

RNN

Главной характеристикой полносвязных и сверточных нейронных сетей, является отсутствие памяти. Каждый вход обрабатывается ими независимо, без сохранения состояния между ними. Чтобы с помощью таких сетей обработать последовательность или временной ряд данных, необходимо передать в сеть всю последовательность целиком, преобразовав ее в единый пакет.

RNN хороша для анализа временных последовательностей (звуковой сигнал или изменение положения объекта во время движения, текст) [важно знание предыдущих элементов последовательности]

Рекуррентная нейронная сеть обрабатывает последовательность, перебирая ее элементы и сохраняя состояние, полученное при обработке предыдущих элементов.

На вход каждого нейрона поступает не только входной сигнал, но и предыдущее состояние сети. Весовые коэффициенты одинаковы на всех итерациях слоя.

- Many to Many – например, для перевода текстов;
- Many to One – например, для анализа эмоциональной окраски текста (на входе текст, на выходе категории: положительный, нейтральный, отрицательный);
- One to Many – например, для генерации описания изображений, когда на вход подаются карты признаков изображения, а на выходе получаем его описание (текст);
- One to One – относительно редкая архитектура для выполнения нелинейных рекуррентных вычислений.

ВРТТ – Обратное распространение ошибки с учетом предыдущих вычислительных шагов, развёрнутых во времени при вычислении локального градиента

Embedding

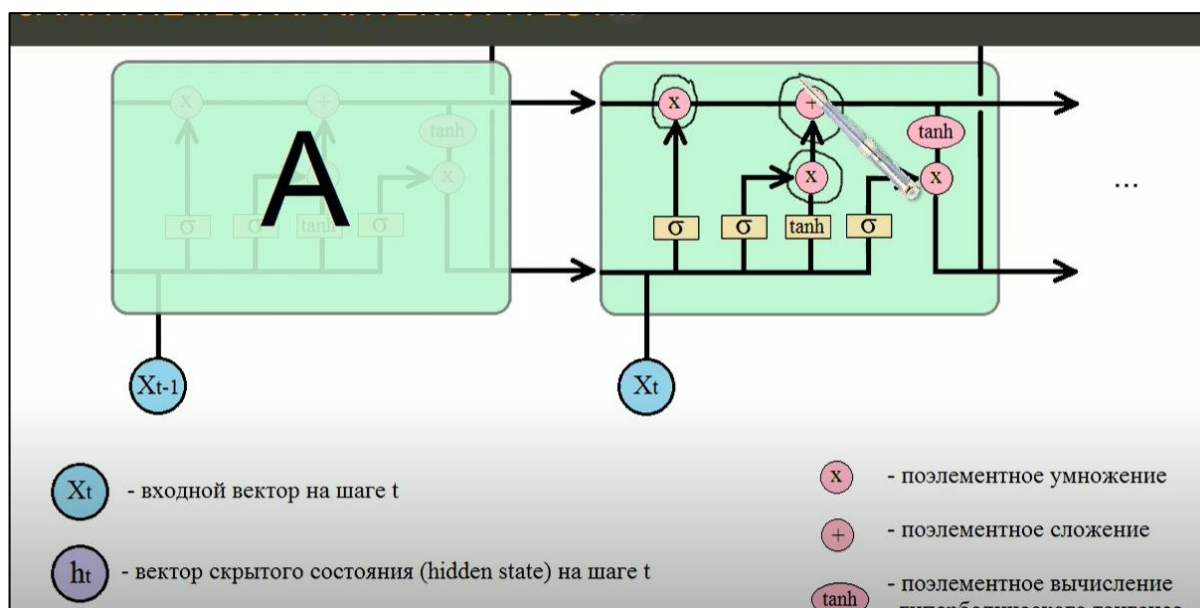
На выходах слоя формируются выходные значения, равные весам связей для переданных значений, т.е. определённому индексу ставится в соответствие определенный вектор.

Мы отобразим каждый обзор фильма в реальную векторную область, популярную технику при работе с текстом, которая называется встраивание слов. Это метод, в котором слова кодируются как действительные векторы в многомерном пространстве, где сходство между словами в смысле смысла приводит к близости в векторном пространстве.

LSTM (долгая краткосрочная память)

У обычной RNN проблема – быстрое забывание прошлого контекста.

В LSTM ячейке памяти присутствует верхний канал, который позволяет сохранять долгосрочный контент. И при движении по этому каналу вектор выполняет сначала поэлементное умножение (забывание ненужного) [долгосрочный контент, пришедший из предыдущего рекуррентного шага и скрытое состояние рекуррентного блока] и поэлементное сложение (запоминание нового)



Разница GRU

В GRU фильтры «забывания» и входа объединяют в один фильтр «обновления» (update gate). Кроме того, состояние ячейки объединяется со скрытым состоянием, есть и другие небольшие изменения. Построенная в

результате модель проще, чем стандартная LSTM, и популярность ее неуклонно возрастает.