

Emre DEMİR



İSKENDERUN TEKNİK

ÜNİVERSİTESİ

MÜHENDİSLİK VE FEN BİLİMLERİ ENSTİTÜSÜ

ELEKTRİK-ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI

**YÜKSEK
LİSANS
TEZİ**

**OSKÜLTASYON SESLERİNDEN
AMPİRİK DALGACIK DÖNÜŞÜMÜ
ANALİZİ KULLANILARAK
KRONİK OBSTRÜKTİF AKCİĞER
HASTALIĞININ TEŞHİSİ**

Emre DEMİR

**ELEKTRİK-ELEKTRONİK MÜHENDİSLİĞİ
ANABİLİM DALI**

OCAK 2020

OCAK 2020



**OSKÜLTASYON SESLERİNDEN AMPİRİK DALGACIK DÖNÜŞÜMÜ
ANALİZİ KULLANILARAK KRONİK OBSTRÜKTİF AKCİĞER
HASTALIĞININ TEŞHİSİ**

Emre DEMİR

**YÜKSEK LİSANS TEZİ
ELEKTRİK – ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI**

**İSKENDERUN TEKNİK ÜNİVERSİTESİ
MÜHENDİSLİK VE FEN BİLİMLERİ ENSTİTÜSÜ**

OCAK 2020

Emre DEMİR tarafından hazırlanan “OSKÜLTASYON SESLERİNDEN AMPİRİK DALGACIK DÖNÜŞÜMÜ ANALİZİ KULLANILARAK KRONİK OBSTRÜKTİF AKCİĞER HASTALIĞININ TEŞHİSİ” adlı tez çalışması aşağıdaki jüri tarafından OY BİRLİĞİ ile İskenderun Teknik Üniversitesi Elektrik-Elektronik Mühendisliği Anabilim Dalında YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

Danışman: Dr. Öğr. Üyesi Ahmet GÖKÇEN

Bilgisayar Mühendisliği Anabilim Dalı, İskenderun Teknik Üniversitesi

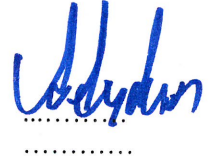
Bu tezin, kapsam ve kalite olarak Yüksek Lisans Tezi olduğunu onaylıyorum.



Başkan: Dr. Öğr. Üyesi Ahmet AYDIN

Biyomedikal Mühendisliği Anabilim Dalı, Çukurova Üniversitesi

Bu tezin, kapsam ve kalite olarak Yüksek Lisans Tezi olduğunu onaylıyorum.



Üye: Dr. Öğr. Üyesi İpek ABASIKELEŞ TURGUT

Bilgisayar Mühendisliği Anabilim Dalı, İskenderun Teknik Üniversitesi

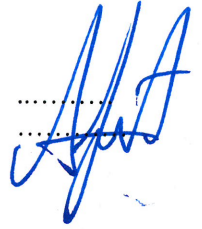
Bu tezin, kapsam ve kalite olarak Yüksek Lisans Tezi olduğunu onaylıyorum.



Üye: Dr. Öğr. Üyesi Ahmet GÖKÇEN

Bilgisayar Mühendisliği Anabilim Dalı, İskenderun Teknik Üniversitesi

Bu tezin, kapsam ve kalite olarak Yüksek Lisans Tezi olduğunu onaylıyorum.



Tez Savunma Tarihi: 10/01/2020

Jüri tarafından kabul edilen bu tezin Yüksek Lisans Tezi olması için gerekli şartları yerine getirdiğini onaylıyorum.

Prof. Dr. Tolga DEPCİ


Mühendislik ve Fen Bilimleri Enstitüsü Müdürü



ETİK BEYAN

İskenderun Teknik Üniversitesi Mühendislik ve Fen Bilimleri Enstitüsü Tez Yazım Kurallarına uygun olarak hazırladığım bu tez çalışmada;

- Yükseköğretim Kuruluna gönderilen kopya ile tarafından Mühendislik ve Fen Bilimleri Enstitüsü'ne verilen basılı ve/veya elektronik kopyaların birebir aynı olduğunu,
 - Tez içinde sunduğum verileri, bilgileri ve dokümanları akademik ve etik kurallar çerçevesinde elde ettiğimi,
 - Tüm bilgi, belge, değerlendirme ve sonuçları bilimsel etik ve ahlak kurallarına uygun olarak sunduğumu,
 - Tez çalışmada yararlandığım eserlerin tümüne uygun atıfta bulunarak kaynak gösterdiğimi,
 - Kullanılan verilerde herhangi bir değişiklik yapmadığımı,
 - Bu tezde sunduğum çalışmanın özgün olduğunu,
- bildirir, aksi bir durumda aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi beyan ederim.



Emre DEMİR

10/01/2020

OSKÜLTASYON SESLERİNDEN AMPİRİK DALGACIK DÖNÜŞÜMÜ ANALİZİ
KULLANILARAK KRONİK OBSTRÜKTİF AKCİĞER HASTALIĞININ TEŞHİSİ
(Yüksek Lisans Tezi)

Emre DEMİR

İSKENDERUN TEKNİK ÜNİVERSİTESİ
MÜHENDİSLİK VE FEN BİLİMLERİ ENSTİTÜSÜ

Ocak 2020

ÖZET

Gelişen teknoloji ile birlikte bilim dünyasında büyük gelişmeler görülmektedir. Bu gelişmelerden klinik tanı ve tedavi uygulamaları da büyük ölçüde etkilenmiştir. Sağlık alanında yapılan çalışmalar sonucu geliştirilen sistemler gerek hastalıkların tanısının konmasında, gerekse hastalıkların tedavisinde kullanılmaktadır. Bu çalışmalarda temel amaç hızlı, güvenilir ve maliyeti az sistemler geliştirerek hata payını azaltmak ve hekimlere yardım etmektir.

Bu tezde, oskültasyon yöntemi ile elde edilmiş veriler kullanılarak Kronik Obstrüktif Akciğer Hastalığı (KOAH) rahatsızlığının hızlı ve etkili bir şekilde sınıflandırılması için yeni bir yöntem geliştirilmiştir. Literatürde oskültasyon tabanlı analizler ile solunum hastalıkları teşhisi üzerine pek çok çalışma vardır. Ancak bu tezi oluşturan çalışmanın diğerlerinden farkı Ampirik Dalgacık Dönüşümü (ADD) tabanlı veriler üzerinde analizler yapılmış olmasıdır. ADD, verilen bir sinyali farklı modlara ayırmak için kullanılan bir yöntemdir. Uyarlanabilir bir dalgacık filtre bankasının tasarımına dayanır. Bu çalışmada kullanılan veri tabanı üzerinde ADD tekniği kullanılarak daha önce hiç çalışma yapılmamıştır. Veriseti olarak yapay zekâ temelli klinik tanı ve tedavi uygulamalarında kullanılacak, 12 kanallı akciğer seslerinden ve 4 kanallı kalp seslerinden oluşan RespiratoryDatabase@TR veri tabanı kullanılmıştır. RespiratoryDatabase@TR veri tabanı içerisinde kalp ve akciğer sesleri dışında göğüs filmi, solunum fonksiyon testi ölçümleri ve SGRQ-C anketinin cevapları da bulunmaktadır. Oskültasyon sesleri üzerinde gerekli ön işlemler yapıldıktan sonra ADD yöntemi uygulanarak sinyallerin farklı modları elde edilmiştir. Dönüşüm sonrasında elde edilen veriler üzerinde istatistiksel öznitelikler hesaplanmıştır. Son olarak geliştirilen sistemin başarısı 9 farklı sınıflandırma modelleri üzerinde uygulanıp, performansları değerlendirilmiştir.

Anahtar Kelimeler : KOAH, ampirik dalgacık dönüşümü, RespiratoryDatabase@TR,
örüntü tanıma, biyomedikal sinyal işleme
Sayfa Adedi : 52
Danışman : Dr. Öğr. Üyesi Ahmet GÖKÇEN

DIAGNOSIS OF CHRONIC OBSTRUCTIVE PULMONARY DISEASE USING
EMPIRICAL WAVELET TRANSFORM ANALYSIS FROM AUSCULTATION SOUNDS
(M. Sc. Thesis)

Emre DEMİR

ISKENDERUN TECHNICAL UNIVERSITY
ENGINEERING AND SCIENCE INSTITUTE

January 2020

ABSTRACT

Along with the developing technology, great developments are seen in the scientific world. Clinical advances in clinical diagnosis and treatment have also been greatly influenced by these developments. The systems developed as a result of the studies in the field of health are used both in the diagnosis of the diseases and in the treatment of the diseases. The main purpose of these studies is to reduce the margin of error and to help physicians by developing fast, reliable and cost-effective systems.

In this thesis, a new method has been developed for the classification of Chronic Obstructive Pulmonary Disease (COPD) disease by using auscultation method. There are many studies in the literature on auscultation-based analysis and diagnosis of respiratory diseases. However, the difference of this study from the others is that Empirical Wavelet Transform (EWT) based data are analyzed. EWT is a method used to divide a given signal into different modes. It is based on the design of an adaptive wavelet filter bank. The database used in this study has never been studied before using the EWT technique. The RespiratoryDatabase@TR database, which consists of 12-channel lung sounds and 4-channel heart sounds, can be used in clinical diagnosis and treatment applications based on artificial intelligence. The RespiratoryDatabase@TR database includes heart and lung sounds as well as chest X-ray, pulmonary function test measurements, and answers to the SGRQ-C questionnaire. After the necessary pre-process were made on auscultation sounds, different modes of the signals were obtained by applying EWT method. Statistical characteristics were calculated on the data obtained after the transformation. Finally, the success of the system was applied on 9 different classification models and their performance was evaluated.

Key Words : COPD, empirical wavelet transform, RespiratoryDatabase@TR, pattern recognition, biomedical signal processing
Page Number : 52
Supervisor : Asst. Prof. Dr. Ahmet GÖKÇEN

TEŞEKKÜR

Tez çalışmamın her aşamasında ilgi ve desteğini esirgemeyen değerli tez danışmanım Dr. Öğr. Üyesi Ahmet GÖKÇEN'e teşekkürlerimi sunarım.

Tez çalışmamda, Türkiye Bilimsel ve Teknolojik Araştırma Kurumu (TÜBİTAK-116E190) desteği ile oluşturdukları veri tabanını paylaşarak hastalarla ilgili verilere ulaşmamda destek sağlayan ve değerli bilgileri ile tecrübelerini esirgemeyen değerli hocalarım Doç. Dr. Yakup KUTLU ve Dr. Öğr. Üyesi Gökhan ALTAN'a teşekkür ederim.

Her zaman yanımda olan, hiçbir desteği esirgemeyen ve bu günlere gelmemde büyük emekleri olan aileme ve kardeşim Berke DEMİR'e teşekkürü bir borç bilir, sonsuz sevgilerimi sunarım.

İÇİNDEKİLER

	Sayfa
ÖZET	iv
ABSTRACT.....	v
TEŞEKKÜR.....	vi
İÇİNDEKİLER	vii
ÇİZELGELERİN LİSTESİ.....	ix
ŞEKİLLERİN LİSTESİ	x
SİMGELER VE KISALTMALAR.....	xi
1. GİRİŞ	1
2. ÖNCEKİ ÇALIŞMALAR	6
3. MATERYAL VE YÖNTEM.....	10
3.1. Kronik Obstrüktif Akciğer Hastalığı (KOAH).....	11
3.2. Veri Tabanı ve Özellikleri.....	13
3.3. Ampirik Dalgacık Dönüşümü (Empirical Wavelet Transform - EWT).....	16
3.4. Öznitelik Çıkarma	20
3.5. Sınıflandırma.....	22
3.5.1. En yakın k komşu (k Nearest Neighbors) Sınıflandırıcısı	23
3.5.2. Naïve Bayes	24
3.5.3. BayesNet	25
3.5.4. Çok katmanlı algılayıcı (Multilayer Perceptron)	26
3.5.5. Destek Vektör Makineleri (Support Vector Machine).....	27
3.5.6. AdaBoost.....	28
3.5.7. Random SubSpace	28
3.5.8. Random Forest	28

	Sayfa
3.5.9. J48.....	29
3.6. Performans Ölçümleri ve İstatistiksel Analiz.....	29
4. ARAŞTIRMA BULGULARI VE SONUÇLARI	31
4.1. En yakın k komşu Sınıflandırıcısı Sonuçları.....	32
4.2. Naïve Bayes Sınıflandırıcısı Sonuçları	33
4.3. BayesNet Sınıflandırıcısı Sonuçları	34
4.4. Çok Katmanlı Algılayıcı Sonuçları	35
4.5. Destek Vektör Makineleri Sonuçları.....	36
4.6. AdaBoost Algoritması Sonuçları	37
4.7. Random Subspace Algoritması Sonuçları.....	38
4.8. Random Forest Algoritması Sonuçları.....	39
4.9. J48 Algoritması Sonuçları.....	40
4.10. Kanalların Genel Başarım Üzerindeki Etkileri.....	41
5. SONUÇ VE ÖNERİLER.....	42
KAYNAKLAR	46
ÖZGEÇMİŞ	50

ÇİZELGELERİN LİSTESİ

Çizelge	Sayfa
Çizelge 3.1. KOAH evreleri ve belirtileri	12
Çizelge 3.2. Sınıflandırma doğruluğu tablo örneği.....	30
Çizelge 4.1. k-NN sınıflandırıcısının değerlendirme sonuçları	32
Çizelge 4.2. Naïve Bayes sınıflandırıcısının değerlendirme sonuçları	33
Çizelge 4.3. BayesNet sınıflandırıcısının değerlendirme sonuçları.....	34
Çizelge 4.4. Çok katmanlı algılayıcı modelinin değerlendirme sonuçları.....	35
Çizelge 4.5. DVM sınıflandırıcısının değerlendirme sonuçları	36
Çizelge 4.6. AdaBoost algoritmasının değerlendirme sonuçları	37
Çizelge 4.7. Random Subspace algoritmasının değerlendirme sonuçları.....	38
Çizelge 4.8. Random Forest algoritmasının değerlendirme sonuçları.....	39
Çizelge 4.9. J48 algoritmasının değerlendirme sonuçları	40
Çizelge 4.10. Kanalların genel başarımlar üzerindeki etkisi.....	41
Çizelge 5.1. Sınıflandırma sonuçları.....	42

ŞEKİLLERİN LİSTESİ

Şekil	Sayfa
Şekil 1.1. Littmann 3200 Digital Stethoscope	2
Şekil 3.1. Örüntü tanıma sistemlerinde temel işlemler	10
Şekil 3.2. Sağlıklı ve KOAH hastalarının solunum yolları ve alveolleri.....	11
Şekil 3.3. RespiratoryDatabase@TR Modeli.....	14
Şekil 3.4. Göğüs ve sırt için kullanılan oskültasyon bölgeleri.....	15
Şekil 3.5. Sağlıklı akciğer ses sinyali ve ADD yöntemi ile çıkarılan modları.....	18
Şekil 3.6. Hasta akciğer ses sinyali ve ADD yöntemi ile çıkarılan modları	19
Şekil 3.7. k-NN sınıflandırma örneği.....	24
Şekil 3.8. Çok katmanlı algılayıcı modeli yapısı	27
Şekil 4.1. Kanalların k-NN sınıflandırıcısı üzerindeki başarımlar grafiği	32
Şekil 4.2. Kanalların Naïve Bayes sınıflandırıcısı üzerindeki başarımlar grafiği.....	33
Şekil 4.3. Kanalların BayesNet sınıflandırıcısı üzerindeki başarımlar grafiği.....	34
Şekil 4.4. Kanalların Çok katmanlı algılayıcılar üzerindeki başarımlar grafiği	35
Şekil 4.5. Kanalların Destek Vektör Makineleri üzerindeki başarımlar grafiği	36
Şekil 4.6. Kanalların AdaBoost sınıflandırıcısı üzerindeki başarımlar grafiği	37
Şekil 4.7. Kanalların Random Subspace sınıflandırıcısı üzerindeki başarımlar grafiği	38
Şekil 4.8. Kanalların Random Forest sınıflandırıcısı üzerindeki başarımlar grafiği.....	39
Şekil 4.9. Kanalların J48 üzerindeki başarımlar grafiği.....	40
Şekil 5.1. Sınıflandırıcı başarımları karşılaştırma grafiği	43
Şekil 5.2. Kanalların performans etkileri.....	44

SİMGELER VE KISALTMALAR

Bu çalışmada kullanılmış simgeler ve kısaltmalar, açıklamaları ile birlikte aşağıda sunulmuştur.

Kısaltmalar

Açıklamalar

ADD	Ampirik Dalgacık Dönüşümü
ARNN	Attractor Recurrent Neural Network
ÇKA	Çok Katmanlı Algılayıcılar
DVM	Destek Vektör Makineleri
EEG	Elektroensefalografi
EEMD	Ensemble Empirical Mode Decomposition
HFD	Hızlı Fourier Dönüşümü
IMF	Intrinsic Mode Function
k-NN	En Yakın k Komşu
KOAH	Kronik Obstrüktif Akciğer Hastalığı
OED	Otomatik Eksternal Defibrilatör
TSA	Tekil Spektrum Analizi
YSA	Yapay Sinir Ağları

1. GİRİŞ

Sağlık ve hastalık kavramları insan hayatında önemli bir yer tutmaktadır. Hastalıkların tanı ve tedavisi ile ilgili çalışmalar çok eskiye dayanmaktadır. Antik çağlardan beri hekimler karşılaştıkları hastalıkların tanı ve tedavisi için çeşitli yöntemler kullanmış ve geliştirmiştir. Solunum hastalıkları ise bu hastalıklar içerisinde önemli bir yere sahiptir. Bugün solunum ve kalp rahatsızlıkları için en yaygın kullanılan teşhis yöntemlerinden olan oskültasyon yöntemi 19. yüzyıla dayanmaktadır. Oskültasyon, fiziki muayenenin stetoskop isimli bir alet yardımıyla göğüs duvarından, vücudun ürettiği sesleri dinlemek için yapılan işlemidir. Akciğer oskültasyonu solunum hastalıklarının teşhisinde ve şiddetinin belirlenmesi aşamasında önemli bir yer almaktadır. Vücut içerisinde oluşan ses dalgaları ve titreşimleri analiz etmek için kullanılan stetoskop, 1816 yılında René-Théophile-Hyacinthe Laennec tarafından icat edilmiştir. Bu sayede sesleri doğrudan kulakla dinleme yerine stetoskop ile dinleme imkânı doğmuştur. İlk zamanlarda tahtadan yapılan küçük bir boru şeklinde olan bu alet zamanla geliştirilmiş ve farklı çeşitleri üretilmiştir. İngilizcesi “stethoscope” şeklinde olan “stetoskop” kelimesi, Yunanca “göğüs” anlamındaki “stetos” ve “bakmak” anlamındaki “skopein” kelimelerinin birleştirilmesi ile türetilmiştir. İnsanlık tarihinin en önemli buluşları arasında yer alan ve hekimlerin vazgeçilmez bir muayene aracı haline gelen bu alet yardımıyla akciğer ve kalp tarafından oluşturulan seslerin hekimler tarafından dinlenip değerlendirilmesi sayesinde hastalığa ait tanı koyulmuştur ve koyulmaya devam etmektedir. Bu tezde kullanılan RespiratoryDatabase@TR veri tabanı üzerindeki veriler Littmann 3200 Digital Stethoscope kullanılarak oluşturulmuştur. Littmann 3200 Digital Stethoscope Şekil 1.1’de gösterilmiştir.



Şekil 1.1. Littmann 3200 Digital Stethoscope (Altan, 2018)

Günümüzde teknoloji hızla gelişirken bunun sayesinde bilim dünyasında da büyük gelişmeler görülmektedir. Bir bütün olarak gelişen bilim ve teknoloji ile birlikte insanların hayatlarını kolaylaştırmaya yönelik çalışmalar hızla artmaktadır. Bu gelişmeler klinik tanı ve tedavi uygulamalarını da büyük ölçüde etkilemiştir. Bu gelişmeler ile birlikte bilimsel araştırmalarda daha hızlı ve doğru sonuç alınmaktadır. Ayrıca yapılan çalışmalar sonucunda geliştirilen klinik tanı ve tedavi uygulamaları insan yaşamını daha kaliteli hale getirmeye yardım etmektedir.

Günümüzde ileri teknoloji ekipmanlar ve biyomedikal sinyal işleme teknikleri ile bilgisayar destekli tanı sistemleri için çeşitli yöntemler geliştirilmeye devam edilmektedir. Bu çalışmalar gerek hastalıkların tanısının konmasında, gerekse hastalıkların tedavisinde başarılı bir şekilde kullanılmaktadır. Bu çalışmalarda temel amaç hızlı, güvenilir ve maliyeti az sistemler geliştirerek hata payını azaltmak ve hekime teşhis konusunda yardımcı olabilmektir. Ayrıca hekimlerin ve diğer sağlık personelinin iş yükünde kısmen azaltılması amaçlanmaktadır.

Biyomedikal sinyallerin izlenmesi ve değerlendirilmesi, insan sağlık durumu ve hastalıkları ile ilgili bilgi edinmek için önemli bir yere sahiptir. Biyomedikal sinyallerin analizi kolay bir işlem değildir. Çünkü sinyal içerisinde bulunan doğru bilgiler gürültü ve parazitlerin oluşmasıyla bozulabilmekte ve bu durum analizi zorlaştırmaktadır. Ayrıca biyomedikal sinyalin bir uzman tarafından yorumlanması; kişinin deneyimine, yorgunluğun veya çevresel etmenlerin neden olduğu sınırlamalara göre değişebilmektedir.

Bilgisayar destekli tanı sistemleri gözden kaçan bulguların uzman tarafından görülmesi ve tekrar analiz edilmesine yardımcı olabilmektedir. Bu nedenle bilgisayar destekli tanı sistemleri hekimlere hata payını azaltmak ve uzun süreli kayıtların analizinde zaman konusunda yardımcı olup tanı koymada önemli bir katkı sağlamaktadır.

Biyomedikal sinyallere örnek olarak, kalbin elektriksel aktivitesi için Elektrokardiyografi (EKG), beyin dalgaları aktivitesi için Elektroensefalografi (EEG), sinir ve çizgili kasların elektrik potansiyeli için Elektromiyografi (EMG) başta olmak üzere Elektrokülografi (EOG), Manyetoensefalografi (MEG), Bilgisayarlı Tomografi (BT/CT), Ultrasonografi (USG) ve Manyetik Rezonans Görüntüleme (MR) gösterilebilir.

Bu tezde ihtiyaç ve eğilimler göz önünde bulundurularak oskültasyon işlemi ile elde edilen sinyaller kullanılarak solunum rahatsızlıklarından en önemlilerinden biri olan Kronik Obstrüktif Akciğer Hastalığı (KOAH) teşhisi üzerine odaklanılmıştır.

KOAH; genellikle zararlı gaz ve partiküllere, özellikle de sigara dumanına uzun süre maruz kalınmasıyla oluşan ve ilerleyen, akciğerdeki hava keseciklerinin tıkanmasına yol açan kronik bir hastalıktır. Kalp-damar hastalıkları için risk oluşturabilen, hava yollarını tıkayan ve solunum ile ilgili en yaygın hastalıklardan olan KOAH en ölümcül solunum bozukluklarından biridir (Erdoğan ve diğerleri, 2010; Demir ve Gökçen, 2019).

Yaklaşık 300 milyon insanı etkileyen KOAH, ekonomik olarak toplumlara ağır yük getiren bir sağlık sorunu olup her yıl 3 milyon ölüm ile sonuçlanan ve 44 milyar dolarlık sağlık maliyeti olan bir hastalıktır (WHO, 2018; Günaydın ve diğerleri, 2019; Demir ve Gökçen, 2019).

Dünya Sağlık Örgütü raporuna göre, KOAH ve benzeri solunum rahatsızlıkları en ölümcül hastalıklar listesinde olup teşhis ve tedavi için gerekli olan bütçeler oldukça yüksektir. Bundan dolayı erken tanı ve tedavi bu hastalıklar için önemlidir. Yapılan çalışmalar ile maliyet ve zamanı düşürerek hekimlere yardımcı olmak amaçlanmaktadır (WHO, 2018; Demir ve Gökçen, 2019). Bu çalışmada, dünya genelinde en büyük mortalite ve morbidite nedenleri haline gelen hastalıklardan biri olan bu hastalığın maliyet ve teşhis süresini düşürerek hekimlere yardımcı olmak, hastalığın teşhisine yardım etmek ve erken teşhis sayesinde tedavi süresini kısaltmak amaçlanmıştır.

KOAH tanı ve tedavi harcamaları oldukça yüksek düzeydedir. Erken tanı ile önlenebilir bir hastalık olan KOAH, ilerleyen aşamalarda KOAH alevlenmeleri ile devam eder. KOAH ilerleme süreci her hastada farklılık gösterebilir. Bazıları dengeli bir yaşam sürdürebilirken bazı hastalar hastalık evresi ilerledikçe daha sık alevlenme yaşar. Alevlenme yaşam kalitesini olumsuz etkiler. KOAH alevlenmeleri, KOAH tanı ve tedavi sürecinin maliyeti en yüksek olan bölümünü kapsamaktadır. Bu nedenle hastalığın önlenmesi, tanısı ve tedavisi hastalar için önemlidir (Sin, Hollander, DeMarco, McManus ve Ng, 2015; X. Liu ve diğerleri, 2019).

Ampirik Dalgacık Dönüşümü (ADD) uyarlanabilir bir dalgacık filtre bankasının tasarımına dayanan ve sinyalin farklı modlarını çıkarmaya yarayan bir sinyal ayrıştırma yöntemidir (Gilles, 2013; Demir, Gökçen, Altan ve Kutlu, 2019).

Bu çalışmada, yapay zekâ temelli klinik tanı ve tedavi uygulamalarında kullanılabilecek bir sistemin geliştirilmesi amaçlanmıştır. Çalışmada öncelikle 12 kanallı akciğer seslerinden ve 4 kanallı kalp seslerinden oluşan RespiratoryDatabase@TR veri tabanından alınan sinyaller üzerinde gerekli ön işlemler yapıldıktan sonra, kullanılan veri tabanı üzerinde daha önce hiç kullanılmayan bir yöntem olan “ADD analizi” yapılmıştır. ADD analizi sonrası elde edilen modlar üzerinde istatistiksel öznitelikler hesaplanmış ve öznitelik vektörü oluşturulmuştur. Son olarak oluşturulan öznitelik vektörleri üzerinde En Yakın k Komşu (k-NN), Naïve Bayes, BayesNet, Çok Katmanlı Yapay Sinir Ağları (ÇKA), Destek Vektör Makineleri (DVM), AdaBoost, Random SubSpace, Random Forest ve J48 olmak üzere 9 farklı sınıflandırma algoritma yöntemleri uygulanarak sınıflandırma işlemi yapılmıştır. Bu işlemler neticesinde bu sınıflandırıcıların performansları karşılaştırılmıştır. Sonuç olarak anormal solunum seslerinden faydalanılarak KOAH teşhisi başarılı bir şekilde yapılmıştır.

Bu tez beş bölümden oluşmaktadır. Birinci bölümde çalışmanın amacı, önemi, solunum hastalıkları arasından neden KOAH'ın tercih edildiği ve konu ile ilgili tanım ve açıklamalar yer almaktadır.

İkinci bölümde yapılan çalışmayla ilgili ve bu çalışmaya paralel olarak yapılan önceki çalışmaların bazıları incelenmiştir. Çalışmalarda ele alınan konunun ne olduğu, kullanılan metotlar ve elde edilen başarı oranları hakkında bilgi verilmektedir.

Üçüncü bölümde KOAH, çalışmada kullanılan veri tabanı, ADD, kullanılan öznitelik çıkarma yöntemi, sınıflandırma yöntemlerine dair tanım ve açıklamalar, performans ölçümleri ve istatistiksel analiz hakkında gerekli temel bilgiler verilmektedir.

Dördüncü bölümde çalışma için gerekli tüm işlemler yapıldıktan sonra önerilen modelin farklı sınıflandırma algoritmaları üzerindeki performans ölçümleri verilmiştir.

Son olarak beşinci bölümde ise önerilen yöntemler ve sınıflandırma algoritmaları sonuçları değerlendirilmiş, gelecekte yapılacak araştırmalar için öneriler sunulmuştur.



2. ÖNCEKİ ÇALIŞMALAR

Günümüzde bilgisayar destekli tanı sistemleri için çeşitli yöntemler geliştirilmeye devam edilmektedir. Bunlarda temel amaç hızlı, güvenilir ve maliyeti az sistemler geliştirmektir. Bu çalışmada; çeşitli sinyal işleme metotları, ayırıştırma teknikleri ve farklı sınıflandırıcılar kullanılarak yapay zekâ temelli klinik tanı ve tedavi uygulamalarında kullanılacak bir sistem geliştirilerek Kronik Obstrüktif Akciğer Hastalığı (KOAH) teşhisine yardımcı olmak amaçlanmaktadır.

Bu bölümde ADD yöntemi, solunum seslerinin analizi, KOAH ve çeşitli biyomedikal uygulamalar ile ilgili yapılan çalışmalar gözden geçirilmiştir.

Literatür incelendiğinde, solunum sesleri üzerinde çeşitli sinyal işleme teknikleri ve yapay zekâ temelli analizler yaparak solunum hastalıklarının teşhis ve tedavi yöntemlerini bilgisayar destekli hale getirmeye çalışan birçok çalışma bulunmakta ve geliştirilmeye devam edilmektedir. Solunum seslerinin analizi ile ilgili yapılan çeşitli çalışmalar aşağıda sunulmuştur.

Islam, Bandyopadhyaya, P. Bhattacharyya ve Saha (2018) yaptıkları çalışmada sırt bölgesinden elde edilen akciğer sesleri üzerinden astım teşhisi ile ilgili bir çalışma yapmışlardır. Oskültasyon sonucunda elde ettikleri 4 kanallı akciğer sesleri üzerinde wheeze sesleri olmadan analiz yapmışlardır. İstatistiksel özellik hesaplamaları sonrasında normal ve astımlı sesleri sınıflandırmak için Yapay Sinir Ağları (YSA) ve DVM yöntemlerini kullanarak sınıflandırma işlemi yapmışlardır.

Sengupta, Sahidullah ve Saha (2016) yaptıkları çalışmada akciğer seslerinin kısa süreli spektral özelliklerini, akciğer bağlantılı hastalıkların tanımlanması için analiz etmişlerdir. Normal, wheeze ve crackle olmak üzere üç tip sesi tespit edebilmek için sınıflandırma aşamasında YSA yöntemini kullanmışlardır.

Becker, Scheffer, Blanckenberg ve Diacon (2013) yaptıkları çalışmada bulaşıcı ve ciddi bir hastalık olan tüberküloz hastalığı üzerinde çalışmışlardır. Çalışmalarında tüberküloz ile ilişkili zaman ve frekans alanlarında bulunan en önemli sinyal özelliklerini, Statistical

Overlap Factor kullanarak belirlemişlerdir. Ardından çıkarılan özellikleri YSA kullanarak sınıflandırmışlardır. Tüberküloz sınıflandırması sonuçlarında hassasiyet, özgüllük ve genel başarımlar için sırasıyla % 80, % 67 ve % 73 oranlarını elde etmişlerdir.

Mondal, Banerjee ve Somkuwar (2017) yaptıkları çalışmada solunum hastalıklarının teşhisinde kullanılan oskültasyon yönteminin doğru yapılabilmesi amacıyla oskültasyon seslerinin içerisinde bulunan kalp seslerinin oluşturduğu parazitin azaltılması ve bu sayede akciğer seslerine ait sinyallerin kalitesinin artırılması için bir yöntem önermişlerdir. Çalışmalarında Ampirik Kip Ayrışımı yöntemini ve Hızlı Fourier Dönüşümü (HFD) temelli tahmin algoritmasını kullanmışlardır.

Khodabakhshi ve Moradi (2017) çok kanallı solunum sinyallerinin otomatik olarak sınıflandırılması üzerine çalışmışlardır. Sinyaller üzerinde Attractor Recurrent Neural Network (ARNN) temelli bir modelleme yapmışlar ve bulanık fonksiyonları temel alan bir ARNN topolojisi geliştirmişlerdir. Çalışma sonucunda ise geliştirilen yöntem kullanılarak % 91 oranında bir sınıflandırma doğruluğu elde ettiklerini belirtmişlerdir.

Hong, Essid, Ser ve Foo (2018) yaptıkları çalışmada akciğer ödemini tespit etmek için kullanılan bir sınıflandırma sistemi tasarlamışlardır. Negatif Olmayan Matris Ayrışımı temelli bir özellik öğrenme yöntemi kullanmışlar ve ardından Lojistik Regresyon ile sınıflandırmışlardır. Farklı denemeler sonucunda sınıflandırma sonuçlarını % 86' ya kadar iyileştirebildiklerini belirtmişlerdir.

Khadar, Tabatabaey-Mashadi ve Daliri (2019) yaptıkları çalışmada Otomatik Eksternal Defibrilatör (OED) cihazı ile ilgili bir yöntem üzerine çalışmışlardır. Çalışmalarında QRS kompleksi içermeyen kardiyak sinyallerin tepe noktalarını ve aritmilerini tanımlayan bir yöntem önermişlerdir. Sınıflandırma aşamasında ise DVM kullanmışlardır. Sonuç olarak OED sistemlerinin tepki süresinin genellikle 10 saniyeden fazla olduğunu ve önerilen yöntemin Raspberry Pi version B üzerinde uygulandığında sinyali algılamak için gereken sürenin bu süreden daha başarılı olduğunu belirtmişlerdir.

Maruf, Azhar, Khawaja ve Akram (2015) yaptıkları çalışmada crackles seslerinin otomatik tespiti için bir sistem geliştirmişlerdir. Gauss Karışımları Modeli kullanmışlar ve sistemin değerlendirilmesinde % 100 özgüllük, % 92,85 hassasiyet, % 100 Pozitif Prediktif Değer

ve %97,56 genel başarı oranlarını elde etmişlerdir.

ADD yöntemi ile ilgili literatür incelendiğinde, bu yöntem kullanılarak mekanik arıza tespiti (Pan, J. Chen, Zi, Y. Li ve He, 2016; J. Chen, Pan, Z. Li, Zi ve X. Chen, 2016; Kedadouche, Thomas ve Tahan, 2016), sismik zaman-frekans analizi (W. Liu, Cao ve Y. Chen, 2016), güç kalitesi endekslerinin tahmini (Thirumala, Umarikar ve Jain, 2014) ve ADD temelli yeni yöntemlerin geliştirilmesi (Amezquita-Sanchez ve Adeli, 2015) gibi farklı alanlarda yapılmış çalışmalar mevcuttur.

ADD yöntemi kullanılarak biyomedikal sinyaller üzerinde yapılan çalışmalar çok kısıtlı olup bunların bazıları aşağıda sunulmuştur.

Babu, Ramkumar ve Manikandan (2019) yaptıkları çalışmada Fonokardiyogram sinyallerinden akciğer seslerinin otomatik tespiti için ADD tabanlı bir yöntem önermişlerdir. Önerdikleri yöntemi Tekil Spektrum Analizi (TSA) ve Ensemble Empirical Mode Decomposition (EEMD) yöntemleri ile karşılaştırmışlardır. Karşılaştırma sonucunda ADD, TSA ve EEMD için sırasıyla % 99,22, % 83,54, % 93,40 doğruluk oranlarını elde etmişlerdir.

Singh ve Sunkaria (2016) yaptıkları çalışmada ADD yöntemine dayanan Arteriyel Kan Basıncı sinyallerinde sistolik tepe noktasının tespiti için yeni bir yaklaşım sunmuşlardır. Çalışmalarında ADD mod sayısını isteğe bağlı olarak beşe ayarlamışlardır. Çalışmalarını çeşitli veri tabanlarından alınan kayıtlar üzerinde test etmişlerdir.

Oung, Muthusamy, Basah, Lee ve Vijejan (2018) parkinson hastalığı üzerine çalışmışlardır. Çalışmalarında ADD ve Ampirik Dalgacık Paket Dönüşümü yöntemlerini uygulamışlardır. Ardından farklı dönüşümler uygulanıp entropi temelli özellikleri çıkarmışlardır. Sınıflandırma aşamasında ise k-NN, Olasılıksal Sinir Ağı ve Aşırı Öğrenme Makinesi algoritmalarını kullanarak sınıflandırma işlemini gerçekleştirmişlerdir.

A. Bhattacharyya ve Pachori (2017) yaptıkları çalışmada epileptik nöbet tespiti için Elektroensefalografi (EEG) sinyallerinin çok değişkenli yapısını ADD yöntemi kullanarak analiz etmişlerdir. Ardından önerilen yöntemin performansını altı farklı sınıflandırma algoritması kullanarak değerlendirmişlerdir.

Maheshwari, Pachori ve Acharya (2016) yaptıkları çalışmada glokom tanısı üzerine odaklanmışlardır. Dijital fundus görüntülerinin ayrıştırılmasında ADD yöntemini, sınıflandırma aşamasında ise En Küçük Kareler Destek Vektör Makinesi algoritmasını kullanmışlardır. Yaptıkları çalışma sonucunda sırasıyla üç ve on kat çapraz doğrulama yöntemi kullanarak % 98,33 ve % 96,67 oranında sınıflandırma doğruluğu elde etmişlerdir.

Maharaja ve Shaby (2017) ise yine dijital fundus görüntüleri kullanarak glokom hastalığının otomatik tanısı üzerine çalışmışlar ve ADD yöntemini kullanmışlardır. Ardından Gri Seviyeli Eş-oluşum Matrisleri kullanarak özellikler çıkartıp YSA ile normal ve glokom sınıflandırması yapmışlardır. Performans değerlendirmesinde % 96 oranında sınıflandırma doğruluğu elde etmişlerdir.

Huang, S. Zhang ve Y. Zhang (2017) beyin-bilgisayar arayüzü üzerine çalışmışlardır. EEG tabanlı sinyaller üzerinde ADD yöntemi uygulayarak duygu tahmini üzerine bir yaklaşım önermişlerdir. Jambholkar, Gurve ve Sharma (2015) ise yaptıkları çalışmada SPECT görüntüleri üzerinde tıbbi görüntü işleme analizi ile beyin tümörü tespiti üzerine odaklanmışlardır. Çalışmalarında ADD yöntemini uygulamışlardır.

3. MATERYAL VE YÖNTEM

Örüntü tanıma sistemleri genel olarak veri toplama, ön işleme, öznitelik çıkarma ve sınıflandırma aşamalarından oluşmaktadır (Şekil 3.1).



Şekil 3.1. Örüntü tanıma sistemlerinde temel işlemler

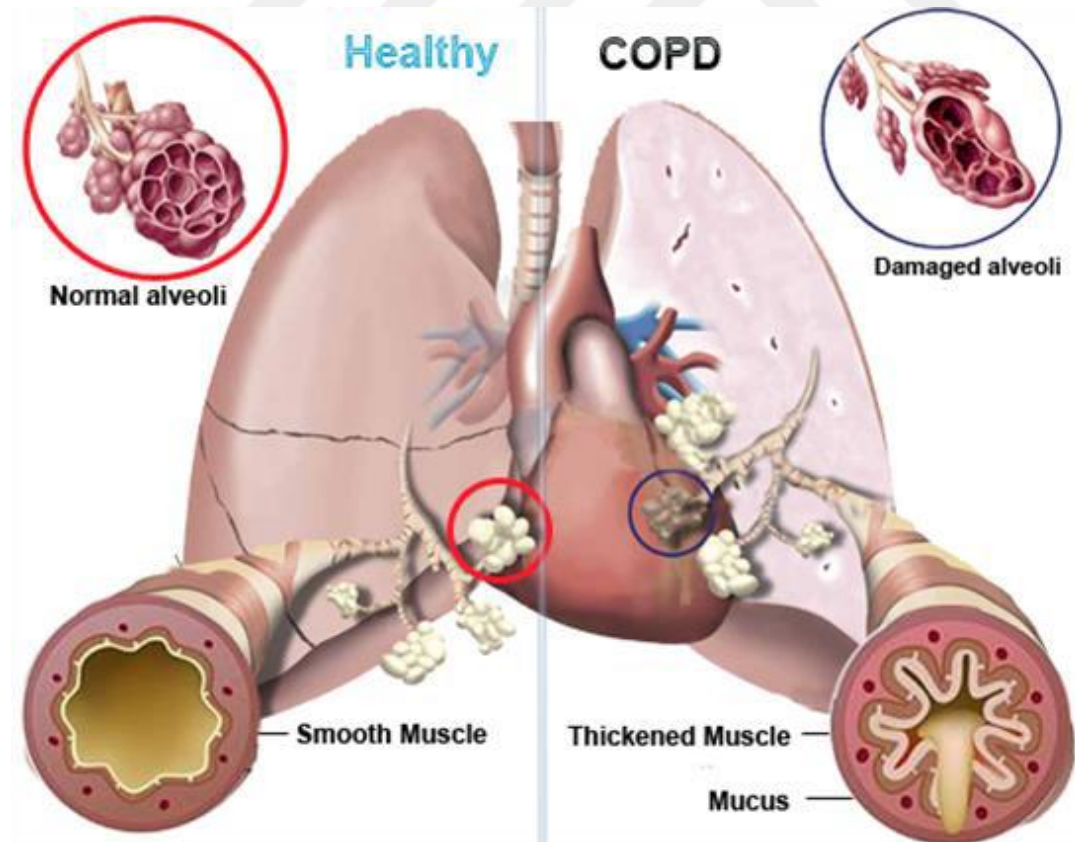
Örüntü tanıma sistemlerinde temel işlemler incelendiğinde; veri toplama aşamasında, veri analizi için gerekli sinyaller sensörler aracılığıyla ölçülür. Sistemin başarılı sonuç verebilmesi için, sinyalin doğru bir şekilde analiz edilmesi gerekir. Bunun için ön işleme aşamasında çeşitli sinyal işleme teknikleri kullanılarak biyomedikal sinyaller için büyük bir sorun olan gürültülerin giderilmesi gibi işlemler yapılır ve veri bir sonraki aşamaya uygun olacak şekilde hazırlanır. Öznitelik çıkarma işleminde ise bir veya daha fazla öznitelik belirlenir ve öznitelik vektörü oluşturulur. Öznitelik vektörleri sinyaller hakkında önemli bilgileri içeren öznitelikler kümesinden oluşur. Öznitelik çıkarma aşaması örüntü tanıma sistemlerinin en önemli adımıdır. Çünkü sistemin özelliklerini temsil eden veriler bu aşamada oluşturulmaktadır. Sınıflandırma aşaması ise örüntü tanıma sisteminin son aşaması olup çeşitli yöntemler kullanılarak karar mekanizmasının geliştirildiği adımdır.

Bu bölümde Kronik Obstrüktif Akciğer Hastalığı (KOAH), çalışmada kullanılan veri tabanı, ADD, sinyal işleme teknikleri, kullanılan öznelik çıkarma yöntemi, sınıflandırma yöntemlerine dair tanım ve açıklamalar, performans ölçümleri ve istatistiksel analiz ile çalışmada kullanılan yöntemler hakkında gerekli temel bilgiler verilmektedir.

3.1. Kronik Obstrüktif Akciğer Hastalığı (KOAH)

Kronik Obstrüktif Akciğer Hastalığı (KOAH); zararlı gaz ve partiküllere, özellikle de sigara dumanına uzun süre maruz kalınmasıyla oluşan ve ilerleyen, akciğerdeki hava keseciklerinin tıkanmasına yol açan kronik bir hastalıktır. Kalıcı hava akımı kısıtlanması ile kendini gösteren ve genellikle ilerleyici bir solunum hastalığı olan KOAH tanısı büyük öneme sahip olan en ölümcül solunum bozukluklarından (Erdoğan ve diğerleri, 2010).

Sağlıklı ve KOAH hastalarının solunum yolları ve alveolleri Şekil 3.2'de gösterilmektedir.



Şekil 3.2. Sağlıklı ve KOAH hastalarının solunum yolları ve alveolleri (Altan,2018)

Dünya Sağlık Örgütü raporuna göre, KOAH ve benzeri solunum rahatsızlıkları en ölümcül hastalıklar listesinde olup teşhis ve tedavi için gerekli bütçeler oldukça yüksektir (WHO, 2018, Demir ve Gökçen, 2019). Dünya genelinde sık görülen ve en büyük mortalite ve morbidite nedenleri haline gelen solunum hastalıklarından olması ve ekonomik olarak toplumlara ağır yük getiren küresel bir problem olması nedeniyle tanısı büyük öneme sahiptir.

Sigara bağımlılığı, genetik faktörler, çevresel ve mesleki maruziyet, iç ve dış ortam hava kirliliği, akciğer gelişimine etkili faktörler, hava yolu hiperreaktivitesi ve astım KOAH gelişiminde aktif rol oynadığı düşünülen risk faktörleri arasında gösterilmektedir. Ancak sigara bağımlılığı KOAH gelişiminde en önemli risk faktörü olarak gösterilmektedir (Erdinç ve diğerleri, 2010).

Risk faktörlerine maruz kalan kişilerde en önemli semptom olan nefes darlığı, öksürük, balgam çıkarma, hışıltılı solunum ve göğüste sıkıntı hissi başta olmak üzere KOAH'tan sorumlu temel risk faktörleri ve diğer parametreler hastalığın belirtileri arasında yer almaktadır (Kocabaş ve diğerleri, 2014).

KOAH şiddetinin belirlenmesi, spirometrik olarak Bronkodilatatör sonrası FEV1 değerinin beklenen değerin yüzdesine göre yapılır (Kocabaş ve diğerleri, 2014). KOAH evreleri ve belirtileri Çizelge 3.1'de gösterilmiştir.

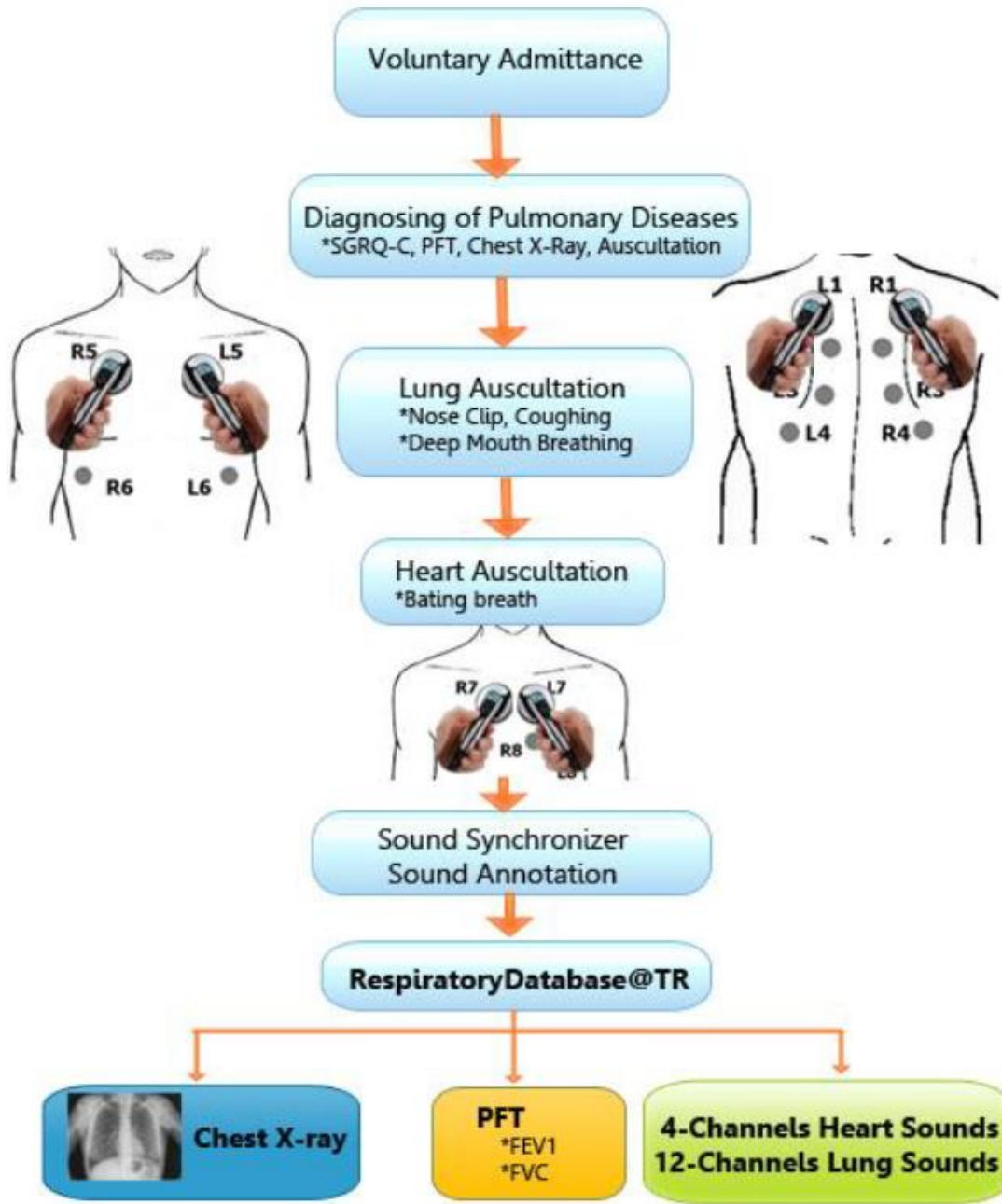
Çizelge 3.1. KOAH evreleri ve belirtileri (Altan, Kutlu, Garbi, Pekmezci ve Nural, 2017)

KOAH Evresi	Semptomlar	
<i>KOAH0 : Risk Altında</i>	<ul style="list-style-type: none"> Solunum Fonksiyon Testi Normal Kronik semptomlar var (kötü veya inatçı öksürük, balgam) 	
<i>KOAH1 : Hafif Seviye</i>	<ul style="list-style-type: none"> FEV1 \geq % 80 Kronik semptomların durumu 	FEV1/FVC < % 70
<i>KOAH2 : Orta Seviye</i>	<ul style="list-style-type: none"> % 50 < FEV1 < % 80 Kronik semptomların durumu 	
<i>KOAH3 : Ağır Seviye</i>	<ul style="list-style-type: none"> % 30 < FEV1 < % 50 Kronik semptomların durumu 	
<i>KOAH4 : Çok Ağır Seviye</i>	<ul style="list-style-type: none"> FEV1 < % 30 ya da FEV1 < % 50 Kronik Solunum Yetmezliği 	

3.2. Veri Tabanı ve Özellikleri

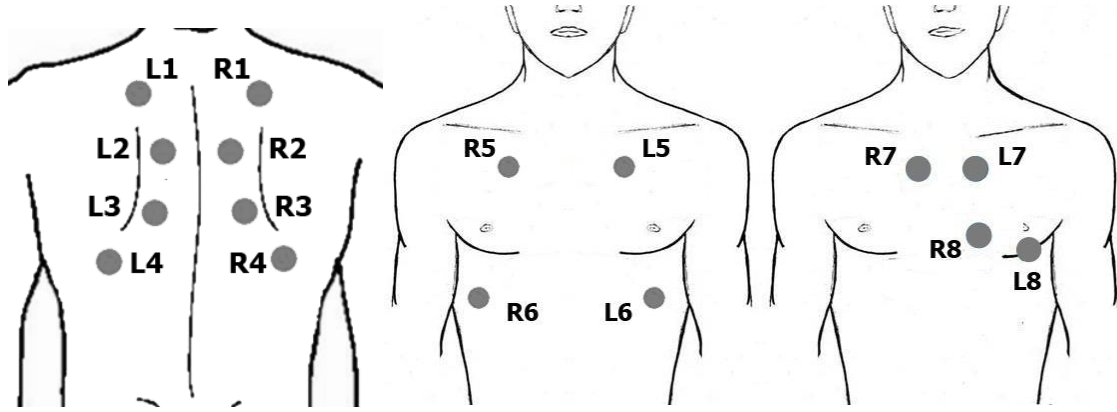
Solunum ve kalp rahatsızlıklarının teşhisi ile ilgili yapılan yeni çalışmalara ve teknolojinin hızla gelişmesine rağmen en yaygın kullanılan teşhis yöntemi oskültasyon olarak bilinmektedir. Bu çalışmada da oskültasyon yöntemi ile elde edilen ses sinyalleri üzerinde çeşitli sinyal işleme yöntemleri ile analizler yapıp farklı yapay zekâ teknikleri ile sınıflandırma yapılmıştır. Çalışma için gerekli olan veriler 12 kanallı akciğer seslerinden ve 4 kanallı kalp seslerinden oluşan RespiratoryDatabase@TR veri tabanından alınmıştır.

Dijital sinyal işleme teknikleri ile solunum analizi için kullanılan bir multimedya solunum veri tabanı olan RespiratoryDatabase@TR içerisinde her hastaya özgü sayısallaştırılmış oskültasyon sesleri, göğüs filmi, St. George Solunum Anket (SGRQ-C) cevapları ve solunum fonksiyon testi ölçümleri bulunmaktadır (Altan ve diğerleri, 2017). RespiratoryDatabase@TR veri tabanı oluşturulurken veri elde etme işleminin şeması Şekil 3.3'te gösterilmektedir.



Şekil 3.3. RespiratoryDatabase@TR Modeli (Altan, 2018)

RespiratoryDatabase@TR veri tabanı oluşturulurken oskültasyon sesleri 16 farklı kanaldan iki dijital stetoskop yardımıyla sağ ve sol odaklardan eşzamanlı olarak kaydedilmiştir. RespiratoryDatabase@TR veri tabanı ile KOAH, astım, kronik bronşit olan kişiler, sağlıklı kişilerden hem akciğer hem de kalp sesleri ile ayrılmıştır. Göğüs ve sırt için kullanılan oskültasyon bölgeleri Şekil 3.4'te gösterilmiştir.



Şekil 3.4. Göğüs ve sırt için kullanılan oskültasyon bölgeleri (Altan ve diğerleri, 2017)

Veri tabanı, Antakya Devlet Hastanesinde göğüs hastalıkları uzmanları tarafından etiketlenmiştir. Ayrıca Mustafa Kemal Üniversitesi tarafından etik kurul onayı vardır (Altan, Kutlu ve Allahverdi, 2019).

RespiratoryDatabase@TR veri tabanı oluşturulurken kullanılan Littmann 3200 Digital Stethoscope, üç frekans aralığı için bir filtreleme teknolojisi sağlar. Frekans aralığı filtreleri Bell (20-200 Hz), Diaphragm (100-500 Hz) ve Extended (20-1000 Hz) olarak sınıflandırılabilir (Altan, 2018). Oskültasyon sırasında gürültüye neden olan ortam seslerini filtreleme ve klasik stetoskoplarla duyulması zor olan oskültasyon seslerini yükseltme yeteneği başta olmak üzere birçok yeteneğe sahip olan Littmann 3200 Digital Stethoscope sinyallerin ön işlenmesi aşamasında büyük avantaj sağlamaktadır. RespiratoryDatabase@TR veri tabanındaki akciğer sesleri, DC offseti kaldırmak için 7.5 Hz'lik (1. Dereceden Butterworth filtre) yüksek geçiren filtreden, yüksek frekanslı gürültüyü azaltmak için ise 10 kHz'lik (8. Dereceden Butterworth filtre) alçak geçiren filtreden geçirilmiştir (Altan, Kutlu, Pekmezci ve Nural, 2018).

Segmentasyon için ortak bir başlangıç noktası oluşturmak gerekir. Öksürük akciğer seslerinde bir tepe noktası oluşturur. Segmentasyonun başlangıç noktası öksürme tepe noktasından sonraki ilk inhalasyon olarak belirlenip tam solunum döngülerini içerecek şekilde sonlandırılmasına dikkat edilmiştir. Biyomedikal analizlerde sinyalin uzunluğu da önemli bir faktördür. Bu çalışmada uygulamanın daha hızlı çalışabilmesi ve ileriki çalışmalarda donanım üzerinde kolay uygulanabilmesi için sinyal uzunluğu 5 saniye olacak şekilde ayarlanmıştır.

3.3. Ampirik Dalgacık Dönüşümü (Empirical Wavelet Transform - EWT)

Bir sinyali analiz etmek için kullanılan uyarlanabilir yöntemler sinyalde yer alan bilgilerin tespiti aşamasında büyük öneme sahiptir. Uyarlanabilir yöntemlerin amacı, sinyalde yer alan bilgilere dayanarak sinyali temsil eden modları tespit edip uygun bir temel oluşturmaktır. Bunlardan biri de Ampirik Kip Ayrışımı'dır (AKA) (Empirical Mode Decomposition - EMD). AKA gibi bazı yöntemler, bir sinyali içerdiği bilgilere göre ayrıştırmayı önermektedir. Temel fikir, sinyali birden fazla IMF'ye ayırmak ve ardından dönüşümleri kullanarak zarf spektrumunu oluşturmak için uygun IMF'yi seçmektir. AKA yöntemi uygulanarak birçok uygulama için başarılı sonuçlar elde edilmiştir. Ancak matematiksel teori eksikliği nedeniyle, AKA yöntemi hala birçok soruna neden olmaktadır. Çünkü algoritmik bir yaklaşımdır ve doğrusal olmaması nedeniyle modellenmesi zordur.

Dalgacık dönüşümü ve AKA'nın avantajları birleştirilerek geliştirilen Ampirik Dalgacık Dönüşümü (ADD), verilen bir sinyali farklı modlara ayırmak için kullanılan bir sinyal ayrıştırma yöntemidir. Fourier bakış açısı göz önüne alındığında bu yapı bant geçiren filtre yapı kümesine eşdeğerdir. Temelde uyarlanabilir bir dalgacık filtre bankasının tasarımına dayanır (Gilles, 2013; Demir ve Gökçen, 2019). Bu işlemleri verilen sinyalin bilgi içeriğine bağlı olarak uyarlanabilir bir hale getiren adaptif dalgacıklar oluşturarak yapar. ADD, önce sinyalde bulunan frekans bileşenlerini tahmin eder, ardından sınırları hesaplar ve hesaplanan sınırlara bağlı olarak sinyalin farklı modlarını çıkarır. Yöntem sinyal ayrıştırma ve analiz sürecinde kullanılmaktadır. Bu tezde oskültasyon yöntemi ile elde edilen akciğer sesleri üzerinde gerekli ön işlemler yapıldıktan sonra ADD yöntemi uygulanarak sinyalin farklı modları elde edilmiştir.

ADD yöntemi uygulanırken öncelikle sinyaller alınır ve HFD değeri hesaplanır. Belirtilen eşikler ile Fourier spektrumunda gerekli matematiksel hesaplamalar yapılarak sinyalin frekansları bulunur ve sınırlar tespit edilir. Tespit edilen sınırlar kullanılarak N dalgacık filtresi bankası oluşturulur ve gerekli formüller kullanılarak matematiksel hesaplamalar yapıldıktan sonra frekans bantları elde edilir ve filtreleme işlemi yapılır. ADD yöntemi uygulanırken kullanılan ölçeklendirme fonksiyonu Eş. 3.1'de ve frekans alanındaki ampirik dalgacıklar ise Eş. 3.2'de verilen denklemler ile hesaplanmaktadır. Yaklaşık katsayıyı elde etmek için sinyalin ve ampirik ölçekleme fonksiyonunun iç çarpımı Eş. 3.3'te verildiği gibi gerçekleştirilir. Ayrıntılı katsayılar ise Eş. 3.4'te verilen denklem ile

elde edilir.

$$\phi_n(\omega) = \begin{cases} 1 & \text{eğer } |\omega| \leq (1 - \gamma)\omega_n \\ \cos \left[\frac{\pi}{2} \beta \left(\frac{1}{2\gamma\omega_n} (|\omega| - (1 - \gamma)\omega_n) \right) \right] & \text{eğer } (1 - \gamma)\omega_n \leq |\omega| \leq (1 + \gamma)\omega_n \\ 0 & \text{aksi takdirde} \end{cases} \quad (3.1)$$

$$\psi_n(\omega) = \begin{cases} 1 & \text{eğer } (1 + \gamma)\omega_n \leq |\omega| \leq (1 - \gamma)\omega_{n+1} \\ e^{-i\frac{\omega}{2}} \cos \left[\frac{\pi}{2} \beta \left(\frac{1}{2\gamma\omega_{n+1}} (|\omega| - (1 - \gamma)\omega_{n+1}) \right) \right] & \text{eğer } (1 + \gamma)\omega_n \leq |\omega| \leq (1 - \gamma)\omega_{n+1} \\ e^{-i\frac{\omega}{2}} \sin \left[\frac{\pi}{2} \beta \left(\frac{1}{2\gamma\omega_n} (|\omega| - (1 - \gamma)\omega_n) \right) \right] & \text{eğer } (1 - \gamma)\omega_n \leq |\omega| \leq (1 + \gamma)\omega_n \\ 0 & \text{aksi takdirde} \end{cases} \quad (3.2)$$

$$W_x(1, t) = \langle x, \phi_1 \rangle = \int x(\tau) \phi(\tau - t) d\tau \quad (3.3)$$

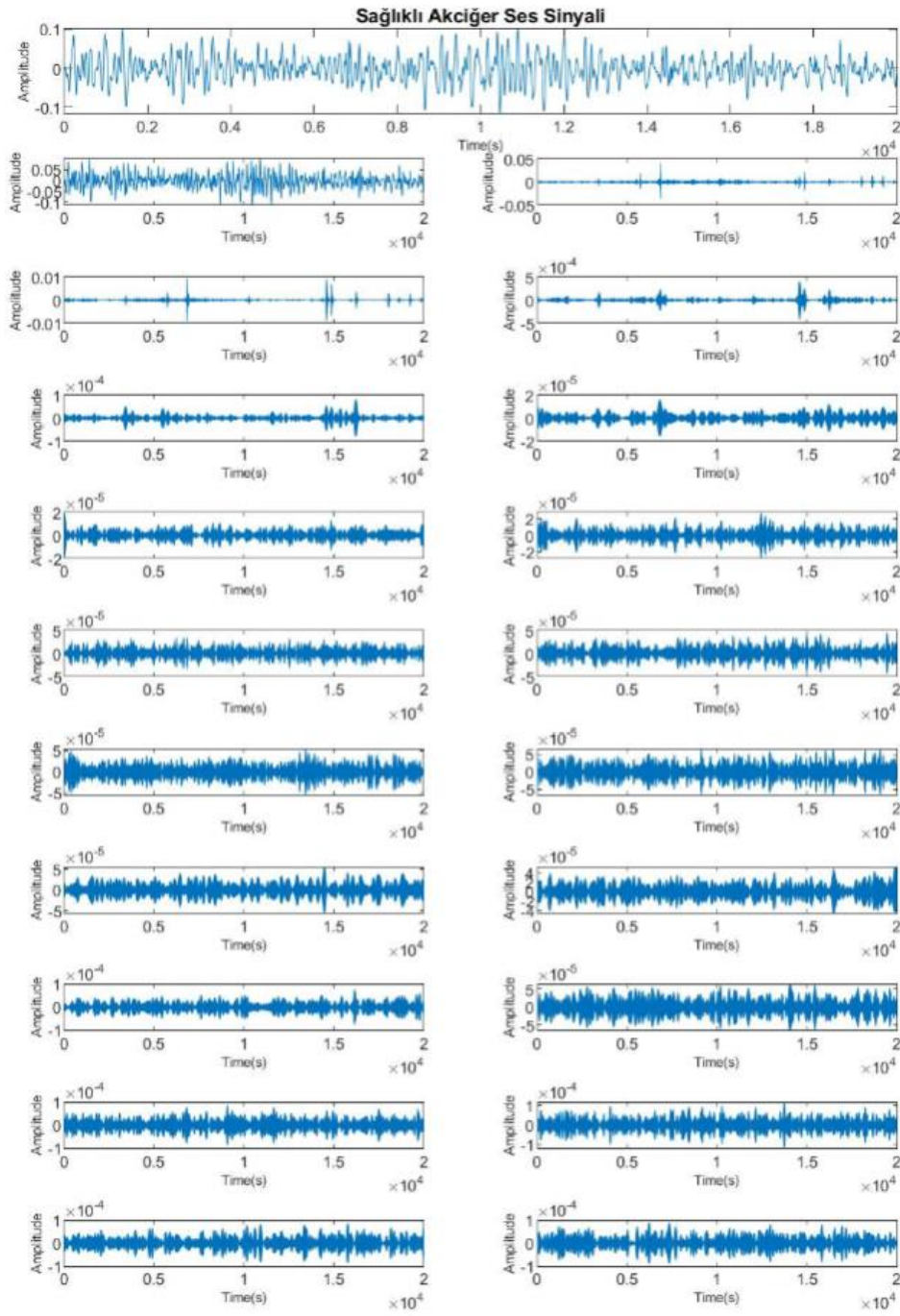
$$= \text{IFFT}(X(\omega) \phi_1(\omega))$$

$$W_x(i, t) = \langle x, \psi_i \rangle = \int x(\tau) \psi_i(\tau - t) d\tau \quad (3.4)$$

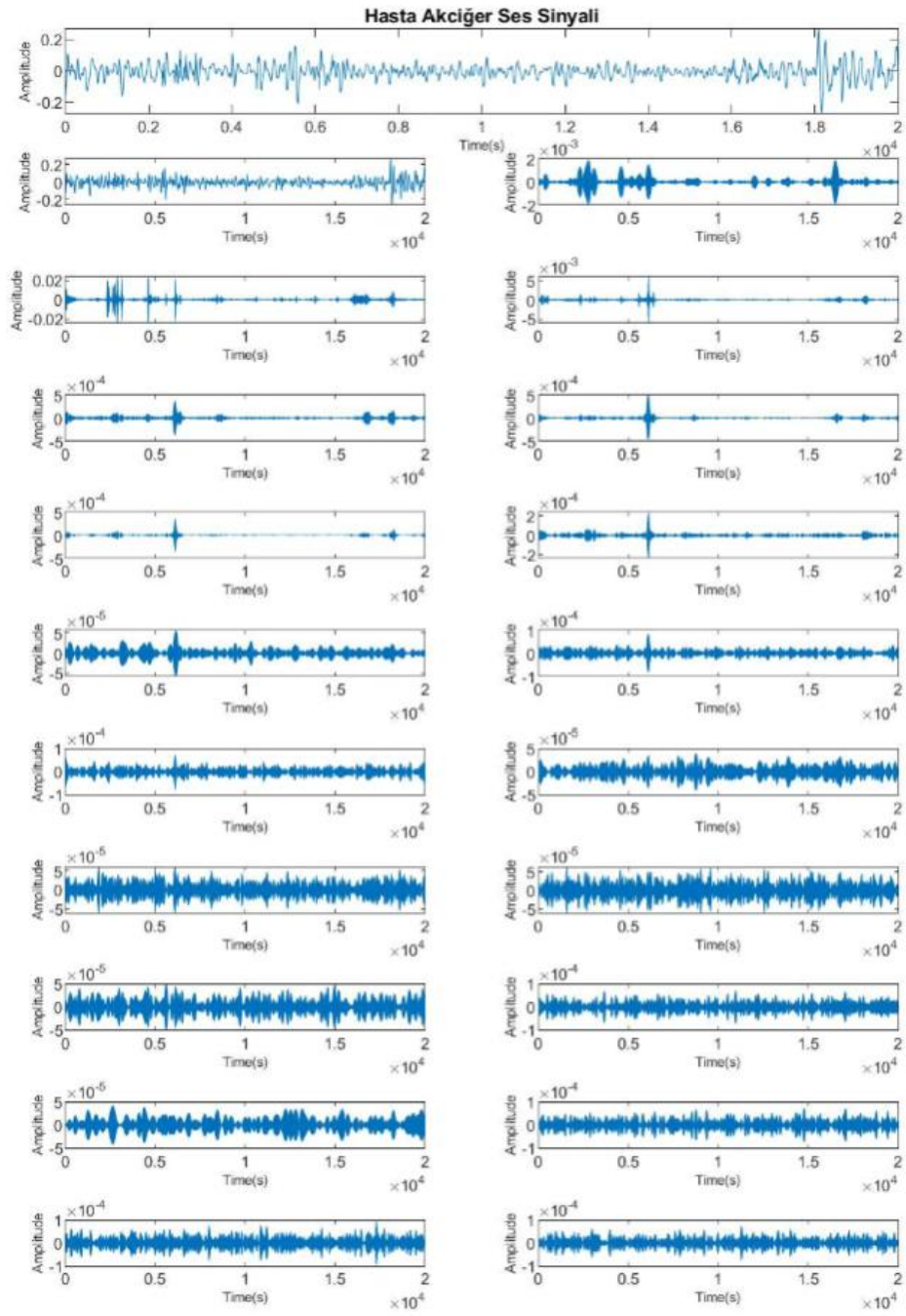
$$= \text{IFFT}(X(\omega) \psi_i(\omega))$$

ADD yöntemi, AKA'ya benzer ancak sağlam bir matematiksel altyapıya sahip olan, hızlı ve tamamen uyarlanabilir farklı bir sinyal analizi yaklaşımıdır. Ölçeklendirme fonksiyonu ve dalgacıklar kendilerini analiz edilen sinyalin içerdiği bilgilere göre uyarlar ve sinyale ilişkin önceden bilgi gerekmez. Kullanılan RespiratoryDatabase@TR veri tabanı üzerinde KOAH tespiti amacıyla çok kanallı solunum ses verilerinin analizinde ADD ilk kez uygulanmaktadır.

Yapılan çalışmada, gerekli ön işlemler yapıldıktan sonra elde edilen veriler üzerinde ADD uygulanmıştır. ADD yöntemi uygulandıktan sonra birçok frekans bileşeni elde edilir. Ancak bu bileşenlerin sayısı sinyale göre değişebilir. İşlemin RespiratoryDatabase@TR veri tabanında bulunan tüm sinyaller üzerinde uygulanabilmesi için mod sayısı; frekans bileşen sayısı en az olan sinyallere göre düzenlenmiştir. ADD yönteminin sinyal üzerinde uygulanması sonucunda elde edilen verilere örnek göstermek amacıyla, oskültasyon yöntemi ile sağlıklı bir kişinin tek kanaldan alınan gerçek akciğer ses sinyali ve sinyalden çıkarılan modların bazıları Şekil 3.5'te gösterilmiştir. KOAH hastası olan bir kişinin yine tek kanaldan alınan gerçek akciğer ses sinyali ve sinyalden çıkarılan modların bazıları Şekil 3.6'da gösterilmiştir.



Şekil 3.5. Sağlıklı akciğer ses sinyali ve ADD yöntemi ile çıkarılan modları



Şekil 3.6. Hasta akciğer ses sinyali ve ADD yöntemi ile çıkarılan modları

3.4. Öznitelik Çıkarma

Öznitelik çıkarma işlemi, makine öğrenmesi, örüntü tanıma, görüntü işleme, ses ve sinyal işleme uygulamaları gibi birçok uygulamada sıkça kullanılan önemli bir aşamadır. Öznitelik çıkarma işlemi, karmaşık ve yüksek boyutlu bir veri içerisindeki anlamlı bilgileri oluşturan özelliklerin çeşitli yöntemler uygulanarak çıkartılması sonucunda istenen amaca yönelik bilgilerin oluşturulması aşamasıdır. En önemli aşamalardan olan bu işlem sayesinde karmaşık olan bir veri, boyutu ve karmaşıklığı azaltılarak daha basit hale getirilir ve yeni oluşturulan temiz veri üzerinde işlemler yapılır. Öznitelik çıkarma işlemi doğru yapılırsa, sınıflandırma algoritmalarının işi kolaylaşacaktır. Bunun sonucunda da işlem sonucunun başarısı ve performansı olumlu yönde etkilenecektir (Demir ve diğerleri, 2019).

Literatürde veri analizi ve sınıflandırma için birçok öznitelik çıkarma yöntemleri bulunmaktadır. Çalışılan probleme ve uygulanan yöntemlerin kullanılış amacına göre tercih değişmektedir. Bu tezde, öncelikle oskültasyon yöntemi ile elde edilen 12 kanallı akciğer seslerinden ve 4 kanallı kalp seslerinden oluşan RespiratoryDatabase@TR veri tabanından alınan akciğer sesleri üzerinde gerekli hesaplamalar ve ön işlemler yapılmıştır. Bu işlemlerden sonra veriler üzerinde istatistiksel öznitelik çıkarma yöntemi uygulanmış ve gerekli hesaplamalar yapıldıktan sonra öznitelikler çıkarılmıştır. Bu aşamada; minimum, maksimum, toplam, medyan, varyans, standart sapma, ortalama, basıklık ve ortalama mutlak sapma olmak üzere 9 farklı istatistiksel öznitelik kullanılmıştır. Son olarak elde edilen özniteliklerle oluşturulan öznitelik vektörleri 9 farklı sınıflandırma algoritmaları ile test edilmiştir. Bu öznitelikler ile ilgili matematiksel formüller, tanım ve açıklamaları ile birlikte aşağıda sunulmuştur.

Minimum, çıkarılan alt bantlardaki değerlerin en küçük olanıdır.

Maksimum, çıkarılan alt bantlardaki değerlerin en büyük olanıdır.

Toplam, çıkarılan alt bantlardaki değerlerin tümünün toplamıdır.

Medyan, merkezsiz konum ölçüsüdür. Tek ve çift sayılar için formüller Eş. 3.5 ve Eş. 3.6'da gösterilmiştir.

$$\text{Medyan} = X_{\left(\frac{n-1}{2}\right)} \quad (3.5)$$

$$\text{Medyan} = \frac{1}{2} \left(x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)} \right) \quad (3.6)$$

Varyans, dağılım ölçüsüdür ve çıkarılan alt bantlardaki değerlerin ortalamaya göre dağılımını gösterir. Formülü ise Eş. 3.7’de gösterilmiştir.

$$\text{Varyans} = \frac{\sum |x - \bar{x}|^2}{(n - 1)} \quad (3.7)$$

Standart sapma, çıkarılan alt bantlardaki değerlerin ortalamadan ne kadar uzaklaştığını ölçmek için kullanılır. Standart sapma varyansın karekökü alınarak hesaplanır. Formülü ise Eş. 3.8’de gösterilmiştir.

$$\text{Standart Sapma} = \sqrt{\frac{\sum |x - \bar{x}|^2}{(n - 1)}} \quad (3.8)$$

Ortalama, çıkarılan alt bantlardaki değerlerin toplamının, bu değerlerin sayısına bölünerek bulunan sayıdır. Formülü ise Eş. 3.9’da gösterilmiştir.

$$\text{Ortalama} = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad (3.9)$$

Basıklık, çıkarılan alt bantlardaki değerlerin basık veya sivri olması ile ilgili bir kavramdır. Formülü ise Eş. 3.10’da gösterilmiştir.

$$\text{Basıklık} = \mathbb{E} \left[\left(\frac{(X - \mu)}{\sigma} \right)^4 \right] = \frac{\mu_4}{\sigma_4} = \frac{\mathbb{E}[(X - \mu)^4]}{(\mathbb{E}[(X - \mu)^2])^2} \quad (3.10)$$

Ortalama Mutlak Sapma (Mean Absolute Deviation - MAD), çıkarılan alt bantlardaki değerlerin yayılımını tanımlamak için kullanılan bir kavramdır. Formülü ise Eş. 3.11’de gösterilmiştir.

$$\text{MAD} = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n} \quad (3.11)$$

3.5. Sınıflandırma

Son yıllarda sıkça karşılaşılan birçok farklı disiplinlerinde ortak çalışma alanlarından olan yapay zekâ kavramı, verilen görevleri insan zekâsını taklit ederek yerine getirmeye çalışan ve topladıkları bilgilerle kendini geliştiren sistemler olarak ifade edilebilir. Yapılan çalışmaların da temel amacı, makinelerin insan gibi düşünmesini sağlamak ve daha akıllı hale getirmeye çalışmaktır. Yapay zekâ uygulamaları birçok alanda kullanılmaktadır. Yapay zekâ ile beraber çeşitli alanlarda büyük değişimler yaşanmıştır.

Sınıflandırma bir veri kümesinde bulunan değerlerin sınıflara ayrılması işlemidir. Verileri ayırmak ve ait olduğu sınıfa doğru bir şekilde atama işlemi yapmayı amaçlayan bu yöntem; biyomedikal sinyallerin analizi, robotik, bilgisayar destekli tanı sistemleri, görüntü işleme gibi birçok alanda kullanılmaktadır. Hesaplama yöntemlerine göre farklı birçok sınıflandırıcı bulunmaktadır. Sınıflandırma algoritmaları seçilirken sınıflandırma aşamasında hatayı en az seviyelere indirecek yöntem araştırılmalı ve tercih edilmelidir.

Sınıflandırıcıların başarılarının sınanması için veri kümesi eğitim ve test kümeleri olarak ayrılmaktadır. K-parçalı çapraz doğrulama yöntemi, uygulanan sınıflandırıcı modelinin başarımını ölçmek için kullanılan yöntemlerden biridir. K-parçalı çapraz doğrulama yönteminde öncelikle bir k değeri belirlenir ve n adet verinin bulunduğu bir veri seti k adet parçaya bölünür. Her bir parçanın hem eğitim hem de test için kullanılması sağlanır. Bu değer geliştirilecek modele bağlı olarak değişmekle birlikte literatürde genellikle k değeri 10 olarak seçilmektedir. Her seferinde farklı bir veri seti kümesi test için ayrılır ve geriye kalan diğer k-1 veri seti eğitim için kullanılır. Test kümesi her defasında değiştirilerek geliştirilen model k defa eğitilir ve modelin performansı değerlendirilir. Bunun sonucunda k-farklı test sonucu elde edilmektedir. Bu sonuçların ortalaması alınarak genel sınıflandırma sonucu elde edilmektedir. Bu tezde 10 parçalı çapraz doğrulama yöntemi kullanılmıştır.

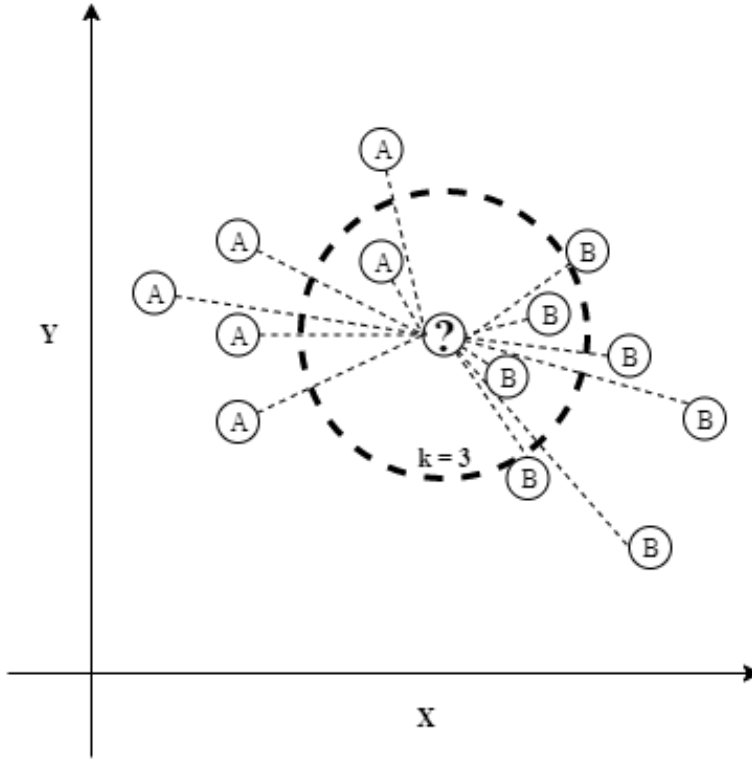
Bu bölümde, önceki aşamalar sonucunda elde edilen öznelik vektörleri üzerinde literatürde kullanımı yaygın olan sınıflandırıcılardan En Yakın k Komşu, Naïve Bayes, BayesNet, Çok Katmanlı Algılayıcılar, Destek Vektör Makineleri, AdaBoost, Random Subspace, Random Forest ve J48 algoritması olmak üzere toplamda 9 farklı yöntem kullanılarak sınıflandırma işlemi yapılmıştır. Bu algoritmalar ile ilgili tanım ve açıklamalar

bu bölümde yer almaktadır.

3.5.1. En yakın k komşu (k Nearest Neighbors) Sınıflandırıcısı

En Yakın k Komşu (k-NN) sınıflandırıcısı, makine öğrenmesi uygulamalarında yaygın olarak kullanılan etkili yöntemlerden biridir. Sınıflandırma problemlerini çözmek için geliştirilmiştir. En yakın komşularını temel alan bir yaklaşıma dayanmaktadır (Mitchell, 1997). Bu yöntem uygulanabilirlik açısından diğer makine öğrenme algoritmaları arasında en basiti olarak kabul edilir. Yöntemin çalışma prensibi incelendiğinde, yeni bir giriş verisinin önceden sınıflandırılmış diğer verilerle arasındaki uzaklığın çeşitli mesafe ölçütleri kullanılarak hesaplanması ardından hesaplanan uzaklıklara bağlı olarak en yakın k adet veriye göre uygun sınıfa karar verildiği bir sınıflandırma yöntemine dayanmaktadır (Demir ve Gökçen, 2019).

k-NN sınıflandırma algoritmasında uzaklık hesaplanırken genelde Öklid uzaklığı kullanılmaktadır. Ancak probleme bağlı olarak Minkowski, Chebyshev, Mahalanobis ve Manhattan uzaklığı gibi çeşitli mesafe ölçütleri de kullanılmaktadır. Şekil 3.7'de görüldüğü gibi k-NN sınıflandırıcısı için yeni bir giriş verisinin, sınıflandırılmış diğer veriler ile arasındaki uzaklık değerlendirilir ve etiketinin belirlenmesi gerçekleştirilir. Buna göre eğer $k = 3$ ise verinin B sınıfına atama işlemi gerçekleştirilir. Çünkü en yakın 3 değer içerisinde B sınıfına ait 2 ve A sınıfına ait sadece 1 veri vardır.



Şekil 3.7. k-NN sınıflandırma örneği

Bu çalışmada mesafe ölçütü olarak Öklid uzaklığı öğrenme örneği sayısı (k) ise 5 olarak belirlenmiş ve öznitelik vektörleri üzerinde k-NN sınıflandırıcısı kullanılarak sınıflandırma işlemi gerçekleştirilmiştir.

3.5.2. Naïve Bayes

Naïve Bayes sınıflandırıcısı, basit varsayımlar temelinde kurulan ancak başarılı sonuçlar elde edilen, olasılık tabanlı ve etkili bir sınıflandırma algoritmasıdır. Naïve Bayes algoritması Bayes teoremine dayalı geliştirilmiş bir sınıflandırma tekniği olup ismini İngiliz matematikçi Thomas Bayes'ten almıştır. Bu yöntemin uygulanması değişkenlerin birbirinden bağımsız olduğu varsayımıyla gerçekleştirilir. Uygulama aşamasında bir veri için oluşabilecek her durumun hesaplanması ve olasılık değeri yüksek olan duruma göre sınıflandırılması sağlanır. Naïve Bayes yöntemi hızlı hesaplama ve yüksek performans göstermesinden dolayı birçok makine öğrenmesi ve veri madenciliği uygulamalarında tercih edilmektedir (Kwon, A. Kwasinski, ve A. Kwasinski, 2019; Demir ve Gökçen, 2019). Bayes teoremi Eş. 3.12, Naïve Bayes sınıflandırıcısına ait fonksiyon ise Eş. 3.13'te

gösterilmiştir.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (3.12)$$

Denkleme göre;

$P(A|B)$; B olayı gerçekleştiği durumda A olayının meydana gelme olasılığını,

$P(B|A)$; A olayı gerçekleştiği durumda B olayının meydana gelme olasılığını,

$P(A)$ ve $P(B)$; A ve B olaylarının gerçekleşme olasılıklarını ifade etmektedir.

$$B' = \arg \max_{b_n \in Y} (P(B = y_n) \prod_{k=1}^l P(A = a_k | B = b_n)) \quad (3.13)$$

Bu çalışmada elde edilen öznitelik vektörleri üzerinde Naïve Bayes algoritması yardımıyla sınıflandırma işlemi gerçekleştirilmiş ve algoritmanın performans analizi yapılmıştır.

3.5.3. BayesNet

BayesNet sınıflandırma algoritması, Bayes teoremine dayanan olasılıksal analiz yöntemlerinden olup değişkenler arasındaki mantıksal ilişkilerin grafiksel olarak gösterilmesini sağlar. BayesNet sınıflandırıcısı, zor problemlerin verimli ve güvenilir bir şekilde çözülebilmesini sağlamak amacıyla mevcut yöntemlerin geliştirilmesi ile oluşturulmuştur. Yöntem birçok alanda kullanılmaktadır. BayesNet sınıflandırma algoritmasının fonksiyonu ise Eş. 3.14'deki gibi hesaplanmaktadır (Pearl, 1985; Friedman, Geiger ve Goldszmidt, 1997; Eraldemir ve E. Yıldırım, 2015).

$$p(X) = \prod_{i=0}^{n-1} p(X_i | \pi_i) \quad (3.14)$$

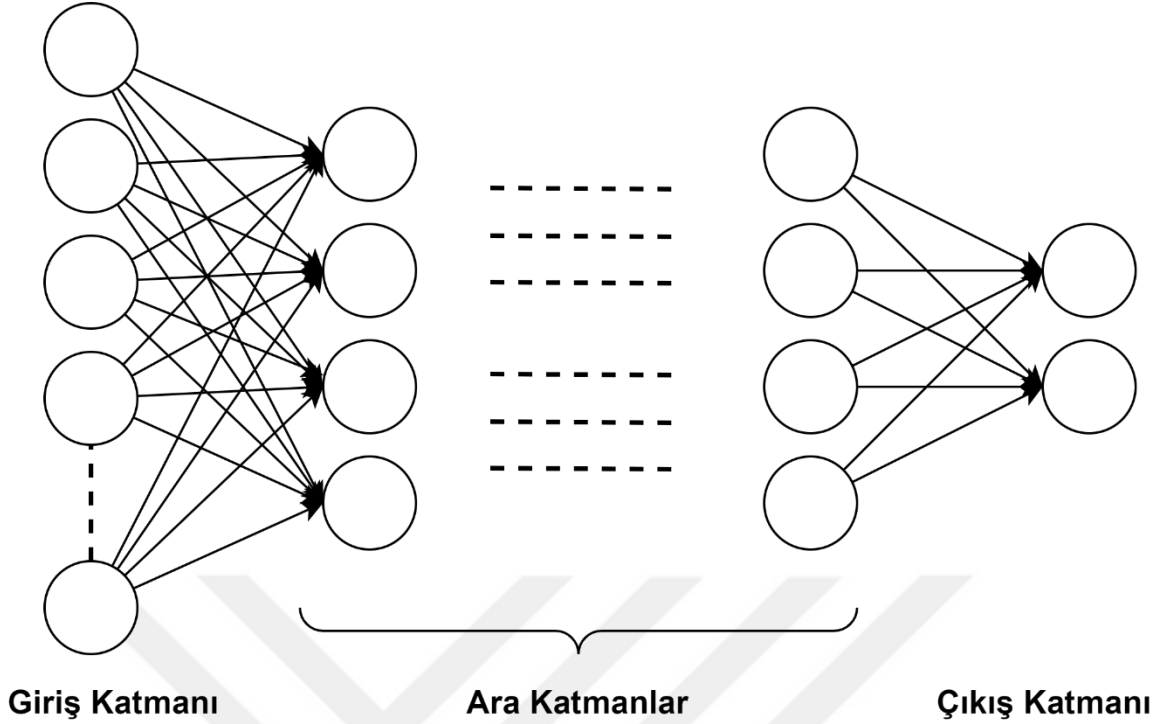
Bu çalışmada öznitelik vektörleri üzerinde BayesNet yöntemi kullanılarak sınıflandırma işlemi gerçekleştirilmiştir. Ardından yöntemin performans analizi yapılmıştır.

3.5.4. Çok katmanlı algılayıcı (Multilayer Perceptron)

Yapay sinir ağı, sinir hücrelerinden oluşan ve insan beyninin çalışmasından esinlenerek geliştirilen matematiksel modelleme yöntemidir. Farklı amaçlarda kullanılmak üzere birçok yapay sinir ağı modeli geliştirilmiştir. Bu modellerden biri olan çok katmanlı algılayıcılar (ÇKA), çalışmalarda oldukça etkili ve başarılı sonuçlar vermesi sebebiyle literatürde yaygın olarak kullanılan sınıflandırma algoritmalarından biridir.

ÇKA; bir giriş katmanı, bir veya daha fazla gizli katman ve bir çıktı katmanına sahiptir. Gizli katmanlar düğümlerden oluşmakta olup katman sayısı ile ilgili kesin bir kural yoktur. Katman ve nöron sayısı probleme ve yöntemin kullanım amacına göre değişiklik göstermektedir. ÇKA yapısındaki, giriş katmanında herhangi bir fonksiyon çalıştırılmaz, bu katmanın görevi dışarıdan alınan veriyi gizli katmanlara aktarmaktır. Gizli katmanlarda, girdi katmanından alınan veriler üzerinde bazı fonksiyonlar uygulanarak işlenmesi ve bu işlemler sonunda çıktı katmanına transfer işlemi gerçekleşmektedir. Son olarak çıktı katmanında ise yine bazı fonksiyonlar uygulanarak önceki katmandan gelen veri işlenir ve çıkış değerlerinin hesaplama işlemi gerçekleştirilir. ÇKA modeli yapısı Şekil 3.8'de gösterilmiştir.

Bu çalışmada öznitelik çıkarımı işlemi sonrasında elde edilen öznitelik vektörleri üzerinde 9 giriş, her biri 5 nörona sahip 2 gizli katman ve 2 çıkış elemanı içeren ÇKA modeli yardımıyla sınıflandırma işlemi gerçekleştirilmiştir. Ayrıca tüm katmanlarda sigmoid aktivasyon fonksiyonu kullanılmıştır.



Şekil 3.8. Çok katmanlı algılayıcı modeli yapısı

3.5.5. Destek Vektör Makineleri (Support Vector Machine)

Destek Vektör Makineleri (DVM), Vapnik tarafından istatistiksel öğrenme teorisi esas alınarak geliştirilmiştir (Vapnik, 1995). DVM yöntemi, makine öğrenmesi, örüntü tanıma, karar verme ve sınıflandırma gibi problemlerin çözümü için kullanılmaktadır. Kullanımı yaygın ve oldukça etkili sınıflandırma algoritmalarından olan DVM, regresyon analizi çalışmalarında da kullanılmaktadır. Birçok alanda kullanılan ve başarılı sonuçlar elde edilen DVM iki sınıfı birbirinden ayırmaya çalışır. Bu işlemde amaç sınıfları ayırabilen en uygun hiperdüzlemin tanımlanmasıdır. DVM’de lineer olarak ayrılabilen veya lineer olarak ayrılamayan yapıda olmak üzere iki durum vardır. Lineer olarak ayrılmış veriler üzerinde doğrudan maksimum sınır bulunması oldukça kolaydır. Ancak karşılaşılan çoğu problem lineer olarak ayrılamayan türden olup lineer olarak ayrılacakları farklı bir uzaya aktarma işlemi yapılması gerekir. Böylece çok boyutlu uzay dönüşümü yapıldıktan sonra işlemler uygulanabilir.

Bu çalışmada öznelik vektörleri üzerinde literatürde en fazla kullanılan sınıflandırma yöntemlerinden olan DVM sınıflandırıcısı yardımıyla sınıflandırma işlemi gerçekleştirilmiştir. Sınıflandırma sonuçlarına bağlı olarak algoritmanın performans analizi

gerçekleştirilmiştir.

3.5.6. AdaBoost

AdaBoost olarak bilinen Adaptive Boosting, Freund ve Schapire tarafından geliştirilen en önemli boosting algoritmalarından birisidir (Freund ve Schapire, 1996). Kullanımı kolay olan AdaBoost algoritması, sağlam teorik altyapısı nedeniyle uygulamalarda doğru ve başarılı sonuçları olan bir yaklaşımdır. AdaBoost algoritması, güçsüz olan öğrencileri bir araya getirip güçlü bir öğrenci oluşturmak için, yani performans geliştirmesi için kullanılır. Bu çalışmada da elde edilen öznelik vektörleri üzerinde AdaBoost algoritması yardımıyla sınıflandırma işlemi gerçekleştirilmiş ve algoritmanın performans analizi yapılmıştır.

3.5.7. Random SubSpace

Random Subspace algoritması, sınıflandırıcıların daha küçük problemler üzerinde çalışabilmesi için alt eğitim setleri oluşturarak eğitim işlemi yapmayı amaçlar. Bu yöntem ile özelliklerin tamamını kullanarak sınıflandırıcıların problemini arttırmak yerine rastgele seçilerek oluşturulan bir altküme üzerinde işlem yaparak sınıflandırıcıların daha küçük alt problemler üzerinde çalışmasını sağlar. Bu sayede daha hızlı ve verimli sonuç vermesine olanak sağlar. Ho (1998) yaptığı çalışmalar sonucunda oluşturulacak olan bu alt kümenin boyutu ile ilgili, toplam özellik sayısının yarısının kullanılmasını önermiştir (Amasyali, 2013). Bu çalışmada öznelik vektörleri üzerinde Random Subspace yöntemi yardımıyla sınıflandırma işlemi gerçekleştirilmiştir.

3.5.8. Random Forest

Random Forest algoritması, Leo Breiman tarafından geliştirilmiştir. Random Forest, içerisinde birbirinden bağımsız birçok karar ağacının bir araya gelmesiyle oluşan ve uygun olanı seçen bir ensemble algoritmadır (Breiman, 2001). Bu yöntemle sınıflandırma işlemi veya regresyon ağacı kurularak kümeleme işlemi yapılabilmektedir. Random Forest algoritması, veriyi sınıflandırma işlemi için rastgele oluşturulmuş birden çok karar ağacı

kullanır. Modelde kullanılacak olan her bir karar ağacı bootstrap tekniği ile çekilen verilerden oluşturulur.

Rastgele seçilen verilerden en iyilerinin tespit edilmesi aşamasından sonra bu veriler içerisinden en iyi bölecek olan veri seçilerek düğümler dallara ayrılır ve rastgele seçilerek ağaç yapısı geliştirilir. Random Forest algoritması sınıflandırma ve regresyon problemleri üzerinde uygulanabilir bir yöntem olmasından dolayı birçok çalışmada tercih edilen algoritmalarından biridir (Witten, Frank, Hall ve Pal, 2016). Bu çalışmada da elde edilen öznelik vektörleri üzerinde Random Forest algoritması yardımıyla sınıflandırma işlemi gerçekleştirilmiş ve algoritmanın performans analizi yapılmıştır.

3.5.9. J48

J48 algoritması, en bilindik karar ağacı algoritmalarından birisidir. Temeli ID3 ve C4.5 algoritmalarına dayanan J48 karar ağacı algoritması, sınıflandırma ve karar verme uygulamalarında yaygın olarak kullanılmaktadır. J48 algoritmasında, her adımda bütün özellikler kontrol edilir. Öznelik en yüksek kazanç oranına göre belirlenir ve en iyi kazanç veren tercih edilir. Bu yöntemin bir diğer özelliği ise budama işlemidir. J48 algoritması karar ağacını uygun şekilde oluşturmak için budama yöntemini kullanır. Bu işlem sayesinde sınıflandırma performansının artırılması, hatalarının minimize edilmesi ve karmaşıklığın azaltılması sağlanır (Salzberg,1994; Holmes, Donkin ve Witten, 1994). Bu çalışmada öznelik vektörleri üzerinde J48 algoritması yardımıyla sınıflandırma işlemi gerçekleştirilmiş ve algoritmanın performans analizi yapılmıştır.

3.6. Performans Ölçümleri ve İstatistiksel Analiz

Makine öğrenmesi ve benzeri alanlarda yapılan çalışmalarda, kullanılan sınıflandırma algoritmalarının performansını değerlendirmek amacıyla genellikle karışıklık matrisi kullanılmaktadır. Karışıklık matrisi, sınıflandırma doğruluk tablosu olarak da bilinmektedir. Bu tabloda veriler sınıflandırılmış değer ile referans değerlerinin karşılaştırılması ile elde edilmektedir (Çizelge 3.2). Performans ölçütleri de bu verilerden yararlanılarak hesaplanmaktadır.

Çizelge 3.2. Sınıflandırma doğruluğu tablo örneği

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

Tabloda yer alan;

True – Positive (TP), Gerçek sınıfta olan bir durumu doğru tahmin etme durumunu,
 True – Negative (TN), Gerçek sınıfta olmayan bir durumu doğru tahmin etme durumunu,
 False – Negative (FN), Gerçek sınıfta olan bir durumu yanlış tahmin etme durumunu,
 False – Positive (FP), Gerçek sınıfta olmayan bir durumu yanlış tahmin etme durumunu ifade etmektedir.

Bu aşamada, tahmin etme durumlarına bağlı olarak çeşitli performans değerlendirme ölçütleri hesaplanır. Bu çalışmada performans ölçütleri olarak en çok kullanılan parametrelerden hassasiyet, özgüllük ve genel başarımların değerleri kullanılmıştır. Performans değerlendirme ölçütleri karışıklık matrisi baz alınarak aşağıda gösterilen denklemler ile hesaplanmıştır.

$$Hassasiyet = \frac{TP}{TP + FN} * 100 \quad (3.15)$$

$$Özgüllük = \frac{TN}{TN + FP} * 100 \quad (3.16)$$

$$Genel Başarımlar = \frac{TP + TN}{Toplam} * 100 \quad (3.17)$$

4. ARAŞTIRMA BULGULARI VE SONUÇLARI

Solunum ve kalp rahatsızlıkları için en yaygın kullanılan teşhis yöntemi oskültasyon sesleridir. Akciğer oskültasyonu ile elde edilen ses sinyalleri tanısı büyük öneme sahip olan ve en ölümcül solunum bozukluklarından olan KOAH için de temel ve etkili sinyallerdir. Bu tezde, çok kanallı akciğer seslerinin ADD analizi ve farklı sınıflandırma algoritmaları yardımıyla sınıflandırılması ile kronik bir hastalık olan KOAH'ın bilgisayar destekli teşhisinde kullanılmak üzere bir uygulama geliştirilmiştir.

Yapılan çalışmada öncelikle 12 kanallı akciğer seslerinden ve 4 kanallı kalp seslerinden oluşan RespiratoryDatabase@TR veri tabanından alınan oskültasyon sinyalleri üzerinde gerekli ön işlemler yapılmıştır. Ardından KOAH teşhisi ile ilgili analizler için bu veri tabanında daha önce hiç kullanılmayan bir sinyal ayrıştırma yöntemi olan ADD uygulanmıştır. Sinyallerin uzunluğu ise 5 saniye olacak şekilde belirlenmiştir. (Bkz. Şekil 3.5) ve (Bkz. Şekil 3.6) örnek vermek amacıyla sağlıklı ve KOAH hastalarında ADD analizinden sonra çıkarılan ve rastgele seçilen modlardan bazılarını göstermektedir. ADD analizi sonucunda çıkarılan modlar üzerinde bazı matematiksel hesaplamalar yapılarak istatistiksel öznitelikler hesaplanmış ve sınıflandırma aşamasında kullanılacak öznitelik vektörleri oluşturulmuştur. Tüm bu işlemler sonucunda elde edilen öznitelik vektörleri üzerinde literatürde kullanımı yaygın olan sınıflandırıcılardan En Yakın k Komşu, Naïve Bayes, BayesNet, Çok Katmanlı Algılayıcılar, Destek Vektör Makineleri, AdaBoost, Random Subspace, Random Forest ve J48 algoritması olmak üzere toplamda 9 farklı sınıflandırma algoritması uygulanarak önerilen yöntemin başarısı test edilmiştir.

Kanalların sınıflandırıcı üzerindeki etkilerini incelemek amacıyla kullanılan tüm sınıflandırıcılar üzerinde her bir kanalın performans analizi yapılmıştır. Her bir kanal için farklı sınıflandırma algoritmaları uygulanarak elde edilen sonuçlar karşılaştırılarak performans etkileri analiz edilmiştir.

Tez kapsamında geliştirilen uygulamadan elde edilen bulgular ile ilgili bilgilere aşağıda yer verilmektedir.

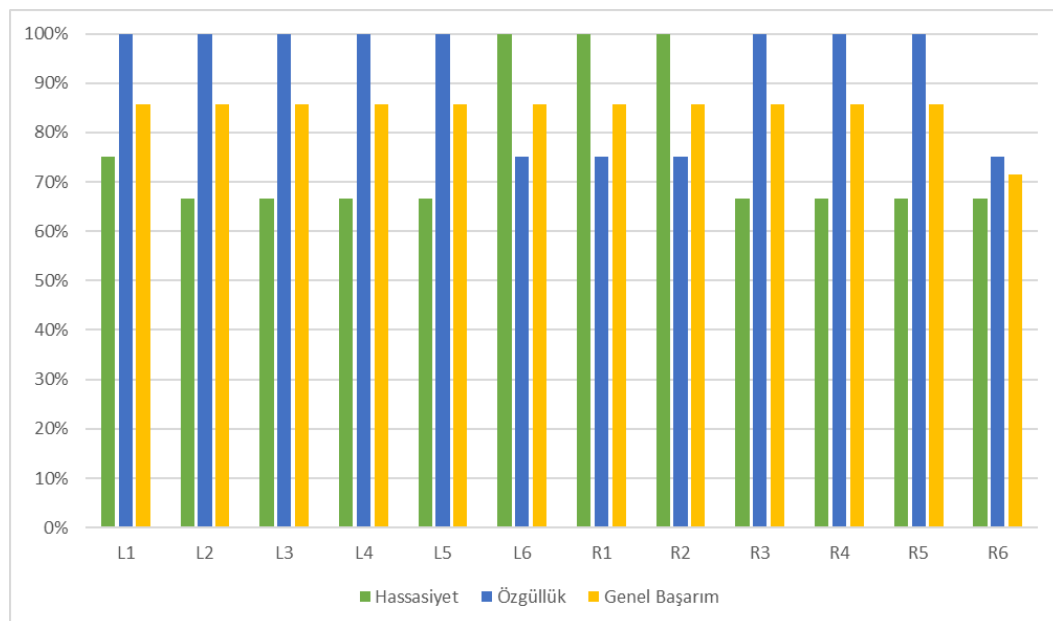
4.1. En yakın k komşu Sınıflandırıcısı Sonuçları

Önerilen yöntem üzerinde k-NN sınıflandırıcısı uygulanarak KOAH tespitinin performans değerlendirme sonuçları incelendiğinde, kanallara ve modlara göre farklı sonuçlar olmakla beraber önerilen yöntemin sınıflandırma sonucunda hassasiyet, özgüllük ve genel başarımları Çizelge 4.1’de gösterilmiştir.

Çizelge 4.1. k-NN sınıflandırıcısının değerlendirme sonuçları

Performans Ölçütü	Başarı Oranı (%)
Hassasiyet	84,00
Özgüllük	87,50
Genel Başarım	85,71

Her bir kanal için k-NN sınıflandırma algoritması üzerindeki performans etkileri analiz edilmiştir. Bu analizde kanalların k-NN sınıflandırıcısındaki performansının değerlendirme sonuçları Şekil 4.1’de gösterilmiştir. Şekil 4.1’deki verilere göre tüm kanalların k-NN sınıflandırıcısı üzerindeki performansları birbirine yakın olup R6 kanalının etkisi diğerlerine göre biraz daha düşük seviyelerdedir.



Şekil 4.1. Kanalların k-NN sınıflandırıcısı üzerindeki başarımları grafiği

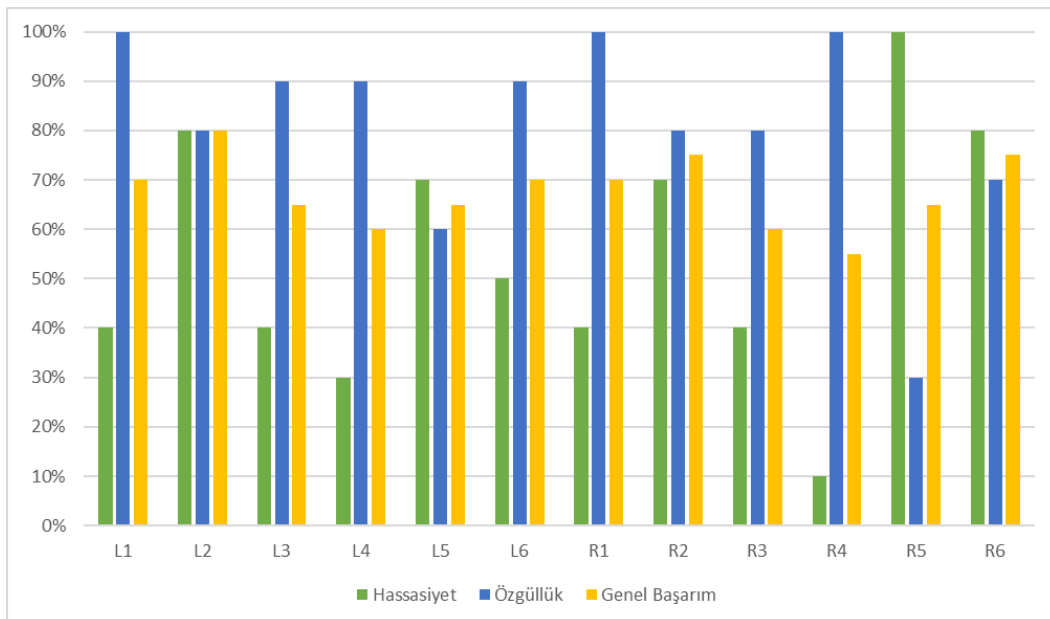
4.2. Naïve Bayes Sınıflandırıcısı Sonuçları

Naïve Bayes sınıflandırıcısının uygulama sonuçları incelendiğinde, kanallara ve modlara göre farklı sonuçlar olmakla beraber önerilen yöntemin sınıflandırma sonucunda hassasiyet, özgüllük ve genel başarımları Çizelge 4.2’de gösterilmiştir.

Çizelge 4.2. Naïve Bayes sınıflandırıcısının değerlendirme sonuçları

Performans Ölçütü	Başarı Oranı (%)
Hassasiyet	80,00
Özgüllük	80,00
Genel Başarım	80,00

Her bir kanal için Naïve Bayes sınıflandırma algoritması üzerindeki performans etkileri analiz edilmiştir. Bu analizde kanalların Naïve Bayes sınıflandırıcısındaki performansının değerlendirme sonuçları Şekil 4.2’de gösterilmiştir. Şekil 4.2’deki verilere göre L2, R2 ve R6 kanallarının Naïve Bayes sınıflandırıcısı üzerindeki performansları diğer kanallara kıyasla daha yüksek olup L4, R3 ve R4 kanallarının etkisi ise diğerlerine göre daha düşük seviyelerdedir.



Şekil 4.2. Kanalların Naïve Bayes sınıflandırıcısı üzerindeki başarımları grafiği

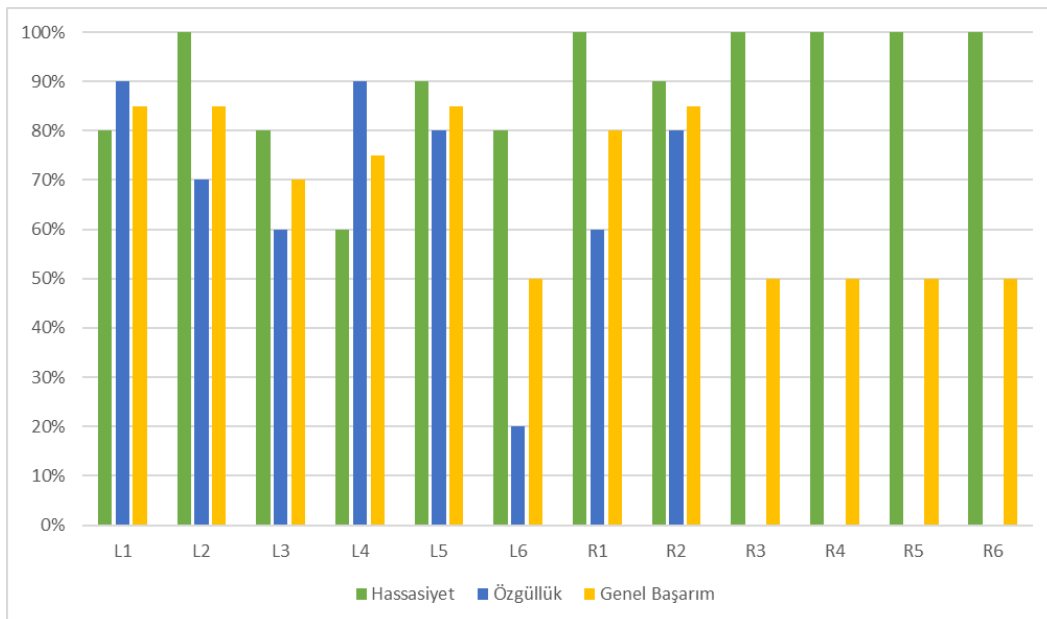
4.3. BayesNet Sınıflandırıcısı Sonuçları

BayesNet sınıflandırıcısının sonuçları incelendiğinde, kanallara ve modlara göre farklı sonuçlar olmakla beraber önerilen yöntemin sınıflandırma sonucunda hassasiyet, özgüllük ve genel başarımları Çizelge 4.3'te gösterilmiştir.

Çizelge 4.3. BayesNet sınıflandırıcısının değerlendirme sonuçları

Performans Ölçütü	Başarı Oranı (%)
Hassasiyet	80,00
Özgüllük	90,00
Genel Başarım	85,00

Her bir kanal için BayesNet sınıflandırma algoritması üzerindeki performans etkileri analiz edilmiştir. Bu analizde kanalların BayesNet sınıflandırıcısındaki performansının değerlendirme sonuçları Şekil 4.3'te gösterilmiştir. Şekil 4.3'teki verilere göre L1, L2, L5 ve R2 kanallarının BayesNet sınıflandırıcısı üzerindeki performansları diğer kanallara göre daha yüksek seviyelerde olup L6, R3, R4, R5 ve R6 kanallarının etkisi diğer kanallara göre daha düşük seviyelerdedir.



Şekil 4.3. Kanalların BayesNet sınıflandırıcısı üzerindeki başarımlar grafiği

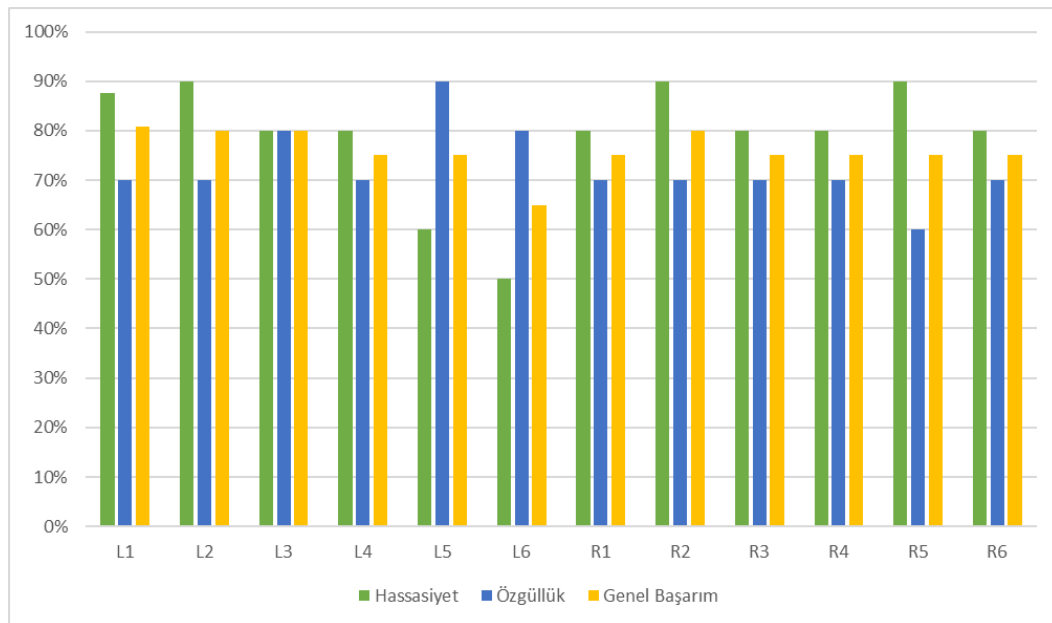
4.4. Çok Katmanlı Algılayıcı Sonuçları

Çok katmanlı algılayıcı modelinin sonuçları incelendiğinde, kanallara ve modlara göre farklı sonuçlar olmakla beraber önerilen yöntemin sınıflandırma sonucunda hassasiyet, özgüllük ve genel başarımları Çizelge 4.4'te gösterilmiştir.

Çizelge 4.4. Çok katmanlı algılayıcı modelinin değerlendirme sonuçları

Performans Ölçütü	Başarı Oranı (%)
Hassasiyet	87,50
Özgüllük	70,00
Genel Başarım	80,77

Her bir kanal için Çok katmanlı algılayıcı modeli üzerindeki performans etkileri analiz edilmiştir. Bu analizde kanalların Çok katmanlı algılayıcı üzerindeki performansının değerlendirme sonuçları Şekil 4.4'te gösterilmiştir. Şekil 4.4'teki verilere göre tüm kanalların Çok katmanlı algılayıcı modeli üzerindeki performansları birbirine yakın olup L6 kanalının etkisi diğer kanallara göre düşük seviyelerdedir.



Şekil 4.4. Kanalların Çok katmanlı algılayıcılar üzerindeki başarımları grafiği

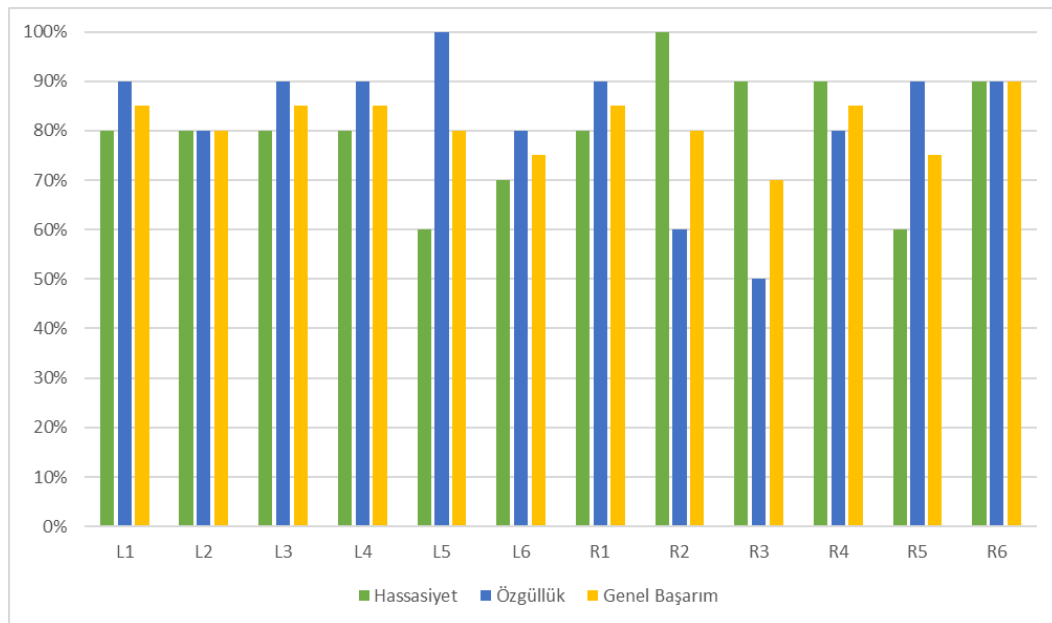
4.5. Destek Vektör Makineleri Sonuçları

DVM sonuçları incelendiğinde, kanallara ve modlara göre farklı sonuçlar olmakla beraber önerilen yöntemin sınıflandırma sonucunda hassasiyet, özgüllük ve genel başarımları Çizelge 4.5'te gösterilmiştir.

Çizelge 4.5. DVM sınıflandırıcısının değerlendirme sonuçları

Performans Ölçütü	Başarı Oranı (%)
Hassasiyet	80,00
Özgüllük	100,00
Genel Başarım	90,00

Her bir kanal için DVM üzerindeki performans etkileri analiz edilmiştir. Bu analizde kanalların DVM üzerindeki performansının değerlendirme sonuçları Şekil 4.5'te gösterilmiştir. Şekil 4.5'teki verilere göre R6 kanalının DVM üzerindeki performansı diğerlerine göre en yüksek olup R3 kanalının etkisi ise diğerlerine göre en düşük olarak görülmektedir.



Şekil 4.5. Kanalların Destek Vektör Makineleri üzerindeki başarımları grafiği

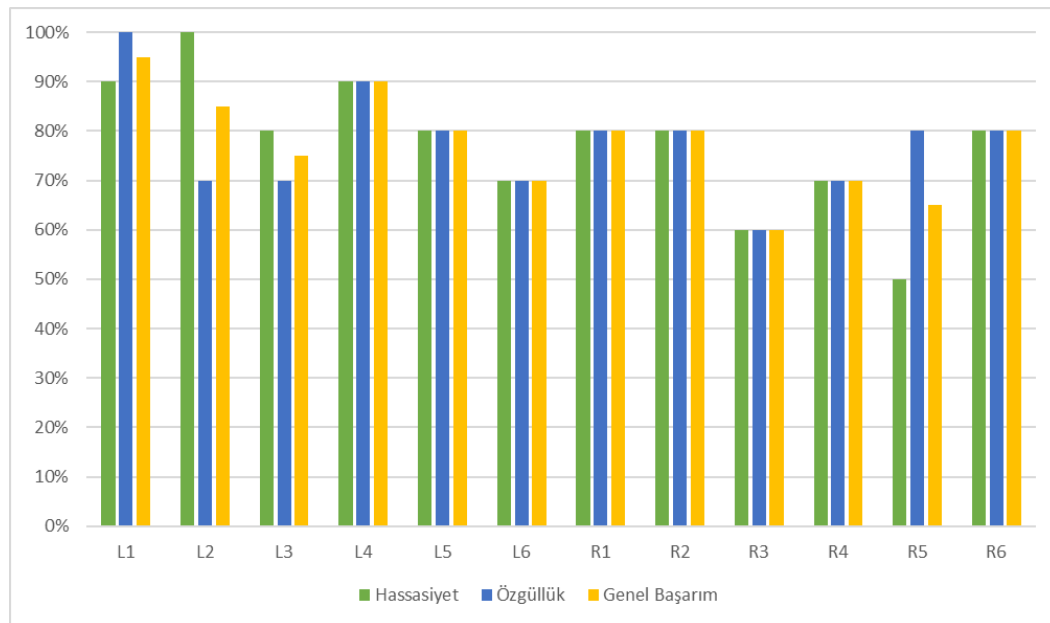
4.6. AdaBoost Algoritması Sonuçları

AdaBoost algoritması uygulandıktan sonra elde edilen sonuçlar incelendiğinde, kanallara ve modlara göre farklı sonuçlar olmakla beraber önerilen yöntemin sınıflandırma sonucunda hassasiyet, özgülük ve genel başarımları Çizelge 4.6’da gösterilmiştir.

Çizelge 4.6. AdaBoost algoritmasının değerlendirme sonuçları

Performans Ölçütü	Başarı Oranı (%)
Hassasiyet	90,00
Özgüllük	100,00
Genel Başarım	95,00

Her bir kanal için AdaBoost üzerindeki performans etkileri analiz edilmiştir. Bu analizde kanalların AdaBoost üzerindeki performansının değerlendirme sonuçları Şekil 4.6’da gösterilmiştir. Şekil 4.6’daki verilere göre sırasıyla L1 ve L4 kanallarının Adaboost üzerindeki performansları diğerlerinde göre yüksek olup R3 kanalının etkisi ise diğerlerine göre en düşük olarak görülmektedir.



Şekil 4.6. Kanalların AdaBoost sınıflandırıcısı üzerindeki başarımları grafiği

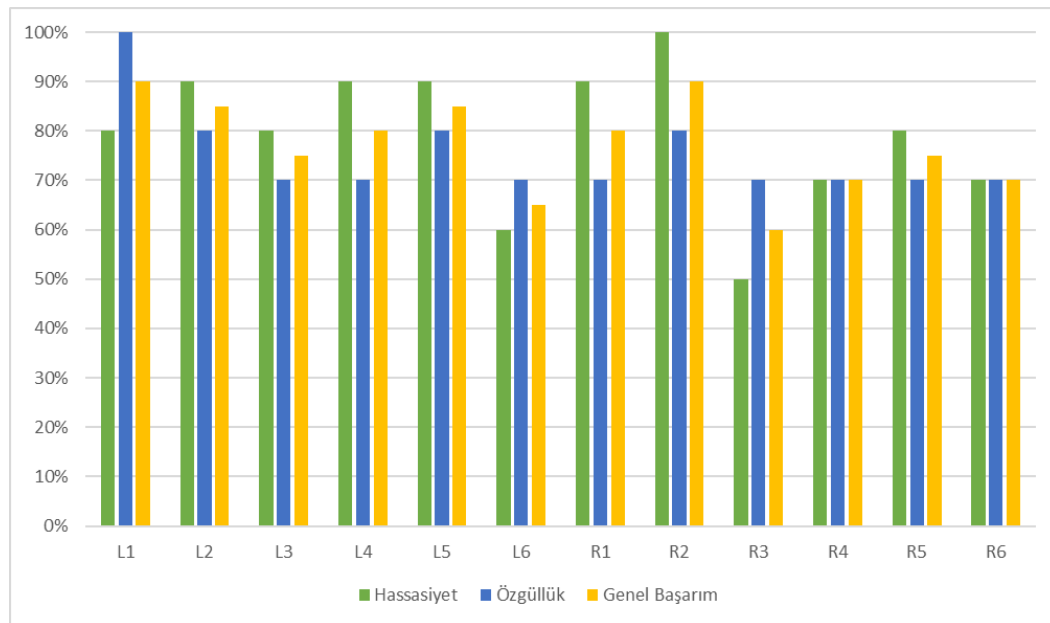
4.7. Random Subspace Algoritması Sonuçları

Random Subspace algoritmasının sonuçları incelendiğinde, kanallara ve modlara göre farklı sonuçlar olmakla beraber önerilen yöntemin sınıflandırma sonucunda hassasiyet, özgüllük ve genel başarımları Çizelge 4.7’de gösterilmiştir.

Çizelge 4.7. Random Subspace algoritmasının değerlendirme sonuçları

Performans Ölçütü	Başarı Oranı (%)
Hassasiyet	80,00
Özgüllük	100,00
Genel Başarım	90,00

Her bir kanal için Random Subspace üzerindeki performans etkileri analiz edilmiştir. Bu analizde kanalların Random Subspace üzerindeki performansının değerlendirme sonuçları Şekil 4.7’de gösterilmiştir. Şekil 4.7’deki verilere göre L1 ve R2 kanallarının Random Subspace algoritması üzerindeki performansları diğer kanallara göre daha yüksek olup L6 ve R3 kanallarının etkisi ise diğer kanallara göre daha düşük seviyelerdedir.



Şekil 4.7. Kanalların Random Subspace sınıflandırıcısı üzerindeki başarımları grafiği

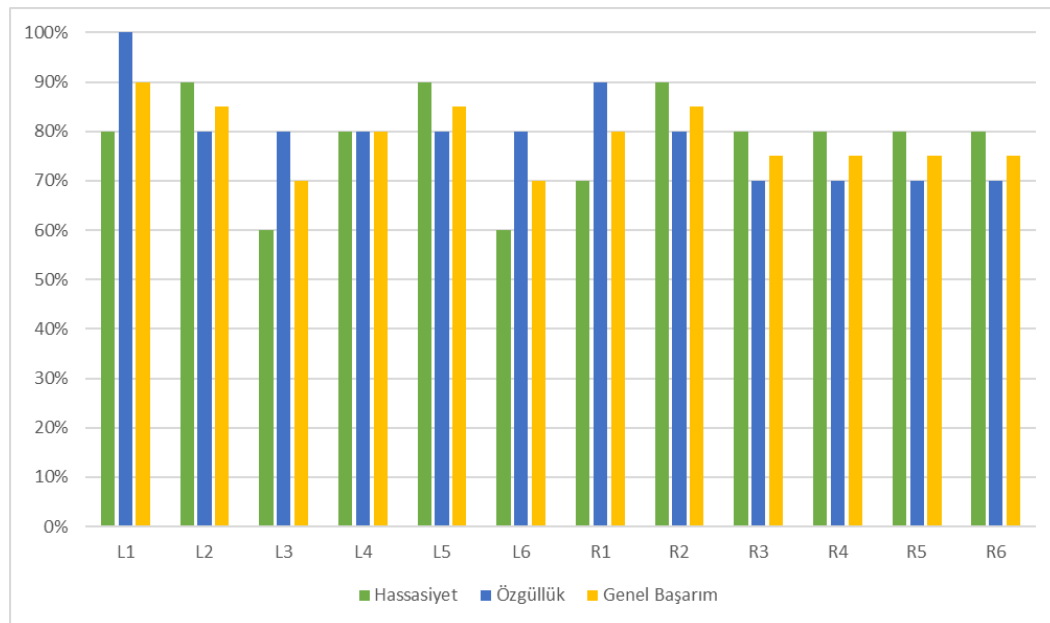
4.8. Random Forest Algoritması Sonuçları

Random Forest algoritmasının sonuçları incelendiğinde, kanallara ve modlara göre farklı sonuçlar olmakla beraber önerilen yöntemin sınıflandırma sonucunda hassasiyet, özgüllük ve genel başarımları Çizelge 4.8’de gösterilmiştir.

Çizelge 4.8. Random Forest algoritmasının değerlendirme sonuçları

Performans Ölçütü	Başarı Oranı (%)
Hassasiyet	80,00
Özgüllük	100,00
Genel Başarım	90,00

Her bir kanal için Random Forest üzerindeki performans etkileri analiz edilmiştir. Bu analizde kanalların Random Forest üzerindeki performansının değerlendirme sonuçları Şekil 4.8’de gösterilmiştir. Şekil 4.9’daki verilere göre L1 kanalının Random Forest algoritması üzerindeki performansı diğer kanallara göre yüksek olup R3 kanalının diğer kanallara göre etkisi daha düşük seviyelerdedir.



Şekil 4.8. Kanalların Random Forest sınıflandırıcısı üzerindeki başarımları grafiği

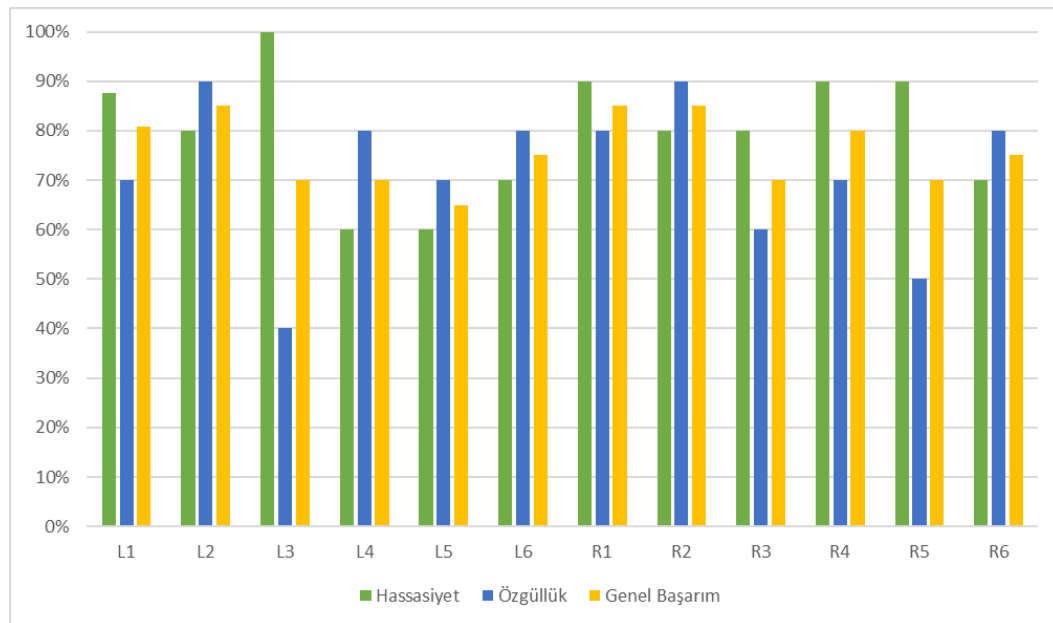
4.9. J48 Algoritması Sonuçları

Son olarak J48 algoritmasından elde edilen sonuçlar incelendiğinde, kanallara ve modlara göre farklı sonuçlar olmakla beraber önerilen yöntemin sınıflandırma sonucunda hassasiyet, özgüllük ve genel başarımları Çizelge 4.9’da gösterilmiştir.

Çizelge 4.9. J48 algoritmasının değerlendirme sonuçları

Performans Ölçütü	Başarı Oranı (%)
Hassasiyet	80,00
Özgüllük	90,00
Genel Başarım	85,00

Her bir kanal için J48 algoritması üzerindeki performans etkileri analiz edilmiştir. Bu analizde kanalların J48 üzerindeki performansının değerlendirme sonuçları ise Şekil 4.9’da gösterilmiştir. Şekil 4.9’daki verilere göre L2, R1 ve R2 kanallarının J48 algoritması üzerindeki performansları diğer kanallara göre daha yüksek seviyelerde olup L5 kanalının diğer kanallar göre etkisi daha düşük seviyelerdedir.



Şekil 4.9. Kanalların J48 üzerindeki başarımları grafiği

4.10. Kanalların Genel Başarım Üzerindeki Etkileri

Çizelge 4.10'da kanalların farklı sınıflandırma algoritmaları üzerindeki etkilerini incelemek amacıyla yapılan çalışma sonucunda her bir kanalın tüm sınıflandırıcılar üzerinde elde edilen başarı ortalaması ve en yüksek değeri gösterilmiştir.

Çizelge 4.10. Kanalların genel başarım üzerindeki etkisi

Kanallar	Ortalama Başarı Oranı (%)	Maksimum Başarı Oranı (%)
L1	84,75	95,00
L2	83,41	85,71
L3	75,08	85,71
L4	77,86	90,00
L5	78,41	85,71
L6	69,52	85,71
R1	80,08	85,71
R2	82,86	90,00
R3	66,19	85,71
R4	71,75	85,71
R5	70,63	85,71
R6	73,49	90,00

5. SONUÇ VE ÖNERİLER

Gelişen teknoloji, veri analizi ve sınıflandırma için çeşitli yöntemlerin kullanımını mümkün kılmıştır. Bu tez kapsamında, çok kanallı akciğer seslerinin ADD yöntemi ile analizi ve sınıflandırılması ile küresel bir sağlık sorunu olan KOAH'ın bilgisayar destekli teşhisinde kullanılmak üzere bir uygulama geliştirilmiştir. Yapay zekâ temelli klinik tanı ve tedavi uygulamalarında kullanılabilecek bu çalışmada ADD tabanlı özniteliklerin kullanılması önerilmiş ve yöntemin sinyaller üzerinde etkisi kapsamlı bir şekilde araştırılmıştır. Bu yöntemin başarısını ispat etmek için çeşitli sınıflandırma algoritmaları kullanılarak yöntem test edilmiştir.

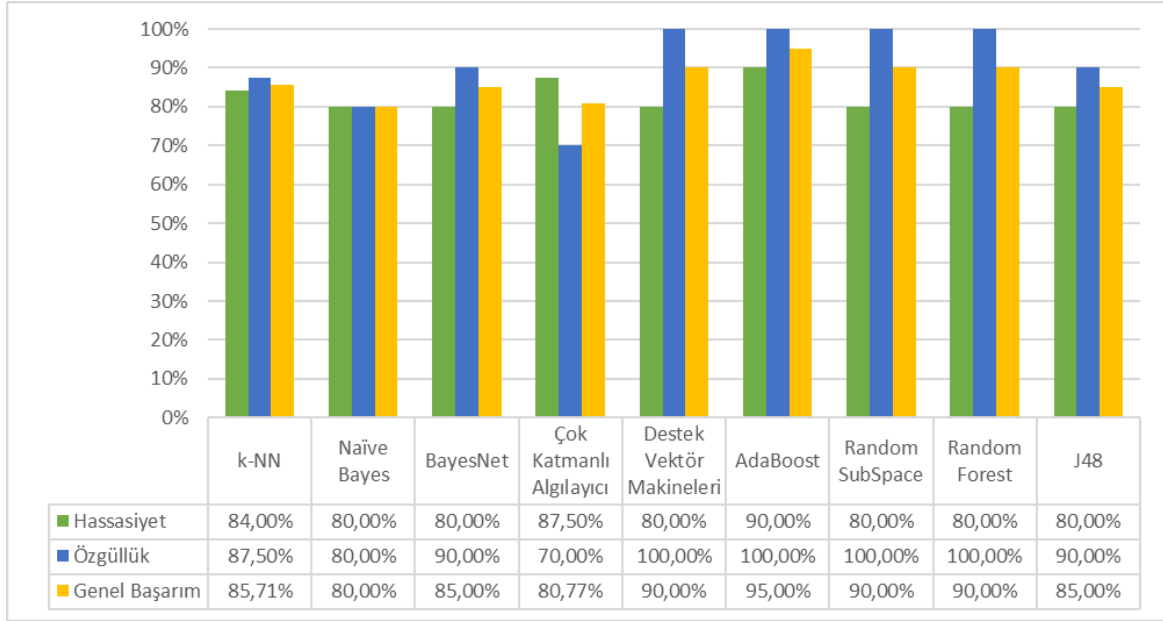
Bu tezde veri tabanından alınan sinyaller üzerinde yapılan çalışmalar sonucunda elde edilen öznitelik vektörlerinin farklı sınıflandırma algoritmaları ile yapılan sınıflandırmada tespit edilen sonuçları Çizelge 5.1'de gösterilmiştir.

Çizelge 5.1. Sınıflandırma sonuçları

Sınıflandırıcı	Hassasiyet (%)	Özgüllük (%)	Genel Başarım (%)
k-NN	84,00	87,50	87,51
Naïve Bayes	80,00	80,00	80,00
BayesNet	80,00	90,00	85,00
Çok Katmanlı Algılayıcı	87,50	70,00	80,77
Destek Vektör Makineleri	80,00	100,00	90,00
AdaBoost	90,00	100,00	95,00
Random Subspace	80,00	100,00	90,00
Random Forest	80,00	100,00	90,00
J48	80,00	90,00	85,00

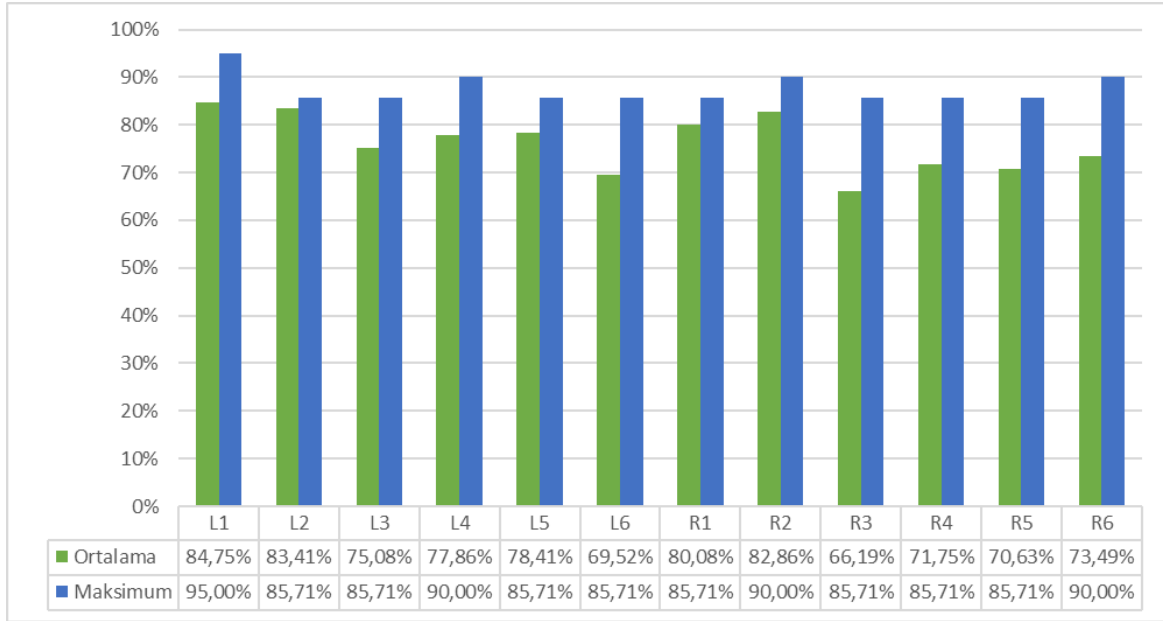
Şekil 5.1'de verilen sonuçlar incelendiğinde önerilen yöntemin tüm sınıflandırıcılarda

oldukça başarılı sonuçlar verdiği görülmektedir. Sınıflandırıcılar karşılaştırıldığında AdaBoost algoritması %90,00 hassasiyet, %100,00 özgüllük ve %95,00 genel başarımla diğerlerine göre daha iyi sonuç vermektedir.



Şekil 5.1. Sınıflandırıcı başarımları karşılaştırma grafiği

Şekil 5.2’de verilen grafiğe göre kanalların performans etkileri incelendiğinde ise %84,75 ortalama ve %95,00 maksimum başarımla sahip olan L1 kanalı diğer kanallara göre daha başarılı sonuç sergilemektedir.



Şekil 5.2. Kanalların performans etkileri

Biyomedikal sinyallerin analizinde sinyal uzunluğu önemli bir faktördür. Biyomedikal analiz çalışmalarındaki temel amaç hızlı ve yüksek performanslı sistemler geliştirmektir. Altan (2018) yaptığı çalışmada sinyal uzunluğunun başarı üzerine etkisini araştırmış ve çalışma sonucunda 5, 10 ve 15 saniye için doğruluk oranlarını %64,23, %93,29 ve %81,71 olarak tespit etmiştir.

Bu tezde sinyal süresi 5 saniye olacak şekilde ayarlanmıştır. Analiz sonucu, geliştirilen sistemin daha kısa sürede, yüksek performanslı ve oldukça başarılı sonuç verdiğini göstermektedir. Burada ADD yönteminin avantajı ortaya çıkmaktadır. Çalışmalar sonucunda elde edilen verilere göre ADD yönteminin kısa süreli sinyalleri anlamlı hale getirdiği ve yöntemin başarılı olduğu görülmektedir.

Bu çalışmanın literatüre sunduğu en önemli avantajlarından biri teşhis koyulması çok güç olan KOAH'ın bilgisayar destekli analizini yaparak 5 saniye gibi kısa bir sürede bile yüksek performans elde edilen bir model olmasıdır. Ayrıca uzman personele bağımlılığın azaltılması, bu sebeplerle az gelişmiş dünya ülkelerinde kullanımı muhtemel olması çalışmanın önemini göstermektedir. Bu çalışmada kullanılan veri tabanı üzerinde ADD tekniği kullanılarak daha önce hiç çalışma yapılmamıştır. Bundan dolayı çalışmanın önemli bir özelliği de kullanılan veri tabanı üzerinde ilgili teknik uygulanarak KOAH analizi ve tespiti için yapılan ilk çalışma olmasıdır. Analiz sonucu önerilen algoritmanın

oldukça başarılı sonuçlar verdiğini göstermektedir.

Araştırma sonucunda elde edilen sonuçlar incelendiğinde önerilen yöntemin başarılı sınıflandırma performansları elde ettiği görülmektedir. Geliştirilen yöntem derin öğrenme modeli veya benzeri yaklaşımlar ile daha iyi sınıflandırma performansı sağlayabilir. İleriki çalışmalarda farklı sinyal ayrıştırma metotları ve sınıflandırıcılar ile başarımlarının yükseltilmesi, yöntemlerin donanım üzerinde gerçekleştirilmesi ve elde edilen sonuçlardan faydalanarak bir cihaz geliştirilmesi amaçlanmaktadır. Önerilen yöntemin dinamik bir yapıya sahip olması nedeniyle farklı biyomedikal sinyallerin analizinde de kullanılarak diğer hastalıkların teşhisi için kullanılmak üzere genişletilebilir.



KAYNAKLAR

- Altan, G., Kutlu, Y., & Allahverdi, N. (2019). Deep Learning on Computerized Analysis of Chronic Obstructive Pulmonary Disease. *IEEE journal of biomedical and health informatics*.
- Altan, G., Kutlu, Y., Pekmezci, A. Ö., & Nural, S. (2018). Deep learning with 3D-second order difference plot on respiratory sounds. *Biomedical Signal Processing and Control*, 45, 58-69.
- Altan, G. (2018). Enhancing deep extreme learning machines: novel multi-kernel autoencoders and implementation for detecting the chronic obstructive pulmonary disease, Doktora Tezi, *Mustafa Kemal Üniversitesi Fen Bilimleri Enstitüsü*, Hatay.
- Altan, G., Kutlu, Y., Garbi, Y., Pekmezci, A. Ö., & Nural, S. (2017). Multimedia Respiratory Database (RespiratoryDatabase@ TR): Auscultation Sounds and Chest X-rays. *Natural and Engineering Sciences*, 2(3), 59-72.
- Amasyali, M. F. (2013, April). A semi-random subspace method for classification ensembles. In *2013 21st Signal Processing and Communications Applications Conference (SIU)* (pp. 1-4). IEEE.
- Amezquita-Sanchez, J. P., & Adeli, H. (2015). A new music-empirical wavelet transform methodology for time–frequency analysis of noisy nonlinear and non-stationary signals. *Digital Signal Processing*, 45, 55-68.
- Babu, K. A., Ramkumar, B., & Manikandan, M. S. (2019, May). Empirical Wavelet Transform Based Lung Sound Removal from Phonocardiogram Signal for Heart Sound Segmentation. In *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 1313-1317). IEEE.
- Becker, K. W., Scheffer, C., Blanckenberg, M. M., & Diacon, A. H. (2013, July). Analysis of adventitious lung sounds originating from pulmonary tuberculosis. In *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 4334-4337). IEEE.
- Bhattacharyya, A., & Pachori, R. B. (2017). A multivariate approach for patient-specific EEG seizure detection using empirical wavelet transform. *IEEE Transactions on Biomedical Engineering*, 64(9), 2003-2015.
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32. doi:10.1023/A:1010933404324
- Chen, J., Pan, J., Li, Z., Zi, Y., & Chen, X. (2016). Generator bearing fault diagnosis for wind turbine via empirical wavelet transform using measured vibration signals. *Renewable Energy*, 89, 80-92.
- Demir, E. & Gökçen, A. (2019), Empirical Wavelet Transform Analysis and Classification

- of Multichannel Lung Sounds. *2nd International Mersin Symposium*, Cilt 6, 218-226.
- Demir, E., Gökçen, A., Altan G. & Kutlu Y. (2019), The Effect of Nonlinear Classifiers On the Diagnosis of Respiratory Diseases Using Empirical Wavelet Transform Based Features. *3rd International Academic Studies Conference*, 178-185.
- Eraldemir, S. G., & Yıldırım, E. (2015, May). Comparison of wavelets for classification of cognitive EEG signals. In *2015 23rd Signal Processing and Communications Applications Conference (SIU)* (pp. 1381-1384). IEEE.
- Erdoğan, E., Polatlı, M., Kocabaş, A., Yıldırım, N., Gürgün, A., Saryal, S., ... & Sevinç, C. (2010). Türk toraks derneği kronik obstrüktif akciğer hastalığı tanı ve tedavi uzlaşma raporu 2010. *Türk Toraks Dergisi Mayıs*, 11(1), 1-66.
- Freund, Y., & Schapire, R. E. (1996, July). Experiments with a new boosting algorithm. In *icml* (Vol. 96, pp. 148-156).
- Friedman, N., Geiger, D., & Goldszmidt, M. (1997). Bayesian network classifiers. *Machine learning*, 29(2-3), 131-163.
- Gilles, J. (2013). Empirical wavelet transform. *IEEE transactions on signal processing*, 61(16), 3999-4010.
- Günaydın, F. E., Kalkan, N., Günlüoğlu, G., Aktepe, E. N., Demirkol, B. ve Altın, S. (2019). The relationship between serum levels of surfactant protein D in COPD exacerbation severity and mortality. *Turkish Journal of Medical Sciences*, 49, 888-893 doi:10.3906/sag-1809-6
- Ho, T. K., "The random subspace method for constructing decision forests", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(8), 1998
- Holmes, G., Donkin, A., & Witten, I. H. (1994). Weka: A machine learning workbench.
- Hong, K. J., Essid, S., Ser, W. ve Foo, D. C.-G. (2018). A robust audio classification system for detecting pulmonary edema. *Biomedical Signal Processing and Control*, 46, 94–103. doi:10.1016/j.bspc.2018.07.004
- Huang, D., Zhang, S., & Zhang, Y. (2017, November). EEG-based emotion recognition using empirical wavelet transform. In *2017 4th International Conference on Systems and Informatics (ICSAI)* (pp. 1444-1449). IEEE.
- Islam, M. A., Bandyopadhyaya, I., Bhattacharyya, P., & Saha, G. (2018). Multichannel lung sound analysis for asthma detection. *Computer methods and programs in biomedicine*, 159, 111-123.
- İnternet: WHO, (2018), The top 10 causes of death, <https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death>, Son Erişim Tarihi: 05.11.2019.
- Jambholkar, T., Gurve, D., & Sharma, P. B. (2015, September). Application of empirical wavelet transform (EWT) on images to explore brain tumor. In *2015 international*

conference on signal processing, computing and control (ISPCC) (pp. 200-204). IEEE.

- Kedadouche, M., Thomas, M., & Tahan, A. (2016). A comparative study between Empirical Wavelet Transforms and Empirical Mode Decomposition Methods: Application to bearing defect diagnosis. *Mechanical Systems and Signal Processing*, 81, 88-107.
- Khadar, S. A., Tabatabaey-Mashadi, N., & Daliri, G. M. (2019). A simple realtime algorithm for automatic external defibrillator. *Biomedical Signal Processing and Control*, 51, 277-284.
- Khodabakhshi, M. B., & Moradi, M. H. (2017). The attractor recurrent neural network based on fuzzy functions: An effective model for the classification of lung abnormalities. *Computers in biology and medicine*, 84, 124-136.
- Kocabaş, A., Atış, S., Çöplü, L., Erdinç, E., Ergan, B., Gürgün, A., ... & Yıldırım, N. (2014). Kronik obstrüktif akciğer hastalığı (KOAH) koruma, tani ve tedavi raporu 2014. *Official journal of the Turkish Thoracic Society*, 15.
- Kwon, Y., Kwasinski, A., & Kwasinski, A. (2019). Solar Irradiance Forecast Using Naïve Bayes Classifier Based on Publicly Available Weather Forecasting Variables. *Energies*, 12(8), 1529.
- Liu, W., Cao, S., & Chen, Y. (2015). Seismic time–frequency analysis via empirical wavelet transform. *IEEE Geoscience and Remote Sensing Letters*, 13(1), 28-32.
- Liu, X., Deng, K., Chen, S., Zhang, Y., Yao, J., Weng, X., Zhang, Y., Gao, T ve Feng, G. (2019). 8-Hydroxy-2'-deoxyguanosine as a biomarker of oxidative stress in acute exacerbation of chronic obstructive pulmonary disease. *Turkish Journal of Medical Sciences*, 49, 93-100 doi:10.3906/sag-1807-106
- Maharaja, D., & Shaby, D. S. M. (2017). Empirical Wavelet Transform and GLCM Features Based Glaucoma Classification from Fundus Image. *International Journal of MC Square Scientific Research*, 9(1), 79-85. doi:10.20894/ijmsr.117.009.001.010
- Maheshwari, S., Pachori, R. B., & Acharya, U. R. (2016). Automated diagnosis of glaucoma using empirical wavelet transform and correntropy features extracted from fundus images. *IEEE journal of biomedical and health informatics*, 21(3), 803-813.
- Maruf, S. O., Azhar, M. U., Khawaja, S. G., & Akram, M. U. (2015, December). Crackle separation and classification from normal Respiratory sounds using Gaussian Mixture Model. In *2015 IEEE 10th International Conference on Industrial and Information Systems (ICIIS)* (pp. 267-271). IEEE.
- Mitchell, T., (1997). *Machine Learning*, McGraw-Hill Higher Education. New York.
- Mondal, A., Banerjee, P., & Somkuwar, A. (2017). Enhancement of lung sounds based on empirical mode decomposition and Fourier transform algorithm. *Computer methods and programs in biomedicine*, 139, 119-136.

- Oung, Q. W., Muthusamy, H., Basah, S. N., Lee, H., & Vijejan, V. (2018). Empirical Wavelet Transform Based Features for Classification of Parkinson's Disease Severity. *Journal of medical systems*, 42(2), 29.
- Pan, J., Chen, J., Zi, Y., Li, Y., & He, Z. (2016). Mono-component feature extraction for mechanical fault diagnosis using modified empirical wavelet transform via data-driven adaptive Fourier spectrum segment. *Mechanical Systems and Signal Processing*, 72, 160-183.
- Pearl, J. (1985). Bayesian networks: A model of self-activated memory for evidential reasoning. In *Proceedings of the 7th Conference of the Cognitive Science Society*, 1985 (pp. 329-334).
- Salzberg, S. L. (1994). C4. 5: Programs for machine learning by j. ross quinlan. morgan kaufmann publishers, inc., 1993. *Machine Learning*, 16(3), 235-240.
- Sengupta, N., Sahidullah, M., & Saha, G. (2016). Lung sound classification using cepstral-based statistical features. *Computers in biology and medicine*, 75, 118-129.
- Sin, D. D., Hollander, Z., DeMarco, M. L., McManus, B. M., & Ng, R. T. (2015). Biomarker development for chronic obstructive pulmonary disease. From discovery to clinical implementation. *American journal of respiratory and critical care medicine*, 192(10), 1162-1170.
- Singh, O., & Sunkaria, R. K. (2016). Empirical wavelet transform-based delineator for arterial blood pressure waveforms. *Bio-Algorithms and Med-Systems*, 12(2), 61-66.
- Thirumala, K., Umarikar, A. C., & Jain, T. (2014). Estimation of single-phase and three-phase power-quality indices using empirical wavelet transform. *IEEE Transactions on power delivery*, 30(1), 445-454.
- Vapnik, V. N. (1995). *The Nature of Statistical Learning Theory*. New York, NY: Springer New York. doi:10.1007/978-1-4757-2440-0
- Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann.

ÖZGEÇMİŞ

Kişisel Bilgiler

Soyadı, adı : DEMİR, Emre
 Uyuğu : T.C.
 Doğum tarihi ve yeri : 01.04.1995, İskenderun
 Medeni hali : Bekâr
 Telefon : 0 (535) 336 36 52
 e-mail : emredemir1995@gmail.com

Eğitim

Derece	Eğitim Birimi	Mezuniyet Tarihi
Yüksek lisans	İskenderun Teknik Üniversitesi / Elektrik - Elektronik Mühendisliği	2020
Lisans	İskenderun Teknik Üniversitesi / Bilgisayar Mühendisliği	2017
Lise	İskenderun Ticaret ve Sanayi Odası Anadolu Lisesi	2013

Yabancı Dil

İngilizce

Yayınlar

Demir, E. & Gökçen, A. (2019), Empirical Wavelet Transform Analysis and Classification of Multichannel Lung Sounds. *2nd International Mersin Symposium*, Cilt 6, 218-226.

Demir, E., Gökçen, A., Altan G. & Kutlu Y. (2019), The Effect of Nonlinear Classifiers On the Diagnosis of Respiratory Diseases Using Empirical Wavelet Transform Based Features. *3rd International Academic Studies Conference*, 178-185.

Demir, E. & Gökçen, A. (2018), Diagnosis of Chronic Obstructive Pulmonary Disease using Empirical Wavelet Transform Analysis from Auscultation Sounds. *International Conference on Artificial Intelligence towards Industry 4.0*, Abstract, 37-38.

DİZİN

A

AdaBoost · 4, 22, 28, 31, 37, 42, 43
 ADD · xi, 4, 5, 6, 8, 9, 11, 16, 17, 18, 19, 31, 42, 44
 akciğer · 4, 6, 7, 8, 12, 13, 16, 17, 18, 19, 20, 31, 42, 47, 48
 Akciğer oskültasyonu · 1, 31
 akciğer sesleri · 6, 16, 20
 Ampirik Dalgacık Dönüşümü · xi, 4, 16
 analiz · 1, 3, 5, 6, 8, 10, 11, 16, 25, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 44
 ayrıştırma teknikleri · 6

B

başarım · 4, 7, 8, 30, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 43, 45
 Bayes · 4, 22, 24, 25, 31, 33, 42, 48
 BayesNet · 4, 22, 25, 31, 34, 42
 bilgisayar destekli tanı sistemleri · 2, 3, 6, 22
 biyomedikal · iv, 2, 6, 8, 10, 22, 45

C

Chronic Obstructive Pulmonary Disease · 46, 50
 COPD · v, 47

Ç

Çok katmanlı algılayıcı · 26, 27, 35
 Çok Katmanlı Algılayıcılar · xi, 22, 31
 Çok Katmanlı Yapay Sinir Ağları · 4

D

dalgacık · iv, 4, 16
 Destek Vektör Makineleri · xi, 4, 22, 27, 31, 36, 42

E

Empirical Wavelet Transform · 16, 46, 47, 48, 49, 50
 En Yakın k Komşu · xi, 4, 22, 23, 31

F

frekans · 6, 8, 16

G

göğüs filmi · 13
 göğüs hastalıkları · 15
 görüntü işleme · 9, 20, 22

H

hassasiyet · 7, 30, 32, 33, 34, 35, 36, 37, 38, 39, 40, 43
 hastalık · 1, 4, 6, 31
 Hızlı Fourier Dönüşümü · xi, 7

İ

istatistiksel öznitelikler · 4, 31

J

J48 · 4, 22, 29, 31, 40, 42

K

kalp · 1, 4, 7, 13, 20, 31
 karar · 10, 23, 27, 28, 29
 karar ağacı · 28, 29
 karışıklık matrisi · 29, 30
 klinik tanı ve tedavi · 2, 4, 6, 42
 k-NN · xi, 4, 8, 23, 24, 32, 42
 KOAH · iv, xi, 3, 4, 5, 6, 11, 12, 17, 31, 32, 42, 44, 48
 kronik · xi, 3, 6, 11, 12, 48
 Kronik Obstrüktif Akciğer Hastalığı · xi, 3, 6, 11

L

Littmann 3200 Digital Stethoscope · 1, 2
 lung sounds · 46, 48

M

makine öğrenmesi · 20, 23, 24, 27
 matematiksel hesaplamalar · 16, 31
 matematiksel modelleme · 26
 mesafe ölçütleri · 23
 modlar · 4, 31
 Multilayer Perceptron · 26

N

Naïve Bayes · 24, 33
 nefes darlığı · 12

O

olasılık · 24
 oskültasyon · 1, 3, 7, 13, 14, 15, 16, 17, 20, 31

Ö

öksürük · 12
 ölçeklendirme fonksiyonu · 16
 ön işlemler · 4, 16, 17, 20, 31
 örüntü tanıma · iv, 10, 20, 27
 özgüllük · 7, 30, 32, 33, 34, 35, 36, 37, 38, 39, 40, 43
 Öznitelik çıkarma · 10, 20
 öznitelik vektörleri · 20, 22, 25, 26, 27, 28, 29, 31

P

performans · 5, 11, 24, 25, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 43, 44
 performans analizi · 25, 27, 28, 29, 31

R

Random Forest · 4, 22, 28, 29, 31, 39, 42
 Random SubSpace · 4, 28
 regresyon ağacı · 28
 regresyon analizi · 27
 RespiratoryDatabase@TR · iv, v, 4, 13, 14, 20, 31
 risk faktörleri · 12

S

sağlık · 2, 3, 42
 ses · 1, 13, 17, 18, 19, 20, 31
 SGRQ-C · 13
 sınıflandırma · 4, 5, 6, 7, 8, 9, 10, 11, 13, 20, 22, 23, 24, 25, 26, 27, 28, 29, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 45

sınıflandırma doğruluk tablosu · 29
 Sigara bağımlılığı · 12
 sinyal işleme teknikleri · 2, 6, 10, 11, 13
 sinyaller · 3, 4, 8, 9, 10, 16, 42
 sistemler · 6, 22, 44
 solunum · 1, 3, 4, 6, 7, 11, 12, 13, 31
 solunum fonksiyon testi · 13
 solunum hastalıkları · 4
 Solunum seslerinin analizi · 6
 stetoskop · 1, 14

T

tanı ve tedavi · 3, 4
 tedavi · 2, 3, 4, 6, 12, 47, 48
 teknoloji · 2, 42
 teşhis · 1, 3, 6, 12, 13, 31

U

uyarlanabilir · 4, 16

V

veri · 1, 4, 5, 8, 10, 11, 13, 14, 20, 22, 23, 24, 26, 29, 31, 42, 44
 veri analizi · 10, 20, 42
 veri toplama · 10

Y

Yapay Sinir Ağları · 6
 yapay zekâ · 4, 6, 13, 22



TEKNOVERSİTE



teknoversite **AYRICALIĞINDASINIZ**

İSTE

