



T.C.
DOĞUŞ ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
ELEKTRONİK ve HABERLEŞME MÜHENDİSLİĞİ
ANABİLİM DALI

MAKİNE ÖĞRENME ALGORİTMALARIYLA KONUŞMACI TANILAMASI

Yüksek Lisans Tezi

Tezi Hazırlayan
KORAY ÖZTÜRK

İstanbul, 2024



T.C.

DOĞUŞ ÜNİVERSİTESİ

LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

ELEKTRONİK ve HABERLEŞME MÜHENDİSLİĞİ

ANABİLİM DALI

MAKİNE ÖĞRENME ALGORİTMALARIYLA KONUŞMACI TANILAMASI

Yüksek Lisans Tezi

Tezi Hazırlayan

KORAY ÖZTÜRK

202191028002

0009-0002-6839-0650

Danışman

Dr. Öğr. Üyesi Temel SÖNMEZOCAK

İstanbul, 2024

YEMİN METNİ

Yüksek Lisans Tezi olarak sunduğum “Makine Öğrenme Algoritmalarıyla Konuşmacı Tanılaması” başlıklı bu çalışmanın, bilimsel ahlak ve geleneklere uygun şekilde tarafımdan yazıldığını, bu tezdeki bütün bilgileri akademik ve etik kurallar içinde elde ettiğimi, yararlandığım eserlerin tamamının kaynaklarda gösterildiğini ve çalışmamın içinde kullanıldıkları her yerde bunlara atıf yapıldığını, patent ve telif haklarını ihlal edici bir davranışımın olmadığını belirtir ve bunu onurumla doğrularım.

Koray ÖZTÜRK

15/09/2024



YÜKSEK LİSANS TEZ SINAV TUTANAĞI

Doküman No	FR.1.26
Yürürlük Tarihi	1.11.2017
Revizyon Tarihi	4.12.2020
Revizyon No	2
Sayfa	1 / 1

LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

Tarih : 24/09/2024

Anabilim/Anasanat Dalı : Elektronik ve Haberleşme Mühendisliği

Öğrencinin Adı Soyadı : Koray ŞİTÖRK

Öğrenci No : 202191028002

Tez Danışmanının Adı Soyadı : Dr. Öğr. Üyesi Temel SÖNMEZÖCAK

İkinci Tez Danışmanının Adı Soyadı : -

Tezin Başlığı : Makine Öğrenme Algoritmalarıyla Konuşmacı Tanılaması

Doğuş Üniversitesi Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliği'nin 33/4 addesi uyarınca yapılan değerlendirmeler sonunda;

☒ tezin kabul edilmesine

☐ tezde düzeltme verilmesine

☐ tezin reddedilmesine

oy birliği / oy çokluğu ile karar verilmiştir. Gereği için arz olunur.

Danışman Üye

Üye

Üye

Üye

Üye

Anabilim/Anasanat Dalı Başkanı Onayı:

Doğuş Üniversitesi ("Üniversite") olarak kişisel verilerinizin güvenliği hakkında azami hassasiyet göstermekteyiz. Üniversite olarak, Üniversite ile ilişkili tüm şahıslara alt her türlü kişisel verinin 6698 sayılı Kişisel Verilerin Korunması Kanunu ("KVKK")'na uygun olarak işlenmesine ve gizliliğinin sağlanmasına büyük önem vermekteyiz. Anayasası 20/3 ile tanınan "kişisel verilerinizin korunmasını isteme" hakkınızın bilincinde olarak, KVKK Kapsamında tanımlı "Veri Sorumlusu" sıfatıyla, kişisel verilerinizi mevzuata uygun şekilde işlemekteyiz. Daha geniş açıklama için Doğuş Üniversitesi <https://www.dogus.edu.tr/hakkimizda/kisisel-verilerin-korunmasi>

As a Doğuş University ("University"), we show maximum sensitivity about the security of your personal data. As a university, we attach great importance to the processing and confidentiality of all personal data of all individuals associated with the University in accordance with the Law No. 6698 on Protection of Personal Data ("KVKK"). Being aware of your right to "ask for the protection of your personal data" recognized by the constitution article 20/3, as "Data Controller" defined within the scope of KVKK, we process your personal data in accordance with the legislation. For a broader explanation you can visit <https://www.dogus.edu.tr/hakkimizda/kisisel-verilerin-korunmasi>

hakkımızda: about us kişisel-verilerin-korunmasi: protection of personal data

(Form No: FR-001, Revizyon Tarihi:30.09.2016, Revizyon No:00)

Adı ve Soyadı : Koray ÖZTÜRK
Danışmanı : Dr. Öğr. Üyesi. Temel SÖNMEZOCAK
Türü ve Tarihi : Yüksek Lisans Tezi /15.09.2024
Alanı : Elektronik ve Haberleşme Mühendisliği
Anahtar Kelimeler : Konuşma Tanıma, Makine Öğrenmesi, Özellik Çıkarımı, Yapay Sinir Ağları, Dalgacık Dönüşümü

ÖZET

MAKİNE ÖĞRENME ALGORİTMALARIYLA KONUŞMACI TANILAMASI

Teknolojinin her geçen gün sürekli gelişmesiyle birlikte, hayatımıza giren otomatik sistem ve cihazların sayısı her geçen gün artmaktadır. Bu yeni teknolojiler ile kontrol edilen sistem ve cihazların güvenliği, kontrolü, sevk ve idare edilmesi gibi durumlar için ayırt edici farklı kontrol mekanizmalarına ihtiyaç duyulmuştur. Bu arayış içerisinde kişiden kişiye değişen, parmak izi tanıma, retina tanıma, yüz tanıma, konuşma tanıma gibi karakteristik özellikleri ayırt etmeye yarayan yöntemler önem kazanmıştır.

Bu çalışmada yaşları 18 ile 39 arasında değişen 10 farklı konuşmacıdan alınan konuşma örnekleri ile bir veri seti oluşturulmuştur. Her konuşmacıya bir donanım aracılığıyla, 25 farklı kelime teker teker söylenilerek kaydedilmiştir. Son bir kayıttaki 25 kelimenin art arda söylenmesi ile oluşturulmuştur. Böylece kişi başı 26 farklı kayıttan oluşan 260 farklı veri elde edilmiştir. Her verinin Matlab programı ile zaman, frekans ve diğer özelliklerine bağlı değerleri çıkarılmıştır. Çıkarılan bu özelliklerin her birinin tanınması istenilen kullanıcı ile olan ilişkisi Pearson korelasyon analizi ile incelenmiştir. Pearson korelasyon analizi sonucu çıkış ile en ilişkili özellikler belirlenerek bir sıralama yapılmıştır. Sonrasında bu sıralamaya göre 14 özelliğin hepsinin, ilk beşinin ve ilk ikisinin kullanılması ile tablolar oluşturulmuştur. Ardından Matlab programının dalgacık dönüşümü arayüzü kullanılarak aynı veri seti için dalgacık dönüşümü analizi gerçekleştirilmiştir. Makine öğrenme algoritmaları bu sürece kadar elde edilen verilerin tamamı ile eğitilerek konuşmacıya yönelik bir tahminde bulunma sürecine tabi tutulmuştur. Bu süreç sonunda farklı makine öğrenme algoritmalarının ses özellik çıkarımı ve dalgacık dönüşümü teknikleriyle eğitimleri sonucu yüzdesel başarı performansları analiz edilerek yorumlanmıştır.

Elde edilen sonuçlara göre özellik sayısı azaldıkça tahminlerin doğruluk oranı da azalmaktadır. Dalgacık dönüşüm ile yapılan analizler sonucu elde edilen verilerde, entropi değerinin de bir özellik sayısı olarak düşünülebileceği ortaya çıkmıştır.

Name and Surname : Koray ÖZTÜRK
Supervisor : Assist. Prof. Temel SÖNMEZOCAK
Degree and Date : Master's Thesis - 15.09.2024
Major : Electronics and Communication Engineering
Key Words : Speech Recognition, Machine Learning, Feature
Extraction, Artificial Neural Networks, Wavelet
Transform.

ABSTRACT

SPEAKER IDENTIFICATION WITH MACHINE LEARNING ALGORITHMS

With the continuous development of technology, the number of automatic systems and devices entering our lives is increasing day by day. Different control mechanisms are needed for the security, control, dispatch and management of systems and devices controlled by these new technologies. In this search, methods that help distinguish characteristic features such as fingerprint recognition, retina recognition, face recognition, and speech recognition, which vary from person to person, have gained importance.

In this study, a data set was created with speech samples taken from 10 different speakers aged between 18 and 39. Each speaker was recorded using equipment to say 25 different words one by one. A final recording was created by saying these 25 words consecutively. Thus, 260 different data consisting of 26 different records per person were obtained. Values of each data based on time, frequency and other characteristics were extracted with the Matlab program. The relationship of each of these extracted features with the desired user was examined with Pearson correlation analysis. As a result of Pearson correlation analysis, the features most associated with the output were determined and a ranking was made. Afterwards, tables were created using all 14 features, the first five and the first two, according to this order. Then, wavelet transform analysis was performed for the same data set using the wavelet transform interface of the Matlab program. Machine learning algorithms have been trained with all the data obtained so far and subjected to a process of making a prediction about the speaker. At the end of this process, the percentage success performances of different machine learning algorithms, as a result of their training with sound feature extraction and wavelet transform techniques, were analyzed and interpreted.

According to the results obtained, as the number of features decreases, the accuracy of the predictions also decreases. In the data obtained as a result of the analysis made with wavelet transform, it was revealed that the entropy value can also be considered as a feature number.

İÇİNDEKİLER

Sayfa No.

ÖZET

ABSTRACT

TABLolar LİSTESİ	iii
ŞEKİLLER LİSTESİ	iv
KISALTMALAR.....	v
GİRİŞ.....	1

BİRİNCİ BÖLÜM

MATERYAL VE YÖNTEMLER

1.1. Ses Dalgası.....	7
1.2. Ses Veri Seti	8
1.3. Ses Sinyallerinin Zaman ve Frekans Özellikleri	9
1.3.1. Ortalama Frekans (Mean Frequency-MNF)	10
1.3.2. Ortanca (Medyan) Frekans (Median Frequency-MDF).....	11
1.3.3. Ortalama Mutlak Genlik (Mean Absolute Value – MAV)	11
1.3.4. Ortalama Karekök (Root Mean Square - RMS)	11
1.3.5. Sıfır Geçiş Sayacı (Zero Crossing - ZC).....	12
1.3.6. Willison Genliği (Willison Amplitude - WAMP)	12
1.3.7. Dalga Boyu (Wavelength - WL).....	13
1.3.8. Eğim İşareti Değişimi (Slope Sign Change - SSC)	13
1.3.9. Log Dedektör (LOG)	13
1.3.10. Varyans (Variance - VAR)	14
1.3.11. Maksimum Fraktal Uzunluğu (Maximum Fractal Length – MFL)	14
1.3.12. Standart Sapma (Standard Deviation – SD).....	14
1.3.13. Mutlak Genlik Değişimi (Average Amplitude Change - AAC)	14
1.3.14. Mutlak Standart Sapma Farkı (Difference Absolute Standard Deviation - DASD).....	15
1.4. Pearson Korelasyon Analizi	15
1.5. Sınıflandırma Algoritmaları.....	16
1.5.1. Karar Ağacı Algoritması.....	16
1.5.2. Lineer Diskriminant Algoritması.....	18
1.5.3. Destek Vektör Makine Algoritması (Support Vector Machine – SVM)	18
1.5.4. Naive Bayes Algoritması	20
1.5.5. En Yakın Komşuluk Algoritması (K-Nearest Neighbors, KNN)	21
1.6. Yapay Sinir Ağları	22
1.7. YSA'nın Avantaj ve Dezavantajları	24
1.8. Yapay Sinir Hücresinin Yapısı	24
1.9. YSA Sınıfları	27
1.9.1. Eğiticişiz Öğrenme.....	27
1.9.2. Eğiticişli Öğrenme.....	28
1.9.3. Takviyeli Öğrenme	28
1.10. YSA'ların Eğitilmesi	29
1.10.1. YSA'lar için Öğrenme Algoritmaları	30

1.11. Sınıflandırmada Performans Kriterleri	30
1.12. Dalgacık Dönüşümü	34
1.12.1. Sürekli Dalgacık Dönüşümü (SDD)	36
1.12.2. Ayrık Dalgacık Dönüşümü (ADD).....	37

İKİNCİ BÖLÜM

ANALİZ BULGULARI

2.1. Ses Sinyaline Ait Genlik ve Frekans Özelliklerinin Analizleri	40
2.2. Ses Sinyalinin Dalgacık Dönüşümlerine Ait Analizler	46

SONUÇ	48
--------------------	-----------

KAYNAKÇA.....	50
----------------------	-----------

EKLER

EK-1. Etik Kurul Raporu.....	55
EK-2. Katılımcı Bilgi ve Onam Formu	66

TABLÖLAR LİSTESİ

	Sayfa No
Tablo 1. Veri Setindeki Konuşmacı Demografisi	9
Tablo 2. Karmaşıklık Matrisi.....	31
Tablo 3. Zamana Bağlı Genlik Frekans Özelliklerine Ait Bulguların Kişilere ait Karşılaştırmalı Sonuçları.....	41
Tablo 4. Pearson Korelasyon Analizi	43
Tablo 5. Zamana Bağlı Genlik Frekans Özelliklerine Ait Bulguların Karşılaştırmalı Sonuçları.....	45
Tablo 6. Entropi (db4) Özelliklerine Ait Bulguların Karşılaştırmalı Sonuçları	47

ŞEKİLLER LİSTESİ

Sayfa No.

Şekil 1. Ses Dalgasının Özellikleri (a) Küresel (b) Boyuna Ses Dalgaları	8
Şekil 2. Dikey Eksende Genlik Yatay Eksende Zamanın Gösterildiği Ses Dalgası Fiziksel Özellikleri	10
Şekil 3. Örnek Bir Karar Ağacı Yapısı.....	17
Şekil 4. Destek Vektör Makine Sınıflandırma Grafiği Örneği.....	19
Şekil 5. KNN Algoritması Çalışma Akış Diyagramı.	22
Şekil 6. Örnek Bir YSA Yapısı	23
Şekil 7. Sinir Hücresi Yapısının Biyolojik Gösterimi	25
Şekil 8. Yapay Sinir Hücresi	26
Şekil 9. Yapay Sinir Ağlarının Sınıflandırılması	27
Şekil 10. Eğiticişiz Öğrenme	28
Şekil 11. Eğiticişiz Öğrenme Yapısı.....	28
Şekil 12. Takviyeli Öğrenme Yapısı	29
Şekil 13. ROC Eğrisi	33
Şekil 14. ROC Eğrisi ve AUC Değeri İlişkisi	34
Şekil 15. Sinyal Analiz Yöntemleri Arasındaki İlişki	35
Şekil 16. Her Bir Deneme için Kişilere Göre Sınıflama Sıralamaları.....	41
Şekil 17. 14 Özellik ile Elde Edilen Maksimum ve Minimum Oranlar.	42
Şekil 18. 5 Özellik ile Elde Edilen Maksimum ve Minimum Oranlar.	44
Şekil 19. 2 Özellik ile Elde Edilen Maksimum ve Minimum Oranlar.	45
Şekil 20. Örnek Bir Sese Ait Dalgacık Shannon Entropi Analizi.	46
Şekil 21. Entropi (db4) ile Elde Edilen Maksimum ve Minimum Oranlar.	47

KISALTMALAR

AAC:	Ortalama Genlik Değişimi (Average Amplitude Change)
CVA:	Ortak Vektör Analizi (Common Vector Analysis)
ÇKA:	Çok Katmanlı Algılayıcı
DASD:	Fark Mutlak Standart Sapma (Difference Absolute Standard Deviation)
DVM:	Destek Vektör Makinesi
DZB:	Dinamik Zaman Bükmesi
DT:	Karar Ağacı (Decision Tree)
HMM:	Saklı Markov Modeli (Hidden Markov Model)
Hz:	Hertz
KNN:	En Yakın Komşuluk (K-Nearest Neighbour)
LD:	Lineer Diskriminant
LOG:	Log Dedektör
MAV:	Ortalama Mutlak Genlik (Mean Absolute Value)
MDF:	Medyan Frekansı (Median Frequency)
MFL:	Maksimum Fraktal Uzunluğu (Maximum Fractal Length)
MNF:	Ortalama Frekans (Mean Frequency)
NB:	Naive Bayes
RMS:	Ortalama Karekök (Root Mean Square)
RNN:	Tekrarlayan Sinir Ağları (Recurrent Neural Networks)
SBS	Ses Basınç Seviyesi
SD:	Standart Sapma (Standard Deviation)
SMM:	Saklı Markov Modeli
SSC:	Eğim İşareti Değişikliği (Slope Sign Change)
SVM:	Destek Vektör Makine (Support Vector Machine)
VAR:	Varyans (Variance)
WAMP:	Willison Genliği (Willison Amplitude)
WL:	Dalgaboyu (Wavelength)
YSA:	Yapay Sinir Ağı
ZC:	Sıfır Geçiş (Zero Crossing)

GİRİŞ

Hayatımızı kolaylaştıran teknolojilerin kullanımı yaşamın her alanında giderek yaygınlaşmaktadır. Bu hızlı gelişim sayesinde hayatımıza hemen her gün farklı teknolojiler ve cihazlar entegre olmaktadır. Bu teknoloji ve cihazların kontrol edilmesi gerekliliği insanları çeşitli arayışlara yönlendirmiştir. Bu arayışa çözüm olabilecek yöntemlerin en başlarında gelen tekniklerden birisi de bu teknolojileri ses veya konuşma tanıma teknikleri ile kontrol etmektir.

Konuşma tanıma; bir makinenin veya programın, insan sesini algılayarak makine diline çevirmesi ve bununla birlikte sesi belirli işlemlerden geçirerek sesin sahibini tahmin etme işlemi olarak yorumlanabilir. Konuşma tanıma ile ilgili çalışmaların tarihteki ilk uygulamalarında sadece sayılar üzerine yoğunlaşmıştır. 1952 yılında ‘Audrey’ isimli uygulama, sadece rakamları algılayabilecek şekilde Bell Laboratuvarları tarafından geliştirilmiştir [1].

Bir başka tanımla ses tanıma teknolojisi; dil bilimi, bilgisayar bilimleri ve elektronik gibi farklı alanları içinde barındıran, konuşma dilini bilgisayar üzerinden bir mikrofon ya da farklı bir kayıt cihazı ile sesin alınıp kaydedildikten sonra işlenmesi ve yorumlanmasını inceleyen bir çalışma alanı olarak tanımlanmaktadır [2].

Konuşma tanıma sistemleri farklı alanlara ayrılmaktadır. Bu alanlar sistemlerin yapıları ve yöntemlerine göre belirlenmektedir. Konuşmacıya ve ses analizine göre şekillendirilen bu alanlar, konuşmacıya göre sınıflandırmada konuşmacıya bağımlı ve konuşmacıdan bağımsız, ses analizine göre sınıflandırmada ise kelime tabanlı ya da fonem olarak belirlenmiştir. Sınıflandırmalar sesin özellikleri zamanda sürekli olarak tamamı ya da ayırık olarak öbekler şeklinde belirlenmektedir. Bu sınıflandırmada konuşmacıya bağımlı olan sistemlerde sadece bir kişinin sesine göre tanıma işlemini gerçekleştirmektedir. Konuşmacıdan bağımsız sistemlerde ise konuşma tanıma işlemi herhangi bir kişiye göre yapılabilir.

Şayet ses analizi fonem tabanlı tercih edildiğinde, karşılaştırma yapılacak ve tanınması gereken en küçük öbek, fonemdir. Fonem tabanlı tercih edilirse, bu durumda tanınması gereken en küçük öbek kelime olmaktadır. Ayırık ya da sürekli sistemlerdeki tanıma ise tamamen sistemin hedefine bağlıdır.

Hedef kısa aralıklardan oluşan konuşmaların ya da seslerin tanınması ise ayrık tabanlı sistem, sürekli devam eden konuşmalardan tanıma isteniyorsa sürekli tabanlı sistem olarak nitelendirilir [3].

Konuşma tanıma sistemleri; ağızdan çıkan sesin elektronik ortama mikrofon gibi bir donanım aracılığıyla aktarılması ile başlamaktadır. Ardından bu ses sinyali, sinyal işleme teknikleri ve matematiksel yöntemlerle sayısal veri ya da öznitelik vektörü şekline dönüştürülür. Yeterli veri kaynağı olduğunda ise yine mikrofondan tanınması istenilen konuşma alınarak sayısal forma dönüştürülür ve önceden bu forma dönüştürülmüş verilerle karşılaştırılır. Bu işlemler sonrasında karşılaştırmalı verilerin sonucu eşleşir ve konuşma tanıma işlemi tamamlanmış olur [4].

Günümüz şartlarında ses ve konuşma tanıma teknolojileri ile ilgili gelişime çok fazla ihtiyaç duyulmaktadır. Bu durumun nedeni sesin tıpkı parmak izi gibi karakteristik bir özellik olması sayesinde kullanıcı ayırt edici sistemlerin yönetimine olanak sağlamasındaki kolaylıktır. Bu sayede ticari açıdan büyük bir pazara hitap eden sistemlerdir. Ancak sesin çeşitlilik göstermesi, konuşulan dillerin farklılığı, cümle ve dil bilgisi yapısı gibi sebepler nedeniyle oldukça zor bir çalışma disiplini gerektirmektedir. Konuşma tanıma sistemlerinin temelleri elektronik, bilgisayar bilimi, dil bilimi, yapay zekâ, matematik ve robot bilimlerine dayanmaktadır. Tüm bu alanlardaki gelişmeler geçmiş yıllarda çok yeterli düzeyde olmadığından konuşma tanıma teknolojisinde tatmin edici gelişmeler kaydedilememiştir. Bu teknolojinin gelişmesindeki en büyük engellerden biri de insan faktörüdür. Yeryüzünde konuşulan diller, şiveler, vurgu ve tonlamalar, yaş ve cinsiyet, gibi çok farklı bileşenleri içerisinde barındıran bir sistemi tasarlamak oldukça zordur. Yapay zekâ, makine öğrenmesi ve değerini öğrenme kavramlarının ortaya çıkmasıyla bu teknolojilerdeki gelişmeler tarih boyunca hiç olmadığı kadar hatırı sayılır düzeyde mesafe kat etmiştir. Verilerin toplanması ve ön işleme adımları bu kavramların ortaya çıkmalarıyla beraber oldukça küçültülmüş ve bir sistematığa bağlanmıştır [5].

Literatür incelemesinde ses tanıma sistemleri ile ilgili yapılmış çeşitli çalışmalar bulunmaktadır. Özkan ve çalışma arkadaşları yaptıkları çalışmada ses sinyalinin spektral yöntemler kullanarak sinyal gürültülerini Tekil Değer Ayrışımı (TDA) ile filtreleyebilen yeni bir yöntem geliştirmiştir [6].

Çakır ve arkadaşlarının 2017 yılında yapmış olduğu çalışmada eş zamanlı ve yüksek kaliteli ses tanıma uygulaması geliştirmiştir. Çalışma kapsamında her konuşmanın sisteme metin etiketlenmesi hedeflenmiştir. Konuşmanın tamamı özellik çıkarımı ve sınıflandırma süreçlerinden geçirilerek tanınmaktadır. Aşamaların seçim süreci daha önce yapılmış çalışmalar daki en kullanışlı tekniklerin belirlenmesi ile gerçekleştirilmiştir. Bu çalışmada kullanılan teknikler Mel Frekanslı Kepstral Katsayılar (MFCC), Vektör Nicemleme, Saklı Markov Modelleri (SMM) modeli ile Özellik çıkarım sürecinde belirlenen teknikten daha verimli sonuçlar elde edilmesi adına çeşitli karşılaştırmalar yapılmıştır En yüksek başarıyı konuşma tanımlanmasında %92,6 performansta doğruluk oranı elde etmişlerdir [7].

SMM tabanlı başka bir çalışma Prakoso ve arkadaşları tarafından 2016 yılında yapılmıştır. Bu çalışmada akustik modelin kelimelerdeki doğruluğunun en iyi başarı ortalamasının %86 olduğunu tespit etmişlerdir [8].

Konuşma tanılması için insan sesinin karakteristik özellikleri kişiden kişiye göre değişmektedir. Lexutan 'ın 2015 yılında yaptığı bir çalışmada bu karakteristik özelliklerin elde edilmesinde zamana bağlı genlik ve frekans özellikleri (Ortalama Genlik, Standart Sapma, Maksimum, Minimum ve Medyan frekans değerleri, regresyon katsayıları v.b. özelliklerini, Karar Ağaçları (Decision Tree-DT) ve Destek Vektör Makine (DVM) ile birlikte kullanarak elde etmiştir. Tanılama oranı başarılı 0,69 ile 0,84 aralığında değiştiği gözlenmiştir [9].

Konuşmada insanın duygu durumu da önemlidir. Hemanta Kumar Palo ve arkadaşları 2018 yılında yaptığı çalışmada, giriş konuşma sinyalinin kaydı sırasında kimliği doğrulanmış konuşmacının duygusunu tanımak için dalgacık tabanlı bir özellik kombinasyonu ile %90,55 doğrulukla tasarlamıştır [10].

Konuşmada insan duygu durumuna ait tanılamada Wei-Hua Cao ve arkadaşları Rastgele Orman algoritması ile sesin akustik özelliklerini türetmiş ve de SVM ile türetilen bu özellikler %78,6'lık bir performans ile sınıflandırılmıştır [11].

İnsan sesinin duygu durumu insan makine etkileşiminde veya güvenlik tabanlı uygulamalarda önem arz etmektedir. Bu sebeple Rajasekar ve arkadaşları 2018 yılında duygu durumunun otomatik olarak tanımlanmasına Mel-Frequency Cepstral Coefficients (MFCC) ile birlikte SVM kullanmışlardır [12].

Yakar ve arkadaşlarının 2016 yılında yaptığı bir çalışmada ise hece ve sözcük tabanlı tanıma sistemini Türkçe dilinde geliştirmişlerdir. Eğitim ve test aşamasında, Dinamik Zaman Bükmesi (DZB), Çok Katmanlı Algılayıcı (ÇKA), Destek Vektör Makinesi (DVM) ve Saklı Markov Modeli (SMM) kullanılmıştır. ÇKA, DVM ve SMM yöntemleriyle her hece ve sözcük için modeller oluşturulmuştur ve buna göre tanıma gerçekleştirilmiştir. Kısaca vektör şeklinde nitelenen hece ve sözcükler ÇKA, DVM ve SMM metotları ile eğitilip akabinde tanıma işlemi gerçekleştirilmiştir. Hece tabanlı sistemlerde en iyi doğrulama %94,2 ile DZB; sözcük tabanlı sistemlerde ise en iyi doğrulama %96 ile yine DZB yöntemi ile elde edilmiştir [13].

Yapılmış bir başka çalışma ise gürültülü ortamlarda konuşma tanıma uygulaması geliştirmeye yöneliktir. Mikrofon ile gürültülü ortamlarda alınan konuşma örnekleri ilk olarak kaydedilmiş ve gürültü temizleme işlemine tabi tutularak veri seti oluşturulmuştur. Bu veri seti Google'a ait bulut sistemine entegre edilerek konuşma tanıma sisteminin başarısı test edilmiştir. Bu sayede özellikle gürültülü ortamlarda, bulut sistemleri ve konuşma tanıma sistemlerinin entegre çalışmaları durumundaki performansları incelenmiştir. Çalışma neticesinde özellikle mobil cihazlar için bulut sistemleri ile konuşma tanıma gerçekleştirilebilmesi için gürültünün oldukça az seviyede ya da gerçek zamanlı olarak temizlenmiş olması gerekmekte olduğu sonucuna varılmıştır [14].

Literatürde MFCC katsayıları kullanılarak konuşma ya da konuşmacı tanınması ayrıca yapılmaktadır. Charisma ve arkadaşları, konuşmacı tanıma modelinde MFCC'yi kullanırken %83,3'lük bir performansa ulaşmaktadır [15]. Winursito ve arkadaşları ise, MFCC tekniğindeki doğruluğa ulaşmadaki hataların üstesinden gelmek için MFCC'nin Temel Bileşen Analizi ile birleştirerek yeni bir model oluşturmuştur. Bu modelde %89,29'a kadar başarılı olmuştur [16].

Yayla tarafından 2018 yılında yapılan bir diğer çalışmada ise konuşma tanıma teknolojisinden faydalanılarak devre tasarımı ve devre analizi yapılmıştır. Sesli komutlar ile laboratuvar ortamında istenilen deneyin seçilebildiği, elde edilen sonuçlara göre analiz yapılabilen ve bu işlemlerin hepsinin sadece konuşarak gerçekleştirildiği bir çalışma elde edilmiştir. Bu sayede elektronik ile ilgili bölümlerin laboratuvar çalışmalarında kolaylık sağlanması hedeflenmiştir [17].

Sistem tasarımı yapılırken ses tanıma işlevi ile farklı sistemlerin çalıştırılması ve yönlendirilmesi gibi kontrol etme işlevlerinin de eklenmesi istenebilir. IBM'in yapay zekâ aracı Watson yapısında çok farklı yapay sinir ağı (YSA) modellerini barındırmaktadır. Böylece sesle verilen komutları çok yüksek oranda doğru anlayarak diğer sistemlerde gerekli tetiklenmelerin gerçekleştirilmesini sağlamaktadır. Watson Araştırma Merkezi'nde sesli komutların bu tarz sistemler üzerinde işlenmesine yönelik gerçekleştirilen çalışmalarda, derin öğrenme metotları ve SMM birlikte kullanıldığı metotlara benzer sonuçlar veren ve uçtan uca çalışan bir sesli komut algılama sistemi geliştirilmiştir. Sesin akustik ve dil modelinin oluşturulması ile tanıma işleminin gerçekleştirilmesi süreçlerinde RNN (Recurent Neural Networks – Tekrarlayan Sinir Ağları) yöntemi kullanılmıştır. Oluşturulan sistemin diğer hibrit modele göre çok daha hızlı ve verimli çalıştığı görülmüştür [18].

Literatürde yapılan araştırmalarda görüldüğü üzere insan seslerinin sözcük tabanlı farklı birçok model ele alınarak sözcüklerin anlamlandırılması, sözcüklerin algılanması, diğer bir deyişle kelime algılamaları, sınıflandırmaları şeklindedir.

Bu tez çalışmasında ise, literatürden farklı olarak kelime ya da sözcük algılamasının yanı sıra farklı kişilere ait konuşmalardan elde edilen farklı ses tonlarıyla kişi seslerini ayırt etmeye dayanan bir çalışmadır. Bu çalışma, özellikle ses taklit seslerin ayırt edilmesi, akıllı ev sistemleri için ev sahibinin sesini tanıma vb. Birçok uygulamalar için ele alınmıştır. Literatürden farklı olarak ses sinyallerine ait zamana bağlı genlik özellikleri, frekans spektrumlarından elde edilen frekans özellikleri olmak üzere toplamda 14 özellik fonksiyonları kullanılarak ses tanınması için sınıflandırma performansları incelenerek maksimum %99,2 doğruluk oranında en yüksek performans değeri elde edilmiştir. Bu özellikler Ortalama Frekans (Mean Frequency-MNF), Ortanca Frekans (Median Frequency-MDF), Ortalama Mutlak Genlik (Mean Absolute Value-MAV), Ortalama Karekök (Root Mean Square-RMS), Sıfır Geçiş Sayacı (Zero Crossing-ZC), Willison Genliği (Willison Amplitude-WAMP), Dalga Boyu (Wavelength-WL), Eğim İşareti Değişimi (Slope Sign Change-SSC), Log Dedektör (LOG), Varyans (VAR), Maksimum Fraktal Uzunluğu (Maximum Fractal Length-MFL), Standart Sapma (Standart Deviation-SD), Mutlak Genlik Değişimi (Average Amplitude-AAC), Mutlak Standart Sapma (Differance Absolute Standard Deviation-DASD) olmaktadır. Bu özelliklerden uygulamada daha hızlı ve etkin bir

performans elde etmek için ayrıca istatistiksel olarak Pearson korelasyon analizi yapılmış ve de en etkili özellikler ortaya konulmuştur. Ayrıca Dalgacık dönüşümü analizi ile sinyallere ait farklı ölçeklemelerden sinyallere ait alt dalgacıkların entropi özellikleriyle de tanılamada kullanılmak için performanslarına bakılmıştır.

Burada ses tanılaması için en etkin performansların elde edilmesi adına Yapay Sinir Ağları, En yakın komşuluk, Destek Vektör Makine, Naive Bayes, Karar Ağacı ve de Lineer Diskriminant olmak üzere toplamda 6 farklı makine öğrenmeleriyle birlikte ele alınmıştır.

Bu tez çalışmasında, 1. Bölüm’de, veri setlerinden ve de kullanılan bu verilerden elde edilen sinyallerin zamana bağlı genlik frekans özellikleri ve de kullanılan makine öğrenme algoritmalarından bahsedilmiştir. 2. Bölüm’de ise yapılan analizlere ait bulgular ve de performans özellikleri ele alınmıştır. 3. ve son bölümde ise elde edilen bulgular tartışılarak çalışmanın sonucu açıklanmıştır.

BİRİNCİ BÖLÜM

MATERYAL VE YÖNTEMLER

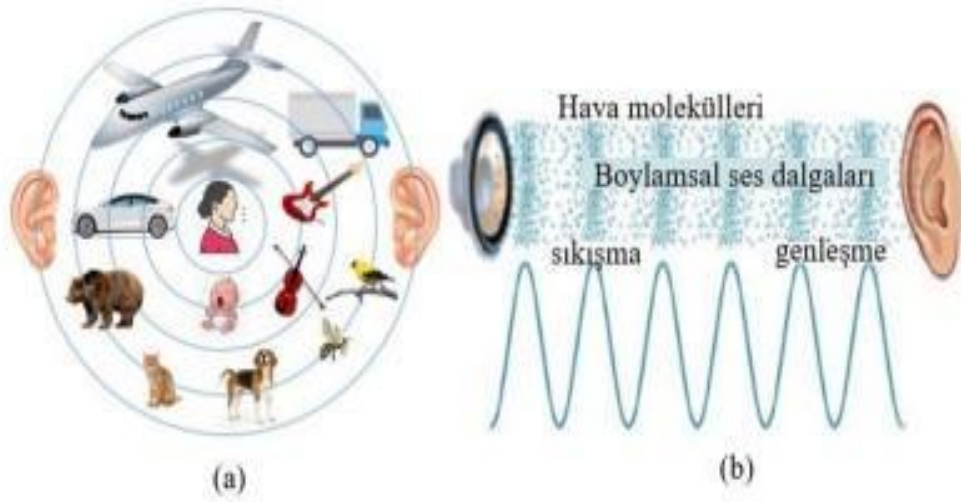
1.1. Ses Dalgası

İnsanın konuşması ve sesinin kişiden kişiye değişen biyometrik bir özellik olduğu bilinmektedir. Farklı konuşmacılara sıralı şekilde birebir aynı kelimelerin yer aldığı cümleler dahi söyletildiğinde, seslerin elektriksel hale dönüştürülmüş sinyal formları arasında çok büyük farklılıklar olduğu görülmektedir. Ses; yaş, duyu durumu, düşünce şekli, cinsiyet, sağlık, şive, aksan, yerel konum, sağlık, kayıt ortamı gibi birçok farklı parametreye göre değişkenlik gösterebilir. Konuşan kişi aynı olsa dahi farklı zaman dilimlerinde kaydedilen ses öbeklerine ait elektriksel sinyallerde farklılıklar gözlenebilir [19].

Konuşma sürekli doğrusal ve durağan formda olmayan akustik değişen bir dalga şeklidir. Akustik ise titreşim yayılım dalgaları ve ses dalgalarını inceleyen fizik dalıdır. Sesin fiziksel nitelikte ölçülebilen tüm özellikleri konuşma içinde geçerlidir [20].

Sesi oluşturan kaynak ise ses kaynağı olarak nitelendirilmektedir. Ses dalgası kaynaktan çıktıktan sonra enerjisi sayesinde içinden geçmekte olduğu katı, sıvı ya da gaz halindeki maddenin moleküllerini titreştirir ve parçacıkların çarpışması bir dalga hareketi oluşturur. Bu dalga hareketleri, boyuna yayılan ses dalgalarıdır. Konuşma sesleri de boyuna dalgalara örnek olarak gösterilebilir [21].

Herhangi bir kaynaktan gelen ses dalgaları, her yöne aynı oranda hareket eden küresel dalgalardır. Sonrasında ise Şekil 1’de görüldüğü gibi insan kulağı üzerinde fiziksel bir baskı oluştururlar. Bu fiziksel baskı sonucu elektriksel sinyaller şeklinde sinir hücreleri aracılığıyla aktararak anlamlandırılırlar.



Şekil 1. Ses Dalgasının Özellikleri (a) Küresel (b) Boyuna Ses Dalgaları

Kaynak : [20, 22].

1.2. Ses Veri Seti

İyi bir veri analizi yapılmasının ön koşulu, ihtiyaca yönelik doğru veri setinin oluşturulmasıdır. Satırlarında gözlem birimleri ve sütunlarında değişkenlerin bulunduğu bir matris olarak tanımlanabilecek veri setlerinde satır ve sütunların kesişim noktalarına hücre ismi verilmektedir. Bu hücrelere ölçme sonuçları sayısal veri ya da sembol olarak işlenmektedir. Veri işlenmeyen hücreler kayıp veri olarak nitelendirilir. Veri setinin hazırlanması öncesi çeşitli gözlem aşamaları gerçekleştirilmelidir. İlk olarak gözlemlenecek özelliklerin ve ilgili değişkenlerin çok iyi tanımlanması gerekir. Bu süreç kuramsal ve kavramsal çerçevenin çizilmesi açısından gereklidir. Sonrasında ise amaçlı gözlem yapılmaya başlanılarak ilgilenilen özelliğe yönelik verilere ulaşılması hedeflenmelidir.

Tez çalışmasının veri seti oluşturma sürecinde hazır veri setleri kullanımı yerine aynı gürültüsüz ortamda kaydedilen konuşma örneklerinden bir veri tabanı oluşturulmuştur. 18-39 yaş aralığında ve farklı cinsiyetteki 5'i Kadın, 5'i Erkek olmak üzere toplamda 10 farklı konuşmacıya 25 farklı kelime söylenmiştir. Bu kelimeler şunlardır; Ansiklopedi, Belli, Benzer, Benzin, Börek, Demokrasi, Dergâh, Elbise, Eldiven, Enteresan, Fevkalade, Hakem, Hoşçakal, Kendi, Lacivert, Mendil, Merhaba, Muhabbet, Mühendis, Portakal, Rüya, Şemsiye, Tolerans, Yarın, Zengin.

Tüm konuşmacılardan bu 25 farklı kelimenin art arda söylendiği, uzun bir konuşma şeklinde bir kayıta alınarak listeye dahil edilmiştir. Böylece 260 tane gözlem birimi ile veri seti oluşturma işlemi gerçekleştirilmiştir. Tüm Ses Kayıtları, 12th Gen Intel(R) Core(TM) i5-1240P 1.70 GHz işlemciye sahip bir laptop bilgisayar ile sestten yalıtılmış bir odada elde edilmiştir. Elde edilen kayıtlar Matlab yazılımı ile Machine Learning Toolbox kullanılarak analiz edilmiştir.

Tablo 1. Veri Setindeki Konuşmacı Demografisi

Konuşmacı No	Cinsiyet	Yaş
1	Erkek	18
2	Erkek	20
3	Kadın	26
4	Erkek	35
5	Kadın	36
6	Erkek	39
7	Erkek	23
8	Kadın	29
9	Kadın	38
10	Kadın	33

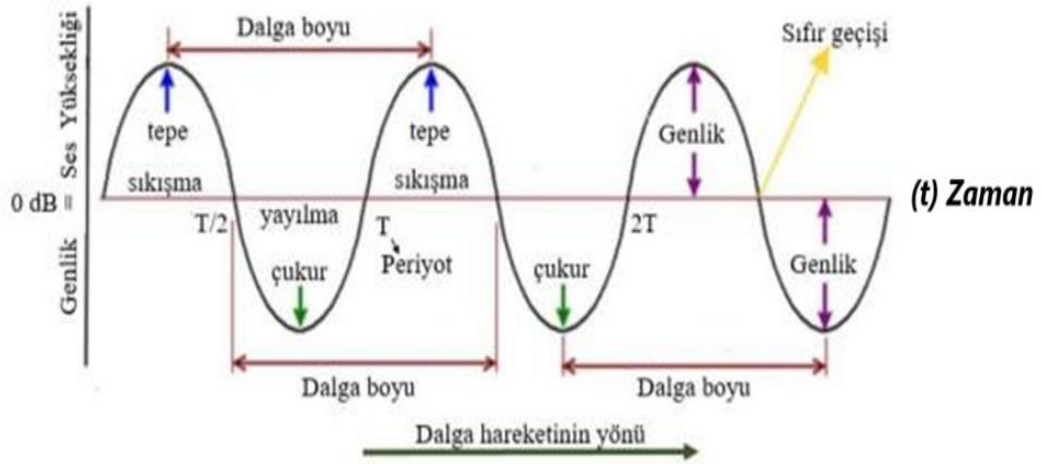
1.3. Ses Sinyallerinin Zaman ve Frekans Özellikleri

Konuşma, yeryüzünde yaşayan tüm canlılar arasında insana ait karakteristik bir özelliktir. Duyma organı olan kulağın hassasiyet eşikleri içerisindeki sınırlar dahilinde duyulabilen ses dalgaları konuşmanın temelini oluşturmaktadır ve bu dalgalar akustik dalgalardır. İnsan kulağı, frekans ve basınca bağlı olarak çevresel sesler, müzik ve konuşma seslerini duyabilmektedir [23]. Yeryüzündeki bu sesler, havadaki molekülleri titreştirerek periyodik bir basınç değişikliği yaratırlar. Bir saniyedeki titreşimlerin sayısı frekans olarak tanımlanır, birimi Hertz'dir (Hz). Her ses sinyalinin oluşma frekansı farklıdır, ses dalgasının ilerleyeceği maddenin titreşiminin hızına göre değişkenlik gösterir.

Bitişik olan iki tepe, sıfır geçiş noktaları ve çukur arasındaki mesafe ise dalga boyu olarak adlandırılmaktadır. Tam bir dalga boyunun olduğu zaman dilimi ise periyot olarak tanımlanmaktadır.

Genlik ise gürültü ya da yoğunluğu ifade eden dalgaların yükseklik değeri olarak tanımlanabilir. Şekil 2'de sese ait fiziksel özellikler gösterilmektedir [20].

Genlik (dB)



Şekil 2. Dikey Eksende Genlik Yatay Eksende Zamanın Gösterildiği Ses Dalgası Fiziksel Özellikleri

Kaynak: [22].

Ses Basınç Seviyesi (SBS), kulak zarı üzerinde oluşmuş mekanik basınca bağlı olarak değişen ses yoğunluğu olarak tanımlanabilir. Birimi Desibel'dir (dB). İnsan kulağı, 20-20.000 Hz (20 kHz) duyabileceği frekans aralığındaki 0-120 dB aralığındaki sesleri duyabilmektedir. Frekansı 20 Hz'den az olan ses dalgaları infrasonik, 20 kHz'den fazla olan ses dalgaları ultrasonik ses dalgaları olarak adlandırılmaktadır. Ultrasonik ses dalgaları köpekler ve yarasalar tarafından duyulabilmektedir. Deprem dalgaları, infrasonik ses dalgalarına örnek olarak gösterilebilir [23].

Bu çalışma kapsamında, sinyallerin frekans ve genlik özellikleri bakımından farklı parametreler ve bu parametreler ile ilişkilendirilmiş özellikler belirlenmiştir. Bu özellikler hiçbir dönüşüm işlemi gerekmeden hesaplanabilmektedir. Bu sayede gerçek zamanlı sinyaller kolaylıkla işlenebilmektedir. Bu özellikler kısaca açıklanmıştır [24,25].

1.3.1. Ortalama Frekans (Mean Frequency-MNF)

Sinyalin her frekanstaki güç spektrum yoğunluğu değerleri ile frekans çarpımlarının toplam güç spektrum yoğunluğuna bölümüdür. Bu şekilde ortalama frekans değeri elde edilir. Denklem (1)'de gösterildiği gibi hesaplanır [24].

$$MNF = \frac{\sum_{j=1}^M f_j P_j}{\sum_{j=1}^M P_j} \quad (1)$$

1 nolu denklemde f_j değeri, j frekans spektrumundaki frekans değerini, P_j değeri, aynı şekilde güç spektrumunu ifade etmektedir.

1.3.2. Ortanca (Medyan) Frekans (Median Frequency-MDF)

Spektrumun her iki yarısının tam ortasına denk gelen frekans değeridir. Spektrumun ortanca frekansı olarak da tanımlanmaktadır. Denklem (2)'deki gibi hesaplanır [24].

$$MDF = \frac{1}{2} \sum_{j=1}^M P_j \quad (2)$$

1.3.3. Ortalama Mutlak Genlik (Mean Absolute Value – MAV)

Sinyallere ait genlik değerleri ile ilgili önemli bilgiler yer almaktadır. Bir veri setindeki her bir elemanın, aritmetik ortalama değerden ne kadar uzak olduğunu gösteren bir ölçümdür.

Bir veri setindeki her bir elemanın pozitif değeri alınır ve ardından bu değerlerin aritmetik ortalaması hesaplanarak elde edilir. Denklem (3)'teki gibi hesaplanır. Bu denklemde X_i , N uzunluğundaki X sinyalinin i noktasındaki mutlak değeri anlamına gelmektedir [26].

$$MAV = \frac{1}{N} \sum_{i=1}^N |X_i| \quad (3)$$

Ortalama mutlak genlik (MAV), bir veri setinin dağılımının ne kadar değişken olduğunu ölçer. Yüksek MAV değerleri, veri setindeki elemanların ortalamadan ne kadar uzak olduğunu gösterir ve bu da veri setinin değişkenliğinin yüksek olduğu anlamına gelir. Düşük MAV değerleri ise veri setindeki elemanların ortalama daha yakın olduğunu gösterir ve bu nedenle değişkenliğin daha düşük olduğu anlamına gelir.

1.3.4. Ortalama Karekök (Root Mean Square - RMS)

Bir sinyalin genlik değerlerinin etkin değeri olarak tanımlanabilir. Denklem (4)'te ifade edildiği gibi örneğin veri setindeki değerlerden N elemanın RMS değeri sırasıyla; veri setindeki her bir elemanın karesinin alınması, bu karelerin toplamının

alınması, toplamın veri setindeki eleman sayısına bölünmesi ve son olarak bu değerin karekökünün alınması ile veya başka bir ifadeyle N uzunluğunda olan bir X sinyalinin, i noktasındaki karelerinin ortalamasının alınması ve bu değerin karekökünün alınmasıyla elde edilen bir değerdir. Özellikle değişken veri setlerinin genel bir ortalama değerini hesaplamak için kullanılır.

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^N X_i^2} \quad (4)$$

RMS, özellikle sinyal işleme, elektrik mühendisliği ve fizik gibi alanlarda kullanılan önemli bir ölçümdür. Bu ölçüm, bir sinyalin veya veri setinin genel enerji seviyesini veya şiddetini göstermek için kullanılmaktadır.

1.3.5. Sıfır Geçiş Sayacı (Zero Crossing - ZC)

Bu parametre, sinyale ait zaman aralığında belirlenen frekans bilgisinin bir ölçüsüdür. Tanımlanan bir eşik seviyesine göre bir sinyalin sıfır geçişi, sinyalin pozitif ve negatif değerler arasında değiştiği noktalardır. Yani sinyalin pozitif değerden negatif değere doğru giderken veya tersi durum söz konusu olduğunda yön değiştirmesi halinde sıfır değerine ulaştığı durumların tespit edilmesidir. Sıfır geçişlerinde gürültüden etkilenmemesi için bir eşik değeri kullanılmalıdır [27, 28].

Aynı zamanda ses sinyallerinin ve diğer sinyallerin özelliklerini analiz etmek için kullanılabilir. Örnek vermek gerekirse, bir müzik parçası incelendiğinde ZC özelliği, parçanın genel ritmi hakkında veya ritimli bir enstrümanın vuruş hızı hakkında bilgi sağlayabilir. Denklem (5)'te hesaplaması verilmiştir.

$$ZC = \sum_{i=1}^{N-1} [sgn(X_i * X_{i+1}) \cap |X_i - X_{i+1}| \geq \text{eşik değeri}]; \quad (5)$$

$$sgn(X) = \{1, \text{Eğer } X \geq \text{eşik değeri } 0, \text{ aksi takdirde}$$

1.3.6. Willison Genliği (Willison Amplitude - WAMP)

Bu parametre, sinyale ait frekans bilgisinin bir ölçüsüdür. Belirli bir zaman aralığındaki sinyalde belirli bir eşik seviyesi aşıldığında, genlik seviyeleri arasında meydana gelen farklardan dolayı oluşan tekrar sayılarının hesaplanması Denklem (6)'da gösterilmiştir.

$$WAMP = \sum_{i=1}^{N-1} [f(|X_i - X_{i+1}|)] \quad (6)$$

$$f(X) = \{1, \text{Eğer } X \geq \text{eşik değeri } 0, \text{ aksi takdirde}$$

1.3.7. Dalga Boyu (Wavelength - WL)

Belirlenen bir aralık boyunca dalgaların bir tepe noktasından diğer tepe noktasına olan uzunluğu diğer bir deyişle genliklerdeki toplam değişimleri açıklamaktadır.

$$WL = \frac{1}{N} \sum_{i=1}^{N-1} |X_{i+1} - X_i| \quad (7)$$

1.3.8. Eğim İşareti Değişimi (Slope Sign Change - SSC)

Sinyale ait frekans bilgilerini gösteren farklı bir yöntemdir. Sinyalin eğimi birkaç kez işaret değiştirdiğinde, sinyalde oluşan arka plan gürültüsünün önüne geçmek amacıyla üç ardışık segment arasındaki negatif ve pozitif eğimler, eşik fonksiyonu kullanılarak değişim sayısının bulunmasıdır.

Aynı zamanda matematiksel bir terim olan SSC bir fonksiyonun türevinin işaret değiştiği noktaları ifade etmektedir. Türevdeki işaret değişiklikleri SSC fonksiyonun orijinal grafiğinde eğimdeki değişim yerlerinin belirlenmesi konusunda yardımcı olur. Eğimdeki bu işaret değişiklikleri genelde fonksiyona ait maksimum-minimum noktalarını bulmak için kullanılmaktadır.

$$SSC = \sum_{i=2}^{N-1} [f(X_i - X_{i-1}) * (X_i - X_{i+1})]; \quad (8)$$

$$f(X) = \{1, \text{Eğer } X \geq \text{eşik değeri } 0, \text{ aksi takdirde}$$

1.3.9. Log Dedektör (LOG)

Kısa süreli bir sinyaldeki frekans bileşenlerini belirlemek için kullanılan bir işlemdir. Bu işlem, sinyalin belirli frekanslardaki bileşenlerini vurgulamak veya tespit etmek için kullanışlı olabilmektedir.

LOG işlemi, bir sinyali logaritmik bir ölçekte işleyerek, sinyalin genel frekans karakteristiğini analiz etmeyi sağlamaktadır. Denklem (9)'da hesaplanmaktadır.

$$LOG = e^{\frac{1}{N} \sum_{i=1}^N \log |X_i|} \quad (9)$$

1.3.10. Varyans (Variance - VAR)

Bu parametre Denklem (10)'da görüldüğü gibi değişken analog bir sinyaldeki sapmaların karesi alınarak elde edilen değerlerin ortalaması olarak hesaplanmaktadır [24].

Aynı zamanda veri setindeki değer dağılımının ne kadar yayıldığını ayrıca veri noktalarının ortalamadan ne kadar uzakta olduğunu gösterir. Denklem (10) kullanılarak hesaplanmaktadır.

$$VAR = \frac{1}{N-1} \sum_{i=1}^N X_i^2 \quad (10)$$

1.3.11. Maksimum Fraktal Uzunluğu (Maximum Fractal Length – MFL)

MFL, veri setinin en büyük fraktal boyutunu temsil etmektedir, yani veri setinin en karmaşık ölçeğini ifade etmektedir. Fraktaller, aynı desenlerdeki şekillerin farklı ölçek ya da boyutlarda tekrar ederek oluşturduğu geometrik şekiller olarak tanımlanır.

Maksimum fraktal uzunluğu ise en düşük sinyal uzunluğunun mutlak değerinin uyarı seviyelerindeki küçük değişiklikleri takip etmesiyle oluşur. Denklem (11)'deki gibi hesaplanır [29].

$$MFL = \log_{10} \sqrt{\sum_{n=1}^{N-1} (x(n+1) - x(n))^2} \quad (11)$$

1.3.12. Standart Sapma (Standard Deviation – SD)

Veri setindeki her bir değer birbirlerine olan uyumu ve yakınlığını gösteren değerdir. Her değer ortalama değerden ne kadar farklı olduğunu gösterir. Standart sapma küçük ise veriler ortalamaya yakın yerlerde, büyük ise ortalamadan uzak yerlerde dağılmışlardır. Tüm değerlerin aynı olması halinde standart sapma sıfırdır. Denklem (12)'deki gibi hesaplanır.

$$SD = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - x)^2} \quad (12)$$

1.3.13. Mutlak Genlik Değişimi (Average Amplitude Change - AAC)

Ses sinyallerinin özellik çıkarım sürecinde kullanılmakta olup belirli bir zaman aralığı boyunca genlikteki değişimi göstererek sesin farklılıklarının analizinde

kullanılmaktadır. Örneğin farklı konuşmacıların veya farklı seslerin genliklerini belirlemek için kullanılabilir. Aynı zamanda sinyalin dalga boyunun ortalamasının alınması dışında sinyalin dalga boyu (WL) ile eşdeğerdir. Denklem (13)'te gösterildiği şekilde hesaplanmaktadır.

$$AAC = \frac{1}{N} \sum_{i=1}^{N-1} |X_{i+1} - X_i| \quad (13)$$

1.3.14. Mutlak Standart Sapma Farkı (Difference Absolute Standard Deviation - DASD)

Mutlak standart sapma farkı, dalga boyunun standart sapma değeridir. Ayrıca RMS özelliğine benzerdir [30]. Denklem (14)'te hesaplanmaktadır.

$$DASDV = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N-1} (X_{i+1} - X_i)^2} \quad (14)$$

Bu tez çalışmasında 10 farklı konuşmacıdan alınan 25 farklı kelimenin her biri ayrı ayrı MATLAB programında işlenerek, yukarıda verilen özellikler çıkartılmıştır.

1.4. Pearson Korelasyon Analizi

Korelasyon, iki değişken arasındaki ilişkinin ölçülmesini içerir. Değişkenlerin pozitif veya negatif ilişkili olup olmadığını ya da birbirleri ile hiçbir şekilde ilişkili olmadığını belirler. Bir değişkendeki değişiklikler, diğer değişkendeki değişiklikleri etkiliyorsa iki değişken ilişkilidir. Değişkenler arasındaki ilişkiyi ölçmek için ilişki derecesini ifade etmek üzere korelasyon katsayılarını kullanırız. Başka bir deyişle, korelasyon katsayıları iki değişken arasındaki ilişkiyi veya ilişkinin gücünü (yön ve büyüklük) ölçer. Korelasyon katsayıları yüksek veya düşük (büyüklük) ve pozitif veya negatif (yön) olabilir. Korelasyon katsayıları -1 ile +1 arasında değişir: -1 ve +1 sırasıyla mükemmel negatif ve mükemmel pozitif korelasyon katsayılarını gösterirken, 0 korelasyon katsayısı korelasyon olmadığını (sıfır ilişki) ifade eder. Ayrıca, 0,40'ın altındaki korelasyon katsayılarının (negatif veya pozitif 0,40) düşük, 0,40 ile 0,60 arasındakilerin orta ve 0,60'ın üzerindeki yüksek olduğu söylenir [31, 32].

Bu çalışmada ses sinyallerine ait özelliklerin veri seti haline getirilmesi sürecinin akabinde çıkış değerleri ile tüm özellik parametreleri arasında bir korelasyon ilişkisi incelemesi gerçekleştirilmiştir. Bu süreçte mühendislik çalışmalarında iki

değişken arasındaki sürekli ilişkiyi incelemede kullanılması nedeniyle pearson korelasyon analizi tercih edilmiştir.

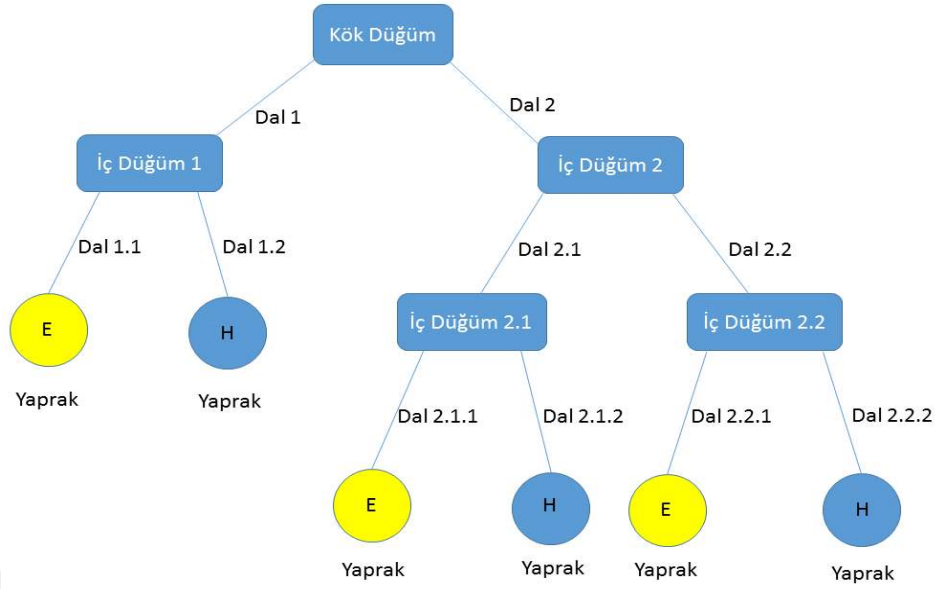
1.5. Sınıflandırma Algoritmaları

Bu tez çalışmasında konuşmacılara ait konuşma örneklerinden elde edilen sinyaller ile yapılan sınıflandırmada farklı sınıflandırıcılar kullanılmıştır. Kullanılan her sınıflandırıcı diğer sınıflandırıcılarla, tüm giriş özellikleri baz alınarak elde edilen performans verilerinin karşılaştırılmasıyla analiz edilmiştir. Kullanılan sınıflandırıcı algoritmalar, Lineer Diskriminant (LD), Karar Ağaçları (Decision Tree DT), En Yakın Komşuluk (K-Nearest Neighbour KNN), Naive Bayes (NB), Destek Vektör Makine (Support Vector Machine SVM) ve Yapay Sinir Ağları (YSA) olarak sıralanabilir. Bu algoritmaların çalışma ilkelerine ayrıca değinilecektir.

1.5.1. Karar Ağacı Algoritması

Karar ağacı (Decision Tree DT) algoritmasında yapılan sınıflandırma bir ağacın yapısı örnek alınarak gerçekleştirilmektedir. Kökten başlayan sınıflandırma, karar düğümlerine doğru ilerler, dal ve yapraklarda sonlanır. Kökten yapraklara doğru ilerlenirken karar verme aşamasında veri dizilerinin her birinin yarısı elenmektedir. Bu sayede çok hızlı bir şekilde sonuç elde edilmektedir. Düğüm sayısının N olması halinde verilecek karar sayısı $\log_2 N$ olur [33].

Karar sürecinde her düğümde verilen karar, bir sonraki dala ya da yaprağa ulaşır ve bu şekilde sonuca gidilir. Şekil 3 'te karar ağacı yapısına ilişkin bir örnek gösterilmiştir.



Şekil 3. Örnek Bir Karar Ağacı Yapısı

Kaynak: [28].

Şekil 3'teki gibi bir ağaç yapısı oluşturma sürecinde farklı algoritmalar geliştirilmiştir. Bunlardan ID3 algoritması karar düğümleri için, C4.5 algoritması ise sayısal özelliklere ait kararlar için geliştirilmişlerdir ve sistemdeki entropi hesabına dayalıdır. $H(T)$ fonksiyonu ise T kümesinin olasılıklarının toplamı olarak nitelenen sistemin entropisidir [34,35]. Denklem (15)'te matematiksel olarak belirtilmektedir.

$$H(T) = -\sum_{i=1}^n p_i \cdot \log_2 p_i \quad (15)$$

Denklem (15)'teki sınıf özelliğini temsil eden T değeri için entropi hesaplanabilmektedir. $H(T)$ ile ifade edilir.

Bu entropi hesaplanırken her özelliğin dahil olduğu sınıfa ait ağırlıklı ortalama dikkate alınarak denklem (16)'daki gibi hesaplanmaktadır. Bilgi kazancı ise T sınıfındaki her bir X Özelliğinin ağırlıklı ortalamalı entropisi ile entropi arasındaki fark olarak tanımlanmaktadır. Denklem (17)'de gösterilmektedir.

$$H(X, T) = -\sum_{i=1}^n \frac{|T_i|}{|T|} H(T_i) \quad (16)$$

$$\text{Bilgi Kazancı}(X, T) = H(T) - H(X, T) \quad (17)$$

Denklem (16)'da T sınıfındaki her X özelliğine ait ağırlıklı ortalamalı entropi belirtilmiştir. T sınıfının i değerindeki entropi değeri ise $H(T_i)$ olarak gösterilmiştir. Bilgi kazancı ise her bir özellik için hesaplanmaktadır.

En yüksek bilgi kazancı değerine sahip olan özellik belirlendikten sonra ağaç yapısının kökü oluşturulur. Karar ağaç yapısı oluşturma sürecinde kalan özellikler kullanılmak suretiyle diğer dal ve düğümler için de sürekli tekrarlanarak devam eder. En yüksek kazançla sahip özellik iç düğüm olarak belirlenir, alt düğüm ve yapraklara ulaşana dek devam eder. Karar ağaç yapısı bu şekilde elde edilmiş olur [34].

1.5.2. Lineer Diskriminant Algoritması

Problem çözme konusunda kullanışlı, oldukça basit ve başarısı yüksek bir algoritmadır. Yeni bir değişken oluşturur ve özelliklere ait doğrusal bir kombinasyon sağlar. Hedef olan iki sınıfın aralarındaki ayrım mesafesi hesaplanır, doğrusal kombinasyonlarıyla tahmin esasına dayandırılır. Mahalanobis mesafesi olarak ifade edilen mesafe, Denklem (18)'de gösterildiği şekilde hesaplanır [35, 36].

$$\Delta^2 = \beta^T (\mu_1 - \mu_2) \quad (18)$$

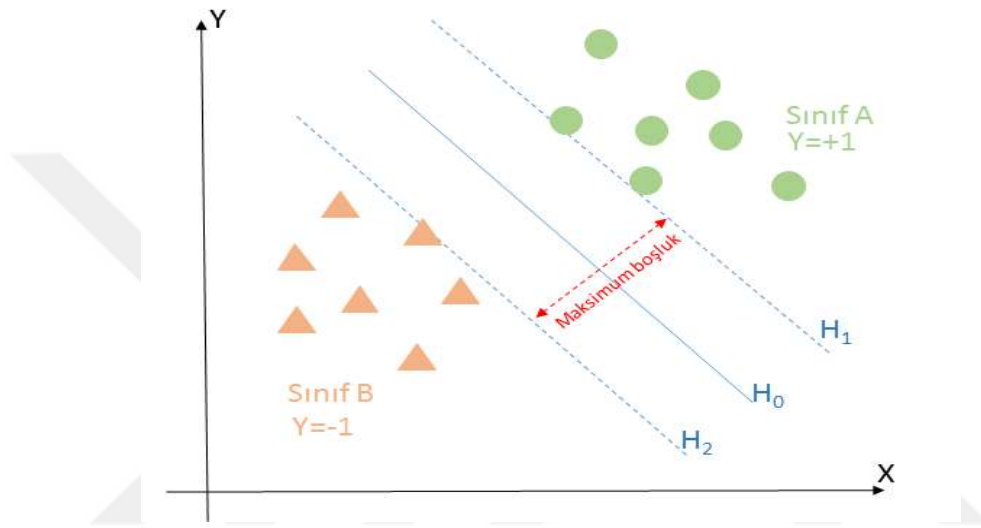
Denklem (18)'de Δ , ayrım ya da Mahalanobis mesafesi olarak tanımlanmaktadır. Aynı zamanda, β katsayı vektörünü, μ_1 ve μ_2 ise ortalama vektörleri ifade etmektedir.

1.5.3. Destek Vektör Makine Algoritması (Support Vector Machine – SVM)

Farklı iki sınıfta kümelenmiş örneklerin birbirlerinden ayrılabilceği en uygun fonksiyonu bularak sınıflandıran algoritmadır. Bir sınıflandırma algoritmasıdır. Sınıflardaki bu örnekler -1 ile $+1$ değerleri arasında etiketlenmektedirler. Örneklerden en yakın olan noktalar veri noktalarıdır. Bu noktalar sınıflar arası hiper düzlemi belirlemek için sınır bölge belirleme aşamasında kullanılmaktadır. Bahsi geçen veri noktaları, destek vektörleri diye adlandırılmaktadır. Sınır genişliğini maksimum yapan fonksiyon, hiper düzlem sınırı içerisinde belirlenmektedir. Veri böylece iki sınıf halinde kümelenmiş olur [37, 38].

Bu fonksiyon, sınır genişliğini maksimum yapmakla birlikte doğrusal olmayan kümelerde iki sınıf ayırıcısı olarak da kullanılabilir. Aynı zamanda lineer de olabilir [31, 39]. Doğrusal olmayan sınıflandırmalarda başarı performansının yüksek olması nedeniyle çok fazla makine öğrenmesi uygulamasında tercih edilmektedir [39].

Şekil 4'te destek vektör makinesinin sınıflandırma grafiği ile ilgili bir örnek verilmiştir. Grafikte aralığı maksimum yapan fonksiyon, H_0 hiper düzlem sınıflandırma fonksiyonu olarak görülmektedir.



Şekil 4. Destek Vektör Makine Sınıflandırma Grafiği Örneği

Kaynak: [38].

Örnek olarak gösterilen sınıflandırmada, H_0 aralığı maksimum yapan ve sınıfları ayıran hiper düzlem fonksiyonunun matematiksel gösterimi Denklem (19)'da gösterilmektedir.

$$H_0 = \sum_{i=1}^n W_i \cdot X_i + b = 0 \quad (19)$$

(19) numaralı denklemde n tane özellik için, W ağırlık vektörü ile ifade edilmektedir. Şekil 4'te gösterilen H_1 ve H_2 hiper düzlemleri, iki ayrı sınıfa ait örnekler arasındaki maksimum aralık mesafeleri olarak nitelendirilmektedir. H_1 ve H_2 düzlemlerine ait gösterimler (20) ve (21) numaralı denklemlerde gösterilmektedir.

$$H_1 = W^T \cdot X + b = +1 \quad (20)$$

$$H_2 = W^T \cdot X + b = -1 \quad (21)$$

Hiper düzlemleri ifade eden H_1 ve H_2 üzerindeki örnekler destek vektörler olarak tanımlanmaktadır. İki vektör arası mesafe olarak ifade edilebilecek d değeri ise (22) numaralı matematiksel eşitlikte gösterilmektedir.

$$d = \frac{|w^T \cdot x + b|}{\|w\|} \quad (22)$$

W vektörünün minimum değer alması halinde, hiper düzlemler olan H_1 ve H_2 arasındaki mesafe maksimum olmaktadır [38].

1.5.4. Naive Bayes Algoritması

İstatistiksel teknikleri kullanarak olasılık sınıflandırma işlemini gerçekleştirmektedir. Belirli bir sınıfa ait olan belirli bir örneğin, ait olma olasılığından yola çıkar ve verinin hangi sınıfta olduğunu tahmin eder. Yapılan çalışmalarda büyük veri setlerinin kullanılması halinde, sinir ağları ve karar ağaçlarına göre daha yüksek doğruluk performansı sergilediği gözlemlenmiştir [40, 41].

Koşullu olasılık kuramına göre, Thomas Bayes tarafından bulunmuş olan bu algorithmada, herhangi bir X örneğinin C sınıfına ait olma olasılığı $P(C|X)$ şeklinde ifade Denklem (23)'deki gibi ifade edilmektedir [39, 34].

$$P(X) = \frac{P(C) \cdot P(C)}{P(X)} \quad (23)$$

Naive Bayes sınıflandırıcı kullanılırken, koşullu olasılık hesaplamalarında her örneğe ait X değerinin birbirinden bağımsız olduğu varsayılarak (24)'deki denklem kullanılmaktadır.

$$P(X) = \prod_{k=1}^n P(C) \cdot P(C) P \quad (24)$$

Naive Bayes sınıflandırıcı kullanımı, konuşma tanıma, kelime tanıma, hastalık teşhisi, şifre kontrolü gibi birçok alanda tercih edilmektedir.

Özellikle büyük veri setlerindeki yüksek sınıflandırma başarısı, üstün bir performansa sahip olması, hızlı ve kolay uygulanabilirliği bu algoritmanın en önemli avantajlarıdır. Dezavantajları ise, dinamik sistemlerde eğitimin defalarca yapılması ve her özellik için eşit değerler verilmesi olarak sıralanabilir [42].

1.5.5. En Yakın Komşuluk Algoritması (K-Nearest Neighbors, KNN)

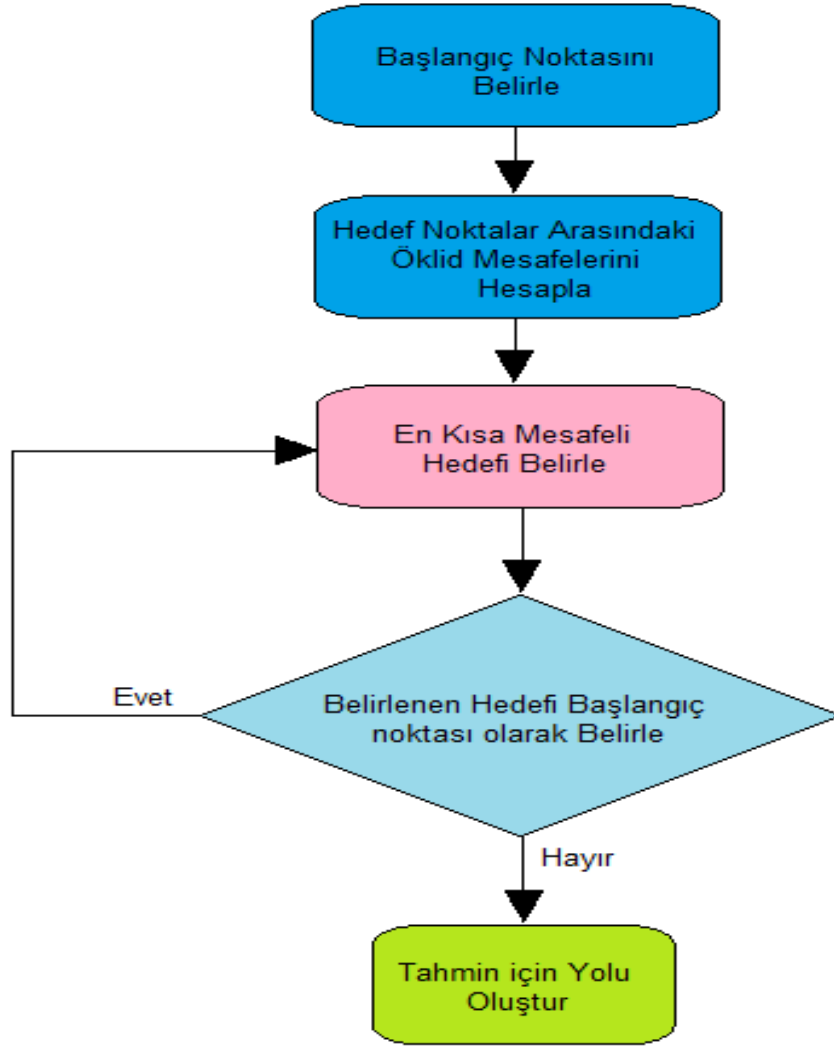
Önceden sınıflandırılmış eğitim verisinde komşularının ya da en yakın komşusunun oluşturduğu sınıfa atanarak, hedef düğümler arasındaki mesafeyi kullanmak suretiyle sınıflandırma yapan bir algoritmadır. Komşu sayısı uzaklıklara göre değerlendirilir ve K ile gösterilir. Öklid uzaklık denklemi ile komşuluk mesafeleri hesaplanabilmektedir [31, 34].

Sınıflandırılacak veri olan X_i ile mesafesi hesaplanacak en yakın N_i arasındaki $D(X,N)$ öklid uzaklığı, Denklem (25)'te n adet özellik için gösterilmektedir.

$$D(X,N) = \sqrt{\sum_{i=1}^n (N_i - X_i)^2} \quad (25)$$

Bu denkleme göre hesaplanan başlangıç düğümünden diğer düğümlere hesaplanan tüm öklid mesafeleri karşılaştırılır, en küçük mesafeye sahip düğüm hedef düğüm olarak belirlenir.

Yani en küçük mesafedeki en yakın K komşunun çoğunluğu hangi sınıfta ise, X örneği bu komşuların sınıfına atanır ve sınıflandırma tahmini bu şekilde gerçekleştirilmektedir [31,34,43]. KNN algoritmasına ait çalışma diyagramı Şekil (5)'te gösterilmektedir.

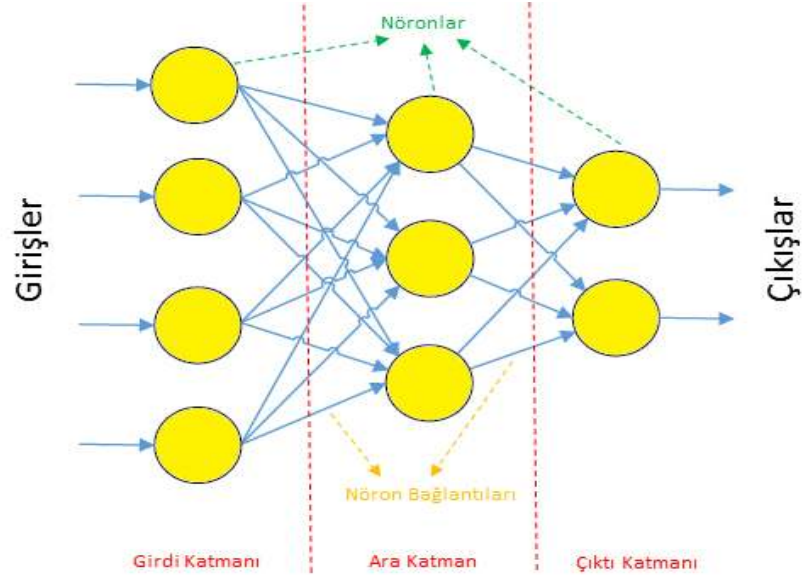


Şekil 5. KNN Algoritması Çalışma Akış Diyagramı.

1.6. Yapay Sinir Ağları

Bilgi işleme, çağımızda genellikle bilgisayarlar aracılığı ile gerçekleştirilmektedir. Buna bağlı olarak da bilgi işleme durumunun yalnız bilgisayarlara bağımlı olduğu fikri oluşmaktadır. Fakat bilgi işleme durumu aslında dünyada canlılığın başlangıcıyla gerçekleşen bir durumdur.

İnsanlar ve doğadaki diğer canlılara ait yapıların benzetilmesi ve bu bağlamda bir sistem geliştirme düşüncesi bilgi işlem teknolojilerinin ilerlemesi hedefiyle oluşmuştur.



Şekil 6. Örnek Bir YSA Yapısı

Kaynak: [34].

Beyin yapısını daha iyi anlayabilmek amacıyla sinirsel psikolojik deneyler, matematik alanında ve fen bilimlerinde yapılan araştırmalar, insan bilgi işlem sürecini taklit edip modellemek için uygun bir yöntem olmuştur. Bu bağlamda YSA insan beynini modelleyerek geliştirilmiştir.

Yapay Sinir Ağları metodunun mantığı kendi kendine öğrenme yoluyla oluşmaktadır. Özellikle sınıflandırırken, tahmin yaparken, veriye ait uygun değeri bulma durumlarında oldukça başarılı ve hızlıdır [44].

YSA'nın kullanıldığı uygulama alanı ve sektörler aşağıda belirtildiği gibi özetlenmiştir.

- Ses Tanıma
- Parmak izi tanıma
- Plaka tanıma
- El yazısı tanıma
- Tıp alanında tahmin, teşhis ve yorumlama
- Robotik

- Sözcük tanıma, tercüme
- Sinyal işleme, görüntü işleme
- Telekomünikasyon uygulamaları, görüntü ve data karşılama
- Kredi uygulamalarında müşteri analizi

1.7. YSA'nın Avantaj ve Dezavantajları

YSA'nın avantajlarını öncelikle sıralarsak:

YSA hücrelerden meydana gelir ve bu hücreler eş zamanlı çalıştığı için karmaşık işlemlerde yapılabilmektedir. İşlemin devam etmesi durumunda hücrelerden birinde sorun olsa bile sistem güvenli olarak çalışmasına devam eder [44].

YSA'lar aynı anda birden fazla işi yapabilme kabiliyetine sahiptir. Yapıları dolayısıyla daha önce hiç karşı karşıya kalmadıkları durumla ilgili de sonuçlar verebilmektedir. YSA'lar hatalara karşı gösterdikleri tolerans sayesinde herhangi bir problemle karşı karşıya kaldıklarında yanlış karar verecek çıktılar üretmezler.

YSA'ların dezavantajları:

Sadece nümerik bilgilerle çalıştığı için, kullanıcıların herhangi bir veriyi doğru olacak şekilde nümerik bilgiye çevirmemesi durumunda YSA hatalı sonuç üretebilmektedir.

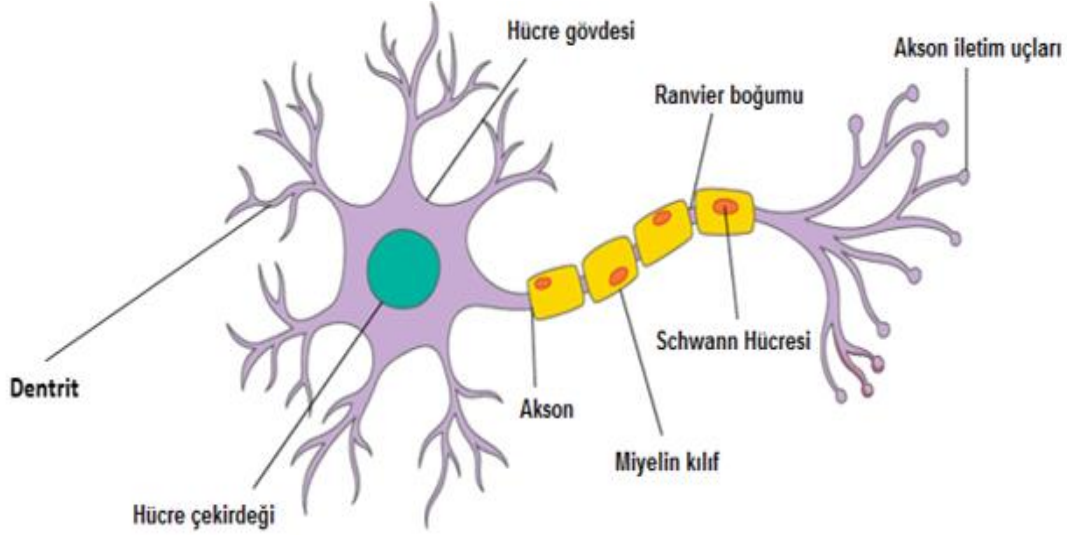
Ağın eğitimi yapılırken ne zaman sonlanacağına dair geliştirilmiş bir yöntem bulunmamaktadır. Genelde eğitimde kullanılan örnekler üzerindeki hatanın belirli bir değerin altına indirilmesiyle eğitim sonlandırılmaktadır. Bu durum kullanıcıyı zor durumda bırakmaktadır. Ağdaki nöron sayısı, ara katman sayısı öğrenme katsayısı gibi parametreler belirlenirken kural olmaması bu durumu ortaya çıkarmaktadır.

Hangi ağ yapısının kullanılacağı konusunda genelde deneme yanılma yolu ile belirlenmektedir. Eğer uygun bir ağ yapısı kurulmazsa problem çözülmemekte ya da performansını düşürerek çalışmaktadır [45,46].

1.8. Yapay Sinir Hücresinin Yapısı

YSA'lar insan beyninden ilham alarak geliştirilmiştir. Nöron hücresi, insan beyninde yer alan bu yapıları temsil etmektedir ayrıca bu yapıların birbirine

bağlanması sonucu sinir ağırları oluşmaktadır. Biyolojik açıdan bu bağlantı, dentrit ve akson uçlarının birbiriyle bağlantısıyla gerçekleşmektedir. Gelen uyarıları alan dentritler, aksonlara iletim sağlar. Aksonlar da dentritlerden aldıkları uyarıları başka sinir hücresine iletimini sağlamaktadır [47].



Şekil 7. Sinir Hücresi Yapısının Biyolojik Gösterimi

Kaynak: [47].

Şekil 7’ de belirtildiği gibi biyolojik sinir hücresine, başka sinir hücrelerinden gelen bilgi ve uyarılar dentritler aracılığıyla alınarak, bilgilerin toplanması amacıyla hücre çekirdeğine iletilmektedir. Bu duruma benzer olarak, yapay sinir hücresinde de, ağırlıklı girişler alınır ve toplama fonksiyonuna iletilmektedir. Hücresinin net girdisini sağlayan fonksiyon, toplama fonksiyonudur. Bu ifadenin matematiksel gösterimi Denklem (26)’da gösterilmiştir.

$$NET = \sum_{i=1}^n w_i X_i \quad (26)$$

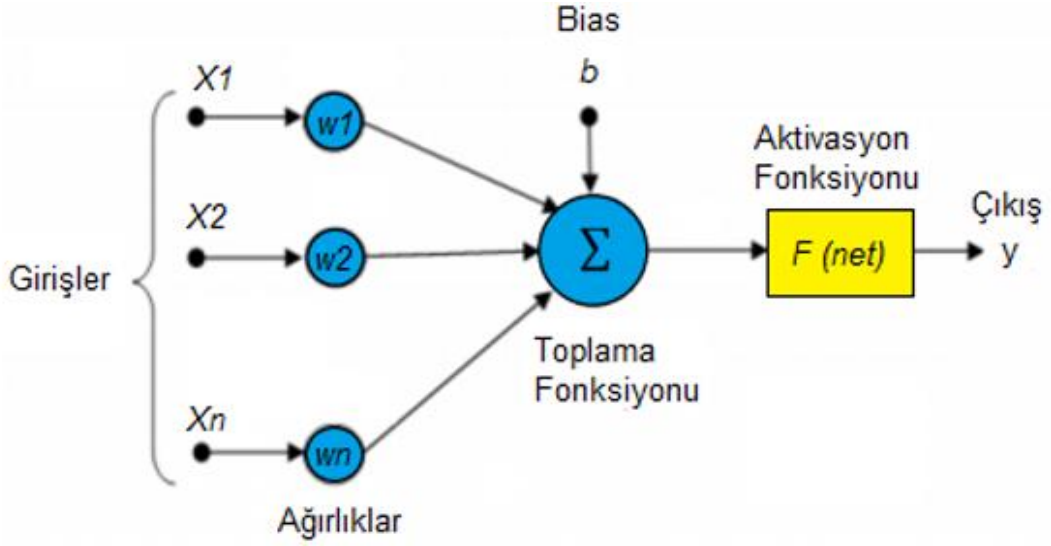
Denklem (26)’ da NET , ağırlıklı girişlerin çarpımının toplamını göstermektedir.

w : Ağırlıklar

X : girişler

n : Toplam giriş sayısı

Yapay sinir hücresinin modelini Şekil (8)’deki gibi göstermek mümkündür.



Şekil 8. Yapay Sinir Hücresi

Kaynak: [48].

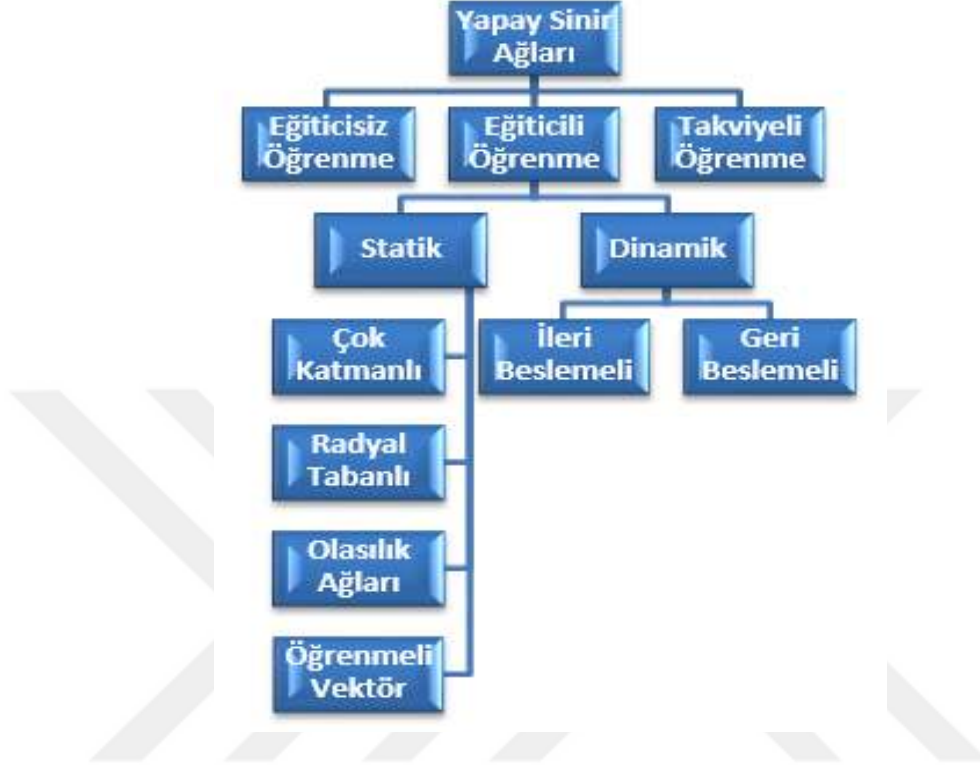
Yapay sinir hücresine gelen veriler ağırlık değerleri ile çarpılır. Hesaplanan bu ağırlık değerleri negatif veya pozitif değer olabilir. Eğer sıfır değerini alırsa, bağlı olduğu girişin etkili olmadığını ifade etmektedir. Şekil (8)' de belirtilen aktivasyon fonksiyonu, elde edilen net girdinin işlenerek sonucu oluşturan fonksiyondur.

YSA modeli seçilirken doğrusal olmayan model seçimi ve türevlenebilir bir fonksiyon olması oldukça önemli bir durumdur. Çok katmanlı YSA'larda sigmoid aktivasyon fonksiyonu genellikle kullanılmaktadır. Denklem (27)'de Toplama fonksiyonundan gelen *NET* toplam değerine göre, sigmoid fonksiyonunun ifadesi gösterilmektedir [47,48,49,50].

$$F(NE T) = \frac{1}{1+e^{-NE T}} \quad (27)$$

1.9. YSA Sınıfları

Yapay sinir ağı sistemlerinin sınıflandırması Şekil (9)'da gösterilmektedir.

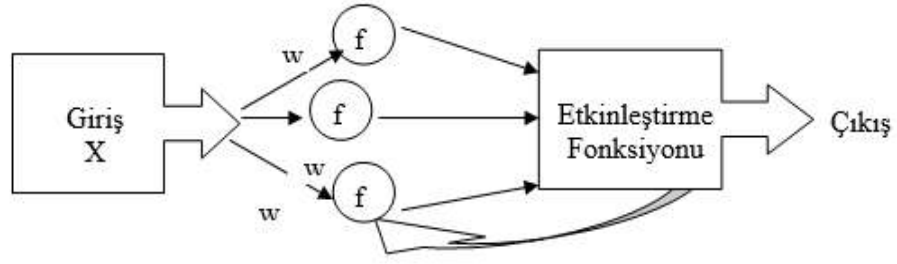


Şekil 9. Yapay Sinir Ağlarının Sınıflandırılması

Kaynak: [51].

1.9.1. Eğitici-siz Öğrenme

Sistem, giriş bilgilerini alır ve bu bilgileri gruplandırarak ağı içinde benzer olanları belirler. Ağırlıkları sistem bu bilgilere göre ayarlar. Giriş bilgilerinin çıkış durumlarına göre kendi sınıflandırmasını yapar. Sistem, verdiği cevaplarla doğru çözümü bulur ve benzer gruplara uygular. Bu süreç, sistem en uygun cevabı bulana kadar döngü şeklinde devam eder. Şekil (10)'da kendi kendine öğrenme yapısını göstermektedir [51].

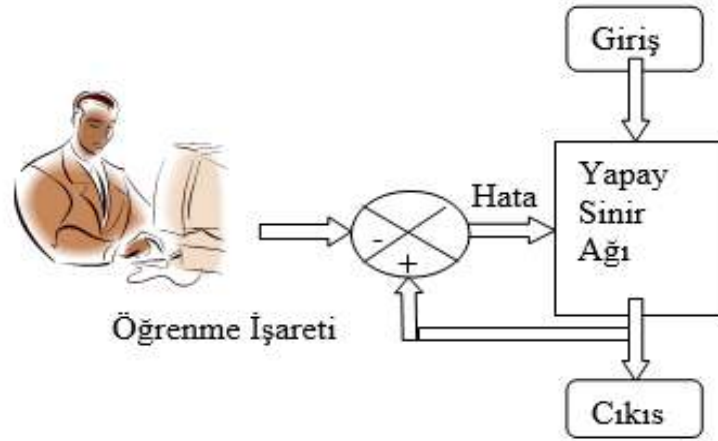


Şekil 10. Eđiticisiz Öğrenme

Kaynak: [46].

1.9.2. Eđiticili Öğrenme

Eđiticili öğrenme modunda, yapay sinir ađı eğitimi için eğitim verileri (eđitim seti) kullanılır. Ađın eğitimi için başlangıçta bağlantı ađırlıkları rasgele atanır. Eğitim süreci, istenilen bilgiler ile ađın çıkışı arasındaki fark (hata) azalana kadar devam eder. Ađdaki hata azaldıkça, ađırlıklar kararlılık kazanır. Ađırlıklar istenen kararlılığa ulaştığında eğitim tamamlanır [52]. Şekil (11)'de eđiticili öğrenme yapısı gösterilmektedir.



Şekil 11. Eđiticili Öğrenme Yapısı

Kaynak: [51].

1.9.3. Takviyeli Öğrenme

Eđiticili öğrenmeye benzer bir modeldir. Eđiticili öğrenme ile farkı, çıkışın uygunluđunu deđerlendirmede öğretmen yerine deneme-yanılma yöntemi kullanmasıdır. Sistem, uygun çıkışı bulana kadar deneme-yanılma yöntemiyle işlemi

hata elde edilene kadar ağ birden fazla kez eğitilebilir. Minimum hatada, gerçek değere en yakın sonuçları elde edebilir ve yeni durumlarla karşılaştığında genelleme yapabilir. Bu süreçte ağ, eğitilmiş bir yapısı haline gelir. Eğitim tamamlandıktan sonra, ağ yapısının en güncel ağırlık değerleriyle daha önce karşılaşmadığı durumlar test edilir ve ağın performansı ölçülür. Bu aşamaya ağın validasyonu denir. Validasyon değeri yeterince yüksek çıkana kadar eğitim devam etmelidir. Validasyon aşamasından sonra ağın eğitimi tamamlanmış olur [48].

YSA'nın eğitimi sırasında üretilen çıktı değerleri ile beklenen çıktı değerleri arasındaki fark, Denklem (28)'de gösterildiği gibi hata olarak hesaplanır. Bu hata değeri kullanılarak ağırlık değerleri ve öğrenme parametreleri güncellenir.

$$E_n = H_n - \zeta_n \quad (28)$$

Denklem (28)' e göre, n. durum için;

E: Hata değeridir.

H: Beklenen çıkış değeridir.

Ç: Ağın hesapladığı çıkış değeridir.

Fakat bu denklemde sadece bir adet işlem elemanı için geçerlidir. Bu durum karşısında hata değeri negatif veya sıfır değerleri alabilir. Bunu engellemek için toplam hata Denklem (29)'daki gibi hesaplanmalıdır [48].

$$E_{toplam} = \frac{1}{2} \sum_m E_n^2 \quad (29)$$

1.10.1. YSA'lar için Öğrenme Algoritmaları

YSA'ların eğitimi için çeşitli öğrenme algoritmaları bulunmaktadır. Literatürde en yaygın tercih edilen algoritmalar arasında geri yayılım (Back Propagation) ve bu algoritmanın geliştirilmesiyle ortaya çıkan Levenberg-Marquardt (LM) algoritması yer almaktadır. Diğer öğrenme algoritmaları arasında Delta Bar Delta (DBD), Quasi Newton (QN) gibi yöntemler bulunmaktadır [54,55].

1.11. Sınıflandırmada Performans Kriterleri

Tez çalışmasında konuşmacılardan alınan ses sinyalleri, genlik ve frekans spektrumlarından elde edilen parametreler ile giriş özellikleri olarak kullanılmıştır.

Çıkış özelliği olarak ise, konuşmacıyı tanınması istenilen doğru kişi ile eşleştirme yapabilmesine bağlı 2 farklı çıkış durumu belirlenmiştir. Tez çalışması kapsamında bu aktiviteleri ayırt etmek için DT, SVM, KNN, NB ve YSA algoritmaları kullanılmıştır. Sınıflandırmada ölçütün performansına ait başarının belirlenmesinde, tablo ve grafikler gibi farklı ölçüt parametreleri kullanılmaktadır.

Bir makine öğrenmesi modeli performans analizi yapılırken her modele göre farklı metrikler kullanılabilir. Bu metriklerden bazıları şu şekilde sıralanabilir: doğruluk, kesinlik, F puanı, ROC (Receiver Operating Characteristic) Eğrisi, karmaşıklık matrisi [56]. Tez çalışmasında, ROC eğrisi ve karmaşıklık matrisleri kullanılarak performans değerlendirmeleri gerçekleştirilmiştir.

Karmaşıklık matrisi; DP, YP, YN ve DN olmak üzere dört öğeden oluşmaktadır. Bu matrisler hücre içindeki sayısal değerlere göre değişen renkler ile görselleştirilirler [57].

Tablo 2. Karmaşıklık Matrisi

<i>Tahmin Sınıfı</i>	0	Doğru Negatif (DN)	Yanlış Negatif (YN)
	1	Yanlış Pozitif (YP)	Doğru Pozitif (DP)
		0	1
		<i>Gerçek Sınıf</i>	

Tablo 2’de gösterilen matrisin öğeleri aşağıda maddeler halinde açıklanmaktadır:

- **DP:** Gerçek pozitif sonuçların sayısını ifade eder. Doğru bir şekilde pozitif olarak sınıflandırılan örneklerin sayısıdır.
- **DN:** Gerçek negatif sonuçların sayısını ifade eder. Doğru bir şekilde negatif olarak sınıflandırılan örneklerin sayısıdır.

- **YP:** Yanlış pozitif sonuçların sayısıdır ancak gerçek negatif olarak sınıflandırılması gerekir. Model tarafından yanlış şekilde pozitif olarak sınıflandırılan örneklerin sayısını vermektedir.

- **YN:** Yanlış negatif olarak verilen sonuçların sayısıdır ancak gerçek pozitif olarak sınıflandırılması gerekmektedir. Modelin yanlış şekilde sınıflandırdığı negatif örneklerin sayısını ifade etmektedir.

Karmaşıklık matrisleri gerçek değerler ile tahmin edilen değerler arasındaki ilişkiyi göstermektedirler [57].

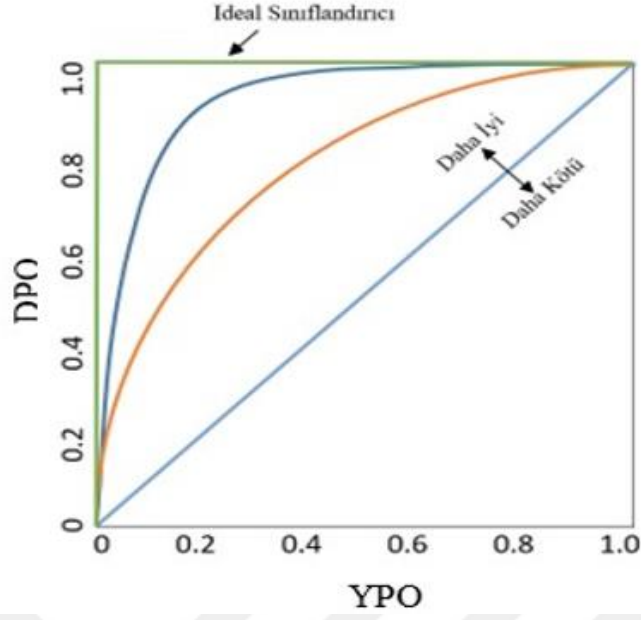
Bu tez çalışmasında Karmaşıklık matrislerinden elde edilen doğruluk performansları ölçülmüştür. Doğruluk performansını hesaplamada doğru sınıflandırılmış örnek sayıları DP ile DN toplamının toplam örnek sayısının diğer bir deyişle doğru sınıfların tamamı ile yanlış sınıfladıklarının toplamlarına oranı şeklindedir. Denklem (30)'da gösterildiği şekilde hesaplanmaktadır.

$$\text{Doğruluk (\%)} = \left(\frac{(DP+DN)}{(DP+YN+YP+DN)} \right) \times 100 \quad (30)$$

Mesela bu durumu bir örnek ile açıklayacak olur isek, DP gerçek hasta kişiler ise DN sağlıklı kişileri temsil etmektedir. Bu durumda YP gerçekte hasta olmayıp hasta olarak sınıflandırılan, YN ise gerçekte hasta olup ancak hasta olarak tahmin edilmeyen kişi sayısı anlamındadır. Şayet verilerin tamamında gerçekte hasta olanlar ve de olmayanları tamamen doğru sınıflamışsa, bu durumda YN ve YP sayıları sıfır olmakla beraber doğruluk oranı %100 olmaktadır.

ROC eğrisi ise herhangi bir sınıflandırma modelinin doğru pozitif oranının, yanlış pozitif oranı ile karşılaştırılarak değişimini gösteren bir grafikdir. ROC Eğrisi altında kalan alan ise AUC değeri olarak nitelendirilmektedir. Bu değer ile öğrenme modelinin performansı incelenmektedir. AUC değeri, problem ve kullanılan algoritmalara bağlı olarak değişik yöntemlerle hesaplanabilmektedir [58].

Kullanılmakta olan en yaygın yöntemde ise Gerçek Pozitif Oranı (DPO) ile Yanlış Pozitif Oranı (YPO) karşılaştırılarak, ROC eğrisi altında kalan alan hesaplanmaktadır [59]. Örnek bir ROC eğrisi ve bu eğri üzerinde dikkat edilmesi gereken noktalar Şekil 13'te gösterilmiştir.



Şekil 13. ROC Eğrisi

Kaynak: [58].

ROC eğrileri, doğru pozitif oranının (DPO) ve yanlış pozitif oranına (YPO) göre çizilir [60]. DPO oranı Denklem (34) ile hesaplanmaktadır:

$$DPO = \frac{DP}{DP+YN} \quad (34)$$

YPO oranı Denklem (35) ile hesaplanmaktadır:

$$YPO = \frac{YP}{YP+DN} \quad (35)$$

AUC değeri Denklem (36) ile hesaplanmaktadır:

$$AUC = \int_0^1 y(x)dx \quad (36)$$

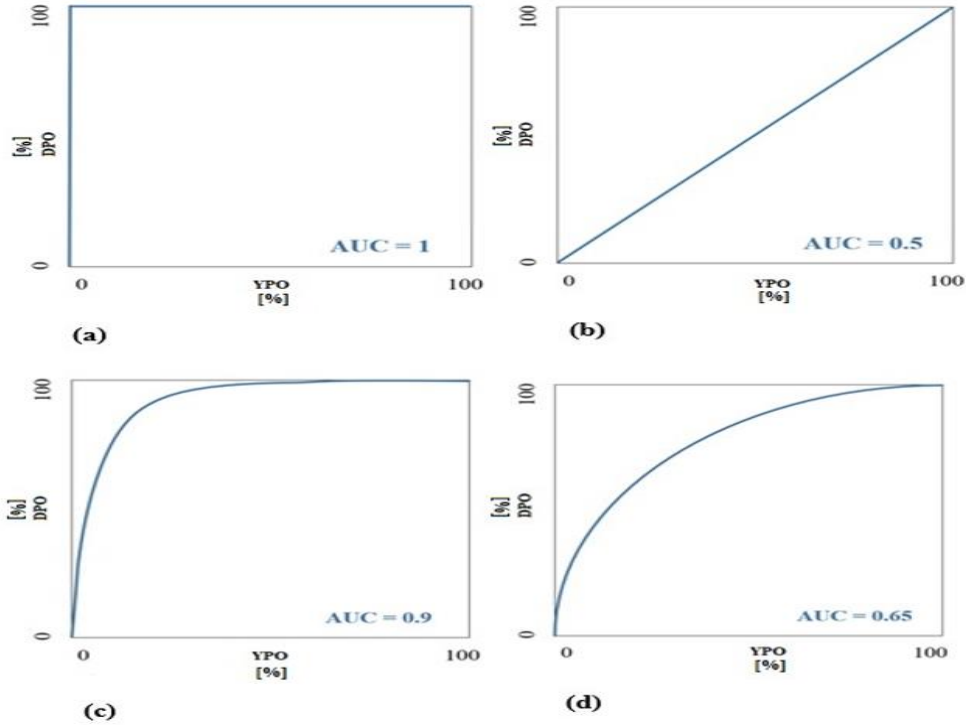
AUC değeri hesaplanırken denklemdeki $y(x)$ fonksiyonu yerine DPO, integrasyon değişkenindeki x yerine ise YPO değerleri koyulduğunda aşağıdaki Denklem (37) elde edilmektedir:

$$AUC = \int_0^1 DPO d(YPO) \quad (37)$$

Şekil 14'te AUC değerinin değişiminin farklı ROC eğrileri ile ilişkileri gösterilmiştir. Grafik (a)'da AUC değeri 1'dir. Model performansı ROC eğrisinin sol

üst köşesine göre arttığından, Grafik (a)'da model, en yüksek performansı göstermektedir. Grafik (b)'de ise AUC değeri 0,5 olarak görülmektedir.

Burada bu değer, DPO ve YPO doğrusal olarak arttığından, modelin doğru çalışmıyor olduğunu göstermektedir. Grafik (c)'deki 0,9 olan AUC değerine göre, model iyi bir performans göstermektedir. Grafik (d)'de ise AUC değeri 0,6'dır ve bu değer modelin orta seviye bir performans sergilediğini ifade etmektedir [60].



Şekil 14. ROC Eğrisi ve AUC Değeri İlişkisi

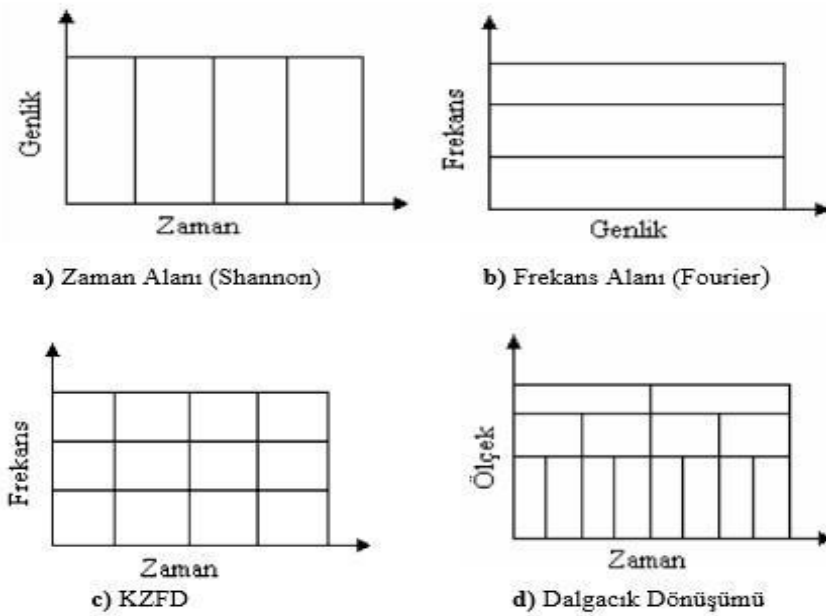
Kaynak: [60].

1.12. Dalgacık Dönüşümü

Dalgacık dönüşümü, durağan olmayan sinyallerin analizlerinde kullanılmaktadır. Özellikle zaman ve frekans analizlerinde fourier dönüşümünün yetersiz kalmasından dolayı tercih edilmektedir. Dalgacık dönüşümü yönteminde, gerçek zamanlı sinyalden zaman ve frekans ölçeklenmesiyle daha fazla özellik elde edilebilmektedir. Bant genişliğinin dar olduğu yerler için ölçeklenebilir ve geniş bir pencere ile alçak frekanslarda yüksek frekans bileşenlerini analiz etmek için kullanılmaktadır [61].

Dalgacık dönüşümü, frekans spektrumundaki farklı bölgelerin frekans çözünürlüklerinin kolaylıkla ayarlanabilmesinde, farklı frekans bantlarının spektrum analizlerinde kullanılması halinde frekans spektrumunun hepsinin hesaba dahil edilmemesi ve bu sayede dönüşüm hızının alçak frekanslarda oldukça yüksek olmasını sağlamaktadır. Böylelikle durağan olmayan sinyaller için daha kullanışlı ve esnek zaman-frekans bilgisi elde edilmesini sağlamaktadır [62].

Dalgacık dönüşümü, geleneksel işaret işleme teknikleri tarafından görülmeyen detayları meydana çıkaran, işaretin ölçeklenebilen zaman-frekans gösterimi ile analizini sağlayan başarılı bir araçtır. İşarete öteleme ve ölçekleme özelliğini kazandırmaktadır. İşaretler analiz edilirken dalgacık dönüşümünde analiz penceresinin konumu ve uzunluğu kendiliğinden değişmektedir, Fourier dönüşümünde ise analiz boyunca analiz penceresi kaydırılabildiği halde pencere uzunluğu sabit kalmaktadır. Dalgacık dönüşüm analizi, farklı ölçek bölgelerine sahip bir pencereleme tekniğidir. Düşük frekans bilgisi önem arz ettiğinde büyük zaman aralıklarının, yüksek frekans bilgisi önem arz ettiğinde ise küçük zaman aralıklarının kullanılmasına izin vermektedir. Dalgacık dönüşümü ile alt bileşenlerine ayrılan durağan olmayan işaret, durağan hale getirilmektedir. Şekil 15'te dalgacık dönüşümü ile sinyal analiz yöntemleri arasındaki ilişki gösterilmektedir [51].



Şekil 15. Sinyal Analiz Yöntemleri Arasındaki İlişki

Kaynak: [51].

Şekil (15a)'da bir işaretin zaman-genlik grafiği gösterilmektedir. Bu gösterim, işaret ile ilgili bir analiz yapabilmek için çok fazla anlam ifade etmemektedir. Böylece genlik-frekans spektrumlarının analizine ihtiyaç duyulmaktadır. Bu durumda Şekil (15b)'de görüldüğü üzere Fourier dönüşümü kullanmak yeterlidir ancak bu da işaretin tamamı ile ilgili bilgi vermektedir. Bir işaretin farklı frekans bantlarında ve farklı bölgelerde incelenebilmesi için Kısa Zaman Fourier Dönüşümü (KZFD) ortaya atılmıştır. Bu dönüşümün temeli belirli bir ölçekteki bir pencere fonksiyonu ile işaretin çarpılmasına dayanmaktadır. Şekil (15c)'de bununla ilgili grafik gösterilmektedir. Bu analiz yönteminde pencerelerin eşit aralıklı olması, işaret analizi yaparken daha büyük ve daha küçük frekansların incelenmesinde sorun teşkil etmektedir. Bu sorunun giderilmesi için dalgacık dönüşümü analizi ortaya atılmıştır. Bu sayede düşük frekans bilgisinin önemli olduğu araştırmalarda büyük zaman aralıkları, yüksek frekans bilgisinin önemli olduğu araştırmalarda da küçük zaman aralıklarının kullanımına olanak sağlanmaktadır [59]. Şekil (15d)'de Dalgacık Dönüşümü ile ilgili grafik gösterilmektedir.

Fourier dönüşümü, sadece sinüs ve kosinüs fonksiyonlarının kullanıldığı tek bir baz fonksiyon kümesine sahiptir. Ancak dalgacık dönüşümleri sonsuz sayıda baz fonksiyonları kümesine sahiptir. Bu nedenle dalgacık dönüşümü analizi ile, fourier analizi gibi diğer zaman-frekans metotlarının kullanılması sonucu saklı kalan bilgiye doğrudan erişilebilmektedir. En fazla kullanılan dalgacık çeşitleri Haar, Meksika şapkası, Coiflet, Meyer, Symlet ve en önemlisi Daubechies olarak sıralanabilir. Daubechies dalgacık ailesi hem sürekli dalgacık dönüşümü hem de ayrık dalgacık dönüşümü için kullanım özelliği taşımaktadır. Aynı zamanda ortogonallik ve biortogonallik özelliklerini sağlamaktadırlar [48, 63].

Mühendislik uygulamalarındaki kullanımlarda en fazla tercih edilen dalgacık türü olduğundan, bu çalışmada da Daubechies dalgacık türleri tercih edilmiştir.

1.12.1. Sürekli Dalgacık Dönüşümü (SDD)

Sürekli Dalgacık Dönüşümü (SDD)'de, ölçeklenebilen ve sinyal boyunca kaydırılabilen pencere modülleri kullanılmaktadır. Her kaydırma işleminin sonunda sinyale ait spektral bileşenler incelenir. Pencereleme yapılırken, fourier dönüşümünde

sabit ölçekleme ile eşit aralıklarda spektrum elde edilmektedir. Dalgacık dönüşümünde ise zaman frekans analizleri farklı ölçeklemeler kullanılarak yapılabilmektedir [64].

Denklem (38)'de bu dönüşüm metodunun matematiksel eşitliği ifade edilmektedir.

$$W(a, b) = \int_{-\infty}^{+\infty} (t) \frac{1}{\sqrt{|a|}} \psi_{a,b} \left(\frac{t-b}{a} \right) dt \quad (38)$$

Denklem (38)'de $X(t)$, sinyalin zamana bağlı değişimi olarak ifade edilmektedir. Dönüşümde kullanılacak olan dalgacık fonksiyonu $\psi(t)$, farklı frekans seviyeleri için dönüşüm faktörü b olarak tanımlanmaktadır. Diğer bir parametre olan a faktörü ise ölçeklemede kullanılmaktadır. Böylece ana dalgacığın kendisi ise; $a=1$, $b=0$ değerleri için $\psi_{a,b}(t) = \psi(t)$ olarak tanımlanmaktadır [61].

1.12.2. Ayrık Dalgacık Dönüşümü (ADD)

Ayrık Dalgacık Dönüşümü (ADD) metodu ise yapı itibariyle ayrık zamanlı Fourier dönüşümüne benzerdir. Giriş sinyali analizler için farklı fonksiyonlar ile parçalı bir şekilde ayrık olarak analiz edilmektedir. Ayrık dalgacık dönüşümü metodunda, ayrık zamanlı fourier dönüşümünden farklı olarak pencereleme yerine ayrı bir dalgacık fonksiyonu kullanılmaktadır ve analizler bu şekilde gerçekleştirilmektedir. Ayrıca bu yöntemde, yüksek frekanslara çıkıldığında zaman çözünürlüğü artmaktadır. Böylece düşük frekans çözünürlüğü elde edilmektedir.

Alçak frekanslara inildiği durumda ise zaman çözünürlüğü azalır, yüksek frekans çözünürlüğü elde edilir. Böylece sinyal zaman-frekans düzleminde etkin bir şekilde işlenmektedir [65]. Bu tez çalışmasında Ayrık Dalgacık Dönüşümü kullanılması tercih edilmektedir.

Ayrık Dalgacık Dönüşümü, durağan olmayan bir sinyale uygulandığında $\phi(t)$ fonksiyonu ve dalgacık $\psi(t)$ fonksiyonu olmak üzere iki önemli fonksiyon karşımıza çıkmaktadır. Ölçekleme fonksiyonu, yaklaşım bileşenlerini oluşturmak üzere giriş sinyali ile konvolüsyona girmektedir. Denklem (39)'da ölçekleme fonksiyonunun matematiksel ifadesi gösterilmektedir.

$$\phi(t) = 2^{\frac{-m}{2}} \phi(2^{-m}t - n) \quad (39)$$

Denklem (39)'da m ölçekleme seviyesi sayısını, n ise öteleme miktarını göstermektedir. Giriş sinyali ile konvolüsyona giren bu fonksiyonun oluşturduğu yaklaşım katsayılarının matematiksel gösterimi Denklem (40)'da ifade edilmektedir.

$$S_{m,n} = \int_{-\infty}^{+\infty} X(t)\phi_{m,n}(t)dt \quad (40)$$

Elde edilen bu yaklaşım katsayıları yaklaşım ve detay bileşenleri hesaplanmaktadır. Ölçekleme fonksiyonu ile tekrar işleme sokulan yaklaşım katsayıları, yaklaşım ve detay bileşenlerinin hesaplanmasını sağlamaktadır. Denklem (41)'de yaklaşım bileşenlerinin, Denklem (42)'de ise detay bileşenlerinin hesaplanmasını gösteren eşitlikler ifade edilmektedir.

$$A(t) = \sum_{n=-\infty}^{\infty} S_{m,n}\phi_{m,n}(t) \quad (41)$$

$$D(t) = \sum_{n=-\infty}^{\infty} T_{m,n}\psi_{m,n}(t) \quad (42)$$

Denklem (42)'de $T_{m,n}$, detay katsayılarını, $\psi_{m,n}(t)$ ise dalgacık fonksiyonunu ifade etmektedir. Detay ve yaklaşım bileşenlerinin hesabı yapılırken ölçekleme fonksiyonunun kayma ve daralma özellikleri önem kazanmaktadır. Denklem (43)'te bu durum ile ilgili eşitlik ifade edilmektedir.

$$\phi(t) = \sum_k c_k \phi(2t - k) \quad (43)$$

Denklem (43)'te zaman ekseninde k kadar ötelenerek elde edilmiş olan $\phi(2t - k)$ fonksiyonunun, her bir k değerine karşılık elde edilmiş olan c_k ölçekleme katsayısı ile çarpılması sonucu $\phi(t)$ fonksiyonu elde edilmektedir. Ortogonallığı sağlamak için her bir k değerine karşılık gelen c_k ölçekleme katsayılarının toplamı ve ölçekleme katsayılarının karesel toplamının ikiye eşit olması gerekmektedir. Dalgacık denklemi, ölçekleme katsayıları kullanılarak Denklem (44)'te gösterildiği şekilde elde edilebilmektedir.

$$\psi(t) = \sum_k (-1)^k c_{1-k} \phi(2t - k) \quad (44)$$

Denklem (44)'te dalgacık katsayısı b_k matematiksel olarak $(-1)^k c_{1-k}$ şeklinde gösterilmiştir. Dolayısıyla eşitlik tekrar düzenlendiğinde Denklem (45)'te gösterildiği gibi bir eşitlik elde edilmektedir [48].

$$\psi(t) = \sum_k b_k \phi(2t - k) \quad (45)$$

Bu katsayıların belirlenmesi sürecinde bahsedildiği üzere farklı dalgacık modelleri mevcuttur. Ancak Daubechies dalgacığı dönüşümü daha kolay hale getirmektedir. Bu dalgacık dbN olarak Böyle bir dalga $2N$ adet ölçekleme katsayısına sahiptir [66].

Bu tez çalışmasında konuşmacılardan alınan ses sinyallerinin dalgacık dönüşümleri $db4$ Daubechies dalgacık fonksiyonu ile ölçekleme yapılarak detay bileşenleri elde edilmiştir. Elde edilmiş olan detay bileşenlerinin her birinden Shannon entropileri, E_1, E_2, E_3 ve E_4 hesaplanarak giriş özellikleri olarak kullanılmıştır. Burada elde edilen her bileşenin dalgacık enerji spektrumları E_{jk} dikkate alınmıştır. Denklem (46)'da bu durum ifade edilmektedir.

$$E_{jk} = |D_j(k)|^2, \{D_j(k): [2^{-(j+1)}fs, 2^{-j}fs] \} (j = 1, 2, \dots, m) \quad (46)$$

Denklem (46)'da $D_j(k)$ ifadesi detay bileşenlerin örnekleme frekansı (fs)'e bağlı olarak j ölçeğinde, k anındaki frekans domenindeki gösterimidir. Böylece her bir detay bileşene ait Shannon entropi E_j değerleri denklem (47)'de gösterildiği şekilde hesaplanmıştır [48].

$$E_j = -\sum_k E_{jk} \log E_{jk} \quad (47)$$

Tez çalışmasında bu işlemlerin tamamı Matlab programının dalgacık dönüşümü arayüzü kullanılarak gerçekleştirilmiştir.

İKİNCİ BÖLÜM

ANALİZ BULGULARI

Bu bölümde, yaşları 18 ile 39 arasında değişen 10 farklı konuşmacıdan alınan konuşma örnekleri kullanılarak veri seti oluşturulmuştur. Bu veri setinde her bir konuşmacıya 25 farklı kelimeler söylenilerek toplamda 260 adet veri Matlab yazılımı ile Machine Learning Toolbox kullanılarak analiz edilmiştir. Sonuçlar grafik ve tablolar halinde sunulmuştur. Verilerin sınıflandırılmalarında K-Katlamalı çapraz doğrulama tekniği (k-fold) değeri 5 olarak seçilmiştir. Bu sayede makine öğrenme modeli 5 kez eğitilmiştir.

2.1. Ses Sinyaline Ait Genlik ve Frekans Özelliklerinin Analizleri

Bu çalışmada ses sinyaline ait Ortalama Frekans (Mean Frequency-MNF), Medyan Frekans (Median Frequency-MDF), Ortalama Mutlak Genlik (Mean Absolute Value-MAV), Ortalama Karekök (Root Mean Square-RMS), Sıfır Geçiş Sayısı (Zero Crossing-ZC), Willison Genliği (Willison Amplitude-WAMP), Dalga Boyu (Wavelength-WL), Log Dedektör (LOG), Varyans (Variance-VAR), Maksimum Fraktal Uzunluğu (Maximum Fractal Length-MFL), Standart Sapma Standart Deviation-(SD), Mutlak Genlik Değişimi (Average Amplitude Change-AAC) ve Mutlak Standart Sapma (Difference Absolute Standard Deviation-DASD) olmak üzere toplamda zamana bağlı genlik ve frekansa ait 14 farklı özellikler, 6 farklı makine öğrenme algoritmalarıyla ayrı ayrı analiz edilmiştir. Bu algoritmalar Karar Ağaçları (Decision Tree-DT), Lineer Diskriminant (LD), Destek Vektör Makine (Support Vector Machine-SVM, Naive Bayes (NB), En Yakın Komşuluk (K-Nearest Neighbors-KNN) ve de Yapay Sinir Ağları (YSA) olmaktadır. Analiz sonuçlarında ROC eğrileri ve de Karmaşıklık Matrisleri elde edilmiştir. Bu analizlerde 2 adet sınıf kullanılmış, makine öğrenmede eğitimler bu iki sınıfa göre gerçekleşmiştir. Bu sınıflamanın amacı gerçek kişiyi diğer bir deyişle istenilen kişiyi (örneğin akıllı ev uygulaması için ev sahibini) tahmin etmektir. Dolayısıyla diğer kişiler (0 sınıfı), gerçek kişi (1 sınıfı) olmaktadır. Ancak bu çalışmada 10 farklı kişi olduğundan dolayı, her bir kişi sırasıyla 1, diğerleri 0 sınıfı olacak şekilde bu çalışma tekrarlanmıştır. Çalışma sıralamasında Şekil 16' da görüldüğü gibi her bir kişi için sırasıyla 1, diğerleri

0 olmak üzere iki sınıfa ayrılmıştır. Elde edilen bulguların tamamı Tablo 3’te ayrıca gösterilmektedir.

	Deneme1	Deneme2	Deneme3	Deneme4	Deneme5	Deneme6	Deneme7	Deneme8	Deneme9	Deneme10
1.ci kişi	1	0	0	0	0	0	0	0	0	0
2.ci kişi	0	1	0	0	0	0	0	0	0	0
3.cü kişi	0	0	1	0	0	0	0	0	0	0
4.cü kişi	0	0	0	1	0	0	0	0	0	0
5.ci kişi	0	0	0	0	1	0	0	0	0	0
6.cı kişi	0	0	0	0	0	1	0	0	0	0
7.ci kişi	0	0	0	0	0	0	1	0	0	0
8.ci kişi	0	0	0	0	0	0	0	1	0	0
9.ci kişi	0	0	0	0	0	0	0	0	1	0
10.cu kişi	0	0	0	0	0	0	0	0	0	1

Şekil 16. Her Bir Deneme için Kişilere Göre Sınıflama Sıralamaları.

Tablo 3. Zamana Bağlı Genlik Frekans Özelliklerine Ait Bulguların Kişilere ait Karşılaştırmalı Sonuçları

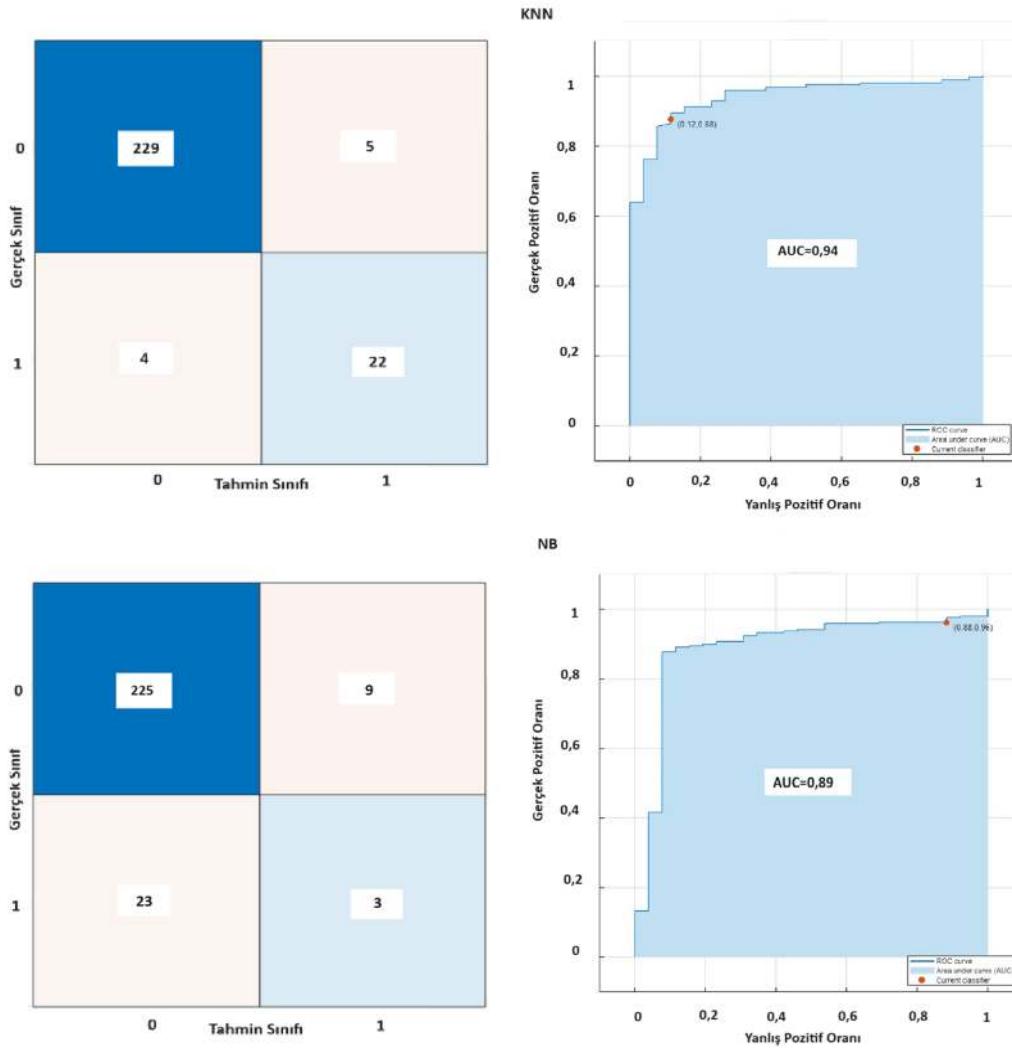
	1. Kişi (ERKEK)	2. Kişi (ERKEK)	3. Kişi (KADIN)	4. Kişi (ERKEK)	5. Kişi (KADIN)	6. Kişi (ERKEK)	7. Kişi (ERKEK)	8. Kişi (KADIN)	9. Kişi (KADIN)	10. Kişi (KADIN)
DT (Fine Tree)	88,8	88,5	88,5	90,4	93,1	93,8	92,7	90	96,5	97,3
DT (Medium Tree)	88,8	88,5	88,5	90,4	93,1	93,8	92,7	90	96,5	97,3
DT (Coarse Tree)	88,8	86,2	90,4	91,5	90,8	93,1	90	90,8	94,6	97,3
LD (Linear Discriminant)	88,5	90,4	89,2	88,1	90,4	90,4	88,8	89,6	89,6	98,5
NB (Kernel Naive Bayes)	82,7	79,6	80,4	85	91,9	87,7	84,6	88,8	86,9	98,8
SVM (Linear SVM)	90	90	89,6	90	90	87,7	88,5	88,8	90	99,2
SVM (Quadratic SVM)	86,2	86,9	88,5	88,5	90,8	95,4	95	84,2	90,4	99,2
SVM (Fine Gaussian SVM)	89,6	89,6	90	88,5	92,7	94,6	93,8	91,9	91,2	93,5
SVM (Medium Gaussian SVM)	90	90,4	90	90	90	91,2	90	90	90	97,7
KNN (Fine KNN)	86,2	90,4	87,7	85,8	92,7	96,5	91,9	87,3	88,5	98,5
KNN (Medium KNN)	89,6	90,4	90	88,8	91,9	93,1	90,8	91,2	91,9	98,8
KNN (Coarse KNN)	90	90	90	90	90	90,0	90	90	90	90
KNN (Weighted KNN)	87,7	91,9	90	89,6	91,9	95,8	91,5	90,8	91,2	98,8
YSA (Narrow Neural Network)	89,2	87,7	90	88,8	92,7	93,5	92,7	89,2	88,5	98,1
YSA (Medium Neural Network)	86,9	90	90,8	89,2	92,7	96,2	93,5	90	88,5	98,5
YSA (Wide Neural Network)	87,7	87,3	89,6	87,3	91,2	95,8	91,2	89,2	88,5	98,5

Tablo 3’te görüldüğü gibi değerler kişiden kişiye değişmekle beraber, maksimum %99,2 ve %98,8 mertebelerinde yüksek doğruluk oranları ile SVM ve KNN sınıflayıcılarda ve minimum değer NB sınıflayıcısıyla %79,6 mertebelerinde elde edilmiştir. Dolayısıyla zamana bağlı genlik ve frekans 14 farklı özelliğin tamamı kullanılarak SVM ve KNN ile yüksek sonuçlar elde edilebilmekle beraber literatürde

%83-%89 doğrulukta konuşmacı tanımlaması için kullanılan MFCC ile yapılan emsal çalışmalara [15,16] kıyasla iyi performans olarak kabul edilebilir.

Ayrıca zamana bağlı genlik ve frekans özellikleri (Ortalama Genlik, Standart Sapma, Maksimum, Minimum ve Medyan frekans değerleri, regresyon katsayıları v.b. özellikler) Karar Ağaçları (Decision Tree-DT) ve Destek Vektör Makine (DVM) ile birlikte kullanarak 0,69 ile 0,84 oranında elde edilmiş [9] çalışmalara kıyasla da sonuçlar tatmin edicidir.

Bu çalışmada her bir katılımcı kişilere ait performanslar 14 farklı özellik kullanılarak Karmaşıklık matrisleri ile ele alınmıştır ve Tablo 3'te gösterildiği gibidir. Örnek olarak 6. kişiden elde edilen sonuçlar irdelendiğinde, Şekil 17'deki grafikler elde edilmiştir.



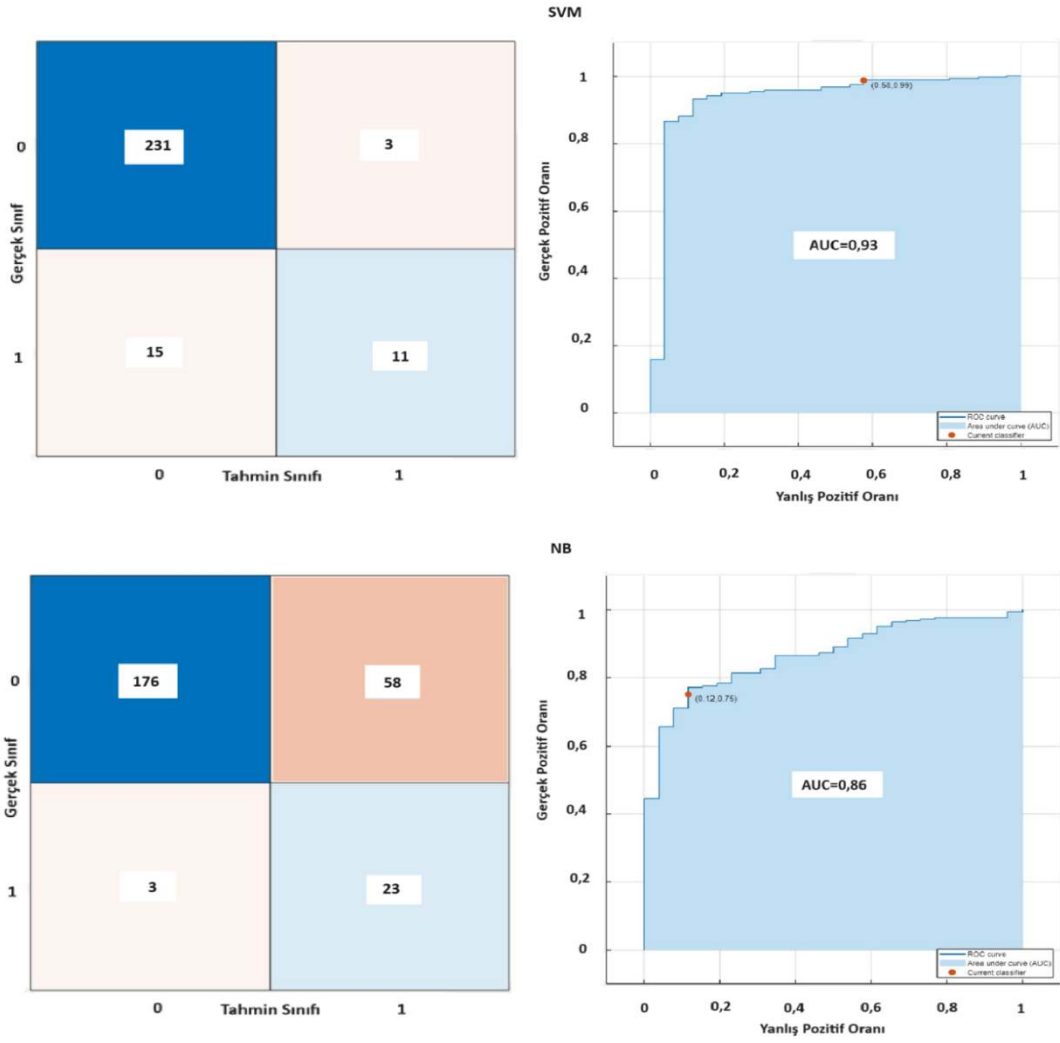
Şekil 17. 14 Özellik ile Elde Edilen Maksimum ve Minimum Oranlar.

Şekil 17’de Farklı 14 özelliğin tamamı kullanılarak en yüksek %96,5 doğruluk oranı ile KNN algoritması, %87,7 ile NB algoritması düşük performans göstermektedir. ROC eğrilerinden elde edilen en yüksek performansta AUC değeri 0,94 ile KNN olmaktadır. Yine de bu 14 özelliğin ayrı ayrı sonuçlara olan etkisini incelemede istatistiksel yöntem olarak pearson korelasyon analizi yapılarak en etkili özellikler belirlenmeye çalışılmıştır. Bu işlemde Kadın ve Erkek katılımcılar 2 sınıfa ayrılarak incelenmiştir.

Tablo 4. Pearson Korelasyon Analizi

Giriş Özelliği	P değeri
MNF	0,1679
MDF	0,1149
MAV	0,6425
RMS	0,6736
ZC	0,1391
WAMP	0,0472
WL	0,2108
SSC	0,0472
VAR	0,5610
MFL	0,3176
SD	0,6744
AAC	0,4766
DASD	0,5134

Tablo 4’te görüldüğü gibi 5 farklı özellik olan MAV, RMS, VAR, SD VE DASD özelliklerin sınıflandırma sonuçlandırmalarına etkisi yüksektir. Bu sebeple sadece bu 5 farklı özellikler kullanılarak analizler tekrarlanmıştır. Buna örnek olarak aynı şekilde 6.cı kişinin 14 özellik ile elde edilen sonuçlarını ele aldığımızda maksimum %93,1 doğruluk oranı ve de 0,93 AUC değeri ile SVM, en kötü performansta ise %76,5 doğruluk oranı ve de 0,86 AUC değeri ile NB algoritması göstermektedir. Bu durum Şekil 18’ de gösterilmektedir.

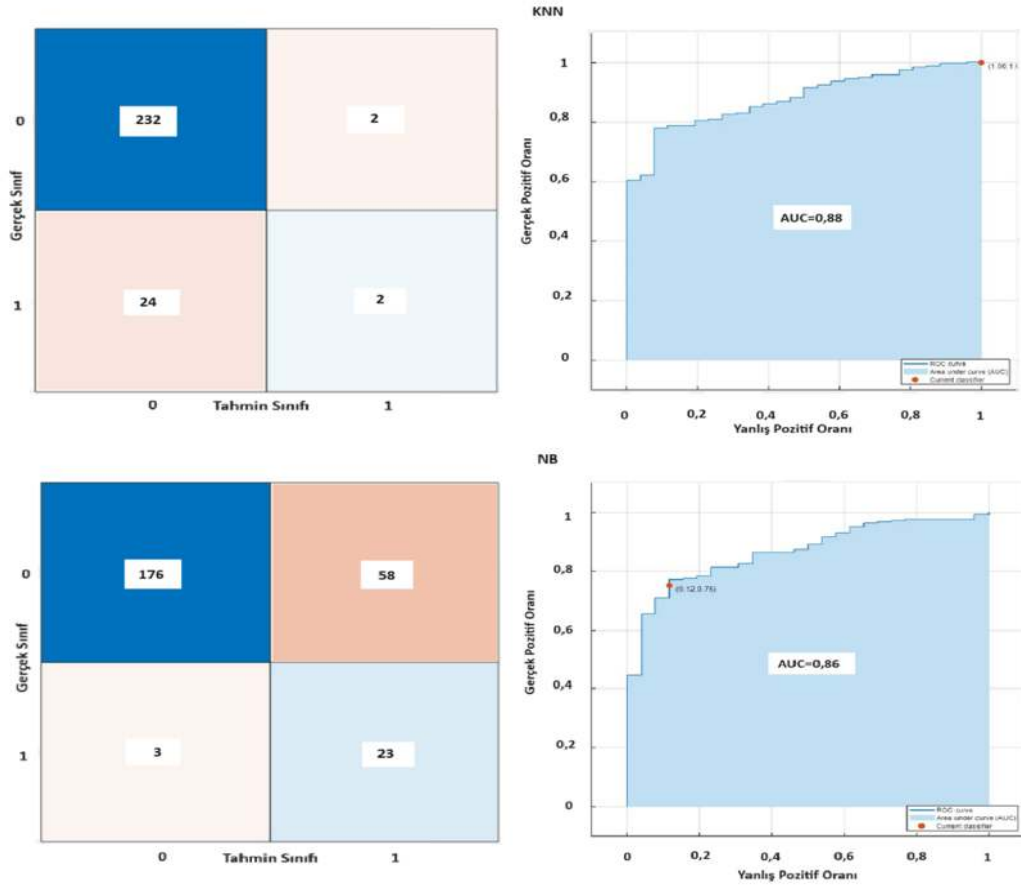


Şekil 18. 5 Özellik ile Elde Edilen Maksimum ve Minimum Oranlar.

Son olarak Tablo 4' e göre en etkin 2 özellik olan RMS değeri ile VAR değeri ile sınıflandırmalar yapıldığında KNN ve SVM algoritmalarının %90 performans verdiği görülmüştür (Şekil 19).

Buraya kadar elde edilen bulgulara ait grafiksel sonuçlar, yukarıda açıklanan özelliklere ait maksimum ve minimum performans değerlerini gösteren sonuçlardır ki bu sonuçlar, ROC ve karmaşıklık matrisleriyle gösterilmiştir. 6. katılımcıya ait farklı algoritmaları içeren tüm doğruluk performans oranları Tablo 5'te detaylandırılmıştır.

Tablo 5'te görüldüğü gibi özellik sayısı azalınca performans değerlerinde düşme gözlenmektedir. Ancak yine de etkin özellikler seçildiğinden dolayı düşüş miktarı yaklaşık %3 mertebelerinde sınırlı kalmaktadır.



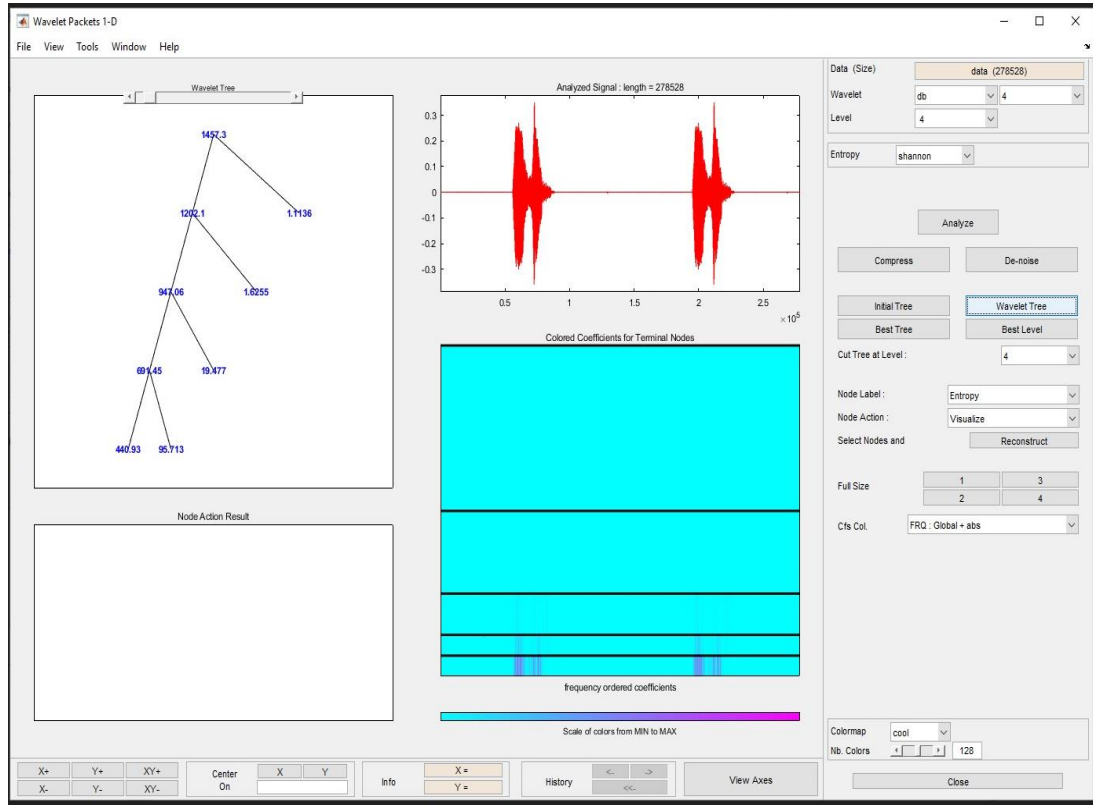
Şekil 19. 2 Özellik ile Elde Edilen Maksimum ve Minimum Oranlar.

Tablo 5. Zamana Bağlı Genlik Frekans Özelliklerine Ait Bulguların Karşılaştırmalı Sonuçları

Makine Öğrenme Algoritmaları	14 Özellik ile Doğruluk Oranları (%)	5 Özellik ile Doğruluk Oranları	
		Doğruluk Oranları (%)	2 Özellik ile Doğruluk Oranları (%)
DT (Fine Tree)	93.8	90.8	84.6
LD (Linear Discriminant)	90,4	93.1	89.2
NB (Kernel Naive Bayes)	87.7	76.5	76,5
VM (Linear SVM)	87.7	92.3	90.0
SVM (Quadratic SVM)	95.4	92.3	82.7
SVM (Fine Gausssian SVM)	94.6	93.1	90.0
SVM (Medium Gausssian SVM)	91.2	91.5	90.0
KNN (Fine KNN)	96.5	91.2	81.2
KNN (Medium KNN)	93.1	90.0	90.0
KNN (Coarse KNN)	90.0	90.0	90.0
KNN (Weighted KNN)	95.8	92.3	82.3

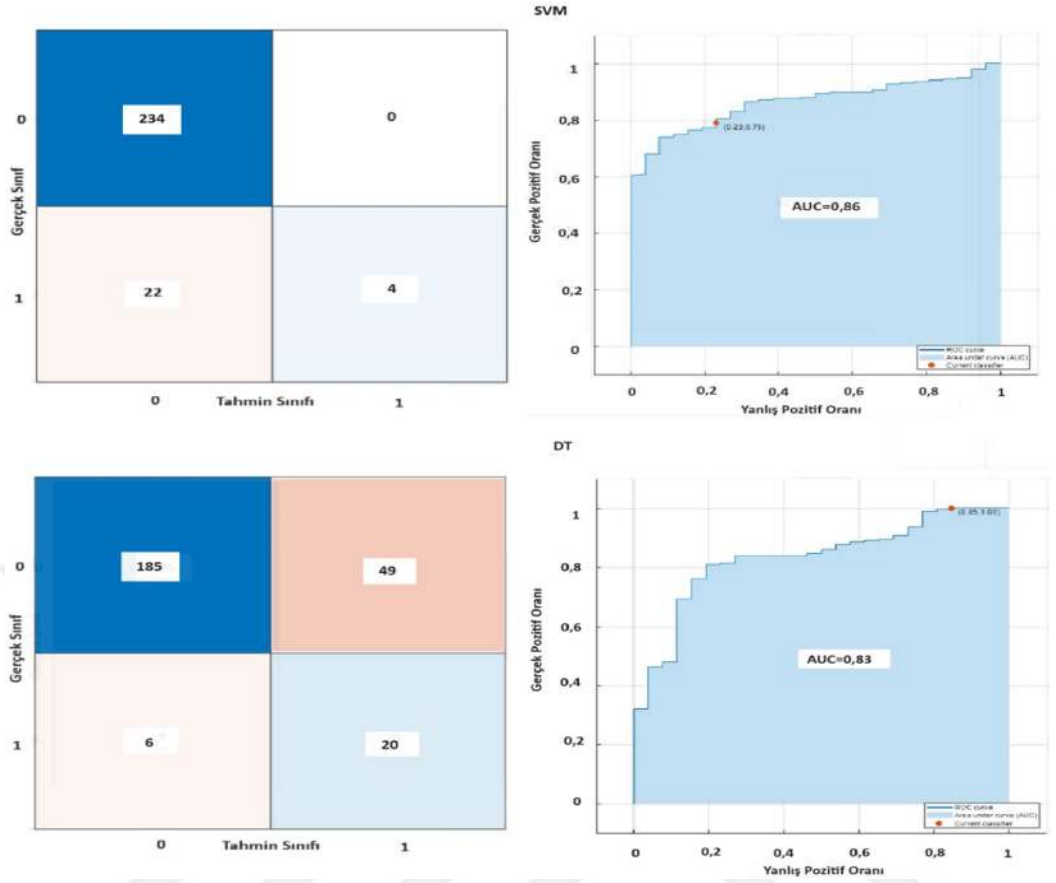
2.2. Ses Sinyalinin Dalgacık Dönüşümlerine Ait Analizler

Bu bölümde ayrıca aynı ses sinyallerinin dalgacık dönüşümleri metodu kullanılmasına yönelik bulgular yer almaktadır. Bu bölümde kullanılan dalgacık fonksiyonu Daubechies (db4) 4 ölçeklemeli fonksiyon kullanılmıştır. Elde edilen dalgacıklara ait Shannon Entropi özellikleri kullanılarak bu kez aynı makine algoritmalarıyla sonuçlar karşılaştırmalı olarak incelenmiştir. Şekil 20’de örnek bir ses sinyaline ait Matlab Wavelet 1-D toolbox kullanılarak elde edilen Shannon Entropileri ekran görüntüsü gösterilmektedir.



Şekil 20. Örnek Bir Sese Ait Dalgacık Shannon Entropi Analizi.

Bu çalışmada, dalgacık dönüşümlerinin 4 özellikli detay katsayıları 46 ve 47 numaralı denklemlere göre hesaplanmış, entropi özellikleri (db4) E1, E2, E3 ve E4 elde edilmiştir. Her bir kişiye ait bu dört özellik verisi makine öğrenme algoritmaları kullanılarak sınıflandırılmıştır. Bu özellikler kullanılarak maksimum %91,5 oranında SVM (Fine Gaussian) doğruluk oranı elde edilmiştir (Şekil 21). Ayrıca AUC değerleri bakımından maksimum 0,86 mertebelerinde olduğu gözlenmiştir (Şekil 21). Tablo 6’da ayrıca diğer algoritmalarla birlikte elde edilmiş performans sonuçları detaylandırılmıştır.



Şekil 21. Entropi (db4) ile Elde Edilen Maksimum ve Minimum Oranlar.

Tablo 6. Entropi (db4) Özelliklerine Ait Bulguların Karşılaştırmalı Sonuçları

Makine Öğrenme Algoritmaları	Dalgacık Shannon Entropi Özellikleri Doğruluk Oranları (%)
DT (Fine Tree)	86,2
DT (Medium Tree)	86,2
DT (RUSBoosted Tree)	78,8
LD (Linear Discriminant)	88,8
NB (Gaussian Naive Bayes)	81,9
NB (Kernel Naive Bayes)	83,5
SVM (Linear SVM)	90,4
SVM (Quadratic SVM)	91,2
SVM (Fine Gausssian SVM)	91,5
SVM (Medium Gausssian SVM)	90,0
KNN (Fine KNN)	86,9
KNN (Medium KNN)	90,0
KNN (Coarse KNN)	90,0
KNN (Weighted KNN)	90,0
YSA (Narrow Neural Network)	90,4

SONUÇ

Bu tez çalışması, ses sinyallerine ait kişi tanılaması için yapılan bir çalışmadır. Bu çalışmada sese ait zamana bağlı genlik ve frekans özellikleri tanılamada DT, SVM, NB, KNN, LD ve YSA algoritmaları kullanılarak detaylıca incelenmiştir. Performans sonuçları Karmaşıklık matrisi, ROC eğri grafikleriyle birlikte tablolar halinde detaylandırılmıştır.

Tablo 3'te görüldüğü üzere 14 özelliğin tamamı kullanılarak %99,2 ve %98,8 mertebelerinde SVM ve KNN sınıflayıcılarda yüksek doğruluk oranları elde edilmiştir. Ancak en etkili 5 özelliği kullandığımızda bu oran %93,1 seviyelerine kadar düşmektedir. Özellik sayısını 2'ye düşürdüğümüzde ise bu oran %90 mertebelerine gerilemektedir. Özellik indirgemesiyle gerçek uygulamalarda hızın artması sonucu zamandan tasarruf edilmesi sağlansa da doğruluk performansını azaltma yönünde bir eğilim söz konusudur. Diğer bir deyişle özellik sayısının artması da doğruluk performansını olumlu etkiler. Böylece bu tez çalışmasında elde edilen sonuçlarla, literatürde (Phinyomark, 2012) ve arkadaşlarının yaptıkları çalışmaların desteklendiği görülmektedir [67]. Ayrıca bu çalışmanın en etkili özelliklere indirgenerek tekrarlanması durumunda da, %90-%93 mertebelerinde etkili bir performans elde edilebilmektedir. Dolayısıyla sadece birkaç özellik kullanılarak, özellikle de endüstriyel ve akıllı ev uygulamalarında kişi tabanlı algılamalar için yeterli ve etkili çözümler üretilebileceği açıktır.

Literatürdeki çalışmalar kelime tabanlı sınıflamaya yöneliktir. Böylece kişilerin söyledikleri kelimeler kolaylıkla algılanabilmektedir. Bu tez çalışması literatürdeki MFCC, regresyon katsayıları, genlik frekans özellikleri (Ortalama Genlik, Standart Sapma, Maksimum, Minimum ve Medyan frekans v.b.) ile konuşmacıya yönelik emsal çalışmalara [9,15,16] kıyasla başarılı bir performans olarak kabul edilebilir.

Bu çalışma kişilerin ses tınlarını algılamaya diğer bir deyişle kişi ayırt etmek için yapılan bir çalışmadır. Ayrıca kullanılan genlik ve frekans özellikleri bakımından özgün ve detaylı bir çalışmadır. Yine de literatüre eksta bir katkı açısından çalışmayı bir adım öne taşımak adına dalgacık dönüşümleri kullanılarak her bir dalgacığın 4.cü seviyede Shannon entropi kazançları da ayrıca 6 farklı makine öğrenme

algoritmalarıyla birlikte ayrı ayrı incelenmiş, performans özellikleri Tablo 6’da detaylandırılmıştır. Burada maksimum SVM ile %91,5 oranında performans elde edildiği gözlenmektedir.

Sonuç olarak bu tez çalışması, endüstriyel ve ev tipi ses tabanlı uygulamaların geliştirilmesi için bizlere hangi özelliklerin daha kullanışlı olduğunu gösterme adına literatüre katkılar sağlamaktadır.



KAYNAKÇA

- [1] H. Kutucu, “*Derin öğrenme algoritmaları kullanarak bir konuşma tanıma uygulaması*”, Yüksek lisans tezi, Sakarya Uygulamalı Bilimler Üniversitesi, 2020.
- [2] C. Özbey ve S. Bayar, "Otomatik Ses Tanıma: Türkçe için Genel Dağarcıklı Akustik Model Oluşturulması ve Test Edilmesi," 19. Akademik Bilişim Konferansı (AB2017), Aksaray, Türkiye, 2017.
- [3] M. Rana, ve E. Saloni, “A review on automatic speech recognition system,” *International Journal of Engineering and Computer Science*, pp. 9849-9852, 2015.
- [4] S. S. Jarande, and S. Waghmare, “A survey on different classifier in speech recognition techniques,” *International Journal of Emerging Technology and Advanced Engineering*, pp. 534-539, 2014.
- [5] B. Korcuklu, “*Derin öğrenme tabanlı konuşma tanıma sistem tasarımı*”, Yayınlanmamış yüksek lisans tezi, Uludağ Üniversitesi, Bursa, 2021.
- [6] K. Özkan, E. Seke and Ş. Işık, “A new Approach for Speech Denoising”, 24th Signal Processing and Communication Application Conference (SIU), Zonguldak, Turkey, 16 -19 May 2016, pp.2109-2112. [Çevrimiçi]. Erişim: <https://api.semanticscholar.org/CorpusID:3209006>
- [7] M. Y. Çakır, “*Gerçek zamanlı yüksek kalitede ses tanıma*”, Yüksek lisans tezi, İstanbul Sabahattin Zaim Üniversitesi, İstanbul, 2017.
- [8] H. Prakoso, R. Ferdiana ve R. Hartanto. “Indonesian Automatic Speech Recognition system using CMUSphinx toolkit and limited dataset. Electronics and Smart Devices (ISESD),” International Symposium, 29-30 Nov. Bandung, Indonesia: IEEE, 2016.
- [9] R. M. Lexutan, “Comparative Study Regarding Characteristic Features of Human Voice” ECAI 2015-International Conference-7th. Edition (2015) Bucharest, Romania.
- [10] H. K. Palo, M. N. Mohanty, “Wavelet based feature combination for recognition of emotions”, *Shams Engineering Journal*, vol. 9, pp. 1799-1806, 2018.
- [11] Wei-Hua Cao and Jian-Ping Xu, “Speaker-independent Speech Emotion Recognition Based on Random Forest Feature Selection Algorithm”, IEEE 36th Chinese Control Conference, pp. 10995-10998, 2017.
- [12] A. Rajasekhar and M. K. Hota, “A Study of Speech, Speaker and Emotion Recognition using Mel Frequency Cepstrum Coefficients and Support Vector Machines”, IEEE 4th International Conference on Advanced Technologies for Signal and Image Processing, pp. 114-118, 2018.
- [13] Ö. Yakar ve R. Aşlıyan, “Saklı Markov Modeli Kullanarak Türkçe Konuşma Tanıma.” *Akademik Bilişim*, cilt 16, s. 1-7, 2016.

- [14] Y. Yurtcan ve B. Günel Kılıç, "Speech recognition on mobile devices in noisy environments," 26th Signal Processing and Communications Applications Conference (SIU), İzmir, Türkiye, 2-5 Mayıs 2018.
- [15] A. Charisma and M. Reza Hidayat, "Speaker Recognition Using Mel-Frequency Cepstrum Coefficients and Sum Square Error", IEEE 3rd International Conference on Wireless and Telematics, pp. 160-163, 2017.
- [16] A. Winursito and R. Hidayat, "Improvement of MFCC Feature Extraction Accuracy Using PCA in Indonesian Speech Recognition", IEEE International Conference on Information and Communications Technology, pp. 379-383, 2018.
- [17] A. Yayla, "*Konuşma tanıma teknolojisi kullanılarak devre tasarım ve analizi*", Doktora tezi, Marmara Üniversitesi, İstanbul, 2018.
- [18] A. Kartik "End to end ASR-free keyword search from speech." 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 4840-4844, 2017.
- [19] Y. Ü. Sönmez ve A. Varol, "A Speech Emotion Recognition Model Based on Multi-Level Local Binary and Local Ternary Patterns." IEEE Access 8, pp. 190784-190796, 2020.
- [20] Y. Ü. Sönmez ve A. Varol, "New Trends in Speech Emotion Recognition." 2019 7th International Symposium on Digital Forensics and Security (ISDFS), pp. 1-7, 2019.
- [21] A. Yılmaz Davutoğlu, "Standart Türkçedeki Ünlülerin Akustik Analizi ve Fonetik Altyapı", Sanatta yeterlik tezi, İstanbul Üniversitesi, 2010.
- [22] Y. Ü. Sönmez ve A. Varol, "In-Depth Analysis of Speech Production, Auditory System, Emotion Theories and Emotion Recognition," 2020 8th International Symposium on Digital Forensics and Security (ISDFS), Beirut, Lebanon, 2020, pp. 1-8. [Çevrimiçi], Erişim: doi:10.1109/ISDFS49300.2020.9116231.
- [23] S., Garima, S. Krishan. "Trends in audio signal feature extraction methods." *Applied Acoustics* 158, 107020, 2020, [Çevrimiçi]. Erişim: <https://doi.org/10.1016/j.apacoust.2019.107020>. [Erişim tarihi: 01.06.2023].
- [24] A. Phinyomark, P. Phukpattaranant, C. Limsakul. "Feature reduction and selection for EMG signal classification." *Expert Syst. Appl.*, vol. 39, pp. 7420-7431, 2012. [Çevrimiçi]. Erişim: <https://doi.org/10.1016/j.eswa.2012.01.102>
- [25] S. Kurt ve T. Sönmezocak, "Machine learning and regression analysis for diagnosis of bruxism by using EMG signals of jaw muscles," *Biomedical Signal Processing and Control*, vol. 69, pp. 102905, 2021
- [26] A. Phinyomark, C. Limsakul, and P. Phukpattaranont, "Application of Wavelet Analysis in EMG Feature Extraction for Pattern Classification." *Measurement Science Review*, vol. 1 issue 2, page 45-52, 2011. <https://doi.org/10.2478/v10048-011-0009-y>
- [27] M. Zecca, S. Micera, M. C. Carrozza, P. Dario, "Control of multifunctional prosthetic hands by processing the electromyographic signal," *Critical Reviews in Biomedical Engineering*, vol. 30, issue 4-6, pp. 459-485, 2002.

- [28] B. Hudgins, P. Parker and R. N. Scott, "A new strategy for multifunction myoelectric control," in *IEEE Transactions on Biomedical Engineering*, vol. 40, no. 1, pp. 82-94, 1993, doi: 10.1109/10.204774.
- [29] S. P., Arjunan, D. K., Kumar, & T. P. Jung, (2009). Changes in decibel scale wavelength properties of EEG with alertness levels while performing sustained attention tasks. Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference, pp. 6288–6291, 2009. <https://doi.org/10.1109/IEMBS.2009.5332801>
- [30] K. S. Kim, H. H. Choi, C.S. Moon, C. W. Mun, "Comparison of k-nearest neighbor, quadratic discriminant and linear discriminant analysis in classification of electromyogram signals based on the wrist-motion directions," *Current App. Phys.*, vol. 11, pp. 740-745, 2011. 10.1016/j.cap.2010.11.051.
- [31] E. L. Obilor, E. C. Amadi, (2018) "Test for significance of Pearson's correlation coefficient (r)", *International Journal of Innovative Mathematics, Statistics & Energy Policies*, vol 6, issue 1, page 11- 23.
- [32] P.C. Ahlgren, B. Jarneving, & R. Rousseau, "Requirements for a cocitation similarity measure, with special reference to Pearson's correlation coefficient". *J. Assoc. Inf. Sci. Technol.*, vol. 54, pp. 550-560, 2003.
- [33] J. Malmivuo, and R. Plonsey, *Bioelectromagnetism: principles and applications of bioelectric and biomagnetic fields*. Oxford University Press, New York, 1995.
- [34] Y. Özkan, *Veri madenciliği yöntemleri*, Papatya Yayıncılık, 2008.
- [35] J. Yang, J. Y. Yang, "Why can LDA be performed in PCA transformed space?," *Pattern Recognition*, vol. 36, pp. 563-566, 2003.
- [36] B. Karlik, "Machine learning algorithms for characterization of EMG signals," *International Journal of Information and Elec. Eng.*, vol. 4, pp. 189-194, 2014.
- [37] M. Coşkun, "Dört rotorlu insansız hava aracı ile sürekli uyarlamalı ortalama kayma algoritması," Yüksek lisans tezi, Fırat Üniversitesi, Elazığ, 2015.
- [38] M. O. Akı, "Sürücü uykuğunun gerçek zamanlı görüntü işleme ve makine öğrenmesi teknikleri ile tespitiye yönelik bir sistem tasarımı ve uygulaması," Doktora tezi, Trakya Üniversitesi, Edirne, 2017.
- [39] E. Alpaydin, *Introduction to machine learning*, MIT press, 2014.
- [40] A. McCallum and K. Nigam, "A comparison of event models for naive bayes text classification," *AAAI-98 workshop on learning for text categorization*, vol. 752, pp. 41-48, 1998.
- [41] I. Rish, "An empirical study of the naive Bayes classifier," *IJCAI 2001 workshop on empirical methods in artificial intelligence*, vol. 3, no. 22, pp. 41-46, 2001.
- [42] G. H. John, and P. Langley, "Estimating Continuous Distributions in Bayesian Classifiers," *Proceedings of the 11th Conference on Uncertainty in Artificial Intelligence*, San Francisco, 1995, pp. 338-345.

- [43] E. Dönmez, “*Designing controllers for path planning applications to mobile robots with head-cameras*,” Ph.D. Thesis, İnönü University, Malatya, 2018.
- [44] Y. Atınç, *Yapay Zeka*, Kodlab Yayınevi, İstanbul, 2022.
- [45] J. J. Hopfield, “Neural Networks and physical systems with emergent collective computational abilities,” *Proceedings National Academy of Science*, Vol. 79, pages 2554-2558, 1982.
- [46] M. Ö. Efe ve O. Kaynak, *Yapay Sinir Ağları ve Uygulamaları*, Boğaziçi Üniversitesi Yayınevi, İstanbul, 2000.
- [47] Ç. Elmas, *Yapay Sinir Ağları*, Seçkin Yayıncılık, İstanbul, 2003.
- [48] T. Sönmezocak, “*Makine öğrenme algoritmalarıyla alt çeneye bağlı kasların yorgunluk analizi ve bruksizm tani yöntemlerinin geliştirilmesi*”, Doktora tezi, Yıldız Teknik Üniversitesi, İstanbul, 2022.
- [49] Mallat, S. A, “Theory for Multiresolution Signal Decomposition: The Wavelet Representation”. *IEEE Trans. Pattern Anal. Mach. Intell.*, 11, 674-693, 1989.
- [50] S. Haykin, “Neural Networks, A Comprehensive foundation”, Prentice Hall, 2004.
- [51] E. Yıldız, “*Doppler işaretlerinin dalgacık dönüşümü ve fraktal boyut kullanarak yapay sinir ağları ile sınıflandırılması*”, Yüksek lisans tezi, Fırat Üniversitesi, Elazığ, 2009
- [52] N. Türker ve F. Güneş, “Mikroşerit Dalga Kılavuzu Transmisyonu için Yapay Sinir Ağı Modeli Geliştirilmesi”, *Akıllı Sistemlerde Yenilikler ve Uygulamaları Sempozyumu*, İstanbul-Türkiye, 2004
- [53] E. Öztemel, *Yapay Sinir Ağları*, Papatya Yayıncılık, İstanbul, 2003.
- [54] Z. Man, H. R. Wu, S. Liu, & X. Yu, (2006). A new adaptive backpropagation algorithm based on Lyapunov stability theory for neural networks. *IEEE Transactions on Neural Networks*, vol. 17(6), Pages 1580-1591.
- [55] T. L. Fine, *Feedforward Neural Network Methodology*, Springer Science & Business Media, 2006.
- [56] K. Akyol ve A. Karacı, “Diyabet Hastalığının Erken Aşamada Tahmin Edilmesi İçin Makine Öğrenme Algoritmalarının Performanslarının Karşılaştırılması”, *DÜBİTED*, cilt. 9, sy. 6, ss. 123–134, 2021, doi: 10.29130/dubited.1014508.
- [57] J. Maria Navin and R. Pankaja, “Performance analysis of text classification algorithms using confusion matrix,” *Int. J. Eng. Tech. Res. IJETR*, vol. 6, pp. 75–78, 2016.
- [58] T. Fawcett. An introduction to ROC analysis. *Pattern Recognit. Lett.*, 27, 861-874, 2006.
- [59] A. Tharwat. *Classification assessment methods*. Applied Computing and Informatics, 2018.

- [60] G. Güler, “Makine öğrenme algoritmaları ile elektroensefalografi analizi ve epilepsi tanı yöntemlerinin geliştirilmesi”, Yüksek lisans tezi, Doğu Üniversitesi, İstanbul, 2023.
- [61] J. L. Semmlow. *Biosignal and biomedical image processing: MATLAB-based applications*. John L. Semmlow. [ed.]. Marcel Dekker., 2004.
- [62] İ. Türkoğlu, “*Durağan olmayan işaretler için zaman-frekans entropilerine dayalı akıllı örüntü tanıma*,” Doktora tezi, Fırat Üniversitesi, Elazığ, 2002.
- [63] Mısıtı M., Missiti Y., Oppenheim G., Pogg J.M., *Wavelet Toolbox For Use With Matlab, User's Guide*, The Mathworks Inc., 1997-2002
- [64] N. Arı, Ş. Özen, Ö. H. Çolak, *Dalgacık Teorisi (Wavelet)*. Palme yayıncılık, Ankara, 2008
- [65] B. Karlık, Y. Koçyiğit, B. C. Fidan, “EMG işaretlerin sınıflamada kullanılan işaret işleme tekniklerinin karşılaştırılması”, SIU 2005, 2005, Kayseri.
- [66] I. Daubechies, “*Ten lectures on wavelets*,” Society for Industrial and Applied Mathematics, Pennsylvania, 1992, doi:10.1137/1.9781611970104.
- [67] A. Phinyomark, P. Phukpattaranont, C. Limsakul, “Feature reduction and selection for EMG signal classification,” *Expert Systems with App.*, vol. 39, pp. 7420-7431, 2012, doi:10.1016/j.eswa.2012.01.102.

EKLER

EK-1: Etik Kurul Raporu

Evrak Tarih ve Sayısı: 17.07.2024-64718



DOĞU ÜNİVERSİTESİ Etik Kurulu

Sayı : E-42435178-050.04-64718
Konu : Koray Öztürk'ün Etik Kurul Raporu

17.07.2024

REKTÖRLÜK MAKAMINA

Etik Kurulu'nun 05.07.2024 tarihinde gerçekleştirmiş olduğu toplantıda, Koray ÖZTÜRK'ün araştırması için yapmış olduğu başvuruya ilgili olmuş olduğu kararı ekte bilgi ve onaylarınıza saygılarımla arz ederim.

Prof. Dr. Mesut KUMRU
Kurul Başkanı

Ek:Koray ÖZTÜRK'ün Etik Kurul Raporu (11 Sayfa)

Bu belge, güvenli elektronik imza ile onaylanmıştır.

Belge Doğrulama Kodu :BS5333Z049 Pin Kodu :97722
Adres:Duhalla OSB Mah. Nato Yolu Cad. 265/1 Ümraniye İstanbul
Telefon:Tel: 444 79 97 Faks:225 327 96 31
e-Posta:info@dogu.edu.tr Web:www.dogu.edu.tr
Kep Adresi: doguuniversitesi@tu01.kep.tr

Belge Takip Adresi : <https://turkiye.gov.tr/dogu-universitesi-ebys>

Bilgi için: Aymer ŞAFAR
Ünvanı Raporlar



Bu belge, güvenli elektronik imza ile onaylanmıştır.

05.07.2024

Doğuş Üniversitesi Rektörlüğüne,

İlgi: Koray ÖZTÜRK'ün araştırması için Etik Kurul Başkanlığına yapılan 04.07.2024 tarihli başvuru.

Lisansüstü Eğitim Enstitüsü Elektronik ve Haberleşme Mühendisliği Anabilim Dalı Elektronik ve Haberleşme Mühendisliği Tezli Yüksek Lisans Programı öğrencisi Koray ÖZTÜRK'ün Dr. Öğr. Üyesi Temel SÖNMEZÖCAK danışmanlığı altında yürütülecek olan **"Makine Öğrenme Algoritmalarıyla Konuşmacı Tanılaması"** başlıklı araştırması için yapacağı ankette kullanacağı soruların bilimsel araştırma etiği açısından **uygun olduğu** tespit edilerek ilişik rapor hazırlanmıştır.

Konuyu bilginize sunar, ilişik raporun araştırmacıya ulaştırılmak üzere Lisansüstü Eğitim Enstitüsü Müdürlüğüne gönderilmesini saygı ile arz ederim.

Prof. Dr. Mesut KUMRU
Etik Kurul Başkanı

Eki: 1 adet Etik Kurul Raporu

05.07.2024

Sayı :2024/107
Konu : Etik Kurul izni

Koray ÖZTÜRK
Doğuş Üniversitesi
Lisansüstü Eğitim Enstitüsü
Elektronik ve Haberleşme Mühendisliği Tezli Yüksek Lisans Programı

Doğuş Üniversitesi Lisansüstü Eğitim Enstitüsü Elektronik ve Haberleşme Mühendisliği Anabilim Dalı Elektronik ve Haberleşme Mühendisliği Tezli Yüksek Lisans Programı çerçevesinde Dr. Öğr. Üyesi Temel SÖNMEZOCAK danışmanlığı altında yürütülecek olan **"Makine Öğrenme Algoritmalarıyla Konuşmacı Tanılaması"** başlıklı araştırmanız için yapacağınız ankete ait sorularla ilgili olarak Doğuş Üniversitesi Etik Kuruluna yapmış olduğunuz başvuru Etik Kurulunda incelenmiş ve bilimsel araştırma etiği açısından **uygun olduğuna** karar verilmiştir.

Bilgilerinizi rica ederiz.

Prof. Dr. Mesut KUMRU
Etik Kurul Başkanı
Mühendislik Fakültesi

Prof. Dr. Hamide ZAFER
Etik Kurul Üyesi
Hukuk Fakültesi

Prof. Dr. Oktay VELİEV
Etik Kurul Üyesi
Mühendislik Fakültesi

Prof. Dr. Nüket SARACEL
Etik Kurul Üyesi
İktisadi ve İdari Bilimler Fakültesi

Prof. Dr. Gülsün Pelin SARIOĞLU ERDOĞDU
Etik Kurul Üyesi
Sanat ve Tasarım Fakültesi

Prof. Dr. Tanık BAYKARA
Etik Kurul Üyesi
Mühendislik Fakültesi

Dr. Öğr. Üyesi Aslı TELSEREN ÖMEROĞLU
Etik Kurul Üyesi
Fen Edebiyat Fakültesi



DOĞUŞ ÜNİVERSİTESİ
Lisansüstü Eğitim Enstitüsü

Sayı : E-22716164-100-64646
Konu : Koray Öztürk'ün Etik Kurul Başvurusu Hk.

04.07.2024

REKTÖRLÜK MAKAMINA

Enstitümüz Elektronik Ve Haberleşme Mühendisliği Anabilim Dalı Elektronik Ve Haberleşme Mühendisliği Yüksek Lisans Programı 202191028002 numaralı öğrencisi Koray ÖZTÜRK'ün "Makine Öğrenme Algoritmalarıyla Konuşmacı Tanılaması" başlıklı tezinin, araştırma sırasında yapmış olduğu güncel çalışmanın bilimsel araştırma etiği açısından uygunluğunun tespiti için etik kurul onayına sunulması hususunda bilgilerinize saygılarımla arz ederim.

Dr. Öğr. Üyesi Ahmet Yekta KAYMAN
Müdür

Ek:Başvuru Evrakları

Bu belge, güvenli elektronik imza ile onaylanmıştır.

Belge Doğrulama Kodu :BSA33DSKJN Pin Kodu :98582
Adres:Dudulla OSB Mah. Nato Yolu Cad. 265/1 Ümraniye İstanbul
Telefon:Tel: 444 79 97 Faks:0216 327 96 31
e-Posta:info@dogus.edu.tr Web:www.dogus.edu.tr
Kep Adresi: dogusuniversitesi@hs01.kep.tr

Belge Takip Adresi : <https://turkiye.gov.tr/dogus-universitesi-ehys>

Bilgi için: Seren ÖZDEMİR
Uzman Öğrenci İşleri Uzmanı



Bu belge, güvenli elektronik imza ile imzalanmıştır.



DOĞUŞ ÜNİVERSİTESİ ETİK KURUL BAŞVURU DOSYASI

03/07/2024

A. ETİK KURUL ONAYI TARAMA FORMU

TEKLİF EDİLEN ARASTIRMA	EVET	HAYIR
Çocukları içermekte midir?		X
Hastalara zarar verebilecek (invaziv) bir tetkik içermekte midir?		X
Muvafakat veremeyecek kişileri içermekte midir?		X
Sağlıklı yetişkin bireyleri mi içermektedir?	X	
Hastaları içermekte midir?		X
İnsan genetik materyali içermekte midir?		X
İnsan biyolojik örnekleri içermekte midir?		X
İnsandan veri toplamayı içermekte midir?	X	
Genetik veya kişisel bilgilerin (sağlık, cinsel hayat tarzı, etnik kimlik, siyasi görüş, dini ya da felsefi kanaat, vb.) işlenmesini içermekte midir?		X
İnsanların yerini belirtmeyi veya insanları gözlemlemeyi içermekte midir?		X

Yukarıdaki formda EVET işaretlenen her nokta hakkında etik gereklilikleri (zarar/yazar dengesi, gizliliğin korunması, hassas grupların – çocuk, zihinsel engelli kişiler, vb.- bilgilendirilmiş karar verme haklarının korunması gibi) yerine getirmek için önerilen yöntemler "Araştırmanın Metodu ve Prosedürü" başlığı altında detaylandırılmalıdır.¹

B. ARASTIRMA BİLGİLERİ

- Projenin Adı (Yüksek Lisans Tezi)
MAKİNE ÖĞRENME ALGORİTMALARIYLA KONUŞMACI TANILMASI
- Projenin Yürütücüsü
Kerem ÖZTÜRK (Yazar)
- Diğer Araştırmacılar

- Araştırmanın Amacı

Teknolojinin her geçen gün sürekli gelişmesiyle birlikte, hayatımıza giren otomatik sistem ve cihazların sayısı her geçen gün artmaktadır. Bu yeni teknolojiler ile kontrol edilen sistem ve cihazların güvenliği, kontrolü, sevk ve idare edilmesi gibi durumlar için ayırt edici farklı kontrol mekanizmalarına ihtiyaç duyulmuştur. Bu arayış içerisinde kişiden kişiye değişen ses özellikleri ayırt etmeye yarayan yöntemler önem kazanmıştır. Özellikle de kişi taklit seslerini kullanarak yapılan ihlaller neticesinde hırsızlık ve suç teşkil eden olaylar yaygınlaşmaktadır. Bu durumun önüne geçebilmek için bu çalışma planlanmaktadır. Dolayısıyla kişilerden alınan seslerin ayırt edici özelliklerini belirlemeye yönelik bir çalışmadır.

- Araştırmanın Süresi
Araştırma 2 yıldır devam etmektedir. Ancak analizler henüz yapılmamıştır. Kalan süre 3 ay olmaktadır.
- Destekleyen Kuruluş

- Araştırmanın Metodu ve Prosedürü

Bu çalışmada yaklaşık 10 ya da 15 farklı kişilerden harici mikrofon aracılığıyla Türkçe de yer alan bir hirinden bağımsız farklı kelimeler kişilere söylenecek. Söylenecek kelimeler sebze , meyve v.b. cins isim şeklinde olacaktır. Kişilerin hakkında ya da siyasal anlam içerebilecek toplumun güvenliğini ya da kişilere ait gizli bilgileri içeren kelimeler kullanılmayacaktır. Kişilerin rızası dikkate alınarak yapılacak, hiçbir kişiye zorlama ya da baskı yapılmayacaktır. Ayrıca fiziksel olarak vücut ile ilgili fiziksel hiçbir bilgi alınmayacaktır. Kişilerden alınacak bu sesler sadece bu tez kapsamında sinyal işleme yöntemleriyle değerlendirilip bilimsel açıdan kişi tanımlaması konusunda literatüre katkıda bulunması hedeflenmektedir. Çıkar çatışması bakımından ticari bir alanda kesinlikle kullanılmayacaktır.

C. TAMAMLAYICI BİLGİLER

1. Gizlilik

Vücutta Lisans Tezi olarak sunulan bu çalışmanın, bilimsel etik ve geleneklere uygun şekilde devam etmektedir, bu tezdaki bütün bilgileri akademik ve etik kurallara dikkate alınarak gizlilik ilkesine bağlıdır ve de patent ve telif haklarını ihlal edici bir davranışın olmuyacağı yazan/yazıların tarafından beyan edilmiştir.

2. Deneklerin Katılımının Potansiyel Riskleri

Deneklerin katılımında potansiyel hiçbir risk yoktur.

3. Deneklerin Katılımının Potansiyel Yararları

Deneklerin katılımının toplum için yararlı olabilecek çalışmaya gönüllü olarak katılımın dışında potansiyel bir maddi yarar ya da beklentisi yoktur.

D. KAYNAKÇA

- [1] H. Kutucu, "Derin öğrenme algoritmaları kullanarak bir konuşma tanıma uygulaması", Yüksek lisans tezi, Elektrik-Elektronik Mühendisliği Bölümü, Sakarya Uygulamalı Bilimler Üniversitesi, Sakarya, 2020.
- [2] C. Ozbey ve S. BAYAR, "Otomatik Ses Tanıma: Türkçe için Genel Dağırcıklı Akustik Model Oluşturulması ve Test Edilmesi," 19. Akademik Bilişim Konferansı (AB2017), Akaraz, Türkiye, 2017.
- [3] Rana, M. ve Saloni, E., A review on automatic speech recognition system, International Journal Of Engineering And Computer Science, 9849-9852, (2015).
- [4] Jarande, S.S. ve Waghmare, S., A survey on different classifier in speech recognition techniques, International Journal of Emerging Technology and Advanced Engineering, 534-539, (2014).
- [5] C. Hanilci ve F. Ertas, "Sürekli Saklı Markov Modelleri ile Metinden Bağımsız Konuşmacı Tanıma Parametrelerinin İncelenmesi", UWF, c. 12, sy. 1, 2007, doi: 10.17482/uwf.95782.
- [6] B. Korkuklu, "Derin öğrenme tabanlı konuşma tanıma sistem tasarımı", Yayınlanmamış yüksek lisans tezi, Bilgisayar Mühendisliği Bölümü, Bursa Uludağ Üniversitesi Fen Bilimleri Enstitüsü, Bursa, 2021.

- [7] K. Özkan, E. Seke and Ş. Işık, "A new Approach for Speech Denoising", 24th Signal Processing and Communication Application Conference (SIU), Zonguldak, Turkey, 16 -19 May 2016, pp.2109-2112. [Çevrimiçi] <https://api.semanticscholar.org/CorpusID:2200006> [Erişim: 01.06.2023].
- [8] M. Y. Çakar, "Gerçek zamanlı yüksek kalitede ses tanıma", Yüksek lisans tezi, Bilgisayar Mühendisliği Bölümü, İstanbul Sabahattin Zaim Üniversitesi Fen Bilimleri Enstitüsü, İstanbul, 2017.
- [9] Prakoso, H., Ferdiana, R., & Hartanto, R. (2016). Indonesian Automatic Speech Recognition system using CMUSphinx toolkit and limited dataset. *Electronics and Smart Devices (ISFSD), International Symposium*, 29-30 Nov. Bandung, Indonesia: IEEE.
- [10] Ö. Yakar ve R. Aşlıyan, (2016). Saklı Markov Modeli Kullanarak Türkçe Konuşma Tanıma. *Akademik Bilişim*, 16, 1-7.
- [11] Yurtcan Y. and Günel Kılıç B.. Speech recognition on mobile devices in noisy environments. 26th Signal Processing and Communications Applications Conference (SIU), İzmir, Türkiye, 2-5 Mayıs 2018.
- [12] A. Yayla, "Konuşma tanıma teknolojisi kullanılarak devre tasarımı ve analizi", Doktora tezi, Elektronik-Bilgisayar Eğitimi Anabilim Dalı, Elektronik-Haberleşme Eğitimi Programı, Marmara Üniversitesi, İstanbul, 2018.
- [13] Audhkhasi, Kartik et al. "End to end ASR-free keyword search from speech." 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (2017): 4840-4844.
- [14] Sönmez, Yeşim Ülgen and Asaf Varol. "A Speech Emotion Recognition Model Based on Multi-Level Local Binary and Local Ternary Patterns." *IEEE Access* 8 (2020): 190784-190796.
- [15] Sonmez, Yesim Ulgen and Asaf Varol. "New Trends in Speech Emotion Recognition." 2019 7th International Symposium on Digital Forensics and Security (ISDFS) (2019): 1-7.
- [16] A. Yılmaz Davutoğlu, "Standart Türkçedeki Ünlülerin Akustik Analizi ve Fonetik Altyapı", Sanatta yeterlik tezi, Sosyal Bilimler Enstitüsü, Sahne Sanatları Anasanat Dalı, Tiyatro Sanat Dalı, İstanbul Üniversitesi, İstanbul, 2010.
- [17] Y. Ü. Sönmez and A. Varol, "In-Depth Analysis of Speech Production, Auditory System, Emotion Theories and Emotion Recognition," 2020 8th International Symposium on Digital Forensics and Security (ISDFS), Beirut, Lebanon, 2020, pp. 1-8.[Çevrimiçi]. <https://doi.org/10.1109/ISDFS49300.2020.9116231>. [Erişim: 02.06.2023].
- [18] Sharma, Garima et al. "Trends in audio signal feature extraction methods." *Applied Acoustics* 158 (2020): 107020. [Çevrimiçi]. <https://doi.org/10.1016/j.apacoust.2019.107020>. [Erişim: 01.06.2023].
- [19] Phinyomark, Angkoon et al. "Feature reduction and selection for EMG signal classification." *Expert Syst. Appl.* 39 (2012): 7420-7431. [Çevrimiçi]. <https://doi.org/10.1016/j.eswa.2012.01.102> [Erişim: 02.06.2023].
- [20] S. Kurt And T. Sönmezocak, "Machine learning and regression analysis for diagnosis of bruxism by using EMG signals of jaw muscles," *Biomedical Signal Processing and Control*, vol.69, pp.102905, 2021.

- [21] Phinyomark, A., Limsakul, C. & Phukpattaranont, P. (2011). Application of Wavelet Analysis in EMG Feature Extraction for Pattern Classification. *Measurement Science Review*, 11(2) 45-52. <https://doi.org/10.2478/v10048-011-0009-y>
- [22] M. Zecca, S. Micera, M. C. Carrozza, P. Dario, "Control of multifunctional prosthetic hands by processing the electromyographic signal," *Critical Reviews in Biomedical Engineering*, pp. 459-485, 2002.
- [23] B. Hudgins, P. Parker and R. N. Scott, "A new strategy for multifunction myoelectric control," in *IEEE Transactions on Biomedical Engineering*, vol. 40, no. 1, pp. 82-94, Jan. 1993, doi: 10.1109/10.204774.
- [24] Arjunan, S. P., Kumar, D. K., & Jung, T. P. (2009). Changes in decibel scale wavelength properties of EEG with alertness levels while performing sustained attention tasks. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society, Annual International Conference, 2009*, 6288-6291. <https://doi.org/10.1109/IEMBS.2009.5332801>
- [25] K. S. Kim, H. H. Choi, C. S. Moon, C. W. Mun, "Comparison of k-nearest neighbor, quadratic discriminant and linear discriminant analysis in classification of electromyogram signals based on the wrist-motion directions," *Current App. Phys.*, vol. 11, pp. 740-745, 2011. 10.1016/j.cap.2010.11.051.
- [26] H. Akpınar, *Data Veri Madenciliği- Veri Analizi*, Papatya Yayıncılık, İstanbul: Marmara Üniversitesi İşletme Fakültesi, 2014.
- [27] Ahlgren, P.C., Jarneving, B., & Rousseau, R. (2003). Requirements for a cocitation similarity measure, with special reference to Pearson's correlation coefficient. *J. Assoc. Inf. Sci. Technol.*, 54, 550-560.
- [28] E. Alpaydm, *Introduction to machine learning*, MIT press, 2014.
- [29] Malmivuo, J. and Plonsey, R. (1995) *Bioelectromagnetism: principles and applications of bioelectric and biomagnetic fields*. Oxford University Press, New York.
- [30] Y. Özkan, *Veri madenciliği yöntemleri*, Papatya Yayıncılık, 2008.
- [31] J. Yang, J. Y. Yang, "Why can LDA be performed in PCA transformed space?," *Pattern Recognition*, vol. 36, pp. 563-566, 2003.
- [32] B. Karlik, "Machine learning algorithms for characterization of EMG signals," *International Journal of Information and Elec. Eng.*, vol. 4, pp. 189-194, 2014.
- [33] M. Coşkun, "Dört rotalı insansız hava aracı ile sürekli uyarılamalı ortalama kayma algoritması," Yüksek lisans tezi, Elektrik Elektronik Mühendisliği Bölümü, Fırat Üniversitesi, Elazığ, 2015.
- [34] M. O. Aka, "Sürücü uykuğunun gerçek zamanlı görüntü işleme ve makine öğrenmesi teknikleri ile tespiti için bir sistem tasarımı ve uygulaması," Doktora tezi, Bilgisayar Mühendisliği Bölümü, Trakya Üniversitesi, Edirne, 2017.
- [35] T. Sönmezocak, "Makine öğrenme algoritmalarıyla alt çeneye bağlı kasların yorgunluk analizi ve brüksizm tanı yöntemlerinin geliştirilmesi", Doktora tezi, Elektronik ve Haberleşme Mühendisliği Bölümü, Yıldız Teknik Üniversitesi, İstanbul, 2022.
- [36] A. McCallum and K. Nigam, "A comparison of event models for naïve bayes text classification," in *AAAI-98 workshop on learning for text categorization*, 1998, vol. 752, pp. 41-48.

- [37] I. Rish, "An empirical study of the naive Bayes classifier," in *IJCAI 2001 workshop on empirical methods in artificial intelligence*, 2001, vol. 3, no. 22, pp. 41-46.
- [38] G. H. John and P. Langley, "Estimating Continuous Distributions in Bayesian Classifiers," *Proceedings of the 11th Conference on Uncertainty in Artificial Intelligence*, San Francisco, 1995, pp. 338-345.
- [39] E. Dönmez, "Designing controllers for path planning applications to mobile robots with head-cameras," Ph.D. Thesis, Computer Engineering Department, İnönü University, Malatya, 2018.
- [40] Y. Atıncı, "Yapay Zeka," Kodlab Yayınevi, İstanbul, 2022.
- [41] J. J. Hopfield, "Neural Networks and physical systems with emergent collective computational abilities," *Proceedings National Academy of Science*, vol. 79, pp. 2554-2558, 1982.
- [42] M. Ö. Efe ve O. Kaynak, "Yapay Sinir Ağları ve Uygulamaları," Boğaziçi Üniversitesi Yayınevi, İstanbul, 2000.
- [43] Ç. Elmas, "Yapay Sinir Ağları," Seçkin Yayıncılık, İstanbul, 2003.
- [44] Mallat, S. (1989). A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 11, 674-693.
- [45] S. Haykin, "Neural Networks, A Comprehensive foundation", Prentice Hall, 2004.
- [46] E. Yıldız, "Doppler işaretlerinin dalgacık dönüşümü ve fraktal boyut kullanarak yapay sinir ağları ile sınıflandırılması", Yüksek lisans tezi, Elektronik ve Bilgisayar Eğitimi Anabilim dalı, Fırat Üniversitesi, Elazığ, 2009
- [47] N. Türker ve F. Güneş, "Mikroşerit Dalga Kılavuzu Transmisyonu için Yapay Sinir Ağı Modeli Geliştirilmesi", *Akıllı Sistemlerde Yenilikler ve Uygulamaları Sempozyumu*, İstanbul-Türkiye, 2004
- [48] E. Öztemel, "Yapay Sinir Ağları," Papatya Yayıncılık, İstanbul, 2003.
- [49] Man, Z., Wu, H. R., Liu, S., & Yu, X. (2006). A new adaptive backpropagation algorithm based on Lyapunov stability theory for neural networks. *IEEE Transactions on Neural Networks*, 17(6), 1580-1591.
- [50] T. L. Fine, *Feedforward Neural Network Methodology*, Springer Science & Business Media, 2006.
- [51] K. Akyol ve A. Karacı, "Diyabet Hastalığının Erken Aşamada Tahmin Edilmesi için Makine Öğrenme Algoritmalarının Performanslarının Karşılaştırılması", *DÜBİTED*, c. 9, sy. 6, ss. 123-134, 2021, doi: 10.29130/dubited.1014508.
- [52] J. Maria Navin and R. Pankaja, "Performance analysis of text classification algorithms using confusion matrix," *Int. J. Eng. Tech. Res. UCTH*, vol. 6, pp. 73-78, 2018.
- [53] Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognit. Lett.*, 27, 861-874.
- [54] Tharwat, A. (2020). Classification assessment methods. *Applied Computing and Informatics*.
- [55] G. Güler, "Makine öğrenme algoritmaları ile elektroensefalografi analizi ve epilepsi tanı yöntemlerinin geliştirilmesi", Yüksek lisans tezi, Elektrik Elektronik Mühendisliği Bölümü, Doğuş Üniversitesi, İstanbul, 2023.

- [56] Semmlow, J. L. (2004). *Biomedical and biomedical image processing: MATLAB-based applications* / John L. Semmlow. [Map]. Marcel Dekker.
- [57] İ. Türkoğlu, "Durağan olmayan işaretler için zaman-frekans entropilerine dayalı akülü görüntü tanıma," Doktora tezi, Fırat Üniversitesi Fen Bilimleri Enstitüsü, Fırat Üniversitesi, Elazığ, 2002.
- [58] Mısırlı M., Mısırlı Y., Oppenheim G., Pogg J.M., Wavelet Toolbox For Use With Matlab, User's Guide. The Mathworks Inc.,1997-2002
- [59] N. An, Ş. Özen, Ö. H. Çolak, "Dalgacık Teorisi (Wavelet)," Palme yayıncılık, Ankara, 2008
- [60] B. Karlık, Y. Koçyiğit, B. C. Fidan, "EMG işaretlerinin sınıflamada kullanılan işaret işleme tekniklerinin karşılaştırılması", SIU2005, 2005, Kayseri.
- [61] I. Daubechies, "Ten lectures on wavelets," Society for Industrial and Applied Mathematics, Pennsylvania, 1992, doi:10.1137/1.9781611970104.

E. EK MATERYALLER

1. Kullanılacak Ölçekler

Ölçek kullanılmayacak olup, sadece mikrofon ses kaydedici cihaz kullanılacaktır. Bu cihaz kişiye fiziksel bir temasta bulunmadan kapalı bir ortamda 30cm mesafeden ortalama 3sn'lik kayıtlar şeklindedir.

2. Katılımcı Bilgi ve Onam Formu

Katılımcı Bilgi ve Onam Formu (EK. 2)'de ayrıca mevcuttur.

3. Katılım Sonrası Bilgi Formu (gerekirse)

¹Bu "Etik Kurul Onayı Tutararı Formu" TÜBİTAK'ın 21 Ekim 2010'da yayınladığı duyuruya istinaden hazırlanmıştır.

**"MAKİNE ÖĞRENME ALGORİTMALARIYLA KONUŞMACI TANILAMASI" KONULU TEZDE SES KAYDI
ALINIRKEN KİŞİLERE SÖYLETİLECEK KELİMELE:**

- 1- Ansiklopedi
- 2- Belli
- 3- Benzer
- 4- Benzin
- 5- Börek
- 6- Demokrasi
- 7- Dergah
- 8- Elbise
- 9- Eldiven
- 10- Enteresan
- 11- Fevkalade
- 12- Hakem
- 13- Hoşçakal
- 14- Kendi
- 15- Lacivert
- 16- Mendil
- 17- Merhaba
- 18- Muhabbet
- 19- Mühendis
- 20- Portakal
- 21- Rüya
- 22- Şemsiye
- 23- Tolerans
- 24- Yarın
- 25- Zengin

EK-2: Katılımcı Bilgi ve Onam Formu

Evrak Tarih ve Sayısı: 17.07.2024-64718

Ek-2 Katılımcı Bilgi ve Onam Formu

Sayın Katılımcı, Katıldığınız bu çalışma bilimsel nitelikte bir araştırma olup konusu, MAKİNE ÖĞRENME ALGORİTMALARIYLA KONUŞMACI TANILAMASI 'dır. Bu araştırma, Doğu Üniversitesi Lisans Üstü Enstitüsü Elektronik ve Haberleşme Mühendisliği Yüksek Lisansının akademik araştırmasını kapsamaktadır.

Bu çalışmaya katılmak gönüllülük esasına dayanmaktadır. Çalışmaya katılmama ve çalışmanın herhangi bir aşamasında, hiçbir cezaya/yaptırıma maruz kalmaksızın, çalışmadan ayrılma hakkına sahipsiniz. Araştırmada yer almanız nedeniyle size hiçbir ödeme yapılmayacak ve sizden de hiçbir ücret talep edilmeyecektir. Anketi yanıtlamanız, araştırmaya katılım için onam verdiğiniz anlamına gelmektedir. Araştırma hakkında daha fazla bilgi almak için araştırmacı Koray ÖZTÜRK'e başvurabilir, kendilerine 0532 221 66 17 numaralı cep telefonlarından erişebilirsiniz.

Araştırmanın bilgilendirilmiş gönüllü olur formundaki tüm açıklamaları okudum/sözlü olarak dinledim. Bana yukarıda konusu ve amacı belirtilen araştırma ile ilgili yazılı ve sözlü açıklama aşağıda belirtilen araştırmacı tarafından yapıldı. Araştırmaya gönüllü olarak katıldığımı istediğim zaman gerekçeli veya gerekçesiz olarak araştırmadan ayrılabileceğimi biliyorum. Söz konusu araştırmaya, hiçbir baskı ve zorlama olmaksızın kendi rızamla katılmayı kabul ediyorum.

Gönüllünün Adı-Soyadı İmzası Tarih:
Araştırma hakkındaki yukarıdaki ve istenen diğer bilgiler katılımcıya tarafından açıklanmış ve yazılı onamını alınmıştır.

Araştırmacının (Adı-Soyadı) İmzası Tarih: Koray ÖZTÜRK

Bu belge, güvenli elektronik imza ile imzalanmıştır.