

Kocaeli Üniversitesi
Bilgisayar Mühendisliği Bölümü
2024-2025 Bahar Dönemi
Yazılım Lab. II
Proje III

Proje Başlangıç Tarihi.....4 Mayıs 2025
Proje Bitiş Tarihi.....25 Mayıs 2025

1. Proje Tanımı

Bu projede, sizlere Google Drive üzerinden sunulacak olan **MultiZoo Görüntü Veri Seti** kullanılacaktır. Bu veri seti, farklı hayvan türlerine ait görüntülerden oluşmaktadır ve çoklu sınıflandırma problemi için hazırlanmıştır. Proje kapsamında, verilen eğitim verisiyle bir model eğitmeniz ve verilen ayrı test verisiyle modelinizi değerlendirmeniz beklenmektedir.

2. Proje Hedefleri

- **Çoklu Sınıflandırma:** Bir görüntüdeki hayvan türünü doğru şekilde tahmin eden bir model geliştirmek. Geliştireceğiniz model transformer tabanlı olmalıdır.
- **Başarı Kriteri:** Test ve doğrulama setinde en az %65 doğruluk skoru (accuracy) elde edilmelidir. 90 üstü ya da en iyi sonuçlar için ayrıca bonus değerlendirmesi yapılacaktır.
- **Değerlendirme Metrikleri:** Doğruluk (Accuracy), kesinlik (Precision), duyarlılık (Recall) ve F1-skoru (F1-score) hesaplanmalıdır.
- **Arayüz:** Eğitimi tamamlayıp kaydettiğiniz modelinizi canlı test edebileceğiniz, vereceğimiz test görüntülerini alıp tahmin döndürebilen bir arayüz geliştirilmelidir.
- **Rapor:** Tüm süreç ve sonuçlar istelere uygun bir şekilde detaylı bir raporla sunulmalıdır.

3. Proje Adımları

Kullanılacak Veri Seti:

- Veri setine aşağıdaki link üzerinden ulaşabilirsiniz:
https://drive.google.com/drive/folders/1uXMJa2CX5eg_kGWghbqg8TWkpeQi-0VI?usp=sharing
- Veri setinden bazı örnek görüntüler;



- Veri seti, eğitim (train) ve test olmak üzere iki ana klasör içermektedir:
 - **Eğitim Seti:** Farklı hayvan türlerine ait görüntülerden oluşur. Modelinizi bu veriyle eğitmelisiniz. Ayrıca, kendi doğrulama (validation) setinizi bu eğitim setinden ayırmanız gerekmektedir.
 - **Test Seti:** Eğitim setinden tamamen bağımsız olarak, tarafımızdan belirlenecek ve ayrı bir klasörde sunulacaktır. Modelinizin nihai başarısı bu test seti üzerinden sunum sırasında ölçülecektir. Projelerinizi teslim ettikten sonra test verisi, veri setinin linki üzerinden erişilebilir olacaktır.
- Size verilen eğitim ve test klasörlerini inceleyiniz. Eğitim setinden kendi doğrulama (validation) setinizi ayırınız (ör. %80 eğitim, %20 doğrulama).

Ön İşleme Adımları

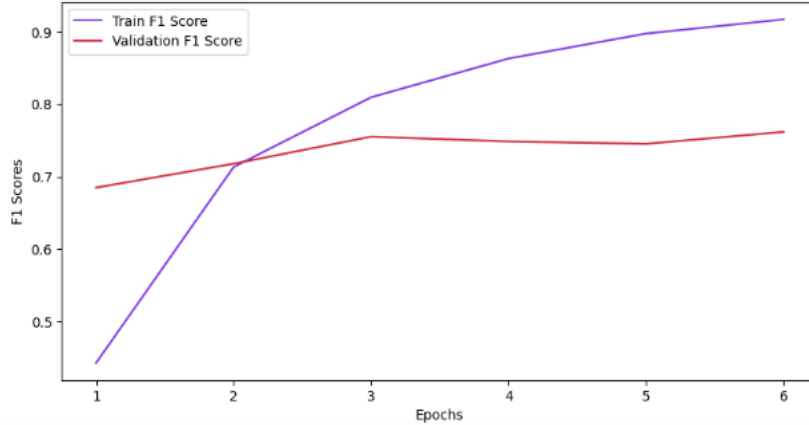
- Modelin başarısını artırmak için veri setine uygun ön işlemleri gerçekleştirebilirsiniz. Uygulanabilecek ön işlemler;
 - Boyutlandırma,
 - Normalizasyon,
 - Histogram eşitleme, filtreleme, gürültü giderme gibi çeşitli görüntü işleme teknikleri,
 - Veri çoğaltma (Data Augmentation)
 - i. yatay çevirme (horizontal flip), rastgele kırpma (random crop), renk jitterlama (color jitter) gibi teknikler uygulanabilir.

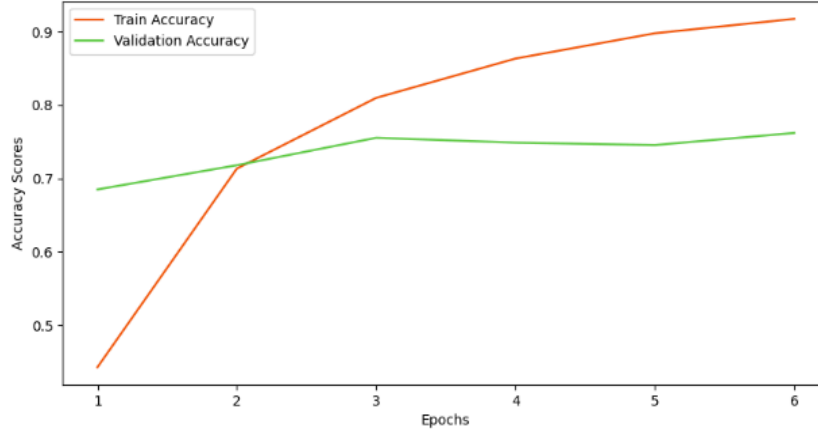
Model Eğitimi

- Model eğitimi Python veya Matlab ile gerçekleştirilebilir. Python kullanacaklar Torch veya Tensorflow kütüphanelerini kullanabilirler.
- Model eğitimini kendi bilgisayarınızda veya Google Colab, Kaggle gibi online GPU kaynakları ile gerçekleştirebilirsiniz.
- Tasarladığınız model transformer tabanlı bir model olmalıdır (Vision Transformer (ViT), Data-efficient Image Transformer (DeiT), Swin Transformer, BEiT gibi modeller bu amaç için uygundur).
- Batch size, epoch, dropout oranı, öğrenme oranı (learning rate) gibi hiperparametrelerin seçimini modelinize uygun olacak şekilde kendiniz belirlemelisiniz. Bu hiperparametrelerin kullanım amacını sunum sırasında açıklayabilmeniz beklenmektedir.
- Overfitting'i önlemek için EarlyStopping gibi yöntemler kullanmanız önerilir.
- Eğittiğiniz modeli geliştireceğiniz arayüzde kullanabilmek için kaydetmelisiniz.
- Test verisi eğitim sürecinde hiçbir şekilde kullanılmamalıdır. Test seti yalnızca modelin son performansını değerlendirmek için kullanılacaktır.

Modelin Değerlendirilmesi

- Modelinizin başarısını değerlendirebilmek için Doğruluk (Accuracy), kesinlik (Precision), duyarlılık (Recall) ve F1-skoru (F1-score) metriklerini kullanmalısınız.
- Sonuçları öğrenme eğrisi (learning curve) ile göstermelisiniz. Örnek öğrenme eğrilerini aşağıda görebilirsiniz:





- Eğitim, doğrulama ve test verileri için elde ettiğiniz metrikleri raporda bir tablo halinde sunmanız beklenmektedir.

Modelin Arayüzle Kullanılması

- Eğitimi tamamlayıp kaydettiğiniz modeli geliştireceğiniz arayüzde kullanmalısınız.
- Arayüz masaüstü ortamında geliştirilmelidir.
- Sunum esnasında test verilerini indirerek kullanmanız istenecektir.
- Kullanıcı, arayüz ile test görüntüsünü girdi olarak verebilmelidir. Arayüzde kullanıcı bir görüntü yüklediğinde, model bu görüntüyü işlemeli, tahmin ettiği hayvan türünü ve güven skorunu (tahmin edilen hayvan türü için doğruluk oranı) ekranda göstermelidir.
- Arayüzde veri setinde olmayan resimler test edilecek ve doğru sonuç vermesi puanı etkileyecektir.
- Arayüz için programlama dili kısıtı bulunmamaktadır.

Proje Teslimi:

- Rapor IEEE formatında (önceki yıllarda verilen formatta) 4 sayfa, akış diyagramı veya yalancı kod içeren, özet, giriş, yöntem, deneysel sonuçlar, sonuç ve kaynakça bölümünden oluşmalıdır. Raporda kullanılan model yapısı açıklanmalı, elde edilen metrikler tablolar halinde sunulmalıdır.
- Dersin takibi projenin teslimi dahil edestek2.kocaeli.edu.tr sistemi üzerinden yapılacaktır. edestek2.kocaeli.edu.tr sitesinde belirtilen tarihten sonra getirilen projeler kabul edilmeyecektir.
- Proje ile ilgili sorular edestek2.kocaeli.edu.tr sitesindeki forum üzerinden Arş. Gör. Dr. Ayşe Gül Eker veya Arş. Gör. Gamze Korkmaz Erdem'e sorulabilir.
- Demo sırasında kullandığınız modelin yapısı, hiper parametrelerin kullanım amacı, arayüz geliştirme ortamı, geliştirdiğiniz kodun çeşitli kısımlarının ne amaçla yazıldığı gibi çeşitli konular hakkında sorular sorulabilir. Kullandığınız herhangi bir satır kodu açıklamanız istenebilir.