Question 1.

Let X contain all real numbers from $[0; 1]$ which can be represented using 256 bits. Let $\mathcal{H} = X$, and let the decision be given by $H \in \mathcal{H}$ as

$$h(x) = 1 \text{ iff } x > H$$

Determine an m such that with probability at least 0.9, $\text{err}(h) < 0.1$, where h is an arbitrary hypothesis from $\mathcal{H}$ consistent with m i.i.d. examples from X. Estimate it

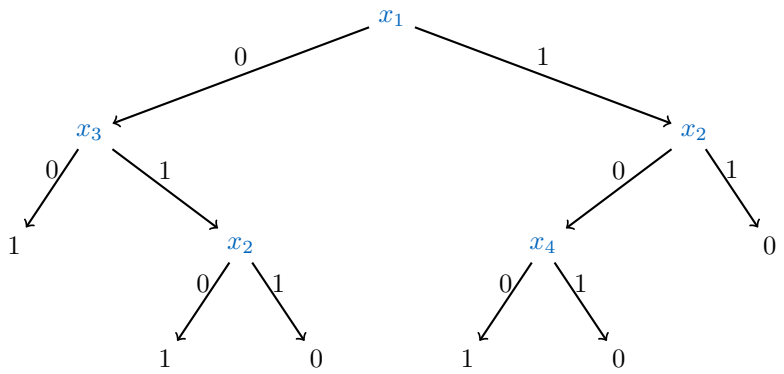
(a) without using any *textbook* lower bounds

(b) using the lower bound $m > \frac{1}{\epsilon} \ln \frac{|\mathcal{H}|}{\delta}$

(c) using the lower bound $m > \frac{8}{\epsilon} \left(\text{VC}(\mathcal{H}) \cdot \ln \frac{16}{\epsilon} + \ln \frac{2}{\delta} \right)$

Question 2.

Consider the following decision tree:



(a) Express the tree as a 3-DNF.

(b) Express the tree as a 3-CNF.

(c) How can we use (modify) the generalization algorithm to learn k-decision trees in the PAC learning model?