



## Yapay Zeka Yöntemleri ile Bebek Seslerinin Tahmin Edilmesi

 Resul BÜTÜNER<sup>\*,a</sup>,

<sup>a,\*</sup> Adil Karaağaç MTAL, Konya, 42130, TÜRKİYE

### MAKALE BİLGİSİ

Alınma: 29.05.2020  
Kabul: 30.06.2020

**Anahtar Kelimeler:**  
Yapay Zekâ, Makina Öğrenmesi, Bebek Sesi Analizi, Tahmin

**\*Sorumlu Yazar**  
e-posta:  
[rbutuner@gmail.com](mailto:rbutuner@gmail.com)

### ÖZET

Bu araştırmada ağlayan bebeklerin veya sağır ebeveynlerin çocuklarının sesini algılayarak ihtiyacın belirlenmesine yönelik platformlar yardımıyla makine öğrenmesi uygulamaları geliştirilmiştir. Bebek sesleri konusunda ebeveynler, çocuk doktorları, sağlık çalışanları ve dijital ortam verilerinden faydalanılacaktır. Bebeklerin ihtiyaçlarına yönelik literatür taraması yapılarak, fikir üretme ve geliştirme olarak analizler yapılacaktır. Bebeklerin genel olarak temel ihtiyaçlarından açım, gazım var, gaz ağrısı var, rahatsızım, uykum var vb. genel ihtiyaçlarına yönelik çıkardıkları sesler toplanarak, ihtiyaç sınıflarına göre veri seti oluşturulacaktır. Veri setleri eğitilerek, öğrenme ve test aşamasından geçerek, platformlar ile makine öğrenmesi çalıştırılacaktır. Makine öğrenmesi algoritmaları olarak Rasgele Orman Sınıflandırma, Çoklu Sınıf Sınıflandırma ve Makine Öğrenmesi (<https://teachablemachine.withgoogle.com/>) web aracı kullanılmıştır. Uygulamalar çalıştırıldığında dışarıdan gelen yeni bebek seslerini, veri setindeki bebek sesleri ile karşılaştırarak, ekrana yazı ve sesli olarak "açım, uykum var, rahatsızım, gazım var, gaz ağrısı var" vb. bilgilendirmeleri yapacaktır. Ayrıca başarı oranları elde edilerek, karşılaştırılması yapılmıştır. Çoklu sınıf sınıflandırma yöntemi kullanılarak %97,5 başarı oranı ile en iyi sonucu vermiştir. Bu çalışmada yapay zekâ da ses ve konuşma tanıma yardımıyla sağır veya sağlıklı ebeveynlerin bebeklerinin veya küçük çocuklarının ses ve konuşma ipuçlarına uygun şekilde hareket edebilmeleri için bebeklerin ağlamalarını anlamalarına yardımcı olmayı hedeflemektedir.

## Estimation of Baby Sounds with Artificial Intelligence Methods

### ARTICLE INFO

Received: 29.05.2020  
Accepted: 30.06.2020

**Keywords:**  
Artificial Intelligence, Machine Learning, Baby Voice Analysis, Prediction

**\*Corresponding Authors**  
e-mail:  
[rbutuner@gmail.com](mailto:rbutuner@gmail.com)

### ABSTRACT

In this study, it is for the platforms for the need of the crying babies or deaf parents by perceiving the voice of their children. Parents, pediatricians, health services and digital media will benefit from baby sounds. Literature review for the use of babies, analysis as to generate and develop ideas will be done. The main needs of babies are hungry, I have gas, I have gas pain, I am disturbed, I have sleep, etc. By collecting the sounds they make for an overview, a data set will be created according to the classes of need. Data sets are trained, passed through the learning and testing phase, and machine learning is carried out with platforms. Random Forest Classification, Multiple Classification and Machine Learning (<https://teachablemachine.withgoogle.com/>) web tool is used as machine learning algorithms. Applications are run, by comparing new baby sounds coming from outside with baby sounds in the data set, "hungry, sleepy, disturbed, I have gas, gas pain" etc. will inform you. It is also for your comparison to achieve success rates. It gave the best results in the multi-class classification method with a success rate of 97.5%. It aims to help children understand the crying of deaf or healthy parents in their artificial intelligence in terms of voice and speech recognition so that their babies or young children can act according to their voice and speech tips.

## 1. GİRİŞ (INTRODUCTION)

Bebekler ihtiyaç duyduğu yemeği, sıcaklığı, rahatlığı vb. kişisel tüm ihtiyaçları anneye bağımlı olarak yaşamaktadırlar. Bebeklerde en önemli bilgi akışı ise ağıladığında bu ihtiyaçlarını iletmek ve annelerden dikkat ve bakım istemenin yolu bu şekilde olmaktadır. Bebeklerin hangi ihtiyacı halletmeniz gerektiğini anlamak bazen zor olabilir. Ancak bebek büyüdükçe anneye iletişim kurmanın başka yollarını da öğrenecektir. Bunlar göz temasında daha iyi olacak, sesler çıkaracak ve gülümseyecektir. Ayrıca bebeklerin ağlamasının ve onları almamız ve ihtiyaçlarını gidermeye çalışmamız büyük önem taşımaktadır. Sağır ebeveynlerde ise çocuklarının ihtiyaçlarını çok büyük sorun haline gelmektedir. Çoğu kez bu problemle beraber bebek ve çocukların yeterli beslenememesine, temel ihtiyaçlarının zamanında yapılamamasına sebep vererek, hastalanmalarına yol açmaktadır.

Günümüz teknolojisinde yapay zekâ ile birlikte sağlık, iletişim, endüstri vb. alanlarda büyük değişimler gerçekleşti ve gerçekleşmeye devam etmektedir. Bu değişimlerden en çok ihtiyaç duyulan alan ise bebek dilini anlamaya çalışmak, bununla ilgili verileri toplayıp veri havuzu oluşturmaktır. Bu çalışmalar oluşturulurken, yapay zekâ öğelerinden makina öğrenmesi, ses tanıma, konuşma tanıma vb. alanlar kullanılmıştır.

Ses tanıma teknolojisi (Speech Recognition Technology), dilbilim, bilgisayar bilimleri, elektronik gibi farklı bölümleri içinde barındıran, doğal konuşma dilini (natural language) bir bilgisayar tarafından mikrofon ile algılanarak otomatik olarak metin verisine dönüştürmeyi sağlayan ve sistemleri inceleyen bir çalışma alanı olarak adlandırılmaktadır[1].

Makina öğrenmesi, ses ve görüntü üzerine literatür taramalarına bakıldığı zaman, şuan ki teknolojilerde, makina öğrenmesi yöntemleri ham verilerden anlamlı bilgilerin tespit edilmesinde başarılı sonuçlar elde etmektedir[2]. Bu yöntemler bebek seslerinin analizinde de başarılı sonuçlar vereceğini göstermektedir.

Bu değişimlerde makina öğrenmesi içeriklerden anlamlı sonuçlar elde etmek ve gerektiğinde kullanmak üzere, devletler, kurumlar veya şahıslarca pek çok veri toplanmaktadır. Her meslekte genellikle sayılar, metinler, sesler ve resimler gibi materyallerin oluşturduğu veriler, bilgisayarlar yardımıyla elektronik ortamlara aktarılmış bulunmaktadır. Bilgisayar, internet ve buna bağlı olan teknolojilerin hayatın tüm alanlarında daha çok yer almasıyla beraber, bu teknolojilerin ürettiği verilerin de depolanması bu hususla anlamlı olmaktadır[3].

Badem (2017) tarafında yapılan çalışmada Parkinson hastalığının ses sinyalleri üzerinden sınıflandırılmasında, KYK, ROS, DVM, NB ve KA makina öğrenmesi tekniklerinin tanımlanması amaçlanmıştır. Bu başarı için PDC veri seti kullanılarak, yüksek doğruluk tahmin değerleri elde edilmiştir[4].

Yapılan bir çalışmada konuşmacının ses örnekleri bilgisayara kaydedilerek bu konuşmacıların kimliklerinin belirlenmesi sağlanmış. Çalışma üç aşamadan meydana gelmektedir. Birinci aşamada konuşmacı sesinin veriye dönüştürülmesi ve özellik çıkarılması, ikinci aşamada konuşmacı seslerinin yapay sinir ağlarıyla eğitilmesi ve üçüncü aşamada ise eğitilmiş sinir ağlarına test verileri göndererek istatistiki bilgiler elde edilmiştir[5].

Çakır ve Okutan (2011) tarafından yapılan çalışmada kullanıcıların web tarayıcısını ses tanıma sisteminin kullanılmasına yönelik çalışmışlardır. Ses tanıma ile web tarayıcısındaki web sitelerinin açılmasını, kaydedilmesi vb. işlemleri taşıyan uygulama geliştirmişlerdir. Uygulamayı test ederek başarılı sonuçlar çıkarmışlardır[6].

Makina öğrenmesi yöntemi ile yaralanmalı trafik kazaların analizi yöntemi kullanılarak Adana ilindeki 2005 ile 2014 yılları arasında gerçekleşen trafik kazaları üzerine araştırma yapılmıştır. Bu kazalarla ilgili sayılar ve meteorolojik verilerin aylık alınarak, birleştirilmesi sonucunda bir veri seti kullanılmıştır. Bu veri setinin işlenmesi ile birlikte trafik kazalarının yaralanmalı olanların sayısı ve yaralı sayısını tahmin edecek modeller geliştirilmiştir. Çalışma sonucunda DVM (Destek Vektör Makinesi) yönteminin her iki tahminde en başarılı olduğu sonucu ortaya çıkmıştır. Yaralanan kaza sayısı tahminlerinin RA (Regresyon Ağacı) dışındaki yaralı tahminlerinden daha başarılı olduğu saptanmıştır[7].

Alex Graves tarafından ses verileri üzerine yaptığı çalışmada, fonetik sunumu kullanılmadan sesi RNN tabanlı olarak direk metinlere çevirebilen bir konuşma tanıma sistemi geliştirilmiştir [8].

Anahtar kelime tespiti üzerine yapılan bir çalışmada çok kapsamlı konuşma verilerinden oluşan Verbmobil veri seti üzerinde %84.5 başarı yüzdesi LSTM ile bulunmuştur[9].

Konuşma spektrogram programlamasına yönelik bir çalışmada ise DAE tabanlı bir mimari uygulanmış ve konuşma verisi ile elde edilen sonuçlarla daha önce görüntü yamalarını kodlamak için elde edilen başarıya yaklaşık olarak aynı olduğunu belirlenmiştir[10].

DBN modeli kullanılarak konuşma tanıma üzerine yapılan çalışmada ses kayıt verileri üzerinde anlamlı sonuçlar ortaya çıkmıştır[11].

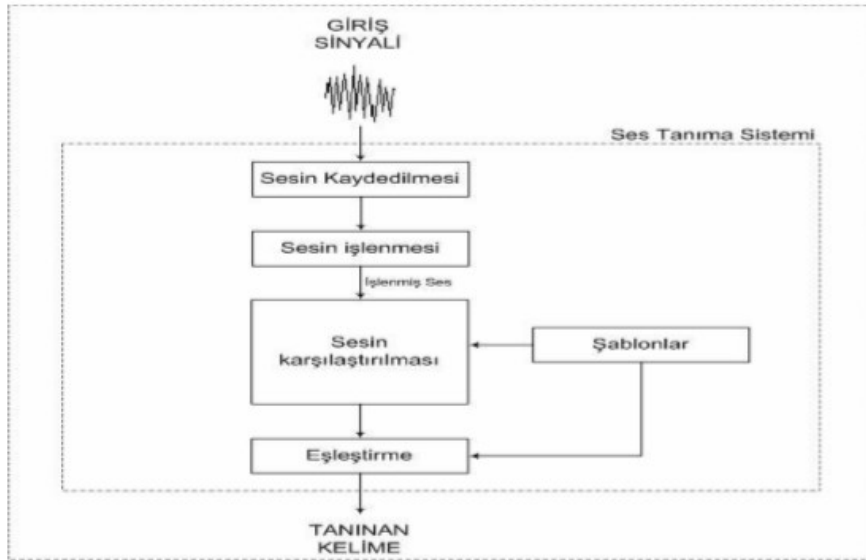
Ses tanıma üzerine durulan bir çok çalışmada RNN ağları kullanıldığı görülmektedir[12]. Hochreiter ve Schmidhuber tarafından (1997) derin öğrenmede özel bir RNN türü olan Uzun Kısa Vadeli Bellek(Long Short Term Memory-LSTM) ağları duyurulmuştur[13].

LSTM mimarisine bakıldığında giriş, unutma ve çıkış olmak üzere 3 kapı blok girişi görülmektedir. Bu mimari de sabit hata döngüsü, çıkış aktivasyon fonksiyonu ve gözetleme (peephole) bağlantıları bulunmaktadır[14]. LSTM mimarileri konuşma, metin işleme, analiz konularında başarılı sonuçlar vermektedir. Ses sınıflandırma üzerine yapılan çalışmada farklı dillerdeki zengin materyal içerikler sunan TIMIT veri seti kullanılarak %70 başarı oranı elde edilmiştir. Bu veri kümesinde yapılan çalışmalar

artırılarak başka bir LSTM yaklaşımı kullanılmış ve aynı veri seti üzerinde %83 başarı oranına ulaşılmıştır[15].

Ses tanıma ile yapılan çalışmalarda internet bankacılığı, sesli arama, veri tabanı ulaşım servisleri, sesle bilgisayarların uzaktan kontrolü, adli uygulamalar, konuşma tanıma ve denetleme uygulamaları, konuşmanın yazılı metin haline çevrilmesi, engelliler için destek, gazetecilik, eğitim, sağlık, görsel medya, robotlar, gibi vb. çokça örnekler verilebilir [16].

Ses tanıma üzerine yapılan bir çalışmada model geliştirilmiş ve aşamaları oluşturulmuştur. Bu aşamalar aşağıdaki gibidir[17]. Çalışmada ses girişi sinyali olarak gelmektedir. Şekil 1'deki bu sistemin akış modeline göre ses kaydedilip, işleniyor devamında veri setindeki seslerle karşılaştırılıyor ve uyan ses ile eşleştirilip, tanınan kelime dışarıya çıktı olarak verilmektedir.



Şekil 1. Ses tanıma sistemi modeli  
(Sound recognition system model)

Şu anki teknolojide ses ve konuşma tanıma üzerine yapılmış çalışmaların yetersiz oluşu ve beraberinde yapılacak olan çalışmaların gereksiniminden dolayı oldukça çok fazla ihtiyaç duyulmaktadır. Bu alanlar ticari açıdan düşünülürse oldukça büyük bir pazara sahip olan bir teknolojidir. İnsan sesinin parmak izi gibi çeşitlilik göstermesi, çalışılan dillerin farklılığı, cümle yapısı vb. sebeplerden dolayı bu alanın oldukça zor bir disipline sahiptir[18]. Bebeklerin gereksinimlerinin çokça oluşu, ihtiyaçlarını anlamadaki güçlükler, hatta engelli olan ebeveynler içinde bebek veya çocuklarının bakımları oldukça zor olmaktadır. Çalışmada kullanılacak olan veriler

bebek seslerinden alınmıştır. Yapay zekâ makina öğrenmesi algoritmaları ile bebeğinizi ağlamasını beş ihtiyaçtan birine çevirmektedir. Bu sesler eğitilerek, test edilmiştir. Bebeğin açım, uykum var, gazım var, gaz ağrım var, rahatsızım gibi ihtiyaçlar ekrana sesli ve metin şeklinde gelebilecektir. Ayrıca sağır olan ebeveynler içinde çocuklarının sesleri metin olarak ekranda gösterilecektir. Böylelikle makina öğrenmesi yöntemi ile hem bebek dilinin anlaşılması hedeflenmektedir. Bu çalışmada, bebeğin bu ihtiyaçları makina öğrenmesi algoritmaları ile ses dosyaları analiz edilerek, ebeveynlerin bebeklerini daha iyi anlamalarına,

ağlamların bebekleri nasıl etkilediğini anlamasına yardımcı olmayı amaçlamaktadır. Kullanılacak olan algoritmalarındaki başarı oranlarının karşılaştırılması amaçlanmaktadır. Ses ve konuşma ile yapılacak diğer kimlik tespiti, robot kontrolü, ev otomasyonu, sesli asistan, vb. çalışmalara örnek olması beklenmektedir.

## 2. YÖNTEM (Method)

Çalışmanın hazırlanması toplamda 5 aşamadan oluşmaktadır. Bu aşamalar;

### 2.1. Verilerin toplanması(Data collection)

Bebeklerin ihtiyaçlarına göre çıkardıkları sesler toplanmıştır. Temelde 5 ihtiyaç belirlenmiştir. Açım, gazım var, gaz ağrısı var, rahatsızım, uykum var vb. gibi. Bu veri setleri (sesler) bebek seslerine ait videolardan toplanarak elde edilmiştir.

### 2.2. Platformların belirlenmesi (Platform determination)

Makine öğrenmesi, blok tabanlı ve seslerle birlikte çalışan platformlar araştırıldı ve en uygun olarak anlaşılması kolay ve yapay zekâ adımları daha düzenli olan <https://machinelearningforkids.co.uk/> olan Çocuklar İçin Makine Öğrenmesi(MLFK) online editörü <https://machinelearningforkids.co.uk/>, Öğretilbilir Makine (Teachable Machine) online editörü <https://teachablemachine.withgoogle.com/> ve Python Konuşma Tanıma(Python Speech Recognizer) kütüphanesi belirlenmiştir.

### 2.3. Veri setinin oluşturulması (Creation of data set)

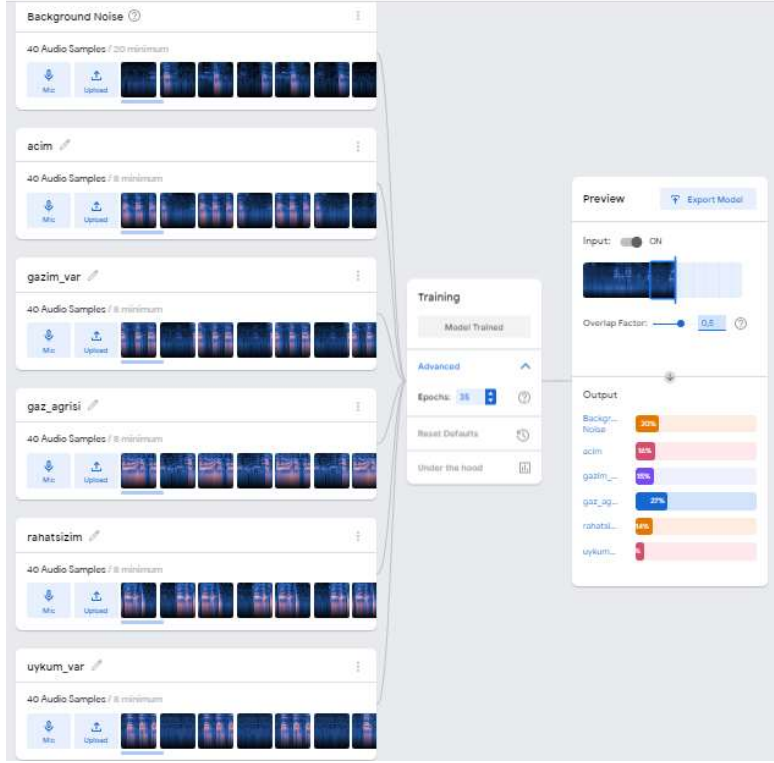
Çalışmayı ses kategorisinde belirledikten sonra, eğitim kısmında arka plan gürültü sesleri de dâhil 6 kategoride veri setlerini oluşturuyoruz. Veri seti oluşturmak için etiketleri <https://machinelearningforkids.co.uk/> web aracından Şekil 2 ve Şekil 3'deki gibi belirtilerek oluşturulmuştur.



Şekil 2. Veri setinin oluşturma ekranı  
(Display on the dataset screen)



Şekil 3. Veri seti oluşturma işlemleri  
(Data set creation operations)



Şekil 4. Makine öğrenmesi web aracı veri seti ekranı  
(Machine learning web tool dataset screen)

Şekil 4’de ise yine farklı platform olan <https://teachablemachine.withgoogle.com/train/audi> o online makine öğrenmesi web aracından 6 kategoride veri seti oluşturulmuştur. Eğitime hazır hale getirilmiştir.

#### Eğitim ve Test Aşaması (Teaching And Testing Stage)

Platformlar ve python ile kodlamada eğitim kısmından bebek seslerini eğiterek test aşamasına uygun hale getirilmiştir.

#### 2.4. Uygulamanın Oluşturulması (Creating the Application)

MLFK için oluştur sekmesinden ise makine öğrenmesi bloklarını kullanarak, uygulamamız tamamlanmıştır.

Python ile eğitim ve test aşamalarını içeren, uygulamayı Pycharm editörü yardımıyla hazırlanmıştır.

Öğretilebilir Makine web aracı ile de veri setleri kullanılarak, eğitim aşaması gerçekleştirilmiştir.

#### 2.5. Çalışma Grubu (Working Group)

Araştırmanın çalışma grubu veri toplama süreci ile ilişkili olarak iki farklı gruba yürütülmüştür.

Birincisi konuşamayan bebeklerdir. İkincisi ise sağır ebeveynlere ait çocuklardır.

#### 2.6. Veri Toplama Araçları Ve Toplanması (Data Collection Tools And Collection)

Araştırma sorularının yanıtlarına ulaşmak amacıyla kullanılacak veri toplama araçlarının tespitinde öncelikle alanyazın taraması yapılmış, bebek seslerine ait veri setlerinin olup olmadığı araştırılmıştır. İnternet üzerinden farklı ülkelerdeki birçok bebek sesi ses ve video dosyalarına ulaşılmıştır. Bu dosyalar dosya indirme programları yardımıyla indirilmiştir. Her bir bebek sesi kategorisine göre sesler kesilerek, bölümlere ayrılmıştır. Bu bebek sesleri yerel bilgisayarın sabit diskine kaydedilmiştir.

#### 2.7. Verilerin Analizi (Analysis of Data)

Bebek seslerine ait veriler makine öğrenmesi editöründe eğitildikten ve test aşamasından geçtikten sonra veriler Çoklu Sınıf Sınıflandırma, Rasgele Orman Sınıflandırma algoritmaları ve Öğretilebilir Makine (Teachable Machine) web aracı ile analiz edilmiştir. Bebek seslerine ait uygun kategoriler bu analiz sonucunda belirlenmiştir.

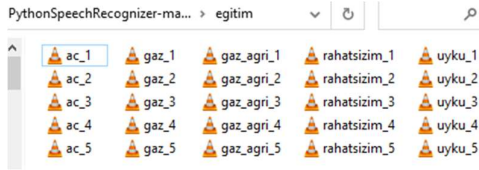
### 3. BULGULAR (Findings)

Bebek seslerinin araştırılması için, video üzerinde çalışma yapılmıştır. Bu video da 5 kategoride bebeklerin 0-3 aylık iken ihtiyaçlarına göre çıkardıkları sesler bulunmaktadır. Bu temel kategoriler:

- ✓ Açım,
- ✓ Uykum var,
- ✓ Rahatsızım,
- ✓ Gazım var,
- ✓ Gaz ağrım var şeklindedir.

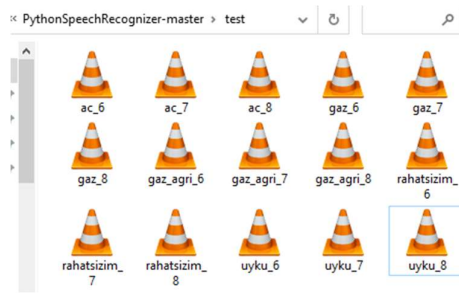
Her kategori için 10 bebekten 10 farklı sesler alınarak veri seti oluşturulmuştur. Bu veri setleri 3 ayrı platformda eğitilmiştir.

Veri setinde 40 adet 5 kategoride bebek sesleri bulunmaktadır. Bu sesler Şekil 5'deki gibidir. Dosyaların hepsi .wav uzantılıdır.



Şekil 5. Eğitim veri setleri  
(Train data sets)

Eğitim ve test için kullanılacak wav dosyalarını oluştururken Audacity programını kullanınız. Burada dikkat edilmesi gereken şey, ses dosyalarını hazırlarken, hem eğitim hemde test için, (Proje Hızı) 16000 Hz, mono (tek kanallı) ve 16 bit pcm şeklinde hazırlanması gerekmektedir.



Şekil 6. Test veri setleri  
(Testing data sets)

Şekil 6'da Audacity programında hazırlanmış 15 adet test verisi bulunmaktadır. Bu veriler her kategoriden 3 adet olmak üzere 5 kategoriye sahiptir. 3 ayrı platformda veriler analiz edilerek, sonuçlar yorumlanmıştır. Bu platformlar:

- I. <https://machinelearningforkids.co.uk/> web aracı,

- II. <https://teachablemachine.withgoogle.com/> web aracı,
- III. Python ile konuşma tanıma uygulaması

### 3.1. Çocuklar İçin Makine Öğrenmesi (Machine Learning For Kids)

Online makine öğrenmesi üzerine çalışan ve metin, ses ve görüntü üzerinde sınıflandırma yapabilen bir web aracıdır. API'ler kendi üzerinde barındırmaktadır. Verilerin analiz edilmesi için, Rasgele Orman Sınıflandırma algoritması kullanılmıştır. Bu algoritma hem karar ağacı gibi hem sınıflandırma hem de regresyon problemlerinde kullanılabilir. Çalışma mantığı birden fazla karar ağacı oluşturur. Bir sonuç üreteceği zaman bu karar ağaçlarındaki ortalama değer alınır ve sonuç üretilir. Çalışmada bebek sesleri verileri için Rasgele Orman Sınıflandırma yöntemi seçilmiştir.

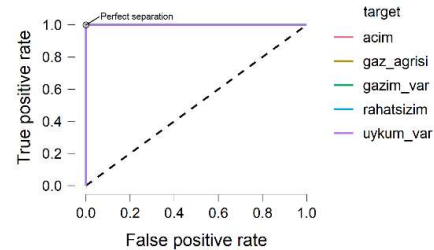
Verilerimizde sınıflandırma yapılarak karar verme sağlanmaktadır. Bunun için aşağıdaki denklemde Rasgele Orman Sınıflandırma algoritması kullanılmaktadır.

$$Gini = 1 - \sum_{i=1}^C (p_i)^2 \quad (1)$$

Şekil 7'de JASP veri analiz programı ile Rasgele Orman algoritması kullanılarak, başarı durumu grafiği gösterilmiştir.



Şekil 7. Sınıflandırma doğruluk grafiği



Şekil 8. Roc eğrisi grafiği  
(Roc curve plot)

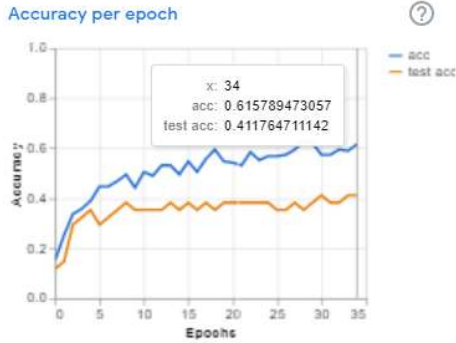
Şekil 8’de ise doğru ve yanlış verilmiştir. Grafiğe bakıldığında %87,5 başarı oranı bulunmuştur.

### 3.2. Öğretilebilir Makine (Teachable Machine)

Ses, görüntü ve poz-hareket noktalarını içeren sinir ağını tarayıcı üzerinde sınıflandırabilen ve sistemde hem test hem de eğitim olanağı sunan bir web aracıdır. Çıktı olarak bir ses veya görüntüye bağlayabilirsiniz.

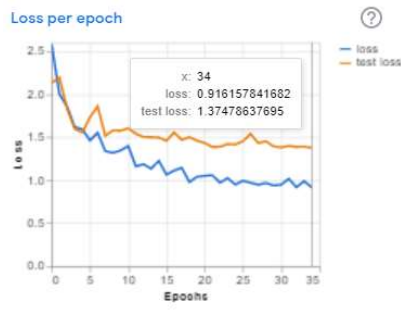
Çalışmalar için makine öğrenme modelleri oluşturmayı hızlı ve kolay hale getiren, kodlama gerektirmeyen bir web aracıdır. Resimlerinizi, seslerinizi ve pozlarınızı tanıması için eğitin, ardından modelinizi siteleriniz, uygulamalarınız ve daha fazlası için bilgisayar ortamında dışa aktarımı sağlamaktadır.

<https://teachablemachine.withgoogle.com/> sitesi, algoritmaları Bulut’tan kullanıcının alanına, masaüstü veya telefonlarına veya diğer akıllı cihazlarına kaydırırken yapay zekâ kişiselleştirme trendini yansıtan bir harekettir. Çalışmada Teachable Machine platformunda veri analizi için Şekil 9’daki başarı grafiği ve Şekil 10’da ise hata



Şekil 9. Doğruluk test sonuçları  
(Accuracy test results)

grafiki verilmiştir. Test sonuçlarına bakıldığında ses tanıma için istenilen başarı elde edilememiştir. Teachable Machine platformu eğitim işleminden sonra başarı ve kayıp oranlarının hesaplamasında %15 olarak test verisi almaktadır.



Şekil 10. Dönem kaybı grafiği  
(loss per epoch)

Şekil 10’da dönem başına kayıp oranları verilmektedir. Dönem kayıpları incelendiğinde test sonuçlarının düşük olmasını doğrulamaktadır.

### 3.3. Python ile Konuşma Tanıma (Python Speech Recognizer)

Speech Recognizer ses ve konuşma tanıma üzerine Python kütüphanesi olarak kullanılmaktadır. Sesleri tanıyarak bizlere, istenilen şekilde (ses, yazı vb.) çıktı vermektedir.

Verilen analizinde Çoklu Sınıf Sınıflandırma algoritması da kullanılmıştır. Bu algoritma ikiden fazla sınıflı sınıflandırma olan veri analizlerinde kullanılmaktadır. Her örnek yalnızca bir sınıf olarak etiketlenebilir. Verilerden elde edilen aşağıdaki sonuçlara bakıldığında %97,5 başarı durumu bulunmuştur.

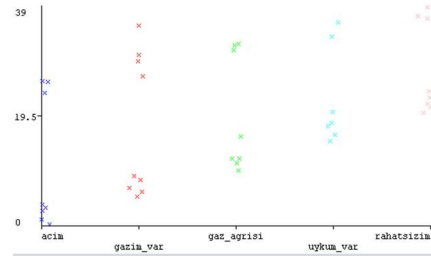
Time taken to build model: 0.13 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	39	97.5 %
Kappa statistic	0.9687	
Mean absolute error	0.0203	
Root mean squared error	0.1225	
Relative absolute error	6.3487 %	
Root relative squared error	30.5428 %	
Total Number of Instances	40	

V1



Şekil 11. Test sonuçları çizim matrisi  
(Test results draqing matrix)

Şekil 11’e bakıldığında test sonuçları sıra numaralarına göre hedef alanları gösterilmiştir. Grafikteki ilk 5 renkteki alan eğitim verilerini belirtmektedir. Toplamda 25 eğitim verisi ile acim, gazim\_var, gaz\_ agrisi, uyku\_var ve rahatsızım değerlerinin her biri farklı renkte grafikte gösterilmiştir.

Şekil 12’de Çoklu Sınıf Sınıflandırma yöntemi ile tüm sınıfların doğruluk oranları verilmiştir. Grafiğe göre uyku\_var sınıfında gerçek değer pozitif iken negatif sınıflandırıldığı görülmektedir.

Kesinlik (Recall) pozitif öngörü de ise gazi\_var sınıfında hatalı sınıflandırma yapılmıştır.

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	1,000	0,000	1,000	1,000	1,000	1,000	1,000	1,000	acim
	0,889	0,000	1,000	0,889	0,941	0,928	0,892	0,915	gazim_var
	1,000	0,000	1,000	1,000	1,000	1,000	1,000	1,000	gaz_agrisi
	1,000	0,030	0,875	1,000	0,933	0,921	0,983	0,909	uykum_var
	1,000	0,000	1,000	1,000	1,000	1,000	1,000	1,000	rahatsızım
Weighted Avg.	0,975	0,005	0,978	0,975	0,975	0,970	0,973	0,965	

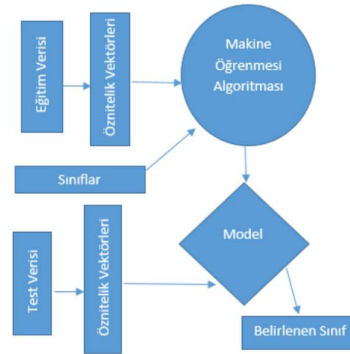
Şekil 12.Çoklu sınıf sınıflandırma yöntemi ile test sonuçları  
(Test results with multi-class classification method)

Tablo 1 Test karşılaştırmaları  
(Test comparisons)

Sıra No	Dosya Adı	Gerçek	MLFK	Python Konuşma Tanıma	Öğretilebilir Makine
			Tahmin	Tahmin	Tahmin
1.	ac_6.wav	Açım	Açım	Açım	Açım
2.	ac_7.wav	Açım	Açım	Açım	Gaz Ağrısı
3.	ac_8.wav	Açım	Açım	Açım	Gazım Var
4.	uyku_6.wav	Uykum Var	Uykum Var	Uykum Var	Uykum Var
5.	uyku_7.wav	Uykum Var	Gaz Ağrısı	Uykum Var	Gaz Ağrısı
6.	uyku_8.wav	Uykum Var	Uykum Var	Açım	Uykum Var
7.	gaz_6.wav	Gazım Var	Gazım Var	Gazım Var	Rahatsızım
8.	gaz_7.wav	Gazım Var	Gazım Var	Gazım Var	Gazım Var
9.	gaz_8.wav	Gazım Var	Gazım Var	Gazım Var	Rahatsızım
10.	gaz_agri_6	Gaz Ağrısı	Gaz Ağrısı	Gaz Ağrısı	Rahatsızım
11.	gaz_agri_7	Gaz Ağrısı	Gaz Ağrısı	Gaz Ağrısı	Gaz Ağrısı
12.	gaz_agri_8	Gaz Ağrısı	Gaz Ağrısı	Gaz Ağrısı	Açım
13.	rahatsızım_6.wav	Rahatsızım	Gazım Var	Rahatsızım	Açım
14.	rahatsızım_7.wav	Rahatsızım	Rahatsızım	Rahatsızım	Rahatsızım
15.	rahatsızım_8.wav	Rahatsızım	Rahatsızım	Rahatsızım	Uykum Var
			Tahmin:13	Tahmin:14	Tahmin:6

Tablo 1’de 3 ayrı platformda veriler üzerinde yapılan test aşamaları görülmektedir. En iyi tahmin Python Speech Recognizer kodlaması ile sonuçlanmıştır. Bu tahminlere göre CSV veri seti dosyası oluşturularak başarı durumları ortaya çıkarılmıştır. Çalışmada Denetimli Öğrenme Modeli kullanılmıştır. Denetimli öğrenme Denetimli Makine Öğrenmesi sistemin etiketi olan veriler kullanılarak eğitilmesi ile öğrenme gerçekleştirilmiştir [19]. Sistem eğitilmesi aşamasında veri setinde bulunan her bir örneğe ait giriş ve çıkışlar verilir. Sistemin doğrulanması aşamasında öğrenme algoritması kategorisi bilinmeyen bir test verisine, eğitim verisinde bulunan çıkışlardan herhangi birini atamaktadır [20]. Denetimli öğrenme modeli süreci Şekil 13’de verildiği gibi gerçekleşmektedir [21]. MLFK, Öğretilebilir Makine web araçları ,Rasgele Orman Sınıflandırma ve Çoklu Sınıf Sınıflandırma

Algoritmaları denetimli öğrenme model sürecini kullanmaktadır. Denetimli Öğrenme Modelinde sınıflar, eğitim verisi ve test verisi daha önceden belirlenmiştir. Bu belirlenen argümanlarla makine öğrenmesi algoritması uygulanarak model oluşturulmuştur. Bu modele göre test verisindeki alana ait veriler, sınıflandırma işlemine girmektedir.



Şekil 13. Denetimli öğrenme model süreci  
(Supervised learning model process)



Tablo 2'deki analiz sonuçları incelendiğinde Bebek seslerinin tanınmasında %97,5 oran ile Çoklu Sınıf Sınıflandırma algoritması daha iyi performans gösteren başarı vermiştir. Çoklu Sınıf algoritması, K ayrı sınıflara bölünmüş bir dizi eğitim örneği sunulur ve daha önce görülmemiş verilerin hangi sınıflara ait olduğunu tahmin etmek için bir makine öğrenme

modeli oluşturulur. Eğitim verilerini görürken, model her sınıfa özgü kalıpları öğrenir ve bu kalıpları gelecekteki verilerin üyeliğini tahmin etmek için kullanır. Çoklu Sınıf algoritmasının diğer algoritma ve platformlara göre daha iyi performans göstererek, doğruluk verdiği görülmektedir.

Tablo 2. Modellerin performanslarının kıyaslanması  
(Of the performance of the models)

Veri Analizi Teknolojisi	Eğitim	Doğrulama	Test	Doğrulama Sonuçları (%)	Test Sonuçları
ML4K-Rasgele Orman Sınıflandırıcısı( <a href="https://machinelearningforkids.co.uk/">https://machinelearningforkids.co.uk/</a> )	25	7	8	%85,7	%87,5
Python Konuşma Tanıma-Çoklu Sınıf Sınıflandırma	25	7	8	%96	%97,5
Öğretilebilir Makine( <a href="https://teachablemachine.withgoogle.com/">https://teachablemachine.withgoogle.com/</a> )	34		6	%61	%41

#### 4. SONUÇLAR (Results)

Yapay zekâ ve bölümlerinden makine öğrenimi günümüz teknolojilerin vazgeçilmez unsurlarıdır. İhtiyaçların değişmesi ve çağa ayak uydurma neticesinde bu hususlara sevk etmiştir. Bu hususlardan birisi de ses tanımadır. Ses tanıma sistemlerinin en önemli amaçlarından biride insan - bilgisayar arasındaki iletişimde, kullanıcının en yaygın olarak kullandığı, sesli iletişimle sağlamak ve insanların işlerini birçok sektörde kolaylaştıran bilgisayar sistemlerinin daha yaygın kullanımına olanak sağlamaktır. Ebeveynler bebeklerin temel ihtiyaçlarına göre bebek seslerini uygulama yardımıyla analiz edebileceklerdir. Bu 5 kategori örnek teşkil edip, uygulamanın geliştirilmesine olanak sağlayacaktır. Rasgele Orman sınıflandırma, Çoklu Sınıf Sınıflandırma ve Teachable Machine (<https://teachablemachine.withgoogle.com/train>) web aracı ile aynı veri seti üzerinde sınıflandırmalar yapılarak, karşılaştırma yapılmış ve %97,5 oran en başarılı sonuç elde edilmiştir. Bu başarı oranının bu çalışmada çok daha iyi performans gösterdiği görülmektedir. Proje şimdilik uygulama üzerinde çalışmaktadır. Ebeveynlerin devamlı bir şekilde kullanması amacıyla bir elektronik kart üzerinde çalıştırılması, programlanmasıyla uygulanmasının daha kullanışlı olacağı gözükmektedir.

Geliştirilen bu uygulamanın başarıya ulaşması ile konuşma tanıma, seslerin analizi, cep telefonu otomasyon uygulamalarında, fiziksel engelli kişilerin bilgisayar kullanımlarının sağlanmasında, araç sürücülerinin belli işlevleri sesli komutlarla gerçekleştirmelerinde, robotların sesle kontrolünde

ve üretim sahasında veri toplama amacıyla kullanılabilmesi düşünülmektedir.

#### KAYNAKLAR (References)

- [1] Özbey, C., & Bayar, S., Otomatik Ses Tanıma: Türkçe için Genel Dağarcıklı Akustik Model Oluşturulması ve Test Edilmesi. *Akademik Bilişim Konferansı*, Aksaray, 2017. s. 1-6.
- [2] Baştürk, A., Baştürk, N. S., & Qurbanov, O. (2018). A comparative Performance analysis of various classifiers for fingerprint Recognition. *Ömer Halisdemir Üniversitesi Mühendislik Bilimleri Dergisi*, 504-511.
- [3] Atalay, M., & Çelik, E. (2017). Büyük Veri Analizinde Yapay Zekâ Ve Makina Öğrenmesi Uygulamaları. *Mehmet akif Ersoy Üniversitesi Sosyal Bilimler Dergisi*, 156-168.
- [4] Badem, H. (2019). Parkinson Hastalığının Ses Sinyalleri Üzerinden Makine Öğrenmesi Teknikleri İle Tanımlanması. *Ömer Halisdemir Üniversitesi Mühendislik Bilimleri Dergisi*, 630-637.
- [5] Küçükler, A., Yüzkollar, C., Sansli, A., & Sen, F. (2006). Güvenlik Sistemleri İçin Mel Frekans Kepstrum Katsayıları ve Yapay Sinir Ağları Kullanılarak Konuşmacı Tanıma. *Electronic Letters on Science & Engineering*, 8-15.
- [6] Çakır, H., & Okutan, B. (2011). Ses Kontrollü Web Tarayıcı. *Bilişim Teknolojileri Dergisi*, 13-18.

- [7] Özden, C., & Acı, Ç. (2016). Makina öğrenmesi yöntemleri ile yaralanmalı trafik kazalarının analizi: Adana örneği. *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, 266-275.
- [8] A. Graves and N. Jaitly, "Towards End-To-End Speech Recognition with Recurrent Neural Networks.," in *ICML*, 2014, pp. 1764–1772.
- [9] S. Fernández, A. Graves, and J. Schmidhuber, "An Application of Recurrent Neural Networks to Discriminative Keyword Spotting," in *International Conference on Artificial Neural Networks*, 2007, pp. 220–229.
- [10] L. Deng, M. Seltzer, D. Yu, A. Acero, A. Mohamed, and G. Hinton, "Binary Coding of Speech Spectrograms Using a Deep Auto-encoder," in *Interspeech 2010*, 2010, p. 1692–1695.
- [11] A.-R. Mohamed, T. N. Sainath, G. Dahl, B. Ramabhadran, G. E. Hinton, and M. A. Picheny, "Deep Belief Networks Using Discriminative Features For Phone Recognition," in *Acoustics, Speech and Signal Processing*, 2011, p. 5060–5063.
- [12] A. Graves, A. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," in 2013 IEEE International Conference on Acoustics, *Speech and Signal Processing*, 2013, pp. 6645–6649.
- [13] Sepp Hochreiter, & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 1-32.
- [14] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: continual prediction with LSTM.," *Neural Comput.*, vol. 12, no. 10, pp. 2451–71, Oct. 2000.
- [15] Graves, A., & Schmidhuber, J. (2005). Framewise Phoneme Classification with Bidirectional LSTM and Other Neural Network Architectures. *Computer Science*, 1-8.
- [16] Cemal Hanilçi, Figen Ertaş. "Sürekli Saklı Markov Modelleri ile Metinden Bağımsız Konuşmacı Tanıma Parametrelerinin İncelenmesi". *Uludağ Üniversitesi Mühendislik- Mimarlık Fakültesi Dergisi*, Cilt 12, Sayı 1, 2007.
- [17] Baygün, M. K., & Yaldır, A. (2006). Linear Predictive Coding ve Dynamic Time Warping Teknikleri Kullanılarak Ses Tanıma Sistemi Geliştirilmesi. *Bilgi Teknolojileri Kongresi IV* (s. 289-294). Pamukkale: Akademik Bilişim 2006.
- [18] Adelli, R., & Ergüzen, A. (2019). Ses Tanıma Teorisi ve Yöntemleri. 1. *Uluslararası İnnopark Dijital Endüstri Sempozyumu*, (s. 356-364). Konya.
- [19] Bilgin, M. (2017). Gerçek Veri Setlerinde Klasik Makine Öğrenmesi Yöntemlerinin Performans Analizi. 19. *Akademik Bilişim Konferansı* (s. 1-6). Aksaray: INETD - 'Internet Teknolojileri Derneği.
- [20] Kotsiantis, S. B. , Zaharakis, I. , Pintelas, P. , 2007. Supervised machine learning: A review of classification techniques. *Informatica* ,31,249-268
- [21] Afrin, F. , Nahar, I. , 2015. Incremental learning based intelligent job search system, Doktora Tezi-BRAC Üniversitesi.