

**T.C.**

**BEYKOZ ÜNİVERSİTESİ**

**MÜHENDİSLİK VE MİMARLIK FAKÜLTESİ**

**YAZILIM MÜHENDİSLİĞİ BÖLÜMÜ**

**DUYGUNU SESLENDİR (TÜRKÇE)**

**Kerem ERGÜNER**

**YAPAYZEKA UYGULAMALARI PROJESİ**

**GİRİŞ**

**DUYGUNU SESLENDİR**

Bu proje, Türkçe konuşulan ses dosyalarından duygu tanıma işlemi gerçekleştirmeyi amaçlamaktadır. Projede kullanılan veri kümesi, veri ön işleme adımları, modelleme teknikleri ve sonuçlar detaylandırılmıştır.

1. **LİTERATÜR ARAŞTIRMASI**

1.1 Algısal Öznitelikler Kullanarak Sesten Otomatik Duygu Durum Tanıma

**Özet**: Bu çalışma, ses verilerinden duygu durumlarını tespit etmek için algısal öznitelikler ve çeşitli sınıflandırma yöntemleri kullanmaktadır. Destek Vektör Makinası, Gauss Karışım Modelleri ve Öğrenimli Vektör Nicemleme algoritmaları, literatürdeki başarımı artırmayı hedeflemektedir.

* Kaynak: [ITU Digital Repository](<https://polen.itu.edu.tr/items/22f25ee2-1892-4ce8-a19f-c8aeb4aaec12>)

1.2 Ses Sinyallerinden Duygu Tanıma İçin Farklı Yaklaşımlar

**Özet:** Bu tez, Berlin Duygu Veri Tabanı kullanarak çeşitli özellik mühendisliği tekniklerini ve sınıflandırma algoritmalarını değerlendirmektedir. Yapay Sinir Ağları, Destek Vektör Makineleri ve k-Nearest Neighbors gibi algoritmalar kullanılarak MFCC ve LPC özellikleriyle sınıflandırma işlemleri yapılmıştır.

* Kaynak: [Konya Technical University Repository [(https://gcris.ktun.edu.tr/handle/20.500.13091/427)](https://gcris.ktun.edu.tr/handle/20.500.13091/427))

1.3 MFCC Özelliklerinin Konuşma Duygu Tanıma Sistemlerinde Performans Analizi

- \*\*Özet:\*\* Bu makale, MFCC özelliklerinin konuşma duygu tanıma sistemlerinde etkinliğini analiz etmektedir. Yapay sinir ağları ve diğer sınıflandırma teknikleri kullanılarak, MFCC özelliklerinin konuşma duygu tanıma üzerindeki etkisi değerlendirilmiştir.

-Kaynak: [ResearchGate([https://d1wqtxts1xzle7.cloudfront.net/38365626/67JSTR\_2015\_183-libre.pdf?1438592156=&response-content-O](https://d1wqtxts1xzle7.cloudfront.net/38365626/67JSTR_2015_183-libre.pdf?1438592156=&response-content-disposition=inline%3B+filename%3DPerformance_Analysis_of_MFCC_Features_On.pdf&Expires=1717336370&Signature=ExIEWIJ7kHQua4lCp2NiyjNZo-4lru3sTypglUtODATcZBTGRusVAsiGrBVh0cwyGaiJ2~rRnS9UAkZ4ZqTNSvDFGrk-M0lazJG3M5Tzn1bd0Xpd5ytclC4F3zoJPdsEwG7UaHsMNVNqAxOwtXejZoehwrHbBeIni2rc8rXF5hxCWl-CitBVY6dXA4Se1rp02k7sbynMvfvSX27LUompX6t3r7HtzdCyOJVK~i5wO5oDIDsA4DdGqVLP5TlD3WJdqHx5EjeRVXiMeiwCo9Hrro8JZ~sb1ogToTJrCzO9zlNBBAQpAEGOqjeQxK9u5DHMbdvQxXIEI0ATANi9xm2nXQ__) )

1. Veri kümesi ve Veri Özellikleri

2.1 Veri Kümesi: Turkish Emotion-Voice Database TurEV-DB

Turkish Emotion-Voice Database (TurEV-DB), Türkçe konuşulan seslerin duygu tanıma çalışmalarında kullanılmak üzere oluşturulmuş bir veri kümesidir. Bu veri kümesi, çeşitli duyguları içeren ses kayıtlarından oluşur ve farklı duygusal durumları yansıtan ses dosyalarını içerir.

Veri Özellikleri:

• Duygular: Mutlu, Üzgün, Kızgın ve Sakin.

• Ses Dosyaları: .wav formatında ses dosyaları.

• Örnekleme Hızı: 44100 Hz.

• Özellikler: Ses spektrum resimleri, csv dosyaları ve ses dosyaları.

1. Veri Ön İşleme, Özellik Mühendisliği

3.1 Veri ön işleme:

Dosya içerisinde bulunan csv dosyaları, spektrum dosyaları ve model output for test data klasörleri çıkarılmıştır. Projeye ses dosyalarından MFCC ses özellikleri çıkartılıp eğitilen model.h5 dosyasıyla devam edilmiştir.

3.2 Özellik Mühendisliği:

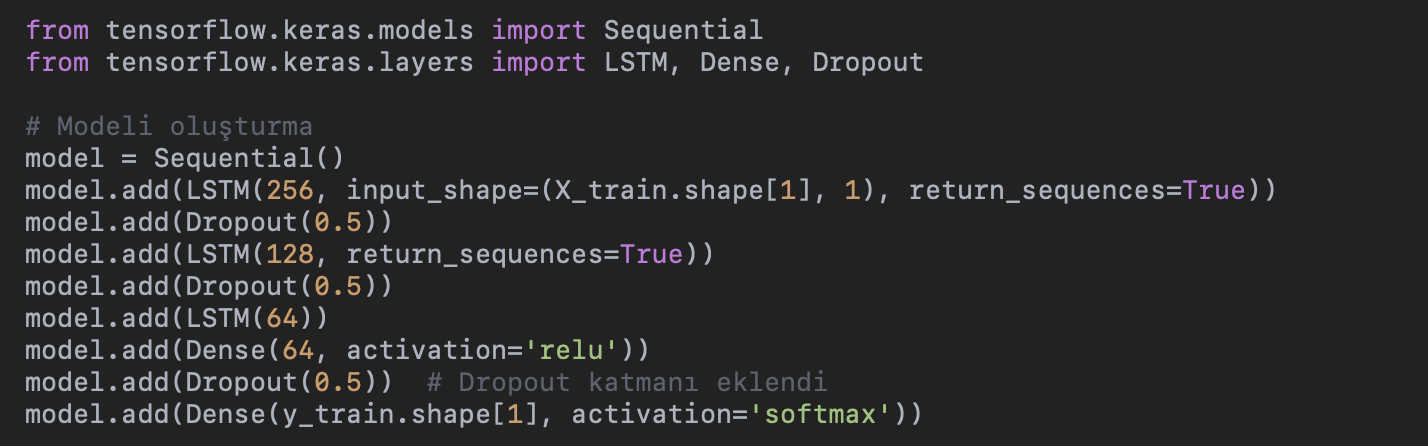
MFCC, ses sinyallerinin kısa süreli güç spektrumu temsilidir ve konuşma duygu tanıma için yaygın olarak kullanılır. Her bir ses dosyası için 13 MFCC katsayısı çıkarılır ve bu katsayıların ortalaması alınarak özellik vektörü oluşturulur. Bu özellikler model.h5 dosyasına eklenip tahmin ettirmek için kullanılır.

1. Modelleme ve Test Aşamaları

**Modelleme**

Modelimiz, zaman serisi verilerindeki ardışık bağımlılıkları modellemek için kullanılan Long Short-Term Memory (LSTM) katmanlarından oluşmaktadır. Aşağıda modelin oluşturulması ve eğitilmesi için kullanılan kod parçacıkları ve açıklamaları bulunmaktadır.

**Model Kodu:**

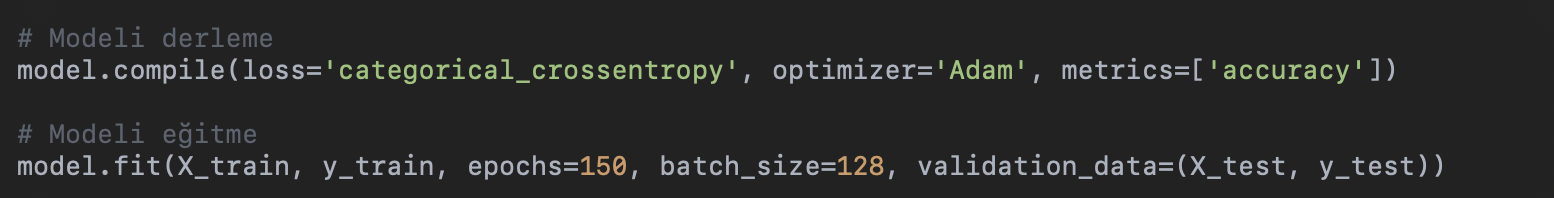


Bu model, ardışık LSTM katmanları ve Dropout katmanlarından oluşur. LSTM katmanları, zaman serisi verilerindeki bağımlılıkları modellemek için kullanılırken, Dropout katmanları aşırı öğrenmeyi önlemek için eklenmiştir.

**Model Eğitimi**

Modeli eğitmek için fit() fonksiyonu kullanılmıştır. Eğitim ve doğrulama verileri kullanılarak model eğitilmiştir.

**Model Eğitimi Kodu:**

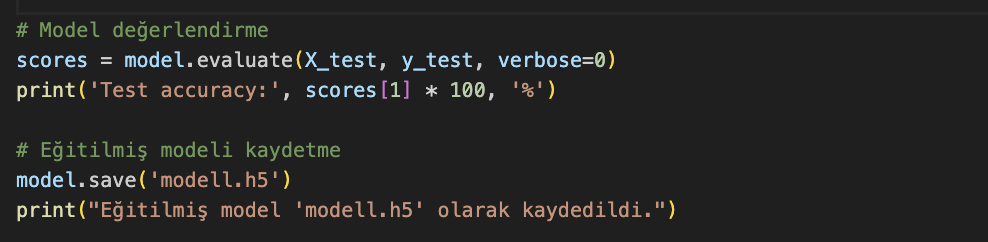


Model categorical\_crossentropy kaybı ve Adam optimizasyon algoritması ile derlenmiş ve 150 epoch boyunca eğitilmiştir. Eğitim sırasında, modelin doğrulama verisi üzerindeki performansı da izlenmiştir.

**Test Aşaması**

Model, test verisi ile değerlendirilmiş ve doğruluk (accuracy) metriği hesaplanmıştır.

**Model Test Kod Parçacığı:**

Model, test verisi ile değerlendirilmiş ve %76 doğruluk elde edilmiştir. Eğitilmiş model, daha sonra kullanılmak üzere ‘modeldeneme123.h5’ dosyasına kaydedilmiştir.

**Kaynaklar, Ortam, API, Kitaplık ve Teknoloji Yığınları**

**Kaynaklar ve Ortam**

**Gerekli Kaynaklar:**

• Python 3.8 veya üstü

• Ses dosyaları (.wav formatında)

• Eğitim ve test verileri (CSV dosyaları ve spektrum görüntüleri)

• Eğitilmiş model dosyası (modeldeneme123.h5)

**Ortam Bilgileri:**

• İşletim Sistemi: Windows 10 / macOS

• Python Virtual Environment veya Conda Environment

**API ve Kitaplıklar**

**Kullanılan API’ler:**

• Flask: Web uygulaması geliştirmek için kullanılır.

• SoundDevice: Mikrofon verisi kaydetmek için kullanılır.

• Librosa: Ses dosyalarını işlemek ve MFCC özelliklerini çıkarmak için kullanılır.

• TensorFlow/Keras: Model oluşturmak, eğitmek ve tahmin etmek için kullanılır.

• NumPy: Sayısal hesaplamalar ve veri işlemleri için kullanılır.

**Gerekli Kitaplıklar ve Kurulum:**

pip install Flask sounddevice librosa tensorflow numpy

**Teknoloji Yığınları**

**Backend:**

• Flask: Web framework

• TensorFlow/Keras: Derin öğrenme kütüphanesi

**Frontend:**

• HTML/CSS: Web arayüzü tasarımı

• JavaScript: Dinamik içerik ve kullanıcı etkileşimleri

**Dağıtım Talimatları**

**Adım 1: Gerekli Dosyaları ve Ortamı Hazırlayın**

1. Python 3.8 veya üstü bir sürümün yüklü olduğundan emin olun.

2. Proje dosyalarını uygun bir klasöre indirin veya klonlayın.

3. Sanal bir ortam oluşturun ve etkinleştirin:

python -m venv env

source env/bin/activate # macOS/Linux

env\Scripts\activate # Windows

4. Gerekli kitaplıkları yükleyin:

pip install Flask sounddevice librosa tensorflow numpy

Uygulamayı Çalıştırma:

a. Github linkini clone alın ve kütüphaneleri yükleyin.

b. Eğitilmiş model dosyasını (modeldeneme123.h5) uygun bir konuma yerleştirin.

1. Web uygulamasını çalıştırmak için app.py dosyasını çalıştırın:

python app.py

2. Tarayıcınızda <http://127.0.0.1:5000/> adresine gidin ve web arayüzünü kullanarak ses kaydedip tahmin sonuçlarını görün.

Github Linki:

git clone <https://github.com/keremerguner/duygunu-seslendir>