

Ensembles for predicting structured outputs

Evaluation measures for classification

Dragi Kocev, Celine Vens, Jan Struyf, Sašo Džeroski

July 19, 2010

1 Performance measures

Nomenclature:

- c_i - class value
- T - number of classes
- TP_i - true positives for class c_i
- FP_i - false positives for class c_i
- FN_i - false negatives for class c_i
- TN_i - true negatives for class c_i
- P_i - precision for class c_i
- R_i - recall for class c_i
- F_i - F-measure for class c_i
- MCC_i - Matthews Correlation Coefficient for class c_i
- $BACC_i$ - Balanced Accuracy for class c_i
- DP_i - Discriminant Power for class c_i
- ω_i - weight for class c_i

Table 1: Evaluation measures - general definitions.

Performance measure	Formula
Precision	$P = \frac{TP}{TP+FP}$
Recall	$R = \frac{TP}{TP+FN}$
F-measure	$F = 2 \cdot \frac{P \cdot R}{P+R}$
MCC	$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP+FP) \cdot (TP+FN) \cdot (TN+FP) \cdot (TN+FN)}}$
BACC	$BACC = \frac{sensitivity+specificity}{2} = \frac{1}{2} \cdot \left(\frac{TP}{TP+FN} + \frac{TN}{TN+FP} \right)$
DP	$DP = \frac{\sqrt{3}}{\pi} \cdot \ln \left(\frac{TP}{FP} \cdot \frac{TN}{FN} \right)$

Note: The formula for *discriminant power* is $BACC = \frac{\sqrt{3}}{\pi} \cdot (\ln X + \ln Y)$, where $X = \frac{sensitivity}{1-sensitivity}$, $Y = \frac{specificity}{1-specificity}$, $sensitivity = \frac{TP}{TP+FN}$ and $specificity = \frac{TN}{TN+FP}$.

1.1 Micro Weighted Average

Table 2: Micro averaged evaluation measures.

Performance measure	Formula
$\mu\text{Precision}$	$P^\mu = \frac{\sum_i^T \omega_i TP_i}{\sum_i^T \omega_i TP_i + \sum_i^T \omega_i FP_i}$
μRecall	$R^\mu = \frac{\sum_i^T \omega_i TP_i}{\sum_i^T \omega_i TP_i + \sum_i^T \omega_i FN_i}$
$\mu\text{F-measure}$	$F^\mu = 2 \cdot \frac{P_i^\mu \cdot R_i^\mu}{P_i^\mu + R_i^\mu}$
μBACC	$BACC^\mu = \frac{1}{2} \cdot \left(\frac{\sum_i^T \omega_i TP_i}{\sum_i^T \omega_i TP_i + \sum_i^T \omega_i FN_i} + \frac{\sum_i^T \omega_i TN_i}{\sum_i^T \omega_i TN_i + \sum_i^T \omega_i FP_i} \right)$
μDP	$DP^\mu = \frac{\sqrt{3}}{\pi} \cdot \ln \left(\frac{\sum_i^T \omega_i TP_i}{\sum_i^T \omega_i FP_i} \cdot \frac{\sum_i^T \omega_i TN_i}{\sum_i^T \omega_i FN_i} \right)$
MCC^μ	$MCC^\mu = \frac{\sum_i^T \omega_i TP_i \cdot \sum_i^T \omega_i TN_i - \sum_i^T \omega_i FP_i \cdot \sum_i^T \omega_i FN_i}{\sqrt{(\sum_i^T \omega_i TP_i + \sum_i^T \omega_i FP_i) \cdot (\sum_i^T \omega_i TP_i + \sum_i^T \omega_i FN_i) \cdot (\sum_i^T \omega_i TN_i + \sum_i^T \omega_i FP_i) \cdot (\sum_i^T \omega_i TN_i + \sum_i^T \omega_i FN_i)}}$

1.2 Macro Weighted Average

Table 3: Macro averaged evaluation measures.

Performance measure	Formula
$M\text{Precision}$	$P^M = \sum_i^T \omega_i P_i$
$M\text{Recall}$	$R^M = \sum_i^T \omega_i R_i$
$M\text{F-measure}$	$F^M = \sum_i^T \omega_i F_i$
$MMCC$	$MCC^M = \sum_i^T \omega_i MCC_i$
$MBACC$	$BACC^M = \sum_i^T \omega_i BACC_i$
MDP	$DP^M = \sum_i^T \omega_i DP_i$