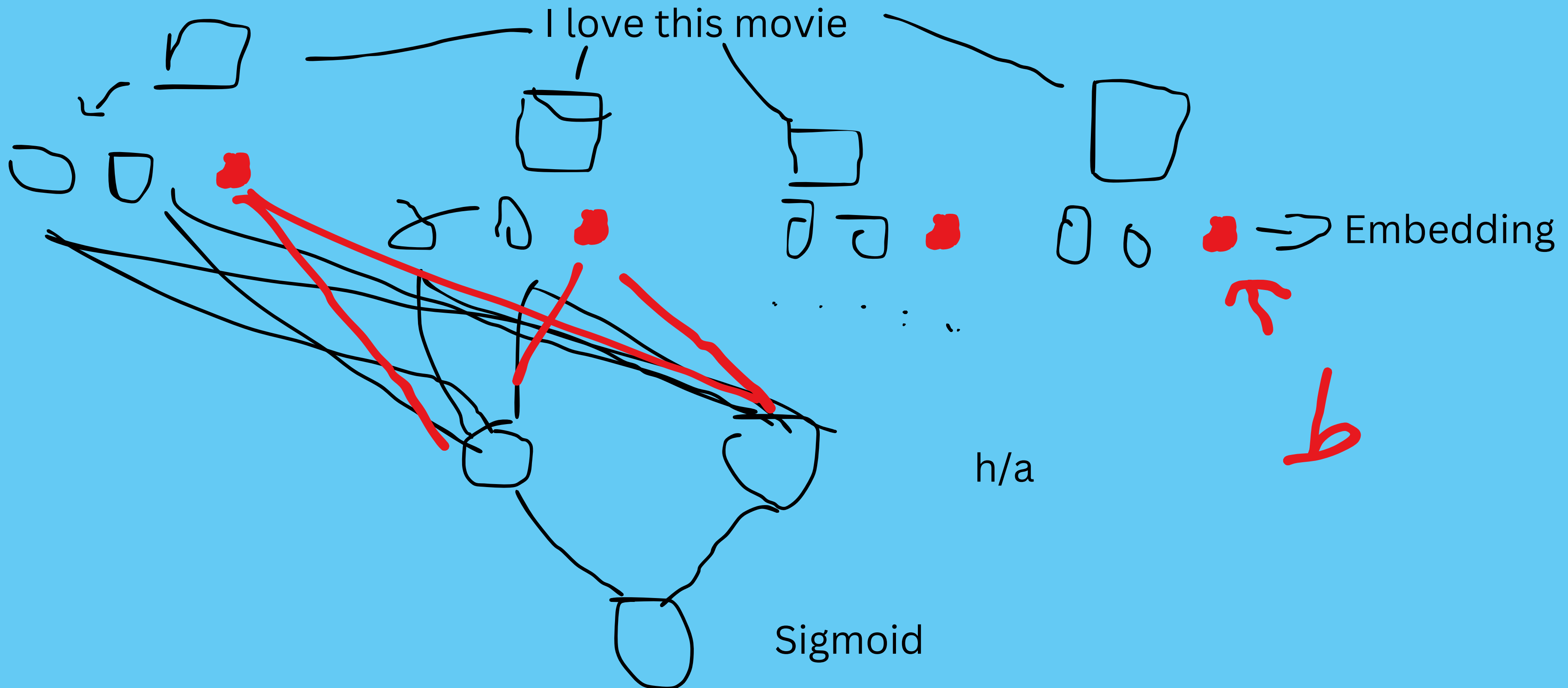


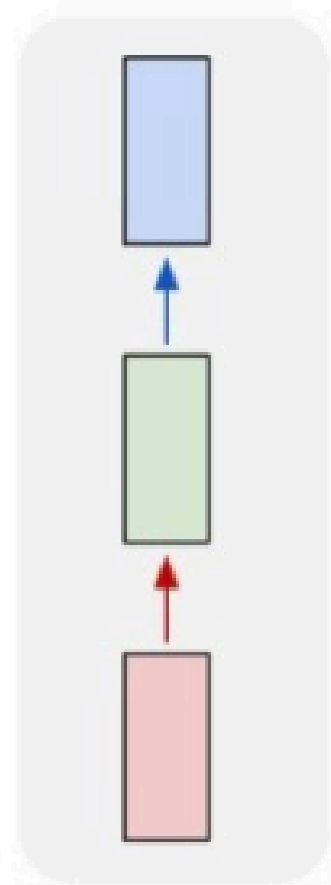
Recurrent Neural Network

) RNN (

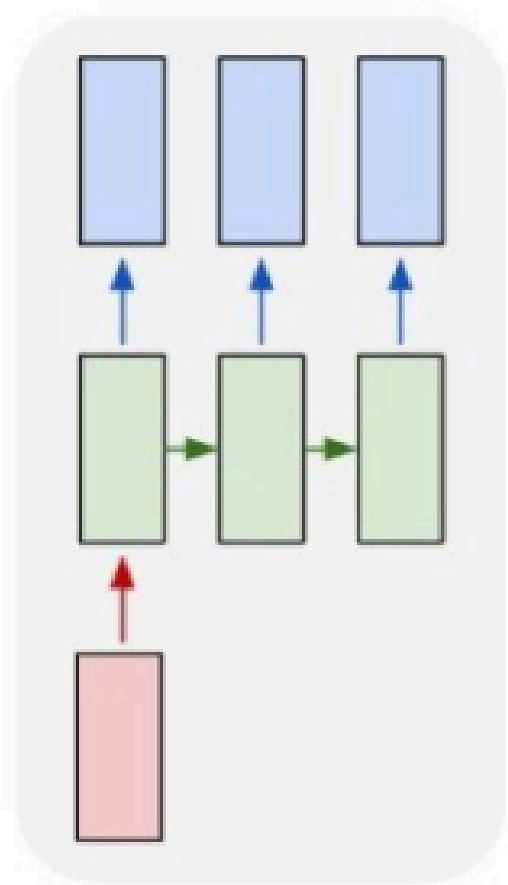


Phân loại bài toán RNN

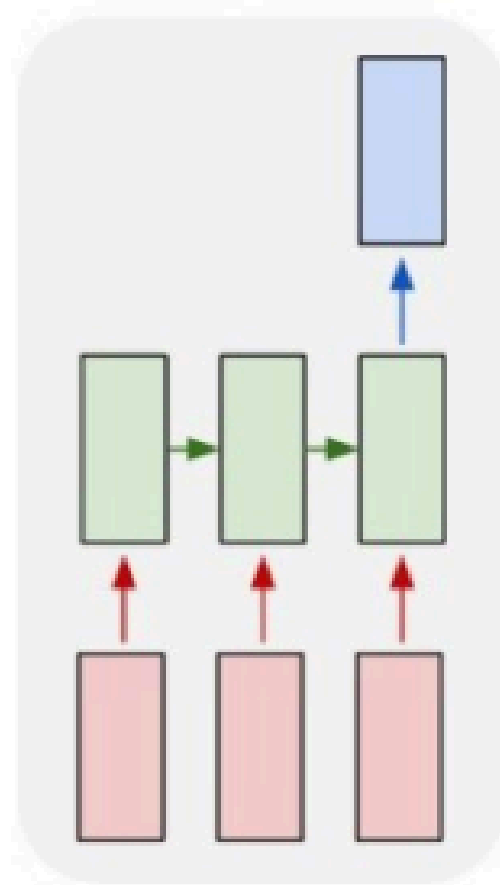
one to one



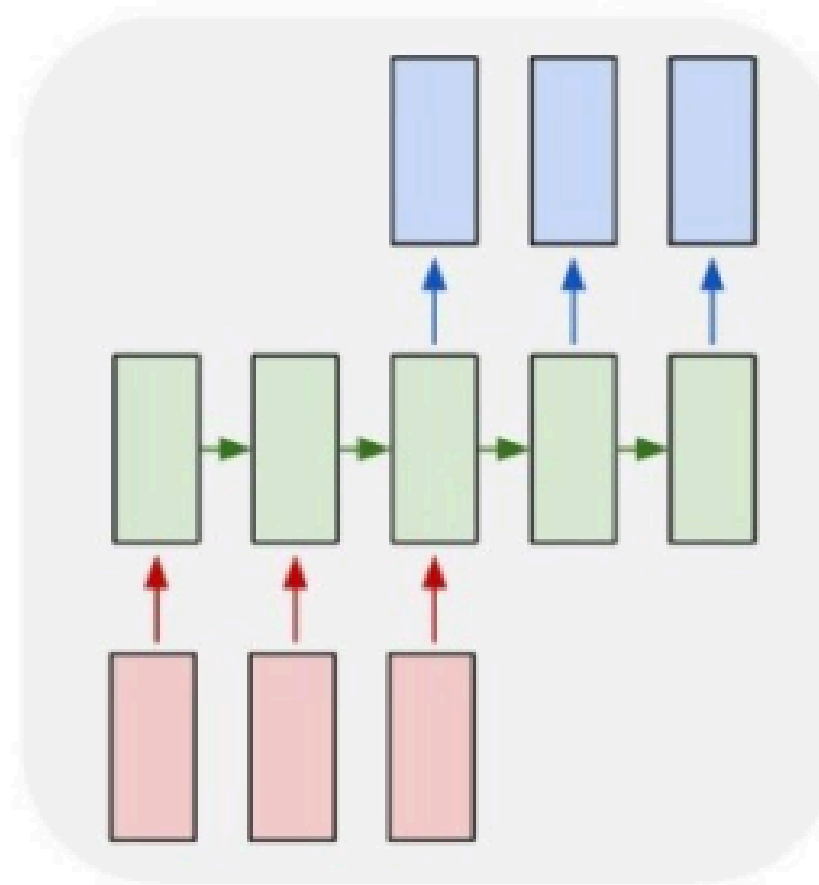
one to many



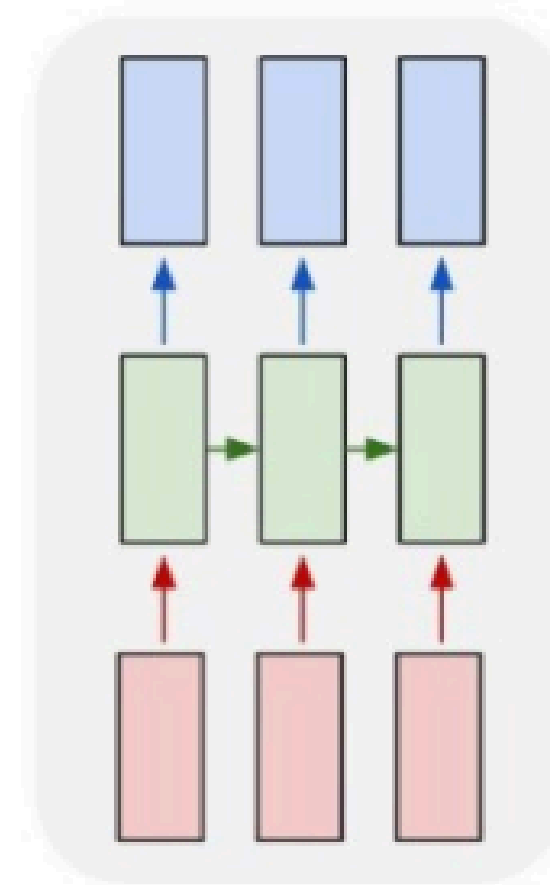
many to one



many to many

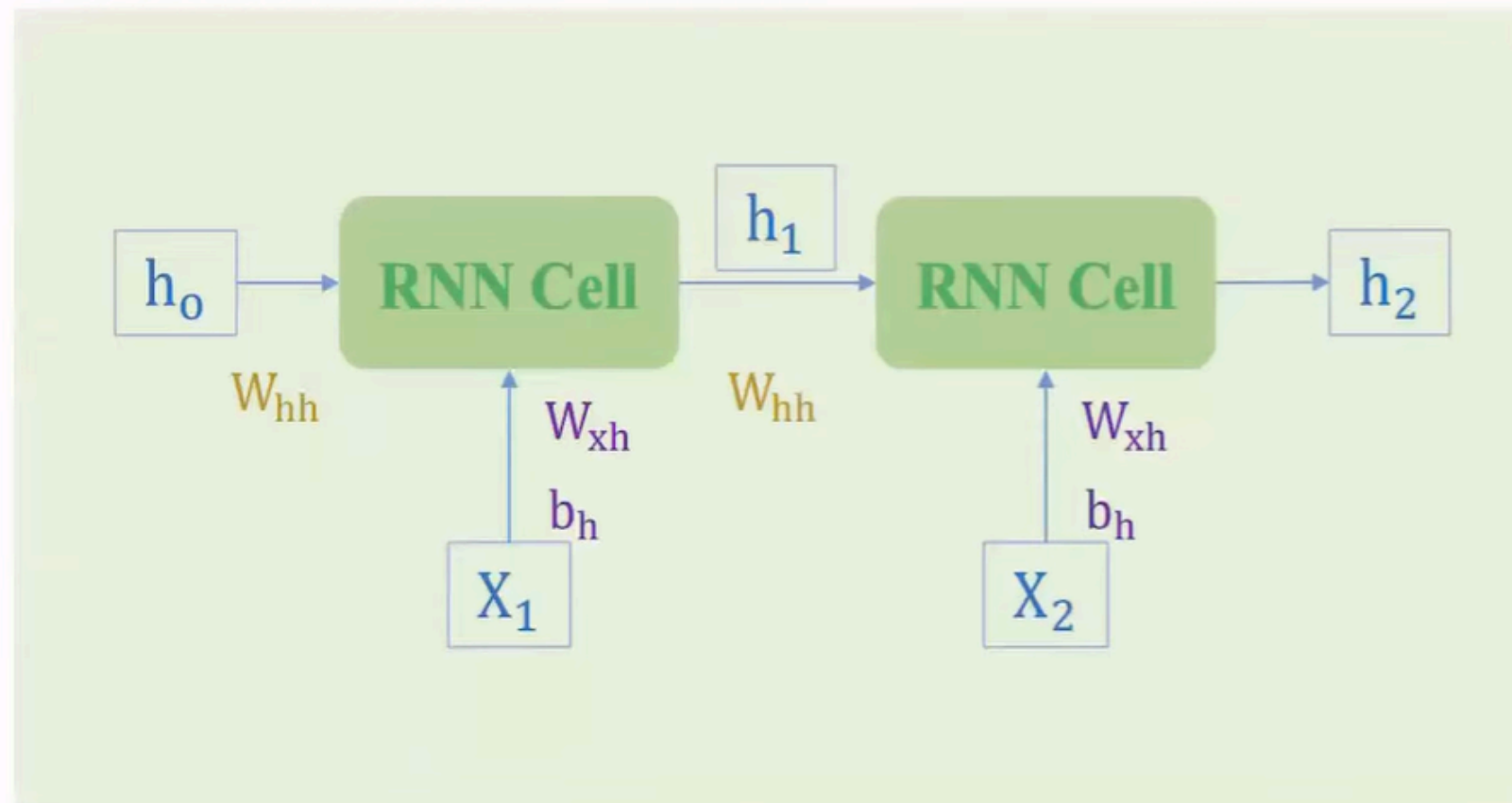


many to many



Example

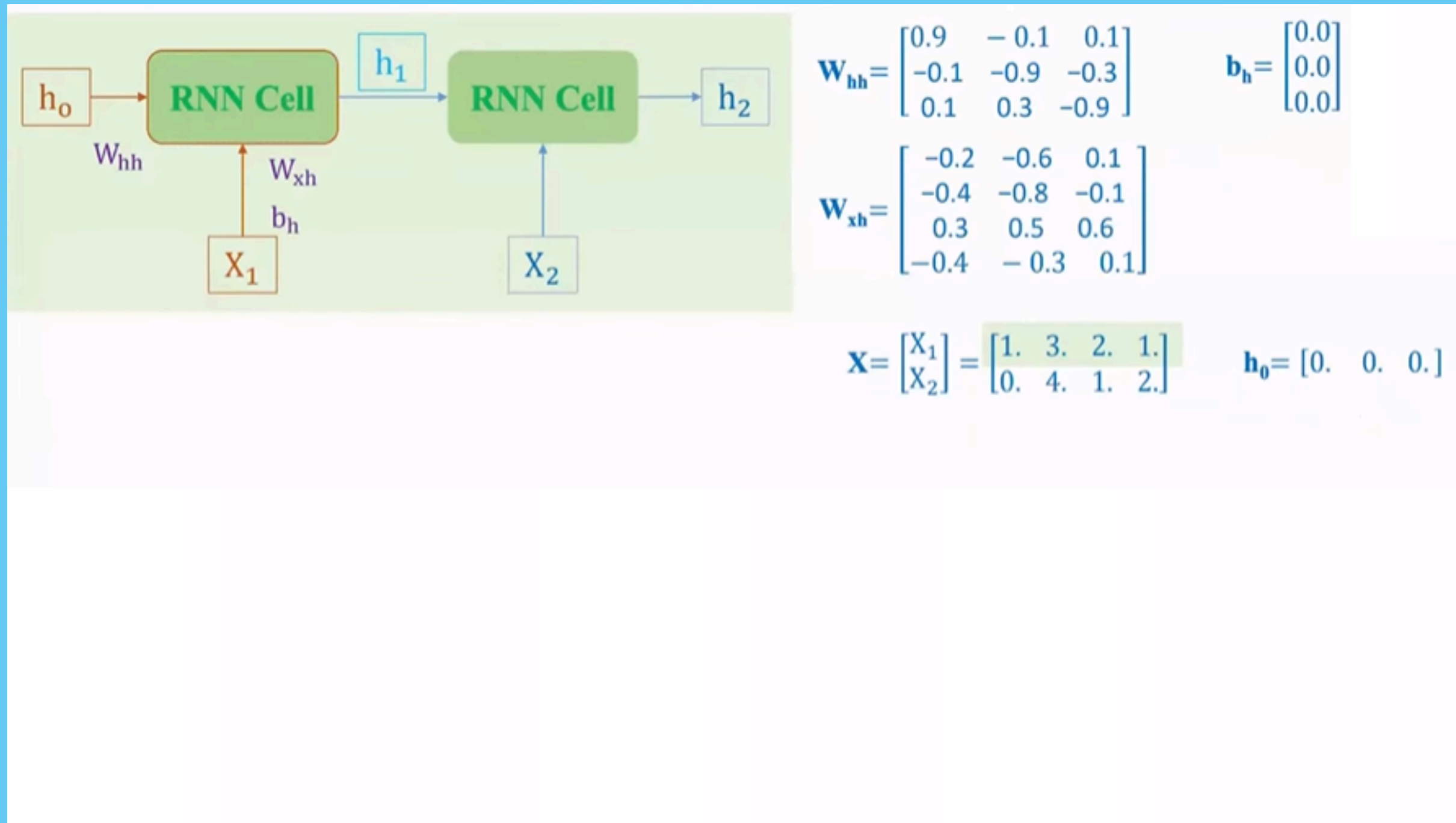
Classification and time-steps=2



$$h_1 = \tanh(X_1 W_{xh} + h_0 W_{hh} + b_h)$$

$$h_2 = \tanh(X_2 W_{xh} + h_1 W_{hh} + b_h)$$

Example



Example

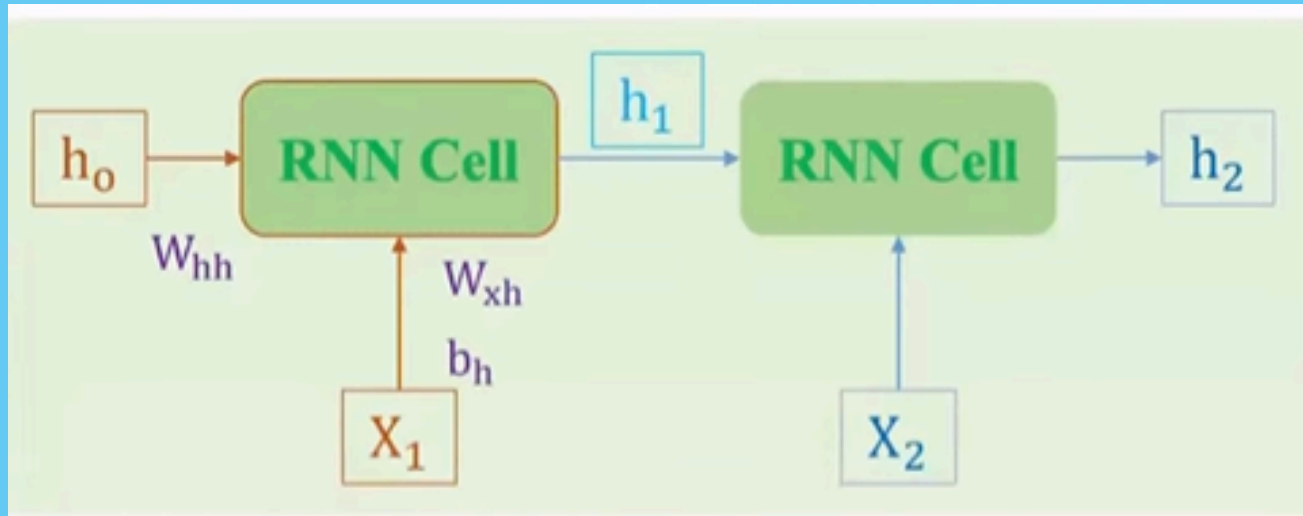


Diagram illustrating an unrolled Recurrent Neural Network (RNN) with two cells. The first cell takes an initial hidden state h_0 and input X_1 to produce hidden state h_1 . The second cell takes h_1 and input X_2 to produce hidden state h_2 . The weights W_{hh} and W_{xh} , and bias b_h are shown connecting the inputs to the cells.

Weights and bias:

$$W_{hh} = \begin{bmatrix} 0.9 & -0.1 & 0.1 \\ -0.1 & -0.9 & -0.3 \\ 0.1 & 0.3 & -0.9 \end{bmatrix} \quad b_h = \begin{bmatrix} 0.0 \\ 0.0 \\ 0.0 \end{bmatrix}$$

$$W_{xh} = \begin{bmatrix} -0.2 & -0.6 & 0.1 \\ -0.4 & -0.8 & -0.1 \\ 0.3 & 0.5 & 0.6 \\ -0.4 & -0.3 & 0.1 \end{bmatrix}$$

Input matrix X and initial hidden state h_0 :

$$X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} 1. & 3. & 2. & 1. \\ 0. & 4. & 1. & 2. \end{bmatrix} \quad h_0 = [0. \quad 0. \quad 0.]$$

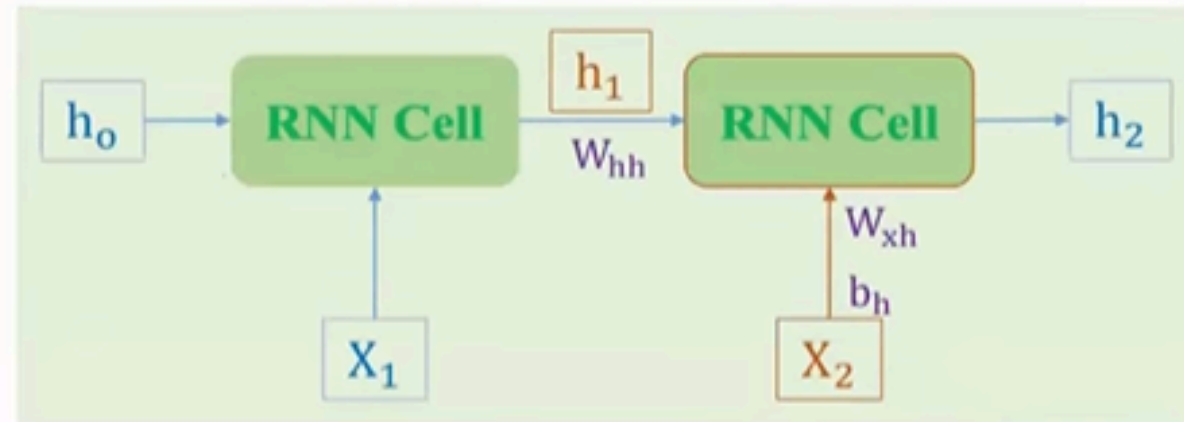
Calculation for h_1 :

$$h_1 = \tanh(X_1 W_{xh} + h_0 W_{hh} + b_h)$$

$$= \tanh \left([1. \quad 3. \quad 2. \quad 1.] \begin{bmatrix} -0.2 & -0.6 & 0.1 \\ -0.4 & -0.8 & -0.1 \\ 0.3 & 0.5 & 0.6 \\ -0.4 & -0.3 & 0.1 \end{bmatrix} + [0. \quad 0. \quad 0.] \begin{bmatrix} 0.9 & -0.1 & 0.1 \\ -0.1 & -0.9 & -0.3 \\ 0.1 & 0.3 & -0.9 \end{bmatrix} + \begin{bmatrix} 0.0 \\ 0.0 \\ 0.0 \end{bmatrix} \right)$$

$$= \tanh([-1.2 \quad -2.3 \quad 1.1]) = [-0.833 \quad -0.98 \quad 0.8]$$

Example



$$W_{hh} = \begin{bmatrix} 0.9 & -0.1 & 0.1 \\ -0.1 & -0.9 & -0.3 \\ 0.1 & 0.3 & -0.9 \end{bmatrix}$$

$$b_h = \begin{bmatrix} 0.0 \\ 0.0 \\ 0.0 \end{bmatrix}$$

$$W_{xh} = \begin{bmatrix} -0.2 & -0.6 & 0.1 \\ -0.4 & -0.8 & -0.1 \\ 0.3 & 0.5 & 0.6 \\ -0.4 & -0.3 & 0.1 \end{bmatrix}$$

$$h_0 = [0. \quad 0. \quad 0.]$$

$$h_1 = [-0.833 \quad -0.98 \quad 0.8]$$

$$X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} 1. & 3. & 2. & 1. \\ 0. & 4. & 1. & 2. \end{bmatrix}$$

$$h_2 = \tanh(X_2 W_{xh} + h_1 W_{hh} + b_h)$$

$$= \tanh \left([0. \quad 4. \quad 1. \quad 2.] \begin{bmatrix} -0.2 & -0.6 & 0.1 \\ -0.4 & -0.8 & -0.1 \\ 0.3 & 0.5 & 0.6 \\ -0.4 & -0.3 & 0.1 \end{bmatrix} + [-0.833 \quad -0.98 \quad 0.8] \begin{bmatrix} 0.9 & -0.1 & 0.1 \\ -0.1 & -0.9 & -0.3 \\ 0.1 & 0.3 & -0.9 \end{bmatrix} + \begin{bmatrix} 0.0 \\ 0.0 \\ 0.0 \end{bmatrix} \right)$$

$$= \tanh([-2.67 \quad -2.09 \quad -0.109]) = [-0.99 \quad -0.97 \quad -0.109]$$

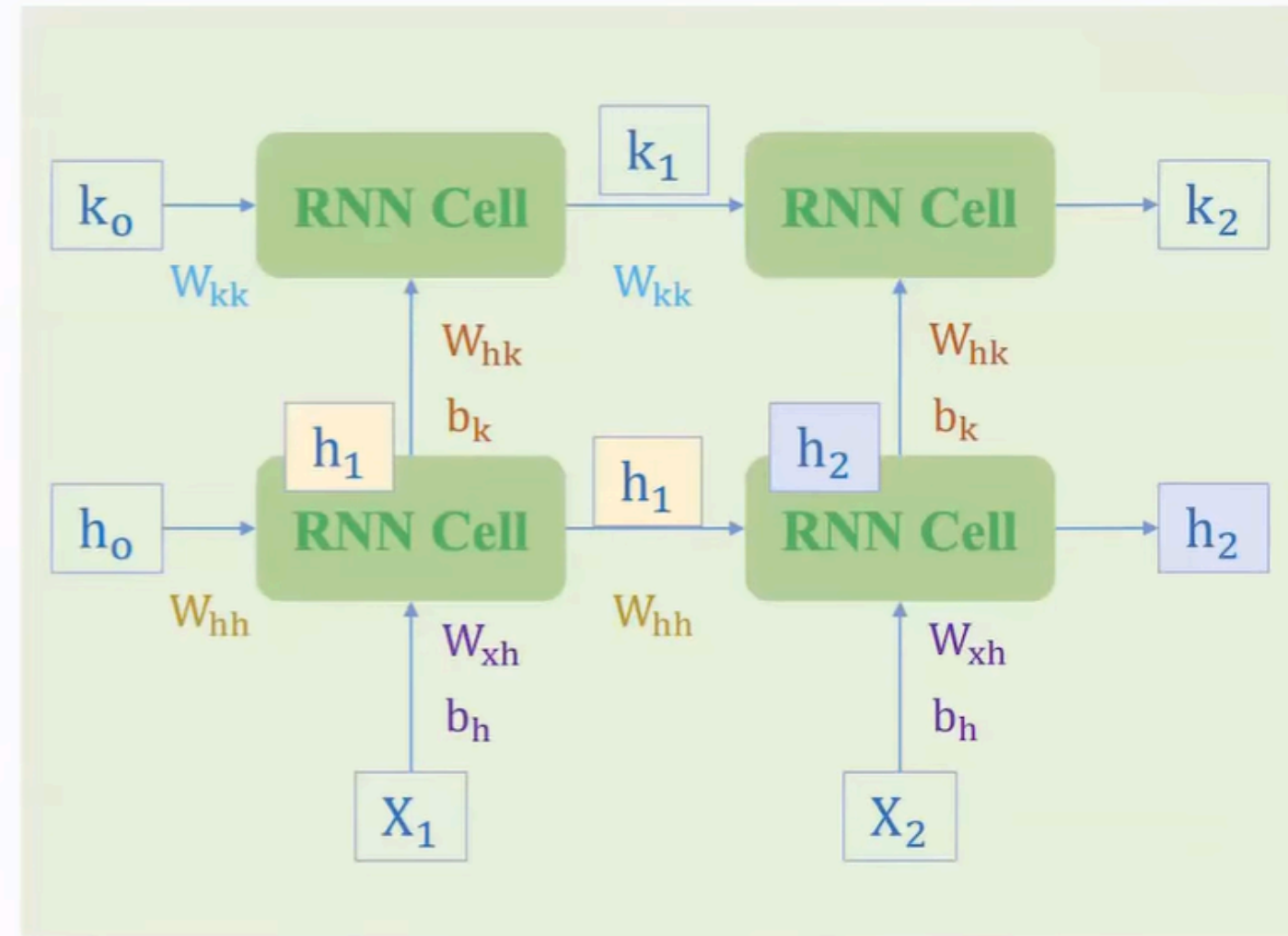
Multi layer

$$k_1 = \tanh(h_1 W_{hk} + h_0 W_{kk} + b_k)$$

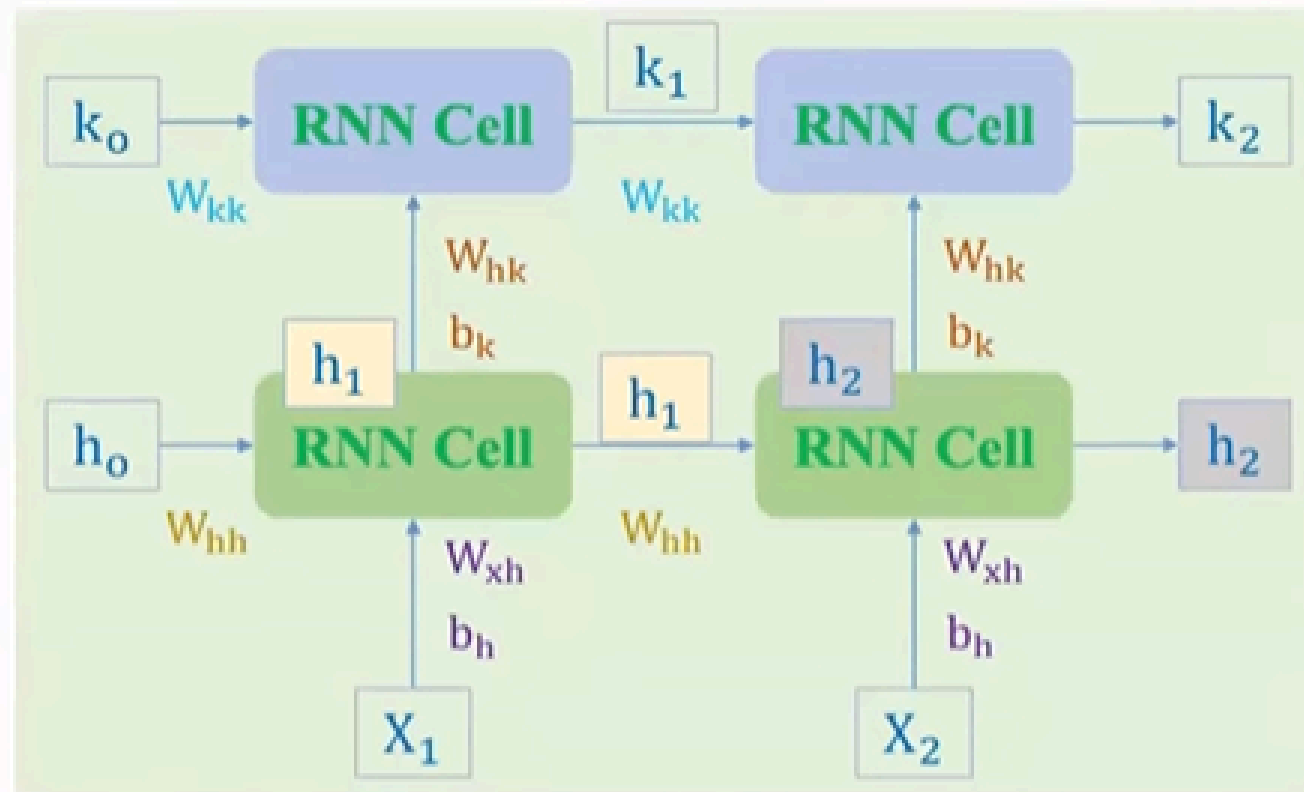
$$k_2 = \tanh(h_2 W_{hk} + h_1 W_{kk} + b_k)$$

$$h_1 = \tanh(X_1 W_{xh} + h_0 W_{hh} + b_h)$$

$$h_2 = \tanh(X_2 W_{xh} + h_1 W_{hh} + b_h)$$



Example



$$W_{hh} = \begin{bmatrix} 0.9 & -0.1 & 0.1 \\ -0.1 & -0.9 & -0.3 \\ 0.1 & 0.3 & -0.9 \end{bmatrix}$$

$$W_{xh} = \begin{bmatrix} -0.2 & -0.6 & 0.1 \\ -0.4 & -0.8 & -0.1 \\ 0.3 & 0.5 & 0.6 \\ -0.4 & -0.3 & 0.1 \end{bmatrix}$$

$$b_h = \begin{bmatrix} 0.0 \\ 0.0 \\ 0.0 \end{bmatrix}$$

$$W_{kk} = \begin{bmatrix} 0.1 & 0.9 \\ 0.9 & -0.1 \end{bmatrix}$$

$$W_{hk} = \begin{bmatrix} -1.1 & 0.1 \\ 0.3 & -0.2 \\ -0.1 & 0.4 \end{bmatrix}$$

$$b_k = \begin{bmatrix} 0.0 \\ 0.0 \\ 0.0 \end{bmatrix}$$

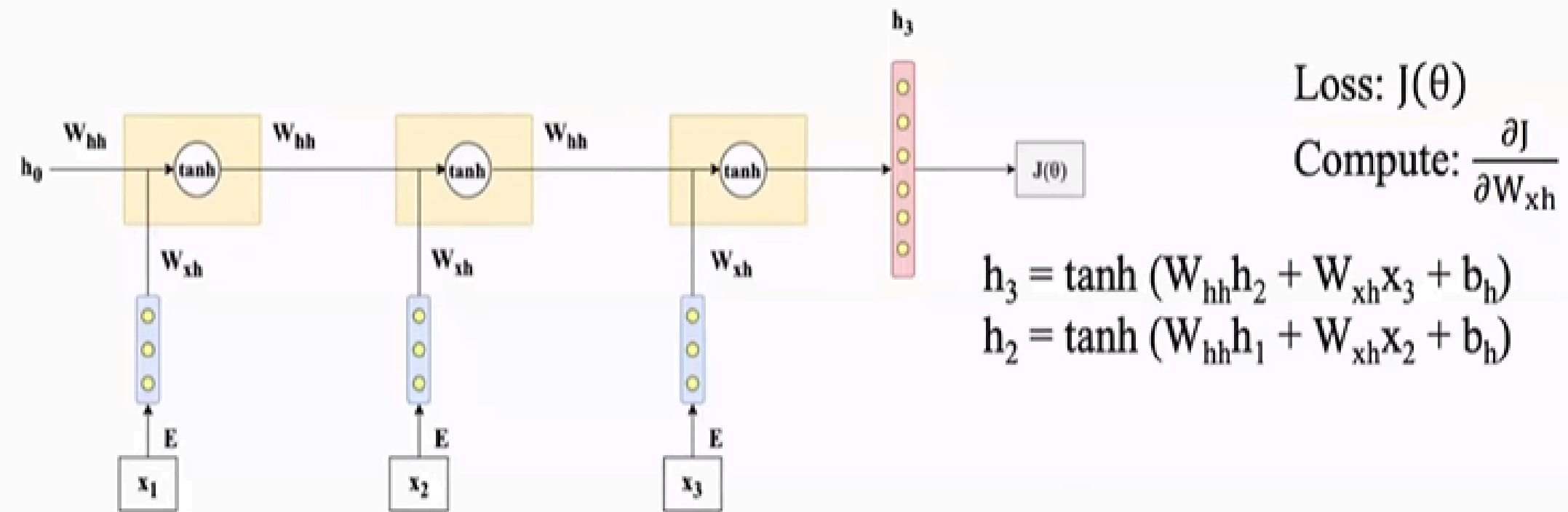
$$k_1 = \tanh(h_1 W_{hk} + h_0 W_{kk} + b_k) \\ = [0.49 \quad 0.407]$$

$$k_2 = \tanh(h_2 W_{hk} + h_1 W_{kk} + b_k) \\ = [0.84 \quad 0.42]$$

$$h_1 = \tanh(X_1 W_{xh} + h_0 W_{hh} + b_h) \\ = [-0.833 \quad -0.98 \quad 0.8]$$

$$h_2 = \tanh(X_2 W_{xh} + h_1 W_{hh} + b_h) \\ = [-0.99 \quad -0.97 \quad -0.109]$$

Backpropagation



$$\frac{\partial J}{\partial W_{xh}} = \frac{\partial J}{\partial h_3} \frac{\partial h_3}{\partial W_{xh}} + \frac{\partial J}{\partial h_3} \frac{\partial h_3}{\partial h_2}$$

Backpropagation

