

○ 回帰モデルにおける代表的な予測誤差集計方法

(1) R^2 (決定係数)

R^2 (決定係数) は、予測誤差を正規化 (数値の大きさなどをそろえること) することで得られる指標で、まったく予測できていない場合を0、すべて予測できている場合を1として大きいほどよい性能であることを示します。予測する値のスケールによらない指標であるため、直観的でわかりやすい指標といえます。

(2) RMSE (平方平均二乗誤差)

RMSE (平方平均二乗誤差) は、予測誤差を二乗して平均したあとに集計する指標で、小さいほどよい性能であることを示します。正規分布の誤差に対して正確な評価ができるため、多くのケースで使われています。 R^2 (決定係数) とは異なり、たとえば予測値が個数であれば指標のスケールも個数となるように、得られる値がそのまま予測値の単位になるため、モデルの具体的な評価がしやすい指標としても知られています。

(3) MAE (平均絶対誤差)

MAE (平均絶対誤差) は、予測誤差の絶対値を平均したあとに集計する指標で、小さいほどよい性能であることを示します。RMSEと比較して外れ値 (通常の誤差よりもかなり大きい誤差をもつ値) に強いため、多くの外れ値が存在するデータセットで評価をする場合に利用されます。また、RMSEと同様に得られる値がそのまま予測値の単位であり、モデルの具体的な評価がしやすい指標と言えます。

■ 3つの評価指標

R^2 (決定係数)	RMSE (平方平均二乗誤差)	MAE (平均絶対誤差)
・ 0から1の範囲内の値を取り、1に近いほど分析の精度が高い	・ 正規分布の誤差に対して正確な評価ができる ・ 局所的な誤差に左右されやすい	・ 多くの外れ値が存在するデータセットで評価をする場合に利用される

○ 分類モデルの性能を評価する

続いて、分類モデルです。分類モデルの評価では、回帰モデルの場合と異なり出力と正解が取りうるパターンが複数考えられるため、それらのパターンを**混合行列**という表にすることが基本となります。たとえば「○」と「×」のラベルを分類する問題であれば、出力と正解が取りうるパターンは2×2で4パターンとなるため、混合行列は下図のように2×2の4マスとなります。

■ 混合行列

	正解が「○」	正解が「×」
「○」と予想	TP	FP
「×」と予想	FN	TN

T: True (予想が当たっている)
F: False (予想が当たっていない)
P: Positive (「○」)
N: Negative (「×」)

2ラベル分類の混合行列の4パターンはそれぞれ、正解が「○」であるものを正しく「○」と予想した回数である**TP (True Positive: 真陽性)**、正解が「○」であるものを間違って「×」と予想した回数である**FN (False Negative: 偽陰性)**、正解が「×」であるものを間違って「○」と予想した回数である**FP (False Positive: 偽陽性)**、正解が「×」であるものを正しく「×」と予想した回数である**TN (True Negative: 真陰性)**と呼ばれます。混合行列のTP、FP、FN、TNの値をベン図に置き換えると下図のようになります。学習させた分類モデルは、青い線で囲った「本当の境界 (データを正しく分けている境界)」を目標としますが、完璧には分けられず、ピンクの線で囲った「推測した境界」のように分類します。緑色で示したデータについては正しい分類ができていますが、橙色で示したデータについては正答とは異なる分類をしています。

■ ベン図で見る混合行列の値

