

1. Что такое LSTM блок? Для чего нужен?

Ответ: LSTM блоки как правило состоят из состояния ячейки(текущий поток информации блока) и 3-4 «вентилей»(по-другому «гейты», логистические функции, которые определяют, сколько информации мы будем обрабатывать). Входные вентили определяют решают, какую новую информацию мы добавим в состояние ячейки, выходные же вентили решают, какую информацию мы передадим в сеть в следующий момент времени. Таким образом LSTM блоки и сети используются с последовательными данными, обработке текста и звуковых данных, для анализа временных данных.

2. Что такое shortcut соединение и для чего оно нужно?

Ответ: у глубоких сверточных сетей есть один недостаток: с увеличением глубины затухает градиент при обратном проходе по причине того, что функция активации relu часто дает 0. Для решения этой проблемы были введены соединения быстрого доступа, явно добавляя тождественное отображение. В итоге обратный проход будет всегда выполняться и затухания градиента не произойдет.