

Problem 1

		Actual class	
		Positive ($y=1$)	Negative ($y=0$)
Predicted class	Positive ($\hat{y}=1$)	TP	FP
	Negative ($\hat{y}=0$)	FN	TN

From the question, we know that

$$P(t=1) = 90\%$$

$$P(t=0) = 10\%$$

Because we always predict $t=1$

$$P(TP) + P(FN) = 0.9$$

$$P(FN) = P(TN) = 0$$

$$P(FP) + P(TN) = 0.1$$

$$P(TP) + P(FP) = 1$$

$$\begin{cases} P(TP) = 0.9 \\ P(FN) = 0 \\ P(FP) = 0.1 \\ P(TN) = 0 \end{cases}$$

$$\begin{aligned}
 \text{Precision (P)} &= \frac{|\{\hat{t}=1\} \cap \{t=1\}|}{|\{\hat{t}=1\}|} \\
 &= \frac{P(\text{TP})}{P(\text{TP}) + P(\text{FP})} \\
 &= \frac{0.9}{0.9 + 0.1} \\
 &= 0.9 = 90\%
 \end{aligned}$$

$$\begin{aligned}
 \text{Recall (R)} &= \frac{|\{\hat{t}=1\} \cap \{t=1\}|}{|\{y=1\}|} \\
 &= \frac{P(\text{TP})}{P(\text{TP}) + P(\text{FN})} \\
 &= \frac{0.9}{0.9 + 0} = 1 = 100\%
 \end{aligned}$$

$$\begin{aligned}
 F_\beta\text{-score} &= \frac{(1 + \beta^2) \cdot P \cdot R}{\beta^2 \cdot P + R} \\
 &= \frac{(1 + 1) \cdot 0.9 \cdot 1}{1^2 \cdot 0.9 + 1} \\
 &= \frac{2 \cdot 0.9 \cdot 1}{1 \cdot 0.9 + 1} = \frac{1.8}{1.9} \approx 0.9474 \\
 &= 94.74\%
 \end{aligned}$$

Problem 2.

$$\sigma(z) = \frac{1}{1 + \exp\{-z\}} \quad \sigma(-z) = \frac{1}{1 + \exp\{z\}}$$

$$1 - \sigma(z) = 1 - \frac{1}{1 + \exp\{-z\}}$$

$$= \frac{\exp\{-z\}}{1 + \exp\{-z\}}$$

$$= \frac{\exp\{-z\} \cdot \exp\{z\}}{(1 + \exp\{-z\}) \cdot \exp\{z\}}$$

$$= \frac{1}{\exp\{z\} + 1} = \sigma(-z)$$

Problem 3.

$$J = -t \log y - (1-t) \log(1-y)$$

$$= -t_1 \log y_1 - t_2 \log y_2$$

$$= t_1 \log \frac{1}{y_1} + t_2 \log \frac{1}{y_2}$$

$$= \sum_{k \in \{1,2\}} t_k \log \frac{1}{y_k}$$

$$KL(P||Q) = \sum_{k=1}^K P_k \log \frac{P_k}{Q_k}, \quad P_k \in \{0,1\},$$

binary classification.

Thus, minimizing the loss $J = -t \log y - (1-t) \log(1-y)$

is equivalent to minimize the kullback-Leibler

divergence between t and y .