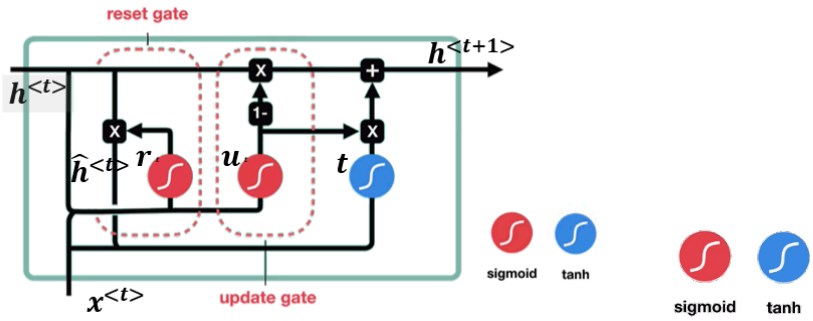
DI504 Foundations of Deep Learning
The Final

Fall 21/22

Duration: 5 hours 02/02/2022

Question 2 (30 Points): Please find below a "Gated Recurrent Unit" (GRU).



At a given time t the weights of the given GRU is as given below:
Wr = [1 1 0 0 0; 1 0 0 1 1; 1 1 0 0 0], $b_r = -1.0$
Wu = [1 0 1 0 1; 1 0 1 0 1; 1 0 1 0 1], $b_u = +1.5$
Wt = [1 1 1 1 1; 0 0 0 0 0; 1 1 1 1 1], $b_t = +0.4$

And the input and the hidden state variables as such:
$x^{<t>} = [+1.82 \quad -0.21]^{\mathrm{T}}$, $h^{<t>} = [+0.36 \quad -1.45 \quad +0.23]^{\mathrm{T}}$

a) **(10 Points)** Find the value of the intermediate of state vector $\hat{h}^{<t>}$ at the output of the "**reset gate**".
b) **(10 Points)** Find the value of the next state $h^{<t+1>}$.
c) **(10 Points)** What exactly happens at time <t> at this gate? Please explain your reasoning.

You may need your calculators. Please remember that the operations inside the GRU are element-wise multiplications/additions.

For the order of the x and h vectors when concatenated, use the formula we saw in class:

$$\sigma\left(\mathbf{W}^{inp1}_{\mathrm{l}\times(\mathrm{k}+\mathrm{n})} \cdot \begin{bmatrix} \mathbf{x}_{\mathrm{n}\times 1}(t_0) \\ \mathbf{a}_{\mathrm{k}\times 1}(t_0) \end{bmatrix} + \mathbf{b}^{inp1}_{\mathrm{l}\times 1}\right)$$