

Рекуррентные нейронные сети

Курс «Интеллектуальные информационные системы»

Кафедра управления и информатики НИУ «МЭИ»

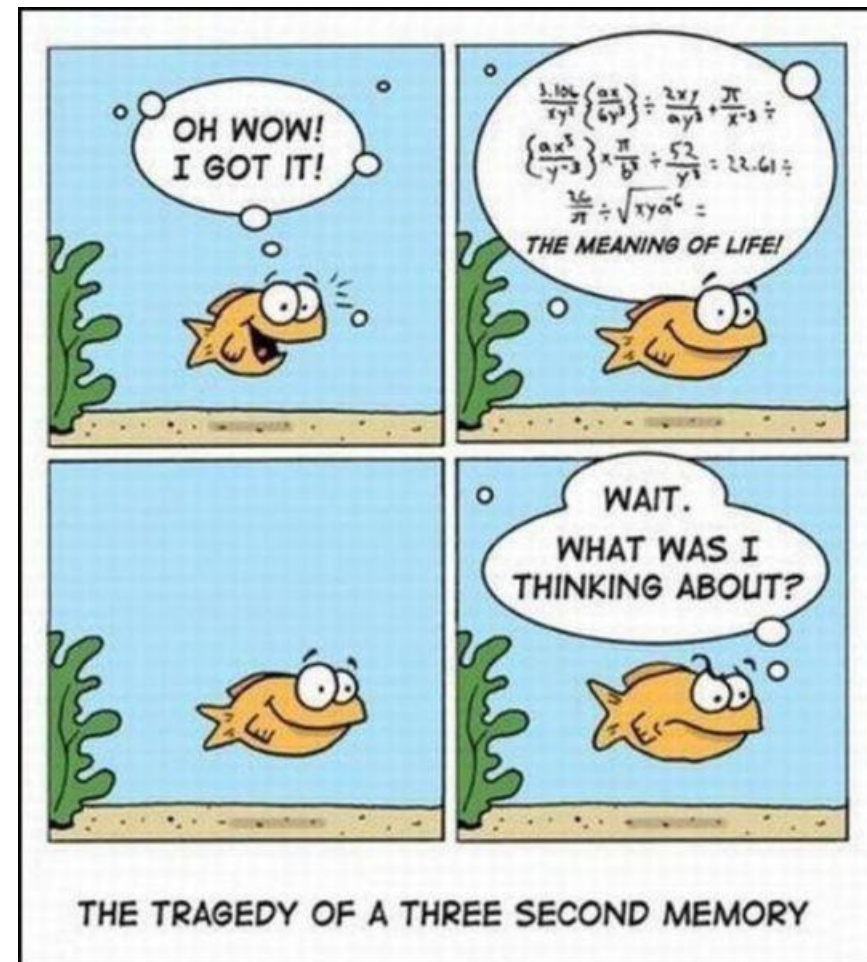
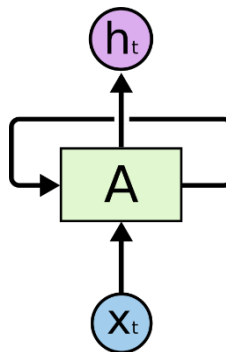
Осень 2018 г.

Рекуррентные нейронные сети*

Люди не начинают думать с чистого листа каждую секунду. Читая этот текст, вы понимаете каждое слово, основываясь на понимании предыдущего слова. Мы не выбрасываем из головы все и не начинаем думать с нуля. Наши мысли обладают постоянством.

Обычные нейросети не обладают памятью, и это их главный недостаток в отличие от рекуррентных нейронных сетей (Recurrent Neural Networks, RNN, PHC).

Это сети, содержащие обратные связи и позволяющие сохранять информацию.



*по мотивам статей:

<https://iamtrask.github.io/2015/11/15/anyone-can-code-lstm/>
<https://habr.com/company/wunderfund/blog/331310/>

Скрытый слой RNN

Классические нейронные сети имеют скрытые слои, и состояние скрытого слоя основывается только на входных данных.

input -> hidden -> output

Наличие памяти меняет эту структуру – теперь состояние скрытого слоя зависит от входных данных на текущем шаге и от состояния того же **скрытого слоя на предыдущем шаге**:

(input + prev_hidden) -> hidden -> output

Скрытый слой RNN (2)

Почему на каждом предыдущем шаге мы смотрим именно скрытый слой а не вход:

(input + prev_input) -> hidden -> output

Рассмотрим 4 шага RNN для обоих случаев:

а) смотрим скрытый слой:

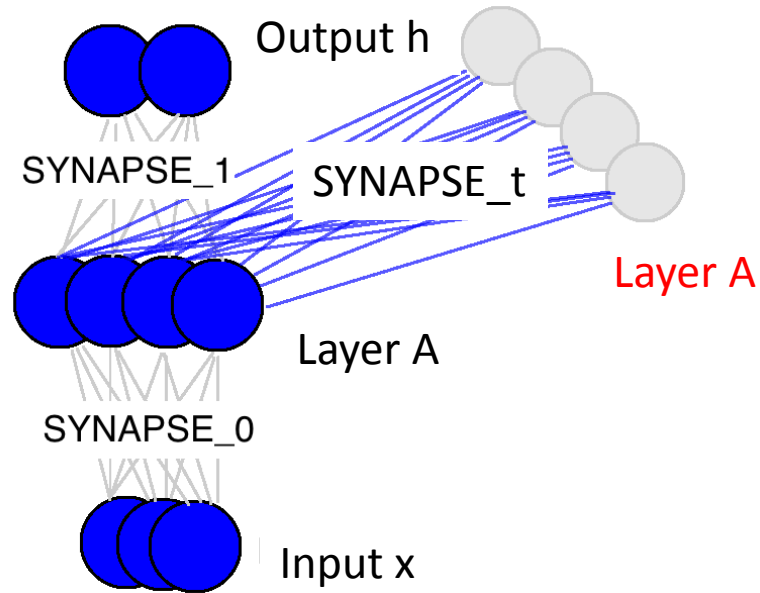
(input + empty_hidden) -> hidden -> output
(input + prev_hidden) -> hidden -> output
(input + prev_hidden) -> hidden -> output
(input + prev_hidden) -> hidden -> output

б) смотрим вход:

(input + empty_input) -> hidden -> output
(input + prev_input) -> hidden -> output
(input + prev_input) -> hidden -> output
(input + prev_input) -> hidden -> output

Память рекуррентной нейросети

Таким образом, «память» формируется за счет комбинации входных данных и скрытого слоя на предыдущем шаге. Такая связь реализуется с помощью весов, которые присваиваются обратной связи.

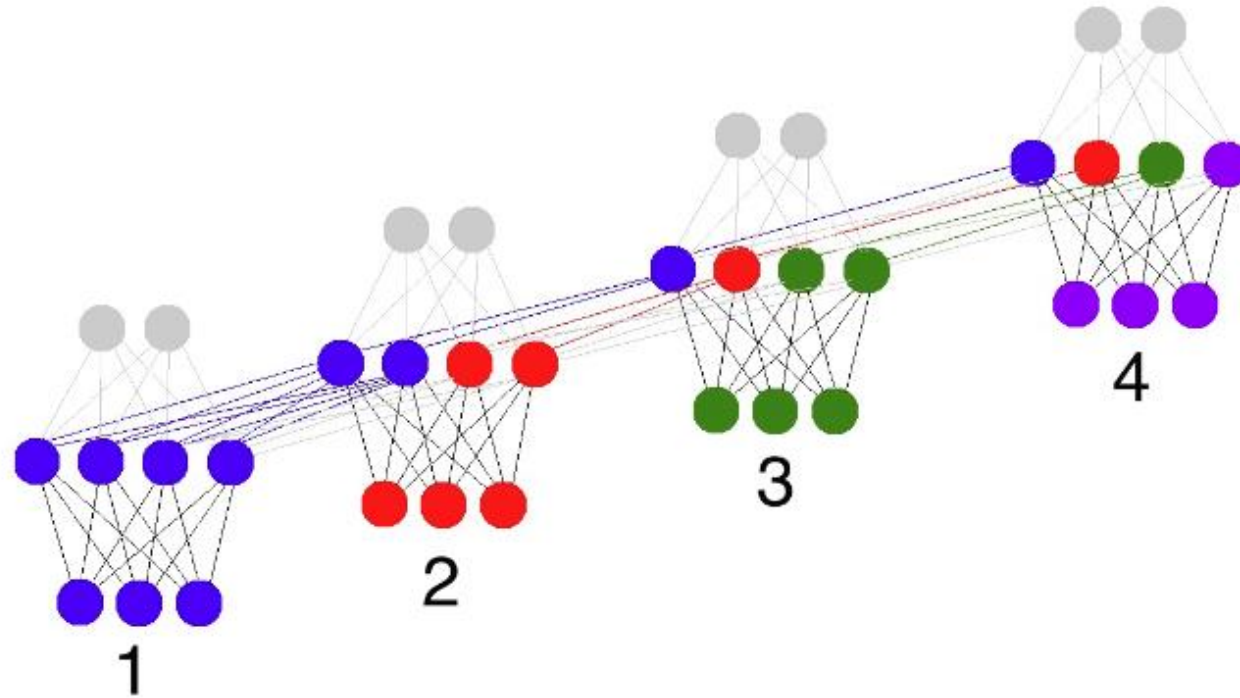


SYNAPSE_0 – матрица весов передачи сигнала от входа к скрытому слою

SYNAPSE_1 – матрица весов передачи сигнала от скрытого слоя к выходу

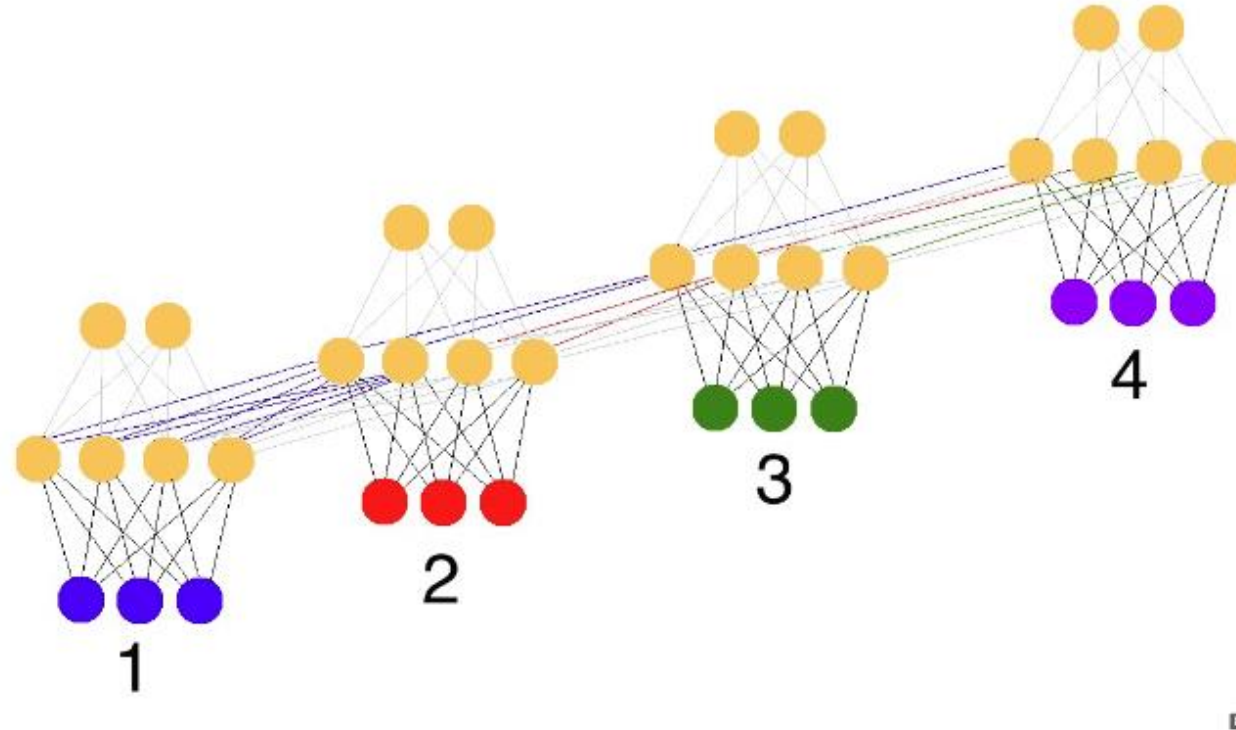
SYNAPSE_t – матрица весов передачи сигнала от скрытого слоя A на текущем шаге к тому же скрытому слою A на следующем шаге.

Память рекуррентной нейросети (2)



- Объем памяти зависит от размера скрытого слоя.
- Выход – не является чистой функцией от входа, входной сигнал изменяет «память», а выход будет зависеть от того, что в этой «памяти» находится.
- Если на 2, 3 и 4 шагах не было бы входных сигналов, состояние скрытого слоя все равно бы менялось на каждом шаге.

Обучение рекуррентной нейросети

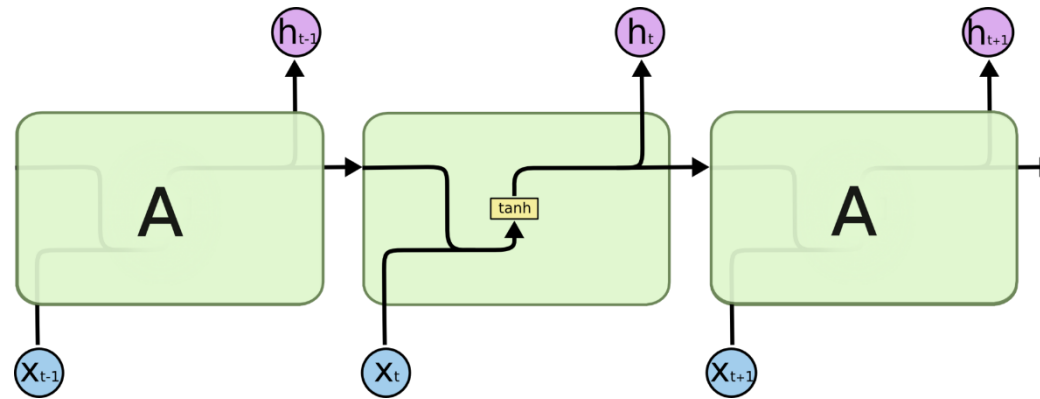


- Объем памяти зависит от размера скрытого слоя.
- Выход – не является чистой функцией от входа, входной сигнал изменяет «память», а выход будет зависеть от того, что в этой «памяти» находится.
- Если на 2, 3 и 4 шагах не было бы входных сигналов, состояние скрытого слоя все равно бы менялось на каждом шаге.

LSTM-сети

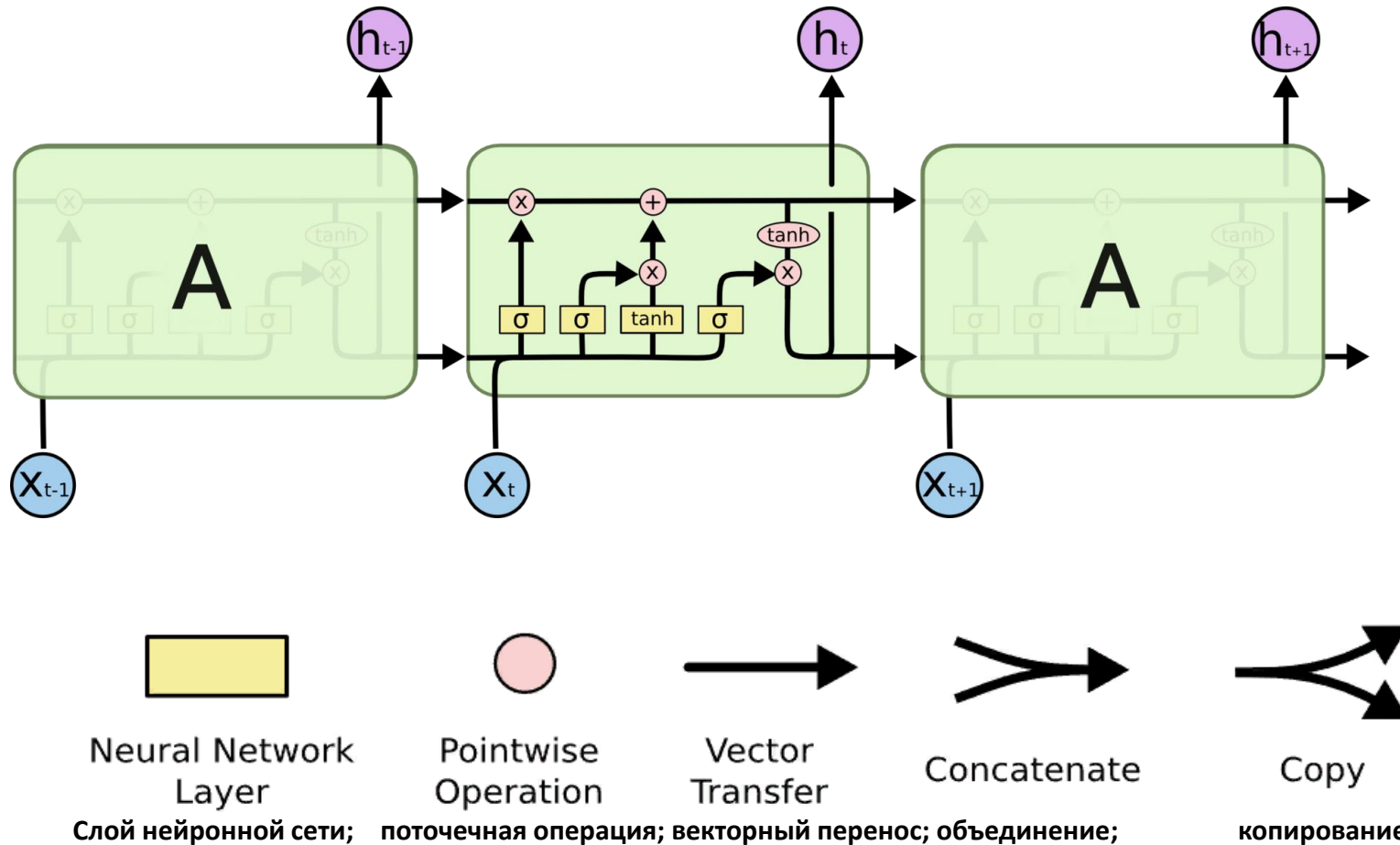
LSTM - Long short-term memory, долгая краткосрочная память - особая разновидность архитектуры рекуррентных нейронных сетей, способная к обучению долговременным зависимостям. Они были представлены Зеппом Хохрайтер и Юргеном Шмидхубером (Jürgen Schmidhuber) в 1997 году, а затем усовершенствованы и популярно изложены в работах многих других исследователей. Особенность LSTM-сетей состоит в запоминании информации в «долгой» памяти.

Повторяющийся модуль в стандартной RNN состоит из одного слоя:

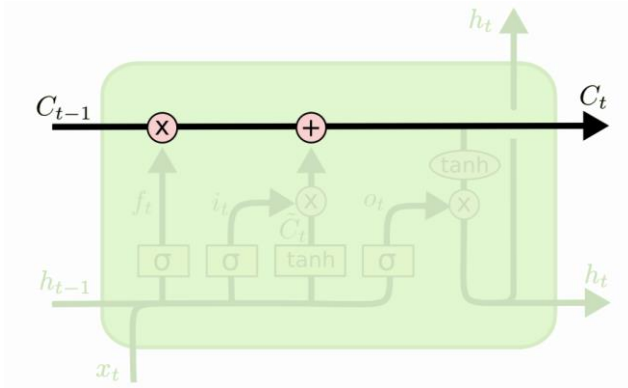


LSTM-сету

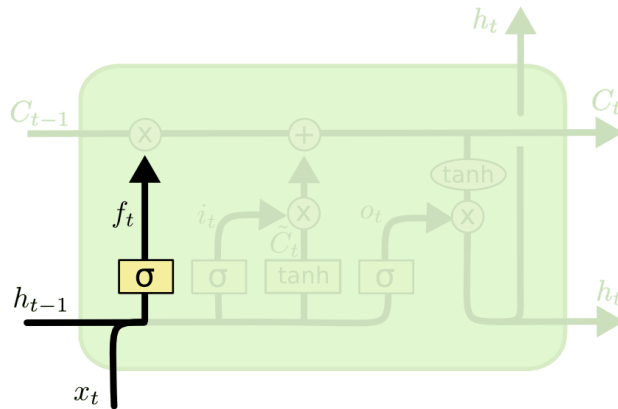
Структура LSTM также напоминает цепочку, но модули выглядят иначе. Вместо одного слоя нейронной сети они содержат целых четыре, и эти слои взаимодействуют особым образом.



LSTM-сети



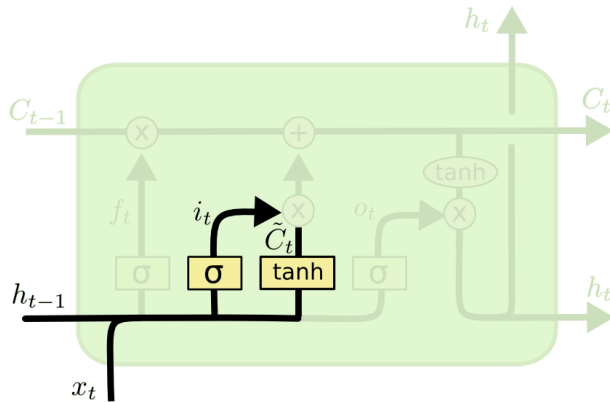
Ключевой компонент LSTM – это состояние ячейки C (cell state) – горизонтальная линия, проходящая по верхней части схемы. Состояние ячейки напоминает конвейерную ленту. Она проходит напрямую через всю цепочку, участвуя лишь в нескольких линейных преобразованиях. Этот процесс регулируется «фильтрами» (gates).



Слой фильтра забывания (forget gate layer) – определяет, какую информацию можно выбросить из состояния ячейки. Возвращает число $[0;1]$, где 0 – полностью забыть, 1 – полностью сохранить

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

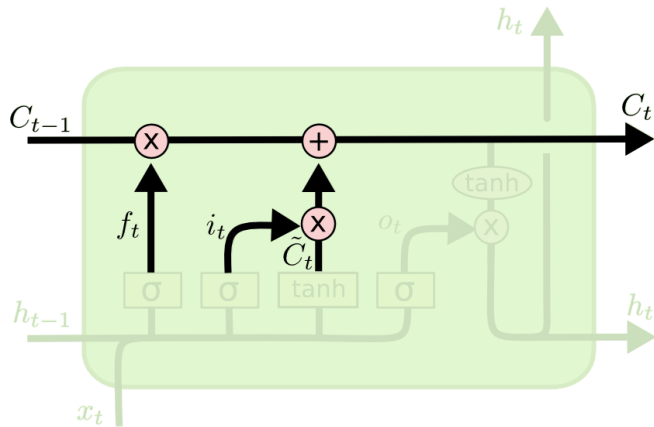
LSTM-сети



Следующий шаг – решить, какая новая информация будет храниться в состоянии ячейки. Этот этап состоит из двух частей. Сначала сигмоидальный слой под названием “слой входного фильтра” (input layer gate) определяет, какие значения следует обновить. Затем tanh-слой строит вектор новых значений-кандидатов \tilde{C}_t , которые можно добавить в состояние ячейки.

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

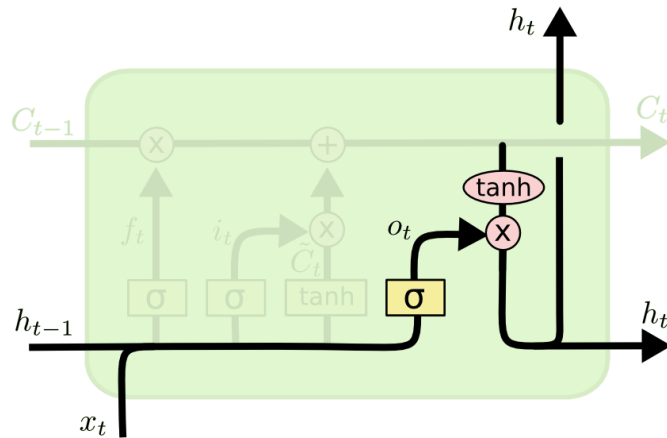
$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$



Меняем состояние ячейки в соответствии с предыдущими слоями forget layer, input layer и tanh layer

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

LSTM-сети



Решаем, какую информацию хотим получать на выходе h . Выходные данные будут основаны на нашем состоянии ячейки, к ним будут применены некоторые фильтры. Сначала мы применяем сигмоидальный слой, который решает, какую информацию от входа и предыдущего шага мы будем выводить. Затем значения состояния ячейки проходят через \tanh -слой, чтобы получить на выходе значения из диапазона от -1 до 1, и перемножаются с выходными значениями сигмоидального слоя, что позволяет выводить только требуемую информацию.

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

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

Пример обучения LSTM

Поставлена задача обучения сети на романе Л.Н. Толстого «Война и мир» и дальнейшего создания сетью осмысленного текста. Результаты после разного количества итераций:

100: tyntd-iafhatawiaoighrdemot lytdws e ,tfti, astai f ogoh eoase rrranbyne 'nhthnee e plia
tklrgd t o idoe ns,smtt h ne etie h,hregtrs nigtike,aoaenns lng

300: "Tmont thithey" fomesscerliund Keushey. Thom here sheulke, anmerenith ol sivh l
lalterthend Bleipile shuwyl on aseterlome coaniogennc Phe lism thond hon at.
MeiDimorotion in ther thize."

500: we counter. He stutn co des. His stanted out one ofler that concossions and was to
gearang reay Jotrets and with fre colt ofp paith thin wall. Which das stimn

<http://karpathy.github.io/2015/05/21/rnn-effectiveness/>

Пример обучения LSTM

700: Aftair fall unsuch that the hall for Prince Velzonski's that me of her hearly, and behs to so arwage fiving were to it beloge, pavu say falling misfort how, and Gogition is so overelical and ofter.

1200: "Kite vouch!" he repeated by her door. "But I would be done and quarts, feeling, then, son is people...."

2000: "Why do what that day," replied Natasha, and wishing to himself the fact the princess, Princess Mary was easier, fed in had oftened him. Pierre aking his soul came to the packs and drove up his father-in-law women.

Сначала модель выявляет общую структуру пространства слов, затем активно начинает изучать слова, начиная с коротких к более длинным. На более поздних этапах начинают выявляться зависимости появления слов и вырисовываться осмысленные предложения