

УДК 004.9

АРХИТЕКТУРА МОДЕЛИ LSTM ДЛЯ ПРОГНОЗИРОВАНИЯ ВРЕМЕННЫХ РЯДОВ РЕДКИХ СОБЫТИЙ

В.В. Якименко

Научный руководитель – И.Л. Савостьянова

Сибирский государственный университет науки и технологий имени академика М. Ф. Решетнева
Российская Федерация, 660037, г. Красноярск, просп. им. газ. «Красноярский рабочий», 31
E-mail: vv.yakimenko@yandex.ru

В статье приведена информация об особой разновидности архитектуры рекуррентных нейронных сетей – LSTM. Были рассмотрены структуры этих нейронных сетей, их основные идеи и вариации.

Ключевые слова: рекуррентные нейронные сети, временные ряды, прогнозирование, LSTM, AI.

LSTM MODEL ARCHITECTURE FOR PREDICTING TIME SERIES OF RARE EVENTS

V.V. Yakimenko

Scientific supervisor – I.L. Savostyanova

Reshetnev Siberian State University of Science and Technology
31, Krasnoyarskii rabochii prospekt, Krasnoyarsk, 660037, Russian Federation
E-mail: vv.yakimenko@yandex.ru

The article provides information about a special kind of recurrent neural network architecture – LSTM. The structures of these neural networks, their main ideas and variations were considered.

Keywords: recurrent neural networks, time series, forecasting, LSTM, AI.

Анализ временных рядов относится к анализу изменения тренда данных за определенный период времени. Анализ временных рядов имеет множество применений, от прогнозирования будущей стоимости товара на основе его прошлых значений до более сложных экономических и информационных явлений [1].

Долгая краткосрочная память (LSTM) – особая разновидность архитектуры рекуррентных нейронных сетей (РНС), способная к обучению долговременным зависимостям. Они могут решать ряд разнообразных задач и в настоящее время обширно используются.

LSTM сети были представлены и разработаны З. Хохрайтер и Ю. Шмидхубером в 1997 году специально, для того, чтобы избегать проблем долговременной зависимости. Их специализация – запоминание информации в течение длительных периодов времени, поэтому они не нуждаются в долгом обучении.

Все рекуррентные нейронные сети имеют форму цепочки повторяющихся модулей нейронной сети. В стандартных РНС этот повторяющийся модуль имеет простую структуру, например, один слой \tanh [2].

На приведенной выше диаграмме каждая линия является вектором. Розовые круги обозначают поточечные операции, например, суммирование векторов. Желтые ячейки – это слои нейронной сети. Совмещение линий – это объединение векторов, а знак разветвления – копирование вектора с последующим хранением в разных местах.

Ключевым понятием LSTM является состояние ячейки (горизонтальная линия, проходящая через верхнюю часть диаграммы).

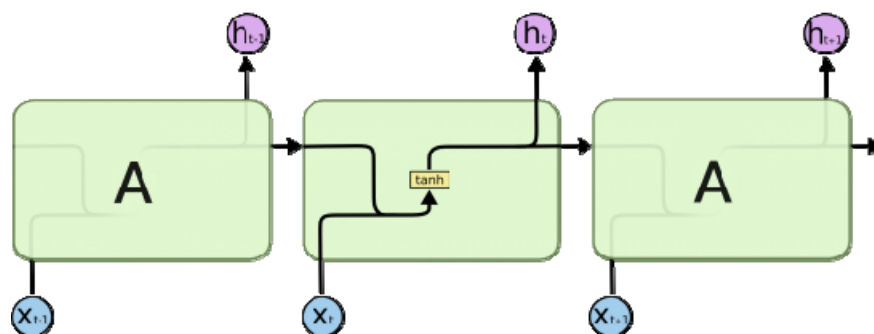


Рис. 1. Структура обычной РНС

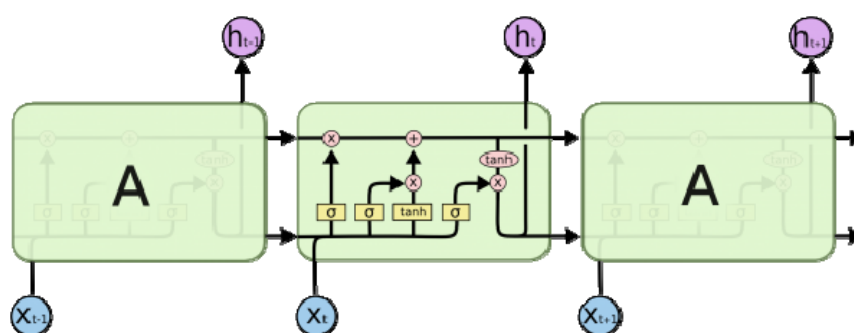


Рис. 2. Структура LSTM

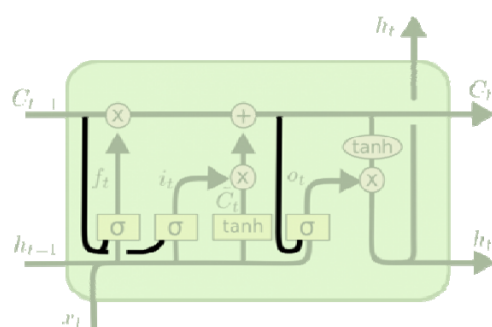
Состояние ячейки похоже на конвейерную линию. Оно проходит через всю цепочку, подвергаясь незначительным линейным преобразованиям.

В LSTM уменьшает или увеличивает количество информации в состоянии ячейки, в зависимости от потребностей. Для этого используются гейты – структуры, которые тщательно настраиваются.

Гейт – это «ворота», пропускающие или не пропускающие информацию. Гейты состоят из сигмовидного слоя нейронной сети и операции поточечного умножения.

Но не все LSTM одинаковы. Существует много вариаций LSTM. Отличия между ними незначительны, но о некоторых из них стоит упомянуть.

Одна из популярных вариаций LSTM, характеризуется добавлением так называемых «смотровых глазков». С их помощью слои фильтров могут видеть состояние ячейки.



$$f_t = \sigma(W_f \cdot [C_{t-1}, h_{t-1}, x_t] + b_f)$$

$$i_t = \sigma(W_i \cdot [C_{t-1}, h_{t-1}, x_t] + b_i)$$

$$o_t = \sigma(W_o \cdot [C_t, h_{t-1}, x_t] + b_o)$$

Рис. 3. Структура LSTM с добавлением «смотровых глазков»

Другие модификации включают объединенные фильтры «забывания» и входные фильтры. В этом случае решения, какую информацию следует забыть, а какую запомнить, принимаются не отдельно, а совместно. Мы забываем какую-либо информацию только тогда, когда необходимо записать что-то на ее место. Мы добавляем новую информацию с состояние ячейки только тогда, когда забываем старую.

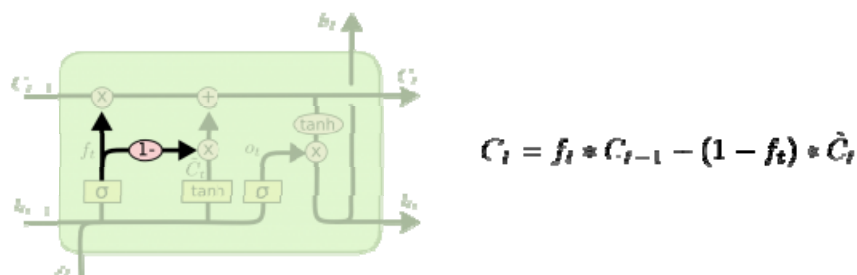


Рис. 4. Структура LSTM с фильтрами «забывания»

Немного больше отличаются от стандартных LSTM управляемые рекуррентные нейроны. В ней фильтры «забывания» и входа объединяют в один фильтр «обновления». Кроме того, состояние ячейки объединяется со скрытым состоянием, есть и другие небольшие изменения. Построенная в результате модель проще, чем стандартная LSTM.

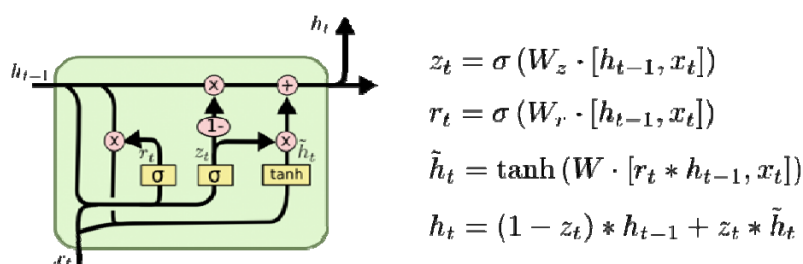


Рис. 5. Структура LSTM с управляемыми рекуррентными нейронами

Это лишь несколько самых примечательных вариаций LSTM. Существует множество других модификаций.

LSTM – большой шаг в развитии РНС. При этом возникает вопрос, каким будет следующий большой шаг. По общему мнению исследователей, следующий шаг заключается в использовании механизма внимания. Идея состоит в следующем: каждый шаг РНС берет данные из более крупного хранилища информации. Например, если мы используем РНС для генерации подписи к изображению, то такая РНС может рассматривать изображение по частям и на основании каждой части генерировать отдельные слова. Последние несколько лет – время расцвета рекуррентных нейронных сетей, и следующие годы обещают принести еще большие плоды.[3]

Библиографические ссылки

1. Brownlee J. Deep learning for time series forecasting: predict the future with MLPs, CNNs and LSTMs in Python. – Machine Learning Mastery, 2018.
2. Understanding LSTM Networks - colah's blog [Электронный ресурс]. URL: <http://colah.github.io/posts/2015-08-Understanding-LSTMs/> (дата обращения: 04.04.2022)
3. Долгая краткосрочная память // Википедия. URL: <https://ru.wikipedia.org/?curid=6768734&oldid=116692933> (дата обращения: 04.04.2022).