

1. Как обычная свертка раскладывается на поканальную свертку?

Сначала входной тензор сворачивается ядром размерности $1 \times 1 \times (\text{кол-во каналов входного тензора})$. Далее делается свертка 3×3 по каждому каналу, потом свертка 1×1 .

2. Что такое рекуррентная сеть?

Рекуррентная нейронная сеть – нейронная сеть, в которой, в отличие от многослойного перцептрона и сверточной нейронной сети, нейроны дополнительно хранят информацию о предыдущем состоянии сети, как будто у нее есть «память». Такие сети хорошо подходят для задачи анализа последовательностей. В рекуррентных сетях разрешены циклы: в этом случае выход нейрона может быть подключен к его входу, ко входу всех нейронов в текущем слое (такой подход используется чаще всего), а также теоретически можно подключить выход нейрона к любому другому нейрону в любому слое. Через эти циклические соединения и поступает информация о том, что было на предыдущем шаге работы сети или даже на нескольких предыдущих шагах. Так сеть анализирует текст не как набор изолированных токенов, а как последовательность.

3. Что такое LSTM блок? Для чего нужен?

LSTM-блоки содержат три или четыре «вентиля», которые используются для контроля потоков информации на входах и на выходах памяти данных блоков. Эти вентили реализованы в виде логистической функции для вычисления значения в диапазоне $[0; 1]$. Умножение на это значение используется для частичного допуска или запрещения потока информации внутрь и наружу памяти. Таким образом, LSTM-блок определяет, как распоряжаться своей памятью как функцией этих значений, и тренировка весов позволяет LSTM-блоку выучить функцию, минимизирующую потери.

4. За счет чего слой Dropout добавляет устойчивость к переобучению?

Dropout исключает нейроны из сети с заданной вероятностью. “Исключение” нейрона означает, что при любых входных данных или параметрах он возвращает 0. Исключенные нейроны не вносят свой вклад в процесс обучения ни на одном из этапов алгоритма обратного распространения ошибки, поэтому исключение хотя бы одного из нейронов равносильно обучению новой нейронной сети. Таким образом, Dropout предотвращает совместную адаптацию нейронов: нейроны не могут полагаться на другие в исправлении собственных ошибок.

5. Распишите, что значат параметры "Convolution2D(conv_depth_2, kernel_size, kernel_size, border_mode='same', activation='relu')(conv_3)"

conv_depth_2 – размерность выходного пространства (количество выходных фильтров в свертке)

kernel_size, kernel_size – целые числа, задающие ширину и высоту окна 2D свертки

border_mode='same' – возвращает то же измерение, что и входное изображение

activation='relu' – функция активации нейронов

6. Каким параметром можно задать, чтобы свертка проводилась с шагом 3?

strides=3 или strides=(3,3)