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)

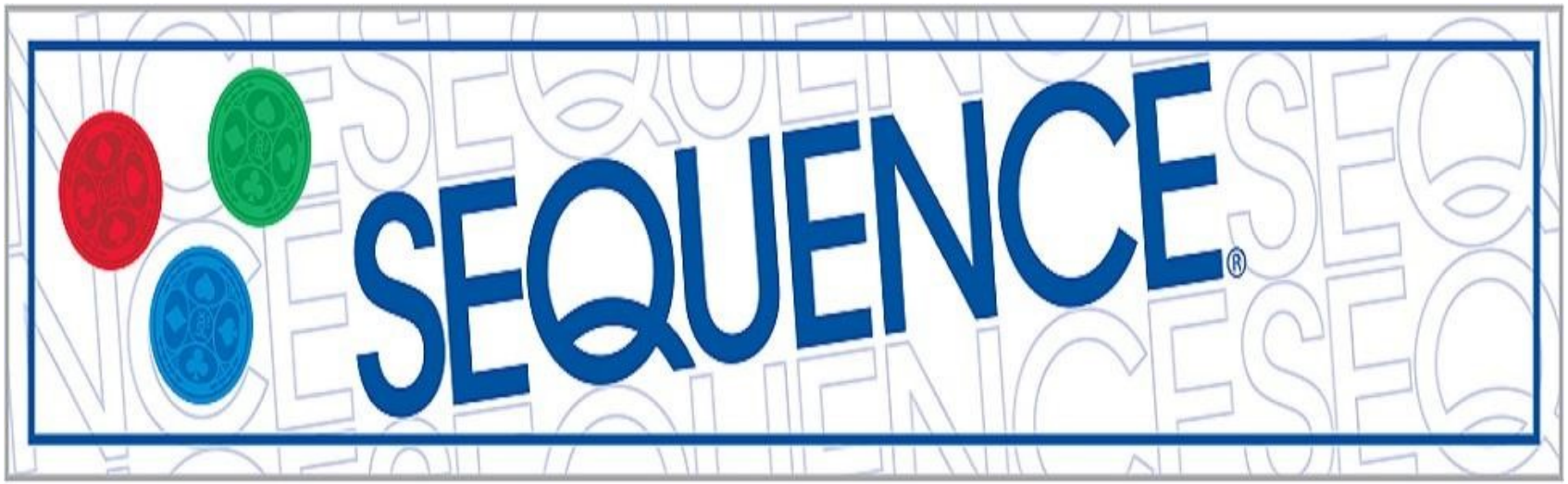
Slide 3: AI Deep Learning: Recurrent Neural Networks (RNN)

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)

What is a Sequence?

- A **sequence** is an ordered list of symbols, things, events, or values.
- The **order** can be determined by **positions** in **space** or **time**.
  - If the elements of a sequence are ordered in time, the sequence is a time series.



Sequence (Source: m.trinitrolley.com)

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

RNN: Examples of Sequence Data

- 1, 2, 3, 4, 5, 6, 7, 8, 9
  - Each data point is a number or value.
    - **NOTES**: *Each element can also be considered as a character or a symbol.*
  - Sequence data points are ordered by their **values** or positions.

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

RNN: Examples of Sequence Data

- A sequence of words spoken by a person
  - Each data point is a word, a symbol or thing.
  - Sequence data points are ordered by the **time** of being spoken.

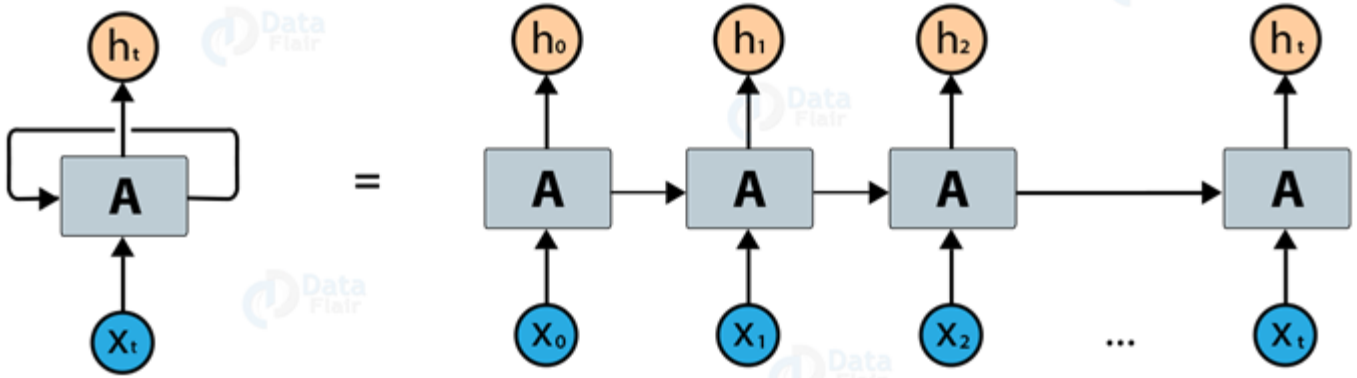
Language as Sequence Data (Source: Background images from pixabay.com)

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

RNN: Sequence Data

- Persistence is a quality that makes humans different from machines.
  - Persistence in the sense that people never start thinking from scratch.
  - A person uses his/her previous **memory** to understand the current learning and makes decisions accordingly.

- For example:
  - Language is an instance of persistence.
  - When a person is talking or writing, the choice of one word is determined both by the words coming before it and those coming after it.



Recurrent Neural Network (Source: data-flair.training)

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

RNN: Overview: Memory

What is Memory?

Memory is human ability to **encode** and **store** information and past experiences in the brain system, then **retain** or **recall** them as necessary.

- Generally, memory can be viewed as the **use of past experience** to **influence present behavior**.

In other words, **memory** is

- What humans can **remember**.
- What gives humans the **capability** to **learn** from **previous experiences**.
- What helps humans **make changes** as necessary to **adapt** to the **new** environments or contexts.

In summary:

- **Memory** is a **mechanism** that **learns** a representation of the past and **help adapt** to the present.

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

RNN: Overview: Memory

Mathematical Model of Memory

Memory is a **mechanism** that **learns** a representation of the past and **help adapt** to the present.

It is assumed:

- There exists a **neural network** represented by the **function h_θ** with shared **parameter θ**.

At time t, project all previous information **1 … t** onto a latent space **C_t**.

- It is defined for C_{t+1} as follows:

$$C_{t+1} = h_\theta (x_{t+1} , C_t)$$

$$C_{t+1} = h_\theta (x_{t+1} , h_\theta(x_t, h_\theta(x_{t-1}, \ldots h_\theta(x_1, C_0))))$$

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

RNN: Overview: Memory

Chairs and Tables in Latent Space (Source: quora.com)

Mathematical Model of Memory: Latent Space

In AI deep learning, the word "**latent**" means "hidden":

- Some data is displayed in the space that can be observed.
- The same data can be mapped to a **latent space** where some "**hidden**" **information or structure** could be exposed, e.g., the similarity between data points or which similar data points are closer together.

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

RNN: Overview: Memory

Chairs and Tables in Latent Space (Source: quora.com)

Mathematical Model of Memory: Latent Space

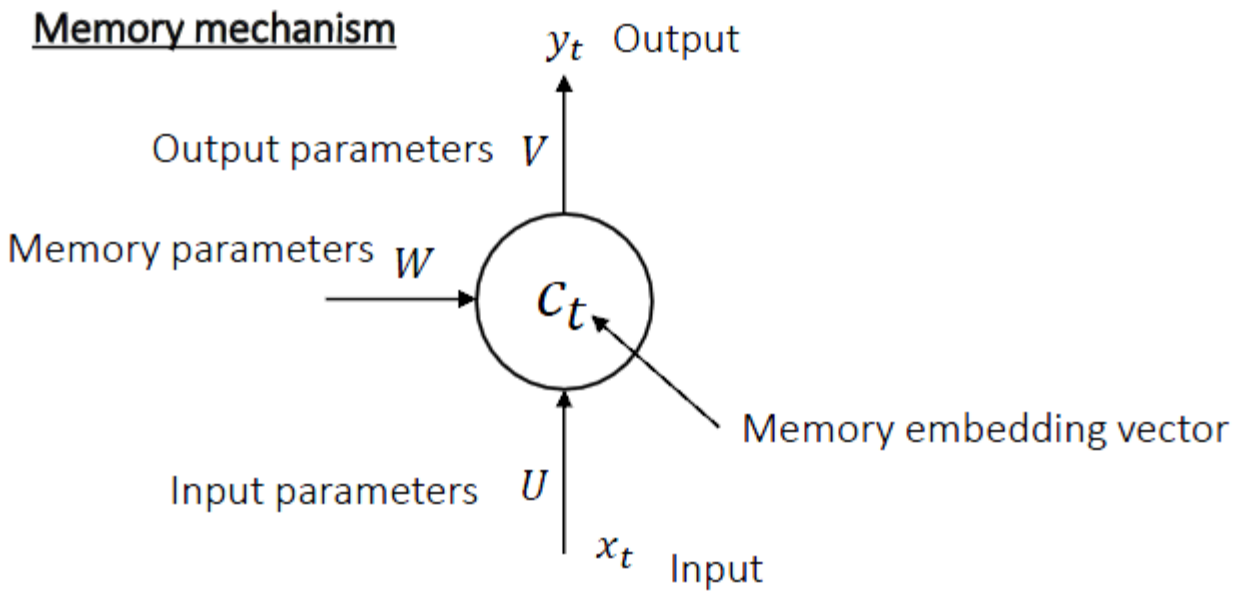
For example, consider the 4 images in the figure to the left:

- In the pixel space, i.e. the physical images, that can be observed, there is no clear similarity between any two images.
- However, if the images were mapped to a **latent space**, it would be found that the images on the left are **closer to each other** in the latent space than to any of the images on the right. So the latent space captures the "hidden" structure of the data.

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

RNN: Memory: A Graphical Representation

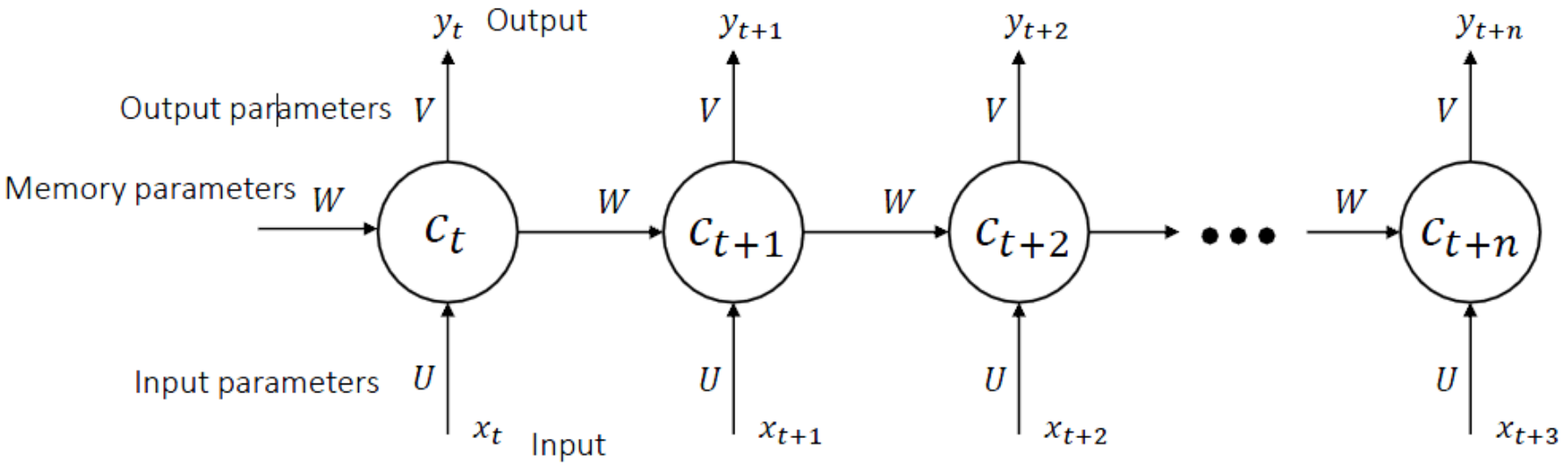
- In the simplest case, data are the Inputs/Outputs of a system.
- Sequence inputs are modeled with parameters **U**.
- Sequence outputs are modeled with parameters **V**.
- Memory I/O is modeled with parameters **W**.



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

RNN: Memory: A Graphical Representation

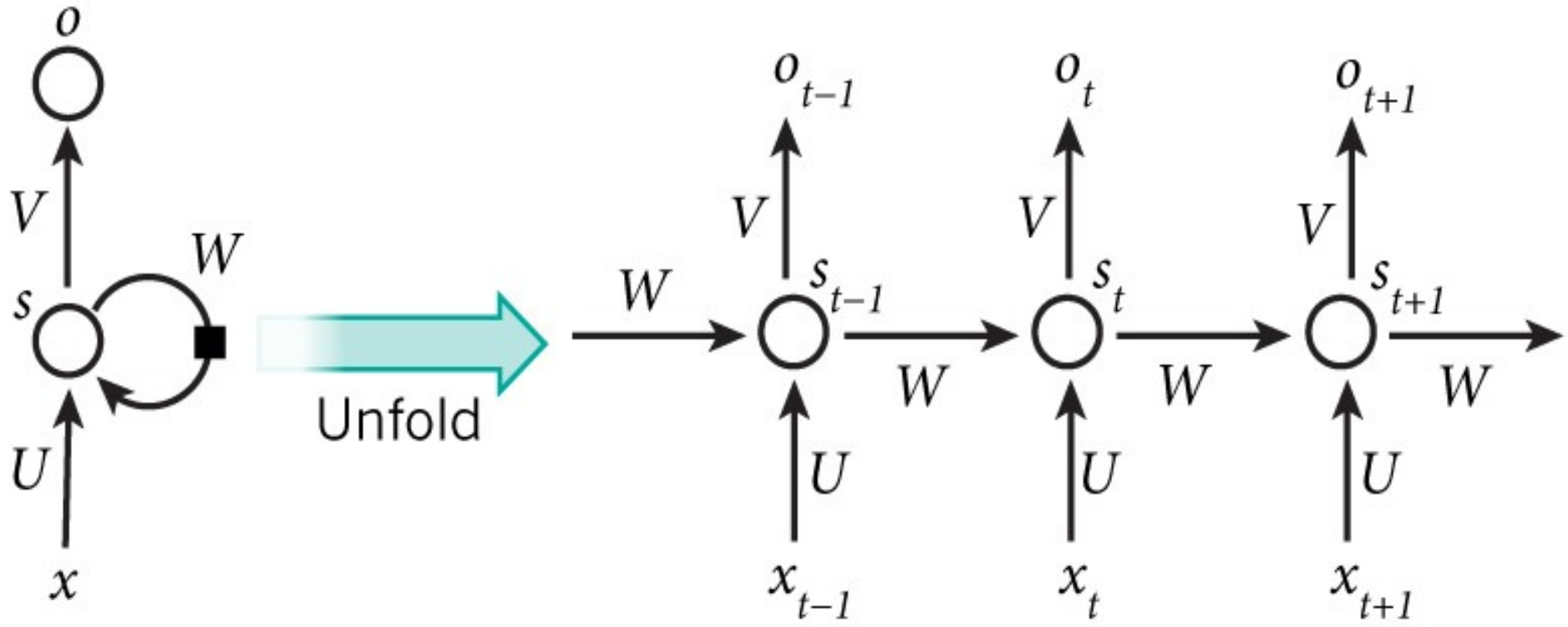
- In the simplest case, data are the Inputs/Outputs of a system.
- Sequence inputs are modeled with parameters **U**.
- Sequence outputs are modeled with parameters **V**.
- Memory I/O is modeled with parameters **W**.



Memory Mechanism

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

RNN: Memory: Folding Memory & Unfolding Memory



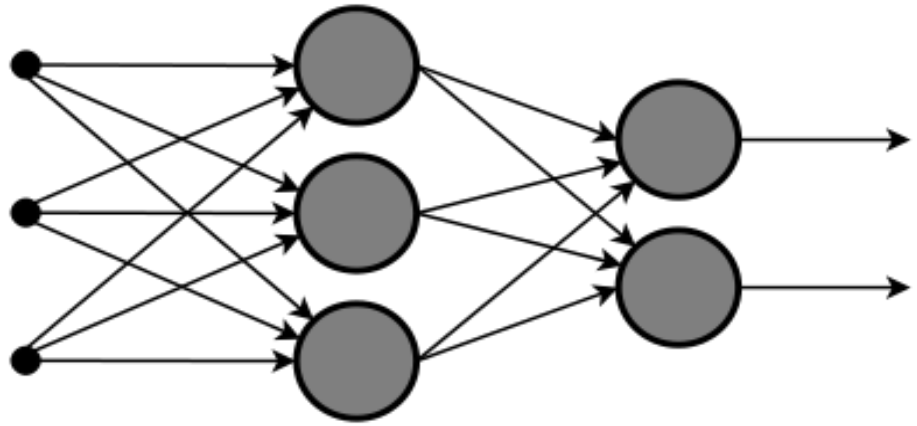
Folding and Unfolding Memory (Source: WILDML.com)

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

RNN: Overview: Feedforward Neural Networks

- Feed-forward neural networks:
  - The output is a function between the inputs and a set of weights.
  - The information moves in only one direction, forward, from the input nodes, through the hidden nodes (if any) and to the output nodes.
  - There are no cycles or loops in the network.

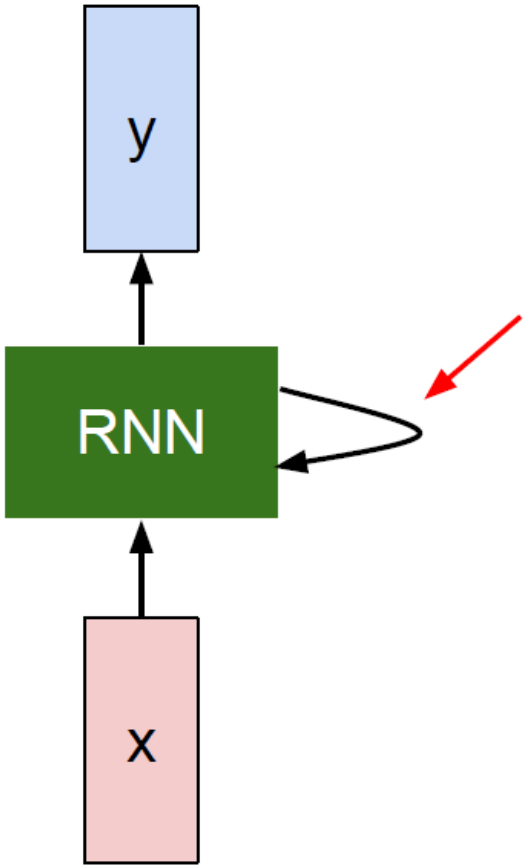
- These networks are primarily used for pattern recognition:
  - It is not efficient to use them for handling sequence data



Feed-Forward Neural Network

Slide 16: 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

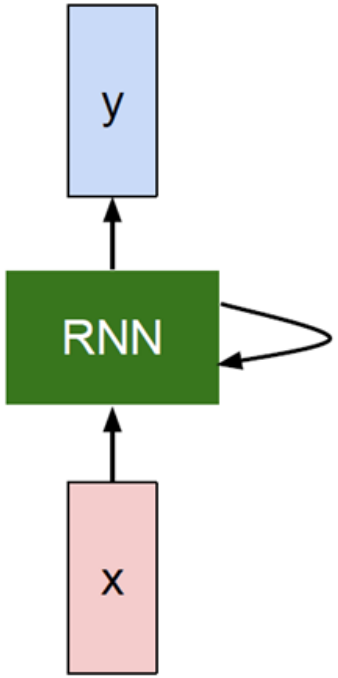
Key words:

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

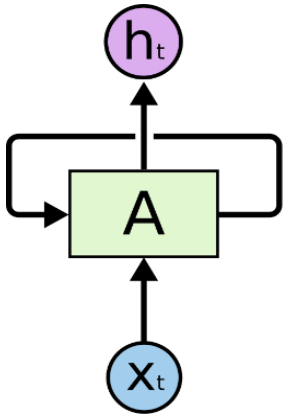
Recurrent Neural Network (Source: Stanford.edu)

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

RNN: Recurrent Neural Network: Rolled & Unrolled

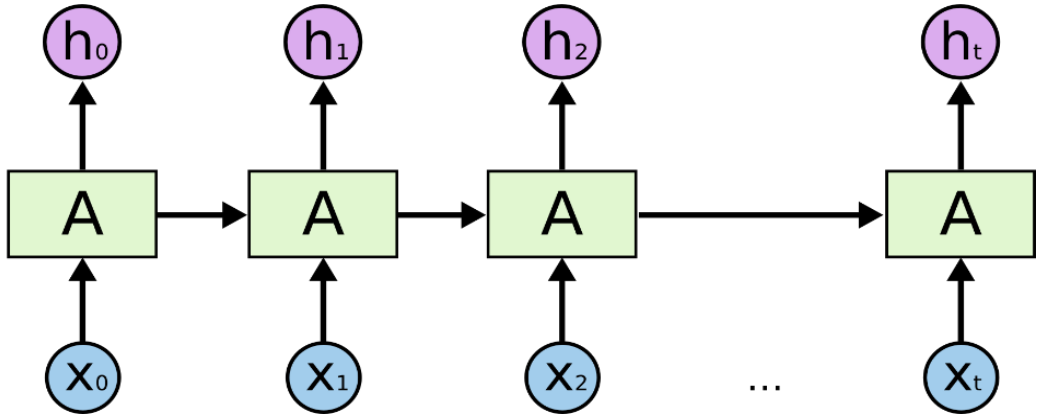


Recurrent Neural Network
(Source: Stanford.edu)



Rolled RNN

=

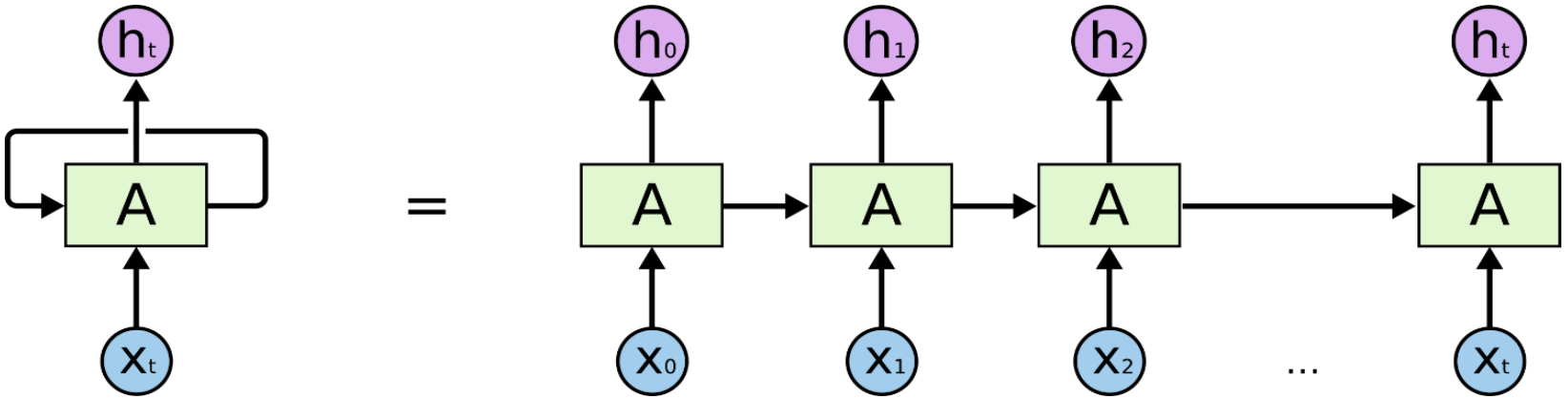


Unrolled RNN

Rolled and Unrolled RNN
(Source: Colah Blogs)

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

RNN: Overview



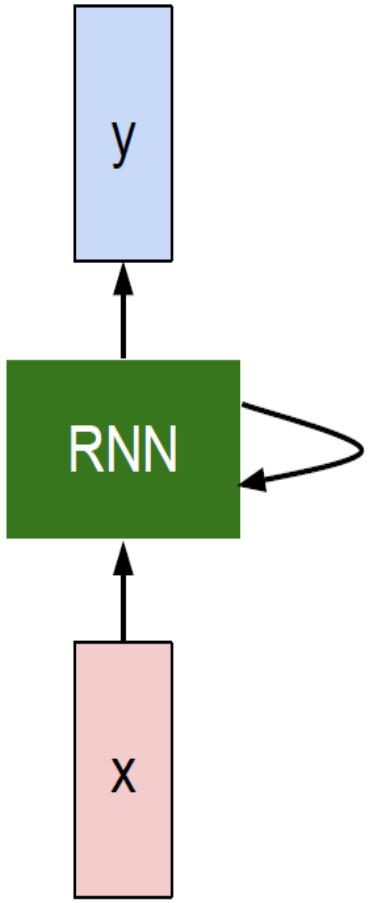
Rolled and Unrolled RNN (Source: Colah Blogs)

The chain-like nature of the recurrent neural network:

- The network is closely related to sequences and lists.
- It is a good fit for such data.

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

RNN: The Mathematical Model



Recurrent Neural Network
(Source: Stanford.edu)

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

new state

some function with parameters W

old state

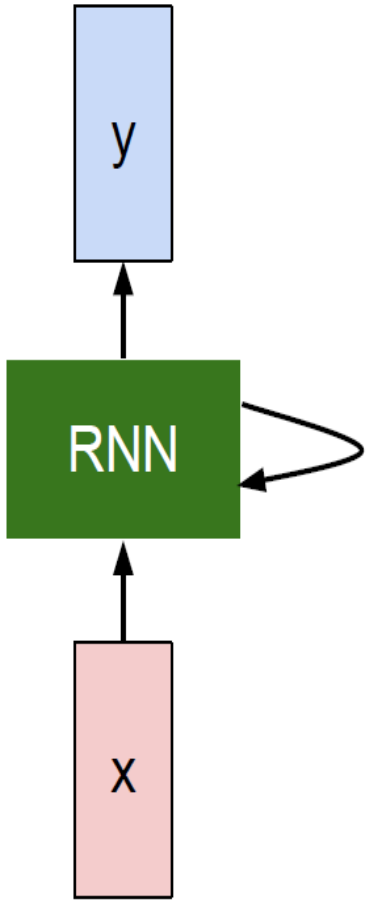
input vector at some time step

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

- The **input sequence x** can be processed by applying a **recurrent formula** at each step.
- $\mathbf{h_t}$: The "hidden" state at the time step t.

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

RNN: The Mathematical Model



Recurrent Neural Network
(Source: Stanford.edu)

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

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.