



Makine Öğrenimi Yaklaşımı ile İşitme Kaybı Derecesinin Cinsiyet ve Yaşa Bağlı Tahmin ve Analizi

Yazılım Mühendisliği Ana Bilim Dalı

Dönem Projesi

Sena Gaga

Proje Danışmanı: Dr.Öğr.Üyesi Mansur Alp Toçoğlu

Ocak 2024

Makine Öğrenimi Yaklaşımı ile İşitme Kaybı Derecesinin Cinsiyet ve Yaşa Bağlı Tahmin ve Analizi

Öz

Bu çalışma, işitme kaybı sağlık sorununu ele alarak ve Makine Öğrenimi yöntemlerini kullanarak cinsiyet ve yaş gibi demografik faktörlere bağlı olarak işitme kaybı derecesinin tahmin edilmesine odaklanmıştır. İlgili veri seti, işitme cihazı merkezine işitme kaybı şikayetiyle başvuran 500 hastanın özel veya kamu hastanelerinde gerçekleştirilen işitme testlerinin sonuçlarına dayanmaktadır. Bu veriler, hastaların sağ ve sol kulak hava yolu işitme eşikleri, cinsiyet, yaş ve işitme kaybı ortalamalarını içermektedir.

Çalışmada kişisel bilgilere yer verilmemekle birlikte, işitme kaybı dereceleri cinsiyet ve yaş faktörlerine bağlı olarak analiz edilecektir. İlgili analizde Makine Öğrenimi algoritmalarından Destek Vektör Makineleri (SVM), K-En Yakın Komşu, Lineer Regresyon ve Random Forest (Rastgele Orman) denetimli öğrenme makine algoritmaları kullanılacaktır.

Projenin temel amacı, bu veri seti üzerinde farklı makine öğrenme algoritmaları kullanarak sağ ve sol kulak işitme kaybı ortalamalarını tahmin etmek ve en iyi performans gösteren algoritmayı belirlemektir. Belirlenen öğrenme modeli, hastaların cinsiyet ve yaş değerlerine dayalı olarak sağ ve sol kulak işitme kaybı ortalamalarını tahmin etmek için kullanılacaktır.

En başarılı sonuç Destek Vektör Makineleri algoritmasıyla elde edilmiştir. İlgili algoritmalarda daha yüksek başarı elde edilebilmesi için hiperparametre optimizasyonu uygulanmıştır.

Bu alıřma, iřitme kaybının olası tahminlenmesinin saęlanması aısından iřitme cihazı uygulama merkezlerinde riskli olabilecek bireylerin ngrlmesini sunarak iřitme saęlıęı alanında olumlu katkılar saęlayarak bireylerin kulak saęlıklarını belli periyotlarda iřitme testi yaptırarak kontrol ettirmesi amacıyla teřvik saęlabilir.

Anahtar Szckler: İřitme Saęlıęı, Destek Vektr Makineleri (SVM), Karar Aęaları (Decision Trees), Lojistik Regresyon, Makine ęrenimi, İřitme Kaybı.

Prediction and Analysis of Hearing Loss Degree Based on Gender and Age Using Machine Learning Approach

Abstract

This study focuses on predicting the degree of hearing loss based on demographic factors such as gender and age, addressing the health issue of hearing loss and utilizing Machine Learning methods. The relevant dataset is derived from the results of hearing tests conducted in private or public hospitals for 500 patients who applied to the hearing aid center with complaints of hearing loss. The data includes air conduction hearing thresholds for the right and left ears, gender, age, and hearing loss averages for the patients.

While personal information is not disclosed in the study, the degrees of hearing loss will be analyzed based on gender and age factors. Support Vector Machines (SVM), K-Nearest Neighbors, Linear Regression, and Random Forest supervised learning machine algorithms will be employed in the analysis.

The primary objective of the project is to predict the averages of hearing loss in the right and left ears using different machine learning algorithms on this dataset and determine the algorithm that performs the best. The identified learning model will be used to predict the averages of hearing loss in the right and left ears based on the patients' gender and age values.

The most successful result was achieved with the Support Vector Machines algorithm. Hyperparameter optimization has been applied to these algorithms to achieve higher success.

This study may encourage individuals to monitor their ear health by periodically undergoing hearing tests, providing positive contributions to hearing health by predicting individuals who may be at risk in hearing aid application centers.

Keywords: Hearing Health, Support Vector Machines (SVM), Decision Trees, Logistic Regression, Machine Learning, Hearing Loss.

Teşekkür

Proje çalışmasına katkılarından dolayı Dr. Öğr. Üyesi Mansur Alp Toçoğlu'na teşekkür ederim.

İçindekiler

Öz	i
Abstract	ii
Teşekkür	iv
Şekiller Listesi	vii
Tablolar Listesi	viii
Kısaltmalar Listesi	ix
Semboller Listesi	x
1 Giriş	1
2 Genel Bilgiler	5
2.1 İşitme	5
2.1.1 İşitme Kaybının Değerlendirilmesi	6
2.2 Makine Öğrenimi	8
2.2.1 Odyolojide Makine Öğrenimi Literatür Özeti	9
3 Gereç ve Yöntem	11
3.1 Veri Seti	11
3.1.1 Veri Ön İşlem Süreci	11
3.1.2 Verinin Görselleştirilmesi	13
3.1.3 Öznitelik Seçimi	21
3.1.4 Veri Setinin Eğitim ve Test Veri Seti Olarak Ayrımı	22
3.1.5 Çoklu Regresyon Modeli	24
3.2 Modelde Kullanılan Makine Öğrenme Algoritmaları	24
3.2.1 Destek Vektör Makineleri	24

3.2.2	Lineer Regresyon	25
3.2.3	Rastgele Ormanlar	26
3.2.4	K-En Yakın Komşu	27
3.3	İşitme Kaybı Tahmin Penceresi	28
4	Bulgular	30
5	Tartışma ve Sonuç	32
	Kaynaklar	36

Şekiller Listesi

Şekil 2.1	Periferik işitme sistemi	5
Şekil 3.1	Odyometri test sonuçları veri setini Python Jupyter görüntüsü	12
Şekil 3.2	Label Encoding Uygulanması	13
Şekil 3.3	Yaş Dağılımının Çubuk Grafiği ile Görselleştirilmesi	14
Şekil 3.4	Cinsiyet Dağılımının Pasta Grafiği ile Görselleştirilmesi	15
Şekil 3.5	Erkeklerin Sağ ve Sol Kulak SSO Histogram Grafiği ile Gösterimi	15
Şekil 3.6	Kadınların Sağ ve Sol Kulak SSO Histogram Grafiği ile Gösterimi	16
Şekil 3.7	Saf ses Ortalamalarının Çizgi Grafiği ile Karşılaştırılması	17
Şekil 3.8	Çok Hafif Dereceli İşitme Kaybının Yaş ve Cinsiyet Histogram Grafiği Gösterimi.....	17
Şekil 3.9	Hafif Dereceli İşitme Kaybının Yaş ve Cinsiyet Histogram Grafiği Gösterimi.....	18
Şekil 3.10	Orta Dereceli İşitme Kaybının Yaş ve Cinsiyet Histogram Grafiği Gösterimi.....	19
Şekil 3.11	Orta-İleri Dereceli İşitme Kaybının Yaş ve Cinsiyet Histogram Grafiği Gösterimi.....	19
Şekil 3.12	İleri Dereceli İşitme Kaybının Yaş ve Cinsiyet Histogram Grafiği Gösterimi.....	20
Şekil 3.13	Çok İleri Dereceli İşitme Kaybının Yaş ve Cinsiyet Histogram Grafiği Gösterimi.....	20
Şekil 3.14	Korelasyon Matrisi	22
Şekil 3.15	Veri Setinin Modellenmesi	23
Şekil 3.16	Destek Vektör Makine Regresyon Modeli (SVR).....	25
Şekil 3.17	Lineer Regresyon Modeli.....	25
Şekil 3.18	K-En Yakın Komşu Modeli	27
Şekil 3.19	Rastgele Orman Modeli	28
Şekil 3.20	İşitme Kaybı Derece Tahmin Pencere Görünümü.....	29
Şekil 3.21	İşitme Kaybı Derece Tahmin Sonuç Görüntüsü.....	29

Tablolar Listesi

Tablo 2.1	Modifiye Goodman skalasına göre işitme kaybının derecelerine göre sınıflandırılması.....	10
-----------	--	----

Kısaltmalar Listesi

Adaptive Boosting	Adaboost
Desibel	dB
Desibel Hearing Level	dB HL
Grafiksel Kullanıcı Arayüzü	GUI
Hertz	Hz
K-En Yakın Komşu	KNN
Sol Kulak	L
Çok Katmanlı Algılayıcı	MLP
Ortalama Kare Hatası	MSE
R-Kare Skoru	R^2
Saf Ses Ortalaması	SSO
Destek Vektör Makineleri	SVM
Tedirgin Edici Ses Seviyesi	UCL
Dünya Sağlık Örgütü	WHO

Bölüm 1

Giriş

İşitme kaybı, dünya genelinde milyonlarca insanı etkileyen önemli bir sağlık sorunudur. Dünya nüfusunun yüzde 5'inden fazlası (466 milyon) işitme kaybından etkilenmektedir (432 milyon yetişkin, 34 milyon çocuk). 2050 yılına kadar 900 milyondan fazla insanın, yani on kişiden birinin işitme kaybı yaşayacağı tahmin edilmektedir [1].

İşitme kaybının sonuçları sıklıkla hafife alınmaktadır. Dünya Sağlık Örgütü (WHO) bu durumu “sessiz salgın” olarak adlandırmaktadır[2]. Şu anda dünya çapında 350 milyon kişi bir tür işitme kaybıyla yaşamaktadır [3].

İşitme kaybının çeşitli gruplar üzerinde farklı etkileri bulunmaktadır. Pediatrik grupta işitme kaybı dil edinimini zorlaştırırken , yetişkin grupta depresyon ve sosyal izolasyon problemleri ortaya çıkmaktadır. Yaşlı yetişkinlerde bazı çalışmalar işitme kaybı ile bilişsel işlevlerde azalma arasında ilişkili olduğunu göstermektedir [4].

İşitme kaybı, bireylerin yaşam kalitesini önemli ölçüde etkilemekte ve bir dizi sosyal ve ekonomik sorunu beraberinde getirmektedir. İşitme kaybının etkili bir şekilde teşhis edilmesi, sınıflandırılması ve tedavi edilmesi, bu alandaki araştırmaların odak noktasını oluşturmuştur.

İşitme kaybının teşhis edilmesi için uygulanan en temel prosodür saf ses odyometri testidir.

Saf ses odyometri testi , odyometri cihazıyla sessiz bir kabinde işitme duyarlılığını ve işitme eşiklerini belirlemek amacıyla 125 ile 8000 Hertz (Hz) frekans aralığında saf tonların ilgili kişiye dinletilmesi ve ilgili kişi tarafından 125-8000 Hz frekanslarında

duyulan seslerin minimum eşik noktasının belirlenmesiyle gerçekleştirilen subjektif bir testtir. Test mümkün olduğunca sessiz ve havalandırma aracının olmadığı yalıtılmış ses kabinlerinde gerçekleştirilmektedir. Test sırasında kişiden alınan cevaplar Odyograma kaydedilmektedir. Odyogram yatay eksen de 125 ile 8000 Hertz frekans bilgilerinin yer aldığı dikey eksen de ise ilgili her bir frekansın ses şiddet seviyelerinin -10 ile 120 desibel (dB) aralığında gösterildiği saf ses odyometri testi esnasında kişiden alınan yanıtlara göre işitme eşiklerinin işaretlendiği bir grafikdir.

İşitme eşiği, kişinin minimum şiddette duyabildiği en son işitilebilen son noktanın dB cinsinden karşılığıdır. Odyometri testi -10 ile 120 dB aralığını kapsamaktadır. Sağ ve sol kulak için toplam 16 frekansta işitme eşikleri belirlendikten sonra saf ses ortalaması hesaplanmaktadır. Saf ses ortalaması her bir kulak için ayrı ayrı 500 Hz, 1000Hz, 2000Hz ve 4000 Hz frekanslardaki işitme eşiklerinin ortalamasıdır. Böylece sağ ve sol kulak saf ses ortalama değerleri dB cinsinden elde edilmektedir. Odyometri testinde her iki kulak için saf ses ortalama değerleri hesaplandıktan sonra konuşmayı anlama eşiği (dB), konuşmayı ayırtma skoru (%) ve tedirgin edici ses seviyesi (UCL) belirleme testleri sırasıyla uygulanmaktadır. Alan uzmanları tarafından işitme kaybının derecelendirilmesi yapılmakta olup gerekli görülmesi dahilinde ileri tetkik test istemleri yapılmaktadır.

İşitme kaybı derecelendirilmesinde günümüzde en yaygın sınıflandırma yöntemi olarak 1965'te Goodman tarafından oluşturulup daha sonra Clark tarafından 1981'de modifiye edilen sınıflandırma kullanılmaktadır.

Yaşın ilerlemesiyle birlikte tüm vücut sisteminde olduğu gibi işitme sisteminde de çeşitli değişiklikler meydana gelmektedir.

İlerleyen yaşla birlikte tüm işitme sistemi de değişime uğramaktadır. Kulak kepçesinde büyüme, dış kulak yolunda katılaşmış kulak kiri, özellikle erkeklerde kılınmada artış, kulak kanal derisinde atrofi, elastikiyet kaybı, yüksek frekanslarda hava yolu iletim özelliklerini bozacak şekilde kıkırdak büyümesi gibi yapısal ve fonksiyonel değişiklikler ileri yaşlarda dış kulakta en çok görülen sorunlardır [5]. Bununla birlikte orta kulak kemikçiklerinde ve kaslarda sertleşme, kulak zarında kalsifikasyon durumları klinikte de sıkça karşılaşılan durumları içermektedir.

Yaşlanmaya bağlı olarak zaman içinde gelişen işitme kaybı presbiakuzi olarak tanımlanır. Presbiakuzi, yaşlılık döneminde romatoid artrit ve hipertansiyondan sonra en sık görülen üçüncü önemli kronik rahatsızlıktır. Yaşa bağlı işitme kayıplarının en önemli özelliği, özellikle saf ses işitme eşiklerinde düşüş ve konuşmayı anlamama sorunlarıyla karakterizedir. İşitme hassasiyeti yaşla doğru orantılı olarak artmaktadır [6].

Yüksek frekanslarda (tiz ses bölgesi) görülen işitmede azalma sorunu, erkeklerde kadınlara göre daha fazladır. İşitme kaybı, yaşlı erkeklerde daha çok hafif derece ile orta-ileri derece arasında değişen bir ortalama ile bilateral ve ani düşüş gösteren bir kayıp şeklindeyken; yaşlı kadınlarda daha çok hafif-orta derecede, yavaş bir eğimi olan (ani düşüş göstermeyen) bilateral, simetrik sensorinöral işitme kaybı ile karakterizedir [7]. Yapılan çalışmalardan da hareketle işitme kaybı derecesinin yaş ve cinsiyet ile korelasyon gösterdiği görülmektedir.

Kadınlarda bulunan östrojen hormonunun kadınları kalp rahatsızlıkları olmak üzere kronik birçok hastalıktan koruduğu ve işitme kaybının da bu hormon etkisiyle kadınlarda erkeklere göre daha az sıklıkla yada daha geç başlangıçlı olabileceği düşünülmektedir. Kılıçdağ ve arkadaşlarının 2004 yılında menapoza giren kadınlar üzerinde yaptıkları çalışmada Östrojen tedavisinin, yaşlanan menopoz sonrası kadınlarda işitme kaybını yavaşlatabileceğine dair bulgular elde edilmiştir [8].

İşitme kaybı ve makine ve derin öğrenme alanında yapılmış birçok çalışma bulunmaktadır. Bu çalışmalar genellikle işitme kaybının teşhisinde, sınıflandırılmasında, tedavi yöntemlerinde ve işitme cihazları optimizasyonunda makine öğrenimi tekniklerinin kullanılmasını içermektedir.

Bu çalışmada ; işitme cihazı merkezine işitme kaybı şikayetiyle başvuran 500 işitme kayıplı hastaların işitme test sonuçlarından yola çıkılarak oluşturulan veri seti üzerinde kişilerin yaş,cinsiyet ve sağ ile sol kulaklarındaki 500 Hz , 1000 Hz , 2000 Hz ve 4000 Hz frekanslarındaki işitme eşik değerlerinin çıktıklarıyla denetimli makine öğrenme yöntemi kullanılarak Destek Vektör Makineleri (SVM), Karar Ağaçları (Decision Trees), Lojistik Regresyon ve Random Forest (Rastgele Orman) makine öğrenimi algoritmalarıyla eğitilen ve değerlendirme metrikleri olan ; Ortalama Kare Hatası (MSE), Kök Ortalama Kare Hatası (RMSE)ve R-Kare skoru çıktılarından en

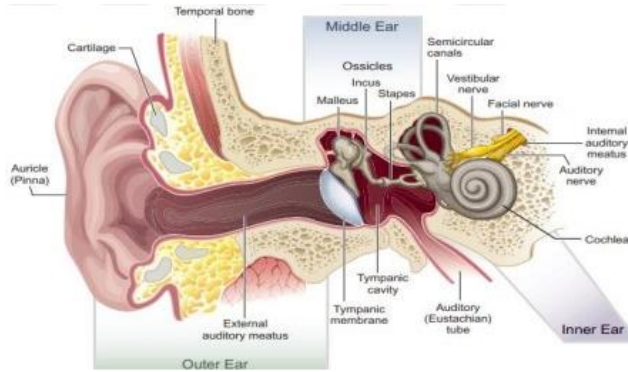
performanslı görülen destek vektör makineleri algoritmasını kullanarak sağ ve sol kulak havayolu işitme kaybı derecesini tahmin etmek amaçlanmaktadır. Özellikle işitme cihazı merkezleri için düşünülen model de kişiden alınan yaş ve cinsiyet bilgisine göre kişide olabilecek muhtemel işitme kaybı derecesi tahmin edilmek istenmektedir.

Bölüm 2

Genel Bilgiler

2.1 İşitme

İşitme , dış kulak kepçesinin ses uyarılarını toplayıp dış kulak yoluna ilemesiyle başlayan daha sonra ses dalgalarının orta kulağa geçişiyle birlikte hava ortamdan sıvı ortama geçen ses dalgalarının amplifiye edilemesi için bir dizi işlem gerçekleştiren orta kulağın ; sesi iç kulakta bulunan işitme sinirine ilemesiyle devam eden bir dizi işlemde meydana gelmektedir.İşitme sinirine ulaşan ses dalgaları sinirsel iletim ile birlikte korteks seviyesinde anlamlandırılmaktadır. İşitme bir dizi karmaşık süreçten meydana gelmektedir.İşitme sisteminin anatomik yapısı Şekil 2.1’de gösterilmiştir.



Şekil 2.1: Periferik işitme sistemi

(Kaynak: Hixon vd., 2022)

İşitme duyusunu sağlayan dış, orta, iç kulak yapıları ile işitme yollarının ve işitme korteksinin herhangi bir inde oluşan patolojiler işitme kaybına neden olur [7].

Bu yapıların hasarlanması birçok faktöre bağlıdır. Yaşlanma, ototoksik ilaç kullanımı, gürültüye maruziyet, kronik hastalıkların varlığı, tinnitus, genetik faktörler olmak üzere birçok sebep işitme kaybının oluşmasına etki etmektedir.

Oluşan işitme kaybı temel olarak odyometri testi ile teşhis edilebilmektedir. İşitme kaybı derecelendirilmekte ve tiplerine göre sınıflandırılmaktadır.

İşitme kaybı tipi belirlenirken işitme kaybının odyogram konfigürasyonu oldukça önemlidir. İşitme kaybı üç tipten meydana gelmektedir. Bunlar Sensörinöral tip, İletim tip ve Mikst tip işitme kayıpları olarak adlandırılmaktadır.

2.1.1 İşitmenin Kaybının Değerlendirilmesi

İşitme kaybının değerlendirilmesi sürecinde iki yöntem kullanılmaktadır. Bunlardan biri olan subjektif yöntemde kişinin aktif katılımıyla yapılan cevapların kişiden alındığı test yöntemleridir. Diğer bir yöntem ise kişinin teste birebir katılmadığı test anında pasif olduğu ve işitsel cevapların korteksten alındığı objektif yöntemlerdir.

İşitme kaybı değerlendirmesinde kullanılan en temel ve birincil yöntem saf ses odyometri testidir. Bu test subjektif bir yöntem olmakla birlikte kişinin aktif katılımını gerektirmektedir. Bu çalışmada saf ses odyometri testinden elde edilen yanıtlarından oluşturulan veri seti kullanılmıştır.

Saf ses odyometri testinde standart odyometre cihazları kullanılmaktadır. Bu cihazlar 125 ile 8000 Hz arasında ölçüm yapmaktadır.

Odyometri testi ses yalıtımlı sessiz kabinlerde havayolu kulaklık kullanılarak gerçekleştirilmektedir. Kişiye bir buton verilerek duyduğu en küçük şiddetteki ses dahil olmak üzere duyduğu seslere yanıt vermesi istenmektedir. Her iki kulak -10 ile 120 dB ses şiddet aralığında 125 ile 8000 Hz frekans bandında çeşitli saf ton ses sinyalleri gönderilerek ölçülmektedir. Modified Hughson-Westlake method kullanılarak işitme eşikleri tespit edilir. Bu yöntem kişinin bir frekanstaki ses şiddet seviyesini duyduğu zaman aynı frekans üzerinde ses şiddet seviyesini 10dB azaltmayı, eğer ilgili frekanstaki ses şiddet seviyesini duyamadıysa aynı frekanstaki ses şiddet seviyesini 5dB artırmayı içeren yöntemler dizisidir. İşitme eşiği belirlemesi bu yöntem ile gerçekleştirilmiştir.

Test sırasında elde edilen işitme eşikleri klinisyen tarafından odyogram adı verilen grafiklere kaydedilmektedir.

İşitme kaybının derecesi saf ses ortalamasına göre yapılmaktadır (WHO, 2009). Her bir frekans için işitme kaybı miktarı decibel Hearing Level (dB HL) cinsinden belirlenir. Dünya Sağlık Örgütü (WHO) 500-4000 Hz. ortalamasının alınmasını önermektedir. Her iki kulak için saf ses ortalamaları bu yöntem ile hesaplanmaktadır.

İşitme kaybının derecesini tanımlayan en yaygın sınıflandırma 1965'te Goodman tarafından oluşturulmuştur ve daha sonra Clark tarafından 1981'de modifiye edilmiştir. Goodman (1965), konuşma frekanslarındaki eşiklerin ortalaması olan saf ses ortalamasının (SSO) işitme kaybının derecesi olarak kullanımını geliştirmiştir ve günümüzde de yaygın olarak kullanılmaktadır (Roup, 2016). İşitme kaybının, derecelerine göre sınıflandırılması Tablo 2.1.1'de gösterilmiştir.

Tablo 2.1: Modifiye Goodman skalasına göre işitme kaybının derecelerine göre sınıflandırılması

Saf Ses Ortalaması	İşitme Kaybı Derecesi
-10-15	Normal işitme
16-25	Çok Hafif Derecede İşitme Kaybı
26-40	Hafif Derecede İşitme Kaybı
41-55	Orta Derecede İşitme Kaybı
55-70	Orta-İleri Derecede İşitme Kaybı
71-90	İleri Derecede İşitme Kaybı
90+	Çok İleri Derecede İşitme Kaybı

(Kaynak: Roup, 2016)

Saf ses ortalaması -10 ile 15 dB aralığı içerisinde ise bu durum Tablo 1 'de de görüldüğü üzere normal işitme olarak adlandırılmaktadır. Saf ses ortalama değeri 16 dB ve daha üzeri seviyelerde olduğu zaman işitme kaybı durumu gözlenmeye başlamaktadır.

İşitme kaybı farklı derecelerde ortaya çıkabilmektedir. Ancak derecesi ne olursa olsun iletişim problemleri , sosyal hayattan kopma , pediatrik grupta konuşma ve dil geriliği problemleri , özgüven problemleri , yaşlı bireyler de demans gelişimine zemin hazırlama veya var olan bilişsel bozuklukta daha da kötüleşmeye neden olmaktadır.

2.2 Makine Öğrenimi

Algoritma, belirli bir problemi çözmek veya belirli bir görevi gerçekleştirmek için adım adım talimatlar içeren mantıksal sıralamadır. Makine öğrenme algoritmaları veri kümesini bir modele dönüştürerek girdilerden tahminler yapmaktadır.

Makine öğrenmesi, tanımlanan ve çözümü çıktı olarak vermesi amaç edinilen sorunu veri kümesindeki bilgilere göre tasarlanıp eğitilen model ile tarayarak çözüme yönelik bilgileri çıktı olarak sunan yapay zekanın alt birimidir . Makine öğrenmesi metodu, veri kümesindeki örüntüleri tespit ederek, akılcı karar verme için istatistik, mantık ve hesaplamayı birlikte kullanan yapay zekâ disiplinler bütününe bağlı olarak çözümü istenen konuda eğitilip modellenen algoritmaya göre çözümü çıktı olarak sunar (McDonald, 2020). Genel olarak üç tür makine öğrenimi vardır:Denetimli öğrenme , Denetimsiz öğrenme ve Pekiştirmeli öğrenme.

Denetimli öğrenme, eğitim veri setindeki giriş ve çıkışları kullanarak bir modelin öğrenmesini sağlar. Model, yeni verilerle test edildiğinde doğru çıktıları üretmeyi amaçlamaktadır.Bu çalışma da denetimli öğrenme modeli kullanılmıştır.Sınıflandırma ve regresyon problemleri denetimli öğrenme sınıfına girmektedir.

Denetimsiz öğrenme algoritmaları veri seti içindeki desenleri ve ilişkileri keşfetmeye çalışarak bir öğrenme gerçekleştirmektedir. Pekiştirmeli öğrenme ise eğitim veri olmadan yapılan öğrenme türüdür. Makine öğrenmesi günlük hayatın birçok farklı noktalarında kullanılmaktadır.Film,müzik öneri sistemleri,fiyat tahmin ve analiz sistemleri , otonom araç sistemleri ,tıp alanında sağlık hizmetlerinin geliştirilmesi , görüntüleme yöntemleri üzerinden hastalık teşhisi gibi farklı disiplinlerde kullanılmaktadır.

Bir problemin makine öğrenmesi ile çözülmesi için temel izlenen makine öğrenme aşamaları bulunmaktadır.Bunlar ;

- 1) Veri seti kullanılarak çözümlenebilen, çözümlenmek istenen problem tanımlanmalıdır.
- 2) Çözülmemek istenen problem ile ilgili veri toplanıp, veri seti oluşturulmalıdır.

- 3) Toplanan veri setinin yapısının makine öğrenme algoritmalarına uygun hale getirilebilmesi, ilgili verinin temizlenmesi aşamalarını içeren veri ön işleme aşaması gerçekleştirilmelidir.
- 4) Veri setinin bir kısmı eğitim veri seti kalan kısmı test veri seti olacak şekilde ayrılarak ilgili algoritmaların gerçekleştirimi sağlanmalıdır.
- 5) Eğitim veri seti ile model oluşturulması, model oluşturulduktan sonra test veri setinin kullanılarak modelin değerlendirilmesi yapılmalıdır.
- 6) Oluşturulan modelin performansının değerlendirilmesi için modele uygun değerlendirme metrikleri kullanılmalıdır.
- 7) Model performansını artırmak için otomatik hiperparametre optimizasyon araçları kullanılabiliyor veya deneme-yanılma yöntemleriyle hiperparametre optimizasyonu gerçekleştirilebilir.
- 8) Elde edilen sonuçların probleme yanıt vermesi, istenilen düzeyde olması durumunda uygulama çeşitli yöntemler ile canlı ortama taşınabilmektedir.

Makine öğrenmesi süreci genellikle bu temel aşamalardan oluşur, Bu adımların her biri, proje gereksinimlerine ve veri setinin özelliklerine bağlı olarak özelleştirilebilmektedir.

2.2.1 Odyolojide Makine öğrenmesi Literatür Özeti

Zhao Y. ve arkadaşları tarafından 2018 yılında “Endüstriyel Gürültüye Maruz Kalan Çalışanlarda İşitme Kaybının Tahminine Yönelik Makine Öğrenimi Modelleri” adlı çalışma yapılmıştır. Destek vektör Makineleri, Random forest, Adaptive Boosting (Adaboost) ve Yapay sinir ağı mimarilerinden biri olan çok katmanlı algılayıcı (MLP) algoritmaları kullanılmıştır. Bu dört makine öğrenimi algoritması uygulanarak gürültüye bağlı işitme kaybını tahmin etmek için hem sınıflandırma hem de regresyon modellerinin geliştirildiği çalışmada dört modelinde performansları incelendiğinin de gürültüye maruz kalan işçilerde gürültüye bağlı işitme kaybının değerlendirmesi ve tahmin edilmesi için potansiyel araçlar olduğunu göstermiştir. Çalışma da gürültü maruziyetinden kaynaklanan işitme kaybını tahmin etmede en etkili aracın, en yüksek puanı elde eden Destek Vektör Makine modeli olduğu gösterilmiştir [9].

Park K.V. ve arkadaşları tarafından 2020 yılında yapılan “Tek Taraflı İdiyopatik Ani Sensörinöral İşitme Kaybında İşitme Prognozunu Tahmin Etmek İçin Makine Öğrenimi Modelleri”adlı başka bir çalışmada da K-En yakın komşu (KNN), Destek Vektör Makinaları, Rastgele Ormanlar, AdaBoost, MLP, çok katmanlı algılayıcı modeller kullanılmış olup destek vektör makine (SVM) modelinin en yüksek doğruluk (%75,36) ve en yüksek F-puanını (0,74) elde ettiğini göstermiştir [10].

Yeni Doğanlarda, Bebeklerde ve Çocuklarda Makine Öğrenimi ile İşitme Kaybı Tahmini” adlı makale, 2022 yılında Venkataramana Pai K. ve arkadaşları tarafından yayınlanmıştır. Makale, işitme kaybını yeni doğanlarda, bebeklerde ve çocuklarda tahmin etmeye odaklanmaktadır. İlk olarak, bir odyolog ile işbirliği içinde odaklanılan nüfus için veri oluşturulmuştur. Ardından, işitme kaybını belirlemek için tahmin modelleri oluşturmak için bu verilere sınıflandırma algoritmaları uygulanmıştır. Naïve Bayes, Destek Vektör Makineleri, XGBoost ve Rastgele Orman, sınıflandırmada kullanılan algoritmalar arasındadır. İki veri seti oluşturulur: birincisi tüm sınıfların eşit sayıda kayda sahip olduğu (dengeli), diğeri ise kayıp prevalansını ve gürültüyü (dengesiz) göz önünde bulunduran bir veri setidir. Destek Vektör Makineleri, dengeli veri seti için %100 maksimum doğruluk elde ederken, dengesiz veri seti için %94,10'luk bir doğruluk elde etmiştir [11].

Hantai Kim ve ark. tarafından 2021 yılında “Konuşmayı Ayır etme Skorlarını Saf Ton Eşikleri Üzerinden Tahmin Etmek: 12,697 Katılımcıdan Alınan Veri Tabanlı Bir Makine Öğrenimi Yaklaşımı” adlı çalışmada Suwon'daki Ajou Üniversite Hastanesi'nde 12,697 denekten elde edilen büyük bir veri kümesi üzerinde çalışılmıştır . Destek Vektör Makineleri (SVM'ler), Çok Katmanlı Algılayıcılar (MLP'ler) ve Rastgele Orman Modelleri gibi yaygın olarak kullanılan makine öğrenimi modellerinin performansı değerlendirilmiştir. Rastgele Orman tabanlı yaklaşımın SVM ve MLP modellerine kıyasla daha yüksek bir performans sergilediğini göstermektedir. Hiperparametre optimizasyonunun hassaslığı, SVM ve MLP modellerinin performansını etkileyebilirken, Rastgele Orman tabanlı model %95'in üzerinde doğrulukla klinik açıdan anlamlı konuşmayı ayır etme skor tahmininde başarılı olmuştur [12].

Çalışmalardan hareketle işitme kaybının etkilediği birçok farklı alanda tahminleme amacıyla çeşitli makine öğrenimi ve derin öğrenme algoritmalarından faydalandığını

görmekteyiz. Bu algoritmaların seçimi, veri setinin özelliklerine, problem bağlamına ve performans kriterlerine bağlı olarak değişmektedir.

Bu çalışmanın amacı ise, işitme cihazı merkezine başvuran 500 işitme kayıplı bireylerin sonuçlarından oluşturulan veri seti üzerinden, yaş, cinsiyet , sağ ve sol kulaklardaki 500 Hz , 1000 Hz , 2000 Hz ve 4000 Hz frekanslardaki işitme eşik değerlerini kullanarak makine öğrenimi ile bireylerin sağ ve sol kulaklarındaki havayolu işitme kaybı derecesini tahmin etmeyi amaçlamaktadır.

Bölüm 3

Gereç ve Yöntem

Bu çalışma, işitme cihazı merkezine başvuran işitme kayıplı kişilerin test sonuçlarından oluşturulan odyometri test sonuçları adlı veri setinin makine öğrenimi algoritmalarıyla denetimli öğrenme yoluyla eğitilip en yüksek performans gösteren algoritmanın seçilerek işitme cihazı merkezlerine başvuran kişilerde hızlı ön değerlendirme yapabilmek için yaş ve cinsiyet demografik özellikleri üzerinden sağ ve sol kulak havayolu işitme kaybı derecelerinin tahminlenmesini sağlayan “İşitme Kaybı Tahmin Penceresi” oluşturmayı amaçlamıştır.

3.1 Veri Seti

500 adet işitme kayıplı bireyin saf ses odyometre testi sonucu veri setine dahil edilmiştir. Excel dosyasında hazırlanmış olan veri seti Python programı için csv formatına çevrilerek programda kullanılacak hale getirilmiştir.

3.1.1 Veri Ön İşlem Süreci

Söz konusu algoritmaların uygulanmasına geçilmeden önce genellikle veri seti üzerinde bazı ön işlemler gerçekleştirilmekte olup, söz konusu ön işlemler ile veri

setinin hazırlanması makine öğrenmesi sürecinde fazla zaman alan aşamalardan biridir (Miksovsky ve ark., 2002).

Bu süreçte belirlenmiş herhangi bir standart bulunmamasına karşın, bu süreçte verilerin eksik değerlerinin doldurulması, veri dönüştürmesi, verinin daha iyi anlamlandırılabilmesi için görselleştirilmesi, veri indirgenmesi gibi teknikler uygulanmaktadır. Böylece ham veri seti anlamlı, yorumlanabilir ve makine öğrenme algoritmaları tarafından okunabilir hale getirilmektedir.

Odyometri test sonuç veri setinin işlenmesi için Jupyter ortamında Python programlama dilinin Python 3.11.1 sürümü kullanılmıştır. Veri setinin ön işleme adımında Pandas ve Numpy kütüphaneleri kullanılmıştır.

Odyometri test sonuç veri seti 500 satır ve 13 sütundan oluşmaktadır. 'AD/SOYAD', 'YAŞ', 'CİNSİYET', 'R500Hz', 'R1000Hz', 'R2000Hz', 'R4000Hz', 'L500Hz', 'L1000Hz', 'L2000Hz', 'L4000Hz', 'R_Safses_ortalama', 'L_Safses_ortalama' adlı sütun isimlerinden oluşmaktadır. İlgili veri setinin ilk beş satırı Şekil 3.1'de gösterilmiştir.

	AD/SOYAD	YAŞ	CİNSİYET	R500Hz	R1000Hz	R2000Hz	R4000Hz	L500Hz	L1000Hz	L2000Hz	L4000Hz	R_Safses_ortalama	L_Safses_ortalama
0	G.O	34.0	E	20	20	10	25	20	20	10	40	18.75	22.50
1	C.D	26.0	K	15	15	5	50	20	10	5	45	21.25	20.00
	click to expand output; double click to hide output					20	55	15	15	25	55	25.00	27.50
3	H.G	29.0	K	10	10	20	60	15	15	20	45	25.00	23.75
4	A.P	34.0	E	10	10	25	35	15	10	20	45	20.00	22.50

Şekil 3.1: Odyometri test sonuçları veri setini Python Jupyter görüntüsü

Şekil 3.1'de sütun isimlerindeki 'R' ifadesi sağ kulak ile ilgili çıktıları gösterirken, 'L' ifadesi sol kulak ile ilgili çıktıları ifade etmek için kullanılmıştır. Tüm sayısal değerler dB cinsinden ses şiddetini ifade etmektedir.

İlgili veri setinin veri setinin tür bilgileri; float64 türünde 3 sütun, int64 türünde 8 sütun ve object türünde 2 sütun bulunmaktadır. Toplam bellek kullanımı ise 50.8 KB'dir. Veri setinde eksik değer bulunmamaktadır.

Makine öğrenme algoritmaları sayısal veriyi işlemeye uygundur. Algoritmaların kullanılabilmesi için veri türü numerik olmayan özniteliklerin veri tipinin dönüştürülmesi gerekmektedir. Kategorik veri türüne sahip özniteliklerin makine öğrenme algoritmalarıyla kullanılabilmesi için veri tipi dönüşümü “label encoder” kullanılarak yapılmıştır. Şekil 3.2’de bu yöntem kullanarak “CİNSİYET” sütunu numerik hale getirilmiştir.

	AD/SOYAD	YAŞ	CİNSİYET	R500Hz	R1000Hz	R2000Hz	R4000Hz	L500Hz	L1000Hz	L2000Hz	L4000Hz	R_Safses_ortalama	L_Safses_ortalama
0	G.O	34.0	0	20	20	10	25	20	20	10	40	18.75	22.50
1	C.D	26.0	1	15	15	5	50	20	10	5	45	21.25	20.00
2	N.A	27.0	1	10	15	20	55	15	15	25	55	25.00	27.50
3	H.G	29.0	1	10	10	20	60	15	15	20	45	25.00	23.75
4	A.P	34.0	0	10	10	25	35	15	10	20	45	20.00	22.50

Şekil 3.2 : Label Encoding Uygulanması

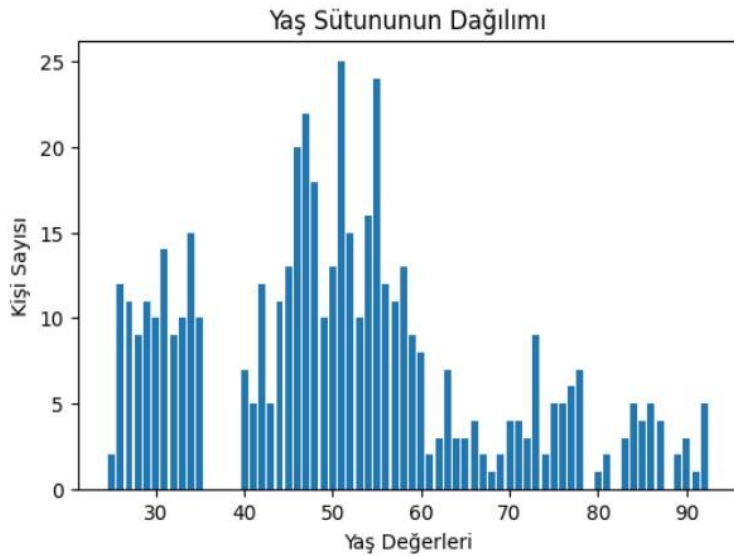
3.1.2 Verinin Görselleştirilmesi

Verileri analiz etmeden önce veri görselleştirme ile verilerin arasındaki dağılım, ilişki, benzerlikler görsel öğeler aracılığıyla daha anlaşılabilir ve kavranabilir olmaktadır. Görselleştirme, kompleks verilerin çok daha etkili şekilde analiz edilmesini ve geniş kitleler tarafından anlaşılabilir olması sağlamaktadır (Uyan Dur 2014). Veri görselleştirme, gözleme dayalı incelemelerin, iki ve üç boyutlu platformlarda görselleştirilerek, veri örüntülerinin keşfedilmesini sağlarlar (Li, 2018, s. 299).

Bu veri setinde, veriyi görselleştirmek amacıyla matplotlib ve seaborn kütüphaneleri kullanılmıştır. Bu kütüphanelerin sağladığı fonksiyonlar aracılığıyla veri setindeki çeşitli özelliklerin daha iyi anlaşılması sağlanmaya çalışılmıştır.

İlgili veri setinde, çubuk grafiği (bar chart), histogram, sütun grafiği, pasta grafiği ve çizgi grafiği (line chart) gibi çeşitli grafik türleri kullanılarak verilerin farklı yönleri görselleştirilmiştir. Çubuk grafiği, kategorik verileri, histogram sayısal verileri, sütun grafiği kategorik verileri, pasta grafiği oransal veri gösterimini, ve çizgi grafiği de iki sayısal değişkenin ilişkisini göstermek için kullanılmıştır.

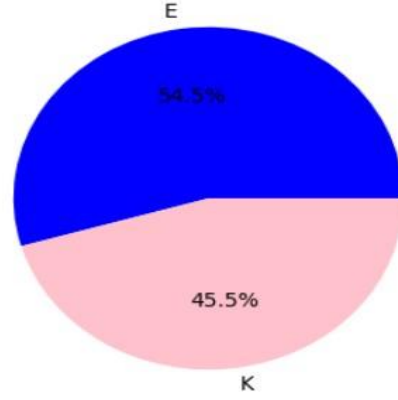
İlgili veri setinde işitme cihazı merkezine işitme kaybı şikayetiyle başvuran kişilerin “YAŞ” sütunundaki değerlerin ortalaması 51.34 olarak elde edilmiştir. Şekil 3.3 ‘te yaş sütunun çubuk grafiği kullanılarak görselleştirilmesi gösterilmiştir. Yaş sütununda, veri setindeki bireyler arasında en genç kişi 25 yaşında iken, en yaşlı kişi 92 yaşındadır.



Şekil 3.3: Yaş Dağılımının Çubuk Grafiği ile Görselleştirilmesi

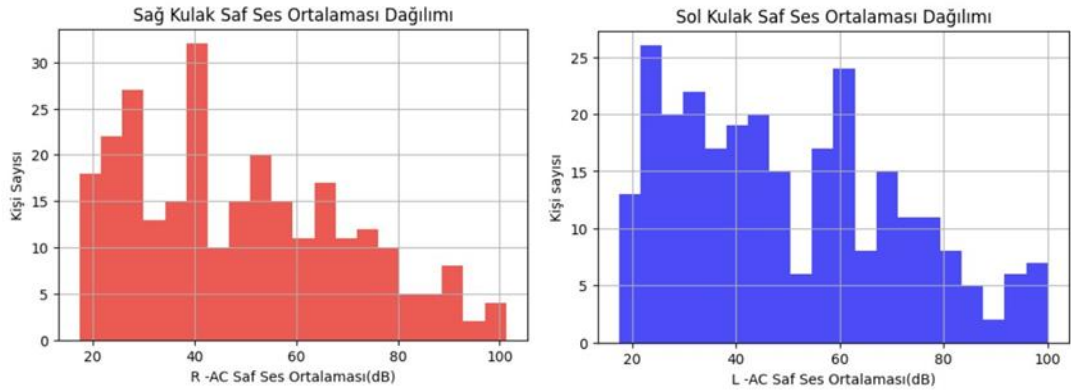
Cinsiyet sütununun dağılımı, şekil 3.4 'te gösterilen pasta grafiği ile gösterilmiştir. Erkek cinsiyetine ait dağılım %54.5 iken, kadın cinsiyetine ait dağılım %45.5 olarak belirlenmiştir.

Cinsiyet Dağılımı



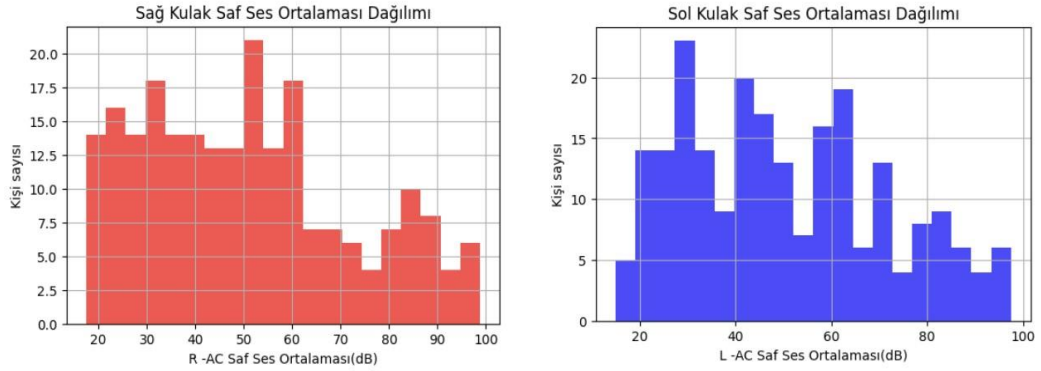
Şekil 3.4 : Cinsiyet Dağılımının Pasta Grafiği ile Görselleştirilmesi

Erkeklerin sağ kulak saf ses ortalaması, ortalama değeri 49.36, sol kulak saf ses ortalaması ise 49.90 olarak belirlenmiştir. Bu değerler, şekil 3.5'te gösterilen histogram grafiği üzerinde görselleştirilmiştir. Saf ses ortalama değerlerinin derecelendirilmesinin gösterildiği Tablo 2.1'e göre bu değerler orta dereceli işitme kaybı değerine denk gelmektedir.



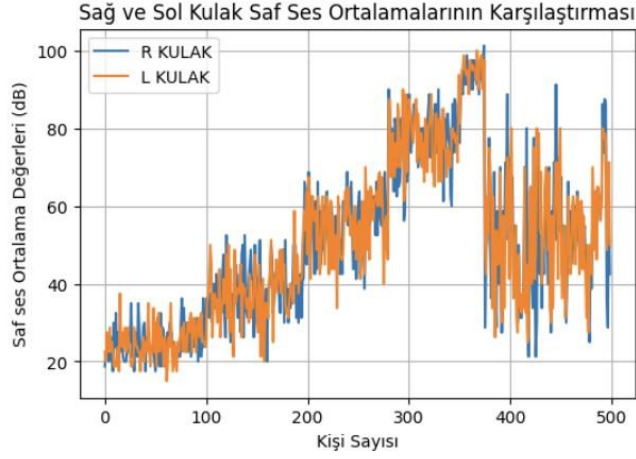
Şekil 3.5: Erkeklerin Sağ ve Sol Kulak Saf Ses Ortalamalarının Histogram Grafiği ile Gösterimi

Kadınların sağ kulak saf ses ortalaması, ortalama değeri 50.40, sol kulak saf ses ortalaması ise 50.41 olarak belirlenmiştir. Bu değerler, şekil 3.6'da gösterilen histogram grafiği üzerinde görselleştirilmiştir. Saf ses ortalama değerlerinin derecelendirilmesinin gösterildiği Tablo 2.1'e göre bu değerler orta dereceli işitme kaybı değerine denk gelmektedir.



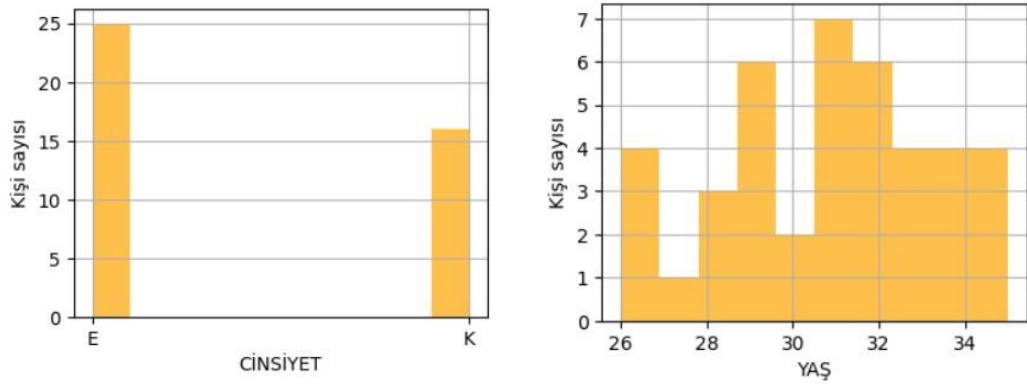
Şekil 3.6: Kadınların Sağ ve Sol Kulak Saf Ses Ortalamalarının Histogram Grafiği ile Gösterimi

Sağ kulak saf ses ortalamasının minimum değeri 17.5 dB , maksimum değeri ise 101.25 dB olarak belirlenmiştir. Sol kulak saf ses ortalamasının minimum değeri ise 16.5 dB, maksimum değeri ise 100.0 dB'dir. Bu değerler, cinsiyetten bağımsız olarak sağ ve sol kulak saf ses ortalamalarını ifade etmektedir. Sağ ve Sol kulakların işitme kaybı ortalamaları açısından karşılaştırılmaları Şekil 3.7 'de gösterilmiştir.



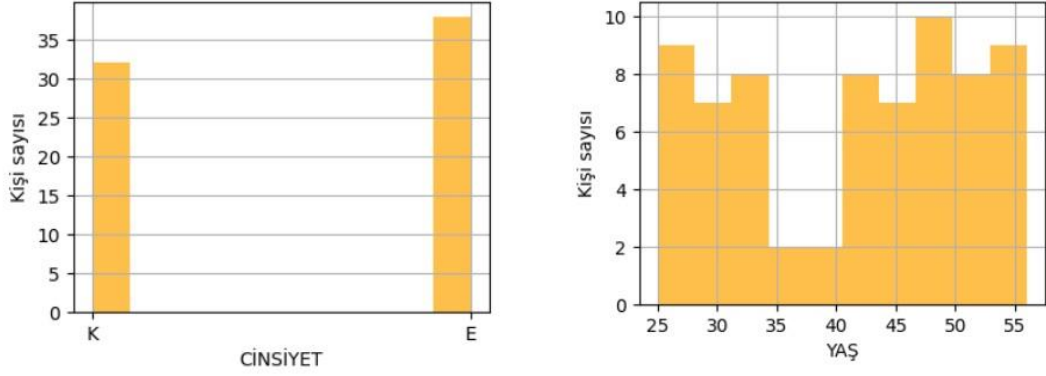
Şekil 3.7: Saf ses Ortalamalarının Çizgi Grafiği ile Karşılaştırılması

Tablo 2.1’de saf ses ortalamalarının değer aralıklarına göre işitme kaybı derecelendirilmesinden bahsedilmiştir. Bu tablodan hareketle saf ses ortalama değerleri sağ ve sol kulakta 16 ile 25 dB aralığında ise çok hafif dereceli işitme kaybı olarak adlandırılmaktadır. Veri setinde bu aralıkta olan 41 kişi bulunmaktadır. Cinsiyetlere göre sınıflandırıldığında 25 kişi erkek, 16 kişi kadındır. Çok hafif dereceli işitme kayıplı bireylerin yaş ve cinsiyet dağılımları Şekil 3.8’de gösterilmiştir.



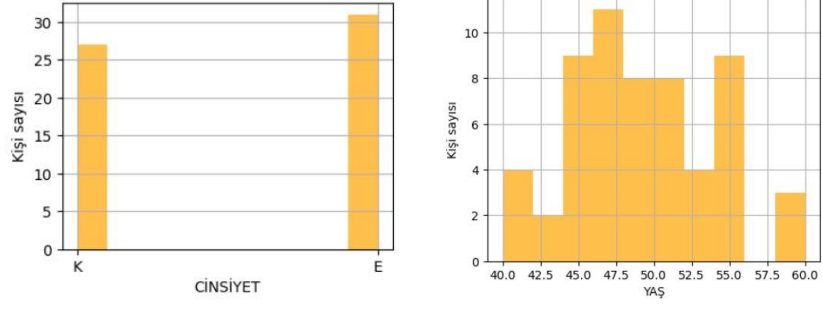
Şekil 3.8: Çok Hafif Dereceli İşitme Kaybının Yaş ve Cinsiyet Histogram Grafiği Gösterimi

Saf ses ortalama deęerleri saę ve sol kulakta 26 ile 41 dB aralıęında ise Hafif dereceli iřitme kaybı olarak adlandırılmaktadır. Veri setinde bu aralıęa giren 70 kiři bulunmaktadır. Cinsiyetlere g6re sınıflandırıldıęında 38 kiři erkek, 32 kiři kadındır. Hafif dereceli iřitme kayıplı bireylerin yař ve cinsiyet daęılımları Őekil 3.9 'da g6sterilmiřtir.



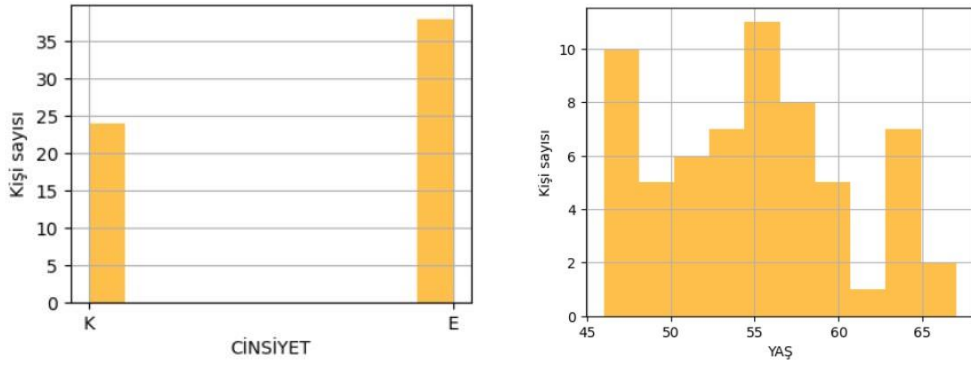
Őekil 3.9: Hafif Dereceli İřitme Kaybının Yař ve Cinsiyet Histogram Grafięi G6sterimi

Saf ses ortalama deęerleri saę ve sol kulakta 41 ile 55 dB aralıęında ise orta dereceli iřitme kaybı olarak adlandırılmaktadır. Veri setinde bu aralıęa giren 58 kiři bulunmaktadır. Cinsiyetlere g6re sınıflandırıldıęında 31 kiři erkek, 27 kiři kadındır. Hafif dereceli iřitme kayıplı bireylerin yař ve cinsiyet daęılımları Őekil 3.10 'da g6sterilmiřtir.



Şekil 3.10: Orta Dereceli İşitme Kaybının Yaş ve Cinsiyet Histogram Grafiği Gösterimi

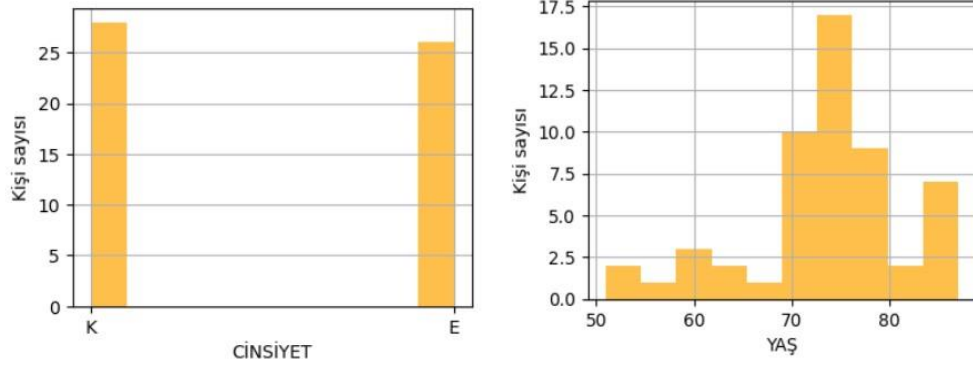
Saf ses ortalama değerleri sağ ve sol kulakta 56 ile 70 dB aralığında ise orta-ileri dereceli işitme kaybı olarak adlandırılmaktadır. Veri setinde bu aralığa giren 62 kişi bulunmaktadır. Cinsiyetlere göre sınıflandırıldığında 38 kişi erkek , 24 kişi kadındır. Orta-ileri dereceli işitme kayıplı bireylerin yaş ve cinsiyet dağılımları Şekil 3.11 'de gösterilmiştir.



Şekil 3.11: Orta-İleri Dereceli İşitme Kaybının Yaş ve Cinsiyet Histogram Grafiği Gösterimi

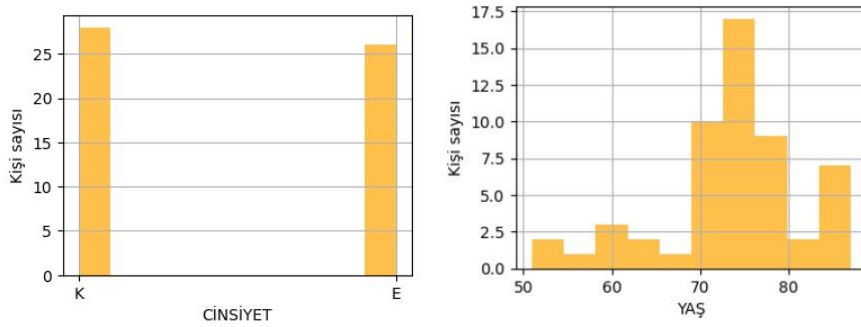
Saf ses ortalama değerleri sağ ve sol kulakta 71 ile 90 dB aralığında ise ileri dereceli işitme kaybı olarak adlandırılmaktadır. Veri setinde bu aralığa giren 54 kişi bulunmaktadır. Cinsiyetlere göre sınıflandırıldığında 28 kişi erkek, 26 kişi kadındır.

İleri dereceli işitme kayıplı bireylerin yaş ve cinsiyet dağılımları Şekil 3.12 'de gösterilmiştir.



Şekil 3.12: İleri Dereceli İşitme Kaybının Yaş ve Cinsiyet Histogram Grafiği Gösterimi

Saf ses ortalama değerleri sağ ve sol kulakta 90 dB ve üzerinde ise çok ileri dereceli işitme kaybı olarak adlandırılmaktadır. Veri setinde bu aralığa giren 14 kişi bulunmaktadır. Cinsiyetlere göre sınıflandırıldığında 8 kişi erkek, 6 kişi kadındır. Çok ileri dereceli işitme kayıplı bireylerin yaş ve cinsiyet dağılımları Şekil 3.13 'te gösterilmiştir.



Şekil 3.13: Çok İleri Dereceli İşitme Kaybının Yaş ve Cinsiyet Histogram Grafiği Gösterimi

Görseller incelendiğinde tam olarak doğru orantılı olmasa da büyük oran da işitme kaybı miktarı artıkça yaş yükselmektedir. Cinsiyet açısından değerlendirildiğinde tüm işitme kaybı derecelendirilmelerinde erkek cinsiyetinin kadın cinsiyetinden sayı olarak fazla olduğu gözlenmektedir.

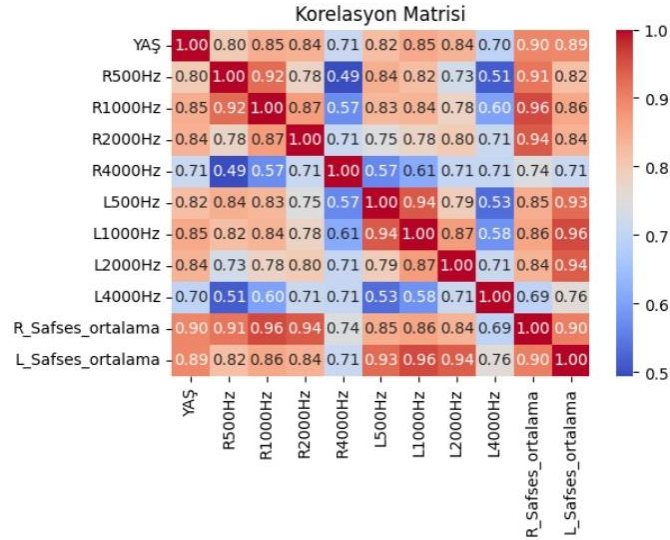
3.1.3 Öznitelik Seçimi

Veri setinde bilgiler öznitelikler ile temsil edilmektedir. Bu sebeple amaca uygun öznitelik seçimi oldukça önemlidir. Özniteliklerin birbirleriyle olan ilişkilerini gözlemleyebilmek için Seaborn kütüphanesindeki Heatmap olarak adlandırılan ısı haritası işlevi kullanılmıştır. Isı haritası sıklıkla Korelasyon Matrisi olarak da adlandırılmaktadır.

Korelasyon matrisi, veri setindeki öznitelikler arasındaki ilişkileri gösteren bir matristir. Genellikle Pearson korelasyon katsayısı kullanılarak hesaplanır ve öznitelikler arasındaki doğrusal ilişkinin gücünü ve yönünü gösterir. Korelasyon matrisi, veri setindeki öznitelikler arasındaki ilişkileri hızlı bir şekilde görselleştirmek ve analiz etmek için kullanılmaktadır.

Korelasyon katsayısı değerleri -1 ile +1 arasında değişmektedir. Korelasyon matrisinde, her bir değişken çifti için bir korelasyon katsayısı hesaplanır ve bu değerler matrisin ilgili hücrelerine yerleştirilir. Matrisin köşegenindeki hücreler genellikle 1 olarak kabul edilir, çünkü bir değişkenin kendisiyle olan korelasyonu her zaman birdir.

Pozitif korelasyon (+1'e yaklaşan değerler), değişkenler arasında doğrusal bir ilişkinin olduğunu gösterirken, negatif korelasyon (-1'e yaklaşan değerler), değişkenler arasında ters yönlü bir ilişki olduğunu göstermektedir. Veri setimizin Korelasyon matrisi Şekil 3.14 'te gösterilmiştir.



Şekil 3.14: Korelasyon Matrisi

Korelasyon matrisi incelendiğinde, yaş faktörü ile sağ kulak saf ses ortalamaları arasında 0.90, sol kulak saf ses ortalamaları arasında ise 0.89 düzeyinde güçlü bir pozitif korelasyon gözlemlenmektedir. Bu değerler, yaşın artmasıyla birlikte hem sağ hem de sol kulak saf ses ortalamalarının arttığını göstermektedir. Cinsiyet kategorik bir öznitelik olması sebebiyle korelasyon matrisinde gösterilememiştir. Korelasyon matrisinde de yalnızca sayısal değerlere sahip öznitelikler arasındaki ilişkiler gösterilmektedir.

3.1.4 Veri Setinin Eğitim ve Test Veri Seti Olarak Ayrımı

Bu çalışmada yaş ve cinsiyet demografik özelliklerine bağlı olarak sağ ve sol kulak işitme kaybı ortalamaları tahminlenmek istenmektedir. Veri setinde bağımlı ve bağımsız özniteliklerin seçimi probleme uygun olarak gerçekleştirilmelidir.

Bağımsız değişkenler, bağımlı değişken üzerinde etki yapabilen ve kendisi üzerinde başka bir değişkenden etkilenmeyen değişkenlerdir

Saf ses ortalamaları sağ ve sol kulak için 500, 1000, 2000 ve 4000 Hz frekanslarındaki işitme eşiklerinin toplanarak 4 'e bölünmesi sonucunda elde edilmektedir. Bu bilgiden hareketle ilgili veri setindeki bağımsız öznitelikler : 'YAŞ', 'CİNSİYET', 'R500Hz', 'R1000Hz', 'R2000Hz', 'R4000Hz', 'L500Hz', 'L1000Hz', 'L2000Hz', 'L4000Hz',

modelimizin bağımsız öznitelikleridir.. Bu değişkenler, tahminleme yapmak için kullanılacak olan girdi veya özelliklerdir.

Bağımlı değişkenler ise bir araştırma veya deneyin odak noktasındaki ana değişkendir. Bu değişken, diğer değişkenlerin etkisi altında kalarak değer alır ve bu nedenle bağımlı olarak adlandırılır.İlgili veri setindeki bağımsız öznitelikler : 'R_Safses_ortalama' ve 'L_Safses_ortalama' öznitelikleri, modelin tahmin etmeye çalıştığı bağımlı değişkenlerdir.

Bir makine öğrenmesi modeli oluşturulurken genellikle X, bağımsız öznitelikleri temsil ederken, y bağımlı öznitelikleri ifade etmektedir. Model, X özniteliklerini kullanarak y özniteliklerini tahmin etmeye çalışmaktadır.Modelimizde de X bağımsız öznitelikleri , y ise bağımlı öznitelikleri temsil etmektedir.Böylece veri seti X ve y olarak ikiye ayrılmıştır.

Makine öğrenme algoritmalarına veri setini vermeden önce veri setinin eğitim ve test veri seti olarak ayrılması gerekmektedir.Sklearn kütüphanesi kullanılarak veri seti X_train, X_test, y_train, y_test olarak eğitim verisi ve test verisi olarak ayrılmıştır. X_train ve y_train verileri eğitim verisi, X_test ve y_test verileri test verileri olarak kullanılmıştır.

Test_size=0.2 parametresi ile veri setinin % 80'inin eğitim verisi %20'sinin test verisi olarak ayrılmıştır. random_state değeri rastgele bölümlendirme veya karıştırma işlemini gerçekleştirmektedir.Şekil 3.15'te oluşturulan model gösterilmektedir.

```
X = df[['YAŞ', 'CİNSİYET', 'R500Hz', 'R1000Hz', 'R2000Hz', 'R4000Hz', 'L500Hz', 'L1000Hz', 'L2000Hz', 'L4000Hz']]
y = df[['R_Safses_ortalama', 'L_Safses_ortalama']]

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

Şekil 3.15: Veri Setinin Modellenmesi

3.1.5 Çoklu Regresyon Modeli

İki temel regresyon vardır. Bunlar, basit doğrusal regresyon ve çoklu doğrusal regresyon yöntemleridir (Freedman, 2009). Çoklu regresyon modelleri, tek bir modelin birden fazla çıkışa sahip olduğu durumları ele almaktadır.

Bu modeldeki amacımız, 'R_Safses_ortalama' ve 'L_Safses_ortalama' değerlerini diğer bağımsız değişkenlerle ilişkilendirerek tahmin etmeye çalışmaktır. Bağımlı değişken olan 'y', 'R_Safses_ortalama' ve 'L_Safses_ortalama' olarak iki sürekli sayısal değişken içermektedir.

Birden fazla bağımlı değişkeni ('R_Safses_ortalama' ve 'L_Safses_ortalama') aynı anda tahminlenmeye çalışıldığından dolayı bu model çoklu regresyon modeli olarak kabul edilmektedir.

3.2 Modelde Kullanılan Makine Öğrenme Algoritmaları

Bu çalışmada kullanılan algoritmalara ilişkin bilgiler aşağıda sunulmaktadır.

3.2.1 Destek Vektör Makineleri

Destek vektör makineleri en sık kullanılan makine öğrenmesi yöntemlerinden biridir [13]. SVM'nin temel amacı, veri setindeki sınıflar arasındaki en iyi ayrımı sağlayacak hiperdüzlemi bulmaktır.

SVM ile sınıflandırma yapılabildiği gibi regresyon analizi de yapılabilmektedir [14]. Bu durumda, SVM regresyonu veya destek vektör regresyonu (SVR) olarak adlandırılır. SVR, regresyon analizi için destek vektör makinelerini uyarlanmış bir versiyonudur. Modelimizde MultiOutputRegressor sınıfı kullanılarak SVR modeli çoklu regresyon modeline (Multi-Output Regression) uyarlanarak eğitilmekte ve ardından test verileri üzerinde tahminler yapmaktadır. Model Şekil 3.16'da gösterilmiştir. İlgili algoritmanın kullanılması için sklearn kütüphanesinden "sklearn.svm" ve "sklearn.multioutput" modülleri dahil edilmiştir.

```

X = df[['YAŞ', 'CİNSİYET', 'R500Hz', 'R1000Hz', 'R2000Hz', 'R4000Hz', 'L500Hz', 'L1000Hz', 'L2000Hz', 'L4000Hz']]
y = df[['R_Safses_ortalama', 'L_Safses_ortalama']]
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

svr_model = SVR()

multioutput_svr_model = MultiOutputRegressor(svr_model)

multioutput_svr_model.fit(X_train, y_train)

y_pred_multioutput_svr = multioutput_svr_model.predict(X_test)

```

Şekil 3.16: Destek Vektör Makine Regresyon Modeli (SVR)

3.2.2 Lineer Regresyon

Lineer regresyon, bir bağımlı özneliğin bir veya daha fazla bağımsız öznelikle lineer ilişkisinin modelini oluşturan bir regresyon algoritmasıdır. Bağımlı özneliğin sürekli bir sayısal değer olduğu regresyon problemlerinde yaygın olarak kullanılmaktadır.

Lineer regresyon, basit olması, yorumlanabilir olması ve geniş bir uygulama alanına sahip olması nedeniyle sıkça tercih edilen bir regresyon tekniğidir. MultiOutputRegressor sınıfı kullanılarak Lineer regresyon modeli çoklu regresyon modeline (Multi-Output Regression) uyarlanarak, model eğitilerek test veri setinde tahminler yapılmaktadır. Sklearn.linear_model kütüphanesinden LinearRegression modülü, sklearn.multioutput kütüphanesinden MultiOutputRegressor modülleri kullanılmıştır. Oluşturulan model Şekil 3.17’de gösterilmiştir.

```

X = df[['YAŞ', 'CİNSİYET', 'R500Hz', 'R1000Hz', 'R2000Hz', 'R4000Hz', 'L500Hz', 'L1000Hz', 'L2000Hz', 'L4000Hz']]
y = df[['R_Safses_ortalama', 'L_Safses_ortalama']]
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

linear_model = LinearRegression()

multioutput_linear_model = MultiOutputRegressor(linear_model)

multioutput_linear_model.fit(X_train, y_train)

y_pred_multioutput_linear = multioutput_linear_model.predict(X_test)

```

Şekil 3.17: Lineer Regresyon Modeli

3.2.3 Rastgele Ormanlar

Rastgele Orman Algoritması, ağaç tabanlı bir modelleme tekniğidir ve birden çok karar ağacını birleştirerek daha güçlü ve genelleyici bir model oluşturmaktadır.

Ağacın kökü kendisi olan X 'tir ve ağaç yapısını oluştururken her adımda bir düğüm (veya denk gelen hücre) iki kısma ayrılır. Düğümler (veya yapraklar), bir araya getirildiklerinde X 'in bir bölünmesini oluşturur. Algoritma, M farklı (rastgele) ağaç oluşturarak çalışır [15]. Her bir ağaç, veri setini belirli bir şekilde bölerek ve karar kuralları oluşturarak örnekleri sınıflandırmaktadır.

Multioutput_random_forest_model adlı değişken, random_forest_model adlı Rastgele Orman modelini MultiOutputRegressor sınıfı ile çoklu regresyon modeline uyarlamak için kullanılmıştır.

En uygun parametre işleminin araştırılması hiperparametre ayarlaması olarak adlandırılmaktadır. İlgili model de hiperparametre ayarlaması da yapılmıştır

GridSearchCV bir hiperparametre arama ve model değerlendirme aracıdır, yani bir algoritma değil, bir sınıftır. Bu sınıf, sklearn kütüphanesinde bulunmaktadır ve modelin hiperparametrelerini belirli bir ızgara (grid) üzerinde deneyerek en iyi hiperparametre kombinasyonunu bulmaya yönelik bir çapraz doğrulama prosedürünü uygulamaktadır. Rastgele Orman algoritmasında bu yöntem kullanılarak hiperparametreler bulunmuştur.

'Param_grid' içinde belirtilen hiperparametre kombinasyonları denenecek en iyi performans gösteren kombinasyon ile model eğitilerek test veri setinde tahminler yapılmıştır. 'Estimator_max_depth' ifadesi Rastgele Orman ağaçlarının maksimum derinliğini belirten bir hiperparametredir. Hiperparametre kombinasyonlarının denemesi ile bu değer 10 olarak alınmıştır 'Estimator_n_estimators' ifadesi Rastgele Orman'ın içindeki ağaç sayısını belirten bir hiperparametredir. Hiperparametre kombinasyonlarının denemesi ile bu değer 150 olarak alınmıştır. Model, toplamda 150 ağaç içerecek şekilde eğitilmiştir. Oluşturulan model Şekil 3.18'de gösterilmiştir.

```

X = df[['YAŞ', 'CİNSİYET', 'R500Hz', 'R1000Hz', 'R2000Hz', 'R4000Hz', 'L500Hz', 'L1000Hz', 'L2000Hz', 'L4000Hz']]
y = df[['R_Safses_ortalama', 'L_Safses_ortalama']]

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

random_forest_model = RandomForestRegressor()

multioutput_random_forest_model = MultiOutputRegressor(random_forest_model)

param_grid = {
    'estimator__n_estimators': [50, 100, 150],
    'estimator__max_depth': [None, 10, 20, 30]
}

grid_search = GridSearchCV(multioutput_random_forest_model, param_grid, cv=5)
grid_search.fit(X_train, y_train)

best_random_forest_model = grid_search.best_estimator_

y_pred_multioutput_rf = best_random_forest_model.predict(X_test)

```

Şekil 3.18: Rastgele Orman Modeli

3.2.4 K-En Yakın Komşu

K-En Yakın Komşu Algoritması (KNN), temel bir sınıflandırma ve regresyon algoritmasıdır. Bu algoritma, veri noktalarının komşuluk yapılarına dayanarak tahminler yapmaktadır. KNN algoritması, yeni bir veri noktasını sınıflandırmak veya tahmin yapmak için bu noktaya en yakın k (bir pozitif tam sayı) sayıda komşuyu belirlemektedir.

KNN algoritması, uygulanan en temel örnek tabanlı öğrenme algoritmaları sınıfında yer almaktadır. Örnek tabanlı öğrenme algoritmalarında, öğrenme işlemi eğitim setinde tutulan verilere dayalı olarak gerçekleştirilmektedir. (Taşçı ve Onan, 2016). Bu yöntemin en önemli avantajı çoklu kategorize edilmiş veri noktaları ile sınıflandırma işleminde başarılı çalışmaların yapılabilmesidir (Kırılıoğlu ve Ceylan, 2014: 19).

Modelimizde KNN algoritması uygulamasında GridSearchCV sınıfı kullanılarak ilgili veri seti üzerinde en iyi hiperparametre kombinasyonu bulunmuştur. 'Estimator__n_neighbors' ifadesi KNN algoritmasındaki komşu sayısını belirten bir hiperparametredir. Bu değer 5 olarak belirlenmiştir. Model, tahmin yaparken gözlemlerin etrafındaki en yakın 5 komşuyu kullanmaktadır. 'Estimator__weights' ifadesi NN'de kullanılan ağırlıklandırma stratejisini belirten bir hiperparametredir. Bu değer 'distance' olarak belirlenmiştir. Bu durum komşuların mesafelerine ters orantılı olarak ağırlıklandırılacağı anlamına gelmektedir. Model en iyi performans gösteren

kombinasyon ile eğitilerek test veri setinde tahminler yapılmıştır.Şekil 3.19’da gösterilmiştir.

```
X = df[['YAŞ', 'CİNSİYET', 'R500Hz', 'R1000Hz', 'R2000Hz', 'R4000Hz', 'L500Hz', 'L1000Hz', 'L2000Hz', 'L4000Hz']]
y = df[['R_Safses_ortalama', 'L_Safses_ortalama']]

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

knn_model = KNeighborsRegressor()

multioutput_knn_model = MultiOutputRegressor(knn_model)

param_grid = {
    'estimator__n_neighbors': [3, 5, 7],
    'estimator__weights': ['uniform', 'distance'],
}

grid_search = GridSearchCV(multioutput_knn_model, param_grid, cv=5)
grid_search.fit(X_train, y_train)

best_multioutput_knn_model = grid_search.best_estimator_

y_pred_multioutput_knn = best_multioutput_knn_model.predict(X_test)
```

Şekil 3.19: K-En Yakın Komşu Modeli

3.3 İşitme Kaybı Tahmin Penceresi

Tahmin penceresi Jupyter Notebook üzerinde Tkinter kütüphanesi kullanılarak penceresi oluşturulmuştur. Tkinter kütüphanesi , pencere, düğme, etiket gibi Grafiksel Kullanıcı Arayüzü (GUI) öğelerini oluşturmak için kullanılan standart bir kütüphanedir. Tkinter kütüphanesi içinde bir dizi modül ve sınıf barındırmaktadır.Aşağıda kullanılan modül ve sınıflar hakkında bilgiler verilmektedir.

Pencere tasarlanırken , Python'un Tkinter kütüphanesine ait olan messagebox modülü kullanılmıştır. Bu modül ile kullanıcıya bir Tkinter penceresi içinde çeşitli mesajlar, uyarılar ve hata bildirimleri gösterilebilmektedir.Gösterilen bir iletişim kutusu (dialog box) olarak düşünülebilir. Bu iletişim kutuları, kullanıcının bilgi girmesini, bir onay almasını veya bir uyarıyı görmesini sağlamak için kullanılmıştır.

Tkinter içinde yer alan ttk modülünden gelen Combobox ise bir açılır menu oluşturmayı sağlamaktadır. Kullanıcıya bir listeden seçim yapma imkanı tanıyan açılır bir menüdür.

Tkinter kütüphanesinde yer alan Spinbox, bir sayı aralığı içinde kullanıcının değer seçmesine olanak tanıyan bir giriş kutusudur. Sayısal değerlerin seçilmesi gereken durumlarda kullanılmaktadır Kullanıcı, bir artış veya azalış düğmesi yardımıyla değeri değiştirebilir veya doğrudan giriş kutusuna bir değer yazabilmektedir.

Bu pencere de kullanıcı yaş ve cinsiyet bilgilerini girerek sağ ve sol kulak işitme kaybı ortalaması ve derecesi hakkında tahmini bir bilgi alabilmektedir. Tahminleme Bulgular bölümünde bahsedilen algortimaların değerlendirme metrikleri açısından karşılaştırıldığında diğerlerine göre daha iyi performans gösteren Destek Vektör Makineleri kullanılarak yapılmaktadır. Açılan Tahminleme penceresinin görüntüsü Şekil 3.20 ve 3.21 'de gösterilmektedir.

Şekil 3.20: İşitme Kaybı Derece Tahmin Pencere Görünümü

Şekil 3.21: İşitme Kaybı Derece Tahmin Sonuç Görünüsü

Bölüm 4

Bulgular

Odyometri test sonuçları adlı 500 adet veri içeren veri setinde 4 farklı algoritma ve ilgili hiperparametre ayarlamaları yapılarak değerlendirme metrikleri hesaplanmıştır. Çoklu regresyon problemlerinde değerlendirme metrikleri olarak kullanılan ortalama Kare Hatası (MSE), Kök Ortalama Kare Hatası (RMSE) ve R-Kare Skoru (R^2) hesaplamaları ile modellerin tahmin performansı değerlendirilmiştir.

R^2 skoru, modelin bağımsız değişkenler tarafından açıklanan toplam değişkenliği yüzde olarak ifade etmektedir. R^2 skoru, modelin veri setindeki değişkenliği ne kadar iyi açıkladığını göstermektedir. En yüksek değer olan 1'e yaklaştıkça, modelin bağımlı değişkenin değişkenliğini tamamen açıkladığı anlamına gelmektedir. R^2 skoru negatif olamaz ve 0 ile 1 arasında bir değer alır. R^2 skoru ne kadar yüksekse, modelin veri setindeki değişkenliği o kadar iyi açıkladığı söylenebilmektedir [16]. Aşağıda modelde kullanılan dört algoritmanın R^2 skor çıktıları sıralanmıştır.

Destek vektör makineleri 0.97, Lineer regresyon 1.0, Random Forest 0.99 ve K-En Yakın Komşu algoritmasıyla 0.98 değerleri elde edilmiştir. Bu çıktıları değerlendirirken veri setimizin büyüklüğü de göz önüne alınmalıdır.

MSE, gerçek ve tahmin edilen değerler arasındaki farkların karesinin ortalama değerini temsil eder. MSE değeri ne kadar düşükse, modelin tahminleri gerçek değerlere o kadar yakındır ve daha iyi bir uyum sağlamış demektir. MSE, tahmin hatalarının büyüklüğünü dikkate alır ve aykırı değerlerin etkisini artırabilir [17]. Aşağıda modelde kullanılan dört algoritmanın MSE skor çıktıları sıralanmıştır.

Destek Vektör Makineleri 10.25, Lineer Regresyon 1.77, Random Forest 4.03, K-En Yakın Komşu algoritmasıyla 5.92 değerleri elde edilmiştir.

RMSE, Tahmin ve gerçek gözlem arasındaki ortalama kare farklarının karekökünün alınması ile elde edilen hata metriğidir. Ortalama mutlak hata gibi sıfır ve sonsuz arası değerler alan bu metrikte de daha iyi sonuçlar için hata puanının düşük olması

beklenmektedir (Fürnkranz vd., 2011; Microsoft, 2019; Willmott ve Matsuura, 2005). Aşağıda modelde kullanılan dört algoritmanın RMSE skor çıktıları sıralanmıştır.

Destek Vektör Makineleri 3.19, Lineer Regresyon 1.27, Random Forest 1.98 ve K-En Yakın Komşu algoritmasıyla 2.43 değerleri elde edilmiştir.

Elde edilen sonuçlara göre Lineer Regresyon (1.0) ve Destek Vektör Makineleri (0.97) en yüksek R^2 skorlarına sahiptir. Lineer Regresyon algoritması (1.77) en düşük MSE'ye sahiptir. Lineer Regresyon modeli düşük MSE, yüksek R^2 ve düşük RMSE değerleri ile dikkat çekmektedir. Ancak burada veri setinin büyüklüğü de göz önünde bulundurulmalıdır. Lineer regresyon işleme kaybı tahmin penceresinin uygulama aşamasında başarısız olmuştur.

K-En Yakın Komşu, orta seviyede bir MSE skoru ve yüksek R^2 skoru ile iyi performans göstermektedir. Ancak, diğer algoritmalara göre RMSE skoru biraz yüksektir. KNN, genellikle veri setinin büyüklüğüne duyarlıdır. Veri setimizin küçük boyutlu olması yönüyle bu algoritma tercih edilmemiştir.

Random Forest, ortalama bir MSE skoru ve yüksek R^2 skoru ile iyi bir performans sergilemektedir. Ayrıca, RMSE skoru da makul düzeydedir. Random Forest, ağaç tabanlı bir ensemble algoritması olduğu için genellikle çok sayıda veri ile kullanılması durumunda daha etkili olabilir.

Destek Vektör Makineleri, diğer algoritmalara göre daha yüksek MSE ve RMSE skorlarına sahiptir, ancak R^2 skoru oldukça iyidir. Bu, modelin verileri iyi açıkladığını, ancak bazı aykırı değerlere karşı hassas olduğunu gösterebilir. Veri setimizin özelliği bağlamında aykırı değerlere karşı hassas olması önemli olduğundan bu algoritma seçilerek ilgili işleme kaybı tahmin penceresi oluşturulmuştur.

Bölüm 5

Tartışma ve Sonuç

Bu çalışma da yaş ve cinsiyet üzerinden işitme kaybı ortalamalarını Destek vektör makineleri modelinin eğitilmesi ve Tinker kütüphanesi ile açılabilir. İşitme kaybı tahmin penceresi oluşturularak tahminleme yapılmıştır.

Tasarlanan model özellikle işitme cihazı merkezleri bağlamında düşünülmüştür. İşitme cihazı merkezleri aynı zaman da satış ve pazarlama odaklı merkezlerdir. İşitme kayıplı olan bireylerin öncelikle kamu ya da özel sağlık kuruluşlarında işitme testi yaptırmaları ve işitme testine uygun cihaz seçiminde bulunmaları gerekmektedir.

İşitme testi zaman alan, sessiz ve havasız bir kabinde gerçekleştirilen bir testtir. Bu sebeple bazı bireyler işitme testi yaptırmayı ertelemekte , işitme testi olmadan işitme cihazı alma talebinde bulunmaktadır. Ayrıca bireyler işitme cihazı almak için öncesinde işitme testinin gerekliliği durumunu süreç içerisinde öğrenmektedir. Yoğun illerde yaşayan bireyler işitme testi yaptırabilmek için aylarca beklemek durumunda kalmaktadır. Bu durumda bireylerin cihazlandırılma süreci uzamaktadır. Sürecin uzaması birey de cihaz kullanmayı ertelemeye de sebep olmaktadır.

Özellikle ikna edilemeyen, vakti sınırlı olan bireyler için cinsiyet ve yaş üzerinden işitme kaybı ortalamalarının tahminlenmesi satış ve pazarlama açısından süreci hızlandırabilir ve diğer rakip işitme cihazı merkezleriyle rekabet açısından önemli olabilir.

Aynı zamanda yaşlı bireylerin zamanla bilişsel fonksiyonlarında zayıflama meydana gelmesi, demans ve unutkanlık gibi problemlerin eşlik ettiği durumlarda bu bireylere uygulanan işitme testleri oldukça zor olmakta ve zaman almaktadır. İşitme testi bireyin verdiği cevaplara bağlı olarak uygulanan subjektif bir test olduğundan özellikle yaşlı bireyler de testin doğruluk oranı azalmaktadır.

İşitme cihazı merkezlerine işitme cihazı bilgisi alma amacıyla randevu oluşturup gelen bireyler de ise işitme testleri mevcut olmadığı zaman ön görüşme sağlanamamaktadır.

Tüm bu durumlar göz önüne alındığında işitme kaybı tahmin penceresi işitme cihazı satış ve uygulama merkezlerine başvuran bireylerin cihazlandırılma sürecinin hızlandırılması, ön görüşme için gelen bireylere yaş ve cinsiyetlerine göre tahmini işitme kaybı ortalamaları üzerinden görüşme sağlanması ve hasta akışının fazla olduğu merkezlerde uygulama da hızlı olunabilmesini sağlayabilir.

Model işitme testinin yerini tam anlamıyla tutmamakla birlikte yakın bir doğruluk sağlayabilir. İşitme kaybının oluşma sebepleri genellikle benzer paternlere bağlıdır. Bu noktada bireyden alınan anamnez bilgileri oldukça önemlidir.

İlgili model 500 adet sınırlı bir veriden yola çıkılarak tasarlanmıştır. Özellikle veri miktarının artırılması bu noktada oldukça önemlidir. Makine öğrenme algoritmaları ne kadar çok veri ile eğitilirse insana yakın sonuç verme oranı da o derece artacaktır.

Veri miktarının artırılmasıyla ilgili model üzerinde yapay sinir ağları modellenilebilir ve hem makine hem yapay sinir ağı modellerinin performansları karşılaştırılarak etkili model oluşturulabilir.

İşitme kaybı tahmin penceresine daha fazla seçim eklenebilir. Bu modelde sadece cinsiyet ve yaş değerleri üzerinden gidilmiştir. İşitme cihazı merkezine başvuran kişilere sorulan sorulardan yola çıkılarak yeni buton seçimleri modele eklenebilir ve böylece daha spesifik bir tahminleme gerçekleştirilebilir. Bireylerin anamnez bilgileri ve test bilgilerinden oluşan veri miktarı daha büyük bir veri seti ile ilgili model tasarlanabilir.

Örneğin kişinin kulak zarı durumu seçilebilir. Yırtık, sağlam, tüp takılı şeklinde seçim eklenebilir. Kafa travması öyküsünün olup olmadığı sorgulanabilir. Sık aralıklarla orta kulak iltihabı geçirme durumu sorgulanarak ilgili yanıtlar da dahil edilebilir. Gürültüye maruz kalıp kalmadığı, meslek grubu, kulağından herhangi bir operasyon geçirip geçirmediği anamnez bilgileri de dahil edilebilir ve tüm bu özelliklerden yola çıkılarak işitme kaybı durumu tahminlenebilir.

Kişiden alınan bu bilgiler veri tabanı sistemi oluşturularak kayıt edilebilir, böylece bireylerin işitme cihazı merkez memnuniyetleri artırılarak bireylere daha spesifik hizmet sağlanabilir. Merkezde çalışan bireylerin değişimi durumunda ise ilgili veri tabanı sisteminden hastaya özgü kontroller sağlanabilir.

Sonu olarak her alanda hayatımıza dahil olmaya bařlayan makine ğrenmesi ve yapay zeka sistemleri bu merkezlere gre tasarlanarak iřitme kayıplı bireylerin konforunu artırabilir.

Kaynaklar

- [1] Bing, D., Ying, J., Miao, J., Lan, L., Wang, D., Zhao, L., Wang, Q. (2018). Predicting the hearing outcome in sudden sensorineural hearing loss via machine learning models. *Clinical Otolaryngology*, 43, 868–874. doi:10.1111/coa.13033
- [2] World Health Organization. *Global Costs of Unaddressed Hearing Loss and Cost-Effectiveness of Interventions*. (World Health Organization, 2017). OCLC: 975492198.
- [3] Olusanya, B. O., Neumann, K. J., & Saunders, J. E. (2014). The global burden of disabling hearing impairment: A call to action. *Bulletin of the World Health Organization*, 92, 367–373. <https://doi.org/10.2471/BLT.13.128728>
- [4] Lin, F. R., et al. (2013). Hearing Loss and Cognitive Decline in Older Adults. *JAMA Internal Medicine*, 173, 293–299. <https://doi.org/10.1001/jamainternmed.2013.1868>
- [5] Maurer, J. F., & Rupp, R. R. (1979). Aged deaf presbycusis. In: *Hearing disorders in old age*. New York: Grune & Stratton; 33-63.
- [6] Erdoğan, A. A. (2016). Yaşlılık Döneminde İşitme Kaybı ve İşitme Kaybına Yaklaşımlar. *Turkish Journal of Family Medicine and Primary Care*, 10(1), 25-33. <https://doi.org/10.5455/tjfmmpc.204524>
- [7] Weinstein, B. E. (2002). Hearing loss in the elderly: A new look at an old problem. In: *Handbook of Clinical Audiology*, 5th ed. Baltimore: Lippincott Williams & Wilkins; 597-607.
- [8] Kılıçdağ, E. B., et al. (2004). Postmenopozal kadınlarda östrojen tedavisinin işitme üzerine etkileri. *Amerikan Kadın Hastalıkları ve Doğum Dergisi*, 190(1), 77-82. <https://doi.org/10.1016/j.ajog.2003.06.001>
- [9] Zhao, Y., Li, J., Zhang, M., & Lu, Y. (2018). Machine Learning Models for the Hearing Impairment Prediction in Workers Exposed to Complex Industrial

Noise: A Pilot Study. *Ear and Hearing*, 40(3), 1.

DOI:10.1097/AUD.0000000000000649

- [10] Park, K. V., et al. (2020). Machine Learning Models for Predicting Hearing Prognosis in Unilateral Idiopathic Sudden Sensorineural Hearing Loss. *Clinical and Experimental Otorhinolaryngology*, 13(2), 148–156. <https://doi.org/10.21053/ceo.2019.01858>
- [11] Pai, K. V. (2022). Hearing Loss Prediction in Newborns, Infants and Toddlers using Machine Learning. In *2022 IEEE North Karnataka Subsection Flagship International Conference (NKCon)* (s. 10127090). IEEE. <https://doi.org/10.1109/NKCon56289.2022.10127090>
- [12] Kim, H., et al. (2021, December 31). Predicting speech discrimination scores from pure-tone thresholds—A machine learning-based approach using data from 12,697 subjects. *PLOS ONE*. <https://doi.org/10.1371/journal.pone.0261433>
- [13] Guyon, I., Weston, J., Barnhill, S., Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Machine Learning*, 46(1-3), 389-422. <https://doi.org/10.1023/A:1012487302797/METRICS>
- [14] Cortes, C., Vapnik, V., Saitta, L. (1995). Support-vector networks. *Machine Learning*, 20(3), 273-297. <https://doi.org/10.1007/BF00994018>
- [15] Biau, G., & Scornet, E. (2016). A random forest guided tour. *TEST*, 25, 197–227. <https://doi.org/10.1007/s11749-016-0481-7>
- [16] Medium. (2021). *Model Performansını Değerlendirmek*. [İnternet]. <https://medium.com/yaz%C4%B1%C4%B1m-ve-bili%C5%9Fimkul%C3%BCb%C3%BC/model-performans%C4%B1n%C4%B1-de%C4%9Ferlendirmek-regresyon-48b4afec8664>

- [17] IBM. (2023). *Ortalama Karesi Hatası*. [Internet].
<https://www.ibm.com/docs/tr/cloud-paks/cp-data/4.6.x?topic=overview-meansquared-error>