

Day 20 – Model Selection & Validation Strategy (Real-World Perspective)

1. Giriş

Makine öğrenmesi projelerinde amaç yalnızca **en yüksek skoru elde etmek** değildir.

Gerçek hayatta asıl soru şudur:

“Hangi modeli production’a almalı ve neden?”

Bu dokümanda, model seçimi ve doğrulama sürecinin **gerçek dünya bakış açısıyla** nasıl ele alınması gerektiği anlatılmaktadır.

2. Model Selection Nedir, Ne Değildir?

✗ Yanlış Yaklaşım

- En yüksek accuracy’ye sahip modeli seçmek
- Tek bir metrik üzerinden karar vermek
- Validation skoruna bakıp modeli sürekli değiştirmek

☑ Doğru Yaklaşım

Model seçimi; **performans, istikrar ve iş etkisinin birlikte değerlendirilmesidir.**

3. Baseline Model’in Önemi

Baseline model:

- Genellikle Logistic Regression veya basit bir karar ağacıdır
- Karmaşık modeller için **referans noktasıdır**

Eğer karmaşık bir model, baseline’dan anlamlı şekilde daha iyi değilse production’a alınması mantıklı değildir.

4. Cross-Validation Sonuçlarını Yorumlama

Cross-validation yalnızca skor üretmez, **model davranışı hakkında bilgi verir.**

4.1 Fold’lar Arası Fark

- Fold skorları birbirine yakınsa → **model stabil**
- Fold skorları çok farklıysa → **yüksek variance**

Bu durumda:

- Daha fazla regularization

- Daha basit model
- Daha fazla veri

gibi önlemler düşünülmelidir.

5. Validation Set vs Test Set

Validation Set

- Model ve hyperparameter seçimi için kullanılır
- Birden fazla kez bakılabilir

Test Set

- **Sadece en sonda kullanılır**
- Model seçimi için geri dönülmez

Test set'e bakarak model değiştirmek, test leakage'a yol açar.

6. Metric, Threshold ve Calibration Birlikte Düşünülmelidir

6.1 Metric Tek Başına Yeterli Değildir

- ROC AUC iyi olabilir
- Ancak precision veya recall iş ihtiyacını karşılamayabilir

6.2 Threshold Kararı

- Varsayılan 0.5 her zaman doğru değildir
- İş maliyetine göre threshold ayarlanmalıdır

6.3 Calibration

- Model olasılıkları güvenilir mi?
- %70 dediğinde gerçekten %70 mi churn oluyor?

Calibration bozursa:

- Threshold kararları yanıltıcı olur
- İş kararları yanlış alınabilir

7. Gerçek Hayatta Model Seçim Kriterleri

Bir modeli production'a almadan önce şu sorular sorulmalıdır:

- Model performansı stabil mi?
- Baseline'a anlamlı katkı sağlıyor mu?
- Açıklanabilir mi?
- Veri drift olduğunda ne olur?

- Yanlış tahminin iş maliyeti nedir?

En iyi model, her zaman en karmaşık olan değildir.

8. Yaygın Hatalar

- Validation skorunu optimize ederken overfitting yapmak
- Test set'i erken kullanmak
- Sadece accuracy'ye bakmak
- Calibration'ı göz ardı etmek
- Modeli iş bağlamından kopuk değerlendirmek

9. Mülakat Perspektifi

Bu konu mülakatta genellikle şu soruyla gelir:

“How do you decide which model to deploy?”

Beklenen cevap:

- Tek metrik değil, çoklu değerlendirme
- İş etkisi vurgusu
- Stabilitate ve risk bilinci

10. Özet

- Model selection bir **optimizasyon problemi değil**, bir **karar problemidir**
- Validation ve test farklı amaçlara hizmet eder
- Metric, threshold ve calibration birlikte ele alınmalıdır
- Gerçek hayatta basit ve stabil modeller çoğu zaman daha değerlidir

11. Kapanış

Day 20 itibarıyla:

- Model değerlendirme süreci bütünsel olarak ele alınmıştır
- Amaç, yalnızca model eğitmek değil **doğru modeli seçebilmektir**