**Makine Öğrenmesi Kullanarak Müzik Enstrümanlarını Tanımak**

Mustafa Sami Şahin

1. **GİRİŞ**

Bu projede Nsynth dataset’i içerisinde verilen 11 farklı enstrümana ait ses kayıtlarını kullanarak yine aynı dataset içerisindeki enstrümanları tanımaya çalıştım.

1. **LİTERATÜR TARAMASI**

Benzer konulardaki projeleri araştırdığımda karşıma bir çok benzer çalışma çıktı. Bunlardan biri Music Instrument Detection Using Lstms And The Nsynth Dataset[1]. Bu çalışmada Nsynth datasetini Uzun Kısa Vadeli Hafıza Ağları algoritmaları kullanarak işlemeyi önermiş fakat işlemleri gerçekleştirmeyip gelecek dönemlere bırakılmış.

Bir diğer çalışma olarak “Can We Guess Musical Instruments With Machine Learning?"[2] isimli

yazıyı bulabildim. Burada da Nsynth datasetini Librosa kütüphanesine sokarak her ses dosyası için 5 ayrı özellik çıkartılmış ve bu özellikler 4 farklı algoritma kullanılarak test edilmiş ve başarı oranları paylaşılmış. Bu 4 algoritmadan Naive Bayes ve Support Vector Machines algoritmalarında başarı oranı yüzde 15’i geçemediği için elenmiş ve kalan 2 algoritmadan Rastgele Orman algoritması Evrişimli Sinir Ağları algoritmasından çok daha hızlı çalıştığı için Rastgele Orman algoritması seçilmiş. Fakat 4 algoritmadan en yüksek başarı yüzdesine sahip olan Rastgele Orman algoritması da çeşitli ayarlamalar ve düzenlemeler sonucunda ancak yüzde 65 başarı oranı göstermiş.

Ben de çalışmamda 2 numaralı çalışmayı baz alarak farklı özellikleri kullanıp benzer başarı oranı yakalamaya çalıştım.

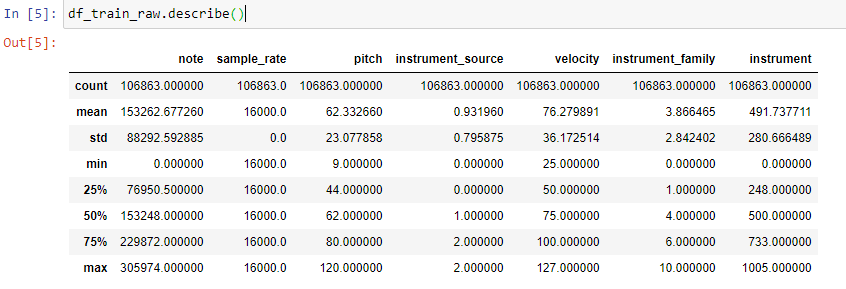
1. **YÖNTEMLER**

Algoritma olarak KNN algoritmasını seçtim. Algoritmayı çalıştırmak için Python dilini Pandas, Numpy ve Scikit-learn kütüphaneleri ile birlikte kullandım. Python versiyonu olarak Python 3.8.5 kullandım. Projeye başlamadan önce dataseti incelemek için Jupyter Notebook kullanıp datasetin içinde neler olduğuna, hangi özelliklerini projede kullanabileceğime baktım.

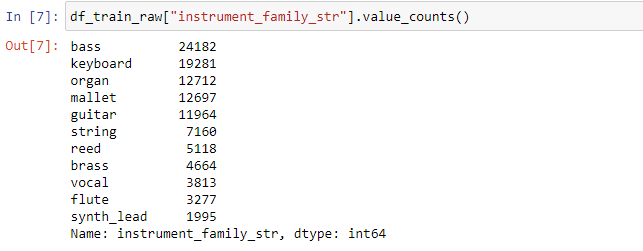
1. **UYGULAMA**

Datasetin Train olarak adlandırılan kısmında 289205 adet 4 saniyelik tek notadan oluşan ses dosyası var ve ayrıca bu dosyalara ait 1 adet JSON dosyası içeriyor. Test kısmında ise aynı düzende 4096 adet dosya var. Bu dosyalar 11 farklı enstrümanın akustik, elektronik ve sentetik versiyonlarını içerebiliyor.

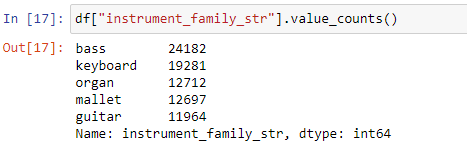
İlk olarak dataseti gereksiz kalabalıktan kurtarmak için kullanmayacağım bilgileri silmem gerekiyordu. Bunun için Jupyter Notebook kullanarak datasetin her bir alanının minimum, maximum, ortamala ve toplam sayılarına baktım. Örnekleme aralığı her dosyada aynı olduğu için örnekleme aralığını çıkarttım.



Sonrasında daha başarılı bir sonuç elde edebilmek için 11 adet enstrümandan sayı olarak yetersiz olduğunu düşündüğüm enstrümanları çıkartmak istedim. Bunun için veri setinin içerisindeki enstrüman adetlerine baktım.



String, reed, brass, vocal, flute ve synth\_lead enstrüman sınıfına ait olan verileri elimdeki veri setinden çıkarttım.



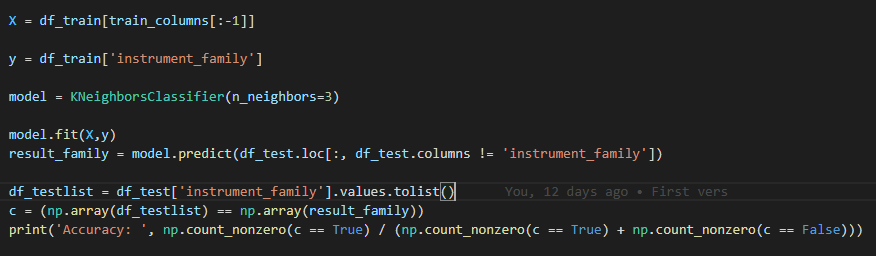
Onun dışında enstrümanın sınıflamaya çalıştığımız sınıf özelliğine dayalı bazı bilgileri de çıkartmam gerekiyordu. Aynı zamanda bazı bilgiler hem numerik hem de isim olarak tutuluyordu. Bunlardan da isim olarak kaydedilmiş olanları çıkarttım.



Sonrasında veri içerisinde gelen Qualities verilerini ayrı ayrı değişkenlere ayırmam gerekiyordu. Bu sayede algoritma içerisinde bu verileri de kullanabildim. Bu sayede algoritmaya verdiğim veriler şu şekilde oldu:

* Nota
* Perde
* Hız
* Frekansların yüksekliği
* Bozukluk
* Hızlı sönme
* Uzun bırakma
* Çok seslilik
* Vuruş
* Yankı
* Tempo uyumu

Daha sonra bu verileri Scikit-learn kütüphanesinin sağladığı KNeighborsClassifier fonksiyonuna çeşitli parametreler ile verdim. Train ettiğim modele test dosyalarını verip karşılaştırma yaptığımda bana yüzde olarak kaçta kaçını doğru bulduğumu gösterdi.



1. **SONUÇLAR**

İlk denemelerimde yüzde 20’lerde başarı sağlarken sonraki farklı argümanlarda yaptığım denemelerde bu oranı en çok yüzde 34.87’ye kadar yükseltebildim. En son olarak 3 komşulu denememde bu oranı yakaladım.



Bundan sonraki süreçte Uygulama başlığında bahsettiğim veri başlıklarından bazılarını çıkartarak daha anlamlı bir veri kümesi oluşturup bu başarı oranını kabul edilebilir bir orana çıkartmayı planlıyorum.

Daha sonraki geliştirme olarak da bir müzik dosyası verdiğim zaman programın bunu dinleyerek içerisindeki enstrümanları saymasını başarmaya çalışacağım.

1. **REFERANSLAR**

[1] Brandi Frisbie, Stanford University, USA

<https://ccrma.stanford.edu/groups/meri/assets/pdf/frisbie2017ISMIR_LBD.pdf>

[2] Nadim Kawwa, Udacity Machine Learning Engineer Nanodegree

<https://medium.com/@nadimkawwa/can-we-guess-musical-instruments-with-machine-learning-afc8790590b8>

1. **EKLER**

Yazdığım kodu github üzerinde bulabilirsiniz.

<https://github.com/mustafasamisahin/InstrumentRecognitionUsingNsynth>