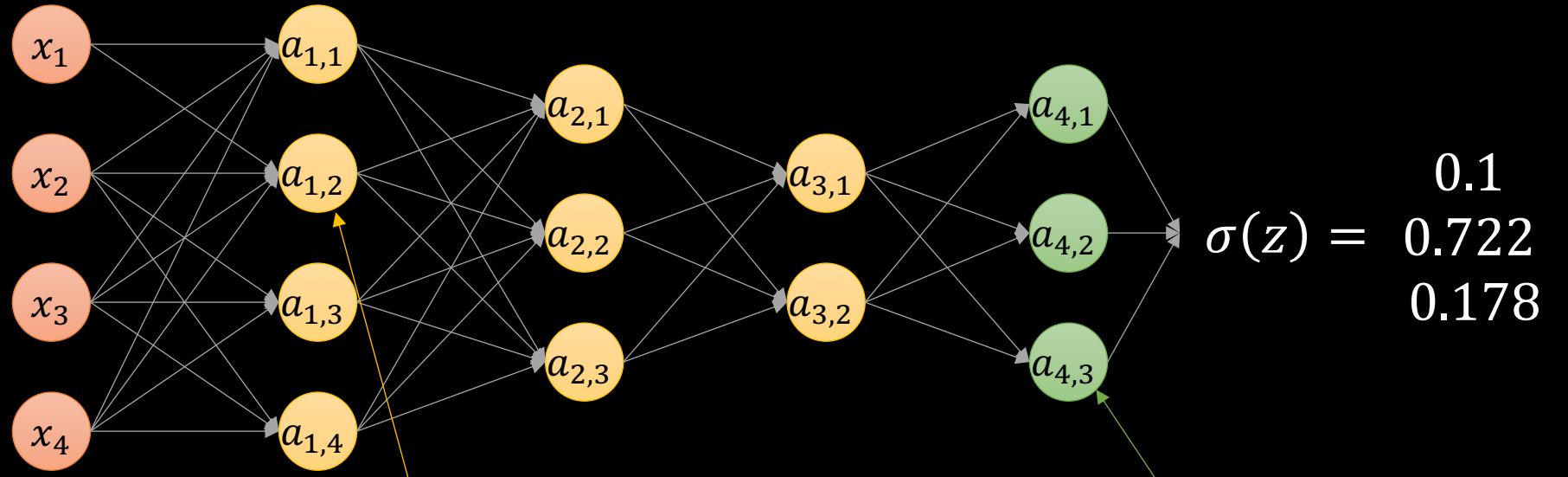


Generating Text with Recurrent Neural Networks

A 'fully connected' neural network



Input layer

Hidden layers

Output layer

$[4.9, 3., 1.4, 0.2]$

input == a set of numbers

$$a_{1,2} = w_1 \times x_1 + w_2 \times x_2 + w_3 \times x_3 + w_4 \times x_4 + b$$

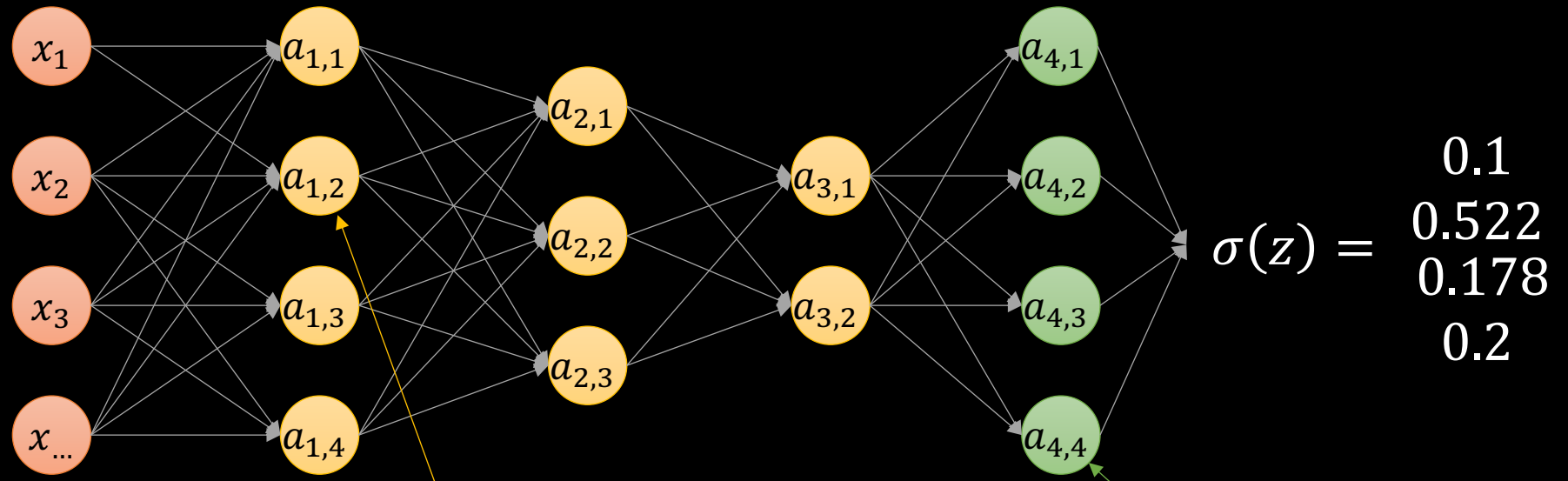
each neuron == a linear function of its inputs

$$a_{4,3} = w_1 \times a_{3,1} + w_2 \times a_{3,2} + b$$

also a neuron

Can we predict the next word in a sentence?

'cat'



Input layer

Hidden layers

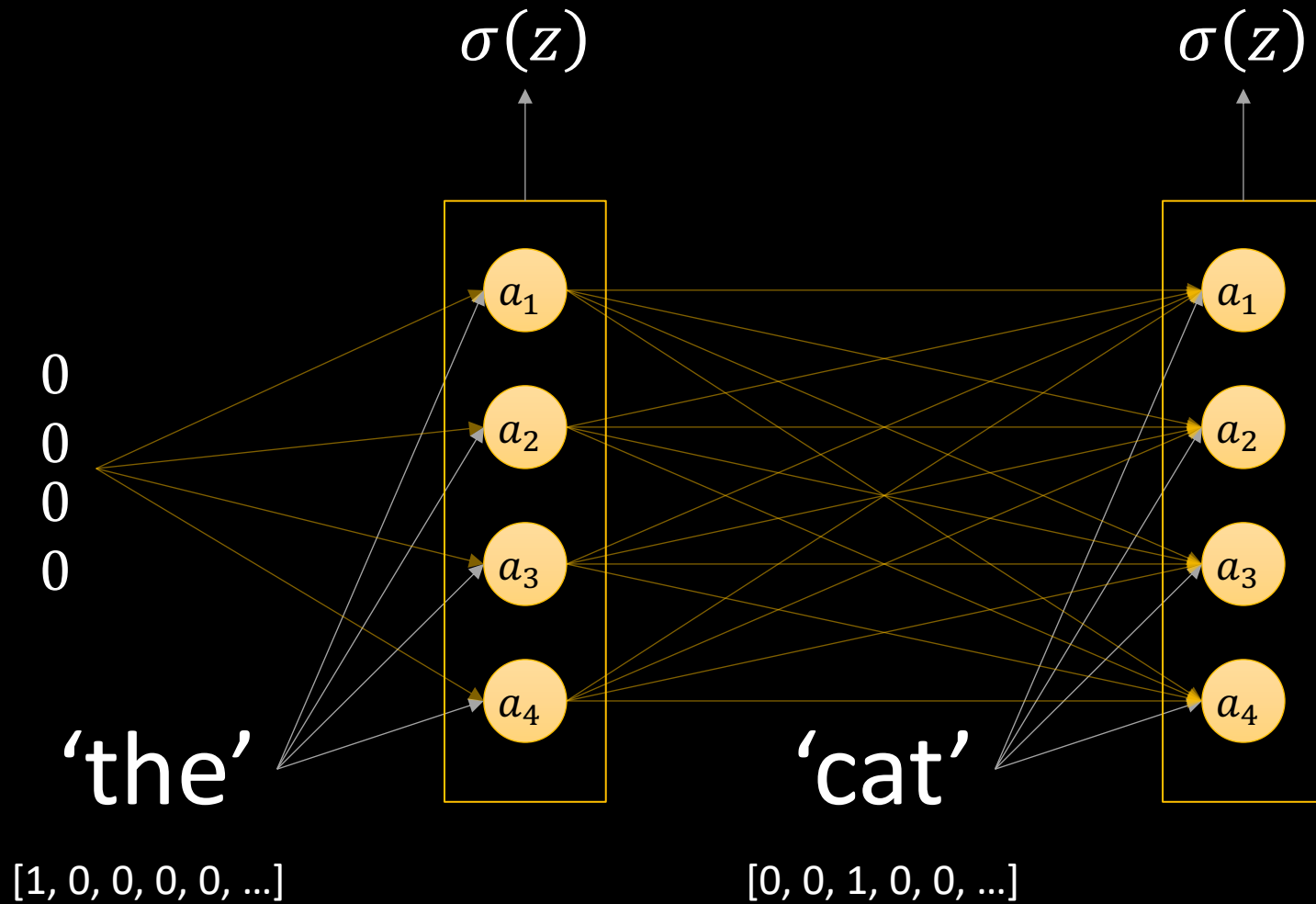
Output layer

$[0, 0, 1, 0, 0, 0, \dots]$

input == a 'one-hot' vector
'cat' is word number 3 in the
dictionary

$a_{1,2} = w_1 \times x_1 + w_2 \times x_2 + w_3 \times x_3 + w_4 \times x_4 + \dots + b$
each neuron == a linear function of its inputs
one 'w' for each word in the dictionary

one output neuron per
word in the dictionary



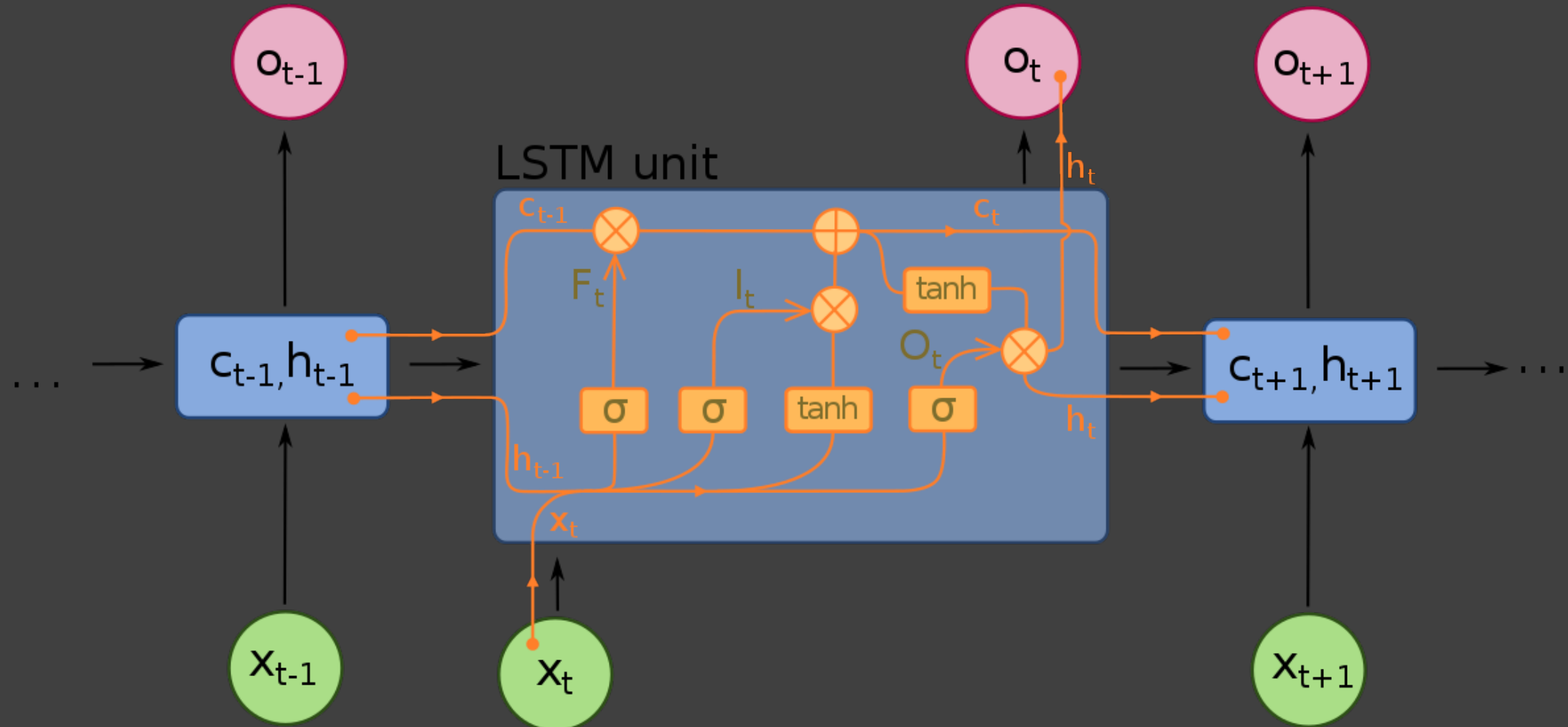
$$h = 4$$

$$v = 10000$$

There are 4 neurons in the RNN 'Cell', and 10000 words in the dictionary.

$$a_2 = \underbrace{w_1 \times a_{1<t-1>} + w_2 \times a_{2<t-1>} + w_3 \times a_{3<t-1>} + w_4 \times a_{4<t-1>}}_{\text{Four weights for the previous outputs of the four neurons}} + \underbrace{w_5 \times x_1 + w_6 \times x_2 + \dots + w_{10004} \times x_{10000}}_{\text{10000 weights for the words (x's are mostly 0)}}$$

Long short-term memory network



Training and Inference

