**Issues with RNNs and their extended models**

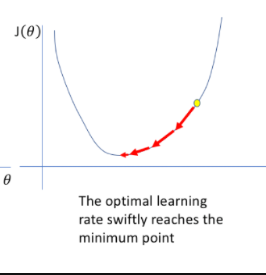
### 1. ****Problems with RNNs****:

The key challenges of RNNs revolve around two main concepts: **Gradient** and **Memory**.

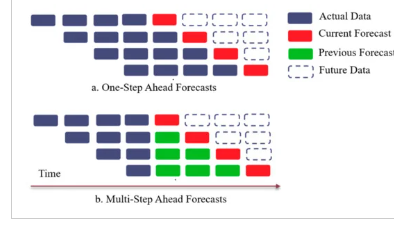
* **Gradient**: A vector that represents the direction and magnitude of the change needed to minimize the loss function when comparing input data with expected values.
* **Memory**: The network's ability to remember information over time steps, which is crucial for processing sequential data.

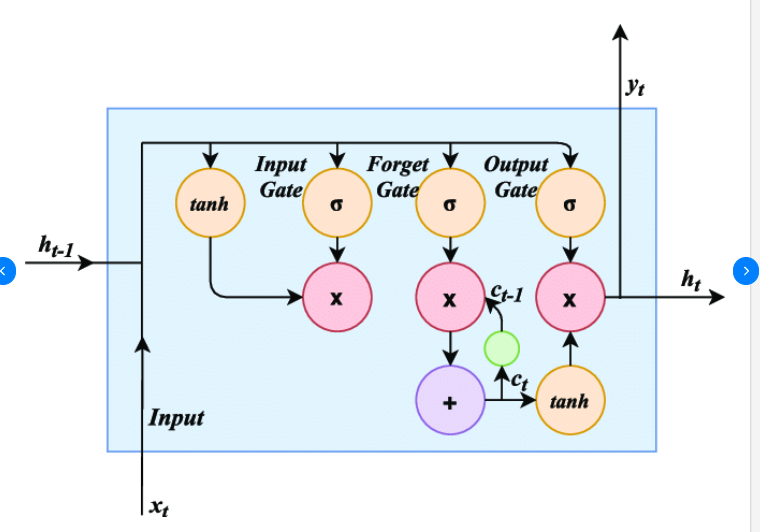
#### Specific Issues:

* **Vanishing Gradient**: One of the biggest challenges of RNNs. When the gradient becomes too small during training, RNNs fail to learn long-term dependencies in the data, which negatively impacts performance.



* **Exploding Gradient**: The opposite of Vanishing Gradient, where the gradient becomes too large, making the training process unstable and reducing the model’s ability to learn effectively.



* **Short-term Memory**: RNNs often struggle to retain information when dealing with long sequences. This limits the model's ability to capture long-term relationships, which can degrade prediction accuracy.
* 

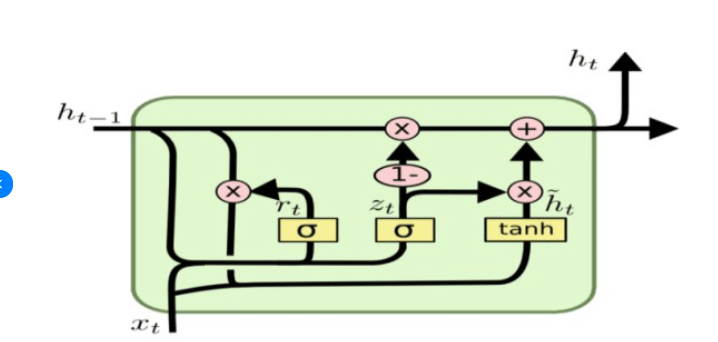
### 2. ****Extended Models of RNNs****

#### ****LSTM (Long Short-Term Memory)****:

LSTM uses "gates" to manage the flow of information through time steps. It addresses the vanishing gradient problem, **Exploding Gradient** and enhances the model's ability to store long-term information.

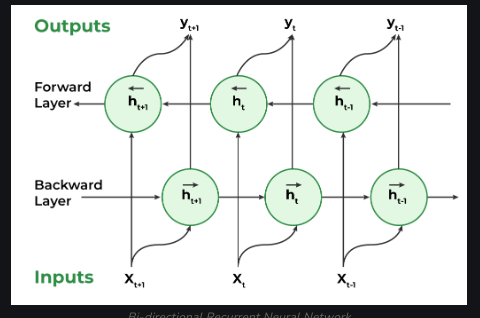
#### ****GRU (Gated Recurrent Unit)****:

GRU works similarly to LSTM but with a simpler architecture, involving fewer parameters. This makes training faster while still being effective for processing long sequences.



#### ****Bidirectional RNN****:

Bidirectional RNNs optimize sequence processing by reading the data in both directions, from past to future and vice versa. This allows the model to better understand context and relationships within the sequence.



**Các vấn đề của RNNs và các mô hình mở rộng**

### 1. ****Các vấn đề của RNNs****:

Các vấn đề chính của RNNs xoay quanh hai khái niệm: **Gradient** và **Memory**.

* **Gradient**: Là một vector thể hiện hướng và độ lớn của sự thay đổi cần thiết để giảm thiểu hàm mất mát khi so sánh đầu vào với giá trị mong muốn.
* **Memory**: Khả năng ghi nhớ thông tin của mạng qua các bước thời gian, đóng vai trò quan trọng trong việc xử lý chuỗi dữ liệu.

#### Các vấn đề cụ thể:

* **Vanishing Gradient**: Đây là một trong những vấn đề lớn nhất của RNNs. Gradient quá nhỏ trong quá trình huấn luyện khiến RNN không thể học được các mối quan hệ dài hạn trong chuỗi dữ liệu, làm giảm hiệu suất của mô hình.
* **Exploding Gradient**: Trái ngược với Vanishing Gradient, Exploding Gradient xảy ra khi gradient quá lớn, khiến quá trình huấn luyện trở nên không ổn định và dẫn đến việc mô hình không học hiệu quả.
* **Short-term Memory**: RNNs thường gặp khó khăn trong việc nhớ thông tin khi chuỗi dữ liệu quá dài. Điều này làm giảm khả năng của mô hình trong việc nắm bắt các mối quan hệ dài hạn, ảnh hưởng đến hiệu quả dự đoán.

### 2. ****Các mô hình mở rộng của RNNs****

#### ****LSTM (Long Short-Term Memory)****:

LSTM sử dụng các "cổng" (gates) để quản lý dòng thông tin qua các bước thời gian. Nó giải quyết vấn đề vanishing gradient,**Exploding Gradient** và cải thiện khả năng lưu trữ thông tin dài hạn.

#### ****GRU (Gated Recurrent Unit)****:

GRU hoạt động tương tự LSTM nhưng với kiến trúc đơn giản hơn, ít tham số hơn, giúp quá trình huấn luyện nhanh hơn mà vẫn hiệu quả trong việc xử lý chuỗi dài.

#### ****Bidirectional RNN****:

Bidirectional RNN tối ưu hóa khả năng xử lý chuỗi dữ liệu bằng cách đọc dữ liệu theo cả hai hướng, từ quá khứ đến tương lai và ngược lại. Điều này giúp mô hình hiểu rõ hơn về ngữ cảnh và các mối quan hệ trong chuỗi dữ liệu.