

分類問題の評価指数(2値問題)

混交行列(Confusion Matrix)とTP, FP, FN, TN

	実際は正 (Positive)	実際は負 (Negative)
予測が正 (Positive)	TP True Positive	FP False Positive 第1種の誤り
予測が負 (Negative)	FN False Negative 第2種の誤り	TN True Negative

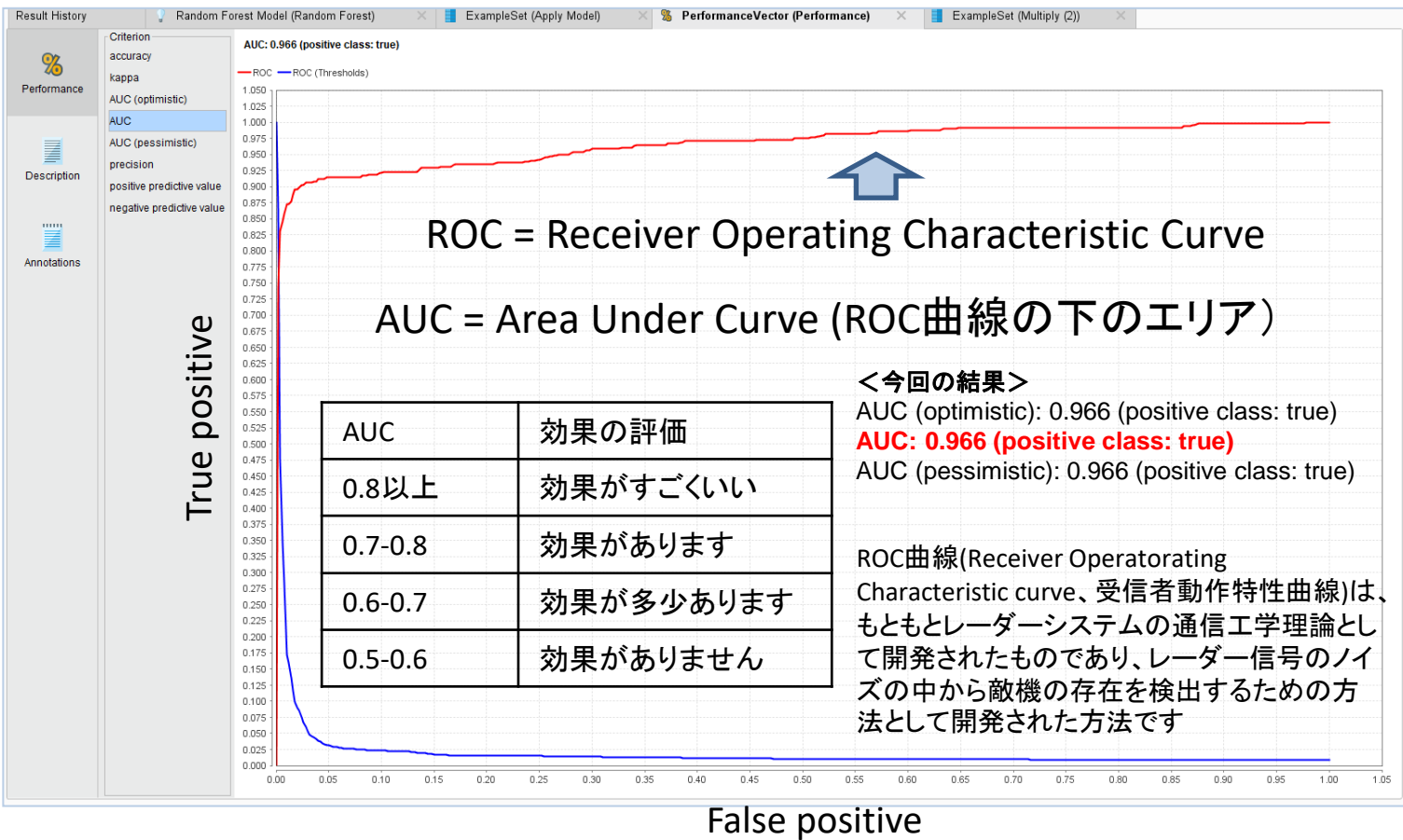
適合率と再現率はトレードオフの関係
両者のバランスを評価するのがF1値

参考:

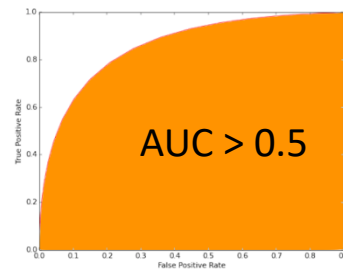
<https://qiita.com/FukuharaYohei/items/be89a99c53586fa4e2e4>

指標	意味	式
正解率(Accuracy)	全予測 正答率	$\frac{TP + TN}{TP + FP + FN + TN}$
適合率(Precision) 感度(Sensitivity)	正予測 の正答率	$\frac{TP}{TP + FP}$
再現率(Recall)	正に対する 正答率	$\frac{TP}{TP + FN}$
特異率(Specificity)	負に対する 正答率	$\frac{TN}{FP + TN}$
F値(F-measure)	適合率と 再現率の 調和平均	$\frac{2 \times \text{適合率} \times \text{再現率}}{\text{適合率} + \text{再現率}}$

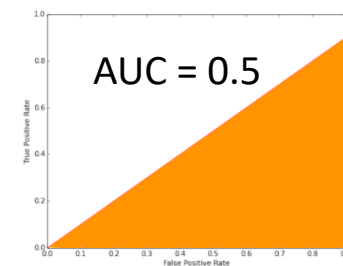
ROC曲線とAUC



分類器の効果



効果がある分類器



効果がない分類器