1. **Teorik Açıklamalar:**

**Doğruluk (Accuracy):**

* Yanıltıcı olmaması amacıyla sınıfların dengeli olduğu durumlarda tercih edilir.
* Bu metrik, tüm tahminlerin toplamına göre ne kadar doğru tahmin yapıldığını gösterir.

**Hassasiyet (Precision):**

* Modelin doğru olarak tahmin ettiği değerler içerisinden, yanlış olan değerleri minimize etmek istendiği durumda kullanılır.
* Yüksek hassasiyetli modelleri daha dengeler.

**Duyarlılık (Recall):**

* Yanlış negatiflerin çok tehlikeli olduğu durumlarda özellikle önemlidir. Negatif değerin pozitif olarak tahmin edilmemesi istendiği durumlarda kullanılır.
* Pozitif örneklerin kaçırılmaması için yüksek duyarlılık sağlar.

**F1 Skoru:**

* Precision ve Recall arasında denge gerektiren veri setlerinde kullanılır.
* Dengesiz veri setlerinde model performansını daha dengeli olarak değerlendirmemizi sağlar.

**Kesinlik-Hatırlama Eğrisi (Precision-Recall Curve)**

* Negatif verilerin çok fazla olduğu durumlarda kullanılır. Grafik olduğundan dolayı formülü yoktur.
* ROC eğrisinden farkı, sınıf dengesizliğine daha duyarlı olmasıdır.

**ROC Eğrisi**

* Modelin tahmin olasılıkları için farklı eşik değerleri kullanılarak çizilen bir grafiktir.
* ROC eğrisi, modelin pozitif ve yanlış pozitif oranını kıyaslayarak ayırt etme yeteneği sağlar.

**AUC-ROC:**

* ROC eğrisinin altında kalan alanın sayısal değeridir. 0-1 arasındadır.
* AUC-ROC değeri arttıkça sınıfları birbirinden ayırma oranı da artmaktadır.

**Hata Matrisi (Confusion Matrix)**

* TP, TN, FP, FN değerlerini gösteren bir tablodur.
* Modelin hangi hataları yaptığını detaylı analiz etmek için kullanılır.

metin, ekran görüntüsü, yazı tipi, sayı, numara içeren bir resim

Yapay zeka tarafından oluşturulan içerik yanlış olabilir.

**Regresyon Metrikleri**

**Ortalama Mutlak Hata (Mean Absolute Error – MAE):**

* Verideki outlier değerlerinden etkilenmez. Tüm hatalar eşit önemdeyse tercih edilir.
* MAE, ortalama hata büyüklüklerini doğrudan ifade ettiğinden dolayı, yorumlaması kolaydır.

**Ortalama Kare Hata (Mean Squared Error – MSE):**

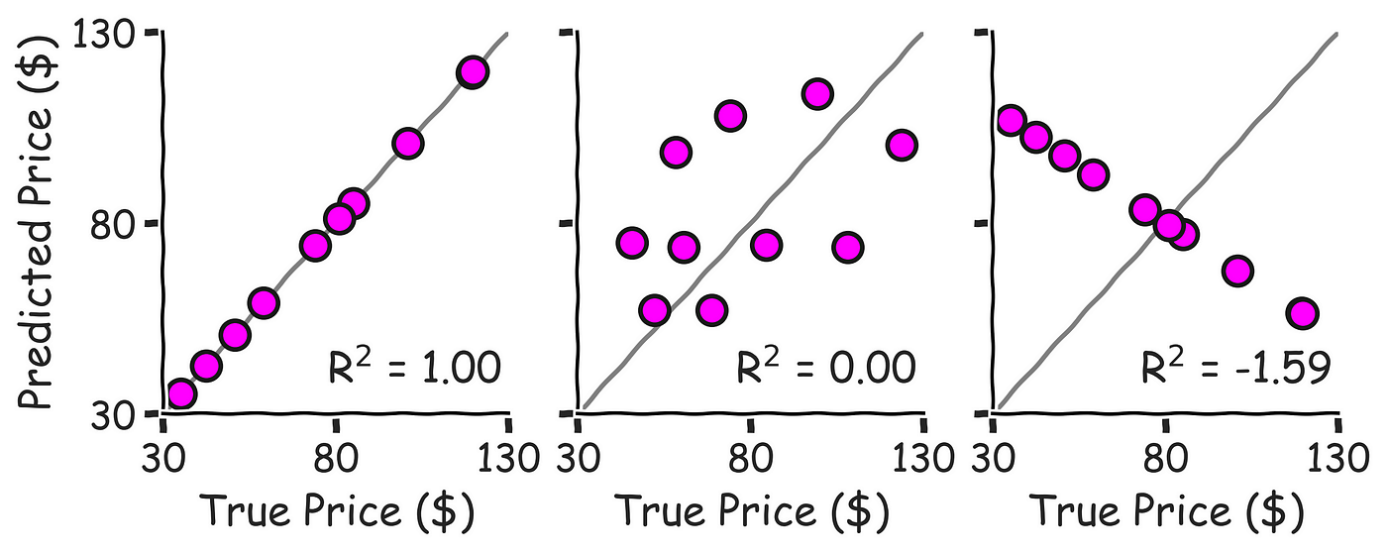
* Büyük hatalara daha fazla ağırlık verilmesi gereken durumlarda tercih edilir.
* Kare alındığından dolayı büyük hatalara daha fazla ağırlık verilir.

**Kök Ortalama Kare Hata (Root Mean Squared Error – RMSE):**

* Büyük hataları öne çıkarırken, hatayı gerçek verinin ölçü birimiyle ifade eder.

**R-kare (R²**) =

* Modelin veriyi ne kadar iyi açıkladığını ölçer. 0-1 arasında değer alır.
* R², modelin bağımlı değişkendeki değişimi ne kadar iyi açıkladığını yüzdesel olarak gösterir



1. **Örnek Senaryolar ve Yorumlama:**

**Örnek**: Spam e-posta tespitinde ise yanlış pozitiflerin (gerçek e-postaların yanlışlıkla spam olarak işaretlenmesi) maliyeti yüksektir, çünkü önemli e-postalar kullanıcıdan kaçabilir. Bu yüzden Precision öncelikli olabilir. Ancak tüm spam e-postalarının doğru bir şekilde yakalanması da çok önemlidir. Bu durumda Recall, yani yanlış negatiflerin (spam e-postalarının normal olarak sınıflandırılması) engellenmesi de kritik hale gelir. F1 Skoru, bu iki metrik arasındaki dengeyi sağlar ve daha güvenilir bir değerlendirme sunar.

* **Kanser teşhisinde yüksek doğruluk oranı neden yanıltıcı olabilir?**

Örneğin, kanser vakalarının çok düşük olduğu bir veri setinde model, çoğunluğu oluşturan "sağlıklı" hastaları doğru bir şekilde tahmin ederek yüksek doğruluk sağlayabilir. Ancak bu durumda kanserli hastaları hiç tespit edemez ve yanlış negatifler (kanserli birinin sağlıklı olarak sınıflandırılması) hayati riskler doğurur. Bu yüzden, doğruluk (accuracy) yerine Recall ve F1 Skoru gibi metrikler daha anlamlıdır.

**Precision mı Recall mı daha önemli?**

**Precision:** Yanlış pozitiflerin yüksek maliyete yol açtığı durumlarda precision daha önemli hale gelir. Örneğin, spam tespitinde, yanlış pozitifleri azaltmak ve gerçek e-postaların kaybolmamasını sağlamak önemlidir.

**Recall:** Yanlış negatiflerin ciddi sonuçlar doğurabileceği durumlarda ise recall daha ön planda olmalıdır. Kanser teşhisi gibi kritik durumlarda, hasta olanları (yanlış negatifleri) kaçırmamak çok daha hayati bir öneme sahiptir.

**Mean Absolute Error (MAE) mı yoksa Root Mean Squared Error (RMSE) mi daha önemli?**

**Mean Absolute Error (MAE):** Hataların ortalama mutlak değeri, hataların büyüklüğünü doğrudan ölçer ve yorumlanması oldukça kolaydır. MAE, aykırı değerlerden RMSE kadar etkilenmez, dolayısıyla hata dağılımı hakkında daha stabil bir fikir verir.

**Root Mean Squared Error (RMSE):** Hataları karesini alarak büyük hatalara daha fazla ağırlık verir. Bu, büyük hataların daha fazla sorun yaratacağı ve daha ağır cezalandırılması gerektiği durumlarda RMSE’yi daha anlamlı kılar.

1. **Karşılaştırmalı Analiz**

**Spam Tespitinde F1 Skoru ve Doğruluk Arasındaki Fark**

**Doğruluk (Accuracy):** Doğruluk, genel başarıyı ölçerken, eğer veri seti dengesizse doğruluk yanıltıcı olabilir.Eğer spam olmayan e-postalar çoğunluktaysa, model tüm e-postaları "spam değil" olarak tahmin edebilir ve yüksek doğruluk oranı elde edebilir. Ancak bu durumda model, spam e-postalarını tespit edemez. Yani, doğruluk oranı yüksek olsa da modelin performansı yanıltıcı olabilir.

**F1 Skoru:** F1 skoru, Precision ve Recall değerlerini birleştirerek dengeli bir performans ölçümü sağlar. Dengesiz veri setlerinde f1 skoru, accuracy değerinden daha faydalı ve bilgilendiricidir.

**Ev fiyatı tahmininde RMSE mi, MAE mi tercih edilmeli?**

**RMSE:** Eğer modelin tahmin ettiği değerlerin bütçemiz aralığında olmasını istiyorsak ve bütçemizi aşmamasını istiyorsak, büyük hataları azaltmak amacıyla RMSE kullanmalıyız.

**MAE**: Seçtiğimiz aralığa göre hataların ortalama büyüklüğünü istiyorsak ve ortalama olarak bütçemize yakın değerler istiyorsak MAE daha uygundur.

**Finansal Tahminlerde Hangi Metrik Daha Anlamlıdır?**

Finansal tahminlerde hata payının düşük olması gerekir çünkü büyük tahmin hataları ciddi maddi kayıplara ve risklere yol açabilir. Bu yüzden, kullanılan modelin hata oranının olabildiğince düşük tutulması önemlidir. RMSE (Root Mean Squared Error) metriği, hataların karesini aldığı için büyük hatalara daha fazla ağırlık verir. Eğer küçük hata sapmalarına odaklanmak ve büyük hataları minimize etmek istiyorsak, RMSE bu hataları belirgin hale getirerek modelin performansını daha hassas bir şekilde değerlendirmemizi sağlar.

**KAYNAKÇA**

<https://towardsdatascience.com/comparing-robustness-of-mae-mse-and-rmse-6d69da870828/>

<https://medium.com/data-science/accuracy-precision-recall-or-f1-331fb37c5cb9>

<https://www.researchgate.net/profile/Shahzad-Ali-8/publication/332429410_Evaluating_Performance_of_Software_Defect_Prediction_Models_Using_Area_Under_Precision-Recall_Curve_AUC-PR/links/5d5e91b3a6fdcc55e81f5d4e/Evaluating-Performance-of-Software-Defect-Prediction-Models-Using-Area-Under-Precision-Recall-Curve-AUC-PR.pdf>

[https://peerj.com/articles/cs-623/#](https://peerj.com/articles/cs-623/)

<https://assets-eu.researchsquare.com/files/rs-3376397/v1/9d39ff47-a2a3-4505-b901-78096bad3d98.pdf?c=1696009342>

https://glassboxmedicine.com/2019/02/17/measuring-performance-the-confusion-matrix/