

# AI Deep Learning: Recurrent Neural Networks

Thuan L Nguyen, PhD

## Slide 2: AI Deep Learning: Recurrent Neural Networks (RNN)



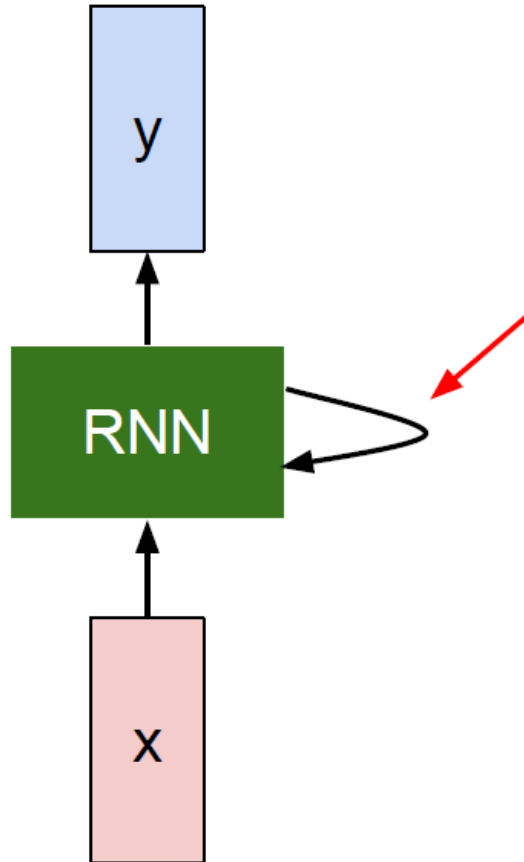
*AI Deep learning (Source: mindovermachines.com)*

## **Slide 3: AI Deep Learning: Recurrent Neural Networks (RNN)**

1. Recurrent Neural Networks: Overview: Sequence Data
2. Recurrent Neural Networks : Overview: Memory
3. Recurrent Neural Networks : Overview: Introduction
4. Recurrent Neural Networks: Overview: Mathematical Model
5. Recurrent Neural Networks: Overview: Simple RNN (Vanilla RNN)
6. Recurrent Neural Networks: Overview: Examples and Applications

# Slide 4: AI Deep Learning: Recurrent Neural Networks (RNN)

## Recurrent Neural Networks: Fundamentals



Key idea: RNNs have an “internal state” that is updated as a sequence is processed

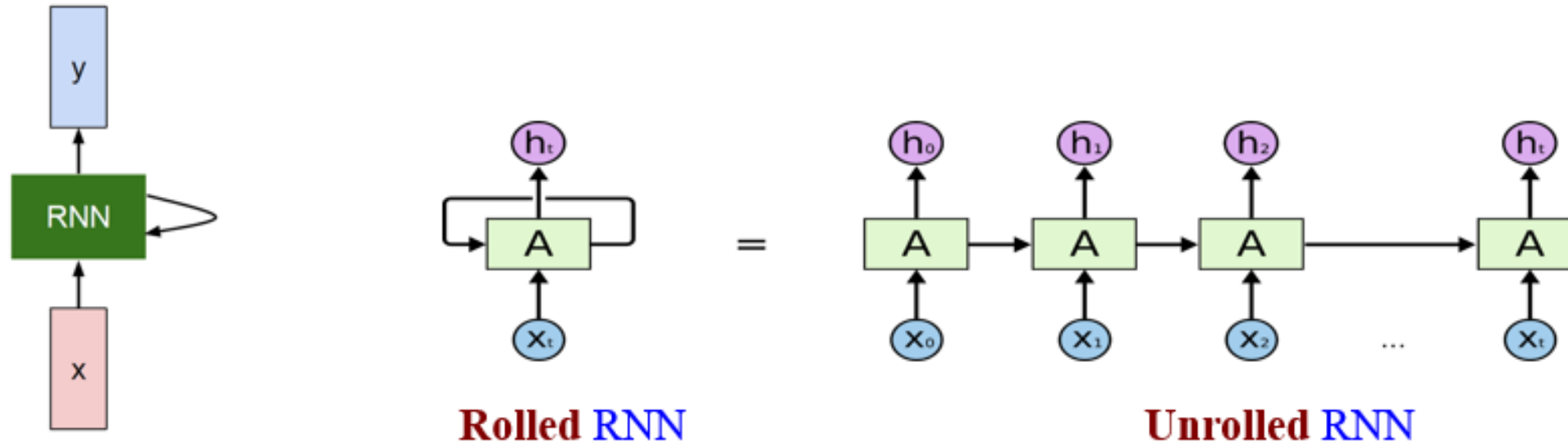
Key words:

- An **internal** (or “**hidden**”) **state** is **updated**
- When a **sequence** is **processed**

*Recurrent Neural Network (Source: Stanford.edu)*

# Slide 5: AI Deep Learning: Recurrent Neural Networks (RNN)

## RNN: Recurrent Neural Network: Rolled & Unrolled



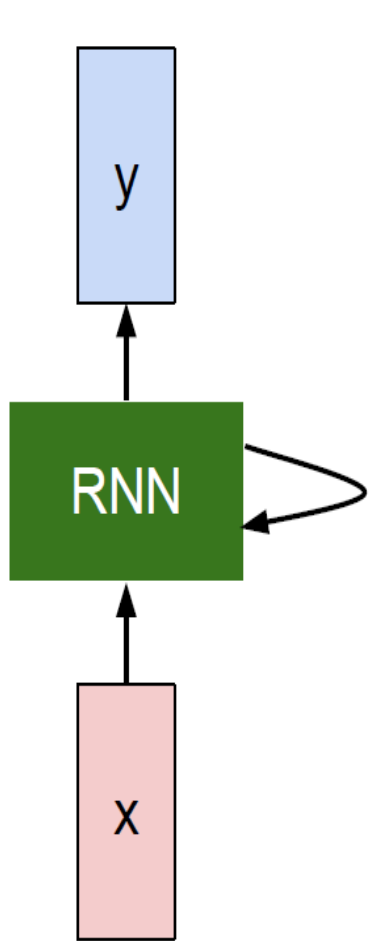
*Rolled and Unrolled RNN (Source: Stanford.edu and Colah Blogs)*

The **fundamental feature** of a Recurrent Neural Network (RNN):

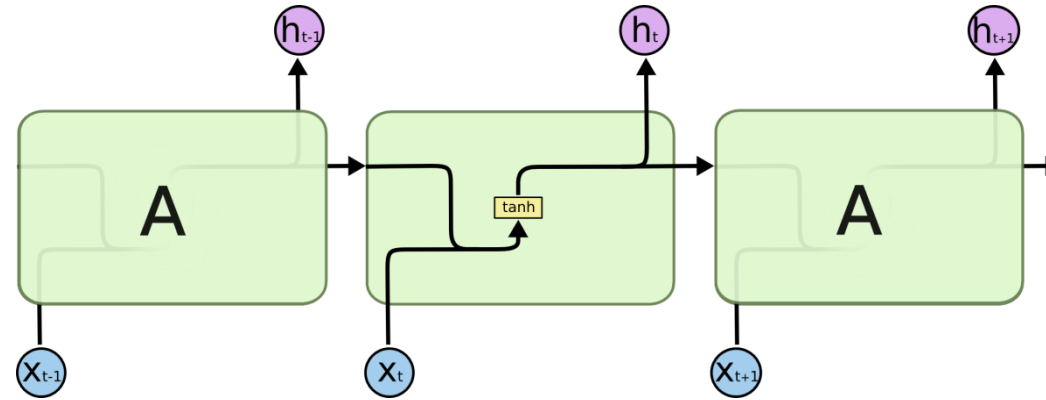
- The **input sequence  $x$**  can be processed by applying a **recurrent formula** at each step.
- In other words, the **same function** and the **same set of parameters** can be used at **each step** of processing the input sequence.

# Slide 6: AI Deep Learning: Recurrent Neural Networks (RNN)

## RNN: Simple Recurrent Neural Network



*Recurrent Neural Network  
(Source: Stanford.edu)*



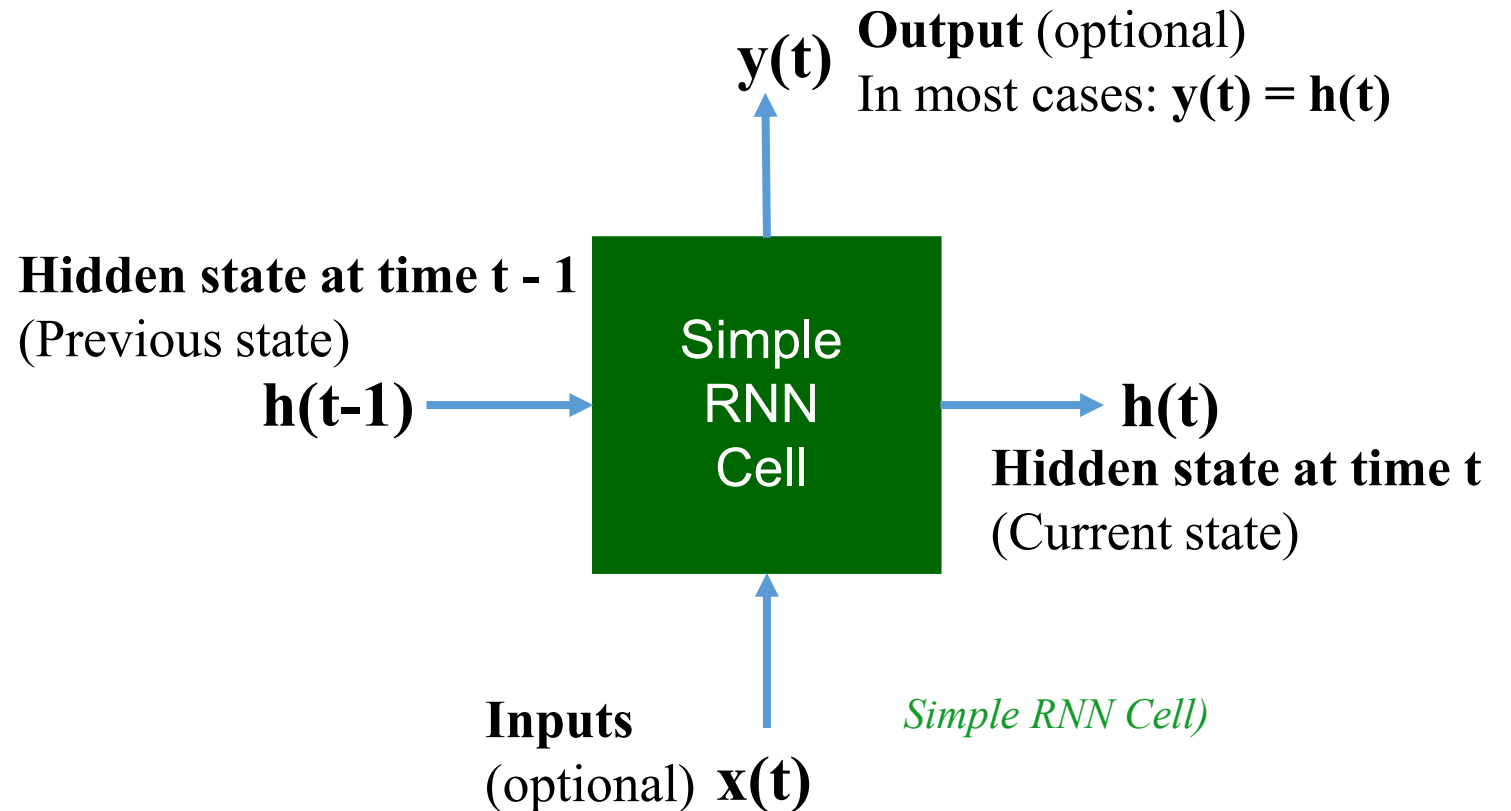
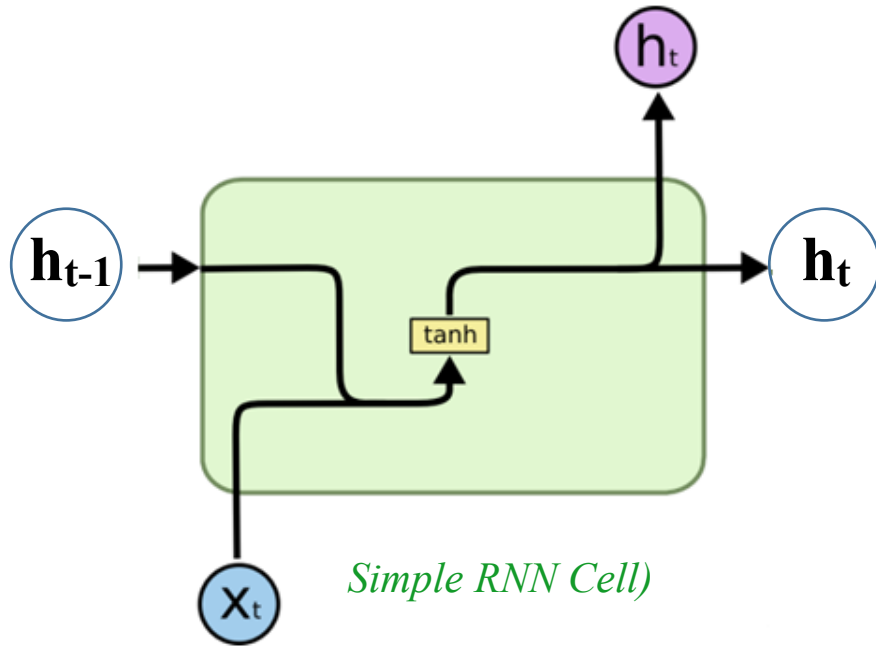
*Simple RNN (Source: Colah Blogs)*

### Simple RNN:

- The **state** consists of only **one “hidden” vector  $h$** .
- It is just a **single-layer** neural network (with feedback)
- Also called:
  - A “**Vanilla**” recurrent neural network.
  - OR an “**Elman RNN**” after Prof. Jeffrey Elman

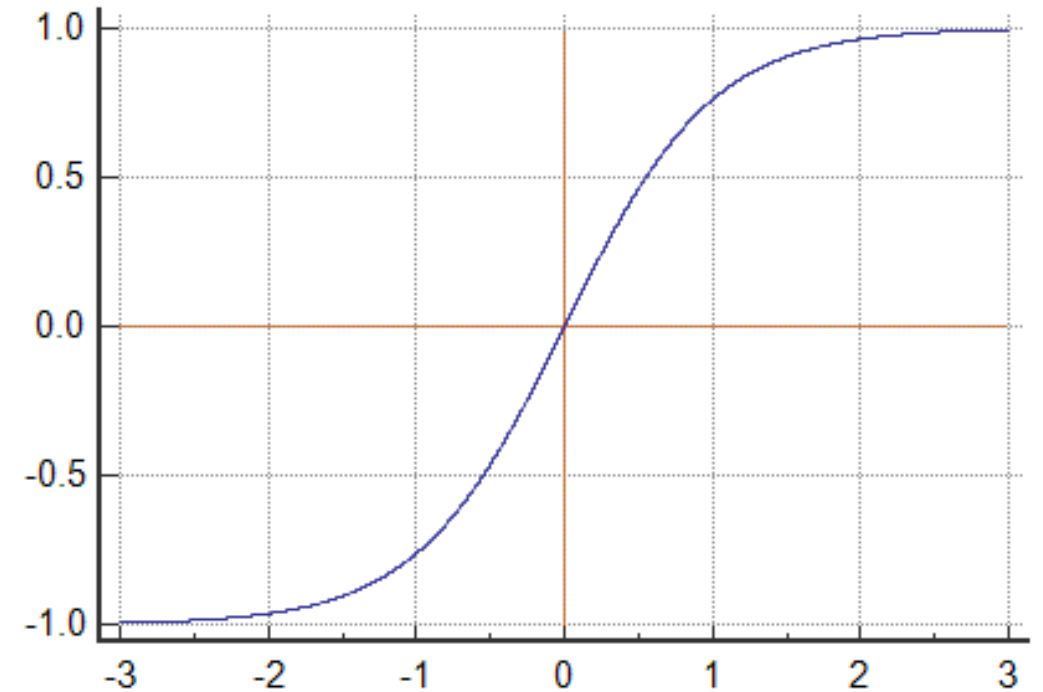
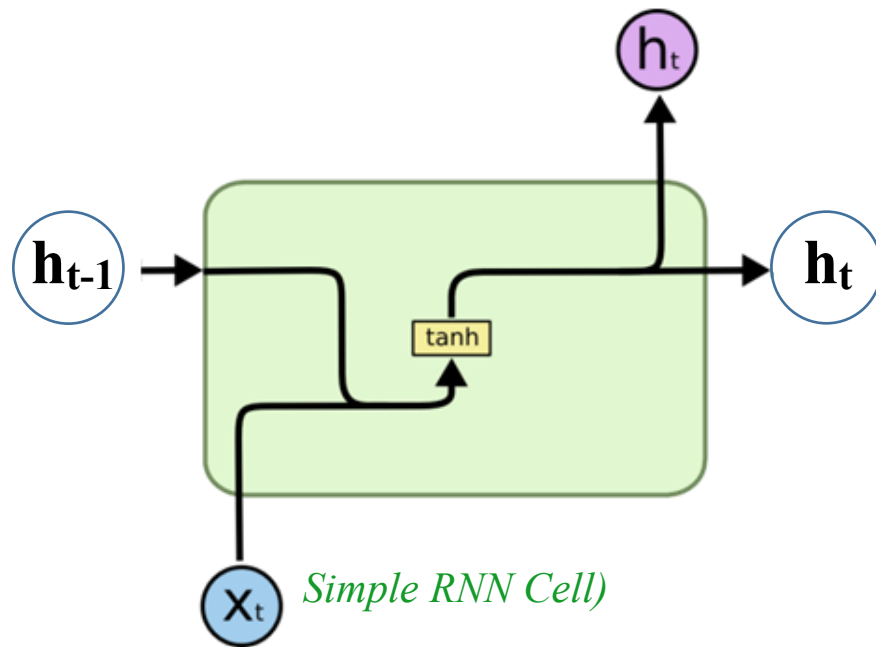
# Slide 7: AI Deep Learning: Recurrent Neural Networks (RNN)

## RNN: Simple Recurrent Neural Network: Simple RNN Cell



# Slide 8: AI Deep Learning: Recurrent Neural Networks (RNN)

## RNN: Simple Recurrent Neural Network: Mathematical Model



Activation function: **Tanh**

Photo source: [medcalc.org](http://medcalc.org)



# Slide 9: AI Deep Learning: Recurrent Neural Networks (RNN)

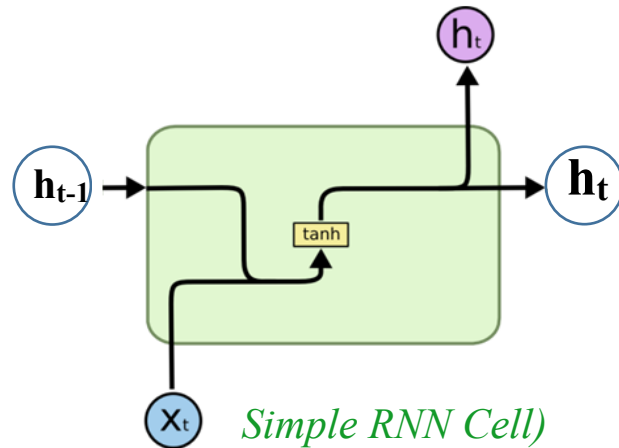
## RNN: Simple Recurrent Neural Network: Mathematical Model

$$h_t = f_W(h_{t-1}, x_t)$$

$$y_t = W_{hy}h_t$$

Applying the activation function tanh:

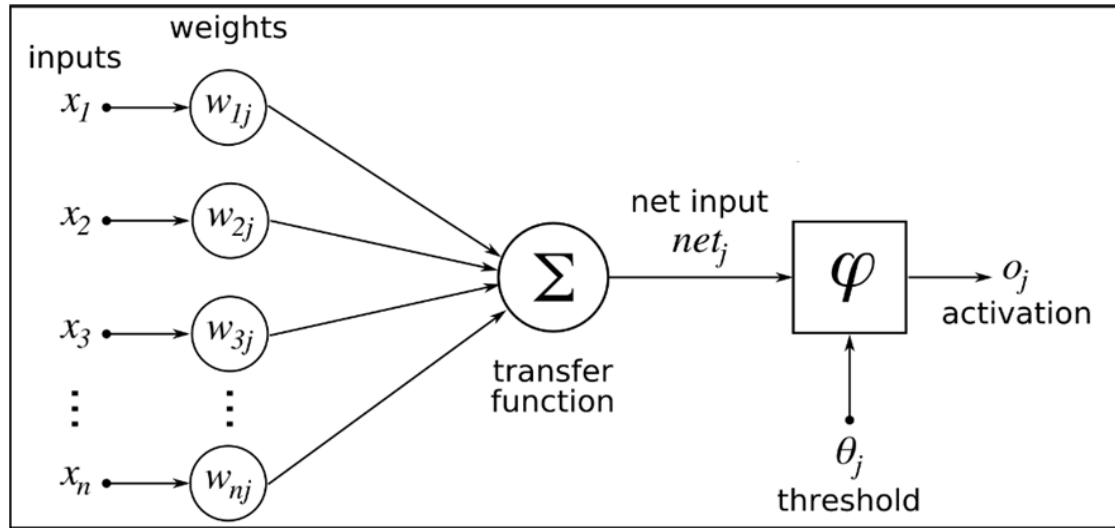
$$\mathbf{h}_t = \tanh(\mathbf{W}_{hh} * \mathbf{h}_{t-1} + \mathbf{W}_{hx} * \mathbf{x}_t)$$



- **W** is weight.
- **h** is the single hidden vector.
- **W<sub>hh</sub>** is the weight at the previous hidden state.
- **W<sub>hx</sub>** is the weight at the current inputs.
- **tanh** is the Activation function.
  - It implements a **Non-linearity** that squashes the activations to the range[-1.1]
- **W<sub>hy</sub>** is the weight at the outputs.
- **h<sub>t-1</sub>** is the previous hidden state
- **h<sub>t</sub>** is the current hidden state.

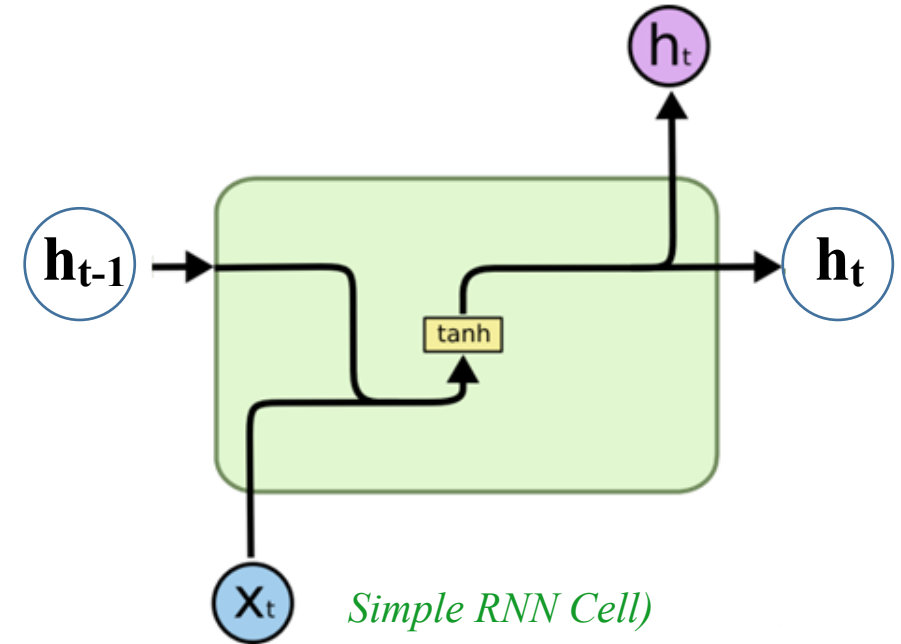
# Slide 10: AI Deep Learning: Recurrent Neural Networks (RNN)

## RNN: Simple Recurrent Neural Network: Simple RNN Cell



*McCulloch-Pitts Neuron Model (Sources: Wikipedia)*

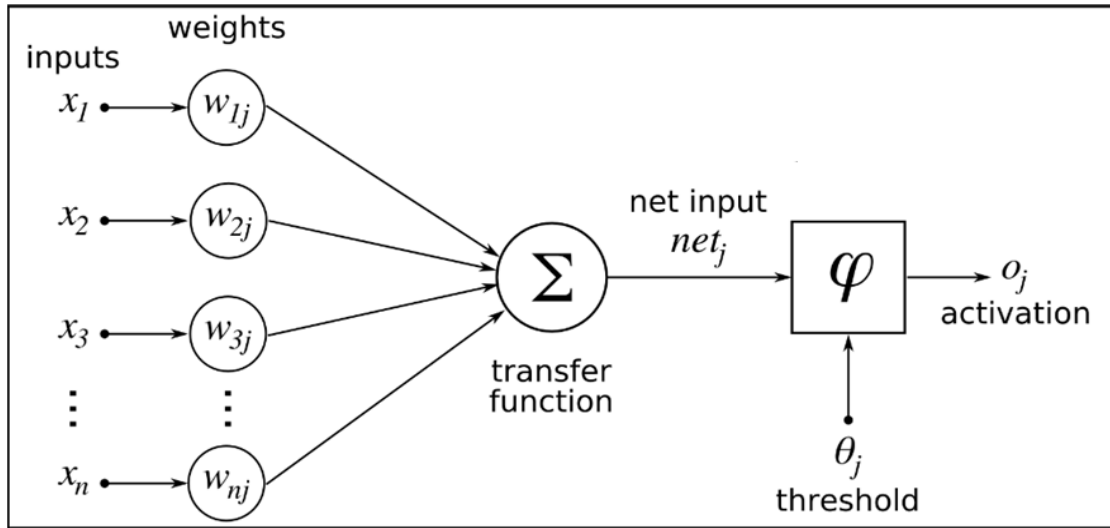
$$\mathbf{h}_t = \tanh (\mathbf{W}_{hh} * \mathbf{h}_{t-1} + \mathbf{W}_{hx} * \mathbf{x}_t)$$



*Simple RNN Cell)*

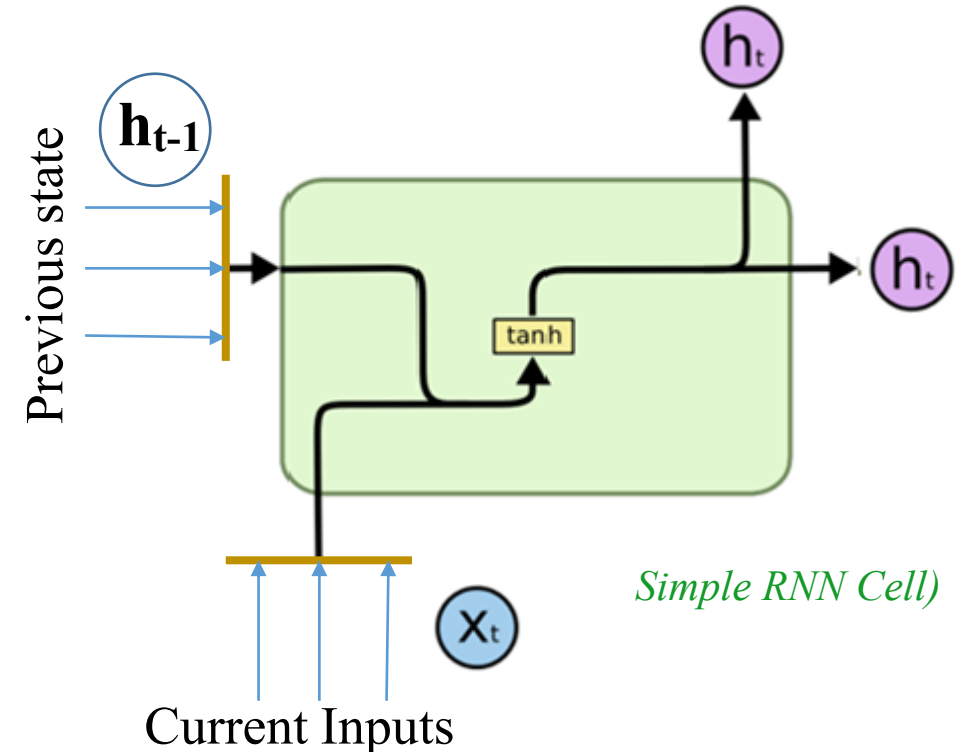
# Slide 11: AI Deep Learning: Recurrent Neural Networks (RNN)

## RNN: Simple Recurrent Neural Network: Simple RNN Cell



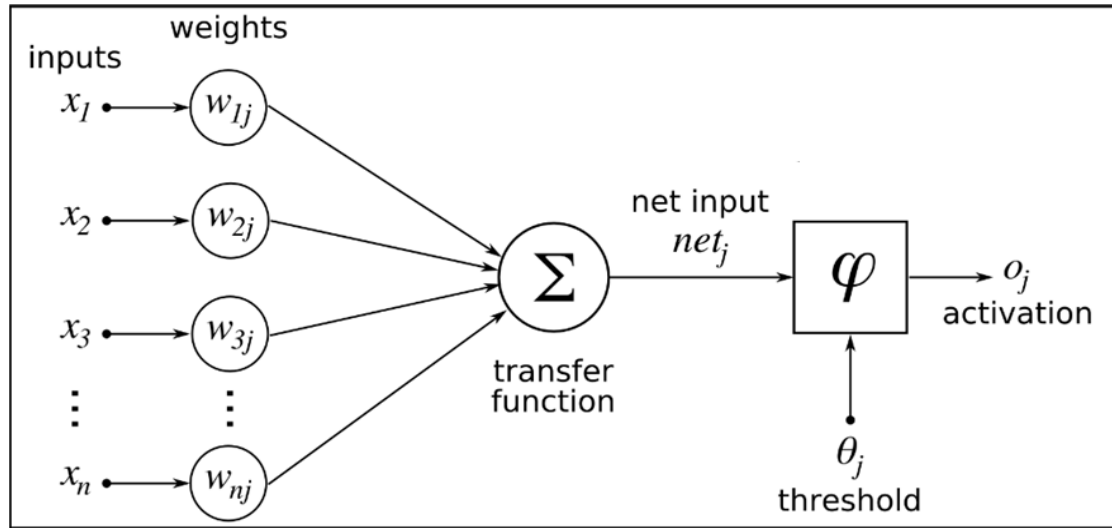
*McCulloch-Pitts Neuron Model (Sources: Wikipedia)*

$$\mathbf{h}_t = \tanh (\mathbf{W}_{hh} * \mathbf{h}_{t-1} + \mathbf{W}_{hx} * \mathbf{x}_t)$$



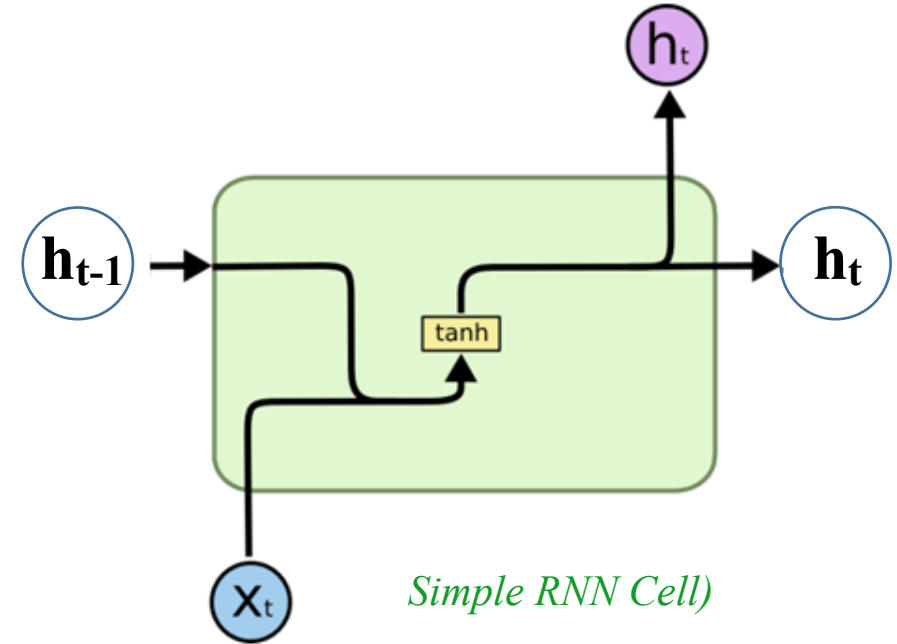
# Slide 12: AI Deep Learning: Recurrent Neural Networks (RNN)

## RNN: Simple Recurrent Neural Network: Simple RNN Cell



*McCulloch-Pitts Neuron Model (Sources: Wikipedia)*

$$h_t = \tanh (W_{hh} * h_{t-1} + W_{hx} * x_t)$$



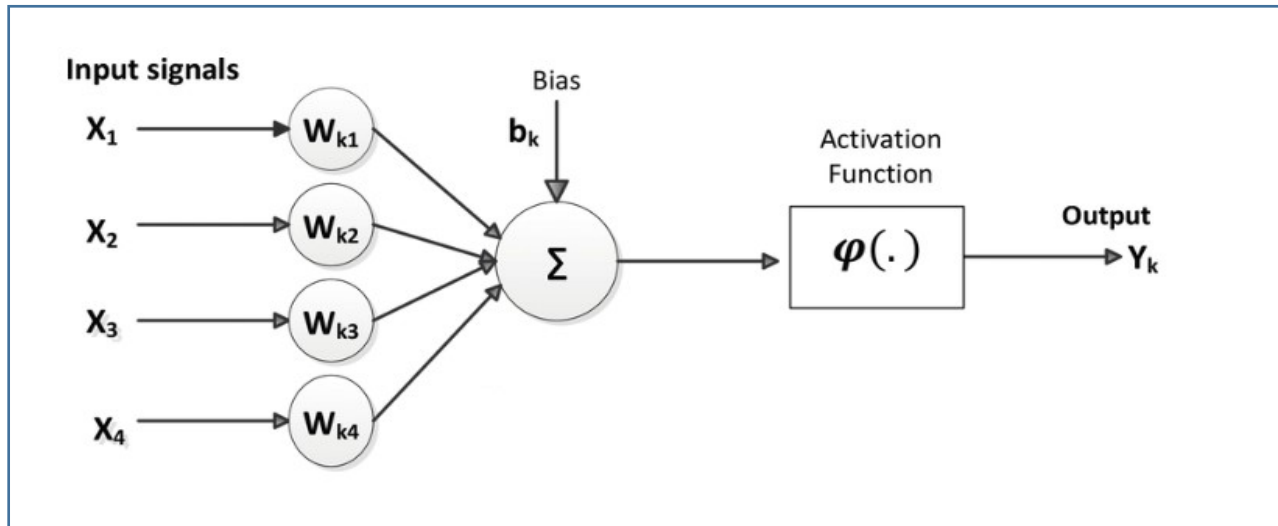
*Simple RNN Cell)*

Is there any **similarity** between the above models?

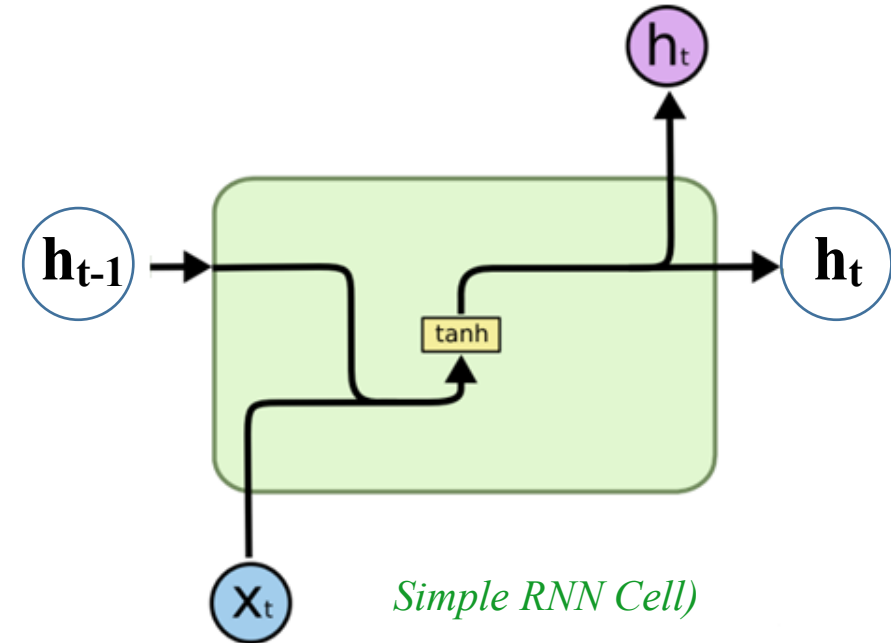
- One: **McCulloch-Pitts** model; Another: **Simple RNN** Cell.
- The **inputs**? The **weights**? The **SUM**?
- The **activation function**? The **output** signal?

# Slide 13: AI Deep Learning: Recurrent Neural Networks (RNN)

## RNN: Simple Recurrent Neural Network: Simple RNN Cell



*McCulloch-Pitts Neuron Model (Sources: Wikipedia)*

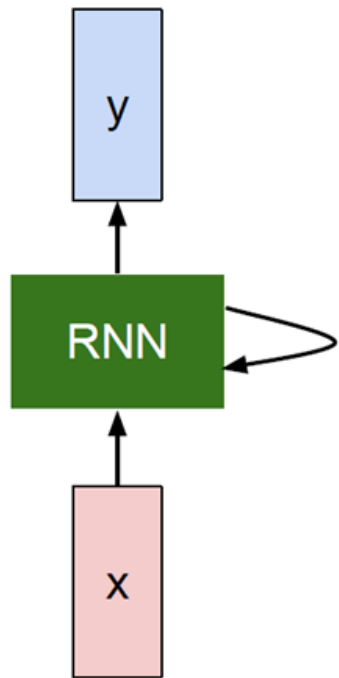


If also considering the **biases  $b$**  beside the **weights**:

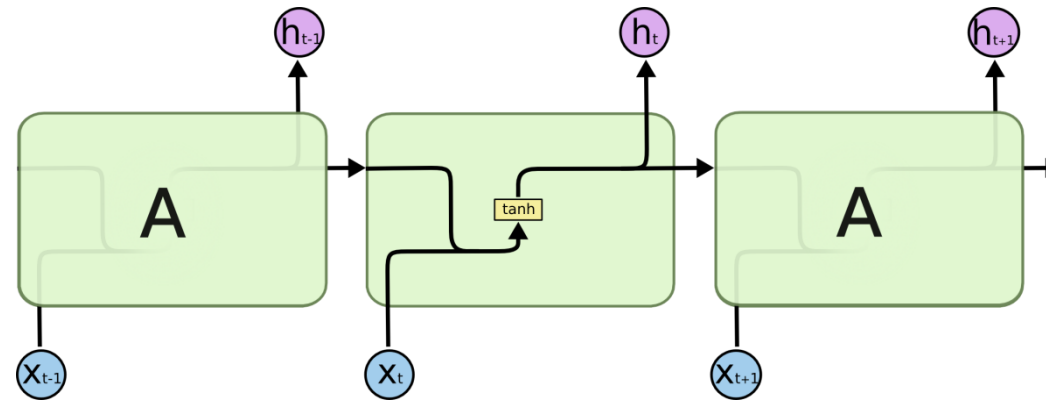
$$h_t = \tanh (W_{hh} * h_{t-1} + W_{hx} * x_t + b)$$

# Slide 14: AI Deep Learning: Recurrent Neural Networks (RNN)

## RNN: Simple Recurrent Neural Network



*Recurrent Neural Network  
(Source: Stanford.edu)*



*Rolled and Unrolled RNN (Source: Colah Blogs)*

### Simple RNN:

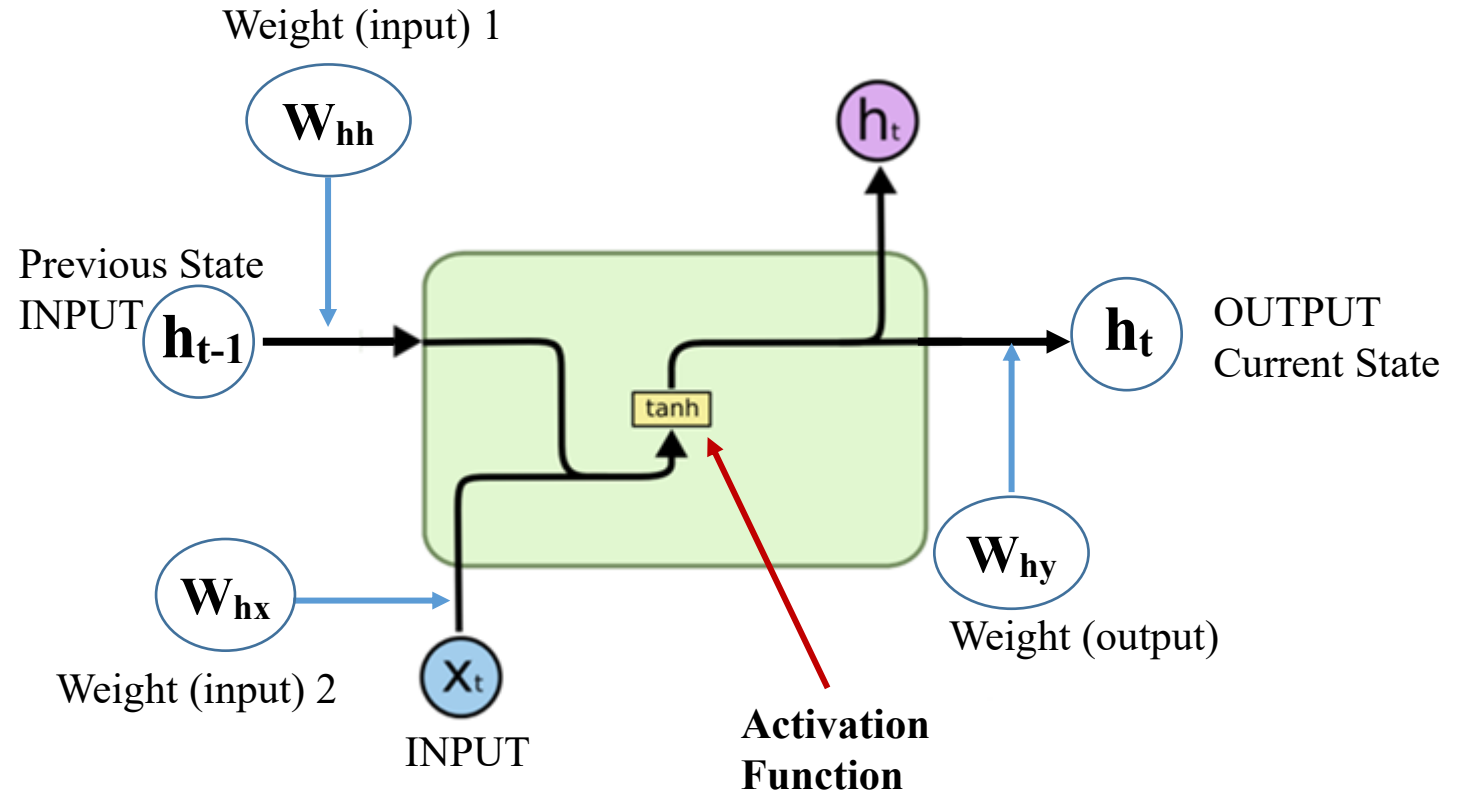
- The **state** consists of only **one “hidden” vector h**.
- It is just a **single-layer** neural network (with feedback)
- Also called:
  - A “**Vanilla**” recurrent neural network.
  - OR an “**Elman RNN**” after Prof. Jeffrey Elman

# Slide 15: AI Deep Learning: Recurrent Neural Networks (RNN)

## RNN: Simple Recurrent Neural Network: Anatomy of Simple RNN Cell

$$\mathbf{h}_t = \tanh(\mathbf{W}_{hh} * \mathbf{h}_{t-1} + \mathbf{W}_{hx} * \mathbf{x}_t)$$

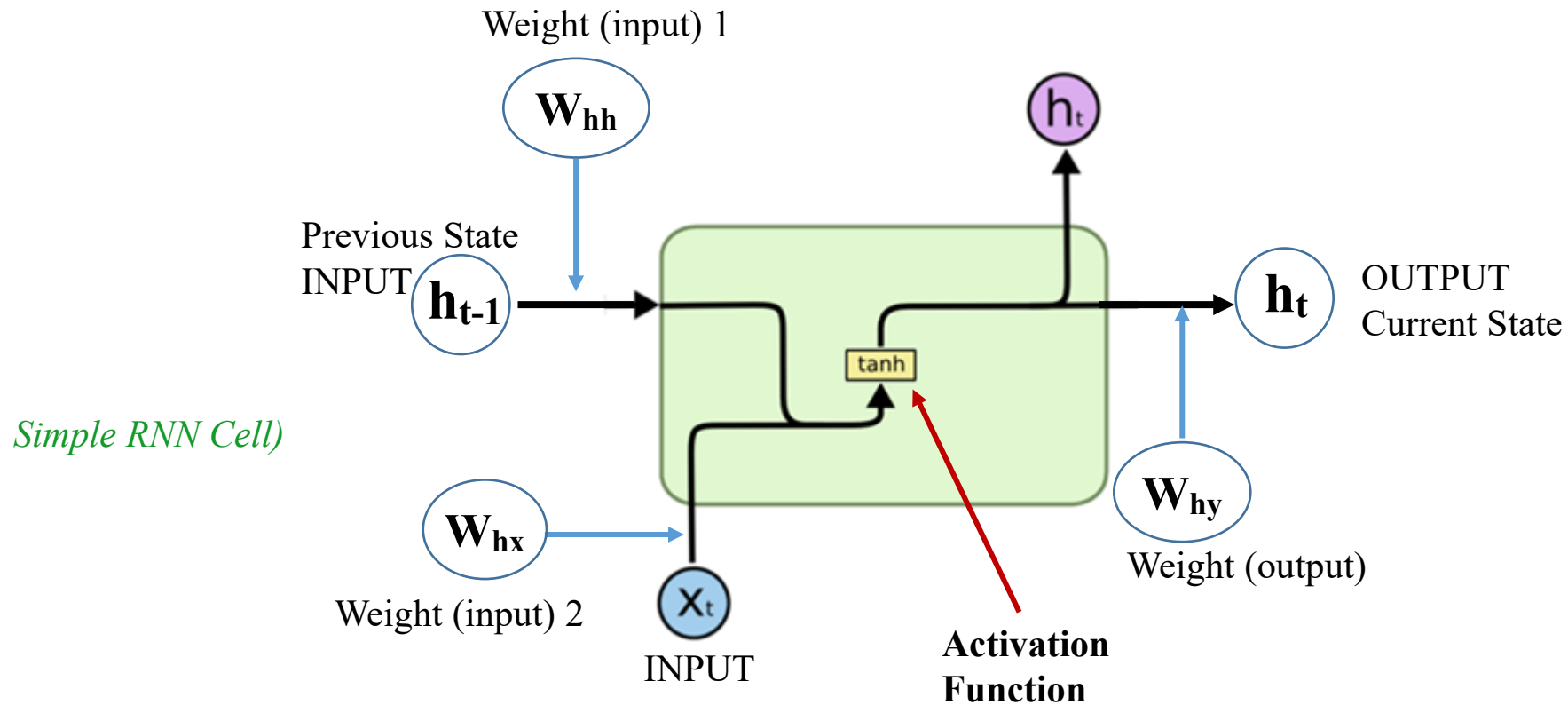
$$y_t = W_{hy} h_t$$



*Simple RNN Cell)*

# Slide 16: AI Deep Learning: Recurrent Neural Networks (RNN)

## RNN: Simple Recurrent Neural Network: Anatomy of Simple RNN Cell



Where is the **memory**?



# Slide 17: AI Deep Learning: Recurrent Neural Networks (RNN)

## RNN: Simple Recurrent Neural Network: Anatomy of Simple RNN Cell

Where is the **memory**?

- Mathematical model of **memory**:

$$\mathbf{C}_{t+1} = h_{\theta} (x_{t+1}, \mathbf{C}_t) \rightarrow \mathbf{C}_{t+1} = h_{\theta} (\mathbf{C}_t, x_{t+1}) \rightarrow \mathbf{C}_t = h_{\theta} (\mathbf{C}_{t-1}, x_t)$$

- Mathematical model of **simple recurrent neural networks**, i.e., vanilla RNN:

$$h_t = f_W(h_{t-1}, x_t)$$

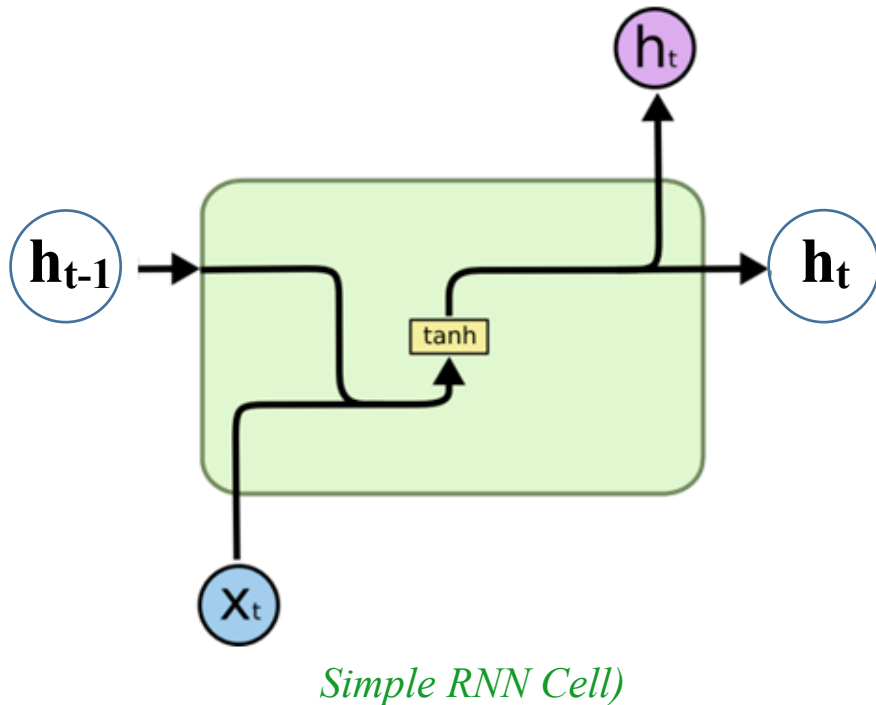
$$\mathbf{h}_t = \tanh (\mathbf{W}_{\mathbf{hh}} * \mathbf{h}_{t-1} + \mathbf{W}_{\mathbf{hx}} * \mathbf{x}_t)$$

### Conclusion:

- In recurrent neural networks, the **hidden state  $\mathbf{h}$**  represents the **memory of the network**.

# Slide 18: AI Deep Learning: Recurrent Neural Networks (RNN)

## RNN: Simple Recurrent Neural Network: Implementation



A recurrent neural network can be implemented with:

- Inputs ( $X$ ): a vector of real values
- Hidden states ( $h$ ): a vector of real values
- Outputs ( $Y$ ): a vector of real values
- Weights ( $W$ ): a vector or a matrix of real values
- Biases ( $b$ ): a vector of real values

# Slide 19: AI Deep Learning: Recurrent Neural Networks (RNN)

## RNN: Simple Recurrent Neural Network: Implementation: Example 1

Given a Recurrent Neural Network (RNN):

- **Inputs (X)**: Let  $m = 3$ , the size/length of the **input vector**, i.e. three inputs.
- **Hidden states (h)**: Let  $n = 2$ , the size/length of the **hidden-state vector**, i.e. two states.

Based on the mathematical model of Simple Recurrence Neural Network, i.e., Vanilla RNN:

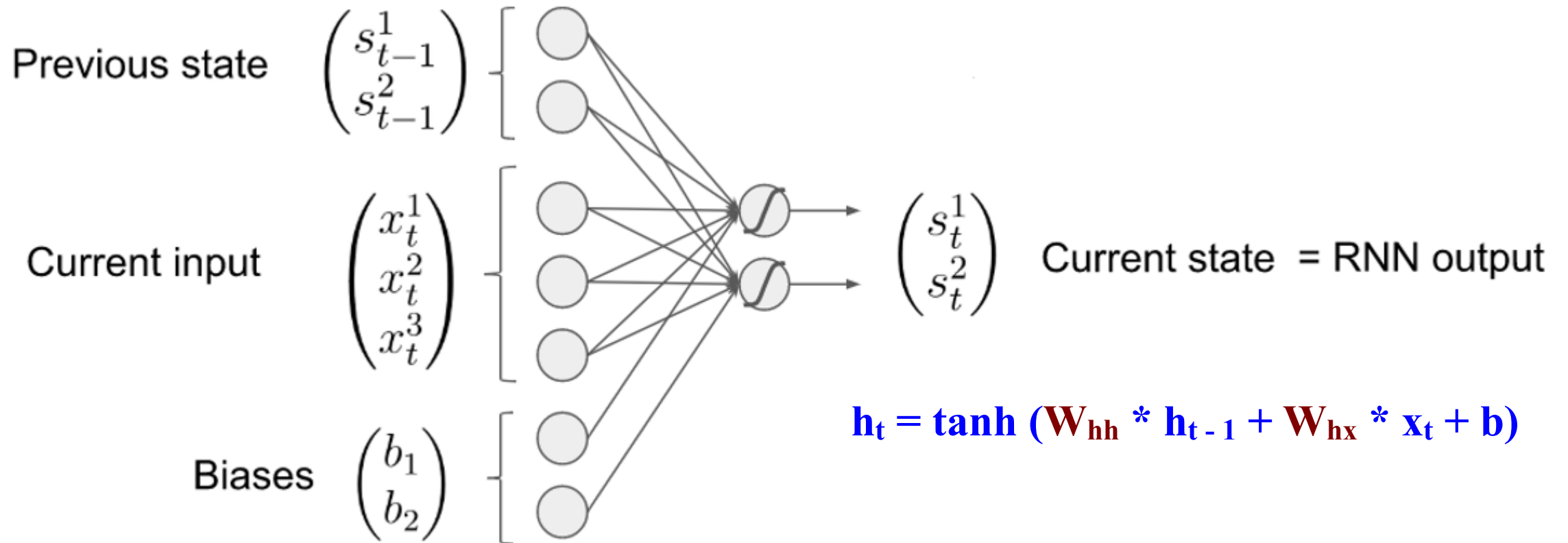
$$\mathbf{h}_t = \tanh (\mathbf{W}_{hh} * \mathbf{h}_{t-1} + \mathbf{W}_{hx} * \mathbf{x}_t + \mathbf{b})$$

We can have:

- **Weight ( $\mathbf{W}_{hh}$  : Inputs (X))**: A matrix  $n \times m$  (2 x 3)
- **Weight ( $\mathbf{W}_{hx}$  : Hidden states (h))**: A matrix  $n \times n$  (2 x 2)
- **Bias (b)**: A **vector** of the size/length of 2, each bias for one state
- **Outputs**: A **vector** of the size/length of 2, each output for one state

## Slide 20: AI Deep Learning: Recurrent Neural Networks (RNN)

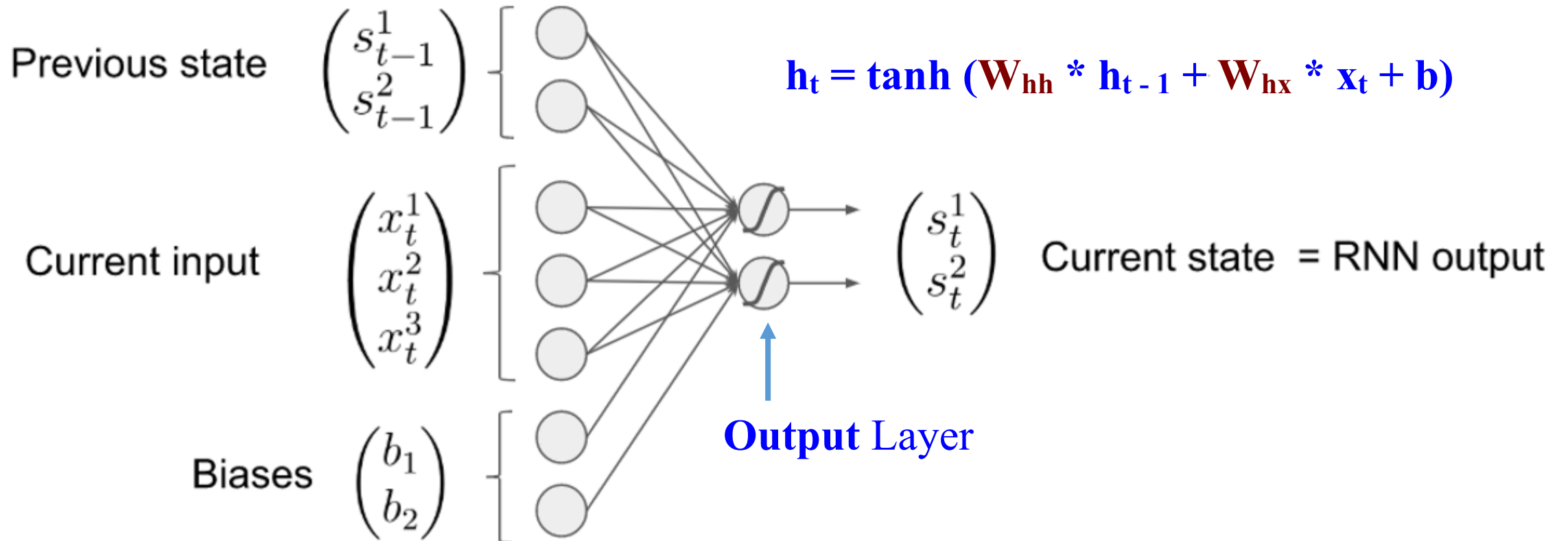
### RNN: Simple Recurrent Neural Network: Implementation: Example 1



*Simple RNN Neural Network (Source: deepsystem.ai)*

# Slide 21: AI Deep Learning: Recurrent Neural Networks (RNN)

## RNN: Simple Recurrent Neural Network: Implementation: Example 1

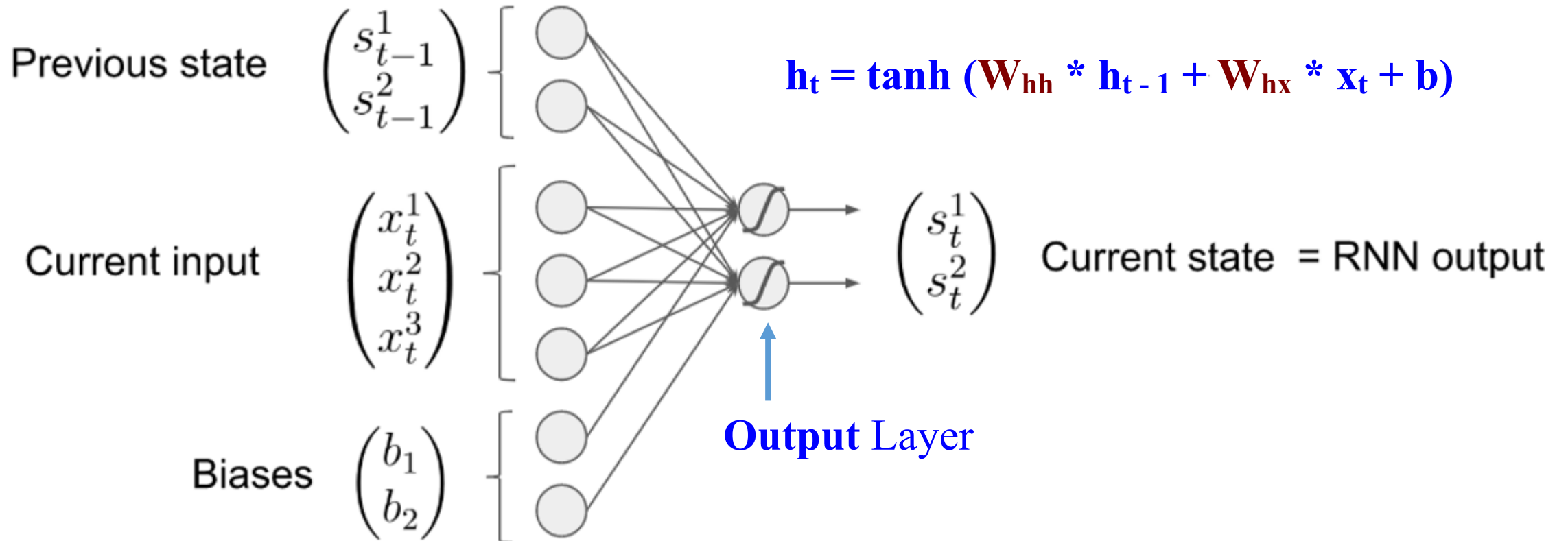


*Simple RNN Neural Network (Source: [deepsystem.ai](https://deepsystem.ai))*

**Vanilla RNN: How Many Layers?**

## Slide 22: AI Deep Learning: Recurrent Neural Networks (RNN)

### RNN: Simple Recurrent Neural Network: Implementation: Example 1

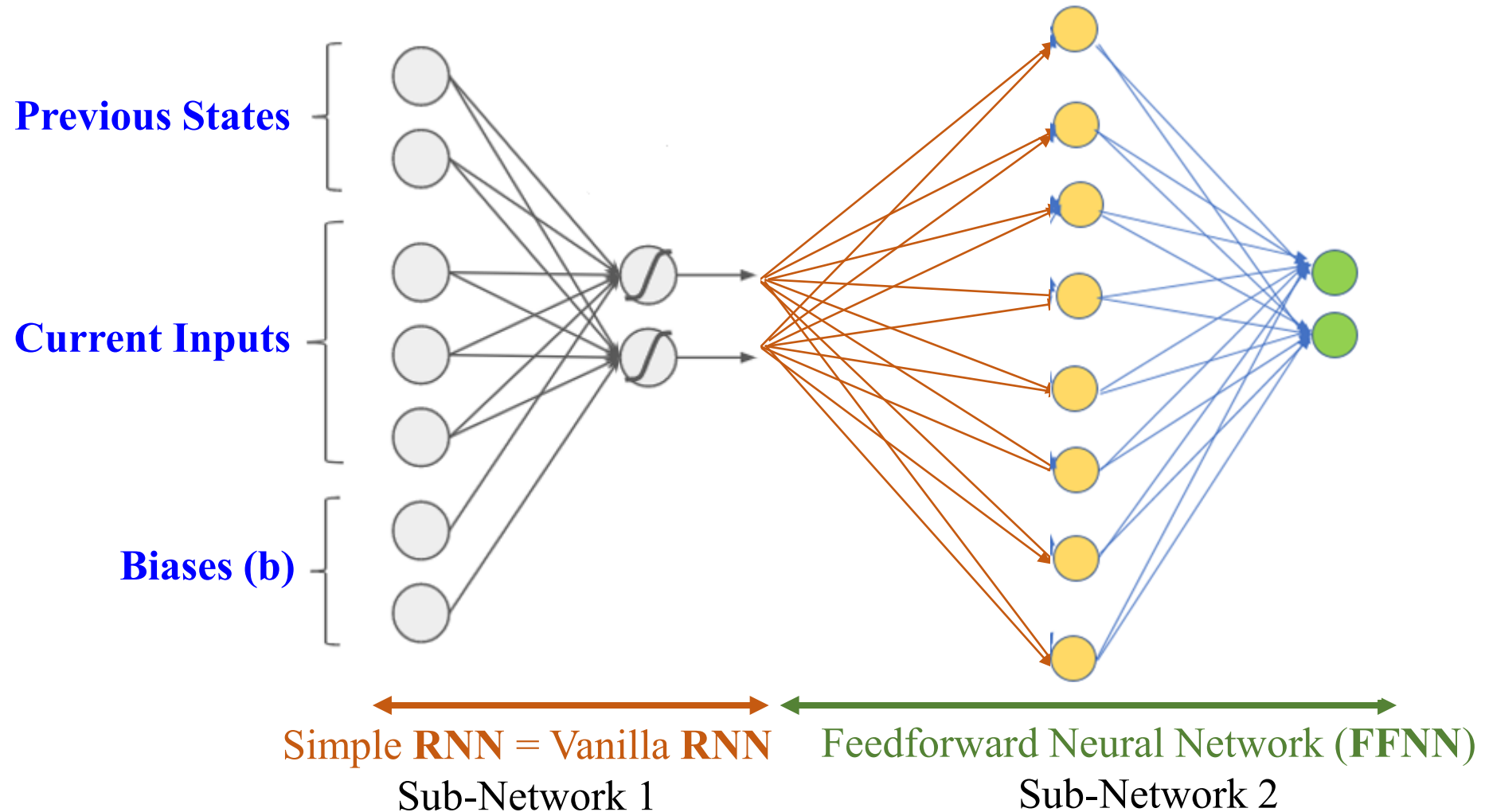


*Simple RNN Neural Network (Source: deepsystem.ai)*

**Vanilla RNN: Single-Layer** Recurrent Neural network

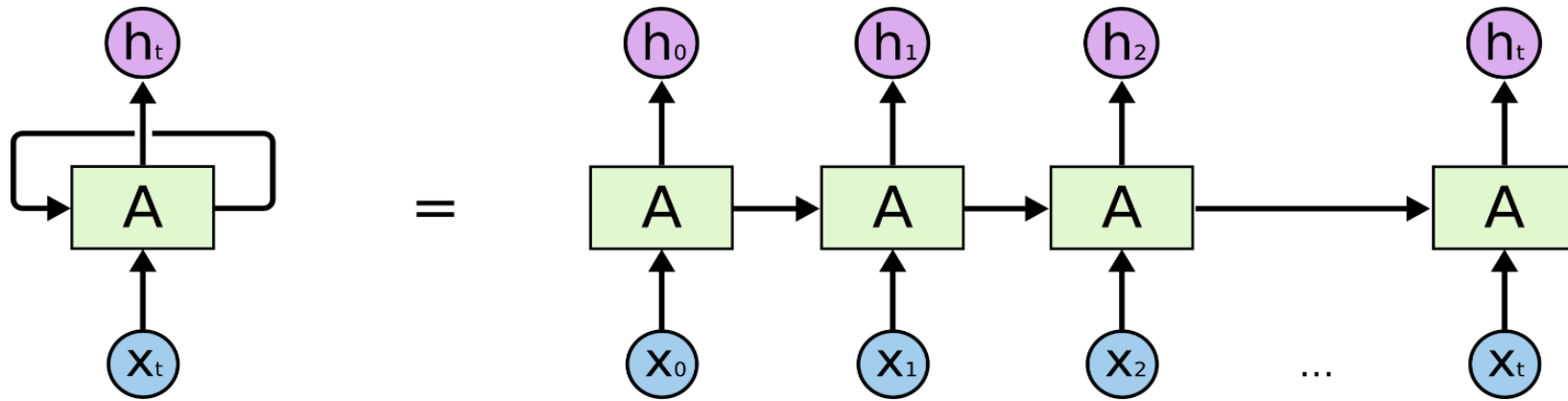
# Slide 23: AI Deep Learning: Recurrent Neural Networks (RNN)

## RNN: Recurrent Neural Network with Simple RNN Cell: Architecture



# Slide 24: AI Deep Learning: Recurrent Neural Networks (RNN)

## RNN: Recurrent Neural Network with Simple RNN Cell: Rolled & Unrolled



*Rolled and Unrolled RNN (Source: Colah Blogs)*

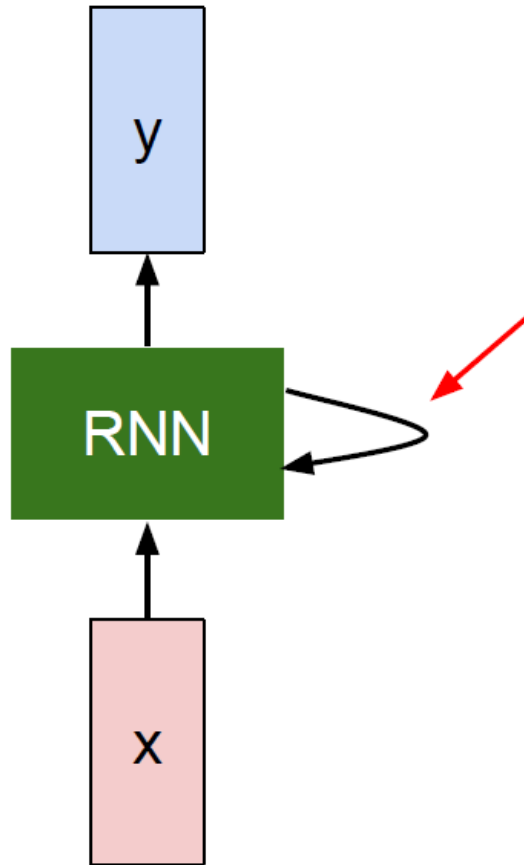
**Rolled** Simple RNN

**Unrolled** Simple RNN



# Slide 25: AI Deep Learning: Recurrent Neural Networks (RNN)

## Recurrent Neural Networks: Fundamentals



Key idea: RNNs have an “internal state” that is updated as a sequence is processed

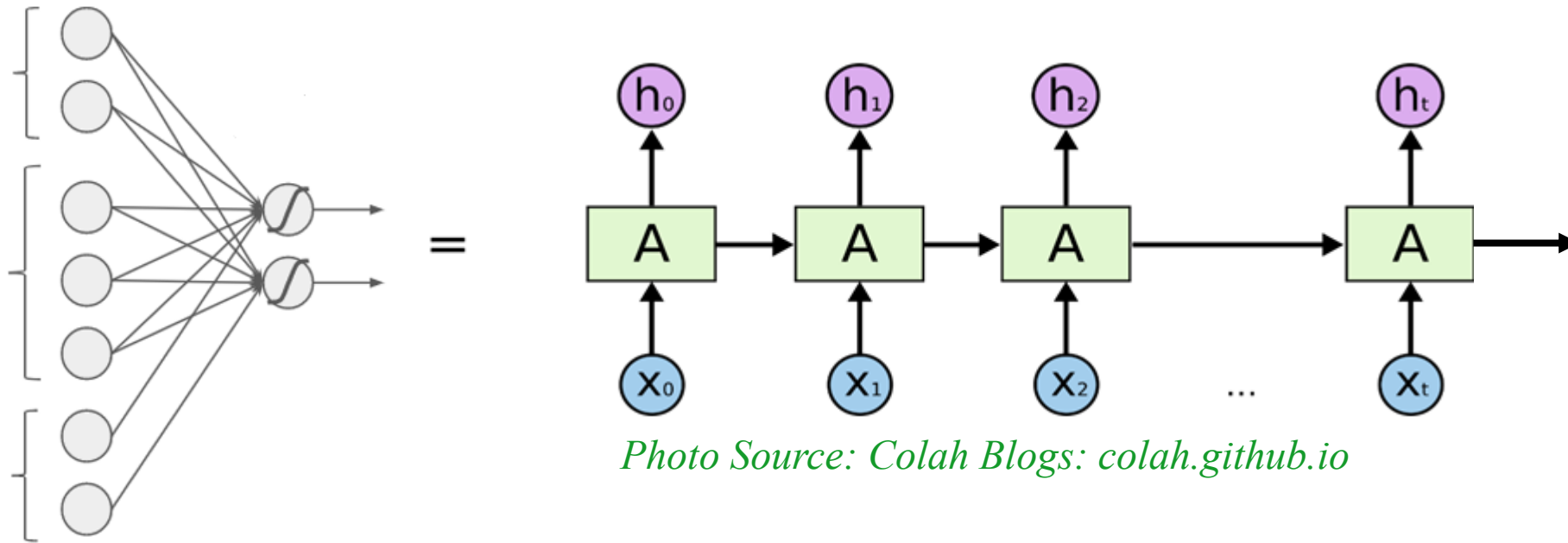
Key words:

- An **internal** (or “**hidden**”) **state** is **updated**
- When a **sequence** is **processed**

*Recurrent Neural Network  
(Source: Stanford.edu)*

## Slide 26: AI Deep Learning: Recurrent Neural Networks (RNN)

### RNN: Recurrent Neural Network with Simple RNN Cell: Rolled & Unrolled



*Photo Source: Colah Blogs: [colah.github.io](http://colah.github.io)*

#### **Rolled** Simple RNN

- > A **single-layer** vanilla RNN
- > **Inputs**: A sequence of **k** data points

#### **Unrolled** Simple RNN

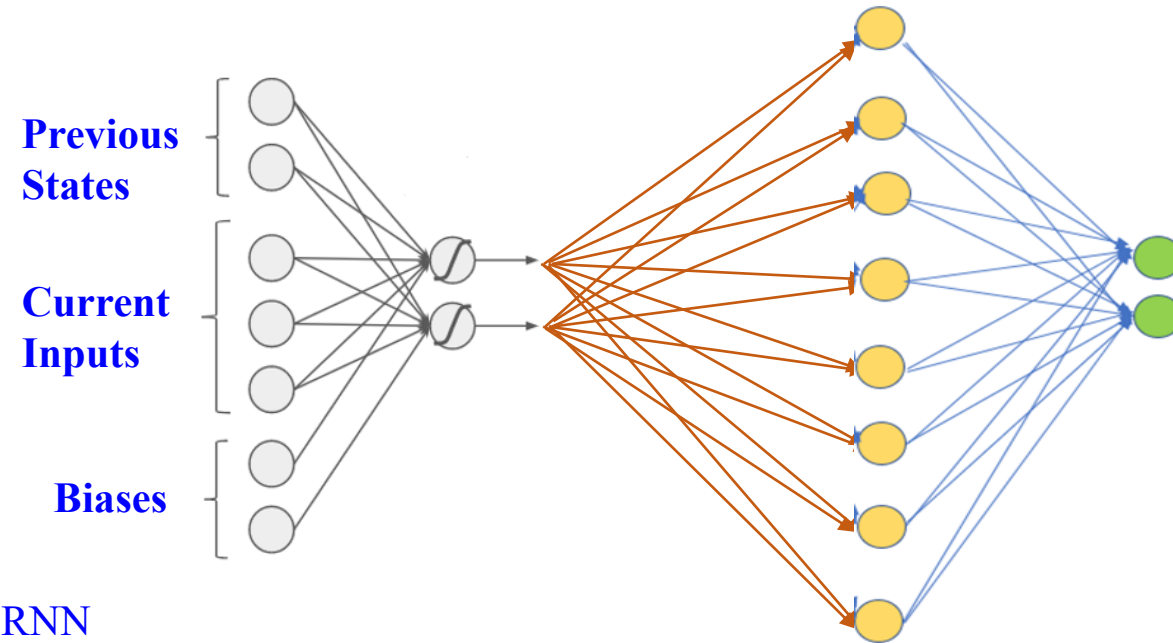
- > Unrolled into **k** steps when the input sequence is processed

# Slide 27: AI Deep Learning: Recurrent Neural Networks (RNN)

## RNN: Recurrent Neural Network with Simple RNN Cell: Example 2

**Simple RNN** for Sentiment Analysis

**Input sequence:** “The movie is very interesting.”



**Rolled** Simple RNN

A **single-layer** simple RNN

**Inputs:** A sequence of **5** data points (5 words)

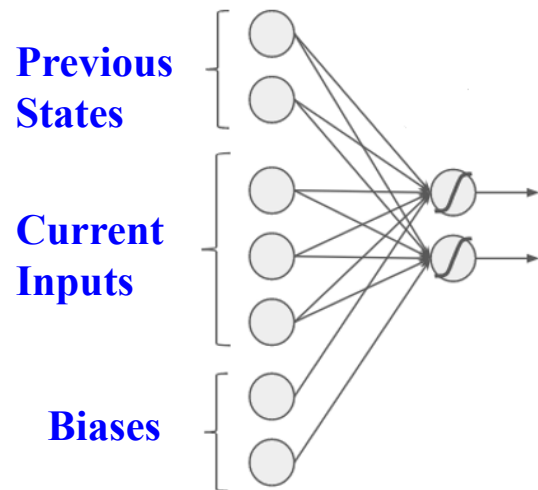
**Feedforward Neural Network (FFNN)**

# Slide 28: AI Deep Learning: Recurrent Neural Networks (RNN)

## RNN: Recurrent Neural Network with Simple RNN Cell: Example 2

### Simple RNN for Sentiment Analysis

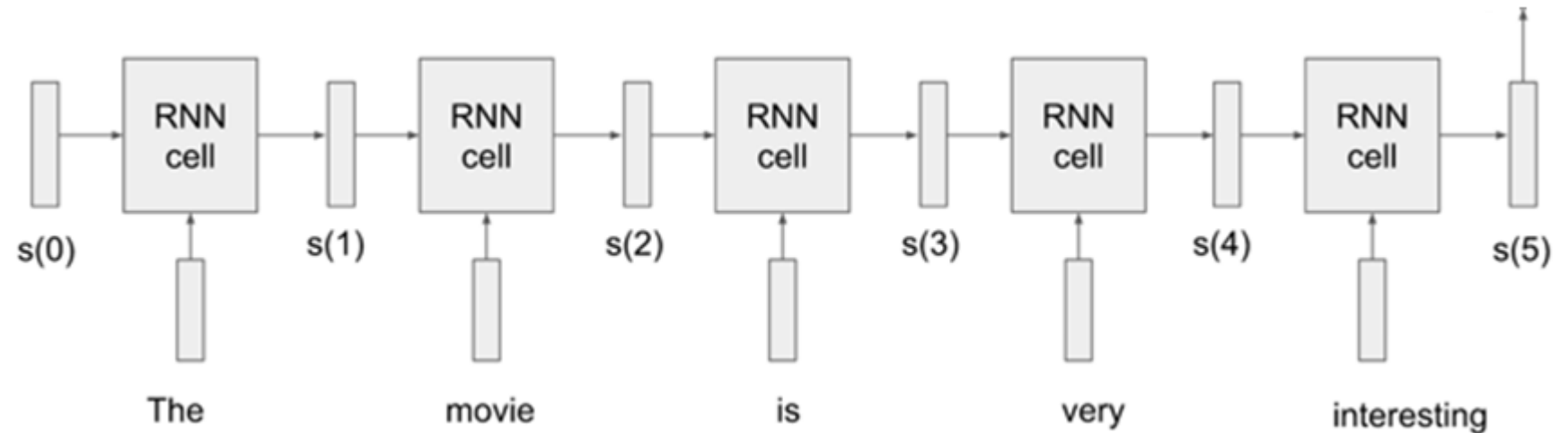
Input sequence: “The movie is very interesting.”



**Rolled** Simple RNN

A **single-layer** simple RNN

Inputs: A sequence of **5** data points (5 words)



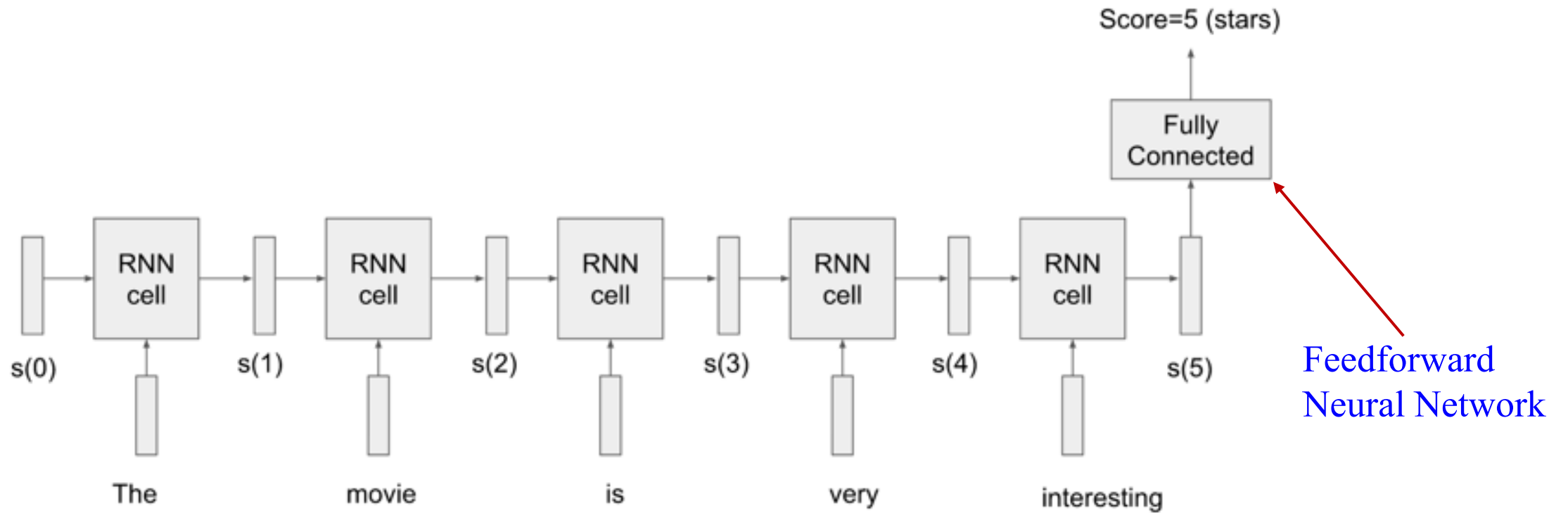
(Source: [deepsystem.ai](https://deepsystem.ai))

**Unrolled** Simple RNN: **Unrolled** into **5** steps

# Slide 29: AI Deep Learning: Recurrent Neural Networks (RNN)

## RNN: Recurrent Neural Network with Simple RNN Cell: Example 2

### Unrolled Simple RNN for Sentiment Analysis



*Unrolled Simple RNN Neural Network (Source: deepsystem.ai)*