

Сети LSTM и GRU

Нейросети для анализа текстов

Сети LSTM и GRU

Рекуррентные нейронные сети

- Сети с циклами

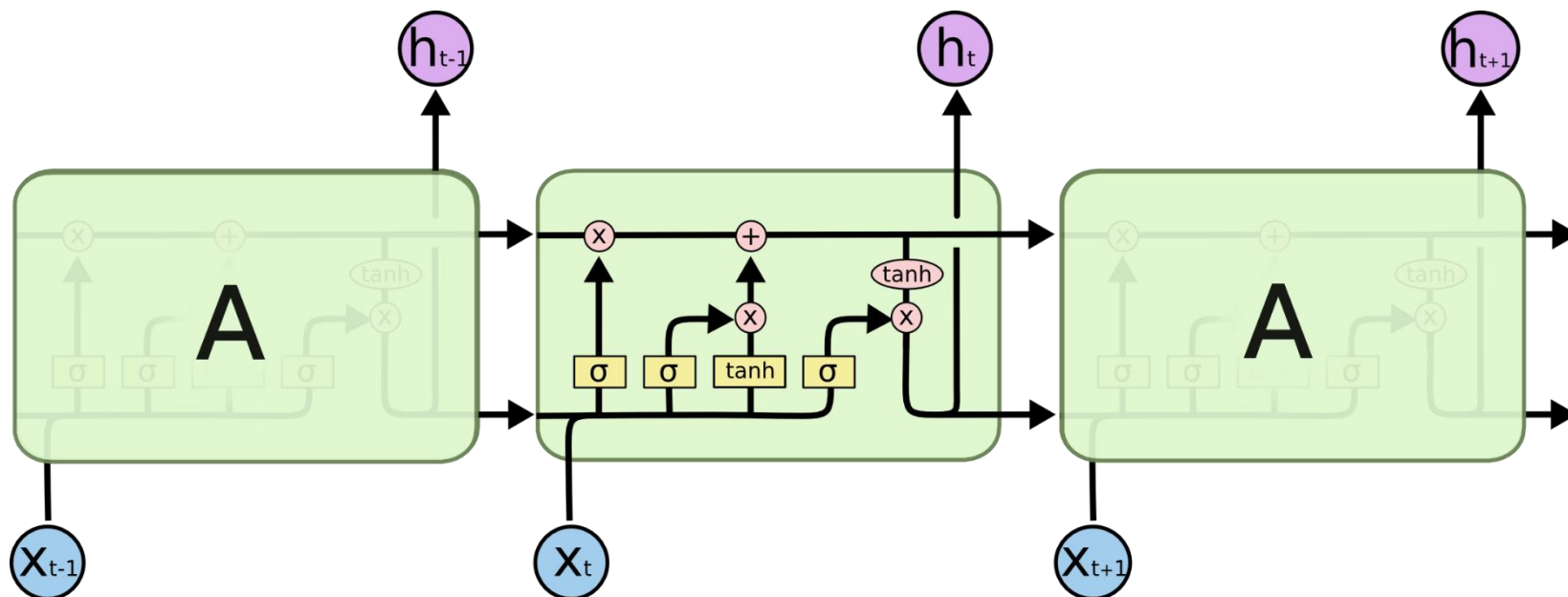
Проблемы рекуррентных нейросетей

- Обучение требует длительного времени
- Проблема исчезающего градиента
- Ограниченная «длительность» запоминания предыдущей информации

Более совершенные архитектуры рекуррентных сетей

- Long Short Term Memory – LSTM (Hochreiter, Schmidhuber, 1997)
- Gated Recurrent Unit – GRU (Cho, 2014)

Ячейка LSTM (LSTM cell)



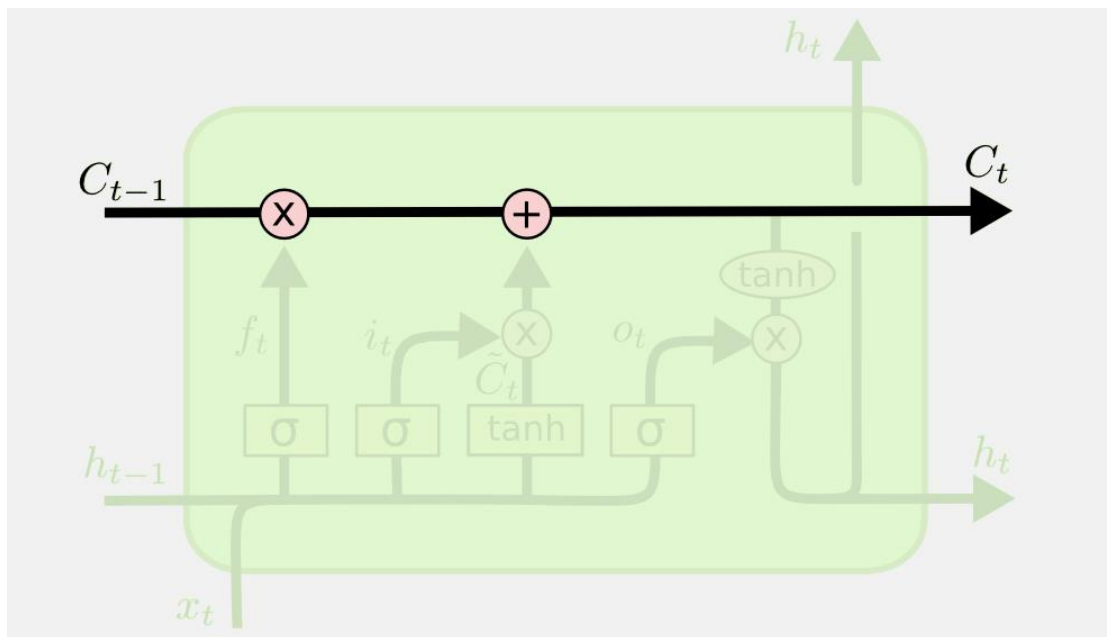
Understanding LSTM Networks

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

Как понять LSTM сети

- <https://alexsosn.github.io/ml/2015/11/17/LSTM.html>

Состояние ячейки

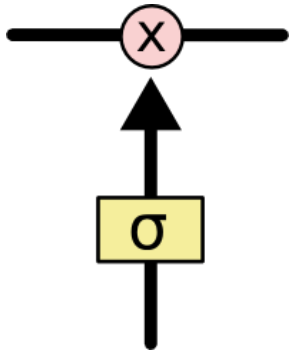


Вентили LSTM

Вентиль (gate) LSTM

- Управляет передачей сигнала внутри ячейки LSTM

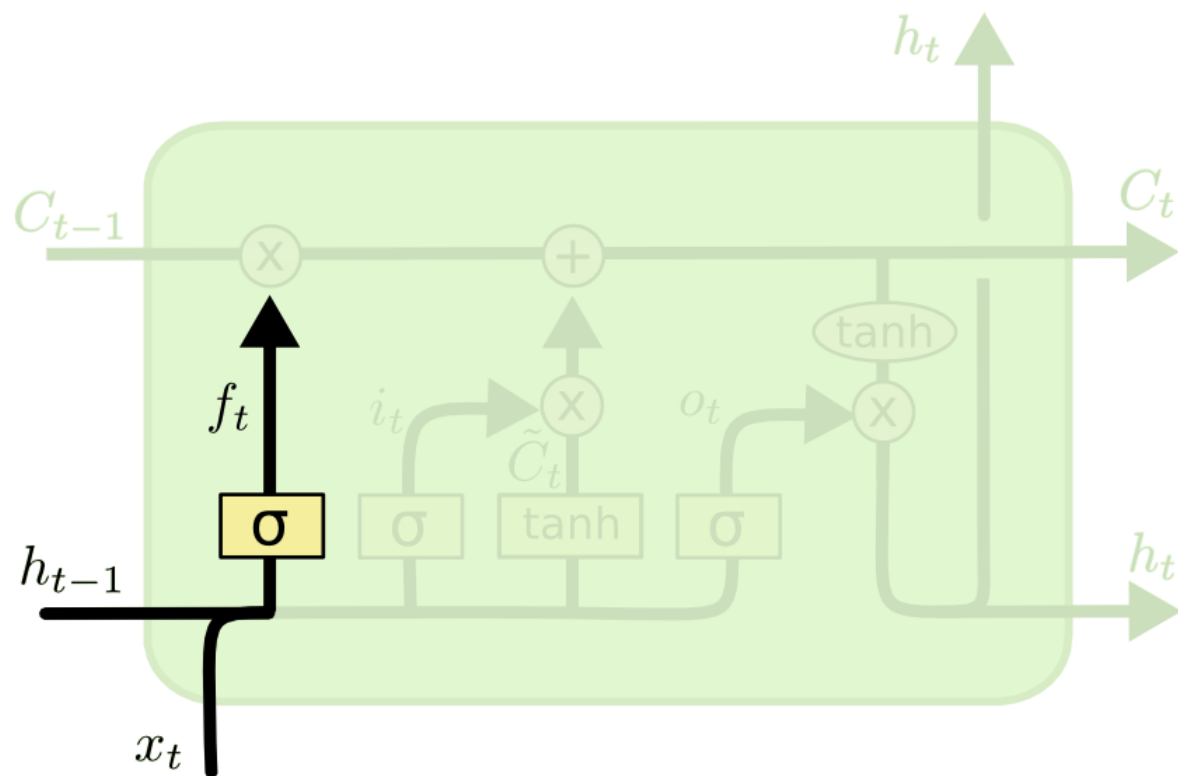
Структура вентилля



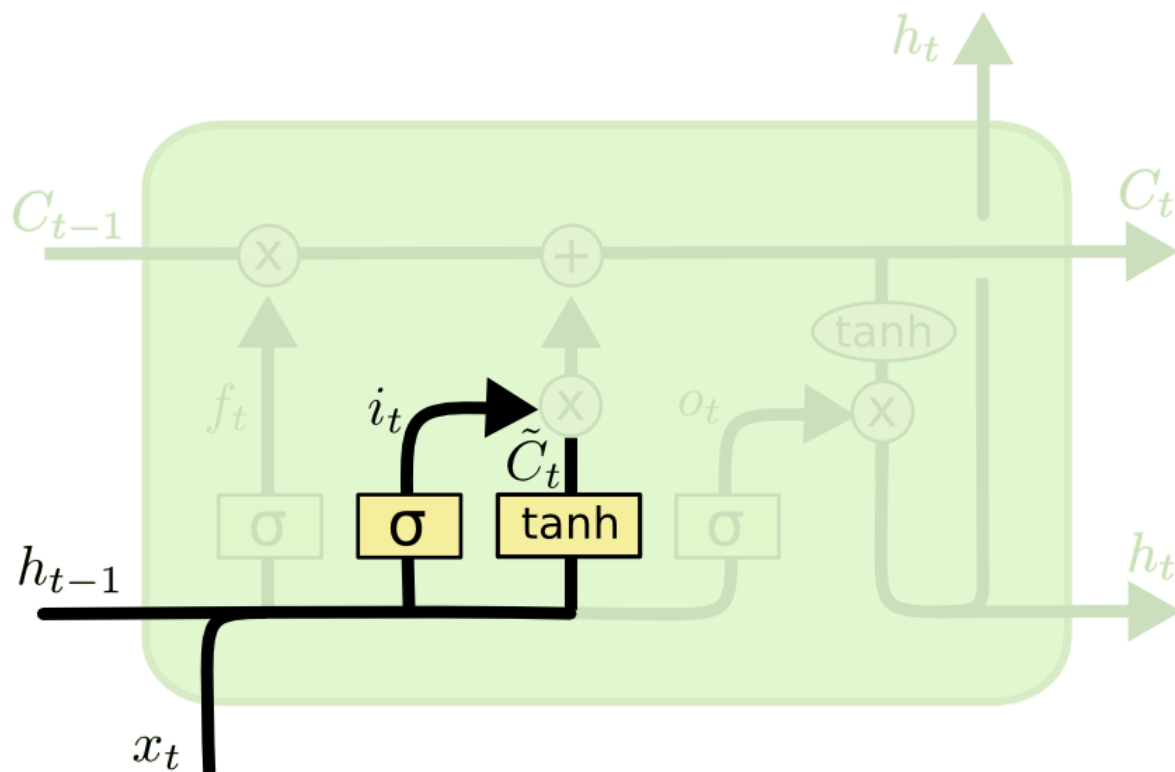
Типы вентилей в LSTM

- Вентиль забвения (forget gate)
- Входной вентиль (input gate)
- Выходной вентили (output gate)

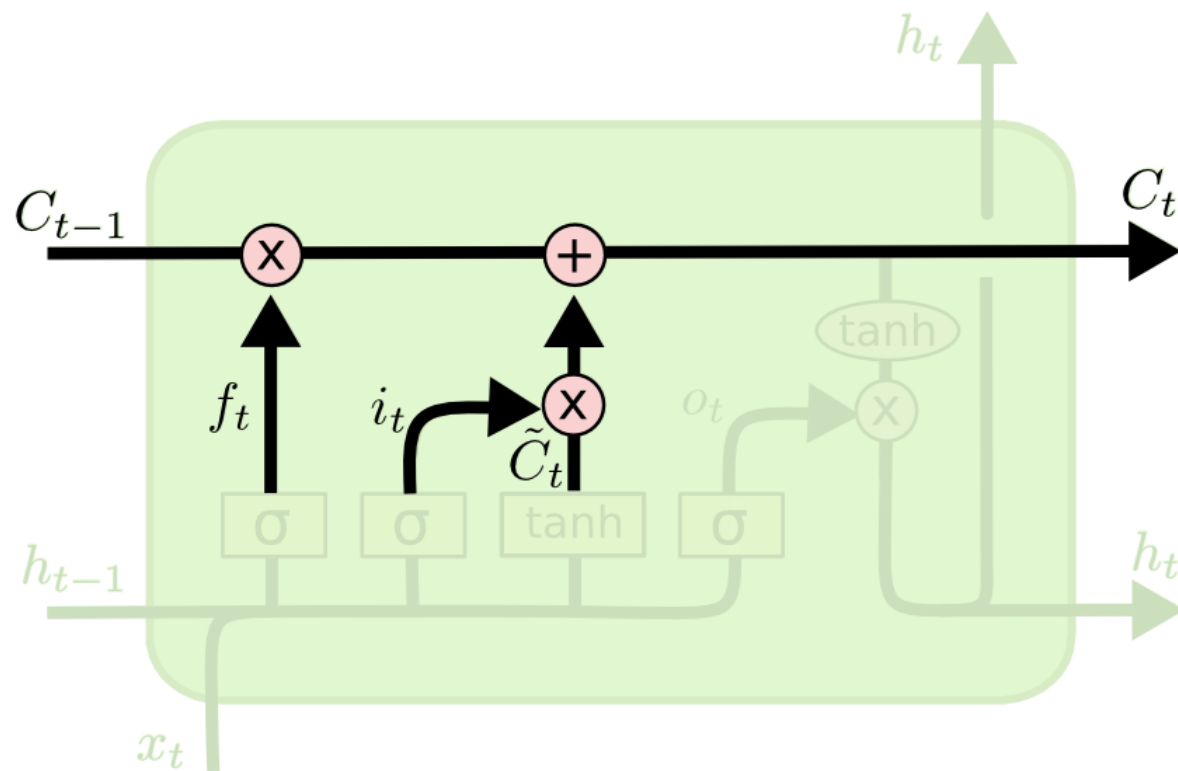
Вентиль забывания



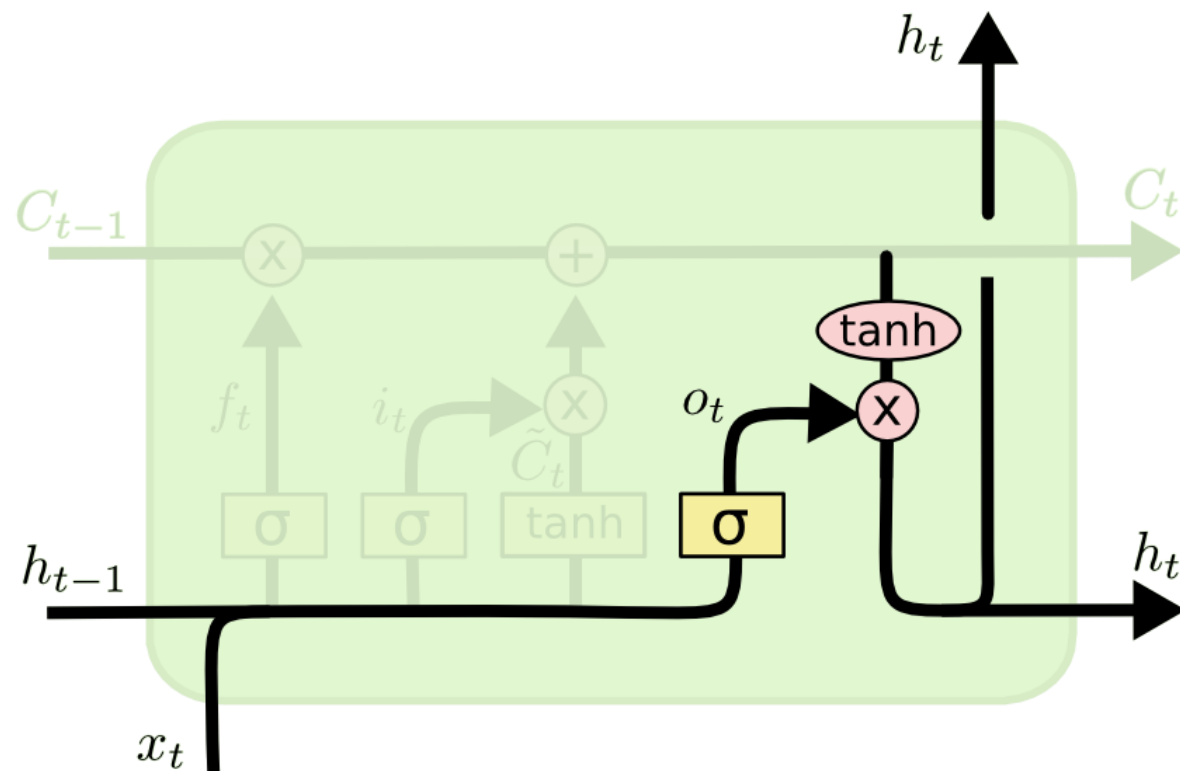
Входной вентиль



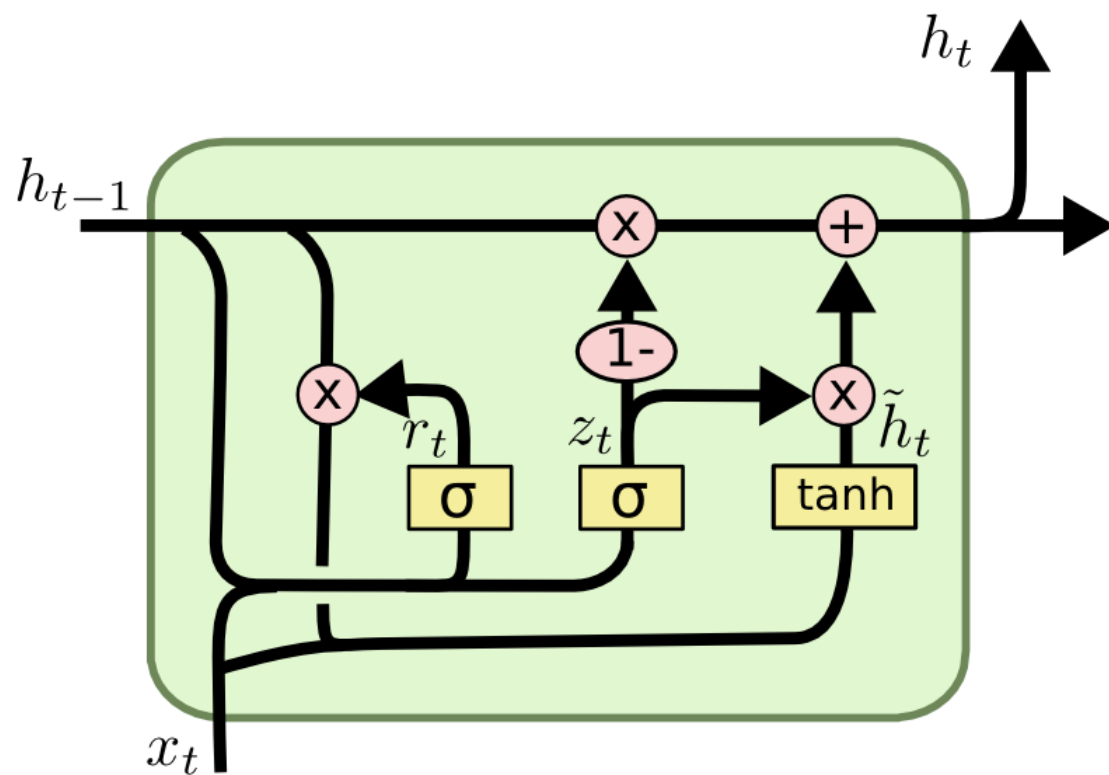
Обновление состояния



Выходной вентиль



Сети GRU



Сети LSTM в Keras

```
model = Sequential()  
model.add(Embedding(input_dim=max_words,  
                    output_dim=50,  
                    input_length=maxlen))  
model.add(LSTM(8))  
model.add(Dense(1, activation='sigmoid'))
```

Сети GRU в Keras

```
model = Sequential()  
model.add(Embedding(input_dim=max_words,  
                    output_dim=50,  
                    input_length=maxlen))  
model.add(GRU(8))  
model.add(Dense(1, activation='sigmoid'))
```

Популярные архитектуры рекуррентных сетей

- Long Short Term Memory – LSTM
- Gated Recurrent Unit – GRU

Ячейки LSTM и GRU

- «Ячейка памяти»
- Вентили

Преимущества

- Могут запоминать информацию на длительное время
- Решена проблема исчезающего градиента

Обучение сетей LSTM и GRU

- Обратное распространение ошибки
- Обучаются вентили и веса входов в нейроны