

**ПРОГНОЗИРОВАНИЕ ПОТРЕБЛЕНИЯ ПРИРОДНОГО ГАЗА ПРИ ПОМОЩИ
РЕКУРРЕНТНОЙ СЕТИ ДОЛГОЙ КРАТКОСРОЧНОЙ ПАМЯТИ LSTM**

Д. М. ВАСИЛЬЕВА

(Полоцкий государственный университет, Беларусь)

Аннотация. В данной статье представлен анализ результаты исследований по достижению лучших результатов в задаче прогнозирования потребления газа на магистральном газопроводе ОАО «Газпром трансгаз Беларусь» при помощи рекуррентной сети долгой краткосрочной памяти LSTM. Выявлены наиболее эффективное количество эпох, слоев и нейронов для обучения и прогнозирования при помощи данной модели, а также сделаны выводы.

Прогнозирование потребления природного газа является основой не только для планирования закупок у газотрейдеров, но, что не менее важно, для разработки мероприятий по управлению энергопотреблением в муниципальном хозяйстве, особенно в отопительный период. Данные о потреблении природного газа являются последовательностью значений наблюдений какой-либо величины, взятых через относительно равные промежутки времени – временным рядом. Следовательно, для решения задачи прогнозирования потребления природного газа возможно использование инструментов анализа временных рядов в том числе нейронных сетей.

При анализе недельного потребления природного газа (рисунок 1) было выявлено, что тренд данного временного ряда нелинеен и сложно выявляемый. Исходя из этого, было сделано решение использовать рекуррентную нейронную сеть, а именно LSTM, поскольку данная сеть хорошо подходит для классификации, обработки и построения прогнозов на основе временных рядов, где взаимосвязанные явления могут происходить с неопределенным временным лагом. Этот временной лаг приводит к затруднениям в использовании классических нейронных сетей в решении данных проблем из-за затухания градиента, в то время как LSTM сети нечувствительны к величине временного лага.

В LSTM сети существуют 3 типа слоев:

1. Слой «забывания» (Forget gate).
2. Слой «запоминания» (Memory gate).
3. Выходной слой (Output gate).

Сперва «слой фильтра забывания» определяет, какую информацию можно забыть или оставить. Значения предыдущего выхода h_{t-1} и текущего входа x_t

пропускаются через сигмоидальный слой. Полученные значения находятся в диапазоне [0; 1]. Значения, которые ближе к 0 будут забыты, а к 1 оставлены.

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

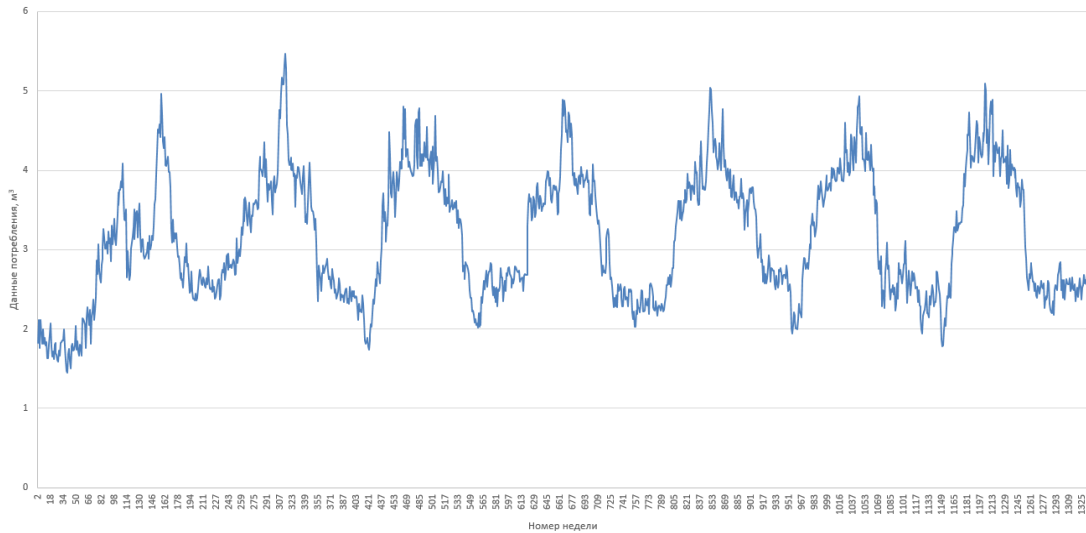


Рисунок 1. – График потребления природного газа

Далее решается, какая новая информация будет храниться в состоянии ячейки. Сначала сигмоидальный слой под названием «слой входного фильтра» определяет, какие значения следует обновить. Затем tanh-слой строит вектор новых значений-кандидатов \tilde{c}_t , которые можно добавить в состояние ячейки.

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i); \quad (2)$$

$$\tilde{c}_t = \tanh(W_c[h_{t-1}, x_t] + b_c). \quad (3)$$

Для замены старого состояния ячейки C_{t-1} на новое состояние C_t . Необходимо умножить старое состояние на f_t , значения оставленные «слоем фильтра забывания». Затем прибавляем $i_t * \tilde{c}_t$. умноженное значение слоев входного фильтра и тангенс-слоя. Это новые значения-кандидаты.

На последнем этапе определяется то, какая информация будет получена на выходе. Выходные данные будут основаны на нашем состоянии ячейки, к ним будут применены некоторые фильтры. Сначала значения предыдущего выхода h_{t-1} и текущего входа пропускаются через сигмоидальный слой, который решает, какая информация из состояния ячейки будет выведена. Затем значения состояния ячейки проходят через tanh-слой, чтобы получить на выходе значения из диапазона от -1 до 1, и перемножаются с выходными значениями сигмоидального слоя, что позволяет выводить только требуемую информацию.

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o); \quad (4)$$

$$h_t = o_t * \tanh(C_t), \quad (5)$$

где x_t – входной вектор;
 h_t – выходной вектор;
 C_t – вектор состояний;
 W и b – матрицы параметров и вектор;
 f_t – вектор вентилля забывания, вес запоминания старой информации;
 i_t – вектор входного вентилля, вес получения новой информации;
 o_t – вектор выходного вентилля, кандидат на выход.

Полученные таким образом h_t и C_t передаются далее по цепочке.

В качестве исходных данных были использованы тренды 220 потребителей природного газа Республики Беларусь за последние 8 лет (2012–2020гг.). Причем данные 2012 г. появляются с мая и по 2020 г. неполны.

Все тренды обладают следующими особенностями:

- в данных наблюдается присутствие ошибок телеметрии, выбросов и нулей, связанных с временным выходом из строя датчиков расхода, или давления и температуры, при вычислении расхода расчетным путем;
- есть пропуски данных на периоды плановых и экстренных ремонтов газопровода;
- отсутствует точное выравнивание меток времени;
- данные отражают не факт потребления газа, а режим работы газотранспортной сети по доставке контрактного газа потребителю, соответственно отражает и нестационарные процессы (закачка газа в систему, разбор газа из системы в режиме баллона, когда движения газа через ГРС или компрессорную станцию не происходит, а потребление газа продолжается);
- расходы газа, фиксируемые датчиками, лишены информации о плотности газа;
- часть расходов получено расчетным путем по стационарной неизотермической модели при граничных условиях, заданных через давление и температуру;

Нейронная сеть реализована на языке программирования Python, с использованием библиотек для обработки и визуализации данных:

- pandas – для обработки и анализа данных;
- numpy – для поддержки многомерных массивов (включая матрицы) и высокоуровневых математических функций, предназначенных для работы с многомерными массивами;
- matplotlib – для визуализации данных двумерной (2D) графикой;
- keras (в качестве основы tensorflow) – для глубокого обучения, представляющая из себя высокоуровневый API;
- sklearn – для решения задач классического машинного обучения.

В качестве функции потерь в процессе обучения используется среднеквадратическая ошибка (Mean Squared Error), оптимизация осуществляется с использованием алгоритм Adam.

Исследования поведутся с учетом поиска максимально эффективного количества нейронов и слоев, а также эпох и временных шагов для обучения нейронной сети.

Спрогнозированы последние 30 недель. Качество обученной сети оценено по результатам ее работы на контрольной выборке и показателем RMSE. RMSE (от англ. Root Mean Square Error, RMS Error — среднеквадратичная ошибка) используется для оценки качества предсказания модели. Он измеряет разницу между истинным и предсказанным значениями.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Predicted_i - Actual_i)^2}{N}}. \quad (6)$$

Всего было проведено 20 экспериментов, часть результатов которых представлены в таблице 1.

Таблица 1. – Значения параметров модели и их среднеквадратическая ошибка

№	Кол-во эпох	Кол-во слоев	Кол-во нейронов	RMSE
1	50	18	50	0,18568
2	250	18	50	0,16447
3	200	18	50	0,14165
4	200	19	50	0,17607
5	200	17	50	0,15736
6	200	14	50	0,10812
7	200	13	50	0,14432
8	200	14	45	0,16866
9	200	14	55	0,16743
10	200	6	50	0,12521
11	200	5	50	0,13874

Актуальным является поиск максимально эффективного количества эпох, нейронов и слоев модели. Первоначально используется модель LSTM состоящая из 18 слоев по 50 нейронов каждый, обученная на 50 эпохах с 30 временными шагами. Обучение производится итеративно в течение 10 эпох для отслеживания возможного переобучения. Среднеквадратическая ошибка такой модели (RMSE) составляет 0,18568, из чего следует, что, хоть модель и имеет небольшие потери, результат все равно близок к реальным значениям. С каждой эпохой потери становились все меньше, поэтому было сделано решение увеличить количество эпох еще на 200, чтобы увидеть, когда модель начнет переобучаться. В результате чего было выяснено, что максимально эффективным количеством эпох для обучения

является 200, т. к. со всеми последующими эпохами количество потерь постепенно начинает увеличиваться. Потери при обучении на 200 эпохах составили 0,14165. При поиске максимально эффективного количества слоев в модели необходимо уменьшить или увеличить количества слоев на 1 и более (если необходимо). Наилучшие результаты были достигнуты при 14 слоев, потери составили 0,10812 (рисунок 2, таблица 2).

При попытке улучшить результаты прогнозирования при помощи уменьшения или увеличения количества нейронов в слое потери незначительно, но возросли, следовательно, 50 нейронов самое оптимальное количество.

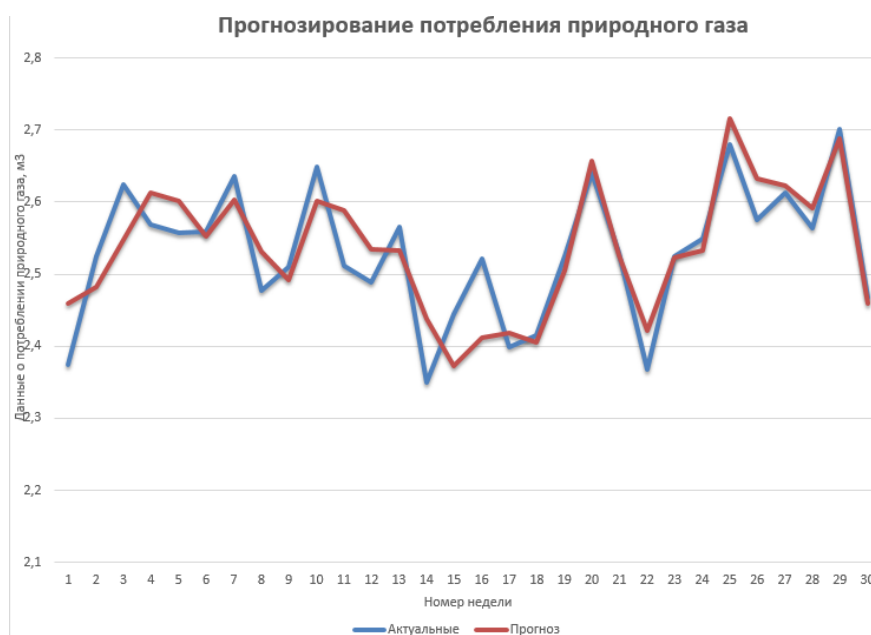


Рисунок 2. – Актуальные и спрогнозированные при помощи модели с максимально эффективными параметрами данные

Таблица 2. – Актуальные и спрогнозированные при помощи модели с максимально эффективными параметрами данные

№	Актуальные	Прогноз	Актуальные	Прогноз
	2	3	4	5
1	2,37426	2,4589016	2,44508	2,3727708
2	2,52446	2,4824038	2,52188	2,4118817
3	2,62482	2,5477538	2,39877	2,4186335
4	2,56934	2,6133864	2,41574	2,4046175
5	2,55824	2,6010296	2,52501	2,5048695
6	2,55921	2,5528083	2,64082	2,6568123
7	2,63618	2,6033185	2,53041	2,5258038
8	2,47801	2,531796	2,36851	2,4212902
9	2,51073	2,4918282	2,52529	2,5233703
10	2,64925	2,6021955	2,54906	2,5332465

Окончание таблицы 2

1	2	3	4	5
11	2,51208	2,5882387	2,67965	2,7166963
12	2,48935	2,533869	2,57579	2,631923
13	2,56598	2,5321705	2,6135	2,622738
14	2,3492	2,438455	2,56379	2,5916238
15	2,37426	2,4589016	2,70191	2,687907

Заключение. При анализе результатов экспериментов, представленных в статье, можно заметить, что прогностическая кривая отображает динамику потребления природного газа и повторяет все его изменения. Уровни ошибки на тренировочной и валидационной выборке показывают, что модель несколько хуже прогнозирует реальные данные, это может быть обусловлено изменяющимися процессами в отложенных данных. В то же время усложнение модели и создание агрегирующих слоев позволяет не использовать большие мощности для оптимизации на 200 и более эпох. Исходя из этого, можно сделать вывод, что рекуррентные нейронные сети долгой краткосрочной памяти LSTM отлично подходят для фиксирования временной динамики данных, и не имеют проблемы исчезающих и взрывных градиентов, следовательно, могут быть использованы для прогнозирования потребления каких-либо природных ресурсов, включая потребление природного газа.

ЛИТЕРАТУРА

1. Раздел «Интервальное прогнозирование временных рядов с помощью рекуррентных нейронных сетей с долгой краткосрочной памятью» на сайте Habr [Электронный ресурс]. – Режим доступа: <https://habr.com/ru/post/505338/>. – Дата доступа: 16.11.2020.
2. УДК 001.891.573 ОТЧЕТ О НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ «Разработка методов анализа больших данных для решения задачи прогнозирования потребления газа на магистральном газопроводе ОАО "Газпром трансгаз Беларусь"». – Дата доступа: 01.03.2022.
3. Раздел «Time Series Prediction with LSTM Recurrent Neural Networks in Python with Keras» на сайте Machine Learning Mastery [Электронный ресурс]. – Режим доступа: <https://machinelearningmastery.com/time-series-prediction-lstm-recurrent-neural-networks-python-keras/>. – Дата доступа: 07.03.2022.
4. Раздел «Анализ временных рядов с помощью python» на сайте habr.com [Электронный ресурс]. – Режим доступа: <https://habr.com/ru/post/207160/>. – Дата доступа: 14.03.2021.
5. УДК 330.4:004.032.26 – Прогнозирование Финансовых Временных Рядов С Использованием Рекуррентных Нейронных Сетей Lstm. – Дата доступа: 16.03.2022.