

Precision, Recall, F1

Precision and recall are commonly used to evaluate binary classification problems. Usually the class of interest is the positive class, and the other classes are the negative class. The prediction of the classifier on the test data set is either correct or incorrect, and the total numbers of the four cases are recorded as:

TP: The number of samples that predict the positive class as a positive class

FN: The number of samples that predict the positive class as a negative class

FP: The number of samples that predict the negative class as a positive class

TN: The number of samples that predict the negative class as a negative class

1 Accuracy

Precision is defined as

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

2 Recall

The recall rate is defined as

$$R = \frac{TP}{TP + FN}$$

3 F1

The F_1 value is the harmonic mean of precision and recall, that is

$$\frac{2}{F_1} = \frac{1}{P} + \frac{1}{R}$$

$$F_1 = \frac{2TP}{2TP + FP + FN}$$

When the precision rate and recall rate are high, the F_1 value will also be high.

References

Li Hang. Statistical Learning Methods. Tsinghua University Press. 2016.