

# Calculating Accuracy, Weighted Precision, Weighted Recall, and Weighted F1-Score for Multi-class Classification

Let  $M$  be a  $n \times n$  confusion matrix for  $n$  classes ( $n \geq 3$ ) where each entry is  $M_{ij}$  for  $i = 0, 1, \dots, n-1$  and  $j = 0, 1, \dots, n-1$ . The rows of  $M$  represent true classes while the columns represent predicted classes. Let  $y\_true$  be a list of  $N$  true classes and  $y\_pred$  be a list of  $N$  corresponding predicted classes. Therefore,  $M_{ij}$  is the number of pairs  $(i, j)$  where  $i = y\_true_k$  and  $j = y\_pred_k$  for all  $k$  where  $k \in \{0, 1, \dots, N-1\}$ . One vs. all multi-class setting is used. For all classes, one class is chosen to represent the positive instance and the rest of the other classes together represent the negative instance. Let  $support_i$  be the number of elements in class  $i$  in  $y\_true$ , so  $support_i = \sum_{j=0}^{n-1} M_{ij}$ . Class  $i$  has a weight ( $w_i$ ) where  $w_i = \frac{support_i}{N}$ . Let  $tp$  be the number of true positives,  $fp$  be the number of false positives, and  $fn$  be the number of false negatives. For class  $i$ ,  $tp_i = M_{ii}$ ,  $fp_i = \sum_{j=0}^{n-1} M_{ji} - M_{ii}$ , and  $fn_i = \sum_{j=0}^{n-1} M_{ij} - M_{ii}$ .

## 1. Accuracy

$$\text{Accuracy } a = \frac{1}{N} \sum_{i=0}^{n-1} tp_i = \frac{1}{N} \sum_{i=0}^{n-1} M_{ii}.$$

## 2. Weighted Precision

$$\text{For class } i, \text{ precision } p_i = \frac{tp_i}{tp_i + fp_i}.$$

$$\text{Thus, } p_i = \frac{M_{ii}}{M_{ii} + \sum_{j=0}^{n-1} M_{ji} - M_{ii}} = \frac{M_{ii}}{\sum_{j=0}^{n-1} M_{ji}}.$$

$$p_{weighted} = \sum_{i=0}^{n-1} w_i p_i = \frac{1}{N} \sum_{i=0}^{n-1} \frac{support_i M_{ii}}{\sum_{j=0}^{n-1} M_{ji}}.$$

## 3. Weighted Recall

$$\text{For class } i, \text{ recall } r_i = \frac{tp_i}{tp_i + fn_i}.$$

$$\text{Thus, } r_i = \frac{M_{ii}}{M_{ii} + \sum_{j=0}^{n-1} M_{ij} - M_{ii}} = \frac{M_{ii}}{\sum_{j=0}^{n-1} M_{ij}} = \frac{M_{ii}}{support_i}.$$

$$r_{weighted} = \sum_{i=0}^{n-1} w_i r_i = \frac{1}{N} \sum_{i=0}^{n-1} M_{ii}.$$

Note that the weighted recall is also the same as accuracy.

## 4. Weighted F1-score

F1-score is the harmonic mean of precision and recall.

For class  $i$ , F1-score  $f_i = \frac{2p_i r_i}{p_i + r_i}$ .

$$\text{Thus, } f_i = \frac{2 \left( \frac{M_{ii}}{\sum_{j=0}^{n-1} M_{ji}} \right) \left( \frac{M_{ii}}{\text{support}_i} \right)}{\frac{M_{ii}}{\sum_{j=0}^{n-1} M_{ji}} + \frac{M_{ii}}{\text{support}_i}} = \frac{2M_{ii}}{\text{support}_i + \sum_{j=0}^{n-1} M_{ji}}.$$

$$f_{\text{weighted}} = \sum_{i=0}^{n-1} w_i f_i = \frac{2}{N} \sum_{i=0}^{n-1} \frac{\text{support}_i M_{ii}}{\text{support}_i + \sum_{j=0}^{n-1} M_{ji}}.$$

## 5. Simple Example

A simple example showing a 3-class ('bird', 'cat', 'dog') classification problem is given below.

$y_{\text{true}} = [\text{'cat'}, \text{'dog'}, \text{'bird'}, \text{'bird'}, \text{'cat'}, \text{'dog'}, \text{'dog'}, \text{'bird'}, \text{'bird'}, \text{'cat'}, \text{'dog'}, \text{'cat'}, \text{'cat'}, \text{'cat'}, \text{'dog'}, \text{'dog'}, \text{'dog'}, \text{'dog'}, \text{'dog'}, \text{'bird'}, \text{'bird'}, \text{'bird'}, \text{'dog'}, \text{'dog'}, \text{'cat'}, \text{'bird'}, \text{'dog'}]$ .

$y_{\text{pred}} = [\text{'bird'}, \text{'dog'}, \text{'bird'}, \text{'cat'}, \text{'cat'}, \text{'bird'}, \text{'dog'}, \text{'bird'}, \text{'bird'}, \text{'dog'}, \text{'dog'}, \text{'cat'}, \text{'cat'}, \text{'cat'}, \text{'cat'}, \text{'dog'}, \text{'dog'}, \text{'dog'}, \text{'dog'}, \text{'dog'}, \text{'bird'}, \text{'bird'}, \text{'bird'}, \text{'cat'}, \text{'dog'}, \text{'dog'}, \text{'bird'}]$ .

These 3 classes ('bird', 'cat', 'dog') are mapped to 0, 1, and 2, respectively.

$y_{\text{true}} = [1, 2, 0, 0, 1, 2, 2, 0, 0, 1, 2, 1, 1, 1, 1, 2, 2, 2, 2, 0, 0, 0, 2, 2, 1, 0, 2]$ .

$y_{\text{pred}} = [0, 1, 0, 1, 1, 0, 2, 0, 0, 2, 2, 1, 1, 1, 1, 2, 2, 2, 2, 2, 0, 0, 0, 1, 1, 2, 2, 0]$ .

$N = 28$ .

The corresponding  $3 \times 3$  confusion matrix is shown below.

$$M = \begin{bmatrix} 6 & 1 & 1 \\ 1 & 5 & 2 \\ 2 & 3 & 7 \end{bmatrix}$$

### (a) Accuracy

$$a = \frac{1}{28}(6 + 5 + 7) = \frac{9}{14} \approx 0.6428571428571429.$$

### (b) Weighted Precision

$$p_{\text{weighted}} = \frac{1}{28} \left( \frac{8(6)}{9} + \frac{8(5)}{9} + \frac{12(7)}{10} \right) = \frac{409}{630} \approx 0.6492063492063492.$$

### (c) Weighted Recall

$$r_{\text{weighted}} = \frac{1}{28} (6 + 5 + 7) = \frac{9}{14} \approx 0.6428571428571429.$$

### (d) Weighted F1-Score

$$f_{\text{weighted}} = \frac{2}{28} \left( \frac{8(6)}{8+9} + \frac{8(5)}{8+9} + \frac{12(7)}{12+10} \right) = \frac{841}{1309} \approx 0.6424751718869367.$$