

BAŞLAMADAN: Neden L1, L2, Elastic'e ihtiyaç var?

Önce şu soruyu soralım:

"Normal **linear regression** (doğrusal regresyon) bize yetmiyor mu?"

👉 Cevap: **Her zaman yetmez.**

Çünkü:

1. Veride **çok fazla özellik** varsa model kararsızlaşabilir.
2. Model **ezberlemeye** başlar (overfitting).
3. Bazı özellikler **önemsizdir**, model onları sıfırlasın isteriz.
4. Bazen veride **çoklu bağlantı** (multicollinearity) olur → model sapıtır.

💡 İşte bu problemleri çözmek için **L1, L2 ve ElasticNet** gibi yöntemler kullanılır. Bunlara genel olarak "**Regularization (Düzenleme)**" denir.

1. L2 Regresyon (Ridge Regression)

✓ **Mantığı:**

Normal regresyon formülüne, **ekstra bir ceza terimi** ekleriz. Amaç:

"Model çok büyük katsayılar kullanmasın."

$$\text{Hata} + \lambda \sum w^2$$

- λ : Ceza gücü (hyperparametre)
- w^2 : Katsayıların kareleri

👉 Yani: Katsayıları çok büyütürsen **ceza verilir** → Model daha "düzgün" kalır.

📦 **Ne İşe Yarar?**

- Özellikler (değişkenler) arasında **benzer ama küçük etkiler** varsa onları korur.

- Özellikleri sıfırlamaz ama **küçültür**.
- **Overfitting'i** azaltır.

2. L1 Regresyon (Lasso Regression)

✓ **Mantığı:**

Ceza olarak bu kez **kare yerine mutlak değerini** ekleriz:

$$\text{Hata} + \lambda \sum |w|$$

✨ **Farkı nedir?**

- Bu ceza türü, bazı katsayıları **direkt sıfır yapabilir**.
- Yani model bazı özellikleri **tamamen eleyebilir**.

👉 Bu yüzden **özellik seçimi** (feature selection) yapar!

📦 **Ne İşe Yarar?**

- Özellik sayısı çok fazla ve bazılarının etkisi yoksa **gereksizleri sıfırlar**.
- **Daha sade modeller** üretir.
- Ama bazen çok agresif olabilir (iyi özellikleri de silebilir).

3. ElasticNet (L1 + L2 birleşimi)

✓ **Mantığı:**

L1 ve L2 yöntemlerini **birleştirir**. Hem:

- $\sum w^2$ (L2 cezası)
- $\sum |w|$ (L1 cezası)


birlikte kullanılır:

$$\text{Hata} + \alpha(r \sum |w| + (1 - r) \sum w^2)$$

-
- α : Ceza gücü
- r : L1 ve L2 arasındaki denge

Ne İşe Yarar?

- Hem **bazı özellikleri sıfırlar** (L1 gibi),
- Hem de kalanları **yumuşakça küçültür** (L2 gibi).

 Çok değişkenli, karmaşık veri setlerinde genellikle en iyi sonucu verir.

Basit Örnekle Anla:

Diyelim ki bir evin fiyatını tahmin ediyorsun.

Özellik	Etkisi
Oda sayısı	Gerçekten önemli
Bahçe boyutu	Orta derecede önemli
Evin rengi	Alakasız
Mahalle puanı	Önemli

- **L2 (Ridge)**: Hepsine küçük küçük katsayılar verir (rengi de dahil).
- **L1 (Lasso)**: Renk gibi önemsiz olanı tamamen sıfırlar.
- **ElasticNet**: Rengi sıfırlar, kalanlara yumuşak ceza verir.