California State University, Dominguez Hills
Department of Computer Science
CSC 595

Professor: Dr. Benyamin Ahmadnia
bahmadniayebosari@csudh.edu

Spring 2025

Copyright Notice

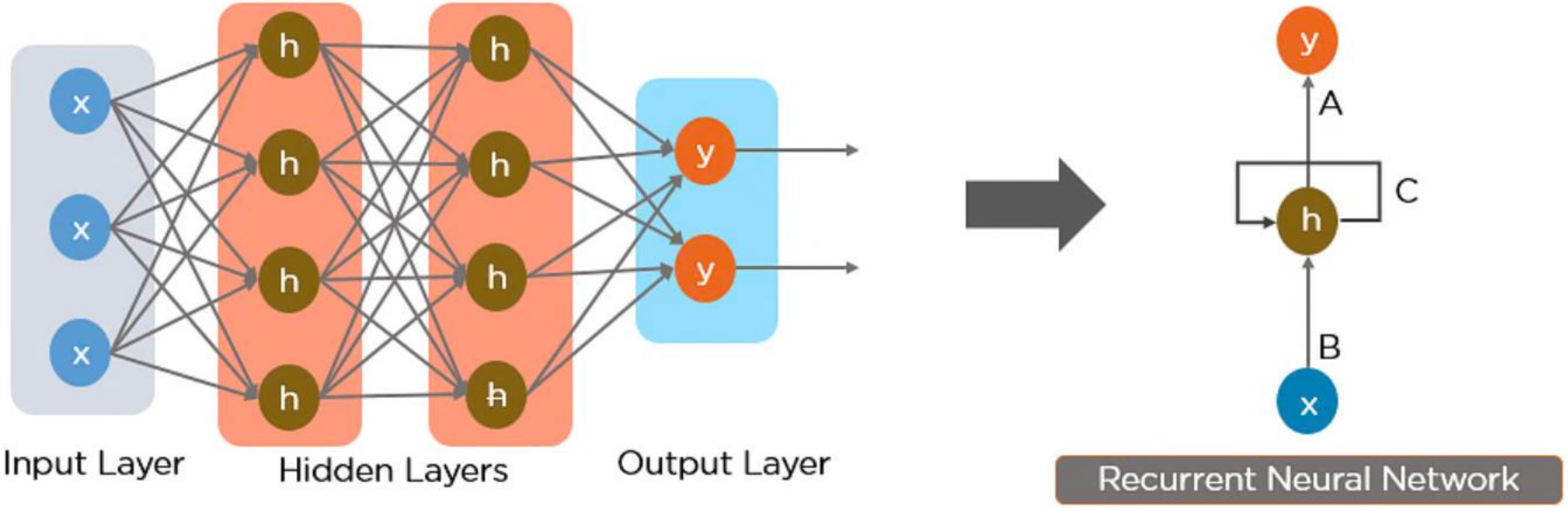
Table of Contents

- RNN Introduction
- RNN Applications
- RNN Advantages and Disadvantages
- RNN Types
- RNN Issues

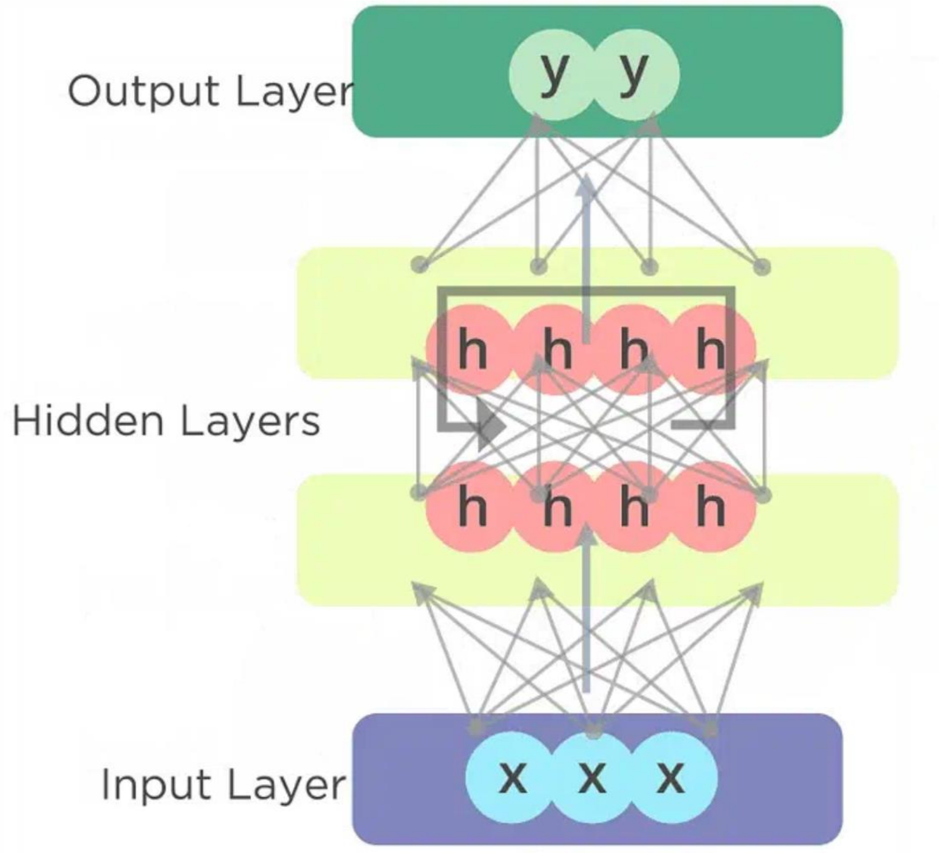
Introduction to RNN

Introduction to RNN

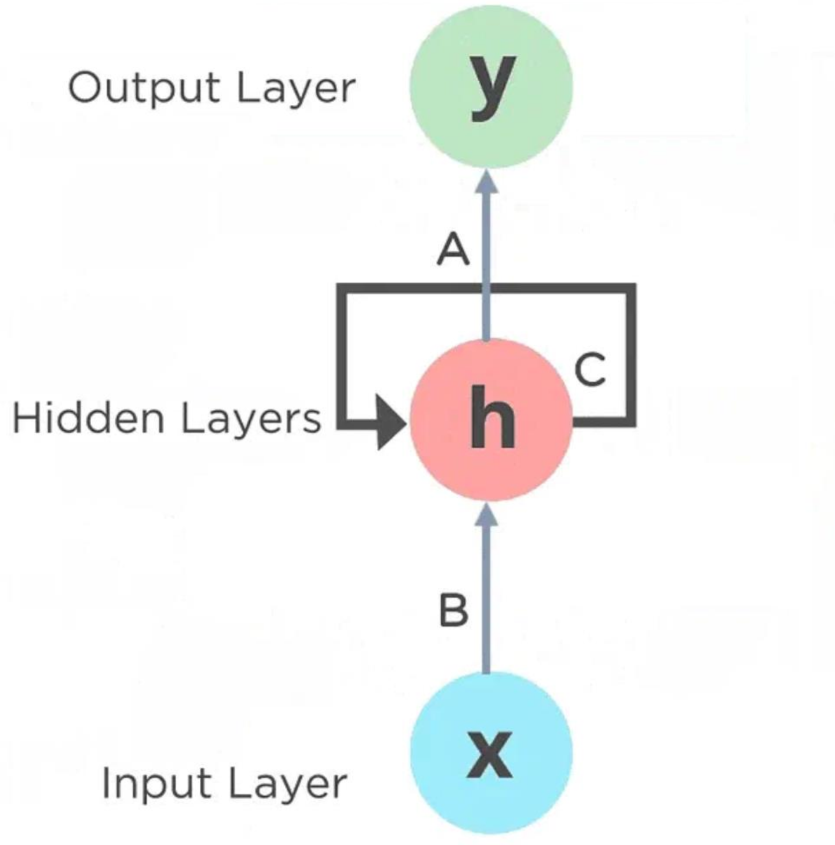
- RNN works on the principle of saving the output of a particular layer and feeding this back to the input in order to predict the output of the layer.



Introduction to RNN

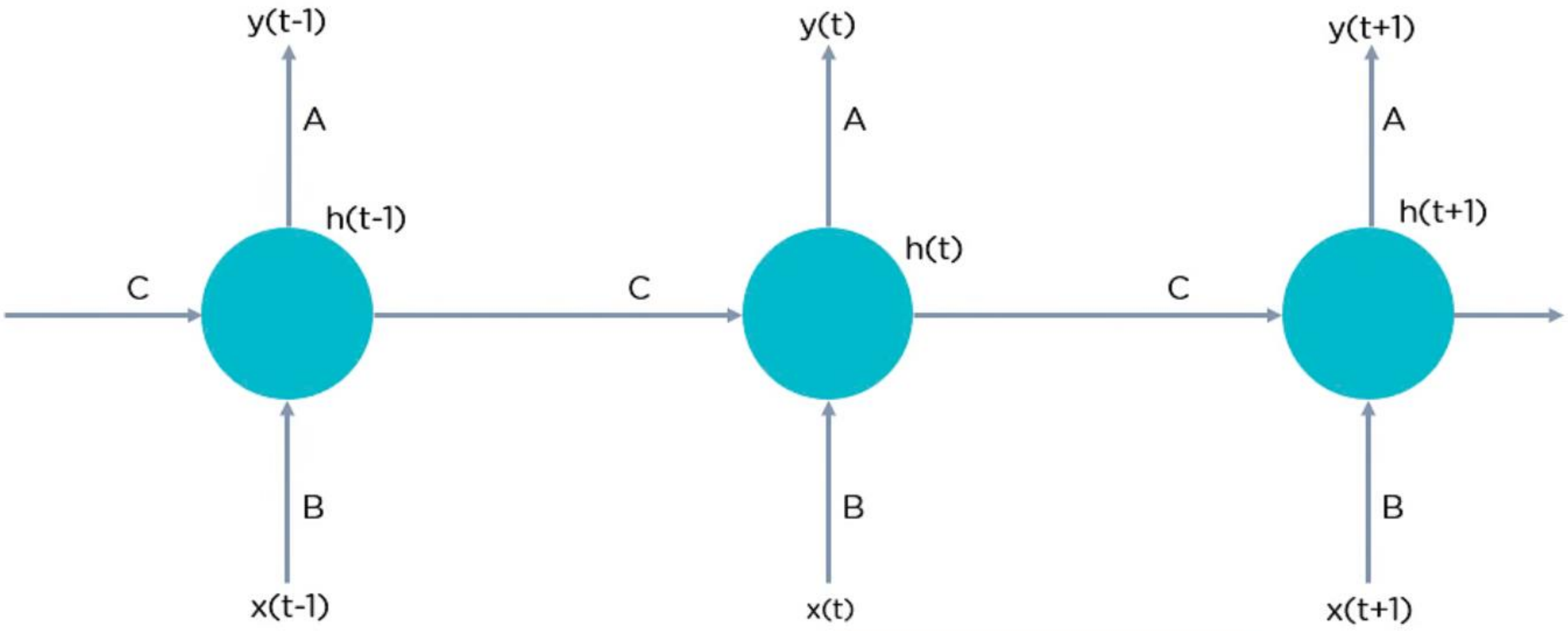


A, B and C are the parameters



A, B and C are the parameters

Introduction to RNN



$h(t) = f_c\ (h(t\text{-}1), x(t))$

$h(t)$ = new state

f_c = function with parameter c

$h(t\text{-}1)$ = old state

$x(t)$ = input vector at time step t

Why RNN?

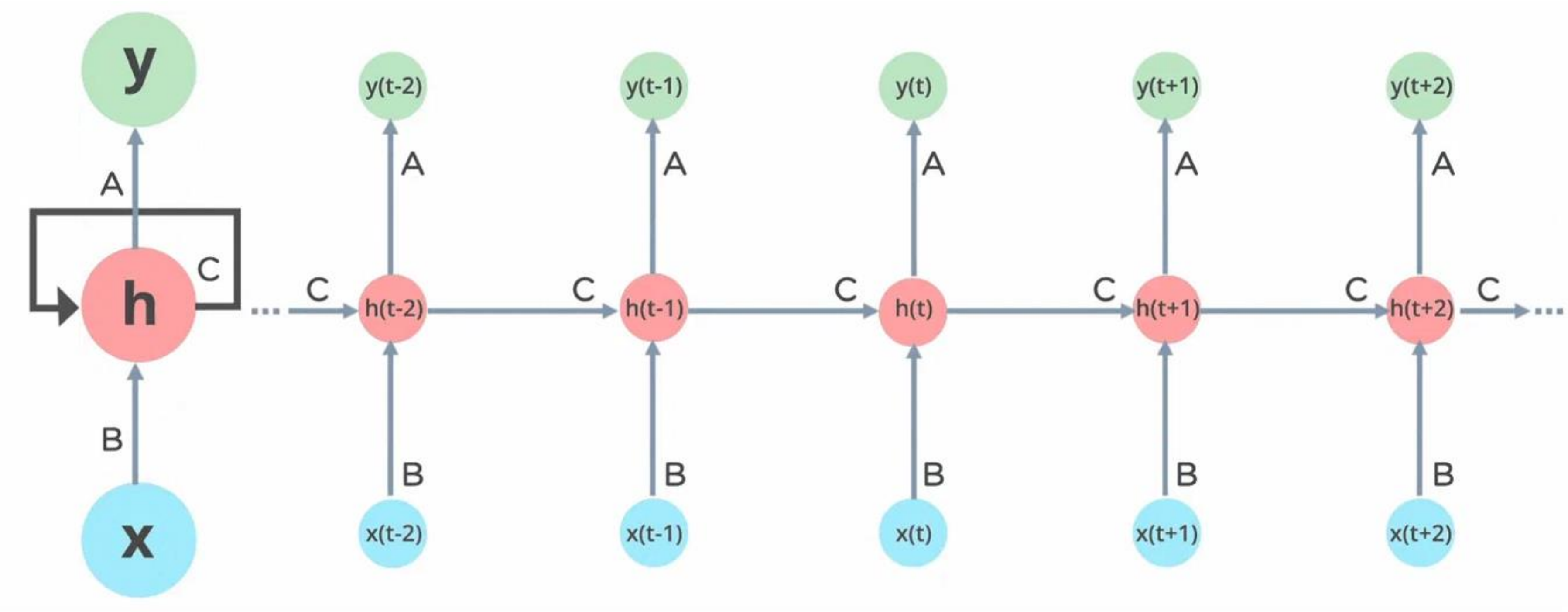
- Feed-forward neural network issues:

  1) Cannot handle sequential data

  2) Considers only the current input

  3)Cannot memorize previous inputs

How does RNN work?



Applications of RNNs

RNN Applications

- Image Captioning
- Time Series Prediction
- Natural Language Processing
- Machine Transaltion

Advantages and Disadvantages of RNNs

RNNs Advantages

- Ability to handle variable-length sequences
- Memory of past inputs
- Parameter sharing
- Non-linear mapping
- Sequential processing
- Flexibility
- Improved accuracy

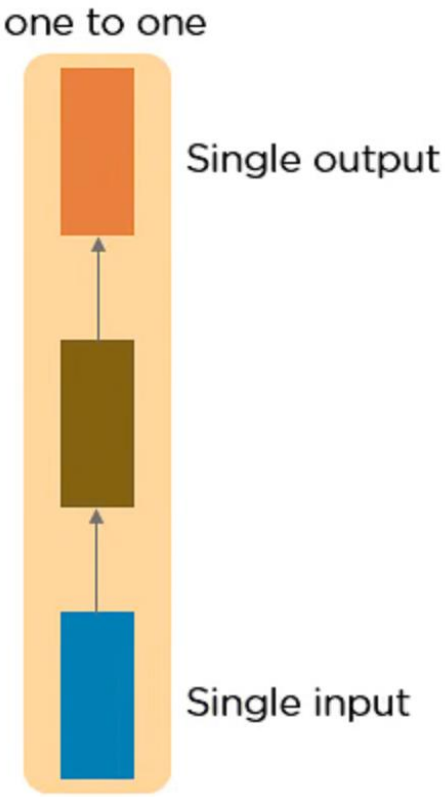
RNNs Disadvantages

- Vanishing and exploding gradients
- Computational complexity
- Difficulty in capturing long-term dependencies
- Lack of parallelism
- Difficulty in choosing the right architecture
- Difficulty in interpreting the output

Types of RNNs

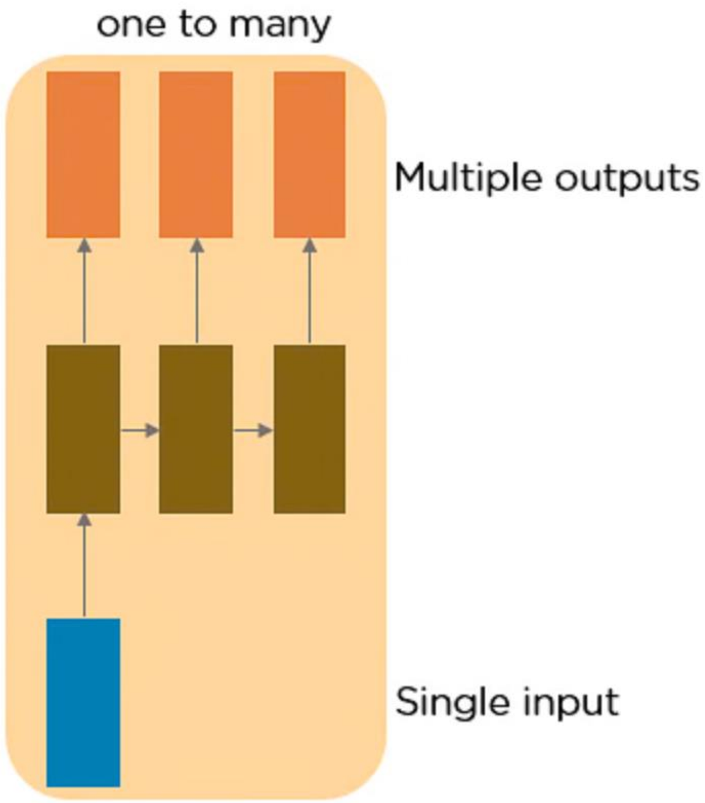
One-to-One RNN

- This type of neural network is known as the Vanilla Neural Network. It's used for general machine learning problems with a single input and output.



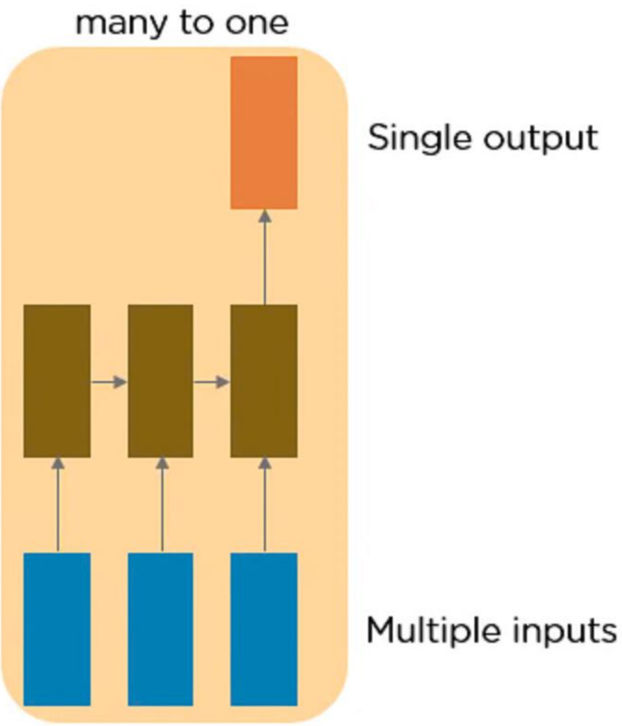
One-to-Many RNN

- This type of neural network has a single input and multiple outputs. An example of this is the image caption.



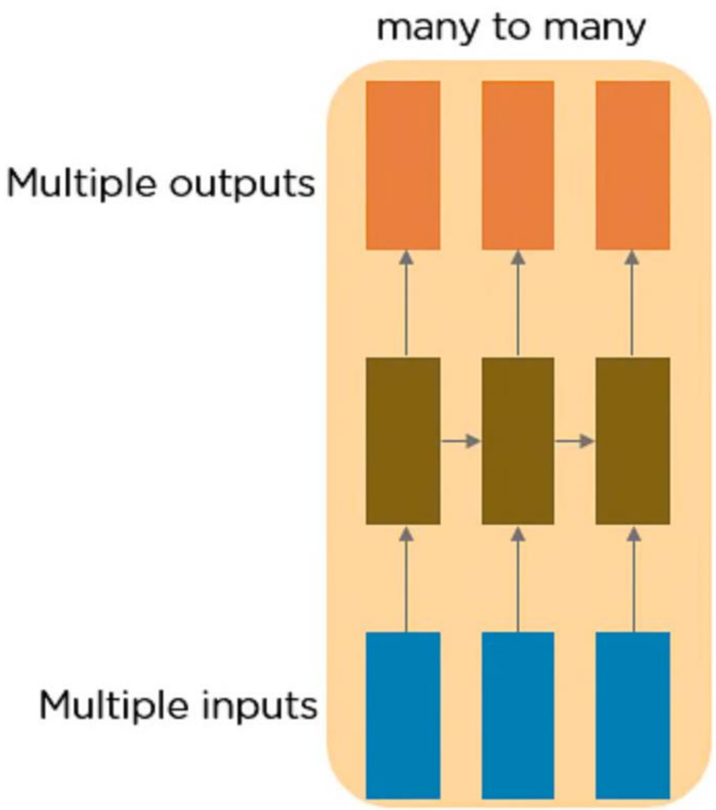
Many-to-One RNN

- This RNN takes a sequence of inputs and generates a single output. Sentiment analysis is a good example of this kind of network, where a given sentence can be classified as expressing positive or negative sentiments.



Many-to-Many RNN

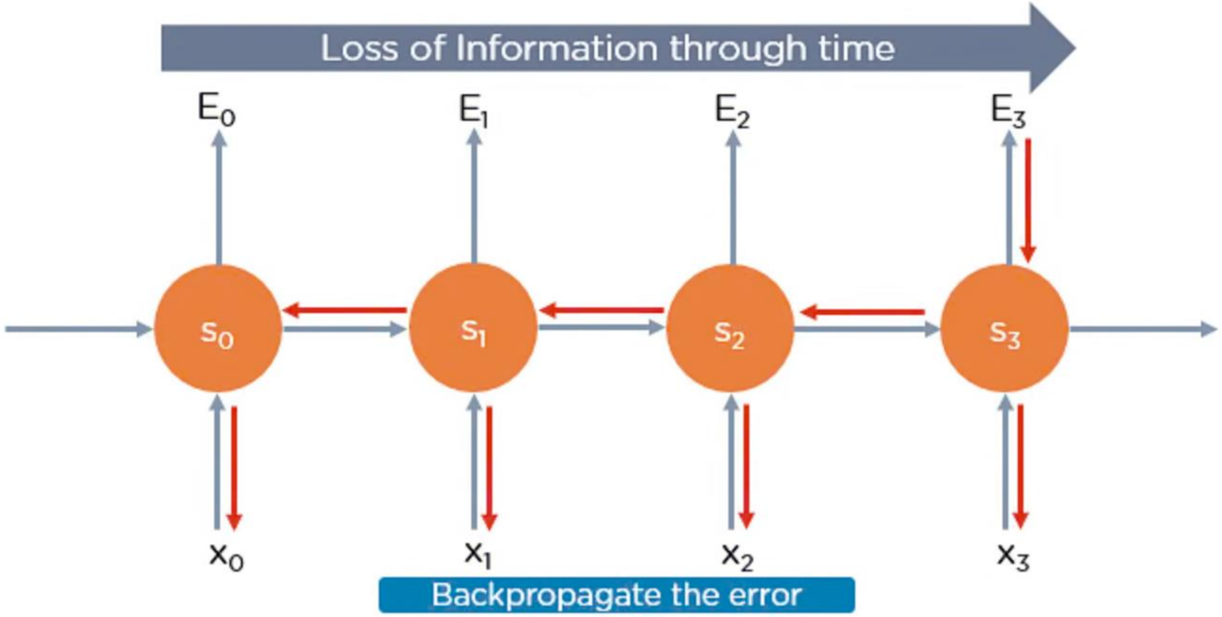
- This RNN takes a sequence of inputs and generates a sequence of outputs. Machine translation is one example.



Two Issues of Standard RNNs

Vanishing Gradient Problem

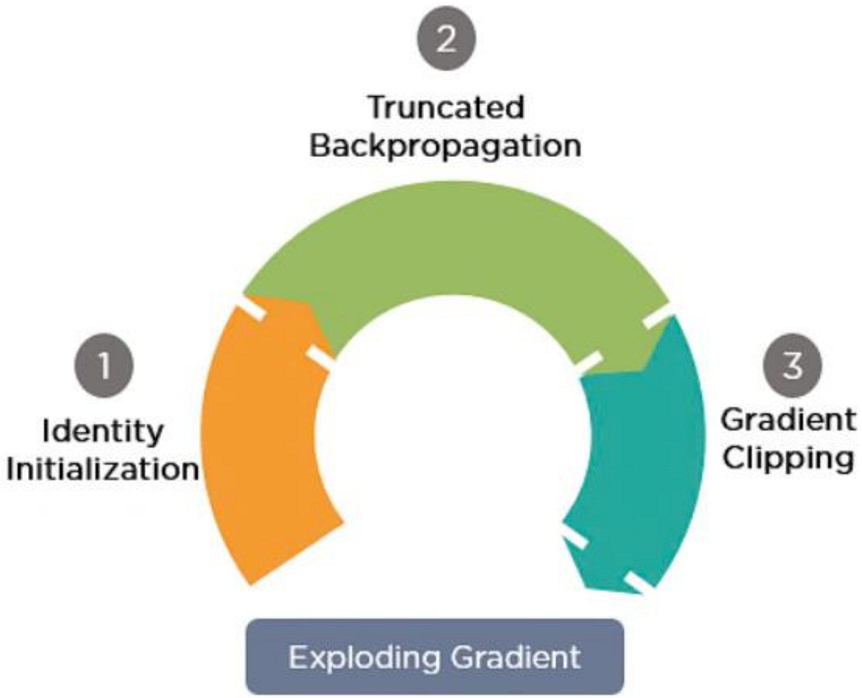
- RNN is hard to train because of the gradient problem.
- The gradients carry information used in the RNN, and when the gradient becomes too small, the parameter updates become insignificant. This makes the learning of long data sequences difficult.



Exploding Gradient Problem

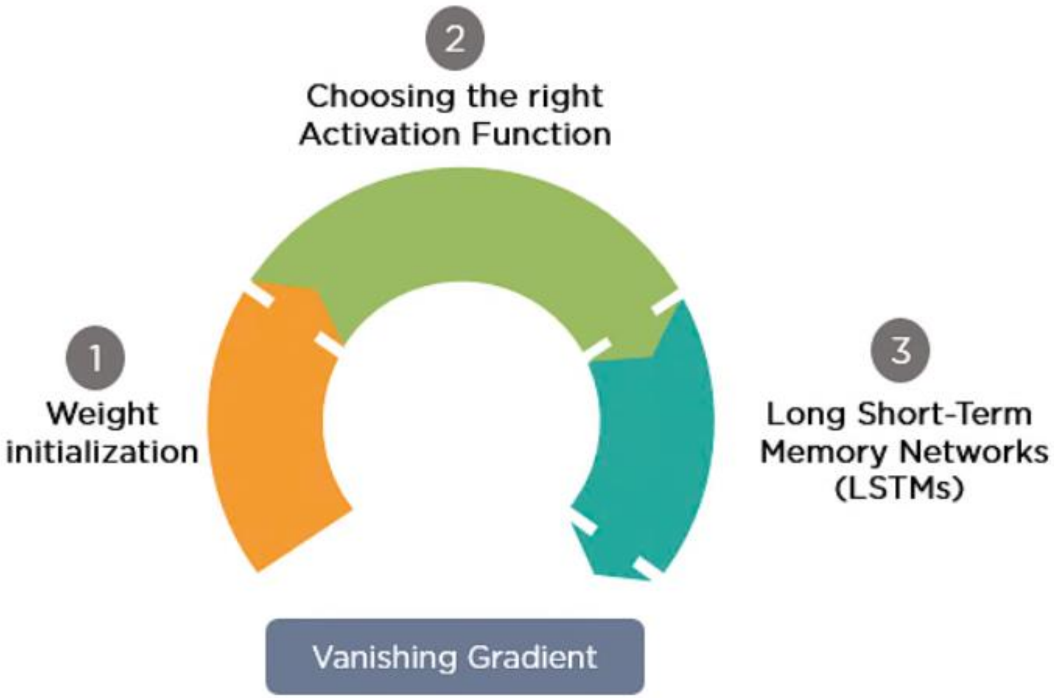
- If the slope of a neural network tends to grow exponentially instead of decaying while being trained, this is called an Exploding Gradient. This problem arises when large error gradients accumulate, resulting in very large updates to the neural network model weights during the training process.
- Long training time, poor performance, and lousy accuracy are the major issues in gradient problems.

Gradient Problem Solutions



Exploding Gradient

1. Identity Initialization
2. Truncated Backpropagation
3. Gradient Clipping



Vanishing Gradient

1. Weight initialization
2. Choosing the right Activation Function
3. Long Short-Term Memory Networks (LSTMs)