

Multi Class Log Loss rename

Search

This is the multi-class version of the [Logarithmic Loss](#) metric. Each observation is in one class and for each observation, you submit a predicted probability for each class. The metric is negative the log likelihood of the model that says each test observation is chosen independently from a distribution that places the submitted probability mass on the corresponding class, for each observation.

$$\text{logloss} = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^M y_{i,j} \log(p_{i,j})$$

where N is the number of observations, M is the number of class labels, \log is the natural logarithm, $y_{i,j}$ is 1 if observation i is in class j and 0 otherwise, and $p_{i,j}$ is the predicted probability that observation i is in class j .

Both the solution file and the submission file are CSV's where each row corresponds to one observation, and each column corresponds to a class. The solution has 1's and 0's (exactly one "1" in each row), while the submission consists of predicted probabilities.

The submitted probabilities need not sum to 1, because they will be rescaled (each is divided by the sum) so that they do before evaluation.

(Note: the actual submitted predicted probabilities are replaced with $\max(\min(p, 1 - 10^{-15}), 10^{-15})$.)

```
id,col1,col2,Indicator
1,1,0,Public
2,1,0,Public
3,1,0,Public
4,0,1,Public
5,0,1,Public
6,0,1,Public
7,1,0,Private
```

```
id,col1,col2
1,0.5,0.5
2,0.1,0.9
3,0.01,0.99
4,0.9,0.1
5,0.75,0.25
6,0.001,0.999
7,1,0
```

```
Assert.AreEqual(1.881797068998267, actual: publicScore, delta: ErrorTolerance);
```

language: php

Note: In this example 'publicScore' is the score for only the rows marked as 'Public'. The observation with ID of 7 is ignored in the calculation of 1.881797068998267 (and $N = 6$).

