

# Metrics to Evaluate ML/DL Algorithms

## Classification Performance Metrics

- Confusion Matrix

Confusion Matrix		Actual	
		Positive	Negative
Predicted	Positive	a	b
	Negative	c	d

- Sensitivity & Specificity

Sensitivity (recall, hit rate, true positive rate (TPR)):  $\frac{a}{a+c}$

Specificity (selectivity, true negative rate (TNR) ):  $\frac{d}{b+d}$

Precision (positive predictive value):  $\frac{a}{a+b}$

negative predictive rate:  $\frac{d}{c+d}$

Accuracy:  $\frac{a+d}{a+b+c+d}$

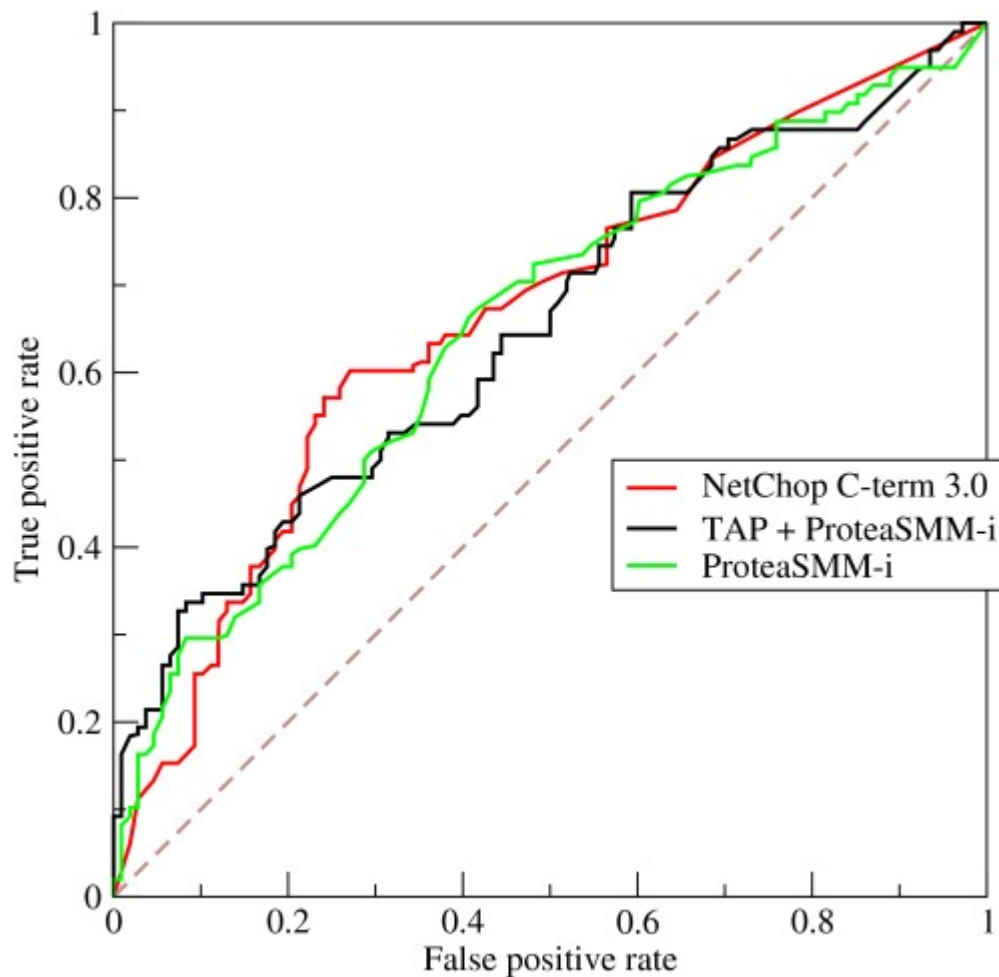
- F1 score

$$F1 = \left( \frac{\text{recall}^{-1} + \text{precision}^{-1}}{2} \right)^{-1} = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$

- ROC & AUC

An ROC curve (receiver operating characteristic curve) is a graph showing the performance of a classification model at all classification thresholds. It plots true positive rate (TPR) vs False positive rate (FPR)

AUC is the area under the ROC curve.



ROC curve (downloaded from wiki [https://en.wikipedia.org/wiki/Receiver\\_operating\\_characteristic](https://en.wikipedia.org/wiki/Receiver_operating_characteristic))

\* ROC is used for identifying the best threshold while the AUC is used for identifying which model is better.

\* Note: people often replace false positive rate with precision.

## Regression Performance Metrics

- Mean Square Error (MSE)
- Root Mean Square Error (RMSE)