



Нейросетевые языковые модели

Евгений Борисов

Нейросетевые языковые модели

Языковая модель

- предсказываем следующее слово на основе предыдущих
- оценка (вероятность) совместимости цепочки слов

Оценка цепочки слов (биграммная модель):

$$p(w_1 \dots w_n) = \prod_{k=1}^n p(w_k | w_{k-1})$$

$$p(w_n | w_{n-1}) = \frac{C(w_{n-1} w_n)}{C(w_{n-1})}$$

P - вероятность совместного использования слов

C(w) – количество слов w в тексте

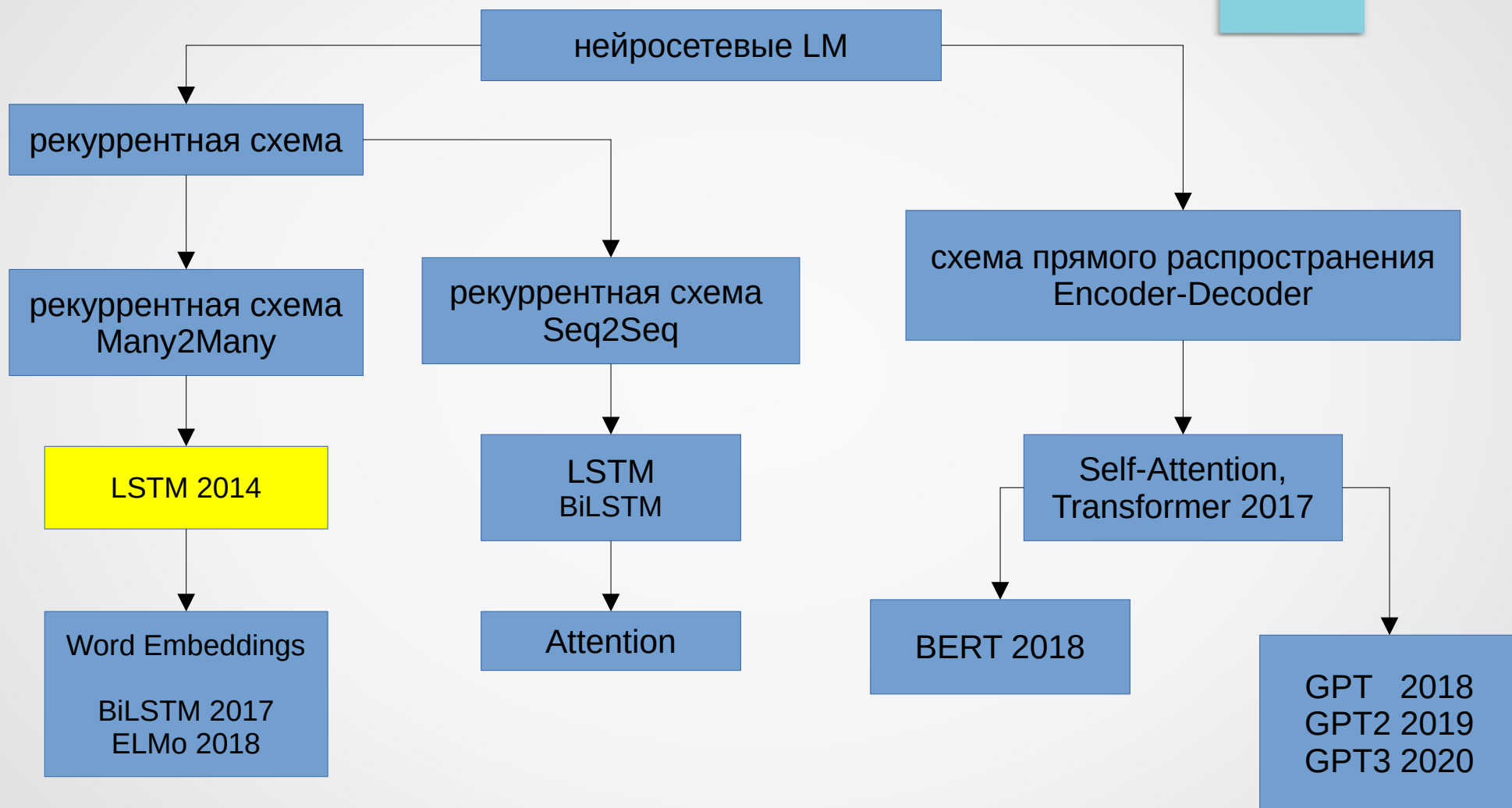
Языковые модели

нейросетевая языковая модель (word based model)

- из текстов собираем пары *[[контекст], слово]*
- обучаем RNN по контексту определять слово

input -> LSTM -> softmax

Нейросетевые языковые модели

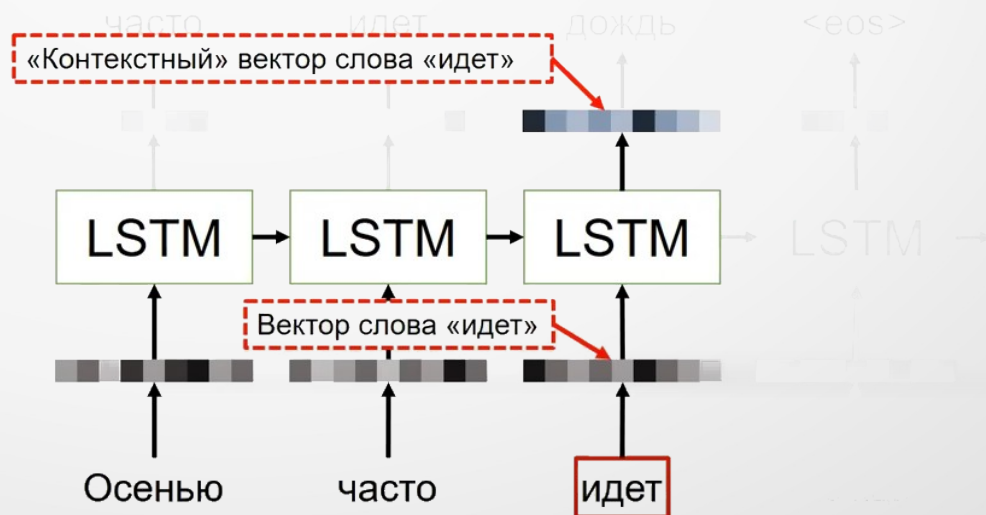
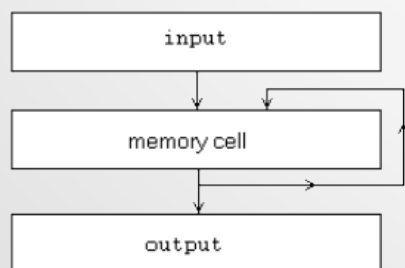
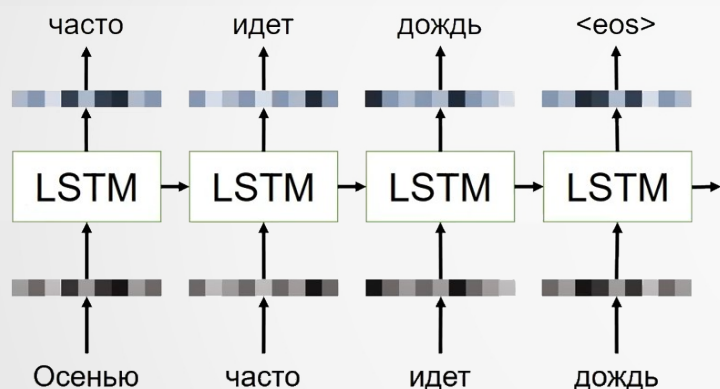


Нейросетевые языковые модели

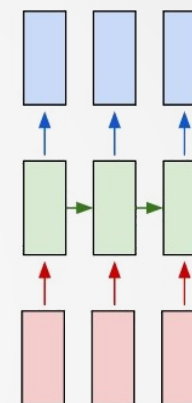
Простая схема с рекуррентной сетью

Предсказываем следующее слово по предыдущему контексту

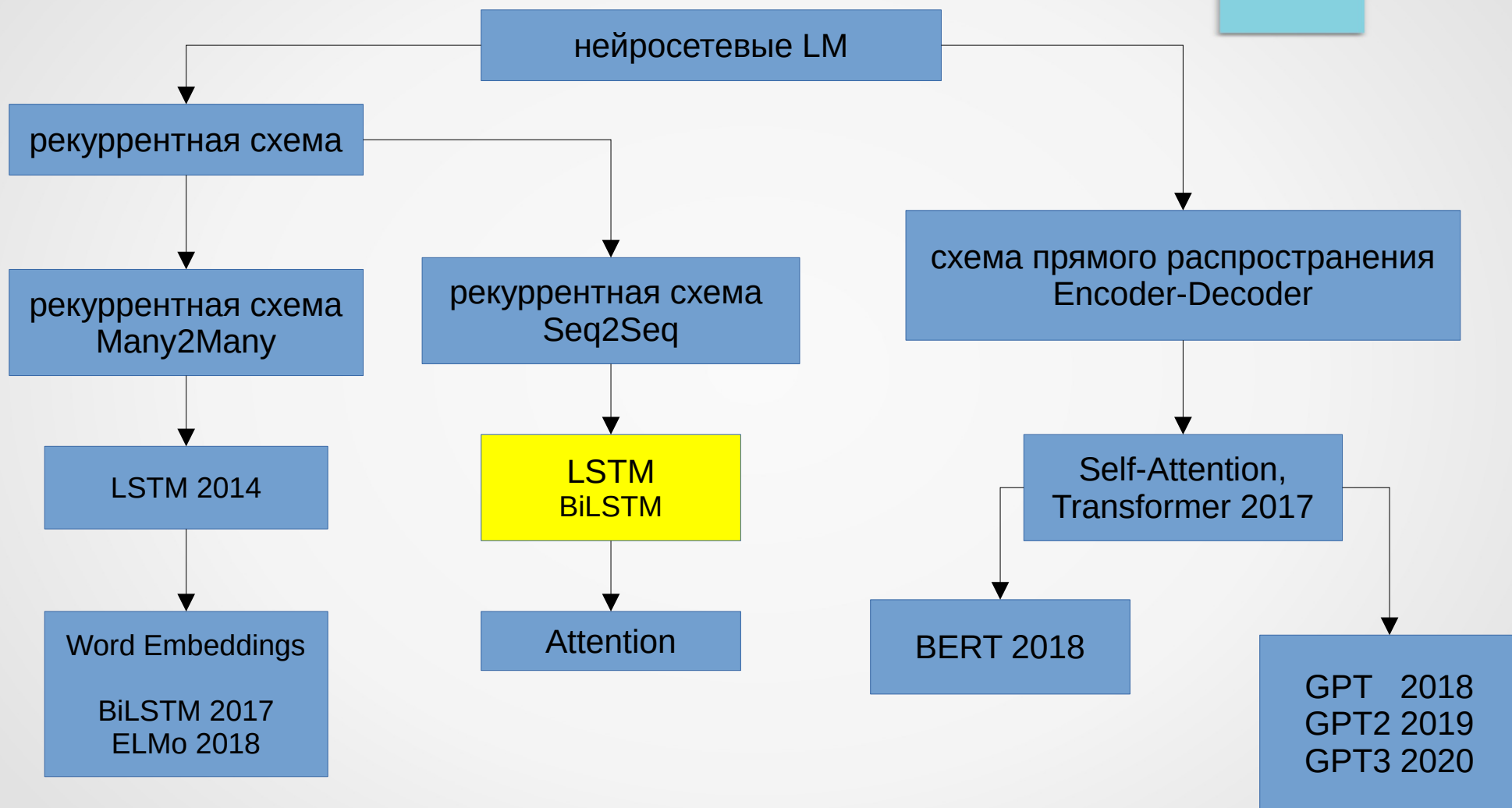
получаем word embedding, который учитывает левый контекст



many to many

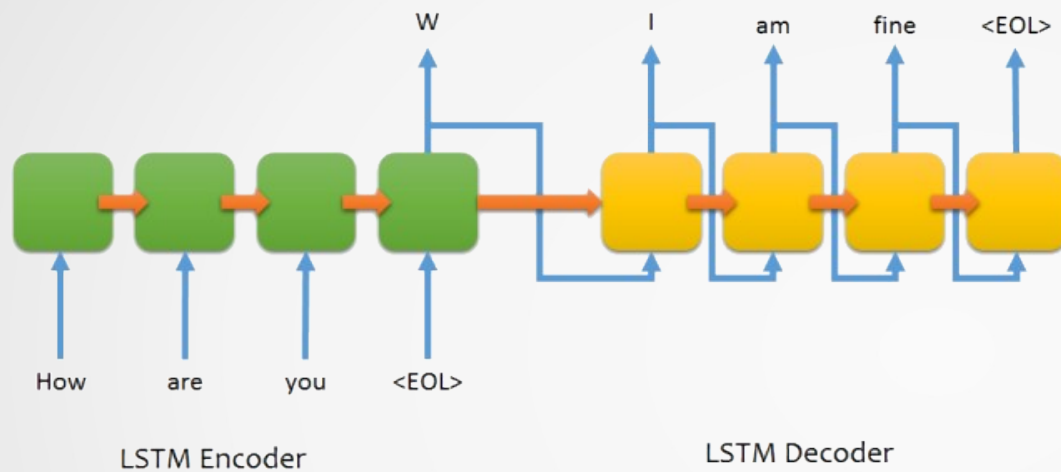


Нейросетевые языковые модели



Нейросетевые языковые модели

Рекуррентная схема SEQ2SEQ

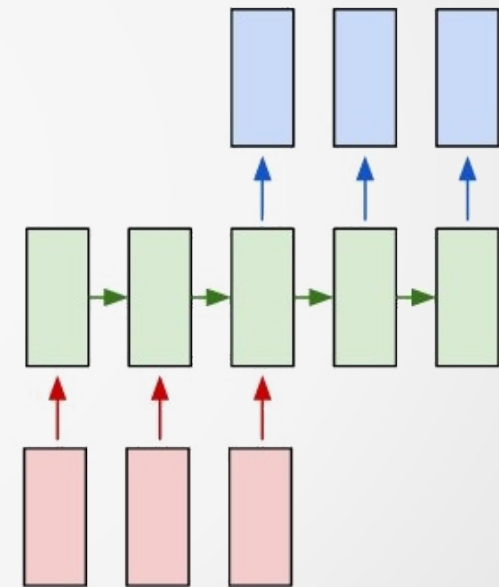


пара рекуррентных неросетей

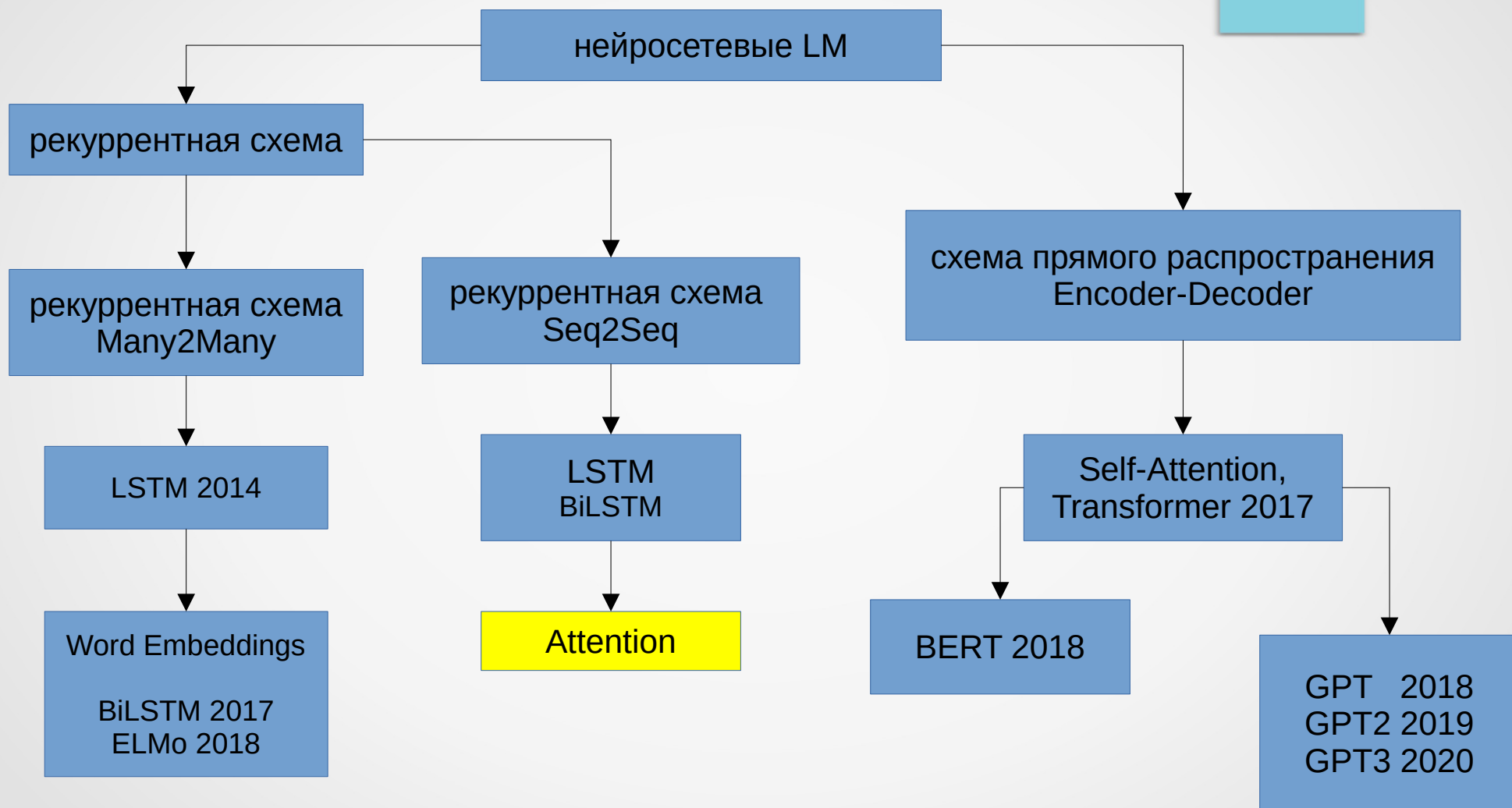
- кодировщик,
формирует внутреннее представление

- декодировщик,
авторегрессионная модель,
разворачивает состояние энкодера

many to many



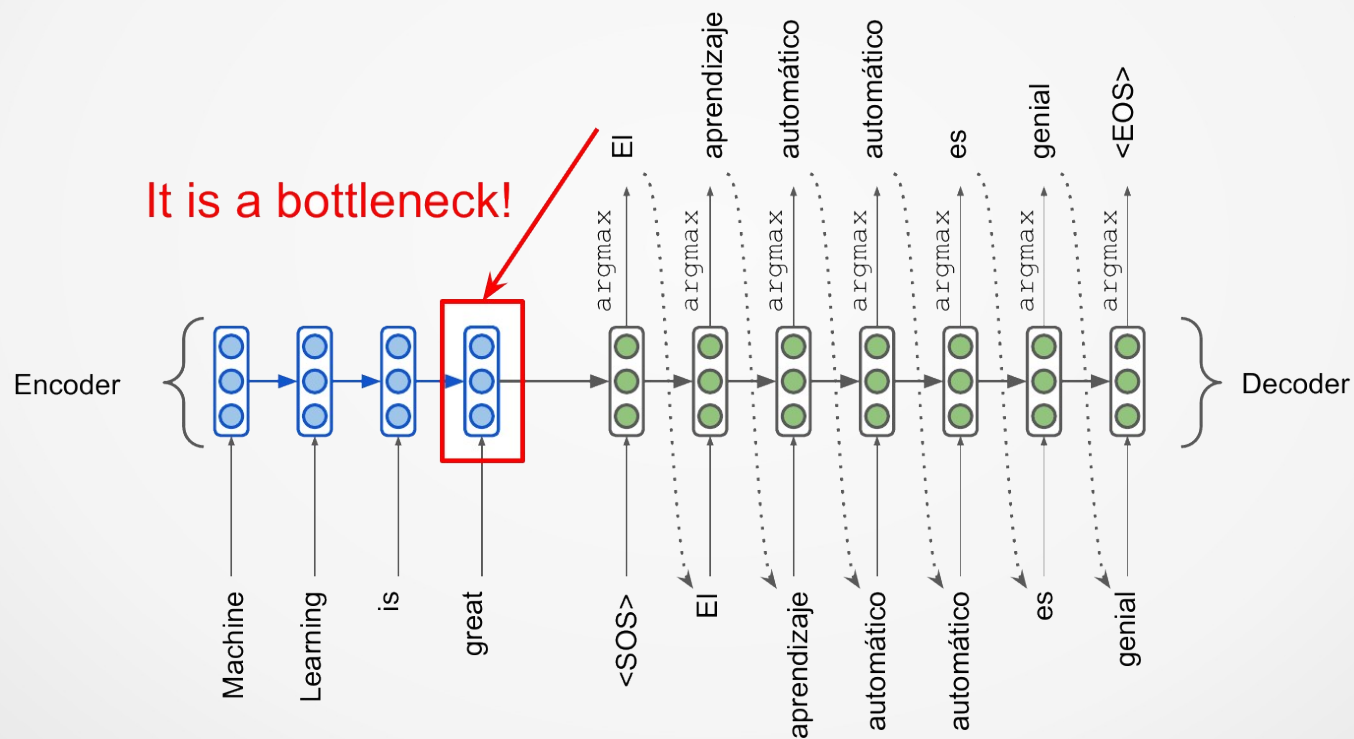
Нейросетевые языковые модели



Нейросетевые языковые модели

Рекуррентная схема SEQ2SEQ

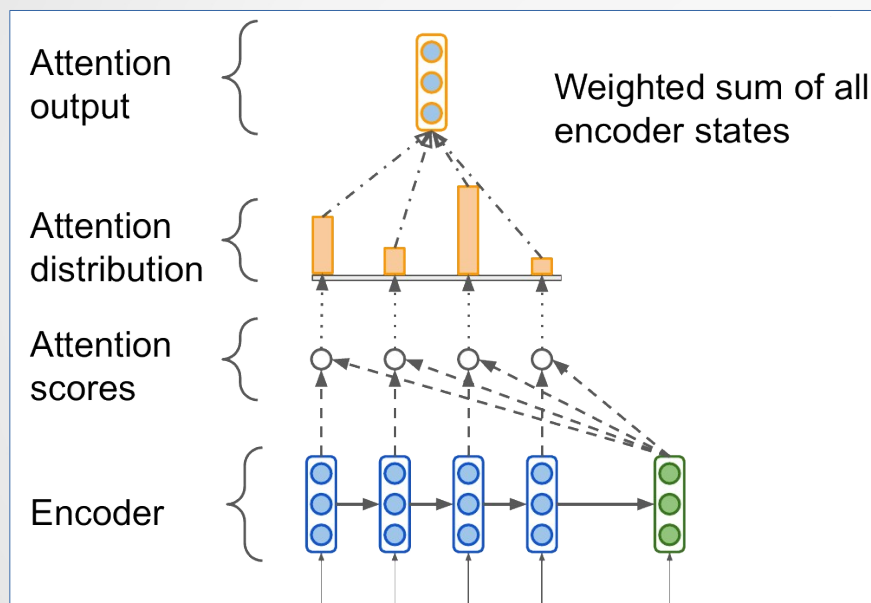
проблема: модель может "забывать" начальный контекст



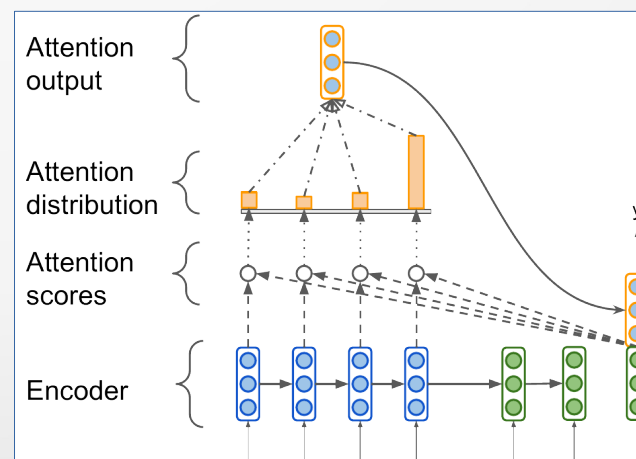
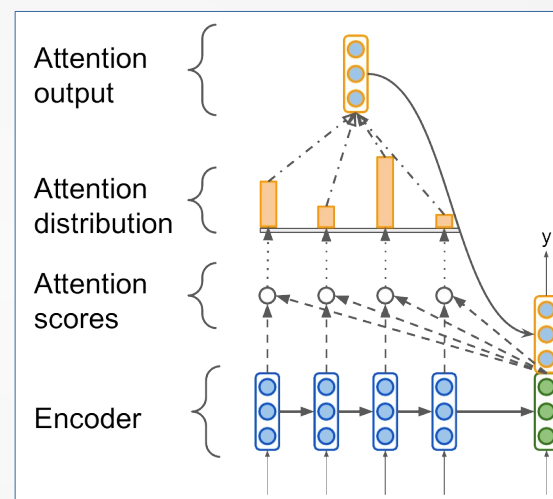
Нейросетевые языковые модели

Рекуррентная схема SEQ2SEQ и механизм внимания (Attention)

- считаем коэффициенты внимания
- добавляем к состоянию декодера взвешенную сумму входов



все слова подаём в модель последовательно,
схему вычислений трудно распараллелить



Литература

Литература

Борисов Е.С. Методы машинного обучения. 2024
https://github.com/mechanoid5/ml_lectorium_2024_I

Борисов Е.С. Методы обработки текстов на естественном языке. 2024
https://github.com/mechanoid5/ml_nlp_2024_I

Турдаков Д.Ю. Основы обработки текстов. лекция 3. Языковые модели. ИСП РАН, 2017
<https://www.youtube.com/watch?v=seAxPaKw33g>

Анатолий Востряков Языковые модели на все случаи жизни, ODS Data Fest 2018
<https://www.youtube.com/watch?v=TaCbj1kaDQY>

Майоров В.Д. Основы обработки текстов. Лекция 10. Языковые модели. ИСП РАН, 2021
https://www.youtube.com/watch?v=_8MGdpt4I9M

Нейчев Радослав Машинное обучение 3. Machine translation. Лекторий ФПМИ, 2020
<https://www.youtube.com/watch?v=6HibilFua-U>