

Makine Öğrenmesi ipuçları ve püf noktaları El Kitabı VIP

Afshine AMIDI and Shervine AMIDI

April 30, 2019

Seray Beşer, Ayyüce Kızrak ve Yavuz Kömeçoğlu tarafından çevrilmiştir

Sınıflandırma metrikleri

İkili bir sınıflandırma durumunda, modelin performansını değerlendirmek için gerekli olan ana metrikler aşağıda verilmiştir.

□ **Karışıklık matrisi** – Karışıklık matrisi, bir modelin performansını değerlendirirken daha eksiksiz bir sonuca sahip olmak için kullanılır. Aşağıdaki şekilde tanımlanmıştır:

| | | Tahmini sınıf | |
|--------------|---|--|---|
| | | + | - |
| Gerçek sınıf | + | TP True Positives | FN False Negatives Type II error |
| | - | FP False Positives Type I error | TN True Negatives |

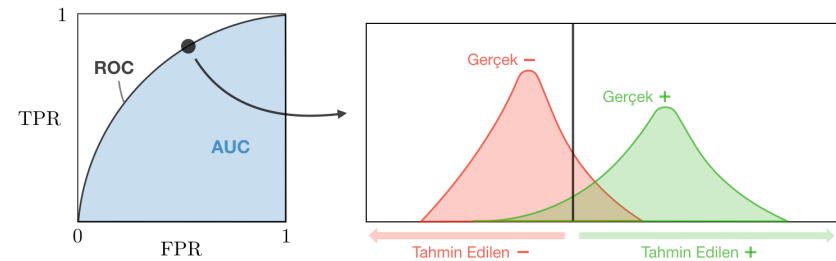
□ **Ana metrikler** – Sınıflandırma modellerinin performansını değerlendirmek için aşağıda verilen metrikler yaygın olarak kullanılmaktadır:

| Metrik | Formül | Açıklama |
|--------------|-------------------------------------|--|
| Doğruluk | $\frac{TP + TN}{TP + TN + FP + FN}$ | Modelin genel performansı |
| Kesinlik | $\frac{TP}{TP + FP}$ | Doğru tahminlerin ne kadar kesin olduğu |
| Geri çağırma | $\frac{TP}{TP + FN}$ | Gerçek pozitif örneklerin oranı |
| Specificity | $\frac{TN}{TN + FP}$ | Gerçek negatif örneklerin oranı |
| F1 skoru | $\frac{2TP}{2TP + FP + FN}$ | Dengesiz sınıflar için yararlı hibrit metrik |

□ **İşlem Karakteristik Eğrisi (ROC)** – İşlem Karakteristik Eğrisi (receiver operating curve), eşik değeri değiştirilerek Doğru Pozitif Oranı-Yanlış Pozitif Oranı grafiğidir. Bu metrikler aşağıdaki tabloda özetlenmiştir:

| Metrik | Formül | Eşdeğer |
|----------------------------|----------------------|---------------|
| True Positive Rate TPR | $\frac{TP}{TP + FN}$ | Geri çağırma |
| False Positive Rate FPR | $\frac{FP}{TN + FP}$ | 1-specificity |

□ **Eğri Altında Kalan Alan (AUC)** – Aynı zamanda AUC veya AUROC olarak belirtilen işlem karakteristlik eğrisi altındaki alan, aşağıdaki şekilde gösterildiği gibi İşlem Karakteristik Eğrisi (ROC)'nin altındaki alandır:



Regresyon metrikleri

□ **Temel metrikler** – Bir f regresyon modeli verildiğinde aşağıdaki metrikler genellikle modelin performansını değerlendirmek için kullanılır:

| Toplam karelerinin toplamı | Karelerinin toplamının açıklaması | Karelerinin toplamından artanlar |
|--|---|---|
| $SS_{\text{tot}} = \sum_{i=1}^m (y_i - \bar{y})^2$ | $SS_{\text{reg}} = \sum_{i=1}^m (f(x_i) - \bar{y})^2$ | $SS_{\text{res}} = \sum_{i=1}^m (y_i - f(x_i))^2$ |

□ **Belirleme katsayısı** – Genellikle R^2 veya r^2 olarak belirtilen belirleme katsayısı, gözlemlenen sonuçların model tarafından ne kadar iyi kopyalandığının bir ölçütüdür ve aşağıdaki gibi tanımlanır:

$$R^2 = 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}}$$

□ **Ana metrikler** – Aşağıdaki metrikler, göz önüne aldıkları değişken sayısını dikkate alarak regresyon modellerinin performansını değerlendirmek için yaygın olarak kullanılır:

| Mallow's Cp | AIC | BIC | Adjusted R ² |
|--|----------------------|---------------------------|----------------------------------|
| $\frac{SS_{\text{res}} + 2(n+1)\hat{\sigma}^2}{m}$ | $2[(n+2) - \log(L)]$ | $\log(m)(n+2) - 2\log(L)$ | $1 - \frac{(1-R^2)(m-1)}{m-n-1}$ |

burada L olabilirlik ve $\hat{\sigma}^2$, her bir yanıtla ilişkili varyansın bir tahminidir.

Model seçimi

□ **Kelime Bilgisi** – Bir model seçerken, aşağıdaki gibi sahip olduğumuz verileri 3 farklı parçaya ayırırız:

| Eğitim seti | Doğrulama seti | Test seti |
|--|--|--|
| <ul style="list-style-type: none"> - Model eğitildi - Genelde veri kümesinin %80'i | <ul style="list-style-type: none"> - Model değerlendirildi - Genelde veri kümesinin %20'si - Ayrıca doğrulama için bir kısmını bekleme veya geliştirme seti olarak da bilinir | <ul style="list-style-type: none"> - Model tahminleri gerçekleştiriyor - Görülmemiş veri |

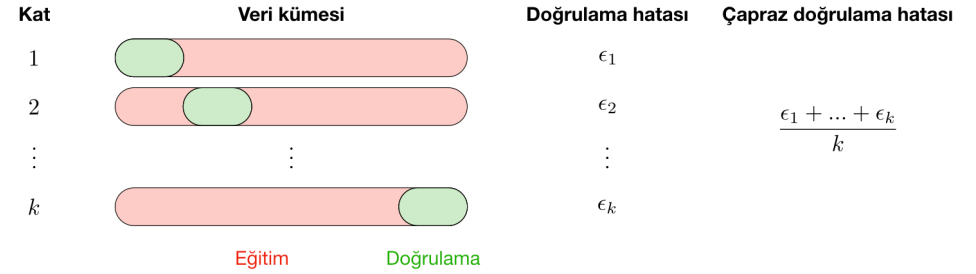
Model bir kere seçildikten sonra, tüm veri seti üzerinde eğitilir ve görünmeyen test setinde test edilir. Bunlar aşağıdaki şekilde gösterilmiştir:



□ **Çapraz doğrulama** – Çapraz doğrulama, başlangıçtaki eğitim setine çok fazla güvenmeyen bir modeli seçmek için kullanılan bir yöntemdir. Farklı tipleri aşağıdaki tabloda özetlenmiştir:

| k -fold | Leave- p -out |
|---|---|
| <ul style="list-style-type: none"> - $k - 1$ katı üzerinde eğitim ve geriye kalanlar üzerinde değerlendirme - Genel olarak $k = 5$ veya 10 | <ul style="list-style-type: none"> - $n - p$ gözlemleri üzerine eğitim ve kalan p üzerinde değerlendirme - Durum $p = 1$'e bir tanesini dışarıda bırak denir |

En yaygın olarak kullanılan yöntem k -kat çapraz doğrulama olarak adlandırılır ve $k - 1$ diğer katlarda olmak üzere, bu k sürelerinin hepsinde model eğitimi yapılırken, modeli bir kat üzerinde doğrulamak için eğitim verilerini k katlarına ayırır. Hata için daha sonra k -katlar üzerinden ortalama alınır ve çapraz doğrulama hatası olarak adlandırılır.



□ **Düzenleştirme (Regularization)** – Düzenleştirme prosedürü, modelin verileri aşırı öğrenmesinden kaçınılmasını ve dolayısıyla yüksek varyans sorunları ile ilgilenmeyi amaçlamaktadır. Aşağıdaki tablo, yaygın olarak kullanılan düzenleştirme tekniklerinin farklı türlerini özetlemektedir:

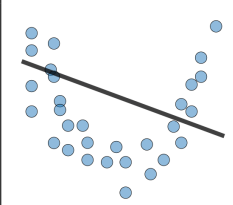
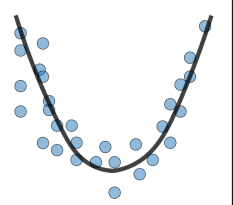
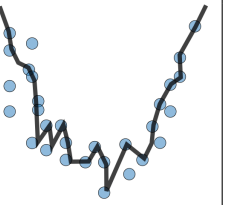
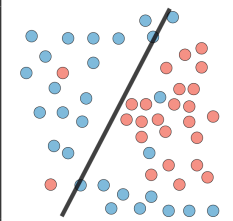
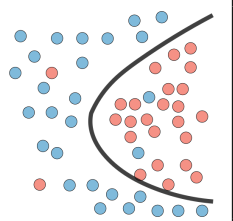
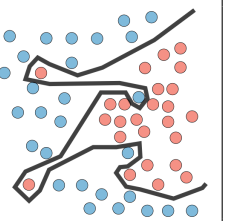
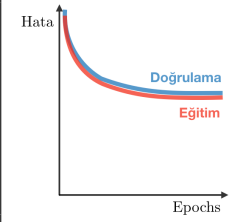
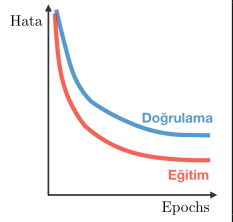
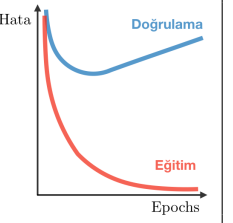
| LASSO | Ridge | Elastic Net |
|---|--|--|
| <ul style="list-style-type: none"> - Değişkenleri 0'a kadra küçült - Değişken seçimi için iyi | Katsayıları daha küçük yap | Değişken seçimi ile küçük katsayılar arasındaki çelişki |
| | | |
| $\dots + \lambda \ \theta\ _1$ $\lambda \in \mathbb{R}$ | $\dots + \lambda \ \theta\ _2^2$ $\lambda \in \mathbb{R}$ | $\dots + \lambda \left[(1-\alpha)\ \theta\ _1 + \alpha\ \theta\ _2^2 \right]$ $\lambda \in \mathbb{R}, \alpha \in [0,1]$ |

Tanı

□ **Önyargı** – Bir modelin önyargısı, beklenen tahmin ve verilen veri noktaları için tahmin etmeye çalıştığımız doğru model arasındaki farktır.

□ **Varyans** – Bir modelin varyansı, belirli veri noktaları için model tahmininin değişkenliğidir.

□ **Önyargı/varyans çelişkisi** – Daha basit model, daha yüksek önyargı, ve daha karmaşık model, daha yüksek varyans.

| | Underfitting | Just right | Overfitting |
|---------------|---|--|---|
| Belirtiler | <ul style="list-style-type: none"> - Yüksek eğitim hatası - Test hatasına yakın eğitim hatası - Yüksek önyargı | <ul style="list-style-type: none"> - Eğitim hatasından biraz daha düşük eğitim hatası | <ul style="list-style-type: none"> - Çok düşük eğitim hatası - Eğitim hatası test hatasının çok altında - Yüksek varyans |
| Regresyon |  |  |  |
| Sınıflandırma |  |  |  |
| Derin öğrenme |  |  |  |
| Olası çareler | <ul style="list-style-type: none"> - Model karmaşıklaştığında - Daha fazla özellik ekle - Daha uzun eğitim süresi ile eğit | | <ul style="list-style-type: none"> - Düzenleştirme gerçekleştir - Daha fazla bilgi edin |

□ **Hata analizi** – Hata analizinde mevcut ve mükemmel modeller arasındaki performans farkının temel nedeni analiz edilir.

□ **Ablatif analiz** – Ablatif analizde mevcut ve başlangıç modelleri arasındaki performans farkının temel nedeni analiz edilir.