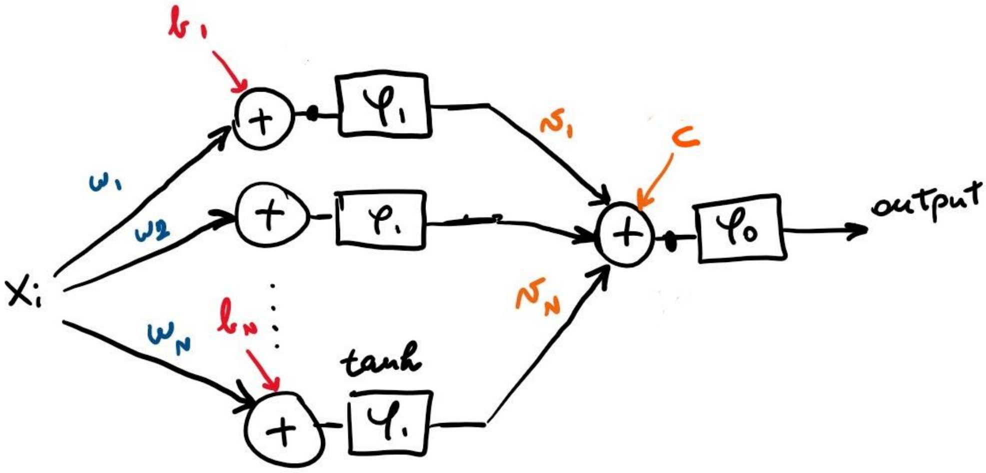
REPORT HM4



FP

$$E = \min \left[\frac{1}{n} \Big(d_1 - \underbrace{\sum^{N} v_i \tanh(w_i x_1 + b_i) + c}_{f_i} \Big)^2 \right]$$

$$\frac{\partial f}{\partial v_i} = -2(d_i - f_i)\tanh(w_i x_1 + b_i)$$

$$\frac{\partial f}{\partial w_1} = -2(d_i - f_i)\, v_i \left(1 - \tanh^2(w_i x_i + b_i)\right) x_i$$

$$\frac{\partial f}{\partial b_i} = -2(d_i - f_i)\, v_i \left(1 - \tanh^2(w_i x_i + b_i)\right)$$

$$\frac{\partial f}{\partial c} = -2(d_i - f_i)$$

With w_i and b_i are indicated the weights and biases of the first layer, while with v_i and c the weights and the bias of the output layer.

PSEUDOCODE based on the draw and the names in the previous page:

While MSE < 0.02:

Increment the number of epoch

For el in Training set -> X_without_mean :

Forward propagation FP -> f (el) passing the network from left to right

Backward propagation BP ->

$$w\_i \leftarrow w\_i - step * \frac{dE}{dw}$$

$$b\_i \leftarrow b\_i - step * \frac{dE}{db}$$

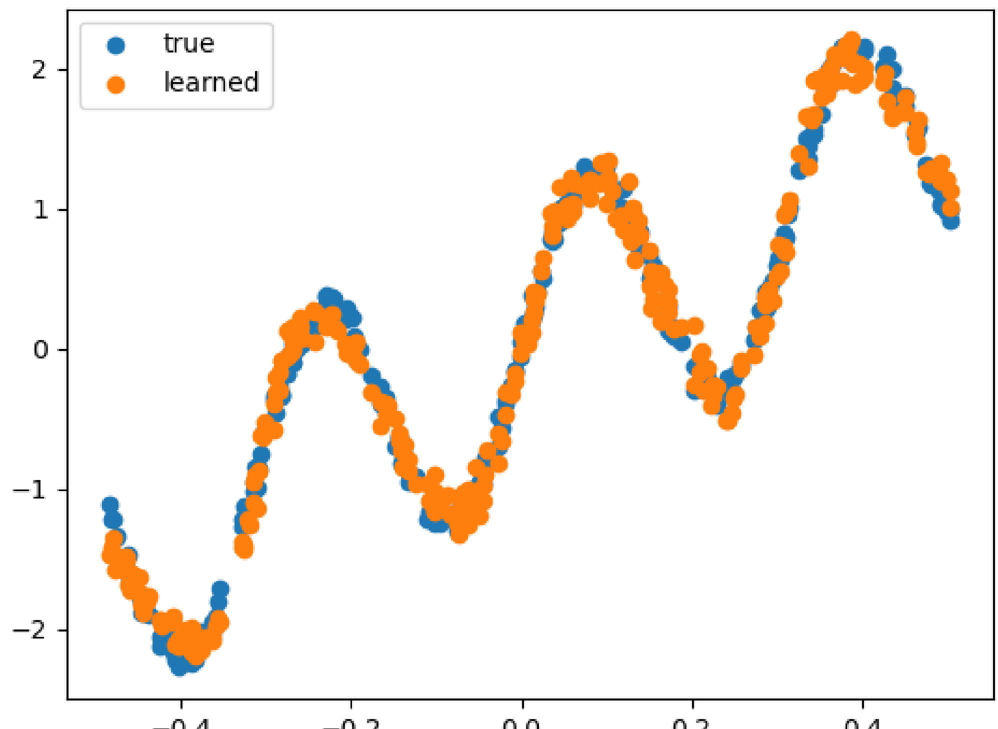
$$v\_i \leftarrow v\_i - step * \frac{dE}{dv}$$

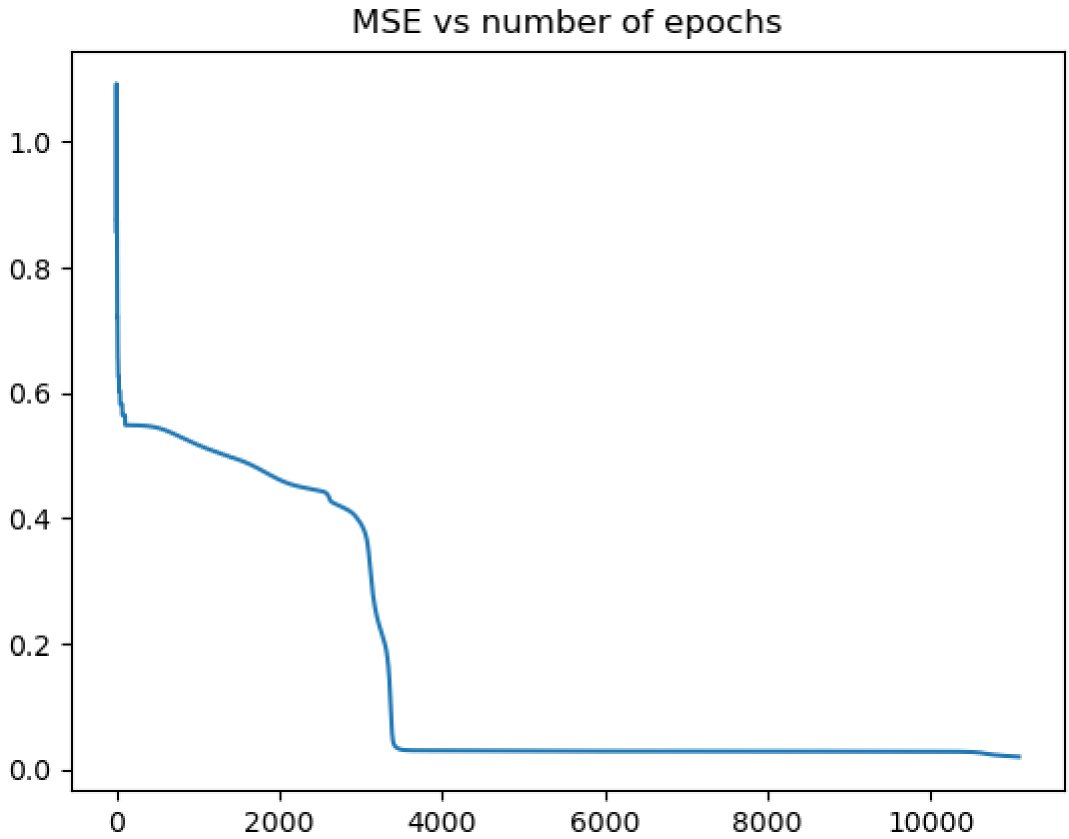
$$c \leftarrow c - step * \frac{dE}{dc}$$

Calculate the MSE of the epoch

STOP -> the network has reached the optimal weights and is trained

The derivative are reported in the upper part, together with a drawing of the network. In the following part I show the plot of the trained network and the plot with the number of epochs and the MSE = $\frac{1}{n}\sum_{i=1}^{n}(d_i - f(x_i, \mathbf{w}))^2$.




MSE vs number of epochs
1.0
0.8
0.6
0.4
0.2
0.0
0
2000
4000
6000
8000
10000