsel değerler 0-1 aralığında normalize edilerek işlemlere eklenmiştir. MR özniteliklerinin yanı sıra yaş, cinsiyet, eğitim seviyesi ve ApoE ε4 sayısı da ilave öznitelikler olarak belirlenmiştir. Öznitelik seçimi için hiçbir sınıflandırıcıya bağlı olmayan filtreleme metodu kullanılmıştır. Sınıflandırma parametreleri optimize edildikten sonra sınıflandırıcılar 10 katlamalı çapraz doğrulama ile test edilmiştir. Bütün sınıflandırma testlerinde ortak olarak kıymetli bulunan öznitelikler yaş, ApoE ε4 sayısı, *sağ ve sol hippocampal*, *sol entorhinal cortex*, *sol amygdale hacimleri* ve *sol orta temporal korteksin ortalama kortikal kalınlığı* olarak listelenmiştir. Ayrıca, kontrol bireyleri v AH sınıflandırmasında *sol inferior* ve *orta temporal korteks*'in hacim ve kalınlığı gibi daha başka özellikler de değerli bulunmuştur. Elde edilen en iyi sonuca SVM sınıflandırıcısı ile kontrol bireyleri v AH sınıflandırmasında karşılaştığı ve ortalama doğruluğun 89.17 ± 5.08 olarak görüldüğü raporlanmıştır [16].

2015 yılında araştırmacıların kendi veri seti ile yapmış olduğu bir çalışmada Pusan Üniversitesi Hastanesi'nin 2010–2012 yılları arasındaki kayıtları içerisinden, 27 subjektif bellek yakınması (SBY), 18 HBB, ve 27 AH hastalarına ait veri kullanılmıştır. MR görüntülerinin *3T Siemens Trio Tim System* (Erlangen, Almanya) cihazından elde edildiği

belirtilmiştir. Açık kaynak kodlu Freesurfer v5.1 beyin analiz yazılım aracı üzerinden 256×256 çözünürlüğünde ve 1 mm kesit kalınlığındaki T1 sekansı görüntü setleri girdi olarak gösterildikten sonra beynin yapısı dijital olarak inşa edilmiştir. Analiz çıktılarından kortikal kalınlıklar ve hacimler, ayrıca beyaz madde bütünlüğü bilgileri öznitelik setini oluşturmuştur. 10 katlamalı çapraz doğrulama ile en uygun sınıflandırma parametreleri belirlendikten sonra veri setinin her grubu için %80'i eğitim seti ve kalan %20'si de test seti olacak şekilde parçalanmıştır. Lineer olmayan SVM sınıflandırıcısı ile sol ve sağ beyin çıkıntılarındaki en önemli ve en gereksiz öznitelikler belirlenmiştir. Her adımda farklı eğitim konfigürasyonları oluşturularak prosedür 1000 kez tekrarlanmış olup ikişerli kıyasta doğruluk ortalaması %84.4 ile %96.3 arasında değişirken, üç sınıfın kıyasında doğruluk ortalaması %70.5 (± 11.5) olarak raporlanmıştır. Gri ve beyaz maddedeki değişimin bilişsel profil ile ilişkili olabileceği ve nörodejeneratif etiyoloji esaslı çalışmalarda bulunan öznitelikler ile karşılaştırılabileceği yorumlanmıştır. Ek olarak, AH araştırmalarında *hippocampus* özneliğinin tek başına ne kadar önemli bir öznitelik olduğu bilinmesine rağmen farklı öznitelik kombinasyonlarının da denendiğini, bu doğrultuda yüksek doğruluk sonuçlarının elde edildiği de çalışma kapsamında belirtilmiştir [17].

2011 yılında, alanının iki büyük standart veri havuzu olarak tanımlanan AddNeuroMed ve ADNI veri tabanlarının ikisinin de kullanıldığı çalışmada 295 AH hastası, 444 HBB hastası ve 335 kontrol bireyine ait toplamda 1074 MR görüntüsü üzerinden numunelere ait öznitelikler Freesurfer v4.5 kullanılarak elde edilmiştir. Analiz sonrasında 34 sağ ve 34 sol beyin bölüm özneliği olmak üzere 68 kortikal kalınlık ve ayrıca 50 adet hacimsel öznitelik beyin analiz yazılım aracı tarafından çıktı olarak sağlanmıştır. Sağ ve sol beyin bölümlerinin ortalaması alınarak öznitelik sayısı azaltılmıştır. Ek olarak, hacim özniteliklerinden *beyaz madde hypointensities*, *optic chiasm*, *sağ ve sol vessel*, *sol ve sağ choroid plexus* bölümleri daha sonra yapılacak detaylı analizler düşünülerek çalışma kapsamından çıkarılmıştır. Bu doğrultuda 23 hacim 34 kortikal kalınlık öznitelikleri olmak üzere toplamda 57 öznitelik üzerinden dik kısmi en küçük kareler yaklaşımı ile çok değişkenli analizi yapılmıştır. Tüm hacim ölçüleri hastanın kafatası içi hacmine göre normalize edilmiştir. Öznitelik seçimi yapılmamıştır, başka bir ifade ile belirtilen toplamda 57 öznitelik üzerinden direkt analiz işlemlerine geçilmiştir. Test aşaması 7 katlamalı çapraz doğrulama ile tamamlanmıştır. Beynin *medial temporal lobe* yapıları, *entorhinal cortex*, *hippocampus* ve *amygdala* bölümlerinin

hastaların sınıflandırmasında en belirleyici öznitelikler olduğu ifade edilmiştir. AH v kontrol bireyleri sınıflandırmasında beynin diğer *temporal* ve *limbic* bölgelerin de önem teşkil ettiği belirtilmiştir. Her iki veri tabanına ait bulunan belirleyici özniteliklerin sırasının nispeten farklı olduğu, genel anlamda yapıların benzer bulunduğu ayrıntıları ile ifade edilmiştir. Bu konuya örnek olarak da beynin *parahippocampal gyrus*, *temporal pole*, *cingulate cortex isthmus*, *inferior*, *superior* ve *middle temporal gyrus* bölümleri gösterilmiştir. Daha da fazlası, beynin *brainstem*, *cerebellum*, *pallidum* ve *corpus callosum* bölümlerinin de sınıflandırma kapsamında neredeyse hiç önem teşkil etmediği açıkça dile getirilmiştir [12].

Wien Center for Alzheimer's Disease ve Memory Disorders, Mount Sinai Medical Center, Miami Beach, FL kurumlarından 2005-2008 yılları arasındaki veriye dayanarak alınan 59 AH hastası, 123 HBB hastası ve 127 kontrol bireyelerine ait toplamda 309 katılımcının beyin öznitelikleri FreeSurfer v5.1.0 beyin yazılım aracı üzerinden hesaplanmıştır. MR görüntülerinin yanı sıra nöropsikolojik test, mini mental durum değerlendirme (MMSE) sonuçları da ilave öznitelikler olarak çalışma kapsamına dâhil edilmiştir. MR görüntüleri, *Siemen's Symphony* (Iselin, NJ, USA) veya *General Electric, HDX* (Milwaukee, Wisc., USA) makinelerinden elde edilmiştir. Her bir görüntü seti 1.5 mm kesit kalınlığına sahip olup görüntü çözünürlüğü 256×256 olarak belirtilmiştir. MR görüntülerinin FreeSurfer programı ile analizi sonrasında 45 kortikal bölge ve 10 yapı bilimsel istatistikleri olmak üzere 55 adet hacim değeri elde edilmiştir. 45 kortikal bölgeden 4 tanesi tüm numuneler için aynı değere sahip olduğundan dışlanmıştır. Dolayısı ile her bir MR görüntü setinden 41 adet bölge ve 10 adet yapı bilimsel istatistik bilgileri olmak üzere 51 öznitelik temin edilmiştir. Sonrasında bu değerler kafatası içi hacmi, yaş ve eğitim seviyesi gibi değerler ile işlenerek nitelendirilmiştir. Güncellenen hacim öznitelikleri ve MMSE skorları ile her bir katılımcıyı tarif eden 52 değere sahip ayırıştırıcı vektörler oluşturulmuştur. Numuneler, SVM algoritması kullanılarak 2 katlamalı çapraz doğrulama tekniği ile sınıflandırılmıştır. Her bir sefer kendi içerisinde bağımsız olacak şekilde, ayrıca eğitim ve test setleri rastgele yaratılacak şekilde prosedür 50 defa tekrarlanmıştır. Çalışma kapsamında MMSE testlerinin önemi, tüm sınıflandırma deneylerinde sınıflandırma doğruluğunu arttıran unsur yorumlaması yapılarak vurgulanmıştır. AH v kontrol bireyleri sınıflandırmasında, MMSE skorları ile birleştirildiğinde en belirleyici hacimsel öznitelikler *sağ hippocampus* ve *sol inferior lateral ventricle* olarak bulunmuştur. Ortalama doğruluk değeri %92.4 olarak not

edilmiştir. Aynı sınıflandırmada en değerli 15 öznelik ise MMSE, *sağ hippocampus*, *sol inferior lateral ventricle*, *sol amygdala*, *sol hippocampus*, *sağ inferior lateral ventricle*, *sağ hemisphere cortex hacmi*, tüm gri hacim, *sağ amygdala*, *cortex hacmi*, *sol hemisphere cortex hacmi*, *sol chroid plexus*, *sağ chroid plexus*, *3rd ventricle* ve *sağ lateral ventricle* olarak sıralanmıştır. Ek olarak, bazı HBB türlerinde *hippocampus atrofisinin* hafıza problemlerine yol açan en önemli etkenlerden biri olarak yorumlandığı belirtilmiştir [18].

Bu tez çalışmasının başlangıç evrelerinde yapay sinir ağları üzerinden yapılan analizlerde, bu çalışmada kullanılan numunelerden 32 tanesi ele alınmıştır; yani bu çalışmada kullanılan AH, FTD, VaD hastalıkları ile oluşan veri setinin yaklaşık olarak yarısı ile çalışılmıştır. Freesurfer v5.3.0 beyin analiz aracından elde edilen beyin bölümlerinin hacimleri, kalınlık ortalaması, yüzey alanı gibi öznelik grupları, herhangi bir öznelik elemesi yapılmadan sınıflandırma algoritmasına ayrı ayrı girdi olarak gösterilmiştir. Test aşaması, tek çıkarımlı çapraz doğrulama test tekniği ile gerçekleştirilmiştir. Detaylı olarak, öznelik matrislerindeki her bir numunenin sınıfı adlandırılmak için işleme sokulduğunda, veri setindeki diğer numuneler eğitim seti olarak ele alınmıştır. Doğruluk sonucu, klinik tanısı doğru tahmin edilen numune sayısının tüm numune sayısına oranı üzerinden hesap edilmiştir. Sınıflandırma başarısı ise %87.5'i bulan doğruluk oranları ile raporlanmıştır [19].

Öznelik seçim metotları öznelik matrisinin boyutunun küçültülmesi, değerli özneliklerin bulunması ve etkisiz özneliklerin çıkarılması amacı ile birçok alanda kullanılmaktadır. Literatürde bu amaç için kullanılan çok fazla yöntem bulunmaktadır. Temel olarak en uygun veya en uyguna yakın sonuç bulan algoritmalar şeklinde ikiye ayrılır. En uygun sonucu bulan *exhausted search*, *branch and bound* gibi yaklaşımlar olası bütün kombinasyonları deneme üzerine kuruludur. En uyguna yakın sonuç bulma algoritmaları ise sezgisel metotlar olarak adlandırılır ve kendi içerisinde belirleyici veya olasılıksal olacak şekilde tek veya çoklu sonuç elde eder. Tek sonuç bulan belirleyici algoritmalara, belirli bir karar ile karşılaşana dek öznelikleri ekleyip çıkararak *max-min*, *sequential forward selection*, *sequential backward selection* gibi öznelik seçim algoritmaları örnek verilebileceği gibi olasılıksal ve çoklu sonuç bulan öznelik seçim algoritmalarına da genetik algoritma (GA) örnek verilebilir [20].

Bu çalışmada da, Freesurfer beyin analiz yazılım aracı ile elde edilen beyin bazı bölümlerine ait hacim, kalınlık, yüzey alanı gibi sayısal veriden farklı öznelik grupları

oluşturulmaktadır. Bu gruplar içerisinde diğerlerine göre daha kıymetli öznitelikleri GA yönteminin uygulanması ile saptamak esas amaçtır.

1.2. Tez Organizasyonu

Bu tez çalışmasında genel olarak öznitelik tabanlı imge sınıflandırma üzerinde araştırma yapılmıştır. Bu kapsamda, öznitelik seçiminin farklı sınıflandırma algoritmaları ile sonuçları elde edilerek karşılaştırmalar sunulmuştur. Karşılaştırma kıstası olarak doğruluk ölçütü kullanılmıştır. Doğruluk sonuçlarını bire bir etkileyen kıymetli öznitelikler listelenmiştir. Önerilen tez çalışması genel olarak aşağıdaki adımlardan oluşmaktadır:

- İlk olarak, yazılan program *script*leri ile girdi imge setleri analiz edilerek öznitelik çıkarımı için kullanılıp kullanılmayacağına göre kendi içerisinde ayrıştırılır.
- Öznitelik çıkarımı için uygun biçimdeki dosyalar yardımcı yazılım aracına girdi olarak gösterilir.
- Yardımcı yazılım aracı ile yapılan analizler sonrasında beyne ait modellemeler ve istatistiksel çıktılar elde edilir.
- Numune esaslı elde edilen tüm istatistiksel çıktılar, yazılan *Linux bash script*leri aracılığıyla birleştirilerek sınıflandırma analizlerinde kullanılacak öznitelik matrisleri oluşturulur.
- Deneysel çalışmaların birinci aşamasında, farklı öznitelik gruplarını niteleyen girdi matrislerinin öznitelik seçim ve sınıflandırma algoritmaları üzerinden farklı parametreler ile birlikte kullanılması sonucunda kendi içerisindeki değerli bilgileri elde edilir.
- Deneysel çalışmaların ikinci aşamasında ise, deneysel çalışmaların birinci aşamasında bulunan değerli bilgilerin birleştirilmesi işlemleri üzerinden nihai sonuçları elde etmek için birinci aşamada izlenen prosedür özgün parametreler ile tekrarlanır.
- Birinci ve ikinci aşamaya ait doğruluk sonuçları ve hata matrisleri kıyaslanmak üzere bulunur.

Tezin bölümleri ise şu şekilde özetlenebilir:

Bölüm 1’de genel olarak, demans hastalıklarının tanımlanmasından, klinik tanılarının konulmasından ve hastalıkların Türkiye ve dünya çapındaki istatistiklerinden; ayrıca,

tıbbi görüntüleme teknikleri detaylandırılarak elde edilen görüntü setlerinin sınıflandırılma yaklaşımında nasıl ele alınacağından bahsedilmektedir.

Bölüm 2’de, çalışma kapsamında kullanılan, öznelik değerlerinin çıkarıldığı veri seti detaylandırılmaktadır.

Bölüm 3’te, öznelik çıkarma, sınıflandırma ve öznelik seçimi süreçlerinde kullanılacak yöntemler tarif edilmektedir.

Bölüm 4’te, deneysel çalışmaların farklı aşamalarında kullanılan algoritmalara ve algoritmalar esnasında kullanılan özgün parametrelere dair izlenen yöntem adımları aktarılmaktadır, sonrasında ise elde edilen değerli öznelikler ve test sonuçları paylaşılmaktadır.

Bölüm 5’te, izlenen yöntemler ve bulgular ile ilgili kıyaslamalar aracılığıyla değerlendirmeler yapılmakta, yorumlamalarda bulunmaktadır. Son olarak, çalışmanın gelecekteki planlarından bahsedilmektedir.

2. VERİ SETİ

Bu çalışmada kullanılan veri seti, Eskişehir Osmangazi Üniversitesi Tıp Fakültesi Radyoloji Bölümü'ne ait *Picture Archiving Communication Systems* (PACS) veri tabanının kısmi verisinden oluşmaktadır. MR modellemesine sahip hasta beyin görüntülerinin anonim olması sağlanarak, başka bir ifade ile kişi bilgisi olmaksızın, beyin bölgelerinin hacim, kalınlık, yüzey alanı gibi ölçülerinin saptanması amacı ile kullanılmıştır.

Ele alınan görüntü setlerine ait çalışma kapsamında dikkat çeken teknik özellikler Tablo 2.1'de, görüntü setlerinin elde edildiği hastaların muayenesine ait genel bilgiler ise Tablo 2.2'de sunulmaktadır.

Tablo 2.1. *Kullanılan Görüntü Setlerinin Teknik Özellikleri*

Özellik	Değer
Format	DICOM
Format Versiyonu	3
Çözünürlük	Minimum 256×256
Renk Tipi	Gri-seviye
Bit Derinliği	12 veya 16
Görüntüleme Cihazı	Discovery MR750w ve Magnetom Vision plus
Görüntüleme Frekansı	127.73 veya 63.61
Modelleme	MR
Sekans	T1
Toplama Tipi	2D
Kesit Kalınlığı	4.5 – 5 mm
Kesitler Arası Boşluk	Değişken

Tablo 2.2. *Muayene Özellikleri*

Özellik	Değer
Muayene Tarihi	2014-2015 yıllarında muayene edilmiş hastalara ait görüntüler
Klinik Tanılar	AH, FTD, VaD
Hastaların Cinsiyetleri	Erkek ve Kadın
Hastaların Yaşları	Minimum 50 ve maksimum 90 olmak üzere değişken yaşlar

Veri seti içerisindeki dilimlenmiş iki boyutlu görüntü setleri, *Discovery MR750w* (GE, Milwaukee) ve *Magnetom Vision plus* (Siemens, Erlangen) tıbbi görüntüleme cihazlarından elde edilen T1 sekanslı MR görüntülerinden; biraz daha detaylı olarak, AH, FTD veya VaD hastalıklarından birine klinik tanı konulmuş her iki cinsiyetten de hastanın bulunduğu 63 örnekten oluşmaktadır. Veri seti içerisindeki örneklere ait klinik tanıların

cihaz tipine göre dağılımı Tablo 2.3'te, cinsiyete göre dağılımı ise Tablo 2.4'te sunulmaktadır.

Tablo 2.3. *Tıbbi Görüntü Cihazına Göre Hastalık Dağılım Tablosu*

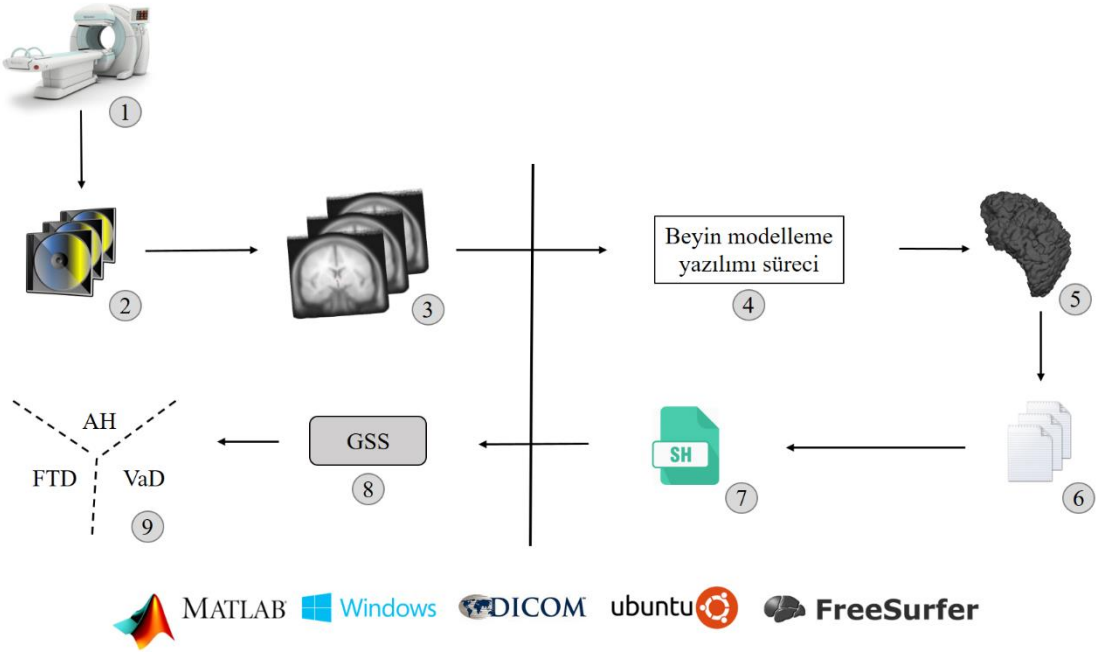
	Discovery MR750w	Magnetom Vision plus	Toplam
AH	13	6	19
FTD	13	6	19
VaD	17	8	25
Toplam	43	20	63

Tablo 2.4. *Cinsiyete Göre Hastalık Dağılım Tablosu*

	Erkek	Kadın	Toplam
AH	5	14	19
FTD	6	13	19
VaD	12	13	25
Toplam	23	40	63

3. YÖNTEM

Klinik tanısı belli olan beyin görüntülerine yönelik sınıflandırma çalışmasında izlenen yöntem adımları genel olarak birkaç basamaktan oluşmaktadır. İki farklı işletim sistemi ve farklı teknolojiler üzerinden ilerletilen çalışmaya ait yöntem basamakları Görsel 3.1’de sunulmaktadır.



Görsel 3.1. Yöntem Adımları

- 1. adım klinik muayeneleri belirtmektedir. Hastalar, manyetik alana sabitlenerek görüntüleme çalışması yapılır.
- 2. adımda, görüntüleme çalışması sonucunda oluşan dijital veri disklere yazdırılır. Hastalara ait farklı görüntü veri setleri elde edilir.
- 3. adımda, disklerden elde edilen farklı modellemelere sahip görüntü setleri, yardımcı beyin analiz yazılım aracına girdi olarak gösterilebilmek adına uygun olup olmadığı durumuna göre ayrıştırılır.
- 4. adımda, uygun formatta olduğu doğrulanan görüntü setleri yardımcı beyin modelleme yazılım aracı üzerinde analiz edilir.

- 5. adımda, yardımcı yazılım aracının çıktısı olarak verdiği beynin 3 boyutlu modelleme dosyaları oluşturulur.
- 6. adımda, modellemeye ilişkin beynin ilgili bölümlerinin ölçülendirmeleri, numunelerin beyin öznitelikleri, yardımcı programın analizi sonucunda istatistiksel veri olarak ifade edilir.
- 7. adımda öznitelikler, istatistiksel veri üzerinden yazılan program parçacıkları ile çekilir.
- 8. adımda, bir önceki adımda elde edilen değerler üzerinden öznitelik matrisleri, ilerleyen kesimlerde anlatılacak genetik algoritma tabanlı sarmalama öznitelik seçim (GSS) algoritması uygulanmak üzere oluşturulur.
- 9. adımda ise, her bir numunenin istatistik bilgileri kullanılarak oluşturulan öznitelik matrisleri sınıflandırma ve öznitelik seçim algoritmaları ile birlikte işlenir.

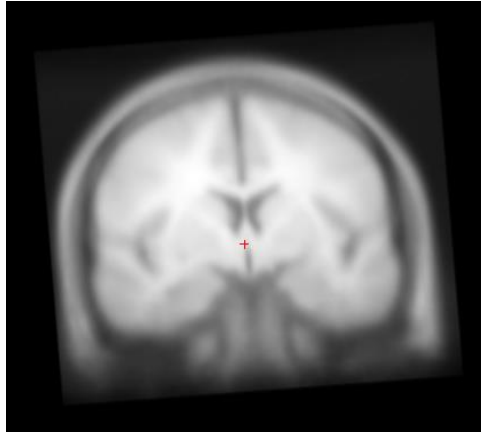
Sarmalama yaklaşımı esaslı öznitelik seçim algoritması olarak GA, Naïve Bayes (NB) ve SVM sınıflandırma algoritmalarının birlikte kullanılması ile gerçekleştirilmiştir. Son olarak, doğru tahmin edilen klinik tanıların tüm tahminlere oranı üzerinden doğruluk sonuçları hesaplanmış, bu adımı gerçekleştirebilmek için doğru ve yanlış tahminleri niteleyen hata matrisleri de belirtilmiştir. Ayrıca, tüm bu süreçler içerisinde bazı beyin bölümlerinin diğerlerine göre ne derece kıymetli bulunduğu da somut olarak gösterilmeye çalışılmıştır.

İzleyen kesimde yöntemin adımları detaylandırılmaktadır.

3.1. Dosyaların Analizi

Görsel dosyalar üzerinden beynin ilgili bölümlerine yönelik ölçülendirmelere ulaşabilmek için daha önceki başlıklarda bahsedilen iki boyutlu dilimlenmiş beyin görüntü dosyalarını görüntü işleme teknikleri ile ele almak gerekir. Ancak bazı durumlarda elde edilen tıbbi görüntü seti içerisinde birbirinden bağımsız özelliklerde alt görüntü setleri ile karşılaşılabilir. Başka bir deyişle beyin T1, T2 sekansları vb. gibi farklı şekillerde ifade eden alt görüntü setleri tek bir görüntü seti içerisinde depolanmış olabilir. Bu bağlamda alt setlerin birbirinden ayrıştırılması gerekebilir. Literatür taramasında da bahsedildiği gibi, içerisindeki görüntülerin birbirine bağımlılığı

ile beyni ifade etmeye çalışan belirli tipteki görüntü setleri yardımcı yazılım araçlarına girdi olarak gösterilebilmektedir. Bu tez çalışması kapsamında da, T1 sekanslı görüntü setleri beynin ilgili bölümlerine ait ölçülendirmelere ulaşmak amacı ile yardımcı yazılım aracının girdisi olarak uygulanmıştır. Tüm numunelere ait farklı farklı modellemelere sahip yüzlerce görüntü seti içerisinde, görüntü dosyalarının başlık kısmındaki bilgileri kullanan MATLAB alt programları ile T1 sekanslı setler diğerlerinden ayrıştırılmıştır. Görsel 3.2'de veri setinden rastgele seçilen bir numunenin T1 sekanslı dilimlenmiş görüntü seti içerisinde herhangi bir kesiti örnek olarak gösterilmektedir.



Görsel 3.2. Bir Numunenin 256×256 Çözünürlüğünde T1 Sekanslı Gri-Seviye Kesiti

Görüntü setlerinin uygunluğuna göre ayrıştırılmasından sonra özniteliklerin elde edilebilmesi için genel yöntemdeki bir sonraki adıma geçilmiştir.

3.2. Öznitelik Çıkarma

Bu tez çalışmasında sınıflandırma sürecine girdi olarak ele alınan beyin görüntüleri üzerinden istenilen beyin bölümlerinin ölçülendirmelerine ulaşabilmek amacı ile Freesurfer v5.3.0 beyin analiz ve sanal olarak modelleme yazılım aracından faydalanılmıştır. Veri setinin her bir numunesi için beyin analiz ve modelleme işlemleri, Freesurfer programı kullanılarak tamamlanmıştır. Freesurfer, birtakım görüntü işleme, numerik, vb. algoritmalar dizisini içerisinde barındıran, insan beyninin fonksiyonel ve yapısal olarak analiz edilmesine olanak sağlayan bir programdır [21]. *Linux, Mac OS X* gibi işletim sistemlerinde açık kaynak kodlu olarak kullanılabilen program, kendi sitesinden ücretsiz olarak indirilip lisanslanabileceği gibi, programın kullanımına veya

herhangi bir aşamasında karşılaşılan hataların çözümüne yönelik dokümanların sayısı da oldukça fazladır.

Freesurfer, ilk olarak aldığı *dicom* veya *nii* türü dilimlenmiş tıbbi görüntü dosyalarını üst bilgileri yardımı ile yapısal olarak birleştirir. Sonrasında, görüntü işleme tekniklerinin sayısal olarak yinelemeli bir şekilde yönetilmesi ile beyin üç boyutlu olarak modellenir [22]. Bir beynin tam olarak modellenmesi için izlenen süreç adımları¹ yazılım aracının kullanım kılavuzundan çıkan bilgilerin derlenmesi ile Tablo 3.1’de sunulmaktadır. Her bir numune için aynı prosedür tekrarlanmaktadır.

Tablo 3.1. *Freesurfer Beyin Yapılandırma Aşamaları*

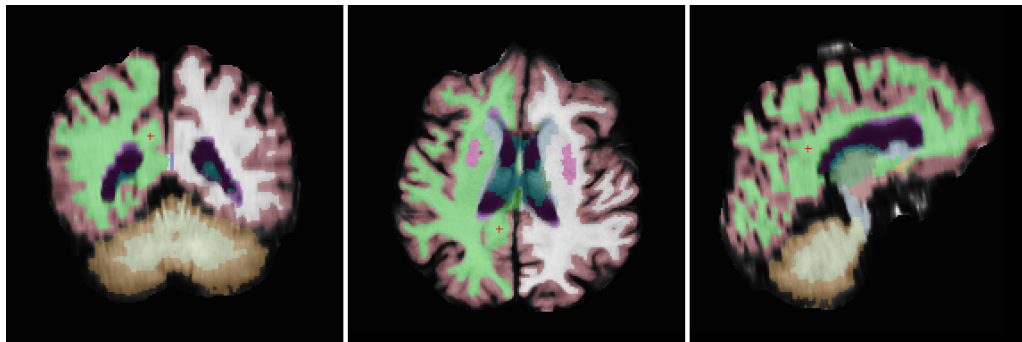
Ana Adım	Alt Adım
Kortikal Yapılandırma #1	Hareket düzenleme ve uygunluk Düzenli olmayan yoğunluk normalizasyonu Talairach hesaplama dönüşümü Yoğunluk normalizasyonu #1 Kafatası çıkarma ve yedekleme
Kortikal Yapılandırma #2	Lineer hacimsel kayıtlama GCA model temelli yoğunluk normalizasyonu GCA model temelli lineer olmayan hacimsel kayıtlama Boyun çıkarma Kafatası ile lineer hacimsel kayıtlama GCA model temelli hacimsel etiketleme ve istatistikler Yoğunluk normalizasyonu #2 Beyaz madde segmentasyonu Beyaz madde düzenlemesi Doldurma Mozaik döşeme Pürüzsüzleştirme #1 Şişme #1 Küreleştirme Otomatik topoloji düzelticisi Son yüzeylerin oluşturulması Pürüzsüzleştirme #2 Şişme #2
Kortikal Yapılandırma #3	Küresel haritalama Küresel kayıtlama Küresel kayıtlama, karşı taraflı yarımküre Ortalama eğrilik haritalaması Kortikal parselasyon Kortikal parselasyon istatistikleri Kortikal şerit maskesi Kortikal parselasyon haritalaması

¹Freesurfer ile beyin modellemek için gerekli tüm süreç adımlarına ve açıklamalarına daha detaylı olarak <https://surfer.nmr.mgh.harvard.edu/fswiki/recon-all> adresinden erişilebilir.

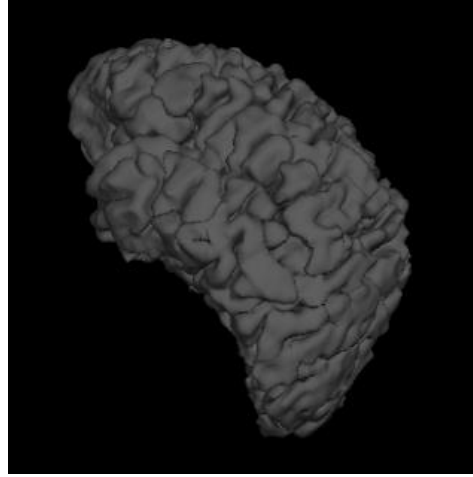
Tablo 3.1'deki bilgilerden de yararlanılarak Freesurfer'in çalışma prensibinin temelinde 3 ana mantıksal aşamadan ibaret olduğu söylenebilir. Beyin, bu üç ana adımlık sürecin ardından sanal olarak inşa edilir. İlk olarak, girdi dosyalarının doğrulanması adımı işleme alınır. Sonrasında, hacim ölçme ile ilgili ana işlemlerin yanı sıra modellemeden boyun bölgesinin çıkarılması, beyaz cisimlerin segmentasyonu, geçiş yumuşatma gibi görsel bazı alt işlemler gerçekleştirilir. Son ana adımda da küresel haritalama, kayıtlama gibi işlemleri takip eden kortikal parselasyon ve haritalama işlemlerinin tamamlanması ile birlikte beyin modellemesi ortaya çıkar [23, 24]. Program çalışması esnasında, her bir alt adımın çıktısı başka bir adımın girdisi olacak şekilde veri aktarılır ve kademeli olarak süreç devam eder.

Çalışmalarda kullanılan bilgisayar donanımına bağlı olarak, tüm analiz ve modelleme işlemleri bir hayli zaman alabilmektedir. 2011 yılında yapılan bir çalışmada, tek bir Freesurfer beyin modellemesinin kabaca bir günü bulabileceği raporlanmıştır [15]. Benzer şekilde 2012'de yapılan bir çalışmada tüm sürecin 30 saati bulan sürelerde tamamlanabileceği; ayrıca, analiz sürecinin program sürümündeki değişikliklerden ya da işletim sistemi farklılıklarından etkilenebileceği dile getirmiştir [25].

Yukarıda anlatılan Freesurfer işlemleri bu çalışmada, Ubuntu 14.04 x64 işletim sistemine sahip Intel® Core™ i7-4700 2.40 GHz CPU'lu ve 1600 MHz 16 GB RAMli bir bilgisayarda gerçekleştirilmiştir. Numuneler çoklu ele alınarak gruplandırıldıktan sonra paralel şekilde Freesurfer v5.3.0 programında işleme alınmıştır. Numune başına yaklaşık 10-15 saatlik süreler sonrasında bölütleme ve üç boyutlu beyin modelleme işlemlerinin tamamlandığı görülmüş ve program çıktıları erişilebilir hale gelmiştir. Görsel 3.3'te bir numunenin beyni içerisinde görülmek istenilen bir noktanın farklı eksenler üzerinden bölütlenmiş beyin kesitleri ve Görsel 3.4'te ise bir numuneye ait Freesurfer analizi sonrasında oluşan üç boyutlu beyin modeli görülmektedir.



Görsel 3.3. Bir Numunenin Koronal-Aksiyel-Sagittal Kesitleri



Görsel 3.4. Bir Numunenin Freesurfer ile Analizi Sonrası Üç Boyutlu Modeli

Bu çalışmada, numuneler için analiz sonrası elde edilen sanal beyin modellemelerine el ile müdahale edilmemiştir.

Freesurfer yazılım aracı üzerinden beyin modelleme sürecinin başarı ile tamamlanması sonucunda, çıkarılan modellemeye ait iki ve üç boyutlu yapısal görsel dosyalar olduğu önceki kesimlerde belirtilmiştir. Ek olarak, istenilen beyin bölümlerinin ölçülerini ifade eden, hacim, kalınlık, yüzey alanı gibi bilgilerin yanı sıra bu ölçülerin hesaplamalarında kullanılan standart sapma, ortalama, minimum, maksimum bulgular gibi sayısal değerleri barındıran istatistiksel dosyalar da analiz çıktısı olarak elde edilebilmektedir. Analizi başarı ile sonuçlanan her bir numune için 18 adet istatistik dosyası oluşur. Farklı istatistik dosyalarında ortak veya değişik beyin bölümlerinin çeşitli kriterlerdeki ölçülendirmeleri program tarafından *stats* uzantılı istatistik dosyalarına metin tabanlı kaydedilmektedir. Örnek olarak Görsel 3.5'te bir numuneye ait *aseg.stats* ve Görsel 3.6'da bir numuneye ait *lh.aparc.stats* istatistik dosyalarındaki beyin bölümlerine ve sayısal değerlerine ait bilgilerin nasıl yer aldığı sunulmaktadır.

#	ColHeaders	Index	SegId	NVoxels	Volume_mm3	StructName	normMean	normStdDev	normMin	normMax	normRange
79	1	4	12279	12278.6	Left-Lateral-Ventricle	27.2754	14.0753	0.0000	101.0000	101.0000	
80	2	5	350	350.3	Left-Inf-Lat-Vent	67.2857	9.8261	41.0000	85.0000	44.0000	
81	3	7	23798	23798.4	Left-Cerebellum-White-Matter	95.4906	5.7365	43.0000	116.0000	73.0000	
82	4	8	39909	39908.7	Left-Cerebellum-Cortex	74.2764	10.9956	31.0000	103.0000	72.0000	
83	5	10	6829	6828.6	Left-Thalamus-Proper	99.7465	8.5935	54.0000	117.0000	63.0000	
84	6	11	3147	3147.1	Left-Caudate	93.6117	9.7077	54.0000	108.0000	54.0000	
85	7	12	4462	4462.2	Left-Putamen	98.1038	5.0399	66.0000	113.0000	47.0000	
86	8	13	1428	1427.7	Left-Pallidum	98.5721	3.5765	75.0000	109.0000	34.0000	
87	9	14	2132	2131.5	3rd-Ventricle	37.3733	16.2586	4.0000	83.0000	79.0000	
88	10	15	1302	1301.8	4th-Ventricle	35.2919	12.2258	8.0000	83.0000	75.0000	
89	11	16	21692	21692.4	Brain-Stem	90.5834	10.8066	15.0000	117.0000	102.0000	
90	12	17	3883	3883.1	Left-Hippocampus	90.1274	7.3114	57.0000	106.0000	49.0000	
91	13	18	880	880.3	Left-Amygdala	88.5854	6.6967	66.0000	102.0000	36.0000	
92	14	24	1792	1791.9	CSF	39.2128	16.2354	6.0000	88.0000	82.0000	
93	15	26	303	303.1	Left-Accumbens-area	87.2384	9.0058	60.0000	104.0000	44.0000	

Görsel 3.5. Bir Numunenin *aseg.stats* Dosyasına Ait Kısmi Ölçülendirme Örneği

#	ColHeaders	StructName	NumVert	SurfArea	GrayVol	ThickAvg	ThickStd	MeanCurv	GausCurv	FoldInd	CurvInd
53	banksts		1085	768	1827	2.336	0.852	0.178	0.112	30	4.5
54	caudalanteriorcingulate		867	534	1245	2.239	0.761	0.140	0.108	32	3.4
55	caudalmiddlefrontal		2708	1768	3976	2.063	0.654	0.132	0.051	32	6.1
56	cuneus		1629	1123	2375	1.839	0.770	0.213	0.516	381	11.4
57	entorhinal		750	434	1995	3.059	0.866	0.137	0.088	12	3.3
58	fusiform		3390	2147	5103	2.194	0.870	0.182	0.242	577	50.3
59	inferiorparietal		5056	3228	7315	2.176	0.805	0.139	0.060	70	13.0
60	inferiortemporal		4152	2740	9450	2.882	1.002	0.159	0.107	127	17.4
61	isthmuscingulate		1050	666	1350	2.078	0.846	0.151	0.134	77	7.3
62	lateraloccipital		6397	4231	10009	2.255	0.824	0.170	0.114	218	27.1
63	lateralorbitofrontal		2663	1813	5396	2.925	0.951	0.136	0.059	34	6.4
64	lingual		2638	1977	3381	1.834	0.682	0.246	0.186	172	20.1
65	medialorbitofrontal		2350	1660	5349	2.904	0.809	0.167	0.096	82	9.5
66	middletemporal		3780	2529	8354	2.754	0.889	0.167	0.100	118	16.7
67	parahippocampal		344	237	573	2.177	0.640	0.109	0.030	2	0.4

Görsel 3.6. Bir Numunenin *lh.aparc.stats* Dosyasına Ait Kısmi Ölçülendirme Örneği

Bahsedilen çıktı dosyaları her bir numunenin beyin analizi sonrasında ayrı ayrı oluşmaktadır. Daha da detaylı olarak, önceki bölümlerde belirtildiği gibi 63 adet örneğin her biri için 18 tane istatistik dosyası oluşmuştur.

Her bir istatistik dosyası kendi içerisinde beyin bölümlerine yönelik hacim, yüzey alanı, gri hacim, kalınlık ortalaması, kalınlık standart sapması, ortalama eğrilik veya Gauss eğriliği, indekisleme ve/veya analiz işlemlerinde iterasyonlar esnasında kullanılacak *normMean*, *normStdDev*, *normMin*, *normMax* gibi farklı kıstaslarda birçok sayısal değerin değişik kombinasyonlarını barındırmaktadır (Görsel 3.5 ve Görsel 3.6). Tüm numunelere ait istatistiksel bilgiler içerisinden segmentasyon ve parselasyon anatomisine ait genel kortikal hacim değerleri *aseg.stats* dosyalarından; sağ ve sol beyin gri madde hacmi, kalınlık, yüzey alanı ile ilgili değerler ise *lh.aparc.stats* ve *rh.aparc.stats* dosyalarından Linux *bash scriptleri*² yardımı ile çekilmiştir. Bahsi geçen bilgiler dosyalarda belirli bir düzende yer aldığından *scriptler* ilgili düzeni çözümleyici şekilde yazılmıştır. Dosyalardan istenilen sayısal verinin çekilmesi ile farklı öznitelik matrisleri yaratılmıştır.

Tüm öznitelik matrisleri yaratılırken, her matrisin sütunları ayrıca incelenmiştir. Başka bir ifade ile incelenen öznitelik, tüm numuneler için aynı değerleri içeriyor ise sınıflandırma aşamasında herhangi bir etkisinin olmayacağından elimine edilmiştir. Bu doğrultuda, *aseg.stats* dosyalarından çekilen bilgiler kapsamında *sol* ve *sağ beyaz madde hypointensities*, *sol* ve *sağ beyaz madde olmayan hypointensities* öznitelikleri tüm numuneler için 0 değerini içerdiğinden oluşturulacak öznitelik matrisine katılmamıştır. Tablo 3.2’de, oluşturulan öznitelik grupları, öznitelik ölçü birimi, hangi istatistik

²Kabuk programlama olarak da bilinen yapı ile komut satır ekranında yazılan farklı komutların bir dosya üzerinden çoklu şekilde çalıştırılmasına olanak sağlanır.

dosyasından elde edildiği ve beynin farklı bölümlerinden toplamda içerdiği öznitelik sayısı ile birlikte sunulmaktadır.

Tablo 3.2. Öznitelik Grupları

Öznitelik Grubu	Ölçü Birimi	İstatistik Dosyası	Toplam Öznitelik Sayısı
Kortikal Hacimler	mm ³	<i>aseg.stats</i>	41
Sol Gri Hacimler	mm ³	<i>lh.aparc.stats</i>	34
Sol Yüzey Alanları	mm ²	<i>lh.aparc.stats</i>	34
Sol Kalınlık Ortalamaları	mm	<i>lh.aparc.stats</i>	34
Sağ Gri Hacimler	mm ³	<i>rh.aparc.stats</i>	34
Sağ Yüzey Alanları	mm ²	<i>rh.aparc.stats</i>	34
Sağ Kalınlık Ortalamaları	mm	<i>rh.aparc.stats</i>	34

MATLAB platformunda işleme alabilmek adına, oluşturulan her bir öznitelik gurubu *mat* uzantılı öznitelik matrisleri şeklinde ifade edilmiştir. Örnek olarak Görsel 3.7’de, tüm numunelerden alınan veri ile oluşturulmuş “Sağ Yüzey Alanları” öznitelik matrisine ait değerlerin bir kısmı MATLAB platformu üzerinden görüntülenmektedir. Veri setinde 63 örnek olduğundan; ayrıca, ilgili grupta 34 öznitelik bulunduğundan, örnek gösterilen matrisin boyutu 63×34 şeklinde gözlemlenebilmektedir.

	1	2	3	4	5	6	7	8	9	10
1	859	654	1644	1510	408	2024	4087	2416	786	4164
2	365	454	1340	855	284	2126	3402	1608	625	2928
3	356	553	1623	688	287	2365	3519	2268	461	2964
4	521	490	1573	1139	298	1464	3591	1686	813	3658
5	345	540	1606	886	242	1687	3307	1608	769	3799
6	879	617	1579	1259	392	2652	5093	2273	949	4589
7	875	631	1587	1321	360	2281	5149	2778	474	5321
8	628	340	1388	996	331	2086	4044	2458	366	3504
9	300	424	1918	1052	521	2563	4245	2207	587	2926
10	831	492	1687	1326	391	2531	3619	2108	715	4145
11	706	368	1499	912	497	1908	4351	2548	448	3244
12	517	384	1146	957	435	1983	3729	1767	426	3007
13	551	407	1624	814	253	2437	3241	2148	673	3084
14	431	131	1360	1331	362	2280	3969	1539	672	3049
15	892	534	1434	1209	401	2184	4522	1937	516	3756
16	538	474	1615	925	357	2277	3748	2192	797	2926
17	505	508	1376	987	343	2342	3507	1563	416	2756
18	678	677	1411	1105	289	1918	4238	2231	884	3968
19	460	486	1151	850	175	1383	3563	1157	653	3292
20	444	392	1659	704	430	2440	3760	2253	457	3747

Görsel 3.7. “Sağ Yüzey Alanları” Öznitelik Grubuna Ait Kısmi Değerler

3.3. Normalizasyon

Normalizasyon, girdi olarak alınan bir matrisin satır veya sütun vektörleri üzerinden değerlerinin güncellenmesidir. Her bir satırdaki veya sütundaki değerlerin kareleri toplamı 1 olacak şekilde ilgili hücreler 0-1 aralığında güncellenir.

Bir vektördeki ($X = [x_1, x_2, \dots, x_{(n-1)}, x_n]$) herhangi bir değer (x_i) normalizasyon denklemi, tüm vektör değerleri üzerinden Denklem (3.1)'de sunulmaktadır.

$$x'_i = \frac{x_i}{\sqrt{\sum_{j=1}^n x_j^2}} \quad (3.1)$$

Değerleri rastgele oluşturulmuş bir A matrisini sütun temelli normalize edilmek üzere örnek alalım. Tablo 3.3'te örnek A matrisinin normalizasyon işlemi uygulanmadan önceki ham değerleri görülmektedir.

Tablo 3.3. 4×4 Boyutlarında Örnek Bir Matrisin Ham Değerleri

	$A(:,1)$	$A(:,2)$	$A(:,3)$	$A(:,4)$
$A(1,:)$	1	2	3	4
$A(2,:)$	1	4	3	2
$A(3,:)$	1	2	9	0
$A(4,:)$	1	4	6	3

Bir A' matrisini, A matrisinin Denklem (3.1) ile sütun temelli normalize edilmiş hali olmak üzere örnek alalım. A matrisinin her bir sütunundaki değerlerinin kareleri toplamı 1 olacak şekilde normalize edildikten sonraki değerler ile A' matrisi oluşmaktadır. Tablo 3.4'te örnek gösterilen A' matrisi sunulmaktadır.

Tablo 3.4. 4×4 Boyutlarında Örnek Normalizasyon

	$A'(:,1)$	$A'(:,2)$	$A'(:,3)$	$A'(:,4)$
$A'(1,:)$	0.5	0.3162	0.2582	0.7428
$A'(2,:)$	0.5	0.6325	0.2582	0.3714
$A'(3,:)$	0.5	0.3162	0.7746	0
$A'(4,:)$	0.5	0.6325	0.5164	0.5571

Bu çalışmada ise, veri setinden çıkarılan öznitelik grupları sınıflandırıcılar ile işleme alınmadan önce, her bir özneliğin kendi içerisindeki etkisi optimize edilmek amacıyla matris sütun temelli olarak normalize edilmektedir. Denklem (3.1)'de ifade

edildiği gibi, öznitelik matrisindeki her bir hücre bulunduğu öznitelik vektörünün tüm numune değerleri üzerinden normalizasyona uğramaktadır.

3.4. Sınıflandırıcılar

Bu çalışmada, makine öğrenmesi araştırmalarında sıkça rastlanılan NB ve SVM sınıflandırma metotları GA ile birlikte kullanılarak sarmalama öznitelik seçim yöntemini oluşturmaktadır. Sarmalama öznitelik seçiminde, uygulanan sınıflandırıcı algoritmasının sonucuna bağlı olarak öznitelikler alt sete eklenir veya alt setten çıkarılır. Değerli öznitelikler ise en yüksek doğruluk sonucunu veren alt setin içeriğine göre isimlendirilir. Deneysel çalışmalar için oluşturulmuş farklı öznitelik gruplarına ait kıymetli öznitelikler, hata matrisleri ve doğruluk sonuçları belirtilen sınıflandırma algoritmaları üzerinden kıyaslanmak üzere elde edilmektedir.

3.4.1. Naïve Bayes sınıflandırıcı

Bayes kuralı dikkate alınarak naïve bağımsızlık varsayımı ile özniteliklerin birbirinden ilintisiz olduğu varsayılır. Sınıflandırma kısmında test örneğinin dâhil olabileceği olası tüm sınıflara yönelik koşullu olasılıklar hesaplanır ve maksimum olasılığın hesaplandığı sınıf, örneğin dâhil edileceği sınıfı belirler. Tamamen olasılık üzerine kurulu bir algoritmadır. Tüm bu özelliklerinden dolayı koşullu sınıflandırıcı olarak da bilinir. Ayrıca NB'nin, metin sınıflandırma, tıbbi teşhis koyma, sistem performans yönetimi gibi birçok uygulamada etkili olduğu daha önce yapılan çalışmalarda vurgulanmıştır [26]. $P(A)$ ve $P(B)$ bağımsız olayların olasılıkları, $P(A/B)$ ise B doğru iken A 'nın olma olasılığı şeklinde yorumlandığında Bayes teoremi matematiksel olarak Denklem (3.2) üzerinden sunulmaktadır.

$$P(A|B) = \frac{P(B|A) P(A)}{P(B)} \quad (3.2)$$

Örnek olarak n öznitelikten oluşan bir X vektörü ($X = [x_1, x_2, \dots, x_{(n-1)}, x_n]$) ve dâhil olabileceği 3 sınıf (C_1, C_2, C_3) olsun. X vektörünün isimlendirilebileceği sınıflara yönelik koşullu olasılık denklemleri $P_1(C_1 / X)$, $P_2(C_2 / X)$ ve $P_3(C_3 / X)$ olarak ifade edilir.

Denklem (3.3) üzerinden belirtilen koşullu olasılıklar, olası her bir sınıf için tek tek hesaplanır.

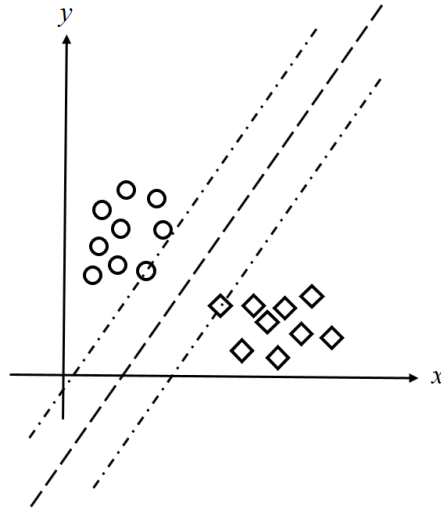
$$P_k(C_k, X) = \prod_{i=1}^n p(x_i | C_k) \quad k = 1, 2, \dots \quad (3.3)$$

Dâhil olunabilecek tüm sınıflara dair olasılıklar tek tek hesaplandıktan sonra, tüm olasılık değerleri içerisinde en yüksek ihtimal değerine (P_k) sahip sınıf, X vektörünün sınıfını belirtir.

Bu çalışmada ise NB ile GA üzerinden öznitelik seçimi yapılırken, ele alınan test numunesi olarak bir hastanın öznitelik vektörü F 'nin 3-Sınıflı deneysel çalışmalarda $P_1(AH | F)$, $P_2(FTD | F)$ ve $P_3(VaD | F)$ koşullu olasılıkları incelenmiştir. 2-Sınıflı deneysel çalışmalarda ise tüm sınıflar üzerinden $CLS = \{AH, FTD, VaD\}$ ve $C \in CLS$ iken $P_1(C | F)$ ve $P_2((CLS \setminus C) | F)$ şeklinde bir sınıfa dâhil olma veya dâhil olmama olasılıkları gözetilmiştir.

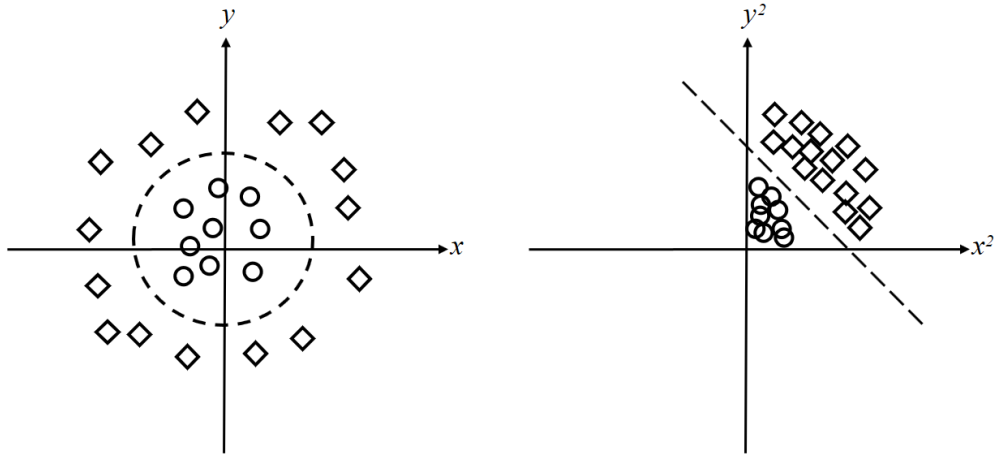
3.4.2. Destek vektör makineleri

SVM, biyomedikal görüntü sınıflandırma çalışmaları arasında literatürdeki en popüler makine öğrenmesi algoritmalarının başında gelmektedir. Temel olarak, düzlemdeki iki sınıfı birbirinden ayırmak için hiperdüzlem bulmayı hedefler. Düzlemdeki gruplanmış noktalar arasında sınır çizgisi çekilirken, sınır çizgisinin her iki grubun noktalarına ve destek vektörlerine en uzak mesafede olması ilkesini temel alır [27]. Bu doğrultuda, maksimum mesafe sınıflandırıcısı olarak da anılabilir ve algoritmanın temel çalışma prensibi *perceptron* algoritmasının metodolojisine dayanır [28]. Şekil 3.1'de lineer şekilde ayrılabilir biçimde dağılmış noktalar kümesi üzerinden SVM ile nasıl sınır çekilebildiği örneği görülmektedir.



Şekil 3.1. *Lineer SVM Modeli Örneği*

Algoritma ne kadar lineer olarak tanımlansa da bazı durumlarda hesaplama yapılamayacak karmaşık şartlar oluşur. *Kernel* fonksiyonları kullanılarak girdi noktaları oluşturulan yeni hiperdüzleme farklı değerler ile ifade edilecek şekilde yerleştirilir. En yakın destek vektörler arasındaki mesafeyi maksimize etme çalışması *kernel* fonksiyonları ile oluşturulmuş yeni lineer düzlem üzerinde gerçekleşir [27, 29]. Şekil 3.2’de lineer olarak ayrıştırılamayan noktalar örneği için, denklemin her iki tarafının karesi alınarak noktalar yeni bir düzlem üzerinde ifade edilmektedir. Sonrasında ise SVM normal işleyişini sürdürüp kümelenmiş noktaları yeni düzlem üzerinden lineer olarak gruplara ayırabilmektedir.



Şekil 3.2. *Kernel Fonksiyonlar ile Lineerleştirilmiş SVM Modeli Örneği*

Örnek olarak n öznitelikten oluşan bir X vektörü ($X = [x_1, x_2, \dots, x_{(n-1)}, x_n]$) ve dâhil olabileceği 2 sınıf olsun. $C_i \in \{-1, 1\}$ iken $S = \{(x_1, C_1), (x_2, C_2), \dots, (x_{(n-1)}, C_{(n-1)}), (x_n, C_n)\}$ ile belirtilen eğitim seti üzerinden test verisinin sınıfı Denklem (3.4)'teki maliyet fonksiyonu aracılığıyla hesaplanır [28].

$$f(x_{test}) = \sum_{i=1}^n \{\alpha_i C_i (x_i^T x_{test}) + b\} \quad (3.4)$$

Destek vektörleri tanımlayan, değeri sıfır olmayan α_i ($i = 1, 2, \dots$) katsayısı ikinci dereceden programlama problemi üzerinden elde edilebilir. Bu bilgiler ışığında hiperdüzelem ve orijin ile bağlantılı dikey ve normal vektörleri de hesap edilebilir. ($|b|/||w||$), orijin ve optimum hiperdüzelem arasındaki dikeydir, w da hiperdüzelemin normal vektörüdür [28].

SVM, iki sınıfı birbirinden ayıran bir algoritma olarak tanındığı gibi, “birine karşı diğerleri” stratejisinin de uygulanması ile çoklu sınıf problemlerinde de kullanılabilir [28].

Bu çalışmada SVM, veri setindeki üç sınıfın deneysel çalışmalar esnasında *AH v diğerleri*, *FTD v diğerleri*, *VaD v diğerleri* şeklinde “birine karşı diğerleri” stratejisi ile gruplandırılması sonrasındaki 2-Sınıflı testlerde GA ile öznitelik seçimi ile birlikte kullanılmıştır.

3.5. Öznitelik Seçimi

Problemin tanımına göre öznitelik sayısının yeteri kadar fazla olmadığı durumlarda tüm ihtimaller, başka bir ifade ile olası bütün alt öznitelik grupları denenerek en iyi çözümün bulunması en iyi senaryodur. Ancak, girdi parametresi olarak çok sayıda öznitelik bulunduğunda, literatürde *brute-force* veya *exhausted* arama olarak adlandırılan değerli özniteliklerin seçiminin, özniteliklerin olası tüm ihtimallerinin denenerek kayıpsız yapılması, çalışma zamanı açısından uzun süreleri gerektirebilmektedir; bu bağlamda bu süreç bazen neredeyse imkansız olarak tanımlanabilmektedir. Bu doğrultuda, tüm ihtimalleri denemeyen, fakat en uygun değerli öznitelik setinin bulunması için sezgisel öznitelik seçim metotları geliştirilmiştir. Denetimli öznitelik seçim metotları kendi içerisinde herhangi bir sınıflandırıcıya bağlı olmayan filtre metotları ve sınıflandırıcı sonucuna bağlı olan sarmalama metotları şeklinde ayrılmaktadır. Filtreleme metotlarında

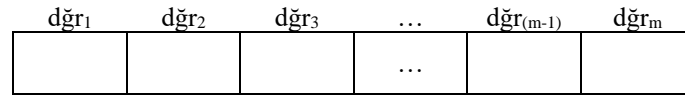
basitçe özniteliklerin kendi içerisinde istatistiksel analizi yapılarak değişimleri incelenir. Bu sayede asıl işlemlere geçmeden önce alakasız öznitelikler filtrelenmiş olur. Sarmalama metotlarında ise seçilen özniteliklere ait sınıflandırıcı sonuçları üzerinden kıyaslamalar yapılarak en iyi öznitelik setine ulaşmak hedeflenir [30].

Bu çalışmada GA tabanlı sarmalama öznitelik seçim metodu kullanılmıştır. Algoritma ve parametreleri izleyen kesimlerde detaylandırılmaktadır.

3.5.1. Genetik algoritma

GA'nın temeli, biyolojik adaptasyon süreci sırasında bir nesilden diğerine evrimleşirken zayıf genlerin doğal seleksiyon ile ortadan yok olması ve iyi genlerin kurtulması üzerine kuruludur [31, 32]. Matematiksel yaklaşım olarak GA ise, tanımlanmış problemler için bir parametre setinden diğerine dolaşarak optimum çözümü arar. Bir *fitness* veya maliyet fonksiyonu üzerinden girdi olarak aldığı parametrelere yönelik bir maliyet çıkarır. Birçok girdi kombinasyonu için farklı farklı maliyetler çıkarıldıktan sonra, problemin tanımına göre en uygun maliyeti sağlayan girdi değerleri kıymetli girdiler olarak nitelendirilir [31, 33].

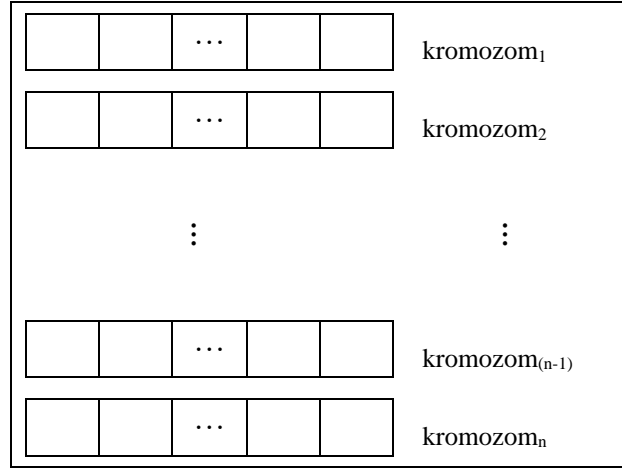
GA metodolojisinde en iyi çözüm, kromozom adı verilen sayısal değerlerin bulunduğu vektörlerden gelir. Kromozomun belirttiği her bir özellik vektör içerisindeki sayısal değerler ile ifade edilir. Genler ise kromozom üzerinde belirli bölümlerde yer alır ve sayısal değerlerin birleşiminden oluşur. Şekil 3.3'te m adet değere sahip bir kromozom modeli sunulmaktadır.



Şekil 3.3. Kromozom Modeli

Her bir kromozomun değerinin sayısal olarak belirlenebilmesi için bir maliyet fonksiyonu kullanılır. Genlere ait değerler, fonksiyonun içerisinde işlemlere dâhil edilerek bir maliyet analizi yapılır. Bir kromozom seti içerisindeki her bir kromozom için aynı prosedür gerçekleştirildiğinde, elde tüm setteki kromozomlara ait maliyet değerleri bulunur. En uygun maliyete sahip kromozomun diğerlerinden daha ön plana çıkması ile problemin çözümü bulunmaya çalışılır.

Probleme en uygun çözüm bulma süreci, nesil adı verilen yinelemeli yapı ile sürdürülür. Her bir nesil, o ana kadarki en iyi maliyet sonucunu bulmak üzere yönelim gösterir. Tüm nesiller kromozomlardan oluşan popülasyonlar üzerine kuruludur. Farklı nesillerdeki her bir popülasyon, aynı sayıda kromozom içerir. Popülasyon içerisindeki kromozomların genleri değişiklik gösterebilir. Şekil 3.4'te n kromozoma sahip bir popülasyon modeli sunulmaktadır.



Şekil 3.4. Popülasyon Modeli

Popülasyon içerisindeki her bir kromozomun maliyet fonksiyonu aracılığı ile değeri hesaplanır. En uygun maliyet fonksiyonunu sağlayan kıymetli kromozomların algoritma tarafından sonraki nesillere genlerini aktarması sağlanır. Aktarılacak genler için kromozom seçim fonksiyonlarının kullanılması ile yeni nesiller inşa edilir. Yeni nesillere ait bireyler oluşturulurken iki adet kromozomu (aile kromozomları) girdi olarak alıp yeni kromozomlar (çocuk kromozomlar) oluşturan fonksiyonlar kullanılır. Neslindeki en uygun maliyete sahip, elit adı verilen kromozomlar diğer nesle direkt olarak geçirilebilir. Bazen ise, gelişme süreci yavaş olduğunda genlerde gelişigüzel ufak değişiklikler olur; eğer bu değişiklikler çözüme olumlu anlamda etki ediyorsa, bu değişiklik ile türeyen yeni bireyler ile eski nesillerden gelen genler yavaş yavaş yok edilmeye başlar [32].

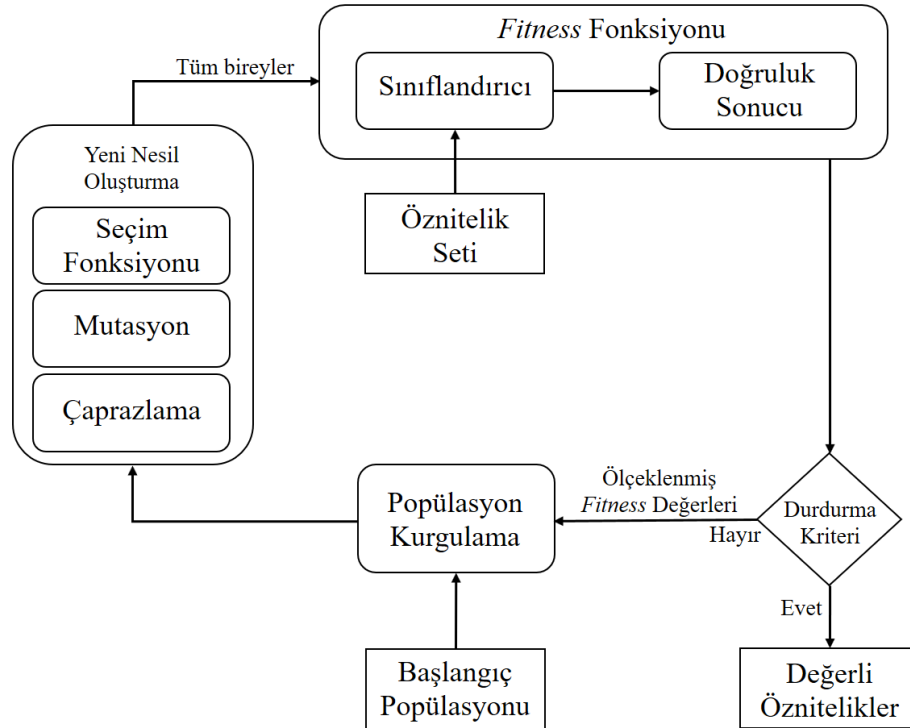
Nesilden nesle en uygun genler aktarılırken çözüm odaklı en iyi popülasyon kurulmaya çalışılır. İstenilen maliyete ulaşılması veya belirtilen maksimum jenerasyon sayısının aşılması gibi durdurma ölçütleri sağlandığında GA süreci sonlandırılır. Son popülasyon içerisindeki en iyi maliyete sahip kromozom, en değerli kromozom olarak belirtilir.

Bu çalışmada GA'nın kullanılmasındaki asıl amaç, sarmalama öznelik seçimi yaklaşımının uygulanması ile kıymetli öznelik setinin bulunması arayışıdır; başka bir ifade ile GA, sınıflandırmalar sırasında etkili bulunmayan özneliklerin tüm setten elenmesi amacı ile kullanılmıştır.

3.5.2. Genetik algoritma parametreleri

Problemin tanımına göre GA ile çözüm bulma süreci, algoritma içerisindeki birçok parametrenin değerine bağlıdır.

Bu çalışmada, girdi olarak alınan öznelik seti, GA süreci içerisinde belirtilen parametrelere bağlı olarak değerli öznelik setini çıktı olarak vermektedir. Çalışmada izlenen GA yöntem adımlarının ve maliyet fonksiyonu kullanımının görselleştirildiği genetik algoritma tabanlı sarmalama öznelik seçim algoritması Şekil 3.5'te sunulmaktadır.



Şekil 3.5. Genetik Algoritma Tabanlı Sarmalama Öznelik Seçimi

İzleyen kesimde, algoritma içerisinde lineer kısıtlamalar olmayacak şekilde parametrelerin prensipleri, GA içerisinde neye hizmet ettiği ve bu çalışma kapsamında izlenen yaklaşımlar detaylandırılmaktadır.

3.5.2.1. Kromozom tipi

GA popülasyonunu oluşturan kromozomlar, özel olarak belirtilmenin haricinde iki tip ile nitelendirilebilir. *Double vector* tipi ile genlere ait değerler rasyonel sayılar ile ifade edilmektedir. *Bit string* tipi ile herhangi bir genin değeri ikili sayı sisteminden oluşmaktadır; başka bir söylem ile bir hücredeki değer ya 0 ya da 1'dir [34]. GA'nın ortaya çıktığı ilk yıllarda kromozomlar sadece *bit string* tipi ile işleme alınmakta iken; sonralarda, problemi ifade etmek için bu kromozom tipinin yeterli gelmediği ve yeni tiplerin ortaya çıktığı belirtilmektedir [32, 35]. Şekil 3.6'da *double vector*, Şekil 3.7'de ise *bit string* kromozom tipi için değerleri rastgele atanmış birer kromozom örneği görülmektedir.

dğr ₁	dğr ₂	dğr ₃	...	dğr _(m-1)	dğr _m
0.0756	0.0768	-6.0356	...	4.8974	-8.8809

Şekil 3.6. 'Double Vector' Tipinde Kromozom Örneği

bit ₁	bit ₂	bit ₃	...	bit _(m-1)	bit _m
1	1	0	...	0	1

Şekil 3.7. 'Bit String' Tipinde Kromozom Örneği

Bu çalışmada, hangi özneliklerin değerli olduğunu bulmak amacıyla kromozomlar *bit string* tipi ile ifade edilmektedir. Kromozomlar, öznelik matrisine bir seçim maskesi olacak şekilde nitelendirilmektedir. Ele alınan kromozomdaki değerleri 1 olan bitler, dolayısı ile bitlerin ifade ettiği öznelikler, muhtemel sonuç hesabında alt öznelik setinde vardır şeklinde yorumlanmaktadır. Sonrasında ise maliyet fonksiyonu, içerisindeki sınıflandırıcıda tüm öznelik seti üzerinden 1 ile ifade edilen kısmi öznelikler ile eğitilecek ve test edilecek şekilde çağırılmıştır.

Şekil 3.8'de GSS algoritmasına girdi olarak verilen örnek bir öznelik matrisinin GA süreci esnasında karşılaşılan bir kromozom üzerinden sınıflandırıcı esaslı maliyet fonksiyonuna nasıl girdi olarak alınacağı sunulmaktadır. Değerleri rastgele atanmış örnek

öznitelik seti 4 numune ve 5 öznitelikten oluşmaktadır. Genler kullanılacak öznitelikleri belirttiğinden maliyeti hesaplanacak kromozomun boyu da öznitelik sayısı ile aynı olmaktadır. Bit değeri 1 olan 3 hücre (bit₁, bit₂ ve bit₄) bulunmaktadır. Kromozomun belirtilen bitlerinin maskelediği, öznitelik matrisindeki vurgulanan 3 sütun (öz₁, öz₂ ve öz₄) ilgili sınıflandırıcı ile doğruluk sonucu hesaplanmak üzere maliyet fonksiyonunda girdi seti olarak kullanılmaktadır.

öz ₁	öz ₂	öz ₃	öz ₄	öz ₅
0.5	0.316	0.258	0.742	0.316
0.5	0.632	0.258	0.371	0.632
0.5	0.316	0.774	0	0.5
0.5	0.632	0.516	0.557	0.5

a. Öznitelik Seti

bit ₁	bit ₂	bit ₃	bit ₄	bit ₅
1	1	0	1	0

b. Kromozom

öz ₁	öz ₂	öz ₄
0.5	0.316	0.742
0.5	0.632	0.371
0.5	0.316	0
0.5	0.632	0.557

c. Kromozomların maskelediği set

Şekil 3.8. Kromozom Aracılığı ile Sınıflandırıcıya Gönderilen Alt Öznitelik Seti Örneği

3.5.2.2. Fitness belirleme ve ölçeklendirme

Popülasyon içerisindeki farklı genlerden oluşan kromozomların birbirine göre kıyaslamasının yapılması, aralarından değerlilerin veya değersizlerin tespit edilmesi asıl hedeftir. Bu doğrultuda ele alınan herhangi bir kromozomun değerinin ölçülebilmesi için *fitness* veya maliyet fonksiyonları adı altında değerlendirici fonksiyonlar kullanılır. Maliyet fonksiyonları girdi parametresi olarak aldığı kromozom vektörü içerisindeki sayısal değerleri işlemde geçirdikten sonra tek bir sayısal çıktı üretir. Çıktı değeri kromozomun kıymetini ifade eder; başka bir deyiş ile maliyetini belirtir.

Bu çalışmada, sarmalama öznitelik seçim yaklaşımı uygulanarak alt öznitelik grupları elde edilmek istendiğinden, maliyet fonksiyonu 1 ve 0 değerlerinden oluşan genleri girdi olarak kullanılmaktadır. Seçim maskeleyesi yaklaşımı ile 1 değerine sahip bitler, matristeki ilgili öznitelikliğin alt öznitelik setinde olduğunu ifade etmektedir. Şekil 3.5'te de sunulduğu gibi maliyet değeri olarak, seçilen özniteliklerin sınıflandırılması sonrasında elde edilen ve hata matrisi üzerinden hesaplanan doğruluk sonucu değeri kullanılmaktadır. *Fitness* değerinin maksimize edilmesi asıl hedeftir; başka bir ifade ile doğruluk sonucu yükseldikçe kromozom değerlenir.

Deneysel çalışmalar esnasında farklı öznitelik matrisleri üzerinden işlemler yapılacağından kromozom uzunluğu herhangi bir değer ile kısıtlanmamıştır. İlgili matris için öznitelik sayısı kromozomların uzunluğunu da belirleyen parametre olmuştur.

Algoritma adımları sırasında tüm bireylerin *fitness* değeri hesaplanmak üzere maliyet analizi yapılır. Bu adımdan sonra ise maliyet fonksiyonu üzerinden dönen ham değerler belirli bir aralıkta ölçeklendirilir. El ile ölçeklendirme tercih edilmiyorsa yaygın olarak kullanılan birkaç ölçeklendirme çeşidi vardır. *Rank* ölçeklendirmesi ile ham değerler sıralandıktan sonra bulunduğu sıranın karekökü ile ters orantılı olarak yeni değerine kavuşur. Bu yöntem ile ham skorların açık ara dağılımı engellenmiş olur. *Proportional* ölçeklendirmesi ile ham skorların oranına göre yeni değerler oluşturulur. *Top* ölçeklendirmesi ile istenilen ilk yüzdelerdeki maliyetler eşit olarak ele alınır, kalan maliyetler ise sıfırlanır. *Shift linear* ölçeklendirmesi ile kullanıcıdan alınacak ekstra bir parametrenin ham skorlar ile çarpılması sonrası yeni değerler elde edilir [34].

Bu çalışmada ise deneysel çalışmaların birinci aşamasında ölçeklendirme çeşidi olarak *rank* ölçeklendirmesi uygulanmıştır. O anki nesilde en iyi maliyetlere sahip olan ve maliyet sıralamasında yukarılarda olan kromozomların yeni nesiller üzerindeki etkisinin daha fazla olması amaçlanmıştır. İkinci aşamada ise *proportional* yaklaşımı kullanılarak tamamen maliyet fonksiyonlarından dönen ham skorlara odaklanılması amaçlanmıştır.

3.5.2.3. Popülasyon

GA sürecinin ilk adımı, başlangıç popülasyonunun belirlenmesi olarak ifade edilebilir. Bir popülasyonun en önemli özellikleri ise, bireylerin değerlerini belirten kromozom tipi ve nesilleri oluşturan büyüklüğüdür.

3.5.2.3.1. Başlangıç popülasyonu

GA sürecinin başlaması için bir popülasyona ihtiyaç vardır. Başlangıç popülasyonuna ait muhtemel çözümü içeren gen değerlerinin istenilen değerler ile oluşturulabileceği gibi, rastgele atanması da mümkündür [36]. Bu yöntem, tüm kromozom tipleri için geçerlidir.

Bu çalışmada, deneysel çalışmaların birinci aşamasında başlangıç popülasyonuna ait kromozomların gen değerleri rastgele oluşturulmuştur. İkinci aşamada ise birinci aşamada elde edilen değerli özniteliklerin birleştirilmesinden sonra kıymetli verinin etkisinin rastgele atama ile kaybedilmemesi amacı ile başlangıç popülasyonu el ile oluşturulmuştur. Kromozomlar üzerinde birinci aşamadan gelen kıymetli alt öznitelik setleri olabildiğince bir bütün halinde tutulacak şekilde ikinci aşamanın başlangıç popülasyonu doğrudan kodlanmıştır. Her bir kromozom için rastlantısal olarak birkaç bitinin değiştirilmesine olanak sağlanmıştır.

3.5.2.3.2. Popülasyon kurgulama

Yeni nesilleri oluşturacak bireylerin nasıl kurgulanacağına dair farklı yaklaşımlar mevcuttur. Elit bireylerin de katıldığı çaprazlama işlemi sonrasında, elit bireylerin direkt olarak aktarılması ve kalan bireylerin ise çocukları oluşturması ile yeni nesil bireyleri ortaya çıkar. Elit seçim işlemi, algoritmanın bu bölümünde gerçekleştirilir. Eski nesillerdeki iyi bireyleri kaybetmenin önüne geçme içgüdüğü ile neslindeki en uygun maliyete sahip olan belirli bir orandaki elit birey yeni nesle taşınır. Elit seçim işlemi ile popülasyon kurgulandıktan sonra tüm bireyler seçim, çaprazlama ve mutasyon işlemlerine tabidir.

Bu çalışmadaki her bir alt deneysel çalışmanın birinci aşaması için elit birey sayısı, popülasyondaki birey sayısının %10'u olacak şekilde belirlenmiştir. Bu sayede oldukça fazla sayıda kıymetli genin bir sonraki nesle aktarılması amaçlanmıştır. Alt deneysel çalışmaların birinci aşamasında bulunan değerli bilgilerin birleştirilerek geçildiği ikinci aşamada ise bu oran %2.5 değerinde tutulmuştur. Bu sayede algoritma ilerleyişi sırasında yeni yüksek doğruluk sonuçlarına erişirebilecek farklı genlerin de çaprazlama veya mutasyon ile ortaya daha yüksek oranda çıkabileceği düşünülmüştür.

3.5.2.3.3. Popülasyon büyüklüğü

Kromozomlar içerisindeki genler nesilden nesle aktarılırken popülasyon büyüklüğü oldukça önemlidir. Nesilleri oluşturan popülasyonlar içerisindeki kromozom sayısının az olması istendiğinde genel süreç kısaltılmış olacaktır. Ancak bu durumda, algoritma belki de genel optimum çözümü bulamadan daha kötü maliyetteki yerel optimum çözüme

takılıp kalacaktır. Bu doğrultuda popülasyon içerisindeki kromozom sayısının fazlalaşması çözümü bulma ihtimalini arttırmak adına daha önemli olarak yorumlanabilmektedir. Bu durumda da algoritma adımları daha fazla zamanda işleneceğinden genel süreç bir hayli uzayabilmektedir [34]. Tüm bu kıstaslar ele alındığında, seçilecek popülasyon büyüklüğünün kromozomdaki gen sayısı ile kromozomdaki gen sayısının iki katı arasında olması önerilmiştir [37].

Bu çalışmada ise ele alınan farklı öznitelik gruplarındaki öznitelik sayıları değişken olduğundan, ayrıca çalışma süresi açısından bir limit konulmadığından, deneysel çalışmalar yüksek değerlerde sabit popülasyon sayıları üzerinden başlatılmıştır. Başlangıç popülasyonunun büyüklüğüne dair bilgiler izleyen kesimde ilgili deneysel çalışmada uygulanan parametreler kısmında detaylandırılmaktadır.

3.5.2.4. Yeni nesil oluşturma

Bulunulan nesildeki bireyler üzerinden yeni neslin çocuklarının nasıl oluşturulacağı işlemi algoritmanın bu aşamasında gerçekleştirilir. Bireyler seçkincilik ilkesi ile nesilden nesle direkt aktarılabilceği gibi, aile bireyleri üzerinden yeni nesil için oluşturulan çocukların soylarını devam ettirmeleri de mümkündür. Mutasyon gibi rastlantısal etkilerin de gerçekleşmesi ile yeni bireyler elde edilebilir.

3.5.2.4.1. Seçim

Hangi bireylerin bir sonraki nesle genlerini aktaracağını tarifler. Başka bir söylem ile çocukları yaratacak aile bireylerinin seçim işleminin nasıl ele alınacağı işlemi gerçekleştirir. Elit diye tabir edilen en iyi maliyete sahip bireylerin direkt olarak yeni nesle aktarılması dışındaki durumlarda bireylere seçim işlemi uygulanır. Kişisel bir seçim fonksiyonu yazılabileceği gibi ölçeklendirilmiş değerler üzerinden aile bireylerini seçen alt fonksiyon çeşitleri de bulunur. *Stochastic uniform* seçimi ile ölçeklendirilmiş değeri ile orantılı olacak şekilde her bir aile bireyi bir bölüme yerleştirilir; algoritma, bölümler üzerinde eşit aralıklarla ilerler ve bu doğrultuda seçim yapılır. *Remainder* seçimi ile ölçeklendirilmiş değerler iki aşamada ele alınır. Değerin tam kısmı kadar listeleme yapılırken, noktalama kısmı ile de *roulette wheel* yaklaşımı uygulanır. Tam kısımlara göre atamalar yapıldıktan sonra kalan işlemler rastlantısal olarak sürdürülür. *Uniform*

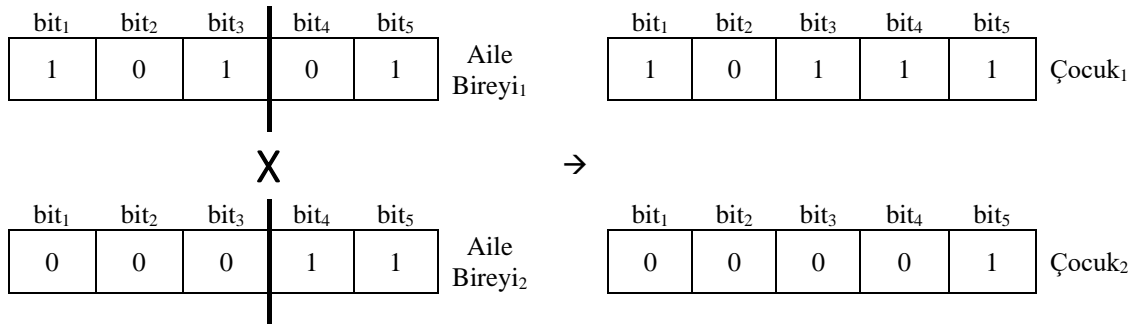
seçim, aile birey sayısını ve beklentileri göz önünde bulundurarak gerçekleşir; etkili bir strateji olmadığı belirtilmektedir. *Roulette wheel* seçimde, bireyler maliyetleri ile doğru orantılı olacak şekilde bir alana yerleştirilir. Ağırlıklı alanlar üzerinden rastlantısal olarak aile bireyleri seçilir. *Tournament* seçim ise belirtilen parametre değeri kadar bireyin rastgele seçildikten sonra aralarından en iyi bireyi seçmek üzerine kuruludur [34].

Bu çalışmada, hangi bireylerin bir sonraki nesle genlerini aktaracağını belirlemek için *roulette wheel* seçim yöntemi uygulanmıştır.

3.5.2.4.2. Çaprazlama

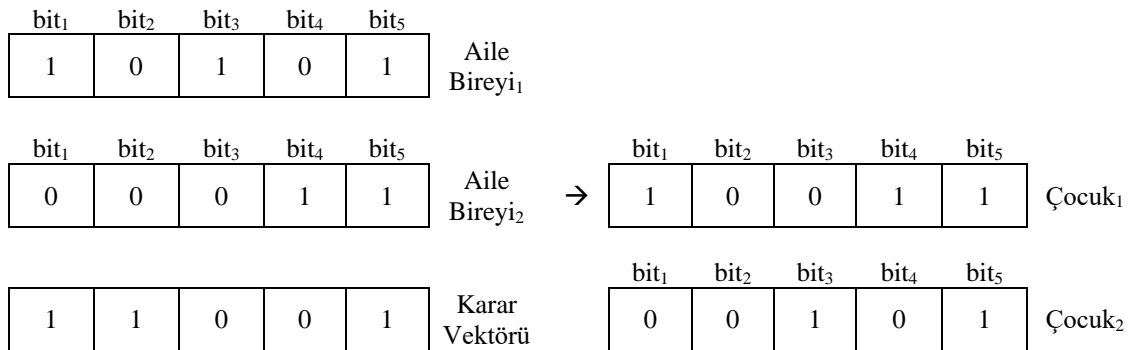
Bireyler üzerinden genlerin değiş tokuşu mantığı ile yapılacak çaprazlama işlemi sonrasında yeni bireyler oluşturulur. Başka bir ifade ile yeni nesil için çocukları oluşturmak üzere aile bireylerinin nasıl işleme alınacağı tariflenir. *Scattered* çaprazlaması ile ikili sistemde sayılara sahip olan bir vektör aracılığı ile ilgili indisteki değerinden anneden veya babadan alınması sağlanır. Yani çocuk genleri aile bireylerinden gen bazlı oluşturulur. *Single point* çaprazlaması ile aile bireylerinin kromozomları rastgele olarak ortak bir noktadan itibaren ikiye bölünür. Bölünen kısımlardan bir tarafı birinci aile bireyinden, diğer tarafı ikinci aile bireyinden alınacak şekilde çocuklar oluşturulur. *Two point* çaprazlaması ile aile bireylerinin kromozomları rastgele olarak ortak iki noktadan bölünür. Bireyler için oluşan üçer yeni alt setin kombinasyonları ile çocuklar oluşturulur [34]. Her yeni bir bireyin genlerinin aile bireylerinin genlerinin lineer bir kombinasyonundan oluşturulması yerine rastgele bir aile bireyinden alınması gibi uyarlamalar da mümkündür [38].

Bu çalışmada ise, deneysel çalışmaların birinci aşamalarında çaprazlama yöntemlerinden *single point* yaklaşımı kullanılmıştır. Standart GA öznitelik çaprazlama yönteminin kullanılması ile birlikte ele alınan kromozomlar öznitelik ayrımı olmadan çocukları meydana getirmiştir. Şekil 3.9'da, kromozom uzunluğu 5 bit olan iki aile bireyinin kromozomlarından rastgele seçilen ortak kırpma noktası (bit₃ | bit₄) sonrası oluşan iki yeni çocuk örneği görülmektedir. Rastgele oluşturulan başlangıç popülasyonu üzerinden, her iki çocuk da genlerinin bir kısmını bir aile bireyinden kalan kısmını ise diğer aile bireyinden alacak şekilde çaprazlanarak elde edilmiştir.



Şekil 3.9. 'Single Point' Çaprazlama Örneği

Deneysel çalışmaların ikinci aşamalarında çaprazlama yöntemlerinden *scattered* yaklaşımı kullanılmıştır. Birinci aşamada elde edilen verinin etkisinin kaybedilmemesi adına öznitelikler başlangıç popülasyonunda olabildiğince bir arada tutulmuştur; başka bir deyiş ile öznitelikler, birinci aşamadan gelen değerli gen gruplarının mümkün mertebe birbirinden ayrılmaması şeklinde ele alınmıştır. Doğrudan kod ile oluşturulan başlangıç popülasyonu üzerinden deneysel çalışmaların birinci aşamalarında elde edilen farklı gruplar içerisindeki kıymetli değerlerin, nesiller ilerledikçe yavaş yavaş bir araya getirilmesi sağlanmıştır. Bu sayede hali hazırda kendi grubu içerisinde kıymetli olan özniteliklerin ardışık olmayan bir kısmının diğer kıymetli özniteliklerin alt kümeleri ile buluşturulması sonucunda yeni nesil bireyleri oluşturulmuştur. Şekil 3.10'da, ikili sistemde rastgele oluşturulan bir vektörü temel alan ve kromozom boyutu 5 olan iki aile bireyine ait kromozomların bu vektör üzerinden yönetilmesi ile oluşan yeni çocukların örneği görülmektedir. Vektör içerisinde 1 değerine sahip olan bitler bir aile bireyinden, 0 değerine sahip olan bitler ise diğer aile bireyinden alınmıştır.



Şekil 3.10. 'Scattered' Çaprazlama Örneği

Bu çalışmada yeni bireylerin daha yüksek oranda çeşitliliğe sahip bir şekilde oluşması adına çaprazlama oranı mümkün olduğunca yüksek tutulmuştur. Bu oran %90 olarak belirlenmiştir. Çaprazlama parametresi, gen çeşitliliğinin sağlanmasında uygulanan parametrelerin en önemlilerinden birisidir.

3.5.2.4.3. Mutasyon

Kromozomları oluşturan genler üzerinden rastgele seçimler yapılır ve ilgili genler rastlantısal olarak ufak değişikliklere uğratılır. Popülasyonda hiç rastlanılmayan ancak diğerlerine göre çok daha etkili bir genin mutasyon sonrasında işlemlere dâhil edilmesi ile oluşan kromozomların daha iyi maliyetlere sahip olmaları sağlanabilir. Başka bir ifade ile başlangıç popülasyonunda optimum bireyin tüm genleri bulunmayabilir. Bu genlerin mutasyon ile genetik aramaya dâhil edilebilmesi ihtimali söz konusu olur. Mutasyon oranı %100 olduğunda, algoritma rastgele arama şeklinde davranır.

Bu çalışmada ise, deneysel çalışmalara geçilmeden önce öznitelik matrisleri ile ön testler yapılmıştır. Hali hazırda değerli diye nitelendirilebilecek farklı özniteliklerin de işleme dâhil olması ile daha iyi doğruluk sonuçlarının elde edilebilmesi adına, yapılan deneysel çalışmalarla en uygun mutasyon oranı %13 olarak belirlenmiştir.

3.5.2.5. Durdurma ölçütleri

GA ile optimum çözümü bulma aşamalarının sonlanması bazı kıstaslara bağlıdır. Maksimum jenerasyon sayısı, algoritmanın çalışma süresi limiti veya maliyet değeri limiti, jenerasyonların ilerlemesi sırasında en iyi maliyet sonuçları arasındaki değişiklik oranı gibi ölçütlere ulaşıldığında genetik algoritma durdurulur. Problemin tanımına yönelik genel olarak en iyi sonuç bulunamamış olsa bile o ana kadarki en iyi sonuç ile algoritma sonlanmış olur.

Bu çalışmada bahsi geçen kıstaslar için herhangi bir süre limiti konulmamıştır. Ancak, jenerasyonlar arası en iyi maliyet sonucunun değişiklik oranı varsayılan kontrol olarak değiştirilmemiştir. Maksimum jenerasyon sayısı deneysel çalışmanın tipine göre değişmektedir. Maliyet limiti olarak ise doğruluk sonucu %100 olduğu an algoritma yinelemelerinin durdurulması şeklinde gerekli ayarlamalar yapılmıştır. Algoritma durdurma ölçütlerine ulaştığında ise, son popülasyonda en yüksek doğruluk sonucunu

sağlayan kromozomu, yani değerli öznelikleri maskeleyen vektörü çıktı olarak vermektedir.

3.6. Çapraz Doğrulama ve Hata Matrisleri

k -katlamalı çapraz doğrulama tekniği, tüm verinin test performansını hesaplama işlemini yerine getirir. İfadedeki k parametresi, her bir çapraz doğrulama iterasyonu içerisinde tüm verinin yaklaşık olarak $1/k$ kadarlık parçasının test verisi, kalan verinin ise eğitim verisi olarak ayrıştırıldığını belirtir. Bir test çalışması esnasında bir örnek sadece bir kez test edilmiş olur [39].

Bu çalışmada GA üzerinden NB ve SVM için sınıflandırma performansı 10 katlamalı çapraz doğrulama üzerinden hesaplanmıştır. Tüm deneysel çalışmalardan önce 2-Sınıflı ve 3-Sınıflı testler için iki ayrı katlama seti rastgele bir kez oluşturulmuş olup, her bir ortak alt kategori testi aynı katlama değerleri ile tamamlanmıştır. 10 katlama için de setin $1/10$ 'u test, kalan $9/10$ 'u deney seti olarak alındığından, parçalı bulunan tahmin sonuçları 10 katlamanın sonunda toplanarak ilgili özneliklerin sınıflandırmasına ait hata matrisleri elde edilmiştir. Sonuç olarak test performansları, doğru şekilde tahmin edilen hücrelerdeki değerlerin tüm hücrelerdeki değerlerin toplamı üzerinden yüzdesel ifade edilmiştir. Örnek olarak Tablo 3.5'te, tüm katlamalar sonrasında elde edilen bir hata matrisi gösterilmektedir.

Tablo 3.5. *Hata Matrisi Örneği*

	AH	FTD	VaD
AH	17	0	2
FTD	1	16	2
VaD	2	3	20

Örnek olarak verilen Tablo 3.5'teki hata matrisinin ilk sütunundaki hastalıklar klinik tanıyı, ilk satırdaki hastalıklar ise sınıflandırma işlemi sonucunda bulunan tahmini ifade etmektedir. Matrisin asal köşegeni üzerindeki değerler doğru tahminleri nitelendirir. 19 AH hastalığına sahip numune içerisinde 17'sinin doğru tahmin edildiği, 2 AH numunesinin ise matematiksel işlemler sonucunda VaD şeklinde nitelendirildiği, herhangi bir AH numunesinin FTD olarak tespit edilmediği örnek bir açıklama olabilir. 63 örnek için, toplamda 53 tahmin doğru yapılmış olup 10 tahmin klinik

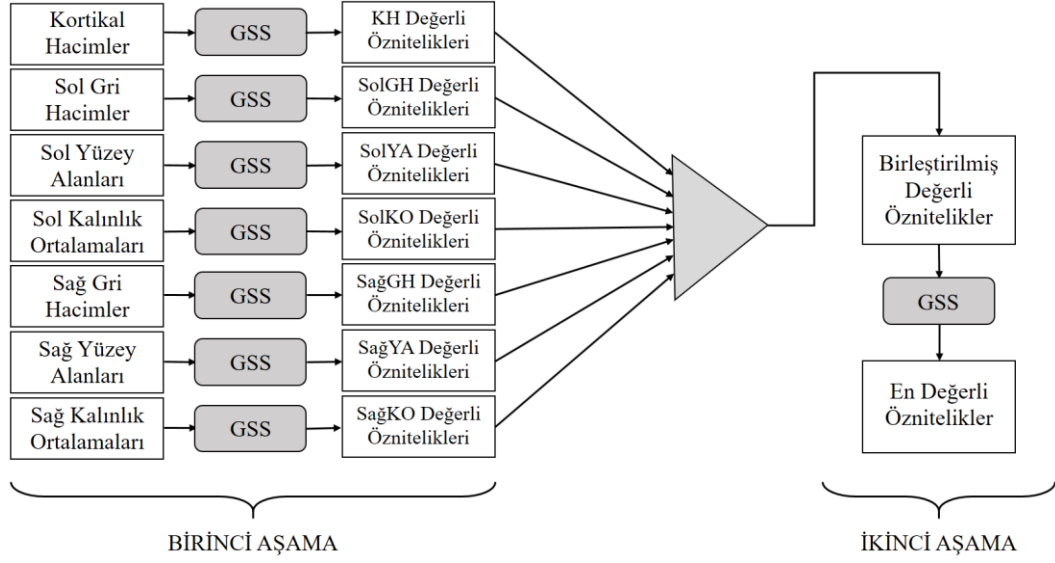
tanı ile uyuşmamaktadır. Bu doğrultuda tüm numuneler üzerinden doğruluk sonucu 53/63 oranı üzerinden %84.1 olarak hesaplanmıştır.

Satırların noktalı virgül ile ayrıldığı ve tek satırda ifade edildiği hata matrisi, Tablo 3.5 örneği için $[17 \ 0 \ 2;1 \ 16 \ 2;2 \ 3 \ 20]$ şeklinde de gösterilebilmektedir. Deneysel çalışmalar sırasında bulunan hata matrisleri, çalışmanın kalan kısmında belirtilen tek satırlı biçimde de ifade edilmektedir.

4. DENEYSEL ÇALIŞMALAR

Önceki bölümlerde anlatılan yöntemler üzerinden deneysel çalışmalar iki türde ilerletilmiştir. AH, FTD, VaD olmak üzere tüm sınıfların bulunduğu 3-Sınıflı ve *AH v diğerleri, FTD v diğerleri, VaD v diğerleri* şeklinde her bir sınıfın kalan tüm sınıflara göre “ilgili sınıftır” veya “ilgili sınıf değildir” yaklaşımı ile karşılaştırılarak ele alındığı 2-Sınıflı deneysel çalışmalar yapılmıştır. 10 katlamalı çapraz doğrulama test aşamasında kullanılacak katlama değerleri her iki deneysel çalışma türü için de çalışmalara başlamadan önce yalnızca bir kez elde edilmiş ve deneysel çalışmaların kalan kısmı aynı katlama değerleri ile sürdürülmüştür. Detaylı olacak şekilde, 3-Sınıflı testler için ayrı bir katlama seti ve 2-Sınıflı testler için ayrı bir katlama seti olmak üzere iki farklı katlama seti tüm çalışmalara başlamadan önce bir kez rastgele oluşturulmuştur. İlgili deneysel çalışma hangi katlama değerleri ile başlatıldıysa, çalışma sonlanana kadar aynı katlama değerleri kullanılmıştır. 3-Sınıflı deneysel çalışmalarda NB sınıflandırıcı, 2-Sınıflı deneysel çalışmalarda ise hem SVM, hem de NB sınıflandırıcının sonuçları GA tabanlı sarmalama öznitelik seçimi yapılarak kıyaslanmıştır.

Her iki türdeki deneysel çalışmalar da kendi içerisinde iki aşamadan oluşmaktadır. Deneysel çalışmaların birinci aşamalarında Freesurfer üzerinden elde edilen öznitelik gruplarının (beynin tümüne yönelik kortikal hacimler, sağ ve sol beyin bölümleri için gri hacimler, yüzey alanları ve kalınlık ortalamaları) kendi içerisindeki değerli özniteliklerinin GA aracılığıyla bulunması amaçlanmıştır. Bu sayede öznitelik grupları içerisinde değerli öznitelikler, tüm grubun boyutuna nazaran daha küçük boyutlu matrisler üzerinden ikinci aşamaya taşınmıştır. İkinci aşamada ise, birinci aşamada elde edilen daha etkili değerler, başka bir ifade ile kendi alt grubunun en iyi öznitelikleri, birleştirilerek en değerli özniteliklerin bulunduğu bir girdi matrisi elde edilmiştir. Deneysel çalışma aşamalarının anlatılan tüm süreci Şekil 4.1’de görsel olarak sunulmaktadır. İkinci aşamada, birinci aşamada kullanılan GA parametrelerinden farklı parametreler uygulanmıştır; ilgili deneysel çalışmaya ait tüm beyin bölümlerine yönelik değerli özniteliklerin bulunması amaçlanmıştır.



Şekil 4.1. Deneysel Çalışma Aşamaları

İzleyen kesimde gerçekleştirilen deneysel çalışmalar esnasında uygulanan algoritma parametreleri ve tüm bu sürecin sonrasında elde edilen bulgular detaylandırılmaktadır.

4.1. 3-Sınıflı Deneysel Çalışmalar ve Sonuçları

3-Sınıflı deneysel çalışmalar, Şekil 4.1’de de belirtildiği gibi farklı öznitelik grupları için oluşturulan ve Denklem (3.1) ile normalize edilmiş öznitelik matrislerinin sıra ile aynı prosedüre sokulması üzerine kurulmuştur. Bu türdeki çalışmaların birinci aşamasında öznitelik seçimi için uygulanan GA parametreleri Tablo 4.1’de sunulmaktadır.

Tablo 4.1. 3-Sınıflı Genetik Algoritma Birinci Aşama Parametreleri

Parametre	Değer
Maksimum Jenerasyon Sayısı	1000
Popülasyon Büyüklüğü	800
Elit Sayısı	%10
Kromozom Tipi	<i>bit string</i>
<i>Fitness</i> Ölçekleme Fonksiyonu	<i>rank</i> ölçekleme fonksiyonu
Çaprazlama Oranı	%90
Çaprazlama Fonksiyonu	<i>single point</i> çaprazlama fonksiyonu
Seçim Fonksiyonu	<i>roulette wheel</i> seçim fonksiyonu
Mutasyon Oranı	%13

Öznitelik seçimi için uygulanan GA parametreleri ile 3-Sınıflı deneysel çalışmalar, birinci aşamada ve ikinci aşamada maksimum jenerasyon sayısına ulaşmadan, maliyet değerindeki değişim kontrolü ile sonlanmıştır. Birinci aşamada öznitelik gruplarının GA ile NB sınıflandırıcısı üzerinden elde edilen sonuçları Tablo 4.2’de sunulmaktadır.

Tablo 4.2. *Naïve Bayes ile 3-Sınıflı Genetik Algoritma Birinci Aşama Sonuçları*

Öznitelik Grubu	Toplam Öznitelik Sayısı	En İyi Öznitelik Sayısı	Doğruluk (%)	Hata Matrisi (AH, FTD, VaD)
Kortikal Hacimler	41	17	84.1	[17 0 2;1 16 2;2 3 20]
Sol Gri Hacimler	34	17	69.8	[10 5 4;0 18 1;4 5 16]
Sol Yüzey Alanları	34	16	74.6	[11 5 3;1 14 4;2 1 22]
Sol Kalınlık Ortalamaları	34	11	69.8	[11 2 6;4 11 4;3 0 22]
Sağ Gri Hacimler	34	16	71.4	[7 4 8;0 16 3;0 3 22]
Sağ Yüzey Alanları	34	17	63.5	[10 8 1;4 12 3;4 3 18]
Sağ Kalınlık Ortalamaları	34	9	73.0	[11 1 7;1 14 4;2 2 21]

Birinci aşama sonucunda her bir öznitelik grubu için elde edilen en iyi öznitelikler birleştirilerek ikinci aşamanın girdi matrisi oluşturulmuştur. 3-Sınıflı deneysel çalışmaların ikinci aşamasında öznitelik seçimi için uygulanan GA parametreleri Tablo 4.3’te sunulmaktadır.

Tablo 4.3. *3-Sınıflı Genetik Algoritma İkinci Aşama Parametreleri*

Parametre	Değer
Maksimum Jenerasyon Sayısı	1000
Popülasyon Büyüklüğü	800
Elit Sayısı	% 2.5
Kromozom Tipi	<i>bit string</i>
<i>Fitness</i> Ölçekleme Fonksiyonu	<i>proportional</i> ölçekleme fonksiyonu
Çaprazlama Oranı	%90
Çaprazlama Fonksiyonu	<i>scattered</i> çaprazlama fonksiyonu
Seçim Fonksiyonu	<i>roulette wheel</i> seçim fonksiyonu
Mutasyon Oranı	%13

Deneysel çalışmanın birinci aşamasında bulunan değerli özniteliklerin birleştirilmesi ile oluşan girdi matrisi üzerinden ikinci aşama parametrelerinin kullanılması ile 3-Sınıflı deneysel çalışmalara ait bulunan nihai sonuçlar Tablo 4.4’te sunulmaktadır.

Tablo 4.4. *Naïve Bayes ile 3-Sınıflı Genetik Algoritma İkinci Aşama Sonuçları*

Toplam Öznitelik Sayısı	En İyi Öznitelik Sayısı	Doğruluk (%)	Hata Matrisi	Kıymetli Öznitelikler		
				Tür	Taraf	Bölge
103	11	90.5	$\begin{bmatrix} 16 & 1 & 2 \\ 1 & 17 & 1 \\ 0 & 1 & 24 \end{bmatrix}$	hacim	sol	<i>putamen</i>
				hacim	tüm	<i>4th ventricle</i>
				hacim	sol	<i>vessel</i>
				hacim	sol	<i>choroid-plexus</i>
				gri hacim	sol	<i>bankssts</i>
				gri hacim	sol	<i>postcentral</i>
				yüzey alanı	sol	<i>entorhinal</i>
				yüzey alanı	sol	<i>inferiortemporal</i>
				yüzey alanı	sol	<i>isthmuscingulate</i>
				yüzey alanı	sağ	<i>middletemporal</i>
yüzey alanı	sağ	<i>supramarginal</i>				

4.2. 2-Sınıflı Deneysel Çalışmalar ve Sonuçları

2-Sınıflı deneysel çalışmalarda, farklı öznitelik grupları için oluşturulan ve Denklem (3.1) ile normalize edilmiş öznitelik matrisleri her bir sınıfın kalan tüm sınıflara göre “ilgili sınıftır” veya “ilgili sınıf değildir” yaklaşımı ile karşılaştırılmıştır. Veri setinde üç farklı sınıf bulunduğundan 2-Sınıflı deneysel çalışmalar kendi içerisinde üçe ayrılmaktadır. Tüm 2-Sınıflı alt deneysel çalışmalarda aynı GA parametreleri kullanılacak biçimde hastalıklar *AH v diğerleri*, *FTD v diğerleri*, *VaD v diğerleri* şeklinde ele alınarak gruplandırılmıştır. Bu türün alt deneysel çalışmalarına başlamadan önce öznitelik matrislerinin sınıfları belirten hücreleri, incelenmek istenilen sınıfa göre güncellenmiştir. 2-Sınıflı deneysel çalışmaların birinci aşamalarında öznitelik seçimi için uygulanan GA parametreleri Tablo 4.5’te, ikinci aşamalarında uygulanan GA parametreleri ise Tablo 4.6’da sunulmaktadır.

Tablo 4.5. *2-Sınıflı Genetik Algoritma Birinci Aşama Parametreleri*

Parametre	Değer
Maksimum Jenerasyon Sayısı	500
Popülasyon Büyüklüğü	400
Elit Sayısı	%10
Kromozom Tipi	<i>bit string</i>
<i>Fitness</i> Ölçekleme Fonksiyonu	<i>rank</i> ölçekleme fonksiyonu
Çaprazlama Oranı	%90
Çaprazlama Fonksiyonu	<i>single point</i> çaprazlama fonksiyonu
Seçim Fonksiyonu	<i>roulette wheel</i> seçim fonksiyonu
Mutasyon Oranı	%13

Tablo 4.6. 2-Sınıflı Genetik Algoritma İkinci Aşama Parametreleri

Parametre	Değer
Maksimum Jenerasyon Sayısı	1000
Popülasyon Büyüklüğü	800
Elit Sayısı	% 2.5
Kromozom Tipi	<i>bit string</i>
<i>Fitness</i> Ölçekleme Fonksiyonu	<i>proportional</i> ölçekleme fonksiyonu
Çaprazlama Oranı	%90
Çaprazlama Fonksiyonu	<i>scattered</i> çaprazlama fonksiyonu
Seçim Fonksiyonu	<i>roulette wheel</i> seçim fonksiyonu
Mutasyon Oranı	%13

Öz nitelik seçimi için uygulanan GA parametreleri ile 2-Sınıflı deneysel çalışmalar, birinci aşamada ve ikinci aşamada maksimum jenerasyon sayısına ulaşmadan, maliyet değerindeki değişim kontrolü ile sonlanmıştır.

4.2.1. AH v diğerleri sınıflandırması deneysel çalışmaları ve sonuçları

Bu deneysel çalışmada test edilecek numunelerin AH olup olmadığı kontrolü yapılmaktadır. Deneysel çalışmaya başlamadan önce veri seti içerisindeki tüm numuneler 19 adet *AH v diğerleri* olacak şekilde gruplanmıştır. Girdi olarak gösterilecek her bir öz nitelik matrisinde sınıf bilgisini tutan hücreler güncellenmiştir. AH tespiti için 2-sınıflı testler NB ve SVM kullanılarak Tablo 4.5 ve Tablo 4.6’da tüm 2-Sınıflı deneysel çalışmalar sırasında kullanılacağı belirtilen GA parametreleri ile gerçekleştirilmiştir.

Öz nitelik grupları üzerinden NB ile bulunan deneysel çalışmanın birinci aşama sonuçları Tablo 4.7’de sunulmaktadır.

Tablo 4.7. Naïve Bayes ile AH v Diğerleri Sınıflandırması Genetik Algoritma Birinci Aşama Sonuçları

Öz nitelik Grubu	Toplam Öz nitelik Sayısı	En İyi Öz nitelik Sayısı	Doğruluk (%)	Hata Matrisi
Kortikal Hacimler	41	18	92.1	[18 1;4 40]
Sol Gri Hacimler	34	18	88.9	[16 3;4 40]
Sol Yüzey Alanları	34	17	79.4	[12 7;6 38]
Sol Kalınlık Ortalamaları	34	18	88.9	[16 3;4 40]
Sağ Gri Hacimler	34	17	81.0	[16 3;9 35]
Sağ Yüzey Alanları	34	18	93.7	[17 2;2 42]
Sağ Kalınlık Ortalamaları	34	18	88.9	[14 5;2 42]

NB ile birinci aşamada bulunan değerli özniteliklerin birleştirilmesi sonrasında elde edilen ikinci aşama sonuçları ise Tablo 4.8’de sunulmaktadır.

Tablo 4.8. *Naïve Bayes ile AH v Diğerleri Sınıflandırması Genetik Algoritma İkinci Aşama Sonuçları*

Toplam Öznitelik Sayısı	En İyi Öznitelik Sayısı	Doğruluk (%)	Hata Matrisi	Kıymetli Öznitelikler		
				Tür	Taraf	Bölge
124	17	93.7	$\begin{bmatrix} 17 & 2 \\ 2 & 42 \end{bmatrix}$	hacim	sol	<i>putamen</i>
				hacim	tüm	<i>3rd ventricle</i>
				hacim	sol	<i>vessel</i>
				hacim	sağ	<i>cerebellum wm</i>
				gri hacim	sol	<i>bankssts</i>
				gri hacim	sol	<i>medialorbitofrontal</i>
				yüzey alanı	sol	<i>posteriorcingulate</i>
				kal. ort.	sol	<i>posteriorcingulate</i>
				gri hacim	sağ	<i>fusiform</i>
				gri hacim	sağ	<i>rostralmiddlefrontal</i>
				yüzey alanı	sağ	<i>caudalanteriorcing...</i>
				yüzey alanı	sağ	<i>postcentral</i>
				yüzey alanı	sağ	<i>precuneus</i>
				yüzey alanı	sağ	<i>transversetemporal</i>
				kal. ort.	sağ	<i>inferiortemporal</i>
				kal. ort.	sağ	<i>parahippocampal</i>
kal. ort.	sağ	<i>insula</i>				

Öznitelik grupları üzerinden SVM ile bulunan deneysel çalışmanın birinci aşama sonuçları Tablo 4.9’da sunulmaktadır.

Tablo 4.9. *SVM ile AH v Diğerleri Sınıflandırması Genetik Algoritma Birinci Aşama Sonuçları*

Öznitelik Grubu	Toplam Öznitelik Sayısı	En İyi Öznitelik Sayısı	Doğruluk (%)	Hata Matrisi
Kortikal Hacimler	41	10	82.5	[8 11;0 44]
Sol Gri Hacimler	34	7	88.9	[13 6;1 43]
Sol Yüzey Alanları	34	8	82.5	[12 7;4 40]
Sol Kalınlık Ortalamaları	34	11	85.7	[10 9;0 44]
Sağ Gri Hacimler	34	9	82.5	[10 9;2 42]
Sağ Yüzey Alanları	34	10	85.7	[10 9;0 44]
Sağ Kalınlık Ortalamaları	34	8	90.5	[14 5;1 43]

SVM ile birinci aşamada bulunan değerli özniteliklerin birleştirilmesi sonrasındaki ikinci aşama sonuçları ise Tablo 4.10’da sunulmaktadır.

Tablo 4.10. SVM ile AH v Diğerleri Sınıflandırması Genetik Algoritma İkinci Aşama Sonuçları

Toplam Öznitelik Sayısı	En İyi Öznitelik Sayısı	Doğruluk (%)	Hata Matrisi	Kıymetli Öznitelikler		
				Tür	Taraf	Bölge
63	8	92.1	$\begin{bmatrix} 14 & 5 \\ 0 & 44 \end{bmatrix}$	gri hacim	sol	<i>bankssts</i>
				gri hacim	sol	<i>entorhinal</i>
				gri hacim	sol	<i>paracentral</i>
				gri hacim	sol	<i>parsopercularis</i>
				kal. ort.	sol	<i>lingual</i>
				kal. ort.	sol	<i>medialorbitofrontal</i>
				gri hacim	sağ	<i>superiortemporal</i>
				kal. ort.	sağ	<i>lingual</i>

4.2.2. FTD v diğerleri sınıflandırması deneysel çalışmaları ve sonuçları

Bu deneysel çalışmada test edilecek numunelerin FTD olup olmadığı kontrolü yapılmaktadır. Deneysel çalışmaya başlamadan önce veri seti içerisindeki tüm numuneler 19 adet *FTD v diğerleri* olacak şekilde gruplanmıştır. Girdi olarak gösterilecek her bir öznitelik matrisinde sınıf bilgisini tutan hücreler güncellenmiştir. FTD tespiti için 2-sınıflı testler NB ve SVM kullanılarak tüm 2-Sınıflı deneysel çalışmalar sırasında kullanılacağı belirtilen GA parametreleri ile gerçekleştirilmiştir.

Öznitelik grupları üzerinden NB ile bulunan deneysel çalışmanın birinci aşama sonuçları Tablo 4.11’de, birinci aşamada bulunan değerli özniteliklerin birleştirilmesi sonrasındaki ikinci aşama sonuçları ise Tablo 4.12’de sunulmaktadır.

Tablo 4.11. Naïve Bayes ile FTD v Diğerleri Sınıflandırması Genetik Algoritma Birinci Aşama Sonuçları

Öznitelik Grubu	Toplam Öznitelik Sayısı	En İyi Öznitelik Sayısı	Doğruluk (%)	Hata Matrisi
Kortikal Hacimler	41	17	90.5	[16 3;3 41]
Sol Gri Hacimler	34	17	90.5	[19 0;6 38]
Sol Yüzey Alanları	34	17	88.9	[14 5;2 42]
Sol Kalınlık Ortalamaları	34	17	85.7	[14 5;4 40]
Sağ Gri Hacimler	34	16	81.0	[9 10;2 42]
Sağ Yüzey Alanları	34	16	82.5	[10 9;2 42]
Sağ Kalınlık Ortalamaları	34	16	87.3	[13 6;2 42]

Tablo 4.12. *Naïve Bayes ile FTD v Diğerleri Sınıflandırması Genetik Algoritma İkinci Aşama Sonuçları*

Toplam Öznitelik Sayısı	En İyi Öznitelik Sayısı	Doğruluk (%)	Hata Matrisi	Kıymetli Öznitelikler		
				Tür	Taraf	Bölge
116	17	95.2	$\begin{bmatrix} 17 & 2 \\ 1 & 43 \end{bmatrix}$	hacim	sol	<i>vessel</i>
				hacim	sol	<i>choroid-plexus</i>
				hacim	sağ	<i>caudate</i>
				hacim	sağ	<i>putamen</i>
				hacim	sağ	<i>ventralDC</i>
				hacim	tüm	<i>5th ventricle</i>
				gri hacim	sol	<i>postcentral</i>
				yüzey alanı	sol	<i>bankssts</i>
				yüzey alanı	sol	<i>paracentral</i>
				yüzey alanı	sol	<i>parstriangularis</i>
				yüzey alanı	sol	<i>precentral</i>
				gri hacim	sağ	<i>medialorbitofrontal</i>
				yüzey alanı	sağ	<i>paracentral</i>
				kal. ort.	sağ	<i>parstriangularis</i>
				kal. ort.	sağ	<i>precuneus</i>
				kal. ort.	sağ	<i>superiorparietal</i>
kal. ort.	sağ	<i>superiortemporal</i>				

Öznitelik grupları üzerinden SVM ile bulunan deneysel çalışmanın birinci aşama sonuçları Tablo 4.13'te sunulmaktadır.

Tablo 4.13. *SVM ile FTD v Diğerleri Sınıflandırması Genetik Algoritma Birinci Aşama Sonuçları*

Öznitelik Grubu	Toplam Öznitelik Sayısı	En İyi Öznitelik Sayısı	Doğruluk (%)	Hata Matrisi
Kortikal Hacimler	41	9	82.5	[8 11;0 44]
Sol Gri Hacimler	34	11	82.5	[8 11;0 44]
Sol Yüzey Alanları	34	9	82.5	[8 11;0 44]
Sol Kalınlık Ortalamaları	34	8	82.5	[9 10;1 43]
Sağ Gri Hacimler	34	12	81.0	[7 12;0 44]
Sağ Yüzey Alanları	34	12	79.4	[6 13;0 44]
Sağ Kalınlık Ortalamaları	34	10	85.7	[11 8;1 43]

SVM ile birinci aşamada bulunan değerli özniteliklerin birleştirilmesi sonrasındaki ikinci aşama sonuçları ise Tablo 4.14'te sunulmaktadır.

Tablo 4.14. SVM ile FTD v Diğerleri Sınıflandırması Genetik Algoritma İkinci Aşama Sonuçları

Toplam Öznitelik Sayısı	En İyi Öznitelik Sayısı	Doğruluk (%)	Hata Matrisi	Kıymetli Öznitelikler		
				Tür	Taraf	Bölge
71	9	90.5	$\begin{bmatrix} 13 & 6 \\ 0 & 44 \end{bmatrix}$	hacim	sağ	<i>putamen</i>
				gri hacim	sol	<i>fusiform</i>
				gri hacim	sol	<i>superiorfrontal</i>
				gri hacim	sağ	<i>bankssts</i>
				gri hacim	sağ	<i>inferiorparietal</i>
				gri hacim	sağ	<i>isthmuscingulate</i>
				gri hacim	sağ	<i>posteriorcingulate</i>
				gri hacim	sağ	<i>transversetemporal</i>
				yüzey alanı	sağ	<i>pericalcarine</i>

4.2.3. VaD v diğerleri sınıflandırması deneysel çalışmaları ve sonuçları

Bu deneysel çalışmada test edilecek numunelerin VaD olup olmadığı kontrolü yapılmaktadır. Deneysel çalışmaya başlamadan önce veri seti içerisindeki tüm numuneler 25 adet *VaD v diğerleri* olacak şekilde gruplanmıştır. Girdi olarak gösterilecek her bir öznitelik matrisinde sınıf bilgisini tutan hücreler güncellenmiştir. VaD tespiti için 2-sınıflı testler NB ve SVM kullanılarak tüm 2-Sınıflı deneysel çalışmalar sırasında kullanılacağı belirtilen GA parametreleri ile gerçekleştirilmiştir.

Öznitelik grupları üzerinden NB ile bulunan deneysel çalışmanın birinci aşama sonuçları Tablo 4.15'te, birinci aşamada bulunan değerli özniteliklerin birleştirilmesi sonrasında elde edilen ikinci aşama sonuçları ise Tablo 4.16'da sunulmaktadır.

Tablo 4.15. Naïve Bayes ile VaD v Diğerleri Sınıflandırması Genetik Algoritma Birinci Aşama Sonuçları

Öznitelik Grubu	Toplam Öznitelik Sayısı	En İyi Öznitelik Sayısı	Doğruluk (%)	Hata Matrisi
Kortikal Hacimler	41	24	87.3	[21 4;4 34]
Sol Gri Hacimler	34	11	88.9	[20 5;2 36]
Sol Yüzey Alanları	34	10	88.9	[19 6;1 37]
Sol Kalınlık Ortalamaları	34	23	82.5	[16 9;2 36]
Sağ Gri Hacimler	34	23	81.0	[16 9;3 35]
Sağ Yüzey Alanları	34	13	79.4	[13 12;1 37]
Sağ Kalınlık Ortalamaları	34	23	82.5	[18 7;4 34]

Tablo 4.16. *Naïve Bayes ile VaD v Diğerleri Sınıflandırması Genetik Algoritma İkinci Aşama Sonuçları*

Toplam Öznitelik Sayısı	En İyi Öznitelik Sayısı	Doğruluk (%)	Hata Matrisi	Kıymetli Öznitelikler		
				Tür	Taraf	Bölge
127	10	95.2	$\begin{bmatrix} 23 & 2 \\ 1 & 37 \end{bmatrix}$	hacim	tüm	4 th ventricle
				hacim	sağ	choroid-plexus
				gri hacim	sol	entorhinal
				gri hacim	sol	inferiorparietal
				gri hacim	sol	isthmuscingulate
				gri hacim	sol	postcentral
				gri hacim	sol	transversetemporal
				yüzey alanı	sol	cuneus
				gri hacim	sağ	supramarginal
				yüzey alanı	sağ	parahippocampal

Öznitelik grupları üzerinden SVM ile bulunan deneysel çalışmanın birinci aşama sonuçları Tablo 4.17’de, ikinci aşama sonuçları ise Tablo 4.18’de sunulmaktadır.

Tablo 4.17. *SVM ile VaD v Diğerleri Sınıflandırması Genetik Algoritma Birinci Aşama Sonuçları*

Öznitelik Grubu	Toplam Öznitelik Sayısı	En İyi Öznitelik Sayısı	Doğruluk (%)	Hata Matrisi
Kortikal Hacimler	41	9	90.5	[19 6;0 38]
Sol Gri Hacimler	34	10	85.7	[17 8;1 37]
Sol Yüzey Alanları	34	10	82.5	[14 11;0 38]
Sol Kalınlık Ortalamaları	34	12	77.8	[15 10;4 34]
Sağ Gri Hacimler	34	11	87.3	[17 8;0 38]
Sağ Yüzey Alanları	34	11	81.0	[14 11;1 37]
Sağ Kalınlık Ortalamaları	34	10	79.4	[15 10;3 35]

Tablo 4.18. *SVM ile VaD v Diğerleri Sınıflandırması Genetik Algoritma İkinci Aşama Sonuçları*

Toplam Öznitelik Sayısı	En İyi Öznitelik Sayısı	Doğruluk (%)	Hata Matrisi	Kıymetli Öznitelikler		
				Tür	Taraf	Bölge
73	8	93.7	$\begin{bmatrix} 21 & 4 \\ 0 & 38 \end{bmatrix}$	hacim	sol	lateral-ventricle
				hacim	sol	putamen
				hacim	sağ	cerebellum-cortex
				hacim	sağ	pallidum
				gri hacim	sol	parsopercularis
				gri hacim	sol	precuneus
				gri hacim	sağ	caudalanteriorcing...
				yüzey alanı	sağ	precentral

4.3. Deneysel Sonuçlar ve Değerlendirme

2-Sınıflı ve 3-Sınıflı deneysel çalışmaların birinci aşamalarına ait test sonuçları analiz edildiğinde, diğer öznitelik gruplarına nazaran hacim öznitelik gruplarının daha yüksek doğruluk sonucu ürettiği görülebilmektedir. Tüm deneysel çalışmalarda birinci aşamanın ardından öznitelik grubu esaslı bulunan en iyi öznitelikler birleştirilerek ikinci aşamanın öznitelik matrisi oluşturulmuştur. İkinci aşamada elde edilen doğruluk sonuçlarının birinci aşamalara göre eşit veya birinci aşamalardaki alt gruplar üzerinden elde edilen sonuçlardan daha iyi bulunduğu gözlemlenmiştir. Eşit olduğu durum NB ile *AH* v *diğerleri* sınıflandırmasında görülmektedir. Bu deneysel çalışmada da ikinci aşamada bulunan en değerli öznitelikler beynin tümüne daha az sayıda yayıldığından birinci aşama sonuçlarını iyi bir şekilde kullandığı yorumlanabilir.

Deneysel çalışmaların genelinde, birleştirilen kıymetli öznitelikler üzerinden elde edilen nihai sonuçlar incelendiğinde, her bir alt deneysel çalışmaya ait değerli öznitelik setleri içerisinde farklı özniteliklerin bulunduğu gözlemlenmektedir. Ancak bazı öznitelikler farklı deneysel çalışmalarda da elde edilen bulgular arasında göze çarpmaktadır. Bu özniteliklerin ilgili hastalığın tanısında ihtiyaç duyulan, değerli bir öznitelik olup olmamasının haricinde, değerine bağlı olarak farklı hastalıklara etki ettiği şeklinde yorumlanmaktadır. Farklı sınıflandırıcılar ile toplamda yapılan 7 deneysel çalışmanın 5'inde, beynin sağ veya sol bölümlerinden *bankssts* gri bölge hacimleri veya yüzey alanları ve *putamen* hacimleri özniteliklerinin kıymetli bulunduğu görülmüştür. Benzer şekilde 4 deneysel çalışmada beynin sağ veya sol bölümlerinden *postcentral* gri bölge hacimleri veya yüzey alanları genetik algoritma süreci sonucunda değerli öznitelikler çıktısında gözlemlenebilmektedir. Gerçekleştirilen 3 deneysel çalışmada ise beynin sol bölümünden *vessel* hacmi göze çarpmaktadır. Benzer şekilde beynin sağ veya sol bölümlerinin gri hacim, kalınlık ortalaması, yüzey alanı kriterlerinden bir veya birkaçı üzerinden *choroid-plexus*, *entorhinal*, *isthmuscingulate*, *medialorbitofrontal*, *paracentral*, *posteriorcingulate*, *precuneus* ve *transversetemporal* beyin bölümleri ile üçer kez karşılaşılmıştır. Bu değerlendirmeler doğrultusunda, beynin *bankssts*, *postcentral* ve *putamen* bölümlerinin bu çalışma kapsamında genel olarak 3 hastalığı belirlemede en değerli öznitelikler olarak nitelendirildiği söylenebilir. İkincil değerliler sıralamasında ise beynin *choroid-plexus*, *entorhinal*, *isthmuscingulate*,

medialorbitofrontal, paracentral, posteriorcingulate, precuneus, transversetemporal ve *vessel* bölümleri sayılabilir.

Her bir 2-Sınıflı deneysel çalışmanın aslında 3-Sınıflı deneysel çalışmaların bir alt versiyonu olarak yorumlanmak istendiğinde sonuçlar da buna destek olmaktadır. 3-Sınıflı ve 2-Sınıflı deneysel çalışmalarda elde edilen sonuçlar incelendiğinde, 3-Sınıflı deneysel çalışmalar sonucunda değerli öznitelik olarak elde edilen *middletemporal* yüzey alanı özneliği hariç tüm özniteliklerin 2-Sınıflı deneysel çalışmalarda da en az bir defa değerli olarak bulunduğu göze çarpmaktadır. Başka bir ifade ile belirtilen beyin bölümünün sadece üç hastalığın sınıflandırılmasında değerli olarak bulunduğu gözlemlenmektedir.

Her iki sınıflandırıcının da hastalık temelli elde ettiği öznitelikler yorumlanmak üzere gruplandırılmıştır. Sonrasında, 3-Sınıflı deneysel çalışmalar ve *AH v diğerleri* ile yapılan 2-Sınıflı deneysel çalışmalar içerisindeki ortak öznitelikler incelendiğinde, beynin sol bölümünden *putamen* ve *vessel* hacimleri, *bankssts* gri hacimleri ile beynin sağ veya sol bölümünden *postcentral, entorhinal* ve *inferiortemporal* bölümlerine ait değerlerin her iki çalışmada da bulunduğu gözlemlenmektedir. Belirtilen beyin bölümlerinin, AH'nin belirlenmesinde ve üç hastalığın sınıflandırmasında gerekli olduğu şeklinde yorumlanmaktadır. Benzer şekilde, 3-Sınıflı deneysel çalışmalar ve *FTD v diğerleri* ile yapılan 2-Sınıflı deneysel çalışmalar içerisindeki ortak öznitelikler incelendiğinde beynin sol veya sağ bölümünden *putamen, vessel* ve *choroid plexus* hacimleri, ayrıca *bankssts, postcentral* ve *isthmuscingulate* bölümlerine ait değerlerin her iki çalışmada da bulunduğu gözlemlenmektedir. Belirtilen beyin bölümlerinin, FTD'nin belirlenmesinde ve üç hastalığın sınıflandırmasında gerekli olduğu şeklinde yorumlanmaktadır. 3-Sınıflı deneysel çalışmalar ve *VaD v diğerleri* ile yapılan 2-Sınıflı deneysel çalışmalar içerisindeki ortak öznitelikler incelendiğinde ise beynin sol veya sağ bölümünden *putamen* ve *choroid plexus* hacimleri, *4th ventricle* hacimleri ile *postcentral entorhinal* ve *supramarginal* bölümlerine ait değerlerin her iki çalışmada da bulunduğu gözlemlenmektedir. Belirtilen beyin bölümlerinin, VaD'nin belirlenmesinde ve üç hastalığın sınıflandırmasında gerekli olduğu şeklinde yorumlanmaktadır.

Ayrıca, 2-Sınıflı deneysel çalışma sonuçlarında 3-Sınıflı deneysel çalışmalara nazaran farklı özniteliklerin de değerli öznitelik olarak bulunduğu görülmektedir. 3-Sınıflı deneysel çalışmalar ve *AH v diğerleri* ile yapılan 2-Sınıflı deneysel çalışmalar kıyaslandığında, beynin *3rd ventricle, sağ cerebellum beyaz madde* hacimleri ile birlikte sol beynin *medialorbitofrontal, paracentral, parsopercularis* bölümlerinin gri hacimleri;

posteriorcingulate yüzey alanı; *lingual*, *medialorbitofrontal*, *posteriorcingulate* kalınlık ortalamaları; ayrıca, sağ beynin *fusiform*, *rostralmiddlefrontal*, *superiortemporal* gri hacimleri; *caudalanteriorcingulate*, *precuneus*, *transversetemporal* yüzey alanları; *parahippocampal*, *insula*, *lingual* kalınlık ortalamaları özniteliklerinin sadece *AH* v *diğerleri* sınıflandırması sonucunda değerli öznitelikler olarak bulunduğu gözlemlenmiştir. Bu öznitelikler üç hastalığın sınıflandırılması dışında, *AH*'nin belirlenmesine has değerli özniteliklerdir şeklinde yorumlanmaktadır. Benzer şekilde 3-Sınıflı deneysel çalışmalar ve *FTD* v *diğerleri* ile yapılan 2-Sınıflı deneysel çalışmalar kıyaslandığında, beynin sağ *caudate*, sağ *ventralDC*, *5th ventricle* hacimleri ile birlikte sol beynin *paracentral*, *parstriangularis*, *precentral* yüzey alanları; *fusiform*, *superiorfrontal* gri hacimleri; ayrıca, sağ beynin *medialorbitofrontal*, *inferiorparietal*, *posteriorcingulate*, *transversetemporal* gri hacimleri; *paracentral*, *pericalcarine* yüzey alanları; *parstriangularis*, *precuneus*, *superiorparietal*, *superiortemporal* kalınlık ortalamaları özniteliklerinin sadece *FTD* v *diğerleri* sınıflandırması sonucunda değerli öznitelikler olarak bulunduğu gözlemlenmiştir. Bu öznitelikler üç hastalığın sınıflandırılması dışında, *FTD*'nin belirlenmesine has değerli özniteliklerdir şeklinde yorumlanmaktadır. 3-Sınıflı deneysel çalışmalar ve *VaD* v *diğerleri* ile yapılan 2-Sınıflı deneysel çalışmalar kıyaslandığında ise, beynin sol *lateral ventricle*, sağ *cerebellum cortex*, *pallidum* hacimleri ile birlikte sol beynin *inferiorparietal*, *isthmuscingulate*, *transversetemporal*, *parsopercularis*, *precuneus* gri hacimleri; *cuneus* yüzey alanı; ayrıca, sağ beynin *caudalanteriorcingulate* gri hacmi; *parahippocampal* ve *precentral* yüzey alanları özniteliklerinin sadece *VaD* v *diğerleri* sınıflandırması sonucunda değerli öznitelikler olarak bulunduğu gözlemlenmiştir. Bu öznitelikler üç hastalığın sınıflandırılması dışında, *VaD*'nin belirlenmesine has değerli özniteliklerdir şeklinde yorumlanmaktadır. Belirtilen tüm hastalık bazlı değerli öznitelikler kısaca, incelenen hastalığa klinik tanı koyma aşamasında farklı beyin bölümlerinin daha etkili olduğu şeklinde yorumlanmaktadır.

5. SONUÇLAR VE YORUMLAR

Nörogörüntüleme biliminde, farklı özellikler barındıran deneysel çalışmaların sınıflandırma sonuçlarının direkt olarak karşılaştırılması, veri setlerine has numune sayısı, içerdiği hastalıkların tipi ve/veya sayısı vb. gibi özelliklerden dolayı, doğru yorumlamaları beraberinde getirmeyebilir. Ek olarak, öznitelik elde etme süreci içerisinde kullanılan görüntü işleme teknikleri vb. gibi izlenen yöntem adımlarının da elde edilen bulguları etkileyebileceğinden daha önce yapılan çalışmalarda bahsedilmiştir. Ayrıca, numunelerin testi kısmında oluşturulan eğitim ve test setlerinin değişkenliği de sonucu etkileyebilmektedir [15, 16]. Dolayısı ile bu tür çalışmalarda veri seti çalışmaya özgü değişebileceğinden, doğruluk sonuçlarının farklı çalışmalara ait çıktılarla bire bir kıyaslanması yerine, izlenen yöntemlere dayanarak kendi içerisindeki tutarlığı daha fazla önem arz etmektedir. Ancak, konusundaki benzer diğer çalışmalarda ulaşılan sonuçların, genel anlamda hangi aralıklarda değiştiğine dair bir referans niteliği taşıyabilir.

Bu çalışmanın başlangıç aşamasında, elde edilen tüm beyin özniteliklerinin birlikte sınıflandırıcılara sokulması sonrasında başarısız sayılabilecek doğruluk sonuçları alınmıştır. Öznitelik seçim ihtiyacı üzerine GA tabanlı sarmalama öznitelik seçim algoritması uygulanmıştır. GA tabanlı sarmalama öznitelik seçim algoritmasının uygulanması ile üç demans tipine ait beyin ölçülendirmelerinin bulunduğu öznitelik grupları üzerinden hastalar, NB ve SVM sınıflandırıcıları ile birlikte başarılı bir şekilde sınıflarına ayrıştırılmıştır. Deneysel çalışmaların birinci aşamalarında gruplar içerisindeki en değerli öznitelikler bulunmuştur; sonrasında, deneysel çalışmaların ikinci aşamasında ise bulunan alt öznitelik setleri birleştirilmiş ve bu set üzerinden beynin tamamına yönelik nihai sonuçlar elde edilmiştir. Algoritmanın işleyişi sırasında hangi hastalığı tespit etme eğilimi olduğuna göre bulunan değerli öznitelikler sıralanmıştır. Tüm deneysel çalışmalar doğrultusunda, kullanılan veri seti üzerinden anlatılan yöntemler ile kıymetli öznitelikler kör arama ile başarılı bir şekilde bulunmuş ve listelenmiştir.

Bulgular literatür ile kıyaslandığında, 3-Sınıflı deneysel çalışmalar sonucunda elde edilen değerli özniteliklerin içerisinde *entorhinal*, *inferiortemporal*, *middletemporal*, *isthmuscingulate*, *choroid-plexus* beyin bölümlerinin daha önce yapılan çalışmalardaki bulgular ile belirgin bir şekilde çakıştığı gözlemlenmektedir [12, 13, 16, 18]. 2-Sınıflı deneysel çalışmalardan *AH* v *diğerleri* sınıflandırmasında elde edilen değerli

özniteliklerin literatürdeki AH hastalarının da veri setine dâhil olduğu çalışmalara bakıldığında ise *parahippocampal*, *superiortemporal* gibi özniteliklerle de karşılaşılmaktadır [12, 13, 16]. Ayrıca literatürde etkisiz öznitelikler şekilde örneklendirilen *brainstem*, *corpus callosum* gibi beyin bölümleri ile bu çalışma sonucunda değerli öznitelik çıktısı olarak karşılaşılmamıştır [12]. Deneysel çalışmaların geneline bakıldığında ise, bu çalışmaya has en değerli şekilde bulunan *bankssts*, *postcentral* ve *putamen* gibi özniteliklerin veri setine bağlı elde edilen sonuçlar olduğu yorumlanabilmektedir. Bu doğrultuda, çalışmaya has en değerli olabilecek özniteliklerin ve ikincil değerliler şekilde nitelendirilebilecek özniteliklerin veri seti büyütülerek daha detaylı bir şekilde araştırılması gerekmektedir.

Geleceğe yönelik planlar ise, ilk ve en önemli olarak veri seti büyüklüğünün ve güvenilirliğinin artırılması üzerine kuruludur. Veri seti boyutunun artırılması ile daha detaylı araştırmalar yapılabilecek ve daha kararlı alt öznitelik setlerinin bulunması sağlanabilecektir. Bu tür araştırmaların sonucuna göre öznitelikler daha kesin bir şekilde seviyelendirilebilecektir. Ayrıca, yüksek sayıda numuneye sahip, uluslararası anlamda standart olarak kullanılabilen veri setleri üzerinden izlenilecek yöntemlerle elde edilen bulgular, literatürde aynı veri setini kullanan farklı çalışmaların sonuçları ile direkt olarak kıyaslanabilecektir. Diğer bir yandan, daha tutarlı sonuçlar verebilecek GA parametreleri üzerinde de deneysel çalışmalar yapılarak elde edilen bulguların daha güvenilir hale getirilmesi planlanmaktadır. Buna ek olarak, farklı öznitelik seçimi ve sınıflandırma algoritmalarının da kullanılması ile birlikte çok daha etkili raporlar hazırlanacak ve literatüre katkı sağlanabilecektir.

KAYNAKÇA

- [1] Güngen, C., Ertan, T., Eker, E., Yaşar, R., ve Engin, F. (2002). Standardize mini mental test'in Türk toplumunda hafif demans tanısında geçerlik ve güvenilirliği, *Türk Psikiyatr. Derg.*, 13 (4), 273–281.
- [2] Kurban, N. K., Zencir, M., Kartal, A., ve Şahiner, T. (2010). Alzheimer hastalığı olan ve olmayan yaşlı bireylerin yaşam kalitesi ve depresyon düzeylerinin karşılaştırılması, *Yaşlı Sorunları Araştırma Derg.*, (1), 34–43.
- [3] İlkin, İ. (2001). Demans ve komorbid durumlar, *Psikiyatr. Dünyası*, 5, 49–54.
- [4] Akdemir, A., Cangöz, B., Örsel, S., ve Selekler, K. (2007). Hafif kognitif bozukluğu olan hastalarla Alzheimer tipi demans hastalarının örtük bellek performansı açısından karşılaştırılması, *Türk Psikiyatr. Derg.*, 18 (2), 118–128.
- [5] Özkay, Ü. D., Öztürk, Y., ve Can, Ö. D. (2011). Yaşlanan dünyanın hastalığı: Alzheimer hastalığı, *SDÜ Tıp Fakültesi Derg.*, 18 (1), 35–42.
- [6] Topçuoğlu, E. S. ve Selekler, K. (1998). Alzheimer hastalığı, *Turkish J. Geriatr.*, 1 (2), 63–67.
- [7] Bıyıklı, A. E. ve Şanlıer, N. (2014). Alzheimer hastalığı ve beslenme, *Genel Tıp Derg.*, 24 (2), 62–65.
- [8] Karaman, Y. (2006). Alzheimer hastalığı, vasküler demans ve hafif kognitif etkilenmede plazma homosistein, vitamin B12 ve folat seviyeleri, *J. Neurol. Sci.*, 23 (3), 175–184.
- [9] Erkol, G. ve Savrun, F. K. (2002). Frontotemporal demanslar, *Cerrahpaşa Tıp Dergisi*, 33 (4), 256–260.
- [10] Pırıldar, Ş., Bora, E., Evyapan, D., ve Özaşkinlı, S. (2001). Frontotemporal demans: Şizofreni benzeri psikoz tablosuyla giden bir olgu sunumu, *Klin. Psikofarmakol. Bülteni*, 11 (2), 116–120.
- [11] Herek, D. and Karabulut, N. (2010). Manyetik rezonans görüntüleme, *TTD Toraks Cerrahisi Bülteni*, 1 (3), 214–222.
- [12] Westman, E., Simmons, A., Muehlboeck, J. S., Mecocci, P., Vellas, B., Tsolaki, M., Kloszewska, I., Soinenen, H., Weiner, M. W., Lovestone, S., Spenger, C., and Wahlund, L. O. (2011). AddNeuroMed and ADNI: Similar patterns of Alzheimer's atrophy and automated MRI classification accuracy in Europe and North America, *Neuroimage*, 58 (3), 818–828.

- [13] Aguilar, C., Westman, E., Muehlboeck, J. S., Mecocci, P., Vellas, B., Tsolaki, M., Kloszewska, I., Soininen, H., Lovestone, S., Spenger, C., Simmons, A., and Wahlund, L. O. (2013). Different multivariate techniques for automated classification of MRI data in Alzheimer’s disease and mild cognitive impairment, *Psychiatry Res. - Neuroimaging*, 212 (2), 89–98.
- [14] Jack, C. R., Bernstein, M. A., Fox, N. C., Thompson, P., Alexander, G., Harvey, D., Borowski, B., Britson, P. J., Whitwell, J. L., Ward, C., Dale, A. M., Felmlee, J. P., Gunter, J. L., Hill, D. L. G., Killiany, R., Schuff, N., Fox-Bosetti, S., Lin, C., Studholme, C., DeCarli, C. S., Krueger, G., Ward, H. A., Metzger, G. J., Scott, K. T., Mallozzi, R., Blezek, D., Levy, J., Debbins, J. P., Fleisher, A. S., Albert, M., Green, R., Bartzokis, G., Glover, G., Mugler, J., and Weiner, M. W. (2008). The Alzheimer’s disease neuroimaging initiative (ADNI): MRI methods, *J. Magn. Reson. Imaging*, 27 (4), 685–691.
- [15] Cuingnet, R., Gerardin, E., Tessieras, J., Auzias, G., Lehéricy, S., Habert, M.-O., Chupin, M., Benali, H., and Colliot, O. (2011). Automatic classification of patients with Alzheimer’s disease from structural MRI: A comparison of ten methods using the ADNI database, *Neuroimage*, 56 (2), 766–781.
- [16] Escudero, J., Zajicek, J. P., and Ifeachor, E. (2011). Machine learning classification of MRI features of Alzheimer’s disease and mild cognitive impairment subjects to reduce the sample size in clinical trials., in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 7957–7960.
- [17] Jung, W. B., Lee, Y. M., Kim, Y. H., and Mun, C.-W. (2015). Automated classification to predict the progression of Alzheimer’s disease using whole-brain volumetry and DTI, *Psychiatry Investig.*, 12 (1), 92–102.
- [18] Zhou, Q., Goryawala, M., Cabrerizo, M., Wang, J., Barker, W., Loewenstein, D. a, Duara, R., and Adjouadi, M. (2014). An optimal decisional space for the classification of Alzheimer’s disease and mild cognitive impairment, *Biomed. Eng. IEEE Trans.*, 61 (8), 2245–2253.
- [19] Okyay, S., Adar, N., Özkan, K., Şaylısoy, S., Adapınar, D. B., and Adapınar, B. (2016). Classification of some dementia types due to feature selection with artificial neural networks, in *IEEE 24th SIU*.
- [20] Jain, A. and Zongker, D. (1997). Feature selection: Evaluation, application, and small sample performance, *IEEE Trans. Pattern Anal. Mach. Intell.*, 19 (2), 153–

158.

- [21] Fischl, B. (2012). FreeSurfer, *Neuroimage*, 62 (2), 774–781.
- [22] Reuter, M., Schmansky, N. J., Rosas, H. D., and Fischl, B. (2012). Within-subject template estimation for unbiased longitudinal image analysis, *Neuroimage*, 61 (4), 1402–1418.
- [23] Dale, A. M., Fischl, B., and Sereno, M. I. (1999). Cortical surface-based analysis: I. segmentation and surface reconstruction, *Neuroimage*, 9 (2), 179–194.
- [24] Fischl, B., Liu, A., and Dale, A. M. (2001). Automated manifold surgery: Constructing geometrically accurate and topologically correct models of the human cerebral cortex, *IEEE Trans. Med. Imaging*, 20 (1), 70–80.
- [25] Gronenschild, E. H. B. M., Habets, P., Jacobs, H. I. L., Mengelers, R., Rozendaal, N., van Os, J., and Marcelis, M. (2012). The effects of FreeSurfer version, workstation type, and Macintosh operating system version on anatomical volume and cortical thickness measurements, *PLoS One*, 7 (6), e38234.
- [26] Rish, I. (2001). An empirical study of the naïve Bayes classifier, in *IJCAI-2001 workshop on Empirical Methods in AI*, 3 (22), pp. 41–46.
- [27] Hearst, M. A., Dumais, S. T., Osuna, E., Platt, J., and Schölkopf, B. (1998). Support vector machines, *IEEE Intell. Syst. Their Appl.*, 13 (4), 18–28.
- [28] Burges, C. C. J. C. (1998). A tutorial on support vector machines for pattern recognition, *Data Min. Knowl. Discov.*, 2 (2), 121–167.
- [29] Smola, A. J., Schölkopf, B., and Müller, K. R. (1998). The connection between regularization operators and support vector kernels, *Neural Networks*, 11 (4), 637–649.
- [30] Sebban, M. and Nock, R. (2002). A hybrid filter/wrapper approach of feature selection using information theory, *Pattern Recognit.*, 35 (4), 835–846.
- [31] Pei, M., Goodman, E., Punch, W., and Ding, Y. (1995). Genetic algorithms for classification and feature extraction, in *Classification Society Conference*, pp. 1–28.
- [32] Konak, A., Coit, D. W., and Smith, A. E. (2006). Multi-objective optimization using genetic algorithms: A tutorial, *Reliab. Eng. Syst. Saf.*, 91 (9), 992–1007.
- [33] Sexton, R., Dorsey, R., and Johnson, J. (1999). Optimization of neural networks: A comparative analysis of the genetic algorithm and simulated annealing, *Eur. J. Oper. Res.*, 114 (3), 589–601.

- [34] Genetic Algorithm Options. <http://www.mathworks.com/help/gads/genetic-algorithm-options.html> (Erişim Tarihi: 19.04.2016).
- [35] Goldberg, D. and Holland, J. (1988). Genetic algorithms and machine learning, *Mach. Learn.*, 3 (5), 95–99.
- [36] Allen, F. and Karjalainen, R. (1999). Using genetic algorithms to find technical trading rules, *J. Financ. Econ.*, 51 (2), 245–271.
- [37] Goldberg, D. E. (1985). *Optimal initial population size for binary-coded genetic algorithms*. Clearinghouse for Genetic Algorithms, Department of Engineering Mechanics, University of Alabama.
- [38] Ludwig, O. and Nunes, U. (2010). Novel maximum-margin training algorithms for supervised neural networks, *IEEE Trans. Neural Networks*, 21 (6), 972–984.
- [39] Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection, in *International Joint Conference on Artificial Intelligence*, 14 (2), pp. 1137–1145.