

**T.C.
ERCIYES ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
ELEKTRİK ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI**

**MAKİNE ÖĞRENMESİ ALGORİTMALARI
KULLANILARAK DİSFONİK SESLERİN İNCELENMESİ**

**Hazırlayan
Pervin GÜLŞEN**

**Danışman
Prof. Dr. Mustafa ALÇI**

Yüksek Lisans Tezi

**Temmuz 2023
KAYSERİ**

**T.C.
ERCIYES ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
ELEKTRİK ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI**

**MAKİNE ÖĞRENMESİ ALGORİTMALARI
KULLANILARAK DİSFONİK SESLERİN İNCELENMESİ
(Yüksek Lisans Tezi)**

**Hazırlayan
Pervin GÜLŞEN**

**Danışman
Prof. Dr. Mustafa Alçı**

**Temmuz 2023
KAYSERİ**

BİLİMSEL ETİĞE UYGUNLUK

Bu çalışmadaki tüm bilgilerin, akademik ve etik kurallara uygun bir şekilde elde edildiğini beyan ederim. Aynı zamanda bu kural ve davranışların gerektirdiği gibi, bu çalışmanın özünde olmayan tüm materyal ve sonuçları tam olarak aktardığımı ve referans gösterdiğimi belirtirim.

Pervin GÜLŞEN

İmza

“Makine Öğrenmesi Algoritmaları Kullanılarak Disfonik Seslerin İncelenmesi” adlı Yüksek Lisans Tezi, Erciyes Üniversitesi Lisansüstü Tez Önerisi ve Tez Yazma Yönergesi’ ne uygun olarak hazırlanmıştır.

Hazırlayan

Pervin GÜLŞEN

Danışman

Prof. Dr. Mustafa ALÇI

Elektrik – Elektronik Mühendisliği ABD Başkanı

Prof. Dr. Enis GÜNAY

İmza



TEŞEKKÜR

Bu tez çalışmasında ve lisans eğitimim boyunca kıymetli bilgi, birikim ve tecrübeleri ile her türlü destek ve yardımı benden esirgemeyen sayın hocam Prof. Dr. Mustafa ALÇI 'ya çok teşekkür ederim.

Hayatım boyunca her daim arkamda olan maddi ve manevi destekleriyle beni hiçbir zaman yalnız bırakmayan sevgili annem Şeminur Kılınç, babam Cemal Kılınç ve eşim Abdulkadir Gülşen'e sonsuz teşekkür ederim. Ayrıca biricik kızım Elif Sare'ye de sonsuz sevgilerimle.

Pervin GÜLŞEN

Temmuz 2023, Kayseri

MAKİNE ÖĞRENMESİ ALGORİTMALARI KULLANILARAK DİSFONİK SESLERİN İNCELENMESİ

Pervin GÜLŞEN

Erciyes Üniversitesi, Fen Bilimleri Enstitüsü

Yüksek Lisans Tezi, Temmuz 2023

Danışman: Prof. Dr. Mustafa ALÇI

ÖZET

Erken tanı ve yönlendirme, ses bozukluğu tedavisinde hayati öneme sahiptir. Son zamanlarda yapılan çalışmalar, ses patolojisi tespit sistemlerinin ses bozukluklarının değerlendirilmesine etkin bir şekilde katkıda bulunabileceğini ve erken teşhisini sağlayabileceğini göstermektedir. Bu sistemler, birçok alanda kullanılan ve özellikle de ses patoloji tespitinde umut verici olarak kabul edilen makine öğrenimi teknikleridir. Yapılan çalışmalarda, sınıflandırmada kullanılan patolojiler, makine öğrenme modelleri ve performans ölçütleri önemli ölçüde farklılık göstermektedir. Bu farklılık ses patolojisinin tespitinde kullanılabilecek en iyi modelin belirlenmesini zorlaştırmaktadır.

Bu tez çalışmasında makine öğrenme modelleri kullanılarak sağlıklı ve patolojik sesler sınıflandırılmıştır. Ses örnekleri "Saarbrücken Voice Database" (SVD)'den alınmıştır. Ses sinyallerinden öznitelik çıkarımı OpenSMILE algoritması ile yapılmıştır. Ayrıca MFCC yöntemi de kullanılmıştır. Modellerin başarımlarını artırmak için Kendall's Tau öznitelik seçim algoritması kullanılmıştır. Bu algoritma ile sınıflandırma başarımına pozitif etkisi olan en iyi öznitelikler seçilerek modeller eğitilip test edilmiştir. Çalışma sonucunda sağlıklı ve patolojik seslerin sınıflandırılmasında makine öğrenme modellerinin performansları sunulmuştur. Aynı zamanda öznitelik seçim algoritmalarının sınıflandırma başarımına olumlu yönde katkı sağladığı ortaya konulmuştur. Sonuçlara göre makine öğrenme modelleri sağlıklı ve patolojik seslerin sınıflandırılmasında iyi bir performans göstermiştir. Ayrıca öznitelik seçimi ile modellerin başarımları önemli ölçüde artırılmıştır.

Anahtar Kelimeler: Patolojik Ses Tanısı, Yapay Sinir Ağları, SVD, Makine Öğrenmesi.

INVESTIGATION OF DYSPHONIC SOUNDS USING MACHINE LEARNING ALGORITHMS

Pervin GÜLŞEN

Erciyes University, Institute of Science

Master's Thesis, July 2023

Supervisor: Prof. Dr. Mustafa ALÇI

ABSTRACT

Early diagnosis and referral are vital in the treatment of voice disorder. Recent studies show that voice pathology detection systems can contribute effectively to the assessment and early diagnosis of voice disorders. These systems are machine learning techniques that are used in many fields and are considered promising, especially in voice pathology detection. Pathologies, machine learning models and performance criteria used in classification differ significantly in studies. This difference makes it difficult to determine the best model that can be used in the detection of voice pathology.

In this thesis study, healthy and pathological voices were classified using machine learning models. Sound samples are taken from "Saar-brucken Voice Database" (SVD). Feature extraction from audio signals is done with OpenSMILE algorithm. In addition, the MFCC method was also used. Kendall's Tau feature selection algorithm was used to increase the performance values of the models. With this algorithm, the best features that have a positive effect on the classification performance are selected and the models are trained and tested. As a result of the study, the performances of machine learning models in the classification of healthy and pathological voices are presented. At the same time, it has been revealed that feature selection algorithms contribute positively to the classification performance. According to the results, machine learning models performed well in classifying healthy and pathological voices. In addition, the performance of the models has been significantly increased with feature selection.

Keywords: Diagnosis of Pathological Voice, Artificial Neural Networks, SVD, Machine Learning

İÇİNDEKİLER

MAKİNE ÖĞRENMESİ ALGORİTMALARI KULLANILARAK DİSFONİK SESLERİN İNCELENMESİ

BİLİMSEL ETİĞE UYGUNLUK.....	ii
YÖNERGEYE UYGUNLUK.....	iii
KABUL VE ONAY	iv
TEŞEKKÜR	v
ÖZET.....	vi
ABSTRACT	vii
İÇİNDEKİLER	viii
KISALTMALAR	xi
TABLolar LİSTESİ.....	xii
ŞEKİLLER LİSTESİ.....	xiii
GİRİŞ	1

1.BÖLÜM

GENEL BİLGİLER VE LİTERATÜR ÇALIŞMASI

1.1. Problem Durumu	3
1.2. Araştırmanın Amacı	3
1.3. Araştırmanın Önemi.....	4
1.4. Literatür Çalışması.....	4
1.5 Ses Nedir, Nasıl Oluşur?	6
1.5.1. Ses Üretim Mekanizmasının Anatomisi.....	6
1.6. Ses Parametreleri	11
1.6.1. Kalite	12
1.6.2. Rezonans	12
1.6.3. Perde	12

1.6.4. Şiddet.....	12
1.7. Ses Yapısındaki Bozulmalar	13
1.7.1. Disfoni	13
1.7.2. Larenjit	13
1.7.3. Spazmodik Disfoni	14
1.7.4. Hiperkinetik Disfoni	14
1.7.5. Hipokinetik Disfoni.....	14
1.8. Disfonide Kullanılan Tanı Yöntemleri.....	15
1.8.1. Subjektif Ses Analizi.....	15
1.8.2. Objektif ses analizi.....	15
1.8.3. Videolaringoskopi	17

2.BÖLÜM

MATERYAL VE YÖNTEM

2.1. Saarbruecken Ses Veritabanı.....	18
2.2. Geliştirme Ortamı	21
2.2.1. Google Colab (Colaboratory)	21
2.2.2. Scikit-Learn Kütüphanesi.....	22
2.2.3. Python	23
2.3. Öznelik Çıkarımı	24
2.3.1. OpenSMILE	25
2.3.2. Mel Frekansı Kepstral Katsayıları (MFCC)	27
2.4. Öznelik Seçimi.....	28
2.4.1. Kendall's Tau	29
2.5. Sınıflandırma Algoritmaları	30
2.5.1. Lojistik Regresyon (LR).....	30
2.5.2. Destek Vektör Makineleri (SVM)	31
2.5.3. Extreme Gradient Boosting (XGB).....	31
2.5.4. Karar Ağaçları (DT).....	32
2.5.5. Rastgele Orman (RF).....	33
2.5.6. Naive Bayes (NB)	33

3. BÖLÜM

BULGULAR

3.1. Araştırma Modeli.....	35
----------------------------	----

3.2. Evren ve Örneklem	38
3.2.1. Erkek Katılımcıların Değerlendirilmesi	38
3.2.1.1. Erkek /a/ sesi sonuçları	39
3.2.1.2. Erkek /i/ sesi sonuçları	42
3.2.1.3. Erkek /u/ sesi sonuçları	45
3.2.1.4. Erkek tüm ses sonuçları	48
3.2.2. Kadın Katılımcıların Değerlendirilmesi	49
3.2.2.1. Kadın /a/ sesi sonuçları	50
3.2.2.2. Kadın /i/ sesi sonuçları	53
3.2.2.3. Kadın /u/ sesi sonuçları	57
3.2.2.4. Kadın tüm ses sonuçları	60
3.2.3. Tüm Katılımcıların Değerlendirilmesi	61

4.BÖLÜM

TARTIŞMA-SONUÇ ve ÖNERİLER

4.1. Tartışma	65
4.2. Sonuç ve Öneriler	69
KAYNAKÇA	70
ÖZGEÇMİŞ	74

KISALTMALAR

YSA: Yapay Sinir Ağları

SVD: Saarbrücken Voice Dataset (Saarbrücken Ses Veri Tabanı)

MFCC: Mel-frequency Cepstral Coefficients (Mel-Frekans Kepstral Katsayıları)

LPC: Linear Predictive Coding (Lineer Öngörü Kodlaması)

LPCC: Linear Prediction Cepstrum Coefficient (Lineer Tahmin Kepstral Katsayıları)

HOS: Higher Order Statistics (Yüksek Mertebeden İstatistik)

FNN: Feed-forward Neural Networks (İleri Beslemeli Sinir Ağı)

CNN: Convolutional Neural Network (Konvolüsyonel Sinir Ağı)

KZFD: Short-Time Fourier Transform (Kısa zamanlı Fourier dönüşümü)

KNN: K-Nearest Neighbors (K-En Yakın Komşu)

RF: Random Forests (Rastgele Ormanlar)

SVM: Support Vector Machines (Destek Vektör Makineleri)

EGG: Elektrolottogram: Vokal foldlar arasındaki temas alanının göreceli ölçüsünü sağlayan aletsel bir ölçüm tekniği.

CSV: Comma Separated Values (Virgülle Ayrılmış Değer), elektronik tablo formatı: Veri girme, derleme, işleme amaçlarıyla tablo içinde kullanıma izin veren dosyalar.

LR: Logistic regression (Lojistik Regresyon)

DT: Decision Tree (Karar Ağaçları)

XGB: Extreme Gradient Boosting

NB: Naive Bayes

HTK: Hidden-Markov Toolkit

FS: Feature Selection (Öznitelik Seçimi)

FE: Feature Extraction (Öznitelik Çıkarımı)

GPU: Graphics Processing Unit (Grafik İşlemci Ünitesi)

TPU: Tensor Processing Unit (Tensör İşlemci Ünitesi)

SD: Spazmodik Disfoni

NNE: (Normalized Noise Energy)

HNR: (Harmonic-to-Noise Ratio)

OpenSMILE: OpenSource Speech and Music Interpretation by Large Space Extraction (Açık Kaynaklı Büyük alan çıkarımı ile Konuşma ve Müzik Yorumlama)

TABLOLAR LİSTESİ

Tablo 1.1. Disfoni yapan nedenler	2
Tablo 2.1. SVD'den alınan veri grupları.	21
Tablo 3.1. Erkekler /a/ sesi OpenSMILE öznitelikleri için sınıflandırma başarımları. ..	39
Tablo 3.2. Erkekler /a/ sesi MFCC öznitelikleri için sınıflandırma başarımları.	41
Tablo 3.3. Erkekler /i/ sesi OpenSMILE öznitelikleri için sınıflandırma başarımları.	42
Tablo 3.4. Erkekler /i/ sesi MFCC öznitelikleri için sınıflandırma başarımları.	44
Tablo 3.5. Erkekler /u/ sesi OpenSMILE öznitelikleri için sınıflandırma başarımları. ..	45
Tablo 3.6. Erkekler /u/ sesi MFCC öznitelikleri için sınıflandırma başarımları.	47
Tablo 3.7. Erkek tüm seslerin MFCC öznitelikleri için sınıflandırma başarımları.	48
Tablo 3.8. Kadınlar /a/ sesi OpenSMILE öznitelikleri için sınıflandırma başarımları. ..	50
Tablo 3.9. Kadınlar /a/ sesi MFCC öznitelikleri için sınıflandırma başarımları.	52
Tablo 3.10. Kadınlar /i/ sesi OpenSMILE öznitelikleri için sınıflandırma başarımları. ..	54
Tablo 3.11. Kadınlar /i/ sesi MFCC öznitelikleri için sınıflandırma başarımları.	56
Tablo 3.12. Kadınlar /u/ sesi OpenSMILE öznitelikleri için sınıflandırma başarımları. ..	57
Tablo 3.13. Kadınlar /u/ sesi MFCC öznitelikleri için sınıflandırma başarımları.	59
Tablo 3.14. Kadın tüm seslerin MFCC öznitelikleri için sınıflandırma başarımları.	60
Tablo 3.15. Tüm seslerin MFCC öznitelikleri için sınıflandırma başarımları.	62

ŞEKİLLER LİSTESİ

Şekil 1.1. Larinksin anatomisi.....	8
Şekil 1.2. Larinksin kıkırdakları.	9
Şekil 1.3. Larinksin kasları.....	10
Şekil 2.1. SVD’de veri seçimi sayfası.....	19
Şekil 2.2. Seçilen verilerin elde edilmesi.	19
Şekil 2.3. Elde edilen verilerin formatının seçilmesi.	20
Şekil 2.4. Öznitelik seçiminde kullanılabilecek algoritmalar.	29
Şekil 3.1. Karmaşıklık matrisi.....	36
Şekil 3.2. Erkek /a/ sesi için öznitelik sayıları doğruluk grafiği.	40
Şekil 3.3. Erkek /a/ sesi MFCC öznitelikleri için modellerin karmaşıklık matrisleri.	41
Şekil 3.4. Erkek /i/ sesi için öznitelik sayıları doğruluk grafiği.....	43
Şekil 3.5. Erkek /i/ sesi MFCC öznitelikleri için modellerin karmaşıklık matrisleri.....	44
Şekil 3.6. Erkek /u/ sesi için öznitelik sayıları doğruluk grafiği.	46
Şekil 3.7. Erkek /u/ sesi MFCC öznitelikleri için modellerin karmaşıklık matrisleri.	47
Şekil 3.8. Erkek tüm sesler MFCC için modellerin karmaşıklık matrisleri.	49
Şekil 3.9. Kadın /a/ sesi için öznitelik sayıları doğruluk grafiği.....	51
Şekil 3.10. Kadın /a/ sesi MFCC öznitelikleri için modellerin karmaşıklık matrisleri..	53
Şekil 3.11. Kadın /i/ sesi için öznitelik sayıları doğruluk grafiği.....	55
Şekil 3.12. Kadın /i/ sesi MFCC öznitelikleri için modellerin karmaşıklık matrisleri. ..	56
Şekil 3.13. Kadın /u/ sesi için öznitelik sayıları doğruluk grafiği.....	58
Şekil 3.14. Kadın /u/ sesi MFCC öznitelikleri için modellerin karmaşıklık matrisleri...	59
Şekil 3.15. Kadın tüm sesler MFCC için modellerin karmaşıklık matrisleri.....	61
Şekil 3.16. Tüm seslerin MFCC öznitelikleri için modellerin karmaşıklık matrisleri....	63
Şekil 3.17. Tüm sınıflandırma gruplarının OpenSMILE için sınıflandırma sonuçları. ..	64
Şekil 3.18. Tüm sınıflandırma gruplarının MFCC için sınıflandırma sonuçları.....	64

GİRİŞ

İnsanoğlunun doğasında var olan konuşma yeteneği, diğer canlılardan ayrıldığı en önemli özelliktir ve iletişimin temel bir parçasıdır. Konuşmanın en belirgin özelliği, sestir. Bu nedenle, herhangi bir ses sorunu iletim kaybına yol açabilir ve insanı sosyal açıdan olumsuz etkileyebilir [1]. Disfoni, larinks (ses kutusu) hastalıklarının en önemli semptomudur. Türkçe 'de ses kısıklığı olarak adlandırılan disfoni, sadece ses kısıklığı değil, sesin genel bozukluklarını da ifade eden daha kapsayıcı bir terimdir. Disfoni, vokal kordların titreşim fonksiyonunun bozulduğu veya vokal kordlar üzerindeki mukozal düzensizliklerin meydana geldiği durumlarda ortaya çıkar. Disfoniye sebep olan birçok faktör bulunduğu için net bir gruplandırma yapmak zordur. Farklı kaynaklarda farklı gruplandırmalar bulunabilir. Tablo 1.1'de disfoniye sebep olan temel faktörlerden bahsedilmektedir [2].

Disfoni, yetişkinlerin %30'unu ve yaşlıların neredeyse %50'sini etkilemektedir. Ses kalitesini değiştirmenin yanı sıra yaşam kalitesi üzerinde de önemli etkileri vardır. Aynı zamanda önemli bir ekonomik yük oluşturmaktadır [3]. Disfonik seslerin tanısında uzmanlar tarafından birçok invaziv (girişimsel olan) ve noninvaziv (girişimsel olmayan) yöntem uygulanmaktadır. Ama henüz kabul görmüş otomatik hastalık tanı sistemi mevcut değildir.

Tüm sinyal işleme uygulamalarında olduğu gibi ses sinyallerinin otomatik analizinde de veri boyutlarının azaltılması, verilerin işlenmesi ve temel özelliklerinin belirlenerek çıkarım yapılması gerekmektedir. Bunun için algoritmik yöntemlere ihtiyaç duyulmaktadır. Bilgisayarlarda bir işi gerçekleştirebilmek ya da bir veriyi işleyebilmek için yapılacak işe uygun bir algoritma seçmek ya da yeni doğru bir algoritma kurmak gerekmektedir [4]. Yapay zekâ bilgisayar sistemlerinin insan benzeri zekaya sahip

olmasını sağlayan bir disiplindir. Bilgisayarlar ve diğer akıllı cihazlar üzerinde insan benzeri düşünme, öğrenme, karar verme, problem çözme, dil anlama, algılama, konuşma tanıma gibi yetenekleri gerçekleştirmeyi hedefler [5]. Bu tez çalışmasında patolojik ve sağlıklı ses işaretleri makine öğrenmesi algoritmaları kullanılarak incelenecek ve bu iki işaret grubu arasındaki farklılıklar sunulacaktır.

Tablo 1.1. Disfoni yapan nedenler.

A. LARİNKS BENİGN KİTLELERİ • Vokal nodül • Vokal polip • Reinke ödemi • Vokal kistler • Granülom B. KONJENİTAL NEDENLER • Larengomalazi • Konjenital larengeal perde • Down sendromu • Kedi miyavlaması sendromu (Cri du Chat) • Subglottik darlık • Larengotrakeal yarık C. ENFEKSİYÖZ ve ENFLAMATUVAR NEDENLER • Akut larenjit • Kronik larenjit • Reflü larenjit • Kronik atrofik larenjit (larinjitis sicca) • Kserolarinks • Perikondrit D. DEJENERATİF BOZUKLUKLAR • Sulkus vokalis • Presbifoni • Skarlaşma	D. LARİNKSİ TUTAN HASTALIKLAR • Nörolojik hastalıklar (Guillain-Barre sendromu, Amyotrofik lateral sklerozis, Parkinson, koreye neden olan hastalıklar, Myopatiler, Miyastenia gravis...) • Larinksi tutan sistemik hastalıklar (Romatoid artrit, pemfigus, sjögren, ankilozan spondilit, amiloidoz...) • Endokrin hastalıklar (Hipotiroidi, akromegali...) • Larinksin granülomatöz hastalıkları ve diğer hastalıklar (Tüberküloz, sarkoidoz, sifiliz....) E. TÜMÖRAL NEDENLER • Papillom • Juvenil laringeal papillom • Benign tümörler • Maligniteler G. TRAVMATİK NEDENLER • Mekanik travmalar • Kimyasal travmalar • Termal travmalar • İyatrojenik travmalar ▪ Tedavinin doğal sonucu (kordotomi, kordektomi....) ▪ Komplikasyon (biopsi, stripping, çatı cerrahileri, vokal kord paralizisi...)
---	---

1. BÖLÜM

GENEL BİLGİLER VE LİTERATÜR ÇALIŞMASI

1.1. Problem Durumu

Disfoni günümüzde özellikle yetişkin ve yaşlılık dönemlerinde sıkça karşılaşılan bir sağlık problemi olarak karşımıza çıkmaktadır. Son yıllarda makine öğrenimi ve derin öğrenme yaklaşımları, mühendislik, bilim ve teknolojinin birçok alanında büyük ilgi görmektedir. Ayrıca biyomedikal ve akıllı sağlık sektöründe de bu yaklaşımlar kullanılarak birçok hastalık erken dönemde teşhis edilip, tedavi edilebilir bir duruma gelmiştir. Bu tıbbi uygulamalardan biri ise ses patolojisinin tespiti. Patolojik ses, normal konuşma akışının bozulması ve ses yolu üzerinde meydana gelen sorunlar sonucunda ortaya çıkar. Ses patolojileri hastanın konuşma performansını etkiler ve diğer durumların yanı sıra tahriş, konuşmada yorgunluk, sözcük telaffuzunda güçlük gibi bozukluklara neden olur. İşitme kaybı, sinir hastalıkları, beyin yaralanmaları, ilaçların yanlış kullanımı, fiziksel yaralanmalar, sesin yanlış kullanımı ve sağlıksız sosyal alışkanlıklar ses patolojilerini önemli ölçüde artırmaktadır.

Bu çalışmanın gerçekleştirilmesinin temel amacı, makine öğrenmesi teknikleri kullanarak patolojik ve sağlıklı ses sinyallerini ayırt etmektir. Bu, tanı ve sağlık hizmetlerinde önemli iyileştirmelere ve insan yaşamında daha ileri gelişmelere yol açabilir.

1.2. Araştırmanın Amacı

Bu çalışmanın amacı disfonik kişilere teşhis amacıyla yapılan muayene prosedürünün girişimsel ve pratisyen deneyimine bağlı olmasıdır. Hasta kişilere genellikle larinks muayenesi yapılmaktadır. Muayene sırasında, indirekt ayna kullanılarak larinks gözlemlenebilir ve daha gelişmiş fiberoptik fleksibl veya rigid endoskoplarla larinks

incelenebilir. Bu tekniklerden bazıları invaziv, hastayı inciten, yöntemler olduğu için uygulama sırasında hasta acı duyabilmektedir. Aslında ses bozuklukları işitsel bir algı analizi ile de teşhis edilebilir. Ancak, bu pratisyen deneyimine bağlı olarak farklı sonuçlara yol açabilmektedir [1]. Bu istenmeyen bir durumdur. Bu tez çalışmasında çeşitli patolojik sebeplerle seslerinde bozulma olan kişilerin (disfonik) kayıtları, sağlıklı kişilerin kayıtlarıyla karşılaştırılacak ve bu işlem makine öğrenmesi algoritmaları kullanarak yapay zekâ tabanlı yöntemlerle gerçekleştirilecektir. Çalışma sonucunda invaziv olmayan bir yöntem ile hasta ve sağlıklı kişilere ait kayıtlar tanımlanacak ve bu işlem normal prosedürden daha hızlı gerçekleşecektir. Doğruluk seviyesi gerçeği en doğru şekilde yansıtan bir algoritma yazılması hedeflenmektedir.

1.3. Araştırmanın Önemi

Patolojik seslerin otomatik tespiti, konuşma bozukluklarının erken tanı ve tedavisi açısından önemli bir araştırma alanı haline gelmiştir. Genellikle, konuşmanın doğru bir şekilde değerlendirilmesi için eğitilmiş personele ihtiyaç duyulmaktadır. Bu kişi ses değerlendirme konusunda tecrübesi doğrultusunda kararlar verecektir. Bu nedenle otomatik patoloji tanı sistemleri hem zaman hem maliyet açısından verimli bir değerlendirme aracı olarak karşımıza çıkmaktadır.

1.4. Literatür Çalışması

Tez konusuyla ilgili yapılan çalışmalar incelendiğinde, Yapay Sinir Ağları (YSA) tekniğinin birçok klinik uygulamada kullanıldığı görülmüştür. Bu çalışmalar, YSA'nın karmaşık hastalıkların teşhisinde hatalı teşhis riskini azaltmada faydalı bir araç olduğunu göstermiştir. Bu nedenle, YSA tekniğinin gereksiz araştırmaları, olumsuz ameliyat oranlarını ve potansiyel maliyeti azaltmaya yardımcı olabileceği söylenebilir [6].

Literatürde, ses analizini algoritmik yöntemler kullanarak gerçekleştiren çalışmalara bakıldığında, Ji-Yeoun Lee (2021) yaptığı bir çalışmada, mel-frekans kepsstral katsayıları (MFCC'ler), lineer tahmin kepsstral katsayıları (LPCC'ler) ve yüksek mertebeden istatistik (HOS) parametrelerini kullanarak, patolojik ses tespiti için beslemeli ileri beslemeli sinir ağı (FNN) ve konvolüsyonel sinir ağı (CNN) kullanmıştır. Çalışmada kadın ve erkekler için sağlıklı ve patolojik ses verileri ayrı ayrı sınıflandırılmıştır. En yüksek doğruluk oranını, erkeklerde /u/ ünlüsünde LPCC parametresiyle CNN sınıflandırıcısıyla %82.69

olarak elde etmiştir. İkinci en iyi performansını ise, kadınlarda /i/ ünlüsü örnekleri için FNN sınıflandırıcısı, MFCC'ler ve HOS'ların kombinasyonu ile %80.77 olarak elde etmiştir. Akustik ölçümlerin HOS parametreleriyle birleştirilmesinin doğruluk açısından daha iyi bir karakterizasyon sağlamak için faydalı olduğunu gözlemlemiştir. Çalışmada çeşitli parametrelerin ve derin öğrenme yöntemlerinin birleştirilmesinin, normal ve patolojik sesleri ayırt etmek için faydalı olduğunu sunmuştur. [7].

Uysal (2014) yüksek lisans tezinde, elektronik stetoskop kullanarak çeşitli akciğer hastalıklarına sahip bireylerden solunum sesleri almış ve daha önce hekimler tarafından teşhis edilen bu hastalıklara ait solunum seslerini sinyal işleme yöntemleri (MFCC, Doğrusal Öngörü Kodlaması (LPC), Kısa Zamanlı Fourier Dönüşümü (KZFD)) kullanarak tespit etmiş ve bilgisayar üzerinde hastalık teşhisi yapmıştır [8].

Shetty ve diğerleri (2018) yapmış oldukları çalışmada MFCC ve YSA yöntemlerini kullanarak ses sinyallerini sağlıklı ve patolojik olarak sınıflandırmışlardır. YSA da kullandıkları katman sayılarını değiştirerek %90 ları aşan doğrulukta sonuç elde etmişlerdir [9].

Dankovicová ve diğerleri (2018), özellik seçimi (FS) ve K-en yakın komşu (KNN), rastgele ormanlar (RF) ve destek vektör makineleri (SVM) gibi makine öğrenimi yöntemlerini kullanarak normal, yüksek, düşük ve düşük-yüksek ses üreten /a/, /i/ ve /u/ sürekli sesleri ile 94 patolojik 100 sağlıklı denek seçtiler. SVM sınıflandırıcısı, özellik seçimi ile 1560'tan 300'e indirilmiş öznelikler ile en yüksek doğruluğu elde etmişlerdir. Özellik seçimine dayalı genel sınıflandırma performansını, karışık örnekler için %80,3, kadın örnekler için %80,6 ve erkek örnekler için %86,2 olarak belirlemişlerdir. [10].

Hedge ve diğerleri (2019) ise, ses bozukluklarının otomatik tespiti üzerinde yapılan araştırma çalışmalarının anketlerini sunmuş ve farklı ses bozukluklarının nasıl tanımlanabileceğini araştırmışlardır. Ayrıca, çeşitli çalışmalarda kullanılan farklı veri tabanlarını, özellik çıkarma tekniklerini ve makine öğrenimi yaklaşımlarını analiz etmişlerdir. Sesler genellikle makalelerin çoğunda normal ve patolojik olarak kategorize edilmiştir; ancak bazı çalışmalar Alzheimer hastalığı ve Parkinson hastalığı gibi hastalıkları da içermektedir. Son olarak, referans makalede, bu alanda yapılan bazı önemli araştırma çalışmalarının performansını gözden geçirmişlerdir.[11].

Huckvale ve Buciuleac tarafından yapılan çalışma ise patolojik ve sağlıklı sesleri ayırt etmek için SVD veri tabanındaki ses kayıtlarından yararlanılmıştır. ResNet-50 mimarisi kullanılarak, ses kayıtlarının spektrogramları giriş olarak verilmiştir. Tüm patolojileri içeren bir durumda, yalnızca bir ses kaydı kullanıldığında %70 doğruluk elde edilmiştir. Ancak, yalnızca organik patolojiler kullanıldığında doğrulukta yaklaşık %2'lik bir artış olduğu gözlemlenmiştir. Ayrıca, yaptıkları deneylerde cinsiyetin performans üzerinde çok az etkisi olduğu tespit edilmiştir [12].

1.5. Ses Nedir, Nasıl Oluşur?

Ses, konuşmanın temel bir ögesidir ve fonasyon adı verilen ses üretme süreci tarafından oluşturulur. Ekspirasyon (soluk verme) sırasında, merkezi sinir sisteminin kontrolüyle vokal kordlar birbirinden ayrılır ve aynı zamanda havayla titreşim oluşturarak primer ses meydana gelir. Bu ses, rezonans boşluklarına yayılarak algıladığımız ve duyabildiğimiz şekilde büyür. Sesle ilgili kaydedilmiş ilk çalışmalar M.Ö. 5. yüzyıla kadar uzanmaktadır. Hipokrat, akciğerin, soluk borusunun, dudakların ve dilin konuşma için önemli olduğunu belirtmiştir. Aristo, ses üzerine bilimsel araştırmalar yapmış ve sesin duyguyla olan ilişkisine odaklanmıştır. Claudius Galen, M.S. 131-201 yılları arasında yaşayan bir hekimdir; fonasyon sürecinde önemli bir rol oynayan larenksi, yutma ve solunum gibi yaşamsal işlevlerle birlikte açıklamıştır. 1854 yılında Manuel Garcia, indirekt laringoskopi tekniğini icat etmiş ve aynı zamanda ses kıvrımlarını görmeye çalışmıştır [13].

1.5.1. Ses Üretim Mekanizmasının Anatomisi

Laringeal mekanizma hakkında ayrıntılı bilgiye sahip olmak, sesin kalitesini anlamak için önemlidir. Fonasyon süreci, larenksi oluşturan kas ve kıkırdak yapılarını içerir ve vokal kıvrımların titreşiminin fizyolojik süreçlerle sonuçlanmasını sağlar. Yetişkinlerde larinks, boyunun ön kısmında servikal 4, 5 ve 6. omurgaların seviyesinde yer alırken, üst sınırı dil köküne ve alt sınırı ise trakeanın birinci halkasına kadar uzanır. Larinksin tüm bileşenleri sağlıklı ve uyumlu bir şekilde çalıştığında normal ses üretimi gerçekleşir [14].

Larenksin içindeki kaslar, laringeal yapılarla ilişkilendirilmiş durumdadır. Bu kaslar, ses üretiminin başlatılması ve sonlandırılması sırasında vokal kordların pozisyonunu ve gerginliğini sağlar. Bu kompleks eylemler, antagonist ve agonist kas kasılmaları

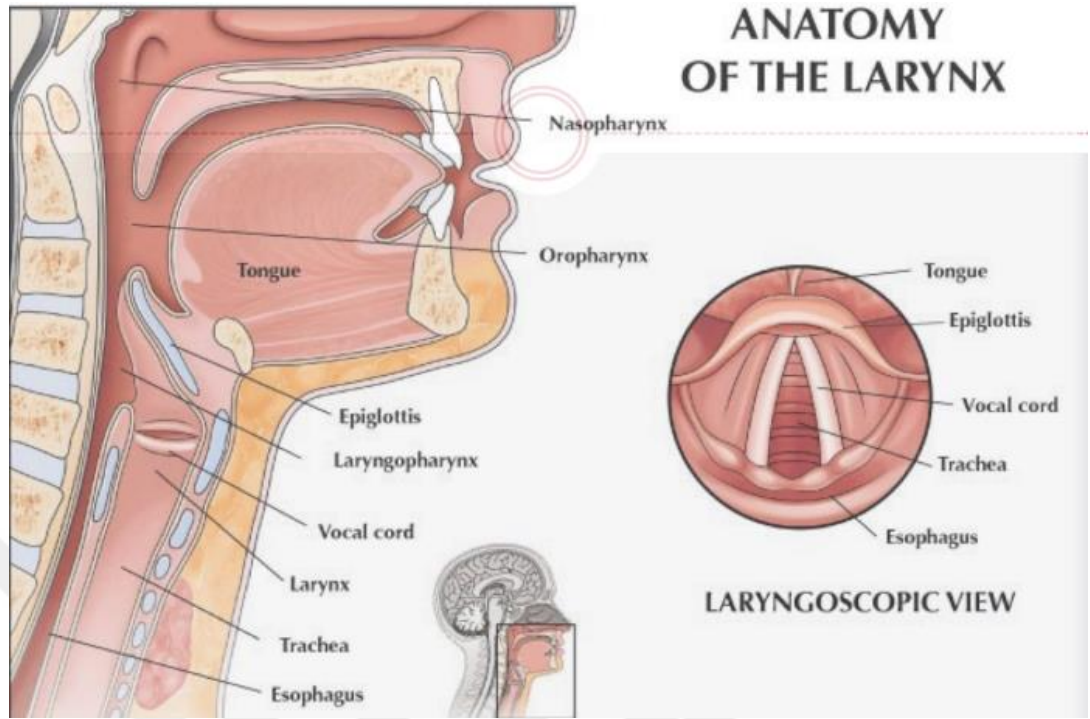
mekanizmasıyla düzenlenir. Krikotiroid kas hariç (bu kas, vagus sinirinin süperior laringeal dalı tarafından innerve edilir), diğer tüm içsel kaslar, vagus sinirinin rekkürrent laringeal dalı tarafından innerve edilir [15].

Larenks, karmaşık bir organdır ve aspirasyondan korunma, solunum ve ses üretimi gibi aktif roller üstlenir. Aynı zamanda öksürme ve yutma süreçlerinde yardımcı bir rol oynar. Ses üretimi sırasında, akciğerler sürekli olarak glottise (ses tellerine) hava akımı sağlar. Bu hava akımı, ses tellerinin titreşimine neden olur [16].

Larenks, konuşmada dört temel fonksiyona sahiptir:

- Fonatuar kontrolü sağlamak için hava akımını açıp kapatabilme yeteneği.
- Sesin temel frekansını kontrol etme.
- Ses şiddetini kontrol etme.
- Ses kalitesini düzenleme.

Larenks, solunum ve koruma işlevlerinin yanı sıra, sözlü iletişimi sağlayan fonasyon işlevine de sahiptir. Larenksin yapısı, kıkırdaklar, bağlar ve membranlar aracılığıyla oluşan çatıdan oluşur. Ayrıca, larenksin iç yüzünü kaplayan mukoza, iç ve dış kaslarla örtülüdür ve ventriküler bant ile ses kıvrımlarını oluşturan yapılar dahil olmak üzere larinksin işlevleri açısından önemlidir [17]. Larinks anatomisi Şekil 1.1'de gösterilmektedir.



Şekil 1.1. Larinksin anatomisi.

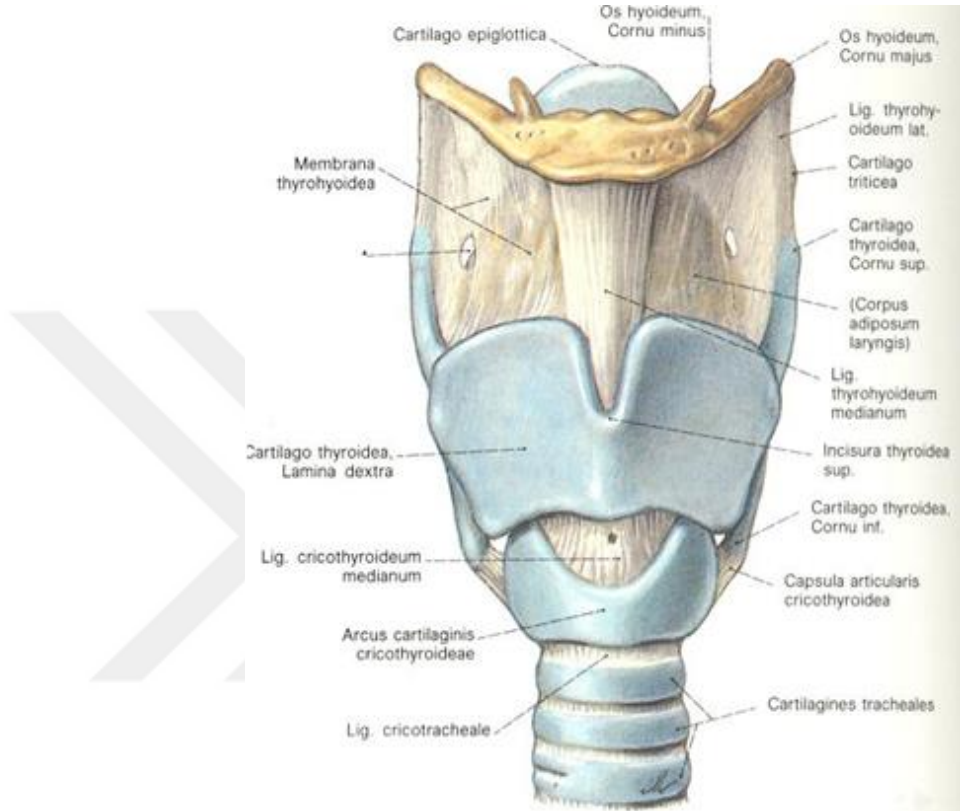
Larenksin çatısını oluşturan yapılar;

Larenks dokuz kıkırdaktan oluşmaktadır. Bu kıkırdaklardan tiroid, krikoid ve epiglot kıkırdaklar tek; ariteooit, komikülat ve kuneiform kıkırdaklar çifttir. Larinks anatomisinde kıkırdakların gösterimi Şekil 1.2’te verilmiştir.

Epiglot, tek bir kıkırdak olan laringeal yapıdır ve yukarı doğru genişleyen bir yaprak şeklindedir. Yutma sırasında gırtlığı kapatarak yiyeceklerin soluk borusuna kaçmasını engeller. Tiroid kıkırdak (Adem elması veya Cartilago tiroidea) kitap yaprağı şeklinde geniş bir kıkırdaktır ve açıklığı arkaya bakar. Ses oluşumu için önemli olan ses telleri ve yanılıcı ses telleri bu kıkırdakta yer alan kaslarla birlikte ön kısımda tutunurlar. Krikoit kıkırdak (Cartilago cricoidea), yüksük şeklinde ve nefes borusunun daha geniş olduğu tiroid kıkırdak ile arasında yer alır.

Aritenoid kıkırdaklar, çift olarak bulunan ve krikoit kıkırdakların arka üst kısmıyla eklem yaparak ses tellerinin arka kısmına tutunan yerdir. Sesin çıkarılması ve ses tellerinin açılıp kapanması bu hareketli kıkırdaklar aracılığıyla gerçekleşir. Aritenoid kıkırdakların üst ve

yan kısımlarına tutunan ve çift olan iki destek kıkırdığı bulunur. Bunlar kornikülat ve kuneiform kıkırdaklardır.



Şekil 1.2. Larinksin kıkırdakları [16].

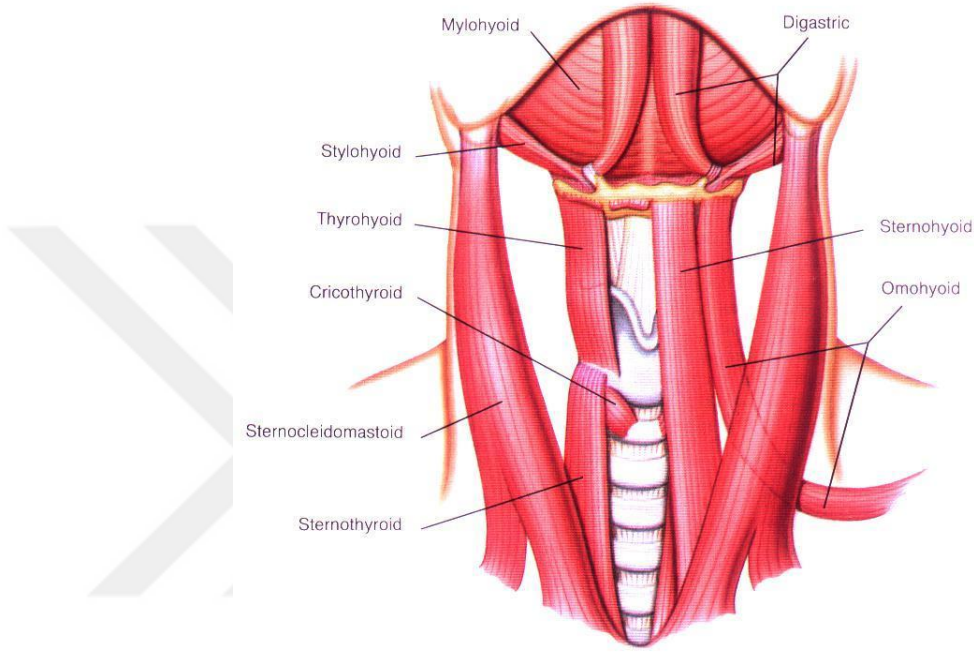
Larenk bağları (ligament), gırtlığın iskelet yapısını elastik yapılarla boyundan bağlar. Ana üst taşıyıcı bağ, temporal kemiğin (kulak kemiği) ile hiyoid kemiği arasındadır. Hiyoid kemiği, geniş bir kemik olan ve çeşitli bağlar ve kaslar aracılığıyla gırtlığın yukarı doğru hareketini sağlayan bir yapıdır [18].

Larenksin eklemleri;

Larenksin içinde, krikotiroid ve krikoaritenoit olmak üzere iki çift sinovyal eklem bulunur. Krikotiroid eklem, krikotiroid kıkırdak ile tiroid kıkırdığın alt komşusu arasındaki eklemdir. Bu eklemden, tiroid kıkırdığın rotasyon ve kayma hareketi gerçekleşir. Krikoaritenoit eklem, aritenoid kıkırdakların tabanı ile krikoid kıkırdak arasında bulunur. Aritenoid kıkırdak, eklem yüzeyi olarak konkav bir yapıya sahipken,

krikoid kıkırdak ise eklem yüzeyi olarak konvektir. Bu eklem, kayma, beşik (eğilme) ve rotasyon olmak üzere üç farklı harekete sahiptir [17].

Larenksin kasları ise ekstrensek ve intrensek olmak üzere iki gruba ayrılmaktadır:



Şekil 1.3. Larinksin kasları [16].

Ekstrensek larenks kasları: Bu kaslar, larenksi çevreleyen dokulara bağlanır ve larenksin boyun bölgesindeki konumunu belirler. Larenksin yukarı-aşağı hareketini, kıkırdaklar arasındaki açılı ve vokal kordların gerginlik durumunu değiştirirler. Ayrıca, intrensek kasların düzgün bir şekilde çalışabilmesi için larenksin boyunda sabitlenmeleri gerekmektedir. Ekstrensek larenks kasları, suprahiyoit (üst hiyoit) ve infrahiyoit (alt hiyoit) olmak üzere iki grupta toplanır. Bu iki grup kasın uyumlu bir şekilde kasılması, larenksin boyun bölgesinde sabitlenmesini sağlar.

İntrensek larenks kasları: Bu kaslar, larenksin temel fonksiyonlarından birincil olarak sorumludur ve vokal kordların hareketlerini etkilerler. Vokal kordların gerilimini ve pozisyonunu düzenleyerek ses üretimini kontrol ederler. İntrensek larenks kasları, vokal

kordların hareketlerini sağlamak için karmaşık bir şekilde birbirleriyle etkileşim halindedir [18].

Larenksin bir diğer önemli bileşeni olan vokal kordlar (titreşen elemanlar) ise insan sesi oluşumunda oldukça önemlidir. Sağ ve sol olmak üzere iki adet vokal kord bulunur. Ön tarafta, vokal kordlar tiroid kıkırdığın iç yüzüne yapışarak ön komissürü oluştururken, arka tarafta aritenoid kıkırdaklara yapışır ve dinlenme durumunda ayrı dururken, fonasyon sırasında birbirine yaklaşır ve arka komissürü oluşturur. Vokal kordlar arasındaki boşluk rima glottis olarak adlandırılır. Bu bölge derin bir nefes alındığında maksimum açıklığa sahiptir, normal solunum ve fisiltı sırasında kısmen açıktır, fonasyon sırasında ise neredeyse tamamen kapalı durumdadır.

Alt solunum yollarından gelen basınçlı hava akımı, glottis seviyesinde kesilerek vokal kordların açılmasını sağlar. Bu sırada vokal kordların titreşimiyle "Primer Glottik Ses" oluşur. Bu noktada, subglottik basıncın, vokal kordların kütlelerinin, uzunluğunun, gerginliğinin ve mukoza hareket kabiliyetinin önemi büyüktür. Primer Glottik Ses, üst rezonatör ve artikülatör yapılar tarafından modüle edilerek ağızdan harf, hece ve kelime olarak çevreye yayılan konuşma şeklinde ifade edilir.

Vokal kordlar erkeklerde yaklaşık 17-20 mm, kadınlarda ise yaklaşık 12-17 mm uzunluğa sahiptir [19].

Mukoza ve kıvrımlar ise, larenksin iç yüzünü kaplayan önemli yapıları oluşturur. Mukoza, larenksin iç yüzünü örter ve ses fonksiyonu açısından önemlidir. Mukoza üzerinde ventriküler bant ve vokal kord olarak adlandırılan iki çift kıvrım bulunur. Bu kıvrımlar, ses fizyolojisi açısından kritik bir rol oynar.

İnsan sesinin oluşumu için tıpkı bir müzik enstrümanında olduğu gibi güç kaynağı, vibratör ve rezonatör gereklidir. Vücudumuzda güç kaynağı görevini ekspiratuvar hava akımını sağlayan diyafram, abdominal kaslar, torakal kaslar ve alt solunum yolları oluştururken vibratör görevini ses kıvrımları; rezonatör görevini ise supraglottik larenks, farinks, ağız ve burun boşluğu görür [14].

1.6. Ses Parametreleri

Sesin; kalite, rezonans, perde ve şiddet olmak üzere 4 temel özelliği dikkat çekmektedir. Bunlar şöyle açıklanabilir:

1.6.1. Kalite

Solunum organları ve vokal kordların birlikte bir uyum içinde çalışması sonuç olarak da supraglottik bölgede hava türbülansı oluşmayacak biçimde eşit aralıklarla ve düzgün bir şekilde vokal kordların titreşmesidir. Anormal ses kalitesi, ses kısıklığı, ses düzensizliği ve solukluluk gibi algısal özelliklerle ifade edilebilir. Fiziksel olarak ise frekans pertürbasyonu (jitter), amplitüt pertürbasyonu (shimmer), NNE (Normalized Noise Energy), HNR (Harmonic-to-Noise Ratio) gibi akustik parametrelerin normalden yüksek olması şeklinde kendini gösterebilir.

1.6.2. Rezonans

Glottis seviyesinde oluşan ses, geçtiği boşluklarda (farinks, ağız gibi) bu boşlukların hacmine ve duvarlarının gerginliğine bağlı olarak değişiklik gösterir. Sesin bazı frekans bölgelerinde, rezonatör organların şekline bağlı olarak ses şiddeti artar ve formantlar oluşur.

1.6.3. Perde

Perde, sesin ince veya kalın oluşunu ifade eden algısal bir terimdir. İki perde arasındaki aralığa semiton denir. Perde, frekansın fiziksel karşılığıdır. Temel frekans, genellikle insan sesinin frekansını ifade eden bir terimdir. Temel frekans, vokal kordların saniyedeki titreşim sayısını temsil eder. Şarkı söylerken bu titreşim sayısı daha yüksek iken, konuşurken daha düşüktür. Konuşma sesinde ortalama temel frekans, erkeklerde yaklaşık 125 Hz iken kadınlarda yaklaşık 215 Hz'dir.

1.6.4. Şiddet

Şiddet, algısal olarak gürültüyü ifade eden bir terim olup, birimi desibeldir (dB). İnsan sesi geniş bir aralığa sahiptir ve 130 dB'ye kadar yükseltilebilir. Sesin şiddeti, glottik direncin ve aynı zamanda subglottik basıncın artmasıyla birlikte artar. Konuşma veya şarkı söyleme sırasında sesin şiddetinin sabit kalabilmesi için subglottik basıncın sabit olması gerekmektedir [17,19].

1.7. Ses Yapısındaki Bozulmalar

Normal bir ses oluşumu için fonasyon organlarının hem anatomik hem de fonksiyonel açıdan kusursuz olması gerekmektedir. Larenksi ilgilendiren ses bozukluklarına disfoni adı verilmektedir. Disfoni, ses kalitesi bozukluğudur.

1.7.1. Disfoni

Disfoni terimi, sesin şiddeti, tınısı, temel frekansı ve maksimum fonasyon zamanı gibi tüm bozuklukları tanımlar. Sesi oluşturan sistemler, solunum sistemi tarafından sağlanan aerodinamik enerji, vokal kordlar (ses telleri) tarafından aerodinamik enerjiyi akustik enerjiye dönüştüren ve larenksten dudaklara kadar olan rezonans motor sistemi tarafından fonotuar karakterleri sağlar. Bu sistemlerle ilgili patolojiler veya fizyolojik olmayan durumlar disfoni (ses bozukluğu) meydana getirebilir [20].

Ses hastalıkları, organik ve fonksiyonel patolojiler olmak üzere iki grupta sınıflandırılır. Fonksiyonel ses hastalıklarında, vokal kordların normal morfolojisi ve hareketiyle ilişkili olan disfoni sıkça görülür. Organik ses hastalıkları ise genellikle larenkteki temel patolojik değişiklikleri içerir. Hem fonksiyonel hem de organik ses hastalıkları, tıbbi, cerrahi veya ses terapisi yöntemleriyle tedavi edilebilir [21].

1.7.2. Larenjit

Larenjit, larinks veya gırtlak adı verilen yapıda meydana gelen iltihaplanmadır. Larinks, solunum yolunun üst kısmında yer alan ve ses üretimine katkıda bulunan bir organdır. Larenjit genellikle virüsler veya bazen bakteriler tarafından neden olunan enfeksiyonlar sonucunda ortaya çıkar. Bunun yanı sıra, sesin aşırı zorlanması, sigara içme, alerjik reaksiyonlar, asit reflüsü gibi faktörler de larenjitin nedenleri arasında yer alabilir.

Larenjit belirtileri arasında ses kısıklığı, boğaz ağrısı, öksürük, boğazda kuruluk ve tahriş hissi bulunabilir. Ayrıca, nefes almada zorluk, yutma güçlüğü veya hırıltılı solunum gibi durumlar da görülebilir.

Larenjit tedavisi genellikle semptomların hafifletilmesine yöneliktir. Ses dinlenmesi, bol sıvı tüketimi, nemli havanın solunması, boğaz pastilleri veya spreylerin kullanımı gibi yöntemler önerilebilir. Bazı durumlarda, altta yatan enfeksiyonun tedavi edilmesi veya ilaçlarla semptomların kontrol altına alınması gerekebilir [22].

1.7.3. Spazmodik Disfoni

Spazmodik disfoni (SD), larengeal kasların etkilendiği fokal bir disfonidir. SD'nin başlangıcı ve oluşumu hakkında çok fazla bilgi olmamakla birlikte, SD, konuşma sırasında larengeal motor nöron kontrolünde düzensizliklere neden olan bir nörobiyolojik mekanizmaya ikincil olarak ortaya çıkar. İstemsiz hareketlerin aksine, istemli konuşma etkilendirken ağlama, kahkaha, öksürme ve esneme gibi hareketlerden etkilenmez. Tipik olarak, hastalar konuşmaya başladıklarında semptomları ortaya çıkar. Konuşma sırasında ses kıvrımları birbirine çok yaklaşır ve ses kesilmeleri veya duraklamaları oluşur. Ses boğuk, kısık ve zorlu olabilir; genellikle sürekli veya sönümsüz, aralıklı veya değişken bir şekilde kendini gösterir [23].

SD'nin doğal seyri ve epidemiyolojisi hakkında çok fazla bilgi bulunmamaktadır. Ancak yapılan çalışmalar bu hastaların çoğunda vücudun başka bir bölgesinin etkilenmediği göstermiştir [24].

1.7.4. Hiperkinetik Disfoni

Hiperkinetik disfoni, sesin yanlış kullanımı sonucunda fonatuar kasların istemsiz olarak aşırı kasılmasıyla karakterizedir. Bu durumda, sesi üreten iç larengeal kaslar yanı sıra ekstra larengeal kaslar ve solunum kasları da aşırı kasılır. Hiperkinetik disfoni, bağırان, hatalı şarkı söyleyen veya agresif ve gergin kişilerde sıkça görülebilir. Bu durum, hiperfonksiyonel (hiperkinetik) disfoni adı verilen bir patolojinin sonucudur. Nodüller, hiperkinetik disfoniye bağlı olarak gelişen bir ses telleri hastalığıdır. Ses terapisi, bu tür durumların tedavisinde yaygın olarak kullanılan bir yöntemdir. Fizyolojik olarak, yalancı ses telleri (ventriküler bandlar) ses üretimine katılmaz. Ancak hiperkinetik disfonide, ventriküler bandların aşırı addüksiyonu ile ventriküler band sesi ortaya çıkabilir. Bu durum, hiperfonksiyonel disfoninin ekstrem bir örneğidir.

1.7.5. Hipokinetik Disfoni

Bu tür disfonide, kas zayıflığı ve buna bağlı olarak vokal kordların iyi birleşmemesi durumu söz konusudur. Larenks kaslarının zayıflığı, genel vücut yapısında zayıflık, yaşlılıkta direnç eksikliği, aşırı kilo kaybı, uzun süreli hiperfonksiyonel disfoni sonucu larenks kaslarının yorgunluğu veya alışkanlıkla yapılan yanlış davranışlar veya korkular gibi faktörlerden kaynaklanabilir. Bu durum hipofonksiyonel disfoni olarak adlandırılır.

Larenks muayenesi sırasında, fonasyon sırasında vokal kordlar orta hatta açık kalır (internus paralizisi) veya poteriorda kapanmaz (posterior üçgen görünümü- transvers yetmezliği) şeklinde gözlemlenir [25].

1.8. Disfonide Kullanılan Tanı Yöntemleri

1.8.1. Subjektif Ses Analizi

Hastanın kendi değerlendirmeleriyle gerçekleştirilen ses analizi yaklaşımdır. Subjektif ses analizinin yapılabilmesi için en yaygın kullanılan yöntemler, Ses Engellilik İndeksi (Voice Handicap Index, VHI) ve Ses Performans Anketi (Vocal Performance Questionnaire, VPQ) olarak bilinir. Bunun yanı sıra, derece, pürüzlülük, nefes alma, asteni, zorlanma ölçeği (Grade, Roughness, Breathiness, Asthenia, Strain, GRBAS Scale) ve Uzlaşmış İşitsel Algısal Ses Değerlendirmesi (Consensus Auditory-Perceptual Evaluation of Voice, CAPE-V) testleri de kullanılmaktadır.

1.8.2. Objektif ses analizi

Objektif ses analizi, hastalardan alınan ses örneğinin çeşitli yazılımlar aracılığıyla incelenmesiyle gerçekleştirilen akustik ses analizini ifade eder.

Akustik ses analizi için aşağıdaki parametreler incelenir:

Temel frekans (Fo): Vokal kordların bir saniyedeki titreşim sayısını temsil eder ve Hertz (Hz) cinsinden ifade edilir. Vokal kordlardaki titreşim hızı değiştiğinde temel frekans da değişir.

Frekans pertürbasyonu parametreleri: Düzgün fonasyonda, ideal olarak frekansta değişiklik olmamalıdır. Ancak fonasyon sırasında istem dışı temel frekans değişiklikleri gözlemlenebilir. Bu düzensizliklere "frekans pertürbasyonu" veya "jitter" denir.

- Mutlak jitter: Ardışık periyotlar arasındaki mutlak farkların ortalaması, mutlak jitter değerini verir ve mikro saniyeler cinsinden ölçülür.
- Jitter (%): Ardışık periyotlar arasındaki mutlak farkın ortalama periyot sayısına bölünmesiyle elde edilir ve yüzde olarak ifade edilir. Mutlak jitter, temel frekansa bağlı olarak değişebilir ve bu bağımlılığı ortadan kaldırmak için Jitter (%)

kullanılır. Ses örneğinde perdedeki kısa süreli değişikliklerin göreceli bir değerlendirmesini sağlar.

Amplitüd pertürbasyonları (Shimmer): Shimmer, ardışık periyotlar arasındaki amplitüd farklarını gösterir ve iki şekilde ifade edilir.

- Shimmer (dB): Ardışık periyotlar arasındaki amplitüd farklarının ortalamasının 10. dereceden logaritması alınarak 20 ile çarpılmasıyla elde edilir ve dB (desibel) olarak ifade edilir. Her bir periyodun tepe amplitüdü, takip eden periyodun tepe amplitüdü ile karşılaştırılarak dB cinsinden hesaplanır.
- Shimmer (%): Birbirini takip eden periyotların amplitüd farklarının amplitüdün ortalamasına bölünmesiyle elde edilir. Bu değer, analiz edilen sesin pik değerleri arasındaki kısa süreli değişimlerin göreceli bir değerlendirmesini sağlar.

Amplitüd pertürbasyonları (Shimmer), sesin titreşim amplitüdündeki kısa süreli değişimleri ölçerek ses bozukluklarının objektif bir şekilde değerlendirilmesine yardımcı olur.

Spektral parametreler: Ses spektrumunda bulunan harmonikler ve harmonikler dışında kalan gürültü ile ilgili parametrelerdir. Harmonikler, temel frekansın katlarından oluşurken, gürültü ise harmonik olmayan ses bileşenlerini temsil eder. Kısa süreli periyot ve amplitüd değişikliklerine ek olarak, vokal kordların kapanmaması veya değişken yapısı glottis boyunca hava kaçağına neden olabilir ve bu durumda akustik karakter gürültü olarak adlandırılır.

- Harmonik-gürültü oranı (HNR): Harmonik spektral enerjinin harmonik olmayan enerjiye oranını ifade eder. Bu değer, analiz edilen sinyalin genel bir değerlendirmesini sağlar. HNR, amplitüd ve frekanstaki pertürbasyonlardan etkilenebilir. Bu ölçüm, sinyalin gürültü seviyesi ve düzensizliğinin en iyi korelasyonunu gösteren bir parametredir.
- Normalize gürültü enerjisi: Total vokal enerjiden harmonik enerjinin çıkarılmasıyla elde edilen bir değerdir.

Spektral parametreler, sesin frekans bileşenlerinin analizini yaparak sesin spektral özelliklerini değerlendirmeye yardımcı olur. Bu parametreler, sesin harmonik yapısı, gürültü seviyesi ve diğer spektral özellikleri hakkında bilgi sağlar [26].

2.8.3. Videolaringoskopi

Normal ses oluşumunun temelini, vokal kordların vibratuar kenarındaki kompleks hareket oluşturur. Fonasyon perdesine bağlı olarak vokal kordlar, saniyede 60 ila 1500 kez açılıp kapanır. Stroboskopi, bir nesnenin hızlı periyodik hareketinin hareketsiz veya yavaş hareket eden bir şekilde optik görüntüsünü ifade eder. Retinaya gelen bir görüntü, 0,2 saniye boyunca kalıcı olabilir. Görüntüler 0,2 saniyeden daha sık aralıklarla gösterilirse hareketli gibi algılanır. Stroboskopi, vokal kordun titreşim frekansı ile illüminasyon frekansı arasındaki senkronizasyonu, kişinin boynuna yerleştirilen bir mikrofona sağlar [27].

2. BÖLÜM

MATERYAL VE YÖNTEM

Bu bölümde öncelikli olarak kullanılan veri tabanı anlatılacak daha sonra bu verilerin işlendiği çalışma ortamlarından bahsedilecektir. Sonrasında sırayla veri setinden öznitelik çıkarımında bulunduğumuz yöntem, öznitelik seçimi ve son olarak sınıflandırma algoritmalarından bahsedilerek tamamlanacaktır.

2.1. Saarbruecken Ses Veritabanı

Saarbruecken Ses Veri tabanı (SVD), 71 farklı fonksiyonel ve organik ses patolojisi için ses kayıtları içeren bir veri tabanıdır. SVD, Saarland Üniversitesi Fonetik Enstitüsü tarafından toplanarak ücretsiz olarak yayımlanmıştır. SVD, 2000’den fazla kişinin ses kayıtlarını içermektedir. Her kişiden “a”, “i”, “u” ünlü harflerinin her biri için “normal”, “düşük”, “yüksek”, “düşük-yüksek-düşük” olmak üzere dört farklı tonda ses kayıtları toplanmıştır. Aynı zamanda “Günaydın, Nasılsın” anlamına gelen Almanca “Guten Morgen, wie geht es Ihnen” cümlesinin de ses kaydı sistemde bulunmaktadır. Bununla beraber toplamda her kişi için 13 ses kaydı bulunmaktadır. Ses kayıtları .wav uzantılı olarak sunulmaktadır.

Ayrıca elektrolottogram (EGG) sinyali her ses için ayrı bir dosya olarak paylaşılmaktadır. Ünlü harflerin ses kayıtlarının uzunluğu 1 ila 3 saniye arasındadır. Tüm ses kayıtları 50 kHz’de örneklenmiş ve çözünürlükleri 16 bittir. Bu veri tabanındaki farklı seslerin ve tonların kaydedilmesi ses patolojilerinin birçok farklı noktadan incelemesi ve analiz yapılmasını mümkün kılmaktadır.

SVD’ ye ait internet sayfasının arayüz görüntüleri aşağıdaki resimlerde gösterilmiştir.

← → ↻ stimmdb.coli.uni-saarland.de/index.php4#target

Version 2.0 **Saarbrücker Stimmdatenbank** 

[Hilfe](#) [Abmelden](#) [Übernehmen](#) [Neue Anfrage](#) [Exportieren](#)

Datenbankanfrage

Sprecher: ☐ männlich ☐ weiblich Alter: Nr.:

Aufnahmesitzung: ☐ gesund ☐ pathologisch Nr.:

Pathologien und Diagnose:

Zwingenden Pathologien:

Ausgewählte Pathologien:

Ausgeschlossene Pathologien:

Anmerkungen zur Diagnose:

Auswahl:

- Amyotrophe Lateralsklerose
- Aryluxation
- Balbuties
- Bulbärparalyse
- Carcinoma in situ
- Chondrom
- Chordektomie
- Cyste
- Diplophonie

[Anfrageerweiterung](#)

Sortierung: Sprecher Nummer Datum der Aufnahme Geschlecht des Sprechers

☐ absteigend ☐ absteigend ☐ absteigend

[Hilfe](#) [Abmelden](#) [Übernehmen](#) [Neue Anfrage](#) [Exportieren](#)

Manfred Pützer  William J. Barry 
 Institut für Phonetik
 Universität des Saarlandes
 Impressum

Zuständiger Administrator 
 Gefördert durch die Deutsche Forschungsgemeinschaft, Bonn
 (BA 737/9-6)

Şekil 2.1. SVD' de veri seçimi sayfası.

Anfrageergebnis

Ergebnisse 1 - 10 von 53

[Auswählen](#) / [Auswahl entfernen](#) [Seite: 1](#) [→](#)

E	ID	T	D	S	G	A	Pathologien	Anmerkungen zur Diagnose	B
☒	495	p	01.04.1998	142	w	54	Dysphonie, Reinke Ödem	Reinke-Ödem beids., re > li; beginnendes Krankheitsbild	
☒	1057	p	18.11.1998	732	w	63	Dysphonie	Cervicogen, asymmetrisches Schwingungsbild	
☒	368	p	19.03.1998	1353	w	47	Dysphonie	Funktionelle ponogene Dysphonie	
☒	445	p	25.03.1998	1354	w	27	Dysphonie	Z. B. hormonelle Dysphonie mit hyperfunktioneller Einschwingung	
☒	449	p	25.03.1998	1358	w	58	Dysphonie	Leichte psychogene, phonasthetisch mäßige Dysphonie	
☒	451	p	25.03.1998	1360	w	56	Dysphonie	Funktionelle Dysphonie	
☒	494	p	01.04.1998	1366	w	30	Dysphonie	Leichte funktionelle Dysphonie bei Stimmlaut	
☒	1276	p	17.03.1999	1372	w	29	Dysphonie	Leichte, unter ponogener Belastung, Chorleiterin	
☒	561	p	29.04.1998	1376	w	46	Dysphonie	psychogene kompensierte Dysphonie bei Zustand nach Ödemabtragung	
☒	674	p	13.05.1998	1401	w	41	Dysphonie	Postlaryngitische Dysphonie	

[Auswählen](#) / [Auswahl entfernen](#) [Seite: 1](#) [→](#)

[Hilfe](#) [Abmelden](#) [Übernehmen](#) [Neue Anfrage](#) [Exportieren](#)

Şekil 2.2. Seçilen verilerin elde edilmesi.

Version 2.0 **Saarbrücker Stimmdatenbank** 

[Hilfe](#) [Abmelden](#) [Übernehmen](#) [Zurücksetzen](#) [Zurück](#)

Export

Dateienauswahl:

Vokaldaten: ☐ Alle Vokale

☒ alle	neutral	hoch	tief	t-h-t
i	☐	☐	☐	☐
a	☐	☐	☐	☐
u	☐	☐	☐	☐

☐ Satzdatei: 'Guten Morgen, wie geht es Ihnen?'
☐ Bemerkungen zu Dateien

Dateiformate:

Sprach-Signal: ☐ NSP ☐ WAV
EGG-Signal: ☐ EGG ☐ WAV

[Hilfe](#) [Abmelden](#) [Übernehmen](#) [Zurücksetzen](#) [Zurück](#)

Manfred Pützer ✉, William J. Barry ✉
Institut für Phonetik
Universität des Saarlandes
Impressum

Zuständiger Administrator ✉
Gefördert durch die Deutsche Forschungsgemeinschaft, Bonn
(BA 737/9-6)

Şekil 2.3. Elde edilen verilerin formatının seçilmesi.

Bu tez çalışmasında “Larenjit”, “Disfoni”, “Hipofonksiyonel (Hipokinetik) Disfoni”, “Hiperfonksiyonel (Hiperkinetik) Disfoni”, “Spazmodik Disfoni” olmak üzere beş farklı ses patolojisi ile tüm sağlıklı ses kayıtları sınıflandırma için kullanılmıştır. Seçilen patolojilerin ses örneği sayısı makine öğrenme modellerinin eğitilmesi için yeterli olduğundan patolojiler kullanılmıştır. Sınıflandırma için sistemdeki bütün ünlü harfler ile bütün tonları 18 ile 65 yaş arası kişiler için seçilmiştir. Eksik veya hasarlı olan kayıtlar incelemeden çıkarılmıştır. Sonuç olarak 307 hasta kadın, 203 hasta erkek, 428 sağlıklı kadın ve 259 sağlıklı erkek olmak üzere toplamda 1207 kişinin ses kayıtları kullanılmıştır. Tablo 2.1’de patolojilere ait örnek sayıları verilmiştir. Her bir kişi için üç ünlü harf ve her harf için dört farklı tonda ses kaydı bulunduğundan her kişinin 12 farklı ses kaydı kullanılmıştır. Böylece toplamda 14484 ses kaydı sınıflandırma için kullanılmıştır.

Tablo 2.1. SVD'den alınan veri grupları.

Hastalık	Kadın	Erkek
Larenjit	51	76
Disfoni	53	48
Spazmodik Disfoni	41	22
Hiperkinetik Disfoni	167	46
Hipokinetik Disfoni	5	11
Toplam Hasta	317	203
Sağlıklı	428	259

2.2. Geliştirme Ortamı

2.2.1. Google Colab (Colaboratory)

Google tarafından sunulan bir bulut tabanlı Jupyter notebook ortamıdır. Colab, kullanıcılara Python kodunu çalıştırabilecekleri ve paylaşabilecekleri bir platform sağlar. Platformun yaygın olarak kullanılma sebepleri şöyle sıralanabilir:

- **Ücretsiz ve Bulut Tabanlı Olması:** Google Colab, kullanıcıların ücretsiz olarak erişebildiği bir hizmettir. Bulut tabanlı olması, herhangi bir yerel kurulum yapmadan tarayıcı üzerinden Python kodu yazıp, çalıştırıp, paylaşmaya olanak sağlar. Ayrıca, Colab'ın otomatik kaydetme özelliği sayesinde kodun kaybedilme riski çok düşüktür.
- **Jupyter Notebook Entegrasyonu:** Colab, Jupyter notebooklarının tamamen bütünleştirilmiş bir sürümünü sunar. Jupyter notebooklar, kod, metin açıklamaları ve çıktıları bir arada tutarak interaktif bir kod geliştirme deneyimi sağlar. Colab, bu interaktif ortamı kullanarak kodu adım adım çalıştırma, sonuçları görselleştirme ve notlar eklemeyi mümkün kılar.
- **GPU ve TPU Desteği:** Google Colab, kullanıcılara ücretsiz olarak GPU (Graphics Processing Unit) ve TPU (Tensor Processing Unit) kaynakları sağlar. Bu, yoğun hesaplama gerektiren makine öğrenimi ve derin öğrenme projeleri için son derece

faydalıdır. GPU ve TPU'lar, işlemleri hızlandırarak model eğitimi ve tahmin sürelerini önemli ölçüde azaltabilir.

- **Kullanıcı Dostu Ortam:** Colab, kullanıcıların Python kütüphanelerini ve araçlarını kolayca yüklemelerine ve kullanmalarına imkân tanır. Popüler Python kütüphaneleri çoğunlukla önceden yüklenmiş olarak gelir, bu da kullanıcıların veri analizi, makine öğrenimi ve görselleştirme gibi işlemleri kolayca gerçekleştirebilmelerini sağlar.
- **Paylaşılabilirlik:** Colab'da oluşturulan Jupyter notebooklar, kolayca ulaşılabilir bağlantılar şeklinde paylaşılabilir. Bu, iş birliği yapma veya kodu başkalarıyla paylaşmayı gerektiren durumlarda son derece faydalıdır.

Google Colab, öğrenme, öğretme, prototip geliştirme veya küçük ölçekli projeler için ideal bir ortam sunar. Erişilebilirliği, kaynakları, kullanıcı dostu arayüzü ve entegre Jupyter notebook özellikleri sayesinde, Python kodlarını kolayca çalıştırmaya olanak tanır [28].

2.2.2. Scikit-Learn Kütüphanesi

Scikit-learn, Python programlama dilinde yaygın olarak kullanılan açık kaynaklı bir makine öğrenimi kütüphanesidir. Scikit-learn, çeşitli makine öğrenimi algoritmalarını, veri işleme araçlarını ve model değerlendirme tekniklerini içeren geniş bir fonksiyonel yelpazeye sahiptir.

Scikit-learn kütüphanesinin temel özellikleri şu şekilde sıralanabilir:

- **Makine Öğrenimi Algoritmaları:** Scikit-learn, sınıflandırma, regresyon, kümeleme, boyut indirgeme, model seçimi ve model değerlendirme gibi çeşitli makine öğrenimi algoritmalarını sunar. Bu çalışmada kullanılan KNN, SVM, DT, RF, XGB, NB gibi algoritmalar scikit-learn'de bulunur.
- **Veri Ön İşleme:** Scikit-learn, veri ön işleme için bir dizi araç sağlar. Özellikle, özellik ölçeklendirme, eksik verilerin ele alınması, kategorik verilerin kodlanması gibi yaygın ön işleme işlemlerini kolaylaştırır.

- **Model Seçimi ve Değerlendirme:** Scikit-learn, model seçimi için çeşitli araçlar sunar. GridSearchCV gibi çapraz doğrulama yöntemleri, hiper parametre ayarını otomatikleştirmek için kullanılabilir. Ayrıca, doğruluk, hassasiyet, geri çağırma gibi metrikleri hesaplamak ve modelin performansını değerlendirmek için fonksiyonlar sağlar.
- **Uygunluk Başlangıç:** Scikit-learn, diğer popüler Python kütüphaneleriyle (NumPy, Pandas) uyumludur ve bu kütüphanelerle birlikte kullanılarak güçlü bir veri analitiği ve makine öğrenimi ekosistemi oluşturur.

Scikit-learn, geniş bir kullanıcı topluluğuna sahip olup, kapsamlı bir dokümantasyon ve zengin örnek koleksiyonuna sahiptir. Ayrıca API tasarımı sayesinde kolay bir şekilde kullanılabilir ve öğrenilmesi kolaydır [29].

2.2.3. Python

Python; genel amaçlı, yüksek seviyeli, nesne tabanlı bir programlama dilidir. Guido van Rossum tarafından 1991 yılında geliştirilmeye başlanmıştır ve o zamandan beri hızla popülerlik kazanmıştır. Python, basit ve anlaşılır sözdizimiyle bilinen ve kullanıcı dostu bir dil olarak kabul edilmiştir. Basit, okunabilir ve anlaşılır bir sözdizimine sahiptir. Bu, programların hızlı bir şekilde yazılmasını ve anlaşılmasını sağlamaktadır. Ayrıca Python'un açık ve düzenli sözdizimi, geliştiricilere kodu kolayca paylaşma ve iş birliği yapma imkânı sunmaktadır. Python'ın temel özellikleri şu şekilde sıralanabilir:

- **Yüksek Seviyeli Dil:** Python, yüksek seviyeli bir programlama dilidir. Bu, geliştiricilerin karmaşık detaylara odaklanmadan kod yazmasına olanak tanımaktadır. Otomatik bellek yönetimi gibi detaylarla uğraşmayarak daha hızlı ve verimli bir şekilde kod yazmayı sağlamaktadır.
- **Nesne Tabanlı Programlama (OOP) Desteği:** Python, nesne tabanlı programlamanın avantajlarını sunmaktadır. Nesne tabanlı programlama, kodun daha modüler, yeniden kullanılabilir ve daha kolay anlaşılabilir olmasını sağlamaktadır. Ayrıca Python, sınıflar, nesneler ve kalıtım gibi nesne tabanlı programlama özelliklerini desteklemektedir.

- **Geniş Kütüphane Desteği:** Python, zengin bir standart kütüphane sunmaktadır. Bu kütüphane, dosya işleme, metin işleme, ağ programlama, veri tabanı bağlantısı, veri analizi, makine öğrenimi ve daha fazlası gibi çeşitli alanlarda kullanılabilen bir dizi fonksiyon ve modül içermektedir.
- **Çapraz Platform Desteği:** Python, çeşitli işletim sistemlerinde (Windows, Linux, MacOS) sorunsuz bir şekilde çalışmaktadır. Bu, aynı kodun farklı platformlarda çalıştırılabilmesini ve taşınabilirliği artırmaktadır.

Python, veri analizi, yapay zekâ, web geliştirme, bilimsel hesaplama, otomasyon ve daha birçok alanda yaygın olarak kullanılan bir programlama dilidir. Ayrıca, büyük bir topluluğa sahiptir ve sürekli olarak geliştirilmekte ve güncellenmektedir. Bu nedenle hem başlangıç düzeyindeki programcılar hem de deneyimli geliştiriciler için tercih edilen bir programlama dilidir [30].

2.3. Öznitelik Çıkarımı

Öznitelik çıkarımı, ham verilerin sayısal özelliklere dönüştürülerek orijinal veri setindeki bilgiyi korurken işlenebilecek hale getirilmesi sürecini ifade eder. Bu, ham verilere doğrudan makine öğrenimi uygulamaktan daha iyi sonuçlar verir.

Öznitelik çıkarımı manuel veya otomatik olarak gerçekleştirilebilir:

Manuel öznitelik çıkarımı yapmak için çalışmanın arka plan ve etki alanıyla ilgili bilgi sahibi uzman kişilerin değerlendirmesi gerekmektedir. Onlarca yıllık araştırmalar ve bilgi birikimi sonucunda bilim insanları görüntüler, sinyaller ve metinler için çeşitli öznitelik çıkarma yöntemleri geliştirmişlerdir. Basitçe bir sinyaldeki bir pencerenin ortalaması bir öznitelik olarak kabul edilebilir.

Otomatik özellik çıkarımı ise insan müdahalesine ihtiyaç duymadan öznitelikleri sinyallerden veya görüntülerden otomatik olarak çıkarmak için özel algoritmalar veya derin ağlar kullanmaktadır. Bu teknik, ham verilerden hızla makine öğrenimi algoritmaları geliştirmeye geçmek istediğimizde çok hızlı sonuç vermektedir. Basitçe, dalgacık saçılımı otomatik öznitelik çıkarımına bir örnektir [31].

2.3.1. OpenSMILE

Açık Kaynaklı Büyük alan çıkarımı ile Konuşma ve Müzik Yorumlama aracı (OpenSMILE), sinyal işleme ve makine öğrenimi uygulamaları için modüler ve esnek bir özellik çıkarıcıdır. Temel odak noktası açıkça ses sinyali özellikleridir. Ancak, yüksek soyutlama kabiliyeti nedeniyle uygun giriş bileşenleri sağlandığında fizyolojik sinyaller, görsel sinyaller ve diğer fiziksel sensörler gibi diğer modalitelerden gelen sinyalleri analiz etmek için de kullanılabilir. C++ dilinde yazılmıştır. Hızlı, verimli ve esnek bir mimariye sahiptir. Aynı zamanda Linux, Windows, MacOS, Android ve iOS gibi yaygın platformlarda çalışabilmektedir.

OpenSMILE gerçek zamanlı çevrimiçi işleme için tasarlanmış olup, büyük veri kümelerinin işlenmesi için ayrıca çevrimdışı olarak da kullanılabilir. Bu, özellik çıkarımı için diğer araçlarda nadiren bulunan bir özelliktir. İlgili projelerin çoğu çevrimdışı çıkarım için tasarlanmış olup, tüm girişin mevcut olmasını gerektirir.

Etkileşimde kalmayı kolaylaştırmak için, OpenSMILE, veri madenciliği ve makine öğrenme alanında yaygın olarak kullanılan çeşitli veri formatlarının okunmasını ve yazılmasını desteklemektedir. Bu formatlar arasında ses dosyaları için PCM WAVE, metin tabanlı veri dosyaları için CSV (Virgülle Ayrılmış Değer, elektronik tablo formatı) ve ARFF (Weka Veri Madenciliği), HTK (Hidden-Markov Toolkit) parametre dosyaları ve ikili özellik verileri için basit bir ikili float matris formatı bulunmaktadır.

OpenSMILE, güçlü ancak sade bir C API ile gelmekte ve diğer uygulamalarla sorunsuz bir entegrasyon sağlamaktadır. Python, .NET, Android ve iOS için sarmalayıcılar mevcuttur. API, ses veya özellik verilerinin OpenSMILE ile gerçek zamanlı olarak iletilmesi de dahil olmak üzere komut satırı aracında mevcut olan tüm işlevselliğe programatik erişim sağlamaktadır.

Dosyalara aktarılan çıkarılan özellikler, open-source yazılım gnuplot ile görselleştirilebilir. OpenSMILE'in son derece modüler mimarisinin sağladığı bir avantaj, özellik çıkarım süreci sırasında oluşturulan neredeyse tüm ara verilerin (pencereyle ayrılmış ses verileri, spektrumlar vb.) erişilebilir olması ve görselleştirme veya daha fazla işleme için dosyalara kaydedilebilmesidir.

OpenSMILE ile her bir ses dosyası için aşağıdaki öznitelikler elde edilebilmektedir [32]:

- Çerçeve Enerjisi
- Çerçeve Yoğunluğu / Ses Yüksekliği (yaklaşık)
- Kritik Bant spektrumları (Mel/Bark/Octave, üçgen maskeleme filtreleri)
- MFCC
- İşitsel Spektrum
- İşitsel spektrumlardan yaklaşık olarak tahmin edilen ses yüksekliği
- Algısal Doğrusal Öngörü Katsayıları
- Algısal Doğrusal Öngörücü Kepstral Katsayılar
- Doğrusal Öngörü Katsayıları
- Çizgi Spektral Çiftleri
- Temel Frekans
- Ses Kalitesi: Jitter ve Shimmer
- Formant frekansları ve bant genişlikleri
- Sıfır ve Ortalama Geçiş oranı
- Spektral özellikler (keyfi bant enerjileri, yuvarlanma noktaları, ağırlık merkezi, entropi, maksimum konumlar, minimum konumlar, varyans (= yayılma), çarpıklık, basıklık, eğim)
- Psikoakustik keskinlik, spektral uyum
- CHROMA (oktav çarpık yarı tonlu spektrumlar) ve CENS özellikleri (enerjiye göre normalleştirilmiş ve yumuşatılmış CHROMA)
- Chord ve Key tanıma için CHROMA'dan türetilmiş özellikler
- F0 Harmonik oranları

2.3.2. Mel Frekansı Kepstral Katsayıları (MFCC)

Herhangi bir otomatik konuşma tanıma sistemindeki ilk adım, özelliklerin çıkarılmasıdır, yani dilsel içeriğin tanımlanması için iyi olan ses sinyalinin bileşenlerinin belirlenmesi ve arka plan gürültüsü, duygu vb. gibi bilgileri taşıyan diğer tüm unsurların atılmasıdır.

Konuşma hakkında anlaşılması gereken temel nokta, insan tarafından üretilen seslerin dil, diş vb. ses yolunun şekli tarafından filtrelenmesidir. Bu şekil, hangi sesin çıkacağını belirler. Şeklin doğru bir şekilde belirlenmesi üretilen ses biriminin doğru bir temsili oluşturur ve oldukça önemlidir. Ses yolunun şekli, kısa süreli güç spektrumu zarfında kendini gösterir ve MFCC'lerin işi, bu zarfı doğru bir şekilde temsil etmektir.

MFCC'ler, otomatik konuşma ve konuşmacı tanımda yaygın olarak kullanılan bir özelliktir. 1980'lerde Davis ve Mermelstein tarafından tanıtılmıştır ve o zamandan bu yana birçok çalışmada kullanılmaya devam edilmektedir.

MFCC aşağıdaki aşamalardan oluşur:

1. Ses Sinyali Alınması: İşlem başlangıcında, ses sinyalinin örneklem değerleri elde edilir. Bu ses sinyali genellikle PCM (Pulse Code Modulation) formatında saklanır.
2. Ön vurgulama: İlk adım olarak, pre-emphasis (ön vurgulama) uygulanabilir. Bu, düşük frekanstaki bileşenlerin zayıflatılmasını sağlayarak yüksek frekanstaki bileşenleri öne çıkarır. Pre-emphasis, sinyaldeki yüksek frekanslı bileşenlerin belirginliğini artırabilir.
3. Çerçeve Bloklama (Frame Blocking): Ses sinyallerinin karakteristikleri, araştırmalar tarafından kanıtlanmıştır ki yeteri kadar küçük bir zaman aralığında kararlı kalmaktadır. Bu nedenle, ses sinyalleri kısa zaman aralıklarında işlenir. Sinyaller genellikle 20 ila 100 milisaniye arasında değişen uzunluklarda çerçevelere bölünür. Ancak çoğu durumda, en etkili zaman aralığı 20 ila 30 milisaniye arasındadır. Ses sinyalleri bu kısa zaman aralıklarında çerçevelere bölünerek analiz edilir.

Ses dalgalarından elde edilen spektrum örnekleri, çerçevelere bölünmüş kısa zaman aralıklarında elde edilir. Her çerçeve, bir önceki çerçevenin belirli bir kısmını örter. Bu örtme yönteminin amacı, bir çerçeveden diğerine geçişin yumuşak olmasını sağlamaktır. Bu, sürekli bir işlem yerine çerçeveler arasında keskin geçişlerin olmasını engelleyerek daha doğal ve sürekli bir ses analizi sağlar.

Çerçevelere bölünmüş ses sinyalleri, her çerçevenin özellikleri ve içeriği hakkında bilgi sağlar. Bu, ses işleme ve konuşma tanıma gibi uygulamalarda ses sinyallerinin daha etkili bir şekilde temsil edilmesini ve işlenmesini sağlar. Özellikle MFCC gibi özellik çıkarma yöntemleri, çerçevelere bölünmüş ses sinyallerinin analizinde yaygın olarak kullanılır ve başarılı sonuçlar elde edilir.

4. Pencereleme (Windowing): Ses sinyalini belirli zaman aralıklarına bölmek için bir çerçeveleme işlemi uygulanır. Bu çerçeveler, genellikle 20-25 ms gibi kısa sürelerde oluşturulur ve ardışık çerçeveler birbiriyle örtüşecek şekilde ilerler. Pencereleme, ses sinyalini daha küçük parçalara ayırarak analizi daha kolay ve etkili hale getirir. Her çerçeve, ses sinyalini temsil eden bir pencere fonksiyonu ile çarpılarak belirlenir. Genellikle, Hamming veya Hanning gibi pencereler kullanılır. Bu çalışmada Hamming pencereleme yöntemi kullanılmıştır.

5. FFT (Hızlı Fourier Dönüşümü) Uygulama: Her pencere çerçevesine FFT uygulanır. Bu, frekans bileşenlerini elde etmek için sinyali frekansta ayırır.

6. Mel Filtre Bankası Uygulama: FFT sonuçları, insan işitmesinin frekans duyarlılığını taklit eden bir dizi Mel filtresi ile çarpılır. Bu adım, frekans bölgelerini daha insan merkezli hale getirir. Frekans değerlerinin mel değerine dönüşümü aşağıdaki formül ile gerçekleştirilmektedir.

$$\text{Mel}(f) = 1125 \times \ln(1 + (f/700))$$

7. Logaritma Alınması: Mel filtre çıktıları üzerine doğal logaritma uygulanır. Bu, ses sinyalinin genlik ölçümlerini daha lineer hale getirir.

8. Kepstral Katsayıların Çıkarılması: Son olarak, kepsral katsayıları hesaplanır. Bu, ayrık fourier dönüşümünün (DFT) logaritmasını alarak gerçekleştirilir.

9. Özellik Vektörü Oluşturma: Kepstral katsayıları bir araya getirilerek MFCC özellik vektörü oluşturulur [33].

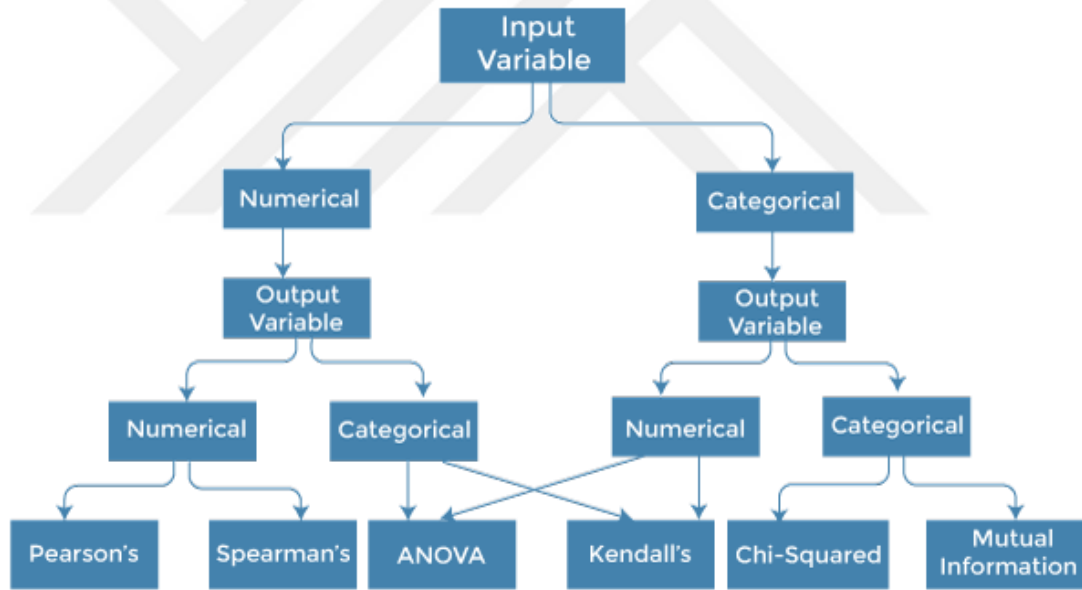
2.4. Öznitelik Seçimi

Öznitelik seçimi; öznitelik çıkarımı ile elde edilen sayısal veri setindeki değişkenlerin önem derecesine göre seçilmesi veya birleştirilmesi olarak ifade edilebilmektedir. Bu

işlem ile veri setini en iyi temsil eden öznitelikler belirlenmekte, gereksiz veya tekrar eden bilgiler ortadan kaldırılmaktadır. Bu sayede işlenecek veri miktarı önemli ölçüde azalmaktadır.

Özellik çıkarımı, veri analizi ve makine öğrenimi alanlarında önemli bir adımdır. İyi seçilmiş ve anlamlı özellikler, daha doğru sonuçlar elde etmek, gereksiz bilgileri ortadan kaldırmak ve veri setinin analizini kolaylaştırmak için büyük bir öneme sahiptir.

Şekil 2.4'te veri setinin yapısına göre seçilmesi uygun olacak öznitelik seçim modelleri gösterilmiştir. Bu tablodan yola çıkarak bu çalışmada OpenSMILE kullanarak elde edilen özniteliklerden seçim yapmak için Kendall's Tau yöntemi kullanılmıştır.



Şekil 2.4. Öznitelik seçiminde kullanılabilecek algoritmalar.

2.4.1. Kendall's Tau

Kendall's Tau, korelasyon derecesinin ölçülmesinde kullanılan parametrik olmayan bir yöntemdir. Maurice Kendall tarafından 1938 yılında tanıtılmıştır. Kendall's Tau özellik seçimi modeli, iki değişken arasındaki ilişkiyi ölçmek için kullanılan bir istatistiksel

yöntemdir. Bu yöntem, iki değişken arasındaki sıralama düzeninin benzerliğini ölçer. Böylece değişkenler arasındaki ilişki gücü ölçülebilir.

Varsayalım ki r_1 ve r_2 sıralamaları karşılaştırılıyor. Veri kümesindeki her bir öznitelik j için, r_1 ve r_2 tarafından üretilen sıralamaları temsil eden $(r_1(j), r_2(j))$ şeklinde düzenlenmiş bir çift vardır. Her bir öznitelik çifti (j_1, j_2) için, $(r_1(j_1), r_2(j_1))$ ve $(r_1(j_2), r_2(j_2))$ sıralamaları karşılaştırılır ve iki sıralamanın uyumlu mu (aynı yönde mi) yoksa uyumsuz mu (farklı yönde mi) olduğuna bağlı olarak +1 veya -1 değeri verilir. r_1 ve r_2 'nin herhangi bir bağlı sıralamalar içermediği varsayılarak, (j_1, j_2) öznitelik çifti, $r_1(j_1) > r_1(j_2)$ ve $r_2(j_1) > r_2(j_2)$ ya da $r_1(j_1) < r_1(j_2)$ ve $r_2(j_1) < r_2(j_2)$ olduğunda uyumlu olarak kabul edilir. Aksi takdirde, (j_1, j_2) uyumsuz olarak kabul edilir. Toplamda $\frac{n(n-1)}{2}$ öznitelik çifti bulunmaktadır. Eğer S , her bir öznitelik çifti için uyumlu ya da uyumsuzluk durumuna göre hesaplanan puanların toplamıysa, Kendall's Tau şu şekilde hesaplanır: $\tau = S / \frac{n(n-1)}{2}$. Eğer tüm çiftler uyumlu ise, iki sıralama tam bir uyum içindedir ve $\tau = 1$ olur. Eğer iki sıralama tamamen zıt ise (yani $\forall j, r_2(j) = n - r_1(j) + 1$), o zaman tüm çiftler uyumsuz olacaktır ve $\tau = -1$ olur. Eğer τ değeri yaklaşık olarak sıfıra yakınsa, iki sıralama arasındaki ilişki çok zayıftır [34].

2.5. Sınıflandırma Algoritmaları

2.5.1. Lojistik Regresyon (LR)

LR, değişkenler arasındaki ilişkileri incelemek için sıkça kullanılan istatistiksel bir yöntemdir. Regresyon analizi, yordanan değişkenle yordayıcı değişkenler arasındaki ilişkiyi en uygun şekilde tanımlayabilen ve en az hata ile uyum sağlayan bir model oluşturmayı amaçlar. Doğrusal regresyon analizi, yordanan değişkenin normal dağılımı, doğrusallık ve eş varyanslılık gibi varsayımlar altında geliştirilmiştir. Ancak LR normallik, süreklilik, eş varyanslılık ve çok değişkenli normallik gibi varsayımları gerektirmeden bir regresyon modeli oluşturmamızı sağlar. LR, yordayıcı değişkenlerin sürekli veya kategorik olabileceği, sürekli ve kategorik değişkenlerin bir arada kullanılabildiği, yordanan değişkenin ise kategorik olduğu bir analiz yöntemidir. İhtiyaç duyulduğunda, LR modelini oluşturmak için sürekli olan bir sonuç değişkeni, yani yordanan değişken, sürekli olmayan bir değişkene dönüştürülebilir [35].

2.5.2. Destek Vektör Makineleri (SVM)

SVM, bir sınıflandırma algoritmasıdır ve istatistiksel öğrenme yöntemlerine dayanır. SVM kullanılarak gerçekleştirilen sınıflandırmalarda genellikle iki sınıfa ait örneklerin sınıf etiketleri $\{-1, +1\}$ şeklinde gösterilir. Amacı, eğitim verisi kullanılarak elde edilen bir karar fonksiyonu yardımıyla bu iki sınıfın birbirinden en iyi şekilde ayrılmasını sağlamaktır.

SVM, bu karar fonksiyonunu kullanarak eğitim verisini en uygun şekilde ayıracak bir hiper düzlem bulmaya çalışır. Bu hiper düzlem, sınıfları birbirinden en iyi şekilde ayırmak için kullanılır. İdeali, sınıfların birbirinden en uzak noktalardan geçen bir hiper düzlemle ayrılmasıdır. Bu hiper düzlem, destek vektörler adı verilen eğitim verisindeki önemli noktalarla belirlenir. Destek vektörler, sınıflar arasındaki ayrımı en iyi şekilde temsil eden ve hiper düzleme en yakın olan noktalardır.

SVM, doğrusal ve doğrusal olmayan (kernel kullanılarak) problemlerde etkili bir şekilde kullanılabilir. Doğrusal olmayan problemlerde, verilerin yüksek boyutlu özellik uzayına taşınması ve bu uzayda bir doğrusal hiper düzlem bulunması sağlanır. Bu, verilerin daha karmaşık bir yapıya sahip olabileceği durumlarda da etkili sınıflandırmalar yapılabilmesini sağlar.

SVM, geniş bir uygulama alanına sahiptir ve sınıflandırma problemlerinin yanı sıra regresyon, anormal veri tespiti ve boyut indirgeme gibi farklı görevlerde de kullanılabilir [36].

2.5.3. Extreme Gradient Boosting (XGB)

XGB, özellikle yapılandırılmış verilerde kullanılan bir makine öğrenimi algoritmasıdır. Gradient Boosting yöntemlerine dayanarak, performansı artırmak ve tahmin yeteneklerini iyileştirmek için özel bir optimize edilmiş hedef fonksiyonu kullanır. Ağaç tabanlı bir öğrenme yöntemidir ve aynı zamanda bir topluluk yöntemi olarak da sınıflandırılır. Bir topluluk yöntemi olarak, birden çok zayıf öğreniciyi (genellikle karar ağaçları) birleştirerek daha güçlü bir öğrenici oluşturur.

XGB'nin temel özellikleri şunlardır:

- Gradient Boosting: XGB, önceki ağaçlardan gelen hataları hedefe yaklaştırmak için gradient boosting tekniğini kullanır. Bu, her bir ağacın hataları düzeltmek için önceki ağaçlardan gelen gradyanları kullanmasını sağlar.
- Regularizasyon: XGB, aşırı öğrenmeyi engellemek için birden fazla regularizasyon yöntemi sağlar. Bunlar arasında ağaç büyüklüğünü sınırlama, düğüm bölünmelerinde minimum kaybı sağlama ve ağaçları düzenleme yer alır.
- Hız ve Verimlilik: XGB, hesaplama hızını artırmak için çeşitli optimizasyon teknikleri kullanır. Ağaç büyüme sürecinde paralel hesaplama, hesaplama sırasında önbellek kullanımı ve çoklu işlemci desteği gibi yöntemler, XGB'nin hızlı ve verimli çalışmasını sağlar.
- Esneklik: XGB, çeşitli veri tiplerini ve hedef fonksiyonlarını destekler. Sınıflandırma, regresyon ve sıralama gibi çeşitli problemler için kullanılabilir [37].

2.5.4. Karar Ağaçları (DT)

DT, bir karar verme yöntemi olup ağaç yapısına sahiptir ve sınıflandırma için farklı bir yaklaşım sunar. Bir karar ağacı, bir dizi "eğer-o zaman" kuralı oluşturur. Her kural, ağacın bir bölümüne yönlendirilir ve bu noktada başka bir "eğer-öyleyse" kuralı uygulanır. İlk bölüm, tüm veri setini içerir ve karar ağacının kök düğümüdür.

DT, nominal sonuçları olan ve birden fazla olası sonuca sahip olan sınıflandırmalarda ve sıralı sonuç değişkenlerinde kullanılır [38].

Sınıflandırma ve regresyon süreçlerinde, karar ağaçlarının ayrıştırma kriterleri genellikle değişkenlere dayanır ve algoritmanın performansını önemli ölçüde etkiler. Ayrıştırmanın nasıl uygulanacağı, karar ağaçlarındaki doğruluğu etkileyen en önemli faktörlerden biridir. Bir düğümü birden çok alt düğüme ayırmak için farklı yöntemler kullanılır. Algoritmanın seçimi, hedef değişkenin türüyle ilgilidir.

Kategorik tabanlı değişkenler için, Gini indeksi, Entropi ve Sınıflandırma Hatası gibi yöntemler yüksek oranda tercih edilir. Bu yöntemler, düğümün saflığını ölçerek bölünme kriterini belirler. Saflık, aynı sınıfa ait örneklerin yoğunluğunu ifade eder. Gini indeksi

ve Entropi, düğümün saflığını ölçmek için kullanılan istatistiksel yöntemlerdir. Sınıflandırma Hatası ise, bir düğümdeki en yaygın sınıfın oranını temel alır.

Sürekli değişkenler için, En Küçük Kareler algoritması sıkça kullanılır. Bu yöntem, düğümdeki sürekli değişkenin değerlerine göre bir eşik değeri belirler ve bu değere göre verileri iki alt düğüme ayırır. En Küçük Kareler algoritması, düğümdeki sürekli değişkenin ortalamasına göre en küçük kare hatasını minimize etmeyi hedefler.

Bu yöntemler, karar ağaçlarının bölünme stratejilerini belirlemek için kullanılan temel kriterlerdir ve seçilen yöntem, veri setinin özelliklerine ve hedefe bağlı olarak değişebilir [39].

2.5.5. Rasgele Orman (RF)

2001 yılında L. Breiman tarafından önerilen RF algoritması, genel amaçlı bir sınıflandırma ve regresyon yöntemi olarak son derece başarılı olmuştur. Bu yaklaşım, birçok rastgele karar ağacını birleştirir ve tahminlerini ortalamayla birleştirerek mükemmel performans gösterir. Özellikle de değişken sayısının gözlem sayısından çok daha büyük olduğu durumlarda etkili bir şekilde çalışır. Ayrıca, büyük ölçekli problemlere uygulanabilir, çeşitli özel öğrenme görevlerine kolayca uyarlanabilir ve değişken önemini ölçebilir [40].

2.5.6. Naive Bayes (NB)

Bayes teoremine dayanarak, bir veri ögesinin belirli bir sınıfa ait olma olasılığını hesaplamak için kullanılır. "Naive" terimi, yöntemin sınıflandırma yaparken her bir özellik arasında bağımsızlık varsayımında bulunmasından kaynaklanır.

NB yöntemi, özellik vektörleri ve bunlara karşılık gelen sınıf etiketlerini içeren bir eğitim veri seti kullanır. Bu veri seti üzerinde özelliklerin olasılık dağılımları tahmin edilir. Sonrasında, yeni bir veri ögesinin sınıflandırılması gerektiğinde, bu özelliklerin verilen sınıfa ait olma olasılıkları hesaplanır ve en yüksek olasılığa sahip sınıf etiketi verilir. Bu yöntem, veri setindeki özelliklerin bağımsız olduğunu varsayar. Bu, bir özelliğin varlığı veya yokluğunun diğer özellikler üzerindeki etkisini dikkate almadığı anlamına gelir. Bu basit varsayım, yöntemi hızlı ve verimli hale getirir.

NB yöntemi, genellikle metin sınıflandırma, spam filtreleme, duygu analizi gibi doğal dil işleme problemlerinde kullanılır. Ayrıca, veri setindeki özellikler arasında gerçek bağımlılıkların bulunmadığı varsayıldığında, hızlı ve etkili bir sınıflandırma çözümü sunar [41].



3. BÖLÜM

BULGULAR

3.1. Araştırma Modeli

Tez çalışmasında SVD veri tabanında bulunan sağlıklı ve patolojik seslere ait kayıtları makine öğrenme yöntemleri kullanılarak sınıflandırılmıştır. Bu kapsamda ilk olarak sesler cinsiyet, patoloji ve ünlü harf bazında ayrıştırılmıştır. “Erkek, a harfi, sağlıklı ve patolojik”, “Erkek, i harfi, sağlıklı ve patolojik”, “Erkek, u harfi, sağlıklı ve patolojik”, “Kadın, a harfi, sağlıklı ve patolojik”, “Kadın, i harfi, sağlıklı ve patolojik”, “Kadın, u harfi, sağlıklı ve patolojik” olmak üzere 12 farklı dosyada toplanmıştır. MFCC ve OpenSMILE olmak üzere iki öznitelik çıkarım yöntemi kullanılmıştır. OpenSMILE ile her bir ses örneği için 6373 farklı öznitelik elde edilmiştir. Modellerin sınıflandırma başarımını arttırmak için Kendall's Tau öznitelik seçim modeli kullanılarak 10 ila 3000 arasında 10'un katları şeklinde artarak algoritma çalıştırılmış ve en iyi öznitelikler seçilmiştir. Her bir öznitelik seçimi için modeller eğitilip test edilerek modelin en iyi başarımlar gösterdiği öznitelik sayısı belirlenmiştir. Ayrıca öznitelik sayılarının başarımlar değerleri grafik halinde sunulmuştur. MFCC ile de her bir ses örneği için 28 farklı öznitelik elde edilmiştir. Modeller MFCC öznitelikleri ile tüm sınıflar için eğitilip test edilmiştir. Modellerin sınıflandırma başarımları ve karmaşıklık matrisleri tablolar halinde sunulmuştur. Tüm modeller için “Doğruluk”, “Kesinlik”, “Duyarlılık” ve “F1” değerleri hesaplanmış ve tablolar halinde sunulmuştur.

Bu değerler şunları ifade etmektedir:

Karmaşıklık Matrisi (Confusion Matrix): Bir sınıflandırma modelinin performansını değerlendirmek için kullanılan bir tablodur. Karmaşıklık matrisi, sınıflandırma modelinin

gerçek sınıf etiketleriyle tahmin ettiği sınıf etiketlerini karşılaştırarak modelin doğruluğunu, hata tiplerini ve performansını anlamaya yardımcı olmaktadır. Şekil 3.1 de verilen karmaşıklık matrisi üzerindeki ifadeler açıklanacak olursa;

- TP (True Positive): Gerçek pozitif örneklerin doğru şekilde tahmin edilme sayısı.
- FP (False Positive): Gerçek negatif örneklerin yanlış şekilde pozitif olarak tahmin edilme sayısı.
- FN (False Negative): Gerçek pozitif örneklerin yanlış şekilde negatif olarak tahmin edilme sayısı.
- TN (True Negative): Gerçek negatif örneklerin doğru şekilde tahmin edilme sayısı.

Karmaşıklık matrisi, modelin performansını değerlendirmek için farklı metriklerin hesaplanmasına olanak tanımaktadır. Bu metrikler arasında doğruluk (accuracy), hassasiyet (precision), duyarlılık (recall), F1 skoru, özgüllük (specificity) gibi değerler yer almaktadır. Bu metrikler, TP, FP, FN ve TN değerlerinin kombinasyonlarına dayanmaktadır.

		Tahmin edilen durum	
Toplam nüfus = P + N		Pozitif (PP)	Negatif (PN)
Gerçek durum	Pozitif (P)	Gerçek pozitif (TP)	Yanlış negatif (FN)
	Negatif (N)	Yanlış pozitif (FP)	Gerçek negatif (TN)

Şekil 3.1. Karmaşıklık matrisi.

Doğruluk (Accuracy): Doğru sınıflandırılmış örnek sayısının tüm örnek sayısına oranı olarak ifade edilebilir. Veri setinin dengeli dağılmış olması hem sınıflandırma yaparken hem de doğruluk ölçütünü belirlerken önemli bir parametredir.

Doğruluk değerinin hesaplanması aşağıdaki formülde belirtildiği gibidir.

$$\text{Doğruluk} = (TP + TN) / (TP + TN + FP + FN)$$

Kesinlik (Precision): Hassasiyet olarak da ifade edilebilmektedir. Doğru pozitifin tüm pozitiflere oranıyla bulunur.

$$\text{Kesinlik} = TP / (TP + FP)$$

Kesinlik, yanlış pozitiflerin önemli olduğu durumlarda modelin yanlış pozitiflerini kontrol etme yeteneğini ölçmek için kullanılır. Düşük Kesinlik değeri, yanlış pozitiflerin yüksek olduğunu gösterir.

Duyarlılık (Recall): Sınıflandırma modelinin gerçek pozitif örnekleri ne kadarını doğru bir şekilde tespit ettiğini ölçen bir performans metriğidir. Duyarlılık, aynı zamanda "gerçek pozitif oranı" veya "duyarlılık oranı" olarak da adlandırılır.

Aşağıdaki formülle hesaplanır:

$$\text{Duyarlılık} = TP / (TP + FN)$$

Duyarlılık, yanlış negatiflerin (gerçek pozitiflerin yanlışlıkla reddedilmesi) önemli olduğu durumlarda modelin duyarlılığını ölçmek için kullanılmaktadır. Düşük duyarlılık değeri, yanlış negatiflerin yüksek olduğunu gösterir.

F1 Skoru: F1 skoru, kesinlik ve duyarlılık arasındaki dengeyi ölçmek için kullanılan bir harmonik ortalama değeridir. F1 skoru, hassasiyet ve duyarlılık değerlerinin birleşik performansını gösterir. Aşağıdaki formülle hesaplanır:

$$\text{F1 Skoru} = 2 * (\text{Kesinlik} * \text{Duyarlılık}) / (\text{Kesinlik} + \text{Duyarlılık})$$

F1 skoru hem yanlış pozitifleri hem de yanlış negatifleri minimize etmeye çalışan bir modelin performansını değerlendirmek için kullanılır.

Özgüllük (Specificity): Özgüllük, gerçek negatif örneklerin doğru olarak tahmin edilen negatif örneklerin oranını temsil eder. Yani, doğru negatiflerin tüm negatif tahminler içindeki oranını gösterir. Aşağıdaki formülle hesaplanır:

$$\text{Özgüllük} = TN / (TN + FP)$$

Özgüllük, negatif sınıfın doğru bir şekilde tanınması önemli olduğunda kullanılır. Düşük özgüllük değeri, yanlış pozitiflerin yüksek olduğunu gösterir.

Bu metrikler, sınıflandırma modellerinin performansını daha ayrıntılı bir şekilde anlamamızı sağlamaktadır ve belirli bir problem bağlamında modelin ne kadar iyi çalıştığını değerlendirmeye yardımcı olur. Hangi metriğin kullanılması gerektiği, uygulama alanına ve problem gereksinimlerine bağlı olarak değişmektedir.

3.2. Evren ve Örneklem

3.2.1. Erkek Katılımcıların Değerlendirilmesi

SVD veri tabanından 18-65 yaş arası erkek katılımcılardan 259 adet sağlıklı 203 adet ise patolojik ses kaydı elde edilmiştir. Her bir katılımcı “a”, “i”, ve “u” ünlü harflerini dört farklı tonda seslendirmiştir. Bunun sonucunda erkek katılımcılar için 3108 sağlıklı ve 2436 patolojik olmak üzere toplamda 5544 ses kaydı kullanılmıştır. Öncelikle her bir ses kaydı için OpenSMILE ile 6373, MFCC ile 28 öznitelik çıkarılmıştır. Her bir ünlü harf kendi aralarında ve tamamı birleştirilerek sınıflandırma yapılmıştır. Veri seti %80 eğitim ve %20 test olacak şekilde bölünmüştür. Eğitim ve test veri setindeki örnekler rastgele ve patolojik ve sağlıklı oranı iki bölümde de eşit seçilmiştir.

OpenSMILE öznitelikleri ile makine öğrenme modelleri tüm öznitelikler ile eğitim veri setinde eğitilmiş ve test veri setinde test edilmiştir. Tüm özellikler için modellerin performans sonuçları tablolarda gösterilmiştir. OpenSMILE öznitelikleri ile elde edilen başarımların artırarak, eğitim ve test sürelerini düşürmek için öznitelik seçimi uygulanmıştır. Öznitelik seçiminde Kendall’s Tau modeli kullanılmıştır. Seçilecek optimum öznitelik sayısını belirlemek için belirli bir aralıktaki bütün öznitelik sayıları tüm modellerde eğitilip test edilmiştir. Bunun için 10 ila 3000 arasında 10’un katları sayıda öznitelik seçilmiş ve tüm modellerin her bir öznitelik sayısı için doğruluk değeri belirlenmiştir. Aşağıdaki grafiklerde Kendall’s Tau öznitelik sayılarına göre modellerin doğruluk değeri gösterilmiştir.

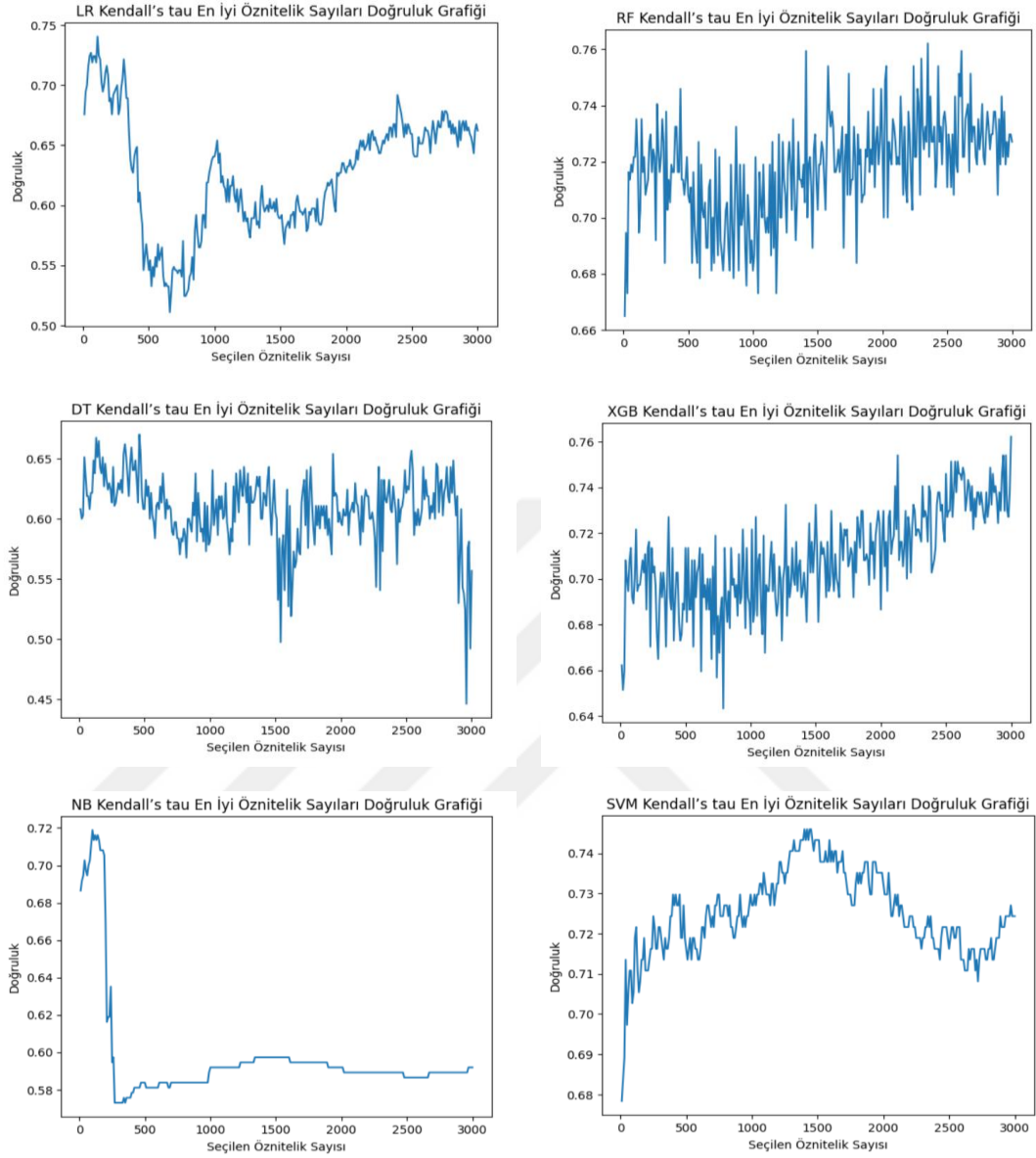
MFCC öznitelikleri ile de modeller eğitilip test edilmiştir. MFCC özniteliklerinde öznitelik seçimi yapılmamıştır. Modellerin başarımları ve karmaşıklık matrisleri tablolar halinde verilmiştir.

3.2.1.1. Erkek /a/ sesi sonuçları

Bu bölümde sadece erkek katılımcıların “a” ünlü harfine ait ses örneklerinin sınıflandırılması yapılmıştır. Bu şartı sağlayan toplamda 259 sağlıklı ve 203 patolojik ses kaydı bulunmaktadır. “a” sesine ait tüm ses tonlarını dahil ettiğimizde toplamda 1036 sağlıklı, 812 patolojik ses kaydı bulunmaktadır. Tablo 3.1’de bu ses kayıtlarının OpenSMILE öznitelikleri için makine öğrenme modellerinin tüm öznitelikler ve Kendall’s Tau öznitelik seçimi modeli kullanılarak seçilen özniteliklerden modellerin en iyi başarımlarını gösterdiği öznitelik sayıları ve başarımları sunulmuştur. Şekil 3.2’de OpenSMILE öznitelikleri içerisinde Kendall’s Tau öznitelik seçimi modeli kullanılarak seçilen 10 ila 3000 arasında 10’un katları sayıda seçilmiş öznitelikler için tüm modellerin her bir öznitelik sayısı için doğruluk değeri grafik halinde gösterilmiştir. OpenSMILE sonuçları incelendiğinde öznitelik seçimi ile başarımların yükseldiği görülmektedir. En iyi sonuç XGB modeli ile Kendall’s Tau öznitelik seçim algoritmasının belirlediği en iyi 3000 öznitelik sayısında %76.21 doğruluk ve %75.69 f1 değeri ile elde edilmiştir.

Tablo 3.1. Erkekler /a/ sesi OpenSMILE öznitelikleri için sınıflandırma başarımları.

Modeller	Tüm Öznitelikler				En İyi Öznitelikler				
	Doğruluk	Kesinlik	Duyarlılık	F1	Doğruluk	Kesinlik	Duyarlılık	F1	Öznitelik Sayısı
LR	71.89	75.36	73.91	71.92	74.05	73.02	85.02	73.52	110
RF	71.62	71.07	83.09	71.02	73.78	73.30	83.57	73.36	290
DT	60.00	63.22	68.11	59.71	67.56	71.01	70.53	67.56	460
XGB	71.35	70.95	82.60	70.77	76.21	74.48	87.43	75.69	3000
SVM	72.70	71.37	85.50	71.96	74.59	72.33	88.40	73.77	1400
NB	59.18	57.95	98.55	48.13	71.89	70.68	85.02	71.10	100

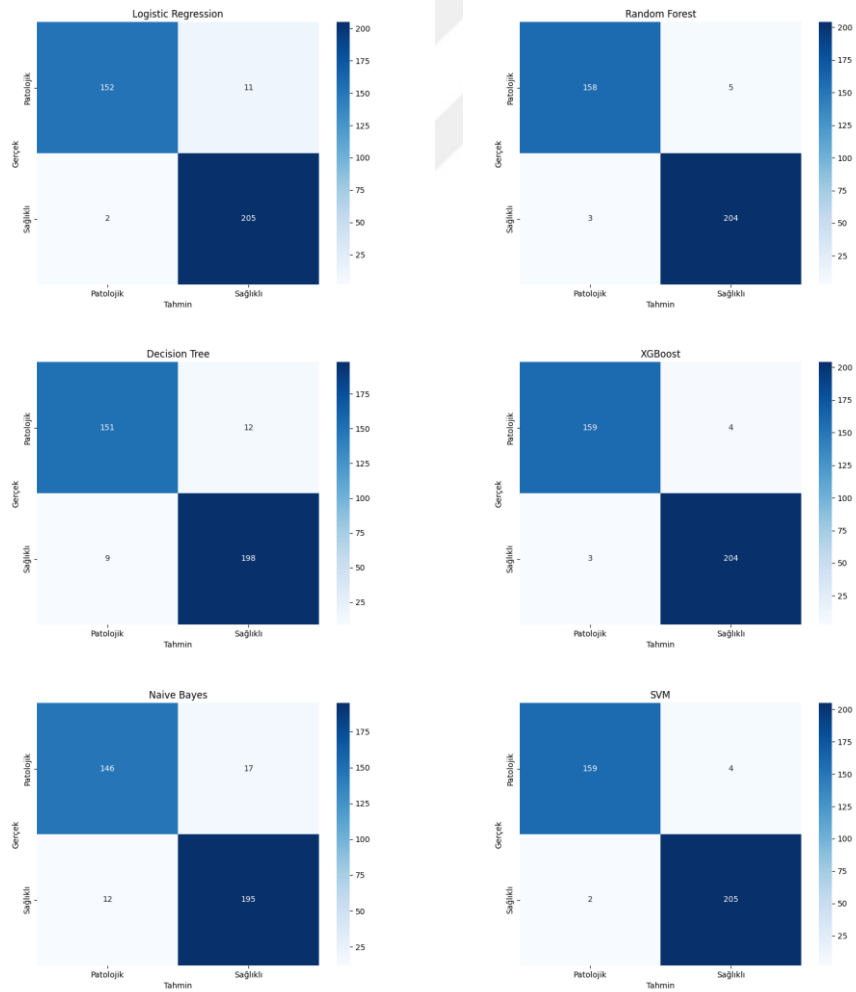


Şekil 3.2. Erkek /a/ sesi için öznitelik sayıları doğruluk grafiği.

MFCC öznitelikleri için modellerin başarımları Tablo 3.2’de gösterilmiştir. Aynı zamanda modellere ait karmaşıklık matrisleri Şekil 3.3’te sunulmuştur.

Tablo 3.2. Erkekler /a/ sesi MFCC öznitelikleri için sınıflandırma başarımları.

Modeller	Doğruluk	Kesinlik	Duyarlılık	F1
LR	96.48	94.90	99.03	96.47
RF	97.83	97.60	98.55	97.83
DT	94.32	94.28	95.65	94.31
XGB	98.10	98.07	98.55	98.10
SVM	98.37	98.08	99.03	98.37
NB	92.16	91.98	94.20	92.14



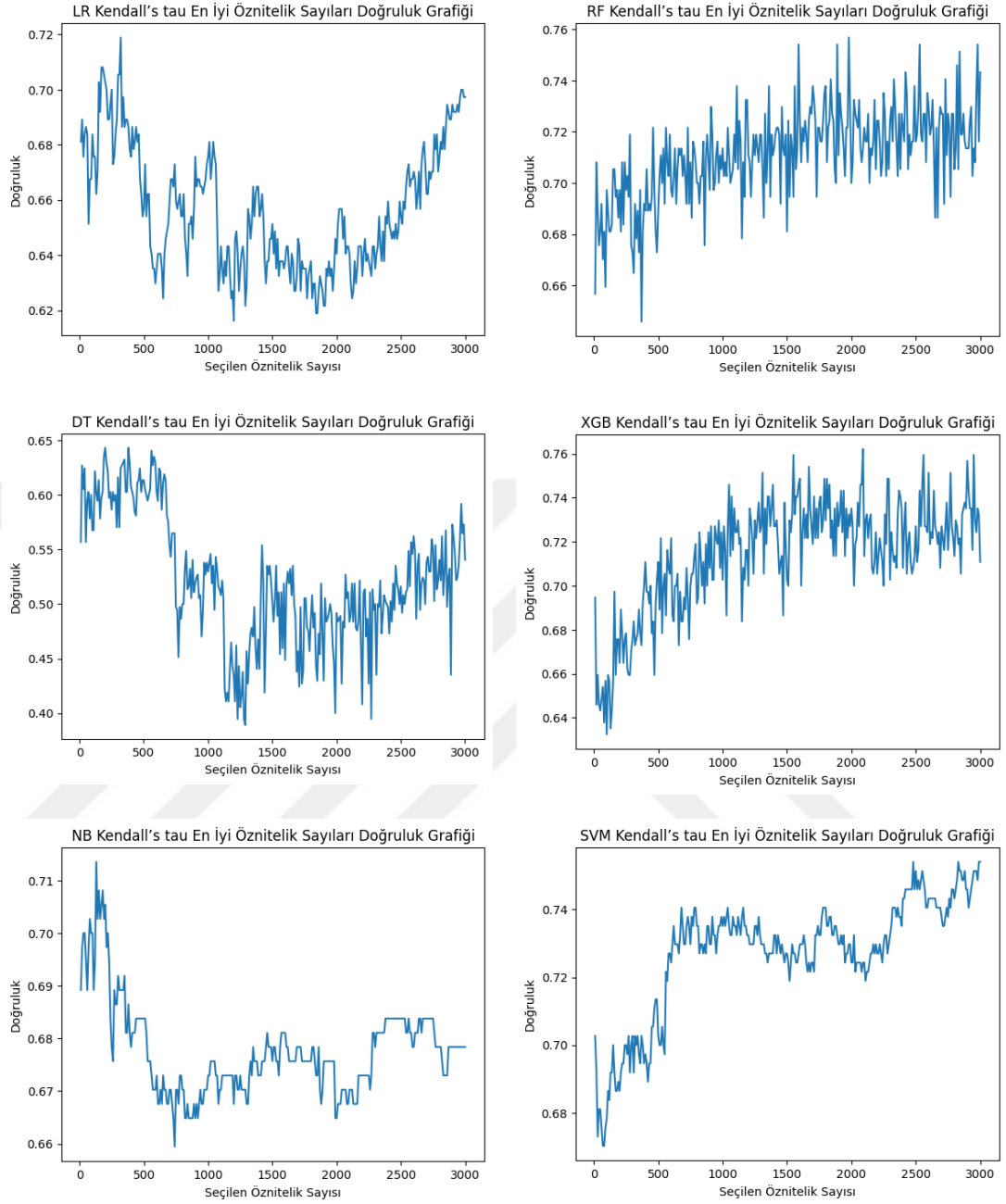
Şekil 3.3. Erkek /a/ sesi MFCC öznitelikleri için modellerin karmaşıklık matrisleri.

3.2.1.2. Erkek /i/ sesi sonuçları

Bu bölümde sadece erkek katılımcıların “i” ünlü harfine ait ses örneklerinin sınıflandırılması yapılmıştır. Bu şartı sağlayan toplamda 259 sağlıklı ve 203 patolojik ses kaydı bulunmaktadır. “i” sesine ait tüm ses tonlarını dahil ettiğimizde toplamda 1036 sağlıklı, 812 patolojik ses kaydı bulunmaktadır. Tablo 3.3’te bu ses kayıtlarının OpenSMILE öznitelikleri için makine öğrenme modellerinin tüm öznitelikler ve Kendall’s Tau öznitelik seçimi modeli kullanılarak seçilen özniteliklerden modellerin en iyi başarımlarını gösterdiği öznitelik sayıları ve başarımları sunulmuştur. Şekil 3.4’te OpenSMILE öznitelikleri içerisinde Kendall’s Tau öznitelik seçimi modeli kullanılarak seçilen 10 ila 3000 arasında 10’un katları sayıda seçilmiş öznitelikler için tüm modellerin her bir öznitelik sayısı için doğruluk değeri grafik halinde gösterilmiştir.

Tablo 3.3. Erkekler /i/ sesi OpenSMILE öznitelikleri için sınıflandırma başarımları.

Modeller	Tüm Öznitelikler				En İyi Öznitelikler				
	Doğruluk	Kesinlik	Duyarlılık	F1	Doğruluk	Kesinlik	Duyarlılık	F1	Öznitelik Sayısı
LR	67.83	70.18	73.91	67.69	71.89	72.29	80.67	71.54	320
RF	70.54	71.68	78.26	70.27	75.67	74.47	85.99	75.22	1980
DT	60.00	63.59	66.66	59.83	64.32	67.12	71.01	64.14	200
XGB	72.43	71.96	83.09	71.92	76.21	77.41	81.15	76.12	2090
SVM	72.43	71.25	85.02	71.71	75.40	73.96	86.47	74.88	2480
NB	44.05	0	0	26.94	71.35	72.24	79.22	71.07	130

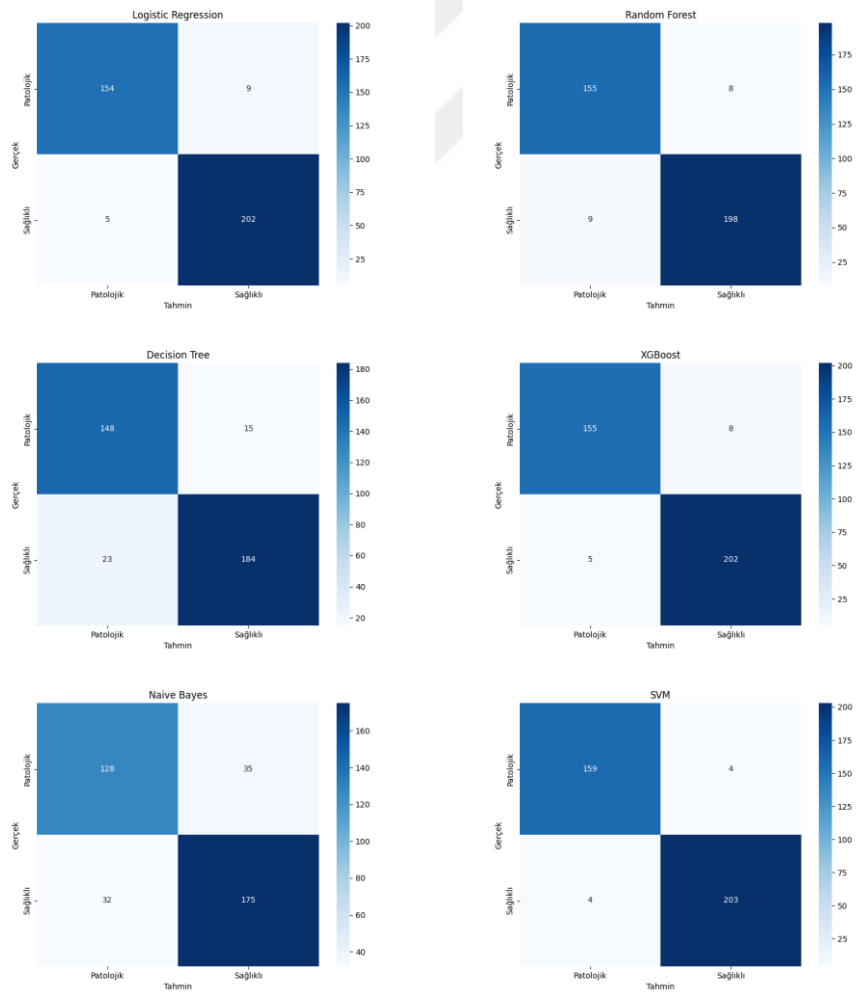


Şekil 3.4. Erkek /i/ sesi için öznitelik sayıları doğruluk grafiği.

MFCC öznitelikleri için modellerin başarımları Tablo 3.4'te gösterilmiştir. Aynı zamanda modellere ait karmaşıklık matrisleri Şekil 3.5'te sunulmuştur.

Tablo 3.4. Erkekler /i/ sesi MFCC öznitelikleri için sınıflandırma başarımları.

Modeller	Doğruluk	Kesinlik	Duyarlılık	F1
LR	96.21	95.73	97.58	96.21
RF	95.40	96.11	95.65	95.40
DT	89.72	92.46	88.88	89.75
XGB	96.48	96.19	97.58	96.48
SVM	97.83	98.06	98.06	97.83
NB	81.89	83.33	84.54	81.87



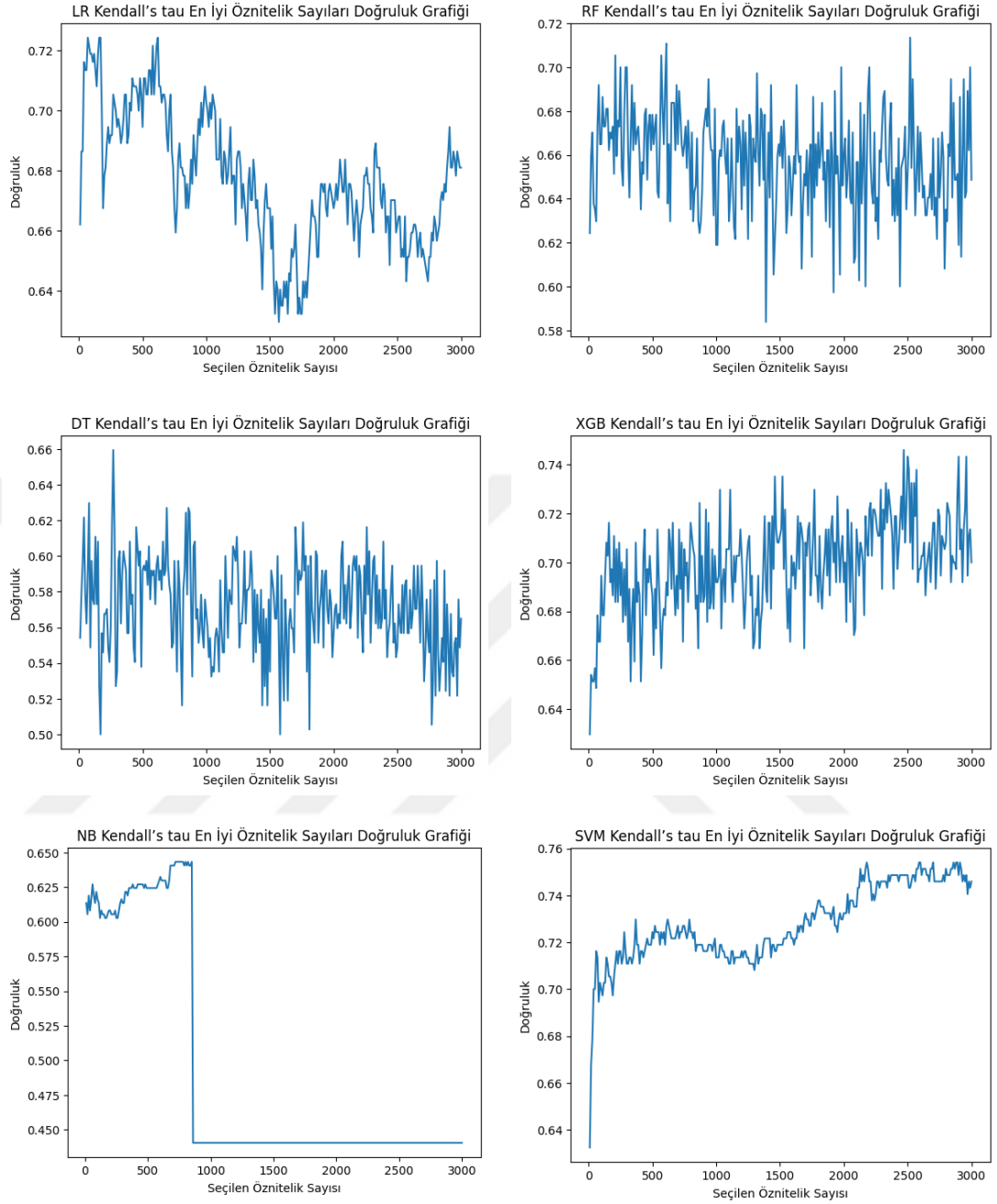
Şekil 3.5. Erkek /i/ sesi MFCC öznitelikleri için modellerin karmaşıklık matrisleri.

3.2.1.3. Erkek /u/ sesi sonuçları

Bu bölümde sadece erkek katılımcıların “u” ünlü harfine ait ses örneklerinin sınıflandırılması yapılmıştır. Bu şartı sağlayan toplamda 259 sağlıklı ve 203 patolojik ses kaydı bulunmaktadır. “u” sesine ait tüm ses tonlarını dahil ettiğimizde toplamda 1034 sağlıklı, 812 patolojik ses kaydı bulunmaktadır. Sağlıklı “u” ses kayıtlarından iki tanesi bozuk çıktığı için sınıflandırmaya dahil edilememiştir. Tablo 3.5’te bu ses kayıtlarının OpenSMILE öznelikleri için makine öğrenme modellerinin tüm öznelıklar ve Kendall’s Tau öznelik seçimi modeli kullanılarak seçilen özneliklerden modellerin en iyi başarımlar gösterdiği öznelik sayıları ve başarımları sunulmuştur. Şekil 3.6’da OpenSMILE öznelikleri içerisinde Kendall’s Tau öznelik seçimi modeli kullanılarak seçilen 10 ila 3000 arasında 10’un katları sayıda seçilmiş öznelikler için tüm modellerin her bir öznelik sayısı için doğruluk değeri grafik halinde gösterilmiştir.

Tablo 3.5. Erkekler /u/ sesi OpenSMILE öznelikleri için sınıflandırma başarımları.

Modeller	Tüm Öznelikler				En İyi Öznelikler				
	Doğruluk	Kesinlik	Duyarlılık	F1	Doğruluk	Kesinlik	Duyarlılık	F1	Öznelik Sayısı
LR	66.75	70.38	70.04	66.76	72.43	70.58	86.95	71.48	70
RF	65.40	70.25	66.18	65.50	71.35	74.39	74.39	71.35	2520
DT	58.37	62.80	62.80	58.37	65.94	69.19	70.53	65.89	270
XGB	69.45	72.16	73.91	69.40	74.59	77.03	77.77	74.57	2470
SVM	74.59	72.69	87.43	73.88	75.40	74.57	85.02	75.01	2180
NB	44.00	0	0	26.94	64.32	61.98	93.71	59.36	720

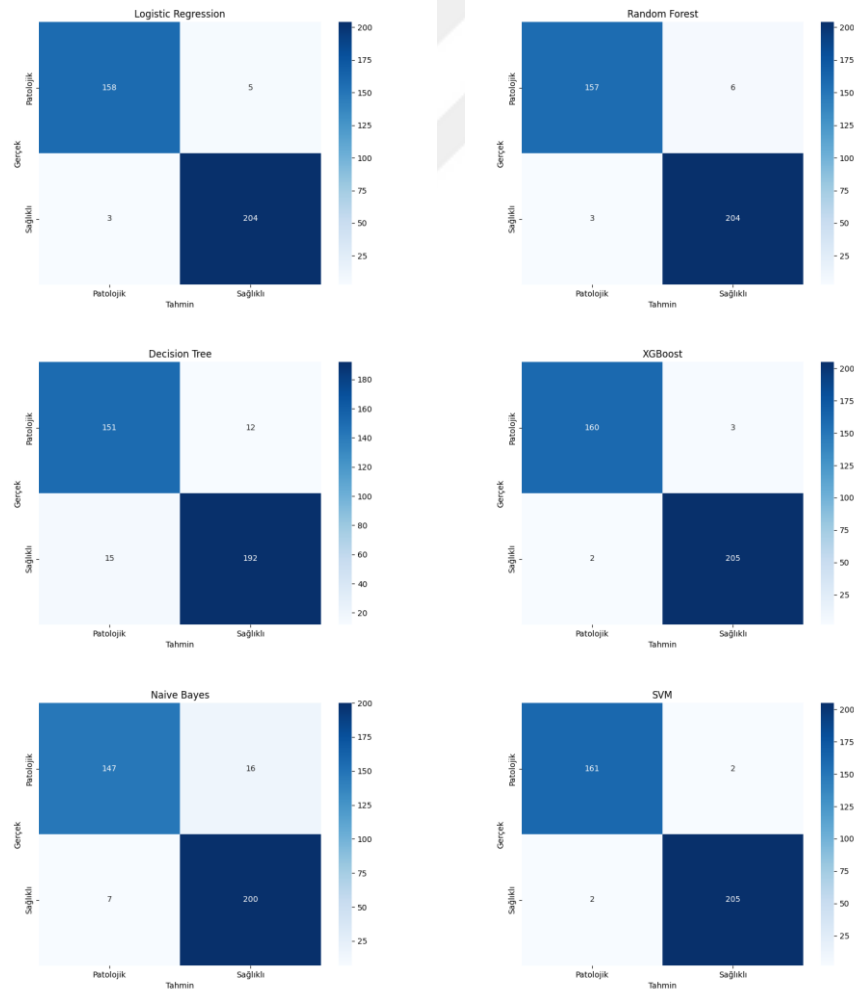


Şekil 3.6. Erkek /u/ sesi için öznitelik sayıları doğruluk grafiği.

MFCC öznitelikleri için modellerin başarımları Tablo 3.6'de gösterilmiştir. Aynı zamanda modellere ait karmaşıklık matrisleri Şekil 3.7'de sunulmuştur.

Tablo 3.6. Erkekler /u/ sesi MFCC öznitelikleri için sınıflandırma başarımları.

Modeller	Doğruluk	Kesinlik	Duyarlılık	F1
LR	97.83	97.60	98.55	97.83
RF	97.56	97.14	98.55	97.56
DT	92.70	94.11	92.75	92.70
XGB	98.64	98.55	99.03	98.64
SVM	98.91	99.03	99.03	98.91
NB	93.78	92.59	96.61	93.76



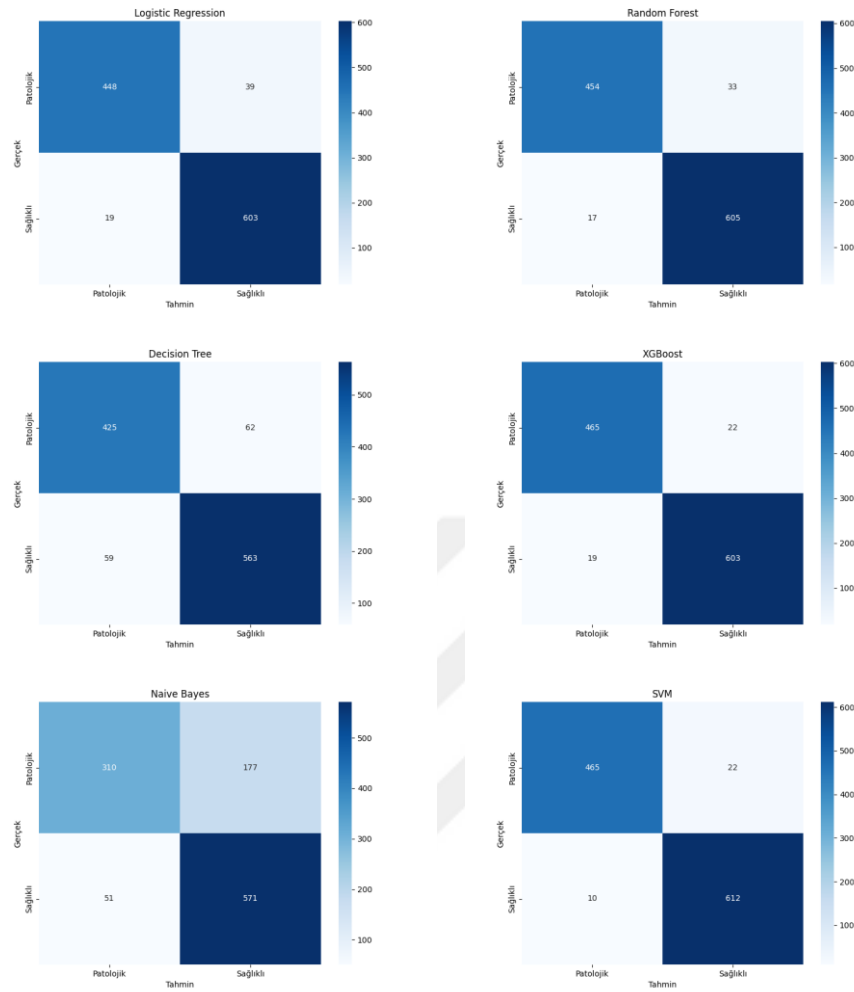
Şekil 3.7. Erkek /u/ sesi MFCC öznitelikleri için modellerin karmaşıklık matrisleri.

3.2.1.4. Erkek tüm ses sonuçları

Bu bölümde erkek katılımcıların tüm ünlü harflere ait ses örneklerinin sınıflandırılması yapılmıştır. Bu şartı sağlayan toplamda 259 sağlıklı ve 203 patolojik ses kaydı bulunmaktadır. Tüm ünlü harflere ait ses tonlarını dahil ettiğimizde 3108 sağlıklı ve 2436 patolojik olmak üzere toplamda 5544 ses kaydı kullanılmıştır. MFCC öznitelikleri için modellerin başarımları Tablo 3.7’de gösterilmiştir. Aynı zamanda modellere ait karmaşıklık matrisleri Şekil 3.8’de sunulmuştur.

Tablo 3.7. Erkek tüm seslerin MFCC öznitelikleri için sınıflandırma başarımları.

Modeller	Doğruluk	Kesinlik	Duyarlılık	F1
LR	94.77	93.92	96.94	94.75
RF	95.49	94.82	97.26	95.48
DT	89.08	90.08	90.51	89.08
XGB	96.30	96.48	96.94	96.30
SVM	97.11	96.52	98.39	97.11
NB	79.44	76.33	91.80	78.85



Şekil 3.8. Erkek tüm sesler MFCC için modellerin karmaşıklık matrisleri.

3.2.2. Kadın Katılımcıların Değerlendirilmesi

SVD veri tabanından 18-65 yaş arası kadın katılımcılardan 428 adet sağlıklı 317 adet ise patolojik ses kaydı elde edilmiştir. Her bir katılımcı “a”, “i”, ve “u” ünlü harflerini dört farklı tonda seslendirmiştir. Bunun sonucunda kadın katılımcılar için 5136 sağlıklı ve 3804 patolojik olmak üzere toplamda 8940 ses kaydı kullanılmıştır. Öncelikle her bir ses kaydı için OpenSMILE ile 6373, MFCC ile 28 öznitelik çıkarılmıştır. Veri seti %80 eğitim ve %20 test olacak şekilde bölünmüştür. Eğitim ve test veri setindeki örnekler rastgele ve patolojik ve sağlıklı oranı iki bölümde de eşit seçilmiştir.

OpenSMILE öznitelikleri ile makine öğrenme modelleri tüm öznitelikler ile eğitim veri setinde eğitilmiş ve test veri setinde test edilmiştir. Tüm özellikler için modellerin performans sonuçları tablolarda gösterilmiştir. OpenSMILE öznitelikleri ile elde edilen başarımların artırarak, eğitim ve test sürelerini düşürmek için öznitelik seçimi uygulanmıştır. Öznitelik seçiminde Kendall's Tau modeli kullanılmıştır. Seçilecek optimum öznitelik sayısını belirlemek için belirli bir aralıktaki bütün öznitelik sayıları tüm modellerde eğitilip test edilmiştir. Bunun için 10 ila 3000 arasında 10'un katları sayıda öznitelik seçilmiş ve tüm modellerin her bir öznitelik sayısı için doğruluk değeri belirlenmiştir. Aşağıdaki grafiklerde Kendall's Tau öznitelik sayılarına göre modellerin doğruluk değeri gösterilmiştir.

MFCC öznitelikleri ile modeller eğitilip test edilmiştir. MFCC özniteliklerinde öznitelik seçimi yapılmamıştır. Modellerin başarımları ve karmaşıklık matrisleri tablolar halinde verilmiştir.

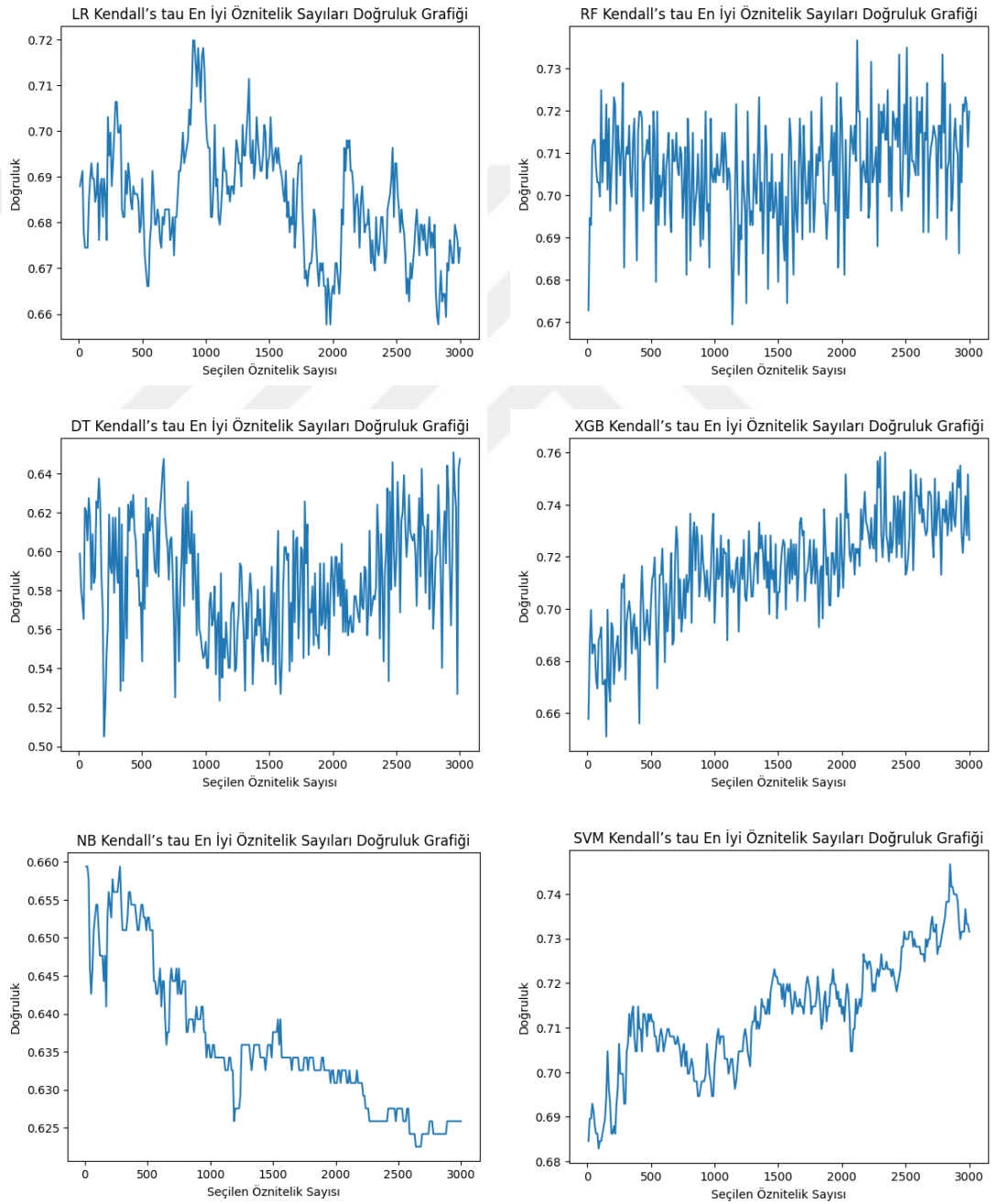
3.2.2.1. Kadın /a/ sesi sonuçları

Bu bölümde sadece kadın katılımcıların "a" ünlü harfine ait ses örneklerinin sınıflandırılması yapılmıştır. Bu şartı sağlayan toplamda 428 sağlıklı ve 317 patolojik ses kaydı bulunmaktadır. "a" sesine ait tüm ses tonlarını dahil ettiğimizde toplamda 1712 sağlıklı, 1268 patolojik ses kaydı bulunmaktadır.

Tablo 3.8. Kadınlar /a/ sesi OpenSMILE öznitelikleri için sınıflandırma başarımları.

	Tüm Öznitelikler				En İyi Öznitelikler				
Modeller	Doğruluk	Keskinlik	Duyarlılık	F1	Doğruluk	Keskinlik	Duyarlılık	F1	Öznitelik Sayısı
LR	67.61	71.84	71.63	67.62	71.97	73.84	79.23	71.74	900
RF	71.30	72.44	80.70	70.88	73.65	74.40	82.45	73.29	2120
DT	60.57	66.56	62.86	60.71	65.10	70.42	67.54	65.20	2950
XGB	73.32	73.16	84.50	72.72	76.00	76.25	84.50	75.67	2340
SVM	71.81	71.64	84.21	71.06	74.66	74.30	85.38	74.12	2850
NB	61.57	63.17	79.23	59.77	65.93	63.44	95.90	60.45	10

Tablo 3.8’de bu ses kayıtlarının OpenSMILE öznitelikleri için makine öğrenme modellerinin tüm öznitelikler ve Kendall’s Tau öznitelik seçimi modeli kullanılarak seçilen özniteliklerden modellerin en iyi başarımlarını gösterdiği öznitelik sayıları ve başarımları sunulmuştur.



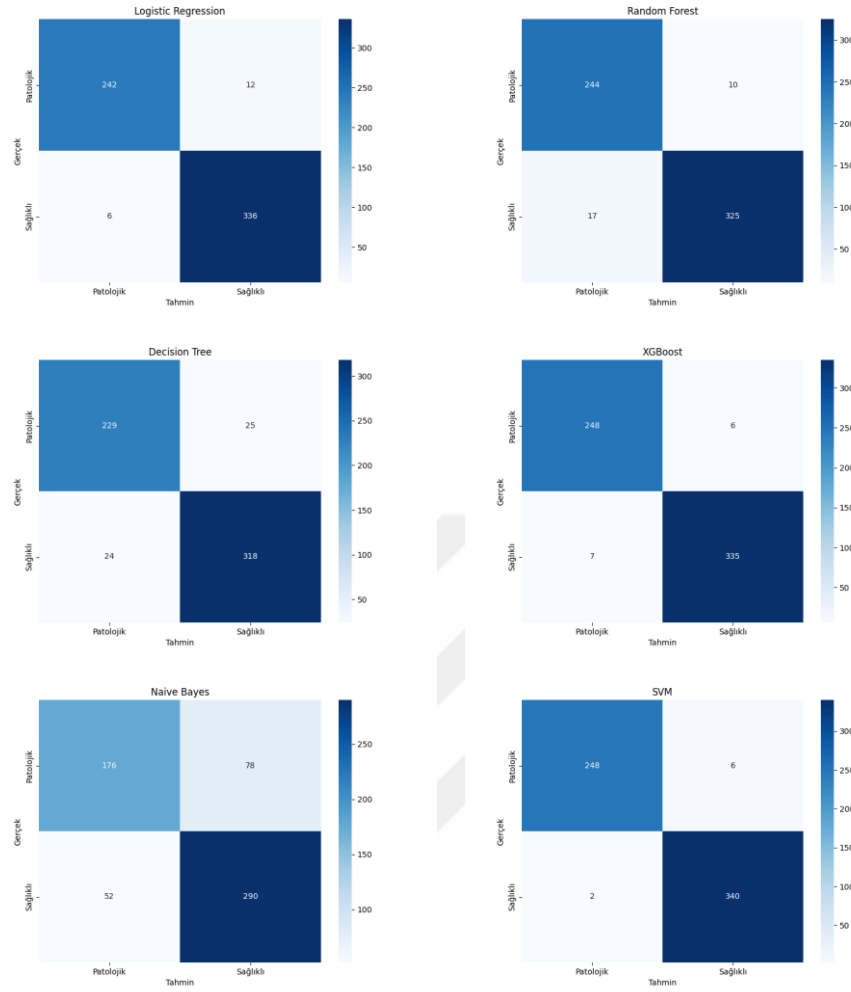
Şekil 3.9. Kadın /a/ sesi için öznitelik sayıları doğruluk grafiği.

Şekil 3.9’da OpenSMILE öznitelikleri içerisinde Kendall’s Tau öznitelik seçimi modeli kullanılarak seçilen 10 ila 3000 arasında 10’un katları sayıda seçilmiş öznitelikler için tüm modellerin her bir öznitelik sayısı için doğruluk değeri grafik halinde gösterilmiştir.

Tablo 3.9. Kadınlar /a/ sesi MFCC öznitelikleri için sınıflandırma başarımları.

Modeller	Doğruluk	Kesinlik	Duyarlılık	F1
LR	96.97	96.55	98.24	96.97
RF	95.46	97.01	95.02	95.47
DT	91.77	92.71	92.98	91.77
XGB	97.81	98.24	97.95	97.81
SVM	98.65	98.26	99.41	98.65
NB	78.18	78.80	84.79	77.99

MFCC öznitelikleri için modellerin başarımlar oranları Tablo 3.9’de gösterilmiştir. Aynı zamanda modellere ait karmaşıklık matrisleri Şekil 3.10’da sunulmuştur.



Şekil 3.10. Kadın /a/ sesi MFCC öz nitelikleri için modellerin karmaşıklık matrisleri.

3.2.2.2. Kadın /i/ sesi sonuçları

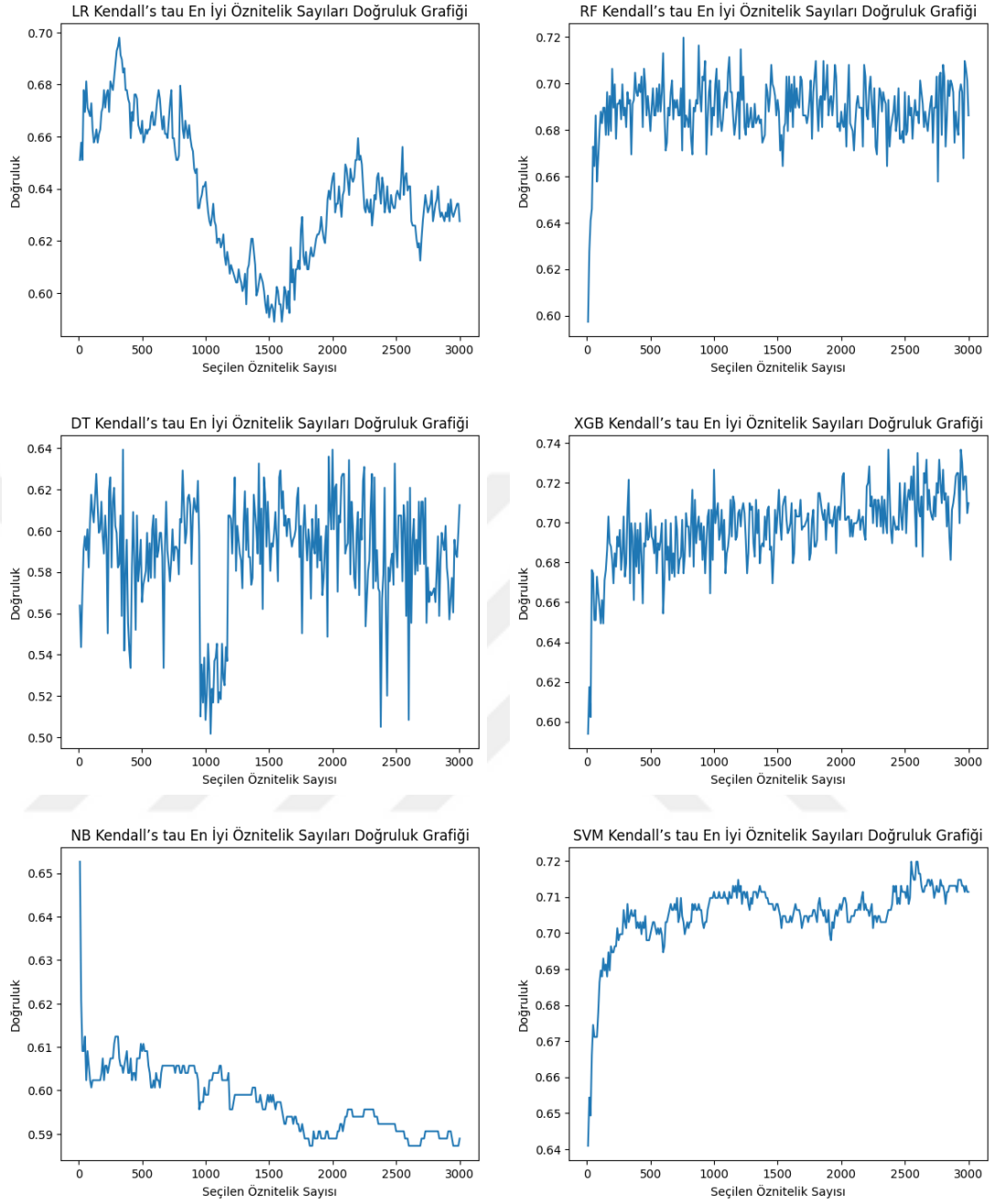
Bu bölümde sadece kadın katılımcıların “i” ünlü harfine ait ses örneklerinin sınıflandırılması yapılmıştır. Bu şartı sağlayan toplamda 428 sağlıklı ve 317 patolojik ses kaydı bulunmaktadır. “i” sesine ait tüm ses tonlarını dahil ettiğimizde toplamda 1712 sağlıklı, 1267 patolojik ses kaydı bulunmaktadır. Patolojik “i” ses kayıtlarından bir tanesi bozuk çıktığı için sınıflandırmaya dahil edilememiştir.

Tablo 3.10. Kadınlar /i/ sesi OpenSMILE öznitelikleri için sınıflandırma başarımları.

Modeller	Tüm Öznitelikler				En İyi Öznitelikler				
	Doğruluk	Kesinlik	Duyarlılık	F1	Doğruluk	Kesinlik	Duyarlılık	F1	Öznitelik Sayısı
LR	64.09	68.06	70.84	63.94	69.79	71.27	79.59	69.33	320
RF	69.12	73.17	73.17	69.12	71.97	73.65	79.	71.69	760
DT	55.03	61.53	58.30	55.20	63.92	67.48	72.01	63.65	350
XGB	69.96	73.56	74.63	69.92	73.65	77.67	76.09	73.70	2370
SVM	71.14	71.86	81.92	70.57	71.97	73.40	80.46	71.64	2550
NB	42.44	0	0	25.29	65.26	67.98	74.92	64.82	10

Tablo 3.10’da bu ses kayıtlarının OpenSMILE öznitelikleri için makine öğrenme modellerinin tüm öznitelikler ve Kendall’s Tau öznitelik seçimi modeli kullanılarak seçilen özniteliklerden modellerin en iyi başarımlarını gösterdiği öznitelik sayıları ve başarımlarları sunulmuştur.

Şekil 3.11’de OpenSMILE öznitelikleri içerisinde Kendall’s Tau öznitelik seçimi modeli kullanılarak seçilen 10 ila 3000 arasında 10’un katları sayıda seçilmiş öznitelikler için tüm modellerin her bir öznitelik sayısı için doğruluk değeri grafik halinde gösterilmiştir.

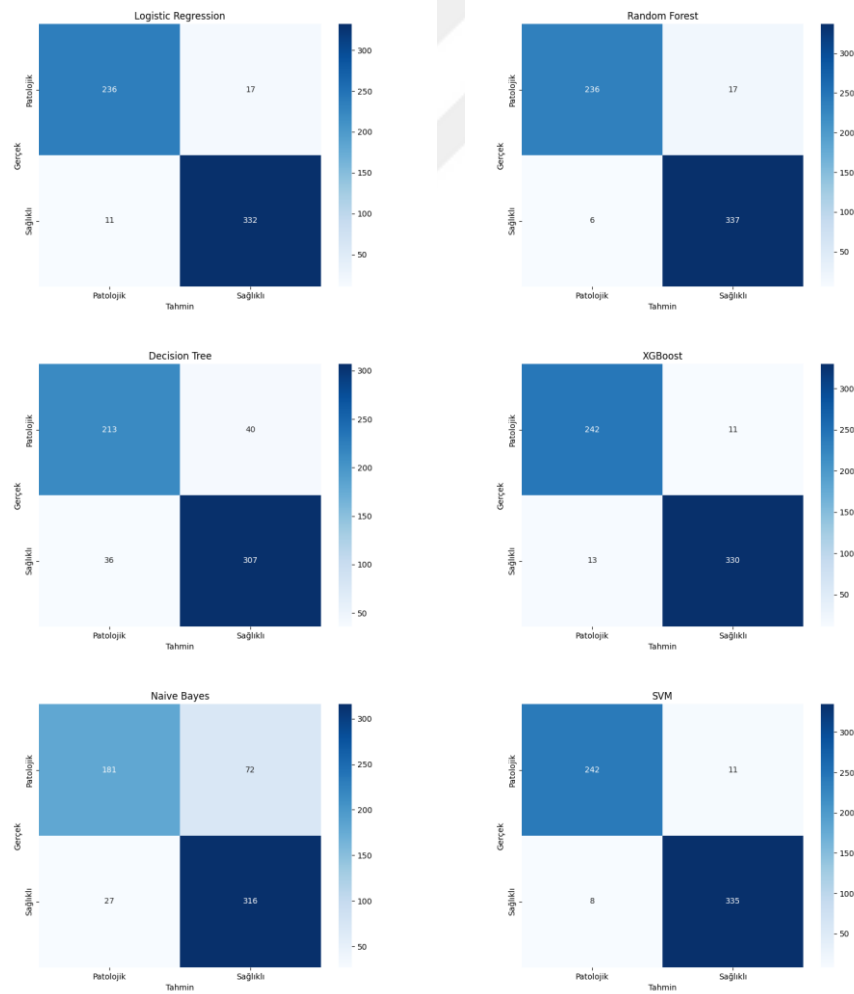


Şekil 3.11. Kadın /i/ sesi için öznitelik sayıları doğruluk grafiği.

MFCC öznitelikleri için modellerin başarımları Tablo 3.11’de gösterilmiştir. Aynı zamanda modellere ait karmaşıklık matrisleri Şekil 3.12’de sunulmuştur.

Tablo 3.11. Kadınlar /i/ sesi MFCC öznitelikleri için sınıflandırma başarımları.

Modeller	Doğruluk	Kesinlik	Duyarlılık	F1
LR	95.30	95.12	96.79	95.29
RF	96.14	95.19	98.25	96.12
DT	87.24	88.47	89.50	87.23
XGB	95.97	96.77	96.20	95.97
SVM	96.81	96.82	97.66	96.80
NB	83.38	81.44	92.12	83.08



Şekil 3.12. Kadın /i/ sesi MFCC öznitelikleri için modellerin karmaşıklık matrisleri.

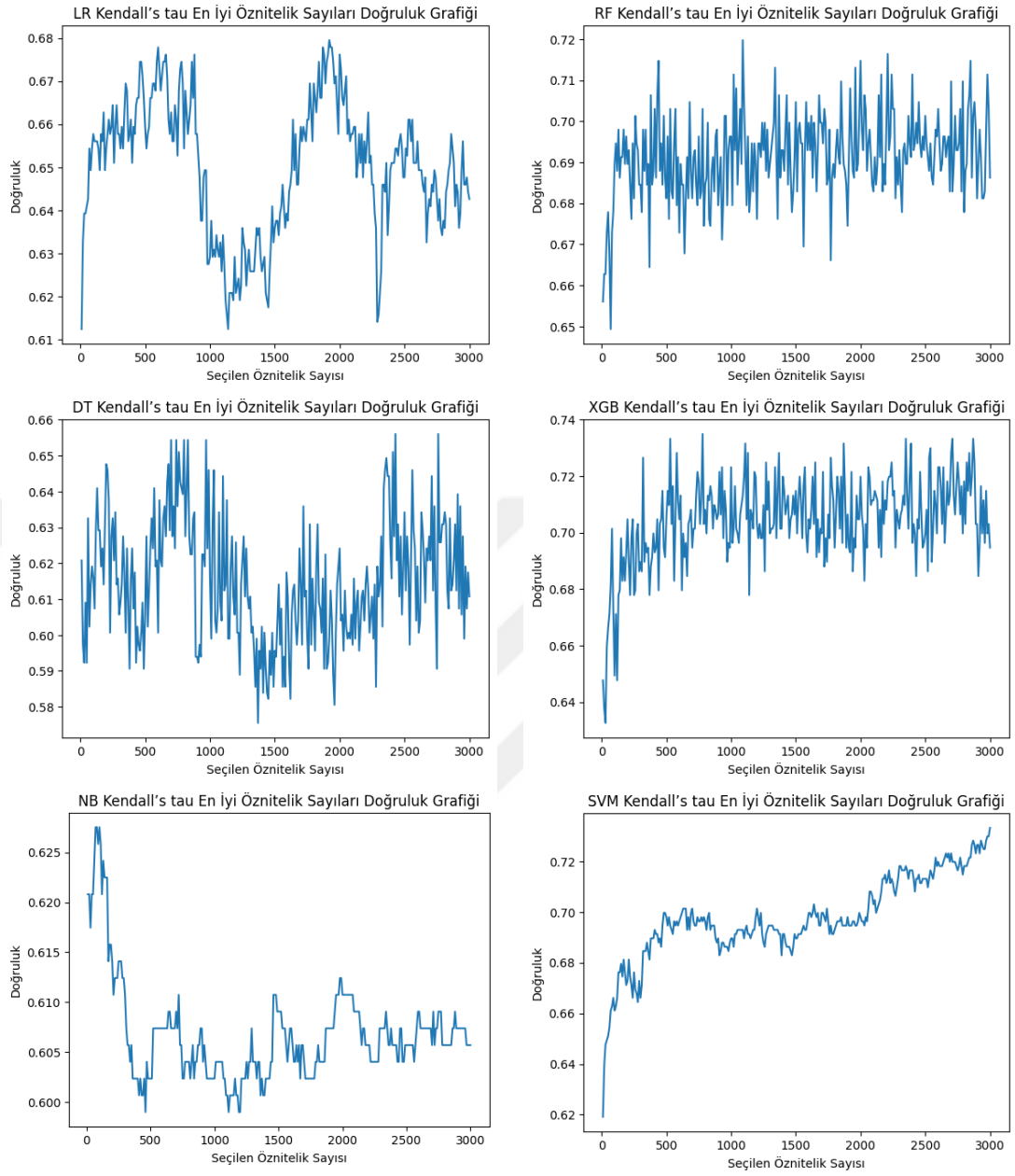
3.2.2.3. Kadın /u/ sesi sonuçları

Bu bölümde sadece kadın katılımcıların “u” ünlü harfine ait ses örneklerinin sınıflandırılması yapılmıştır. Bu şartı sağlayan toplamda 428 sağlıklı ve 317 patolojik ses kaydı bulunmaktadır. “u” sesine ait tüm ses tonlarını dahil ettiğimizde toplamda 1712 sağlıklı, 1268 patolojik ses kaydı bulunmaktadır. Tablo 3.12’de bu ses kayıtlarının OpenSMILE öznitelikleri için makine öğrenme modellerinin tüm öznitelikler ve Kendall’s Tau öznitelik seçimi modeli kullanılarak seçilen özniteliklerden modellerin en iyi başarımlar gösterdiği öznitelik sayıları ve başarımları sunulmuştur. Şekil 3.11’de OpenSMILE öznitelikleri içerisinde Kendall’s Tau öznitelik seçimi modeli kullanılarak seçilen 10 ila 3000 arasında 10’un katları sayıda seçilmiş öznitelikler için tüm modellerin her bir öznitelik sayısı için doğruluk değeri grafik halinde gösterilmiştir.

Tablo 3.12. Kadınlar /u/ sesi OpenSMILE öznitelikleri için sınıflandırma başarımları.

Modeller	Tüm Öznitelikler				En İyi Öznitelikler				
	Doğruluk	Kesinlik	Duyarlılık	F1	Doğruluk	Kesinlik	Duyarlılık	F1	Öznitelik Sayısı
LR	65.77	70.17	70.17	65.77	67.95	72.67	70.76	68.01	1920
RF	67.44	67.96	81.87	66.35	71.97	72.37	82.74	71.14	1090
DT	59.22	65.81	60.23	59.42	65.60	70.44	69.00	65.65	2430
XGB	71.81	73.01	80.70	71.44	73.48	73.58	83.91	72.97	780
SVM	71.64	71.04	85.38	70.71	73.32	73.76	83.04	72.87	3000
NB	60.73	61.11	86.84	56.38	62.75	61.19	95.90	55.36	70

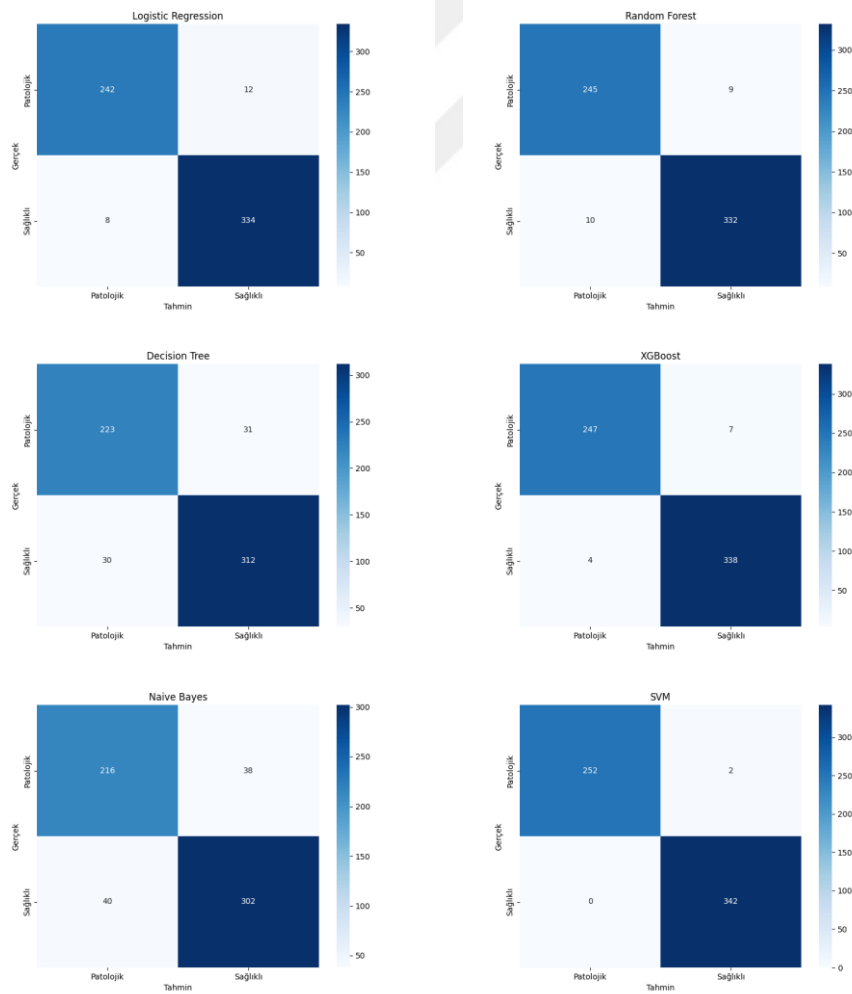
MFCC öznitelikleri için modellerin başarımları Tablo 3.13’te gösterilmiştir. Aynı zamanda modellere ait karmaşıklık matrisleri Şekil 3.14’te sunulmuştur.



Şekil 3.13. Kadın /u/ sesi için özellik sayıları doğruluk grafiği.

Tablo 3.13. Kadınlar /u/ sesi MFCC öznitelikleri için sınıflandırma başarımları.

Modeller	Doğruluk	Kesinlik	Duyarlılık	F1
LR	96.64	96.53	97.66	96.64
RF	96.81	97.36	97.07	96.81
DT	89.76	90.96	91.22	89.76
XGB	98.15	97.97	98.83	98.15
SVM	99.66	99.41	100.00	99.66
NB	86.91	88.82	88.30	86.91



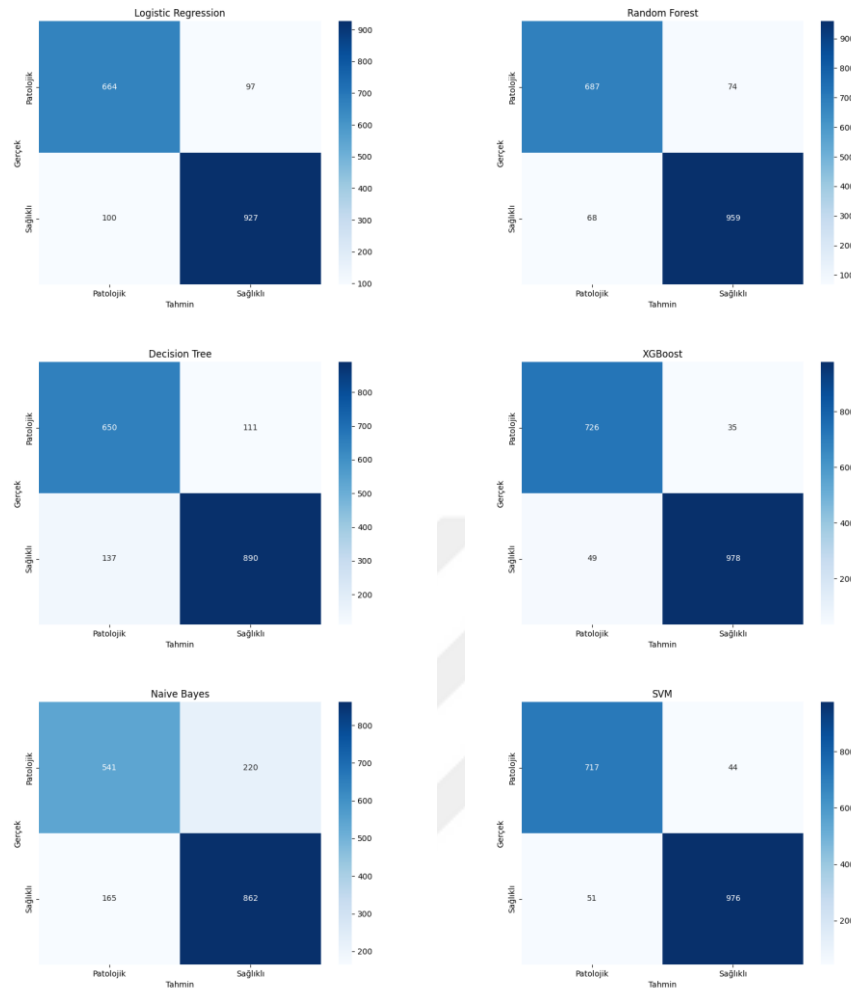
Şekil 3.14. Kadın /u/ sesi MFCC öznitelikleri için modellerin karmaşıklık matrisleri.

3.2.2.4. Kadın tüm ses sonuçları

Bu bölümde kadın katılımcıların tüm ünlü harflere ait ses örneklerinin sınıflandırılması yapılmıştır. Bu şartı sağlayan toplamda 428 sağlıklı ve 317 patolojik ses kaydı bulunmaktadır. Tüm ünlü harflere ait ses tonlarını dahil ettiğimizde 5136 sağlıklı ve 3803 patolojik olmak üzere toplamda 8939 ses kaydı kullanılmıştır. MFCC öznitelikleri için modellerin başarımları Tablo 3.14'te gösterilmiştir. Aynı zamanda modellere ait karmaşıklık matrisleri Şekil 3.15'te sunulmuştur.

Tablo 3.14. Kadın tüm seslerin MFCC öznitelikleri için sınıflandırma başarımları.

Modeller	Doğruluk	Kesinlik	Duyarlılık	F1
LR	88.98	90.52	90.26	88.98
RF	92.05	92.83	93.37	92.05
DT	86.12	88.91	86.66	86.15
XGB	95.30	96.54	95.22	95.30
SVM	94.68	95.68	95.03	94.68
NB	78.46	79.66	83.93	78.34



Şekil 3.15. Kadın tüm sesler MFCC için modellerin karmaşıklık matrisleri.

3.2.3. Tüm Katılımcıların Değerlendirilmesi

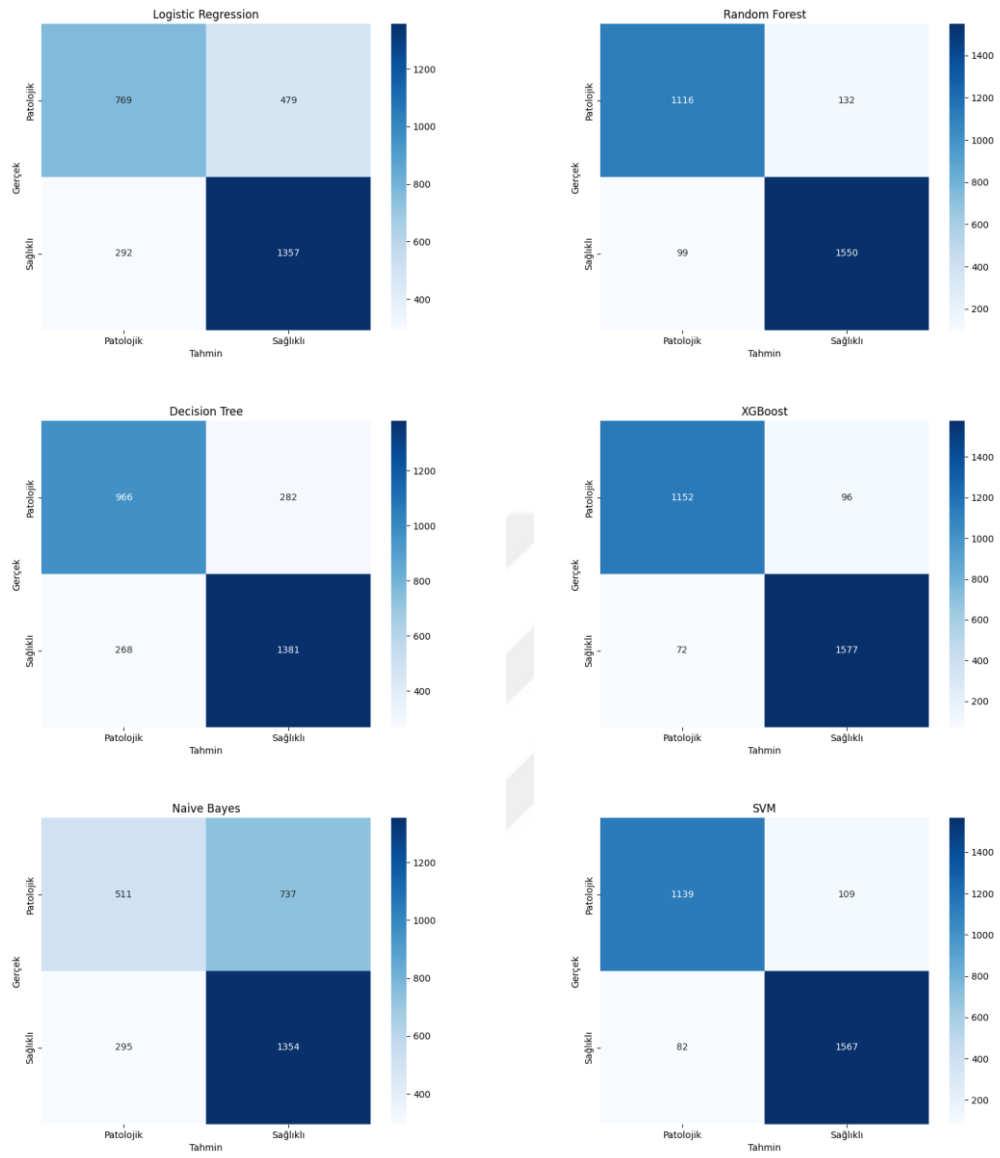
Bu bölümde kadın ve erkek tüm katılımcıların tüm ünlü harflere ait ses örneklerinin sınıflandırılması yapılmıştır. Bu şartı sağlayan toplamda 687 sağlıklı ve 520 patolojik ses kaydı bulunmaktadır. Tüm ünlü harflere ait ses tonlarını dahil ettiğimizde 8242 sağlıklı ve 6239 patolojik olmak üzere toplamda 14481 ses kaydı kullanılmıştır. MFCC öznitelikleri için modellerin başarımları Tablo 3.15’te gösterilmiştir. Aynı zamanda modellere ait karmaşıklık matrisleri Şekil 3.16’da sunulmuştur.

Tablo 3.15. Tüm seslerin MFCC öznitelikleri için sınıflandırma başarımları.

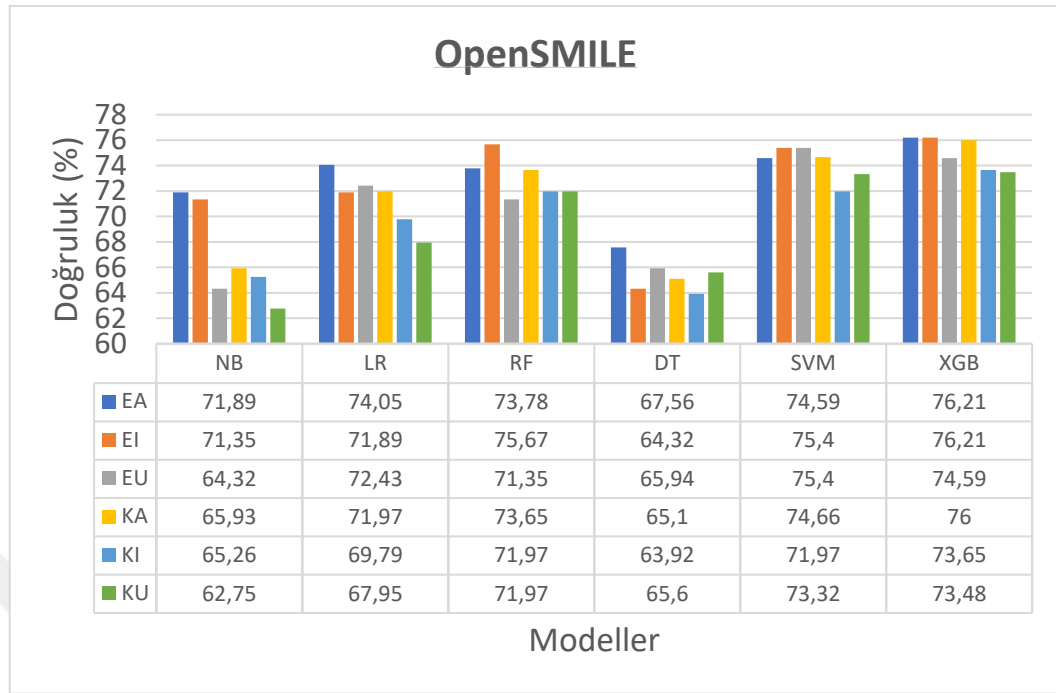
Modeller	Doğruluk	Kesinlik	Duyarlılık	F1
LR	73.38	73.91	82.29	73.02
RF	92.02	92.15	93.99	92.01
DT	81.01	83.04	83.74	81.00
XGB	94.20	94.26	95.63	94.19
SVM	93.40	93.49	95.02	93.39
NB	64.37	64.75	82.11	62.64

Şekil 3.17’de tüm sınıflandırma gruplarının OpenSMILE öznitelikleri ile sınıflandırma sonuçları gösterilmektedir. EA erkek /a/ sesi, EI erkek /i/ sesi, EU erkek /u/ sesi, KA kadın /a/ sesi, KI kadın /i/ sesi ve KU kadın /u/ sesini göstermektedir.

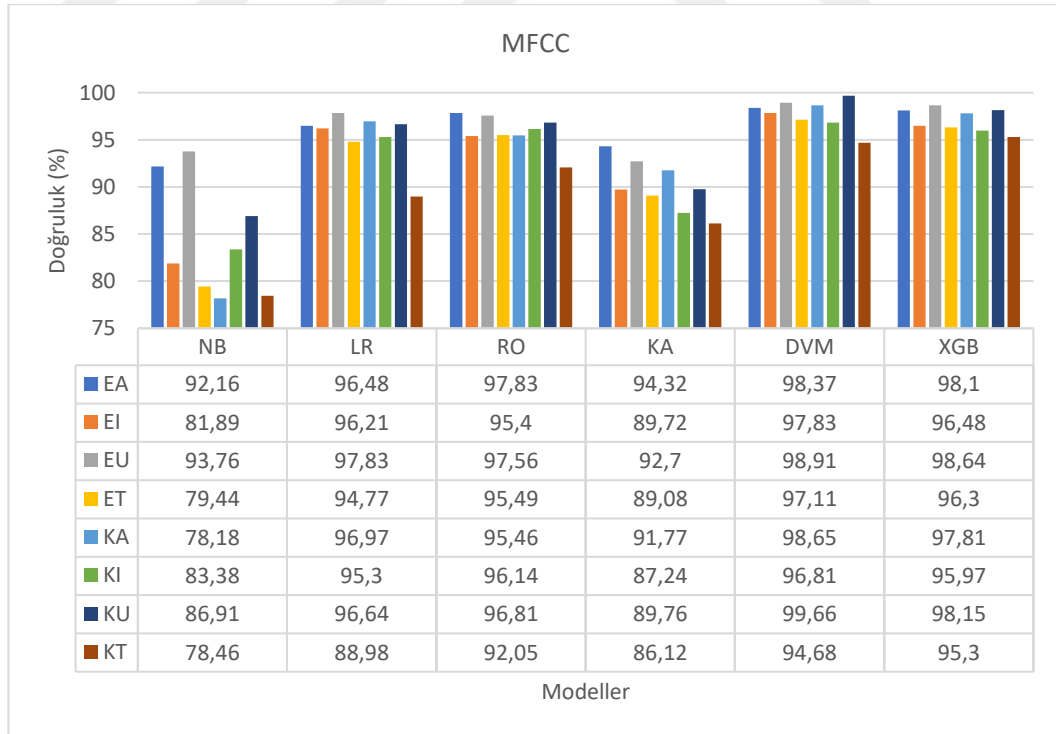
Şekil 3.18’de tüm sınıflandırma gruplarının MFCC öznitelikleri ile sınıflandırma sonuçları gösterilmektedir. EA erkek /a/ sesi, EI erkek /i/ sesi, EU erkek /u/ sesi, ET erkek tüm sesler, KA kadın /a/ sesi, KI kadın /i/ sesi, KU kadın /u/ sesi ve KT kadın tüm sesleri göstermektedir.



Şekil 3.16. Tüm seslerin MFCC öznitelikleri için modellerin karmaşıklık matrisleri.



Şekil 3.17. Tüm sınıflandırma gruplarının OpenSMILE için sınıflandırma sonuçları.



Şekil 3.18. Tüm sınıflandırma gruplarının MFCC için sınıflandırma sonuçları.

4. BÖLÜM

TARTIŞMA-SONUÇ VE ÖNERİLER

4.1. Tartışma

Bu tez çalışmasında, makine öğrenme yöntemleri kullanılarak patolojik seslerin belirlenmesi için farklı algoritma modelleri üzerine çalışılmıştır. Literatürde ses analizinde popüler olarak kullanılan algoritmaların yanı sıra farklı modellerde denenmiş ve sonuçları karşılaştırılmıştır. Çalışmada kullanılan sesler SVD veritabanından alınmıştır. 687'si sağlıklı ve 520 patolojik olmak üzere 1207 kişiye ait kayıtlar kullanılmıştır. Kullanılan patolojik sesler larenjit, disfoni, hiperkinetik disfoni, hipokinetik disfoni ve spazmodik disfoni gruplarına aittir. Bu çalışma için her bir bireyin dört farklı tonda ses kaydı /a/, /i/, /u/ ünlüleri için kullanılmıştır. Veriler, sırasıyla %80 ve %20 örnek içeren eğitim seti ve test seti olarak ayrılmıştır. Patolojik seslerin belirlenmesine yönelik çalışmalar yalnızca teknik açıdan değil, aynı zamanda hem analiz için seçilen ses kayıtları hem de sınıflandırma için patolojilerin alt kümelerinin nasıl kullanıldığı açısından farklılık göstermektedir. Literatürde günümüze kadar yapılmış çalışmalarda araştırmacıların en büyük problemi veri azlığı sebebiyle homojen bir veri seti oluşturmamak olmuştur. Bu tez çalışmasında diğer çalışmalara göre oldukça homojen bir veri seti oluşturulmuştur.

Çalışmada beş farklı sınıflandırma algoritması kullanılarak performansları değerlendirilmiştir. Kadın ve erkek sesleri kendi arasında farklı özellikler gösterdiği için çalışma esnasında önce ayrı ayrı işleme alınmış daha sonra birleştirilerek anlamlı bir fark olup olmadığı araştırılmıştır.

Öncelikle erkeklere ait ses kayıtları OpenSMILE öznitelik çıkarım yöntemi ve Kendall's Tau öznitelik seçimi işlemine tabi tutulmuş ve her bir ünlü harf için ayrı ayrı incelenmiştir. /a/ ünlü harfi için en iyi sonuç %75.69 ile XGB algoritmasında alınmıştır. Bu değer öznitelik seçimi yapmadan önce %70,77'dir.

Tablo 3.1'de erkek /a/ sesinin tüm OpenSMILE öznitelikler ve Kendall's Tau ile belirlenmiş en iyi öznitelikler ile sınıflandırma sonuçları verilmiştir. Sonuçlar incelendiğinde Kendall's Tau uygulanması ile tüm modellerin sınıflandırma başarımları arttırılmıştır. En iyi doğruluk XGB algoritması ile en iyi 3000 öznitelikte %76,21 olarak elde edilmiştir. NB ve LR algoritmaları ise sırasıyla 100 ve 110 öznitelik ile en az sayıda öznitelik ile sınıflandırma başarımlarını arttırmıştır. Şekil 3.2'de her modelin 10 ile 3000 arasında 10'un katları olacak şekilde Kendall's Tau öznitelikleri için sınıflandırma başarımlarını göstermektedir. Grafikler incelendiğinde LR, NB ve SVM modellerinin başarımları stabil bir şekilde değişirken diğer modellerde sürekli artıp azalan bir başarımlar görülmektedir. Bu durum Kendall's Tau öznitelik seçim modelinin OpenSMILE öznitelikleri üzerinde tam anlamıyla doğru seçim yapamadığından kaynaklanmaktadır. Tablo 3.2'de erkek /a/ sesinin tüm MFCC öznitelikleri ile sınıflandırma sonuçları verilmiştir. Tablo incelendiğinde bütün modellerin %90'ların üzerinde bir başarımlar elde ettiği görülmektedir. En iyi başarımlar %98,37 ile SVM modeli ile elde edilmiştir. Bu sonuçlar erkek /a/ sesi için MFCC özniteliklerinin patolojik seslerin tespitinde önemli bir öznitelik kümesi olduğunu göstermektedir.

Tablo 3.3'te erkek /i/ sesi için tüm OpenSMILE öznitelikleri ve Kendall's Tau öznitelikleri için sınıflandırma sonuçları verilmiştir. En iyi doğruluk XGB algoritması ile en iyi 2090 öznitelikte %76,21 olarak elde edilmiştir. Tablo 3.4'te erkek /i/ sesinin tüm MFCC öznitelikleri ile sınıflandırma sonuçları verilmiştir. En iyi doğruluk SVM algoritması ile %97,83 olarak elde edilmiştir. LR, RF ve XGB modelleri de %95'in üzerinde doğruluklar elde ederken en düşük doğruluk değerini %81,89 ile NB modeli elde etmiştir.

Tablo 3.5 incelendiğinde erkek /u/ sesi için en iyi doğruluk SVM algoritması ile en iyi 2180 öznitelikte %75,4 olarak elde edilmiştir. Tablo 3.6'da erkek /u/ sesinin tüm MFCC öznitelikleri ile sınıflandırma sonuçları verilmiştir. En iyi doğruluk SVM algoritması ile %98,91 olarak elde edilmiştir. Diğer modellerin de %90'ların üzerinde bir başarımlar elde

ettiği görülmektedir. MFCC öznitelikleri erkek /a/ sesinde olduğu gibi /u/ sesinde de patolojik seslerin tespitinde önemli bir öznitelik kümesi olduğunu göstermektedir.

Tablo 3.7’de bütün erkek seslerinin MFCC öznitelikleri ile sınıflandırma sonuçları verilmiştir. En iyi doğruluk SVM algoritması ile %97,11 olarak elde edilmiştir. LR, RF ve XGB modelleri de %94’ün üzerinde doğruluklar elde ederken en düşük doğruluk değerini %79,44 ile NB modeli elde etmiştir. Erkeklerin tüm harflerine ait sesler birleştirilerek yapılan sınıflandırmada harfler ayrı olarak yapılan sınıflandırmaya göre genel olarak doğruluklarda küçük bir düşüş görülmektedir ancak yine de patolojik seslerin tespitinde önemli doğruluk elde edilmiştir.

Kadın /a/ sesinin tüm OpenSMILE öznitelikleri ve Kendall’s Tau ile belirlenmiş en iyi öznitelikler ile sınıflandırma sonuçları Tablo 3.8’de verilmiştir. Sonuçlar incelendiğinde Kendall’s Tau uygulanması ile tüm modellerin sınıflandırma başarımları arttırılmıştır. En iyi doğruluk XGB algoritması ile en iyi 2340 öznitelikte %76 olarak elde edilmiştir. NB algoritması %65,93 doğruluk değerinde kalmasına rağmen 10 öznitelik ile en az sayıda öznitelik ile sınıflandırma başarımını arttırmıştır. Şekil 3.9’da her modelin 10 ile 3000 arasında 10’un katları olacak şekilde Kendall’s Tau öznitelikleri için sınıflandırma başarımlarını göstermektedir. Grafikler incelendiğinde NB ve SVM modellerinin başarımları nispeten stabil bir şekilde değişirken diğer modellerde sürekli artıp azalan bir başarımlar görülmektedir. Tablo 3.9’da kadın /a/ sesinin tüm MFCC öznitelikleri ile sınıflandırma sonuçları verilmiştir. Tablo incelendiğinde en iyi başarımlar %98,65 ile SVM modeli ile elde edilmiştir. LR, RF ve XGB modelleri ile %95’in üzerinde doğruluk elde edilirken DT modeli ile %91,81 ve NB modeli ile %78,18 doğruluk elde edilmiştir.

Tablo 3.10’de kadın /i/ sesi için tüm OpenSMILE öznitelikleri ve Kendall’s Tau öznitelikleri için sınıflandırma sonuçları verilmiştir. En iyi doğruluk XGB algoritması ile en iyi 2370 öznitelikte %73,65 olarak elde edilmiştir. Tablo 3.11’de kadın /i/ sesinin tüm MFCC öznitelikleri ile sınıflandırma sonuçları verilmiştir. En iyi doğruluk SVM algoritması ile %96,81 olarak elde edilmiştir. LR, RF ve XGB modelleri ile %95’in üzerinde doğruluklar elde edilirken en düşük doğruluk değeri %83,38 ile NB modeli ile elde edilmiştir.

Tablo 3.12 incelendiğinde kadın /u/ sesi için en iyi doğruluk XGB algoritması ile en iyi 780 öznitelikte %73,48 olarak elde edilmiştir. Tablo 3.13’te kadın /u/ sesinin tüm MFCC

öznitelikleri ile sınıflandırma sonuçları verilmiştir. En iyi doğruluk SVM algoritması ile %99,66 olarak elde edilmiştir. LR, RF ve XGB modelleri ile %96'nın üzerinde doğruluk elde edilirken en düşük doğruluk %86,91 ile NB modeli ile elde edilmiştir. Kadın /u/ MFCC öznitelikleri erkek /a/ sesinde olduğu gibi /u/ sesinde de patolojik seslerin tespitinde önemli bir öznitelik kümesi olduğunu göstermektedir.

Tablo 3.14'te bütün kadın seslerinin MFCC öznitelikleri ile sınıflandırma sonuçları verilmiştir. En iyi doğruluk XGB algoritması ile %95,3 olarak elde edilmiştir. RF ve SVM modelleri %92'nin üzerinde doğruluklar elde ederken en düşük doğruluk değerini %78,46 ile NB modeli elde etmiştir. Kadınların tüm harflerine ait sesler birleştirilerek yapılan sınıflandırmada harfler ayrı olarak yapılan sınıflandırmaya göre genel olarak doğruluklarda önemli ölçüde bir düşüş görülmektedir. Bu durum kadınlarda harflerin ayrı olarak sınıflandırılmasının patolojik seslerin tespitinde daha etkin olduğunu göstermektedir. Özellikle /u/ sesi kadınlarda patolojik seslerin tespitinde en yüksek başarıyı sağlamıştır.

Şekil 3.17 incelendiğinde OpenSMILE ile en iyi sonuç erkeklerde /i/ sesinde %76,21 doğruluk ve %76,12 F1 skoru ile XGB modelinde elde edilmiştir. Kadınlarda ise en iyi sonuç /a/ sesinde %76,00 doğruluk ve %75,67 F1 skoru ile yine XGB modeli ile elde edilmiştir.

Şekil 3.18 incelendiğinde MFCC ile en iyi sonuç erkeklerde /a/ sesinde %98,91 doğruluk ve F1 skoru ile SVM modeli ile elde edilmiştir. Kadınlarda ise en iyi sonuç /u/ sesinde %99,66 doğruluk ve F1 skoru ile yine SVM modeli ile elde edilmiştir.

MFCC ve OpenSMILE dan elde edilen özniteliklerle modellerin başarımları karşılaştırıldığında MFCC'de %95 leri aşan doğruluk değerleri elde edilirken OpenSMILE ile en yüksek %80 lere yakın doğruluk elde edilmiştir. Bu durum MFCC özniteliklerinin ses patolojisi tanısında oldukça yüksek doğruluklarda sonuçlar verdiğini göstermiştir.

OpenSMILE ile elde edilen öznitelikler Kendall's Tau öznitelik seçim modeli kullanılarak en iyi öznitelikler seçilmiştir. Bulgularda paylaşıldığı üzere doğruluklarda nispeten iyileştirme sağlamasına rağmen kayda değer bir artış göstermemiştir.

Sonuçlar ses patolojisinin tanısında tek bir ünlü harf üzerinden yapılan sınıflandırmaların tüm harflerin birleştirilerek yapılmasına göre daha yüksek doğruluk verdiği gözlemlenmiştir. Aynı zamanda cinsiyetleri ayrı olarak sınıflandırmanın birleştirerek sınıflandırmaya göre daha yüksek doğruluk verdiği görülmüştür.

Literatürde yapılan çalışmalarda u sesinin girdi olarak uygulandığı algoritmaların daha yüksek doğruluk oranına sahip olduğunu gösteren çalışmalar mevcuttur. Bu çalışmada da bunu destekler nitelikte sonuçlar bulunmuştur. Ayrıca XGB algoritmasının bazı gruplarda ses sınıflandırma uygulamalarında sıklıkla kullanılan SVM den daha yüksek doğruluk değerine ulaştığı gözlemlenmiştir. Öznitelik seçimi yapılarak algoritmaların performansı artırılmış ve daha iyi sonuçlar elde edilmiştir.

4.2. Sonuç ve Öneriler

Patolojik ses bozukluklarının erken tespiti, tedavi süreci ve tedavinin başarısı için oldukça önemlidir. Bu nedenle, patolojik seslerin tespiti için bir karar destek sisteminin geliştirilmesi büyük önem taşır. Son yıllarda, araştırmacılar tarafından patolojik seslerin belirlenmesi için makine öğrenme yöntemleri ilgi çekici bir konu haline gelmiştir.

Bu tez çalışmasında, patolojik seslerin belirlenmesi amacıyla makine öğrenmesi yöntemleri kullanılmıştır. Çalışma kapsamında patolojik ve sağlıklı ses kayıtlarının bulunduğu bir veri tabanından disfonik ses kayıtları seçilmiştir. Ses kayıtları öncelikle sayısal verilere dönüştürülmüş, öznitelik çıkarımı ve seçimi yapılarak sınıflandırma algoritmalarına girdi olarak verilmiştir.

Öncelikle SVM, RF gibi yaygın olarak kullanılan sınıflandırma algoritmaları seçilmiş daha sonra literatürde az rastladığımız NB ve XGB algoritmaları çalıştırılmıştır.

XGB algoritmasının SVM'den daha iyi sonuçlar vermesi daha sonra yapılacak çalışmalarda farklı veri tabanları ile genişletilerek umut verici sonuçlar elde edileceğini göstermiştir.

KAYNAKÇA

1. Şişman, B., 2014. Sesini Kaybetmiş Hastaların, Eski Ses Kayıtlarından Özgün Ses Üretme Algoritmasının (SÜA) Geliştirilmesi. İstanbul Üniversitesi, Doktora Tezi, İstanbul, 128 s.
2. Neighbors, C., Song, SA., 2022. Dysphonia. (Web sayfası: <https://www.ncbi.nlm.nih.gov/books/NBK565881/>), (Erişim tarihi: Haziran 2023).
3. Teixeira, P., Fernandes, O., Alves, N., 2017. Vocal acoustic analysis- classification of dysphonic voices with artificial neural networks. **Procedia Computer Science**, **121**: 19-26.
4. Şakar, B., 2014. Çoklu Ses Kayıtları İçeren Bir Parkinson Konuşma Veri Kümesinin Toplanması ve Analizi. İstanbul Üniversitesi, Doktora Tezi, İstanbul, 127 s.
5. Moe, M. C., Westerlund, U., Varghese, M., Berg-Johnsen, J., Svensson, M., & Langmoen, I. A., 2005. Development of neural networks from single stem cells harvested from the adult human brain. **Neurosurgery**, **56(6)**: 1182–1190.
6. Gönül, Y., Ulu, Ş., Bucak, A., Bilir, A., 2015. Yapay sinir ağları ve klinik araştırmalarda kullanımı, **Selçuk Üniversitesi Genel Tıp Dergisi**, **25**: 104-111.
7. Lee, J. Y., 2021. Experimental evaluation of deep learning methods for an intelligent pathological voice detection system using the saarbruecken voice database. **Applied Sciences**, **11(15)**: 7149.
8. Uysal, S., 2014. Ses Analizi İle Hastalık Teşhisi. Yıldız Teknik Üniversitesi, Yüksek Lisans Tezi, İstanbul, 98 s.
9. Shetty, S., Hegde, S., Dodderi, T., 2018. Classification of healthy and pathological voices using mfcc and ann. **IEEE International Conference on Advances in Electronics, Computers and Communications (ICAIECC)** : 1-5.
10. Dankovičová, Z., Sovák, D., Drotár, P., & Vokorokos, L., 2018. Machine learning approach to dysphonia detection. **Applied Sciences**, **8(10)**: 1927.
11. Huckvale, M., Buciuleac, C., 2021. Automated detection of voice disorder in the Saarbrücken voice database: effects of pathology subset and audio materials. **In**

Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH (6) : 4850-4854.

12. Koca Kayıkçı, E., 2015. Türk Sanat Müziği Korosu Ses Sanatçılarındaki Elektrolottografi ile Ses Analizi Değerlendirilmesi. Turgut Özal Üniversitesi, Yüksek Lisans Tezi, Malatya, 98 s.
13. Boone, D. R., MacFarlane, S. C., Von Berg, S. L., Zraick, R. I., 2009. The Voice And Voice Therapy, Allyn and Bacon, 360 pp.
14. Duffy, J. R., 2012. Motor Speech Disorders-E-Book: Substrates, Differential Diagnosis, And Management. Elsevier Health Sciences, 512 pp.
15. Cevanşir, B., Gürel, G., 1982. Foniatri Sesin Oluşumu, Bozuklukları ve Korunmasında Temel İlkeler, İstanbul Üniversitesi Tıp Fakültesi Yayınları. 112 s.
16. Kılıç, M., 2002. Larenksin fonksiyonel anatomisi ve ses fizyolojisi, **Türkiye Klinikleri KBB Dergisi (Ses Bozuklukları Özel Sayısı), 2: 1-8**
17. Şanlı, A., 2023. Gırtlak anatomisi. (Web sayfası: https://www.drarifsanli.com/tr/girtlak-anatomisi_a.html), (Erişim tarihi: Haziran 2023).
18. Bigat, İ., 2022. Patolojik Seslerin Tanısı İçin Derin Öğrenme Tabanlı Tıbbi Karar Destek Sisteminin Geliştirilmesi. Tobb Ekonomi Ve Teknoloji Üniversitesi, Yüksek Lisans Tezi, Ankara, 108 s.
19. Koç, C., 2004. KBB Hastalıkları ve Baş Boyun Cerrahisi, Güneş Yayıncılık, Ankara, 1504 s.
20. Karamürsel A., Dursun G., 2003. Organik ve fonksiyonel disfonilerde tedavi sonrası ses kalitesindeki değişikliklerin akustik olarak ölçümü, **KBB ve BBC Dergisi, 11(3): 99–106.**
21. Çağlayan, V., 2023. Akut larenjit. (Web sayfası: <https://www.tkbbv.org.tr/menu/48/akut-larenjit-op-dr-volkan-caglayan>), (Erişim tarihi: Haziran 2023).
22. Blitzler, A., Brin, M. F., 1991. Laryngeal dystonia: a series with botulinum toxin therapy. **Annals of Otology, Rhinology & Laryngology, 100(2): 85-89.**
23. Schweinfurth, J. M., Billante, M., Courey, M. S., 2002. Risk factors and demographics in patients with spasmodic dysphonia. **The Laryngoscope, 112(2) : 220-223**

24. Başerler, N., Ertaş, B., 2005. Disfoni nedenleri. **Klinik Gelişim Dergisi**, **18(1)** : 22-26.
25. Erdinç, M., 2009. Vokal Kord Poliplerinde Cerrahi Sonrası Ses Kalitesinin Araştırılması. Eskişehir Osmangazi Üniversitesi, Tıpta Uzmanlık Tezi, Eskişehir, 68 s.
26. Bless, D. M., Hirano, M., Feder, R. J., 1987. Videostroboscopic evaluation of the larynx. **Ear, nose, & throat journal**, **66(7)**: 289-296.
27. Bisong, E., 2019. Google colab, building machine learning and deep learning models on google cloud platform. **Apress, Berkeley, CA** : 59-64.
28. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Duchesnay, É., 2011. Scikit-learn: Machine learning in python. **the Journal of machine Learning research**, **12**: 2825-2830.
29. Van Rossum, G., Drake, F. L., 1995. Python reference manual. **Amsterdam: Centrum voor Wiskunde en Informatica**, **111**: 1-52.
30. MATLAB & Simulink, 2023, Feature extraction. (Web sayfası: <https://www.mathworks.com/discovery/feature-extraction.html>), (Erişim tarihi: Haziran 2023).
31. Eyben, F., Wöllmer, M., Schuller, B., 2010. Opensmile: the munich versatile and fast open-source audio feature extractor. **In Proceedings of the 18th ACM international conference on Multimedia**: 1459-1462.
32. Eskidere, Ö., Ertaş, F., 2009. Mel frekansı keştrüm katsayılarındaki değışimlerin konuşmacı tanımayaya etkisi. **Uludağ Üniversitesi Mühendislik Fakültesi Dergisi**, **14(2)**: 93-110.
33. Van Hulse, J., Khoshgoftaar, T. M., Napolitano, A., Wald, R., 2012. Threshold-based feature selection techniques for high-dimensional bioinformatics data. **Network modeling analysis in health informatics and bioinformatics**, **1**: 47-61.
34. Bircan, H., 2004. Lojistik regresyon analizi: tıp verileri üzerine bir uygulama. **Kocaeli Üniversitesi Sosyal Bilimler Dergisi**, **8**: 185-208.
35. Choi, S., Jiang, Z., 2010. Cardiac sound murmurs classification with autoregressive spectral analysis and multi-support vector machine technique. **Computers in Biology and Medicine** **40(2010)**: 8-20.
36. Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., Zhou, T., 2015. Xgboost: extreme gradient boosting. **R package version 0.4-2**, **1(4)**: 1-4.

37. Das, R., 2010. A comparison of multiple classification methods for diagnosis of parkinson disease. **Expert Systems with Applications**, **37(2)**: 1568- 1572
38. Kantardzic, M., 2011. Data Mining: Concepts, Models, Methods, And Algorithms, John Wiley & Sons, 672 pp.
39. Biau, G., & Scornet, E., 2016. A random forest guided tour. **Test**, **25**: 197-227.
40. Leung, K. M., 2007. Naive bayesian classifier. **Polytechnic University Department of Computer Science/Finance and Risk Engineering**, **2007**: 123-156.



ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı: Pervin Gülşen

Uyruğu: Türkiye (T.C)

EĞİTİM

Derece	Kurum	Mezuniyet Tarihi
Lisans	Erciyes Üniversitesi Elk.-Elkronik Mühendisliği	2018
Lisans	Erciyes Üniversitesi Biyomedikal Mühendisliği	2017
Lise	Mustafa Eminoglu Anadolu Lisesi	2014

İŞ DENEYİMLERİ

Yıl	Kurum	Görev
2017-2018	ERKAM	2017

YABANCI DİL

İngilizce

YAYINLAR

1. Kiliç, P., Doğan, R. S., Altinkaynak, M., & Güven, A. (2017, October). Assessment of attention before and after exam with reaction time. In *2017 Medical Technologies National Congress (TIPTEKNO)* (pp. 1-3). IEEE.