RNN LSTM, её устройство и применение для анализа макроэкономических рядов

Зехов Матвей

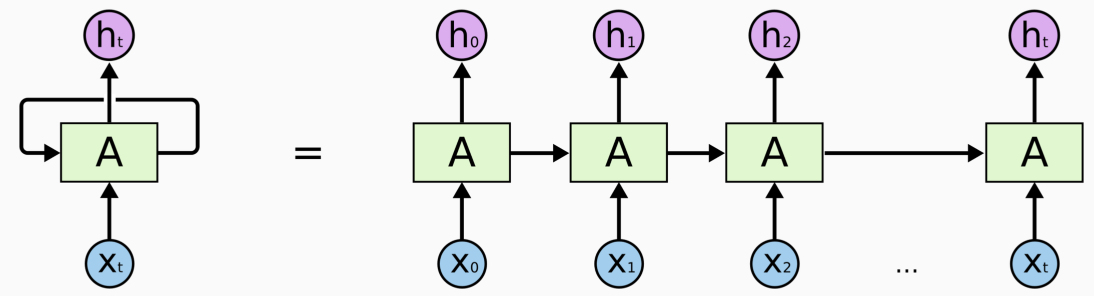
Факультет экономических наук НИУ ВШЭ

Что это такое?

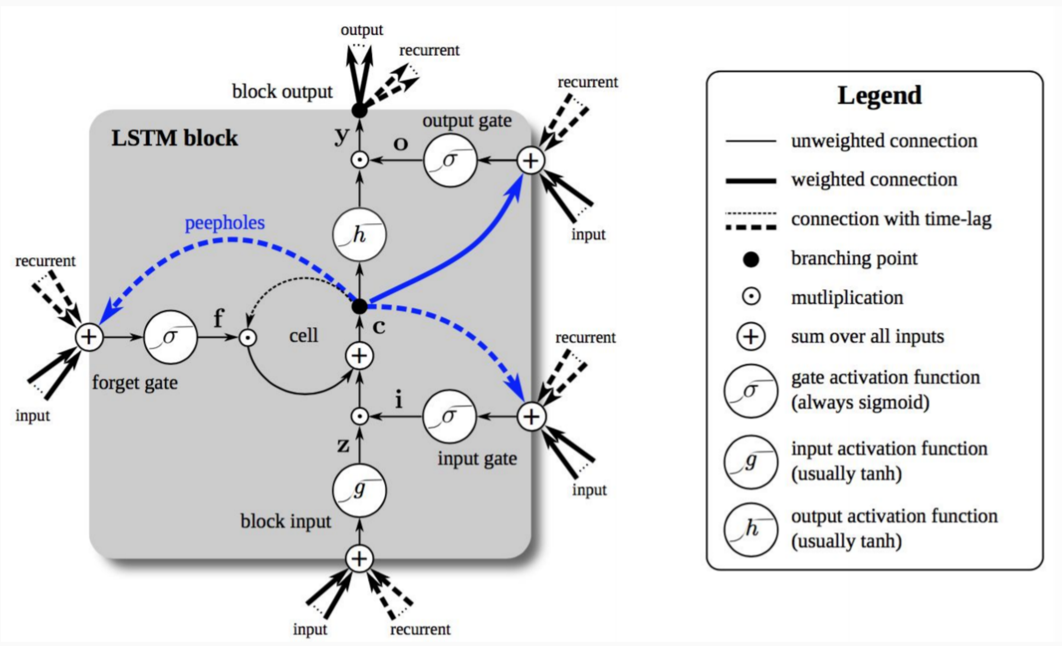
Сети долгой краткосрочной памяти(Long Short-Term Memory) представляют собой специальный подкласс архитектур рекуррентных нейронных сетей, предназначенных для обработки различных форм последовательностей. Они были представлены Зеппом Хохрайтером и Юргеном Шмидхубером в 1997 году. Наиболее яркими примерами применения таких сетей являются проекты по обработке текстов и временных рядов.

Теоретические принципы

Не секрет, что для обработки долговременных зависимостей классические RNN модели слабо подходят, так как их архитектура не подразумевает какого-либо вида памяти. Структура LSTM решает эту проблему.



В том случае, когда расстояние между актуальной информацией и местом, где она пригодилась, невелико, RNN вполне адекватно справляются с задачей, однако при увеличении расстояния эта связь теряется. Ключевым понятием для построения блока является состояние ячейки. Этот вектор проходит через все итерации внутри клетки и несёт в себе информацию о всей последовательности в целом. Схематично блок сети LSTM можно представить следующим образом:



$$f^t = \sigma(W_f x^t + R_f y^{t-1} + b_f) \quad (1)$$

$$i^t = \sigma(W_i x^t + R_i y^{t-1} + b_i) \quad (2)$$

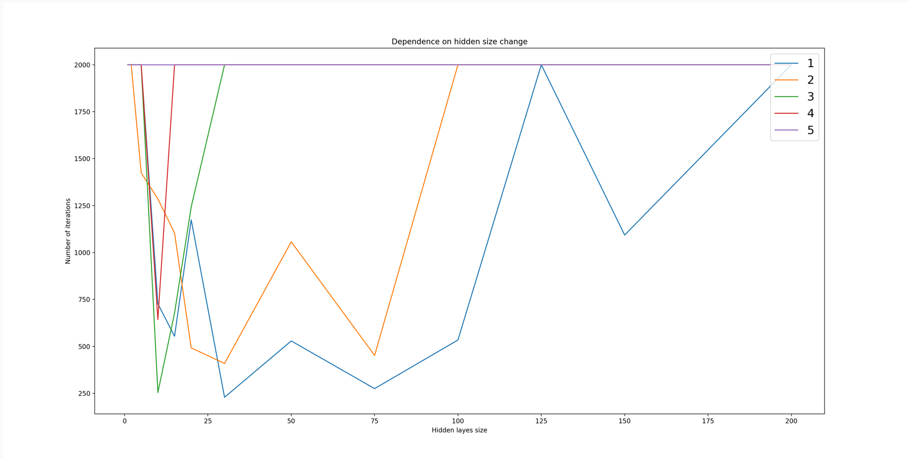
$$z^t = \sigma(W_z x^t + R_z y^{t-1} + b_z) \quad (3)$$

$$C^t = f^t \odot C^{t-1} \oplus i^t \odot z^t \quad (4)$$

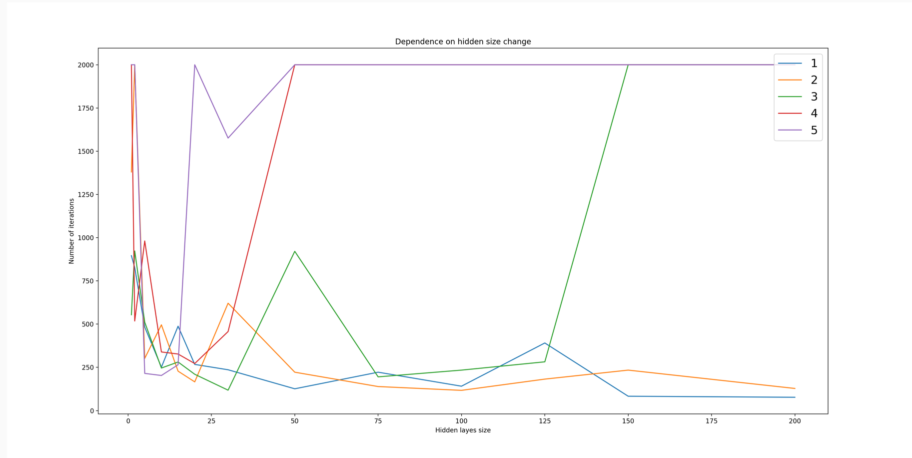
$$o^t = \sigma(W_o x^t + R_o y^{t-1} + b_o) \quad (5)$$

$$y^t = o^t \odot tanh(C^t) \quad (6)$$

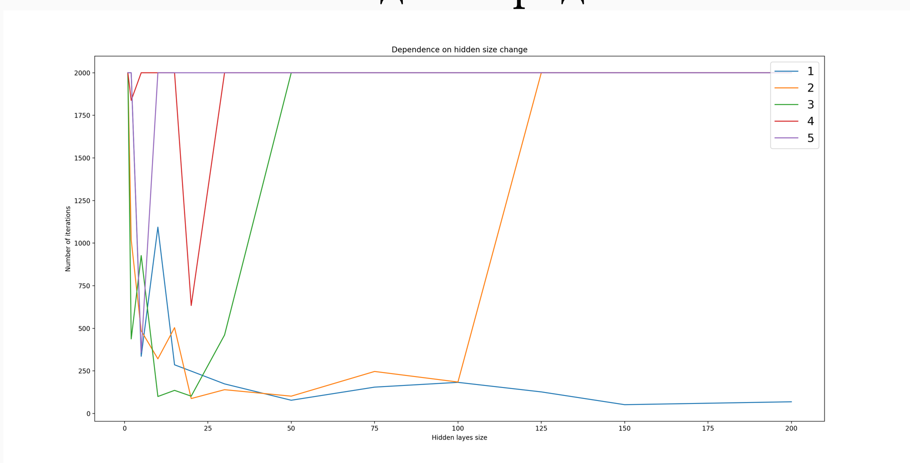
Зависимость сходимости от количества слоёв



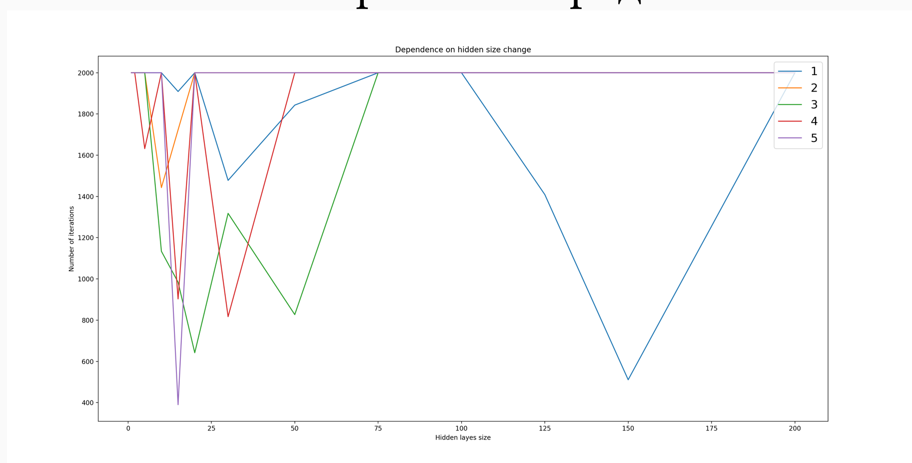
Годовой ряд



Квартальный ряд



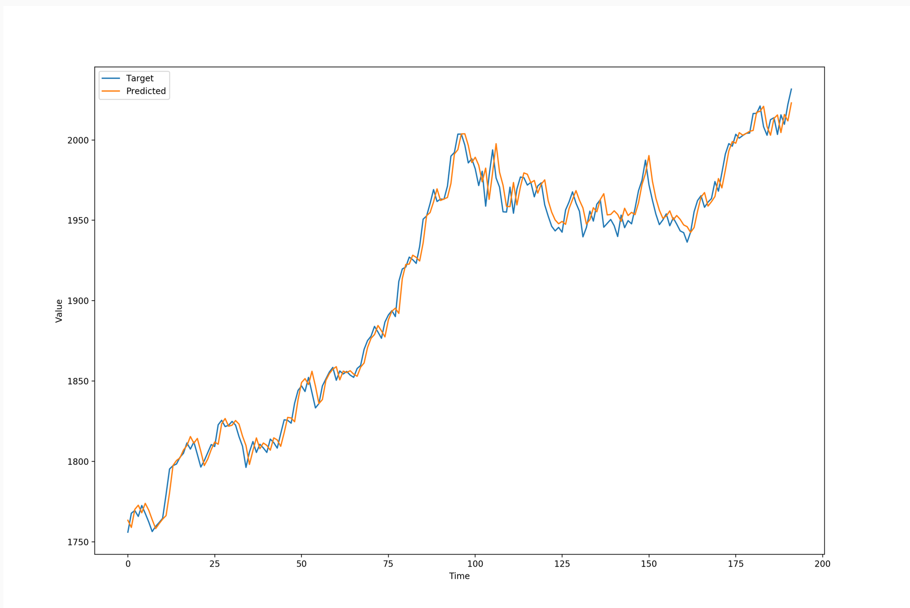
Месячный ряд



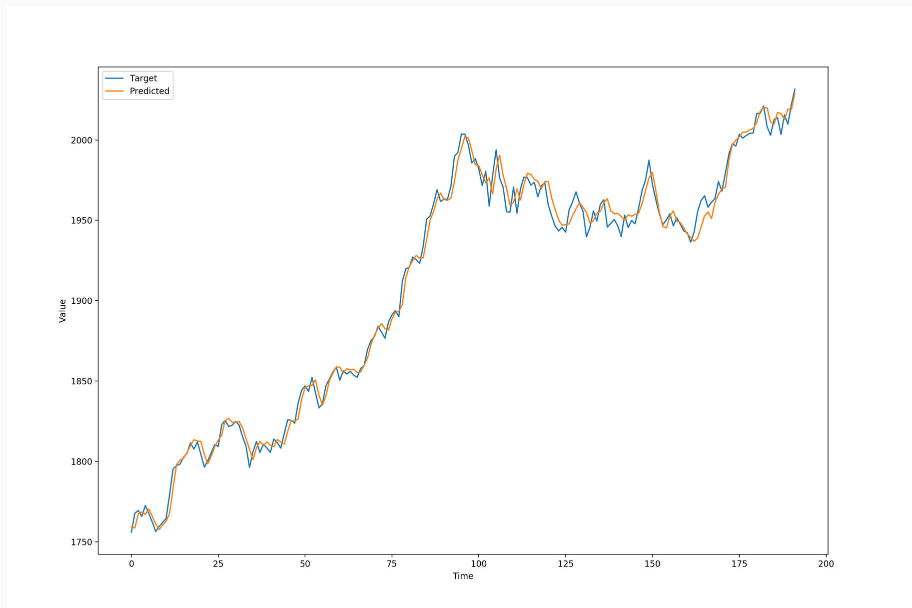
Недельный ряд

Очевидно, что самый эффективный алгоритм - однослойный. При этом тонкая подстройка выявила, что корреляция длины ряда с оптимальной размерностью скрытого слоя равна 0.76!

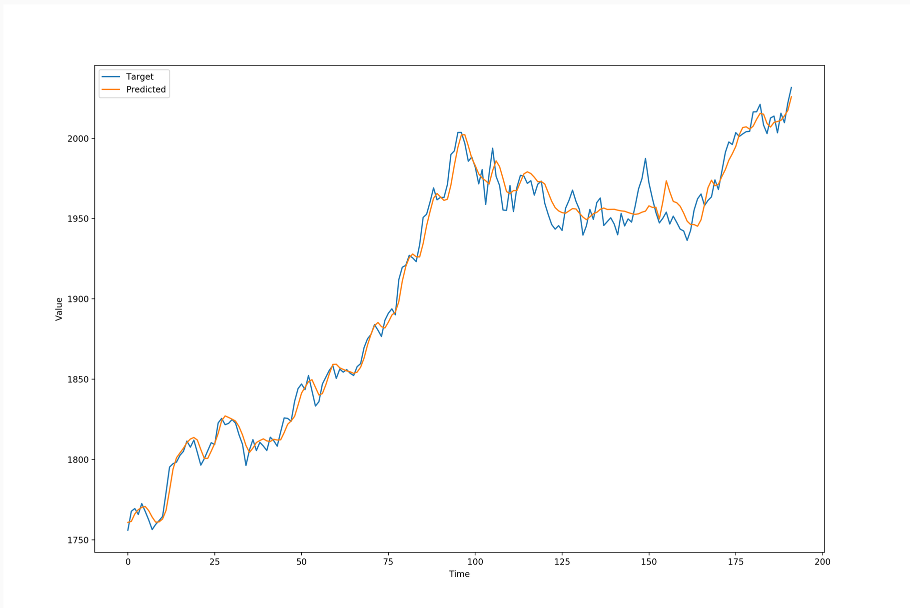
Предсказания методов



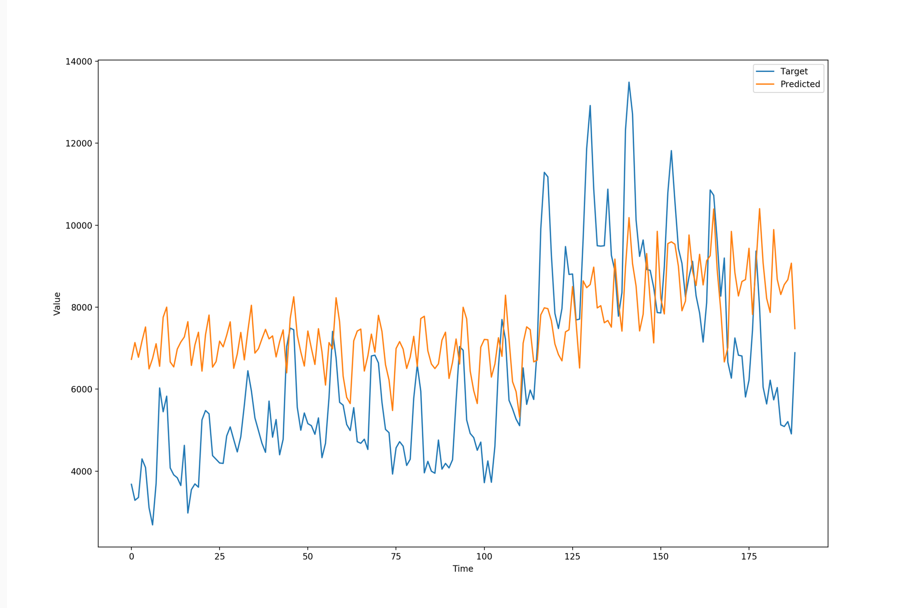
Один пример



50 примеров



Весь сет



Расширяющееся окно

Методы построения модели

При анализе данной архитектуры было использовано два различных метода: расширяющееся окно и сдвигающееся окно.

$$train \longrightarrow \boxed{x_1, x_2, x_3, x_4,}\boxed{x_5, x_6, x_7} \longleftarrow test$$

Получив на основе части train предсказание относительно элемента x_5, сравним его с маркером x_5. После этого перенесём элемент x_5 из test в train:

$$train \longrightarrow \boxed{x_1, x_2, x_3, x_4, x_5,}\boxed{x_6, x_7} \longleftarrow test$$

Метод сдвигающегося окна почти аналогичен, только перемещаются и правая и левая границы окна:

$$\boxed{x_1, x_2, x_3, x_4,}\boxed{x_5, x_6, x_7}$$

$$\boxed{x_1,}\boxed{x_2, x_3, x_4, x_5,}\boxed{x_6, x_7}$$

На графиках слева отображается зависимость сходимости алгоритма расширяющегося окна от количества скрытых слоёв нейросети. По оси абсцисс - количество нейронов скрытого слоя, по оси ординат - итерация, на которой сошёлся алгоритм. Для анализа были взяты шесть рядов различной длины:

Year: 30 | Quarter: 24 | Month: 50 | Week: 70 | Day: 90 | Hour: 100

Анализ предсказаний модели

Обратите внимание на графики слева и таблицу внизу. Ошибка наивного прогноза равна 65.62. Ошибка в скобках - на тренировочных данных.

1. Модель расширяющегося окна явно переобучается. Длины ряда недостаточно, но повысить её нельзя из-за вычислительной сложности.
2. Модель полного градиентного спуска на всём сете эффективна, но не победила наивный прогноз. Кроме того, время исполнения довольно велико.
3. Стохастический спуск по одному примеру эффективен, но подстраивается под наивный прогноз, что очевидно из графика.
4. Алгоритм с пятьюдесятью примерами за итерацию - самый эффективный. Притом он подстраивается не под наивный прогноз, а под реальную картину.

Сравнение показателей

Тип модели	Сдвигающееся окно		Расширяющееся окно	
	Time	MSE	Time	MSE
Один пример	0.0348 (0.3291)	60.4925 (0.5553)	0.0857 (0.3695)	$\geq 10^6$ (≤ 0.01)
50 примеров	0.1132 (0.4313)	43.3558 (0.2564)	— —	— —
Весь сет	0.1226 (0.3590)	69.6659 (16.2584)	— —	— —

Learn more: https://github.com/Pyatachokk/LSTM-guide