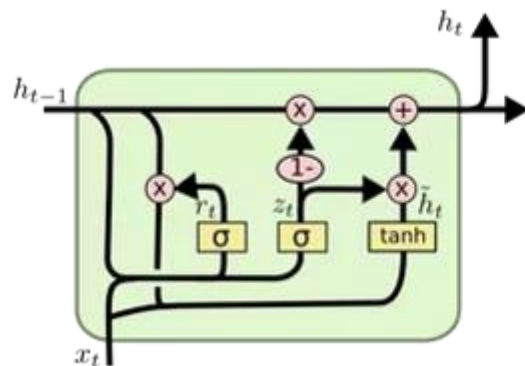


1. Как меняется расчет внутреннего состояния в рекуррентных нейронных сетях?

Есть разные алгоритмы расчета внутреннего состояния, например, характерные для GRU, LSTM. В общем случае суть в том, что учитываются предыдущие элементы входной последовательности. Для этого обычно используют текущее внутреннее состояние и текущий элемент входной последовательности, к ним применяют определенные фильтры и операции произведения и суммы.

2. Что такое блок GRU?

*Gated Recurrent Unit. Тип архитектуры рекурсивной сети. Принцип в том, что у сети есть внутреннее состояние, которое зависит от обработанных элементов входной последовательности. Это реализуется так: при вычислении выходного сигнала для очередного элемента учитывается выходной сигнал предыдущего элемента, который определяет состояние сети в момент обработки элемента. Но поскольку последний выходной сигнал зависел от предпоследнего, состояние сети определяется всеми обработанными элементами последовательности и их порядком. Принципиальная схема показывает, как именно элемент входной последовательности и внутреннее состояние определяют выходной сигнал.*



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

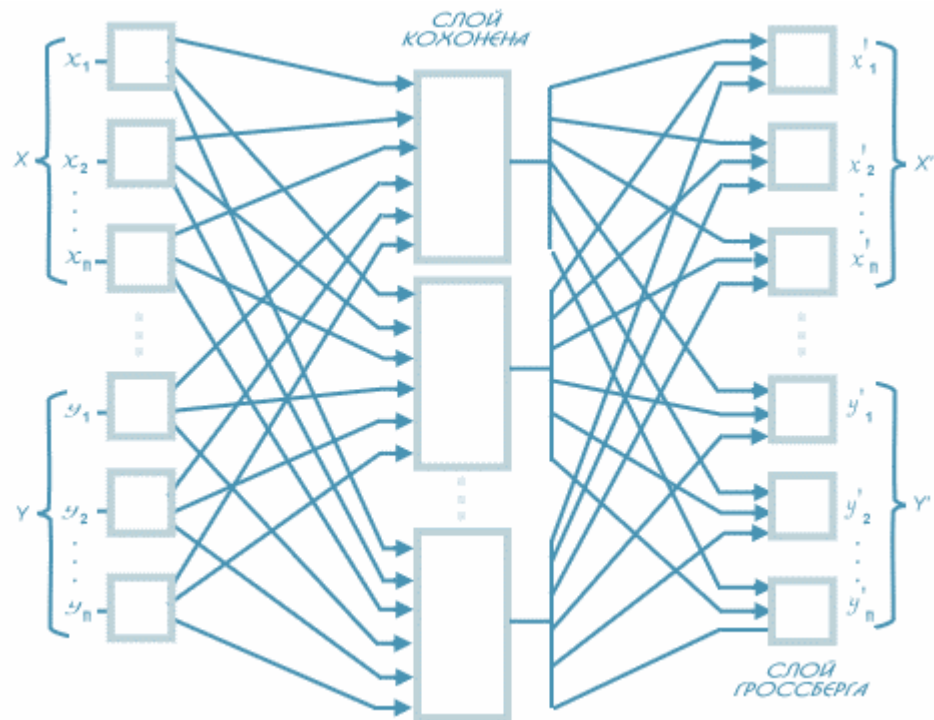
$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t])$$

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

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

3. В каких случаях может быть необходимо использовать 2 входных слоя в ИНС?

Такое может случиться, например, в НС встречного распространения, где на вход подается X и Y.



4. За оформлению отчета снимаю 0,5 балла

Скажите пожалуйста, в чем конкретно дело (не хочу повторить ошибку в следующей работе)?

5. Можно ли применять слой Convolution2D после слоя Flatten?

Нет, потому что Convolution2D предполагает двумерность карты. Возможно, можно как-то извернуться и заставить его интерпретировать вектор  $n$  элементов как матрицу  $(1, n)$  или заставить сеть применить reshape. Но это лишено смысла, потому что удобнее тогда отказаться от Flatten.

6. Как изменится тензор  $[1 \ 1 \ 2]$ , если был применен слой Dropout с шансом отклонения 0.5, и был отброшен 3 элемент?

$[1 \ 1 \ 0]$ . Я знаю, что такой ответ вы уже не приняли, поэтому, пожалуйста, объясните или назовите верный ответ.

