



Akbank Derin Öğrenme Bootcamp: Yeni Nesil Proje Kampı

Bu proje kapsamında, temel özellikleri kullanarak bir Derin Öğrenme projesi geliştirmeniz beklenecektir.

Katılımcılara, Derin Öğrenme alanında başlangıç düzeyinde sınıflandırma, veri analizi, model geliştirme ve değerlendirme teknikleri konusunda pratik deneyim kazandırılması amaçlanmaktadır.

Proje Konusu & Veri Seti: <https://www.kaggle.com/datasets/crowww/a-large-scale-fish-dataset/data>

Bu veri seti ile, bir sınıflandırma projesi geliştirmeniz beklenecektir.

1 - Geliştirme Ortamı (Kaggle&GitHub&README.md)

Projenizi, bu veri setinden oluşturacağınız yeni notebook üzerinde ("New Notebook" butonuna tıklayarak) Kaggle ortamında geliştirmeniz gerekmektedir.

Not 1: Veri seti boyutu 3GB gibi görece büyük bir boyutta olduğu için download ederek lokalinizde çalışmamanız önerilir.

Not 2: Dileyen katılımcılar projelerini Colab ortamında(veya arzu ettiğiniz başka bir ortamda) da geliştirebilir ancak projenin finalinde Kaggle'da yer almasını istiyor olacağız. Bundan dolayı tüm geliştirme süreçlerini Kaggle üzerinde yürütmeniz önerilir.

* Projenize ait teknik anlatımlara notebook'unuzun içerisinde **markdown** formatında hücrelerle yer vermeniz beklenmektedir.

* GitHub reponuzda yer alacak olan **README.md** dosyasında ise özet bilgilerle projenizden bahsetmelisiniz.

* README.md dosyanızın sonunda Kaggle'daki notebook linkinize yer vermelisiniz. Hücrelerinizin çıktıları zaten görünüyor olacak ve gerekli durumlarda Kaggle üzerinde projenizi çalıştırabilir olacağız, bundan dolayı **veri setinizi reponuza eklemenize gerek yok.**

2 - Veri Önışleme

Veri setiniz, klasörlerin içinde yer alan .png formatında fotoğraflardan oluşmaktadır. Bundan dolayı ML uygulamalarında yaptığımız gibi *pd.read_csv* ile okuyamıyoruz. Öncesinde bu veriyi düzenlemeniz ve bir Pandas DataFrame oluşturmanız gerekmektedir. Aşağıdaki kod parçası buna bir örnektir:

```
label = []
path = []
fish_dir = '/kaggle/input/a-large-scale-fish-dataset/Fish_Dataset/Fish_Dataset'
for dir_name, _, filenames in os.walk(fish_dir):
    for filename in filenames:
        if os.path.splitext(filename)[-1]=='.png':           # If filename contains .png
            if dir_name.split()[-1]!='GT':                   # If directory doesn't contain GT
                label.append(os.path.split(dir_name)[-1])     # Append the directory name to label
                path.append(os.path.join(dir_name,filename))  # Append all the png files to path of that directory

data = pd.DataFrame(columns=['path', 'label'])
data['path']=path
data['label']=label
```

Bu adımın ardından **data** değişkeniniz veri setiniz olarak hazırdır. En iyi projeyi çıkarabilmek için, önışleme aşamasını **pandas** komutlarıyla veri setini inceleyerek, veya **matplotlib** kütüphanesiyle veri setinizdeki fotoğraf verisini görselleştirerek zenginleştirebilirsiniz.

Not: Fikir edinmek için veri setinin Kaggle sayfasındaki “Code” sekmesinde yer alan notebookları inceleyebilirsiniz.

Bu adımın ardından verinizi uygun şekilde train, test olarak ayırıp train edilmeye hazırlayarak bu bölümü bitirmeniz beklenmektedir.

3 - Modelin Eğitilmesi

ANN mimarisi kullanarak bir model oluşturacaksınız, Eğitim platformunda ilgili başlıktaki videoyu izleyip notebook’u inceleyerek fikir edinebilirsiniz. Eğitim yayınında daha detaylı değinilecektir.

Modelinizde katmanlarınız, aktivasyon fonksiyonlarınız, dropout katmanınız vs. gibi model performansınızı iyileştirecek detaylara yer verebilirsiniz.

4 - Modelin Değerlendirilmesi

Bu adımda modelin başarı metriklerine yer vermelisiniz. Accuracy hesabı, kayıp fonksiyonu gibi ölçümler kullanabilirsiniz.

Daha iyi bir proje için:

*** Kayıp fonksiyonunuzun iterasyonlar arası grafiğini çizdirebilir, confusion matrix ve classification report oluşturabilirsiniz.**

5 - Hiperparametre Optimizasyonu

- Katman sayısı,
- Düğüm sayısı,
- Dropout rate,
- Optimizer

Gibi parametrelerle oynayarak hiper parametre optimizasyonu yapmanız ve mümkün olan en iyi performans gösteren modeli oluşturmanız beklenmektedir.

Not: Model performansı en iyi projenin belirlenmesinde önemlidir. Modelinizin overfit etmediğine emin olmanız, kod ve markdown hücrelerinizde bunu göstermeniz ve anlatmanız gerekmektedir.