

Alarm eşiği: precision, recall ve işletme maliyeti

Tek bir doğruluk yüzdesi, seyrek arızalarda neden yanıltıcıdır?

Kısaca

- Yalnızca on arıza olan bin gözlemde her zaman normal diyen model yüzde 99 doğruluk gösterir ama hiçbir arızayı yakalamaz.
- Precision verilen uyarıların ne kadarının doğru olduğunu, recall ise gerçek olayların ne kadarının yakalandığını gösterir.
- Tek bir arıza etrafında üretilen otuz ardışık uyarı otuz bağımsız başarı sayılmaz, bu yüzden uyarılar olay pencerelerinde gruplanmalıdır.

01 / Paydayı görünür yapın

Bin gözlemin yalnızca onunda arıza varsa, her zaman normal diyen model yüzde 99 doğruluk gösterir ve bütün arızaları kaçıır. Bu örnek gerçek bir saha ölçümü değildir; sınıf dengesizliğinin etkisini açıklayan bir hesap örneğidir. Kabul tablosunda gerçek pozitif, yanlış pozitif, yanlış negatif ve gerçek negatif sayıları bulunsun.

02 / İki soruyu ayırın

Precision, verilen uyarıların ne kadarının doğru olduğunu; recall ise gerçek olayların ne kadarının yakalandığını sorar. Payda sıfırsa sonuç tanımsızdır; yüzde sıfır veya yüz yazmak yerine hesaplanamadığını gösterin. F1 bu iki değeri tek sayıda birleştirir, fakat operasyonel maliyeti kendiliğinden içermez.

03 / Eşiği doğrulamada seçin

Eşiği düşürmek daha çok uyarı üretir; bunun yararı ve maliyeti beraber değerlendirilmelidir. Önerilen pilotta bakım kapasitesi ve kaçırılan olayın etkisi için iki ayrı maliyet yazın. Seçilen eşiği son test dönemine dokunmadan sabitleyin. Kalibre edilmemiş model skorunu arıza ihtimali diye adlandırmayın.

04 / Olay düzeyini raporlayın

Tek arıza için 30 ardışık uyarı üretmek 30 ayrı başarı değildir. Uyarıları olay pencerelerinde gruplayın ve tekrar bildirim politikasını belirtin. Vardiya başına yanlış alarm ve olay başına ilk uyarı zamanı, bakım ekibinin kullanabileceği ölçülerdir. Nokta düzeyi metrikleri de koruyun; iki görünüm farklı soruları yanıtlar.

Uygulayın

Kutup Laboratuvarı eşik deneyinde sürgüyü değiştirin. Uyarı sayısı ile kaçırılan örnekleri aynı anda izleyin.

Kaynaklar ve çevrimiçi rehber

Güncel çevrimiçi rehberi açın

scikit-learn: classification metrics

Önerilen pilot adımları saha doğruluk iddiası içermez. Sentetik örnekler gerçek tesis ölçümü değildir.