Group 6: Timm Behner, Philipp Bruckschen, Patrick Kaster, Markus Schwalb
MA-INF 4111 - Intelligent Learning and Analysis Systems: Machine Learning
Exercise Sheet 4

1. Confidence Intervals

(a) given: $\text{error}_S(h) = \frac{10}{65}, n = 65$

$$\text{error}_D^{95\%}(h) \leq \text{error}_S(h) + z_N \sqrt{\frac{\text{error}_S(h)(1 - \text{error}_S(h))}{n}}$$

$$= \frac{10}{65} + 1.96 \sqrt{\frac{\frac{10}{65}(1 - \frac{10}{65})}{65}}$$

$$\approx 0.24$$

(b)

$$\text{error}_D^{90\%}(h) \leq \frac{10}{65} + 1.64 \sqrt{\frac{\frac{10}{65}(1 - \frac{10}{65})}{65}}$$

$$\approx 0.23$$

∎

2. Sample Size

given: $\text{error}_D(h) \in [0.2, 0.6] \Rightarrow \sigma = 0.2$
wanted: bound on error $E < 0.1$

$$E = z_N \frac{\sigma}{\sqrt{n}} \Rightarrow$$

$$n = \left(z_N \frac{\sigma}{E}\right)^2$$

$$n = \left(1.96 \frac{0.2}{0.1}\right)^2$$

$$= \lceil 15.3664 \rceil = 16$$

∎

4. AUC

(a)

$$\text{Accuracy}_S(h) = \frac{\#\,\text{Correctly Classified}}{\#\text{Test Examples}}$$

$$= \frac{990}{100} = 0.99$$

∎