

Urban Sounds Classification

Koç Holding Derin Öğrenme Bootcamp

Bu projede “UrbanSounds8K” veri setini kullanarak şehirlerde duyulan seslerin sınıflandırılmasıyla ilgili bir proje yapacağız. Bu projede hedefimiz, derin öğrenme algoritmalarını kullanarak yapay zeka modeli hazırlamaktır. Veri setini indirecek, ses dosyalarından spectrogram oluşturacak ve bunları ait oldukları kategorilere göre bilgisayara kaydedeceğiz. Daha sonra bu görüntülere önişleme yaparak CNN modeli hazırlayacak ve bu modeli eğiteceğiz. Değerlendirme aşamasında model performansı bir ölçüt olmayacaktır. Önemli olan modeli hazırlamak, eğitmek ve hiperparametre optimizasyonu yapmaktır 😊

Kullanabileceğimiz kütüphanelerin dokümanları:

- Librosa (Ses İşleme): <https://librosa.org/doc/latest/index.html>
- OpenCV (Görüntü İşleme): <https://librosa.org/doc/latest/index.html>
- TensorFlow/Keras (Derin Öğrenme): https://www.tensorflow.org/api_docs/python/tf

Projelerin Değerlendirilmesi:

- Değerlendirme sadece “Önişleme” ve “Model Hazırlanması ve Eğitilmesi” notebookları üzerinden olacaktır! Sizden 2 farklı notebook beklenmektedir. Spectrogram oluşturma kısmını isteyen arkadaşlarım yapmayarak, hazırlanmış spectrogramları aşağıdaki linkten indirerek çalışmalarına bunun üzerinden devam edebilirler.

Spectrograms: https://drive.google.com/drive/folders/1xey3vAVNDjWxnSfhuUsf_5dGANZWmCog?usp=sharing

- İsteddiğiniz çalışma ortamını kullanabilirsiniz. Fakat sonuçları bir *Notebook* dosyası olarak yüklemenizi bekleyeceğiz. Planlamanızı lütfen buna göre yapın 😊
- Derin öğrenme eğitimleri sırasında kullanacağımız GPU elinizde yoksa, Google Colab tarafından ücretsiz olarak sağlananları kullanabilirsiniz.

(<https://colab.research.google.com>)

0. Veri Setinin İncelenmesi

- a. “UrbanSounds8K” veri setini bilgisayarınıza indirin. (5,6 GB)
<https://urbansounddataset.weebly.com/urbansound8k.html>
- b. İndirdiğiniz .tar dosyasını açın. (Mentorlarınızdan destek alabilirsiniz 😊)
- c. *README* dosyasını okuyarak elimizdeki verinin ne olduğunu ve nasıl isimlendirildiğini iyice anlayın.

1. Spectrogram Oluşturma - Notebook

- a. Ses dosyalarının spectrogramlarını oluşturun. Kolaylık olması için aşağıdaki fonksiyonu kullanabilirsiniz. **create_spectrogram** fonksiyonuna, okuduğunuz ses dosyasını gönderdiğinizde size hazırlanmış spectrogramı dönecektir.

```
def create_spectrogram(y):  
    spec = librosa.feature.melspectrogram(y=y)  
    spec_conv = librosa.amplitude_to_db(spec, ref=np.max)  
    return spec_conv
```

- b. Oluşturduğunuz spectrogramları ait olduğu sınıfın klasörüne kaydedin.

2. Önileme - Notebook

- a. Görüntüleri (spectrogramları) sırasıyla okuyarak, *grayscale* dönüşümü, *resizing* ve normalizasyon yapın.
- b. Görüntüleri ait oldukları etiketlerle birlikte, *[görüntü, etiket]* formatında bir listeye ekleyin.
- c. Bu listeyi kullanarak, *X_train*, *y_train*, *X_val*, *y_val*, *X_test* ve *y_test* veri setlerini oluşturun.
- d. Bu veri setlerini bilgisayarınıza kaydedin.

3. Model Hazırlanması ve Eğitimi - Notebook

- a. Bir CNN modeli hazırlayın.
- b. Modeli hazırlamış olduğunuz veriyi kullanarak eğitin.
- c. Modelin performans metriklerini, loss ve accuracy grafiklerini ekrana yazdırın.
- d. Aldığınız sonuçlara göre *hyperparameter optimization* yapın.

4. Sonuçların Yüklenmesi

- a. Hazırladığınız notebookları, çıktı hücrelerini **silmeden** GitHub hesabınıza yükleyin.
- b. <https://forms.gle/tk9m8cRXTkLV2Phv8> linkinde bulunan formu doldurun.
- c. Dosyaları **herkes görüntüleyebilir** şeklinde ayarlamayı unutmayın 😊