Artificial intelligence
University of Rennes 1
ESIR
2021/22

Exercise 1: Properties of activation functions Recall the definitions of the popular activation functions: sigmoid function $\sigma(x) = \frac{1}{1+e^{-x}}$, hyperbolic tangent $tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$, rectified liner unit $relu(x) = max(0, x)$.

- What is their value at $x = 0$?
- What is the range of the values (of $f(x)$)? (where $f(x)$ is one of the activation function)
- Draw the graphs of these functions.
- For sigmoid, and tanh, express $f(-x)$ as a function of $f(x)$
- Compute the derivative functions of the activation functions.

Exercise 2: Activation functions: relationship to the derivative functions Demonstrate the following properties of the activation functions.

$$\sigma(x) = \frac{1}{1+e^{-x}}$$

$$tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

- Sigmoid function $\sigma'(x) = \sigma(x)(1 - \sigma(x))$
- Tanh function $tanh'(x) = 1 - tanh^2(x)$

Exercise 3: Single layer perceptron: Decision surface Decision boundary: a hypersurface that partitions the underlying vector space into two sets, for each class (all points on the one side of the decision boundary are classified to one class, while those on the other side, to another class)

Let us suppose that we have a single layer perceptron with two input nodes: x_1 and x_2. (The perceptron also has a bias term, with an imaginary 1 node). We use a binary threshold activation function. A binary threshold function $f(z)$ returns 1 (i.e. $f(z) = 1$), if $z \geq 0$, otherwise, if $z < 0$, it returns 0, that is $f(z) = 0$. Analyse the decision boundary of the prediction.

Exercise 4: Single layer perceptron: XOR gate Try to design a single layer perceptron that computes the XOR of the two inputs x_1 and x_2 (see Table). Let us assume also for this exercise, that we use the binary threshold activation function. Is this possible? Explain how to chose the weights of the network or argument why is it not possible.

x_1	x_2	x_1 XOR x_2
0	0	0
0	1	1
1	0	1
1	1	0

Table 1: XOR