



# Bebek Ağlamalarının Makine Öğrenmesi Yöntemleriyle Sınıflandırılması

Bengisu Şeyda Durak<sup>1\*</sup>, Levent Bayındır<sup>2</sup>

<sup>1\*</sup> Atatürk Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Erzurum, Türkiye, (ORCID: 0000-0002-0716-6455), [bengisudurak2@gmail.com](mailto:bengisudurak2@gmail.com)

<sup>2</sup> Atatürk Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Erzurum, Türkiye (ORCID: 0000-0001-7318-5884), [levent.bayindir@atauni.edu.tr](mailto:levent.bayindir@atauni.edu.tr)

(İlk Geliş Tarihi 9 Ağustos 2021 ve Kabul Tarihi 10 Kasım 2021)

(DOI: 10.31590/ejosat.979526)

**ATIF/REFERENCE:** Durak, B., Bayındır, L. (2021). Bebek Ağlamalarının Makine Öğrenmesi Yöntemleriyle Sınıflandırılması. *Avrupa Bilim ve Teknoloji Dergisi*, (27), 784-791.

## Öz

Bu çalışmada bebek ağlama seslerinden yararlanılarak, bebeklerin ihtiyaçlarını belirleyebilmek ve herhangi bir hastalığa maruz olup olmadıklarını öğrenmek için çeşitli makine öğrenmesi yöntemleri kullanılmıştır. Bebek ağlama sesleri için iki farklı veri kümesinden yararlanılmıştır. Veri kümeleri eğitim, öğrenme ve test aşamalarından geçtikten sonra ağlama türünün bulunması hedeflenmiştir. Ağlama seslerinin özniteliklerini çıkarmak için Mel Frekans Kepstrum Katsayıları (MFCC) ve Doğrusal Öngörü Kepstrum Katsayıları (LPCC) yöntemleri karşılaştırılmış ve MFCC yönteminin doğruluğu arttırmada LPCC'ye göre daha etkili olduğu bulunmuştur. Öznitelikleri çıkarılan ses sinyallerinin sınıflandırılma aşamasında makine öğrenme algoritmalarından k-En Yakın Komşuluk (k-NN) algoritması, Çok Katmanlı Algılayıcılar, Karar Ağacı ve Rastgele Orman algoritmaları kullanılmış ve başarı oranları karşılaştırılmıştır. Sonuçta MFCC ve Çok Katmanlı Algılayıcılar yöntemi kullanılarak %93 başarı oranı bulunmuştur.

**Anahtar Kelimeler:** Bebek ağlaması sınıflandırma, Makine öğrenmesi, MFCC, LPCC.

## Classification of Infant Cries with Machine Learning Methods

### Abstract

In this study, various machine learning methods were used to determine the needs of babies and to find out whether they were exposed to any disease by making use of baby crying sounds. Two different data sets were used for baby crying sounds. It is aimed to find the type of crying after the data sets go through the training, learning and testing stages. Mel Frequency Cepstral Coefficients (MFCC) and Linear Prediction Cepstral Coefficients (LPCC) methods were compared to extract the attributes of crying sounds, and the MFCC method was found to be more effective than LPCC in increasing accuracy. In the classification phase of the audio signals whose attributes were extracted, k-Nearest Neighborhood (k-NN) algorithm, Multilayer Perceptron, Decision Tree and Random Forest algorithms from machine learning algorithms were used and success rates were compared. As a result, a 93% success rate was found using MFCC and Multilayer Perceptron.

**Keywords:** Infant cry classification, Machine learning, MFCC, LPCC.

\* Sorumlu Yazar: [bengisudurak2@gmail.com](mailto:bengisudurak2@gmail.com)

## 1. Giriş

Yeni doğan bebeklerin iletişim kurabildiği ve duygularını ifade edebildiği tek yol ağlamaktır. Bu nedenle bebekler ihtiyaçlarının karşılanmasını istediklerinde ebeveynlerine ağlayarak belli ederler. Ağlama nedenleri açlık, uyku, yorgunluk, gaz sancısı, ağrı gibi nedenlerden kaynaklanabilir (Bănică vd., 2016). Pediatristler gibi uzman kişiler bebeğin ağlayış şekline gönderdiği mesajı kolay bir şekilde anlayabilir. Ancak ebeveynler için bebeğin ağlama sebebini belirleyebilmek pek kolay değildir. Bu, bebeklerin ihtiyaçlarına cevap vermek isteyen anne ve babalar için gerçek bir sorundur. Bebek ağlaması ile ilgili yapılan çalışmalar yıllardır devam etmekte ve farklı alanlarda ele alınmaktadır. Bebeğin psikolojik, fizyolojik, pediatrik durumunu öğrenmek için birçok farklı çalışma yapılmıştır.

Saha ve arkadaşlarının yaptığı çalışmada gömülü bir cihaz yardımıyla bebek ağlama sesleri sınıflandırılmıştır. Cihaza bağlı olan harici bir mikrofon, bebeğin ağlama seslerini kaydeder ve cihaz ağlama nedenini açlık, ağrı, ıslak bebek bezi ve diğer ağlama türleri olarak göstermektedir (Saha vd., 2013). Başka bir çalışma yine bebeklerin ihtiyaçlarını tahmin etmek için 0-3 aylık 127 bebekten 6 tür ağlama içermektedir. Gauss Karışım modeli ile birlikte i-vektör algoritması (Dehak vd., 2011) kullanılarak yapılan deneylerde sınıfların (kolik, açlık, rahatsızlık, gaz sancısı, ağrı ve yorgunluk) birbirlerinden ayrıştırıcı olmadığı bulunmuş ve çalışmanın devamında birbirlerine daha çok karıştırılan rahatsızlık, açlık ve yorgunluk sınıfları kullanılmıştır. Sonuçta i-vektörlerin kullanıldığı deneyler ile rahatsızlık ağlamaları %5.7, açlık ağlamaları %67.9 ve ağrı çığlıklarının %70.4 oranında doğru bir şekilde sınıflandırıldığı bulunmuştur (Bănică vd., 2016).

2018 yılında yapılan bir çalışmada, Raspberry Pi kullanılarak bebeğin ağlama nedeninin açlık, ağrı ya da ateş olduğu belirlenmiştir. Çalışmada bir mikrofon yardımıyla ham ağlama sinyali kaydedilmiş ve sinyali işlemek için Raspberry Pi aygıtı kullanılmıştır. Ağlama sinyalinin tespitinde ses aktivitesi algılama (VAD) algoritması kullanılmakta ve sinyal sınıflandırıldıktan sonra sonuçlar bir ekranda görüntülenmektedir (Patil ve Kamerikar, 2018).

Başka bir çalışmada ise bebeğin neden ağladığını bulmak için bebeğin yüzünün görüntüsünden ve ağlama sesinden yararlanılan bir sistem tasarlanmıştır. Bu çalışmada acı, açlık, öfke, üzüntü ve korku nedeniyle ağlamalar sınıflandırılmıştır. Çalışmada bebeğin yüzü ve sesi bağımsız analiz edilip, bir füzyon sistemi ile birleştirilmiştir. Görüntü işleme sistem; ağız, gözler ve kaşlar gibi temel özelliklerden yararlanarak bebeğin ağlama nedenini tanımlamış ve sonuçta %64 doğruluk bulunmuştur. Ses işleme modülünde ise ağlama sesinin temel frekans (F0) ve formant frekans (F1, F2, F3) değerlerine bakılmış ve %74 doğruluk bulunmuştur. Sonuçta bu iki sistem birleştirilerek %75,2 doğruluk bulunmuştur (Pal vd., 2006).

Galaviz ve Garcia'nın yaptıkları çalışmada ise bebeğin duygu durumu yerine pediatrik durumunun belirlenmesi amaçlanmıştır. Normal, sağır ve asfiksi hastası bebeklerin sınıflandırılması yapılmış ve bebeklerin ağlama seslerinden doğdukları anda patolojik bir hastalığa sahip olup olmadıkları bulunmuştur. Normalde bebeklerin hastalıkları doğumdan sonra ilk bir yıl içinde tespit edilirken, bu çalışma sayesinde bebek doğduğu anda hastalık tespiti yapılabilmiş ve erken teşhis ile büyük ölçüde hastalıkların önüne geçilmiştir. Çalışmada Mel Frekans Kepstrum Katsayıları (MFCC) ve Doğrusal Öngörü Katsayıları (LPC)

yöntemleri karşılaştırılmış ve MFCC yöntemi ile %86 doğruluk oranı bulunmuştur (Reyes-Galaviz ve Reyes-Garcia, 2004).

Başka bir çalışmada yine bebek ağlama sesleri ve hastalıkların ilişkileri tespit edilmiş ve 0-7 aylık bebekler arasından normal ve asfiksi hastası bebeklerin sınıflandırılması yapılmıştır. Çalışmada çok katmanlı algılayıcıların sınıflandırma performansına en küçük kareler yönteminin etkisi tartışılmıştır. Sonuçta en küçük kareler yönteminin hesaplama yükünü azalttığı ve çok katmanlı algılayıcıların doğruluk oranını artırarak %94 doğruluk performansı bulunmuştur (Zabidi vd., 2010).

Yapılan diğer bir çalışmada normal ve sağır bebeklerin sınıflandırılması için doktorlar tarafından ICD-67 Sony dijital kayıt cihazlarıyla 31 bebekten toplanan ses kayıtları kullanılmıştır. İki farklı deney yapılmış ve ses kayıtları bir saniyelik ve üç saniyelik numunelere ayrılmıştır. MFCC yöntemi kullanılan çalışmada üç saniyelik numunelerle sınıflandırma doğruluğu, aynı parametre setini kullanan bir saniyelik numunelerden daha büyük çıkmış ve sonuçta %97.43 doğruluk oranı bulunmuştur (Garcia ve Reyes Garcia, 2003).

Lavner ve arkadaşlarının çalışmalarındaki amaç ise bebek ile bakıcı arasındaki iletişimi kolaylaştırmak ve aralarında bağ kurmalarını sağlamaktır. Çalışmada 0-6 aylık bebeklerin sesleri ev ortamında 7/24 kayıt altına alınmıştır. Lojistik Regresyon ve Evrişimli Sinir Ağları (CNN) olmak üzere iki farklı makine öğrenmesi algoritması karşılaştırılmıştır. Sonuçlar Evrişimli Sinir Ağlarının daha avantajlı olduğunu göstermektedir (Lavner vd., 2016).

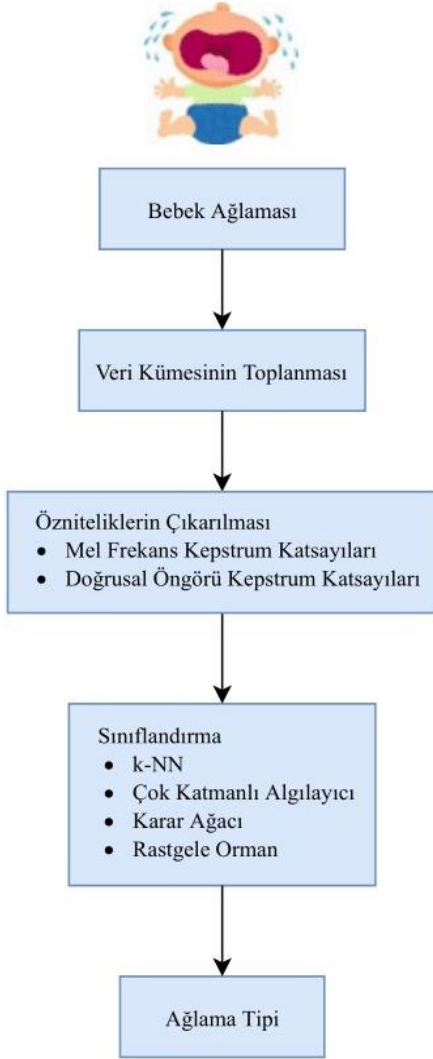
Myakala ve arkadaşlarının yaptığı Akıllı Ağlama Algılama Sistemi bir mikrofon ve Raspberry Pi 3 ile gerçek zamanlı olarak ağlama sinyallerini yakalayıp, sinyal işleme yapmaktadır. Sistem sınıflandırma işleminden sonra Wi-Fi üzerinden ebeveynlerin telefonlarına SMS gönderir. Bebeğin ihtiyacına göre SMS içeriği acil/çok acil gibi geri bildirimler de sunmaktadır. Bu da bebeklerin yanlarında olmayan ebeveynler için uzaktan bebeğin ihtiyaçlarını anlamada büyük kolaylık sağlamıştır (Myakala vd., 2017).

2020 yılında Bütüner'in yaptığı çalışmada sağır ve sağlıklı ebeveynlerin çocuklarının ihtiyaçlarını karşılayabilmeleri için bir uygulama geliştirilmiştir. Bu uygulamada makine öğrenmesi yöntemleri ile birlikte ses ve konuşma sistemlerinden yararlanılmıştır. Uygulama ses tanıma sisteminin sonucunda ekranda yazılı ve sesli olarak bebeğin ihtiyacını göstermekte ve engelli ebeveynler için büyük kolaylık sağlamaktadır (Bütüner, 2020).

Bugüne kadar yapılan çalışmalara bakıldığında teknolojinin gelişmesiyle birlikte ses tanıma sistemlerinin birçok alanda kullanıldığı ve yapılan çalışmaların makine öğrenmesi yöntemleri ile başarılı sonuçlar verdiği görülmektedir. Bu çalışmada, farklı tür ağlama seslerini makine öğrenmesi yöntemleri kullanarak sınıflandıran otomatik bebek ağlama tanıma sistemi oluşturulmuştur. İki ayrı bebek ağlaması veri kümesi üzerinde çalışmalar yapılmıştır. Ağlama sesleri çeşitli öznelik çıkarma ve sınıflandırma aşamalarından geçerek bebeğin neden ağladığının bulunması ve bebek dilinin anlaşılması hedeflenmektedir. Kullanılacak olan öznelik çıkarma ve sınıflandırma yöntemlerinin başarı oranlarının karşılaştırılması amaçlanmaktadır.

## 2. Materyal ve Metot

Bu çalışmada bebek ağlama seslerinin sınıflandırılması için Python programlama dili tercih edilmiş ve Jupyter Notebook platformu kullanılmıştır. Çalışma bebek ağlaması veri kümesinin oluşturulması, özniteliklerin çıkarılması ve sınıflandırma olmak üzere üç temel aşamadan oluşmaktadır. Bu aşamalar Şekil 1’de gösterilmekte ve alt başlıklar halinde aşağıda açıklanmaktadır.



Şekil 1. Bebek Ağlama Tipi Sınıflandırma Sistemi Adımları

### 2.1. Veri Kümesinin Toplanması

Bu çalışmada bebek ağlama tipi sınıflandırılması için iki ayrı veri kümesi kullanılmıştır. Birinci veri kümesi “.wav” uzantılı, 1 sn uzunluğundaki ses dosyalarından oluşan hazır bir veri kümesidir ve Baby Chillanto (Reyes-García, 2006) veri kümesi olarak bilinmektedir. Veri kümesinde toplam 5 sınıf ve 2268 bebek ağlama sesi bulunmaktadır. İkinci veri kümesi ise İnternet üzerinde bulunan videolarda tespit edilen bebek ağlama seslerinin kesilmesi ile “.wav” uzantılı, 3-5 sn uzunluk aralıklarındaki ses dosyalarından oluşturulmuştur. Veri kümesinde toplam 3 sınıf ve 134 bebek ağlama sesi bulunmaktadır. Oluşturulan veri kümelerinin ağlama türlerine göre dağılımı Tablo 1’de gösterilmiştir.

Tablo 1. Veri Kümelerinin Özellikleri

Veri Kümesi-1	
Sınıf	Örnek Sayısı
Aç	350
Acı çeken	192
Normal	507
Asfiksi	340
Sağır	879
<b>Toplam</b>	<b>2268</b>
Veri Kümesi-2	
Sınıf	Örnek Sayısı
Aç	30
Uykulu	47
Kolik	57
<b>Toplam</b>	<b>134</b>

### 2.2. Özniteliklerin Çıkarılması

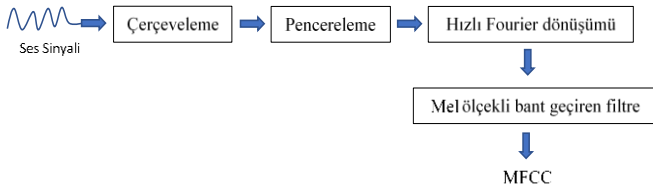
Bebek ağlama seslerini sınıflandırma yapmadan önce ses kayıtlarının özniteliklerinin oluşturulması gereklidir. Özniteliklerin çıkarılmasının amacı sinyaldeki akustik özelliklerden ödün vermeden yüklü miktardaki karmaşık veriyi özetlemektir (Başbuğ, 2019). Bu çalışmada özniteliklerin çıkarılması için Mel Frekans Kepstrum Katsayıları (MFCC) ve Doğrusal Öngörü Kepstrum Katsayıları (LPCC) yöntemleri kullanılmaktadır ve bu yöntemler aşağıda alt başlıklar halinde açıklanmıştır.

#### 2.2.1. Mel Frekans Kepstrum Katsayıları (MFCC)

MFCC yöntemi, ses tanıma problemlerinde en çok kullanılan öznitelik çıkarma yöntemlerinden biridir. Bu yöntem, insan kulağının sesi algılama şeklini taklit eder. İnsan kulağının hassasiyeti 1 kHz’e kadar doğrusal daha yüksek değerler için ise logaritmik olarak devam etmektedir. Gerçek frekans birimi Hertz ile frekans birimi Mel arasındaki dönüşüm Denklem 1 ile sağlanmaktadır (Karasartova, 2011).

$$mel(f) = 2595 * \log\left(1 + \frac{f}{700}\right) \quad (1)$$

MFCC yöntemi ile öznelikleri çıkarma aşamaları Şekil 2'deki blok diyagramında gösterilmiştir.



Şekil 2. MFCC Öznelik Çıkarma Yönteminin Aşamaları

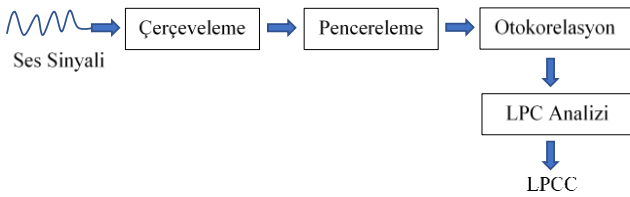
MFCC'nin ilk adımı olan çerçeveleme adımında, ses sinyali belirli zaman aralıklarına bölünür. Bunun nedeni yapılan çalışmalara göre sesin, 10-30 msn arasında karakteristik özelliklerini göstermesinden kaynaklanmaktadır. Zaman aralıklarına bölünen sese bir örtüşme oranı uygulanır. Örtüşme, çerçevelerin sonundaki önemli bilgilerin kaybolmaması için uygulanır.

İkinci adım olan pencerelemede amaç çerçeveler arasındaki geçişte sürekliliği sağlamaktır. Yaygın olarak kullanılan pencereleme yöntemleri; Kaiser, Hamming, Hanning, Dikdörtgen, Blackman yöntemidir (Eray, 2008). Pencerelemeden sonraki aşama her çerçeveye Hızlı Fourier Dönüşümü (FFT) uygulamaktır. FFT ile n örnekten oluşan zaman alanındaki her çerçeve, frekans alanına çevrilir (Ahmed ve Koçer, 2019).

FFT alınmış olan ses sinyali son aşamada üçgenel pencereler şeklindeki Mel ölçekli filtreden geçirilir ve Mel frekans kepstrumu hesaplanır (Karasartova, 2011). Mel frekans kepstrumu, bir ses sinyalinin güç spektrumunun kısa vadeli bir temsilidir (Ahmed ve Koçer, 2019). Son aşamada bulunan katsayılar ile öznelik vektörü oluşturulur.

### 2.2.2. Doğrusal Öngörük Kepstrum Katsayıları (LPCC)

LPCC yöntemi, ağlama sinyalinin geçmiş örneklerine bakılarak yeni gelen ağlama sinyalinin belirli aşamalardan geçerek tanımlanmasıdır. LPCC yönteminin aşamaları Şekil 3'deki blok diyagramında gösterilmiştir.



Şekil 3. LPCC Öznelik Çıkarma Yönteminin Aşamaları

LPCC yönteminin çerçeveleme ve pencereleme adımları MFCC yöntemi ile aynıdır. Pencerelemiş her bir ses sinyaline otokorelasyon analizi yapılır. Otokorelasyon bir çerçevenin enerjisini belirlemede önemli rol oynar ve ses tanıma sistemlerimde önemli bir işlemdir. Otokorelasyon adımından sonra Doğrusal Öngörük Kodlama (LPC) analizi yapılır. LPC analizi, sıradaki ses örneğinin geçmiş p örneklerine bakarak ağırlıklandırılmış doğrusal kombinasyonu ile Denklem 2'deki gibi hesaplanmaktadır (Eray, 2008).

$$S(n) = \sum_{i=1}^p a_i s(n-i) \quad (2)$$

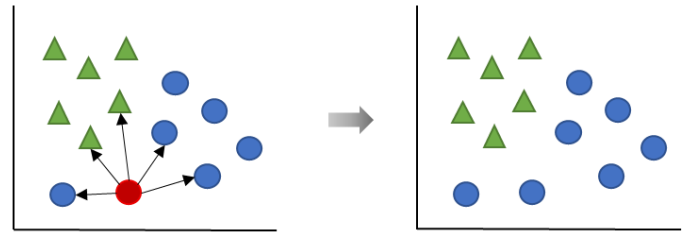
Burada S (n), n zamandaki ses örneği, p LPC'nin derecesi ve  $a_i$  LPC katsayıları olarak ifade edilir. LPC analizi sonucunda oluşan katsayılar ile öznelik vektörü oluşturulur.

## 2.3. Sınıflandırma

Öznelikleri çıkarılan ağlama seslerinin son aşamada hangi sınıfa ait olduklarını bulmak için dört farklı sınıflandırma algoritması üzerinde çalışmalar yapılmıştır. Bu algoritmalar alt başlıklar halinde aşağıda açıklanmaktadır.

### 2.3.1. K En Yakın Komşuluk (k-NN) Algoritması

K en yakın komşuluk (k-NN) algoritması, sınıflandırma ve regresyon problemlerinde sık kullanılan algoritmalarındandır. Algoritma, sınıfları belirli olan örnek veri kümesi kullanır ve veri kümesine yeni katılacak olan verinin en yakın komşuluklarına bakarak sınıfını belirler. Örneğin k=5 olursa yeni gelen veri sınıflandırılırken en yakın 5 nokta belirlenir ve bunların içinde en fazla hangi sınıf varsa yeni verinin sınıfı o olur. k-NN algoritmasının çalışmasına dair sembolik bir örnek Şekil 4'de gösterilmektedir. Yeni gelen kırmızı örneğin hangi sınıfa ait olduğunu bulmak için en yakın 5 komşusuna (k=5) bakıldığında komşuların 2 tanesinin yeşil üçgen, 3 tanesinin ise mavi yuvarlak sınıfı olduğu görülecektir. Bu durumda en çok komşu mavi sınıfta olduğundan yeni gelen örnek mavi yuvarlak olarak sınıflandırılır.



Şekil 4. k-NN En Yakın Beş Komşuluk

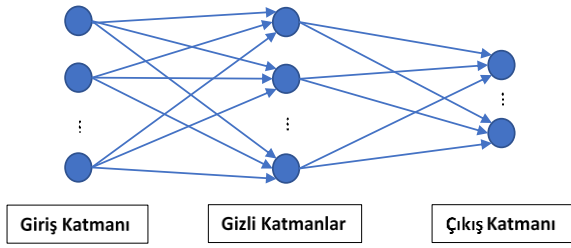
Yeni gelen verinin en yakınındaki komşuluklarına olan uzaklıklarının hesabı için farklı mesafe fonksiyonları kullanılır: Euclidean (Öklid), Manhattan, Minkowski ve Chebyshev. Bu yöntemlerden iki nokta arasındaki mesafe ölçümünde kullanılan Öklid fonksiyonu Denklem 3'deki gibi hesaplanmaktadır.

$$d(i, j) = \sqrt{\sum_{k=1}^n (x_{ik} - x_{jk})^2} \quad (3)$$

### 2.3.2. Çok Katmanlı Algılayıcı Algoritması

Çok Katmanlı Algılayıcılar, en az üç farklı katmandan oluşan ileri beslemeli yapay sinir ağlarıdır. Çok katmanlı algılayıcı algoritması giriş katmanı, gizli katmanlar ve çıkış katmanından oluşur. Sisteme verilen girdi, ara katmanlar olan gizli katmanlardan geçerek çıktı katmanına ulaşır. Girdi katmanı verilerin okunduğu katman, gizli katmanlar ise verilerin ara işlemlerden geçtiği katmandır. Bu katmanların sayısı ve her bir gizli katmanda bulunan nöronların sayısı eğitim aşamasında elde edilen ara sonuçlara göre optimize edilmelidir. Son katman olan çıktı katmanı ise sınıfların olduğu katmandır ve sınıf sayısı kadar nöron içerir (Kaynar vd., 2016). Çok katmanlı algılayıcılara ait sembolik mimari Şekil 5'de gösterilmektedir.





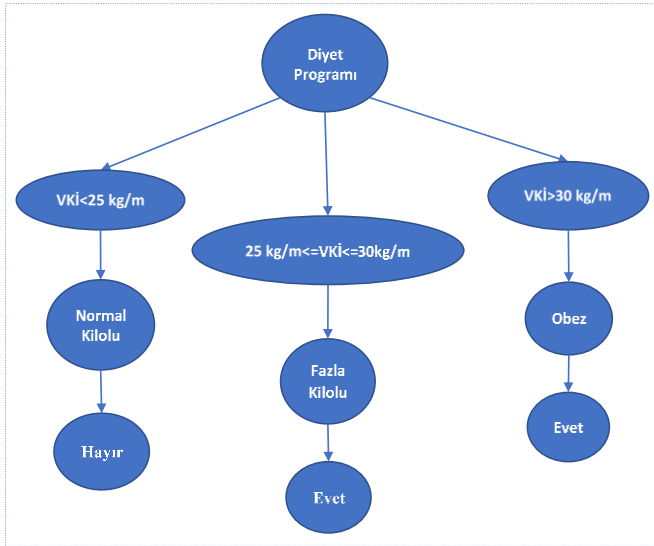
Şekil 5. Çok Katmanlı Algılayıcıların Yapısı

### 2.3.3. Karar Ağacı Algoritması

Yaygın olarak kullanılan başka bir sınıflandırma algoritması Karar Ağacı algoritmasıdır. Bu algoritmada sınıf etiketleri bilinen veri örnekleri kullanılarak kural tabanlı bir ağaç yapısı oluşturulmaya çalışılır. Ağacın en üstteki bileşenine kök, uçlardaki bileşenlerine yaprak ve aradaki bileşenlere ise dal adı verilir.

Bu yöntem, en iyi tahmini elde edebilmek için bağımlı ve bağımsız değişkenler arasında olan tüm ilişkileri araştırır ve en kuvvetli ilişkiye sahip olan bağımsız değişkeni bulduğunda, veri kümesi bu bağımsız değişkenin değerlerine göre ikiye ayrılır. Söz konusu bu süreç olası tüm bölümler tamamlanıncaya kadar devam ettirilir. (Büyükarıkan, 2020).

Şekil 6’da karar ağacı algoritmasına örnek bir problem gösterilmiştir. Bu örnekte vücut kitle indeksine (VKİ) göre diyet programına uyması gereken kişilerin bulunduğu karar ağacı verilmiştir. Burada diyet programı 3 farklı koşul içerir: Vücut kitle indeksi 25’ten küçük olanlar, 25-30 arası ve 30’dan fazla olanlar. Örneğin kişi “Normal Kilolu” ise diyet programına uymaz, ancak “Fazla Kilolu” veya “Obez” ise diyet programına uyması gerekir.



Şekil 6. Karar Ağacı Örneği

### 2.3.4. Rastgele Orman Algoritması

Rastgele Orman algoritması, her ağacın bağımsız olarak örneklenen rastgele bir vektörün değerlerine bağlı olarak üretildiği ve ormandaki tüm ağaçlar için aynı dağılıma sahip olduğu ağaç sınıflandırıcılarının birleşimidir (Breiman, 2001).

Algoritma, her düğümde kullanılacak değişken sayısı (m) ve sınıflandırma ağaçlarının sayısı (n) olmak üzere 2 parametreye ihtiyaç duyar. Rastgele orman algoritması, yeni gelen veriyi sınıflandırmak için istenen sınıflandırma ağaçlarının sayısında (m) rastgele seçilmiş öngörücü değişken kullanır ve veri kümesindeki her örnek n sayısı kadar ağaç tarafından sınıflandırılır. Her bir karar ağacının sınıflandırma sonuçlarına bakıldığında en sık bulunan değere göre yeni gelen verinin sınıfı belirlenir (Rodriguez-Galiano vd., 2012).

## 3. Araştırma Sonuçları ve Tartışma

Bu çalışmada bebek ağlama seslerinin özniteliklerini çıkarmak için MFCC ve LPCC öznitelik yöntemleri karşılaştırılmıştır. MFCC ve LPCC öznitelik sayıları performans olarak daha iyi sonuç verdiği için 40 olarak alınmıştır. Deneysel bulgular MFCC algoritmasının sınıflandırma aşamasında daha verimli performans verdiğini göstermektedir.

Öznitelikleri çıkarılan ağlama seslerini sınıflandırma aşamasında k-NN, çok katmanlı algılayıcılar, karar ağacı ve rastgele orman yöntemleri ile iki farklı veri kümesi üzerinde deneyler yapılmıştır. Deneysel olarak en iyi sonuçların bulunduğu parametreler Tablo 2’de gösterilmektedir. k-NN yönteminde k değeri için 1, 2, 3, 4 ve 5 sayıları denenmiş ve en yüksek performans k=3 için bulunduğu için çalışmanın devamında bu değer kullanılmıştır.

Tablo 2. Sınıflandırma Yöntemlerinde Kullanılan Parametreler

Sınıflandırma Yöntemlerinin Parametreleri		
Sınıflandırma Yöntemi	Parametre	Parametre Değeri
k-NN	k	3
Çok Katmanlı Algılayıcılar	Katman Sayısı	3
	Giriş Katmanı Nöron Sayısı	40
	Gizli Katman Nöron Sayısı	256
	Çıkış Katmanı Nöron Sayısı	Sınıf Sayıları Kadar
	Optimizasyon Yöntemi	Adam
Rastgele Orman	Ağaç sayısı	400

MFCC ve LPCC yöntemleri ile veri kümeleri üzerinde yapılan deneylerin performanslarını ölçmek için doğruluk (accuracy), kesinlik (precision), duyarlılık (recall), F1 skoru (F1 score) ve dengeli doğruluk (balanced accuracy) değerleri karmaşıklık matrisi (confusion matrix) kullanılarak hesaplanmıştır. Karmaşıklık matrisinin yapısı Tablo 3’te gösterilmiştir.

**TP:** Tahmin edilen değer pozitif ve doğru  
**TN:** Tahmin edilen değer negatif ve doğru  
**FP:** Tahmin edilen değer pozitif ve yanlış  
**FN:** Tahmin edilen değer negatif ve yanlış

Tablo 3. Karmaşıklık Matrisi Tablosu

		Gerçek Değerler	
		Pozitif	Negatif
Tahmin Edilen Değerler	Pozitif	TP	FP
	Negatif	FN	TN

$$Doğruluk = \frac{TP+TN}{TP+FP+TN+FN} \quad (4)$$

$$Kesinlik = \frac{TP}{TP+FP} \quad (5)$$

$$Duyarlılık = \frac{TP}{TP+FN} \quad (6)$$

$$F1 = 2 \times \frac{Kesinlik \times Duyarlilik}{Kesinlik + Duyarlilik} \quad (7)$$

$$Dengeli Doğruluk = \frac{1}{2} \times \left( \frac{TP}{TP+FN} + \frac{TN}{TN+FP} \right) \quad (8)$$

Tablo 4 ve Tablo 5'te iki farklı veri kümesi kullanılarak farklı sınıflandırma yöntemleri ile hesaplanan doğruluk, kesinlik, duyarlılık, F1 skoru ve dengeli doğruluk sonuçları gösterilmiştir. Veri kümesi-1 ile yapılan deneylerde, MFCC yöntemi ile çok katmanlı algılayıcılar sınıflandırıcının %93 doğruluk, %93 kesinlik, %92 duyarlılık, %92 F1 skoru %87 dengeli doğruluk değeri ile en iyi şekilde çalıştığını görebiliriz.

Tablo 4. MFCC Yöntemi Kullanarak Farklı Sınıflandırma Yöntemlerinin Karşılaştırılması

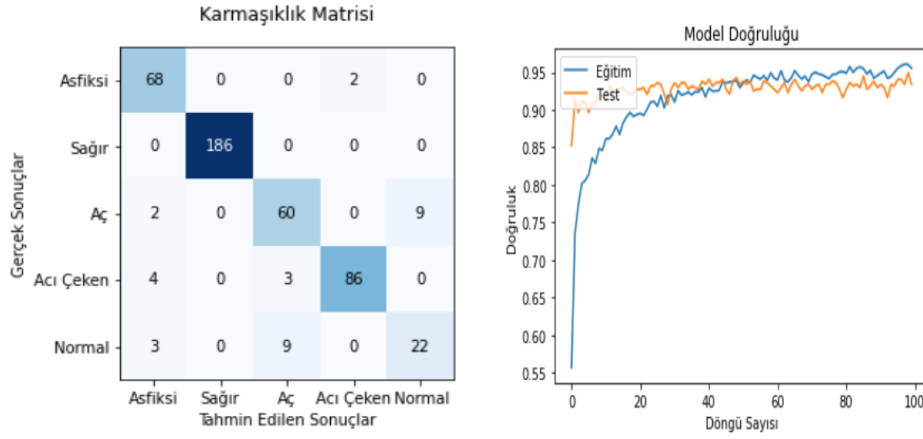
Veri Kümesi-1					
Yöntem	Doğruluk	Kesinlik	Duyarlılık	F1	Dengeli Doğruluk
3-NN	%91	%92	%91	%91	%85
Çok Katmanlı Algılayıcılar	%93	%93	%92	%92	%87
Karar Ağacı	%87	%87	%87	%87	%88
Rastgele Orman	%90	%96	%90	%93	%84
Veri Kümesi-2					
Yöntem	Doğruluk	Kesinlik	Duyarlılık	F1	Dengeli Doğruluk
3-NN	%88	%88	%88	%88	%90
Çok Katmanlı Algılayıcılar	%88	%85	%88	%87	%85
Karar Ağacı	%62	%62	%62	%62	%67
Rastgele Orman	%62	%94	%62	%75	%76

Tablo 5. LPCC Yöntemi Kullanarak Farklı Sınıflandırma Yöntemlerinin Karşılaştırılması

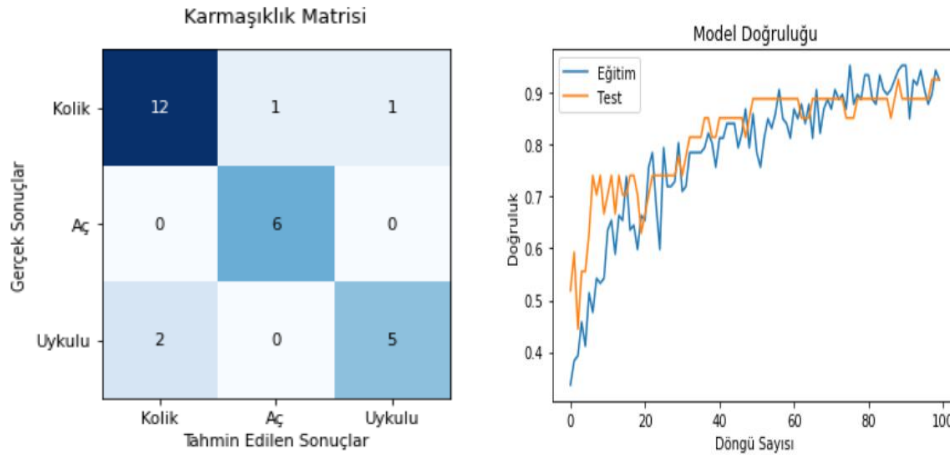
Veri Kümesi-1					
Yöntem	Doğruluk	Kesinlik	Duyarlılık	F1	Dengeli Doğruluk
3-NN	%63	%71	%63	%66	%55
Çok Katmanlı Algılayıcılar	%60	%77	%64	%70	%54
Karar Ağacı	%58	%58	%58	%58	%48
Rastgele Orman	%53	%87	%53	%66	%52
Veri Kümesi-2					
Yöntem	Doğruluk	Kesinlik	Duyarlılık	F1	Dengeli Doğruluk
3-NN	%66	%69	%66	%67	%62
Çok Katmanlı Algılayıcılar	%70	%75	%81	%78	%75
Karar Ağacı	%62	%62	%62	%62	%68
Rastgele Orman	%70	%79	%70	%74	%64

İki farklı veri kümesi üzerinde yapılan deneyler sonucunda en yüksek performansları veren MFCC öznelilik çıkarım yöntemi ile çok katmanlı algılayıcılar sınıflandırma yönteminin iki farklı veri

kümesinde hesaplanan karmaşıklık matrisi ve doğruluk performansları Şekil 7'de ve Şekil 8'de gösterilmektedir.



Şekil 7. Veri Kümesi-1 ile Hesaplanan Karmaşıklık Matrisi ve Doğruluk Test Sonucu



Şekil 8. Veri Kümesi-2 ile Hesaplanan Karmaşıklık Matrisi ve Doğruluk Test Sonucu

#### 4. Sonuç

Bu çalışmada bebek ağlama seslerinin sınıflandırılıp bebeğin ihtiyacının belirlenmesi ve aynı zamanda bebeğin doğduğunda sağrlık, kolik, asfiksi gibi hastalıklara maruz olup olmadığını belirlemek için makine öğrenmesi yöntemleri ile deneyler yapılmıştır. İlk aşamada öznelikleri çıkarmak için MFCC ve LPCC yöntemleri karşılaştırılmıştır. Sonuçta LPCC yönteminin aksine MFCC yöntemi, insan kulağının duyma yönteminden esinlendiği için performansının LPCC'ye göre daha yüksek olduğu bulunmuştur. Ayrıca LPCC yönteminin daha küçük veri kümesi olan Veri Kümesi-2'de daha yüksek performans gösterdiği bulunmuştur. Çalışmanın devamı olan sınıflandırma aşamasında ise k-NN, çok katmanlı algılayıcılar, karar ağacı ve rastgele orman algoritmaları karşılaştırılmıştır. İki ayrı veri kümesi üzerinde yapılan deneylere göre en yüksek performansları çok katmanlı algılayıcılar yönteminin verdiği bulunmuştur. Sonuçta Veri Kümesi-1 kullanılarak, MFCC yöntemi ve çok katmanlı algılayıcılar yöntemi ile yapılan deneyde %93 başarı oranı bulunmuştur. Bu sonuç ebeveynlerin bebek ağlama seslerini tanımlayabilmeleri ve aynı zamanda sağrlık, kolik, asfiksi gibi hastalıkların erken teşhisi için büyük oranda kolaylık sağlamaktadır.

Bu sistemi ebeveynlerin daha kolay bir şekilde kullanabilmesi için mikrofon vb. donanımlar yardımı ile projenin geliştirilmesi ve sınıf sayılarının artırılması planlanmaktadır. Diğer yandan farklı öznelik ve sınıflandırma yöntemleri de kullanılarak daha iyi sonuçlar elde edilebilir.

#### Kaynakça

- Ahmed, M. C. and Koçer, H. E. (2019). Turkish Speech recognition using Mel-frequency cepstral coefficients(MFCC) and Hidden Markov Model (HMM). *Veri Bilimi Dergisi*, 2(2), 39-44.
- Bănică, I.-A., Cucu, H., Buzo, A., Burileanu, D. and Burileanu, C. (2016). Baby cry recognition in real-world conditions. *2016 39th International Conference on Telecommunications and Signal Processing (TSP)*, 315-318. Vienna. doi:10.1109/TSP.2016.7760887
- Başbuğ, A. M. (2019). Ses olayı tanıma ve akustik sahne geri getirme. Yüksek Lisans Tezi, Başkent Üniversitesi Fen Bilimleri Enstitüsü, Ankara.
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45, 5-32. doi:https://doi.org/10.1023/A:1010933404324
- Bütüner, R. (2020). Yapay zeka yöntemleri ile bebek seslerinin tahmin edilmesi. *Journal of Information Systems and Management Research*, 2(1), 32-41.

- Büyükarıkan, U. (2020). Finansal performansa etki eden finansal değişkenlerin CHAID karar ağacıyla belirlenmesi: tekstil sektörü örneği. *Aydın İktisat Fakültesi Dergisi*, 5(1), 1-10.
- Dehak, N., Kenny, P., Dehak, R., Dumouchel, P. and Ouellet, P. (2011). Front-End Factor Analysis for Speaker Verification. *IEEE Transactions on Audio, Speech, and Language Processing*, 19(4), 788-798. doi:10.1109/TASL.2010.2064307
- Eray, O. (2008). Destek vektör makineleri ile ses tanıma uygulaması. Yüksek Lisans Tezi, Pamukkale Üniversitesi Fen Bilimleri Enstitüsü, Denizli.
- Garcia, J. O. and Reyes Garcia, C. A. (2003). Mel-frequency cepstrum coefficients extraction from infant cry for classification of normal and pathological cry with feed-forward neural networks. *Proceedings of the International Joint Conference on Neural Networks*, 3140-3145. Portland. doi:10.1109/IJCNN.2003.1224074
- Güldal, H. ve Çakıcı, Y. (2017). Ders yönetim sistemi yazılımı kullanıcı etkileşimlerinin sınıflandırma algoritmaları ile analizi. *Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 21(4), 1355-1367.
- Karasartova, S. (2011). Metinden bağımsız konuşmacı tanıma sistemlerinin incelenmesi ve gerçekleştirilmesi. Yüksek Lisans Tezi, Ankara Üniversitesi Fen Bilimleri Enstitüsü, Ankara.
- Kaynar, O., Görmez, Y., Yıldız, M. ve Albayrak, A. (2016). Makine öğrenmesi yöntemleri ile duygu analizi. *International Artificial Intelligence and Data Processing Symposium*, 234-241. Malatya.
- Lavner, Y., Cohen, R., Ruinskiy, D. and IJzerman, H. (2016). Baby cry detection in domestic environment using deep learning. *2016 IEEE International Conference on the Science of Electrical Engineering (ICSEE)*, 1-5. Eilat. doi:10.1109/ICSEE.2016.7806117
- Myakala, P. R., Nalumachu, R., Sharma, S. and Mittal, V. K. (2017). An intelligent system for infant cry detection and information in real time. *2017 Seventh International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*, 141-146. San Antonio. doi:10.1109/ACIIW.2017.8272604
- Pal, P., Iyer, A. N. and Yantorno, R. E. (2006). Emotion detection from infant facial expressions and cries., 721-724. Toulouse. doi:10.1109/ICASSP.2006.1660444
- Patil, M. S. and Kamerikar, U. A. (2018). Cause estimation of neonatal cry using raspberry pi. *International Journal of Applied Engineering Research*, 13(12), 10722-10726.
- Reyes-Galaviz, O. F. and Reyes-Garcia, C. A. (2004). A system for the processing of infant cry to recognize pathologies in recently born babies with neural networks. *9th International Conference on Speech and Computer*. St. Petersburg.
- Reyes-García, C. A. (2006). Baby Chillanto Database. Mexican National Institute of Author Rights (INDAUTOR).
- Rodriguez-Galiano, V. F., Ghimire, B., Rogan, J., Chica-Olmo, M. and Rigol-Sanchez, J. P. (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 67(0924-2716), 93-104. doi:10.1016/j.isprsjprs.2011.11.002
- Saha, B., Purkait, P. K., Mukherjee, J., Majumdar, A. K., Majumdar, B. and Singh, A. K. (2013). An embedded system for automatic classification of neonatal cry. *2013 IEEE Point-of-Care Healthcare Technologies (PHT)*, 248-251. Bangalore. doi:10.1109/PHT.2013.6461331
- Zabidi, A., Khuan, L. Y., Mansor, W., Yassin, I. M. and Sahak, R. (2010). Classification of infant cries with asphyxia using multilayer perceptron neural network. *2010 Second International Conference on Computer Engineering and Applications*, 204-208. Bali. doi:10.1109/ICCEA.2010.47