

AYGAZ YAPAY ZEKAYA GİRİŞ BOOTCAMP

PROJE SUNUMU

İçeriği görüntülerden oluşan veri setini çeşitli makine ve onun kümesinde bulunan derin öğrenme teknikleri ile daha iyi okunabilir ve gözlemlenebilir sonuçlara ulaşabileceğimiz hale getirmek projenin başlıca amaçlarından biridir.

Projeyi geliştirirken kullanılan materyal , metot ve adımlar ise şunlardır:

- Veriyi okuma ,işleme ,görselleştirme ve öğrenme tekniklerini uygulayabilmemiz için kullanılan kütüphaneler.(pandas,numpy,matplotlib,seaborn,randomforestclassifier vb.)
- Eğitim ve test kümeleri oluşturarak verinin işlenmesini ve öğrenme- test etme işlemini kolaylaştırmak.
- Görselleştirme ile veri setinin içeriğini gözlemlemek.(Eğitim ve test kümesi arasındaki farka bakabilmek)
- KNN, SVC, Random Forest ve Decision Tree makine öğrenmesi algoritmalarını kullanarak ve accurasy ,recall ,f1 score ve precision değerlerini inceleyerek hangi algorithmada en iyi performansı verdiği gözlemlemek.
- Aynı uygulamayı YSA ve CNN gibi derin öğrenme teknikleriyle yapıp sonucu gözlemlemek.

Veri setinin adı fashion_mnist içeriğinde ise ayakkabı ,gömlek ,çanta vb. giyim ürünleri vardır. Bu giyim ürünleri hem test hem train kümesine ayırarak birbirinden öğrenme ve test etme uygulamaları ile makinemizi eğitmeyi amaçladık.

Kullandığımız makine öğrenmesi modellerinden KNN,SVC,Random Forest ve Decision Tree arasından en iyi performansı gösteren modelimiz SVC olurken bu modelin zaman bakımından en uzun sürede işlemi tamamlayan da model olduğunu söylememiz gerekir. En kısa zaman ve diğer modellerle karşılaştırıldığında iyi bir performans sergileyen modelimiz ise KNN'dir. Decision Tree ise en kötü performansı göstererek hem uzun hem verimsiz bir sonuç almamızı sağladı.

Kullandığımız derin öğrenme modellerinden YSA ve CNN arasında ise en iyi performansı gösteren modelimiz çeşitli değişiklikler yaparsak yine de YSA ile sağlanmış oluyor. YSA modelimizde parametremizi "adam" ve "adagrad" yaparsak kötü sonuç almışken "adadelta" ile yaptığımızda neredeyse eğitim ve test arasında farksız bir sonuç elde etmiş oluyoruz. CNN modelimizde ise "adadelta" çok kötü bir performans gösterirken "adam" parametresi ona nazaran daha iyi bir sonuca ulaşmamızı sağladı.

Referans ve Kaynak olarak Global AI Hub mentör hocamızın projesinden ve eski kod kütüphanelerimden yararlandım.