0 Formulas

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

$$F_{score} = \frac{2TP}{2TP + FN + FP} \tag{3}$$

$$F_{score(weighted)} = \frac{(1 + \beta)TP}{(1 + \beta)TP + \beta + FN + FP} \tag{4}$$

If $\beta = 0$, $F_{score(weighted)} = \frac{TP}{TP + FP}$ so *Precision* is emphasized.
For n subgroups, the **Macro Average** is the average of the subgroups.

$$\text{Macro Average} = \frac{\sum_{i=1}^{n} \text{Evaluation}_i}{n} \tag{5}$$

1 Terms

Definition 1.1. False Positive: Looks like a winner, but is not.

Definition 1.2. False Negative: Result is true, but predicted to be false.

Definition 1.3. Precision: Emphasizes *correct* results. consider a Google Search.

Definition 1.4. Recall: "Better safe than sorry". Emphasizes more positives, even if incorrect since they can be checked later. Consider a cancer screening.

Definition 1.5. Overfitting: Model fits training data, but does not show overall tred, i.e. does not accurately predict labels for test data.

Definition 1.6. Underfitting: Model does not fit training data, and is too general for an accurate prediction.