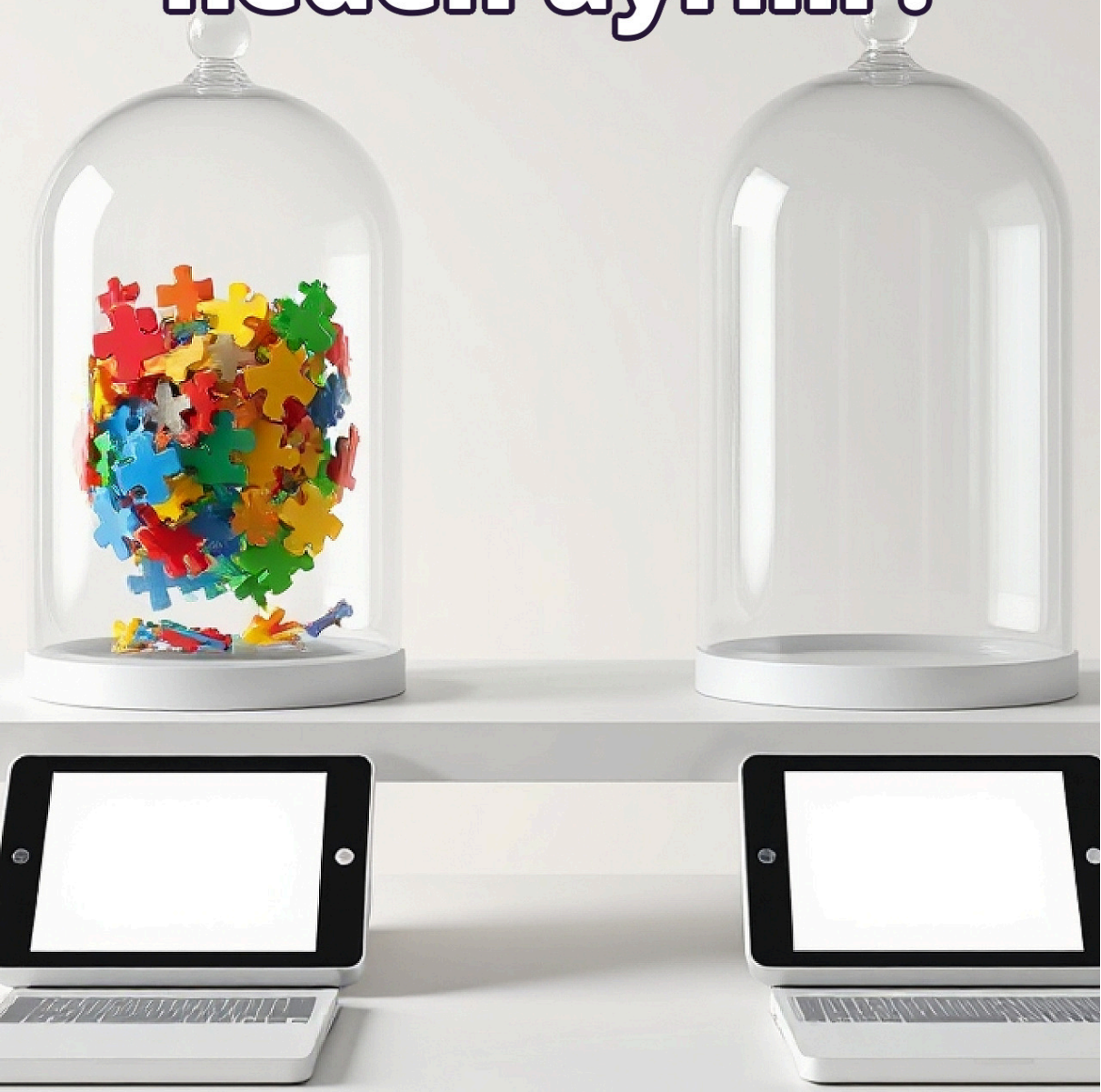
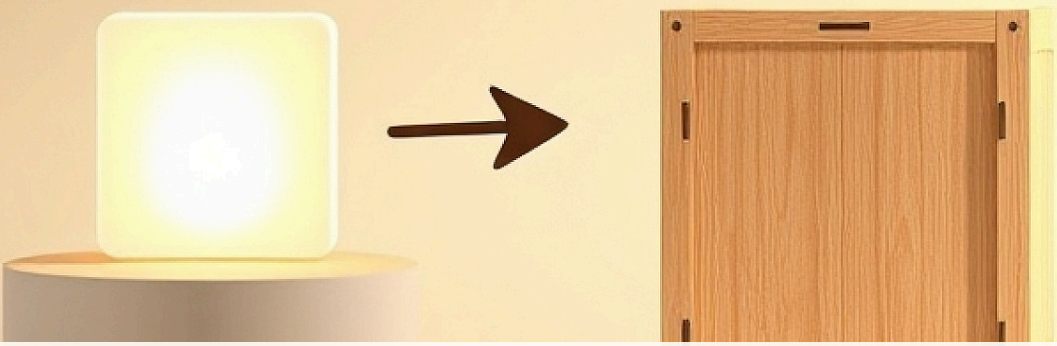
# Eğitim verisi ile test verisi neden ayrılır?



www.yzokul.com.tr

# Eğitim verisi ile test verisi neden ayrılır?



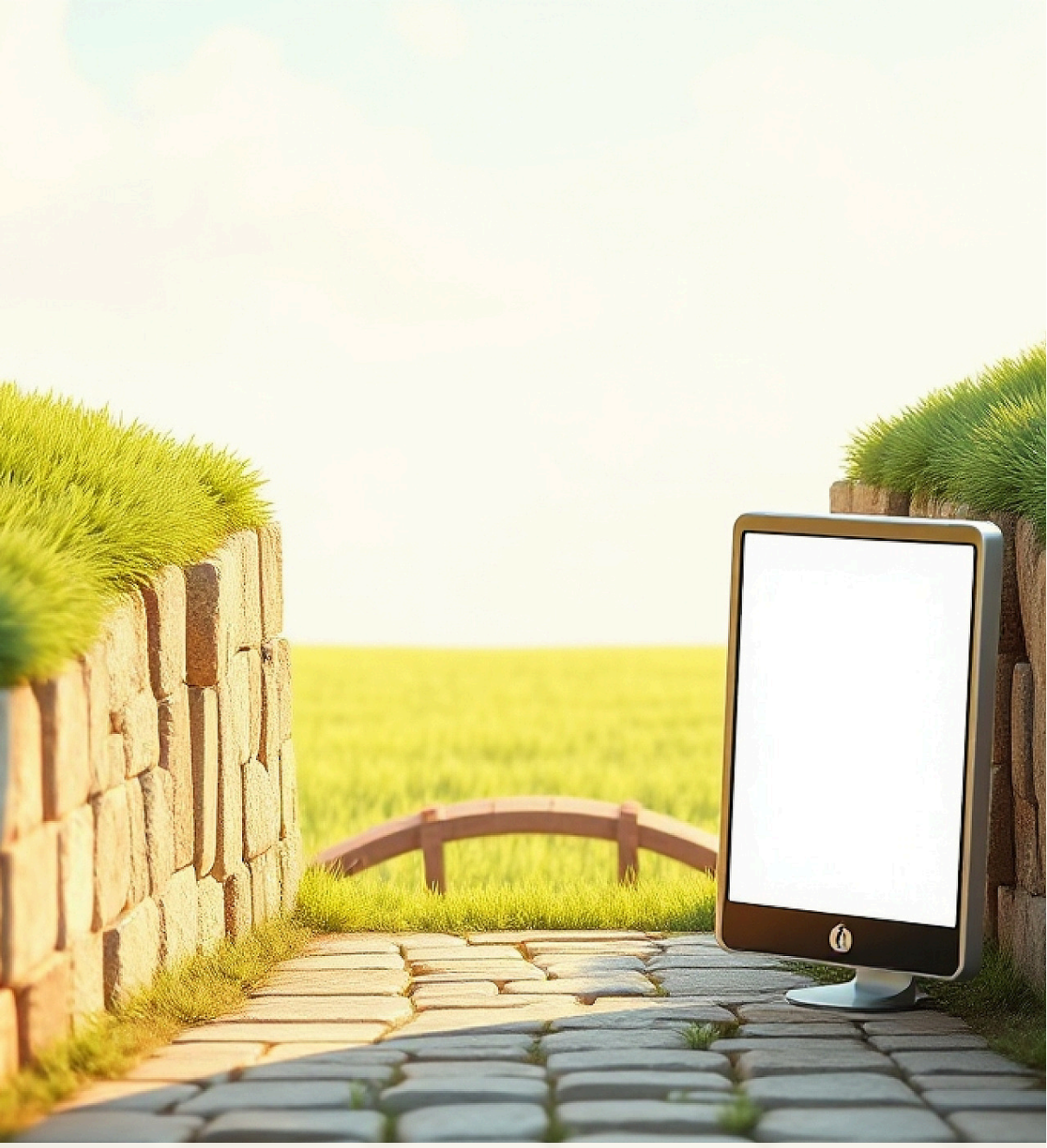
Bir model, eğitim verisi adı verilen bir dizi örnekten öğrenir. Modeli yalnızca bu aynı örnekler üzerinde değerlendirirsen, gurur okşayan ama yanlış bir tablo elde edersin. Kısa yanıt şudur: Öğrenilen örneklerde iyi performans göstermek, yeni örneklerde iyi performans göstereceğini kanıtlamaz. Bu yüzden bazı veriler son bir kontrol için saklanır. Bu kitap, eğitim verisi ile test verisinin neden ayrı tutulduğunu, tam bir yeni başlayan için adım adım açıklar.



Eğitim verisi, bir modelin öğrendiği örnekler kümesidir. Model bu örneklere tekrar tekrar bakar ve her seferinde kendini ayarlar. Bunu, sınava çalışmadan önce incelediğin alıştırma materyali gibi düşün. Modelin iç ayarları, bu veride gördüklerine göre değişir. Eğitim verisi olmadan modelin öğreneceği bir şey yoktur.



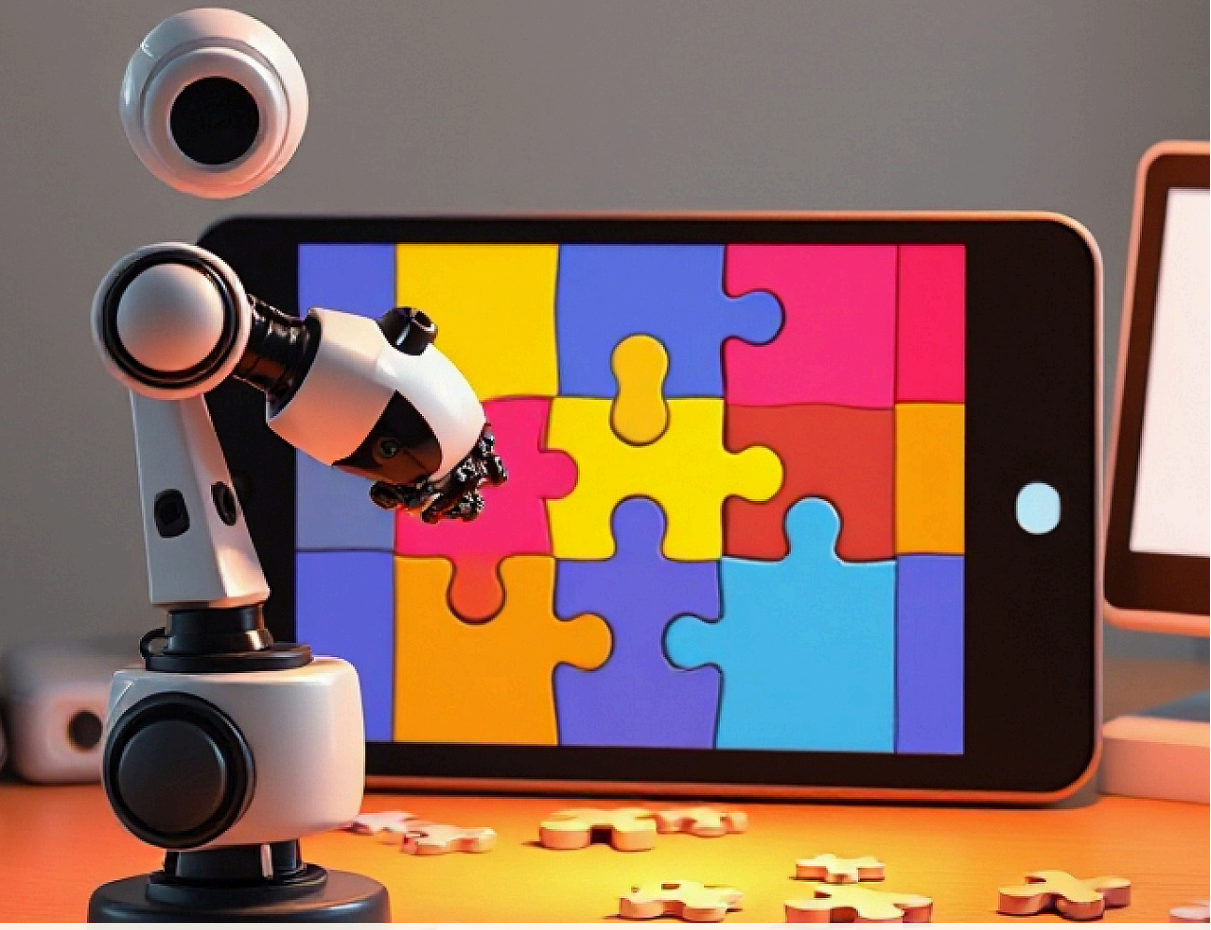
Test verisi, modelin hiçbir zaman öğrenmediği ayrı bir örnekler kümesidir. Bunu yalnızca sonunda, modelin görmediği örneklerle ne kadar iyi başa çıktığını kontrol etmek için kullanırsın. Bu, çalıştıktan sonra girdiğin final sınavı gibidir. Modelin bu veriye uyum sağlama şansı yoktur. Tek görevi, yeni örneklerdeki gerçek yeteneği ölçmektir.



Modelin öğrendiği örneklerdeki başarı, yeni örneklerde başarılı olacağını göstermez. Bir model eski veride iyi görünüp yine de yeni veride başarısız olabilir. Kümeleri ayırmanın arkasındaki temel fikir budur. Model, yalnızca gördüğü belirli örneklere uyum sağlamış olabilir. Yeni örnekler, genel bir örüntü mü öğrendiğini yoksa yalnızca eğitim kümesinin ayrıntılarını mı öğrendiğini test eder.



Daha önce çözdüğü soruları harfi harfine ezberleyen bir öğrenci düşün. Bu öğrenci o sorularda tam puan alabilir ama yeni sorularda zorlanabilir. Bir örnekler kümesi üzerinde eğitilen bir model de aynı şekilde davranabilir. Altında yatan kuralı anlamadan gördüğü yanıtları yeniden üretebilir. Yalnızca eğitim verisindeki yüksek puanlar yanıltıcı olabilir.



Bunu, hiç görmediği soruları çözebilen bir öğrenciyle karşılaştır. İstediğimiz gerçek beceri budur ve test verisinin bir modelde ölçmesi gereken şey de budur. İyi bir model genelleme yapmalıdır, yani yeni örneklerle iyi başa çıkmalıdır. Test verisi bu yeteneğe dair gerçekçi bir görünüm sunar çünkü model ondan öğrenmemiştir.



Modelin yeteneğinin dürüst bir son kontrolüne ihtiyacımız var. Test verisi bu kontrolü sağlar çünkü modelin ondan öğrenme şansı olmamıştır. Ayrı bir test kümesi olmadan, modelin gerçekten öğrenip öğrenmediğini yoksa yalnızca ezberleyip ezberlemediğini anlayamazsın. Son kontrol, modelin gerçek kullanımda yeni örneklerde nasıl performans göstereceğini yansıtmalıdır. Bu yüzden test kümesi ayrı kalmalıdır.



Test verisi eğitime sızarsa ne ters gider? Model yanıtları zaten görmüştür. Son kontrol o zaman gerçek beceriyi değil, belleği ölçer. Buna veri sızıntısı denir ve test örnekleri kazara eğitim kümesine girdiğinde ortaya çıkar. Küçük bir sızıntı bile sonuçları güvenilmez kılabilir. Model iyi performans gösteriyor gibi görünebilir, ama yalnızca test örneklerini daha önce gördüğü için.



Test verisi eğitime sızdığında, bildirilen başarı modelin gerçek yeteneğinden yüksek görünür. Sayılar yanıltıcı hale gelir ve onlara güvenemezsin. Model gerçek kullanıma hazır sanabilirsin, oysa değildir. Bu abartılı tahmin tehlikelidir çünkü ona dayanan kararlar pratikte başarısız olabilir. Kümeleri ayrı tutmak bu yanlış güveni önler.



Basit kural şudur: Test verisi eğitim sırasında el sürülmeden kalır ve yalnızca son değerlendirme için açılır. Onu hiçbir şekilde modeli ayarlamak için kullanma. Ayrı bir yerde tut ve model tamamen oluşturulana kadar ona bakma. Test kümesini temiz tutmak, son sonucu dürüst tutar. Bu kuralı söylemek kolaydır ama uymak önemlidir.



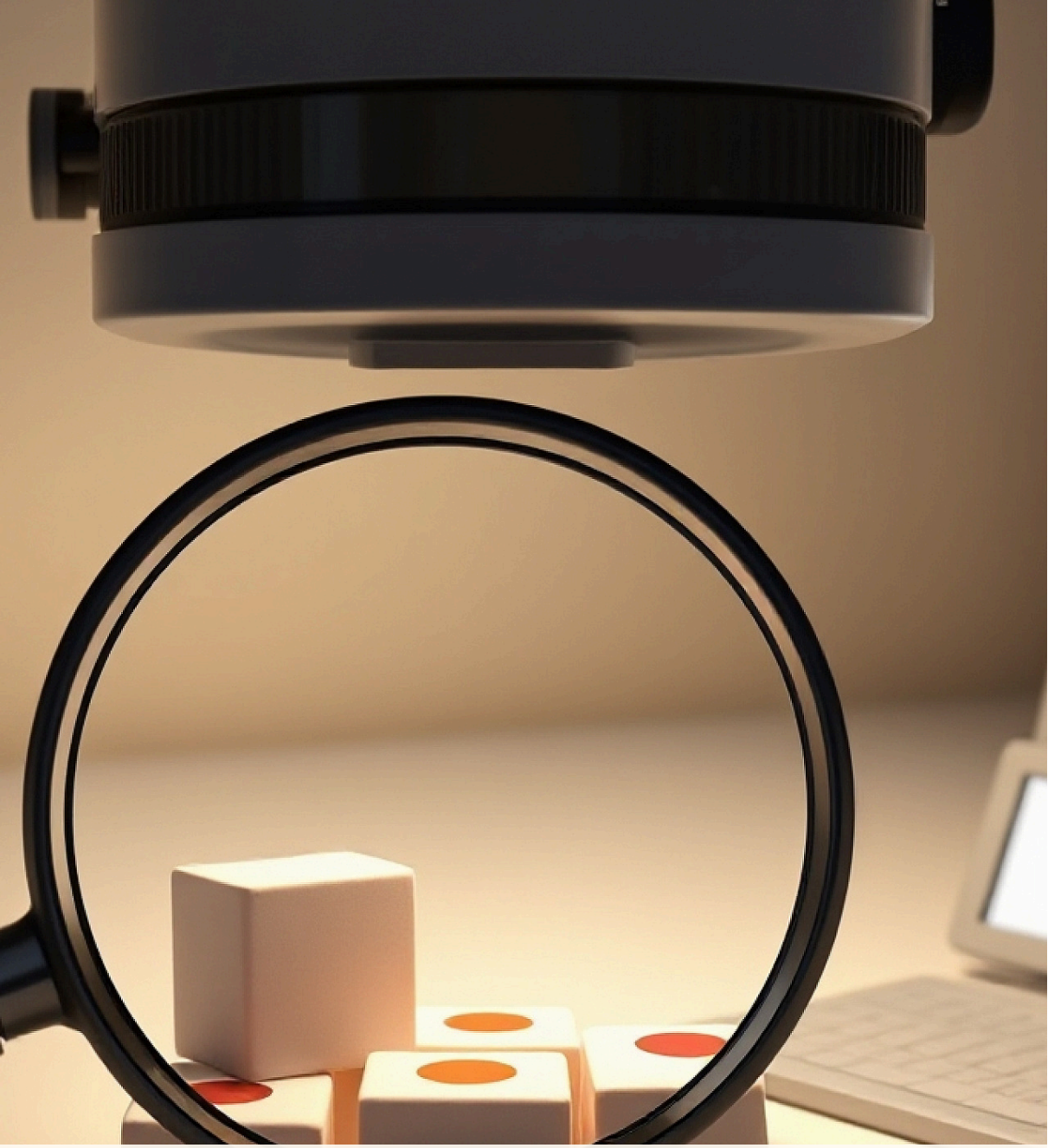
Doğrulama verisi, model oluşturulurken kullanılan bir başka ayrı kümedir. Hangi model sürümünü tutacağın gibi seçenekler arasında seçim yapmana yardımcı olur. Modeli doğrudan eğitmek için değil, modelleri karşılaştırmak veya ayarları düzenlemek için kullanırsın. Doğrulama verisi, hazırlanmak için kullandığın bir deneme sınavı gibi işlev görür. Son test kümesine dokunmadan geliştirme sırasında sana geri bildirim verir.



Doğrulama verisi de aynı nedenle eğitimden ayrı tutulmalıdır. Doğrulama örnekleri eğitime sızarsa, oluşturma sırasında yapılan seçimler ezberlenmiş örneklere dayanırdı. Daha önce gördüğü veride iyi görünen bir modeli seçerdin. Bu da hangi modelin gerçekten daha iyi olduğu konusunda seni yanıltırdı. Doğrulamayı ayrı tutmak, seçimlerinin yeni örneklere dayanmasını sağlar.



Üç ayrı yığın düşün: öğrenmek için eğitim verisi, seçim yapmak için doğrulama verisi ve son kontrol için test verisi. Her yığının tek ve net bir işi vardır. Eğitim yığını modele öğretir. Doğrulama yığını en iyi sürümü seçmene yardımcı olur. Test yığını son ve dürüst puanı verir. Yığınları karıştırmak amaçlarını bozar.



Yaygın bir yeni başlayan hatası, modeli öğrendiği aynı veri üzerinde değerlendirmektir. Bu, ne kadar iyi çalıştığına dair gurur okşayan ama yanlış bir tablo verir. Model eğitim örneklerini ezberlemiş olabilir, bu yüzden puanı yüksek görünür. Gerçek performansı görmek için her zaman ayrı veri üzerinde değerlendir. Eğitim verisi üzerinde test etmek, bir öğrenciyi ezberlediği sorulardan sınav yapmak gibidir.



Bir başka hata, test verisine çok erken bakmak veya modeli ona göre ayarlamaktır. Bunu yaptığın anda test kümesi artık adil bir son kontrol olmaktan çıkar. Modeli test örneklerinde iyi performans gösterecek şekilde ayarlayabilirsin; bu da bir tür sızıntıdır. Test kümesi yalnızca bir kez, en sonda kullanılmalıdır. Onu geliştirme sırasında her türlü etkiden koru.



Öğrenilen örnekler yeni örneklerdeki başarıyı kanıtlamaz, bu yüzden veri ayrı tutulur. Üç ayrı kümeyi hatırla: öğrenmek için eğitim, seçmek için doğrulama ve son kontrol için test. Temiz bir test kümesi dürüst bir son sonuç verir. Test verisini sona kadar el sürülmeden tut ve bir modeli asla öğrendiği veri üzerinde değerlendirme. Bu basit alışkanlıklar değerlendirmeni güvenilir kılar.