

Cifar-10 Veri Seti ve Model Deneyleri

Projenin amacı: En iyi en başarılı öğrenme modelini bulmak ve daha sonra bu modeli kullanarak görsel tanıma uygulaması yapmak.

Cifar-10 Veri Seti Hakkında Bilgi

- **Cifar-10** veri seti, 10 farklı sınıftan oluşan 60,000 renkli (32x32 piksel) görüntü içerir. Uçak, otomobil, kuş, kedi, geyik, köpek, kurbağa, at, gemi, kamyon olarak 10 farklı sınıf mevcuttur.

Kullanılan Modeller ve Mimariler

Convolutional Neural Network (CNN)

- **Mimari:** Çeşitli konvolüsyonel ve pooling katmanları ile derin ağ yapısı.
- **Accuracy:** 0.7144

K-Nearest Neighbors (KNN)

- **Yöntem:** En yakın komşu algoritması.
- **Accuracy:** 0.3398

Support Vector Machine (SVM)

- **Yöntem:** Düşük boyutlu özellik uzayında optimal bir hiper-düzlem bulma.
- **Accuracy:** 0.318

Logistic Regression

- **Yöntem:** Sınıflandırma için lineer bir model.
- **Accuracy:** 0.2985

Decision Tree

- **Mimari:** Karar ağaçları yapısı.
- **Accuracy:** 0.2435

Random Forest

- **Mimari:** Birçok karar ağacının bir araya gelmesiyle oluşturulan ensemble model.
- **Accuracy:** 0.423

Yapay Sinir Ağı (YSA)

- **Mimari:** Giriş, gizli ve çıkış katmanlarından oluşan derin öğrenme modeli.
- **Accuracy:** 0.534

Deneyde Kullanılan Yöntemler

- **Veri Ön İşleme:** Görüntülerin normalize edilmesi, boyutlandırılması.
- **Eğitim ve Test Ayrımı:** Veri setinin eğitim ve test olarak ayrılması.
- **Model Eğitimi:** Her modelin belirlenen mimarisi ile eğitilmesi.
- **Performans Ölçümü:** Accuracy metriği kullanılarak modellerin performansının değerlendirilmesi.

Deney Sonuçları

- **CNN:** En yüksek doğruluk oranıyla (0.7144) en başarılı model olarak öne çıktı.
- **YSA:** Diğer modellerden (0.534) sonra ikinci en yüksek doğruluk oranına sahipti.
- **Random Forest:** Ensemble yöntemiyle (0.423) iyi bir sonuç elde etti.
- **SVM, KNN, Logistic Regression, Decision Tree:** Daha düşük doğruluk oranları ile sırasıyla sıralandılar.

Sonuçlar ve Değerlendirme

- Cifar-10 veri setinde CNN ve YSA gibi derin öğrenme modelleri, diğer geleneksel makine öğrenimi modellerine göre daha yüksek doğruluk sağlamıştır.
- Random Forest gibi ensemble yöntemleri ise orta seviyede performans göstermiştir.
- Geleneksel yöntemler (SVM, KNN, Logistic Regression, Decision Tree) ise düşük doğruluk oranlarıyla sınırlı kalmıştır.

Sonuç

Kullanılan modellerden mimari olarak sadece ysa da düzenleme yapılmıştır. Bu düzenleme bize gösteriyor ki bir modelin başarılı olması için epoch sayısının nöron sayısının ve batch_size'ını önemi vardır. Datasetin çeşitliliği görsel mi sayısal mı olması da hangi modelin başarılı olacağını gösteriyor. Görsel veri setlerinde derin öğrenme modelleri daha başarılı olmuştur. Sayısal veri seti olsaydı belki makine öğrenmesi modelleri de başarılı olabilirdi. Tabi burda kullanılan mimari yapının da önemi var. En başarısız modelin mimari yapısı ile oynayarak en başarılı hale de getirebiliriz.