

GİRİŞ

Proje Adı: Örüntü Tanıma ile Ses Sınıflandırması

Proje Konusu: Proje; yapay zekâ, makine öğrenmesi, veri madenciliği, örüntü tanıma alanlarını içeren bilimsel bir araştırma uygulamasıdır.

Problem Durumu: Örüntü tanıma teknikleri kullanılarak insan konuşma sesi ile müzik sesi ayırt edilebilir mi?

Proje Amacı: Belirlenen örneklem kümesi sayısında ses kaydı yoluyla alınan verileri örüntü tanıma teknikleri kullanarak doğru bir şekilde sınıflandırmak.

Alt Amaçlar

- Sınıflandırma algoritmaları uygulanması hakkında genel bilgi sahibi olmak
- KNN (K Nearest Neighborhood) sınıflandırma yönteminin teorisini ve uygulamasını öğrenmek
- SVM(Support Vector Machine) sınıflandırma yönteminin teorisini ve uygulamasını öğrenmek
- MATLAB, MIRTtoolbox araç kutusu ve WEKA programlarının kullanılmasını öğrenmek
- Yapay zeka teknolojilerine sınıflandırma algoritmaları üzerinden yeni bir bakış kazanmak

YÖNTEM

Araştırma Modeli

Ses verilerini müzik ya da konuşma olarak doğru sınıflandırmayı amaçlayan bu proje nicel bir araştırma olup "bağıntısal araştırma" modeli kullanılmıştır.

Evren-Örneklem (Çalışma Grubu)

Araştırmanın evreni tüm konuşma sesleri ile tüm müzik sesleridir. Araştırmanın örneklemini ise çeşitli dillerden, farklı yaş grupları ve farklı cinsiyetlere ait toplam 600 adet ses kaydı oluşturmaktadır.

Veri Toplama Araçları

Mp3 formatındaki ses dosyaları, belirlenen öznelikler ile MATLAB'ın MIRTtoolbox araç kutusu üzerinde işletilerek veri seti elde edilmiştir.

Mp3 formatındaki 300 konuşma, 300 müzik ses dosyası MATLAB üzerinde uygulama kodları üzerinden işletilerek veri seti oluşturulmuştur. Veri seti oluşturulurken MATLAB'ın ses sinyallerinden müzikal özelliklerin çıkarılması için hizmet veren araç kutusu MIRTtoolbox kullanılmıştır. Oluşturulan veri seti üzerinde sınıflandırma algoritmaları kullanılarak optimizasyon yapılmıştır. Optimizasyon sonucunda elde edilen verilerle WEKA uygulaması kullanılarak sistemin örüntüyü öğrenmesi için eğitim ve test işlemleri gerçekleştirilmiştir. Sınıflandırmanın doğruluğunu güvence altına almak için iki ayrı yöntem kullanılmıştır.

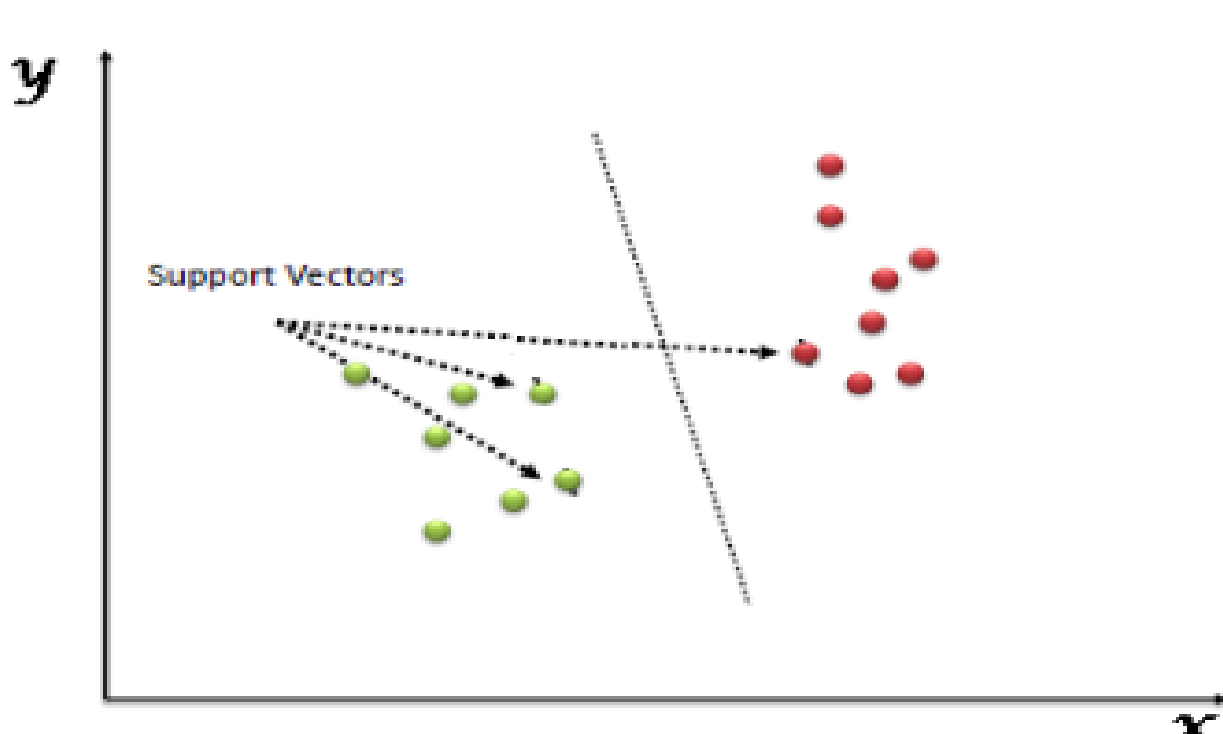
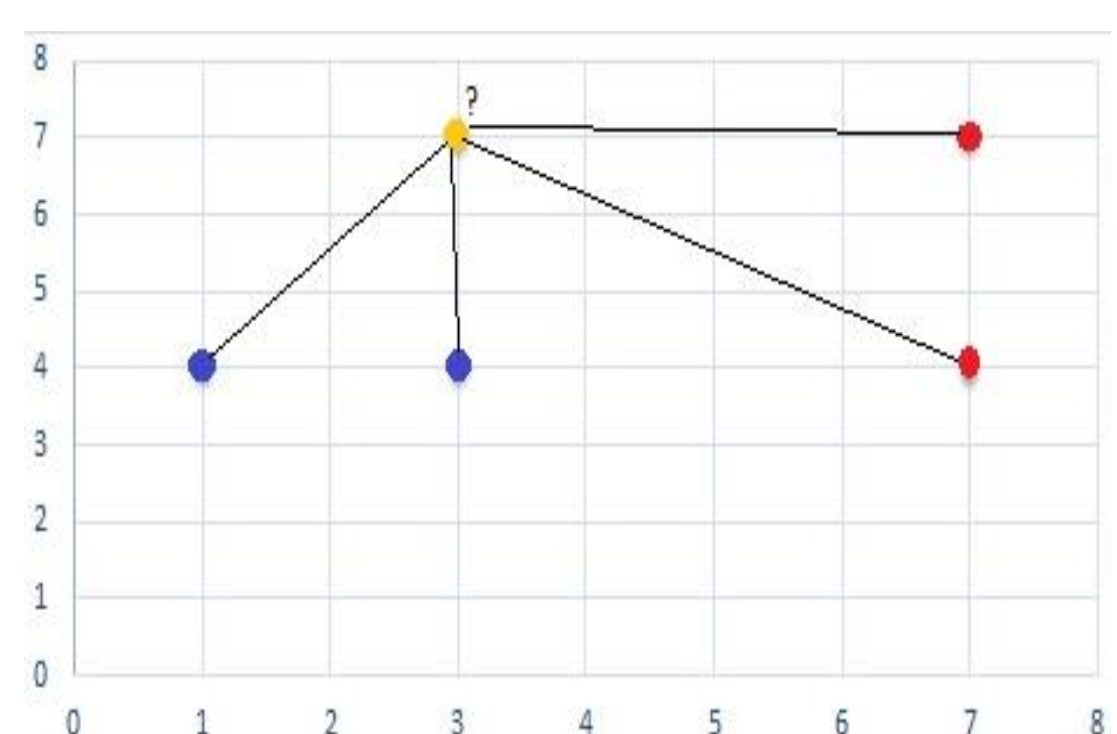
K nearest neighborhood (En Yakın K Komşu)

KNN ile temelde yeni noktaya en yakın noktalar aranır. K, bilinmeyen noktanın en yakın komşularının miktarını temsil eder. Yeni gelen her örneğin eğitim veri setindeki her örneğe olan uzaklığı tek tek hesaplanır. K-NN, eğitim veri setini doğrudan kullanarak tahmin yapar.

SVM (Destek Vektör Makineleri)

Destek Vektör Makineleri, temel olarak iki sınıfa ait verileri birbirinden en uygun şekilde ayırmak için kullanılır. Bunun için karar sınırları belirlenir.

Karar doğrusunun yeni katılacak olan veriye karşı dayanıklı olabilmesi için sınır çizgisinin, iki sınıfın sınır çizgilerine en yakın uzaklıkta olması gerekmektedir. Bu sınır çizgisine en yakın noktalar, destek noktaları olarak adlandırılmaktadır.



BULGULAR

KNN Sınıflandırma Yöntemine Göre ;

Correctly Classified Instances	398
Kappa statistic	0.3291
Mean absolute error	0.4131
Root mean squared error	0.456
Relative absolute error	82.6135 %
Root relative squared error	91.2091 %
Total Number of Instances	599

K=3 belirlenerek çalıştırılan verilerin bulguları yukarıdaki gibidir. Görüldüğü üzere doğru sınıflandırmış örnek sayısı 398 olup **doğruluk oranı: 66.4441 %** olarak bulunmuştur.

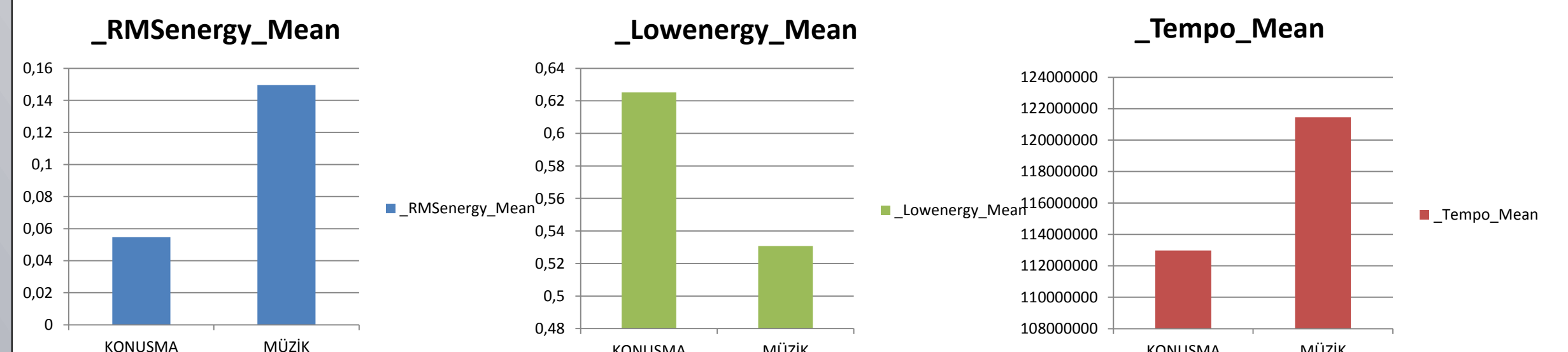
SVM Sınıflandırma Yöntemine Göre ;

Correctly Classified Instances	399
Kappa statistic	0.3325
Mean absolute error	0.3339
Root mean squared error	0.5778
Relative absolute error	66.778 %
Root relative squared error	115.5662 %
Total Number of Instances	599

SVM sınıflandırmasına göre çalıştırılan verilerde **doğruluk oranı: 66.611 %** bulunmuştur. Bu yöntem ile sınıflandırma işlemi KNN'e göre çok az miktarda değişiklik göstermiş olup sistemin tutarlı çalıştığı görülmüştür.

Confusion Matrix (Karşılaştırma Matrisi)

a b
223 76 | a = Konuşma
125 175 | b = Müzik



SONUÇ VE TARTIŞMA

Araştırma, konuşma sesi ile müzik sesini ayırt edebilen makine öğrenmesini gerçekleştirme amacıyla başlatılmış olup proje planında kabul edilebilir oran olarak belirlenen %65'in üzerine çıkmıştır. Projenin sonuçları oluşturulan hipotezleri desteklemektedir. KNN ve SVM yöntemlerinin her ikisinde de doğruluk oranı %66 oranında çıkmıştır. Veriler arasında belirleyici olarak gördüğümüz RMS, lowenergy, tempo, fluctuation gibi öznelikler beklenen şekilde farklılık göstermiştir ve bu farklılıklar grafikler üzerinde de gösterilmiştir.

Proje; yapay zekâ, makine öğrenmesi, veri madenciliği, örüntü tanıma alanlarını öğrenmeye ve uygulamaya yönelik bir araştırma çalışmasıdır. Bu araştırmanın, kapsamı genişletilip detaylandırılarak daha gelişmiş bir sistem üretmek için yüksek lisans veya doktora çalışmasıyla ilerletilmesi beklenmektedir.

KAYNAKÇA

- C. Ünsalan and A. Erçil, "Comparison of feature selection algorithms a new performance criteria for feature selection", Proceedings of IEEE SIU'98, pp. 60-65, May 1998
- Bozkurt Barış, Gedik Ali Cenk, Karaosmanoğlu M. Kemal, «1 TMVB: Klasik Türk Müziği İcra Analiz Çalışmaları İçin Bir Veri Bankası» www.researchgate.net, 2014
- Olivier Lartillot, «MIRtoolbox 1.7 User's Manual» University of Oslo, Norway Department of Musicology, June 2017
- www.mathworks.com