



DevCon School

Технологии будущего

Рекуррентные нейронные сети и анализ последовательностей

Михаил Бурцев,
iPavlov.ai, МФТИ

Теория RNN

Рекуррентные нейросети
(RNN)

Ячейка нейросетевой
памяти
(LSTM)

RNN в приложениях

Архитектура системы
машинного перевода

Google Neural Machine
Translation

LSTM демо

Разбор примера
имплементации LSTM с
использованием библиотеки
TensorFlow.

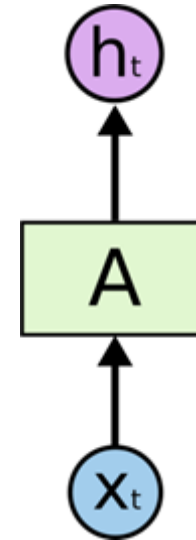
Теория рекуррентных нейросетей

#msdevcon

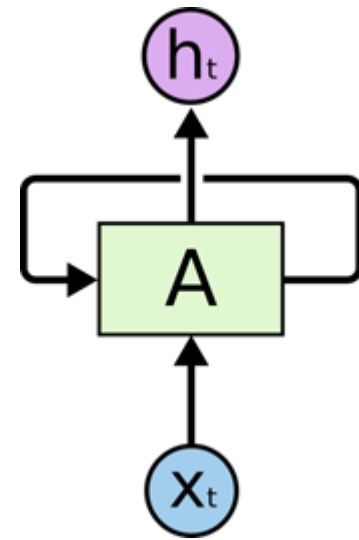
В поисках памяти

- «Классические» нейронные сети — сети прямого распространения (feedforward network) не обладают памятью.
- «Активации» проходят всю сеть насквозь.
- Нейросети «с памятью» — рекуррентные сети.
- За счет наличия обратных связей «активации» циркулируют в сети.

сеть прямого распространения

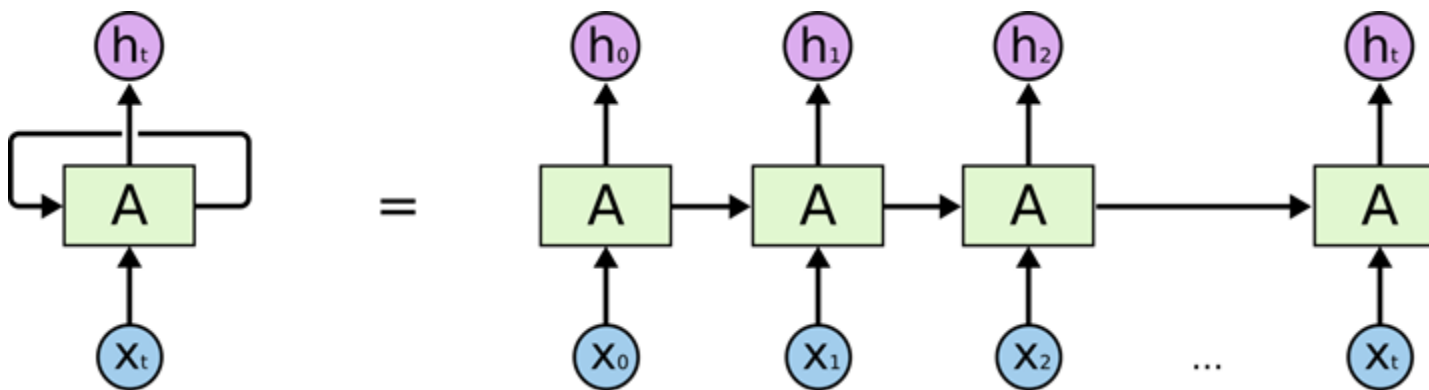


рекуррентная сеть



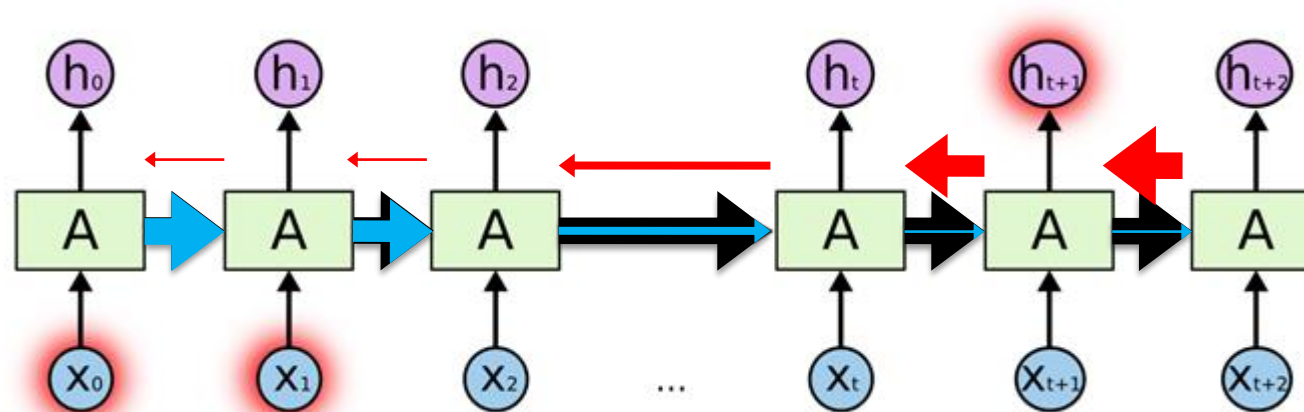
Из глубины времен

- Рекуррентную нейронную сеть на шаге t можно представить в виде многослойной сети глубины t
- Эта операция называется «разворачивание» - unrolling



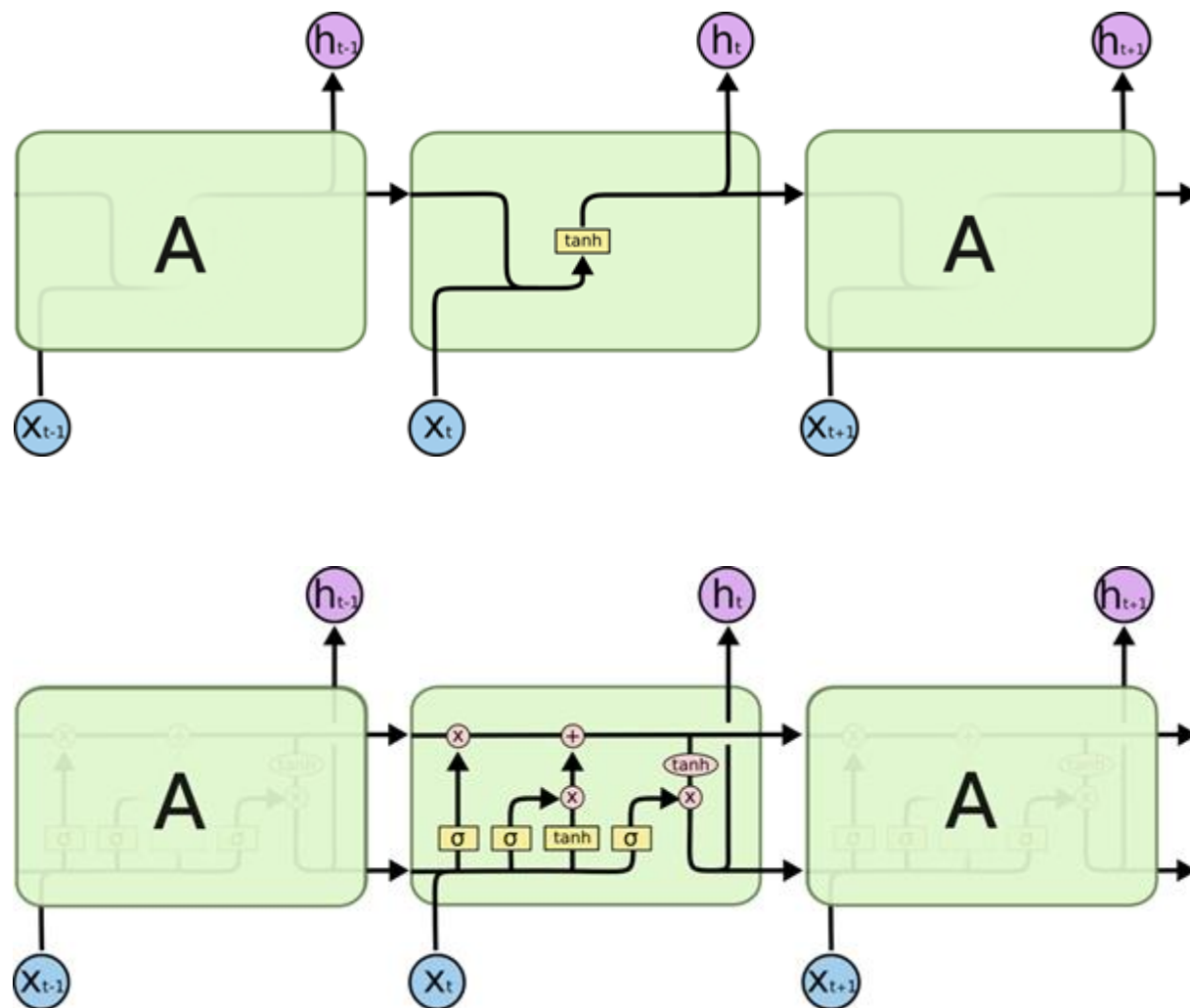
Обучение рекуррентной сети

- Сеть «разворачивается» во времени
- Обучение рекуррентной сети производится стандартным методом обратного распространения ошибки
- Проблема затухания градиента (vanishing gradient problem)



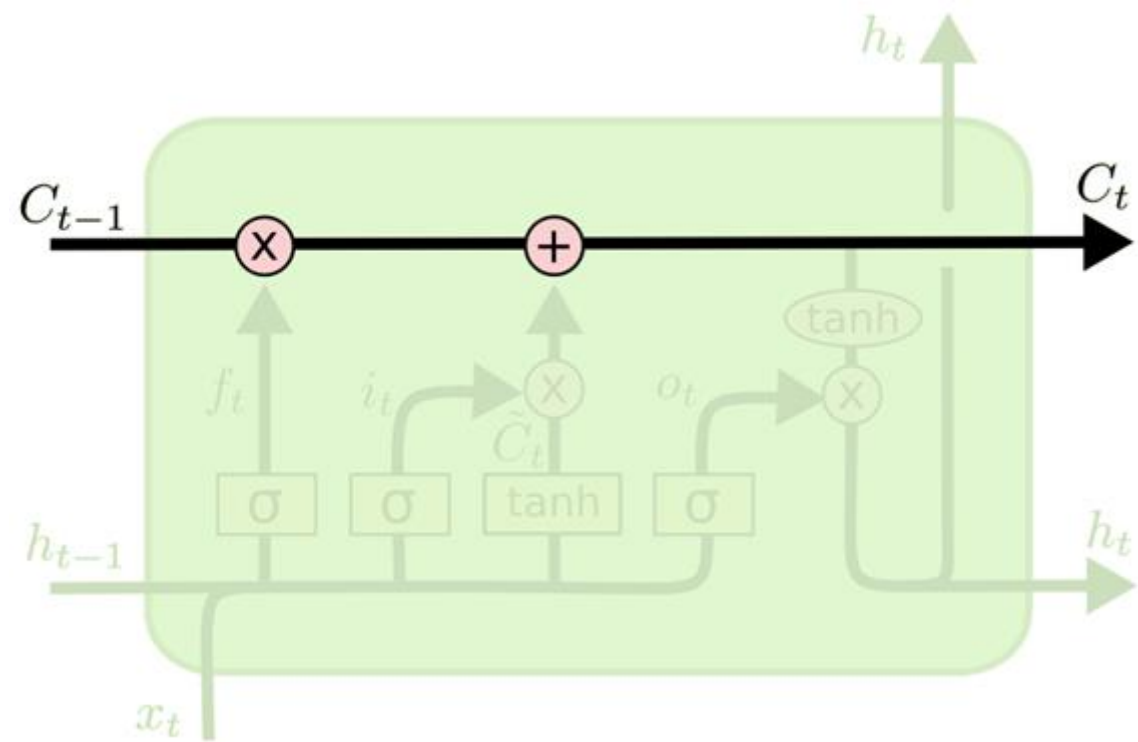
LSTM

- «Однослойная» рекуррентная сеть
- **LSTM** «Ячейка памяти», управляемая затворами



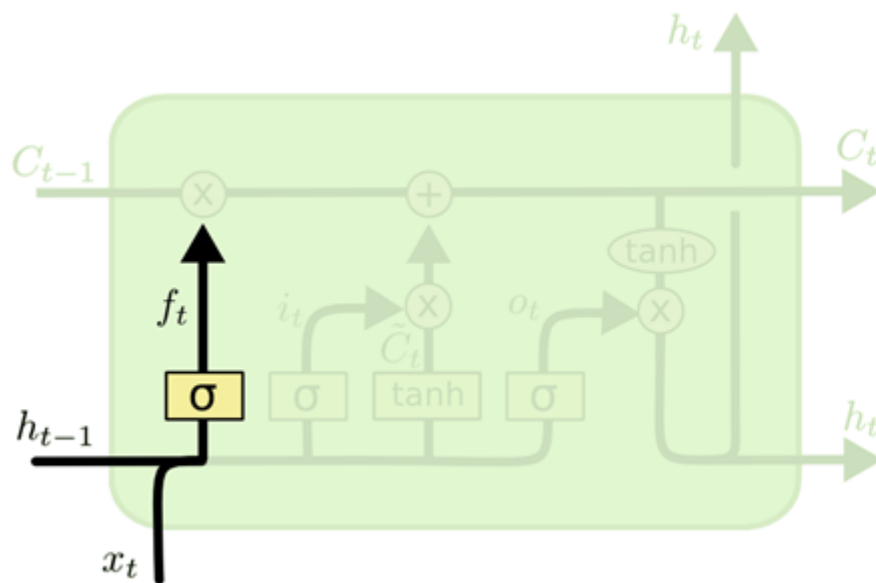
LSTM

C – значение хранимое «ячейкой памяти».



LSTM

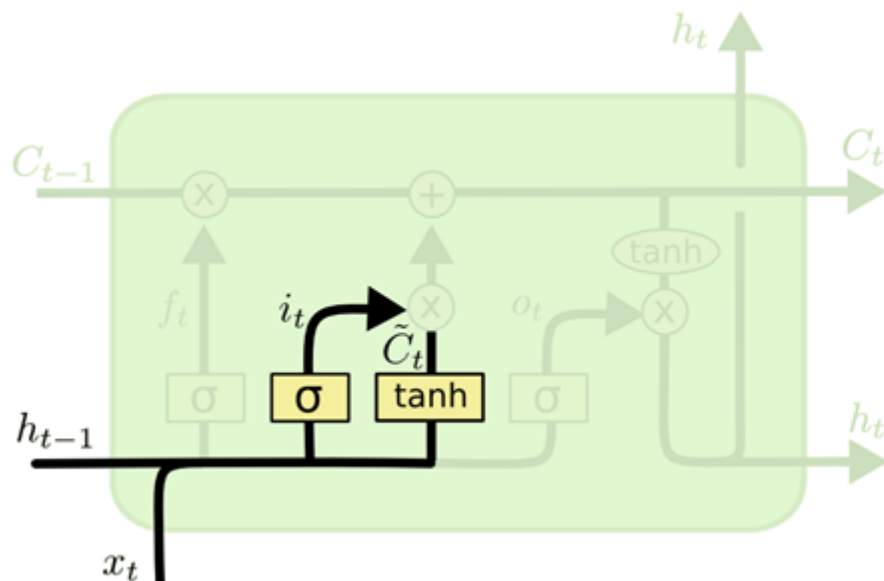
Текущий вход x_t и активации скрытого слоя на предыдущем такте h_{t-1} подаются на вход нейросети «забвения» f_t выход которой определяет затухание значения ячейки памяти C на текущем такте.



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

LSTM

1. Текущий вход x_t и активации скрытого слоя на предыдущем такте h_{t-1} подаются на вход нейросети «гипотезы» выход которой определяет текущее обновление значения ячейки памяти C .
2. Текущий вход x_t и активации скрытого слоя на предыдущем такте h_{t-1} подаются на вход нейросети «обновления» выход которой управляет затвором для обновленного значения ячейки памяти C .

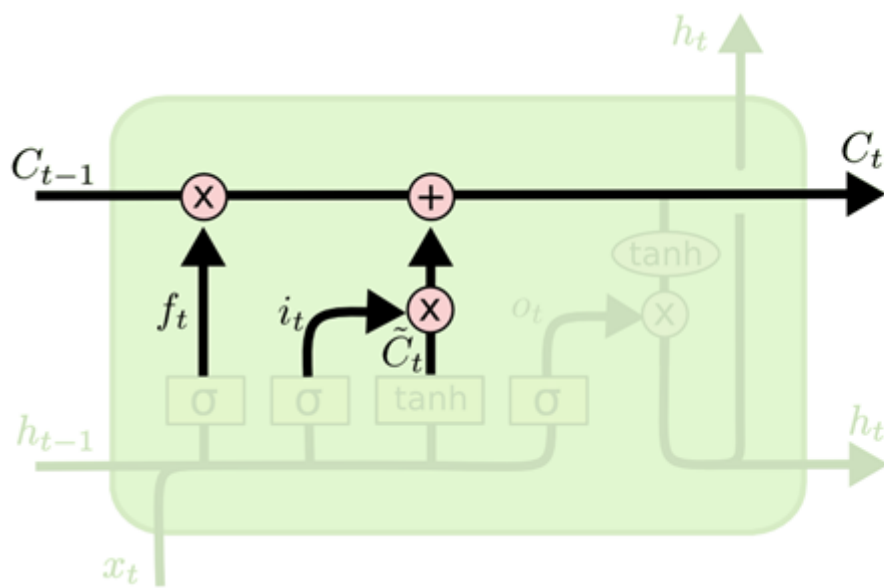


$$i_t = \sigma (W_i \cdot [h_{t-1}, x_t] + b_i)$$

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

LSTM

1. Значение ячейки памяти C с предыдущего такта умножается на коэффициент «забвения».
2. Результат предыдущего шага складывается с новым значением, прошедшим через затвор «обновления».

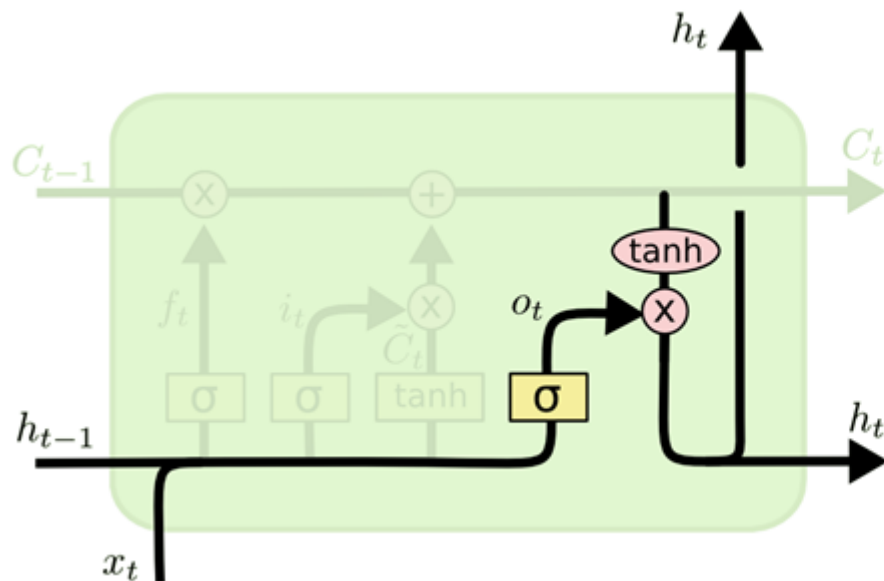


$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

LSTM

Текущий вход x_t и активации скрытого слоя на предыдущем такте h_{t-1} подаются на вход нейросети «выходного затвора» выход которой определяет значение выхода элемента.

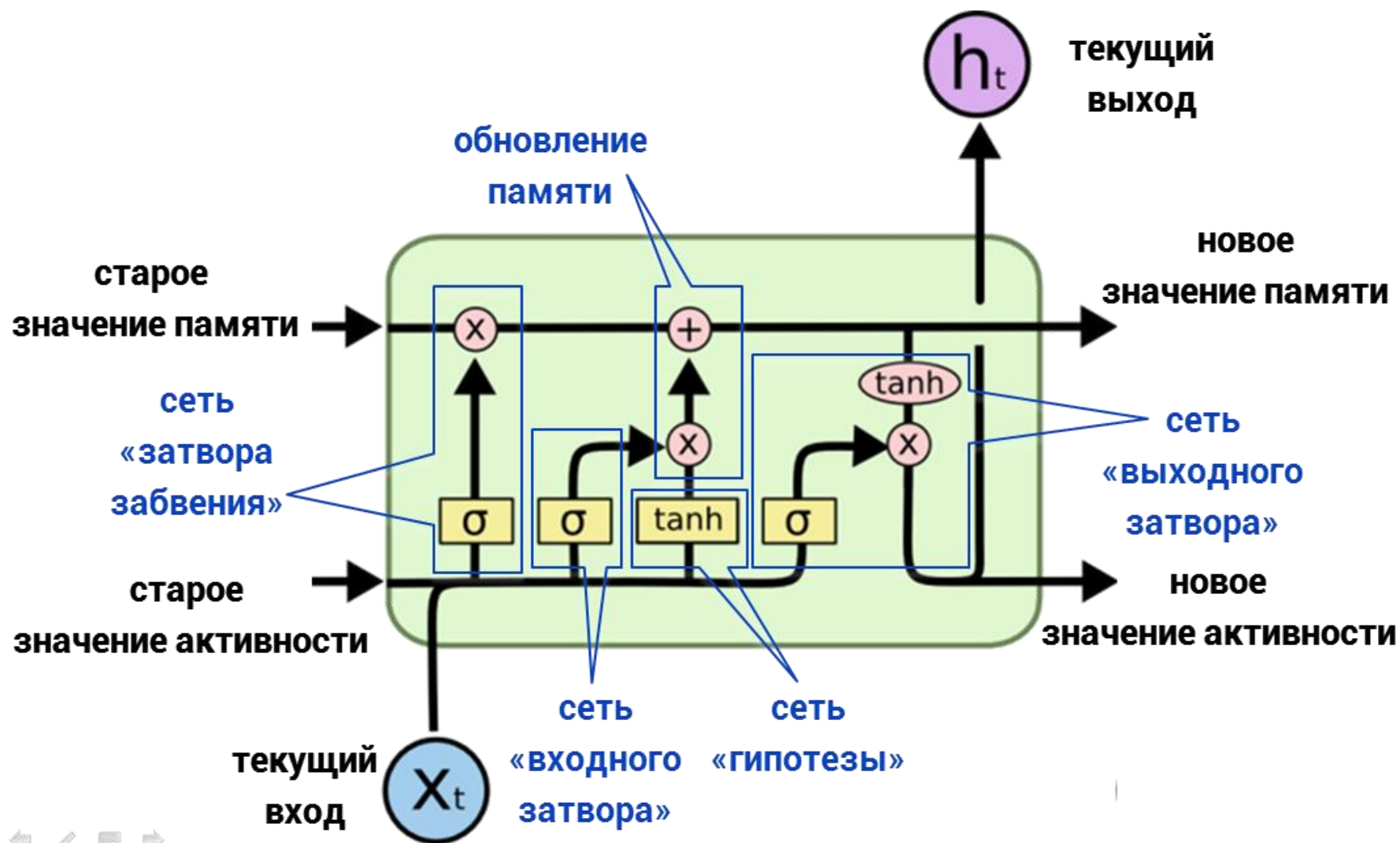
Значение выхода элемента не обязательно совпадает со значением хранимым в памяти!



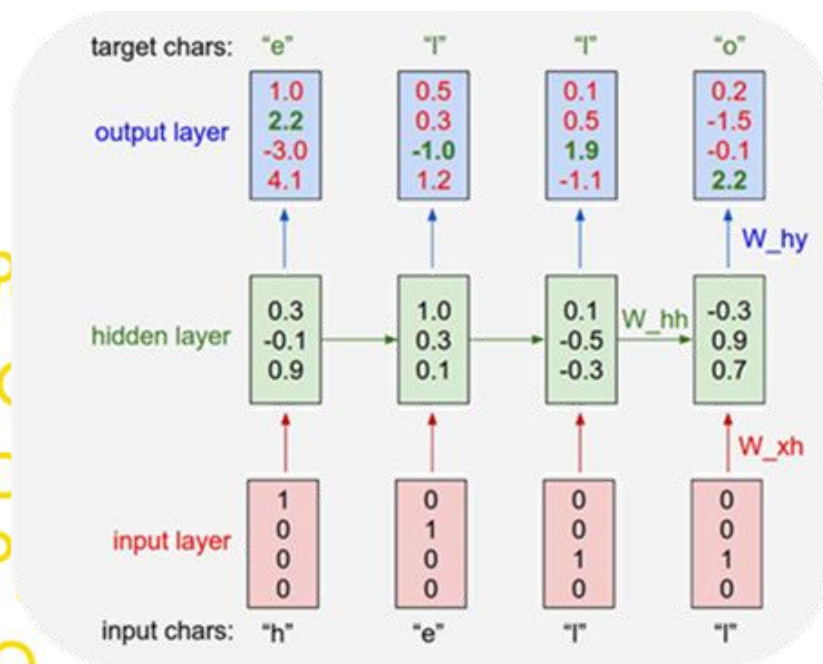
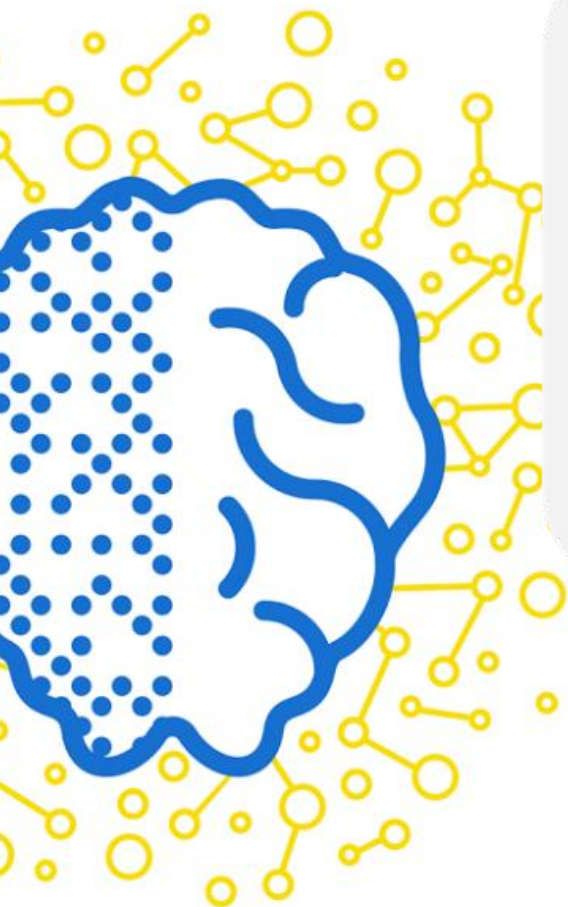
$$o_t = \sigma (W_o [h_{t-1}, x_t] + b_o)$$

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

Повторение – мать ученья



Пример задачи



Человек: Ты глупый

Машина: Как насчет «привет».

Вопрос: Какой чудесный вечер. Не правда ли?

Ответ: Ты действительно думаешь, что я буду помнить об этом?

Человек: Ты дура.

Машина: так почему тебе не все равно?

Человек: Потому что мне с тобой жить!

Машина: Ты совсем не такой!

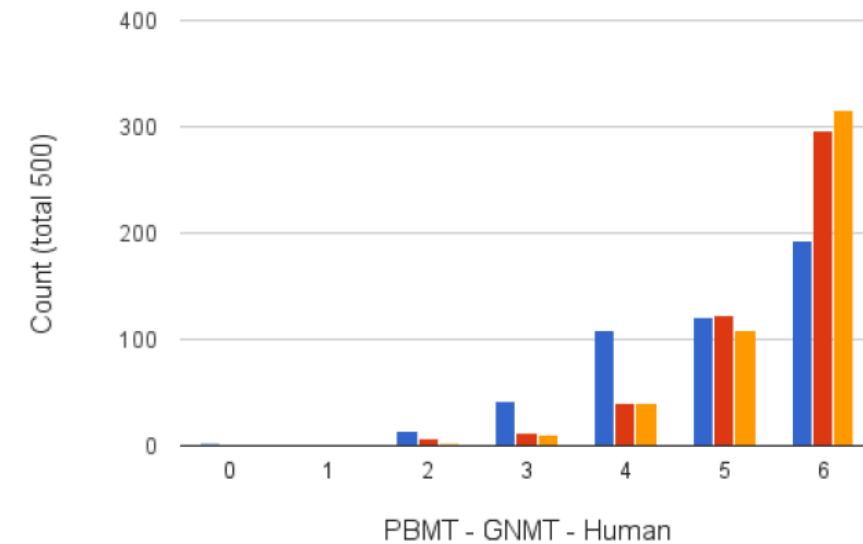
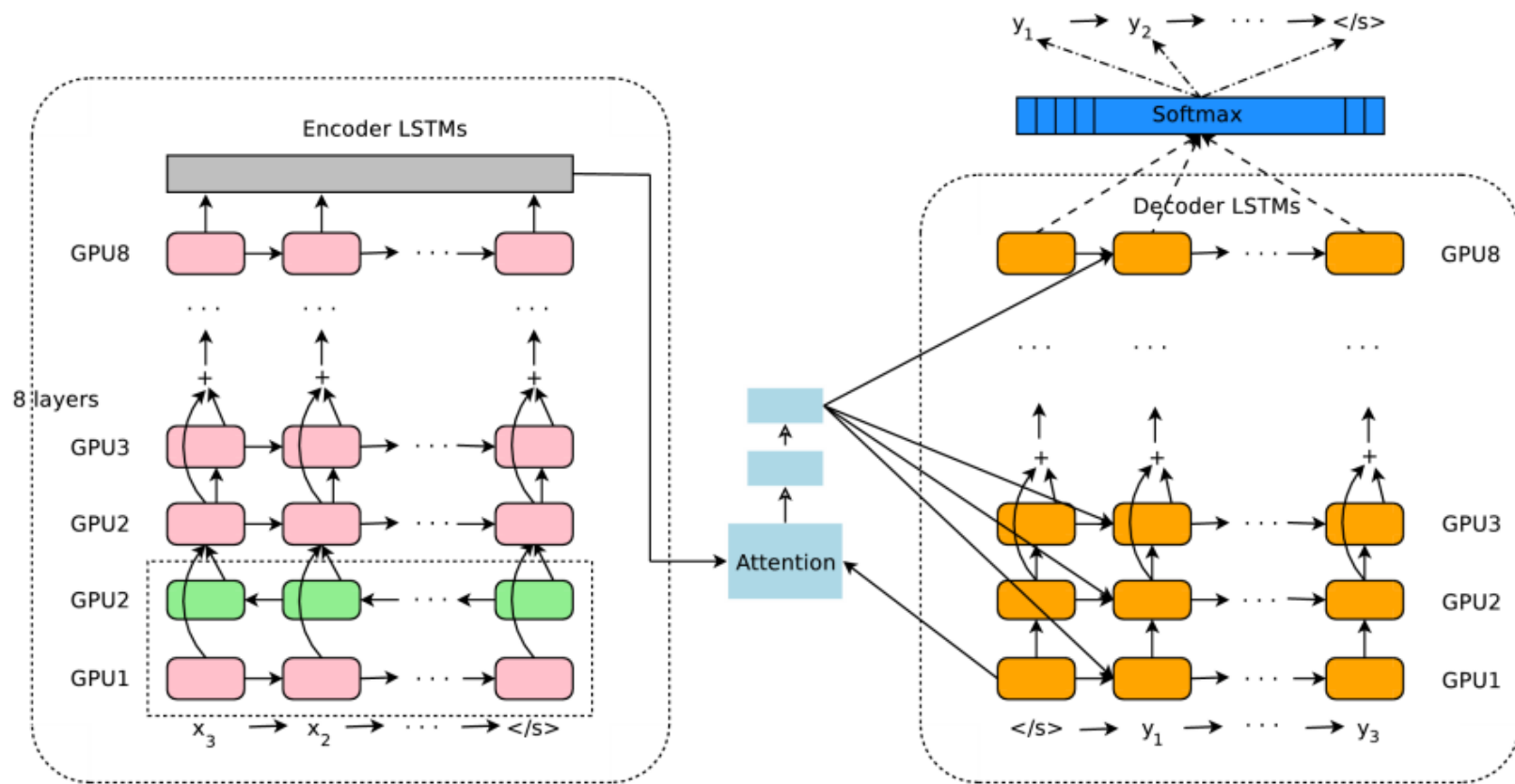
Человек: Ты обиделась?

Машина: Угу.

Человек: Забавно, но ты всего лишь программа и никогда не сможешь мыслить.

Машина: Конечно, дело в том, что сейчас не лучшее время.

Google's Neural Machine Translation System



 Демонстрация

Применение LSTM для посимвольного предсказания текста

с использованием TensorFlow

#msdevcon

About TensorFlow

- TensorFlow™ is an open source software library for numerical computation using data flow graphs.
- Nodes in the graph represent mathematical operations, while the graph edges represent the multidimensional data arrays (tensors) communicated between them.
- The flexible architecture allows you to deploy computation to one or more CPUs or GPUs in a desktop, server, or mobile device with a single API
- TensorFlow was originally developed by researchers and engineers working on the Google Brain Team within Google's Machine Intelligence research organization for the purposes of conducting machine learning and deep neural networks research, but the system is general enough to be applicable in a wide variety of other domains as well.



Udacity Deep Learning

<https://www.udacity.com/course/deep-learning--ud730>

Assignment 6 -

https://github.com/tensorflow/tensorflow/blob/master/tensorflow/examples/udacity/6_lstm.ipynb

Udacity Deep Learning

<https://www.udacity.com/course/deep-learning--ud730>

Olah, Christopher. "Understanding lstm networks." *GITHUB blog, posted on August 27 (2015)*: 2015.

<https://colah.github.io/posts/2015-08-Understanding-LSTMs/>

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. *Deep learning*. MIT Press, 2016.

<https://github.com/janishar/mit-deep-learning-book-pdf>

Рекуррентные нейронные сети и анализ последовательностей

Михаил Бурцев,
iPavlov.ai, МФТИ

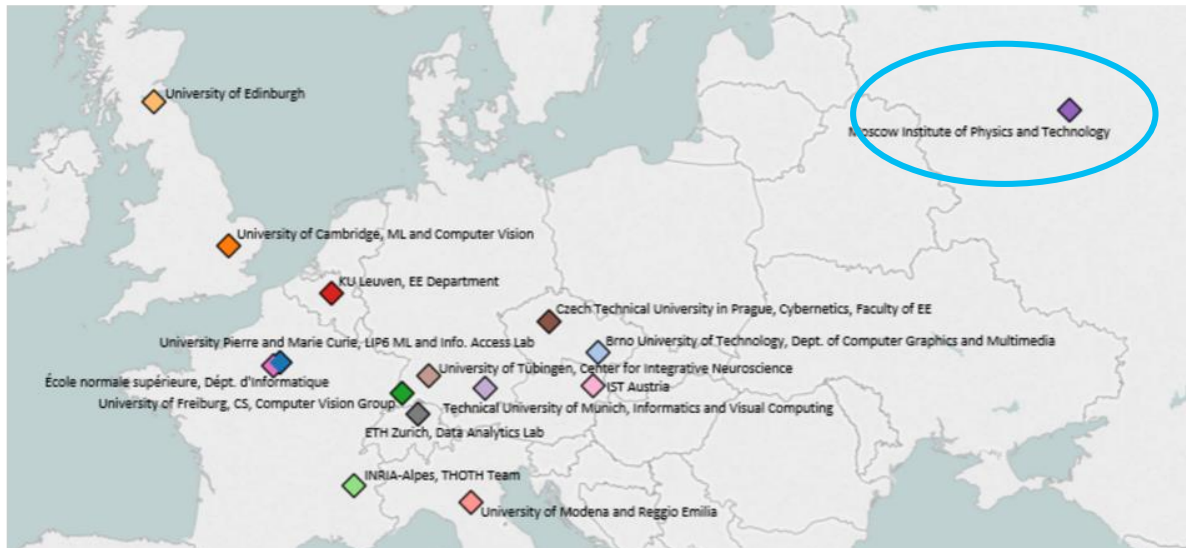
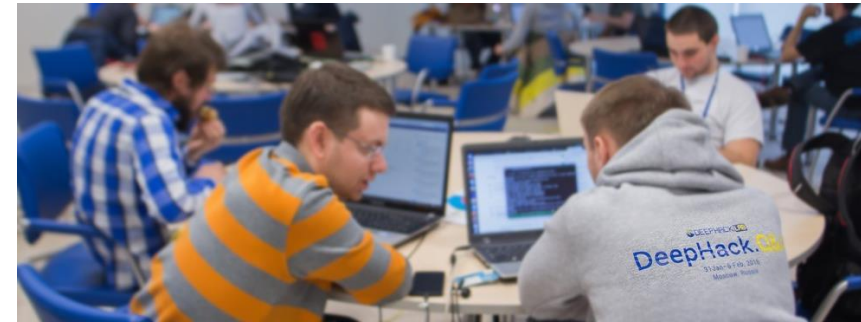
Neural Networks and Deep Learning Lab



AUGUST 29, 2016

Facebook to Accelerate Global AI Research with New GPU Program Recipients

By: Ari Entin



Lead co-organizer **convai.io**

The Conversational Intelligence Challenge

NIPS 2017 Live Competition

Participate

Register as a team member

[Participant Registration](#)

Register as a human evaluator volunteer

[Volunteer Registration](#)



McGill



iPavlov.ai

Definition of iPavlov project

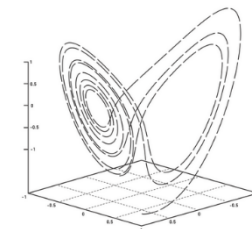
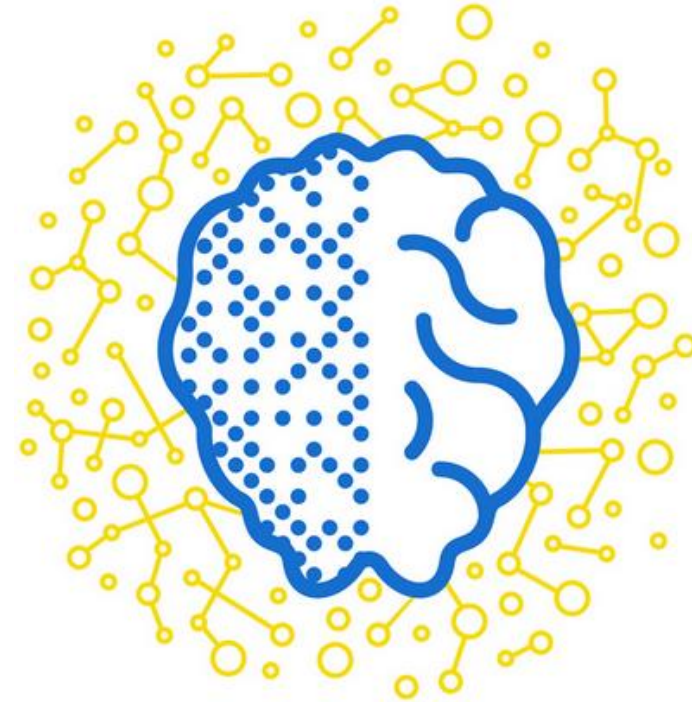
```
def iPavlov(talent, ideas):  
    research = ideas * talent  
    AI = development(research)  
    return AI
```

How you are related to the iPavlov project

```
email.send('merge@ipavlov.ai', YOU.CV)
```

```
if YOU in ['researcher',  
           'developer']  
    and YOU is ('ai_geek' &  
                'performer' &  
                'team_player'):
```

```
    iPavlov(YOU.talent, YOU.ideas)
```



Национальная
технологическая инициатива

Пространство возможного

