

Let Encoder Hidden States

$$h_1 = [1, 2, 3]$$

$$h_2 = [3, 1, 1]$$

$$h_3 = [4, 2, 1]$$

and decoder state

$$s_1 = [3, 4, 5]$$

Find attention scores.

$$e^t = [s_t^T h_1, \dots, s_t^T h_n]$$

$$\begin{aligned} s_1^T h_1 &= \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix} \cdot [1, 2, 3] \\ &= 3 + 8 + 15 = 26 \end{aligned}$$

$$s_1^T h_2 = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix} \cdot [3, 1, 1]$$

$$= 9 + 4 + 5 = 18$$

$$s_1^T h_3 = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix} \cdot [4, 2, 1]$$

$$= 12 + 8 + 5 = 25$$

$$e^t = [26, 18, 25]$$

28431

$$\alpha' = \text{softmax}(e^t)$$

$$\alpha' = \left[\frac{e^{26}}{e^{26+18+25}}, \frac{e^{18}}{e^{26+18+25}}, \frac{e^{25}}{e^{26+18+25}} \right]$$

$$\alpha' = [0.73, 0.0002, 0.2688]$$

$$a_t = \sum_{i=1}^N \alpha_i^t h_i^o$$

$$= 0.73 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + 0.002 \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}$$

$$+ 0.2688 \begin{bmatrix} 4 \\ 2 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} \\ \\ \end{bmatrix}$$

// calculate this