Çukurova Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi, 35(3), ss. 687-698, Eylül 2020
Çukurova University Journal of the Faculty of Engineering and Architecture, 35(3), pp. 687-698, September 2020

Farklı Kültürlere Ait Farklı Türdeki Müziklerden Duygu Tanıma

Serhat HIZLISOY[*1], Zekeriya TÜFEKCİ[1]

[1]*Çukurova Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Adana*

Geliş tarihi: 22.09.2020 Kabul tarihi: 23.10.2020

Öz

Bu çalışmada, klasik makine öğrenme yöntemleri farklı kültürlere ait farklı türdeki müziklerden oluşmuş veri tabanları üzerinde duygu tanıması yapmak için kullanılmışlardır. Bu veri tabanlarında bulunan müziklerden öznitelik çıkarmak için çalışmalarda yaygın olarak kullanılan araçlar tercih edilmiştir. Çıkarılan bütün özniteliklere korelasyon tabanlı öznitelik seçme yöntemi uygulanmıştır. Makine öğrenmesi yöntemleri olarak Bayes Ağları, Sıralı Minimal Optimizasyon, Lojistik Regresyon ve Karar Ağaçları kullanılmıştır. Öznitelik seçim işlemi sonrasında kalan özniteliklere Bayes Ağları yöntemi uygulandığında, Türkçe Duygusal Müzik Veri Tabanı için %94,35, Bi-Modal Veri Tabanı için %79,62 ve Soundtrack Veri Tabanı için ise %75,45 tanıma oranı elde edilmiş ve karşılaştırılan sınıflandırıcılardan daha iyi sonuç alınmıştır. Daha sonra, araçlardan çıkarılan öznitelikler bir araya getirilmiş ve yine seçim işlemi yapılmıştır. Bu işlemden sonra ise, sırasıyla bu veritabanları için %95,96, %80,24 ve %82,72 tanıma oranları elde edilmiştir.

Anahtar Kelimeler: Müzik, Duygu tanıma, Makine öğrenmesi

Emotion Recognition From Different Types of Music From Different Cultures

Abstract

In this study, various machine learning methods were used to recognize emotions on databases of different types of music belonging to different cultures. In order to obtain features from the music in these databases, widely used toolboxes were preferred. Correlation-based feature selection method was applied to all the obtained features. BayesNet, Sequential Minimal Optimization, Logistic Regression and Decision Trees are used as machine learning methods. When BayesNet was applied to the remaining features after the feature selection process, %94,35 recognition accuracy rate was obtained for Turkish Emotional Music Database, %79,62 for Bi-Modal Database, and %75,45 for Soundtrack Database, and better results were achieved than other classifiers. Then, the features obtained from the toolboxes were combined and the selection process was made again. After this process, recognition rates of %95,96, %80,24 and %82,72 were obtained for these databases, respectively.

Keywords: Music, Emotion recognition, Machine learning

[*]Sorumlu yazar (Corresponding author): Serhat HIZLISOY, *shizlisoy@cu.edu.tr*

1. GİRİŞ

Müzik, kişide çeşitli duygu ve düşünceler uyandıran ve insanlığın başlangıcından bu yana hayatımızda bulunan evrensel bir dildir. Müzik dinlemek insanların günlük hayatta çok sık yaptığı aktivitelerdendir. Müziğin, insan psikolojisi ve fizyolojisi üzerinde önemli etkileri vardır. İnsanlar çoğu zaman duygu durumlarını değiştirmek için müziğe yönelirlerken, bazen de hissettikleri olumlu ya da olumsuz duygu oranlarını daha da artırmak için de müziğe başvurabilirler.

Teknolojinin gelişmesi, cihaz ve uygulamaların geliştirilmesi ile duygu tanıma üzerine yapılan çalışmalar her geçen gün artmaktadır. Bilgisayar bilimi, müzikoloji, sosyoloji ve hatta psikoloji gibi farklı disiplinlerde yapılan birçok çalışma, bu duyguları farklı yöntemler ve ölçme araçları kullanarak ortaya çıkarmayı amaçlamıştır. Son zamanlarda ise bu çalışmalar, müziğin insanda uyandırdığı duygular düşünülerek, müzik ile duygu tanıma üzerine yoğunlaşmıştır. Yapılan çalışmalarda genellikle ses ve görüntü verilerinden faydalanılır.

Müzikten duygu tanıma yapmak için, genellikle müzik sinyalleri kullanılır. Duygu tanıma için, bazı çalışmalarda ham sinyal ya da bu sinyallerden çıkarılan öznitelikler, bazılarında ise bu özniteliklerden elde edilen spektogram gibi görsel ögeler kullanılır. Son zamanlarda yapılan çalışmalarda müziklerin, insanların duygularına yön vererek hormonlarını etkilediği de düşünüldüğünden müzik sinyalinin kalp ritmi, nabız ve beyin dalgaları gibi fizyolojik parametrelerle birlikte uygulandığı da görülmektedir [1].

Günümüzde Spotify gibi insanların dinledikleri müziklerden duygu analizi yapan sistemler barındıran çok sayıda müzik öneri sisteminin ve otomatik çalma listesi oluşturma uygulamalarının geliştirildiğini görmekteyiz. Bu uygulamalar dışında müzikten duygu tanıma sistemleri ayrıca, müziğin rehabilite etme ve sosyalleştirme özelliği sayesinde müzik terapisi ve duygusal bozuklukların tedavisi gibi birçok amaç için de kullanılmaktadır [2].

Müziğin insanda hissettirdiği duygular; üzüntü, mutluluk, öfke, depresyon, heyecan ve sakinlik gibi çeşitli kategoriler altında toplanabilir. Ancak günümüzde duygusal müzik kategorilerinin belirlenmesi birçok nedenden dolayı oldukça zordur ve bu durum müzikten duygu tanıması yapılması sürecini zorlaştırır. Bu nedenlerden biri, insanların duygu algısı öznel olduğu için aynı müzikten farklı duyguları hissedebilmesidir. Hissedilen duygu; yaşa, cinsiyete, karaktere, şarkı sözlerine, hatta çevresel faktörlere göre bile değişebilir [3].

Müzikten duygu tanımayı zorlaştıran diğer etmenleri kaliteli veri tabanı seçiminin zorluğu, müziklerin doğru etiketlenmemesi, öznitelik çıkarımı ve uygun sınıflandırma yöntemleri seçimi olarak sayabiliriz. Bu çalışma ile yapılış amaçları ayrı olan (Album, Film) farklı kültürlerdeki (Türk Müziği, Batı Müziği) müziklerden oluşan veri tabanları üzerinde duygu tanıma yapılması için genel problemler tespit edilmiştir. Bu problemlerin üstesinden gelmek ve sınıflandırma başarısını artırmak için çeşitli yaklaşımlar geliştirilmiştir.

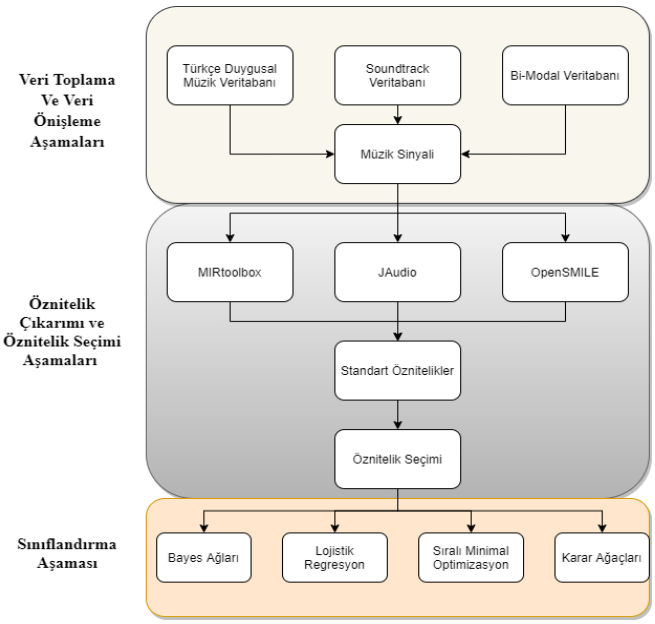
Müzikten duygu tanıma yapabilmek için, öncelikle doğru etiketlenmiş duygusal müzik veri tabanına ihtiyaç vardır. Çünkü verileri etiketleme kalitesi arttıkça, bu veriler üzerinde eğitilen yöntemlerin başarısı da o derece artmaktadır. Dinleyiciler tarafından duyguların etiketlenmesi için kategorik ve boyutsal olmak üzere iki yaklaşım vardır. Kategorik yaklaşımda, müzikteki duygular ayrık etiketlerle değerlendirilir. Bu alandaki en eski çalışma Hevner [4] tarafından, sekiz kategoride 66 duygu etiketi kullanılarak yapılmıştır. Duyguları sınıflandırmak için Y. Feng ve arkadaşları [5] ise üzgün, mutlu, korku, öfkeyi kullanmışlardır. Paul Ekman [6] ise öfke, tiksinti, korku, mutluluk, üzüntü ve şaşkınlıktan oluşan 6 temel duygulu bir model sunmuştur. İkinci yaklaşımda ise duygular boyutsal uzayda çeşitli ve bağımsız eksenlerle temsil edilir. Russels [7] tarafından değerlik ve aktivasyon olmak üzere iki boyuttan oluşan bir model önerilmiştir. Ayrıca, müzikal tanım için iyi bilinen bir diğer boyutsal model, Thayer'in enerji-stres modelidir [8]. Bu model Russell'ın döngüsel modelinden uyarlanmıştır. Bazı çalışmalarda

üçüncü boyut olarak baskınlık kullanılsa da genellikle iki boyutlu model tercih edilir. Kategorik modellerde kategori sayısı üzerinde fikir birliği yoktur. Boyutsal modellerin avantajı, kategorik yaklaşıma göre belirsizliğin az olmasıdır.

Müzikten duygu tanımadaki bir diğer önemli faktör kullanılan veri tabanlarıdır. Kaliteli veri tabanlarının sayısının az olması, birçok araştırmacıyı kendi veri tabanlarını oluşturmaları konusunda mecbur bırakmıştır. Yang ve Chen [9] kendi veri tabanlarını oluşturmuşlardır. Bu veri tabanı boyutsal modeller üzerinden etiketlenmiş 1240 adet pop müzikten (Çince) oluşmaktadır. Diğer bir büyük veri tabanı, M. Soleymani ve arkadaşları [10] tarafından hazırlanmıştır ve 744 adet müzik parçasından oluşmaktadır.

Sınıflandırma başarısını artırmak için gerekli aşamalardan biri de, müzikten alınan duyguyu tam olarak karşılayan akustik öznitelikleri çıkarmayı sağlayan öznitelik çıkarma işlemidir. Araştırmacılar tarafından bu zamana kadar müzik ve duygu arasında ilişki kurabilecek harmoni, tını, ritim ve enerji gibi çok çeşitli akustik öznitelikler bulunmuştur. Artikülasyon, ses perdesi, melodi, uyum, tonalite ve ritim gibi öznitelikler kullanılmıştır [10]. Ek olarak bütün bu öznitelikleri bir arada getiren MIRToolbox [11], Psysound [12], Jaudio [13], OpenSMILE [14] ve Marsyas [15] gibi çok sayıda araç da mevcuttur. Bu çalışmada, yaygın olarak kullanılan ve duygusal özellikleri temsil etme yeteneğine sahip öznitelikler MIRtoolbox, OpenSMILE ve JAudio araçları kullanılarak çıkarılmıştır.

Müzikten duygu tanımada belki de en zorlu süreç doğru sınıflandırıcıyı seçme aşamasıdır. Müzikte duyguyu tanımaya yönelik yapılan çalışmalar incelendiğinde çoğunlukla Destek Vektör Makineleri (Support Vector Machine) [16], Lojistik Regresyon [17], Sıralı Minimal Optimizasyon [18], k tane en yakın komşu (k-NN) [19], Rastgele Ormanlar [20], Karar Ağaçları (J.48) [21], ve Bayes Ağları gibi makine öğrenme yöntemlerinin kullandığı görülmektedir. Bu yöntemlerin dışında Evrişimli Sinir Ağları (Convolutional Neural Networks, CNN), Uzun Kısa Vadeli Hafıza Ağları (Long Short Term Memory, LSTM) [22], ve derin evrişimli sinir ağları (Deep Convolutional Neural Network, DCNN) [23] gibi derin öğrenme metodları da araştırmacılar tarafından kullanılmaktadır.



Şekil 1. Sistemin akış şeması

Çalışmanın şu şekilde organize edilmiştir: İkinci bölümde kullanılan veri tabanları ile ilgili bilgiler, özniteliklerin nasıl çıkarıldığı, nasıl seçildiği, kullanılan sınıflandırıcılar ile ilgili bilgiler ve gerçekleştirilme aşamaları ele alınmıştır. Üçüncü bölümde çalışma ile elde edilen bulgular analiz edilirken, çalışmanın son bölümünde elde edilen sonuçlar ve tartışmalar bulunmaktadır.

2. MATERYAL VE METOT

2.1. Kullanılan Veri Tabanları

2.1.1. Türkçe Duygusal Müzik Veri Tabanı

Yapılan çalışmada üç farklı veri tabanı kullanılmıştır. İlk olarak, müzikten duygu tanıma çalışmalarında kullanmak için geliştirmiş olduğumuz geleneksel Türkçe müziklerden oluşan Türkçe Duygusal Müzik Veri Tabanı [24] kullanılmıştır. Bu veri tabanı her biri 30 saniye uzunluğunda farklı türlerde 124 parçanın bir araya getirilmesiyle oluşturulmuştur. Bu parçalar, şarkı sözlerinin değerlendiricinin şarkının duygusu ile ilgili fikirlerini etkileyeceği düşünüldüğü için sadece müzik olacak şekilde meydana getirilmiştir. Her bir parça sadece dinleyici olmayan aynı zamanda enstrüman da kullanabilen 21 (5 kadın, 16 erkek) değerlendirici tarafından, AnnoEmo [25] programı kullanılarak 2 boyutlu modele göre değerlendirilmiştir. Daha sonra bu değerlendirmelerin her bir parça için ortalama değerleri alınmış ve 2 boyutlu düzlemde 3 farklı bölgede toplandığı görülmüştür. Son olarak bu değerlendirmeler mutlu, üzgün ve öfke olmak üzere 3 sınıfa adapte edilmiştir. Veri tabanındaki müzikler Çizelge 1'de gösterildiği gibi, 3 sınıfa dağılmıştır.

Çizelge 1. Veri tabanlarına göre müziklerin duygu dağılımı

Duygu	Türkçe Veri Tabanı	Bi-Modal Veri Tabanı	Soundtrack Veri Tabanı
Mutlu	75	52	23
Kızgın	11	45	39
Üzgün	38	31	27
Sakin	-	34	21
Toplam	124	162	110

2.1.2. Bi-Modal Veri Tabanı

Malheiro ve arkadaşları [26], tarafından yayınlanan Bi-Modal veri tabanı, farklı dönemlere ait çeşitli müzik türlerini içeren 162 şarkıdan oluşmaktadır. Her şarkı 30 saniye uzunluğundadır ve şarkılar iki boyutlu modele göre 4 duygu olarak (mutlu, öfke, üzgün ve sakin) sırasıyla 52, 45, 31 ve 34 sayıda şarkı ile dağıtılmıştır. Şarkılar, farklı özelliklere sahip 39 kişi tarafından manuel olarak değerlendirilmiştir. [24]'deki veri tabanının aksine, bu veri tabanındaki şarkılar şarkı sözleri içermekte ve farklı dillerdeki şarkılardan oluşmaktadır. Deneysel çalışmalarda sadece müzik sinyali ele alınmış ve sözler ihmal edilmiştir. Kullanılan veri tabanı, boyutu, hazırlanışı ve değerlendirme şekli gibi birçok açıdan [24]'deki veri tabanı ile benzerlik göstermektedir.

2.1.3. Soundtrack Veri Tabanı

Soundtrack (Stimulus Set 2) veri tabanı [27] farklı film türlerine sahip 110 film müzik parçasından oluşur. Her bir müzik parçası beş farklı duygu kategorisi (öfke, üzgün, mutlu, korku, sakin) ve üç boyutlu model (değerlik, enerji, baskınlık) ile değerlendiriciler tarafından etiketlenmiştir. Değerlendiriciler, 18-42 yaş arası 116 üniversite öğrencisidir (%68 kadın ve %32 erkek).

Her bir müzik parçası 15 saniyedir ve şarkı sözü, diyalog ya da ses efekti içermez. İlk 50 müzik parçasının duygu durumu 5 farklı kategoriye göre, kalan 60 müzik parçası ise 3 boyutlu düzlem ile değerlendirilmiştir. Değerlendiricilerin 60 parça için verdiği notların ortalaması alınmış ve en çok hangi duygu ile eşleşiyorsa o duyguyu temsil ettiği kabul edilmiştir. Verilen değerlere bakıldığında, öfke ve korku duygusu arasında yüksek bir korelasyon olduğunu fark edilmiş ve bu yüzden bu iki sınıf birleştirilip tek bir sınıf olarak kabul edilmiştir. Sonuç olarak, müzik parçaları mutlu, kızgın, üzgün ve sakin olarak 4 farklı duyguya sırasıyla 23, 39, 27 ve 21 sayılarında dağılmıştır.

2.2. Öznitelik Çıkarımı

Ses sinyallerinde, eğitim modelleri için sinyalin karakteristiğini yansıtacak basit ve anlamlı veriler

elde etmek gerekir. [11,12,13]'deki gibi çalışmalar duygu sınıflandırması için en bilgilendirici öznitelikleri aramaya odaklanmış olsa da duyguyu belirlemede baskın tek bir öznitelik ortaya çıkmamıştır. Duygu tanıma sürecinin başarısı, müziklerden elde edilen özelliklerin, sınıf özelliklerini ne kadar iyi temsil ettiğine bağlıdır. Müziksel özellikler genellikle dinamik (enerji), tını (spektral), harmoni (melodi) ve ritim (tempo) gibi dört ila sekiz farklı başlık altında toplanılabilir [20]. Yüksek enerjili müziklerin mutluluk ve öfke duyguları ile ilgili olduğu, düşük enerjili müziklerin ise sakin ve üzgün duygularıyla ilgili olduğu söylenebilir [18]. Diğer bir taraftan hızlı tempolu müzikler genellikle mutluluk veya öfke duygusuyla ilişkilendirilirken, yavaş tempolu müzikler üzgün veya sakin duyguları ile ilişkilendirilir. Yine mutlu duyguları ifade eden müzikler, genellikle üzgün duygulara sahip müziklerden daha büyük bir spektral enerjiye sahiplerdir. Son olarak, basit ve uyumlu, harmoniler genellikle mutlu veya sakin gibi duygularla ilgili iken, karmaşık, uyumsuz harmoniler, müzikal bir harekette istikrarsızlık yarattıkları için öfke veya üzgün gibi duygularla ilgilidir [18]. Bu çalışmada müziklerden almamız gereken öznitelikler, çalışmalarda sık karşılaştığımız ve çok sayıda öznitelik çıkarmamızı sağlayan MIRtoolbox, OpenSMILE, ve JAudio araçları kullanılarak toplanmıştır.

Yaygın olarak kullanılan bir yazılım olan MIRtoolbox, kendine özgü entegre bir dizi fonksiyona sahip olmakla birlikte, müzik verilerinden bilgi çıkarma alanındaki hesaplama yaklaşımlarına genel bir bakış sağlamak için hazırlanmıştır. MATLAB'a entegre bu yazılımın avantajları, kullanımı kolay ve kapsamlı dökümantasyona sahip olmasıdır. Bu yazılım sayesinde hem düşük hem de yüksek seviyeli çok sayıda öznitelik çıkarılabilmektedir. Örneğin, bu çalışmada 348 öznitelik çıkarılmıştır.

OpenSMILE yazılımı, sinyal işleme ve makine öğrenim teknikleri için modüler ve esnek bir öznitelik çıkarıcıdır. Bu yazılım ilk olarak, konuşma işleme için tasarlanmış olsa da, duygu tanıma konusundaki birçok çalışmada da iyi performans göstermiştir. Büyük OpenSMILE (emo-large) duygu öznitelik seti, düşük düzeyli tanımlayıcılara, çeşitli işlevler, filtreler ve dönüşümler uygulanarak çok sayıda öznitelik çıkarmamızı sağlar. Bu çalışmada OpenSMILE yazılımı kullanılarak 6553 öznitelik çıkarılmıştır.

Son olarak, JAudio kütüphanesi kullanılmış ve MIRtoolbox'tan daha fazla öznitelik (468 adet) çıkarılmıştır. Öznitelik çıkarma sırasında karşılaşılan sorunlara kapsamlı bir çözüm sağlayan JAudio, Java ortamında yazılan, öznitelikleri hem kullanıcı arayüzü hem de komut satırı kullanılarak çıkarabilen açık kaynaklı bir kütüphanedir. OpenSMILE ve MIRtoolbox'tan farklı olarak, karekök ortalama, düşük enerji, en güçlü vuruş, gibi çok sayıda özelliğin kullanımına olanak sağlar. Ayrıca standart sapma, türev ve ortalama gibi meta öznitelikler herhangi bir özniteliğe uygulanabilir ve öznitelik sayısı genişletilebilir.

2.3. Öznitelik Seçimi

Müziklerden çıkarılan özniteliklerin kalitesi, başarı oranını ve sınıflandırma yöntemlerini etkileyen önemli bir faktördür. Örneğin, öznitelik çıkarma işlemi bittikten sonra yüzlerce veya binlerce özelliğe sahip olma şansımız vardır. Ancak, modele verilecek çok fazla öznitelik olması modelin çalışmasını zorlaştırır, eğitim süresini ve süreç yükünü artırır. Çünkü öznitelikler eşit önemde değildir ve mevcut öznitelikler arasından en iyileri seçmek gerekir. Eğitim sırasında duygu tanıma performansını olumsuz yönde etkileyen gereksiz öznitelikleri ortadan kaldırmak için veri setlerinde çeşitli öznitelik seçme yöntemleri kullanılır. Öznitelik seçim yöntemleri, tahmin doğruluğunu en üst düzeye çıkararak, en uygun öznitelik setini bulmayı sağlar. Bu nedenle öznitelik seçimi müzikten duygu tanıma için önemli bir adımdır. Çalışmamızda, öznitelik uzayımızın boyutu, Korelasyon Tabanlı Öznitelik Seçimi Algoritması (Correlation-based Feature Selection, CFS) [28] kullanılarak küçültülmüş ve daha kolay çalışılabilir hale dönüştürülmüştür. Bu algoritma, birbiriyle bağlantısı olmayan ama sınıfla ilişkisi yüksek öznitelik alt kümelerini sezgisel fonksiyon aracılığıyla sıralayarak işlemini gerçekleştirir.

2.4. Sınıflandırıcılar

Sınıflandırma, eldeki verilerle, belirli öznitelikler dikkate alınarak, sınıfı belli olmayan verilerin çeşitli sınıflara ayrılmasına denir. Sınıflandırma problemlerinde çıktı uzayındaki her elemana "sınıf", sınıflandırma problemini çözen algoritmaya ise "sınıflandırıcı" adı verilir. Sınıflandırma makine öğreniminin popüler ve temel görevlerinden biridir ve denetimli öğrenme problemidir. Makine öğrenme yöntemleri, modeli iyi etiketlenmiş veriler üzerinden öğrenebilen ve bilinmeyenler hakkında tahminlerde bulunabilen sınıflandırıcılardır. Sınıflandırma işleminde, örneklerden çıkarılan öznitelik veri tabanı "eğitim" ve "test" olmak üzere iki gruba ayrılır, bazıları sınıflandırıcıyı eğitmek için, diğerleri ise test etmek için kullanılır. Eğitim için önceden sınıfları belirli veri seti olmalıdır. Test için veri setine yeni örnekler eklenmesindeki amaç, bunları mevcut sınıflar içinden en yüksek doğrulukla uygun sınıfa yerleştirmektir. Yani sınıflandırmanın amacı, hangi sınıfa ait oldukları bilinen örneklerin özniteliklerini kullanarak bir model oluşturmak ve bu modeli, özellikleri bilinmeyen yeni örneklerin sınıflandırılmasında kullanmaktır. Modelin başarı oranı yeterli ise, elde edilen modelin mevcut problemin çözümünde etkili olduğu söylenebilir. Tanıma sistemlerinin başarısında, sınıflandırma yöntemi öznitelik çıkarma kadar önemlidir ve çoğu zaman hangi sınıflandırma yönteminin en iyi veya en uygun olduğu kesin olarak söylenemez. Örneklerin yapısı ve sayısı, işlem süresi ve karmaşıklık gibi faktörler, bir sınıflandırma yöntemini diğerinden daha başarılı hale getirebilir. Sınıflandırma aşamasında WEKA [29] programı kullanılmıştır. Bu program içerisinde çok sayıda sınıflandırma ve kümeleme algoritmaları bulunmaktadır. Duygu analizi ile ilgili çalışmalar incelendiğinde, makine öğrenimi yöntemlerinin birçok kez tercih edildiği görülmektedir. Bu çalışmada, oldukça popüler olan ve yaygın olarak uygulanan Karar Ağaçları (J.48), Bayes Ağları, Sıralı Minimal Optimizasyon, ve Lojistik Regresyon, kullanılmıştır. Bu sınıflandırıcılar Weka'nın varsayılan parametreleri ile kullanılmıştır.

2.4.1. Karar Ağaçları

Karar ağaçları yorumlanması kolay, veri yapılarını ifade etmede başarılı, hızlı, ve güvenli olması nedeniyle yaygın kullanılan makine öğrenme algoritmalarındandır. Deneysel çalışmalar çoğu zaman Weka da gerçekleştirildiği için karar ağacı olarak J.48 algoritması [29] kullanılmaktadır [21]. Bu algoritma çalışmalarda sık karşımıza çıkan C4.5 [30] algoritmasına dayanır ki bu algoritma da ağaç oluşturma algoritması ID3'ün geliştirilmiş versiyonudur. J.48, ilk olarak entropi değerini hesaplar ve karar ağacı oluşturma işlemine dallanmanın hangi niteliğe göre olacağını belirler. Daha sonra bilgi kazançlarını da hesaplayarak devam eder. Bu hesaplamaların amacı, en yüksek bilgi kazancını sağlayan sınıfı bulmaktır. Bilgi kazancı en yüksek olan değişken, köke yerleştirilir ve ağaç bu değişkenden itibaren yukarıdan aşağıya doğru özyinelemeli olarak dallandırılır. Eğer kayıp bir veri varsa, bu veri diğer veriler kullanılarak tahmin edilebilir. Amaç, bütün düğümlerde en az sayıda örnek olana kadar ağacı büyütmektir. Ayrıca anlamlı olmayan, zayıf düğümlere budama işlemi yapılarak ağaç optimize edilir.

2.4.2. Bayes Ağları

Judea Pearl [31] tarafından önerilen Bayes ağları, son dönemde sık karşılaştığımız, veri modelleme ve durum geçişi ifade etmek için kullanılan olasılıksal ağlardır. Çıkarım yapma ve yorumlama açısından başarılı olduğu için bir çok alanda kullanılmaktadır. Bayes Teoremi'nin üzerine inşa edilmiştir ve olasılık ilişkisi bulunan küme elemanlarının ilişkilerini modellemeye yarar. Grafikler ve tablolar sayesinde, anlaşılması zor matematiksel denklem ve ifadeleri daha anlaşılabilir hale getirir. Bayes ağlarında her düğüm bir raslantı değişkenini ifade eder ve daire biçiminde gösterilir. Düğümler arasındaki yollar değişkenler arasındaki olasılıksal bağımlılıkları açıklar ve düğümler arasına çizilen oklarla gösterilirler. Bu ilişkileri, hem ilişkili düğümleri hem de ilişkisiz düğümleri anlatan yönlü, çevrimsiz grafikleri kullanarak yapar.

2.4.3. Sıralı Minimal Optimizasyon

Sıralı Minimal Optimizasyon Algoritması (SMO), John Platt [32] tarafından geliştirilmiştir. Destek vektör sınıflandırıcısını (SVM) eğitmek için kullanılan kavramsal olarak basit, uygulaması kolay ve popüler bir algoritmadır. Her aşamada mümkün olan en küçük optimizasyon sorununu çözerek sonucu elde etmeyi hedefler. Çünkü SVM, eğitim aşamasında büyük miktarda karesel programlama hesaplamaları gerektirir ve SVM'nin eğitim algoritmaları da bu büyük problemlerde yavaştır ve karmaşıktır. SMO, SVM'nin standart eğitim algoritmalarının aksine, bu hesaplamaları, ekstra matris depolaması olmadan, küçük problemlere bölerek, hızlı ve daha iyi bir şekilde çözmek için geliştirilmiştir. İlk adım, başlangıç parametrelerini ayarlamaktır. İkinci adım ise döngü prosedürüdür. Döngü prosedürü, iki Lagrange çarpanını seçmek, bunların iyileştirmelerini takip etmek ve SVM parametrelerini güncellemek olan üç temel işlemi gerçekleştirir. Döngüler iç ve dış olmak üzere 2 döngüden oluşur. Bütün örnekler istenilen seviyeye çıkana kadar döngüler çalıştırılmaya devam eder. SMO sınıflandırma problemlerinin dışında ayrıca regresyon işlemlerinde de kullanılır.

2.4.4. Lojistik Regresyon

Lojistik Regresyon, verilerin sınıflandırılmasında ve regresyon işlemlerinde kullanılır. Standart regresyon problemlerinde olduğu gibi bağımlı ve bağımsız değişkenler arasındaki ilişkileri araştırır ve en az değişken ile en çok uyumu sağlayan matematiksel bir model kurar. Kullanımı rahat, esnek ve aynı zamanda doğrusal olmayan verilerle çalışıldığında da başarılı bir yöntemdir [33]. Bu metod, birçok istatistiksel analiz metoduyla benzerlikler gösterir. Lojistik regresyon, logit fonksiyonunu kullanarak, olayın olma olasılığının olmama olasılığına bölünmesiyle elde edilen değerin e tabanında logaritmasının alınmasıyla bulunur. Bu fonksiyon ile negatif ve pozitif değerler alabilen ve 0-1 aralığında bir olasılığa çevrilebilen metrik bir değişken elde edilir.

2.5. Gerçekleştirim

Bu çalışmada OpenSMILE aracı kullanılarak 6553 öznitelik, MIRtoolbox aracı kullanılarak 348 öznitelik ve son olarak JAudio aracı kullanılarak 468 öznitelik olmak üzere toplamda 7369 öznitelik çıkarılmıştır. Bu araçlardan çıkarılan öznitelikler hem bir arada, hem de ayrı ayrı olarak çalışmalarda kullanılmıştır. Deneysel çalışmalar, Java dilinde geliştirilmiş, açık kaynak kodlu WEKA programı kullanılarak yapılmıştır. Sınıflandırma modellerinin eğitilmesi ve sonuçların değerlendirilmesi için K Katlamalı Çapraz Doğrulama (K-Fold Cross Validation) yöntemi kullanılmıştır. Bu yönteme göre veri seti 10 eşit parçaya ayrılır (K=10) ve bu parçalardan 9'u eğitim verisi olarak kullanılarak model oluşturulur. Oluşturulan model geriye kalan 1 test verisi üzerinden test edilir ve doğruluk oranı hesaplanır. Her bir parça hem eğitim verisi hem de test verisi olarak kullanılana kadar bu işlem 10 kez tekrarlanır. Modelin doğruluğu, 10 kez yapılan bu işlem sonucunda bulunan doğruluk oranlarının aritmetik ortalaması alınarak hesaplanır. Oluşturulan modeller doğruluk ölçütü kullanılarak değerlendirilmiş ve bu değerleri bulmak için hata matrisi oluşturulmuştur. Doğruluk ölçütü, başarılı sınıflandırılmış örnek sayısının (DN+DP), toplam örnek sayısına oranına (DN+DP+YP+YN) bölünmesi ile Eşitlik 1'de gösterildiği gibi bulunur.

		Tahmin Edilen	
Gerçek		Doğru Pozitif (DP)	Yanlış Pozitif (YP)
		Yanlış Negatif (YN)	Doğru Negatif (DN)

Şekil 2. Hata matrisi

Oluşturulan modellerin başarısını ölçmek için, doğruluk ölçütü dışında yaygın olarak kullanılan başka ölçütler de vardır. Hata oranı, duyarlılık, kesinlik ve f-ölçütü, bu ölçütlerin en bilinenlerindendir.

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

3. BULGULAR

Müzikteki duyguyu belirlemek için yapılan çalışmalar incelendiğinde, araştırmacıların neredeyse tamamına yakınının batı müziğindeki örnekleri kullanarak çalışma yaptıkları görülmüştür. Bu çalışmada ise farklı kültüre ait ve farklı türdeki müziklerden oluşan 3 veri tabanındaki müziklerden duygu tanıma yapılmak istenmektedir.

Türkçe müziklerden oluşan veri tabanında elde edilen duygu tanıma oranının diğer veri tabanlarından elde edilenlere oranla daha yüksek olduğu görülmektedir. Buradan, Türkçe müziklerden insana geçen duyguların, diğer dillere ve kültürlere ait müziklerle kıyaslandığında daha net bir şekilde aktarıldığı sonucu çıkarılabilir. Ek olarak, bu veri tabanının daha az sınıftan oluşması ve sınıflara ait müziklerin eşit şekilde dağılmamasının da etkisi olduğu düşünülebilir. Bi-modal veri tabanında ise, Soundtrack veri tabanına oranla daha az tanıma yapıldığı görülmektedir. Şarkı sözleri bunun gerçekleşmesinde önemli bir etken olabilir. Çünkü müzikleri etiketleyen değerlendiriciler şarkı sözlerinden etkilenebilirler.

Çizelge 2. Türkçe duygusal veri tabanı üzerinde yapılan sınıflandırma sonuçları

Sınıflandırıcılar	MIRtoolbox		Jaudio		OpenSMILE	
	Öznitelik Seçimi Öncesi	Öznitelik Seçimi Sonrası	Öznitelik Seçimi Öncesi	Öznitelik Seçimi Sonrası	Öznitelik Seçimi Öncesi	Öznitelik Seçimi Sonrası
Bayes Ağları	%87,09	%94,35	%81,45	%91,93	%83,87	%94,35
Karar Ağaçları	%82,25	%87,09	%83,87	%84,67	%81,45	%82,25
Lojistik Regresyon	%77,41	%87,09	%75	%85,48	%82,25	%91,12
SMO	%89,51	%90,32	%89,51	%89,51	%90,32	%93,54

Çizelge 3. Bi-Modal veri tabanı üzerinde yapılan sınıflandırma sonuçları

Sınıflandırıcılar	MIRtoolbox		Jaudio		OpenSMILE	
	Öznitelik Seçimi Öncesi	Öznitelik Seçimi Sonrası	Öznitelik Seçimi Öncesi	Öznitelik Seçimi Sonrası	Öznitelik Seçimi Öncesi	Öznitelik Seçimi Sonrası
Bayes Ağları	%61,72	%66,04	%59,25	%63,58	%63,58	%79,62
Karar Ağaçları	%58,02	%61,72	%51,85	%56,79	%55,55	%60,49
Lojistik Regresyon	%43,20	%55,55	%41,35	%49,38	%49,38	%52,46
SMO	%63,96	%64,81	%61,11	%63,58	%70,98	%72,22

Ayrıca Soundtrack veri tabanı belirli sahnelerde belirli duyguları vurgulamak için yapılan film müziklerinden oluştuğu için, değerlendiriciler tarafından daha iyi etiketlenme ihtimali vardır.

Müziklerden çıkarılacak öznitelikler ise 3 (MIRtoolbox, OpenSMILE, JAudio) farklı araçtan sağlanmıştır. Çıkarılan öznitelikler klasik makine öğrenme metotları olan 4 farklı sınıflandırıcı ile kullanılmıştır. Çizelgelerdeki sonuçlara bakıldığında öznitelik seçim işleminin tanıma oranlarını artırmada ne kadar etkili olduğu görülmüştür. Çünkü sınıflandırma metotlarının hepsi bu işlem uygulandıktan sonra başarılarını önemli ölçüde artırmıştır.

Ayrıca verilen çizelgeler incelenirse, her bir veri tabanı için araçların ayrı ayrı kullanılmasıyla çıkarılan öznitelikler ile en iyi duygu tanıma yapan sınıflandırıcılar görülmektedir. Hemen hemen bütün sonuçlarda, öznitelik seçimi öncesinde SMO sınıflandırıcısı, öznitelik seçimi sonrasında ise Bayes ağları sınıflandırıcısı en iyi sonucu elde etmiştir. Son çizelgede ise, bütün veri tabanlarında en iyi sonuçları almamızı sağlayan Bayes ağları metodunun, farklı araçlarla ve bu araçların

hepsinden çıkarılan özniteliklerin bir araya getirilmesi ile elde edilen doğruluk oranları gösterilmiştir. Bütün öznitelikler bir araya gelince, öznitelik sayısının artışına da bağlı olarak tanıma oranları Türkçe Duygusal Müzik veri tabanında %1,6 artış ile %95,96'ya, Bi-Modal veri tabanında %0,62 artış ile %80,24'e ve son olarak Soundtrack veri tabanında %7,27 artış ile %82,72'ye çıktığı gözlemlenmiştir. Bu oranlar, deneysel çalışmalarda bu veri tabanları için elde edilen en yüksek tanıma oranlarıdır.

4. SONUÇLAR VE TARTIŞMALAR

Kullanılan araçlar sayesinde çıkarılan öznitelikler ile yapılan duygu tanıma işlemi sonuçlarına bakıldığında en başarılı sonuçlara OpenSMILE aracı kullanıldığında ulaşıldığı görülüyor. Bu aracın öznitelik sayısı bakımından diğer araçlara oranla çok fazla olması ve çok daha farklı öznitelikleri kazandırması bu başarıdaki temel etken olarak düşünülebilir.

Çizelge 4. Soundtrack veri tabanı üzerinde yapılan sınıflandırma sonuçları

Sınıflandırıcılar	MIRtoolbox		Jaudio		OpenSMILE	
	Öznitelik Seçimi Öncesi	Öznitelik Seçimi Sonrası	Öznitelik Seçimi Öncesi	Öznitelik Seçimi Sonrası	Öznitelik Seçimi Öncesi	Öznitelik Seçimi Sonrası
Bayes Ağları	%53,63	%67,27	%70,00	%70,90	%44,54	%75,45
Karar Ağaçları	%59,09	%63,63	%53,63	%59,09	%46,36	%70,00
Lojistik Regresyon	%54,54	%66,36	%45,45	%53,63	%48,18	%57,27
SMO	%61,81	%69,09	%53,63	%56,36	%61,81	%70,90

Çizelge 5. Farklı veri tabanları üzerinde ayrı ayrı ve birlikte kullanıldığı durumlarda elde edilen sonuçlar

Araçlar	Türkçe Duygusal Müzik Veri Tabanı	Bi-Modal Veri Tabanı	Soundtrack Veri Tabanı
	Bayes Ağları Sonuçlar	Bayes Ağları Sonuçlar	Bayes Ağları Sonuçlar
MIRtoolbox	%94,35	%66,04	%67,27
JAudio	%91,93	%63,58	%70,90
OpenSMILE	%94,35	%79,62	%75,45
MIRtoolbox + JAudio + OpenSMILE	%95,96	%80,24	%82,72

Daha önce bahsedildiği gibi farklı araçların kullanılmasının nedeni farklı özniteliklerin çıkarılmasıdır. Dolayıyla araçlar bir arada kullanıldığında kendilerinde bulunmayan öznitelikler de öznitelik havuzuna eklenir. Havuzdaki farklı öznitelik sayısının artması ile doğru orantılı olarak tanıma oranlarının da artacağı düşüncesinin doğru olduğu Çizelge 5 incelendiğinde görülmüştür. Öznitelik seçim işlemi sonrasında sınıflandırma için kullanılacak öznitelik sayısının Türkçe müziklerden oluşan veri tabanında 81'e, Bi-Modal veri tabanında 106'ya ve son olarak Soundtrack veri tabanında 80'e düştüğü görülmektedir. Ancak, bu özniteliklerin temsil edilen sınıflarla yüksek ilintili olması ve sistemin sınıflarla ilişkisi olmayan gereksiz özniteliklerden arındırılması tanıma oranını önemli ölçüde artırır.

Veri tabanları ile ilgili sonuçlar incelendiğinde, neredeyse bütün veri tabanlarında, öznitelik seçim işlemi öncesinde SMO'nun, öznitelik seçim işlemi sonrasında ise Bayes ağlarının en başarılı sonuçları aldığı gözlenmektedir. Bu sonuçlardan büyük miktarda verilerin kullanıldığı ve fazla hesaplama

gerektiren durumlarda SMO'nun, veri sayılarının az olduğu durumlarda ise Bayes Ağları'nın, diğer yöntemlerden daha iyi neticeler elde ettiği söylenebilir.

Bu çalışmada yeni bir yöntem önerilmemiş, klasik makine öğrenmesi yöntemleri kullanılarak farklı kültürlere ait farklı türdeki müziklerden oluşan veri tabanları üzerindeki duygu tanıma performansları ölçülmüş ve daha sonraki çalışmalara bir temel oluşturulmak istenmiştir. Yöntem olarak Sıralı Minimal Optimizasyon, Lojistik Regresyon, Bayes Ağları ve Karar Ağaçları kullanılmıştır. Bu yöntemlerin, öznitelik seçim işlemi uygulandıktan sonrasındaki performansları karşılaştırıldığında, Bayes Ağlarının diğer metotlardan çok daha başarılı tanıma oranları elde ettiği gözlenmiştir. Bu yöntemi sırasıyla Sıralı Minimal Optimizasyon, Karar Ağaçları ve Lojistik Regresyon yöntemleri takip etmektedir.

Türkçe müziklerden oluşan veri tabanında Bayes Ağları, öznitelik seçimi sonrasında duygu tanıma için %94,35 doğruluk oranı elde ederken, bütün araçlar bir arada kullanılarak öznitelik sayısının artmasıyla, doğruluk oranını %95,96'ya çıkarmıştır. Bi-modal veri tabanında ise öznitelik seçimi sonrasında %79,62 doğruluk oranı elde ederken, bütün araçlar bir arada kullanıldığında %80,24 tanıma yapmıştır. Son olarak Soundtrack veri tabanında ise öznitelik seçimi sonrasında %75,45 tanıma yaparken, bütün araçlar bir arada kullanıldığında %82,72 doğruluk oranı elde etmiştir. Deneysel sonuçlar kullanılan veri tabanlarının kalitesini, çıkarılan özniteliklerin tanıma oranlarına katkısını ve kullanılan sınıflandırıcıların ne kadar etkili sonuçlar ortaya koyduğunu göstermiştir.

5. KAYNAKLAR

1. Alakuş, T.B., Türkoğlu, İ., 2018. EEG Tabanlı Duygu Analiz Sistemleri. Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi, 11(1), 26-39.

2. Kaya, A.İ., 2018. Üniversite Öğrencilerinde Dini İçerikli Müzik Terapinin Kaygı, Duygu Durumu ve Algılanan Stres Üzerindeki Etkisi. Uluslararası Din & Felsefe Araştırmaları Dergisi 1,1.

3. Erdal, B., 2015. Hissedilen ve Algılanan Duygular Bağlamında Arabesk Müzik Beğenisi Etkileyen Faktörler Üzerine Bir Araştırma. International Journal of Human Sciences, 12, 1: 1016-1055 doi: 10.14687/ijhs.v12i1.3199.

4. Hevner, K., 1936. Experimental Studies of the Elements of Expression in Music. The American Journal of Psychology, 48(2), 246-268.

5. Feng, Y., Zhuang, Y., Pan, Y., 2003. Popular Music Retrieval by Detecting Mood. In: SIGIR Forum (ACM Spec. Interes. Gr. Inf. Retrieval), 375–376.

6. Ekman, P., 2005. Basic Emotions. In Handbook of Cognition and Emotion, (3), 45-60, doi: 10.1002/0470013494.ch3.

7. Russell, J.A., 1980. A Circumplex Model of Affect. Journal of Personality and Social Psychology, 39, 1161–1178.

8. Thayer, R.E., 1989. The Biopsychology of Mood and Arousal. New York: Oxford University Press.

9. Yang, Y.H., Chen, H.H., 2011. Ranking-based Emotion Recognition for Music Organization and Retrieval. IEEE Transactions on Audio Speech and Language Processing, 19, 762–774.

10. Soleymani, M., Caro, M.N., Schmidt, E.M., Sha, C.Y., Yang, Y.H., 2013. 1000 Songs for Emotional Analysis of Music. CrowdMM-Proceedings of the 2nd ACM International Workshop on Crowdsourcing for Multimedia; 1-6, doi:10.1145/2506364.2506365

11. Lartillot, O., Toiviainen, 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.

12. Cabrera, D., 1999. PsySound: A Computer Program for Psychoacoustical Analysis. Australian Acoustical Society Conference, 47-54, doi: 10.1002/asi.

13. McKay, C., 2009. JAudio: Towards a Standardized Extensible Audio Music Feature Extraction System. Course Paper, McGill University, Canada.

14. 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.

15. Tzanetakis, G., Cook, P., 1999. MARSYAS: A Framework for Audio Analysis. Organised Sound, 4(3), 169-175.

16. 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.

17. Kim, Y., Schmidt, E., Migneco, R., Morton, B., Richardson, P., Scott, J., Speck, J., Turnbull, D., 2010. State of the Art Report: 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. 255-266.

18. 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.

19. Delbouys, R., Hennequin, R., Piccoli, F., Royo-Letelier, J. ve Moussallam, M., 2018. Music Mood Detection Based on Audio and Lyrics with Deep Neural Net. CoRR, abs/1809.07276.

20. Zhang, F., Meng, H., Li, M., 2016. Emotion Extraction and Recognition from Music. 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery, 1728-1733, doi: 10.1109-FSKD.2016.7603438.

21. 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.

22. 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.

23. 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.

24. Hızlısoy, S., Tüfekci, Z., 2020. Türkçe Müzikten Duygu Tanıma. 1st International Conference on Computer, Electrical, and Electronic Sciences, 8-10 October 2020.

25. 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.

26. Malheiro, R., Panda, R., Gomes, P., Paiva, R.P., 2016. Bi-Modal Music Emotion Recognition: Novel Lyrical Features and Dataset. 9th International Workshop on Music and Machine Learning-MML'2016-in Conjunction with the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases-ECML/PKDD, Riva del Garda, Italy.

27. Eerola, T., Vuoskoski, J., 2011. A Comparison of the Discrete and Dimensional Models of Emotion in Music. Psychology of Music, 10.1177/0305735610362821.

28. 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.

29. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I., 2008. The WEKA Data Mining Software: An Update. SIGKDD Explor. Newsl.. 11. 10-18.

30. Quinlan, J.R., 1993. C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers.

31. 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.

32. Platt, J., 1998. Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines. Microsoft Research Technical Report: MSRTR, 98-14.
33. Le Cessie, S., van Houwelingen, J.C., 1992. Ridge Estimators in Logistic Regression. Applied Statistics, 41(1), 191-201.