

Makine Öğrenmesi Yöntemleri ile GTZAN Veri Kümesine Ait Müzik Türlerinin Sınıflandırılması

Classification of Music Genres of GTZAN Dataset with Machine Learning Methods

¹Mehmet Cüneyt ÖZBALCI , ²Hasan ŞAHİN , ³Turgay Tugay BİLGİN 

^{1,2,3}Bursa Teknik Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, Yıldırım/Bursa, Türkiye

mehmet.ozbalci@btu.edu.tr, hasan.sahin@btu.edu.tr, turgay.bilgin@btu.edu.tr [Araştırma Makalesi/Research Article](#)

ARTICLE INFO

Article history

Received : 13 February 2024

Accepted : 21 March 2024

Keywords:

Random Forest, Artificial Neural Network, Support Vector Machines, Music Classification

ABSTRACT

Over time, the field of computer science research on music has substantial growth. Music analytic techniques and digital transformation have made music classification more useful. The categorization procedure in this study utilizes the numerical values obtained from the audio signal. The GTZAN dataset is utilized to classify music genres based on a wide range of factors. The classification approaches used were Random Forest (RF), Support Vector Machines (SVM), and Artificial Neural Networks (ANN). The study investigated the performance of these widely-used algorithms in the field of music categorization using the GTZAN dataset, which is commonly used in research. The RF algorithm achieved an accuracy rate of 81%, while the SVM algorithm achieved an accuracy rate of 72.33%, and the ANN algorithm achieved an accuracy rate of 67.67%. This study proved the efficacy of the RF approach in accurately categorizing music by utilizing audio signal-derived attributes.

© 2024 Bandırma Onyedi Eylül University, Faculty of Engineering and Natural Science. Published by Dergi Park. All rights reserved.

MAKALE BİLGİSİ

Makale Tarihleri

Gönderim : 13 Şubat 2024

Kabul : 21 Mart 2024

Anahtar Kelimeler:

Rastgele Orman, Yapay Sinir Ağları, Destek Vektör Makineleri, Müzik Sınıflandırma

ÖZET

Zaman içerisinde müzik, bilgisayar bilimleri için önemli ölçüde gelişme göstermiştir. Müzikte analiz teknikleri ve dijital dönüşüm müzik sınıflandırmasını kullanışlı hale getirmiştir. Bu çalışmada ses sinyalinden elde edilen sayısal değerler ile sınıflandırma işlemi gerçekleştirilmiştir. Müzik türlerini sınıflandırmak için oldukça çeşitli öznitelikler içeren GTZAN veri kümesi kullanılmıştır. Sınıflandırma algoritmaları olarak Rastgele Orman (RF), Destek Vektör Makineleri (SVM) ve Yapay Sinir Ağları (YSA) kullanılmıştır. Müzik sınıflandırmada oldukça popüler olan bu algoritmaların literatürde sıklıkla kullanılan bir veri kümesi olan GTZAN ile birlikte nasıl sonuçlar vereceği gözlemlenmiştir. Bu algoritmalarla elde edilen doğruluk oranları RF için %81, SVM için %72,33 ve YSA için %67,67 olarak elde edilmiştir. Yapılan çalışmada müzik sınıflandırmada, ses sinyallerinden elde edilen öznitelikler ile RF algoritmasının oldukça başarılı sonuçlar üretebileceği gösterilmiştir.

© 2024 Bandırma Onyedi Eylül Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi. Dergi Park tarafından yayınlanmaktadır. Tüm Hakları Saklıdır.

ORCID ID: ¹0000-0003-4499-0061

²0000-0002-8915-000X

³0000-0002-9245-5728

1. GİRİŞ

Müzik, insanların duygularını yansıtan etkili bir sanattır. Armonik notalar ile anlam kazanan müzik temelde zaman ve frekans bilgisinin ön plana çıktığı, çok çeşitli karmaşık bileşenlerden oluşan bir ses sinyalidir [1]. Farklı ulusların ve toplulukların kendi müzik kültürlerine sahip olması, müziği kategorize etme zorluğunu ortaya çıkarmıştır. Müzik kategorizasyonu algısal bir eylemdir. Yapay zekâ alanındaki teknolojiler, müzikte insan algısına en yakın performansı göstererek gelişmektedir. Müziğin hemen her bireyin hayatında önemli bir yere sahip olduğu düşünüldüğünde, yapay zekânın birçok farklı disiplinde olduğu gibi müzikte de önemli bir çalışma alanı olduğu açıktır [2].

Yapay zekâ ve makine öğrenmesi, müzikte iki temel alanda kullanılmaktadır: müzik sentezi ve müzik analizi. Müzik sentezi müzik üretimini ifade ederken, müzik analizi mevcut müzik verilerinin işlenerek müziğe dair bilgilerin değerlendirilmesi ve anlamlı sonuçlar elde edilmesine dayanır. Müzik analizi, duygu analizi, enstrüman analizi, söz analizi, tını analizi ve melodi analizi gibi çalışma alanlarını içerir. Müzik sınıflandırmasında, sınıflandırılacak unsura bağlı olarak birçok parametre devreye girebilir. Bu parametreler tını, melodi, söz, tempo, perde, akor gibi özellikleri içerebilir. Duygu, enstrüman, tını, sanatçı ve perde bilgisine göre sınıflandırma çalışmaları günümüzde de devam etmektedir. Müzik sınıflandırmada çeşitli derin öğrenme ve makine öğrenmesi algoritmaları kullanılmaktadır. Bu algoritmalar ön plana çıkan SVM, ANN, K-NN, RNN ve LSTM gibi algoritmalar başarıyla kullanılmıştır. Optimum sonucun elde edilmesi belirlenecek olan sınıflandırma algoritmasına ve seçilecek olan parametrelere bağlıdır. Ses sinyalinden elde edilen özelliklerin müzik sınıflandırmasında sıkça kullanıldığı görülmektedir. Bu özellikler müzik ve ses analizinde insan kulağının algısal özelliklerini başarılı bir şekilde yansıttıkları için tercih edilir.

Ses sinyalinin kalitesini artırmak ve en yüksek verimi elde etmek için çeşitli filtreleme ve vurgulama yöntemleri geliştirilmiştir. Bu yöntemlerden bazıları; alçak geçiren filtre, yüksek geçiren filtre, IIR (sonsuz dürtü yanıtı) filtresi, FIR (sonlu dürtü yanıtı) filtresi ve Mel filtre bankası olarak görülmektedir. Bu yöntemler çeşitli matematiksel işlemleri içermektedir. Alçak geçiren filtre, belirli bir Hz (Hertz) değerinin üstündeki sinyalleri susturmaktadır. Genlik değerinin azaltılması ile bu işlem gerçekleştirilmektedir. Bu yöntem dış ortam seslerinin yalıtılmasında etkilidir [3]. Yüksek geçiren filtre, düşük frekanstaki sapsmaların optimize edilmesinde kullanılmaktadır. IIR ve FIR filtrelerde, filtre mertebesi belirlenip giriş sinyaline bağlı olarak bir çıkış sinyali elde edilmektedir. IIR filtresinin FIR filtresinden temel farkı özyinelemeli bir yapıda olmasıdır. Mel filtre bankasında ses, genlik ve frekans bilgileri kullanılarak işlenir. Konuşma tanımada etkili bir yöntemdir [4].

2. LİTERATÜR ARAŞTIRMASI

Müzik sınıflandırması ve yapay zekâ arasındaki ilişki, son yıllarda giderek artan bir ilgi görmektedir. Müzik, insan duygularını ifade eden güçlü bir sanat formu olduğundan, yapay zekâ tekniklerinin müzik analizi ve sınıflandırmasında kullanılması, müzik endüstrisinde ve müzik dinleme deneyiminde önemli etkiler yaratabilir. Literatürdeki çalışmalar incelendiğinde, müzik sınıflandırmasında yapay zekâ ve makine öğrenmesi tekniklerinin çeşitli uygulamalarının olduğu görülmektedir. Bu uygulamalar arasında farklı özneliklerin kullanılması, farklı sınıflandırma algoritmalarının karşılaştırılması ve farklı müzik türlerinin tanımlanması gibi konular bulunmaktadır.

Örneğin yapılan bir çalışmada, derin öğrenme yöntemlerinin müzikal duyguların sınıflandırılmasında etkili olduğu gösterilmiştir. Bu çalışma, müzikal duyguların tanımlanması ve analiz edilmesi üzerine odaklanarak müzikal özneliklerin derin öğrenme algoritmalarıyla ilişkilendirilmesini incelemiştir. Bununla birlikte müzik sınıflandırmasında kullanılan özneliklerin belirlenmesi ve sınıflandırma algoritmalarının seçimi gibi konularda hala bazı zorluklar bulunmaktadır. Bu zorlukların üstesinden gelmek için farklı yaklaşımların ve yöntemlerin geliştirilmesi gerekmektedir.

Genel olarak literatürdeki çalışmaların incelenmesi, müzik sınıflandırması ve yapay zekâ alanında yapılan araştırmaların mevcut durumunu anlamamıza ve gelecekteki çalışmalar için ilham alabileceğimiz fikirler sunmamıza yardımcı olur. Bu nedenle bu alandaki literatürü dikkatle incelemek, araştırmanın temelini oluştururken önemlidir.

Lojistik regresyon, K-NN, SVM, RF, YSA ve Naive Bayes yöntemleri ile gerçekleştirilen müzik sınıflandırma çalışmasında GTZAN veri seti kullanılmıştır. Çeşitli eğitim zamanı ve hiper-parametre kombinasyonları ile elde edilen en yüksek doğruluk oranları sırasıyla lojistik regresyon için %81, K-NN için %92,69, SVM için %80,80, RF için %80,28 ve Naive Bayes için %54,50 olarak görülmüştür [5]. Veri seti ön işlemeden geçtikten sonra GTZAN veri setine ilk olarak boyut azaltma teknikleri uygulanmış ardından lojistik regresyon, RF, K-NN, SVM ve YSA sınıflandırma yöntemleri uygulanmıştır. En yüksek doğruluk oranları sırasıyla lojistik regresyon ile %67,56, RF ile %77,41, K-NN ile %72,07, SVM ile %74,17 ve YSA ile %75,17 olarak elde edilmiştir [6]. GTZAN veri seti ile verilerin ön işlenmesi gerçekleştirilmiş ardından çeşitli sınıflandırma algoritmaları ile müzik türü sınıflandırması gerçekleştirilmiştir. Kullanılan sınıflandırma algoritmaları ve bu algoritmaların doğruluk oranları sırasıyla YSA için %70, SVM için %68,90, Çok Katmanlı Algılayıcılar için %68,70, Karar Ağaçları için %74,30 ve CNN için %91 olarak gözlemlenmiştir [7]. Spotify'dan elde edilen veri kümesi ile gerçekleştirilen müzik sınıflandırmasında verilerin ön işlenerek sınıflandırmaya hazır hale getirilmiştir. Filtreleme yöntemleri ile öznelik seçimi işlemi gerçekleştirilmiştir. K-NN, RF ve lojistik regresyon ile gerçekleştirilen sınıflandırmada en

yüksek doğruluk oranları sırasıyla K-NN ile %95,5, RF ile %96,3, lojistik regresyon ile %95,5 olarak elde edilmiştir [8]. Dört farklı kategoride müzik sınıflandırma gerçekleştirilmiştir. MFCCs özellikleri öznitelik olarak seçilmiştir. 400 ses dosyasının her birinin 60 saniyelik kısımları kırılarak DNN algoritması ile başarılı bir sonuç elde edilmiştir. Veri seti olarak özgün bir veri seti kullanılmış olup %98,8 doğruluk oranı elde edilmiştir [9]. 16 bar MIDI dosyaları elde edilip 3 farklı model uygulanmıştır. Modeller: çift yönlü uzun kısa süreli bellek (BiLSTM), çarpımsal uzun kısa süreli bellek (mLSTM) ve tekrarlayan kapı birimi (GRU) modelleridir. Modellerin performansları birbirine yakın seviyede görülmüştür. Veri seti Wikifonia'dan elde edilen melodili müziklerle oluşturulmuştur. 1,0 AUC değeri yakalanmıştır [10]. Uygulanan Co-MRMR algoritması; MRMR ve Rasco algoritmalarının tümleştirilmesi ile elde edilmiş olup öznitelikler Databionic Music Miner ve MARYSAS yazılımları kullanılarak elde edilmiştir. Veri seti olarak özgün bir veri seti oluşturulup kullanılmıştır. %77,8 doğruluk oranına ulaşılabilmektedir [11]. MFCC'ler ile heat map'ler elde edilip bu map'lerden duyguya göre, enstrümana göre, cinsiyete göre, tempoya göre, türe göre ve çeşitli parametrelere göre sınıflandırma yapılmıştır. Veri seti olarak MagnaTagATune ve Million Song Dataset kullanılmıştır. Self-attention based deep sequence ve CNN algoritmaları kullanılmış olup %90,8 AUROC değeri elde edilebilmiştir [12]. Boyut indirgeme için PCA kullanılmış ve bunun yanı sıra tSNE ve MDS teknikleriyle renklere ayrıştırma işlemi gerçekleştirilmiştir (mavi, mor, yeşil ve sarı). Daha sonra renkler ile sınıflandırma yapılmıştır. Özgün bir veri seti oluşturulup kullanılmıştır. Random forest, C4.5, k-star, naive bayes, simple logistic regression, multilayer perceptron, SVM ve AdaBoost algoritmaları kullanılmış olup %72 doğruluk oranı elde edilebilmiştir [13]. Dijital ses sinyalinin işlenerek ilk olarak görüntü elde etme daha sonra ise görüntüyü parçalarına ayırma işlemi uygulanmıştır. Ses verilerinin işlenmesinde; gürültü ekleme, perde kaydırma, ses yüksekliği değişimi ve zaman uzatma gibi teknikler uygulanmıştır. Yalnızca CNN'ye dayalı yaklaşımla ve tek tonlu perde kaydırma ile en yüksek başarımlar elde edilmiştir. Veri seti olarak LMD (Latin Müzik Veritabanı) kullanılmıştır. %89,45 doğruluk oranı elde edilebilmiştir [14]. GTZAN veri seti ile müziklerden veri ön işleme ile çeşitli renkli görüntüler elde edilmiştir. Ardından bu görüntüler ile CNN kullanılarak sınıflandırma işlemi gerçekleştirilmiştir. %60 doğruluk oranı elde edilebilmiştir [15]. Özgün bir veri seti oluşturulup öznitelik olarak MFCCs kullanılmıştır. Pop, rock ve klasik olmak üzere 3 farklı türe müzik sınıflandırması yapılmıştır. SVM algoritması ile yapılan çalışmada farklı çekirdek fonksiyonlarının başarımları incelenerek kıyaslanmıştır. %91 doğruluk elde edilebilmiştir [16]. MIDI türünde veriler ile duyguya göre müzik sınıflandırması yapılmıştır. Duygu belirleme hususundaki deneyler lisans düzeyindeki öğrencilerle yapılmıştır. 3 farklı algoritmanın sonuçları bulunarak karşılaştırmalı analizi gerçekleştirilmiştir. Youtube'dan elde edilen veriler ile RNN, FNN, Geliştirilmiş FNN ile gerçekleştirilen çalışmada %75,4 doğruluk oranı elde edilebilmiştir [17]. GTZAN veri seti ile MFCCs özniteliği kullanılarak 10 farklı türe müzik sınıflandırması gerçekleştirilmiştir. CNN algoritması ile gerçekleştirilen çalışmada %76 doğruluk oranı elde edilmiştir [18].

3. MATERYAL VE METOTLAR

3.1. Knime Aracı

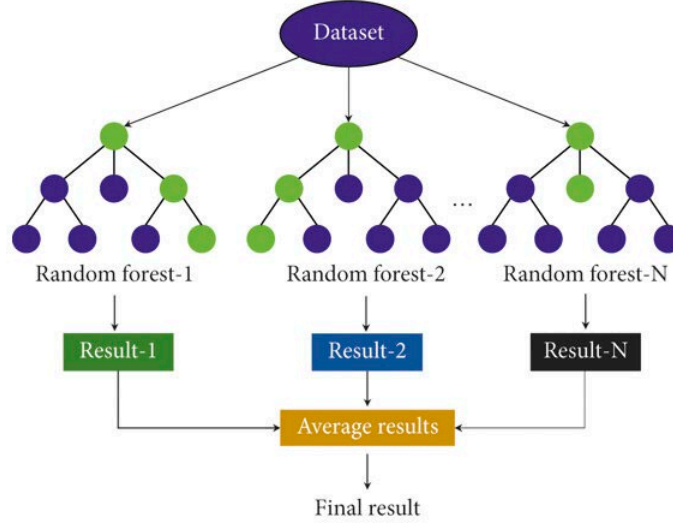
Açık kaynak kodlu bir yazılım olan KNIME veri biliminde yaygın olarak kullanılan bir platformdur. Bu program ile birçok veri işleme tekniği uygulanabilmektedir. Regresyon analizi, sınıflandırma, istatistiksel değerlendirme gibi işlemler düğümle vasıtasıyla gerçekleştirilmektedir. KNIME birçok algoritma için algoritmanın en yakın hali ile kullanılmasına olanak sağlayan bir arayüzdür. Kullanıcıyı kod karmaşasından uzaklaştırmak gibi bir imkân sunar. KNIME'da birçok dosya türü okunabilmektedir. Bunun için çeşitli modüller mevcuttur. Veri analizinde çeşitli modüller yardımıyla detaylı bir şekilde hem veriler ön işlenebilmekte hem de üretilen sonuçlar çeşitli görseller ve grafiklerle analiz edilebilmektedir. Verilerin ön işlenmesi süreci modüller yardımıyla detaylı bir şekilde gerçekleştirilebilmektedir.

3.2. Veri Kümesi

GTZAN veri kümesindeki tüm parçalar 30 saniyelik kesitlerden oluşmaktadır. Veri kümesi farklı türde müziklerden oluşmaktadır. Bu türler sırasıyla blues, classical, country, disco, hiphop, jazz, metal, pop, reggae ve rock türleridir. GTZAN veri kümesi, müzik türlerinin sınıflandırılması için yaygın olarak kullanılan bir veri kümesidir, ancak bazı kısıtlamaları vardır. Örneğin, veri kümesi yalnızca 10 müzik türünü içermekte, her türden 100 parça bulunmakta ve her parça 30 saniyelik bir kesittir. Bu nedenle, daha fazla müzik türü, daha fazla parça ve daha uzun süreli kesitler içeren farklı veri kümeleri kullanmak, sınıflandırma performansını artırabilecektir. Müziklerin sinyallerinden elde edilen birçok özellik mevcuttur. Bunlardan başlıcası ve en belirleyici olanı MFCCs özelliğidir. Veri kümesinde bu bilgilere ek olarak müzik sinyallerinden elde edilmiş görüntüler de mevcuttur. Veri kümesinin boyutu yaklaşık olarak 1,35 gigabyte civarındadır. Veri kümesindeki öznitelik sayısı 57'dir. Bu öznitelikler sırasıyla chroma_stft, rms, spectral_centroid, spectral_bandwidth, rolloff, zero crossing rate, harmony, perceptr, tempo ve MFCCs (Mel Frekans Kepstral Katsayıları) olarak belirlenmiştir. Bu özniteliklerden tempo haricindeki tüm öznitelikler için hem ortalama hem varyans bilgileri mevcuttur. Tempo için ise yalnızca 1 sayısal değer mevcuttur. Ayrıca MFCCs özniteliği için 20 vektörlü bir hesaplama yapılmıştır.

3.3. Rastgele Orman

Rastgele orman algoritması, esnek bir makine öğrenmesi algoritmasıdır ve çeşitli amaçlar için kullanılabilir. Hem sınıflandırma hem de regresyon problemlerine çözümler sunabilir. Rastgele orman, karar ağaçlarının geliştirilmiş bir türüdür ve bölütleme işlemi rastgele olarak gerçekleştirilir. Amacı, en uygun dalları bulmak ve bu dallara dayanarak alt dallara geçmektir [19-20]. Rastgele orman algoritması, topluluk öğrenme algoritmalarından biridir. Topluluk öğrenme yöntemlerinde, modellerin birleştirilmesiyle performansın artırılması hedeflenir. Rastgele orman algoritması, sürekli ve sürekli olmayan değişkenlerin bir arada kullanılmasına izin verir. Ayrıca büyük boyutlu veri setlerinde etkili bir şekilde kullanılabilir [21]. Rastgele orman algoritmasının işleyiş mantığı aşağıdaki şekilde verilmiştir.

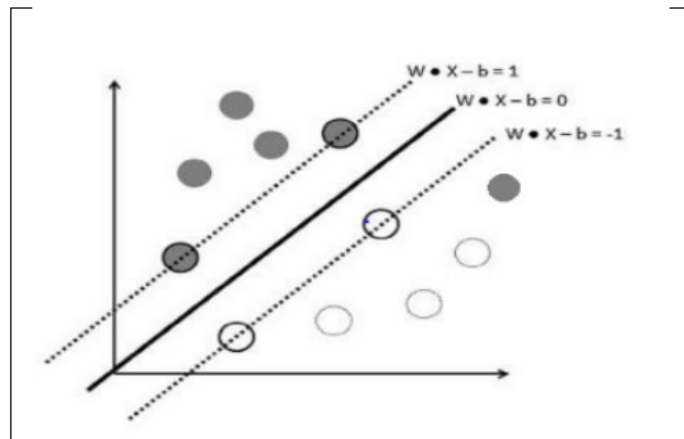


Şekil 1. Rastgele Orman Algoritmasının uygulanması [22].

Ağaç tabanlı ve özyinelemeli bir algoritma olan Rastgele Orman, önceden belirlenmiş bir durdurma koşulu karşılancaya kadar belirli bir kriterle dayalı olarak verilen veri kümesini özyinelemeli olarak iki gruba ayırmayı içermektedir [23]. Birçok sınıflandırma ve regresyon probleminde başarılı sonuçlar veren rastgele orman algoritması, müzik alanında da iyi sonuçlar verdiği görülmektedir.

3.4. Destek Vektör Makineleri

Destek vektör makineleri, sınıflandırma için etkili bir yöntemdir ve genellikle başarılı sonuçlar verir. Bu yöntem, bir veya daha fazla özellik vektörü kullanarak etiketleri tahmin etmeye yardımcı olur. Sınıflandırma işlemi, sınıflar arasında bir karar sınırı oluşturarak gerçekleştirilir [24]. Radyal tabanlı çekirdek, lineer çekirdek ve polinomal çekirdek olmak üzere üç yaygın olarak kullanılan çekirdek fonksiyonu bulunur. Veri ön işleme ve benzer işlemler, başarı için büyük öneme sahiptir ve seçilen çekirdek fonksiyonu da büyük ölçüde başarıyı etkiler. Genellikle ideal çekirdek fonksiyonun seçimi, tüm çekirdek fonksiyonlarını deneyerek belirlenir.

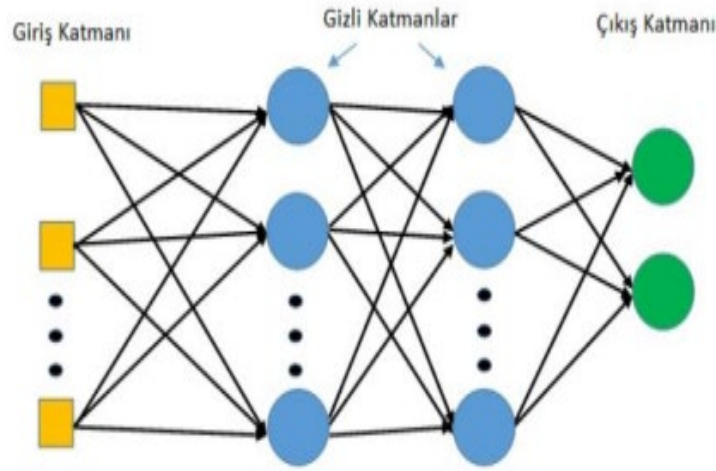


Şekil 2. SVM algoritmasında hiper düzlemler [25].

Verilerin sınıflandırılması, hiper düzlemler aracılığıyla gerçekleştirilir. En iyi ayrıştırıcı, ideal hiper düzlem olarak kabul edilebilir. Hiper düzlemler arasındaki mesafe, marjin olarak adlandırılır. SVM algoritmasında, öğrenme aşamasındaki işlem sayısının az olması, performansa olumlu etki eder. SVM algoritması doğrusal SVM ve doğrusal olmayan SVM olarak ikiye ayrılır [26]. Doğrusal olmayan SVM daha çok, veri kümelerinin karmaşık ve verilerin ayırt edilmesi zor olduğu durumlarda tercih edilmektedir. Doğrusal olmayan SVM algoritmasında çizilecek olan hiper düzlem doğrusal olmayıp optimum sonucu elde edebilecek biçimde çeşitli şekillerde olabilmektedir.

3.5. Yapay Sinir Ağları

Yapay sinir ağları (YSA), insan beyninin işleyiş mantığını taklit eden bir algoritmadır. Nöronlar arasındaki iletişim benzer şekilde bu algorithmada uygulanmaktadır. Nöronlar girdi, çıktı ve gizli katman olmak üzere 3 katmandan oluşmaktadır [27]. Bir YSA, yapay nöronlardan oluşan bir grup düğümlerden oluşmaktadır. Sinapslar arasındaki bağlantıya benzer şekilde, nöronlar arası sinyal aktarımı söz konusudur. Sinyali alan nöron, gerek işlem yaparken gerekse iletişime geçerek sinyal ulaştırabilmektedir [28]. Bu algoritma sınıflandırmada oldukça başarılı sonuçlar vermektedir. YSA'nın işleyiş yapısı aşağıdaki şekilde verilmiştir.



Şekil 3. YSA katmanlar arası yapısal ilişki [29].

Giriş katmanındaki veriler gizli katmanlara iletilerek işlenilip nihai olarak çıkış katmanına aktararak sonuç elde edilmektedir. YSA'da eğitim için ayrılan veriler ve test için ayrılan verilerin miktarı elde edilecek sonuçların başarımı açısından ciddi önem arz etmektedir [30]. Daha yüksek veri hacmi birçok durumda daha olumlu sonuçları beraberinde getirecektir.

3.6. Başarım Ölçütleri

Gerçekleştirilen çalışmada 10 farklı sınıf için 5 ayrı metrik hesaplanmıştır. Hesaplanan metrikler accuary (doğruluk), hassasiyet (recall), kesinlik (precision), özgünlük (specifity), F1 Skor (F1 Score) metrikleridir. Metriklerin hesaplanması, TP (True Positive), TN (True Negative), FP (False Positive) ve FN (False Positive) değerleri ile gerçekleştirilmektedir. Bu metriklerin her biri farklı şekillerde; TP, TN, FP, FN değerlerinin ne kadar tutarlı bir şekilde ölçüldüğü hakkında bilgi vermektedir [31]. Bu başarımler metriklerinin üreteceği sonuçlar, sınıflar arasındaki ilişkiler hakkında da fikir verici olmaktadır [32-33].

$$\text{Doğruluk} = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$\text{Hassasiyet} = \frac{TP}{TP+FN} \quad (2)$$

$$\text{Kesinlik} = \frac{TP}{TP+FP} \quad (3)$$

$$\text{Özgünlük} = \frac{TN}{TP+FN} \quad (4)$$

$$\text{F1 Skor} = \frac{2 * \text{Hassasiyet} * \text{Kesinlik}}{\text{Hassasiyet} + \text{Kesinlik}} \quad (5)$$

4. BULGULAR

Bu çalışmada, GTZAN veri kümesi kullanılarak 10 farklı müzik türünün sınıflandırılması için çeşitli öznelik kombinasyonları deneyerek üç farklı algoritma uygulanmıştır. Elde edilen sonuçlara göre, RF algoritması %81

doğruluk oranı ile en yüksek başarıma ulaşmıştır. SVM algoritması %72,33, YSA algoritması ise %67,66 doğruluk oranı ile diğer algoritmalarından daha düşük bir performans göstermiştir. Özniteliklerin seçimi, sınıflandırma başarısını önemli ölçüde etkilemiştir. Her bir algoritma için farklı optimum öznitelik kümesi belirlenmiştir. Müzik sinyallerinden elde edilen özniteliklerin sınıflandırmaya olan katkısı gözlemlenmiştir. Klasik müzik türü, tüm algoritmalarda en yüksek sınıflandırma başarımına sahip olmuştur. Rock müzik türü ise tüm algoritmalarda en düşük sınıflandırma başarımına sahip olmuştur.

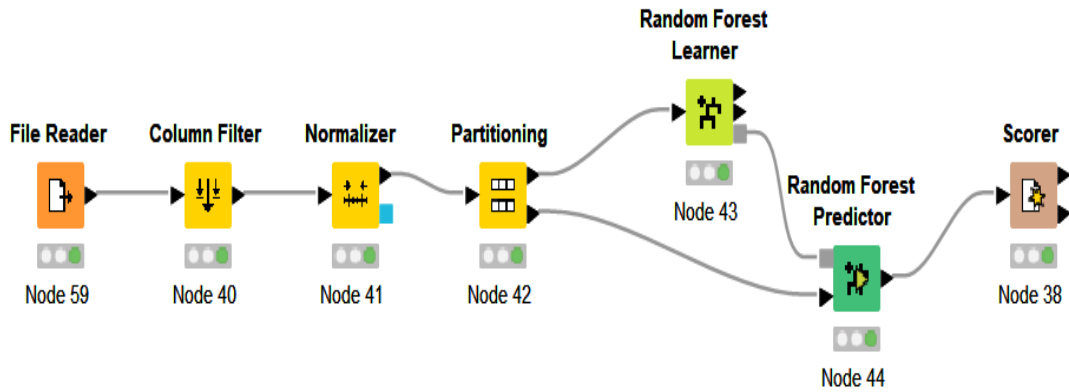
Öznitelik seçimi, sınıflandırma başarımını etkileyen önemli bir faktördür. Bu çalışmada, ses karakteristik özellikleri arasından en uygun olanlarının seçilmesi için çeşitli kombinasyonlar denemiştir. Çeşitli öznitelikler arasında en belirleyici olanının MFCC özniteliği olarak görülmüştür.

Sınıflandırılan müzik türleri arasında, klasik müzik türüne ait sonuçlar diğer türlere göre daha yüksek olmuştur. Bu durumun, klasik müziğin diğer müzik türlerinden daha belirgin ve ayırt edici özelliklere sahip olmasından kaynaklandığı düşünülmektedir. Rock müzik türü ise en düşük sonuçlar elde edilen tür olmuştur. Bu durumun ise rock müziğin diğer müzik türleriyle karıştırılmasına neden olan benzerlikler içermesinden dolayı olduğu düşünülmektedir. Bu çalışma, GTZAN veri kümesi ile müzik türlerini sınıflandırmak için üç farklı makine öğrenmesi algoritması (RF, SVM ve YSA) kullanmıştır. Bu konularda çalışma yapmayı planlayan araştırmacılara şu konu başlıklarına yönelik çalışmaları önerilebilir: GTZAN veri kümesi, müzik türlerinin sınıflandırılması için yaygın olarak kullanılan bir veri kümesidir, ancak bazı kısıtlamaları vardır. Örneğin, veri kümesi yalnızca 10 müzik türünü içermekte, her türden 100 parça bulunmakta ve her parça 30 saniyelik bir kesittir. Bu nedenle, daha fazla müzik türü, daha fazla parça ve daha uzun süreli kesitler içeren farklı veri kümeleri kullanmak, sınıflandırma performansını ve genelleştirilme gücünü artırabilecektir. Bu çalışmada, müzik sinyallerinden elde edilen 57 öznitelik kullanılmıştır. Ancak bu özniteliklerin hepsi sınıflandırma için eşit derecede önemli veya uygun olmayabilir. Bazı öznitelikler gürültü, aykırı korelasyon veya çok boyutluluk gibi sorunlara neden olabilmektedir. Bu nedenle, farklı öznitelik seçim veya çıkarma yöntemleri kullanmak sınıflandırma için en uygun öznitelik kümesini belirlemeye yardımcı olabilmektedir.

Bu çalışmada kullanılan algoritmaların her biri farklı parametreler veya hiper-parametreler gerektirir. Bu parametrelerin değerleri, sınıflandırma sonucunu etkileyebilir. Bu nedenle, farklı algoritmalar veya parametreler denemek, sınıflandırma için en uygun modeli bulmaya yardımcı olacaktır.

4.1. Rastgele Orman ile Elde Edilen Sonuçlar

Sınıflandırıcı olarak Rastgele Orman seçildiği durum için ilk olarak veri kümesine min-max normalizasyonu uygulanmıştır. Ardından veri kümesi, eğitim için %70 test için ise %30 olacak şekilde ayrılmıştır. Daha sonra rastgele orman algoritması uygulanarak sınıflandırmanın işlemi gerçekleştirilmiştir. Elde edilen başarımlar diğer algoritmalarla kıyasla daha yüksektir. Tüm sınıflar için elde edilen başarımların sonuçları ve akış diyagramları aşağıda detaylarıyla verilmiştir.



Şekil 4. KNIME ile RF algoritmasının uygulanması.

RF algoritması ile gerçekleştirilen sınıflandırmada, spektral özniteliklerden yalnızca spectral_centroid özniteliği değerlendirilecek öznitelik kümesinden çıkarıldığı durumda en yüksek başarımlar elde edilmiştir. Elde edilen doğruluk oranı %81 olarak görülmüştür.

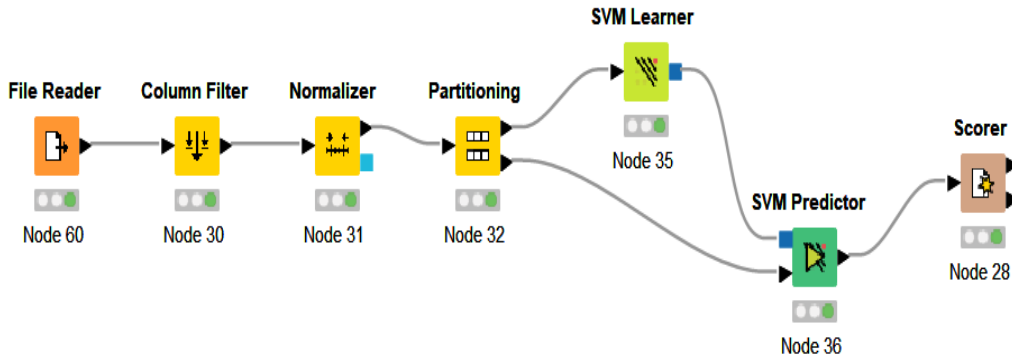
RF ile elde edilen başarımların sonuçlarına bakıldığında, klasik müzik türüne ait sonuçlar oldukça yüksektir. Rock müzik türünde ise en düşük değerler elde edilmiştir.

Tablo 1. RF ile başarımların sonuçları.

Metrikler					
Türler	Recall	Precision	Sensitivity	Specificity	F1-Score
blues	0,81	0,71	0,81	0,97	0,76
classical	0,91	0,95	0,91	0,99	0,93
country	0,69	0,78	0,69	0,97	0,74
disco	0,91	0,81	0,91	0,97	0,85
hiphop	0,83	0,79	0,83	0,98	0,81
jazz	0,81	0,83	0,81	0,98	0,82
metal	0,96	0,86	0,96	0,99	0,91
pop	0,88	0,81	0,88	0,97	0,85
reggae	0,75	0,87	0,75	0,98	0,81
Rock	0,64	0,72	0,64	0,97	0,68

4.2. SVM ile Elde Edilen Sonuçlar

Sınıflandırıcı olarak SVM seçildiği durum için ilk olarak veri kümesine min-max normalizasyonu uygulanmıştır. Ardından veri kümesi, eğitim için %70 test için ise %30 olacak şekilde ayrılmıştır. Daha sonra SVM algoritması uygulanarak sınıflandırmanın işlemi gerçekleştirilmiştir. Tüm sınıflar için elde edilen başarımların sonuçları ve akış diyagramları aşağıda detaylarıyla verilmiştir.

**Şekil 5.** KNIME ile SVM algoritmasının uygulanması.

SVM ile uygulanan sınıflandırma işleminde, rms ve spektral öznitelikleri öznitelik kümesinden çıkarıldığı durumda en yüksek başarımlar elde edilmiştir. Elde edilen doğruluk oranı %72.33 olarak görülmüştür. SVM ile elde edilen başarımların sonuçlarına bakıldığında görülmektedir ki klasik müzik türüne ait sonuçlar oldukça yüksektir. Rock müzik türünde ise en düşük değerler elde edilmiştir.

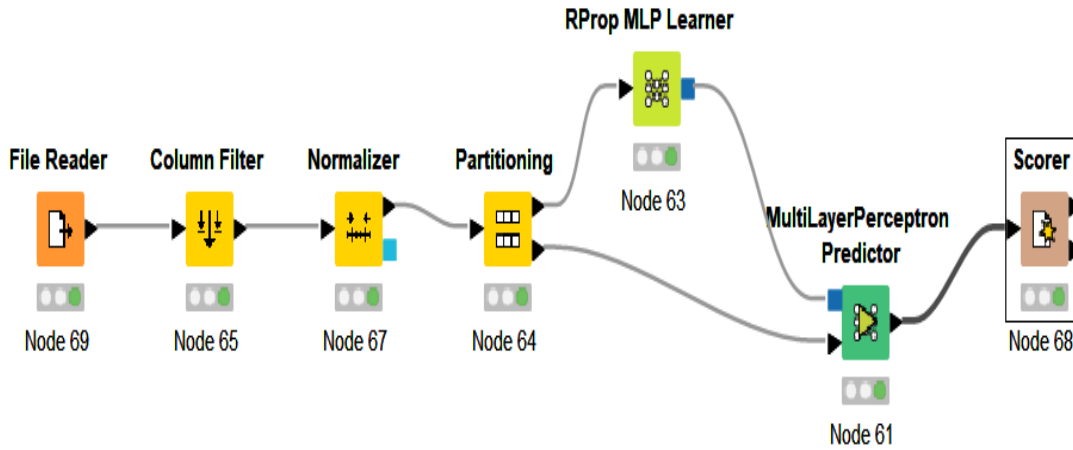
4.3. YSA ile Elde Edilen Sonuçlar

Sınıflandırıcı olarak YSA seçildiği durum için ilk olarak veri kümesine min-max normalizasyonu uygulanmıştır. Ardından veri kümesi, eğitim için %70 test için ise %30 olacak şekilde ayrılmıştır. Daha sonra yapay sinir ağları algoritması uygulanarak sınıflandırmanın işlemi gerçekleştirilmiştir. Tüm sınıflar için elde edilen başarımların sonuçları ve akış diyagramları aşağıda detaylarıyla verilmiştir.

YSA algoritması ile gerçekleştirilen sınıflandırmada, özniteliklerden yalnızca rms özniteliği değerlendirilecek öznitelik kümesinden çıkarıldığı durumda en yüksek başarımlar elde edilmiştir. Elde edilen doğruluk oranı %67.67 olarak görülmüştür.

Tablo 2. SVM ile başarımların metriklerinin sonuçları.

Metrikler					
Türler	Recall	Precision	Sensitivity	Specificity	F1-Score
blues	0,72	0,61	0,72	0,94	0,66
classical	0,94	0,89	0,94	0,98	0,91
country	0,63	0,74	0,63	0,98	0,68
disco	0,5	0,71	0,5	0,98	0,59
hiphop	0,48	0,83	0,48	0,99	0,61
jazz	0,86	0,82	0,86	0,97	0,84
metal	0,92	0,76	0,92	0,96	0,83
pop	0,88	0,77	0,88	0,97	0,82
reggae	0,83	0,54	0,83	0,94	0,66
rock	0,34	0,53	0,34	0,97	0,42

**Şekil 6.** KNIME ile YSA algoritmasının uygulanması.

YSA ile elde edilen başarımların metrikleri sonuçlarına bakıldığında diğer sınıflandırma algoritmalarındaki gibi yine klasik müzik türüne ait sonuçlar oldukça yüksektir. Rock müzik türünde de yine diğer algoritmalarındaki benzer şekilde en düşük değerler elde edilmiştir.

5. SONUÇ

Bu çalışmada, GTZAN veri kümesi kullanılarak müzik türlerinin sınıflandırılması için farklı özellikler içeren bir yöntem geliştirilmiştir. Rastgele Orman (RF), Destek Vektör Makineleri (SVM) ve Yapay Sinir Ağları (YSA) gibi sınıflandırma algoritmaları kullanılarak elde edilen sonuçlar incelenmiştir.

3 farklı algoritma ile gerçekleştirilen müzik türlerini sınıflandırma çalışmasında çeşitli özellik kombinasyonu deneyleri neticesinde RF ile %81, YSA ile %67,67, SVM ile %72,33 doğruluk oranı elde edilmiştir. Elde edilen sonuçlar gösteriyor ki RF algoritması diğer 2 algoritmadan az bir farkla daha yüksek bir başarıma ulaşmıştır. Bu çalışma, müzik sınıflandırmasında çeşitli algoritmaların ne oranda başarıma ulaşacağını incelemiştir. Ayrıca özelliklerin etkileri de irdelenmiştir. Algoritmalar arasında kayda değer bir fark olduğu ve müzik sinyallerinden elde edilen özelliklerin sınıflandırmaya olan etkisi gözlemlenmiştir. GTZAN veri kümesindeki oldukça çeşitli özelliklerin kullanılarak hangisinin başarıma olan etkisinin ne kadar yüksek olduğunun ölçülebilmesi imkânı elde edilmiştir.

Özelliklerin seçimi, sınıflandırma için en uygun özellik kümesini belirlemek için önemli bir faktördür. Bu çalışmada, ses karakteristik özellikleri arasından en uygun olanlarının seçilmesi için çeşitli kombinasyonlar denemiştir. Çeşitli özellikler arasında en belirleyici olanının MFCC özelliği olarak görülmüştür. Sınıflandırılan

Tablo 3. YSA ile başarımların metriklerinin sonuçları.

Metrikler					
Türler	Recall	Precision	Sensitivity	Specificity	F1-Score
blues	0,62	0,55	0,62	0,95	0,58
classical	0,86	0,91	0,86	0,99	0,89
country	0,71	0,57	0,71	0,94	0,63
disco	0,72	0,62	0,72	0,95	0,67
hiphop	0,73	0,76	0,73	0,98	0,75
jazz	0,85	0,65	0,85	0,96	0,73
metal	0,81	0,79	0,81	0,98	0,80
pop	0,74	0,74	0,74	0,97	0,74
reggae	0,54	0,71	0,54	0,97	0,61
rock	0,15	0,27	0,15	0,96	0,20

müzik türleri arasında, klasik müzik türüne ait sonuçlar diğer türlere göre daha yüksek olmuştur. Bu durumun, klasik müziğin diğer müzik türlerinden daha belirgin ve ayırt edici özelliklere sahip olmasından kaynaklandığı düşünülmektedir. Rock müzik türü ise en düşük sonuçlar elde edilen tür olmuştur. Bu durumun ise rock müziğin diğer müzik türleriyle karıştırılmasına neden olan benzerlikler içermesinden dolayı olduğu düşünülmektedir.

RF algoritmasıyla elde edilen başarı oranı %81 olarak belirlenmiştir. Bu sonuç, RF'nin müzik türlerini sınıflandırmak için oldukça etkili bir yöntem olduğunu göstermektedir. SVM algoritmasıyla elde edilen başarı oranı %72,33 olarak belirlenmiştir. SVM algoritmasının da müzik türlerini sınıflandırmada başarılı olduğu ancak RF kadar etkili olmadığı görülmektedir. YSA algoritmasıyla elde edilen başarı oranı ise %67,67 olarak belirlenmiştir. YSA'nın da müzik türlerini sınıflandırmada kullanılabilecek ancak diğer iki algoritma kadar etkili olmayabileceği sonucuna varılmıştır.

Bu sonuçlar, müzik türlerinin sınıflandırılmasında RF ve SVM gibi makine öğrenimi algoritmalarının etkili bir şekilde kullanılabileceğini göstermektedir. Ancak, daha kapsamlı ve hassas bir sınıflandırma için farklı özelliklerin ve algoritmaların da incelenmesi gerektiği unutulmamalıdır. Bu çalışma, müzik ve bilgisayar bilimleri arasındaki etkileşimi anlamak ve müzik türlerinin analizini geliştirmek için önemli bir adım olarak değerlendirilebilir.

Literatürdeki mevcut çalışmalarda GTZAN veri kümesi ile yapılan çalışmalarda genellikle CNN algoritması uygulanmıştır. Bu çalışmada, ses sinyallerinden elde edilen özellikler ile müzik sınıflandırmada sıkça kullanılan 3 farklı makine öğrenmesi algoritmasının performansı ölçülerek başarımları karşılaştırılmıştır. Bunun yanı sıra bu çalışma gerek makine öğrenmesi algoritmalarının gerekse derin öğrenme algoritmalarının çalıştırılabildiği KNIME platformunda gerçekleştirilmiştir.

Bu çalışmada, müzik sinyallerinden elde edilen 57 özellik kullanılmıştır. Ancak, bu özelliklerin hepsi sınıflandırma için eşit derecede önemli veya uygun olmayabilir. Bazı özellikler gürültü, aykırı korelasyon gibi sorunlara neden olabilmektedir. Bu nedenle, farklı özellik seçim veya çıkarma yöntemleri kullanmak sınıflandırma için en uygun özellik kümesini belirlemeye yardımcı olabilmektedir. Bu çalışmada kullanılan algoritmaların her biri farklı parametreler veya hiper-parametreler gerektirir. Bu parametrelerin değerleri, sınıflandırma sonucunu etkileyebilir. Bu nedenle farklı algoritmalar veya parametreler denemek, sınıflandırma için en uygun modeli bulmaya yardımcı olacaktır.

6. TARTIŞMA

Müzik sınıflandırmada belirli türler arasında ciddi benzerlikler bulunmaktadır. Örneğin rock müzik türü ve metal müzik türü arasında ciddi benzerlikler bulunmaktadır. Bu çalışmada da bu 2 tür arasında ciddi benzerlikler gözlemlenmiştir. Bu durumun aşılması ve daha isabetli sonuçların üretilebilmesi için araştırmacılar hibrit türler üzerinde çalışmalara yönelebilir. Bu durumda belirli oranlarda örtüşen türlere ait parçaların hibrit kabul edilerek tür etiketlerinin güncellenmesi ve bu sayede sınıflandırma çalışmalarının daha belirleyici olabileceği öngörülmektedir.

Gerçekleştirilen sınıflandırma çalışması ile makine öğrenmesi yöntemlerinden öne çıkan yöntemlerin başarımlarının karşılaştırılması hedeflenmiştir. Bu yöntemlerden en başarılı sonucu veren yöntem Rastgele Orman olarak tespit edilmiştir. Bu yöntemin müzik sınıflandırma çalışmalarında oldukça etkili bir yöntem olduğu ve dikkate değer bir algoritma olacağı düşünülmektedir. Bunun yanı sıra farklı algoritmaların uygulanması ile

özellikle derin öğrenme algoritmaları ve çeşitli hiper-parametre kombinasyonları ile başarılı sonuçlar elde edilebileceği de göz ardı edilmemelidir. Kullanılan veri kümesi oldukça geniş bir öznitelik havuzuna sahiptir. Bu durumun ileriye dönük çalışmalarda benzer bir veri kümesi oluşturulmasında ilham verici olabileceği unutulmamalıdır.

Bu makine öğrenme yöntemlerinin yanı sıra çeşitli derin öğrenme yöntemleri de eklenerek çalışmanın kapsamı genişletilebilir. Bunun için KNIME platformu araştırmacılara gerek verinin ön işlenmesi aşamasında gerekse sınıflandırma aşamasında çeşitli imkanlar sunmaktadır.

Yazar Katkıları

Her bir yazar eşit seviyede katkı sunmuştur.

Çıkar Çatışması

Makale yazarları aralarında herhangi bir çıkar çatışması olmadığını beyan ederler.

KAYNAKÇA

- [1] M.B. Er, H. Çiğ, "Türk müziği uyarınları kullanılarak insan duygularının makine öğrenmesi yöntemi ile tanınması", Gazi University Journal of Science Part C: Design and Technology, vol. 8, no. 2, pp. 458-474, 2020.
- [2] M. Yılmaz, H. Şahin, A. Yıldız, "Sectoral application analysis of studies made with deep learning models" Electronic Letters on Science and Engineering, vol. 17, no. 2, pp.126-140, 2021.
- [3] S. Çeven, R. Bayır, "Ortam Sesinden İnsan Sesinin Ayırıştırılması için Filtre Geliştirilmesi", Avrupa Bilim ve Teknoloji Dergisi, pp. 331-337, 2020.
- [4] H.E. Kocer, M.C. Ahmed, "Turkish Speech recognition using Mel-frequency cepstral coefficients (MFCC) and Hidden Markov Model (HMM)", Veri Bilimi, vol. 2, no. 2, pp. 39-44, 2019.
- [5] N. Ndou, R. Ajoodha, A. Jadhav, "Music genre classification: A review of deep-learning and traditional machine-learning approaches", In 2021 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS), IEEE, pp. 1-6, 2021.
- [6] J. SuriyaPrakash, S. Kiran, "Obtain Better Accuracy Using Music Genre Classification System on GTZAN Dataset", In 2022 IEEE North Karnataka Subsection Flagship International Conference (NKCon), IEEE, pp. 1-5, 2022.
- [7] A. Ghildiyal, K. Singh, S. Sharma, "Music genre classification using machine learning", In 2020 4th international conference on electronics, communication and aerospace technology (ICECA), IEEE, pp. 1368-1372, 2020.
- [8] F. Khan, I. Tarimer, H.S. Alwageed, B.C. Karadağ, M. Fayaz, A. B. Abdusalomov, Y.I. Cho, "Effect of feature selection on the accuracy of music popularity classification using machine learning algorithms", Electronics, vol. 11, no. 21, pp. 3518, 2022.
- [9] G. Jawaharlalnehru, S. Jothilakshmi, "Music genre classification using deep neural networks", International Journal of Scientific Research in Science, Engineering and Technology, vol. 4, pp. 935, 2018.
- [10] J. Guo, A. Liu, J. Xiao, "Melody classification based on performance event vector and BRNN", arXiv preprint arXiv:2010.07562, 2020.
- [11] Y. Yaslan, Z. Cataltepe, "Audio genre classification with co-mrmm", IEEE 17th Signal Processing and Communications Applications Conference, pp. 408-411, 2009.
- [12] M. Won, S. Chun, X. Serra, "Toward interpretable music tagging with self-attention", arXiv preprint arXiv:1906.04972, 2019.
- [13] C. Ritter, C. Altenhofen, M. Zeppelzauer, A. Kuijper, T. Schreck, J. Bernard, "Personalized Visual-Interactive Music Classification", In EuroVA@ EuroVis, pp. 31-35, 2018.
- [14] R.L. Aguiar, Y.M. Costa, C.N. Silla, "Exploring data augmentation to improve music genre classification with convnets", International Joint Conference on Neural Networks (IJCNN), IEEE, pp. 1-8, 2018.
- [15] Z. Durdağ, P. Erdoğan, "Müzik türlerinin derin öğrenme ağları ile sınıflandırılması", Sakarya University Journal of Computer and Information Sciences, vol. 2, no. 1, pp. 53-60, 2019.
- [16] R. Thiruvengatanadhan, "Music Classification using MFCC and SVM", International Research Journal of Engineering and Technology, vol. 5, pp. 922-924, 2018.
- [17] W. Zhao, Y. Zhou, Y. Tie, Y. Zhao, "Recurrent neural network for MIDI music emotion classification", IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), IEEE, pp. 2596-2600, 2018.

- [18] S. Vishnupriya, K. Meenakshi, "Automatic music genre classification using convolution neural network", International Conference on Computer Communication and Informatics (ICCCI), IEEE, pp. 1-4, 2018.
- [19] K.J. Archer, R.V. Kimes, "Empirical characterization of random forest variable importance measures", Computational statistics & Data analysis, vol. 52, no. 4, pp. 2249-2260, 2008.
- [20] A. Atalan, H. Şahin, Y.A. Atalan, "Integration of machine learning algorithms and discrete-event simulation for the cost of healthcare resources" In Healthcare, MDPI, vol. 10, no. 10, pp. 1920, 2022.
- [21] P. Akın, Y. Terzi, "Dengesiz veri setli sağkalım verilerinde cox regresyon ve rastgele orman yöntemlerin karşılaştırılması" Veri Bilimi, vol. 3, no. 1, pp. 21-25, 2020.
- [22] H. Fu, K. Qi, "Evaluation model of teachers' teaching ability based on improved random forest with grey relation projection", Scientific Programming, pp. 1-12, 2022.
- [23] M. Schonlau, R.Y. Zou, "The random forest algorithm for statistical learning" The Stata Journal, vol. 20, no. 1, pp. 3-29, 2020.
- [24] S. Huang, N. Cai, P.P. Pacheco, S. Narrandes, Y. Wang, W. Xu, "Applications of support vector machine (SVM) learning in cancer genomics", Cancer genomics & proteomics, vol. 15, no. 1, pp. 41-51, 2018.
- [25] S. Ghosh, A. Dasgupta, A. Swetapadma, "A study on support vector machine based linear and non-linear pattern classification", International Conference on Intelligent Sustainable Systems (ICISS), IEEE, pp. 24-28, 2019.
- [26] T.T. Bilgin, S.B. Altınışık, N.A. Adıgüzel, "A Comparative Study of Classification and Clustering Methods for Data Analysis in Digital Transformation and IoT Systems", Orclever Proceedings of Research and Development, vol. 3, no. 1, pp. 1-18, 2023.
- [27] J. Zou, Y. Han, S.S. So, "Overview of artificial neural networks", pp. 14-22, 2009.
- [28] R.M. Sadek, S.A. Mohammed, A.R.K. Abunbehan, A.K.H.A. Ghattas, M.R. Badawi, M.N. Mortaja, S.S. Abu-Naser, "Parkinson's disease prediction using artificial neural network", International Journal of Academic Health and Medical Research, vol. 3, no. 1, pp. 1-8, 2019.
- [29] M.E. Tabar, A.C. Başara, Y. Şişman, "Çoklu regresyon ve yapay sinir ağları ile Tokat ilinde konut değerlendirme çalışması", Türkiye Arazi Yönetimi Dergisi, vol. 3, no. 1, pp. 1-7, 2021.
- [30] S. Budak, B. Akbal, "Görüntü işleme ve yapay sinir ağları ile iletişim hatlarında arıza yeri belirleme", Konya Mühendislik Bilimleri Dergisi, vol. 8, no. 3, pp. 678-692, 2020.
- [31] F. Aydemir, S. Arslan, "A System Design with Deep Learning and IoT to Ensure Education Continuity for Post-COVID", IEEE Transactions on Consumer Electronics, 2023.
- [32] S. Dörterler, "Hybridization of k-means and meta-heuristics algorithms for heart disease diagnosis", New Trends in Engineering and Applied Natural Sciences, vol. 55, 2022.
- [33] S. Dörterler, "Kanser Hastalığı Teşhisinde Ölüm Oyunu Optimizasyon Algoritmasının Etkisi", Mühendislik Alanında Uluslararası Araştırmalar VIII, pp. 15, 2023.