

Polinom Regresyon: Doğrusal Modellerin Ötesine Geçmek

Atıl Samancıoğlu

1 Giriş

Makine öğrenmesinde doğrusal regresyon, iki değişken arasındaki ilişkiyi modellemek için kullanılan en basit yöntemlerden biridir. Ancak verideki ilişki doğrusal değilse, bu yöntem yetersiz kalabilir. Bu gibi durumlarda **polinom regresyon** kullanılır.

2 Doğrusal ve Doğrusal Olmayan İlişkiler

Doğrusal regresyonda hedefimiz, şu denklemle ifade edilen en iyi doğruyu bulmaktır:

$$\hat{y} = \beta_0 + \beta_1 x$$

Ancak veriler eğrisel (non-linear) bir desen izliyorsa, bu doğru hiçbir zaman yeterli olmaz. Bu durumda daha karmaşık, eğrisel bir fonksiyon gerekir.

3 Polinom Regresyon Nedir?

Polinom regresyon, girdiyi (bağımsız değişkeni) daha yüksek dereceli terimlerle modelleyerek daha esnek tahminler yapılmasını sağlar. Örneğin:

$$\hat{y} = \beta_0 + \beta_1 x + \beta_2 x^2 + \dots + \beta_n x^n$$

Buradaki n , polinomun derecesini belirtir. Derece arttıkça model daha esnek hale gelir ancak **aşırı öğrenme (overfitting)** riski artar.

4 Polinom Derecesi Nedir?

Polinomsal regresyonun gücü, modelin derecesinde (degree) yatar. Polinom derecesi, bağımsız değişkenin üst dereceli terimlerinin modele dahil edilip edilmeyeceğini belirler.

Derece 0

$$h(x) = \beta_0$$

Sabit bir modeldir. Tüm tahminler aynı değeri verir.

Derece 1 (Basit Doğrusal Regresyon)

$$h(x) = \beta_0 + \beta_1 x$$

Derece 2

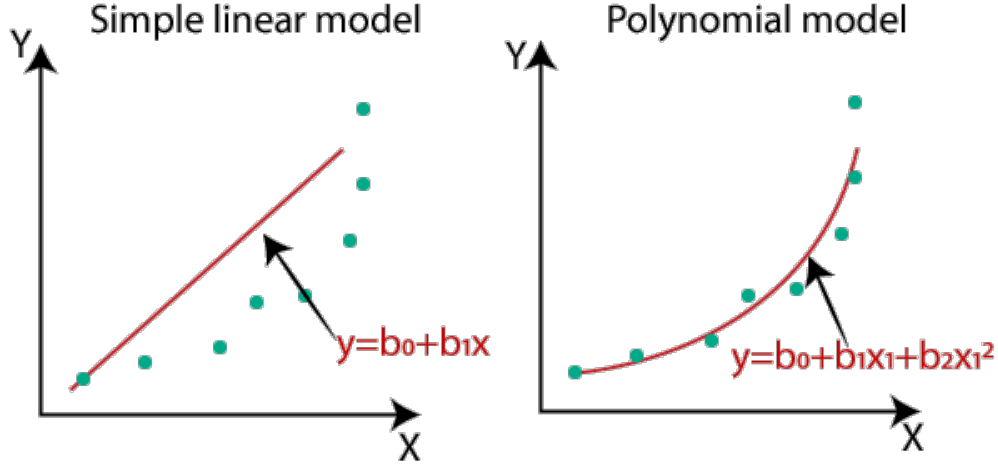
$$h(x) = \beta_0 + \beta_1 x + \beta_2 x^2$$

Genel Polinomsal Model

$$h(x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \dots + \beta_n x^n$$

5 Neden Polinom Regresyona İhtiyacımız Var?

Aşağıdaki görselde, veri noktaları eğrisel bir dağılım göstermektedir. Kırmızı çizgi doğrusal regresyonu, mavi eğri ise polinom regresyonu temsil eder:



Şekil 1: Doğrusal regresyon (kırmızı) eğrisel ilişkiyi yakalayamazken, polinom regresyon daha başarılıdır.

Görüldüğü gibi doğrusal model veriye uyum sağlayamazken, polinom model eğrisel yapıyı yakalayabilmektedir. Bu durum, polinom regresyonun gücünü ortaya koyar.

6 Sonuç

Polinom regresyon, doğrusal olmayan ilişkileri modellemek için güçlü bir araçtır. Ancak model karmaşıklığı arttıkça aşırı öğrenme riski de artar. Bu nedenle:

- Düşük dereceler genellikle yeterlidir.
- Modelin başarısı **doğrulama (validation)** verisi ile test edilmelidir.

Doğru dereceyi seçmek, modelin genelleme kabiliyeti açısından kritiktir.