

Artificial Intelligence

Unit 09

Recurrent Neural Networks (RNN)

By:

Syeda Saleha Raza



AL NAFI,
A company with a focus on education,
wellbeing and renewable energy.

اَللّٰهُمَّ اِنِّیْ اَسْأَلُكَ عِلْمًا نَّافِعًا ،
وَرِزْقًا طَیِّبًا ، وَعَمَلًا مُّتَقَبَّلًا ،

(O Allah, I ask You for beneficial knowledge,
goodly provision and acceptable deeds)

اے اللہ ، میں آپ سے سوال کرتی ہوں نفع بخش علم کا، طیب رزق کا، اور اس عمل کا

(Sunan Ibn Majah: 925)

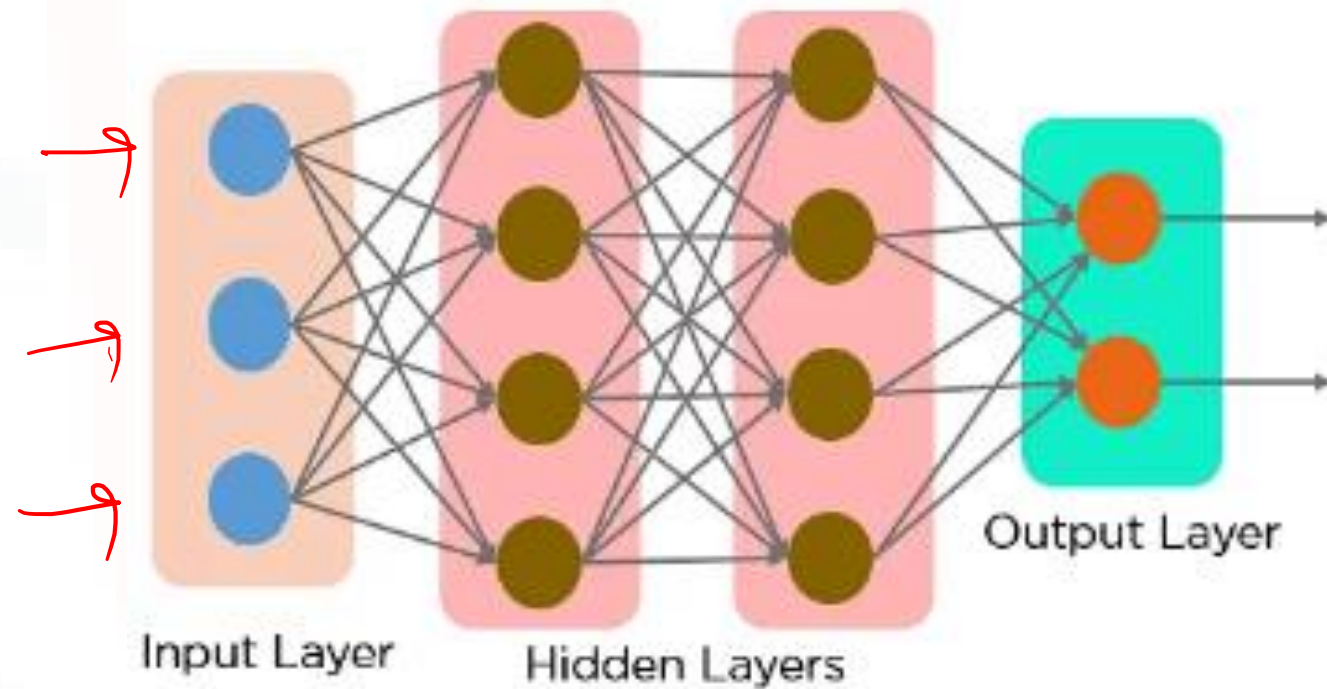
Outline

- Feed-forward Networks – A recap
- Motivation for RNN
- What are RNNs?
- Working of RNN
- Limitations of RNN
- Bi-directional Neural Networks
- Code Walkthrough

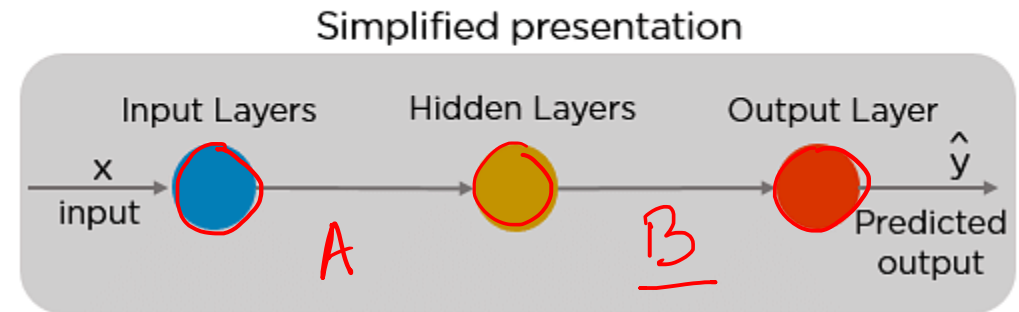
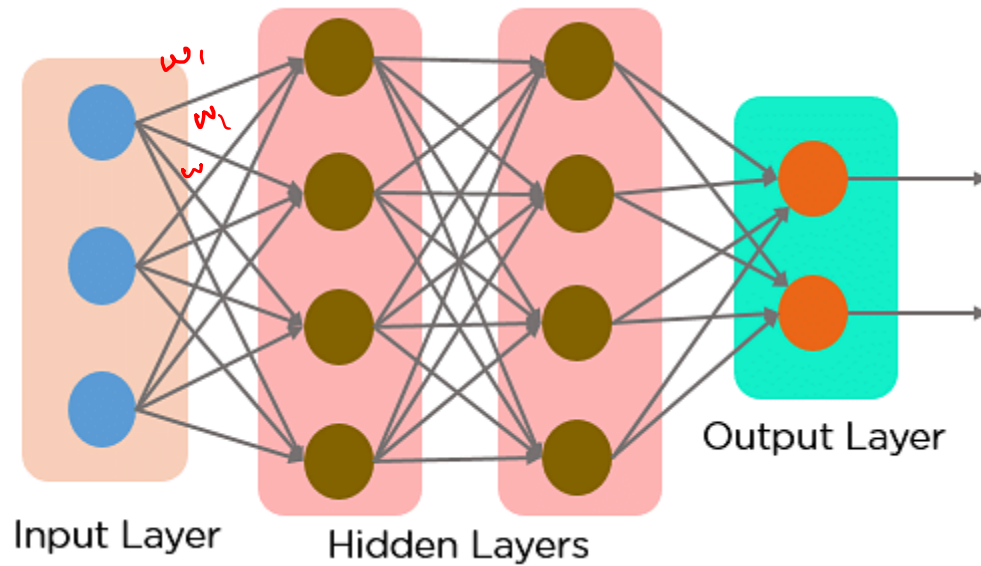
Acknowledgement

- <https://www.v7labs.com/blog/recurrent-neural-networks-guide>
- [Illustrated Guide to Recurrent Neural Networks | by Michael Phi | Towards Data Science](#)

Recap – Feed-Forward Neural Networks



A simplified representation



Some limitations

- Cannot handle sequential data
- Considers only the current input
- Cannot memorize previous inputs

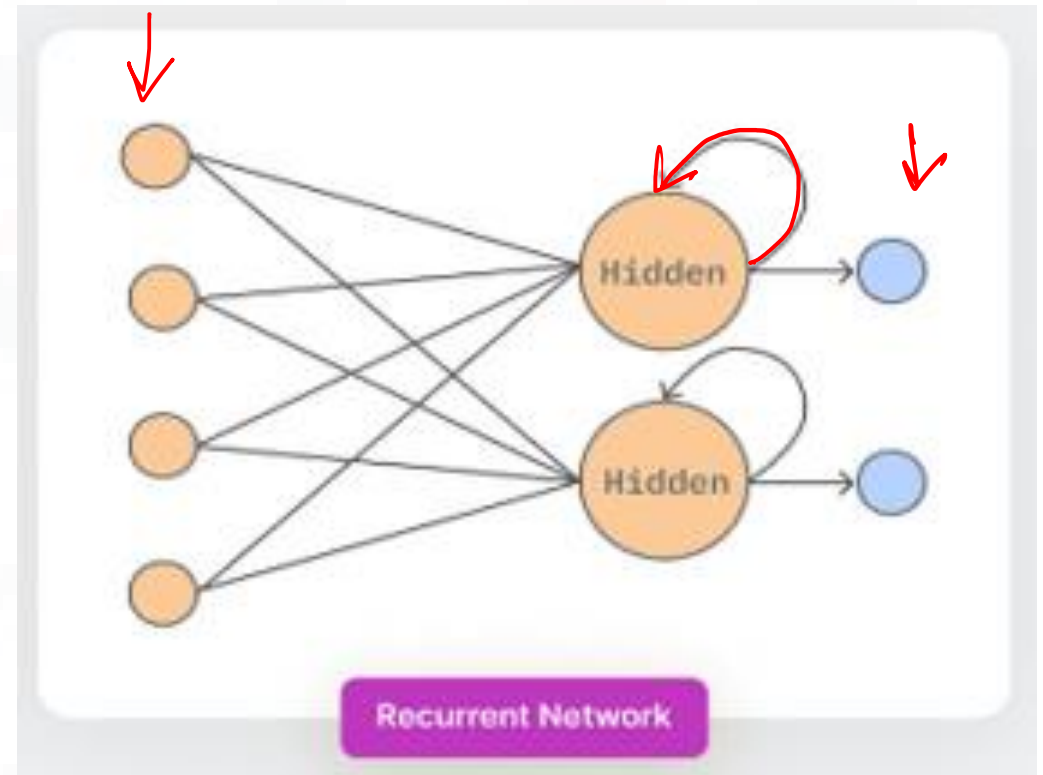
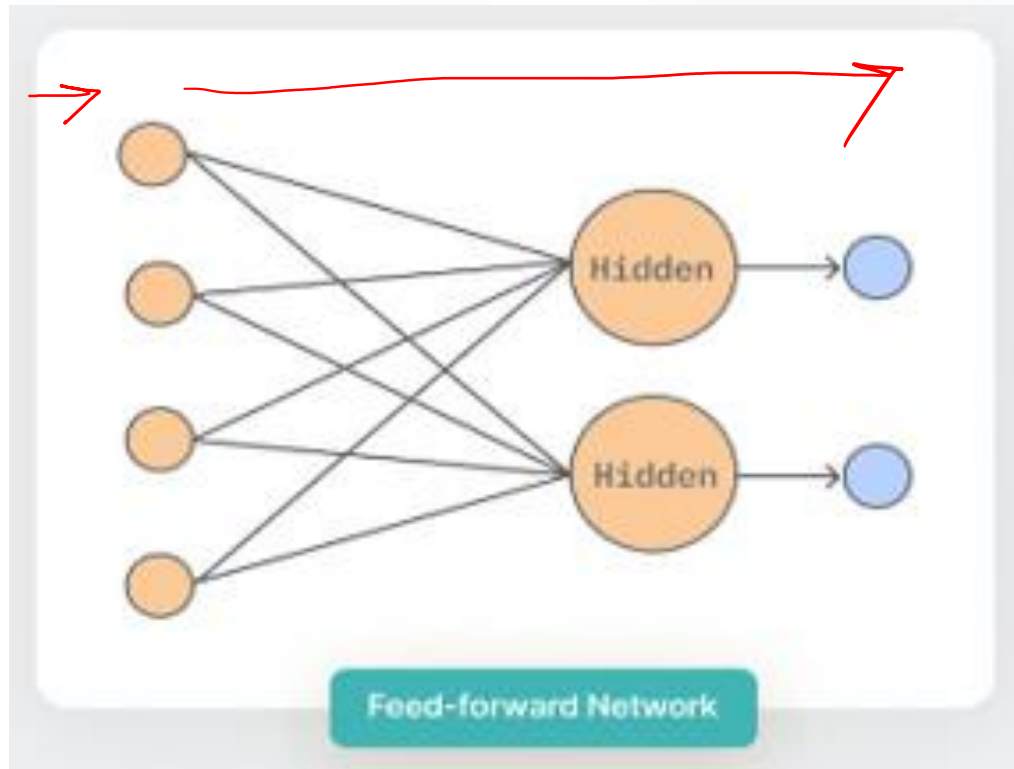
What is RNN?

- Recurrent Neural Network is a type of Artificial Neural Network that are good at modeling sequential data.



"One of the other reviewers has mentioned that after watching just 1 Oz episode you'll be hooked. They are right, as this is exactly what happened with me. The first thing that struck me about Oz was its brutality and unflinching scenes of violence, which set in right from the word GO. Trust me, this is not a show for the faint hearted or timid. This show pulls no punches with regards to drugs, sex or violence. Its is hardcore, in the classic use of the word. It is called OZ as that is the nickname given to the Oswald Maximum Security State Penitentiary. It focuses mainly on Emerald City, an experimental section of the prison where all the cells have glass fronts and face inwards, so privacy is not high on the agenda. Em City is home to many..Aryans, Muslims, gangstas, Latinos, Christians, Italians, Irish and more...so scuffles, death stares, dodgy dealings and shady agreements are never far away. I would say the main appeal of the show is due to the fact that it goes where other shows wouldn't dare. Forget pretty pictures painted for mainstream audiences, forget charm, forget romance...OZ doesn't mess around. The first episode I ever saw struck me as so nasty it was surreal, I couldn't say I was ready for it, but as I watched more, I developed a taste for Oz, and got accustomed to the high levels of graphic violence. Not just violence, but injustice (crooked guards who'll be sold out for a nickel, inmates who'll kill on order and get away with it, well mannered, middle class inmates being turned into prison bitches due to their lack of street skills or prison experience) Watching Oz, you may become comfortable with what is uncomfortable viewing...thats if you can get in touch with your darker

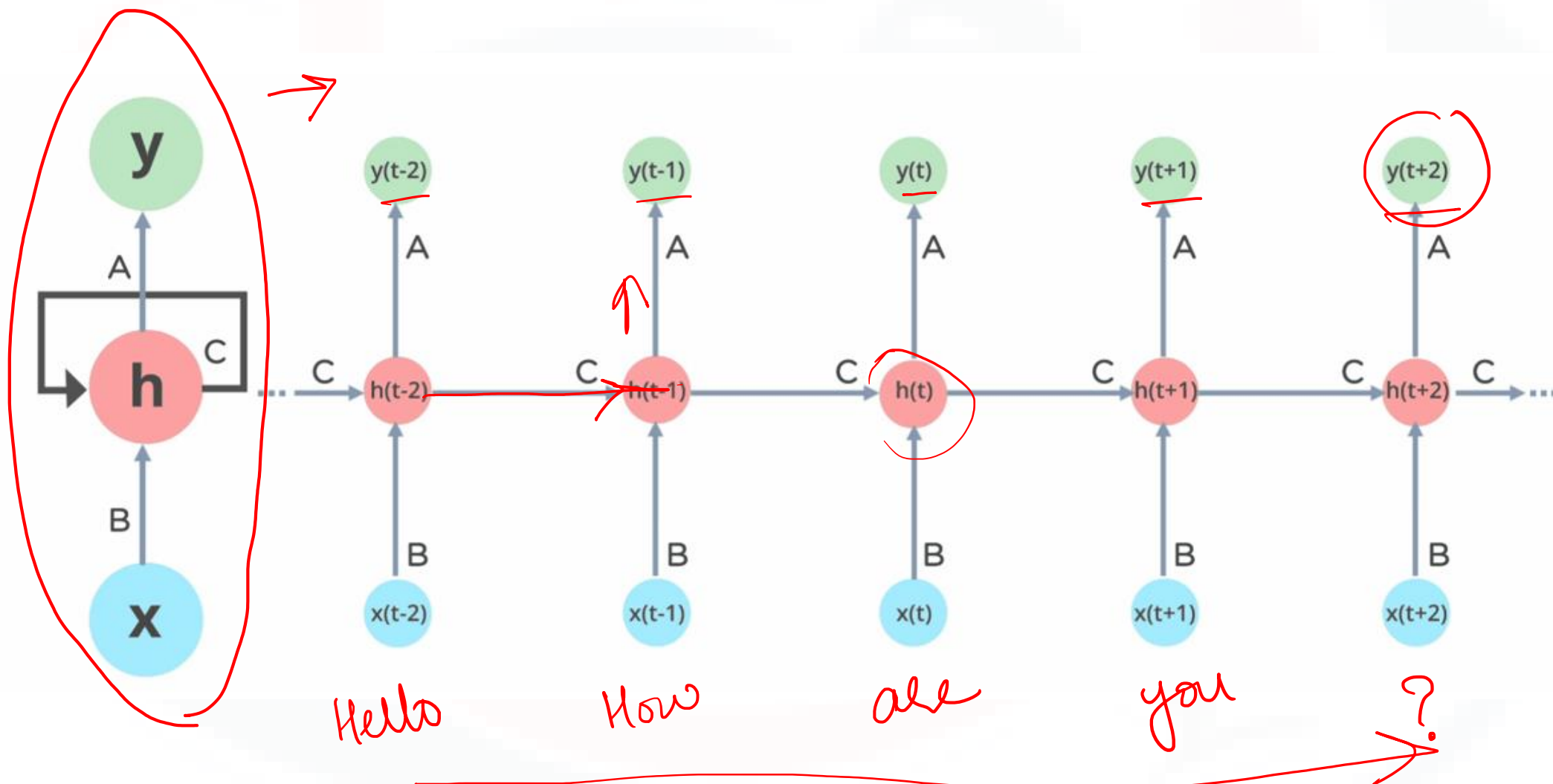
Feed-forward vs Recurrent Neural Network



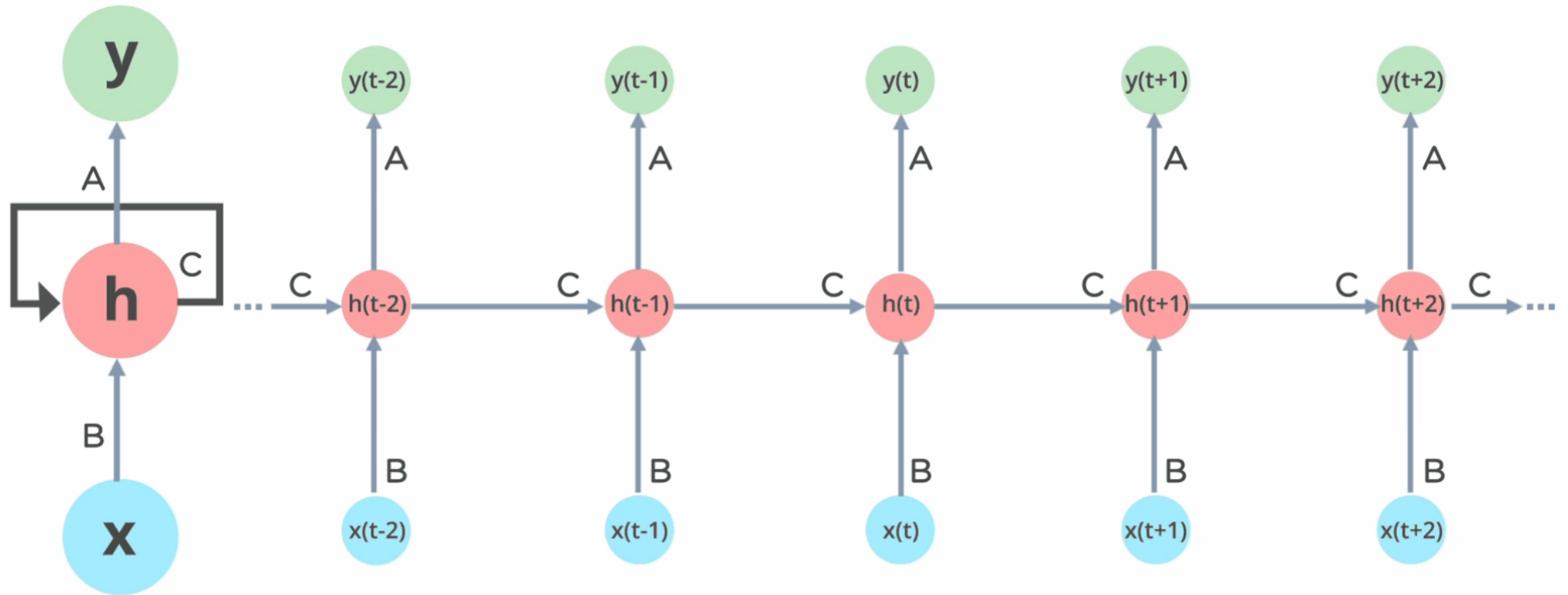
Recurrent Neural Networks (RNN)

- With recurrent connections, RNN has inherent “*memory*” as they take information from prior inputs to influence the current input and output. One can think of this as a hidden layer that remembers information through the passage of time.

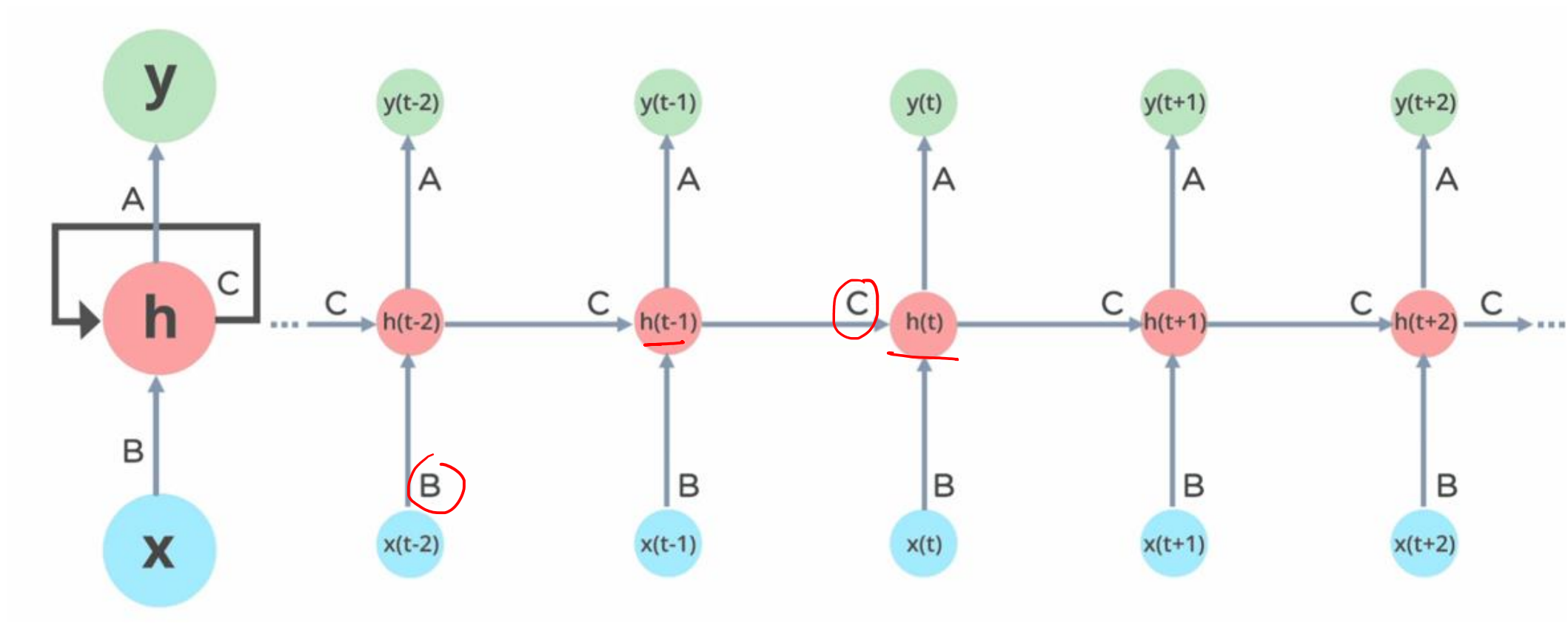
Unfolding RNN



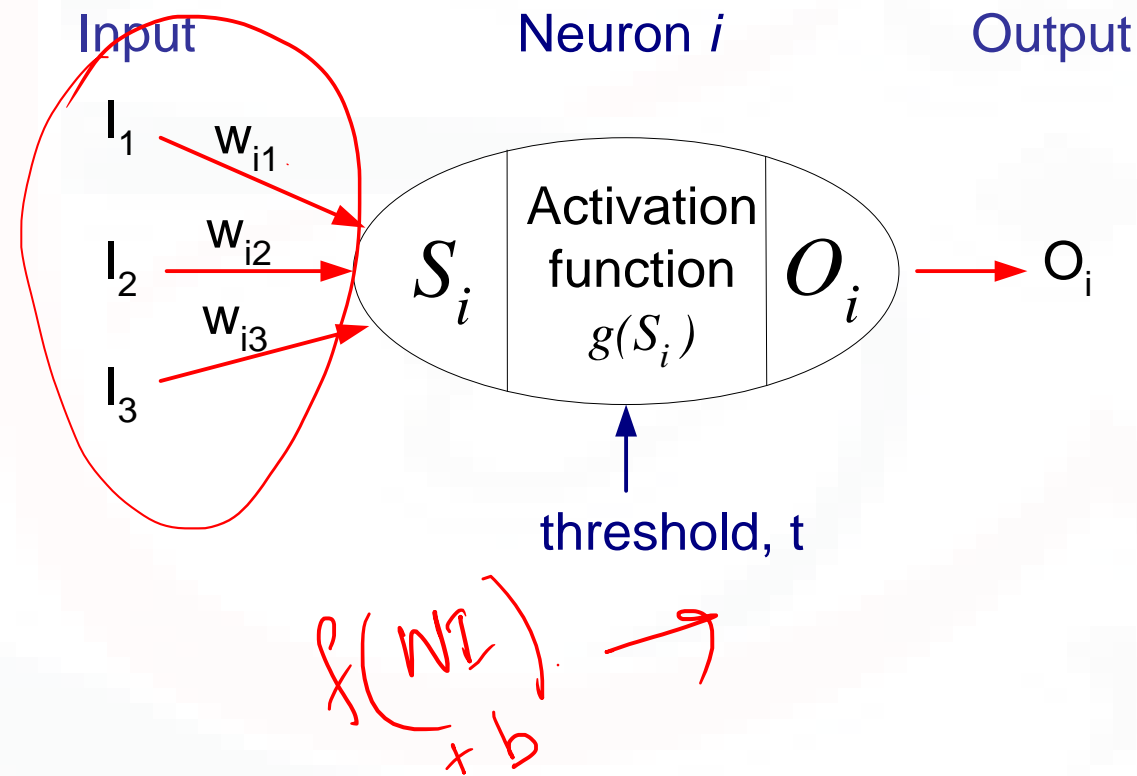
How do RNNs work?



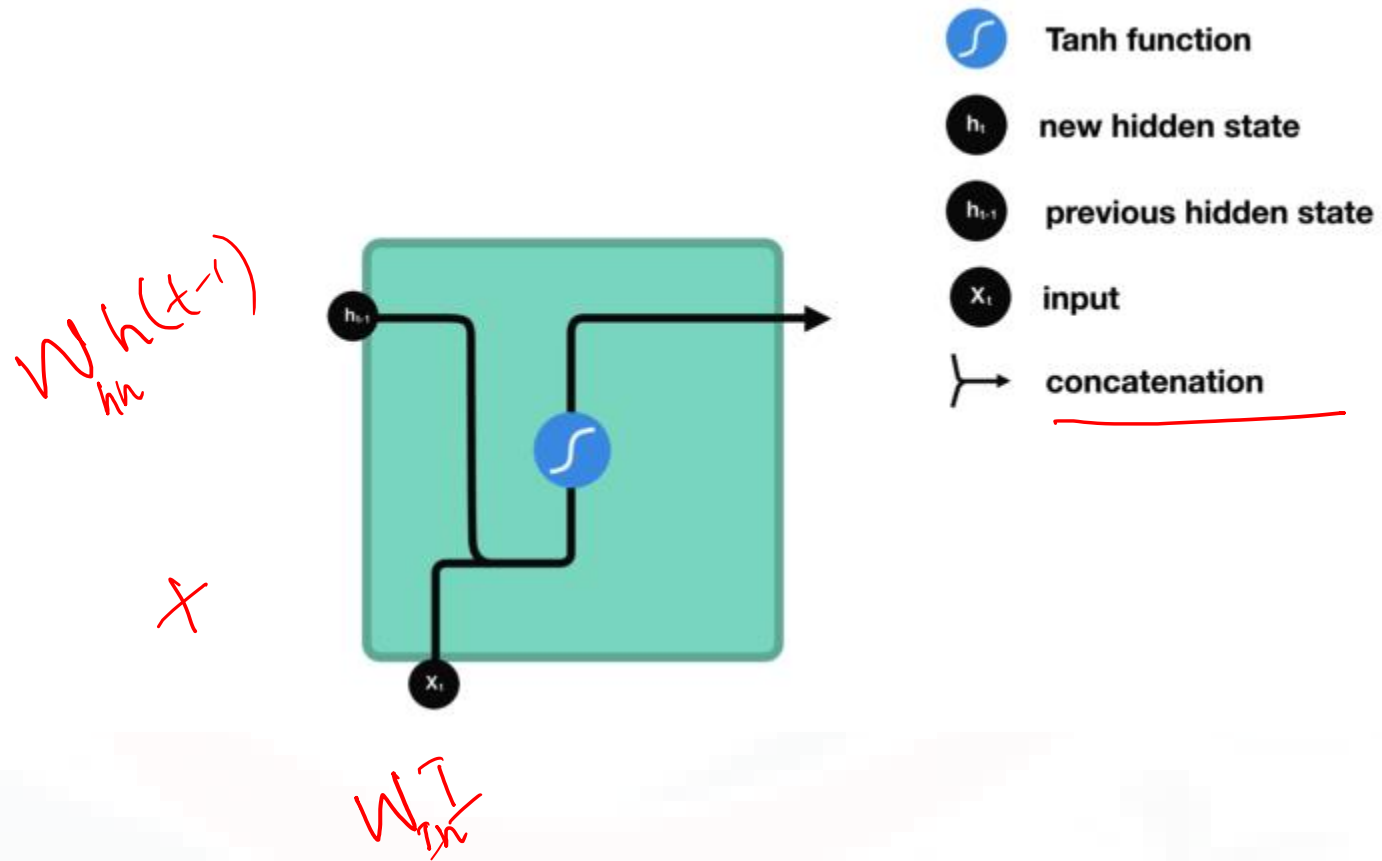
RNN



Neuron in a conventional feed-forward network

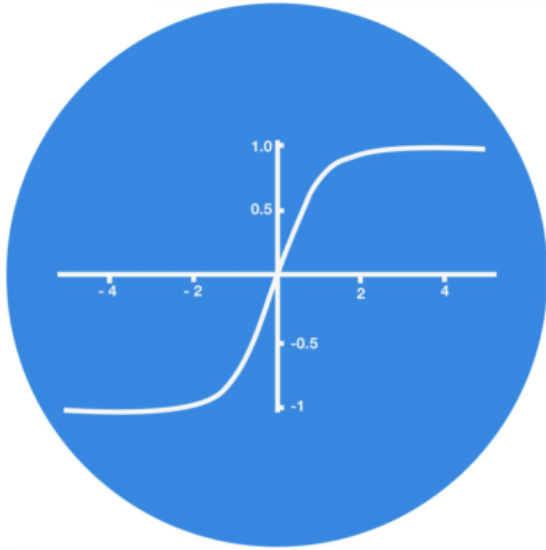


RNN Cell



Activation Functions – tanh vs sigmoid

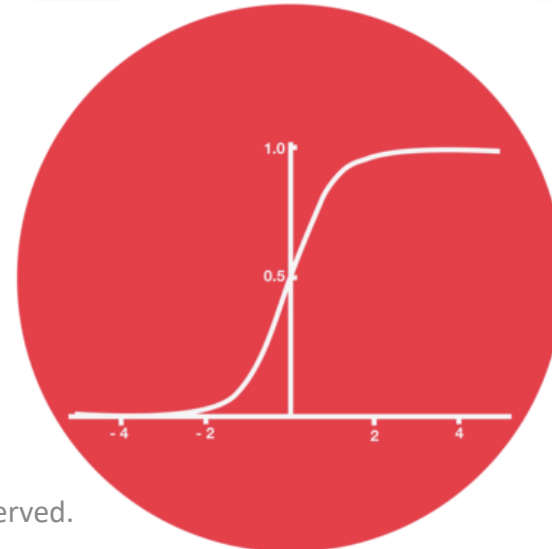
5
0.1
-0.5



-1

1

5
0.1
-0.5



0-1

Some applications in NLP



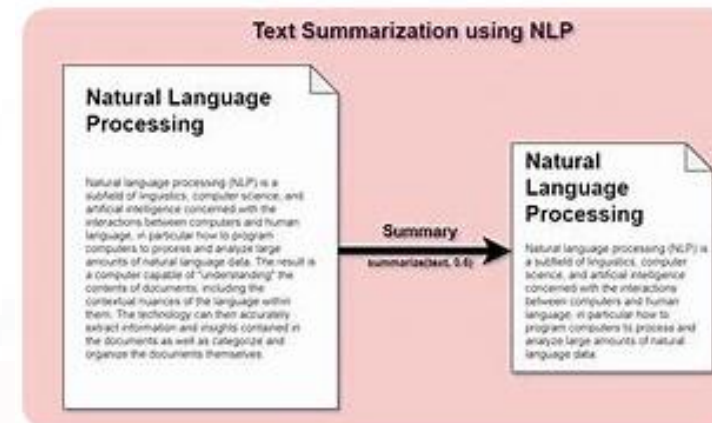
Machine Translation



Speech Recognition



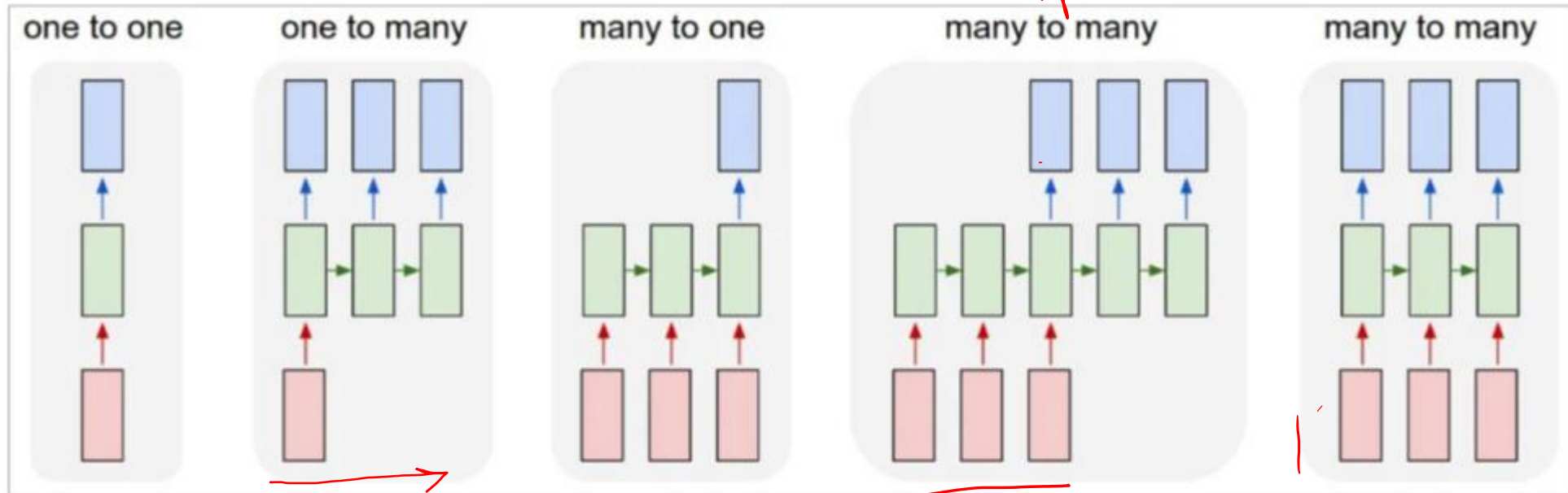
Image Captioning



Text Summarization

Types of RNN

Machine Translation



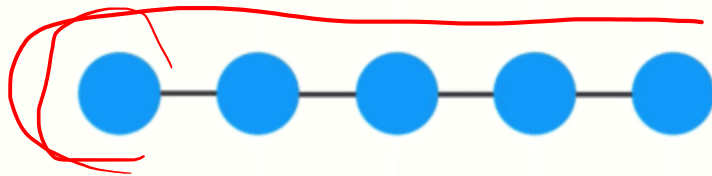
Each rectangle is a vector and arrows represent functions (e.g. matrix multiply). Input vectors are in red, output vectors are in blue and green vectors hold the RNN's state (more on this soon). From left to right: **(1)** Vanilla mode of processing without RNN, from fixed-sized input to fixed-sized output (e.g. image classification). **(2)** Sequence output (e.g. image captioning takes an image and outputs a sentence of words). **(3)** Sequence input (e.g. sentiment analysis where a given sentence is classified as expressing positive or negative sentiment). **(4)** Sequence input and sequence output (e.g. Machine Translation: an RNN reads a sentence in English and then outputs a sentence in French). **(5)** Synced sequence input and output (e.g. video classification where we wish to label each frame of the video). Notice that in every case are no pre-specified constraints on the lengths sequences because the recurrent transformation (green) is fixed and can be applied as many times as we like.

Challenges of RNN

- Training of RNN models can be very difficult and time-consuming as compared to other Artificial Neural Networks.
- RNNs are not able to keep track of long-term dependencies.
- Prone to problems such as *exploding* and *gradient* vanishing.

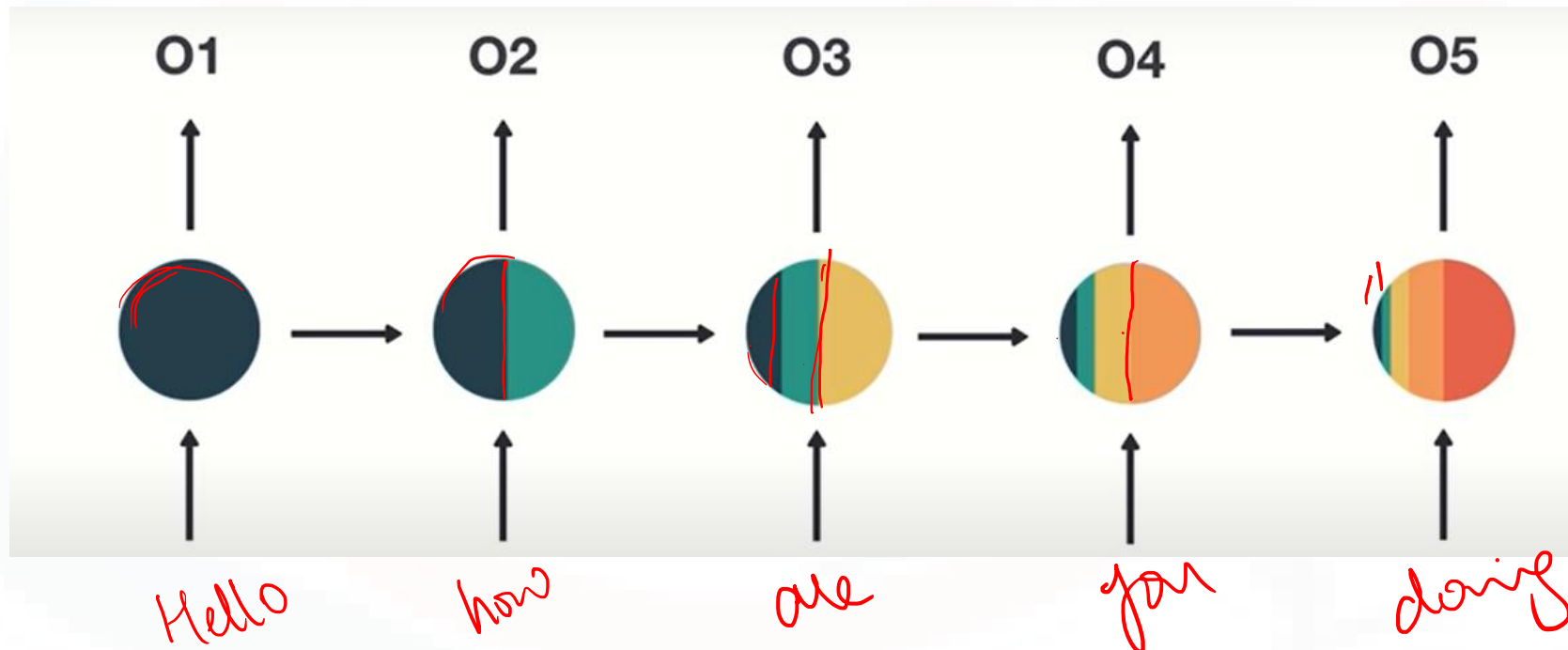
Limitations of RNN – Vanishing gradient

- One main limitation of RNN is that the gradient either explodes or vanishes; The network doesn't learn much from the data which is far away from the current position.



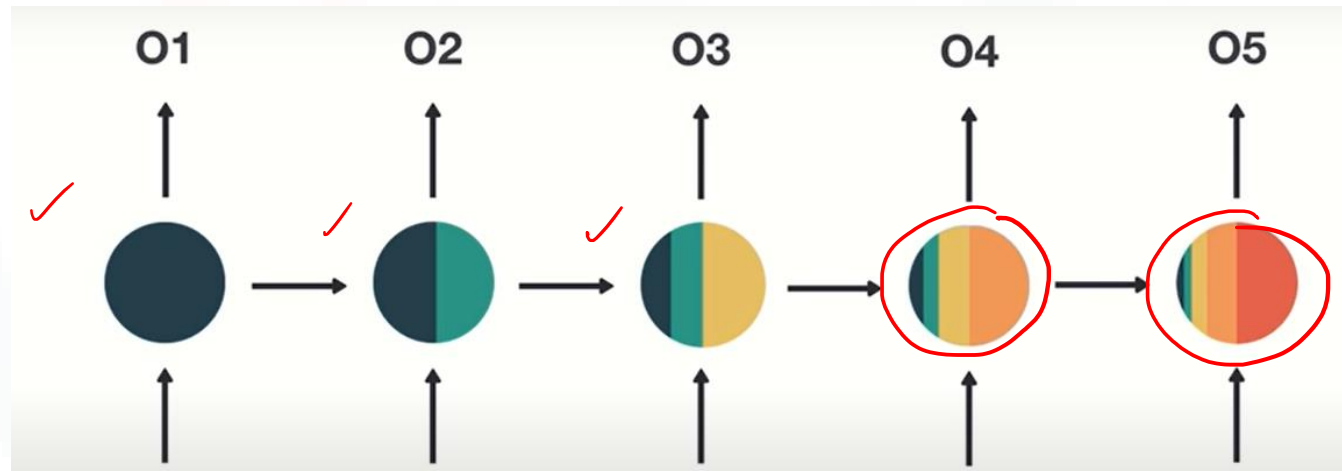
Limitations of RNN – short term memory

- RNNs have short term memory problem. To overcome this problem specialized versions of RNN are created like LSTM, GRU.



Limitations of RNN – temporal order

- Another limitation of RNN is that it processes inputs in a strict temporal order. This means current input has context of previous inputs but not the future.
- Bidirectional RNN (BRNN) duplicates the RNN processing chain so that inputs are processed in both forward and reverse time order.

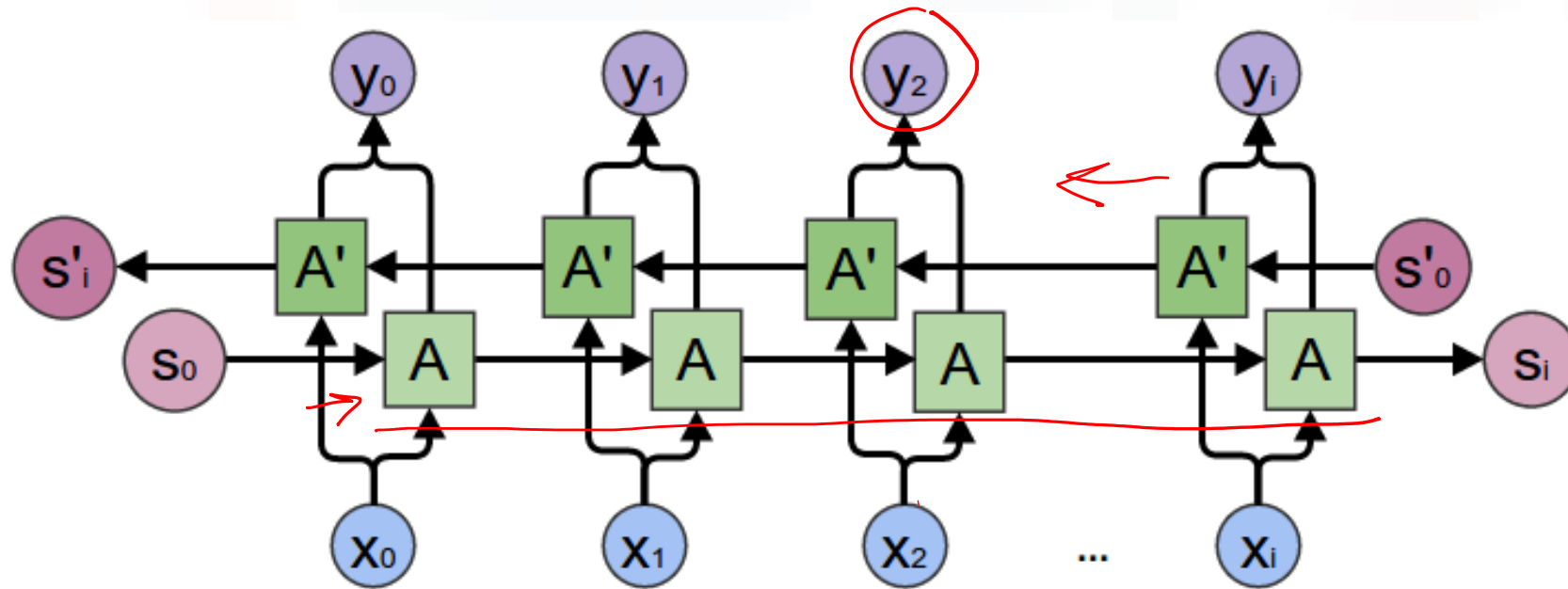


*Machine
Translation*

Bidirectional Recurrent Neural Networks (BRNN)

- An architecture of a neural network called a bidirectional recurrent neural network (BRNN) is made to process sequential data.
- In order for the network to use information from both the past and future context in its predictions, BRNNs process input sequences in both the forward and backward directions.
- This is the main distinction between BRNNs and conventional recurrent neural networks.

BRNN



Code Walkthrough

Coming next...

- Long-Short Term Model (LSTM)
- Gated Recurrent Unit (GRU)

References

- <https://www.v7labs.com/blog/recurrent-neural-networks-guide>
- https://medium.com/@humble_bee/rnn-recurrent-neural-networks-lstm-842ba7205bbf
- [Illustrated Guide to Recurrent Neural Networks | by Michael Phi | Towards Data Science](#)

جزاك الله

To ask questions, Please use communities link
(for respective course) within portal

<https://portal.alnafi.com/enrollments>