

Exercises - Solutions

§6.8: Classification Performance

Problem 1 Total: $P = 30 + 10 + 20 + 140 = 200$. Accuracy: $\frac{30+140}{200} = 0.85$. Precision: $\frac{30}{30+10} = 0.75$. Recall: $\frac{30}{30+20} = 0.60$. Specificity: $\frac{140}{140+10} = \frac{140}{150} \approx 0.933$. F_1 : $\frac{2(0.75)(0.60)}{0.75+0.60} = \frac{0.90}{1.35} \approx 0.667$.

Problem 2 Every observation is predicted negative. Thus $TN = 5000$, $FN = 50$, $TP = 0$, $FP = 0$. Accuracy: $\frac{5000}{5050} \approx 0.9901$ or about 99.01%. Recall: $\frac{0}{0+50} = 0$. The classifier misses every positive observation. The high accuracy reflects only the extreme class imbalance.

Problem 3 Recall: $\frac{90}{90+10} = 0.90$. Precision: $\frac{90}{90+180} = \frac{90}{270} = \frac{1}{3} \approx 0.333$. If the main purpose is to avoid missing disease, recall is the more directly relevant metric because it measures the fraction of actual disease cases detected.

Problem 4 At $t = 0.5$: predictions: 1, 0, 1, 0, 0. Therefore $TP = 1$, $FN = 1$, $FP = 1$, $TN = 2$. Recall: $\frac{1}{2} = 0.50$. Specificity: $\frac{2}{3} \approx 0.667$. At $t = 0.3$: predictions: 1, 1, 1, 1, 0. Thus $TP = 2$, $FN = 0$, $FP = 2$, $TN = 1$. Recall: $\frac{2}{2} = 1$. Specificity: $\frac{1}{3} \approx 0.333$. Lowering the threshold increased recall but reduced specificity.

Problem 5 Specificity: $\frac{TN}{TN+FP}$. Therefore $1 - \text{Specificity} = 1 - \frac{TN}{TN+FP} = \frac{TN+FP-TN}{TN+FP} = \frac{FP}{TN+FP} = \text{FPR}$. Similarly: $1 - \text{Recall} = 1 - \frac{TP}{TP+FN} = \frac{TP+FN-TP}{TP+FN} = \frac{FN}{TP+FN} = \text{FNR}$.

Problem 6 $F_1 = 2 \frac{(0.75)(0.60)}{0.75+0.60}$. Thus $F_1 = \frac{0.90}{1.35} = 0.667$. Arithmetic mean: $\frac{0.75+0.60}{2} = 0.675$. The harmonic mean is slightly smaller. In general it penalizes imbalance between precision and recall more strongly.

Problem 7

```
y = np.array([1, 0, 1, 1, 0, 0, 1, 0])
p = np.array([.8, .7, .6, .4, .3, .2, .9, .1])

threshold = 0.5
y_pred = (p >= threshold)

TP = np.sum((y_pred == 1) & (y == 1))
FP = np.sum((y_pred == 1) & (y == 0))
FN = np.sum((y_pred == 0) & (y == 1))
TN = np.sum((y_pred == 0) & (y == 0))
```

Predictions: (1, 1, 1, 0, 0, 0, 1, 0). Thus $TP = 3$, $FP = 1$, $FN = 1$, $TN = 3$. Accuracy: $\frac{3+3}{8} = 0.75$. Precision: $\frac{3}{4} = 0.75$. Recall: $\frac{3}{4} = 0.75$. Specificity: $\frac{3}{4} = 0.75$. F_1 : 0.75.

Problem 8 (a) Training cross-entropy is the loss used to fit a probabilistic classifier on the training observations. It evaluates the probabilities assigned to the observed labels. (b) Test-set accuracy is the fraction of correct hard predictions on observations not used to fit the model, at a specified classification threshold. (c) ROC AUC summarizes

how well the classifier's scores rank positive observations above negative observations across all thresholds. Two classifiers can have similar accuracy at one particular threshold but different ranking quality across the full range of thresholds. Therefore their AUC values can differ.