Exercise 1:

Given the recurrent neural network

$$h_t = \tanh(W h_{t-1} + U x_t)$$
$$y_t = V h_t$$

Draw the computation graph for two time steps with outputs y_1, y_2. We assume all variables to be scalar. Mark variables and parameters with circels and computations with rectangles. Indicate where h_1, h_2 are computed in the graph.

$$h_t = \tanh(W h_{t-1} + U x_t)$$
$$y_t = V h_t$$

Exercise 2:

Let $z = y_1 \cdot y_2$ and $\tau_i = Wh_{i-1} + Ux_i$. Add z to your computation graph of exercise 1. Then derive the gradient $\frac{dz}{dW}$ as a function of $y_1, y_2, h_0, h_1, \tau_1, \tau_2, W, V$.

Exercise 2:

Using $\frac{d\tanh(x)}{dx}|_{x_{\text{in}}} = 1 - \tanh(x_{\text{in}})^2$, mark the variables that have to be stored in memory (additional to the parameters and inputs) during the forward pass in order to be able to compute thegradients during the backwards pass. How many additional variables would have to be stored if we backpropagated through 100 steps instead of 2.

Now reusing the compuation from the forward pass $y = \tanh(x)$ and $\frac{dy}{dx} = 1 - y^2$, mark the variables needed for backpropagation. How many additional variables would now need to be stored when backpropagating thorugh 100 steps instead of 2.

Exercise 3:

Looking at your results from exercise 2, where does the dependency on the inputs x_1, x_2 appear in $\frac{dz}{dW}$?

Exercise 4:

Evaluation – Step by Step

1. Get out your smartphones
2. Open an app for QR-code Reading
3. Read the following QR-code on the right side →
4. Open the website
5. Answer the questions
6. Send the evaluation

OR

1. Open the following website in your browser: https://evasys.zv.tum.de/evasys/online.php?p=AIAT-10
2. Answer the questions
3. Send the evaluation