Exercise 1

Assume that $X \sim (\text{Unif}[0,1])^3$. Consider a supervised learning problem with

$$\mathbb{E}[Y|X = (x_1, x_2, x_3)^T]] = m^*(x_1, x_2, x_3) = \mathbb{1}(x_1 \geq 0.5) + \mathbb{1}(x_2 \leq 0.3) + \mathbb{1}(x_2 \geq 0.5)\mathbb{1}(x_3 \leq 0.5).$$

(a) Build a binary tree that can represent the function m^*. How many leaves does the tree have?

(b) Construct multiple binary trees such that their sum can represent m^*. Construct the trees in way such that each tree has less leaves than the tree constructed in (a).

(c) Discuss why an estimator based on (b) is expected to have lower prediction error compared to an estimator based on (a).