

YÜKSEK DÜZEY PROGRAMLAMA PROJESİ

Talha TÜRK

202113171005

Bilgisayar Mühendisliği / Normal Öğretim

Digit Recognizer Projesi

Digit Recognizer Proje Raporu

Bu rapor, **Lojistik Regresyon** ve **Convolutional Neural Network (CNN)** modellerinin aynı veri seti üzerinde performanslarını inceleyerek ve karşılaştırarak model sonuçlarını değerlendirmeyi amaçlamaktadır.

1) Veri Seti ve İşleme

- Eğitim ve test veri setlerinden eğitim veri seti, train.csv dosyasından yüklenmiştir.
- Görsellerin boyutu 28x28 piksel olup gri tonlamalıdır.
- Veri setindeki etiketler 0'dan 9'a kadar olan rakamları temsil etmektedir.
- Görsellerin pikselleri 0-255 aralığında olduğu için 0-1 arasına normalize edilmiştir.

2) Modeller

Lojistik Regresyon

- Veriler düzleştirilmiş ve Lojistik Regresyon algoritması ile eğitim gerçekleştirilmiştir.

```
label pixel0 pixel1 pixel2 pixel3 pixel4 ... pixel778 pixel779 pixel780 pixel781 pixel782 pixel783
0      1      0      0      0      0      0 ...      0      0      0      0      0      0
1      0      0      0      0      0      0 ...      0      0      0      0      0      0
2      1      0      0      0      0      0 ...      0      0      0      0      0      0
3      4      0      0      0      0      0 ...      0      0      0      0      0      0
4      0      0      0      0      0      0 ...      0      0      0      0      0      0
```

[5 rows x 785 columns]

CNN

- Model, TensorFlow ile oluşturulmuş bir CNN modelidir.
- Flatten Layer: Görselleri düzleştirir.
- Dense Layer: 128 nöronlu, ReLU aktivasyon fonksiyonlu tam bağlantılı katmandır.
- Output Layer: 10 sınıf için softmax aktivasyon fonksiyonlu çıkış katmanıdır.

3) Model Performansı

Lojistik Regresyon

- Eğitim doğruluğu: %[logreg_accuracy]

Lojistik Regresyon Doğruluğu: 94.27%

- Yanlış tahmin edilen örneklerin görselleştirilmesi, yanlış tahmin edilen her rakamdan 3 örnek ile sağlanmıştır.

CNN

- Eğitim doğruluğu: %[cnn_accuracy]

←[1m1313/1313←[0m ←[32m—————←[0m←[37m←[0m ←[1m2s←[0m 1ms/step
CNN Doğruluğu: 98.14%

- Eğitim sırasında 5 epoch boyunca model eğitilmiştir.

```
Epoch 1/5
1/5 [1m52s/52s] 0m 32s 0m 37m 0m 1m2s 0m 3ms/step - accuracy: 0.8256 - loss: 0.6306 - val_accuracy: 0.9388 - val_loss: 0.2114
Epoch 2/5
1/5 [1m52s/52s] 0m 32m 0m 37m 0m 1m1s 0m 2ms/step - accuracy: 0.9494 - loss: 0.1862 - val_accuracy: 0.9538 - val_loss: 0.1510
Epoch 3/5
1/5 [1m52s/52s] 0m 32m 0m 37m 0m 1m1s 0m 2ms/step - accuracy: 0.9626 - loss: 0.1266 - val_accuracy: 0.9611 - val_loss: 0.1272
Epoch 4/5
1/5 [1m52s/52s] 0m 32m 0m 37m 0m 1m1s 0m 2ms/step - accuracy: 0.9736 - loss: 0.0948 - val_accuracy: 0.9663 - val_loss: 0.1172
Epoch 5/5
1/5 [1m52s/52s] 0m 32m 0m 37m 0m 1m1s 0m 2ms/step - accuracy: 0.9775 - loss: 0.0803 - val_accuracy: 0.9686 - val_loss: 0.1045
```

- Eğitim/doğrulama doğrulukları ve kayıpları sunulmuştur.
- Yanlış tahmin edilen örnekler, her rakamdan 3 örnek olmak üzere görselleştirilmiştir.

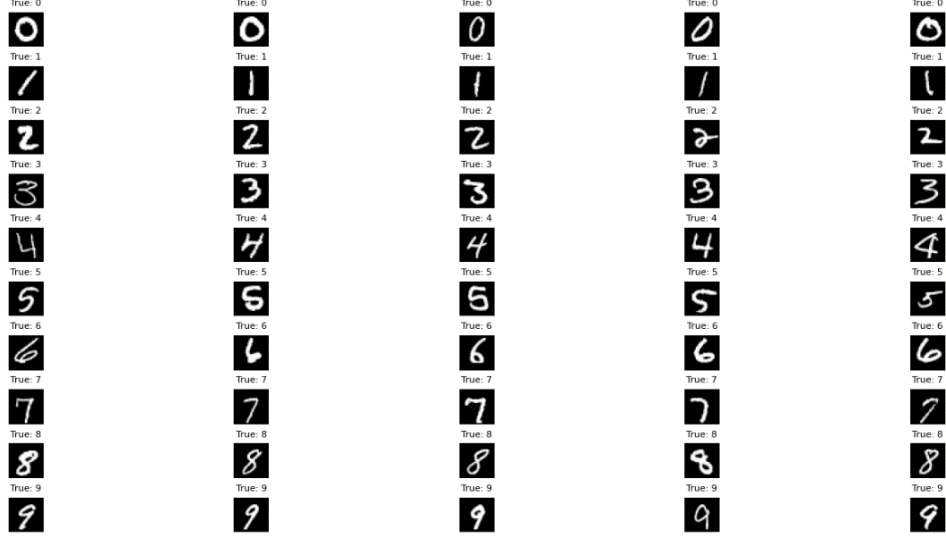
4. Çıkarımlar

- Lojistik Regresyon, basit ve hızlı bir yöntem olmasına rağmen doğruluğu CNN'e göre daha düşüktür. Yanlış tahmin edilen görsellerden de görülebileceği üzere basit doğrusal bir model karmaşık görsel örüntüleri öğrenmekte sınırlıdır.
- CNN, görsellerdeki mekansal ilişkileri öğrenebilme yeteneği sayesinde daha yüksek doğruluğa ulaşmıştır. Ancak CNN'in eğitim süresi daha uzundur.
- Eğitim ve doğrulama kayıplarındaki fark, değerlendirilmek üzere incelenmiştir.

5. Görseller

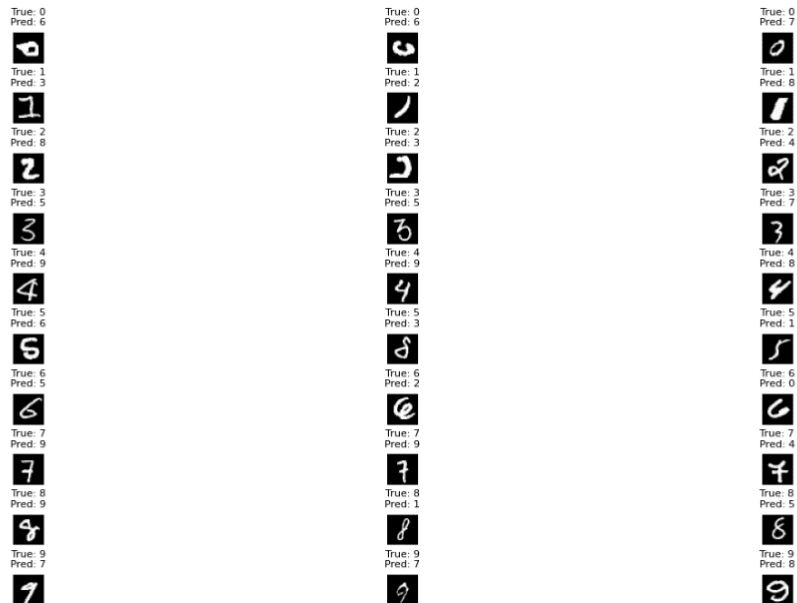
- Eğitim verilerinden örnekler: Her rakamdan 5 örnek görselleştirilmiştir.

Eğitim Verilerinden Örnekler (0'dan 9'a)

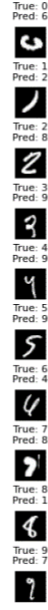


- Yanlış tahmin edilen görseller: Lojistik Regresyon ve CNN'in yanlış tahmin ettiği her rakamdan 3'er örnek gösterilmiştir.

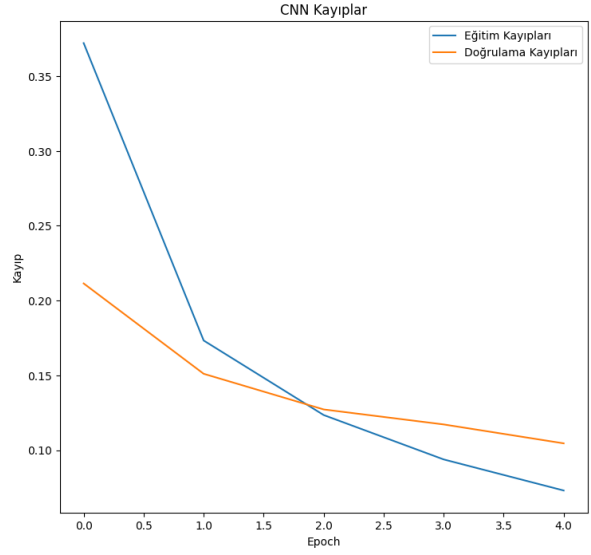
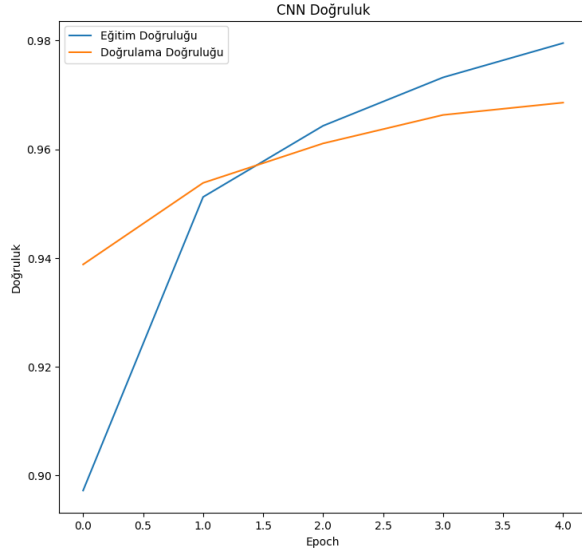
Lojistik Regresyon Yanlış Tahmin Edilen Görseller



CNN Yanlıř Tahmin Edilen Grseller



- Eđitim ve dođrulama grafikleri: CNN'in eđitim srecindeki dođruluk ve kayıplar detaylı řekilde incelenmiřtir.



6) Geliştirme Önerileri

- Daha karmaşık CNN mimarileri ile doğruluğu artırmak.
- Daha fazla veri ile modelin genel performansını iyileştirmek.
- Diğer makine öğrenimi algoritmaları ile karşılaştırmalar yapmak.

Sonuç

Bu proje, Lojistik Regresyon modeli ile CNN modelinin performanslarını incelemiş, görsel olarak sınıflandırma üzerinde karşılaştırmış ve CNN modelinin üstünlüğünü göstermiştir.