

## WEKA'DA MODEL OLUŞTURMA

Kullanılacak olan makine öğrenmesi yazılımı ne olursa olsun öncelikle veriler bir elektronik tablolama programı veri tabanı vb. ile yapılandırılmış hale getirilerek kaydedilmelidir. Bunun yanında veriler üzerinde indirgeme, temizleme, aykırılık kontrolü, normalizasyon vb. işlemler yapılarak veri seti analiz edilebilecek son hale getirilmelidir. Veri bu forma dönüştürüldükten sonra sırasıyla aşağıdaki tanımlar kullanılır.

1) WEKA yazılımı açılarak Explorer girilir burada;

Open File → Dosyadan (.csv,.txt,.xls)

Open URL → Bir web adresinden veri almak için

Generate → Alıştırma yapmak için sentetik veri oluşturur.

Seçeneklerinden uygun olan bölüme girilir. Veriler WEKA'ya alınır.

2) Bu veriler üzerinde sınıflandırıcı model oluşacaksa "Classify" sekmesine geçilir.

Aşağıdaki adımlar uygulanır.

a) Classify sekmesinden choose bölümüne girilerek belli başlı algoritmalar (birçok sınıflandırma probleminde yüksek performans gösterdiği diğer araştırmacılar tarafından test edilmiş) kullanılarak denemeler yapılır. Algoritmaları birbirleri ile karşılaştırmak için Correctly Classified Instance, Kappa artistic, mean absolute error... metriklerinin değerlerine göre kullanılır. Buna göre başarıları arttırmak amacıyla algoritmaların parametreleri choose butonunun yanındaki isimlere direkt tıklatılarak parametreler değiştirilebilir. Test options bölümünde yer alan eğitim ve test seti kullanma şekilleri değiştirilebilir.

Başarı oranı tatmin edici ise bir sonraki adıma geçilir. Değil ise önceki makine öğrenmesi aşamalarına geri gönderilir.

3) Seçilen modele sağ tıklatılarak "Save model" seçilir ve model kaydedilir.

4) WEKA kapatılmadan test verileriyle test yapılacaksa

a) Veri seti .arff formatında kaydedilir.

b) Veri seti .arff bir text editöründe açılır. @data'dan sonra genel kısım silinerek buraya modelin sınıflandırması istenilen test verileri yazılır. Test verilerinin zaten eğitim verisi içinde olup olmadığı kontrol edilir. Çünkü eğitim verisi içerisinde bulunan bir satır test verisindeki satır ile örtüşürse model %100 doğru sınıflandıracaktır. Ancak gerçekçi bir test olmaktan uzaklaşacaktır. Test verisi içerisinde modelin cevaplamasını istediğimiz ve genelde son virgülden önce sınıf bilgisinin yerine soru işareti konulabilir. Bu yapılırsa model test verisi geçirildikten sonra hataları görmek mümkün olmaz. Soru işareti konulmayıp, sınıf bilgisi verilirse WEKA hata raporu hazırlayabilecektir.

c) Test options → supplied test set → set butonuyla .arff formatındaki test videosu gösterilir.

WEKA kapatıldıktan sonra test yapılacaksa;

a) Bir önceki maddede anlatıldığı üzere test verisi .arff formatında oluşturulur.

b) WEKA açılarak Explorer ekranına geçilir.

c) Classify sekmesine geçmek gerekir. Ancak bu sekme herhangi bir veri WEKA yazılımına aktarılmamışsa pasiftir. Bu yüzden geçilmez. Bu engeli aşmak için herhangi bir veri hiç

kullanılmayacak olsa da preprocess sekmesinden açılır. (generate seçeneğiyle WEKA'nın sentetik veri üretmesi sağlanarak bu yapılabilir.)

- d) Classify sekmesine geçilir result list bölümündeki boşluğa sağ tıklatılır Load model seçilerek daha önce kaydedilmiş olan model gösterilir.
  - e) Test options bölümündeki supplied test set → set .arff formatındaki test videosu gösterilir.
- 5) Test options → more options → açılan pencerede output predictions → plain text vb. istenilen türe ayarlanır. Output predictions modelin tahminlerinin ve göstergelerinin nereye hangi formatta çıkacağını belirler. plain text seçilirse WEKA'nın kendi çıktı ekranında görülecektir. Bu seçildikten sonra result list sağ tıklatılır. Re-evaluate model on current test set seçilir. Bu aşamadan sonra WEKA kendi tahminlerini ve raporlarını ekrana basacaktır.