import some libraries

basic parameters

Load Data

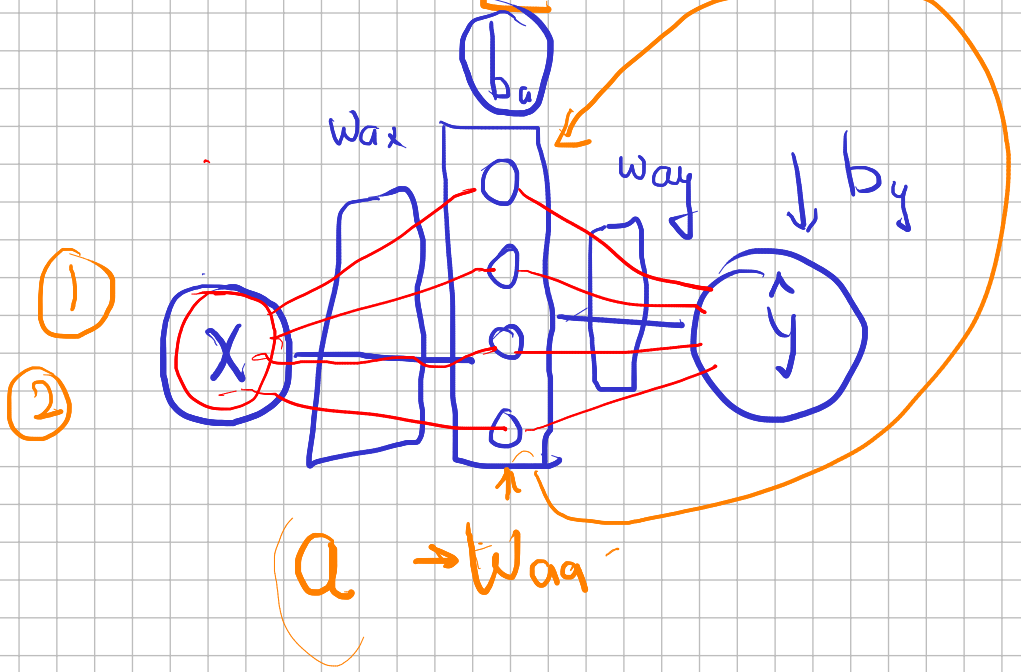
Train Test Split

Transform data

Prepare model

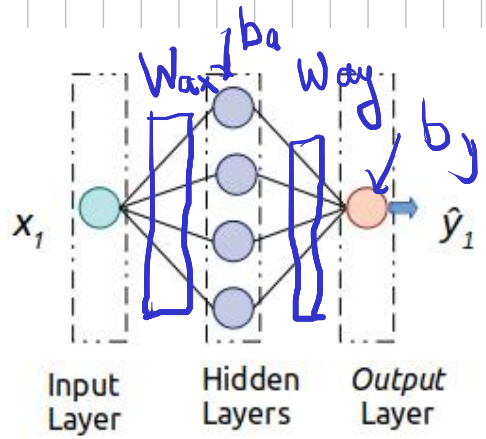
Evaluate

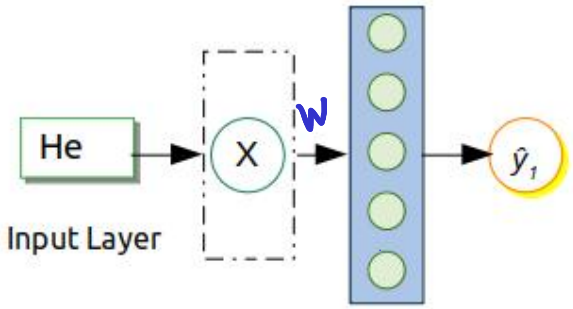
$$a_2 = \tanh\left[[a; x] \times W + b\right]$$



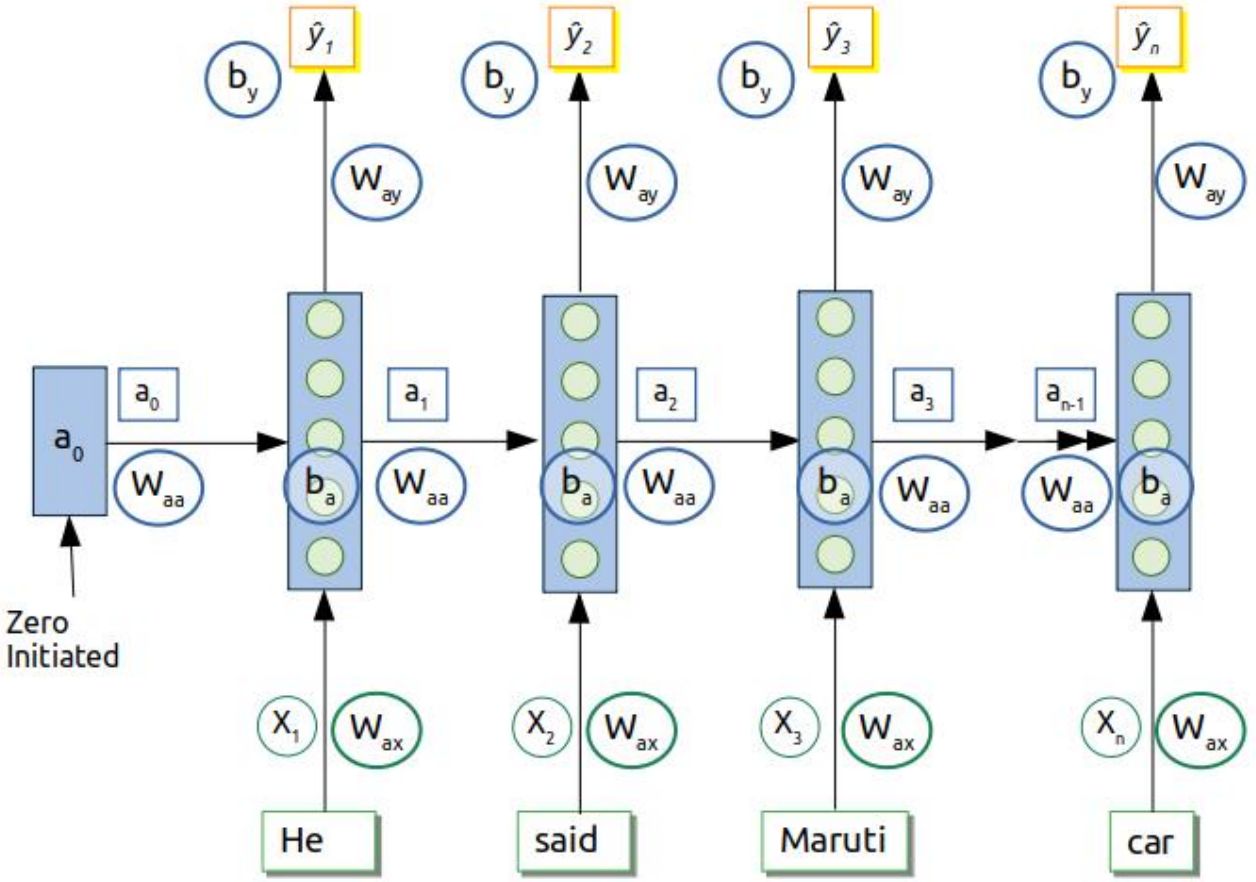
$$a_2 = \tanh\left[x_1 \cdot W_{aa} + x_2 \cdot W_{ax} + b_a\right]$$

$$\hat{y} = \sigma\left[a_2 \cdot W_{ay} + b_y\right]$$





Simple Feed- Forward Network



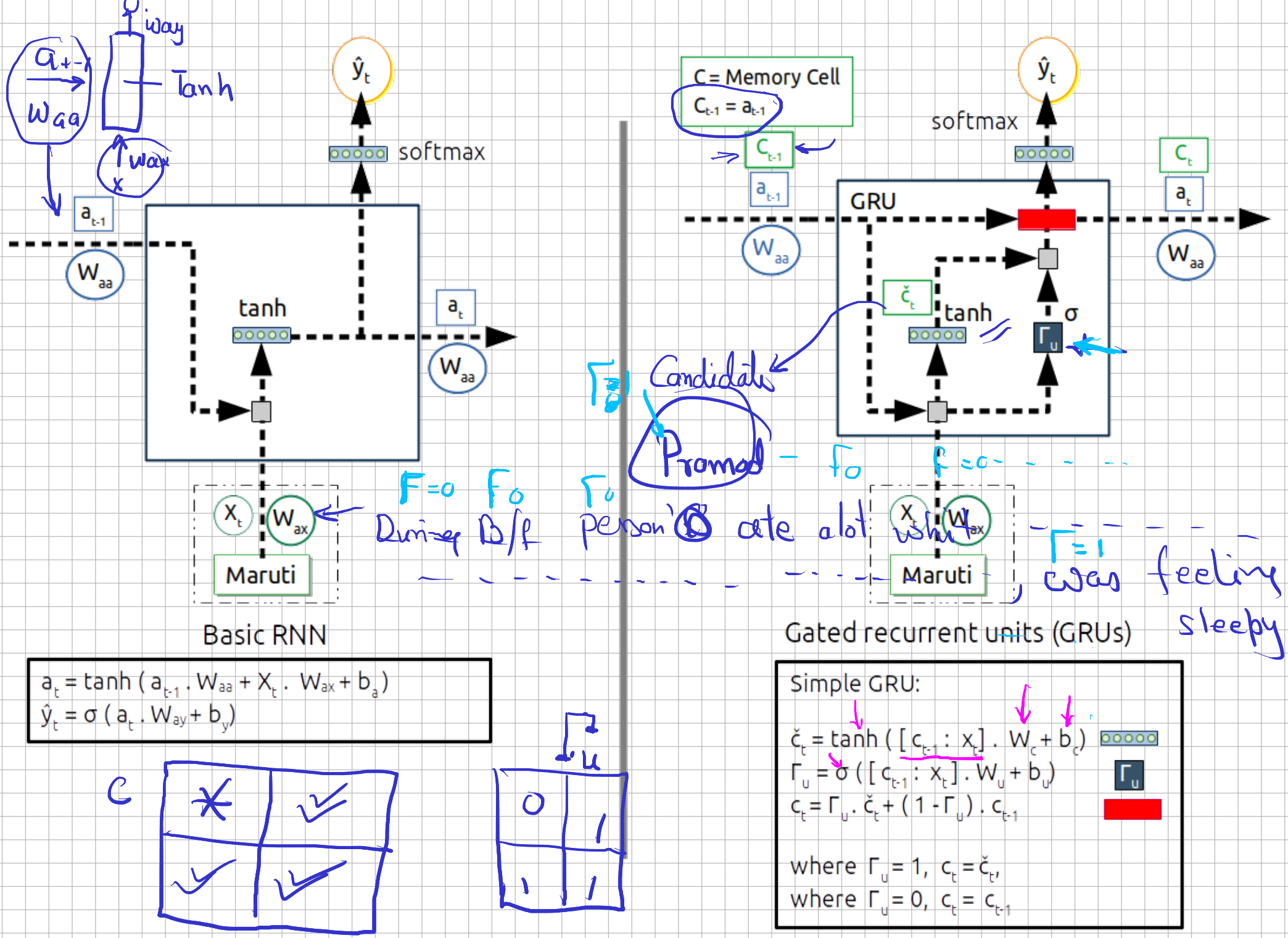
Recurrent Neural Network

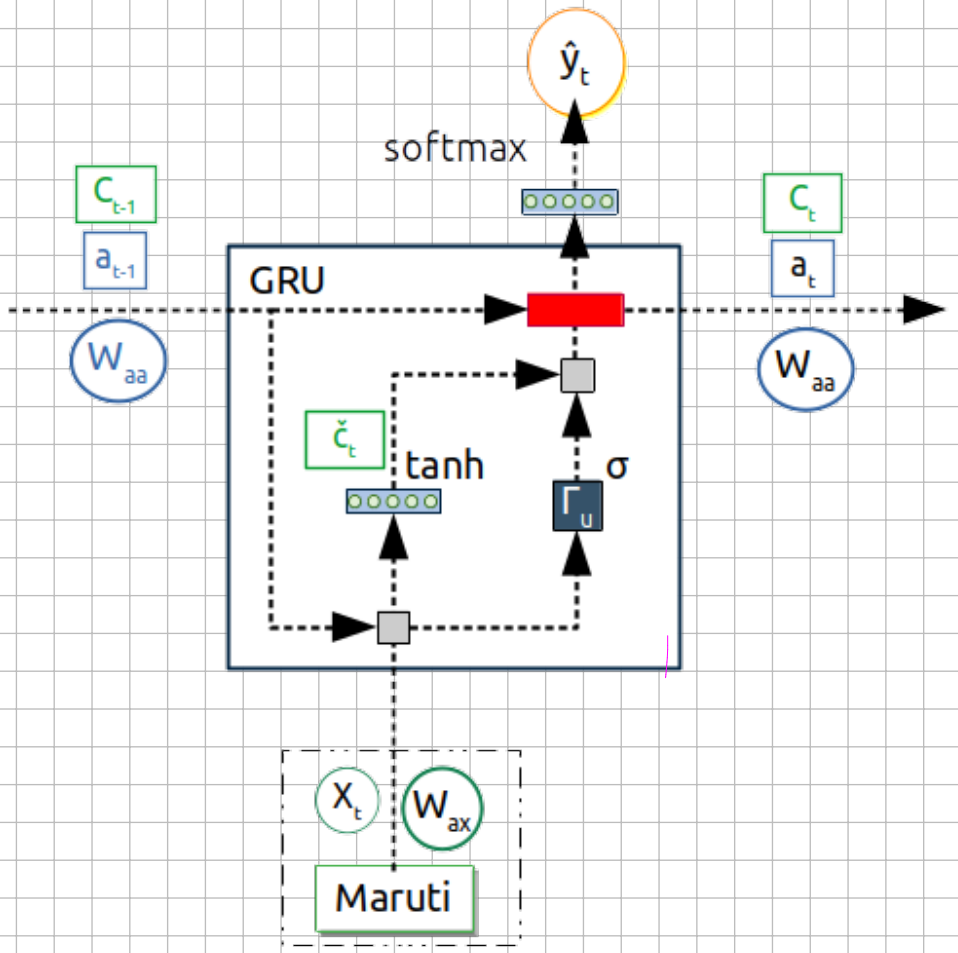
$$a_t = g_1(a_{t-1} \cdot W_{aa} + X_t \cdot W_{ax} + b_a)$$
$$\hat{y}_t = g_2(a_t \cdot W_{ay} + b_y)$$

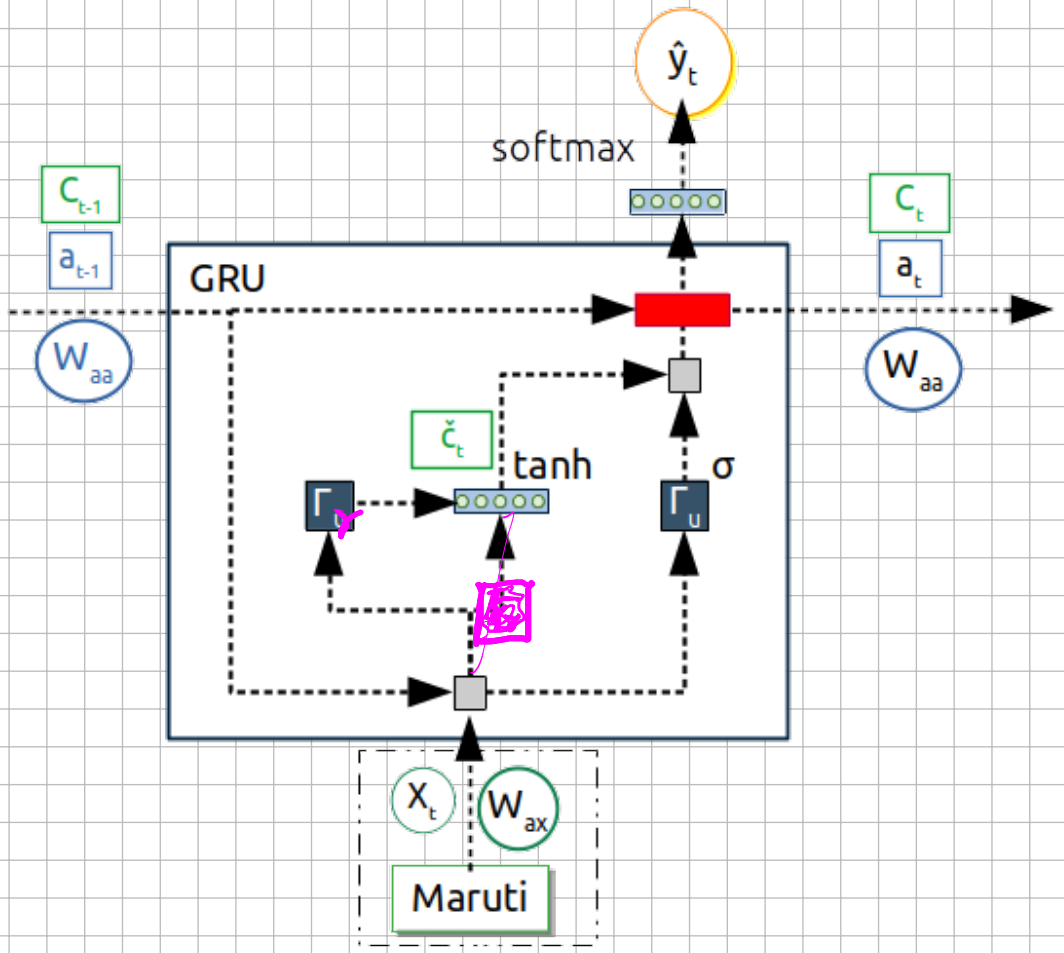
$$a_t = g_1([a_{t-1} : x_t] \cdot W_a + b_a)$$
$$\hat{y}_t = g_2(a_t \cdot W_y + b_y)$$

$$w_0 + x_1 \cdot w_1 + x_2 \cdot w_2 + b$$

$$[x_1, x_2] \cdot \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$


C = Memory Cell
$C_{t-1} = a_{t-1}$
softmax
GRU
tanh
σ
Candidate
Promod
Maruti
Basic RNN
Gated recurrent units (GRUs)
Simple GRU:
$a_t = \tanh(a_{t-1} \cdot W_{aa} + X_t \cdot W_{ax} + b_a)$
$\hat{y}_t = \sigma(a_t \cdot W_{ay} + b_y)$
$\check{c}_t = \tanh([c_{t-1} : x_t] \cdot W_c + b_c)$
$\Gamma_u = \sigma([c_{t-1} : x_t] \cdot W_u + b_u)$
$c_t = \Gamma_u \cdot \check{c}_t + (1 - \Gamma_u) \cdot c_{t-1}$
where $\Gamma_u = 1$, $c_t = \check{c}_t$,
where $\Gamma_u = 0$, $c_t = c_{t-1}$





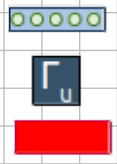
Gated recurrent units (GRUs)

Simple GRU:

$\check{c}_t = \tanh([c_{t-1} : x_t] \cdot W_c + b_c)$

$\Gamma_u = \sigma([c_{t-1} : x_t] \cdot W_u + b_u)$

$c_t = \Gamma_u \cdot \check{c}_t + (1 - \Gamma_u) \cdot c_{t-1}$



where $\Gamma_u = 1$, $c_t = \check{c}_t$,

where $\Gamma_u = 0$, $c_t = c_{t-1}$

Extended GRU:

$\check{c}_t = \tanh([\Gamma_r * c_{t-1} : x_t] \cdot W_c + b_c)$

$\Gamma_u = \sigma([c_{t-1} : x_t] \cdot W_u + b_u)$

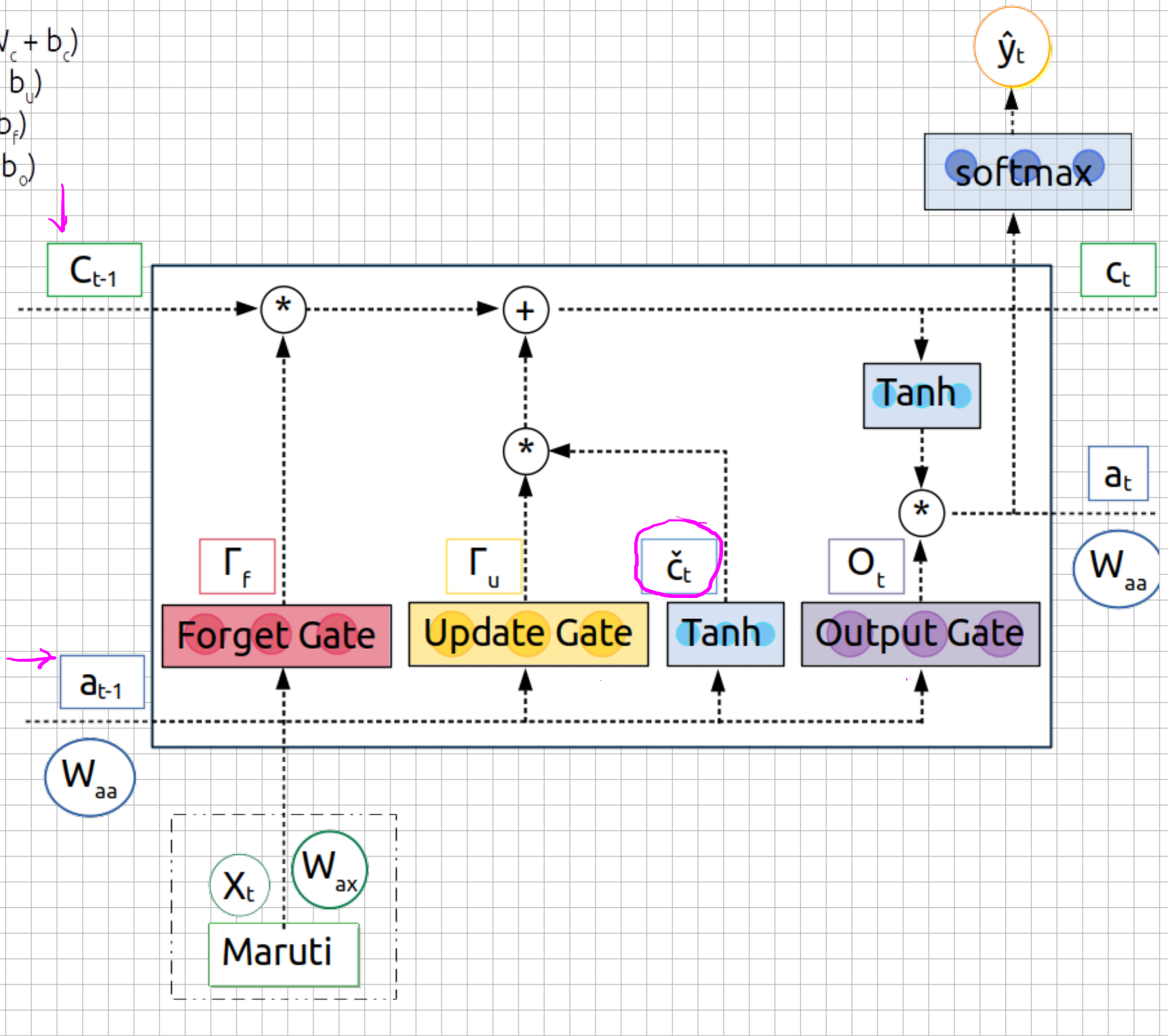
$\Gamma_r = \sigma([c_{t-1} : x_t] \cdot W_r + b_r)$

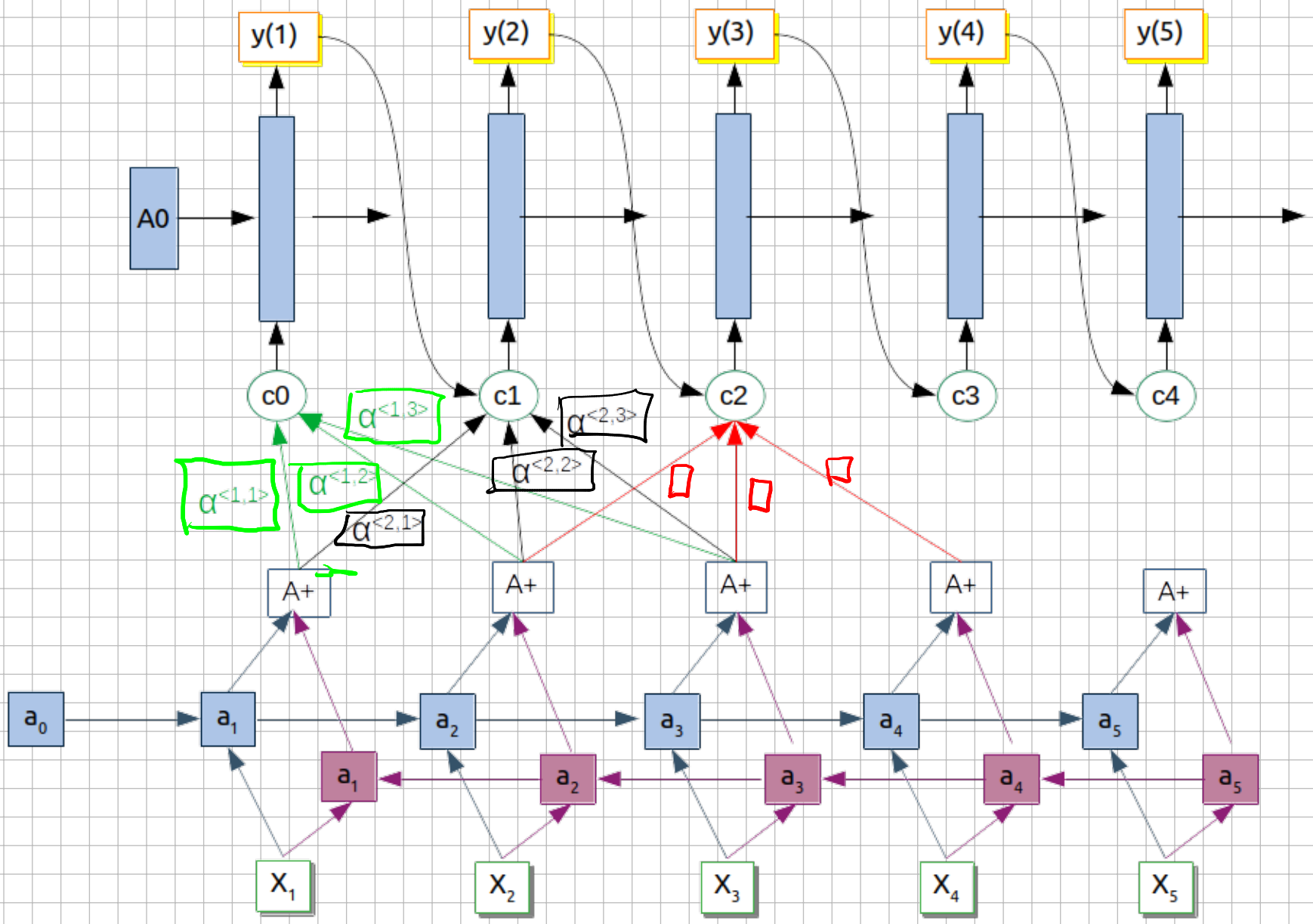
$c_t = \Gamma_u \cdot \check{c}_t + (1 - \Gamma_u) \cdot c_{t-1}$

Γ_r = how relevant C_{t-1} is in current context! Thus,

Γ_r is element wise multiplication to C_{t-1}.

$$\check{C}_t = \tanh([a_{t-1} : X_t] \cdot W_c + b_c)$$
$$\Gamma_u = \sigma([a_{t-1} : X_t] \cdot W_u + b_u)$$
$$\Gamma_f = \sigma([a_{t-1} : X_t] \cdot W_f + b_f)$$
$$\Gamma_o = \sigma([a_{t-1} : X_t] \cdot W_o + b_o)$$
$$C_t = \Gamma_u \cdot \check{C}_t + \Gamma_f \cdot C_{t-1}$$
$$a_t = \Gamma_o * \tanh C_t$$




y(1)
y(2)
y(3)
y(4)
y(5)
A0
c0
c1
c2
c3
c4
α<1,3>
α<2,3>
α<2,2>
α<1,1>
α<1,2>
α<2,1>
A+
A+
A+
A+
A+
a0
a1
a2
a3
a4
a5
a1
a2
a3
a4
a5
X1
X2
X3
X4
X5