Haonan Hu

3/9/2021

2863545

1. List the three inputs to a LSTM module and describe them.

   x_t: input feature vector for time t

   h_{t-1}: output vector from time t-1

   c_{t-1}: cell state from time t-1

2. List the two outputs from a LSTM module and describe them.

   c_t: cell state from time t

   h_t: output vector from time t

3. Which of the inputs to a LSTM module "remembers" previous inputs and outputs?

   c_{t-1}

4. Which of the outputs from a LSTM module "remembers" previous inputs and outputs?

   c_t

5. Which of the inputs to a LSTM module are used in the?
   a. forget gate function?

      x_t, h_{t-1}

b. input gate function?

$\boldsymbol{x_t, h_{t-1}}$

c. output gate function?

$\boldsymbol{x_t, h_{t-1}}$

6. Which of the inputs to a LSTM module are used in the cell state function (c_t)?

$\boldsymbol{c_{t-1}, h_{t-1}, x_t}$

7. Which of the inputs and outputs to a LSTM module are used in the output function (h_t)?

Need x_t and h_{t-1} to get o_t, Then use c_t and o_t to get h_t

8. If a is the vector [1, 4, 5] and b is the vector [2, 6, 9], what does a ∘ b =?

[1, 4, 5] ∘ [2, 6, 9] = [1 * 2, 4 * 6, 5 * 9] = [2, 24, 45]

9. Write the forget gate function in terms of the sigmoid function S(x) instead of σ_g.

$$\mathbf{S}(\boldsymbol{W_f x_t + U_f h_{t-1} + b_f}) = \frac{\mathbf{1}}{\left(\mathbf{1}+\boldsymbol{e}^{-(\boldsymbol{W_f X_t + U_f h_{t-1} + b_f})}\right)}$$

10. Write the cell state function in terms of the tanh function instead of σ_h.

$$C_t = f_t \circ c_{t-1} + i_t \circ \tanh(W_c x_t + U_c h_{t-1} + b_c)$$