

ARTIFICIAL INTELLIGENCE AND DEEP LEARNING



WEEK 10
2021 SPRING

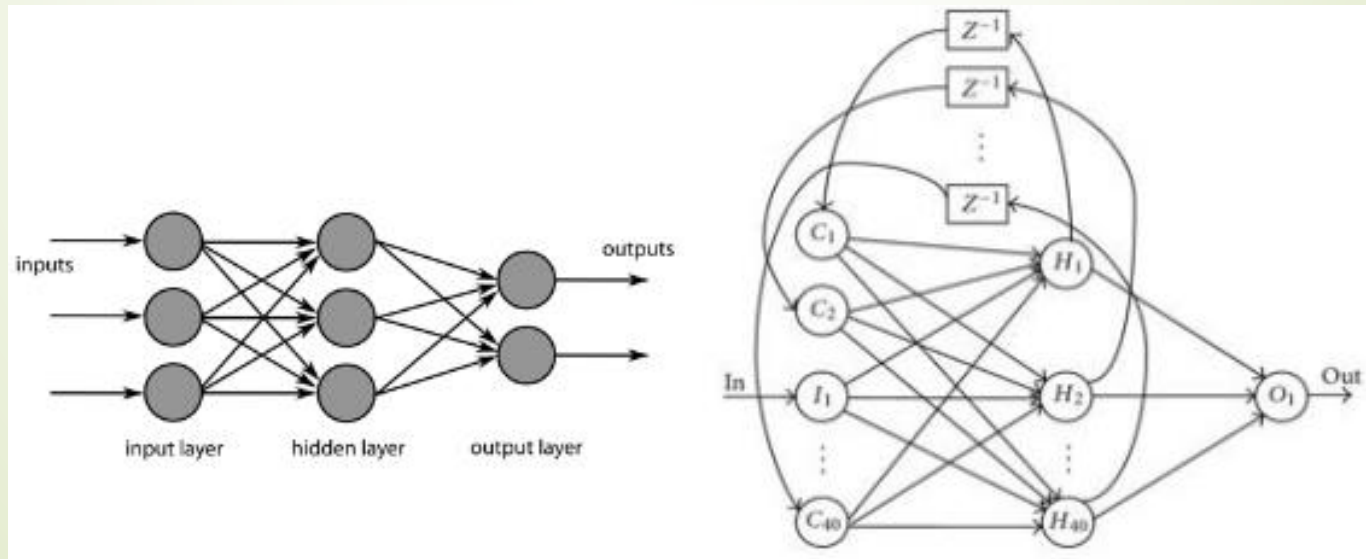


RNN, Recurrent Neural Networks

- Recurrent Neural Networks (RNN) are a class of artificial neural networks (ANN) in which the connections between nodes form a directed loop. This allows the network to exhibit temporary behavior.
- Differently than feed Forward neural networks, RNNs can use their own input memory to process the random order of the inputs. The main aim of recurrent neural networks is to utilize consecutive information.
- In a conventional neural network, we assumed that all inputs (and outputs) are independent of each other. RNNs are named as recurrent because they perform the same task for each element of a sequence; the output depends on previous calculations.

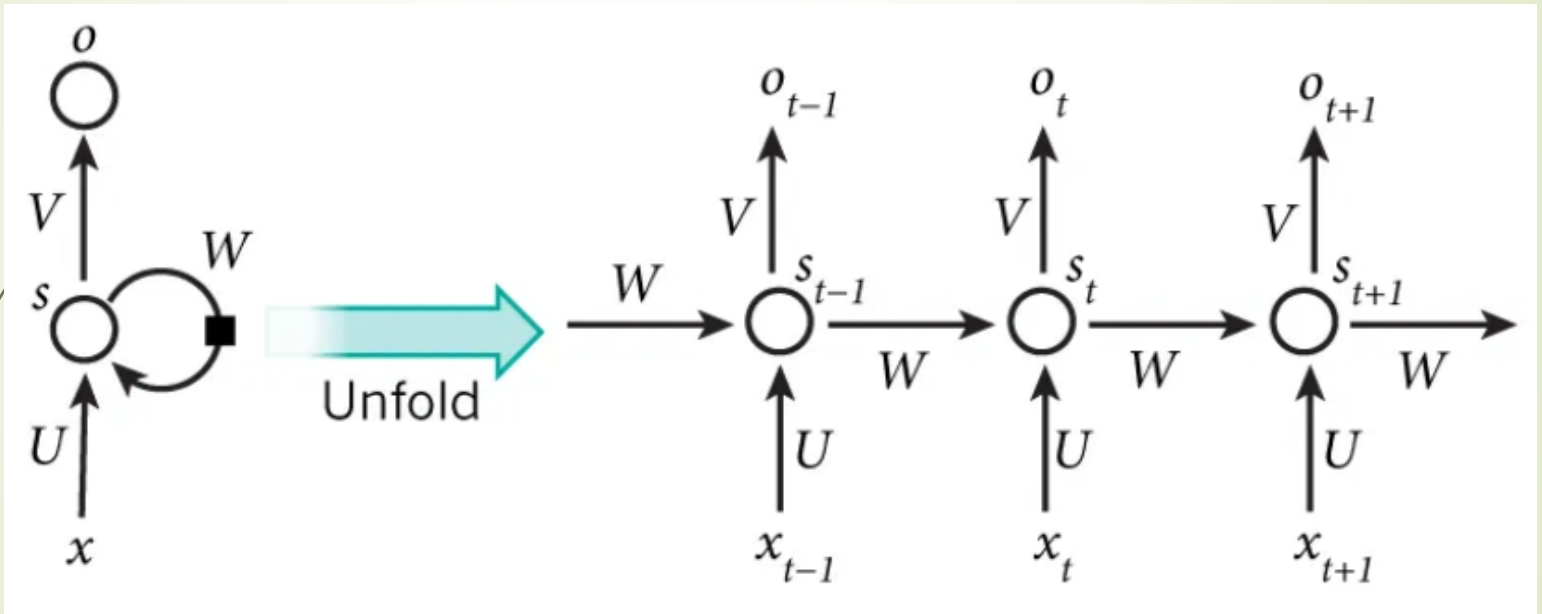
RNN, Recurrent Neural Networks

- In classical ANNs no associations between previous states or inputs is made.
- However, in RNNs which are one of the deep learning algorithms, associations between previous inputs or states is carried out.



RNN, Recurrent Neural Networks

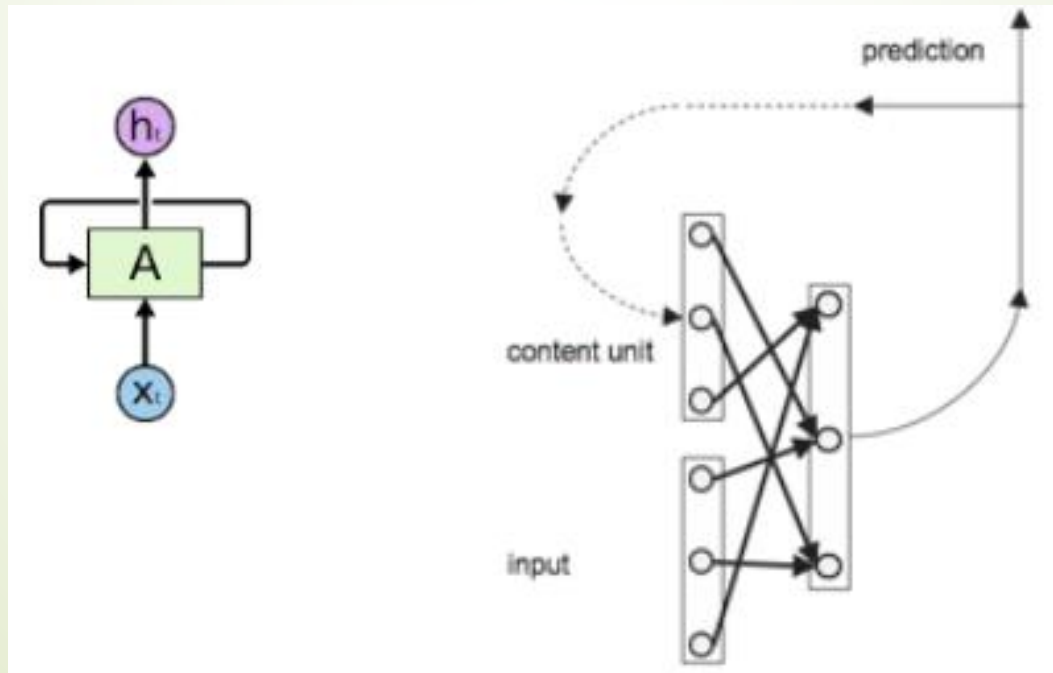
- Theoretically, RNNs can use the information from arbitrarily long series, however in practice, they are limited to only a few steps backward. A typical RNN:



- The diagram above shows the unfolding of a RNN into a complete network.

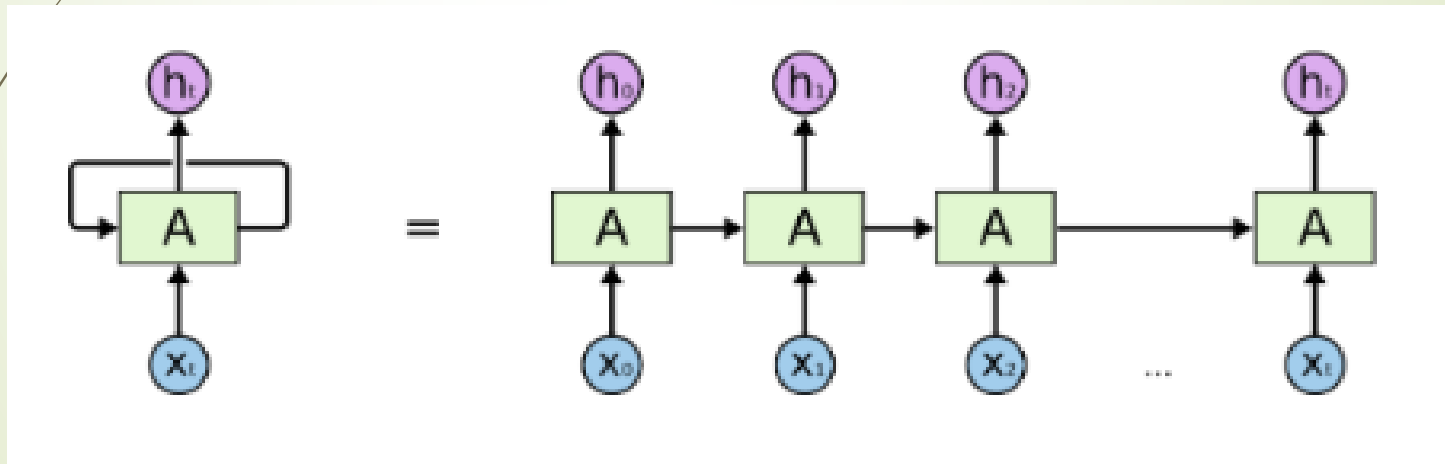
RNN, Recurrent Neural Networks

- RNNs can be said to have a loop inside them.
- In the figure below, A represents a neural network while x and h represent the input and output, respectively.



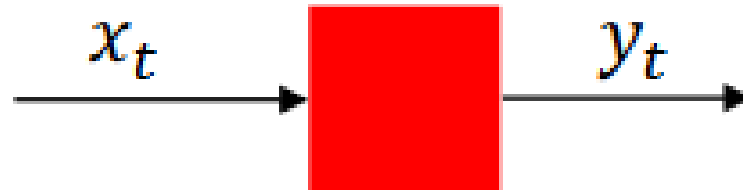
RNN, Recurrent Neural Networks

- When a RNN is analyzed, it is thought that it consists of many neural network copies.
- Here, each neural network transfers the information to the next one (input).



RNN, Recurrent Neural Networks

- In simple feed Forward networks each output is calculated for its own input.



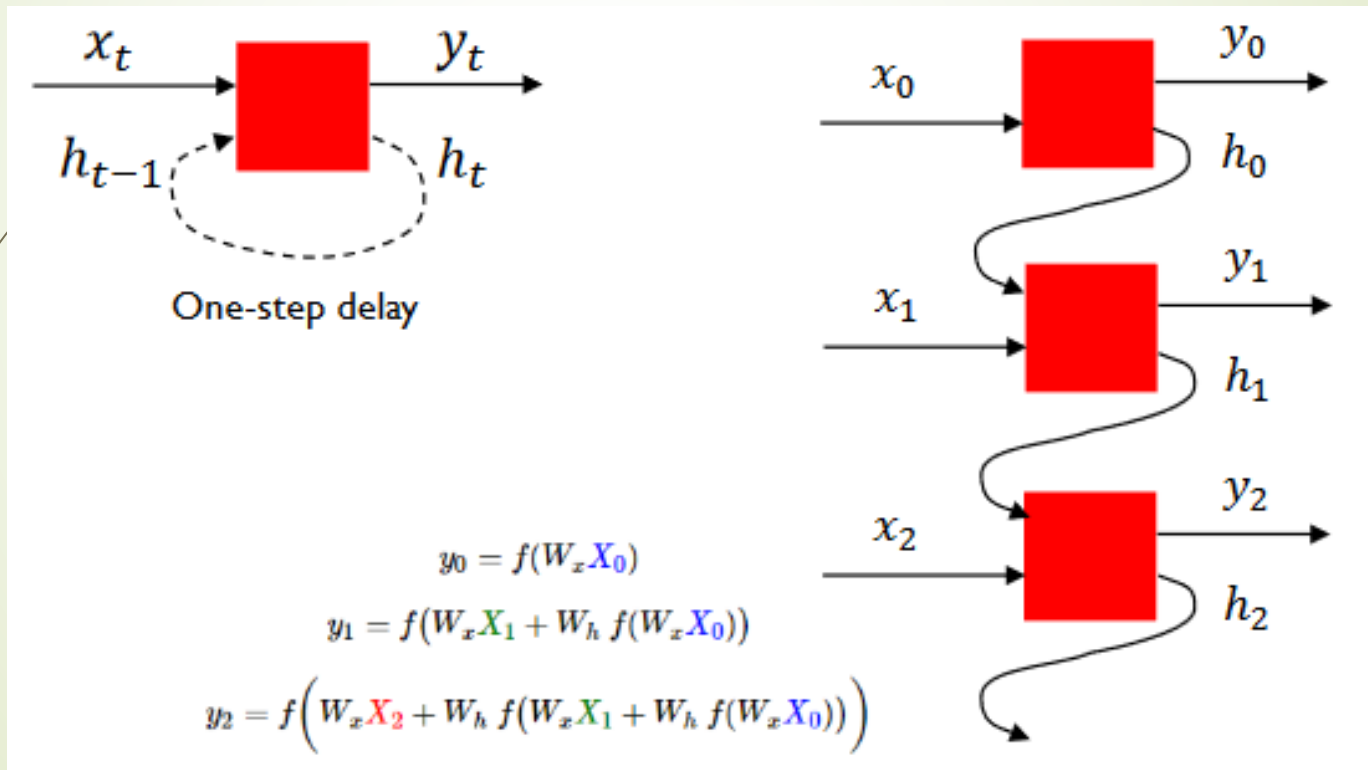
$$y_0 = f(W_x X_0)$$

$$y_1 = f(W_x X_1)$$

$$y_2 = f(W_x X_2)$$

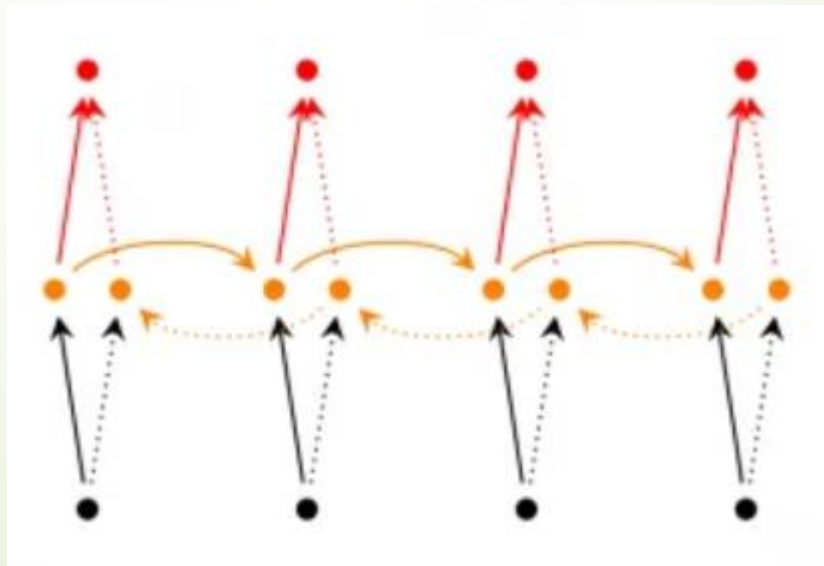
RNN, Recurrent Neural Networks

- In RNNs, on the other hand, each output is calculated depending on its input and previous inputs.



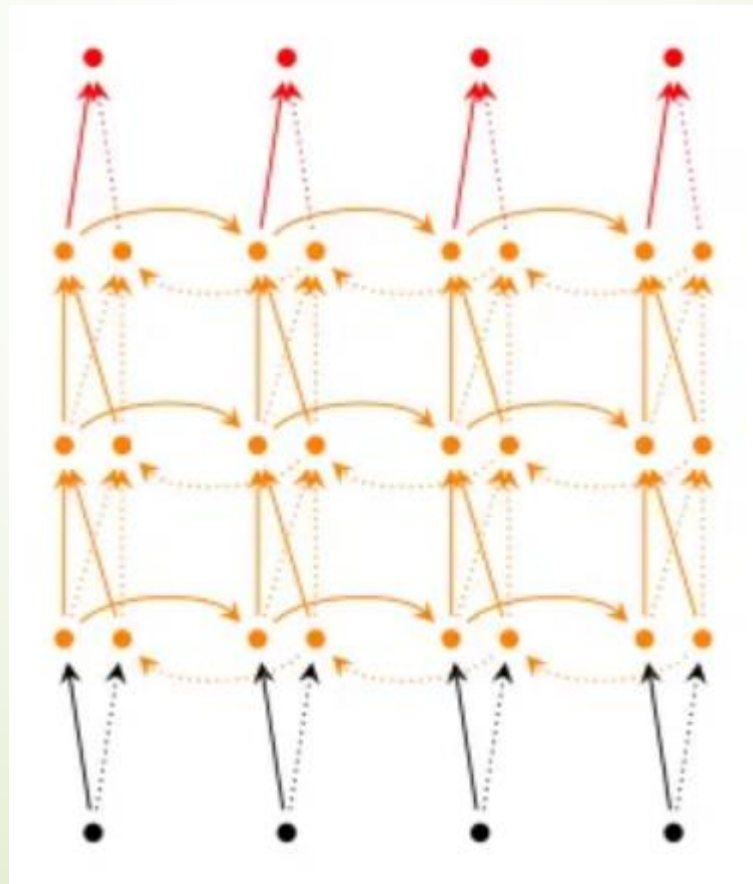
RNN, Recurrent Neural Networks

- **Bidirectional RNNs:** Based on the idea that the outputs at time t may not only depends on previous elements. in the sequences but also future elements. For example, in order to predict a missing word in a sequence, you want to look at both left and right content. Bidirectional RNNs can be used for this purpose. These are nothing but two RNNs stacked on top of each other. The outputs are calculated later, considering the hidden state of both RNNs.



RNN, Recurrent Neural Networks

- **Deep Bidirectional RNNs:** They are similar to Bidirectional RNNs, the only difference is they consist of multiple layers at time step. In practice, this provides us a higher learning capacity, but it requires more learning data.





RNN, Recurrent Neural Networks

- **LSTM (Long-Short Term Memory) Networks:** An analogy can be drawn to the way that people don't start thinking from scratch at any moment: our thoughts today are actually a continuation of what happened in our minds in the previous days. LSTM networks also have the property of long and short term memory and aim to produce new outputs continuously using earlier information.
- Although they have both long and short term memory qualities, LSTMs show their superiority over long term memory to a greater extent.
- LSTMs also handle the diminishing gradients problem encountered by conventional RNNs.
- The property of using information from the past also exists in RNNs. However, a RNN may not be able to determine at which point in the past the information is useful.

RNN, Recurrent Neural Networks

- For example: «The clouds in the sky are» sentence can be completed by any RNN as 'white' and they will be successful in simple sentences like this.
- However, if the previous sentence was a sentence like «It'll be rainy today too», the prediction of RNN will be a false one. In this case, a network must also consider this sentence in order to make a correct prediction.
- Ordinary RNNs are usually successful in situations where the information (hint) to deliver the correct output is located near the output. As the distance (time or location) between the hint and the output to be predicted grows, the accuracy of RNNs decrease.
- At this point LSTM can solve this problem with its property of understanding at which point in the past information is useful.
- The cells within LSTM determine which information to keep and which information will be deleted. Later, they merge the previous state, current memory and input. Such units are understood to be very effective in capturing long term dependencies.