MACHINE LEARNING 1: ASSIGNMENT 2

Tom Nick 340528
Niklas Gebauer 340942

Exercise 1

(a) With the definition $P(x_k \mid \sigma)$:

$$P(x \mid \sigma) = \begin{cases} \sigma & \text{if} x = \text{head} \\ 1 - sigma & \text{if} \quad x = \text{tail} \end{cases}$$

We can state the likelihood function $P(\mathcal{D} \mid \sigma)$:

$$\begin{aligned} P(\mathcal{D} \mid \sigma) &= \prod_{k=i}^{n} P(x_k \mid \sigma) \\ &= \sigma^5 \cdot (1-\sigma)^2 \end{aligned}$$

(b) The maximum likelihood solution for σ is simply the sample mean:

$$\sigma = \frac{\#\{x = \text{head} \mid x \in D\}}{\#(D)} = \frac{5}{7}$$

With this we can compute $P(x_8 = \text{head}, \quad x_9 = \text{head} \mid \sigma)$:

$$P(x_8 = \text{head}, \quad x_9 = \text{head} \mid \sigma) = P(x_8 = \text{head} \mid \sigma) P(x_9 = \text{head} \mid \sigma) = \sigma^2 = \frac{25}{49} = 0.51$$

(c) With the definition of $p(\sigma) = 1$, the posterior is computed as:

$$\begin{aligned} p(\sigma \mid \mathcal{D}) &= \frac{p(\mathcal{D} \mid \sigma) p(\sigma)}{p(x)} \\ &= \frac{p(\mathcal{D} \mid \sigma)}{p(x)} \\ &= \frac{\sigma^5 \cdot (1-\sigma)^2}{p(x)} \end{aligned}$$