

FASHION-MNIST İLE DERIN ÖĞRENME TABANLI GÖRÜNTÜ SINIFLANDIRMA

Ders : Derin Öğrenme
Öğrenci: amine sayed
Numara: 23120205075

1) Proje Konusu

Proje Konusu

Bu projede, **Fashion-MNIST** veri seti kullanılarak, giyim ürünlerinin **derin öğrenme tabanlı bir Evrişimsel Sinir Ağı (Convolutional Neural Network – CNN)** ile otomatik olarak sınıflandırılması amaçlanmıştır. Model, bir görüntünün hangi giyim kategorisine (tişört, ayakkabı, çanta vb.) ait olduğunu tahmin etmektedir.

Seçilme Gerekçesi

Görüntü sınıflandırma, derin öğrenmenin en temel ve en yaygın uygulama alanlarından biridir. Fashion-MNIST veri seti, klasik MNIST rakam veri setine göre daha zor ve gerçek hayatı daha yakın olduğu için tercih edilmiştir. Bu sayede modelin yalnızca basit desenleri değil, daha karmaşık görsel özellikleri de öğrenmesi hedeflenmiştir.

Ayrıca bu proje, derin öğrenme mimarilerinin (özellikle CNN'lerin) **görüntülerden otomatik özellik çıkarma yeteneğini** göstermek için oldukça uygundur.

Daha Önce Yapılan Çalışmalar

Literatürde Fashion-MNIST veri seti üzerinde;

- Çok Katmanlı Algılayıcılar (MLP)
- Destek Vektör Makineleri (SVM)
- Evrişimsel Sinir Ağları (CNN)

gibi yöntemler kullanılmıştır. Yapılan çalışmalar CNN tabanlı modellerin, geleneksel yöntemlere kıyasla daha yüksek doğruluk sağladığını göstermektedir.

İlgili Alanın Önemi

Bu tür görüntü sınıflandırma sistemleri;

- E-ticaret sitelerinde ürün kategorizasyonu
- Otomatik etiketleme sistemleri
- Görüsel arama motorları
- Moda sektöründe öneri sistemleri

gibi birçok gerçek dünya uygulamasında aktif olarak kullanılmaktadır. Bu nedenle proje, hem akademik hem de endüstriyel açıdan önemlidir.

2) Veri Setinin Belirlenmesi

Bu projede **Fashion-MNIST** veri seti kullanılmıştır.

Veri Seti Özellikleri

- Toplam **70.000 görüntü**
 - 60.000 eğitim verisi
 - 10.000 test verisi
- Görüntüler:
 - 28×28 piksel
 - Gri tonlamalı (tek kanal)
- Toplam **10 sınıf** bulunmaktadır:
 - T-shirt/Top
 - Trouser
 - Pullover
 - Dress
 - Coat
 - Sandal
 - Shirt
 - Sneaker
 - Bag
 - Ankle Boot

Veri Setinin Avantajları

- Küçük boyutlu olduğu için hızlı eğitim imkânı sağlar
- Görüntü sınıflandırma problemleri için standart bir benchmarktir
- CNN mimarilerinin performansını değerlendirmek için uygundur

Ön İşleme (Preprocessing)

Eğitim aşamasında:

- Görüntüler `ToTensor()` ile $[0,1]$ aralığına dönüştürülmüştür
- Ek bir normalizasyon uygulanmamıştır

Tahmin aşamasında:

- Girdi görüntüler 28×28 boyutuna getirilmiştir
- Fashion-MNIST formatına uygun olması için gri tonlama ve piksel ters çevirme (invert) işlemi uygulanmıştır

3) Yöntem / Algoritma Seçimi

Kullanılan Yöntem

Bu projede **Evrişimsel Sinir Ağı (CNN)** kullanılmıştır.

Neden CNN?

CNN'ler görüntü verileri için özel olarak tasarlanmış yapılardır ve şu avantajlara sahiptir:

- Manuel özellik çıkarımına gerek duymaz
- Kenar, şekil ve desen gibi görsel özellikleri otomatik öğrenir
- Görüntü sınıflandırma problemlerinde yüksek doğruluk sağlar

Alternatif Yöntemlerle Karşılaştırma

Yöntem	Avantaj	Dezavantaj
Logistic Regression Basit		Görüntüler için yetersiz
SVM	Orta seviye başarı	Yüksek boyutta yavaş
MLP	CNN'den daha basit	Görsel yapıyı iyi yakalayamaz
CNN	Yüksek doğruluk	Daha fazla hesaplama

Literatürde yapılan çalışmalar CNN'lerin Fashion-MNIST üzerinde %90'ın üzerinde doğruluk elde edebildiğini göstermektedir. Bu nedenle CNN tercih edilmiştir.

4) Model Eğitimi ve Model Değerlendirilmesi

Model Mimarisi

Kullanılan CNN mimarisi şu katmanlardan oluşmaktadır:

- 2 adet Convolution + ReLU + MaxPooling
- Fully Connected katmanlar
- Dropout ile overfitting önleme

Eğitim Detayları

- Optimizer: Adam
- Learning rate: 0.001
- Epoch: 5
- Batch size: 100
- Loss fonksiyonu: Cross Entropy Loss

Model Performansı

Test verisi üzerinde elde edilen sonuçlar:

- **Accuracy: %91**
- Precision, Recall ve F1-score değerleri dengelidir

Değerlendirme Metrikleri

- Accuracy
- Confusion Matrix
- Classification Report (Precision, Recall, F1-score)

Ayrıca eğitim sırasında **loss** ve **accuracy** grafikleri oluşturulmuştur.

Karşılaşılan Sorun ve Çözüm

Gerçek MNIST tarzı görüntüler ilk başta yanlış sınıflandırılmıştır. Bunun sebebi, eğitim verisi ile tahmin sırasında kullanılan görüntüler arasındaki **arka plan / ön plan tersliği (domain mismatch)** olmuştur. Girdi görüntülerini ters çevirdikten sonra model doğru tahminler yapmaya başlamıştır.