T.C Kütahya Dumlupınar Üniversitesi

Bilgisayar Mühendisliği Bölümü

Yüksek Düzey Programlama

- Digit Recognizer -

Ad-Soyad : Berfin MOGUL

Öğrenci No: 2022113171809

Öz

Bu çalışmada, el yazısı rakamları otomatik olarak tanıyabilen bir görüntü sınıflandırma modeli geliştirilmiştir. Modelin temelinde, karmaşık görsellerden anlamlı özellikler çıkarabilen bir derin öğrenme yöntemi olan Convolutional Neural Network (CNN) bulunmaktadır. 28x28 piksellik gri tonlamalı rakam görüntülerinden oluşan veri seti, piksellerin normalizasyonu ve yeniden boyutlandırılması ile model için uygun hale getirilmiştir. Modelin eğitimi sırasında, doğru sınıflandırma performansı elde edebilmek için Adam optimizasyon algoritması ve sparse categorical crossentropy kayıp fonksiyonu kullanılmıştır. Sonuç olarak, model hem eğitim hem de test veri setlerinde yüksek doğruluk oranları elde ederek, el yazısı rakamları güvenilir şekilde tanıyabilmiştir.

Giriş

El yazısı gibi kişisel ve değişken özellikler taşıyan görsel verileri doğru şekilde tanıyıp sınıflandırabilmek, modern yapay zeka uygulamalarının önemli hedeflerinden biridir. Bu hedef doğrultusunda geliştirilen algoritmalar, eğitimden sağlık sektörüne, dijital belgelerden güvenlik uygulamalarına kadar geniş bir yelpazede çözüm sunma potansiyeline sahiptir. Bu çalışmada, el yazısı rakamların doğru şekilde tanınmasını sağlayan bir Convolutional Neural Network (CNN) modeli tasarlanmıştır. CNN tabanlı modeller, görsellerdeki karmaşık pikseller arası ilişkileri öğrenip analiz ederek, görsel verilerin otomatik sınıflandırılmasını mümkün kılar.

Projede kullanılan veri seti, her biri 28x28 piksel boyutunda gri tonlamalı el yazısı rakamlardan oluşmaktadır. Modelin doğru çalışabilmesi için, veri ön işleme adımları kapsamında pikseller 0-1 aralığında normalleştirilmiş ve CNN modeli için gerekli boyutlara getirilmiştir. Modelin eğitim sürecinde, Adam optimizasyon algoritması ile modelin hızla en iyi sonuçlara ulaşması sağlanırken, sınıflandırma başarımını ölçmek için çok sınıflı sınıflandırmaya uygun sparse categorical crossentropy kayıp fonksiyonu kullanılmıştır.

Materyal ve Metot

Bu çalışmada, el yazısı rakamları sınıflandırmak için Kaggle‘dan alınan Digit Recognizer veri seti kullanılmıştır. Bu veri seti, her biri 28x28 piksel boyutunda gri tonlamalı rakam görüntülerinden oluşmaktadır. Veri seti iki ana dosyadan oluşmaktadır:

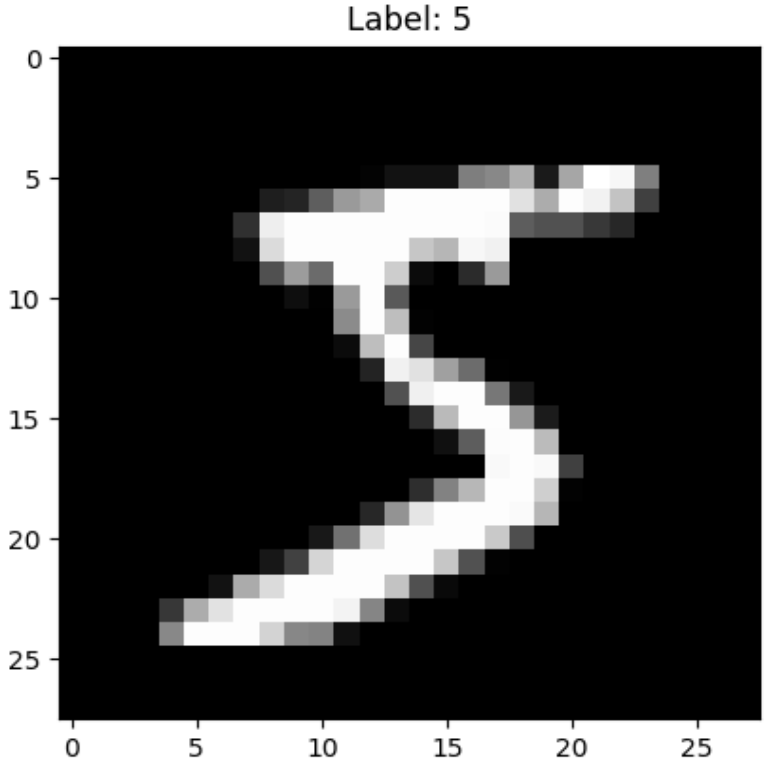
-train.csv : İlk sütun, görüntünün doğru sınıfını (0-9 arasındaki rakamları) temsil eden etiketleri içerirken, geri kalan sütunlar her bir pikselin değerini içerir.

-test.csv: İçerisinde 28.000 örnek içerir, veriler modelin performansını test etmek amacıyla kullanılır.

Proje, el yazısı rakamların sınıflandırılması için derin öğrenme tabanlı bir Convolutional Neural Network (CNN) modeli ile geliştirilmiştir.Veri seti, modelin eğitimi ve doğrulanması için %80'i eğitim ve %20'si doğrulama olacak şekilde ayrılmıştır.

Öncelikle, Digit Recognizer veri setini yüklüyoruz. MNIST veri seti 28x28 boyutunda gri tonlamalı resimlerden oluşur ve her resim bir rakamı temsil eder (0-9). Daha sonra resimlerdeki piksel değerleri 0-255 arasında olduğu için bunların normalize edilmesi gerekiyor.

Normalize edildikten sonra ilk görüntüyü elde ediyoruz.(şekil 1)



Daha sonra modeli KNN modeli ile tanıyıp eğitiyoruz, sonrasında tahmin yapıyoruz.

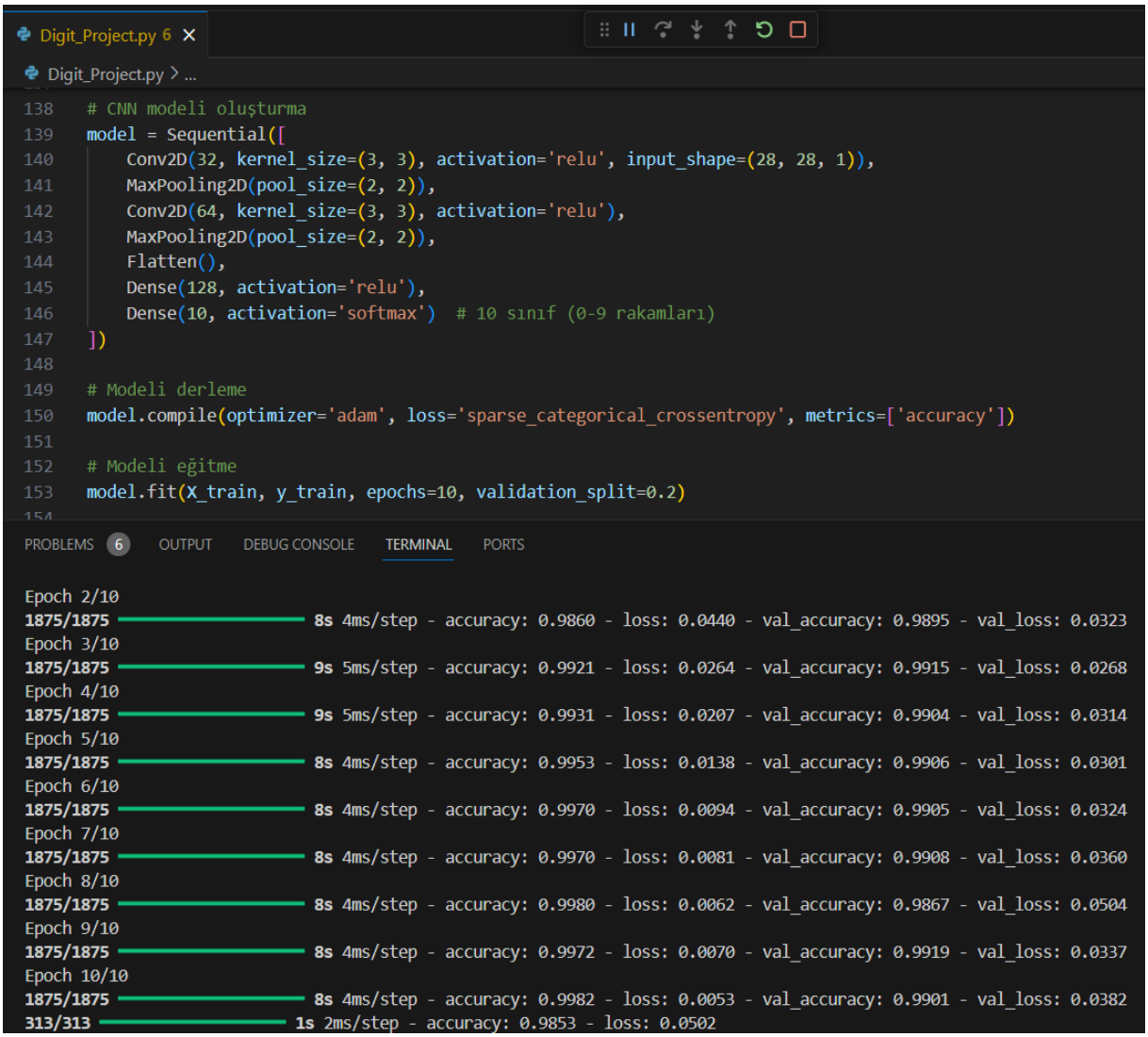
```
X_train shape: (60000, 28, 28), y_train shape: (60000,)
X_test shape: (10000, 28, 28), y_test shape: (10000,)
KNN model accuracy: 97.05%
```

Doğruluk oranı %97.05 olarak görülüyor.

CNN

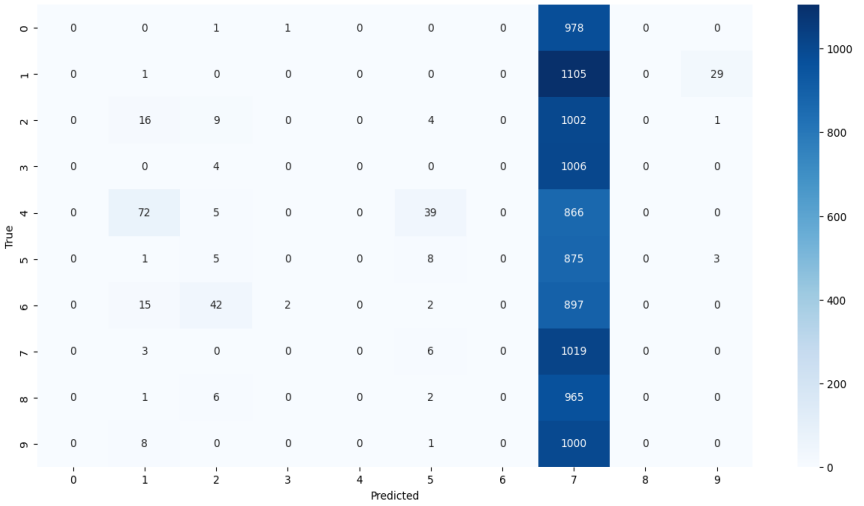
Evrişimli sinir ağları (CNN'ler), uygulanan verileri doğrudan işleyerek sonuç elde etmek için kullanılan bir makine öğrenme modelidir. Bu bir özellik öğrenme modelidir. CNN, özellikle sınıflandırma, algılama ve segmentasyon amaçları için görüntü işlemede etkileyici sonuçlar gösterdiğinden son zamanlarda çok popüler hale gelmiştir.

CNN mimarisi oluşturulduktan sonra ağın eğitimi gerçekleştirilir. Eğitim sonucunda alınan tahmin verileri doğru etiketli veriler ile karşılaştırılarak bir hata değeri hesaplanır. Geriye yayılım algoritması ile her eğitim adımında ağdaki ağırlık değerleri güncellenerek hatanın minimize edilmesi sağlanmaktadır.



KNN, veriyi sınıflandırmak için en yakın komşulara dayalı bir algoritmadır. İlk olarak, sklearn kütüphanesini kullanarak KNN modelini eğittik ve daha sonra daha derin bir öğrenme modeli olan CNN modeli oluşturduk. Katman sayısını 10 olarak belirledik. Sonuç olarak Test Verisi **0.9853** doğruluk oranı verdi.

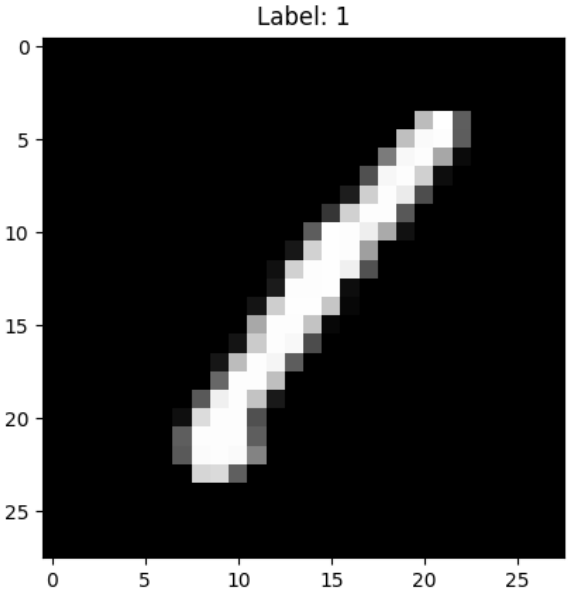
Daha sonra geliştirilen modelin hangi sınıflarda daha iyi veya kötü performans gösterdiğini görmek için bir karışıklık matrisi (confusion matrix) oluşturuldu. Bu matrisin analizi, modelin hangi sınıflarda güçlü olduğunu ve hangi sınıflarda hata yapma eğiliminde olduğunu, her bir sınıf için doğru ve hatalı tahminlerini ayrıntılı olarak gözlemlememizi sağlar. Matrisin satırlarında gerçek sınıflar, sütunlarında ise modelin tahmin ettiği sınıflar yer alır.



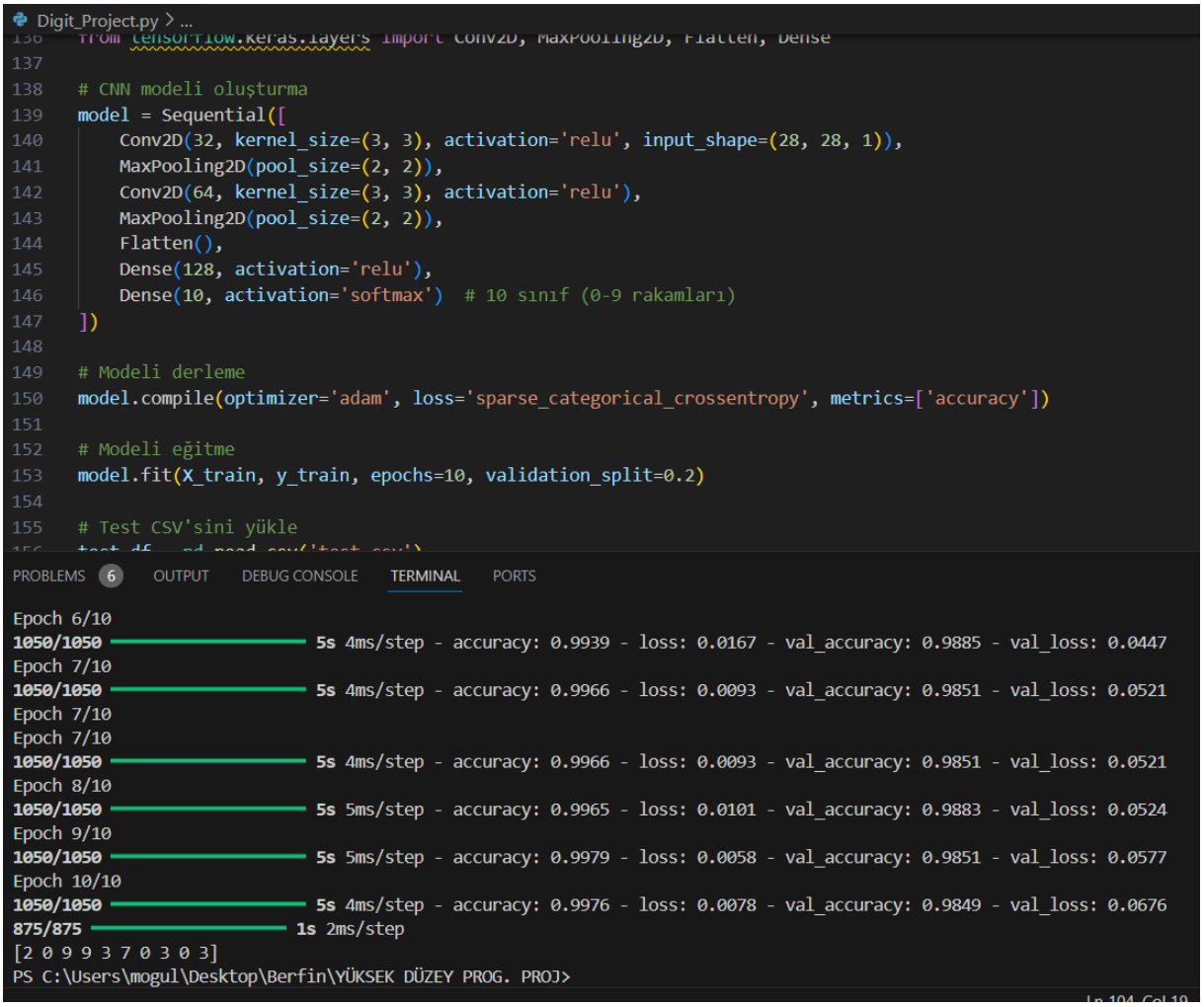
Eğitim veri setini(train.csv) yüklüyoruz, ilk satırın çıktısını alıyor ve ilk görüntüyü elde ediyoruz.

```
   label  pixel0  pixel1  pixel2  pixel3  pixel4  pixel5  pixel6  ...  pixel776  pixel777  pixel778  pixel779  pixel780  pixel781  pixel782  pixel783
0      1       0       0       0       0       0       0       0  ...         0         0         0         0         0         0         0         0
1      0       0       0       0       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         0         0         0         0
3      4       0       0       0       0       0       0       0  ...         0         0         0         0         0         0         0         0
4      0       0       0       0       0       0       0       0  ...         0         0         0         0         0         0         0         0

[5 rows x 785 columns]
Veri şekli: (42000, 28, 28, 1), Etiket şekli: (42000,)
```



CNN performansını arttırmak modeli daha çok geliştirmemize yarar sağlayacak. CNN performansını arttırmak için belli başlı şeyler yapılabilir, bunlardan birisi Epoch Artırımı, yani katman sayısını arttırmak; bir diğeriyse Veri Artırımı'dır(görüntüleri tekrardan ölçeklendirmek gibi).



Görüldüğü üzere CNN modeli tekrar oluşturuyoruz ve eğitiyoruz. Eğitim sonucunda elde edilen doğruluk oranının **0.9976** olduğu görülüyor.

Son olarak ilk 10 tahmini makineden istiyoruz ve ekrandaki sonucu ediniyoruz.

```
[2 0 9 9 3 7 0 3 0 3]
```