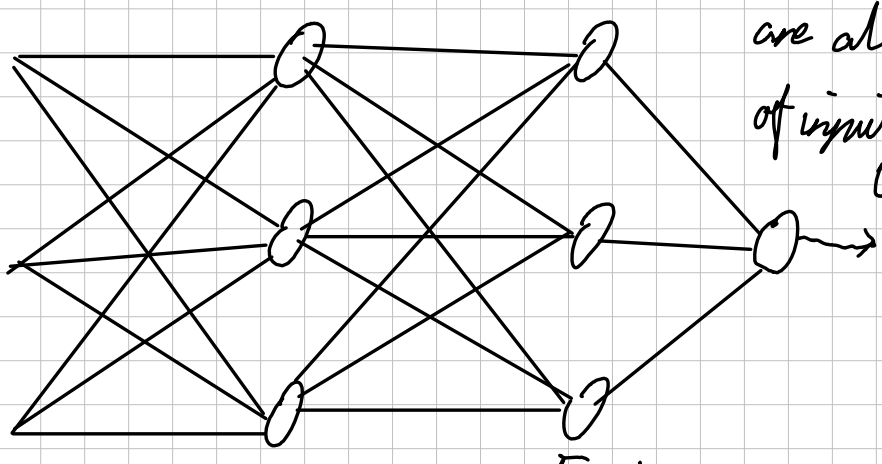
8.1.)

The memory equivalent capacity of a is:



For this node there are also just 3 bits of input, although 4 bits of capacity

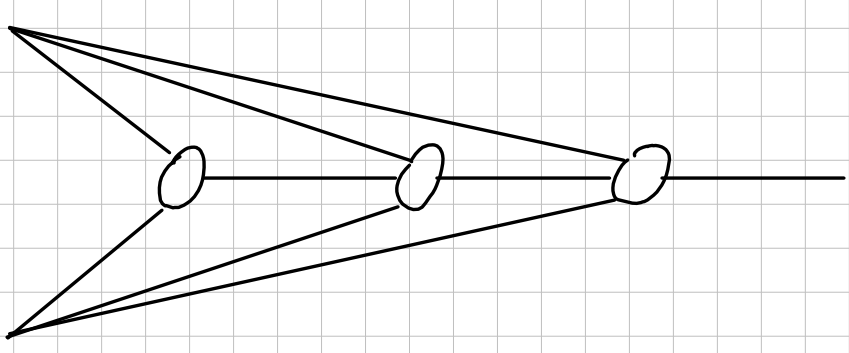
Each of these nodes can save 3+1 bits, that is 3·4 bits

Each node in this part can also save 3·4, but there are really just 3 distinct outputs with 3 distinct bits

Total capacity is therefore:

$$\underline{MEC = 12 + 3 + 3 = 18}$$

b) Then we have this network:



- None of these nodes have their output restricted by the previous layer. We therefore got

$$\underline{MEC = 3 + 2(3+1) = 3 + 8 = 11 \text{ bits}}$$

c) The maximum amount of rows any network can "memorize" can be said to be infinity if f.ex. all labels are 0, depending on the definition of "memorize".

However the maximum number of rows we can guarantee to memorize for binary classification are 18 for network a., and 11 for network b.

d) MEC is the number of bits needed to make 1 binary decision per row. For 4 classes, $\log_2(4) = 2$ binary decisions have to be made per row instead, such that one could now guarantee to memorize half the amount of rows compared to binary classification.

That is:

a.) $\left\lfloor \frac{18}{2} \right\rfloor = \underline{9 \text{ rows}}$

b) $\left\lfloor \frac{11}{2} \right\rfloor = \underline{5 \text{ rows}}$