



МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
Национальный исследовательский
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

ВСЕ ГРАНИ МАТЕМАТИКИ И МЕХАНИКИ

**Сборник статей
Всероссийской молодежной
научной конференции студентов**

Томск, 27 мая – 1 июня 2021 г.



ТОМСК
«Издательство НТЛ»
2021

Прогноз приземной температуры воздуха на основе модели рекуррентной нейронной сети

И.В. Дель, А.В. Старченко

*Национальный исследовательский
Томский государственный университет, г. Томск, Россия*

В данной работе реализована рекуррентная нейронная сеть типа Long Short-Term Memory для прогнозирования метеорологических величин. На основе известного распределения метеорологических значений ставилась задача прогнозировать значения приземной температуры воздуха. Общая среднеквадратическая ошибка прогноза составила 0.12 С. Сравнение результатов прогнозирования рекуррентной нейронной сети с результатами многовариантной и одновариантной модели многослойного перцептрона показало преимущество использования LSTM-сети. Сравнение результатов прогнозирования рекуррентной нейронной сети типа LSTM с результатами прогнозирования численной модели Weather Research and Forecasting показало перспективность использования нейронных сетей для предсказания метеорологических параметров.

Ключевые слова: *прогнозирование, нейронная сеть, RNN, LSTM.*

Опасные метеорологические условия оказывают серьезное воздействие на различные сферы человеческой жизни, такие, как экономическая и хозяйственная. Во избежание финансовых потерь требуется своевременное, а также точное предсказание метеорологических величин. Например, для сельского хозяйства, которое напрямую зависит от метеорологических параметров и для которого прогноз заморозков, аномально холодных/теплых температур, обильных осадков имеет первоочередное значение. А опасные метеорологические явления, такие, как ураганный ветер, могут повлечь за собой человеческие жертвы. Ответственность может извлечь выгоду из своевременных прогнозов метеорологических величин.

В настоящее время прогностические модели все время совершенствуются, повышается точность прогнозирования различных процессов. Таким образом, тема моделирования метеорологических величин будет оставаться актуальной.

Существуют различные способы прогноза метеорологических параметров: синоптическое прогнозирование, численный и статистический

методы прогноза, методы машинного обучения и нейронные сети. Чаще всего используются численные методы прогноза погоды, но методы нейросетевого прогнозирования вполне могут конкурировать со стандартными методами прогнозирования, так как это локальный, менее затратный способ прогнозирования метеорологических параметров.

Целью данной работы является реализация и применение модели рекуррентной нейронной сети для прогноза приземной температуры воздуха. По измеренным стационарной метеостанцией историческим значениям давления, температуры и относительной влажности, скорости и направления ветра поставлена задача предсказать значения приземной температуры воздуха и реализовать сравнение результатов прогнозирования с результатами прогнозирования модели многослойного перцептрона.

Архитектура рекуррентных нейронных сетей

Разработанная модель искусственной нейронной сети (ИНС) основана на одной из архитектур нейронных сетей, называемой рекуррентной нейронной сетью типа LSTM.

Рекуррентные нейронные сети – нейронные сети с обратной связью между различными слоями нейронов. Их характерная особенность – передача сигналов с выходного или скрытого слоя во входной слой. Рекуррентная нейронная сеть может состоять из любого числа слоев [1, 2].

Для обучения рекуррентных нейронных сетей применяется алгоритм обратного распространения ошибки по времени, который является вариантом алгоритма обратного распространения ошибки, используемый для нейронных сетей прямого распространения сигнала. Данный алгоритм определяет стратегию подбора весов нейронной сети с применением градиентных методов оптимизации и также решается задача минимизации суммарных потерь на всех данных из обучающей выборки.

Обучение сети с использованием алгоритма обратного распространения ошибки проходит следующим образом. На каждом шаге обучения рекуррентной нейронной сети предъявляется некоторый входной вектор x , для которого известно ожидаемое выходное значение y . Для этого входного вектора рассчитываются веса нейронов скрытых слоев, далее – веса нейронов выходного слоя.

Выходной вектор скрытого слоя рекуррентной нейронной сети на каждом i -м шаге обучения вычисляется следующим образом:

$$t_i = \sigma_t(W_t \cdot x_i + U_t \cdot t_{i-1} + b_t),$$

где σ_t – функция активации скрытого слоя; x_i – входной вектор на i -м шаге; W_t – матрица весов для входного вектора x_i ; U_t – матрица весов для нейронов скрытого слоя; b_t – вектор смещения скрытого слоя.

В простой рекуррентной нейронной сети выходной вектор вычисляется следующим образом:

$$h_i = \sigma_h(W_h \cdot t_i + b_h),$$

где W_h – матрица весов для нейронов выходного слоя; b_h – вектор весов для нейронов выходного слоя; σ_h – функция активации для выходного слоя.

На рис. 1 представлена простая рекуррентная нейронная сеть в развернутом виде.

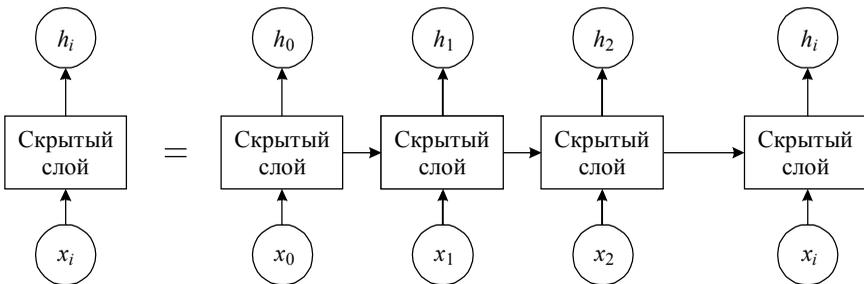


Рис. 1. Архитектура простой рекуррентной нейронной сети

Long short-term memory (LSTM) – тип рекуррентных нейронных сетей с долгой кратковременной памятью. Все рекуррентные нейронные сети представляют собой цепочки повторяющихся модулей нейронной сети. LSTM специально разработаны для устранения проблемы долгосрочной зависимости. Их специализация – запоминание информации в течение длительных периодов времени, поэтому их практически не нужно обучать [3].

Схематично архитектура LSTM сети представлена на рис. 2.

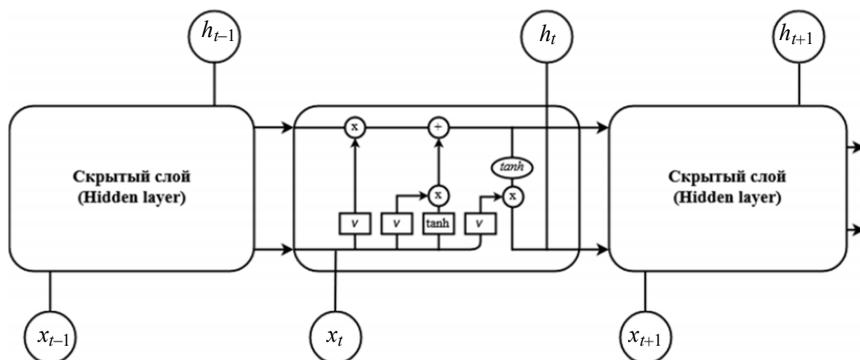


Рис. 2. Архитектура рекуррентной сети типа LSTM

На приведенном рис. 2 каждая линия является вектором. Круг означает поточечные операции, например суммирование векторов. Под прямоугольными ячейками понимаются слои нейронной сети (повторяющийся модуль LSTM состоит из четырех взаимодействующих слоев). Совмещение линий – это объединение векторов, а разветвления – копирование вектора с последующим хранением в разных местах.

Ключевым компонентом сети LSTM является состояние ячейки, которому сопоставляется горизонтальная линия в верхней части центрального блока. Оно участвует в некоторых линейных преобразованиях и отвечает за процесс обучения, обратное распространение ошибки и обновление весов. В LSTM можно уменьшать или увеличивать количество информации в состоянии ячейки при помощи структур, называемых гейтами. Гейт – это «ворота» либо пропускающие, либо не пропускающие информацию [3].

Гейты состоят из сигмоидального слоя (то есть предполагают использование сигмоидальной функции активации) и операции поточечного умножения. Сигмоидальный слой возвращает число в диапазоне от «0» до «1», определяющее, какую долю каждого блока информации следует пропустить дальше по сети. Число «0» означает «не пропускать ничего», а число «1» – «пропустить все». LSTM имеет три таких гейта для контроля состояния ячейки [3].

На первом этапе сеть решает, какую информацию можно выбросить из состояния ячейки. Это решение принимается сигмоидальным слоем, называемым «слоем гейта утраты». Он выдает число от 0 до 1 для каж-

дого номера в состоянии ячейки; 1 означает «полностью сохранить», а 0 – «полностью удалить» [3].

Следующий шаг – решить, какая новая информация будет храниться в состоянии ячейки. Этот этап состоит из двух частей. Сначала сигмоидальный слой, называемый «слоем гейта входа», решает, какие значения требуется обновить. Затем слой \tanh создает новый вектор состояния ячеек [3].

На третьем шаге LSTM-сеть изменяет старое состояние ячейки на новое состояние.

Наконец, нужно решить, что требуется получить на выходе. Результатом будет являться отфильтрованное состояние ячейки. Сначала используется сигмоидальный слой, который решает, какие части состояния ячейки выводить. Затем состояние ячейки пропускается через \tanh (чтобы разместить все значения в интервале $[-1, 1]$) и умножается на выходной сигнал сигмоидального гейта [3].

Построение модели рекуррентной нейронной сети типа LSTM

Задача состоит в предсказании приземной температуры воздуха на 1 час.

Исходные метеорологические данные для исследования были получены в ЦКП «Атмосфера» Института оптики атмосферы СО РАН за период времени с 01.01.2020 по 30.10.2020. Размер набора данных 6336 ежечасных наблюдений.

После сбора данных проводится процедура предварительной обработки данных: восстановление пропусков. Отсутствующие данные восстанавливаются с помощью линейной интерполяции. Затем все параметры нормализуются. Далее выбранные данные разделяются на две выборки: обучающую выборку, соответствующую 80 % данных от основной выборки, и тестовую выборку, соответствующую 20 %, чтобы можно было проверить способность сети к прогнозированию после фазы обучения.

Формой ввода будет 1 временной шаг с 8 параметрами. Скрытый слой рекуррентной нейронной сети LSTM будет содержать 50 нейронов. Выходной слой – 1 нейрон [4].

В этой нейронной сети осуществляется оптимизация целевой функции с подбором весовых коэффициентов методом адаптивной инерции Adam (Adaptive Moment Estimation) (вариант стохастического градиентного спуска) [5].

Метрикой для рассматриваемой задачи в данной работе будет RMSE, рассчитываемая для каждого прогнозного временного шага.

За функцию потерь принята среднеквадратичная ошибка (MSE), поскольку она соответствует масштабу принятой ранее метрике (RMSE):

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - d_i)^2,$$

где y_i – предсказанное выходное значение; d_i – реальное выходное значение. Функция потерь используется для расчета ошибки между реальными и полученными значениями. Глобальная цель – минимизировать эту ошибку.

Сеть обучалась за фиксированное количество эпох (epochs = 100). Количество эпох показывает, сколько раз модель подвергается воздействию обучения. Число эпох выбрано экспериментальным путем.

Степень близости графиков на рис. 3 свидетельствует об удовлетворительной настройке сети.

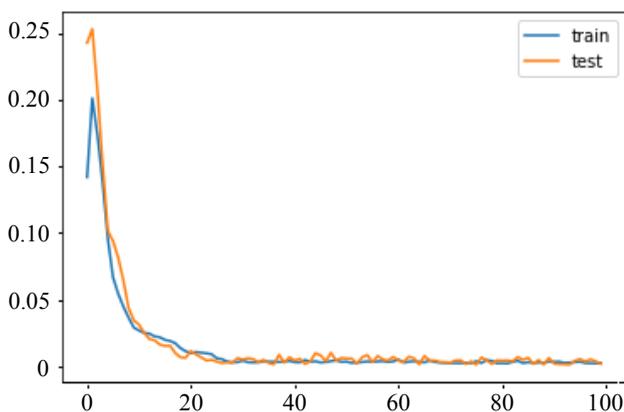


Рис. 3. Потери на этапах обучения и проверки

Результаты прогнозирования

Общее значение RMSE по всему прогнозу составила 0.115 С. Результат более чем удовлетворительный. Ошибка довольно мала.

Сравнение результатов прогнозирования двух типов нейронных сетей

Сравним значение ошибки LSTM-сети с ошибками одновариантного и многовариантного многослойного персептрона.

Многослойный персептрон – это класс искусственных нейронных сетей, состоящих из нескольких уровней персептронов (искусственных нейронов) с пороговой активацией. Для получения значений весовых коэффициентов используется метод обучения с учителем, называемым обратным распространением ошибки [2].

Общее значение ошибки модели многовариантного многослойного персептрона составила 0.186 С.

Общее значение ошибки модели одновариантного многослойного персептрона составила 3.12 С.

При сравнении рис. 4, *a–в*, становится ясно, что LSTM-сеть лучше всего справилась с задачей прогнозирования. Одновариантный персептрон показал худший результат. Причина кроется в том, что прогноз производился на 24 часа с использованием только значений приземной температуры воздуха.

Сравнение результатов прогнозирования рекуррентной нейронной сети с прогнозом численной модели WRF

WRF (Weather Research and Forecasting) – численная модель прогноза погоды, система, предназначенная как для проведения атмосферных исследований, так и для оперативного прогнозирования [6].

WRF – набор программ, реализующих расчёт модели. WRF поддерживает два вида динамических решателей, систему на ассимиляции данных и программные компоненты для обеспечения параллельных вычислений и расширяемости системы. Модель WRF имеет широкий спектр применений в метеорологии и используется в масштабах от метров до тысяч километров.

Из рис. 5 можно сделать вывод, что прогноз почасовой температуры за 15.09.20, полученный при помощи модели рекуррентной нейронной сети типа LSTM, является куда более точным, чем прогноз численной модели WRF.

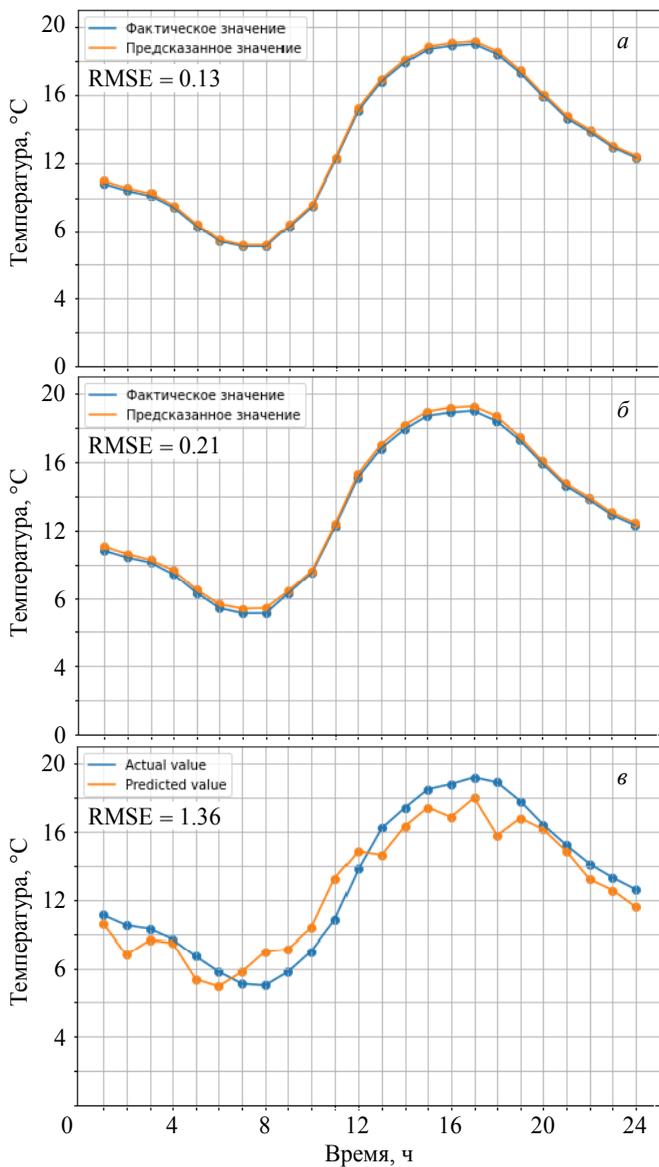


Рис 4. Сравнения прогнозируемой температуры при помощи LSTM-сети (*а*), многовариантного многослойного перцептрона (*б*) и одновариантного многослойного перцептрона (*в*) с наблюдениями за 15.09.20

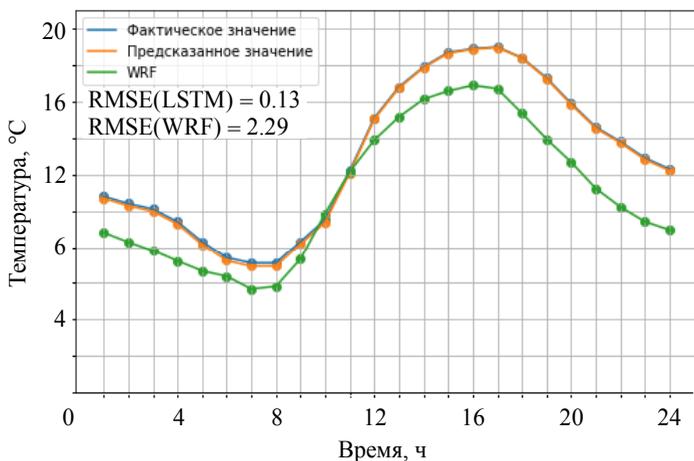


Рис. 5. Сравнение результатов прогноза LSTM-сети с прогнозом численной модели WRF

Заключение

В данной работе была изучена и разработана модель рекуррентной нейронной сети LSTM с дальнейшим ее применением для прогноза приземной температуры воздуха в городе Томске. По измеренным стационарной метеостанцией историческим значениям давления, температуры и относительной влажности, скорости и направления ветра предсказывалось значение приземной температуры воздуха на один час с использованием библиотеки Keras языка программирования Python.

Сравнительный анализ реальных данных и результатов, полученных с помощью модели рекуррентной нейронной сети LSTM, свидетельствует о том, что метод нейросетевого прогнозирования может конкурировать со стандартными методами прогнозирования.

СПИСОК ЛИТЕРАТУРЫ

1. Hochreiter S., Schmidhuber J. Long Short-Term Memory // Neural Computation. 1997. V. 9. No. 8. P. 1735–1780.
2. Хайкин С. Нейронные сети. Полный курс. М: Вильямс, 2008. 1104 с.
3. Гафаров Ф.М., Галимянов А.Ф. Искусственные нейронные сети и их приложения. Казань: Изд-во Казан. ун-та, 2018. 121 с.
4. Jason Brownlee. Deep Learning for Time Series Forecasting. Predict the Future with MLPs, CNNs and LSTMs in Python. 2018. 557 p.

5. *Diederik P. Kingma, Jimmy Ba. Adam. A Method for Stochastic Optimization // ICLR 2015. 2015.*
6. *Powers J.G., Klemp J.B., Skamarock W.C., et al. The weather research and forecasting model: Overview, system efforts and future directions // Bull. Am. Meteor. Soc. 2017. V. 98. P. 1717–1737.*

Дель Ирина Васильевна, магистрантка ТГУ; irina.del@mail.tsu.ru

Старченко Александр Васильевич, д.ф.-м.н., проф., зав. кафедрой вычислительной математики и компьютерного моделирования ММФ ТГУ; starch@math.tsu.ru