

RMSE (Root Mean Square Error) ve RMSLE (Root Mean Square Logarithmic Error), istatistiksel ve veri analizi alanlarında kullanılan hata ölçümleri veya metrikleridir. Her ikisi de tahminlerin gerçek değerlerle ne kadar uyumlu olduğunu değerlendirmek için kullanılırlar.

1. \*\*RMSE (Root Mean Square Error)\*\*:

RMSE, tahminlerin gerçek değerlere olan ortalama kare farkını ölçer. RMSE hesaplama formülü şu şekildedir:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

- $n$ , veri noktalarının sayısını temsil eder.
- $y_i$ , gerçek değeri temsil eder.
- $\hat{y}_i$ , tahmin edilen değeri temsil eder.

RMSE, büyük hataların daha fazla vurgulanmasını sağlar ve bir tahmin modelinin performansını değerlendirmek için sıkça kullanılır. Değer ne kadar düşükse, model o kadar iyidir.

Örnek: Bir ev fiyat tahminleme modeli kullanarak bir dizi evin fiyatlarını tahmin ediyorsunuz ve RMSE 10.000 TL olarak hesaplandı. Bu, tahminlerin gerçek fiyatlardan ortalama olarak 10.000 TL sapma gösterdiği anlamına gelir.

2. \*\*RMSLE (Root Mean Square Logarithmic Error)\*\*:

RMSLE, tahminlerin ve gerçek değerlerin logaritmaları alındıktan sonra hesaplanan RMSE'nin bir türevidir. RMSLE hesaplama formülü şu şekildedir:

$$RMSLE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\log(y_i + 1) - \log(\hat{y}_i + 1))^2}$$

- $n$ , veri noktalarının sayısını temsil eder.
- $y_i$ , gerçek değeri temsil eder.
- $\hat{y}_i$ , tahmin edilen değeri temsil eder.

RMSLE, genellikle özellikle regresyon problemleri için kullanılır ve büyük hataların daha fazla vurgulanmasını sağlar. RMSLE'nin logaritmik ölçek kullanmasının nedeni, tahminlerin ve gerçek değerlerin büyük farklarının daha az cezalandırılmasını sağlamaktır. Bu, verilerin geniş bir aralığa sahip olduğu durumlarda faydalı olabilir.

Örnek: Bir satış tahminleme modeli kullanarak bir dizi ürünün satışlarını tahmin ediyorsunuz ve RMSLE 0.2 olarak hesaplandı. Bu, tahminlerin gerçek satışlardan logaritmik ölçekte 0.2 birim sapma gösterdiği anlamına gelir.

Bu metrikler, tahmin modelinin performansını değerlendirmek için kullanılırken, hangi metriğin kullanılacağı veri ve problem türüne bağlı olarak değişebilir.