# Yorum

Stats Model ile Lojistik Regresyon

Stats models kütüphanesi, veriler üzerinden detaylı olarak istatistiksel anlamlılıkların elde edilmesini sağlayan bir kütüphanedir. Scikit learn kütüphanesinden en belirgin farkı ise modele dahil olan özellikler hakkında detaylı sonuçlar elde edebilme imkanı sunmasıdır. Stats models kütüphanesinde ise yaşadığım en önemli deneyim ise, bağımlı ve bağımsız değişkenleri ayırırken yaptığım dikkatsizlik sonucu stats model kütüphanesinden aldığım hatadır.

Lojistik regresyonu ilk olarak scikit learn kütüphanesinden denediğimde %100 tahmin performansı ile karşılaştım. İşlerin yolunda gitmediğini fark etsem de analizlere devam ettim. Bir istatistikçi olarak modelin ve katsayıların anlamlı olup olmadığı benim için çok önemli Bu nedenle bana çeşitli istatistiksel testleri birlikte verebilecek olan stats model kütüphanesine yöneldim. Ancak ne yaparsam yapayım "perfect separation error" hatasını almayı durduramadım. En son eğitim ve test veri setlerini kontrol ettiğimde bağımsız değişkenlerin arasında bağımlı değişkenin de yer aldığını gördüm. Stats models bana bu hatayı mükemmel ayrım hatası olarak verdi. Stats model kütüphanesi her ne kadar yaygın olarak kullanılmasa da bu hata benim oldukça ilgimi çekti. Çünkü ne olursa olsun bana anileze devam etmeme izin vermedi. Ayrıca bu hatayı, modele çok fazla değişken eklediğimde de aldım. Bu durum bana, her zaman tek bir yöntemin sonuca ulaşmak konusunda yeterli olmadığını tekrar hatırlattı.