Performance Measures

Classification/Regression

Classifier evaluation – Errors

- Let´s consider θ as the real value and $\hat{\theta}$ as the estimated value.

Mean absolute error is:

$$MSE = \frac{1}{N}\sum_{i=1}^{N} |\hat{\theta}_i - \theta_i|$$

Root mean square error is:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \left(\hat{\theta}_i - \theta_i\right)^2}$$

Relative absolute error:

$$RAE = \frac{\sum_{i=1}^{N} |\hat{\theta}_i - \theta_i|}{\sum_{i=1}^{N} |\bar{\theta} - \theta_i|}$$

where $\bar{\theta}$ is a mean value of θ.

Root relative squared error:

$$RRSE = \sqrt{\frac{\sum_{i=1}^{N} \left(\hat{\theta}_i - \theta_i\right)^2}{\sum_{i=1}^{N} \left(\bar{\theta} - \theta_i\right)^2}}$$

Kappa Statistic– a measure of agreement

Let´s consider a binary classification, and the table bellow containing probabilities.

		Rater #1		
		1	2	Total
Rater #2	1	p_{11}	p_{12}	$p_{1.}$
	2	p_{21}	p_{22}	$p_{2.}$
	Total	$p_{.1}$	$p_{.2}$	1

- To compute Kappa, you first need to calculate the observed level of agreement

$$p_o = p_{11} + p_{22}$$

- This value needs to be compared to the value that you would expect if the two raters were totally independent,

$$p_e = p_{.1}p_{1.} + p_{.2}p_{2.}$$

- The value of Kappa is defined as

$$\kappa = \frac{p_o - p_e}{1 - p_e}$$

Possible interpretation of Kappa:

- Poor agreement = Less than 0.20
- Fair agreement = 0.20 to 0.40
- Moderate agreement = 0.40 to 0.60
- Good agreement = 0.60 to 0.80
- Very good agreement = 0.80 to 1.00

Measures derived from the confusion matrix

		prediction outcome		
		p	**n**	**total**
actual value	**p'**	True Positive	False Negative	P'
	n'	False Positive	True Negative	N'
	total	P	N	

true positive (TP)

eqv. with hit

true negative (TN)

eqv. with correct rejection

false positive (FP)

eqv. with false alarm, Type I error

false negative (FN)

eqv. with miss, Type II error

sensitivity or true positive rate (TPR)

eqv. with hit rate, recall

$$TPR = \frac{TP}{P} = \frac{TP}{TP + FN}$$

specificity (SPC) or true negative rate (TNR)

$$SPC = \frac{TN}{N} = \frac{TN}{FP + TN}$$

precision or positive predictive value (PPV)

$$PPV = \frac{TP}{TP + FP}$$

negative predictive value (NPV)

$$NPV = \frac{TN}{TN + FN}$$

fall-out or false positive rate (FPR)

$$FPR = \frac{FP}{N} = \frac{FP}{FP + TN} = 1 - SPC$$

false discovery rate (FDR)

$$FDR = \frac{FP}{FP + TP} = 1 - PPV$$

miss rate or false negative rate (FNR)

$$FNR = \frac{FN}{P} = \frac{FN}{FN + TP}$$

---

accuracy (ACC)

$$ACC = \frac{TP + TN}{P + N}$$

F1 score

is the harmonic mean of precision and sensitivity

$$F1 = \frac{2TP}{2TP + FP + FN}$$

Matthews correlation coefficient (MCC)

$$\frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

Informedness = Sensitivity + Specificity - 1

Markedness = Precision + NPV - 1

Source: Fawcett, Tom (2006). "An Introduction to ROC Analysis". *Pattern Recognition Letters* **27** (8): 861 – 874. doi: 10.1016/j.patrec.2005.10.010.

Correlation coeficient – (for regression)

- The most commonly used is Pearson's product moment correlation.
- A correlation coefficient shows the degree of linear dependence between two variables (x and y). How close two variables lie along a line.
- If the coefficient is equal to 1 or -1, all the points lie along a line. If the correlation coefficient is equal to zero, there is no linear relation between x and y.
- A positive relationship means that the two variables move into the same direction. A higher value of x corresponds to higher values of y, and vice versa.
- A negative relationship means that the two variables move into the opposite directions. A lower value of x corresponds to higher values of y, and vice versa.

Correlation coeficient (r)