# *Problems with RNN*

1.) Problem of Long-Term Dependency
2.) Problem of Unstable Gradients or Stagnetted Training

Becoz of this 2 Problems, RNN aren't used , Dono problem 1 hie fact se aati hai

That is unstable gradients

1.) Problem of Long-Term Dependency

Yeah Problem Vanishing Gradient issue se hota hai

2.) Problem of Unstable Gradients or Stagnetted Training Means kahi baar hota hai ki RNN barabr se Train hie nahi hota hai iska 1 Reason Exploding Gradient Problem hota hai

1. **Problem of Long-Term Dependency:**
   - In traditional RNNs, it can be challenging for the network to capture long-term dependencies in sequences. When processing a sequence of data, information from earlier time steps may become diluted or lost as it propagates through the network, making it difficult for the RNN to learn and remember dependencies that exist across long sequences.
   - This issue is often referred to as the "vanishing gradient problem," where gradients become very small during backpropagation, leading to minimal updates to the weights of the network's earlier layers. As a result, the network has difficulty learning relationships that depend on information from distant time steps.
2. **Problem of Unstable Gradients:**
   - The problem of unstable gradients is closely related to the vanishing gradient problem. In some cases, gradients can become extremely large during training, leading to instability and difficulties in converging to an optimal solution. This is known as the "exploding gradient problem."
   - Unstable gradients can cause the weights of the network to be updated too drastically, leading to oscillations or divergence during training. This instability makes it challenging to train deep networks, including deep RNNs.

To address these issues, various improvements and architectures have been proposed, such as:

- **Long Short-Term Memory (LSTM) networks:** LSTMs are designed to better capture long-term dependencies by introducing memory cells and gating mechanisms that regulate the flow of information through the network.
- **Gated Recurrent Unit (GRU) networks:** Similar to LSTMs, GRUs use gating mechanisms to control the flow of information, but they have a slightly simpler architecture compared to LSTMs.
- **Gradient clipping:** This technique is used to prevent the exploding gradient problem by scaling gradients if they exceed a certain threshold during training.
- **Skip connections:** These connections allow gradients to flow more easily through the network, mitigating the vanishing gradient problem.

## ***Solution:***

## ***Solution of Long Term Dependency ( Vanishing GD Problm )***

1.) To use different Activation Function
    As tanh is 0 to 1 , Relu or Leaky Relu use karskte hai
2.) Better Weight Initialization
3.) We can use slight different Architectures like
    Skip RNN
4.) RNN Chodo , use different Architectures like LSTM

## ***Solution of Unstable Gradient (Exploding Gradient)***

*1.) Gradient Clipping*
*2.) Controlled Learning Rate*
*3.) Choddo RNN, Move to LSTM*