Эксплуатационные ошибки

Терехов Олег

Московский физико-технический институт
Факультет управления и прикладной математики
Кафедра интеллектуальных систем

10 апреля 2019 г.

Precision-Recall

Основные метрики

1. Recall or Sensitivity or TPR (True Positive Rate): Number of items correctly identified as positive out of total true positives- TP/(TP+FN)
2. Specificity or TNR (True Negative Rate): Number of items correctly identified as negative out of total negatives- TN/(TN+FP)
3. Precision: Number of items correctly identified as positive out of total items identified as positive- TP/(TP+FP)
4. False Positive Rate or Type I Error: Number of items wrongly identified as positive out of total true negatives- FP/(FP+TN)
5. False Negative Rate or Type II Error: Number of items wrongly identified as negative out of total true positives- FN/(FN+TP)

F_β-measure

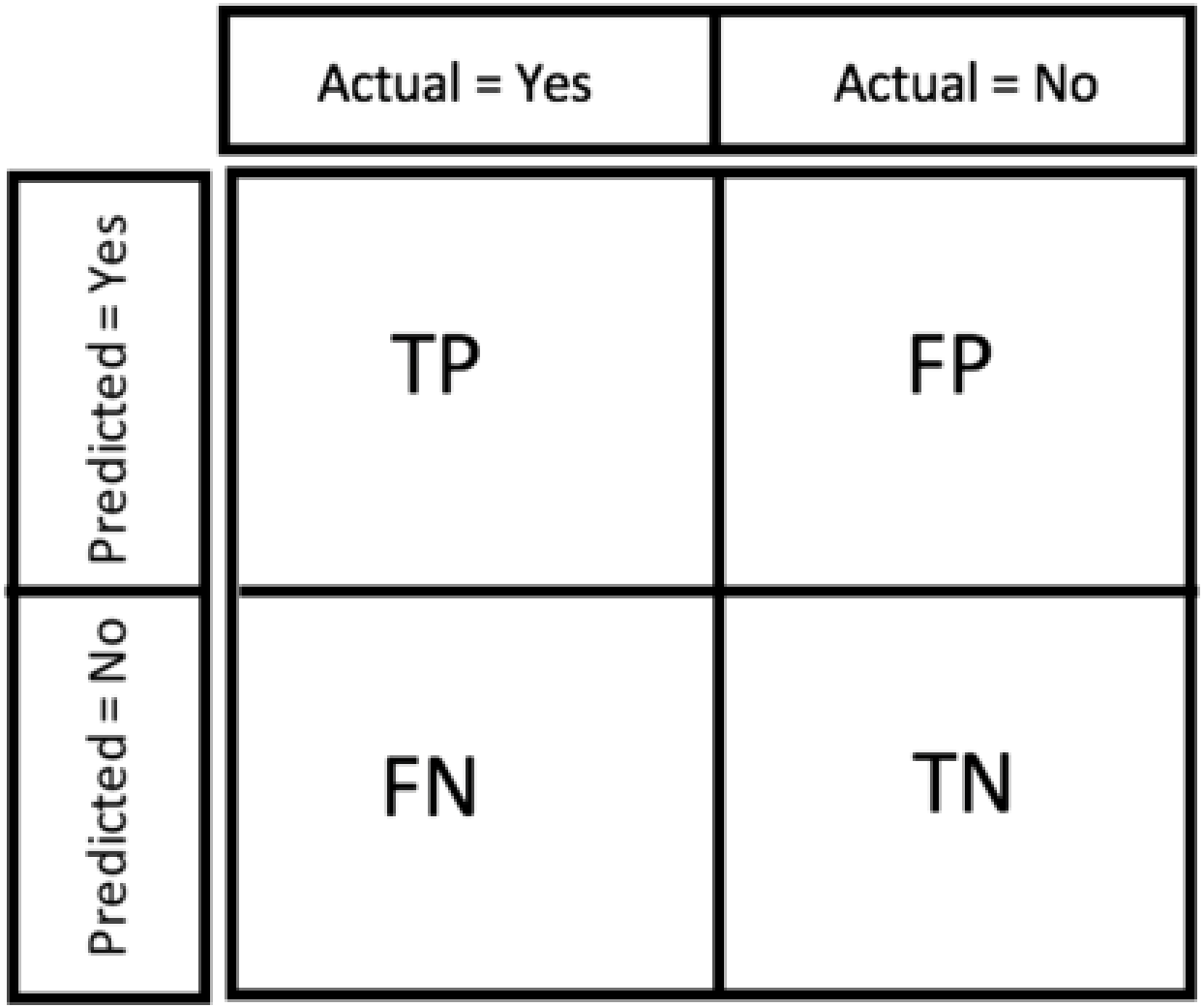
$$F_\beta = (1 + \beta^2)\frac{precision \times recall}{(\beta^2 precision) + recall}$$

The F-measure reaches a maximum with completeness and accuracy equal to one, and is close to zero if one of the arguments is close to zero.

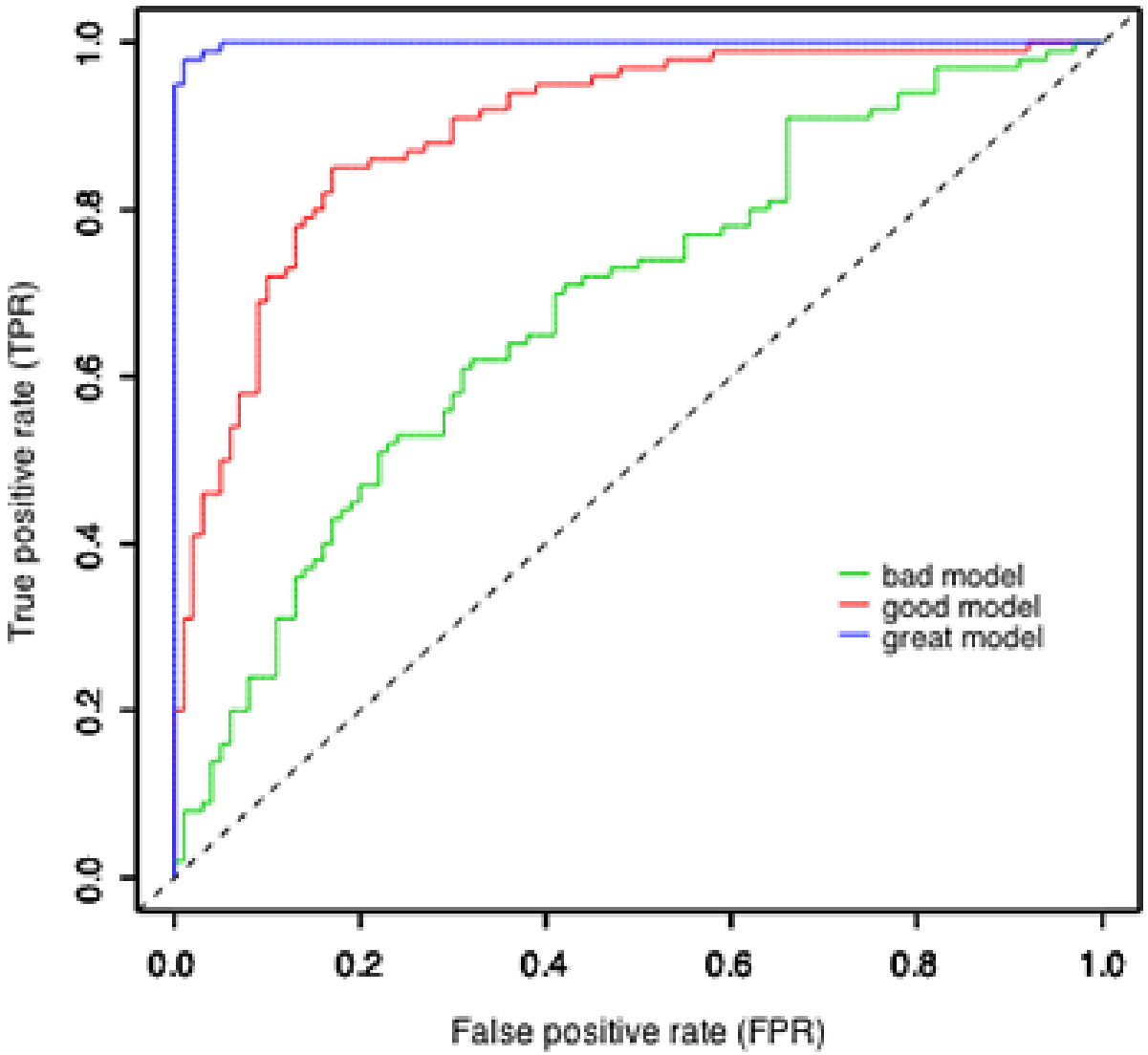
F1 Score: a harmonic mean of precision and recall

$$F1 = 2\frac{Precision \cdot Recall}{Precision + Recall}$$

Confusion Matrix



ROC-curve


True positive rate (TPR)
False positive rate (FPR)
0.0
0.2
0.4
0.6
0.8
1.0
bad model
good model
great model

Accuracy

Accuracy: Percentage of total items classified correctly

$$\text{Accuracy} = \frac{\#\text{correctly classified items}}{\#\text{all classified items}}$$

Log-Loss

Log-loss is a measurement of accuracy that incorporates the idea of probabilistic confidence given by following expression for binary class

$$logloss = -\frac{1}{n}\sum_{i=1}^{n}(y_i \log \hat{y}_i + (1 - y_i) \log (1 - \hat{y}_i))$$

RMSE, MAE

RMSE

It represents the sample standard deviation of the differences between predicted values and observed values (called residuals). Mathematically, it is calculated using this formula:

$$RMSE = \sqrt{1n \times \sum_{j=1}^{n} (y_j - \hat{y}_j)^2}$$

MAE

MAE is the average of the absolute difference between the predicted values and observed value. The MAE is a linear score which means that all the individual differences are weighted equally in the average. For example, the difference between 10 and 0 will be twice the difference between 5 and 0. However, same is not true