ARTIFICIAL INTELLIGENCE (E016350A)
GHENT UNIVERSITY
AY 2024/2025
Prof. Aleksandra Pizurica
Asst. Prof. Nicolas Nadisic
E-mail: ai@lists.ugent.be

Neural networks – Part 1

1. **[Old exam question] Neural networks**

   We consider a binary classification problem where the input data are two-dimensional. The table below shows the first five examples in the training set.

Example	x_1	x_2	Class
1	10	20	+
2	4	15	-
3	1	15	+
4	1	10	-
5	5	20	+

   (a) Assign to the class '+' the label $y = 1$ and to the class '-' the label $y = 0$. Initialize the weight vector as $\mathbf{w} = [1, 2, 3]^\top$, where the first entry is for the bias term. Run the perceptron learning rule on the five examples, with the learning rate of 1, and fill in the following table and show your calculation:

Example	Score	Predicted class	Updated weights
1			
2			
3			
4			
5			

(b) Now keep all the same as in (a) but use a different convention for the class labels y: for the class '+' use $y = 1$ and for the class '-' use $y = -1$. Will now the perceptron learning rule lead to different updated weights? Explain your answer and show the calculation for the updated weights based on the first two examples from the given dataset.

The updated weights will ◯ change ◯ not change

because

For Example 1:

updated weights: ______________________________

For Example 2:

updated weights: ______________________________

For Example 3:

updated weights: ______________________________

For Example 4:

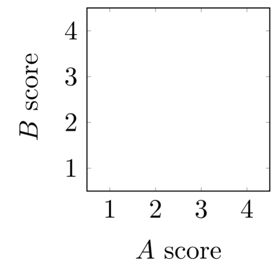
updated weights: ______________________________

For Example 5:

updated weights: ______________________________

2. You want to predict if movies will be profitable based on their screenplays. You hire two critics A and B to read a script you have and rate it on a scale of 1 to 4. The critics are not perfect; here are five data points including the critics' scores and the performance of the movie

#	Movie name	A	B	Profit?
1	Pellet Power	1	1	-
2	Ghosts!	3	2	+
3	Pac is Bac	2	4	+
4	Not a Pizza	3	4	+
5	Endless Maze	2	3	-



(a) First, you would like to examine the linear separability of the data. Plot the data on the 2D plane above, label profitable movies with + and non-profitable movies with - and determine if the data are linearly separable.

(b) Now you decide to use a single-layer perceptron to classify your data. Suppose you directly use the points given above as features. That is f_1 = points given by A and f_2 = points given by B.

Run one pass through the data with the perceptron learning algorithm, filling out the table below. Initialize the weights as $(-1, 0, 0)$, where the first entry denotes the bias, and set the learning rate to 1. Go through the data points in the given order using data point #1 at step 1.

Step	Current weights	Score (show your calculation here)	Updated weights
1			
2			
3			
4			
5			

(c) Have weights been learned that separate the data?

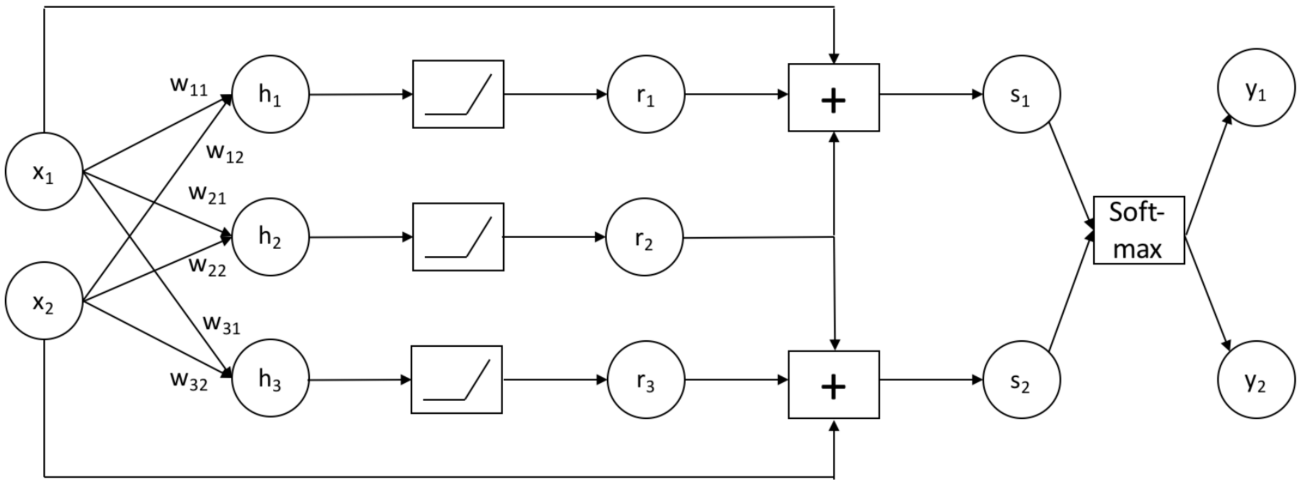
(d) More generally, irrespective of the training data, you want to know if your features are powerful enough to allow you to handle a range of scenarios. Circle the scenarios for which a perceptron using the features above can indeed perfectly classify movies which are profitable according to the given rules:

(1) Your reviewers are awesome: if the total of their scores is more than 8, then the movie will definitely be profitable, and otherwise it won't be.

(2) Your reviewers are art critics. Your movie will be profitable if and only if each reviewer gives either a score of 2 or a score of 3.

(3) Your reviewers have weird but different tastes. Your movie will be profitable if and only if both reviewers agree.

3. The network below is a neural network with inputs x_1 and x_2, and outputs y_1 and y_2. The internal nodes are computed below. All variables are scalar values.
Note that $\text{ReLU}(x) = \max(0, x)$.



The expressions for the internal nodes in the network are given here for convenience:

$$h_1 = w_{11}x_1 + w_{12}x_2 \qquad h_2 = w_{21}x_1 + w_{22}x_2 \qquad h_3 = w_{31}x_1 + w_{32}x_2$$

$$r_1 = \text{ReLU}(h_1) \qquad r_2 = \text{ReLU}(h_2) \qquad r_3 = \text{ReLU}(h_3)$$

$$s_1 = x_1 + r_1 + r_2 \qquad s_2 = x_2 + r_2 + r_3$$

$$y_1 = \frac{\exp(s_1)}{\exp(s_1) + \exp(s_2)} \qquad y_2 = \frac{\exp(s_2)}{\exp(s_1) + \exp(s_2)}$$

(a) **Forward Propagation:** Suppose for this part that

$$x_1 = 3,\ x_2 = 5,\ w_{11} = -10,\ w_{12} = 7,\ w_{21} = 2,\ w_{22} = 5,\ w_{31} = 4,\ w_{32} = -4.$$

What are the values of the following internal nodes? Please simplify any fractions.

(1) $h_1 =$ ____________

(2) $s_1 =$ ____________

(3) $y_2 =$ ____________

(b) **Back Propagation:** Compute the following gradients analytically. The answer should be an expression of any of the nodes in the network $(x_1, x_2, h_1, h_2, h_3, r_1, r_2, r_3, s_1, s_2, y_1, y_2)$ or weights w_{11}, w_{12}, w_{21}, w_{22}, w_{31}, w_{32}. In the case where the gradient depends on the value of nodes in the network, please list all possible analytical expressions, caused by active/inactive ReLU, in form of a set $\{\dots\}$.

(1) $\dfrac{\partial h_2}{\partial x_1} =$ ____________ (3) $\dfrac{\partial r_3}{\partial w_{31}} =$ ____________ (5) $\dfrac{\partial s_1}{\partial x_1} =$ ____________

(2) $\dfrac{\partial h_1}{\partial w_{21}} =$ ____________ (4) $\dfrac{\partial s_1}{\partial r_1} =$ ____________ (6) $\dfrac{\partial y_2}{\partial s_2} =$ ____________