

Answers 2021 ML exam (Exercise 1,3, 4, 5)

Jonas Nordqvist

May 26, 2023

Exercise 1

To answer a) and b) right away we have for instance

$$W^{(2)} = \begin{pmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \end{pmatrix}, \quad b^{(2)} = \begin{pmatrix} -2.5 \\ 0.5 \end{pmatrix}$$

and

$$W^{(3)} = \begin{pmatrix} -1 & -1 \end{pmatrix}, \quad b^{(3)} = 0.5.$$

Regarding c) the answer is *no* as the ‘true’ outputs of the NN should be the half plane in the unit cube of all points with at least two active signals except for (1, 1, 1). This can be compared to the XOR function.

Exercise 3

At the parent node the Gini index is given by

$$G_{\text{parent}} = 1 - \left(\frac{1}{2}\right)^2 - \left(\frac{1}{2}\right)^2 = 1 - 0.25 - 0.25 = 0.5.$$

We have

$$\begin{aligned} \Delta_{S_1} &= 0.5 - \left(\frac{9}{10} \left(1 - \left(\frac{4}{9}\right)^2 - \left(\frac{5}{9}\right)^2 \right) + \frac{1}{10}(1-1) \right) \\ &= 0.5 - \frac{4}{9} = \frac{1}{18}. \end{aligned}$$

Also, considering S_2 we have

$$\begin{aligned} \Delta_{S_2} &= 0.5 - \left(\frac{1}{2} \left(1 - \left(\frac{1}{5}\right)^2 - \left(\frac{4}{5}\right)^2 \right) + \frac{1}{2} \left(1 - \left(\frac{4}{5}\right)^2 - \left(\frac{1}{5}\right)^2 \right) \right) \\ &= 0.5 - \left(1 - \left(\frac{1}{5}\right)^2 - \left(\frac{4}{5}\right)^2 \right) \\ &= 0.5 - \frac{8}{25} \\ &= \frac{9}{50}. \end{aligned}$$

We conclude that the preferred split is S_2 , as $\Delta_{S_2} > \Delta_{S_1}$.

Exercise 4

In a) three classifiers each in b) (ii) would yield 4 classifiers and (i) would give $\binom{4}{2} = 6$.

Regarding c) the answer is no in (ii) there is for instance a small (upside down) triangle in the middle of the points which in which no class is dominating, but for (i) the different decision boundaries are the x and y axis, see Figure 1. Hence, for d) one could argue that (i) is better as in the entire feature space there is no ambiguity for the algorithm. We also note that the decision regions may vary from that of the figure slightly depending on

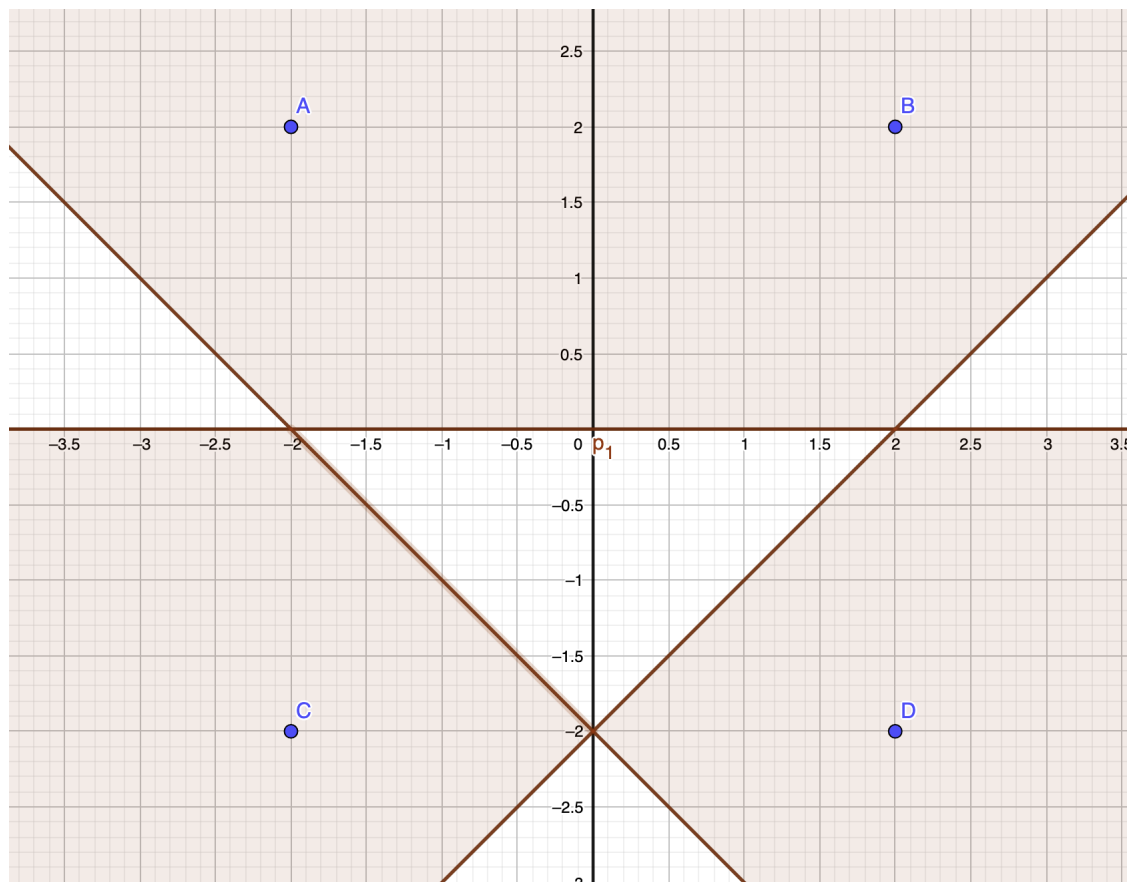


Figure 1: The x and y -axes are the decision boundaries for one vs one, and the red areas are the decision boundaries for one vs all.

the implementation. But it is clear that in the region of the upside-down triangle, none of the classifiers in (i) are voting for these points.

Exercise 5

The update are computed as follows

$$\text{tmp}_1 = \beta_1 - \gamma \frac{\partial J}{\partial \beta_1} = 0.5 - 0.01 \cdot 2 \cdot \frac{5}{6}$$

$$\text{tmp}_2 = \beta_2 - \gamma \frac{\partial J}{\partial \beta_2} = 2 - 0.01 \cdot 2 \cdot \frac{11}{6}$$

So all in all we have after the update

$$(\beta_1, \beta_2) \approx (0.4833, 1.9633).$$