

Multimedia Information Systems

Term Project

Eray Samet Gündüz
Ersan Ergin

Giriş

Günümüzde multimedya içeriğinin patlamasıyla, özellikle müzik gibi ses verilerini verimli ve doğru bir şekilde sınıflandırmak ve anlamak büyük bir zorluk haline gelmiştir.

Müzik, evrensel bir dildir ve kültürel sınırları aşar. Ancak sınıflandırılması önemli bir hesaplama sorununu ortaya çıkarır.



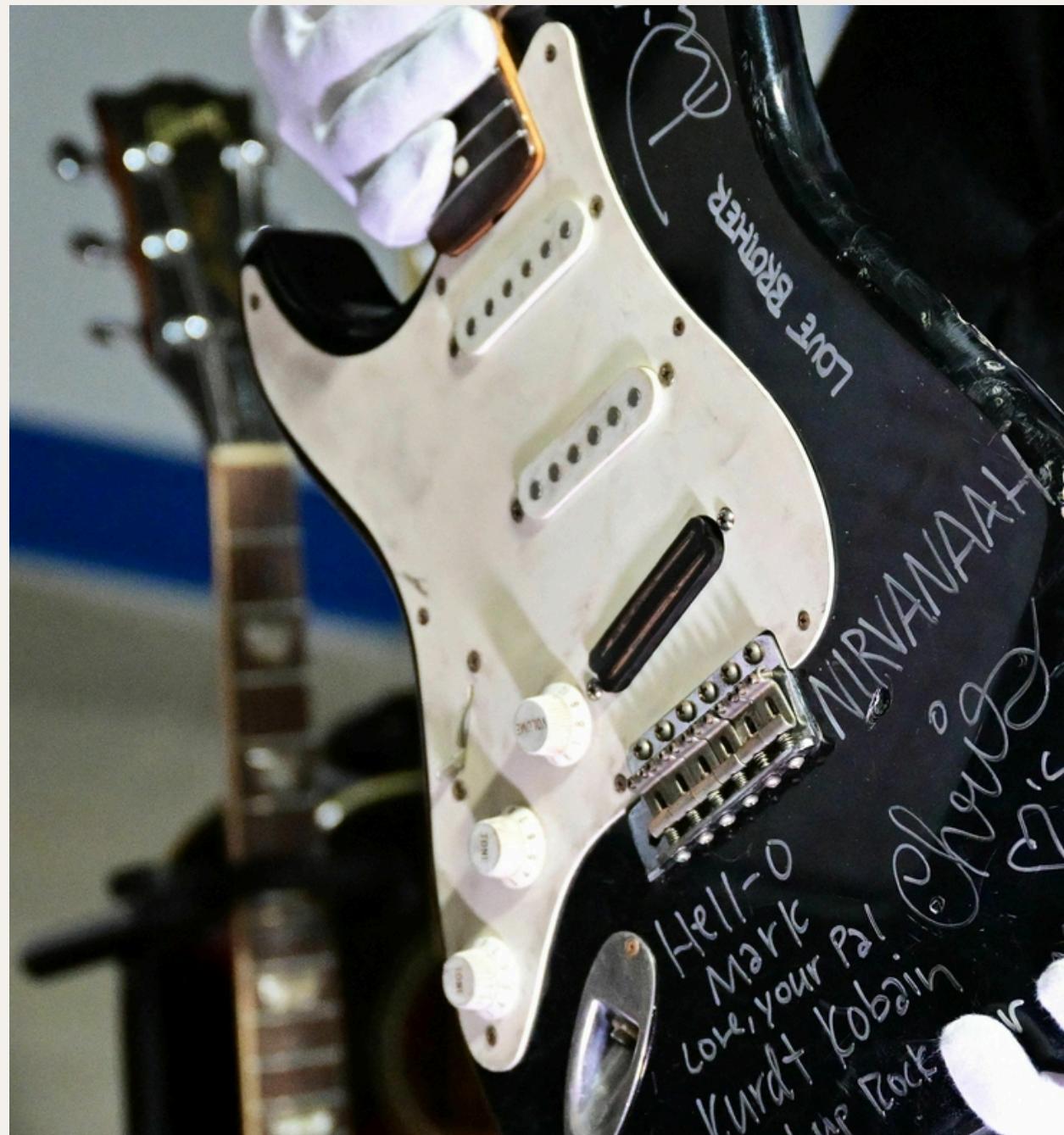
PROBLEM TANIMI

Müzik türlerini ayırt etmek ve kategorize etmek, yalnızca müzik tavsiye sistemleri için değil, aynı zamanda müzik analitiği, dijital arşivleme ve otomatik içerik etiketleme gibi birçok uygulama için de çok önemlidir.

Gelişmiş ses işleme araçlarının mevcut olmasına rağmen, müzikal unsurların nüanslı ve çok yönlü yapısı nedeniyle müzik türlerini doğru bir şekilde sınıflandırmak karmaşık bir iştir.



MOTIVASYON



Bu projenin arkasındaki motivasyon, sesin temel özelliklerini daha iyi anlamak ve müziği görselleştirmek ve anlamak için sağlam yöntemler geliştirmektir.

Bu projeden elde edilen bilgiler, gelişmiş müzik tavsiye motorlarına, müzik akış platformlarında daha iyi kullanıcı deneyimlerine ve ses analizi teknolojilerinde gelişmelere yol açabilir.

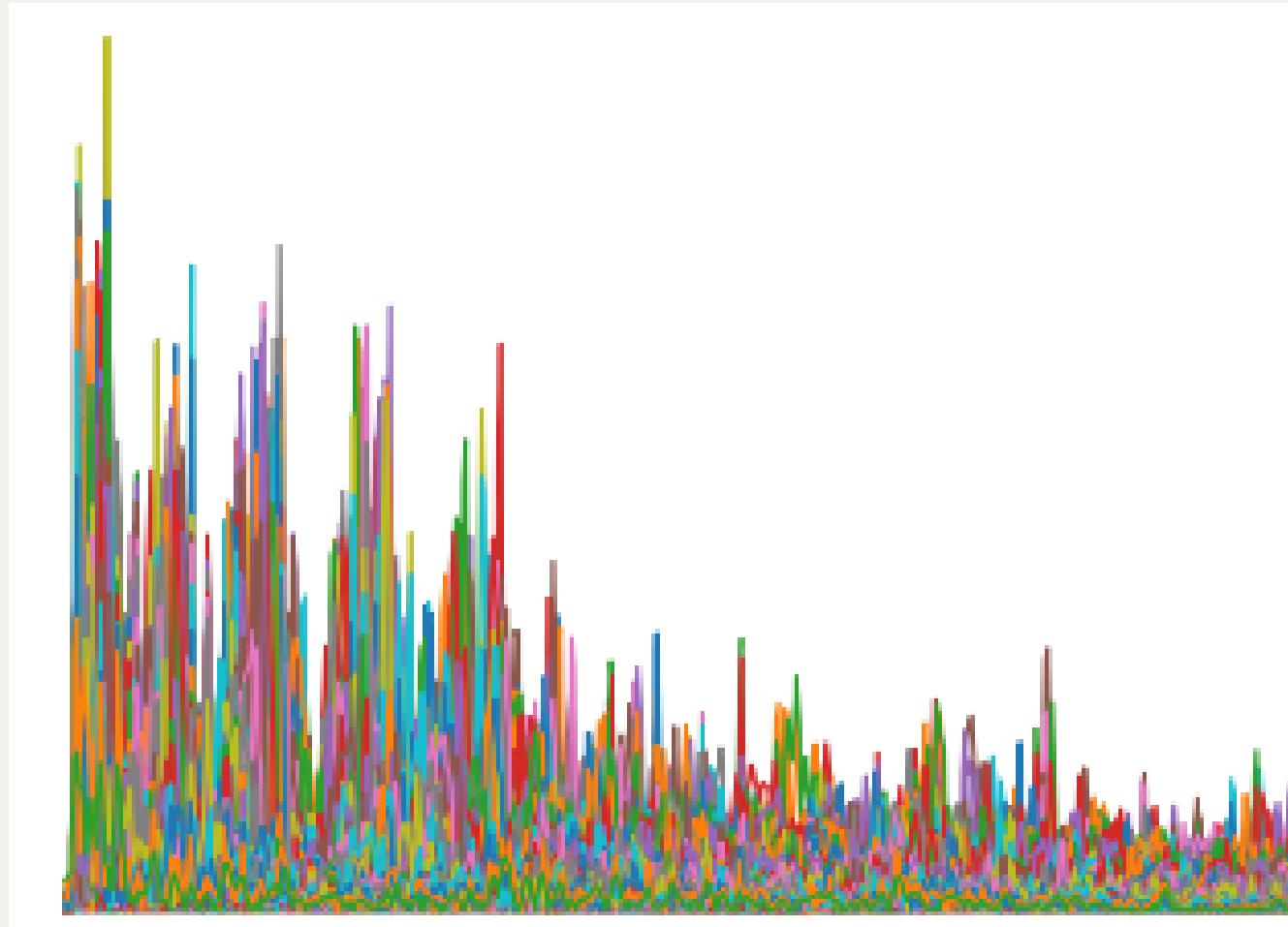
Müzik bilgi topluluğunda iyi bilinen bir kıyaslama olan GTZAN Veri Setinin kullanılması, metodolojilerimizi değerlendirmek için sağlam bir temel sağlar ve bulgularımızın alaka düzeyini ve uygulanabilirliğini sağlar.

YAKLAŞIM

Kapsamlı bir müzik sınıflandırma ve görselleştirme anlayışına ulaşmak için, yerleşik metodolojilerden yararlanan yapılandırılmış bir yaklaşım benimsedik.

- 1.Ses Verisi Yükleme ve Ön İşleme: Librosa ile ses dosyasını yükle, kırp ve normalize et.
- 2.Dalga Formu Görselleştirme (Sonograf): Ses verisini zaman-genlik grafiği olarak görselleştir.
- 3.Kısa Zamanlı Fourier Dönüşümü (STFT): Ses sinyalini frekans ve zaman bileşenlerine ayır.
- 4.Spektrogram Oluşturma: STFT sonucunu frekans-zaman düzleminde görselleştir.
- 5.Mel Spektrogram Oluşturma: Frekans eksenini Mel ölçüğine dönüştürerek spektrogramı insan algısına uyarlayın.
- 6.Kroma Özellikleri Çıkarma: Müzikteki perde sınıflarını (notaları) belirle ve görselleştir

VERİ TOPLAMA VE ÖN İŞLEME



Çeşitli türlerden müzik parçalarından oluşan bir veri seti topladık (GTZAN Veri Seti).

Ses dosyalarını tutarlı bir biçimde (WAV) dönüştürdük ve tüm parçalar için örnekleme hızını standartlaştırdık.

Kayıt koşullarındaki farklılıklardan kaynaklanan tutarsızlıklarını ortadan kaldırmak için ses seviyelerini ortak bir ses düzeyine normalleştirdik.

SESİ ANLAMA VE ÖZELLİK ÇIKARMA

Librosa (Python paketi) ile sesleri analiz ettik ve özellik çıkarma yöntemlerini kullandık:

Fourier Dönüşümü: Zaman alanından frekans alanına dönüşüm.

Spektrogram: Frekans spektrumunun zamanla değişiminin görsel temsili.

Mel Frekanslı Kepstral Katsayılar (MFCC'ler): Sesin spektral özelliklerinin kompakt temsili.

Kroma Frekansları: Müzikteki on iki farklı perde sınıfının temsili.

Spektral Kontrast: Ses spektrumundaki tepe noktaları ve çukurlar arasındaki genlik farkı.

