

# MACHINE LEARNING AND AI

## ЛЕКЦИЯ 9

# Sequential data and time series

## Что такое последовательные данные?

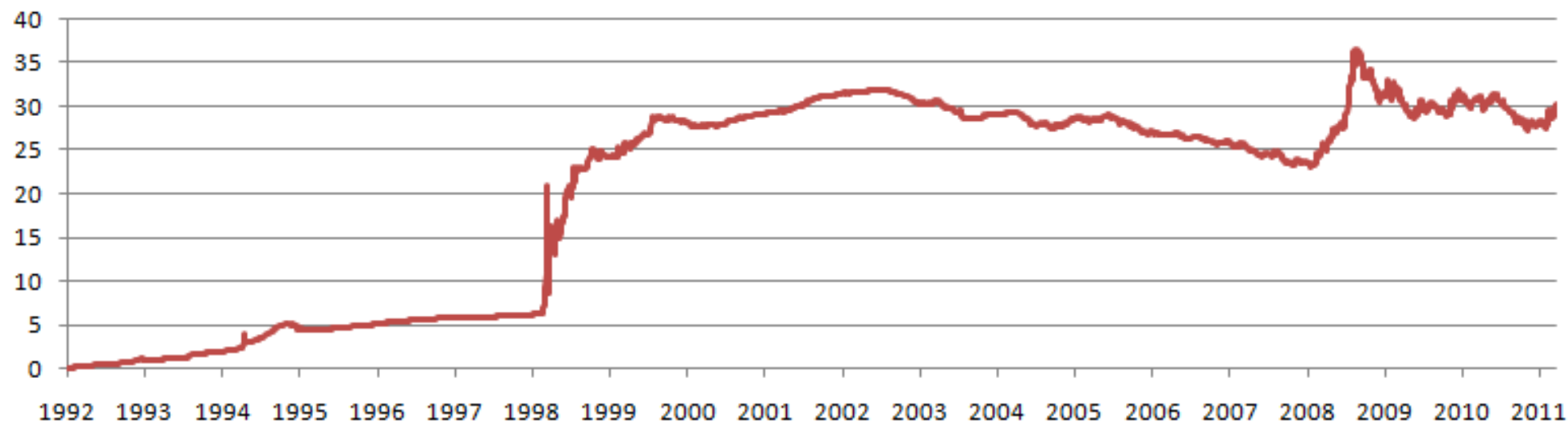
Всякий раз, когда примеры в наборе данных зависят от других точек в наборе данных, данные называются последовательными данными. Типичным примером являются временные ряды, такие как цена акций или данные датчика, где каждая точка представляет собой наблюдение в определенный момент времени.

Ещё один распространённый пример – последовательность слов в естественном человеческом языке.

# ПРИМЕР ВРЕМЕННОГО РЯДА

## Приведенный курс доллара

значения до 1 янв. 1998 разделены на 1000



# ПРИМЕР ПОСЛЕДОВАТЕЛЬНЫХ ДАННЫХ

DETECT LANGUAGE

ENGLISH

FRENCH

**RUSSIAN**

↕

↔

RUSSIAN

**ENGLISH**

SPANISH



↕

Рекуррентные нейронные сети — вид нейронных сетей, где связи между элементами образуют направленную последовательность


×


Recurrent neural networks are a type of neural networks where connections between elements form a directed sequence




☆

118 / 5000

 P<sub>Y</sub> ▾



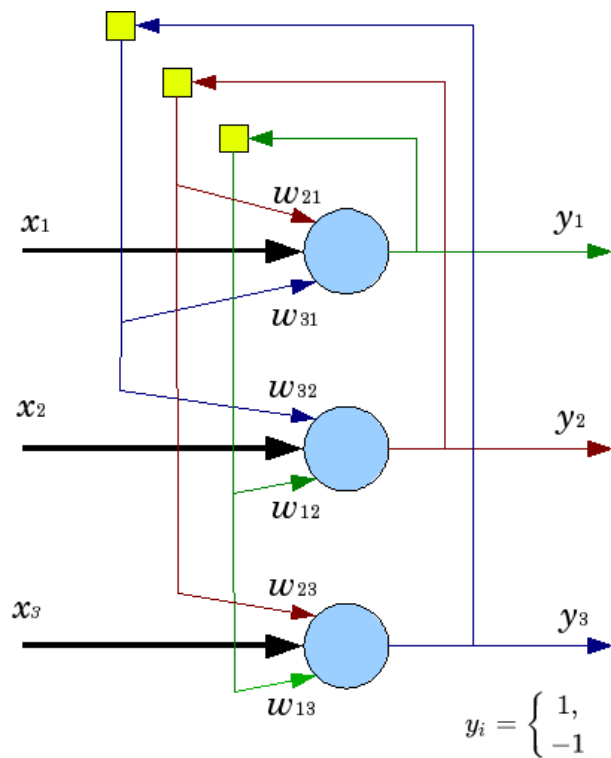
# РЕКУРРЕНТНЫЕ НЕЙРОННЫЕ СЕТИ

## Recurrent Neural Networks (RNN)

рекуррентные сети могут использовать свою внутреннюю память для обработки последовательностей произвольной длины

RNN применимы в таких задачах, где нечто целостное разбито на части, например: распознавание рукописного текста или распознавание речи

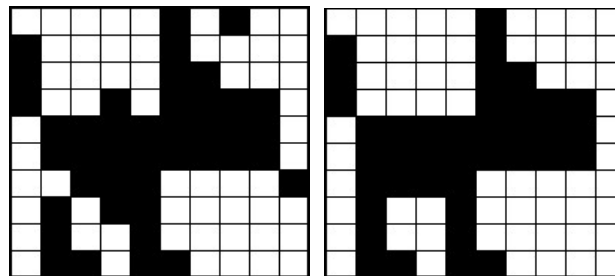
# СЕТЬ ХОПФИЛДА



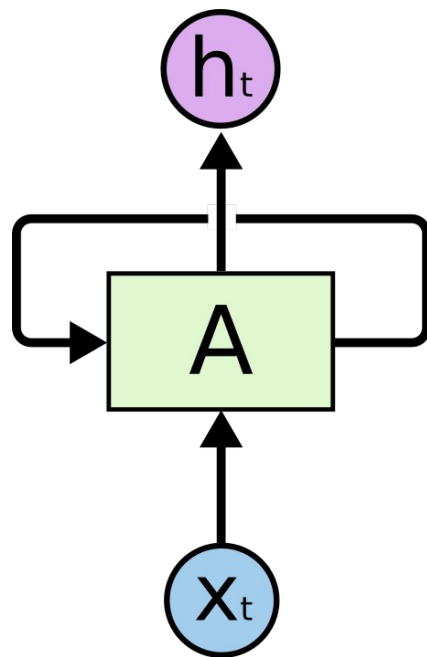
$$X_i = WX_i$$

$$w_{ij} = \frac{1}{N} \sum_{d=1..m} X_{id} X_{jd}$$

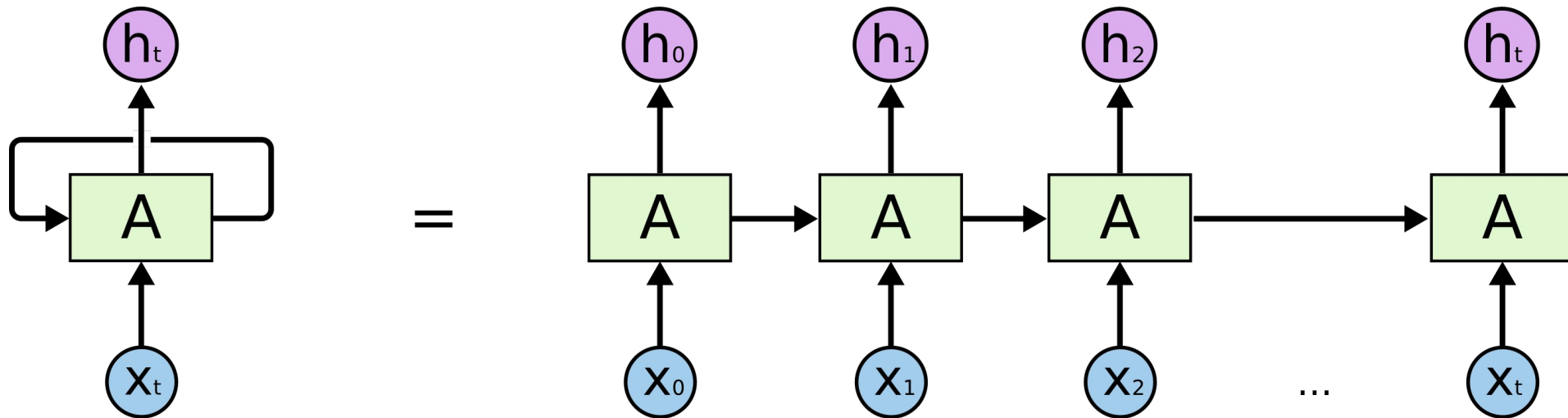
$$W = \frac{1}{N} \sum_i X_i X_i^T$$



# РАЗНОВИДНОСТИ RNN

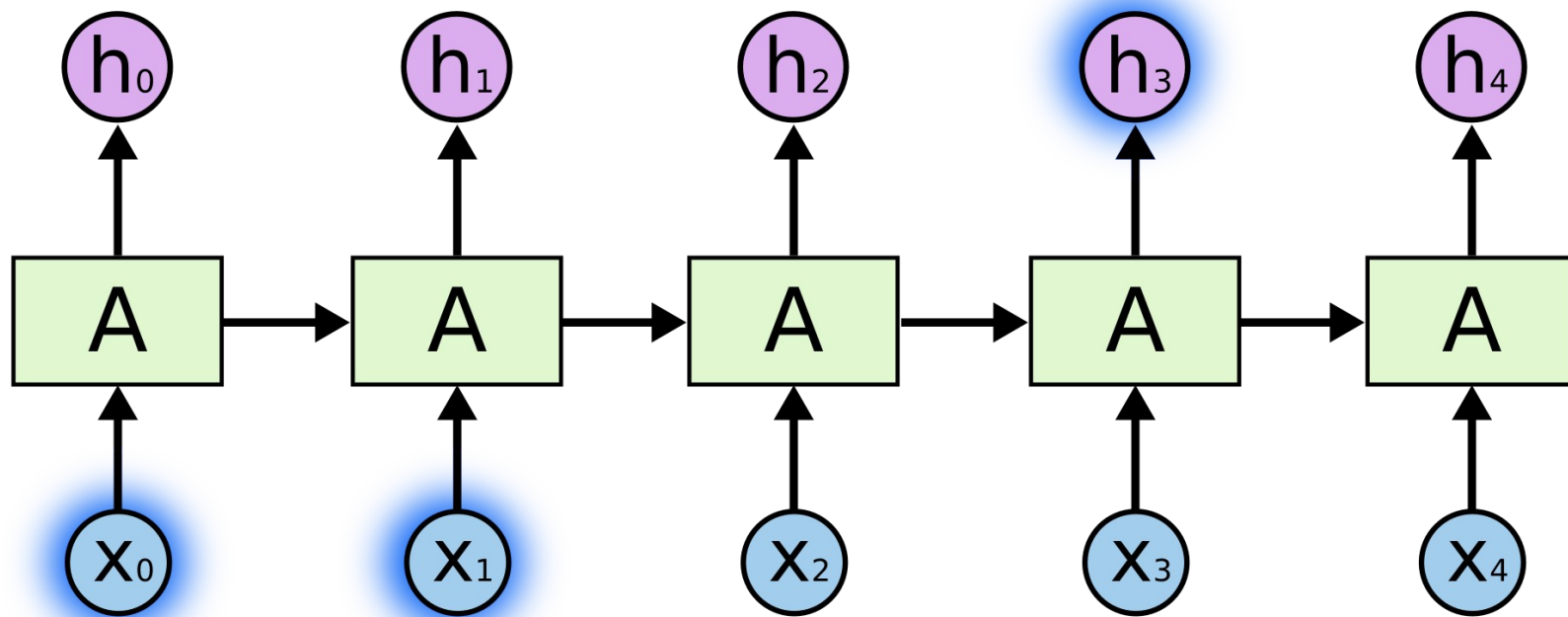


# РАЗНОВИДНОСТИ RNN



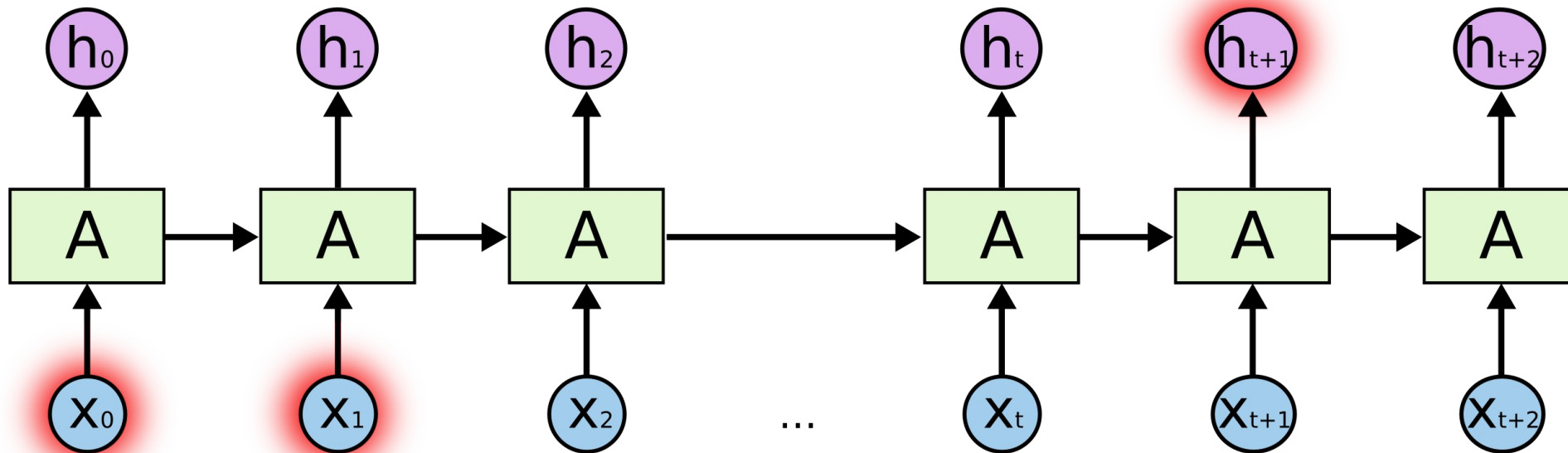


# ИСПОЛЬЗОВАНИЕ СТАРОЙ ИНФОРМАЦИИ



The clouds are in the *sky*

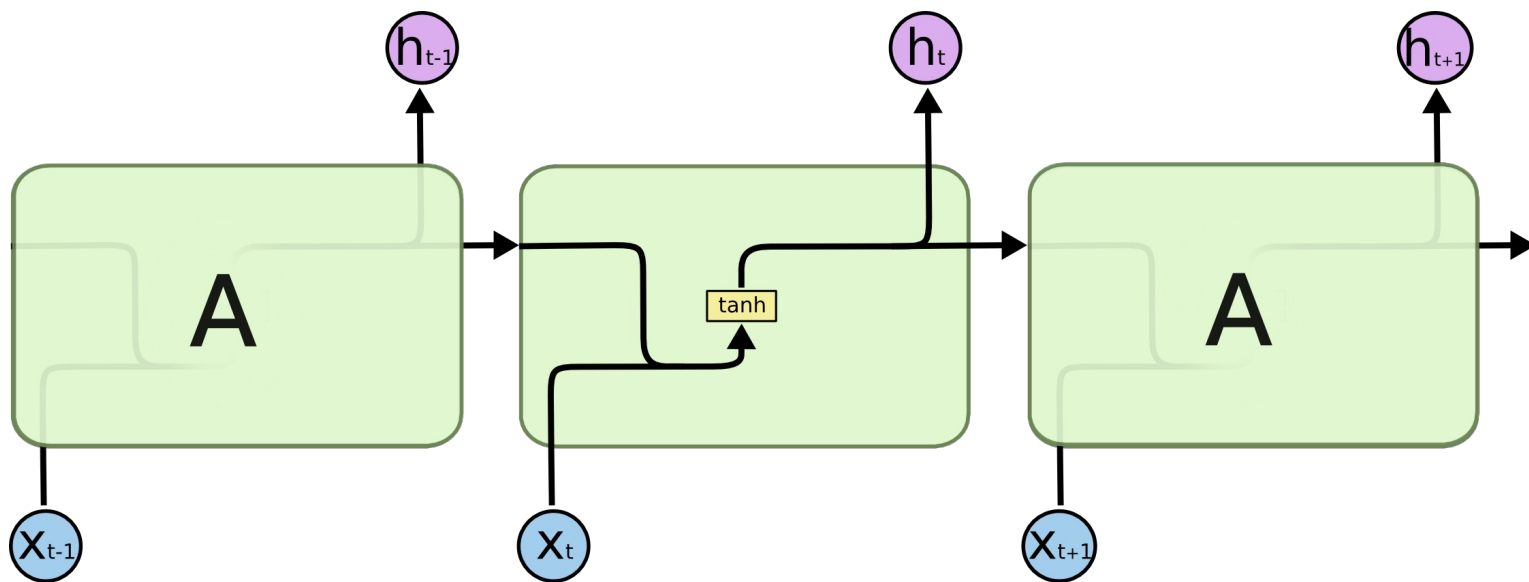
# ИСПОЛЬЗОВАНИЕ СТАРОЙ ИНФОРМАЦИИ



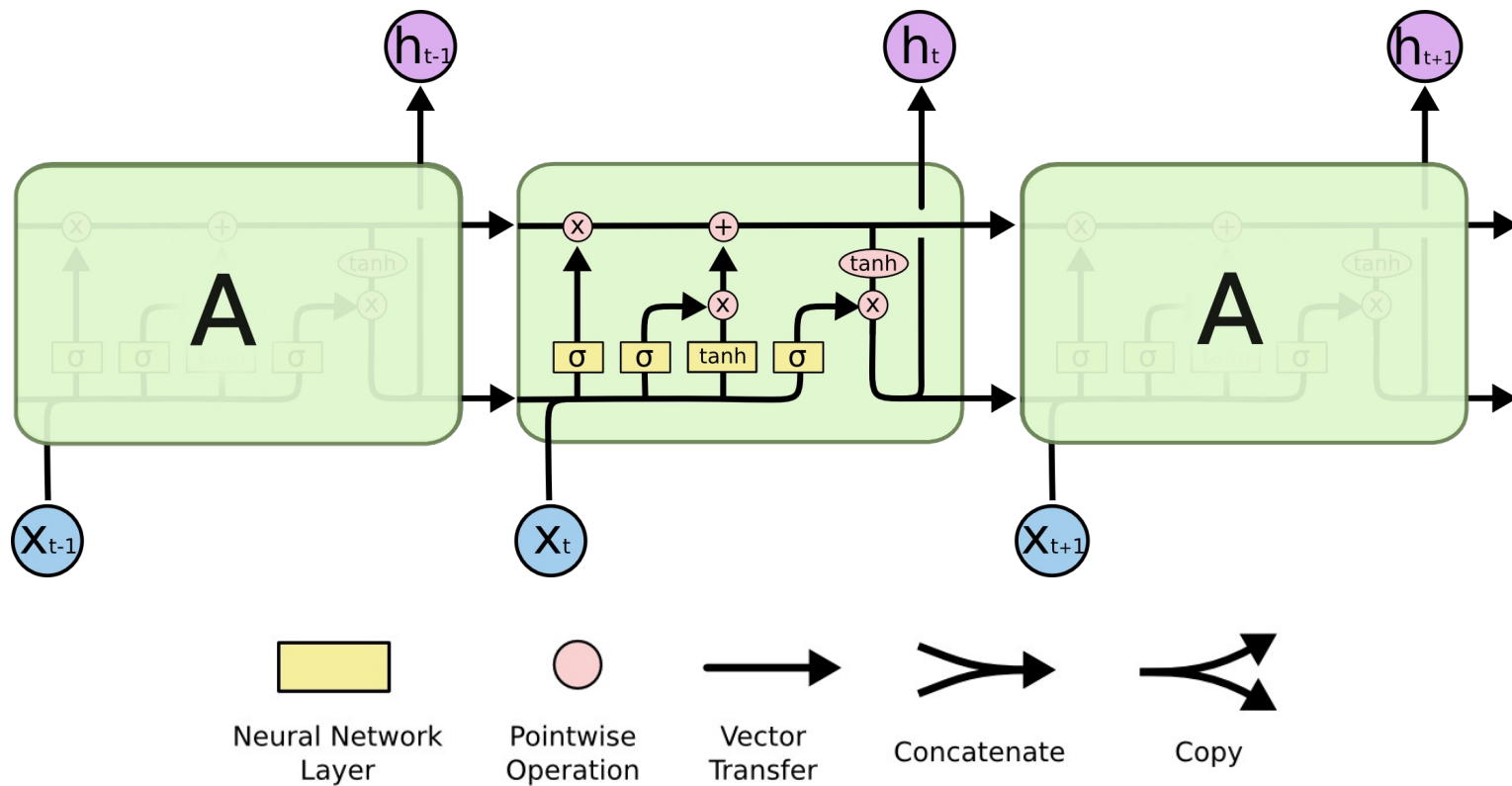
I grew up in France... I speak fluent *French*.

# Long short-term memory (LSTM)

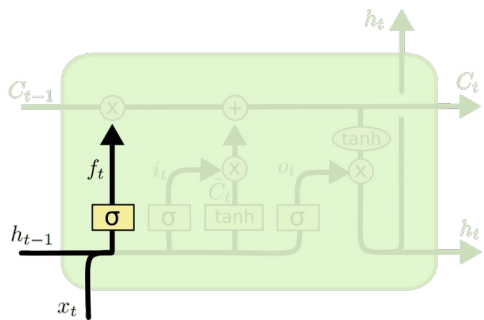
Долгая краткосрочная память



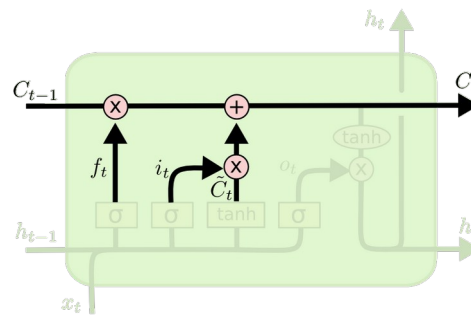
# Более сложная версия LSTM



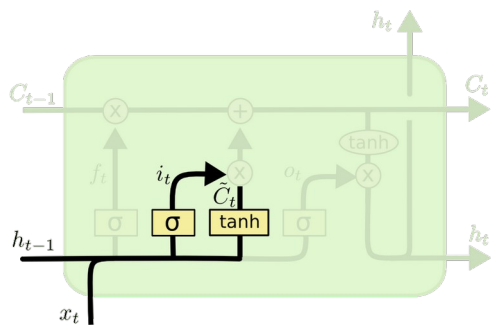
# STEP-BY-STEP LSTM WALK THROUGH



$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$

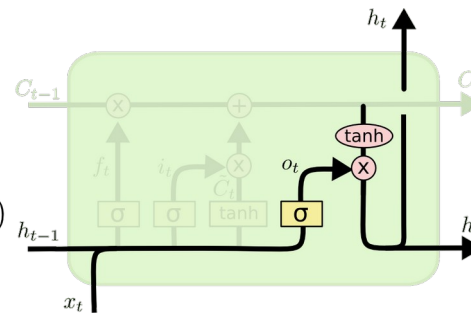


$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$



$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

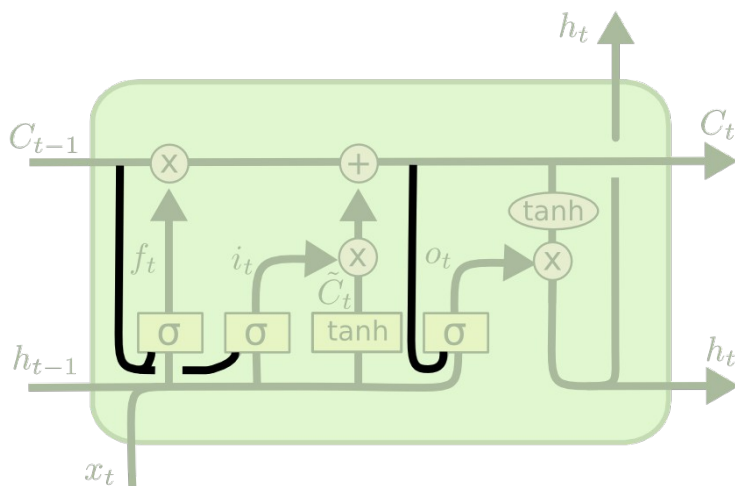


$$o_t = \sigma(W_o [h_{t-1}, x_t] + b_o)$$

$$h_t = o_t * \tanh(C_t)$$

# PEERHOLE CONNECTIONS

Позволяем слоям наблюдать за состоянием ячейки



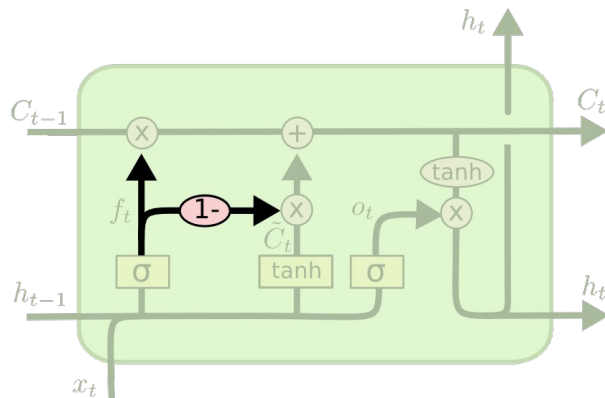
$$f_t = \sigma(W_f \cdot [C_{t-1}, h_{t-1}, x_t] + b_f)$$

$$i_t = \sigma(W_i \cdot [C_{t-1}, h_{t-1}, x_t] + b_i)$$

$$o_t = \sigma(W_o \cdot [C_t, h_{t-1}, x_t] + b_o)$$

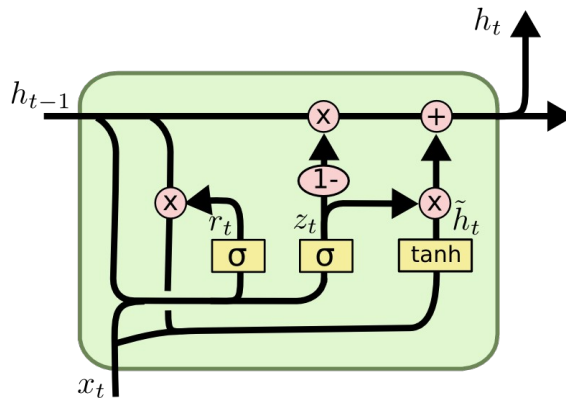
# ДРУГИЕ ВИДЫ LSTM

Вариант забывания



$$C_t = f_t * C_{t-1} + (1 - f_t) * \tilde{C}_t$$

Gated Recurrent Unit (GRU)



$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t])$$

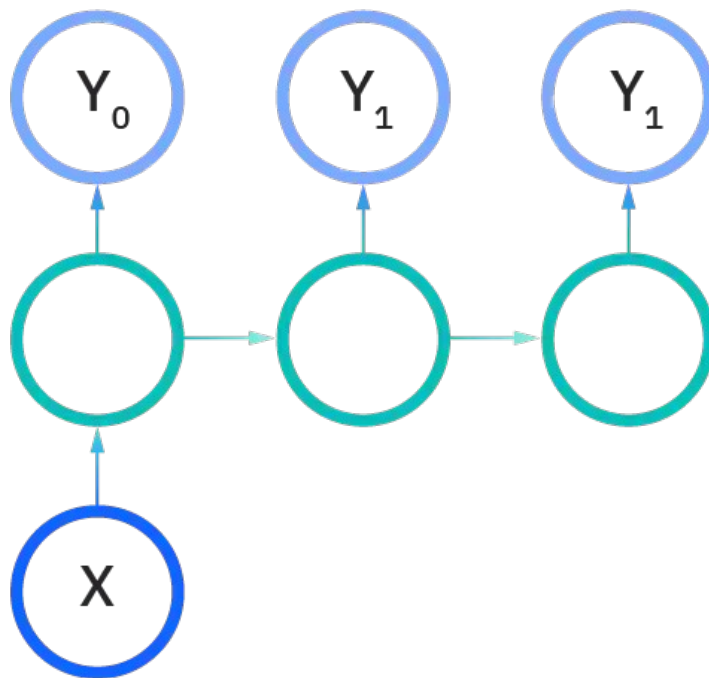
$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t])$$

$$\tilde{h}_t = \tanh(W \cdot [r_t * h_{t-1}, x_t])$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t$$

# ВИДЫ RNN

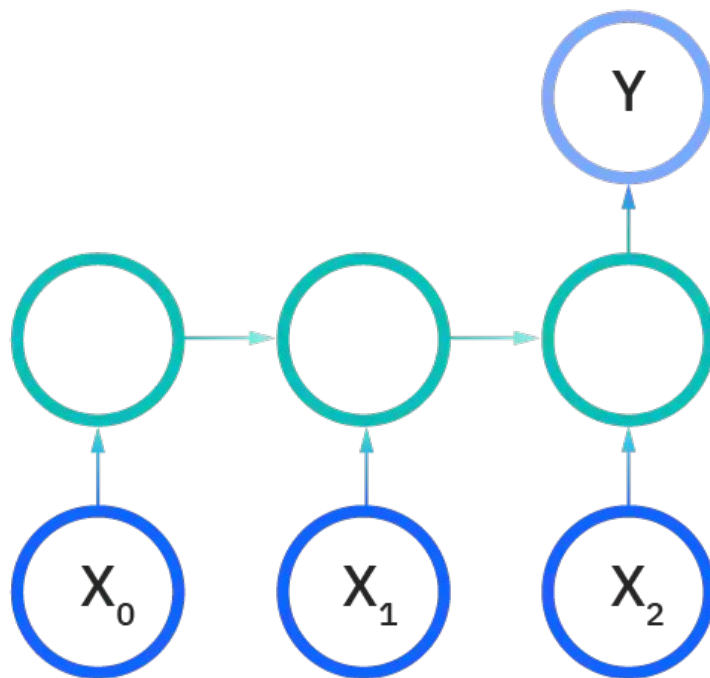
**One-to-many:**





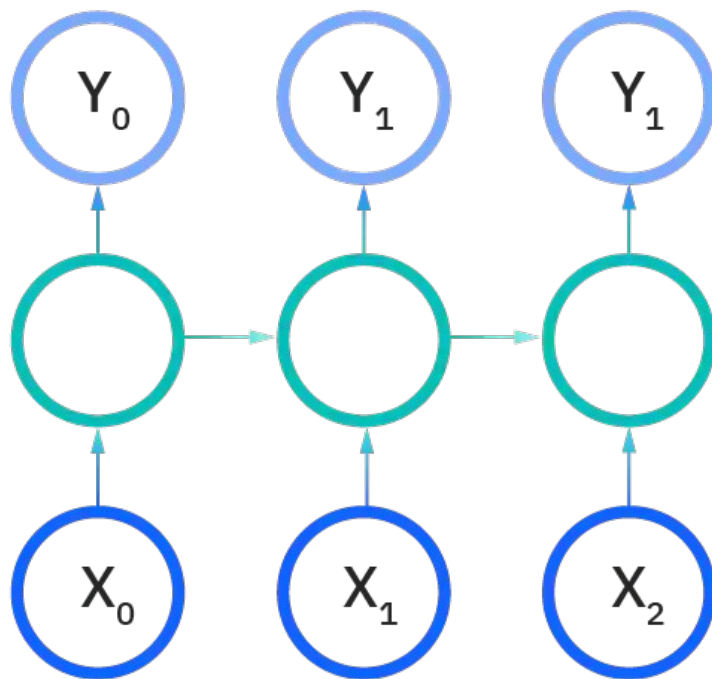
# ВИДЫ RNN

Many-to-one:



# ВИДЫ RNN

Many-to-many:



# ВИДЫ RNN

Many-to-many:

