

T.C.
ÇANAKKALE ONSEKİZ MART ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

MAKİNE ÖĞRENME ALGORİTMALARI KULLANARAK SES VERİLERİNDE HASTALIK TESPİTİ

YÜKSEK LİSANS TEZİ

DUYGU ÇOKAY

Tez Danışmanı
DOÇ. DR. ENGİN ŞAHİN

ÇANAKKALE – 2024

T.C.
ÇANAKKALE ONSEKİZ MART ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

MAKİNE ÖĞRENME ALGORİTMALARI KULLANARAK SES VERİLERİNDE HASTALIK TESPİTİ

YÜKSEK LİSANS TEZİ

DUYGU ÇOKAY

Tez Danışmanı
DOÇ. DR. ENGİN ŞAHİN

Bu çalışma, Çanakkale Onsekiz Mart Üniversitesi Bilimsel Araştırma Projeleri Koordinasyon Birimi tarafından desteklenmiştir.
Proje No: FYL-2023-4610

ÇANAKKALE – 2024

ETİK BEYAN

Çanakkale Onsekiz Mart Üniversitesi Lisansüstü Eğitim Enstitüsü Tez Yazım Kuralları'na uygun olarak hazırladığım bu tez çalışmasında; tez içinde sunduğum verileri, bilgileri ve dokümanları akademik ve etik kurallar çerçevesinde elde ettiğimi, tüm bilgi, belge, değerlendirme ve sonuçları bilimsel etik ve ahlak kurallarına uygun olarak sunduğumu, tez çalışmasında yararlandığım eserlerin tümüne uygun atıfta bulunarak kaynak gösterdiğimi, kullanılan verilerde herhangi bir değişiklik yapmadığımı, bu tezde sunduğum çalışmanın özgün olduğunu, bildirir, aksi bir durumda aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi taahhüt ve beyan ederim.

(İmza)
DUYGU ÇOKAY
26/06/2024

TEŞEKKÜR

Bu tezin gerçekleştirilmesinde, çalışmam boyunca benden bir an olsun yardımlarını esirgemeyen saygı değer danışman hocam Doç. Dr. Engin ŞAHİN'e, çalışma süresince tüm zorlukları benimle göğüsleyen eşim Barış İlke ÇOKAY ve kızım Derin Ada ÇOKAY'a, hayatımın her evresinde bana destek olan anne ve babama sonsuz teşekkürlerimi sunarım.

Duygu ÇOKAY

Çanakkale, Haziran 2024

ÖZET

MAKİNE ÖĞRENME ALGORİTMALARI KULLANARAK SES VERİLERİNDE HASTALIK TESPİTİ

Duygu ÇOKAY
Çanakkale Onsekiz Mart Üniversitesi
Lisansüstü Eğitim Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı Yüksek Lisans Tezi
Danışman: Doç. Dr. Engin ŞAHİN
26/06/2024, 59

Hastalıkların ileri evrelerde tespit edilmesi iyileşme oranlarını düşürür, tedavi sürecini zorlaştırır, iyileşme sürecindeki maliyeti artırır. Bu nedenle hastalıkların erken evrelerde tespit edilmesi önemlidir. Günümüzde makine öğrenmesi yöntemleri; sağlık alanında, tıbbi veri analizi ve hastalık tespiti amacıyla yaygın şekilde kullanılmaktadır. Ses kısıklığı, toplumda yaygın görülen şikayetlerden biridir. Ses kısıklığına neden olan ses teli polipi, gırtlak kanseri, akut larenjit, ses teli felci gibi çeşitli hastalıklar mevcuttur. Bu çalışmada, ses kısıklığına neden olan farklı hastalıkların, ses verileri kullanılarak, patolojik ve endoskopik muayeneden önce temel sınıflandırılmasının yapılması amaçlanmıştır. Çalışmada Saarland Üniversitesi Fonetik Enstitüsü tarafından oluşturulan Saarbruecken veri tabanından elde edilen ses verileri kullanılmıştır. Farklı hastalıklara sahip olan ve sağlıklı kişilerden oluşan 652 kişiye ait 5865 ses verisi üzerinde K En Yakın Komşu, Naive Bayes, Karar Ağacı, Destek Vektör Makinesi, Rassal Orman olarak beş farklı makine öğrenme algoritmasıyla sınıflama yapılmış, elde edilen sonuçlar karşılaştırılmıştır. Elde edilen sonuçlara göre hastalıklı-sağlıklı ses ayrımında en yüksek doğruluk oranı 0,79 ile Destek Vektör Makinesi algoritmasından elde edilmiştir. Hastalıkları patolojilerine göre sınıflamada en iyi sonucu K En Yakın Komşu algoritması vermiştir. Sınıflamada en başarısız algoritma Naive Bayes algoritmasıdır.

Anahtar Kelimeler: Makine Öğrenmesi, Ses İşleme, Sınıflandırma Algoritmaları, Hastalık Tespiti, Polip, Larenjit

ABSTRACT

DISEASE DETECTION IN VOICE DATA BY USING MACHINE LEARNING ALGORITHMS

Duygu ÇOKAY
Çanakkale Onsekiz Mart University
School of Graduate Studies
Master of Science Thesis in Computer Engineering
Advisor: Assoc. Prof. Dr. Engin ŞAHİN
26/06/2024, 59

Detecting diseases at advanced stages reduces recovery rates, complicates the treatment process, and increases the cost of recovery. Therefore, early detection of diseases is crucial. Today, machine learning methods are widely used in the health field for medical data analysis and disease detection. Hoarseness is a common complaint in the community. There are various diseases that cause hoarseness, such as vocal cord polyps, laryngeal cancer, acute laryngitis, and vocal cord paralysis. This study aims to perform basic classification of different diseases causing hoarseness using voice data, prior to pathological and endoscopic examination. The voice data used in this study were obtained from the Saarbruecken database, created by the Institute of Phonetics at Saarland University. Classification was performed on 5865 voice recordings from 652 individuals, including both healthy individuals and those with various diseases, using five different machine learning algorithms: K-Nearest Neighbors, Naive Bayes, Decision Tree, Support Vector Machine, and Random Forest. The obtained results were then compared. According to the results, the highest accuracy rate in distinguishing between healthy and diseased voices was achieved with the Support Vector Machine algorithm, at 0.79. The best results for classifying diseases according to their pathologies were obtained with the K-Nearest Neighbors algorithm. The least successful algorithm in classification was the Naive Bayes algorithm.

Keywords: Machine Learning, Speech Processing, Classification Algorithms, Disease Detection, Polyp, Laryngitis

İÇİNDEKİLER

Sayfa No

JÜRİ ONAY SAYFASI……..……... i
ETİK BEYAN…… ii
TEŞEKKÜR…….. iii
ÖZET ……..…….…….. iv
ABSTRACT …….. v
İÇİNDEKİLER ……..…… vi
SİMGELER ve KISALTMALAR…….....……. viii
TABLOLAR DİZİNİ…….…. ix
ŞEKİLLER DİZİNİ……..……...…….…….….. x

BİRİNCİ BÖLÜM
GİRİŞ 1

İKİNCİ BÖLÜM
KURAMSAL ÇERÇEVE/ÖNCEKİ ÇALIŞMALAR 3

2.1. Ses Kısıklığına Neden Olan Hastalıklar ……..... 3
2.1.1. Reinke Ödemi……. 4
2.1.2 Akut Larenjit……... 4
2.1.3 Ses Teli Polipi……. 5
2.1.4 Gırtlak Kanseri…… 6
2.1.5 Ses Teli Felci……... 6
2.2. Makine Öğrenme Algoritmaları…….. 7
2.2.1. K En Yakın Komşu……. 9
2.2.2. Naive Bayes…… 10
2.2.3 Karar Ağaçları……. 12
2.2.4 Destek Vektör Makinesi…….. 13
2.2.5 Rassal Orman…….. 15
2.3. Veri Seti…….. 16
2.4. Literatürdeki Çalışmalar……. 18

2.5. Değerlendirme Kriterleri…… 22

ÜÇÜNCÜ BÖLÜM
ARAŞTIRMA YÖNTEMİ/MATERYAL YÖNTEM 24

3.1. Kullanılan Yazılım ve Donanım……. 24
3.2 Verilerin Toplanması…….. 24
3.3. Veri Ön İşleme…… 26
3.3.1 Öznitelik Çıkarımı…… 27
Ön Vurgulama (Pre-emphasis)……. 28
Sinyalinin Kesilmesi (Framing)……... 28
Pencereleme (Windowing)……... 29
Hızlı Fourier Dönüşümü (Fast FourierTransform)……... 30
Mel-Frekans Filtre Bankası Uygulaması (Mel Filter Bank)……. 31
Logaritmik Dönüşüm……... 31
Ayrık Kosinüs Dönüşümü…… 32
3.4. Veri Setinin Bölünmesi ve Modelin Eğitilmesi……... 33
3.5. Modellerin Değerlendirilmesi……. 35

DÖRDÜNCÜ BÖLÜM
ARAŞTIRMA BULGULARI 37

4.1. K En Yakın Komşu Algoritmasından Elde Edilen Sonuçlar……... 37
4.2. Naive Bayes Algoritmasından Elde Edilen Sonuçlar…….. 40
4.3 Karar Ağaçları Algoritmasından Elde Edilen Sonuçlar……... 43
4.4 Destek Vektör Makineleri Algoritmasından Elde Edilen Sonuçlar……. 45
4.5. Rassal Orman Algoritmasından Elde Edilen Sonuçlar…… 48

BEŞİNCİ BÖLÜM
SONUÇ ve ÖNERİLER 51

KAYNAKÇA…….. 54

SİMGELER VE KISALTMALAR

k-NN	En yakın komşu algoritması
NB	Naive bayes algoritması
SVM	Destek vektör makineleri
RF	Random forest
DT	Decision tree
MFCC	Mel-frekanslı cepstral katsayıları
FEMH	Far eastern memorial hastanesi
GDA	Gradyan arttırma
CNN	Evrişimli sinir ağları
OSELM	Çevrimiçi sıralı aşırı öğrenme makinesi
YSA	Yapay sinir ağı
DNNs	Derin sinir ağları
LR	Lojistik regresyon
ANN	Yapay sinir ağı
LSTM	Uzun kısa süreli bellek
TP	True pozitif
FP	False pozitif
TN	True negatif
FN	False negatif
GNB	Gaussian naive bayes
MNB	Multinominal naive bayes
BNB	Bernoulli naive bayes
MRI	Manyetik rezonans görüntüleme
CT	Bilgisayarlı tomografi
KBB	Kulak burun boğaz
EMG	Elektromiyografi
BT	Bilgisayarlı tomografi
NGB	Gaussian naive bayes
MNB	Multinominal naive bayes
BNB	Bernoulli naive bayes
FIR	Sonlu dürtü yanıtı

TABLOLAR DİZİNİ

Tablo No	Tablo Adı	Sayfa No
Tablo 1	Tez kapsamında kullanılan veri seti	25
Tablo 2	k-NN algoritmasının sağlıklı-hasta verilerini ayırmasında elde edilen sonuçlar	37
Tablo 3	k-NN algoritmasının patolojilere göre sınıflamada elde ettiği değerler	39
Tablo 4	NB algoritmasının sağlıklı-hasta verilerini ayırmasında elde edilen sonuçlar	40
Tablo 5	NB algoritmasının patolojilere göre sınıflamada elde ettiği değerler	42
Tablo 6	DT algoritmasının sağlıklı-hasta verilerini ayırmasında elde edilen sonuçlar	43
Tablo 7	DT algoritmasının patolojilere göre sınıflamada elde ettiği değerler	44
Tablo 8	SVM algoritmasının sağlıklı-hasta verilerini ayırmasında elde edilen sonuçlar	45
Tablo 9	SVM algoritmasının patolojilere göre sınıflamada elde ettiği değerler	47
Tablo 10	RF algoritmasının sağlıklı-hasta verilerini ayırmasında elde edilen sonuçlar	48
Tablo 11	RF algoritmasının patolojilere göre sınıflamada elde ettiği değerler	49
Tablo 12	Algoritmaların sağlıklı-hasta sınıflamasında elde ettiği başarı ölçütleri	51
Tablo 13	Algoritmaların hastalıkları sınıflamadaki doğruluk oranları	52

ŞEKİLLER DİZİNİ

Şekil No	Şekil Adı	Sayfa No
Şekil 1	Ses teli	3
Şekil 2	Reinke ödemi	4
Şekil 3	Akut larenjit	5
Şekil 4	Makine öğrenmesi adımları	8
Şekil 5	k-NN algoritması	10
Şekil 6	Karar ağacı	13
Şekil 7	SVM algoritması gösterimi	14
Şekil 8	Kernel fonksiyonu	14
Şekil 9	Rassal orman	16
Şekil 10	Saarbruecken ses veritabanı	17
Şekil 11	Saarbruecken ses veritabanı	18
Şekil 12	Veri tabanındaki seslerin seçimi	25
Şekil 13	Veri tabanındaki hastalıkların seçimi	26
Şekil 14	Verinin standartlaştırılması ve eğitim-test verilerinin seçilmesi	27
Şekil 15	MFCC gösteriminin Python kodu	32
Şekil 16	MFCC spektrumu	32
Şekil 17	Ses verilerinin hasta-sağlıklı olarak ayrılması	34
Şekil 18	Ses verilerinin sınıflarına göre ayrılması	35
Şekil 19	Karmaşıklık matrisi ile doğruluk hesaplanması	36
Şekil 20	k-NN algoritmasının sağlıklı-hasta verilerini ayırmasında elde edilen karmaşıklık matrisi	38

Şekil 21	k-NN algoritmasının patolojilere göre sınıflamada elde ettiği karmaşıklık matrisi	39
Şekil 22	NB algoritmasının uygulanması	40
Şekil 23	NB algoritmasının sağlıklı-hasta verilerini ayırmasında elde edilen karmaşıklık matrisi	41
Şekil 24	NB algoritmasının patolojilere göre sınıflamada elde ettiği karmaşıklık matrisi	42
Şekil 25	DT algoritmasının sağlıklı-hasta verilerini ayırmasında elde edilen karmaşıklık matrisi	43
Şekil 26	DT algoritmasının patolojilere göre sınıflamada elde ettiği karmaşıklık matrisi	45
Şekil 27	SVM algoritmasının sağlıklı-hasta verilerini ayırmasında elde edilen karmaşıklık matrisi	46
Şekil 28	SVM algoritmasının patolojilere göre sınıflamada elde ettiği karmaşıklık matrisi	47
Şekil 29	RF algoritmasının sağlıklı-hasta verilerini ayırmasında elde edilen karmaşıklık matrisi	48
Şekil 30	RF algoritmasının patolojilere göre sınıflamada elde ettiği karmaşıklık matrisi	50

BİRİNCİ BÖLÜM
GİRİŞ

Hastalıkların erken evrelerde tespit edilmesi, tedavi süreçleri açısından oldukça önemlidir. Erken teşhis hastalıkların ilerlemesini engelleyebilir, daha etkili tedavilerin uygulanmasını sağlayabilir. Erken teşhisin önemi özellikle kanser gibi hastalıkların tedavisinde belirgindir (Sağlık Bakanlığı, 2016). Kanser gibi hastalıkların erken aşamalarda tespit edilmesi tedavi başarısını arttırabilir. Bu nedenle Dünya Sağlık Örgütü hastalıkların erken evrede tespitine önem vermektedir.

Günümüzde makine öğrenmesi, yapay zeka, derin öğrenme yöntemleri sağlık alanında tıbbi veriler içerisinde hastalıkları yakalamak ve analiz etmek için yaygın şekilde kullanılmaktadır. Hasta verileri karmaşık ve büyük boyutlu verilerdir. Çoklu veriler üzerinde sınıflama yaparken makine öğrenme algoritmaları hızlı ve yüksek doğruluk oranları elde etmemizi sağlar (Lundberg vd., 2018).

Ses kısıklığı, seste perde daralması toplumumuzda gözardı edilen şikayerlerdendir. Ses kısıklığı uzman hekim tarafından değerlendirilmesi gereken önemli bir belirtidir. Ses kısıklığına neden olan hastalıklar ses teli polipi, gırtlak kanseri, akut larenjit, ses teli felci gibi çok çeşitli olabilir. Bu hastalıkların ortak noktası ses kısıklığı olmasına rağmen nedenleri, tedavi süreçleri, riskleri farklılık gösterir.

Araştırmalar çalışan nüfusun yaklaşık %10'unun ve genel nüfusun %3-%9'unun hayatının herhangi bir döneminde ses kısıklığı ile ilgili bir sorunla karşılaştığını göstermektedir (Verdolini ve Ramig, 2001).

Ses teli hastalıklarının tanısı uzman doktor tarafından fiziki muayene, endoskopik muayene, akustik analiz gibi farklı muayeneler ayrı ya da birlikte yapılarak konulur. Sadece ses kısıklığı üzerinden tanı koymak hekim için zordur ve yanlış tanıya neden olabilir. İnsan kulağının algılayabileceği ve ayırt edebileceği ses özellikleri sınırlıdır. Bu nedenle ses kısıklığı şikayeti ile gelen hasta için hasta hikayesi ve fiziki muayene dışında farklı testler, muayeneler gerekebilir. Endoskopik muayene için gerekli tıbbi donanımın maaliyeti yüksektir ve kullanılması uzmanlık gerektirir. Birinci basamak sağlık hizmetleri veren kuruluşlarda (Aile hekimleri vb.) tanısının konması zordur. Ses verisini tanıyıp olası hastalıkları analiz edebilen ve tanı koymada destek sağlayan algoritmalar hekimlerin tanı koymasını kolaylaştırabilir, hatalı tanı oranının azalmasında yardımcı olabilir.

Patolojik ses tespitinin akustik analizi hekimin algısal analizini tamamlamak için kullanılabilir. Makine öğrenmesi, derin öğrenme, bilgisayarlı görü ve konuşma tanıma tanı koyma alanında kullanılan popüler çalışmalardır. Erken teşhis, düşük maliyet, yüksek doğruluk oranları bu algoritmaların sağlık alanında kullanılmasının en önemli nedenleridir.

Tıbbi tanılamada, biyokimyasal ve klinik testlerin yorumlanmasında, radyolojik görüntülerin yorumlanmasında, endoskopik incelemelerde, patolojik incelemelerde, yoğun bakım süreçlerinin takibinde, salgın hastalıkların yönetiminde yapay zeka, makine öğrenme algoritmaları yaygın olarak çalışılmaktadır.

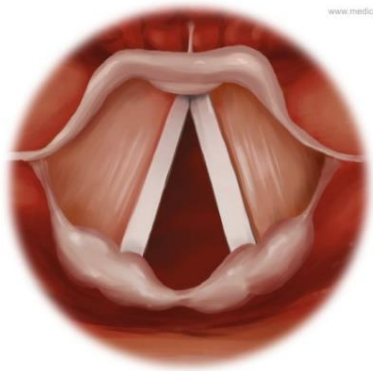
Bu çalışmada ses kısıklığına neden olan reinke ödemi, akut larenjit, kanser ve kanser öncesi lekelenme, polip, ses teli felci hastalıklarına sahip hastaların ses verileri makine öğrenme algoritmalarıyla sınıflandırılmıştır. Sınıflandırmada K En Yakın Komşu (k-NN) (Cover ve Hart, 1967), Naive Bayes (NB) (Flach ve Lachiche, 2004), Karar Ağacı (DT) (Breiman, 2017), Destek Vektör Makinesi (SVM) (Cortes ve Vapnik, 1995) ve Rassal Orman (RF) (Breiman, 2001) algoritmaları kullanılmıştır. Algoritmaların başarı sonuçları karşılaştırılmıştır.

Çalışmada Saarland Üniversitesi Fonetik Enstitüsü tarafından oluşturulan ve http://www.stimmdatenbank.coli.uni-saarland.de/ adresinde çevrimiçi paylaşılan ses verileri kullanılmıştır (William ve Manfred, 2007).

İKİNCİ BÖLÜM
KURAMSAL ÇERÇEVE/ÖNCEKİ ÇALIŞMALAR

2.1. Ses Kısıklığına Neden Olan Hastalıklar

Ses telleri larenks adı verilen boğazın üst kısmında bulunan iki küçük yapılı kastır. Birbirlerine temas ederek titreşir ve ses oluşturur. Ses tellerinin titreşimi ses dalgası olarak yayılır ve sesin temel frekansını belirler. Oluşan ses dalgaları ağız boşluğundan geçerken şekil değiştirir. Bu durum ses tonumuzu, sesimizin yüksekliğini ve ses rengimizi belirler (Titze, 1994).



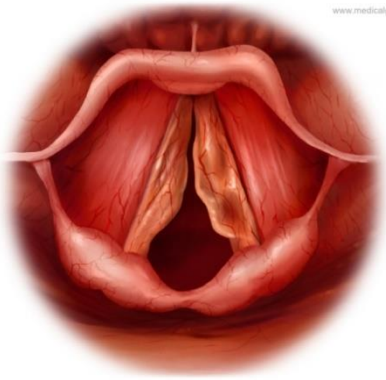
Şekil 1. Ses teli (Medical Graphics, 2023)

Çeşitli nedenlerle ses telleri hasar görebilir ve bu durum ses kısıklığına yol açar. Ses kısıklığı birçok farklı durumun belirtisi olabilir. Ses kısıklığına neden olan hastalıklar arasında nodüller, polipler, gırtlak enfeksiyonları, larenjit, gırtlak kanseri, reflü, sinir hasarları bulunabilir. Ses kısıklığı geçici bir durum olabilir, uygun tedavi ile kısa sürede iyileşebilir. Ancak kalıcı ses kısıklığı durumunda altta yatan nedenin tıbbi olarak değerlendirilmesi ve uygun tedavinin başlanması çok önemlidir. Tez çalışmasında ele alınan, ses kısıklığına neden olan ve sınıflandırılması amaçlanan hastalıklar şunlardır:

2.1.1. Reinke Ödemi

Reinke ödemi, ses tellerinin ortasındaki iç yüzeyi kaplayan ve titreşimleri oluşturan yumuşak doku tabakasında ortaya çıkan bir hastalıktır. Ses tellerinin şişmesine neden olur ve bu durum da ses kısıklığına yol açar (Jones ve Song, 2023).

Reinke ödeminin sigara kullanımı, reflü hastalığı, solunum yolu enfeksiyonları, sesi kötüye kullanma gibi çeşitli sebepleri olabilir. Reinke ödemi tedavisinde hastalığın nedenine göre sigarayı bırakma, ses istirahati, ses terapisi, cerrahi müdahale gibi farklı yöntemler izlenebilir (Deem ve Miller, 2000). Reinke ödemi tedavisinde bazı durumlarda cerrahi müdahale gerekebilir. Cerrahi işlem sonrasında ses terapisi uygulanmaktadır (Mathieson, 2013).

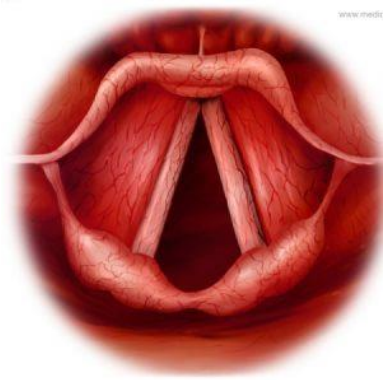


Şekil 2. Reinke ödemi (Medical Graphics, 2023)

2.1.2. Akut larenjit

Akut larenjit genellikle virüs nedenli ses tellerinin ve üst solunum yollarının iltihaplanmasıdır. Soğuk algınlığı veya viral enfeksiyonlarla ortaya çıkar. Soğuk algınlığı, grip gibi hastalıklarla birlikte ortaya çıktığı için kış mevsimlerinde sıklıkla görülür ve tekrar eder. Ses kısıklığı, ses kaybı, boğaz ağrısı, kuru öksürük, boğazda yanma gibi semptomları vardır. Akut larenjit genellikle semptomlar temel alınarak teşhis edilir. Fiziksel muayene esnasında doktor genellikle boğazı kontrol eder. Tanı koyma süreci diğer ses kısıklığı hastalıklarına göre nispeten kolaydır. Tedavisi genellikle semptomları hafifletmeye yöneliktir.

Sigara kullanımı, alkol ve hava kirliliği larenjitin kronikleşmesine neden olabilir. Larenjitin kronikleşmesi durumunda hastalar sıklıkla seslerini daha yüksek kullanmaya çalışır. Bu durum gırtlakta daha fazla tahrişe neden olur (Deem ve Miller, 2000).



Şekil 3: Akut Larenjit (Medical Graphics, 2023)

2.1.3. Ses Teli Polipi

Ses teli polipleri genellikle ses tellerinin temas ettiği yüzeylerde oluşan, dokular üzerinde gelişen küçük yumuşak kitlelerdir. Genellikle ses tellerinin aşırı kullanımı ve kötü kullanımı ile gelişir. Ses teli polipleri larinksin en yaygın iyi huylu lezyonlarıdır (Kleinsasser, 1982).

Sigara kullanmak, reflü hastalığı da ses teli poliplerinin nedeni olabilir. Ses kısıklığı, orta kulakta ağrı, seste perde daralması, ses değişmesi gibi belirtileri olabilir. Genellikle iyi huylu tümörler olarak bilinmesine rağmen kansere evrilebilir ya da boyutlarının büyümesi insan sağlığını olumsuz etkileyebilir (Martins vd., 2011). Bu nedenle erken evrede teşhis ve tedavisi önemlidir. Kulak burun boğaz (KBB) hekimleri tarafından gırtlak bölgesi endoskopisiyle tanısı konur. Birinci basamak sağlık hizmetleri veren kuruluşlarda (Aile hekimliği vb.) tanısının konması zordur. Endoskopik muayene için gerekli tıbbi malzeme maliyeti yüksektir ve kullanması uzmanlık gerektirir.

Ses teli poliplerinin tedavisi hastaya ve polipin boyutunda bağlı olarak değişiklik gösterir. Polipler küçük ve belirtiler azsa ses dinlenmesi ve yaşam tarzı değişikliği tedavi süreci için yeterli olabilir. Ancak belirtiler şiddetli ve uzun süreli ise polipler endoskopik cerrahi veya mikro cerrahi ile çıkarılabilir.

2.1.4. Gırtlak Kanseri

Gırtlak kanseri larenks denilen solunum yollarının üst bölgesinde oluşan bir tür kanserdir. Kanser larenks bölgesinin herhangi bir yerinde gelişebilir. Gırtlak kanseri akciğer kanserinden sonra solunum yollarındaki ikinci yaygın kanserdir (Cattaruzza vd., 1996). Erkeklerde daha yaygın görülen bir kanser türüdür. Gırtlak kanserinin en belirgin belirtileri ses kısıklığı veya ses değişiklikleri, kulak ve boğaz ağrıları, yutmada ve nefes almada zorlanmadır.

Sigara ve alkol tüketimi gırtlak kanserinin başlıca risk faktörleridir. Yapılan çalışmalar sigara tüketenlerin sigara tüketmeyenlere oranla 10-15 kat risk altında olduğunu göstermiştir. Beslenme alışkanlıklarının da gırtlak kanseri riskini artırabileceği, kırmızı etin riski arttırırken meyve ve sebzelerle oluşturulmuş diyetlerin gırtlak kanseri riskini azaltabileceği belirtilmiştir (Steuer vd., 2017).

Tüm kanserlerde olduğu gibi gırtlak kanserinde de tedavi öykü alımı ve klinik muayene ile başlar. Tümör evresinin değerlendirilmesi için görüntü alınır. Bunun yanında hastadan tomografi, endoskopik muayenede alınan biyopsi gibi çeşitli veriler elde edilir. Hastalığın erken evrelerde tespitinde radyoterapi, küçük cerrahi müdahale, açık kısmi cerrahi gibi yöntemler uygulanabilir. Hastalığın erken evrelerde tespit edilmesi iyileşme şansını arttırır (Obid vd., 2019).

2.1.5. Ses Teli Felci

Ses teli felci sinir hasarı sonucu ses tellerinin normal hareketlerini kaybetmesi durumudur (Terris vd., 1992). Bu sinir hasarı ses tellerini kontrol eden farklı sinirlerden kaynaklanabilir. Ses telleri normalde nefes alırken açılır, konuşma halinde kapanır. Ancak ses teli felci durumunda ses tellerinden biri ya da ikisi de hareket edemez.

Ses teli felcinin en yaygın belirtisi ses kısıklığı veya ses kaybıdır. Ses teli felci sesin yoğunluğunu, kalitesini de olumsuz etkileyebilir. Ses teli felcinin yutmada ve nefes almada zorlanma, boğaz ağrısı ve boğazda hassasiyet gibi farklı belirtiler de gözlemlenebilir (Tucker, 1980).

Ses teli felci tanısı KBB uzmanları tarafından hastanın belirtileri, tıbbi geçmişi, fiziksel muayenesi değerlendirilerek konur. Tanı koyma sürecinde hastanın ses tellerinin hareketlerinin değerlendirilmesi için fonasyon testleri uygulanabilir. Ses telleri kontrol eden sinirlerin işlevlerini değerlendirebilmek için elektromiyografi (EMG) kullanılabilir. Doktor kimi zaman bilgisayarlı tomografi (BT), veya magnetik rezonans görüntüleme (MRI) gibi görüntüleme tekniklerine ihtiyaç duyabilir.

2.2. Makine Öğrenme Algoritmaları

Makine öğrenmesi insan zekasını taklit ederek tasarlanmış, çevredeki ortam verilerinden öğrenerek yeni örnekler için çıkarsama yapan hesaplamalı bir algoritma dalıdır. Makine öğrenme algoritmalarının genelleme yeteneği örüntü tanıma, bilgisayarla görü, hesaplamalı biyoloji, tıbbi tanılama gibi birçok alanda başarılı şekilde kullanılmaktadır (El Naqa ve Murphy, 2015).

Makine öğrenme algoritmaları denetimli öğrenme, denetimsiz öğrenme, yarı denetimli öğrenme olarak sınıflandırılabilir.

Denetimli öğrenme etiketlenmiş veri setleriyle çalışır. Denetimli öğrenmede giriş verileriyle doğru çıktıları içeren eğitim veri seti kullanılarak bir model oluşturulur. Bu model giriş verileri ve eğitim verileri arasındaki ilişkiyi öğrenir ve bu ilişkileri kullanarak yeni verilerde tahminler yapar (Cunningham vd., 2008).

Denetimsiz makine öğrenme algoritmaları verilerdeki ilişkileri tespit etmek için kullanılan bir tür öğrenme yaklaşımıdır. Etiketlenmiş verilere ihtiyaç duymadan veri kümesindeki ilişkileri tespit etmeye odaklanır. Önceden belirlenmiş bir çıktıya göre eğitilmezler, bunun yerine verilerin ilişkilerine odaklanarak kendi sınıflamalarını oluştururlar. Denetimsiz. öğrenme algoritmaları veri setlerindeki ilişkileri tanımlamak için farklı teknikler kullanır. Bu teknikler arasında kümeleme, boyut azaltma vb. gibi yöntemler bulunur (Ghahramani, 2003).

Makine öğrenmesi süreci veri toplama ve hazırlama, veri ön işleme, model seçimi, seçilen modelin eğitilmesi, modelin değerlendirilmesi ve iyileştirme adımlarından oluşur. İlk adım olarak işlenecek verinin elde edilmesi ve gerekirse verilerin temizlenmesi adımıdır. Toplanan veriler eksik ya da gürültülü olabilir. Bu nedenle veriler ön işleme adımından

geçirilir. Ön işleme adımında eksik veriler doldurulur ve verilerin normalizasyonu yapılır. Model seçimi adımında veri setine ve çözülmesi amaçlanan probleme bağlı olarak eğitilecek model seçilir. Sınıflama, regresyon, kümeleme gibi farklı problemler için farklı modeller kullanılabilir. Model eğitimi sürecinde seçilen model eğitim verileri kullanılarak eğitilir. Eğitilen model test verileri kullanılarak değerlendirilir. Son adım olarak modelin performansını arttırmak için gerektiğinde modelin yapılandırılması, daha fazla verinin toplanması, veri ön işleme adımlarının iyileştirilmesi gerekebilir.



Şekil 4. Makine öğrenmesi adımları

Günümüzde tıbbi alandaki veriler hastanın yaşı, cinsiyeti, yaşadığı bölge özellikleri, aile genetik bilgileri, hastanın genel sağlık durumu, geçirdiği hastalıklar, ilaç geçmişi, kan tahlilleri, muayene bulguları, laboratuvar test sonuçları, şikayetleri, ölçüm bilgileri(nabız, vücut sıcaklığı vb.), görüntüleme sonuçları (röntgen, MRI, CT taramaları) gibi çok geniş bir veri kümesini oluşturmaktadır. Bu geniş veri kümesi üzerinde değerlendirme yapmak, çoklu verileri kullanarak çıkarsama yapmak hekim açısından zor bir süreçtir. Makine öğrenme algoritmaları bu büyük veri kümelerinin hızlı şekilde analiz edilmesinde, hastalıkların ve risk faktörlerinin belirlenmesinde başarılı sonuçlar vermektedir.

Makine öğrenme algoritmaları sınıflama ve regresyon problemlerinde kullanılmaktadır. Regresyon problemlerinde bir veya birden fazla bağımsız değişkenin etki ettiği bir bağımlı değişkeni tahmin etmek amaçlanır. Bir hastanın yaşı, kan değerleri, cinsiyet ve laboratuvar sonuçlarına göre hastanede kalma süresini tahmin etmek regresyon problemlerine örnektir. Sınıflama problemlerinde ise amaç eldeki verileri belirli kategorilere göre sınıflara atamaktır. Hasta tarafından belirtilen semptomların bir hastalıkla ilişkilendirilmesi sınıflama problemlerine örnektir.

Yapılan tez çalışmasında makine öğrenme algoritmaları ile sınıflama yapılması amaçlanmıştır. Tez çalışmasında kullanılan makine öğrenme algoritmaları aşağıda açıklanmıştır.

2.2.1. K En Yakın Komşu

Basit ve kolay uygulanan bir denetimli öğrenme algoritmasıdır. Sınıflandırma ve regresyon problemlerinde kullanılır. Algoritma yeni bir veri noktasını sınıflandırmak veya tahmin etmek için çevresindeki en yakın komşuları kullanır. Sınıflandırılması planlanan veri k sayısı kadar komşularına bakılarak çoğunluğun üyesi olduğu sınıfa dahil edilir (Cover ve Hart, 1967).

Komşularının yakınlık hesaplaması Euclid mesafesi, Manhattan mesafesi, Minkowski mesafesi gibi farklı mesafe ölçütleriyle elde edilebilir.

Euclid mesafesi hesaplamasında iki komşuluk arasındaki doğrudan mesafe ölçülür. Bu ölçüm iki nokta arasındaki en kısa mesafeyi temsil eder. Euclid mesafesi ölçümü şu şekilde hesaplanır:

$$d(x,y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \tag{2.1}$$

Manhattan mesafesi iki nokta arasındaki doğrusal mesafeyi ölçer. Formülü şu şekildedir:

$$d(x,y) = \sum_{i=1}^{n}|x_i - y_i| \tag{2.2}$$

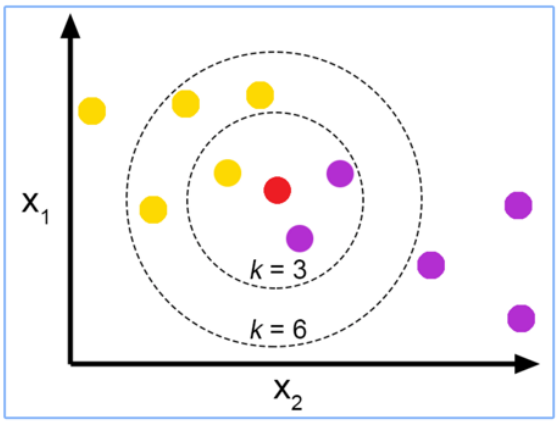
Minkowski mesafesi iki nokta arasındaki farkların bir fonksiyonudur. Formülü şu şekildedir:

$$d(x,y) = \left(\sum_{i=1}^{n}|x_i - y_i|^p\right)^{\frac{1}{p}} \tag{2.3}$$

Formülde x ve y noktaları sınıflamada kullanılan noktaların konumunu belirtir. P değeri ise Minkowski mesafesinin parametresidir ve farklı değerler alabilir (Rodrigues, 2018).

k-NN algoritmasında bir diğer önemli değer k sayısıdır. K sayısının farklı değerleri sınıflamayı değiştirebilir. K sayısının belirlenmesinde veri setinin boyutu, veri noktalarının dağılımı önemli rol oynar. K sayısının çift olması sınıfın belirlenememesine neden olabilir. Bu nedenle k sayısının tek sayı seçilmesi önemlidir (Dudoit vd., 2002).

k-NN algoritması basit, anlaşılır, uygulaması kolay bir algoritmadır. Çok sınıflı problemlerde başarılı sonuçlar verir. Ancak büyük veri setleri üzerinde yavaş çalışabilir (Cover ve Hart, 1967).



Şekil 5. k-NN algoritması (Towards Data Science, 2018)

2.2.2. Naive Bayes

NB algoritması bayes teorisine dayanan olasılık tabanlı bir sınıflama algoritmasıdır. Bayes teoremi bir durumun gerçekleşmesi olasılığını, durumun gözlemlenmesinden önceki olaylara bakarak hesaplar (Flach ve Lachiche, 2004).

Bayes teoreminin formülü şu şekildedir:

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)} \tag{2.4}$$

Formülde,

P(A|B) : B olayının gerçekleştiği durumda A olayının gerçekleşme olasılığını temsil eder.

P(B|A) : A olayının gerçekleştiği durumda B olayının gerçekleşme olasılığını temsil eder.

P(A) : A olayının olasılığını temsil eder.

P(B) : B olayını temsil eder.

NB algoritması eğitim ve veri setlerini kullanarak sınıflar arası ilişkileri öğrenir. Sınıflama yapabilmek için eğitim veri setinin özelliklerini analiz eder. Bu özellikler kelime frekansları, belirli kelime öbekleri gibi metinsel ya da sayısal değerler olabilir. Sonrasında test verileri noktaları için sınıflar belirlenir. Algoritma her bir sınıf için ait olma olasılığını hesaplar. En yüksek olasılığa sahip olan sınıf seçilir.

Bayes teoremi uygulamaları, veri analizi, doğal dil işleme, tıbbi teşhis, robotik, finansal tahminler ve daha pek çok alanda yaygın olarak kullanılır.

NB algoritması kullanımı veri setinin yapısına ve analiz edilen problemin özelliklerine göre farklılık gösterebilir. Naive Bayes türleri şunlardır:

Gaussian Naive Bayes (GNB):

GNB, temel varsayım olarak her özelliğin normal (Gaussian) dağılıma sahip olduğunu kabul eder. Bu, özelliklerin her sınıfta normal bir dağılıma sahip olduğu anlamına gelir (Kamel vd., 2019).

Multinominal Naive Bayes (MNB):

MNB, özelliklerin multinomial dağılıma sahip olduğunu varsayar. Bu, özelliklerin kategorik veya sayısal ancak sıklıkla belirli bir kategori içinde sıklıkla tekrarlanan (örneğin, kelime sayıları) olduğu durumlar için uygundur. Genellikle metin sınıflandırması gibi problemlerde kullanılır. Örneğin, belirli bir belgede belirli bir kelimenin belirtilen kategorideki belgelerdeki sıklığı gibi özelliklerle çalışır (Kibriya vd., 2005).

Bernoulli Naive Bayes (BNB):

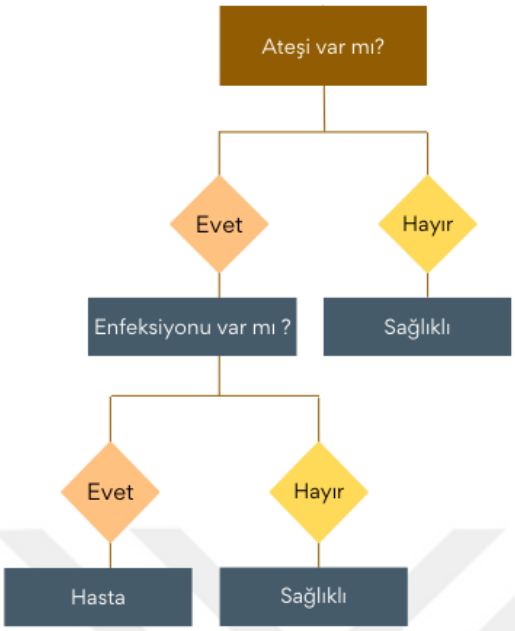
BNB, özelliklerin Bernoulli dağılıma sahip olduğunu varsayar. Bu, özelliklerin ikili (binary) olduğu durumlar için uygundur. Genellikle metin sınıflandırması gibi ikili verilerle çalışır. Örneğin, belirli bir kelimenin bir belge içinde bulunup bulunmadığı gibi özelliklerle çalışır (Singh vd., 2019).

2.2.3. Karar Ağaçları

DT yaygın kullanılan, basit ve kolay anlaşılır makine öğrenme algoritmalarından biridir. DT etiketlenmiş veri kümesi üzerinden oluşturulur. Etiketlenmiş verilerin desen analizleri yapılır ve yeni verileri sınıflandırmak ya da tahmin etmek için ağaç yapısı şeklinde modeller oluşturulur, her düğümde belli bir kurala dayalı karar alınır.

Model oluşturma sürecinde öncelikle tüm veri kümesi kök düğüm olarak tanımlanan bir düğümde toplanır. Ardından veriler belli özelliklerine göre alt gruplara bölünür. Bu işlem homojen gruplar elde edilinceye dek devam eder. Alt düğümler yaprak düğümler olarak ifade edilir. Yaprak düğümler karar ağacının sonuçlarını temsil eder (Charbuty ve Abdulazeez, 2021).

DT algoritmasında ilk aşama kök düğümlerin seçimidir. Kök düğüm veri kümesindeki en önemli özelliği seçerek bir veya daha fazla sınıf oluşturur. İlerleyen her düğümde belirli özelliklere göre veri kümesi bölünür. Bu süreç yaprak düğümlere ulaşıncaya kadar devam eder. Algoritmanın durma kriteri maksimum ağaç derinliği, minimum örnek sayısı gibi kriterlere göre belirlenebilir. Yeni bir veri girdi olarak verildiğinde ağaçtaki karar noktaları ve dallar boyunca ilerlenir. Algoritma hem sınıflama hem de regresyon problemlerinde kullanılabilir.

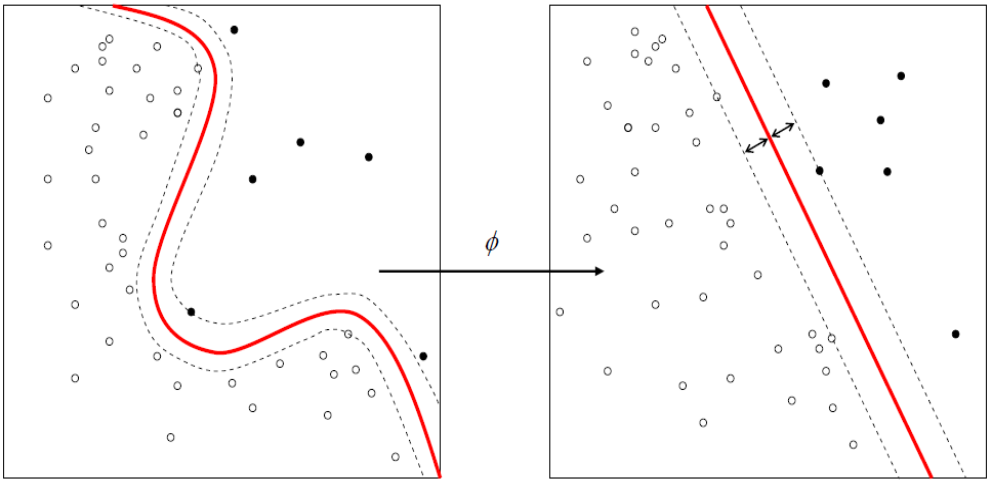


Şekil 6. Karar ağacı

2.2.4. Destek Vektör Makinesi

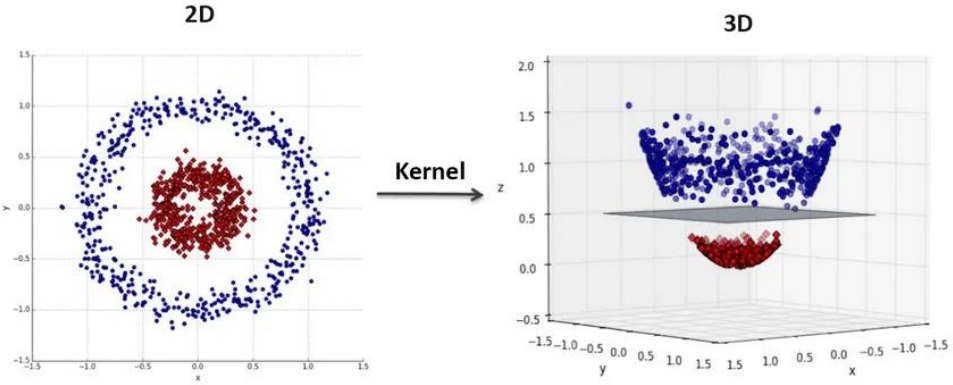
Sınıflandırma ve regresyon problemleri için kullanılan güçlü bir makine öğrenimi modelidir. Denetimli bir öğrenme modelidir. SVM algoritmasında amaç doğrusal olarak ayrılabilecek iki sınıfı birbirinden ayırmaktır.

SVM algoritması, verileri en iyi şekilde ayıran veya tahmin eden bir hiper-düzlem elde etmeye çalışır. Destek vektör makinesi algoritmasının temel fikri, sınıflar arasındaki en geniş aralığı (maksimum margin) elde etmektir (Vapnik, 2013).



Şekil 7. SVM algoritması gösterimi (Şirin, 2017)

Bu algoritmada veriler özellikleriyle uzayda bir koordinat değeri olarak ifade edilir. Ardından, noktalar arasında iki sınıfı ayıracak bir hiper-düzlem oluşturulur. SVM algoritması, doğrusal olarak ayrılabilen veri kümeleri için kullanılabileceği gibi, doğrusal olarak ayrılamayan veri kümelerini ayrıştırmak için de kullanılabilir. Gerçek dünya problemlerinde çoğunlukla iki sınıf doğrusal olarak birbirinden ayrılmaz. Kernel fonksiyonları ile veriyi yüksek boyutlu uzaylara taşıyarak doğrusal olarak ayrılabilir bir yapı oluşturabilir. Genelleme yeteneğinin yüksek olması, aykırı değerlere karşı dirençli olması, yüksek boyutlu veri kümeleriyle çalışabilmesi bu algoritmanın avantajlarıdır. Büyük veri kümelerinde hesaplama maliyetinin yüksekliği, özellikle doğrusal olmayan verileri sınıflamada kernel parametrelerinin ayarlanmasındaki hassasiyet algoritmanın dezavantajlarıdır.



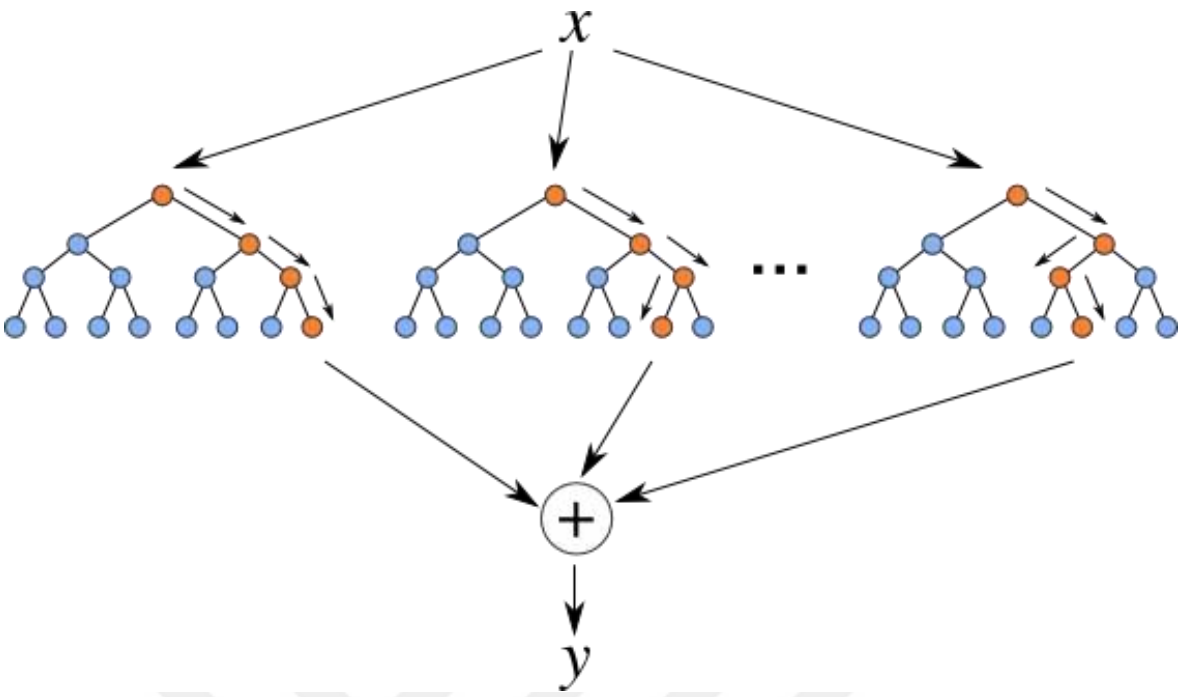
Şekil 8. Kernel fonksiyonu (Akça, 2020)

2.2.5. Rassal Orman

RF algoritması Leo Breiman ve Adele Cutler tarafından geliştirilen, sınıflandırma ve regresyon problemlerinde kullanılabilen bir denetimli öğrenme modelidir (Grömping, 2009).

RF algoritması birden fazla karar ağacını (decision tree) bir araya getirerek çalışır ve bu ağaçların sonuçlarını bir araya getirerek bir tahminde bulunur.

Algoritmada ilk olarak karar ağaçlarını oluşturmak için veri setinden rastgele örneklemler alınır. Bu örnekleme işleminde, veri setinden rastgele örnekler alarak yeni alt veri setleri oluşturur. Bu alt veri setleri, orijinal veri setinden örnekleme yaparken bazı veri noktalarının tekrarlanmasına izin verir. Sonrasında karar ağaçlarının oluşturulması adımına geçilir. Her alt veri seti üzerinde bir karar ağacı oluşturulur. Karar ağacı oluşturma işlemi, veri setindeki özelliklerin (features) kullanılarak verilerin sınıflarını veya çıktılarını tahmin eden bir yapı oluşturmayı içerir. Her ağaç, veri setindeki özellikler arasında en iyi bölünmeyi (split) bulmak için özyinelemeli olarak çalışır. Her karar ağacı, her bir düğümde verileri bölerek en iyi bölünmeyi bulmaya çalışır. En iyi bölünme, verileri en homojen alt gruplara ayıran ve en az hata ile sınıflandıran bölünmedir. Bölünmeler genellikle özelliklerin değerleri üzerinde yapılan karşılaştırmalarla belirlenir. Random Forest, bu ağaçları bir araya getirerek bir ensemble (topluluk) oluşturur. Her bir ağacın tahminini kullanarak, genellikle sınıflandırma durumunda sınıf oylarının (class votes) veya regresyon durumunda tahminlerin ortalaması alınır. Bu sayede, birden fazla modelin tahminleri bir araya gelerek daha doğru ve güvenilir bir tahmin sağlar. Random Forest, her bir ağacın rastgele örnekleme ve rastgele özellik seçimi yaparak oluşturulmasını sağlar. Bu özellikler, her ağacın farklı bir alt kümesini kullanmasını sağlayarak aşırı uydurmaya karşı direnç sağlar. Ayrıca, birden fazla ağacın bir araya gelerek ortalaması veya çoğunluk oyuyla karar vermesi, modelin genelleme yeteneğini artırır (Breiman, 2001).



Şekil 9. Rassal orman (Viswanathan, 2023)

2.3. Veri Seti

Çalışmada Saarland Üniversitesi Fonetik Enstitüsü tarafından oluşturulan ve https://stimmdb.coli.uni-saarland.de/ sitesinde çevrimiçi paylaşılan Saarbruecken Ses Veritabanı kullanılmıştır. Saarbruecken veritabanı 71 farklı patolojiyi içeren zengin bir veritabanıdır. Veritabanında 2000'den fazla bireyin ses kayıtları vardır. Veri tabanında hastalar cinsiyet, yaş, hastalık ismi ya da sağlıklı hasta şeklinde sınıflandırılabilmektedir.

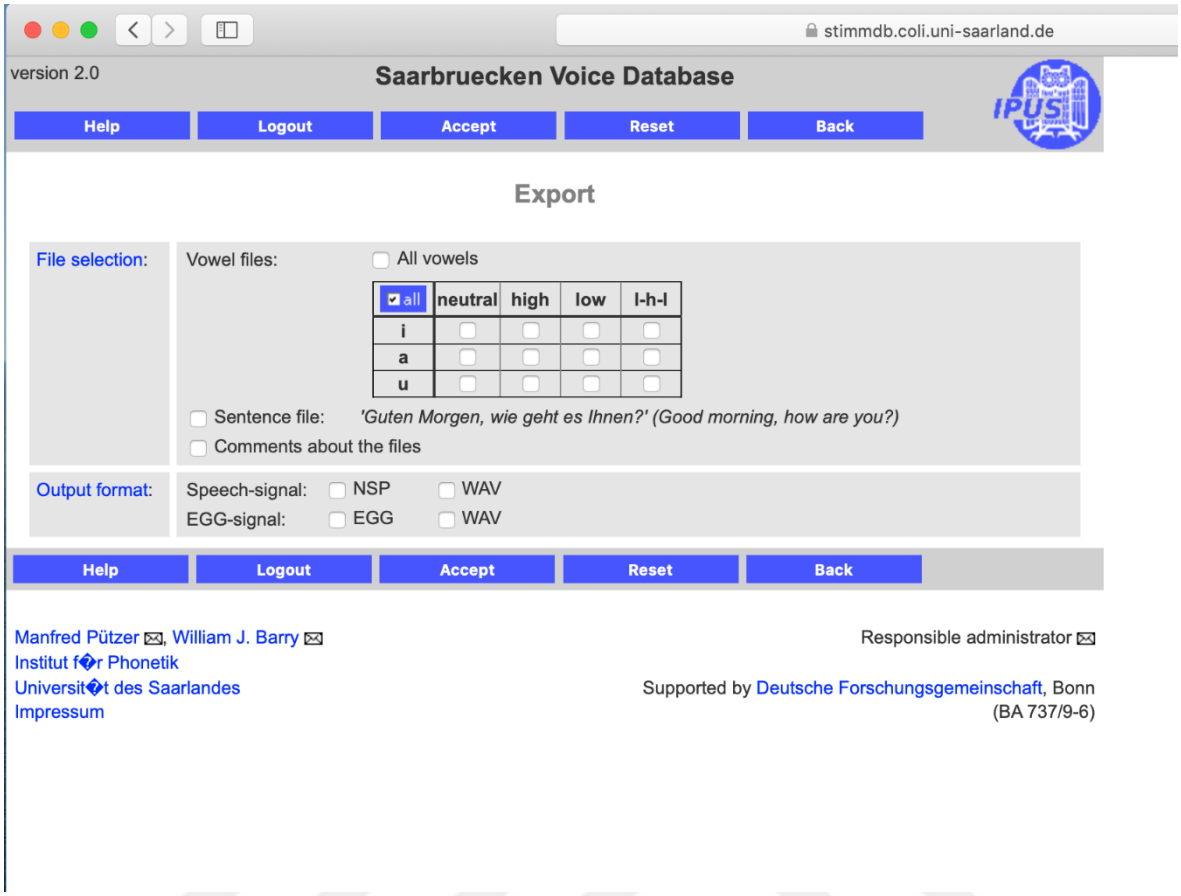
Hastalardan alınan ses kayıtları /i/, /a/, /u/ sesleri şeklinde normal yüksek ve alçak ses tonlarında wav formatında kaydedilmiştir. Sürekli konuşma cümlesi olarak "Guten Morgen,wie geth es Ihnen?" ("Günaydın, nasılsınız?") şeklinde kayıt mevcuttur. Sürekli konuşma cümlesi çalışmada kullanılmamıştır. Kayıtlar 50 kHz örnekleme frekansında ve 16-bit çözünürlükte, "wav" uzantısı ile kaydedilmiştir.

Çalışmada, 176 ses teli felci hastasına ait 1581 ses verisi, 82 larenjit tanılı hastaya ait 738 ses verisi, 35 kanser öncüsü lezyon veya kanser hastasına ait 315 ses verisi, 27 polip tanılı hastaya ait 243 ses verisi, 34 reinke ödemi tanılı hastaya ait 306 ses verisi ve 298 sağlıklı hastadan alınan 2682 ses verisi olmak üzere toplamda 652 hastaya ait 5865 ses verisi

veri seti olarak kullanılmıştır. Çalışmada kullanılan sesler cinsiyet ve yaşa göre ayrılmamış patolojik sınıftaki tüm sesler veri setine dahil edilmiştir.



Şekil 10. Saarbruecken ses veritabanı (William, J.B.ve Manfred, P., 2007)



Şekil 11. Saarbruecken ses veritabanı (William, J.B. ve Manfred, P., 2007)

Veri tabanında bazı hastalarda sınıflamaya konu olan birden fazla hastalık aynı anda bulunmaktadır. Ses verilerinin seçiminde her ses verisi için tek patolojik durum seçilerek veriler indirilmiştir.

2.4. Literatürdeki Çalışmalar

Zuzana Dankovichova ve arkadaşları, yaptıkları çalışmada nörolojik hastalıklar, enfeksiyonlar, doku değişiklikleri gibi nedenlerle seste meydana gelen bozulmaları tespit etmeyi amaçlamışlardır. Çalışmada Saarland Üniversitesi tarafından hazırlanmış ses verilerini kullanmışlardır. Kullanılan 194 sesin 94 ü disfoni (bozuk ses), 100'ü sağlıklı bireylere ait seslerdir. Tüm katılımcılar 18 yaşın üzerindedir. Çalışmada sınıflama yapmak için SVM, RF, k-NN algoritmaları kullanılmıştır. Yapılan çalışmada algoritmaların başarı sonuçları %80,3, %80,6, %86,2 olarak bulunmuştur. Elde edilen sonuçların doğruluk oranlarının yüksek olmasına rağmen çalışmanın eksik yanları, örneklem sayısının küçük

olması ve algoritmaların sesleri patolojik konuşma bozukluğu var (disfoni) ve sağlıklı ses olarak ayırmasıdır (Dankovičová vd., 2018).

Minh Pham ve arkadaşlarının yaptığı çalışmada kullanılan ses verileri Tayvan'da bulunan Far Eastern Memorial Hastanesi'nden (FEMH) elde edilmiştir. Ses örnekleri, bireylerin /a:/ sesli harfini söyledikleri kayıtlardır ve waveform audio file format (.wav) olarak kaydedilmiştir. 50 normal ses örneği ve vokal nodüller, polipler ve kistler, vokal paralizi gibi yaygın ses bozuklukları örneği bulunmaktadır. Eğitim veri setinin etiketleri, cinsiyet, yaş, konuşmacının sağlıklı olup olmadığı ve ilgili ses hastalıklarını içermektedir. Çalışmada SVM, RF, k-NN ve Gradyan Arttırma (GDA) (Khaneja vd., 2005) algoritmaları kullanılmıştır. Elde edilen sonuçlarda, SVM algoritmasının doğruluk oranı %64, RF algoritmasının doğruluk oranı %66, k-NN algoritması doğruluk oranı %66, GDA algoritmasının doğruluk oranı %67 olarak tespit edilmiştir. Yapılan çalışma hastalıkları sınıflamada başarılı sonuçlar vermiştir. Veri setinin küçük olması çalışmanın eksik yönüdür (Pham vd., 2018).

Huiyi Wu ve arkadaşları, yaptıkları çalışmada evrişimli sinir ağları (CNN) (Hershey vd., 2017) kullanılarak sağlıklı ve sağlıksız sesi ayırmayı amaçlamıştır. Çalışmada Saarbrüecken Ses Veritabanı kullanılmıştır. 482 sağlıklı 482 patolojik tanılı hastaların doğal tonda ve yalnız /a/ sesleri kullanılmıştır. Patolojik ses verileri 140 larenjit, 41 lökoplaki, 68 reinke ödemi olan, 213 tekrarlayan larinks sinir felci, 22 ses teli felci karsinomu ve 45 polipli hastaya ait verilerdir. Çalışmanın değerlendirilmesi karmaşıklık matrisiyle yapılmıştır. Elde edilen sonuçlara göre sınıflama başarı oranları eğitim veri setinde, doğrulama veri setinde ve test veri setinde sırasıyla %88, %66, %77 olarak ölçülmüştür. Eğitim veri setinde elde edilen doğruluk doğrulama ve test veri setinden daha iyidir. Bu durum aşırı uyum (overfitting) sorununu ortaya koymaktadır (Wu vd., 2018).

Fahad Taha AL-Dhief ve arkadaşları, yaptıkları çalışmada Saarland Üniversitesi Fonetik Enstitüsü tarafından hazırlanan Saarbrücken Ses Veritabanını kullanmışlardır. Çalışmada yalnızca /a/ sesinden oluşan 600 ses verisi kullanılmıştır. Çalışmada sağlıklı ve patolojik (hastalıklı ses) in birbirinden ayrılması amaçlanmıştır. Veri tabanı, eğitim aşaması için %80 ve test aşaması için %20 olarak bölünmüştür. Çalışmada Çevrimiçi Sıralı Aşırı Öğrenme Makinesi (OSELM) (Gandhi vd., 2021) algoritması kullanılmıştır. Çalışmada, ses özellikleri Mel-Frekanslı Cepstral Katsayıları (MFCC) kullanılarak çıkarılmıştır.

Çalışmanın değerlendirilmesi karmaşıklık matrisi kullanılarak yapılmıştır. Çalışmada elde edilen sonuçlar, maksimum doğruluk, duyarlılık ve özgünlük oranlarının sırasıyla %85, %87 ve %87 olduğunu göstermektedir (AL-Dhief vd., 2020).

Wei Wang ve arkadaşları, ses teli poliplerinin ses verisi üzerinden sıkıştırılmış algılama (Compressed sensing) (Donoho, 2006) ve SVM algoritması kullanarak tahmin edilmesini sağlayan bir çalışma gerçekleştirmişlerdir. Çalışmada, boğaz polipi tespiti için sıkıştırıcı algılama kullanarak sıkıştırılmış /a:/ ve /i:/ ses sinyalleri toplanmıştır. Çalışmada 13 ses teli polipli ve 13 ses teli polipi olmayan hasta olmak üzere 26 hastanın ses verisi kullanılmıştır. Yapılan çalışmada ses sinyali sıkıştırılmış algılama teorisi ile yeniden oluşturulmuş ve elde edilen veri destek vektör makineleri yöntemi ile sınıflandırılmıştır. Sağlıklı ve polip sahibi hastaların frekans bileşenlerinin farklı olduğu ve hastanın boğaz polipi olduğunda ses sinyallerindeki düşük frekans bileşeninin daha belirgin şekilde değiştiği görülmüştür. Hastaların frekans bileşenleri boğaz polipi olduğunda değişebilmektedir. Dolayısıyla, ses sinyallerinin frekans bileşen enerjisi, boğaz polipi tespiti için özellik olarak kullanılabilir. Deneysel sonuçlara göre çalışma polip varlığını tespit etmede başarılıdır. Çalışmanın eksik yönü kullanılan veri setinin küçük olmasıdır (Wang vd., 2014).

Hao-Chun Hu ve arkadaşları, yaptıkları çalışmada yapay zeka kullanarak patolojik ses tanıma yoluyla çeşitli ses teli hastalıklarını tespit etmeyi amaçlamaktadır. Çalışmada 189 normal ses kullanılmıştır. Ses teli atrofisi olan bireylerden 224, tek taraflı ses teli felci olan hastalardan 50, ses teli lezyonları bulunan hastalardan 248 ve addüktör spazmodik disfonili 30 ses olmak üzere toplam 552 hastalıklı ses kullanılmıştır. Sağlıklı ve hastalıklı toplam 741 örnek, 593 örneklik eğitim seti ve 148 örneklik test seti olarak ikiye ayrılmıştır. Modeli eğitmek için konvolüsyonel sinir ağı yaklaşımı uygulanmıştır ve bulgular insan uzmanların bulgularıyla karşılaştırılmıştır. Çalışma sonucunda eğitilen sinir ağı modelinin normal ses, vokal atrofi, tek taraflı vokal felç ve vokal kord lezyonlarını ayırt etme başarısı %66,9 olarak saptanmıştır. Aynı örneklem üzerinde 2 kbb uzmanı ve 2 laringoloğun doğru sınıflandırma oranları %51,4 ve %43,2 olarak tespit edilmiştir. Sonuç olarak elde edilen çalışma başarılı olarak değerlendirilmiş ve doktorlar için tanı koymada yardımcı olarak kullanılabileceği belirtilmiştir (Hu vd., 2021).

Manisha Milani ve arkadaşları çalışmalarında sağlıklı ve hastalıklı insan seslerini gerçek zamanlı ayırt edebilen bir çalışma gerçekleştirmişlerdir. Çalışmada gerçek zamanlı insan sesi mikrofonla bilgisayar ortamına alınmıştır. Elde edilen ses sinyalinin özelliklerinin

çıkarılmasında MFCC yöntemi kullanılmıştır. Bu yöntemle ayıklanan ses sınıflandırmaya hazır hale getirilmiştir. DT, SVM ve yapay sinir ağı (YSA) kullanılmıştır. Sonuç olarak YSA %87,5 ile en yüksek sınıflandırma doğruluğunu vermiştir. DT ve SVM algoritmaları sırasıyla %62,5 ve %50 performans göstermiştir (Milani vd., 2020).

Nursen Keleş, yüksek lisans tez çalışmasında covid tanısını ses verileri üzerinden makine öğrenmesiyle yakalamayı amaçlayan bir çalışma yapmıştır. Çalışmada covid testi pozitif ya da negatif çıkan bireylerin ses verileri kullanılmıştır. 750 örnekle bir veri kümesi oluşturulmuştur. Oluşturulan veri kümesi üzerinde en sık kullanılan MFCC, spektrogram, sıfır geçiş oranı, spektral merkez, spektral roll off, renk tonu, karekök ortalama olmak üzere 7 tane öznitelik çıkarma yöntemi kullanılmıştır. Her bir öznitelik kümesi ayrı ayrı ele alınarak sınıflandırma gerçekleştirilmiştir. Sınıflandırmada SVM, RF, k-NN, NB ve derin sinir ağları (DNNs) (Sze vd., 2017) kullanılmıştır. Pozitif ve negatif bireylerin sınıflandırılmasında SVM %64,02, RF %96,97, k-NN %65,98, NB %83,68, DNNs, %98,48 doğruluk sonucu vermiştir (Keleş, 2022).

Raife Göksu KÜÇÜK yüksek lisans tezinde, parkinson hastalığının makine öğrenmesi yöntemleri ile tespit edilmesi üzerinde çalışmıştır. Çalışmada python dilinde yapılmış, pandas ve numpy, scikit-learn, matplotlib ve seaborn kütüphaneleri kullanılmıştır. Çalışmada 23'ü parkinson hastası olan 31 denekten alınan ses ölçümleri veri seti olarak kullanılmıştır. Veri setinin %30 u eğitim verisi, %70 i test verisi olarak uygulanmıştır. Veri seti üzerinde öncelikle Gauss süreci yöntemi (GP) (Williams ve Rasmussen, 1995) uygulanmış bu yöntem ile elde edilen doğruluk oranı DT, RF, SVM, k-NN ve NB yöntemlerinden elde edilen doğruluk oranlarıyla karşılaştırılmıştır. Elde edilen sonuçlara göre GB ile yapılan analiz %98, NB algoritması ile elde edilen doğruluk %74, SVM algoritması ile elde edilen sonucun doğruluğu %89, DT algoritmasının doğruluğu %89, k-NN algoritmasının doğruluğu %93 olarak bulunmuştur (Küçük, 2019).

Abdullah H. Al-nefaie ve arkadaşları, yaptıkları çalışmada, parkinson hastalığının ses verisi üzerinden tespit edilmesi amaçlanmıştır. Araştırmada k-NN, SVM, RF, Lojistik Regresyon (LR) (Menard, 2002) ve AdaBoost (Schapire, 2013) algoritmaları kullanılmıştır. Bu sınıflandırıcılar, konuşma verilerinden çıkarılan birçok fonetik özellik kullanılarak eğitilmiştir. Elde edilen sonuçlara göre LR algoritması %86, k-NN algoritması %92, SVM algoritması %95, RF algoritması %95, AdaBoost algoritması %93 başarı göstermiştir. Bu

sonuçlar, makine öğrenme yöntemlerinin parkinson bireyleri sağlıklı bireylerden ayırt etmede yüksek performans sergilediğini göstermektedir (Al-nefaie vd., 2024).

Michele Giuseppe Di Cesare ve diğerleri, akıllı telefonlarla kaydedilen ses örnekleri üzerinden çeşitli makine öğrenimi yöntemleri kullanarak sağlıklı ve patolojik sesleri ayırt etmeyi incelemişlerdir. VOICED veri seti kullanılarak yapılan çalışmada, MFCC yöntemi ile ses öznitelikleri çıkarılmış ve k-NN yöntemi %98,3 doğrulukla en yüksek performansı göstermiştir. Bu bulgular, ses bozukluklarının erken tespiti ve sürekli izlenmesi için akıllı telefonların etkin kullanım potansiyelini vurgulamaktadır (Di Cesare vd., 2024).

Verma ve diğerleri, yaptıkları çalışmada MFCC ve temel frekans ve spektral centroid kullanarak ses bozukluğu tespiti yapan yenilikçi bir yöntem üzerinde çalışmışlardır. Bu yöntemde, akustik özellikler üzerine eğitilmiş bir yapay sinir ağı (ANN) (Han ve Wang, 2021) ve MFCC özellikleri üzerine eğitilmiş bir uzun kısa süreli bellek (LSTM) (Yu vd., 2019) modeli birleştirilmiştir. Ardından, ANN ve LSTM modelleri tarafından üretilen olasılıklar bir araya getirilerek XGBoost'a (Chen vd., 2015) girdi olarak kullanılmıştır. Bu yöntem, bir sesin bozuk olup olmadığını daha doğru bir şekilde tespit ederek yüksek doğruluk sağlamıştır. Bu yaklaşım, mevcut tekniklerin ötesine geçerek %95,67 doğruluk, %95,36 duyarlılık, %96,49 özgüllük ve %96,9 F1 skoru ile umut verici sonuçlar elde etmiştir (Verma vd., 2023).

2.5. Değerlendirme Kriterleri

Tez çalışmasında elde edilen sonuçlar karmaşıklık matrisi kullanılarak değerlendirilmiştir. Karmaşıklık matrisi algoritmaların performansını değerlendirmek için kullanılan bir değerlendirme aracıdır. Karmaşıklık matrisi gerçek sınıf tanımlarını ve çalışılan modelin tahminlerini kullanarak dört kategori elde etmemizi sağlar:

1. True Pozitif (TP): Gerçekte pozitif olan örneklerin, model tarafından doğru şekilde pozitif olarak tahmin edilmesidir.
2. False Pozitif (FP): Gerçekte negatif olan örneklerin model tarafından yanlış tanımlanarak pozitif olarak tahmin edilmesidir.
3. True Negatif (TN): Gerçekte negatif olan örneklerin model tarafından doğru olarak negatif sınıfında tanımlanmasıdır.

4. False Negatif (FN): Gerçekte pozitif olan örneklerin model tarafından yanlış tahmin edilerek negatif olarak tanımlanmasıdır.

Karmaşıklık matrisi doğruluk (accuracy), hassasiyet (precision), duyarlılık (recall) ve F1 skoru gibi performans ölçümleri hesaplamak için kullanılır. Aşağıda formülleri verilmiştir (Memiş vd., 2022).

$$\text{Doğruluk (accuracy)} = \frac{1}{k}\sum_{i=1}^{k}\frac{TP_i + TN_i}{TP_i + TN_i + FP_i + FN_i} \tag{2.5}$$

$$\text{Hassasiyet (precision)} = \frac{1}{k}\sum_{i=1}^{k}\frac{TP_i}{TP_i + FP_i} \tag{2.6}$$

$$\text{Duyarlılık (recall)} = \frac{1}{k}\sum_{i=1}^{k}\frac{TP_i}{TP_i + FN_i} \tag{2.7}$$

$$\text{F1} = 2 \times \frac{\text{Hassasiyet} \times \text{Duyarlılık}}{\text{Hassasiyet} + \text{Duyarlılık}} \tag{2.8}$$

Doğruluk hesaplanırken, doğru tahmin edilen örneklerin toplam örnek sayısına oranı alınır. Hassasiyet, pozitif olarak tahmin edilen örneklerin gerçek pozitif sayısına oranıdır. Duyarlılık, gerçekte pozitif olan örneklerin model tarafından kaçının doğru olarak sınıflandırıldığını gösteren ölçümdür. F1 skoru ise hassasiyet ve duyarlılık ölçümlerinin harmonik ortalamasıdır. F1 skoru 0-1 arasında değer alır. 1 ya da 1'e yakın F1 skoru, modelin hem pozitif sınıfı doğru şekilde tanımladığını (yüksek hassasiyet), hem de gerçek pozitiflerin doğru şekilde tahmin edildiğini (yüksek duyarlılık) gösterir.

Elde edilen bu metrikler sınıflandırma modelinin farklı yönlerini değerlendirir. Doğruluk (accuracy) modelin genel performansını, hassasiyet (precision) modelin pozitif tahminlerinin doğruluğunu, duyarlılık (recall) modelin gerçek pozitifleri ne kadar tespit edebildiğini gösterir.

ÜÇÜNCÜ BÖLÜM
ARAŞTIRMA YÖNTEMİ/MATERYAL VE YÖNTEM

3.1. Kullanılan Yazılım ve Donanım

Çalışma veri alma ve veri işlemede sağladığı kolaylıklar nedeniyle Python programlama dilinde yazılmıştır. Ses dosyalarını yükleme, ses verilerini dönüştürme, sesin özelliklerini çıkarma işlemleri için librosa kütüphanesi kullanılmıştır. Verilerin analiz edilmesi, veri çerçevelerinin oluşturulması, verilerin tablolar, grafikler ve görseller şeklinde ifade edilebilmesi için Pandas kütüphanesi, veriler üzerinde matematiksel işlemlerin yapılabilmesi için NumPy kütüphanesi kullanılmıştır. Makine öğrenme algoritmaları için Scikit-learn kütüphanesi kullanılmıştır. Verilerin görselleştirilmesi için Matplotlib kütüphanesi kullanılmıştır. Çalışmada 12th Gen Intel(R) Core(TM) i5 2.50 GHz işlemci ve 32 gb ram kapasiteli bilgisayar kullanılmıştır.

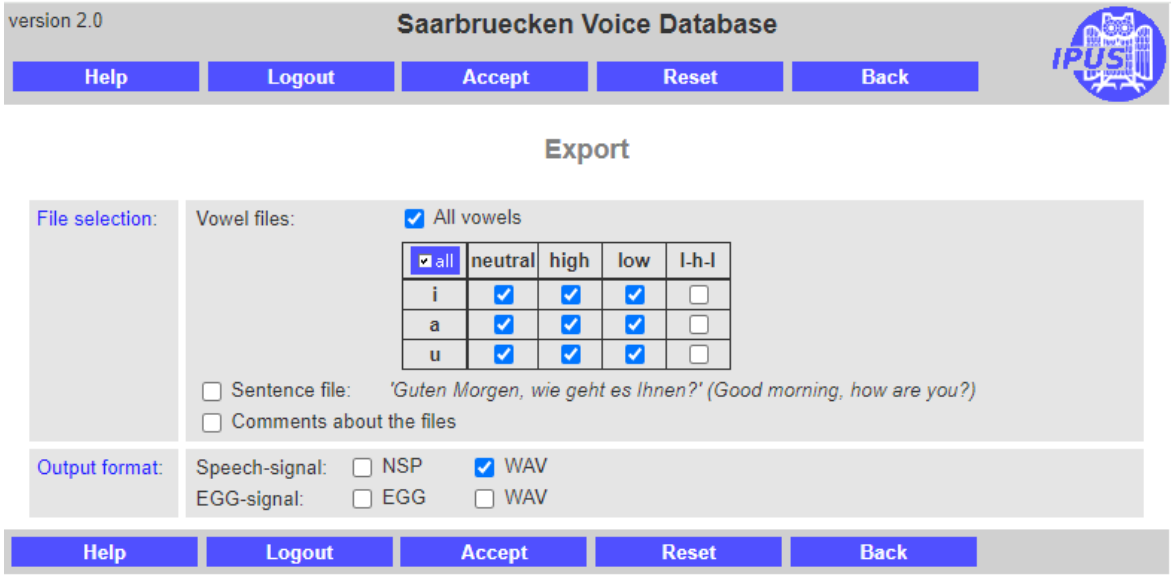
3.2. Verilerin Toplanması

652 hastaya ait 5865 ses verisi https://stimmdb.coli.uni-saarland.de/ adresinden wav uzantılı şekilde indirilmiştir. Hastalara ait indirilen sesler /a/, /i/, /u/ seslerinin normal yüksek ve alçak tonlarını içerir. İndirilen ses verileri patolojilerine göre “saglikli”, “felc”, “larenjit”, “lokoplazi_kanser”, “polip”, “reinke” isimli 6 klasöre ayrılmıştır. Veri seti uygun formata getirildikten sonra uygulamaya yüklenmiştir. Veri tabanında bazı kişilerin birden fazla hastalığa sahip olduğu görülmektedir. Yalnızca bir patolojik tanısı olan hastalara ait ses verileri çalışmada kullanılmıştır. Hastalar yaş cinsiyet olarak ayrılmamış veri tabanındaki uygun patoloji sınıfındaki tüm sesler çalışmaya dahil edilmiştir.

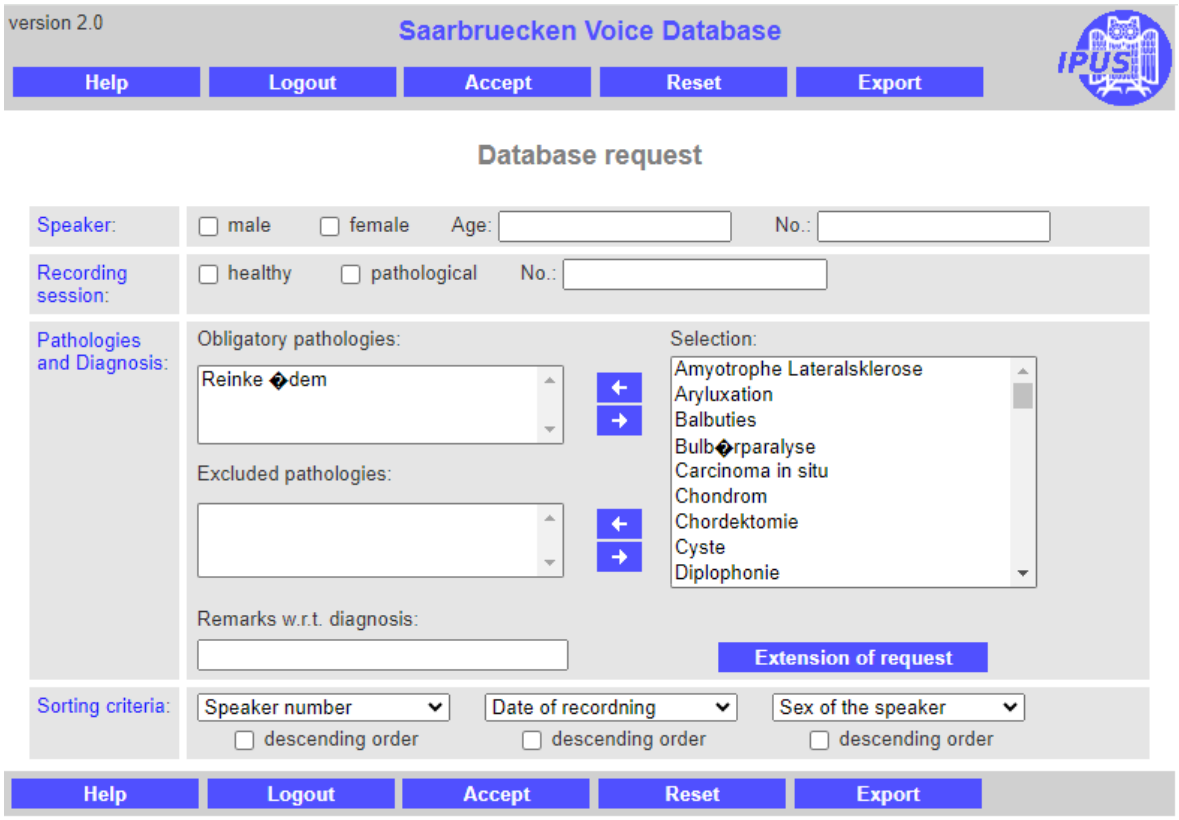
Tablo 1

Tez kapsamında kullanılan veri seti

Hastalıklar	Hasta Sayısı	Ses Dosyası Sayısı
Ses teli felci	176	1581
Larenjit	82	738
Kanser öncüsü lezyon veya kanser	35	315
Polip	27	243
Sağlıklı	298	2682
Reinke ödemi	34	306
Toplam	652	5865



Şekil 12. Veri tabanındaki seslerin seçimi



Şekil 13. Veri tabanındaki hastalıkların seçimi

3.3. Veri Ön İşleme

Veri ön işleme adımı makine öğrenmesi sürecinde verilerin toplanmasından sonra ele alınan ikinci adımdır. Elde edilen veri setlerinin temizlenmesi, düzenlenmesi ve hazırlanması sürecidir. Bu adım, modelin verimli bir şekilde eğitilmesini sağlamak ve modelin performansını artırmak için önemlidir. Veri setinde eksik veriler olabilir. Bu eksik veriler, model performansını olumsuz etkileyebilir. Eksik verilerin tamamlanması veri ön işleme adımında gerçekleştirilir. Veri setinde normal dağılımı etkileyen ve modelin genelleme becerisini düşüren aykırı noktalar olabilir. Aykırı noktaların temizlenmesi işlemi veri ön işleme adımında yapılır. Veri setindeki kategorik verilerin de veri ön işleme adımında numerik verilere dönüştürülmesi gerekir. Bu sayede kategorik veriler makine öğrenme sürecine uygun hale getirilir. Kategorik verilerin numerik hale getirilmesi One-Hot Encoding veya Label Encoding gibi kodlama teknikleri kullanılarak yapılabilir (Brownlee, 2020).

Makine öğrenimi modellerini eğitmeden önce, veri özellikleri standartlaştırmak veri setindeki özelliklerin dağılımını standart bir hale getirir ve modelin daha iyi performans

göstermesini sağlar. Özellikler arasındaki büyük farklar, modelin doğruluğunu olumsuz etkileyebilir. Farklı ölçeklerdeki özellikler arasında model karşılaştırması yaparken, özelliklerin standartlaştırılmış olması, modellerin doğru bir şekilde karşılaştırılmasını sağlar. Bu nedenle makine öğrenme algoritmaları öncesi veri standardizasyonu yaygın bir uygulamadır. Tez çalışmasında da yüklenen veri seti üzerinde eğitim ve test verisi olarak bölünmeden önce standartlaştırma işlemi yapılmıştır.

```
# accuracy artırmak için verileri scale etme
scaler = StandardScaler()
X = scaler.fit_transform(np.array(X))

# train_test_split ile test ve train setleri oluşturma
X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.2, random_state=0)
```

Şekil 14. Verinin standartlaştırılması ve eğitim-test verilerinin seçilmesi

3.3.1. Öznitelik Çıkarımı

Ses verisinde öznitelik çıkarımı, ses sinyallerinden anlamlı özelliklerin çıkarılması işlemidir. Bu özellikler, ses verisini sayısal bir formata dönüştürerek makine öğrenimi modellerinin anlayabileceği bir şekle getirir. Sesin özelliklerini elde etmemizde kullanılan bazı ses özellikleri vardır. Örneğin ses dosyaları genellikle dalga formundadır ve bu dalga formu, örnekleme frekansı ve bit derinliği gibi özelliklere sahiptir. Örnekleme frekansı, saniyedeki örnekleme sayısını, bit derinliği ise her bir örneğin alabileceği değerlerin aralığını belirtir. Zaman alanı öznitelikleri ses dalgasının zamana bağlı değişimlerini temsil eder. Bir diğer önemli öznitelikler sınıfı, ses dalgasının frekans bileşenlerine odaklanır. Bu tür öznitelikler, sesin spektral içeriğini temsil eder. Genellikle Fourier dönüşümü gibi dönüşümler kullanılarak elde edilirler. Bu öznitelikler arasında sesin frekans spektrumu, temel frekans (pitch), spektral genlik, spektral tarama (spectral centroid) ve spektral bant genişliği bulunur.

Tez çalışmasında ses verisi elde edildikten sonra öznitelik çıkarımı için MFCC yöntemi kullanılmıştır. MFCC sesin frekans içeriğini temsil etmek için birçok konuşma tanıma sisteminde yaygın olarak kullanılan bir özelliktir. (Matsui ve Furui, 1995). Özellikle

konuşma tanıma ve ses işleme alanlarında yaygın olarak kullanılır. MFCC, bir ses sinyalinin özelliklerini temsil etmek için kullanılan bir dizi katsayıdır ve bu katsayılar sesin frekans ve zaman özelliklerini temsil eder. MFCC, insan kulağı ölçeğine dayanan Mel ölçeğini kullanarak frekans alanında temel alan konuşma tanıma sistemlerinde kullanılan en popüler özellik çıkarım tekniklerinden biridir (Shen vd., 2011). MFCC, insan kulağının kritik işitme bant genişliği ve frekansındaki değişmeleri, düşük frekansta doğrusal süzgeçler ile yüksek frekanslarda ise logaritmik süzgeçler kullanılarak modellenmesi prensibine dayanır (Eskidere ve Ertaş, 2009).

MFCC ses sinyallerinin zaman-frekans alanındaki özelliklerini yakalamak için bir dizi adımdan geçerler. Bu adımlar şu şekildedir:

Ön Vurgulama (Pre-emphasis):

Ön vurgulama, ses sinyallerindeki yüksek frekans bileşenlerini güçlendirerek sinyalin spektral dengesini iyileştirir. Bu adım özellikle konuşma sinyallerindeki düşük frekans bileşenlerinin güçlendirilmesi ve genel netliğin arttırılması için kullanılır. Ön vurgu işlemi birinci dereceden bir FIR (Finite Impulse Response) filtre kullanılarak gerçekleştirilir. Bu filtre, sinyalin türevini yaklaşık olarak hesaplayarak yüksek frekanslı bileşenlerin vurgulanmasını sağlar. Matematiksel olarak giriş sinyali x[n] ve çıkış sinyali $y[n]$ arasındaki ilişki şu şekildedir:

$$y[n]=x[n]-\alpha x[n-1] \quad (3.1)$$

Burada α, vurgulamanın derecesini kontrol eden bir sabittir ve tipik olarak 0.95 ile 0.97 arasında bir değer alır. Bu değer, genellikle tasarım amacına ve uygulamanın gereksinimlerine bağlı olarak seçilir (Rabiner vd., 1978).

Sinyalinin Kesilmesi (Framing):

Sinyalin kesilmesi adımı, giriş ses sinyalinin küçük, sabit süreli parçalara bölünmesidir. Bu parçalara genellikle "çerçeveler" denir. Çerçeveleme adımı, ses sinyalinin

zaman boyutunu frekans boyutuna dönüştürmeyi sağlar ve sonraki işlemlerin uygulanabilmesi için giriş verisini düzenler. Bu adımın sonunda, giriş ses sinyali belirlenen çerçeve uzunluğuna göre parçalara bölünmüş olur (Karpov, 2003). Bu çerçeveler, ses sinyalinin farklı zaman dilimlerindeki özelliklerini yakalamak için kullanılabilir haldedir.

Pencereleme (Windowing):

Pencereleme adımında bir önceki adımda oluşturulan her çerçeve bir pencere fonksiyonu ile çarpılır. Bu yöntem, çerçevenin başlangıç ve bitiş noktalarındaki ani geçişlerin neden olduğu sinyal bozulmalarını azaltmaya yardımcı olur (Rabiner ve Juang, 1999). Pencereleme adımında Hamming, Hanning, Blackman gibi farklı matematiksel fonksiyonlar kullanılabilir (Lim vd., 1978).

Bu fonksiyonlar aşağıdaki gibidir:

Hamming Fonksiyonu:

$$w(n) = 0.54 - 0.46 \cos \frac{2\pi n}{N-1} \tag{3.2}$$

Formülde; n çerçevedeki örnek indeksi, N ise çerçeve uzunluğudur. Bu formülle, her örneğe bir ağırlık atanır ve bu ağırlık, örneğin çerçevenin başlangıç ve bitiş noktalarına yaklaştıkça azalır. Bu sayede, çerçevedeki örneklerin başlangıç ve bitiş noktalarındaki ani geçişlerin neden olduğu sinyal bozulmaları azaltılır.

Hanning Fonksiyonu:

$$w(n) = 0.5 - 0.5 \cos \frac{2\pi n}{N-1} \tag{3.3}$$

Formülde; n çerçevedeki örnek indeksi, N ise çerçeve uzunluğudur. Hanning fonksiyonunun grafiği, genellikle bir merkezde maksimum değere sahip ve kenarlara doğru düşen bir dalga şeklinde görünür. Bu görüntü sinyalin kenar etkilerini azaltmaya ve Fourier dönüşümü sonuçlarını iyileştirmeye yardımcı olur.

Blackman Fonksiyonu:

Blackman pencere fonksiyonunun formülü şu şekildedir:

$$w(n) = 0.42 - 0.5.\cos\frac{2\pi n}{N-1} + 0.08.\cos\frac{4\pi n}{N-1} \tag{3.4}$$

Formülde; w(n) Blackman pencere fonksiyonunu, n pencere içindeki örnekleme indeksini, N pencere uzunluğunu temsil eder. Formülde cos fonksiyonları, pencerenin içindeki örneklerin ağırlıklarını belirler.

Hızlı Fourier Dönüşümü (Fast FourierTransform):

Her çerçeve için, Fourier dönüşümü uygulanır. Fourier dönüşümü, bir zaman alanındaki sinyali frekans alanına dönüştürmek için kullanılan matematiksel bir işlemdir. Bu, ses sinyalinin frekans bileşenlerine ayrılmasını sağlar ve çerçevenin frekans içeriğini gösteren bir güç spektrumunu elde edilir. Bu sayede, bir sinyalin zaman alanında nasıl değiştiğini gösteren bir fonksiyonun, o sinyalin frekans bileşenlerine ayrılmasını sağlar. Kısaca bir sinyalin zamana göre nasıl değiştiğini anlamak yerine, sinyalin hangi frekanslardaki bileşenlere sahip olduğunu anlamamıza olanak tanır. Matematiksel olarak bir sinyalin Fourier dönüşümü şu şekilde ifade edilir:

$$X(f) = \int_{-\infty}^{\infty} x(t).e^{-i2\pi ft}\ dt \tag{3.5}$$

Formülde;

- X(t), zaman alanında verilen sinyali temsil eder.
- X(f), frekans alanında sinyalin dönüşmüş halini temsil eder.
- f, frekansı temsil eder.
- t, zamanı temsil eder.
- e, Euler sabitini temsil eder.

Bu formül, zaman domainindeki bir sinyali frekans domainine dönüştürmek için integral hesaplamasını temsil eder. Sonuç olarak, Fourier dönüşümü, bir sinyalin frekans bileşenlerini belirlememizi sağlar. Böylece çerçevenin frekans içeriğini gösteren bir güç spektrumunu elde ederiz.

Mel-Frekans Filtre Bankası Uygulaması (Mel Filter Bank):

Mel Filtresi Bankası, ses sinyallerinin spektral içeriğini insan işitme sisteminin özelliklerine daha yakın bir şekilde temsil etmek için kullanılan bir tekniktir (Rabiner vd., 1978).

Bir önceki adımda elde edilen güç spektrumu, insan işitme sisteminin frekans algısına daha uyumlu bir şekilde temsil edilmek üzere mel ölçeğine dönüştürülür.

Mel Filtresi Bankası, frekans spektrumunu lineer olmayan bir şekilde bölerek, insan kulağının algısal özelliklerine daha uygun hale getirir. İnsan kulağı, düşük frekanslı seslere daha duyarlıdır ve yüksek frekanslı seslere daha az duyarlıdır. Dolayısıyla, Mel Filtresi Bankası, frekans bandlarını bu insan işitme özelliklerine dayalı olarak düzenler.(Rabiner vd., 1978).

Mel frekansını hesaplamak için şu formül kullanılır:

$$m = 2595 \log_{10}\left(1 + \frac{f}{700}\right) \tag{3.6}$$

Formülde f Hertz cinsinden frekansı temsil eder. Bu formülle, Hertz cinsinden bir frekans Mel ölçeğine dönüştürülür (Davis ve Mermelstein, 1980).

Logaritmik Dönüşüm:

Bir önceki adımda Mel filtresi bankasındaki her bir filtreden elde edilen enerji logaritmik olarak dönüştürülür. Bu adım insan işitmesinin özelliklerini daha iyi temsil etmek için yapılır ve ayrıca özellik vektörünün genlik ölçeğindeki büyük değişimleri dengelemeye yardımcı olur (Rabiner vd., 1978).

Ayrık Kosinüs Dönüşümü:

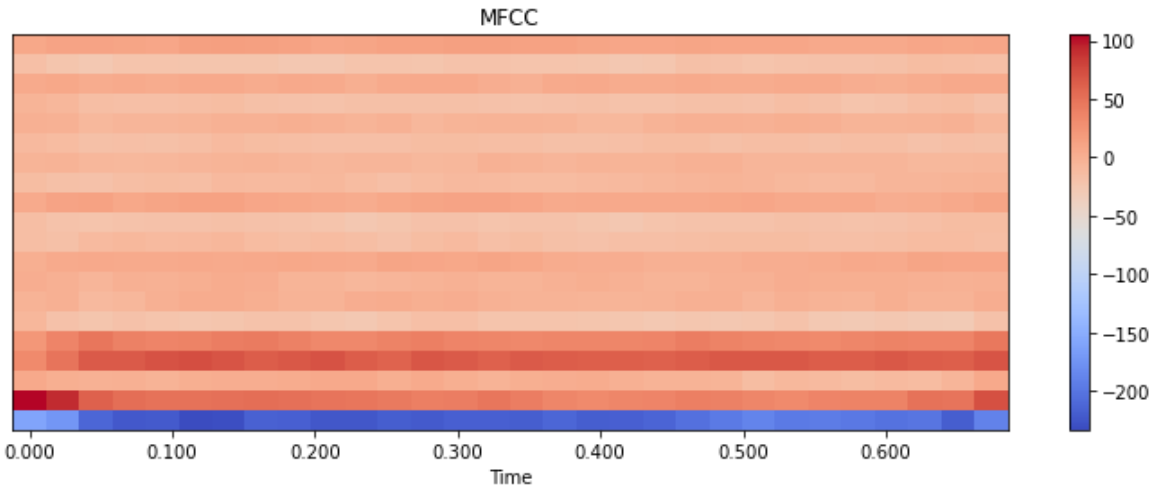
MFCC elde edilmesinde en son olarak kepstrum katsayıları hesaplanır. Kepstral gösterimi ile kayıt ve iletim ortamından dolayı oluşan spektral şekil değişimleri kaldırılır. Ayrıca kepstral katsayılar, yüksek derecede istatiksel bağımsızlık gösterip genlik spektrum gösteriminden daha yüksek tanıma oranı verirler. Gerçel kepstrum, logaritmik genlik spektrumun ters fourier dönüşümü olarak tanımlanıp, gerçel işaretler için kosinüs dönüşümü kullanılarak hesaplanır (Eskidere ve Ertaş, 2009).

Sonuç olarak, her çerçeve için bir MFCC vektörü elde edilir. Bu vektörler, ses sinyalinin zaman içindeki değişimini temsil eder ve makine öğrenimi modelleri için girdi olarak kullanılabilir. MFCC, ses sinyalinin özünü temsil eden kompakt bir formda bilgi sağlar.

```
import matplotlib.pyplot as plt

plt.figure(figsize=(10, 4))
librosa.display.specshow(mfcc, x_axis="time")
plt.colorbar()
plt.title("MFCC")
plt.tight_layout()
plt.show()
```

Şekil 15. MFCC gösteriminin Python kodu



Şekil 16. MFCC spektrumu

3.4. Veri Setinin Bölünmesi ve Modelin Eğitilmesi

Eğitim ve veri setinin hangi oranlarda bölüneceği veri setinin büyüklüğü, problemin karmaşıklığı, veri setinin kalitesi gibi birçok etkene bağlıdır. Bu tez çalışmasında veri seti eğitim ve test verisi olarak öncelikle %70 ve %30 oranlarında bölünmüştür, sonrasında %80 ve %20 oranında denenmiştir. %80 ve %20 oranları daha başarılı sonuç verdiği için çalışmada %80 ve %20 oranları kullanılmıştır. Verilerin bölünmesinde scikit-learn kütüphanesinde 'train_test_split' fonksiyonu kullanılmıştır. Eğitim ve test verisi oluşturulurken rastgelelik kontrolü için kullanılan 'random _state' parametresi 0 olarak belirlenmiştir.

Tez çalışmasında kullanılan, k-NN, NB, SVM ve RF olmak üzere 5 algoritma reinke ödemi, akut larenjit, kanser ve kanser öncesi lekelenme, polip, ses teli felci ve sağlıklı sınıflarına ayrılan ses verileri üzerinde kullanılmıştır. Algoritmaların her biri öncelikle tüm ses verileri içerisinden yalnızca sağlıklı sesleri ayırmıştır. Başka bir ifadeyle her bir makine öğrenme algoritması ile öncelikle sağlıklı-sağlıksız ses ayrımı yapılmıştır. Sonrasında her bir makine öğrenme algoritması ses verilerini ses patolojilerine göre ayırmıştır.

```
for folder in os.listdir(curr_path):
    if folder in [
        "felc",
        "larenjit",
        "lokoplazi_kanser",
        "polip",
        "reinke",
        "saglikli",
    ]:
        for audio_file in os.listdir(join(curr_path, folder)):
            if ".wav" in audio_file:
                y, sr = librosa.load(join(join(curr_path, folder), audio_file))
                mfcc = librosa.feature.mfcc(y=y, sr=sr)
                mfcc_scaled = np.mean(mfcc.T, axis=0)

                if folder == "saglikli":
                    category = "saglikli"
                else:
                    category = "hasta"

                df = pd.concat(
                    [
                        df,
                        pd.DataFrame.from_records(
                            [{"MFCCS": mfcc_scaled, "category": category}]
                        ),
                    ]
```

Şekil 17. Ses verilerinin hasta-sağlıklı olarak ayrılması

```
# Audio klasöründeki her klasör için
for folder in os.listdir(curr_path):

    if folder in [
        "felc",
        "larenjit",
        "lokoplazi_kanser",
        "polip",
        "reinke",
        "saglikli",
    ]:

        # her klasör içindeki her dosya için
        for audio_file in os.listdir(join(curr_path, folder)):

            # .ipynb gizli dosyası hata verdiği için .wav kontrol ediliyor
            if ".wav" in audio_file:

                # librosa kütüphanesi ile wav yükleme
                y, sr = librosa.load(join(join(curr_path, folder), audio_file))

                # mfcc alma ve numpy mean alma
                mfcc = librosa.feature.mfcc(y=y, sr=sr)
                mfcc_scaled = np.mean(mfcc.T, axis=0)

                df = pd.concat(
                    [
                        df,
                        pd.DataFrame.from_records(
                            [{"MFCCS": mfcc_scaled, "category": folder}]
                        ),
                    ]
                )
```

Şekil 18. Ses verilerinin sınıflarına göre ayrılması

3.5. Modellerin Değerlendirilmesi

Tez çalışmasında algoritmalardan elde edilen sonuçlar karmaşıklık matrisi kullanılarak değerlendirilmiştir. Çalışmada karmaşıklık matrisi (confusion matrix) sklearn.metrics modülünün altında confusion_matrix fonksiyonu kullanılarak uygulanmıştır. Her bir algoritma için öncelikle sağlıklı-hasta ayrımı yapan kodların doğruluk (accuracy), hassasiyet (precision), duyarlılık (recall) ve F1 skorları hesaplanmıştır. Sonrasında her algoritma için veri setindeki reinke ödemi, akut larenjit, kanser ve kanser öncesi lekelenme, polip, ses teli felci ve sağlıklı sınıflarının ayrımı yapılmış ve doğruluk (accuracy), hassasiyet (precision), duyarlılık (recall) ve F1 skorları hesaplanmıştır. Elde edilen sonuçlar dördüncü bölümde detaylı bir şekilde incelenmiştir.

```
cnf_matrix = np.array(cnf_matrix)
# TP ve TN değerlerini belirle
TP = cnf_matrix[0, 0]
TN = cnf_matrix[1, 1]

# Toplam doğru sınıflandırılan örnek sayısını hesapla
total_correct_predictions = TP + TN

# Toplam örnek sayısını hesapla
total_samples = np.sum(cnf_matrix)

# Doğruluk değerini hesapla
accuracy = total_correct_predictions / total_samples
print("Doğruluk:", accuracy)
```

Şekil 19. Karmaşıklık matrisi ile doğruluk hesaplanması

DÖRDÜNCÜ BÖLÜM
ARAŞTIRMA BULGULARI

1.1. K En Yakın Komşu Algoritmasından Elde Edilen Sonuçlar

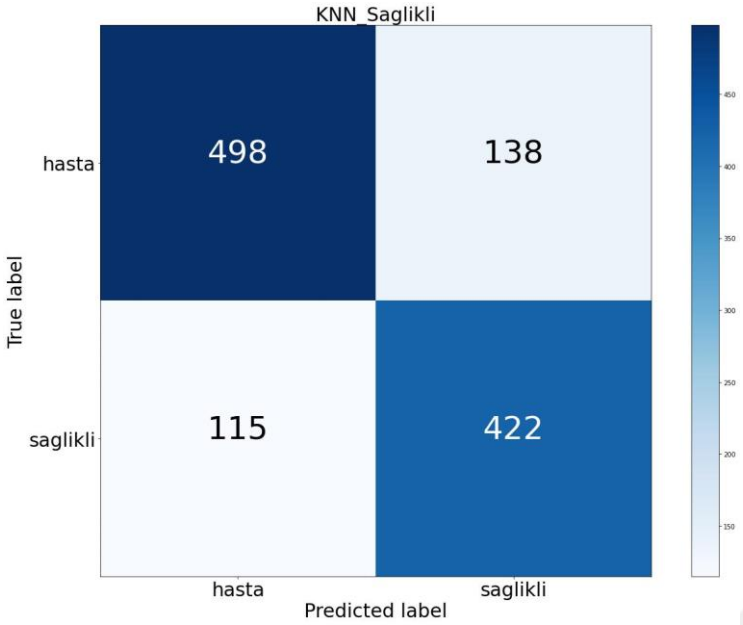
Çalışmada k-NN algoritmasından en iyi sonucu almak için Grid Search yöntemi kullanılmıştır. Grid Search, önceden belirlenen hiperparametre aralığında muhtemel tüm olasılıkları deneyerek en iyi sonucu veren bir arama algoritmasıdır (Ghawi ve Pfeffer, 2019). Model değerlendirmede k-fold çarpraz doğrulama kullanılmıştır. Katman sayısı 10 olarak belirlenmiştir.

k-NN algoritmasının sağlıklı ve hasta sesleri ayırmasında elde edilen doğruluk (accuracy), hassasiyet (precision), duyarlılık (recall) ve F1 skorları aşağıda tabloda verilmiştir. Uygulanan Grid Search sürecinde elde edilen en iyi parametreler, k = 4, ağırlıklandırma fonksiyonu “distance”, mesafe metriği “manhattan” olarak belirlenmiştir.

Tablo 2

K-NN algoritmasının sağlıklı-hasta verilerini ayırmasında elde edilen sonuçlar

Algoritma Adı	**Doğruluk**	**Hassasiyet**	**Duyarlılık**	**F1 skoru**
k-NN	0,78	0,81	0,78	0,79



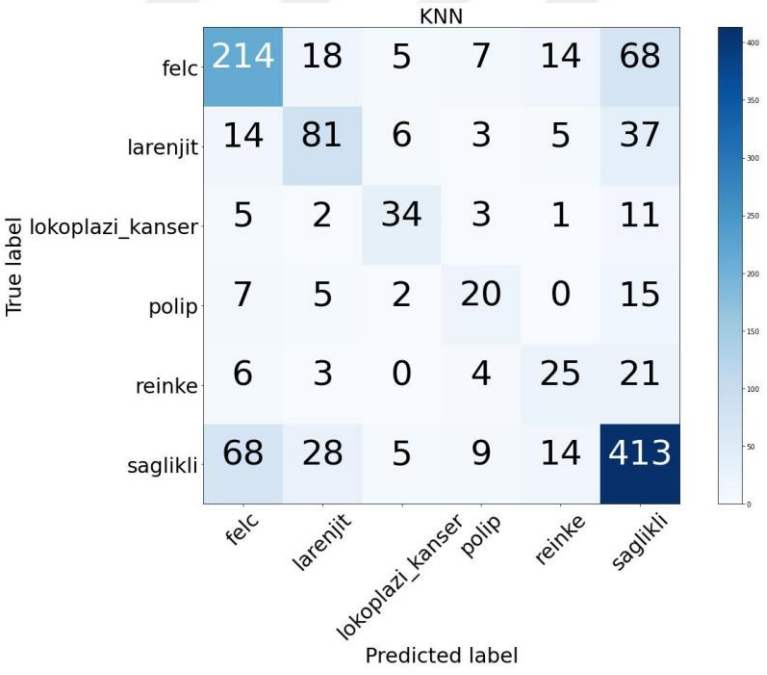
Şekil 20. k-NN algoritmasının sağlıklı-hasta verilerini ayırmasında elde edilen karmaşıklık matrisi

k-NN algoritmasının veri setindeki reinke ödemi, akut larenjit, kanser ve kanser öncesi lekelenme, polip, ses teli felci ve sağlıklı sınıflarının sesleri ayırmasında elde edilen doğruluk (accuracy), hassasiyet (precision), duyarlılık (recall) ve F1 skorları aşağıda tabloda verilmiştir. Uygulanan Grid Search sürecinde elde edilen sonuçlara göre en iyi parametreler, k = 1, ağırlıklandırma fonksiyonu ‘uniform’, mesafe metriği ‘euclidean’ olarak belirlenmiştir.

Tablo 3

k-NN algoritmasının patolojilere göre sınıflamada elde ettiği değerler

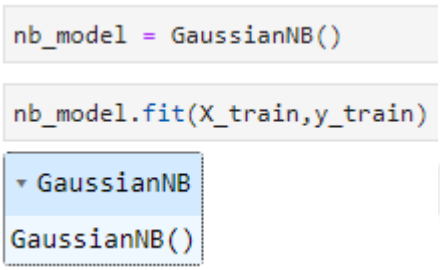
Hastalık	Doğruluk	Hassasiyet	Duyarlılık	F1 skoru
Felç	0,65	0,68	0,65	0,66
Larenjit	0,55	0,59	0,55	0,57
Kanser	0,60	0,65	0,60	0,62
Polip	0,40	0,43	0,40	0,42
Reinke Ödemi	0,42	0,42	0,42	0,42
Sağlıklı	0,76	0,73	0,76	0,74



Şekil 21. k-NN algoritmasının patolojilere göre sınıflamada elde ettiği karmaşıklık matrisi

4.2. Naive Bayes Algoritmasından Elde Edilen Sonuçlar

NB algoritmasının sağlıklı ve hasta sesleri ayırmasında elde edilen doğruluk (accuracy), hassasiyet (precision), duyarlılık (recall) ve F1 skorları aşağıda tabloda verilmiştir. Çalışmada Gaussian Naive Bayes sınıflandırıcısı kullanılmıştır. Model değerlendirmede k-fold çarpraz doğrulama kullanılmıştır. Katman sayısı 5 olarak belirlenmiştir.

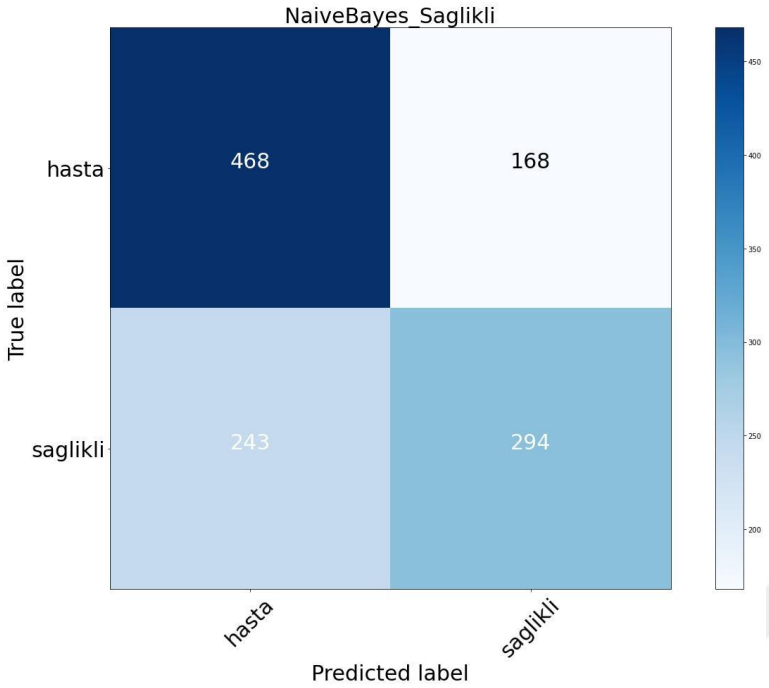


Şekil 22.NB algoritmasının uygulanması

Tablo 4

NB algoritmasının sağlıklı-hasta verilerini ayırmasında elde edilen sonuçlar

Algoritma Adı	Doğruluk	Hassasiyet	Duyarlılık	F1 skoru
Naive Bayes	0,64	0,65	0,73	0,69



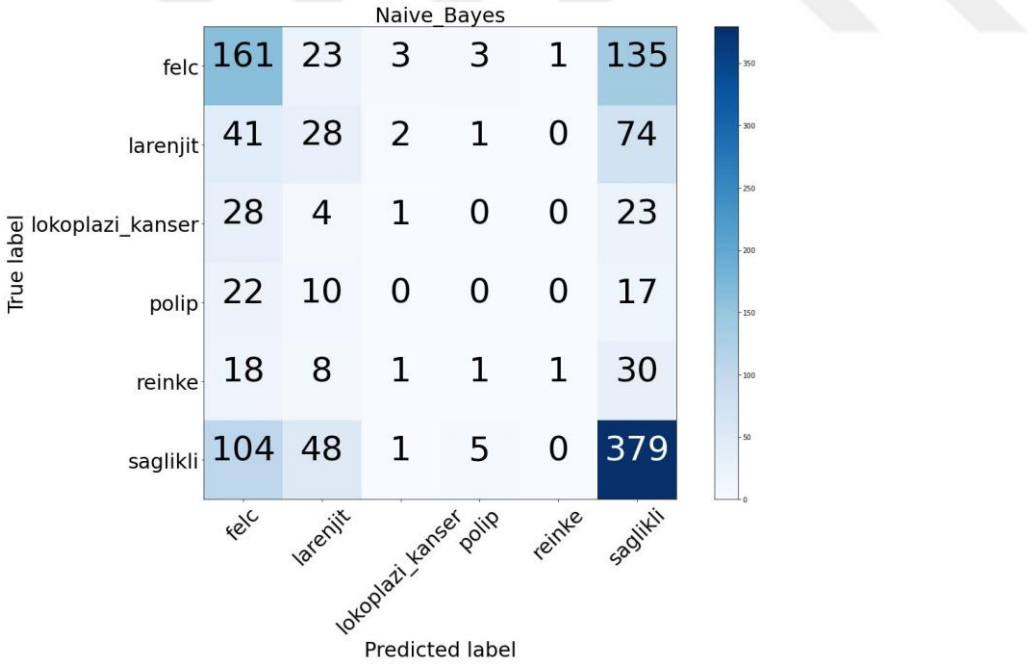
Şekil 23. NB algoritmasının sağlıklı-hasta verilerini ayırmasında elde edilen karmaşıklık matrisi

NB algoritmasının veri setindeki reinke ödemi, akut larenjit, kanser ve kanser öncesi lekelenme, polip, ses teli felci ve sağlıklı sınıflarının sesleri ayırmasında elde edilen doğruluk (accuracy), hassasiyet (precision), duyarlılık (recall) ve F1 skorları aşağıda tabloda verilmiştir. Model değerlendirmede k-fold çarpraz doğrulama kullanılmıştır. Katman sayısı 5 olarak belirlenmiştir.

Tablo 5

NB algoritmasının patolojilere göre sınıflamada elde ettiği değerler

Hastalık	Doğruluk	Hassasiyet	Duyarlılık	F1 skoru
Felç	0,49	0,43	0,49	0,46
Larenjit	0,19	0,23	0,19	0,20
Kanser	0,01	0,12	0,01	0,03
Polip	0,0	0	0	-
Reinke Ödemi	0,01	0,5	0,01	0,03
Sağlıklı	0,70	0,57	0,70	0,63



Şekil 24. NB algoritmasının patolojilere göre sınıflamada elde ettiği karmaşıklık matrisi

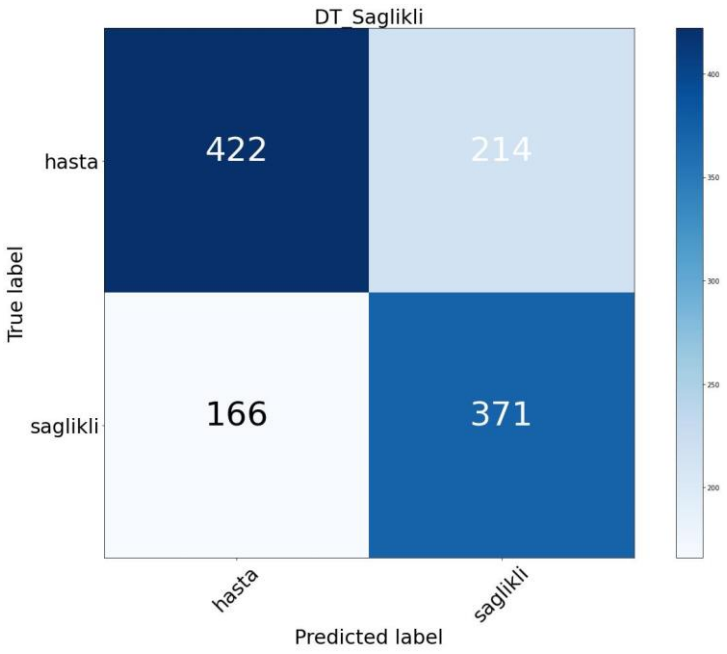
4.3.Karar Ağaçları Algoritmasından Elde Edilen Sonuçlar

DT algoritmasının sağlıklı ve hasta sesleri ayırmasında elde edilen doğruluk (accuracy), hassasiyet (precision), duyarlılık (recall) ve F1 skorları aşağıda tabloda verilmiştir. Model değerlendirmede k-fold çarpraz doğrulama kullanılmıştır. Katman sayısı 5 olarak belirlenmiştir. Uygulanan Grid Search sürecinde elde edilen sonuçlara göre en iyi parametreler, max depth = 10, min samples leaf = 1, min samples split = 2, criterion = 'gini' olarak belirlenmiştir.

Tablo 6

DT algoritmasının sağlıklı-hasta verilerini ayırmasında elde edilen sonuçlar

Algoritma Adı	Doğruluk	Hassasiyet	Duyarlılık	F1 skoru
(DT)	0,67	0,71	0,66	0,68



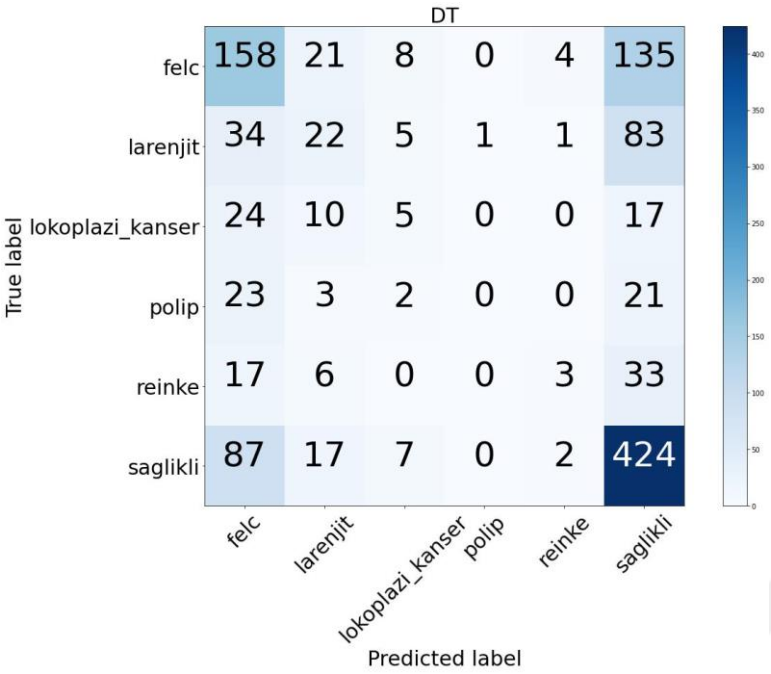
Şekil 25. DT algoritmasının sağlıklı-hasta verilerini ayırmasında elde edilen karmaşıklık matrisi

DT algoritmasının veri setindeki reinke ödemi, akut larenjit, kanser ve kanser öncesi lekelenme, polip, ses teli felci ve sağlıklı sınıflarının sesleri ayırmasında elde edilen doğruluk (accuracy), hassasiyet (precision), duyarlılık (recall) ve F1 skorları aşağıda tabloda verilmiştir. Model değerlendirmede k-fold çarpraz doğrulama kullanılmıştır. Katman sayısı 5 olarak belirlenmiştir. Uygulanan Grid Search sürecinde elde edilen sonuçlara göre en iyi parametreler, max depth = 7, min samples leaf = 4, min samples split = 10, criterion = 'gini' olarak belirlenmiştir.

Tablo 7

DT algoritmasının patolojilere göre sınıflamada elde ettiği değerler

Hastalık (DT)	**Doğruluk**	**Hassasiyet**	**Duyarlılık**	**F1 skoru**
Felç	0,48	0,46	0,48	0,47
Larenjit	0,15	0,27	0,15	0,19
Kanser	0,08	0,1	0,08	0,12
Polip	0,0	0	0	-
Reinke Ödemi	0,05	0,3	0,05	0,08
Sağlıklı	0,78	0,59	0,78	0,67



Şekil 26. DT algoritmasının patolojilere göre sınıflamada elde ettiği karmaşıklık matrisi

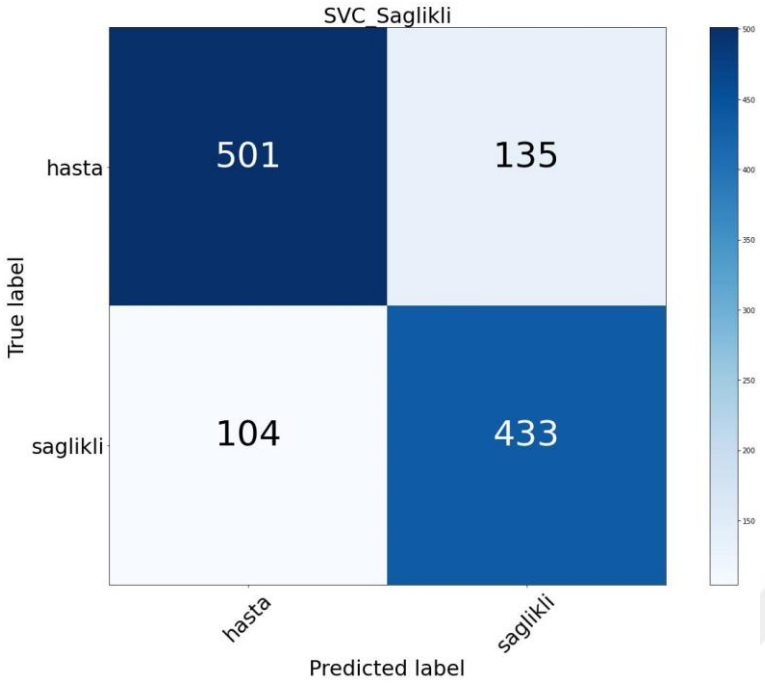
4.4.Destek Vektör Makinaları Algoritmasından Elde Edilen Sonuçlar

SVM algoritmasının sağlıklı ve hasta sesleri ayırmasında elde edilen doğruluk (accuracy), hassasiyet (precision), duyarlılık (recall) ve F1 skorları aşağıda tabloda verilmiştir. Model değerlendirmede k-fold çarpraz doğrulama kullanılmıştır. Katman sayısı 5 olarak belirlenmiştir. Uygulanan Grid Search sürecinde elde edilen sonuçlara göre en iyi parametreler, kernel = 'rbf', c = 10, gamma = 'scale' olarak belirlenmiştir.

Tablo 8

SVM algoritmasının sağlıklı-hasta verilerini ayırmasında elde edilen sonuçlar

Algoritma Adı	Doğruluk	Hassasiyet	Duyarlılık	F1 skoru
SVM	0,79	0,82	0,78	0,80



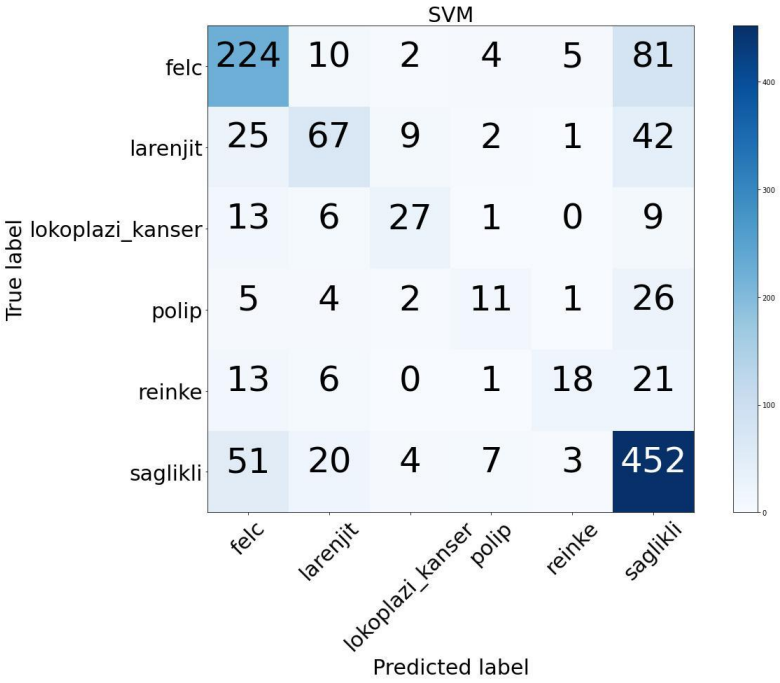
Şekil 27. SVM algoritmasının sağlıklı-hasta verilerini ayırmasında elde edilen karmaşıklık matrisi

SVM algoritmasının veri setindeki reinke ödemi, akut larenjit, kanser ve kanser öncesi lekelenme, polip, ses teli felci ve sağlıklı sınıflarının sesleri ayırmasında elde edilen doğruluk (accuracy), hassasiyet (precision), duyarlılık (recall) ve F1 skorları aşağıda tabloda verilmiştir. Model değerlendirmede k-fold çarpraz doğrulama kullanılmıştır. Katman sayısı 5 olarak belirlenmiştir. Uygulanan Grid Search sürecinde elde edilen sonuçlara göre en iyi parametreler, kernel = ‘rbf’, c = 10, gamma = 0,1 olarak belirlenmiştir.

Tablo 9

SVM algoritmasının patolojilere göre sınıflamada elde ettiği değerler

Hastalık (SVM)	Doğruluk	Hassasiyet	Duyarlılık	F1 skoru
Felç	0,68	0,67	0,68	0,68
Larenjit	0,45	0,59	0,45	0,51
Kanser	0,48	0,61	0,48	0,54
Polip	0,22	0,42	0,22	0,29
Reinke Ödemi	0,30	0,64	0,30	0,41
Sağlıklı	0,84	0,71	0,84	0,77



Şekil 28. SVM algoritmasının patolojilere göre sınıflamada elde ettiği karmaşıklık matrisi

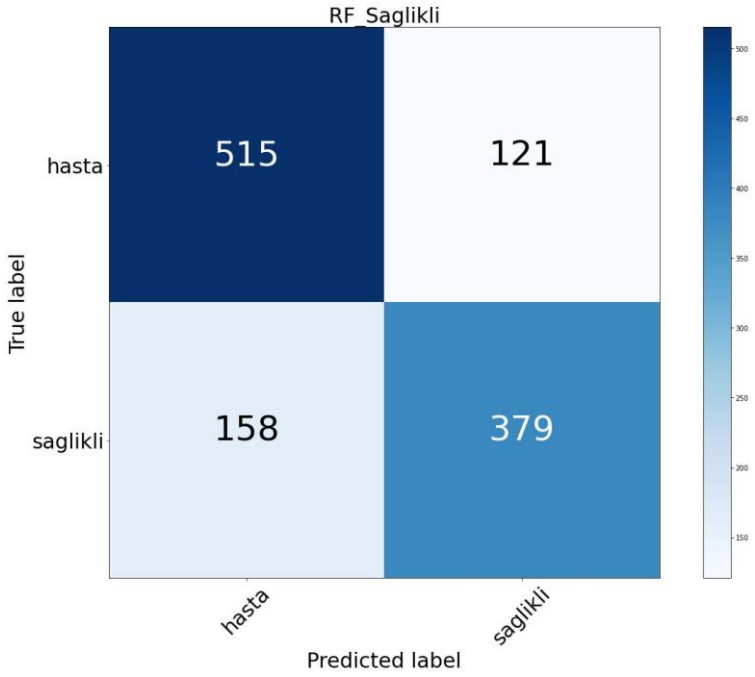
4.5. Rassal Orman Algoritmasından Elde Edilen Sonuçlar

RF algoritmasının sağlıklı ve hasta sesleri ayırmasında elde edilen doğruluk (accuracy), hassasiyet (precision), duyarlılık (recall) ve F1 skorları aşağıda tabloda verilmiştir. Model değerlendirmede k-fold çarpraz doğrulama kullanılmıştır. Katman sayısı 5 olarak belirlenmiştir. Algoritmada kullanılan parametreler, min samples split = 5, max depth = 'none', min samples leaf = 1, n_estimator = 100 olarak belirlenmiştir.

Tablo 10

RF algoritmasının sağlıklı-hasta verilerini ayırmasında elde edilen sonuçlar

Algoritma Adı	Doğruluk	Hassasiyet	Duyarlılık	F1 skoru
Random Forest	0,76	0,76	0,8	0,78



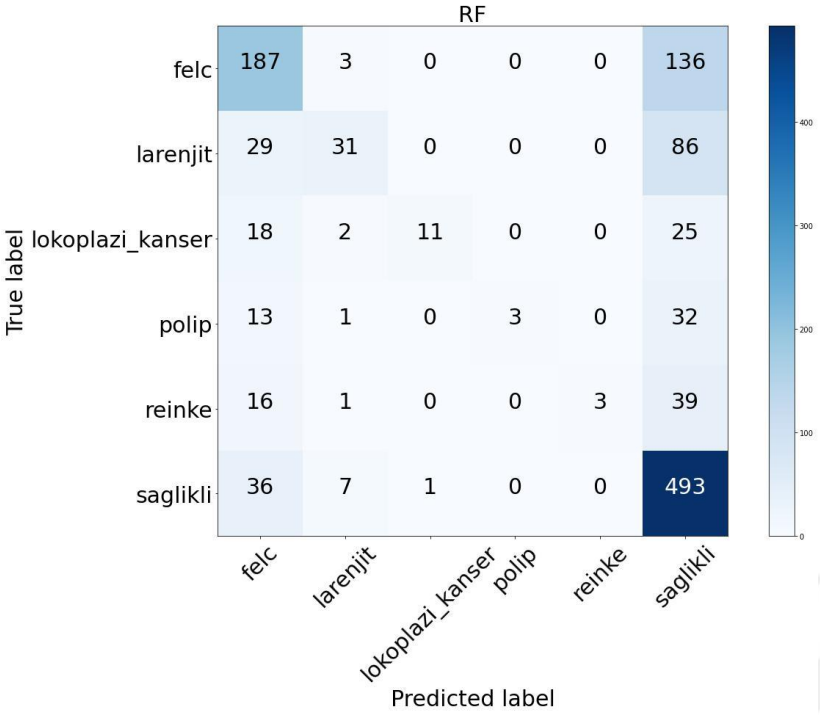
Şekil 29. RF algoritmasının sağlıklı-hasta verilerini ayırmasında elde edilen karmaşıklık matrisi

RF algoritmasının veri setindeki reinke ödemi, akut larenjit, kanser ve kanser öncesi lekelenme, polip, ses teli felci ve sağlıklı sınıflarının sesleri ayırmasında elde edilen doğruluk (accuracy), hassasiyet (precision), duyarlılık (recall) ve F1 skorları aşağıda tabloda verilmiştir. Çalışmada RF algoritmasından en iyi sonucu almak için Grid Search yöntemi kullanılmıştır. Model değerlendirmede k-fold çarpraz doğrulama kullanılmıştır. Katman sayısı 5 olarak belirlenmiştir. Algoritmada kullanılan parametreler, min samples split = 2, max depth = 'none', min samples leaf = 1, n_estimator = 100 olarak belirlenmiştir.

Tablo 11

RF algoritmasının patolojilere göre sınıflamada elde ettiği değerler

Hastalık (RF)	**Doğruluk**	**Hassasiyet**	**Duyarlılık**	**F1 skoru**
Felç	0,57	0,62	0,57	0,60
Larenjit	0,21	0,68	0,21	0,34
Kanser	0,19	0,9	0,19	0,32
Polip	0,06	1	0,06	0,11
Reinke Ödemi	0,05	1	0,05	0,09
Sağlıklı	0,91	0,60	0,91	0,73



Şekil 30. RF algoritmasının patolojilere göre sınıflamada elde ettiği karmaşıklık matrisi

BEŞİNCİ BÖLÜM BÖLÜM
SONUÇ VE ÖNERİLER

Yapılan çalışmada öncelikle ses kısıklığına neden olan reinke ödemli, larenjit, kanser ya da kanser öncesi lekelenme tanısı bulunan, polip, ses teli felçli ve sağlıklı hastalardan oluşan toplamda 652 kişiye ait 5865 ses verisi kullanılmıştır. Bu veriler üzerinde k-NN, NB, DT, SVM, RF olarak beş farklı makine öğrenme algoritmaları uygulanmıştır. Her bir algoritma için sağlıklı-hastalıklı ses sınıflandırması yapılmış ve sonuçları karşılaştırılmıştır. Elde edilen sonuçlara göre hastalıklı-sağlıklı ses ayrımında en yüksek doğruluk oranı 0,79 ile Destek Vektör Makinaları algoritmalarından elde edilmiştir.

Yaptığımız çalışmada ikinci olarak k-NN, NB, DT, SVM, RF algoritmaları ile reinke ödemli, larenjit, kanser ya da kanser öncesi lekelenme tanısı bulunan, polip, ses teli felçli ve sağlıklı şeklinde sınıflama yapılmıştır. Elde edilen sonuçlara göre algoritmaların doğruluk oranları aşağıdaki çizelgede verilmiştir.

Tablo 12

Algoritmaların sağlıklı-hasta sınıflamasında elde ettiği başarı ölçütleri

Hastalık	**k-NN**	**NB**	**DT**	**SVM**	**RF**
Doğruluk	0,78	0,64	0,67	0,79	0,76
Hassasiyet	0,81	0,65	0,71	0,82	0,76
Duyarlılık	0,78	0,73	0,66	0,78	0,80
F1	0,79	0,69	0,68	0,80	0,78

Tablo 13

Algoritmaların hastalıkları sınıflamadaki doğruluk oranları

Hastalık	k-NN	NB	DT	SVM	RF
Felç	0,65	0,49	0,48	0,68	0,57
Larenjit	0,55	0,19	0,15	0,45	0,21
Kanser	0,60	0,01	0,08	0,48	0,19
Polip	0,40	0,0	0,0	0,22	0,06
Reinke	0,42	0,01	0,05	0,30	0,05
Sağlıklı	0,76	0,70	0,78	0,84	0,91

Elde edilen sonuçlara göre hastalıkları sınıflamada en iyi sonucu k-NN algoritması vermiştir. Sınıflamada en başarısız algoritma NB algoritmasıdır. Ses verilerinin patolojilerine göre sınıflanmasında tüm algoritmalarda sağlıklı sınıfı en yüksek doğruluk oranı ile ayrılmıştır. Patolojilerine göre sınıflamada tüm algoritmalarda ikinci yüksek doğruluk oranı ses teli felci sınıfını ayırmada tespit edilmiştir. En başarısız sınıflanan ses verisi polip hastalarına aittir. Elde edilen sonuçlara göre doğruluk oranının veri setinde kullanılan ses verilerinin sayısı ile doğru orantılı olduğu gözlemlenmiştir. En yüksek doğruluk oranıyla sınıflanan ‘sağlıklı’ sesler 2682 adet ses verisiyle veri setinin en büyük sınıfını oluşturmaktadır. Benzer şekilde veri setinin en küçük sınıfı olan 243 ses verisiyle polip sınıfı çalışmada en düşük doğruluk oranıyla sınıflandırılmıştır.

Yapılan çalışma farklı makine öğrenme algoritmalarının ses verileri üzerindeki performansını sunmaktadır. Çalışma sağlıklı-hastalıklı ses verilerini ayırmada başarılı sonuçlar vermiştir.

Çalışmanın sınırlılıkları, hazır veri seti kullanmak ve veri setindeki patolojik sınıflardaki ses sayılarının eşit olmamasıdır. Bu nedenle sağlıklı hastalara ve ses teli felci olan hastalara ait ses verileri diğer hastalıklara göre sınıflamada üstünlük sağlamıştır.

Gelecekteki çalışmalarda veri seti genişletilip zenginleştirilerek algoritmaların tekrar çalıştırılması planlanmaktadır. Çalışmada öznitelik çıkarımı yalnızca MFCC ile yapılmıştır.

Gelecekteki çalışmalarda farklı öznitelik çıkarım yöntemlerinin sınıflama başarısına etkisi araştırılacaktır. Bunun yanında, mevcut sınıflandırma algoritmalarının performansını artırmak için hiperparametre optimizasyonu yapılması planlanmaktadır. Ayrıca gelecekteki çalışmalar için, veri setinin gerçek hastalardan, uygun ve yalıtılmış bir ortamda, etik izinleri alınarak oluşturulması planlanmaktadır. Son olarak ise literatürdeki fuzzy (bulanık) parametreli fuzzy esnek sınıflandırıcılar üzerinde araştırmalar gerçekleştirilip, çalışmanın bu açıdan analizleri yapılacaktır.

KAYNAKÇA

Akça F.M., (2020). Deep learning Türkiye, https://medium.com/deep-learning-turkiye/ (Erişim tarihi:15.04.2024).

AL-Dhief, F. T., Latiff, N. M. a. A., Malik, N. N. N. A., Sabri, N., Baki, M. M., Albadr, M. A. A., Abbas, A. F., Hussein, Y. M., ve Mohammed, M. A. (2020). Voice Pathology Detection Using Machine Learning Technique. 2020 IEEE 5th International Symposium on Telecommunication Technologies (ISTT). 09-11 Kasım, Malaysia.

Al-nefaie, A. H., Aldhyani, T. H., ve Koundal, D. (2024). Developing System-based Voice Features for Detecting Parkinson's Disease Using Machine Learning Algorithms. *Journal of Disability Research*, *3*(1), 20240001.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5-32.

Breiman, L. (2017). *Classification and Regression Trees*, Routledge.

Brownlee, J. (2020). D*ata Preparation for Machine Learning Data Cleaning, Feature Selection, and Data Transforms in Python*. Machine Learning Mastery.

Cattaruzza, M., Maisonneuve, P., ve Boyle, P. (1996). epidemiology of laryngeal cancer. *European Journal of Cancer Part B: Oral Oncology*, *32*(5), 293-305.

Charbuty, B., ve Abdulazeez, A. (2021). Classification Based on Decision Tree Algorithm For Machine Learning. *Journal of Applied Science and Technology Trends*, *2*(01), 20-28.

Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., Chen, K., Mitchell, R., Cano, I., ve Zhou, T. (2015). Xgboost: extreme gradient boosting. *R package version 0.4-2*, *1*(4), 1-4.

Cortes, C., ve Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*, 273-297.

Cover, T., ve Hart, P. (1967). Nearest Neighbor Pattern Classification. *IEEE Transactions on Information Theory*, *13*(1), 21-27.

Cunningham, P., Cord, M., ve Delany, S. J. (2008). Supervised Learning. In *Machine learning Techniques for Multimedia: Case Studies on Organization and Retrieval* (pp. 21-49). Berlin, Heidelberg: Springer Berlin Heidelberg.

Dankovičová, Z., Sovák, D., Drotár, P., ve Vokorokos, L. (2018). Machine learning approach to dysphonia detection. *Applied Sciences*, *8*(10), 1927.

Davis, S., ve Mermelstein, P. (1980). Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *28*(4), 357-366.

Deem, J. F., ve Miller, L. (2000). *Manual of Voice Therapy*. PRO-ED. https://books.google.com.tr/books?id=_osEAAAACAAJ.

Di Cesare, M. G., Perpetuini, D., Cardone, D., ve Merla, A. (2024). Assessment of Voice Disorders Using Machine Learning and Vocal Analysis of Voice Samples Recorded through Smartphones. *BioMedInformatics*, *4*(1), 549-565.

Donoho, D. L. (2006). Compressed sensing. *IEEE transactions on information theory*, *52*(4), 1289-1306.

Dudoit, S., Fridlyand, J., ve Speed, T. P. (2002). Comparison of discrimination methods for the classification of tumors using gene expression data. *Journal of the American statistical association*, *97*(457), 77-87.

El Naqa, I., ve Murphy, M. J. (2015). *What is machine learning?* (pp.3-11), Springer.

Eskidere, Ö., ve Ertaş, F. (2009). Mel frekansı kepstrum katsayılarındaki değişimlerin konuşmacı tanımaya etkisi. *Uludağ Üniversitesi Mühendislik Fakültesi Dergisi*, *14*(2).

Flach, P. A., ve Lachiche, N. (2004). Naive Bayesian classification of structured data. *Machine learning*, *57*, 233-269.

Gandhi, A. M., Shanmugan, S., Gorjian, S., Pruncu, C. I., Sivakumar, S., Elsheikh, A. H., Essa, F., Omara, Z., ve Panchal, H. (2021). Performance enhancement of stepped basin solar still based on OSELM with traversal tree for higher energy adaptive control. *Desalination*, *502*, 114926.

Ghahramani, Z. (2003). Unsupervised learning. In *Summer school on machine learning* (pp. 72-112). Springer.

Ghawi, R., Pfeffer, J. (2019). Efficient hyperparameter tuning with grid search for text categorization using kNN approach with BM25 similarity. *Open Computer Science*, *9*(1), 160-180.

Grömping, U. (2009). Variable importance assessment in regression: linear regression versus random forest. *The American Statistician*, *63*(4), 308-319.

Han, K., ve Wang, Y. (2021). A review of artificial neural network techniques for environmental issues prediction. *Journal of Thermal Analysis and Calorimetry*, *145*(4), 2191-2207.

Hershey, S., Chaudhuri, S., Ellis, D. P., Gemmeke, J. F., Jansen, A., Moore, R. C., Plakal, M., Platt, D., Saurous, R. A., ve Seybold, B. (2017). CNN architectures for large-scale audio classification. *2017 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, (pp. 131-135). IEEE.

Hu, H.-C., Chang, S.-Y., Wang, C.-H., Li, K.-J., Cho, H.-Y., Chen, Y.-T., Lu, C.-J., Tsai, T.-P., ve Lee, O. K.-S. (2021). Deep learning application for vocal fold disease prediction through voice recognition: preliminary development study. *Journal of medical Internet research*, *23*(6), e25247.

Jones, N. N., ve Song, S. A. (2023). Reinke Edema. In: StatPearls. StatPearls Publishing, Treasure Island (FL); 2023. PMID: 37603666.

Kamel, H., Abdulah, D., ve Al-Tuwaijari, J. M. (2019). Cancer classification using gaussian naive bayes algorithm. *2019 international engineering conference (IEC) (pp. 165-170)*. IEEE.

Karpov, E. (2003). Real-Time Speaker Identification. Master's Thesis. University of Joensuu, Department of Computer Science.

Keleş, N. (2022). Öksürük Ses Kayıtları Kullanılarak COVID-19 Tahmini, Yayımlanmamış Yüksek Lisans Tezi. Atatürk Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı, Erzurum.

Khaneja, N., Reiss, T., Kehlet, C., Schulte-Herbrüggen, T., ve Glaser, S. J. (2005). Optimal control of coupled spin dynamics: design of NMR pulse sequences by gradient ascent algorithms. *Journal of magnetic resonance*, *172*(2), 296-305.

Kibriya, A. M., Frank, E., Pfahringer, B., ve Holmes, G. (2005). Multinomial naive bayes for text categorization revisited. *AI 2004: Advances in Artificial Intelligence: 17th Australian Joint Conference on Artificial Intelligence, Cairns, Australia, December 4-6, 2004. Proceedings 17* (pp. 488-499). Springer Berlin Heidelberg.

Kleinsasser, O. (1982). Pathogenesis of vocal cord polyps. *Annals of Otology, Rhinology & Laryngology*, *91*(4), 378-381.

Küçük, R. G. (2019). Makine Öğrenmesi Yöntemleri Ile Parkinson Hastalığının Teşhis Edilmesi, Yüksek Lisans Tezi. Istanbul Aydın University, Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği Anabilim Dalı, Istanbul.

Lim, J., Oppenheim, A., ve Braida, L. (1978). Evaluation of an adaptive comb filtering method for enhancing speech degraded by white noise addition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *26*(4), 354-358.

Lundberg, S. M., Nair, B., Vavilala, M. S., Horibe, M., Eisses, M. J., Adams, T., Liston, D. E., Low, D. K.-W., Newman, S.-F., ve Kim, J. (2018). Explainable machine-learning predictions for the prevention of hypoxaemia during surgery. *Nature biomedical engineering*, *2*(10), 749-760.

Martins, R. H. G., Defaveri, J., Domingues, M. A. C., ve e Silva, R. d. A. (2011). Vocal polyps: clinical, morphological, and immunohistochemical aspects. *Journal of voice*, *25*(1), 98-106.

Mathieson, L. (2001). *Greene and Mathieson's the Voice and its Disorders*. John Wiley & Sons.

Matsui, T., ve Furui, S. (1995). Speaker recognition technology. *NTT review*, *7*(2), 40-48.

Memiş, S., Enginoğlu, S., ve Erkan, U. (2022). Fuzzy parameterized fuzzy soft k-nearest neighbor classifier. *Neurocomputing*, *500*, 351-378.

Menard, S. (2002). *Applied logistic regression analysis* (Vol. 106). Sage.

Milani, M. M., Ramashini, M., ve Krishani, M. (2020). A real-time application to detect human voice disorders. *2020 International Conference on Decision Aid Sciences and Application (DASA)* (pp.979-984). IEEE.

Obid, R., Redlich, M., ve Tomeh, C. (2019). The treatment of laryngeal cancer. *Oral and Maxillofacial Surgery Clinics*, *31*(1), 1-11.

Pham, M., Lin, J., ve Zhang, Y. (2018). Diagnosing voice disorder with machine learning. *2018 IEEE International Conference on Big Data (Big Data)* (pp. 5263-5266). IEEE.

Rabiner, L., Rosenberg, A., ve Levinson, S. (1978). Considerations in dynamic time warping algorithms for discrete word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *26*(6), 575-582.

Rabiner, L. R., ve Juang, B.-H. (1999). *Fundamentals of speech recognition*. Tsinghua University Press.

Rodrigues, É. O. (2018). Combining Minkowski and Chebyshev: New distance proposal and survey of distance metrics using k-nearest neighbours classifier. *Pattern Recognition Letters*, *110*, 66-71.

Schapire, R. E. (2013). Explaining adaboost. In *Empirical Inference: Festschrift in Honor of Vladimir N. Vapnik* (pp. 37-52). Springer.

Shen, P., Changjun, Z., ve Chen, X. (2011). Automatic speech emotion recognition using support vector machine. *Proceedings of 2011 international conference on electronic & mechanical engineering and information technology*. (Vol. 2, pp. 621-625). IEEE.

Singh, G., Kumar, B., Gaur, L., & Tyagi, A. (2019, April). Comparison between multinomial and Bernoulli naïve Bayes for text classification. In *2019 International conference on automation, computational and technology management (ICACTM)* (pp. 593-596). IEEE.

Steuer, C. E., El-Deiry, M., Parks, J. R., Higgins, K. A., ve Saba, N. F. (2017). An update on larynx cancer. *CA: a cancer journal for clinicians*, *67*(1), 31-50.

Sze, V., Chen, Y.-H., Yang, T.-J., ve Emer, J. S. (2017). Efficient processing of deep neural networks: A tutorial and survey. *Proceedings of the IEEE*, *105*(12), 2295-2329.

Şirin E., (2017). https://www.veribilimiokulu.com/support-vector-machine-svm-ile-siniflandirma-python-ornek-uygulamasi/ (Erişim tarihi:15.04.2024).

Terris, D. J., Arnstein, D. P., ve Nguyen, H. H. (1992). Contemporary evaluation of unilateral vocal cord paralysis. *Otolaryngology–Head and Neck Surgery*, *107*(1), 84-90.

Titze, I. R. (1994). *Principles of Voice Production*. Prentice Hall. https://books.google.com.tr/books?id=m48JAQAAMAAJ.

TC. Sağlık Bakanlığı, (2016). 'Erken Teşhis Hayat Kurtarır'. https://www.saglik.gov.tr/TR,19826/erken-teshis-hayat-kurtarir.html (Erişim tarihi:15.04.2024).

Tucker, H. M. (1980). Vocal cord paralysis—1979: etiology and management. *The Laryngoscope*, *90*(4), 585-590.

Vapnik, V. (2013). *The nature of statistical learning theory*. Springer science & business media.

Verdolini, K., ve Ramig, L. O. (2001). Occupational risks for voice problems. *Logopedics Phoniatrics Vocology*, *26*(1), 37-46.

Verma, V., Benjwal, A., Chhabra, A., Singh, S. K., Kumar, S., Gupta, B. B., Arya, V., ve Chui, K. T. (2023). A novel hybrid model integrating MFCC and acoustic parameters for voice disorder detection. *Scientific Reports*, *13*(1), 22719.

Wang, W., Chen, Z., Mu, J., ve Han, T. (2014). Throat polyp detection based on compressed big data of voice with support vector machine algorithm. *Eurasip Journal on Advances in Signal Processing*, *2014*, 1-6.

Williams, C., ve Rasmussen, C. (1995). Gaussian processes for regression. *Advances in neural information processing systems*, *8*.

William, J. B., Manfred, P. (2007). Saarbrucken Voice Database: Institute of Phonetics, University of Saarland. http://www.stimmdatenbank.coli.uni-saarland.de/ (Erişim Tarihi: 13 Mayıs 2024).

Wu, H., Soraghan, J., Lowit, A., ve Di Caterina, G. (2018, July). Convolutional neural networks for pathological voice detection. In *2018 40th annual international conference of the ieee engineering in medicine and biology society (EMBC)* (pp.1-4). IEEE.

www.MedicalGraphics.de, license (CC BY-ND 4.0 EN), https://www.medicalgraphics.de/en/product/vocal-cords-open/ (Erişim tarihi:15.04.2024).

Yu, Y., Si, X., Hu, C., ve Zhang, J. (2019). A review of recurrent neural networks: LSTM cells and network architectures. *Neural computation*, *31*(7), 1235-1270.