

Day 18 – ML Theory

Decision Threshold & Cost-Sensitive Evaluation

17. gün: "Modeli nasıl ölçerim?" **18. gün:** "Modelle nasıl karar veririm?"

Bu ayırım **mülakatta seni ayıran çizgi**.

1 Decision Threshold Nedir?

Çoğu classification modeli:

- Doğrudan **0 / 1** üretmez
- **Olasılık (probability)** üretir

Örnek:

$$P(y=1) = 0.73$$

🔗 Bu değeri **1 mi 0 mı** yapacağımıza karar veren şey:

Decision Threshold

Varsayılan:

$$\text{threshold} = 0.5$$

Ama ⚠️:

0.5 = teknik varsayım, business kararı değil

2 Threshold Değişirse Ne Olur?

Threshold ↓ (örn: 0.3)

- Daha fazla **1** tahmini
- Recall ↑
- Precision ↓
- False Positive ↑

Threshold ↑ (örn: 0.7)

- Daha az **1** tahmini
- Precision ↑

- Recall ↓
- False Negative ↑

🔗 **Bu bir matematik değil, iş kararıdır.**

3 False Positive vs False Negative (İŞİN KALBI)

Hata	Anlamı
False Positive (FP)	Gerek yokken alarm
False Negative (FN)	Kritik durumu kaçırma

🔄 Örnek Senaryolar

🏥 Sağlık (Kanser Tarama)

- FN = Ölüm riski
- FP = Gereksiz test

➡ **Recall odaklı**, threshold ↓

🏦 Credit Risk

- FN = Batacak müşteriyi kaçırmak
- FP = İyi müşteriyi reddetmek

➡ FN daha pahalı → threshold ↓

📧 Spam Filter

- FP = Önemli mail spam'e düşer
- FN = Spam inbox'a gelir

➡ FP daha pahalı → threshold ↑

4 Cost-Sensitive Thinking (ÇOK ÖNEMLİ)

Accuracy şunu sorar:

“Kaç tane doğru bildim?”

Business şunu sorar:

“Kaç para kaybettim?”

Örnek Cost Matrix

Gerçek / Tahmin	1	0
Gerçek 1	0	-1000₺
Gerçek 0	-100₺	0

➡ FN maliyeti FP'den **10 kat fazla**

🔗 Böyle bir durumda:

- Accuracy yüksek olsa bile
- FN fazla ise model **kötüdür**

5 Precision–Recall Curve Neden Önemli?

ROC:

- Sınıf oranlarından daha az etkilenir
- Ama **business maliyeti göstermez**

Precision–Recall Curve:

- Imbalanced data'da daha anlamlı
- Threshold seçimini **net** gösterir

🔗 Churn, fraud, risk → **PR Curve > ROC**

6 Gerçek Projede Nasıl Kullanılır?

Tipik akış:

1. Modeli train et
2. Validation set'te:
 - Precision
 - Recall
 - F1
3. Farklı threshold'ları dene
4. Business maliyetine göre seç
5. Test set'te raporla

8 Day 18 – ML Theory Kısa Özet

- Threshold = karar noktasıdır
- 0.5 kutsal değildir
- Precision–Recall trade-off business'a bağlıdır
- Accuracy ≠ Business success

- İyi model = **doğru threshold + doğru maliyet bakışı**
-