evaluation methods for classification

A **Confusion Matrix** is constructed to determine the quality of a **Classification Algorithm**:

	Actual Value $\rightarrow$ $y = 1$	**Actual Value** $\rightarrow$ $y = -1$
Predicted Value $\rightarrow$ $\hat{y} = 1$	**True Positive (+)**	**False Positive (Type I Error)**
Predicted Value $\rightarrow$ $\hat{y} = -1$	**False Negative (Type II Error)**	**True Negative (-)**

Misclassification Error (also Misclassification Rate or Accuracy), most used in Machine Learning:

$$\frac{\text{FP} + \text{FN}}{n} = \frac{1}{n}\sum_{i-1}^{n} 1_{[y_i \neq \hat{y}_i]} = \frac{\text{False Positives} + \text{False Negatives}}{\text{Total number of observations}}$$

True Positive Rate TPR (also Sensitivity or Recall):

$$\frac{\text{TP}}{\#\text{Positive}} = \frac{\text{TP}}{\text{TP} + \text{FN}} = \frac{\sum_i^n 1_{[y_i = \hat{y}_i \text{ and } y_i = 1]}}{\sum_i^n 1_{[y_i = 1]}} = \frac{\text{True Positives (+)}}{\text{True Positives (+)} + \text{False Negatives}}$$

True Negative Rate TNR (also Specificity):

$$\frac{\text{TN}}{\#\text{Negative}} = \frac{\text{TN}}{\text{TN} + \text{FP}} = \frac{\sum_i^n 1_{[y_i = \hat{y}_i \text{ and } y_i = -1]}}{\sum_i^n 1_{[y_i = -1]}} = \frac{\text{True Negatives (}-\text{)}}{\text{False Postives} + \text{True Negatives (}-\text{)}}$$

False Positive Rate FPR:

$$\frac{\text{FP}}{\#\text{Negative}} = \frac{\text{FP}}{\text{TN} + \text{FP}} = \frac{\sum_i^n 1_{[y_i \neq \hat{y}_i \text{ and } y_i = -1]}}{\sum_i^n 1_{[y_i = -1]}} = \frac{\text{False Positives}}{\text{False Positives} + \text{True Negatives (}-\text{)}}$$

Precision:

$$\frac{\text{TP}}{\#\text{Predicted Positive (+)}} = \frac{\text{TP}}{\text{TP} + \text{FP}} = \frac{\sum_i^n 1_{[y_i = \hat{y}_i \text{ and } y_i = 1]}}{\sum_i^n 1_{[\hat{y}_i = 1]}} = \frac{\text{True Positives (+)}}{\text{True Positives (+)} + \text{False Positives}}$$

F-1 Score:

$$F1 = 2\frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} = 2\frac{\left(\frac{\text{TP}}{\#\text{Predicted Positive (+)}} = \frac{\text{TP}}{\text{TP} + \text{FP}}\right) \times \left(\frac{\text{TP}}{\#\text{Positive}} = \frac{\text{TP}}{\text{TP} + \text{FN}}\right)}{\left(\frac{\text{TP}}{\#\text{Predicted Positive (+)}} = \frac{\text{TP}}{\text{TP} + \text{FP}}\right) + \left(\frac{\text{TP}}{\#\text{Positive}} = \frac{\text{TP}}{\text{TP} + \text{FN}}\right)}$$