

# Eğitim Modeli Seçimi ve Makine Öğrenmesi

**Eğitim Modeli Seçimi Nedir?** Eğitim modeli seçimi, verileri analiz etmek ve belirli görevleri yerine getirmek için uygun matematiksel modelleri seçme sürecidir. Bu, istatistik, makine öğrenimi ve veri madenciliği gibi alanlarda uygulanır.

**Makine Öğrenmesi Modelleri ve Veri Gereksinimleri** Makine öğrenmesi modellerinin iyi performans gösterebilmesi için çok sayıda veri gereklidir. Eğitim verileri, metinler, resimler veya hizmet kullanıcılarından toplanan veriler gibi çeşitli kaynaklardan gelebilir. Bir modeli eğitmek, bu verilerden anlamlı bağlantılar öğrenmeyi ve tahmin yapabilmek için doğru ağırlıkları ve eşik değerlerini belirlemeyi içerir.

**Denetimli Öğrenme ve Ampirik Risk Minimizasyonu** Denetimli öğrenmede, algoritma, etiketli örnekler üzerinde çalışarak ve hatayı (kayıp) en aza indiren bir model oluşturarak öğrenir. Bu sürece "ampirik risk minimizasyonu" denir. Kayıp, modelin ne kadar kötü tahmin yaptığını gösteren bir sayıdır. İyi bir modelin amacı, tüm örnekler üzerinde ortalama olarak düşük kayıplı bir tahmin yapmaktır.

## Kayıp Fonksiyonu ve Doğrusal Regresyon

Doğrusal regresyon modellerinde yaygın olarak "karesel kayıp" fonksiyonu kullanılır. Tek bir örnek için kayıp, tahmin edilen değer ile gerçek değer arasındaki farkın karesidir:

$$\text{Karesel Kayıp} = (y - \hat{y})^2$$
$$\text{Karesel Kayıp} = (y - \hat{y})^2$$
 Burada  $y$ , gerçek değer ve  $\hat{y}$ , tahmin edilen değerdir.

**Ortalama Kare Hata (MSE)** MSE, tüm veri kümesi için örnek başına düşen ortalama karesel kayıptır: 
$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$
 Burada  $N$ , toplam örnek sayısıdır.

**Örnek:** Bir veri setindeki gerçek değerler  $y = [1, 2, 3, 4, 5]$  ve tahmin edilen değerler  $\hat{y} = [1, 2, 7, 2, 3]$  olsun. MSE şu şekilde hesaplanır:

$$\text{MSE} = \frac{(1-1)^2 + (2-2)^2 + (3-7)^2 + (4-2)^2 + (5-3)^2}{5} = \frac{0 + 0 + 16 + 4 + 4}{5} = 4.8$$

Farklı bir tahmin seti ile:  $\hat{y} = [1, 2, 4, 3, 5]$

$$\text{MSE} = \frac{(1-1)^2 + (2-2)^2 + (3-4)^2 + (4-3)^2 + (5-5)^2}{5} = \frac{0 + 0 + 1 + 1 + 0}{5} = 0.4$$

İkinci tahmin seti daha düşük bir MSE'ye sahiptir, yani daha iyi bir modeldir.

## Toplu Öğrenme (Ensemble Learning)

Toplu öğrenme, birden fazla modelin birleştirilmesiyle daha güçlü ve doğru tahminler yapmayı amaçlar. İki ana toplu yöntem türü vardır:

- Sıralı Topluluk Yöntemleri:** Modeller ardışık olarak eğitilir, her yeni model önceki modelin hatalarını düzeltmeye çalışır (örneğin, boosting).

2. **Paralel Topluluk Yöntemleri:** Modeller bağımsız olarak ve paralel olarak eğitilir, ardından sonuçlar birleştirilir (örneğin, bagging).

## **Önyargı ve Varyans**

Bir modelin hata oranı, önyargı ve varyans olarak ikiye ayrılabilir:

- **Önyargı:** Modelin ortalama tahminlerinin ne kadar doğru olduğunu gösterir.
- **Varyans:** Modelin farklı veri setleri üzerindeki tahminlerinin ne kadar değişken olduğunu gösterir.