



Türkçe Müzikten Duygu Tanıma

Serhat Hızlısoy^{1*}, Zekeriya Tüfekci²

¹ Çukurova Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Adana, Türkiye (ORCID: 0000-0001-8440-5539)

² Çukurova Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Adana, Türkiye (ORCID: 0000-0001-7835-2741)

(1st International Conference on Computer, Electrical and Electronic Sciences ICCEES 2020 – 8-10 Ekim 2020)

(DOI: 10.31590/ejosat. 802169)

ATIF/REFERENCE: Hızlısoy, S. & Tüfekci, Z. (2020). Türkçe Müzikten Duygu Tanıma. *Avrupa Bilim ve Teknoloji Dergisi*, (Özel Sayı), 6-12.

Öz

Müzikten duygu tanıma yapılması, günümüzde hala oldukça zor bir görevdir. Bu çalışmada, müzikten duygu tanıma yapılması için genel problemler tespit edilmiş, bu problemlerin üstesinden gelmek ve sınıflandırma başarısını artırmak için yaklaşımlar geliştirilmiştir. Bu amaçla, çeşitli makine öğrenmesi yöntemleri ve farklı araçlardan elde edilen öznelilikler kullanılarak Türkçe müziklerden duygu tanınması yapılmak istenmiştir. Yöntem olarak Bayes Ağları, Sıralı Minimal Optimizasyon (SMO), Karar Ağaçları (J.48) ve Lojistik Regresyon kullanılmıştır. Bu yöntemler, duygu tanıma yapmak için oluşturulan bir veri tabanı üzerine uygulanmış ve performansları ölçülmüştür. Bu veri tabanı her biri 30 saniyelik 124 müzik alıntısından oluşan Türkçe Duygusal Müzik Veri Tabanı'dır. Müzik sinyallerinden öznelilik elde etmek için ise, yapılan çalışmalarda sık sık karşımıza çıkan ve öznelilik çıkarma sırasında karşılaşılan sorunlara kapsamlı çözüm sağlayan çeşitli araçlar kullanılmıştır. Bu araçlar çok sayıda farklı öznelilik elde etmemize olanak sağlar. Buna ek olarak gereksiz olan öznelilikleri çıkarmak ve sınıflandırıcı performansını artırmak amacıyla korelasyon tabanlı öznelilik seçme yöntemi (Correlation-based Feature Selection) kullanılmıştır. Her bir araçtan elde edilen özellikler ayrı ayrı kullanılarak, makine öğrenmesi yöntemleri ile birlikte sınıflandırma işlemi yapılmıştır. Sınıflandırma aşamasında sonuçları değerlendirmek ve karşılaştırmak için 10 kat çapraz doğrulama yöntemi uygulanmıştır. Yapılan çalışmada, elde edilen özneliliklere öznelilik seçim yöntemi uygulanarak ve Bayes Ağları sınıflandırıcısı kullanılarak %94.35 oranında doğruluk ile duygu tanıma gerçekleştirilmiş ve diğer sınıflandırıcıların hepsinden daha iyi sonuç alınmıştır. Son olarak, bütün araçlardan elde edilen öznelilikler bir araya getirilmiş ve bu özneliliklere yine seçim işlemi yapılmıştır. Bu işlemden sonra ise, Bayes Ağları kullanılarak elde edilen duygu tanıma oranı %1.6 artarak, %95.96 olmuştur.

Anahtar Kelimeler: Müzik, Duygu Tanıma, Makine Öğrenmesi.

Emotion Recognition From Turkish Music

Abstract

Recognizing emotion from music is still a very difficult task today. In this study, general problems were determined for emotion recognition from music, and approaches were developed to overcome these problems and to increase classification success. For this purpose, emotion recognition from Turkish music was aimed by using various machine learning methods and features obtained from different toolboxes. BayesNet, Sequential Minimal Optimization (SMO), Decision Trees (J.48) and Logistic Regression were used as methods. These methods were applied on a database constructed for emotion recognition and their performance was measured. This database is the Turkish Emotional Music Database consisting of 124 music excerpts of 30 seconds each. In order to obtain features from music signals, various toolboxes have been used that provide comprehensive solutions to the problems encountered frequently during feature extraction. These toolboxes allow us to obtain a large number of different features. In addition, the correlation-based feature selection method (CFS) was used to remove unnecessary features and to increase classifier performance. The classification

* Sorumlu Yazar: Çukurova Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Adana, Türkiye, ORCID: 0000-0001-8440-5539, shizlisoy@cu.edu.tr

was made with machine learning methods, using the features obtained from each toolbox separately. 10-fold cross validation method was applied to evaluate and compare the results at the classification. Accuracy measure was used to evaluate the success of the system. In the study, %94.35 emotion recognition was achieved by using the feature selection method and BayesNet classifier which yielded better results than all other classifiers. Finally, all features are combined and the selection process is made for these features again. After this process, the emotion recognition rate obtained by using BayesNet classifier increased by %1.6 to %95.96.

Keywords: Music, Emotion Recognition, Machine Learning.

1. Giriş

Müziği, birbiriyle uyumlu sesleri bir araya getirerek ritmik bir yapıya sahip anlamlı titreşimler olarak tanımlayabiliriz. Müzik, insanlığın başlangıcından bu yana hayatımızda var olan hem sosyal hem de bireysel olarak önemli bir olgudur. Müzik, insan psikolojisi üzerindeki etkisi ve bir eğlence aracı olarak görülmesi nedeniyle de neredeyse tüm insanlar tarafından sevilmektedir. Müziğin insanlar üzerinde birçok etkisi vardır. Bu etkilerin başında duygu gelir. Çünkü müzik, duyguların ifade edildiği bir sanat olarak da bilinir.

Duygular hayatımızda çok önemli bir rol oynar. Davranışımıza yön vermeye ve o doğrultuda hareket etmemize yardımcı olur. Bu duygular mutluluk, üzüntü, öfke, heyecan ve sakinlik gibi çeşitli başlıklar altında toplanabilir. Duyguların tanımlanması duygu tanıma sistemleri ile gerçekleştirilebilir. İşitsel ve görsel öğeler, duygu tanıma sistemlerinde önemli veriler olarak kullanılır. Bunlara ek olarak, kalp hızı, kan basıncı, beyin dalgaları gibi birçok fizyolojik parametre de kullanılabilir. Konuşmadan duygu tanıma sistemleri veya müzikten duygu tanıma sistemleri, ses öğelerinden duygu tanıma sistemlerine örnek olarak verilebilirler.

Müzik dinlemek insanların duygusal durumunu etkileyebilir. İnsanların duygusal durumları da dinleyecekleri müziğe rehberlik edebilir. Bu nedenle, insanlar mutlu olduklarında hareketli müzik dinleme veya üzgün olduklarında sakin müzik dinleme eğiliminde olabilirler. Son zamanlarda Spotify gibi bu amaca hizmet eden birçok müzik öneri uygulamasının geliştirildiğini görebiliriz. Bu uygulamalar, sinyal işleme veya makine öğrenme tekniklerini kullanarak oluşturulan tanıma sistemleri sayesinde bir müziğin duygusal içeriğini otomatik olarak analiz eder, tanır ve sınıflandırır. Ayrıca müzikten duygu tanıma sistemleri, müzik terapisi ve duygusal bozuklukların tedavisi gibi birçok araştırma için de kullanılır.

Müzikten duygu tanıma yapılması, günümüzde hala oldukça zordur. Çünkü insanların duygu algısı özünde öznel olduğu için aynı müzikten farklı duyguları hissedebilirler. Hissedilen duygu; yaş, cinsiyet, karakter, şarkı sözleri veya çevresel faktörlere göre değişebilir. Dolayısıyla bu tür etmenler dinleyicilerin müzikleri farklı değerlendirmesine neden olur. Bunun dışında, kaliteli veritabanı seçimi, öznel çıkarımı ve doğru sınıflandırma yöntemleri kullanımı gibi ele alınması gereken birkaç önemli konu daha vardır. Bu çalışmada müzikten duygunun tanınması için genel problemler tespit edilmiş, bu problemlerin üstesinden gelmek ve sınıflandırma başarısını artırmak için yaklaşımlar geliştirilmiştir.

İlk olarak, bir müzik duygu tanıma sistemi oluşturmak için etiketlenmiş bir duygusal müzik veritabanına ihtiyaç vardır. Etiketleme kalitesi, bu veriler üzerinde eğitilen yöntemlerin başarısında önemli rol oynamaktadır. Müzikteki duyguları etiketlemek için kategorik ve boyutsal olmak üzere iki yaklaşım vardır. Kategorik yaklaşımda, duygular ayrı etiketler kullanılarak karakterize edilir. Bu alandaki en eski çalışmalardan biri Hevner (1936) tarafından yapılmıştır. Bu çalışmada sekiz kategoride kümelenmiş 66 duygu etiketi vardır. Duyguları sınıflandırmak için müzikten duygu tanıma ile ilgili yapılan ilk çalışma da ise Y. Feng ve ark. (2003) kullanılmak için üzüntü, mutlu, korku, öfke gibi duyguları seçmiştir. Kategorik modeller sorunludur çünkü kategori sayısı üzerinde fikir birliği yoktur. İkinci yaklaşım olan boyutsal modeller söz konusu olduğunda ise bu sorun ortadan kalkar. Duygular boyutsal alanda bağımsız eksenlerle temsil edilirler. Çalışmalarda genellikle iki boyutlu model kullanılır. Boyutsal modellerin avantajı, kategorik yaklaşımla karşılaştırıldığında belirsizliğin azalmasıdır. Russels (1980) tarafından iki boyuttan oluşan bir model önerilmiştir: değerlik ve aktivasyon. Bu model insanların, duyguları iki farklı boyutta ölçmeleri için güvenilir bir yol sağlamıştır. Aktivasyon eksenini, müzikten hissedilen enerji seviyesini göstermek için kullanılırken, değerlik eksenini, haz veya hoşnutsuzluk seviyesini göstermek için kullanılır.

Oluşturulan veritabanları da müzikten duygu tanımadaki önemli bir rol oynamaktadır. Li ve Ogihara (2003), bu konuda çalışmış ve kendi veritabanlarını oluşturmuşlardır. Çalışmalarındaki temel sorunlarından biri, veritabanındaki müzikleri sınıflandırmak için tek bir değerlendiricinin kullanılmasıydı. Panda ve Paiva (2012) ise 5 duygu kümesi ile etiketlenmiş 903 müzikten oluşan bir veritabanı önermişlerdir. Müziğin duygusal analizi için yapılan bir başka veritabanı ise değerlik ve aktivasyon ile etiketlenmiş 1802 parçadan oluşur (Aljanaki, Yang, ve Soleymani, 2017).

Müzikten duygu tanımanın başarısını etkileyen bir diğer önemli faktörde, müzikten algılanan duyguyu tam olarak karşılayan akustik öznelikler elde etmemizi sağlayan öznelik çıkarma işlemidir. Bu zamana kadar araştırmacılar tarafından çok çeşitli akustik öznelikler bulunmuştur. Müzik ve duyguyu ilişkilendirebilecek artikülasyon, ses perdesi, melodi, uyum, tonalite ve ritim gibi öznelikler kullanılmıştır. Bu özneliklere ek olarak enerji, sıfır geçiş hızı, mel frekans kepsral katsayıları, spektral şekil gibi klasik ses öznelikleri de kullanılmıştır. Ayrıca bütün bu öznelikleri toplu olarak elde etmemizi sağlayacak OpenSMILE (Eyben, ve Schuller, 2015), MIRToolbox (Lartillot, ve Toivainen, 2007), YAAFE (Mathieu, ve ark., 2010), JAudio (McKay, 2009) ve Marsyas (Tzanetakis, ve Cook, 1999) gibi araçlar da mevcuttur. Çalışmamızda sınıflandırma başarısını artırmak için gereken öznelikler, yukarıda bahsi geçen araçlardan OpenSMILE, MIRtoolbox ve JAudio araçları kullanılarak elde edilmiştir.

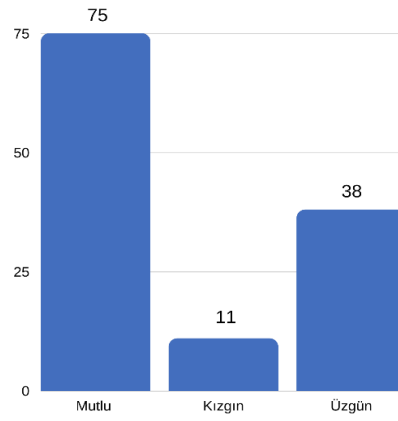
Müzikten duygu tanımadaki bir başka zorlu süreç de doğru sınıflandırıcıyı seçmektir. Bu alanda yapılan çalışmalar incelendiğinde, çoğunun makine öğrenme yöntemlerini kullandığı görülmektedir (Öztürk ve Özkaya, 2020). Destek vektör makineleri (SVM) (Song, Dixon, ve Pearce, 2012), lojistik regresyon (Kim ve ark., 2010), sıralı minimal optimizasyon (Rocha, Panda, ve Paiva, 2013), karar

ağaçları (J.48) (Grekow, 2015), uzun kısa vadeli hafıza ağları (LSTM) (de Benito-Gorron, 2019), ve derin evrişimli sinir ağı (DCNN) (Sarkar, 2019) gibi sınıflandırıcıların müzikten duyguların tanınması için yapılan birçok çalışmada araştırmacılar tarafından kullanıldığını görmekteyiz. Çalışmanın II. bölümünde kullanılan veritabanı, özniteliklerin nasıl elde edildiği, nasıl seçildiği, kullanılan sınıflandırıcılar ile ilgili bilgileri ve gerçekleştirilme aşamalarını, III. bölümünde çalışma ile elde edilen bulguları ve çalışma ile ilgili tartışmaları, son olarak IV. bölümünde ise çalışmada elde edilen sonuçları bulabilirsiniz.

2. Materyal ve Metot

2.1. Veri Tabanı

Yapılan çalışmada geleneksel Türkçe müziklerden oluşturulmuş Türkçe Duygusal Müzik Veriseti kullanılmıştır. Veritabanı şarkıların bir duygu ile temsil edilebileceği düşünülen, farklı türlerde her biri 30 saniye uzunluğunda 124 adet parçanın biraraya getirilmesiyle oluşturulmuş. Şarkı sözlerinin kişide uyandıracak duyguyu etkileyeceği düşünüldüğü için parçalar sadece müzikten ibaret olacak şekilde meydana getirilmiş. Her bir müzik enstrümanda kullanabilen 21 (5 kadın, 16 erkek) dinleyici tarafından AnnoEmo (Yang ve ark., 2007) programı kullanılarak 2 boyutlu modele göre değerlendirilmiştir. Bu değerlendirmelerin ortalama değerleri alındığında 3 farklı bölgede toplanma olduğu görülmüş. Daha sonra bu değerlendirmeler mutlu, üzgün ve kızgın olmak üzere 3 sınıfa adapte edilmiştir. Veritabanındaki müziklerin duygu kategorilerine göre dağılımı Şekil 1’de verilmiştir.



Şekil 1. Müziklerin duygu kategorilerine göre dağılımı

2.2. Özniteliklerin Elde Edilmesi

Yapılan bu çalışmada müzik sinyallerine ait öznitelikler, çalışmalarda en çok karşımıza çıkan ve çok sayıda farklı öznitelik elde etmemizi sağlayan OpenSMILE, MIRtoolbox ve JAudio yazılımları kullanılarak toplanmıştır. OpenSMILE kullanılarak logaritmik enerji, sıfır geçiş oranı, spectral geçirgenlik, mel frekansı kestrum katsayısı ve seslilik olasılığı gibi düşük seviye tanımlayıcıları, bu tanımlayıcıların fonksiyonları (min, maks, ortalama vs.) ve hepsinin birinci ve ikinci türevlerine ait istatistik değerleri kullanılarak 6553 öznitelik elde edilmiştir.

MIRToolbox, MATLAB’da yazılan, ses perdesi, tını, tonalite ve diğerleri gibi müzik bilgilerinin çıkarılmasını sağlayan kendine özgü entegre bir dizi fonksiyona sahiptir. Bu araç hem düşük hem de yüksek seviyeli çok sayıda öznitelik çıkardığı için yaygın olarak kullanılır ve bu araç kullanılarak 348 öznitelik elde edilmiştir.

Son olarak, öznitelik çıkarma sırasında karşılaşılan sorunlara kapsamlı bir çözüm sağlayan JAudio kütüphanesi kullanılmış ve 468 öznitelik elde edilmiştir. Java ortamında yazılan bu kütüphane, OpenSMILE ve MIRtoolbox’tan farklı olarak, karekök ortalama, düşük enerji, en güçlü vuruş, doğrusal önkestirim kodlayıcı, kompaktlık, spektral değişkenlik, spektral pürüzsüzlük gibi hem genişletilebilir, hem de kullanımı kolay çok sayıda özelliğin genel kullanımına izin verir. Ayrıca standart sapma, türev ve ortalama gibi herhangi bir öznitelige uygulanarak yeni öznitelikler oluşturan meta özelliklerde bulunur.

2.3. Öznitelik Seçimi

Öznitelik çıkarma işlemi bittikten sonra müziklerden elde ettiğimiz yüzlerce belki de binlerce özniteliğimiz olabilir. Ancak bu özniteliklerin hepsi eşit öneme sahip değildir ve mevcut öznitelikler arasından en iyi olanı seçmek gerekir. Duygu tanıma işlemi için gereksiz olan bu öznitelikleri çıkarmak ve sınıflandırıcı performansını artırmak amacıyla öznitelik seçim işlemi yapılır. Bu işlem için de çalışmamızda, birbiriyle bağlantısı olmayan ama sınıfla ilişkisi yüksek öznitelik alt kümelerini sezgisel fonksiyon aracılığıyla sıralayan korelasyon tabanlı öznitelik seçimi algoritması (Correlation-based Feature Selection) (Hall ve Smith, 1997) kullanılmıştır.

2.4. Sınıflandırıcılar

Eldeki veriler kullanılarak ve belirli özellikler baz alınarak sınıfı belli olmayan verilerin, birbirlerinden yüksek doğruluk oranları ile çeşitli sınıflara ayırt edilmeye çalışılması işlemine sınıflandırma denir. Sınıflandırma bir denetimli öğrenme problemidir. Bu çalışmanın sınıflandırma aşamasında WEKA (Hall ve ark., 2009) aracı kullanılmıştır. Bu araç içerisinde sınıflandırma ve kümeleme algoritmaları gibi çok sayıda veri madenciliği yöntemi bulunmaktadır. Yapılan müzikten duygu tanıma çalışmasında oldukça popüler

olan sıralı minimal optimizasyon, lojistik regresyon, karar ağaçları (J.48) ve Bayes ağları kullanılmıştır ve bu sınıflandırıcılar Weka'da varsayılan parametreleri ile kullanılmıştır.

2.4.1. Sıralı Minimal Optimizasyon

Sıralı minimal optimizasyon algoritması, John Platt (1998) tarafından geliştirilmiş, destek vektör sınıflandırıcısını (SVM) eğitmeyi amaçlayan basit bir algoritmadır. Bu eğitim büyük miktarda karesel programlama hesaplamaları gerektirir. SMO, SVM'nin aksine bu hesaplamaları ekstra matris depolaması olmadan, küçük problemlere dönüştürerek daha hızlı bir şekilde çözmek için geliştirilmiştir. İç ve dış olmak üzere 2 döngüden oluşur. Dış döngüde en iyi veriler seçilirken, iç döngüde bu verilere göre 2 lagrange çarpanı bulunur. SMO bu iki çarpan için analitik olarak çözüm bulabilir. Böylece karesel programlama probleminin önüne geçilir. Bütün örnekler istenilen seviyeye gelene kadar döngüler çalıştırılır. SMO sınıflandırma problemlerinin dışında regresyon işlemlerinde de kullanılır.

2.4.2. Lojistik Regresyon

Lojistik regresyon verilerin sınıflandırılmasında kullanılan, standart regresyon analizlerinde olduğu gibi bağımlı ve bağımsız değişkenler arasındaki ilişkileri araştırarak ve en az değişken ile en iyi uyumu sağlayan matematiksel bir model kuran, esnek, kullanımı rahat, doğrusal olmayan verilerle çalışıldığında başarılı bir yöntemdir (Le Cessie ve Houwelingen, 1992). Lojistik regresyon, birçok istatistiksel analiz metoduyla bezerlikler gösterir. Lojistik regresyon, logit fonksiyonunu kullanır. Logit fonksiyonu, olayın olma olasılığının olmama olasılığına bölünmesiyle elde edilen değer e tabanında logaritmasının alınmasıyla hesaplanır. Logit fonksiyonunun hesaplanması ile pozitif ve negatif değerler alabilen, 0-1 arasında bir olasılığa geri çevrilebilen metrik bir değişken elde edilmiş olur.

2.4.3. Karar Ağaçları

Karar ağaçları veri yapılarını ifade etmede başarılı ve hızlı, kurulumu ve yorumlanması kolay, ve güvenli olması nedeniyle en çok kullanılan makine öğrenme algoritmalarından biridir. Çalışma Weka da yapıldığı için karar ağacı olarak J.48 algoritması kullanılmıştır. J.48 (Hall ve ark., 2009) çalışmalarda sık kullanılan C4.5 (Quinlan, 1993) algoritmasına dayanır. Bu algoritma karar ağacı oluşturma işlemine dallanmanın hangi niteliğe göre olacağını belirlemek için ilk olarak entropi değerini hesaplayarak başlar ve daha sonra bilgi kazançlarını hesaplayarak devam eder. Bu hesaplamaların amacı en yüksek bilgi kazanımını sağlayan tahmin edici sınıfı bulmaktır. Bilgi kazancının en yüksek olduğu değişken köke yerleştirilir ve ağaç bu değişkenden itibaren dallandırılarak özinelemeli olarak oluşturulur. Kayıp veri varsa, diğer veriler kullanılarak tahmin edilebilir. Amaç, her düğüm az sayıda örnek içerece kadar ağacı büyütmek ve gereksiz düğümleri budama yaparak daha az karmaşık ve anlaşılır bir ağaç elde etmektir.

2.4.4. Bayes Ağları

Bayes ağları, son dönemde sık sık karşımıza çıkan ve birçok alanda veri modelleme ve durum geçişi ifade etmek için kullanılan bir metottür. Judea Pearl (1985) tarafından önerilen bu ağlar, yaşanmış olayları analiz ederek gelecekte yaşanacak olaylarla ilgili çıkarım yapabilmemize olanak sağlar. Bayes Teoremi'nin üzerine inşa edilmiş, olasılık ilişkisi bulunan küme elemanlarının ilişkilerini modellemeye yarayan, belirsiz değerlerin ya da kararların grafiksel olarak gösterimini sağlayan, anlaşılması zor matematiksel ifadeleri ve denklemleri daha kolay incelenebilir hale getiren ağlardır. Bu ağlar sonlu sayıda duruma sahip değişkenleri ifade eden düğümlerden oluşur. Düğümler arasındaki yollar ilgili değişkenler arasındaki olasılıksal ilişkileri açıklar. Değişkenler arasındaki bu ilişkiler, hem ilişkili düğümleri hem de ilişkisiz düğümleri anlatacak şekilde bir yönlendirilmiş döngüsüz diyagram ile gösterilir.

2.5. Gerçekleştirim

Bu çalışmada MIRtoolbox aracından 348, JAudio aracından 468 ve OpenSMILE aracından 6553 olmak üzere toplam 7369 öznitelik elde edilmiştir. Bu öznitelikler hem ayrı ayrı hem de bir arada kullanılarak tanıma yapılmıştır.

Deneysel çalışmalar, Waikato Üniversitesi tarafından açık kaynak kodlu olarak JAVA dili üzerinde geliştirilmiş olan WEKA programı ile yapılmıştır. Sonuçlarını değerlendirmek ve karşılaştırmak için 10 kat çapraz doğrulama yöntemi kullanılmıştır. Bu yöntemde veritabanı eşit 10 verisetine ayrılır, 9'u eğitim verisi olarak kullanılarak model oluşturulur. Oluşturulan model doğrulama için ayrılan 1 test verisi üzerinden test edilir ve doğruluk oranı hesaplanır. Herbir veriseti, hem eğitim verisi hem de test verisi olarak kullanana kadar süreç 10 defa tekrar eder. Modelin doğruluk oranı, ayrı ayrı hesaplanan 10 doğruluk oranının aritmetik ortalaması alınarak bulunur.

Oluşturulan bu modellerin başarısını değerlendirmek için doğruluğun dışında, hata matrisi, hata oranı, duyarlılık, kesinlik ve f-ölçütü gibi ölçütlerde kullanılabilir. Bunlar arasında basit ve en popüler olan yöntem modele ait doğruluk oranıdır. Doğruluk oranı Denklem (1) de görüldüğü gibi başarılı sınıflandırılmış örnek sayısının (DN+DP) toplam örnek sayısına oranına (DN+DP+YP+YN) bölünmesi ile hesaplanır.

$$\text{Doğruluk} = \frac{DP + DN}{DP + YN + YP + DN} \quad (1)$$

3. Araştırma Sonuçları ve Tartışma

3.1. Araştırma Sonuçları

Bu çalışmada 3 farklı araç kullanılarak öznelikler elde edilmiştir. Elde edilen bu özellikler, 4 farklı sınıflandırıcıda kullanılarak sonuçlar elde edilmiştir. İlk 3 tablodaki sonuçlara bakıldığında öznelik seçimi işleminin tanıma oranını artırmada önemli ölçüde etkili olduğu ve bütün sınıflandırma metodlarının bu işlem sonucunda başarısını artırdığı görülmektedir. Yine bu tablolara baktığımızda, her bir aracın ayrı ayrı kullanılmasıyla elde edilen özelliklerle en iyi doğruluğu veren sonuçların öznelik seçimi öncesinde SMO ile, öznelik seçimi sonrasında ise Bayes ağları ile elde edildiği görülmektedir.

Tablo 1. MIRtoolbox Aracı Kullanılarak Elde Edilen Sonuçlar

Sınıflandırıcı	MIRtoolbox	
	Öznelik Seçimi Öncesi Doğruluk	Öznelik Seçimi Sonrası Doğruluk
Bayes Ağları	87.09	94.35
Karar Ağaçları	82.25	87.09
Lojistik Regresyon	77.41	87.09
SMO	89.51	90.32

Tablo 2. JAudio Aracı Kullanılarak Elde Edilen Sonuçlar

Sınıflandırıcı	Jaudio	
	Öznelik Seçimi Öncesi Doğruluk	Öznelik Seçimi Sonrası Doğruluk
Bayes Ağları	81.45	91.93
Karar Ağaçları	83.87	84.67
Lojistik Regresyon	75	85.48
SMO	89.51	89.51

Tablo 3. OpenSMILE Aracı Kullanılarak Elde Edilen Sonuçlar

Sınıflandırıcı	OpenSMILE	
	Öznelik Seçimi Öncesi Doğruluk	Öznelik Seçimi Sonrası Doğruluk
Bayes Ağları	83.87	94.35
Karar Ağaçları	81.45	82.25
Lojistik Regresyon	82.25	91.12
SMO	90.32	93.54

Son tabloda ise, Bayes ağlarının farklı araçlarla elde ettiği tanıma oranlarına, bu araçların hepsinden elde edilen özneliklerin bir araya getirilmesi ve bu özneliklere seçim işlemi uygulanması sonrasında elde edilen doğruluk oranı eklenmiştir. Öznelik sayısını artışına da bağlı olarak % 1.6 artış gözlemlenerek % 95.96'ya çıktığı görülmüştür. Bu oran, bu çalışmada elde edilen en yüksek müzikten duygu tanıma oranıdır.

Tablo 4. Araçlar Birlikte Kullanıldığında Elde Edilen Sonuçlar

Araçlar	Bayes Ağları Doğruluk Sonuçları
MIRtoolbox	94.35
JAudio	91.93
OpenSMILE	94.35
MIRtoolbox + Jaudio + OpenSMILE	95.96

3.2. Tartışma

Araçlar içerisinde en iyi tanımın OpenSMILE aracı kullanılarak elde edilen özniteliklerle yapıldığı görülmektedir. Öznitelik sayısının diğer araçlardan daha fazla olması bu başarıda etkili olabilir.

Çalışmada araçların bir arada kullanılmasıyla öznitelik sayısının artması ve buna bağlı olarak tanıma oranlarında artacağı düşüncesi tablo 4 incelendiğinde de doğrulanmıştır. Öznitelik seçimi sonrasında her ne kadar öznitelik sayısı düşse de, öncesinde sınıfla ilintili öznitelik sayısı arttığından başarı oranı da buna bağlı olarak artmaktadır.

Sonuçlara bakıldığında bütün tablolarda öznitelik seçimi öncesinde SMO sınıflandırıcısının, öznitelik seçimi sonrasında ise Bayes ağlarının en iyi tanıma oranını elde ettiği görülmektedir. Aynı durum lojistik regresyon ile karar ağaçları içinde geçerlidir. Buradan SMO'nun büyük miktarda karesel programlama gerektiren durumlarda, Bayes ağlarının ise, olasılık ilişkisi bulunan küme elemanlarının sayısı daha az olduğunda, ilişkilerini daha iyi modellediği söylenebilir.

4. Sonuç

Bu çalışma kapsamında farklı makine öğrenmesi yöntemleri kullanılarak müzikten duygu tanıma işlemi gerçekleştirilmiştir. Bu çalışmada yeni bir yöntem önerilmemiş olsa da makine öğrenmesi yöntemlerinin bir veriseti üzerindeki performansları ölçülmüş ve bundan sonraki çalışmalara bir temel oluşturulmuştur.

Makine öğrenmesi yöntemleri olarak Lojistik Regresyon, Sıralı Minimal Optimizasyon, Karar Ağaçları ve Bayes Ağları kullanılmıştır. Kullanılan sınıflandırıcıların doğruluk ölçütü açısından performansları karşılaştırıldığında ise öznitelik seçimi sonrası Bayes Ağlarının diğer metodlara oranla daha başarılı tanıma yaptığı söylenebilir, bu yöntemi sırasıyla Sıralı Minimal Optimizasyon, Lojistik Regresyon, ve Karar Ağaçları yöntemleri takip etmektedir.

Bayes Ağları kullanılarak öznitelik seçimi sonrasında duygu tanıma için % 94.35 doğruluk oranı elde edilmişken, bütün araçlar bir arada kullanılarak öznitelik sayısının artmasıyla, Bayes ağlarının doğruluk oranının % 95.96'ya çıktığı görülmektedir. Deneysel sonuçlar kullanılan veritabanının kalitesini, elde edilen özniteliklerin başarıya etkisini ve kullanılan sınıflandırıcıların etkili sonuçlar ortaya koyduğunu göstermiştir.

Kaynakça

- Aljanaki, A., Yang, Y. H., & Soleymani, M., (2017). Developing a benchmark for emotional analysis of music. PLoS ONE 12(3), 1–22, doi: 10.1371-journal.pone.0173392
- Benito-Gorron, D. de. Lozano-Diez, A., Toledano, D. T., & Gonzalez- Rodriguez, J., (2019). Exploring convolutional, recurrent, and hybrid deep neural networks for speech and music detection in a large audio dataset. Eurasip Journal on Audio, Speech, and Music Processing, (1), 1–18.
- Eyben, F., & Schuller, B., (2015). OpenSMILE – The Munich Versatile and fast open-source audio feature extractor. ACM SIGMultimedia Records, 6(4), 4–13, doi: 10.1145/2729095.2729097.
- Feng, Y., Zhuang, Y., & Pan, Y., (2003). Popular music retrieval by detecting mood. in: SIGIR Forum (ACM Spec. Interes. Gr. Inf. Retrieval), 375–376.
- Grekow, J., (2015). Audio features dedicated to the detection of four basic emotion. Computer Information Systems and Industrial Management: CISIM'2015: 14th IFIP TC8 International Conference, September 24-26, Warszawa, Poland.
- Hall, M., & Smith, L., (1997). Feature subset selection: A correlation-based filter approach. Proceedings of the 4th International Conference on Neural Information Processing and Intelligent Information Systems, New Zealand, 855–858.
- Hall, M., Frank, E., Holmes, G, Pfahringer, B., Reutemann, P., & Witten, I., (2009). The WEKA data mining software: An update. SIGKDD Explorations, 11, 10-18, doi: 10.1145/1656274.1656278.
- Hevner, K., (1936). Experimental studies of the elements of expression in music. The American Journal of Psychology, 48, 2: 246-268.
- Kim, E. Y., Schmidt, M. E., Migneco, R., Morton, B. G., Richardson, P., Scott, J., Speck, A. A., & Turnbull, D., (2010). Music emotion recognition: A state of the art review. Proceedings of the 11th International Society for Music Information Retrieval Conference, 9-13 August, Utrecht, Netherlands.
- Lartillot, O., & Toivainen, P., (2007). Mir in matlab (II): A toolbox for musical feature extraction from audio. Proceedings of the 8th International Conference on Music Information Retrieval, September 23-27, Vienna, Austria, 127–130.
- Le Cessie, S., & van Houwelingen, J., C., (1992). Ridge estimators in logistic regression. Applied Statistics, 41(1), pp. 191-201.
- Li, T., & Ogihara, M., (2003). Detecting emotion in music. Proceedings of the International Symposium on Music Information Retrieval, (3), 239-240.
- Mathieu, B., Essid, S., Fillon, T., Prado, J., & Richard, G., (2010). Yaafe, an easy to use and efficient audio feature extraction software. Proceedings of the 11th International Society for Music Information Retrieval Conference August 9-13, Utrecht, Netherlands, 441–446.
- McKay, C. (2009). JAudio: Towards a standardized extensible audio music feature extraction system. Course Paper, McGill University, Canada.
- Öztürk, Ş., & Özkaya, U. (2020). Gastrointestinal tract classification using improved LSTM based CNN. *Multimedia Tools and Applications*, 1-16.

- Panda, R., & Paiva, R. P., (2012). Music emotion classification: Dataset acquisition and comparative analysis. Proc. of the 15th Int. Conference on Digital Audio Effects, September 17-21, York, UK.
- Pearl, J., (1985). Bayesian networks: A model of self-activated memory for evidential reasoning. Proceedings of the Seventh Conference of the Cognitive Science Society, California, USA.
- Platt, J., (1998). Sequential minimal optimization: A fast algorithm for training support vector machines. Microsoft Research Technical Report: MSRTR, 98-14.
- Quinlan, J. R., (1993). C4.5: Programs for machine learning. Morgan Kaufmann Publishers.
- Rocha, B., Panda, R., & Paiva, R. P., (2013). Music emotion recognition: The importance of melodic features. 6th International Workshop on Music and Machine Learning in conjunction with the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases, September, Prague, Czech Republic.
- Russell, J. A., (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39, 1161–1178.
- Sarkar, R., Choudhury, S., Dutta, S., Roy, A., & Saha, S. K., (2019). Recognition of emotion in music based on deep convolutional neural network. *Multimedia Tools and Applications*, 79:765–783.
- Song, Y., Dixon, S., & Pearce, M., (2012). Evaluation of musical features for emotion classification. Proceedings of the 13th International Society for Music Information Retrieval Conference, October, Porto, Portugal.
- Tzanetakis, G., & Cook, P., (1999). MARSYAS: A Framework for audio analysis. *Organised Sound*, 4(3), 169-175.
- Yang, Y. H., Su, Y. F., Lin, Y. C., & Chen, H. H., (2007). Music emotion recognition: The role of individuality. In Proceedings of the ACM International Workshop on Human-Centered Multimedia, 13-21.