

## 1.1

a)

$$TP_A \text{ (True Positive)} = 100$$

$$TN_A \text{ (True Negative)} = 5$$

$$FP_A \text{ (False Positive)} = 8$$

$$FN_A \text{ (False Negative)} = 2$$

$$TP_B \text{ (True Positive)} = 5$$

$$TN_B \text{ (True Negative)} = 100$$

$$FP_B \text{ (False Positive)} = 2$$

$$FN_B \text{ (False Negative)} = 8$$

Insgesamt Teil von Klasse A: 102

Insgesamt Teil von Klasse B: 13

$$TP_A \text{ (True Positive)} = 100$$

$$TP_B \text{ (True Positive)} = 5$$

$$TN_A \text{ (True Negative)} = 5$$

$$TN_B \text{ (True Negative)} = 100$$

$$FP_A \text{ (False Positive)} = 8$$

$$FP_B \text{ (False Positive)} = 2$$

$$FN_A \text{ (False Negative)} = 2$$

$$FN_B \text{ (False Negative)} = 8$$

Insgesamt Teil von Klasse A: 102

Insgesamt Teil von Klasse B: 13

b)    • **Formeln:**

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Accuracy} = \frac{TP + TN}{\text{Gesamtanzahl der Daten}}$$

$$F_1\text{-Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

• **Berechnung für Klasse A:**

$$\text{Recall}_A = \frac{100}{100 + 5} = 0.95$$

$$\text{Precision}_A = \frac{100}{100 + 8} = 0.926$$

$$\text{Accuracy}_A = \frac{105}{115} = 0.913$$

$$F_{1,A}\text{-Score} = 2 \cdot \frac{0.926 \cdot 0.95}{0.926 + 0.95} = 0.938$$

• **Berechnung für Klasse B:**

$$\text{Recall}_B = 0.38$$

$$\text{Precision}_B = 0.71$$

$$\text{Accuracy}_B = 0.91$$

$$F_{1,B}\text{-Score} = 0.50$$

- c) Man kann mit dem  $F_1\text{-Score}$  die Qualität des Klassifizierers beschreiben. Der Score gibt eine Kombination aus Precision und Recall an, wobei 1 als gut und 0 als schlecht angesehen wird.
- d) Bei einer Gleichverteilung wird ein Datenpunkt zu 50% zu A und zu 50% zu B zugeordnet.  
 $\Rightarrow$  wir haben also eine Accuracy von  $\approx 50\%$ , da wir im Schnitt  $\frac{57.5}{115} = 0.5$  haben.
- e) Ein Klassifizierer der nur ein mal bei einem extrem sicheren wahren Ergebnis true gibt und sonst immer nur false.  
 Hat durch die fehlenden False Positives eine Precision von 100%. Setzt sich (trivialer Weise) zusammen aus  $\frac{1}{1} = 1$
- f) Mehr Daten für Klasse B sammeln und B im allgemeinen besser einordnen (Recall verbessern).