

УПРАВЛЕНИЕ СЛОЖНЫМИ СИСТЕМАМИ

DOI: 10.18137/RNU.V9I187.24.04.P.64

УДК 004.852

М.В. Раскатова, Е.Е. Терелецкова, А.А. Сало, Э.А. Чельшев

ПРОГНОЗИРОВАНИЕ МЕТЕОРОЛОГИЧЕСКИХ ПОКАЗАТЕЛЕЙ С ИСПОЛЬЗОВАНИЕМ НЕЙРОСЕТЕВЫХ ТЕХНОЛОГИЙ

Аннотация. В статье рассматривается применение нейросетевых технологий для прогнозирования метеорологических показателей в Москве. Актуальность исследования обусловлена необходимостью повышения точности и надежности прогнозов температуры, что становится особенно важным в условиях изменения климата и роста экстремальных погодных явлений. Методология включает сбор и обработку метеорологических данных, настройку и обучение нейросетевой модели с помощью языка программирования Python и его библиотек. В статье представлена работоспособная модель прогнозирования метеорологических показателей, верифицированная с использованием следующих метрик: средняя абсолютная ошибка (MAE) и коэффициент детерминации (R^2).

Ключевые слова: нейросетевая модель, прогнозирование, анализ данных, метеорология, погода.

M.V. Raskatova, E.E. Tereletsikova, A.A. Salo, E.A. Chelyshev

FORECASTING METEOROLOGICAL INDICATORS USING NEURAL NETWORK TECHNOLOGIES

Abstract. The article discusses the use of neural network technologies for forecasting meteorological indicators in Moscow. The relevance of the study is grounded on the need to improve the accuracy and reliability of temperature forecasts, which is becoming especially important in the context of climate change and the growth of extreme weather events. The methodology includes collecting and processing meteorological data, setting up and training a neural network model using the Python programming language and related libraries. As a final result, the article provides a working model for forecasting meteorological indicators, tested on the average absolute error and the determination coefficient.

Keywords: neural networks, forecasting, data analysis, meteorology, weather.

Введение

Москва как один из крупнейших мегаполисов мира сталкивается с уникальными климатическими вызовами, что делает эффективное прогнозирование погоды особенно актуальным¹.

Традиционные методы прогнозирования, разработанные на статистических моделях и численных методах, имеют свои ограничения, особенно в условиях быстрых изменений метеорологических условий. В последние годы нейросетевые технологии, основанные на принципах глубокого обучения, продемонстрировали значительные успехи в различных областях, включая метеорологию. Эти методы способны обрабатывать большие объемы данных и выявлять сложные паттерны, что открывает новые перспективы для повышения точности прогнозов погоды.

¹ Миронов Н. Столица субтропиков: как изменится климат в Москве в ближайшие десятилетия // Вечерняя Москва. 2024. 22 июля. URL: <https://vm.ru/moscow/1150643-stolica-subtropikov-kak-izmenitsya-klimat-v-moskve-v-blizhajshie-desyatiletija> (дата обращения: 22.10.24).

Раскатова Марина Викторовна

кандидат технических наук, доцент кафедры вычислительных машин, систем и сетей, Национальный исследовательский университет «МЭИ», Москва. Сфера научных интересов: нейронные сети, вычислительная техника, прогнозирование временных рядов, сравнительный анализ моделей прогнозирования. Автор более 100 опубликованных научных работ. ORCID: 0000-0001-7671-3312, SPIN-код: 8053-5041, AuthorID: 609945.

Электронный адрес: RaskatovaMV@mpei.ru

Терелецкова Есения Евгеньевна

студент, Национальный исследовательский университет «МЭИ», Москва. Сфера научных интересов: цифровое общество, экономическое моделирование, нейронные сети, методы прогнозирования. Автор пяти опубликованных научных работ. ORCID: 0009-0003-6409-3939.

Электронный адрес: esenia2018.t@gmail.com

Сало Анна Андреевна

студент, Национальный исследовательский университет «МЭИ», Москва. Сфера научных интересов: нейросетевое прогнозирование, эконометрические модели, Python. Автор одной опубликованной научной работы. ORCID: 0009-0009-2640-8614.

Электронный адрес: anyasalo@yandex.ru

Чельшев Эдуард Артурович

аспирант, Национальный исследовательский университет «МЭИ», Москва. Сфера научных интересов: машинное обучение по прецедентам, искусственные нейронные сети, машинная обработка текстов на естественном языке, C++. Автор более 20 опубликованных научных работ. ORCID: 0000-0001-8417-8823, SPIN-код: 5357-7604, AuthorID: 1088395.

Электронный адрес: chel.ed@yandex.ru

Нейронная сеть

Нейронные сети являются мощным инструментом для анализа данных, обнаружения паттернов и автоматизации процессов в различных областях [1]. Профессор Саймон Хайкин в своей книге «Нейронные сети» отмечает, что «нейронная сеть – это громадный распределенный параллельный процессор, состоящий из элементарных единиц обработки информации, накапливающих экспериментальные знания и предоставляющих их для последующей обработки» [2, с. 32].

Учитывая мнения различных ученых, можно представить следующую классификацию нейронных сетей:

- «... по характеру обучения – нейронные сети, использующие обучение с учителем и без учителя» (Ф.М. Гафаров и А.Ф. Галимянов [3, с. 12]);
- «... по характеру настройки синапсов – сети с фиксированными связями. В этом случае весовые коэффициенты нейронной сети выбираются сразу исходя из условия задачи; сети с динамическими связями. Для них в процессе обучения происходит настройка синаптических связей» (В.А. Головкин и В.В. Краснопрошин [4, с. 24]);
- «три основных типа нейронных сетей:
 - полносвязные нейронные сети – каждый нейрон передает свой выходной сигнал остальным нейронам, в том числе и самому себе;
 - многослойные нейронные сети – нейроны объединяются в слои;

– слабосвязные нейронные сети» (Д.В. Сысоев, О.В. Курипта и Д.К. Проскурин [5, с. 20]).

Язык программирования Python

Для представления «кодирования информации» можно использовать различные языки программирования, в том числе Python [6, с. 2].

С.К. Буйначев и Н.Ю. Боклаг подчеркивают, что «Python – это язык программирования с динамическим контролем типа, в котором имена во время выполнения программы могут представлять значения различных типов» [7, с. 8]. То есть образуется связь между именем и значением, где данные представлены объектами.

Анализ данных в Python

Для анализа метеорологических данных в Python используются следующие библиотеки: Tensor Flow, Num Py, scikit-learn, Pandas, Matplotlib, statsmodels – «для очистки и подготовки данных, анализа и визуализации данных, а также для создания нейронной модели» [8, с. 25].

Данные о метеорологических значениях были взяты с сайта «Прогноз погоды в Москве» (<https://pogoda.mail.ru/prognoz/moskva/october-2024/>).

Первым этапом решения задачи является предварительный анализ данных (см. Рисунок 1). Набор данных состоит из 57 записей с интервалом в 1 день и включает в себя следующие признаки: «Дата», «Температура», «Давление (мм)», «Ветер (м/с)», «Влажность (%)». Все перечисленные параметры являются количественными переменными, что позволяет использовать разнообразные статистические методы для их анализа. К ним относятся:

- проверка данных на дубликаты;
- пропущенные значения;
- повторяющиеся значения;
- очистка данных;
- статистический анализ;
- визуализация.

	Дата	Температура	Давление(мм)	Ветер(м/с)	Влажность (%)
0	01.09.2024	29	755	1	38
1	02.09.2024	27	757	2	32
2	03.09.2024	27	757	1	31
3	04.09.2024	20	759	2	52
4	05.09.2024	21	761	2	42
5	06.09.2024	23	761	1	39
6	07.09.2024	24	760	1	36
7	08.09.2024	25	758	1	34
8	09.09.2024	26	754	1	40
9	10.09.2024	26	750	2	36
10	11.09.2024	24	749	2	38
11	12.09.2024	21	751	1	34
12	13.09.2024	27	754	3	30
13	14.09.2024	26	755	2	32
14	15.09.2024	26	757	2	24

Рисунок 1. Исходные данные

Источник: здесь и далее схемы, рисунки и графики составлены авторами

Прогнозирование метеорологических показателей с использованием нейросетевых...

На Рисунке 2 представлены результаты проверки на дубликаты (используется метод `duplicated`), пропущенные значения (можно выявить с помощью метода `isnull`), уникальные значения (`nunique`).

	Температура	Давление(мм)	Ветер(м/с)	Влажность (%)
count	57.000000	57.000000	57.000000	57.000000
mean	17.298246	752.964912	2.859649	54.333333
std	7.176263	5.223453	1.705439	21.429563
min	4.000000	740.000000	1.000000	24.000000
25%	11.000000	749.000000	2.000000	35.000000
50%	19.000000	753.000000	2.000000	53.000000
75%	23.000000	757.000000	4.000000	72.000000
max	29.000000	761.000000	7.000000	95.000000
Дата	57			
Температура	20			
Давление(мм)	19			
Ветер(м/с)	7			
Влажность (%)	42			

Рисунок 2. Предварительный анализ данных

В результате предварительного анализа не выявлены дубликаты и пропущенные значения. Было найдено количество уникальных значений для каждого столбца и проведен статистический анализ. Все это обеспечивает отсутствие неточностей при выполнении машинного обучения. По проведенной обработке данных можно сделать вывод, что показатели корректны для дальнейших этапов создания модели прогнозирования.

«Визуализация данных – это набор методов, которые позволяют использовать визуальное представление для изучения, анализа и коммуникации количественных данных, в том числе видеть тенденции и закономерности»². В работе использовались следующие способы визуализации: гистограмма, диаграмма размаха и корреляционный анализ.

«Конечная цель визуализации данных – способствовать принятию более эффективных решений и мер»³.

Гистограмма – это своего рода столбчатая диаграмма, которая дает дискретное отображение частоты значений [9, с. 18].

Диаграмма размаха – особый вид диаграммы, компактно показывающий распределение элементов. На ней видно медиану, размах, можно оценить дисперсию, асимметрию и определить выбросы.

Корреляционный анализ – это метод математической статистики, который позволяет определить степень взаимосвязи между различными параметрами [10, с. 2].

² Few S. What is data visualization? Perceptual Edge [блог]. 2017. May 4. URL: <https://www.perceptualedge.com/blog/?p=2636> (дата обращения: 22.10.24).

³ Инструменты для качественной визуализации данных: искусство использования диаграмм. Копенгаген : Европейское региональное бюро ВОЗ, 2021. Лицензия: CC BY-NC-SA 3.0 IGO. С. 1. URL: https://evidence-impact.org/storage/364/Tools-for-data-visualizations_rus.pdf (дата обращения: 27.10.24).

По матрице корреляций можно увидеть, что на прогнозируемую переменную «Температура» больше всего влияет переменная «Влажность (%)», однако переменные «Давление (мм)» и «Ветер (м/с)» также влияют на эту зависимую переменную, но в меньшей степени (см. Рисунок 3).

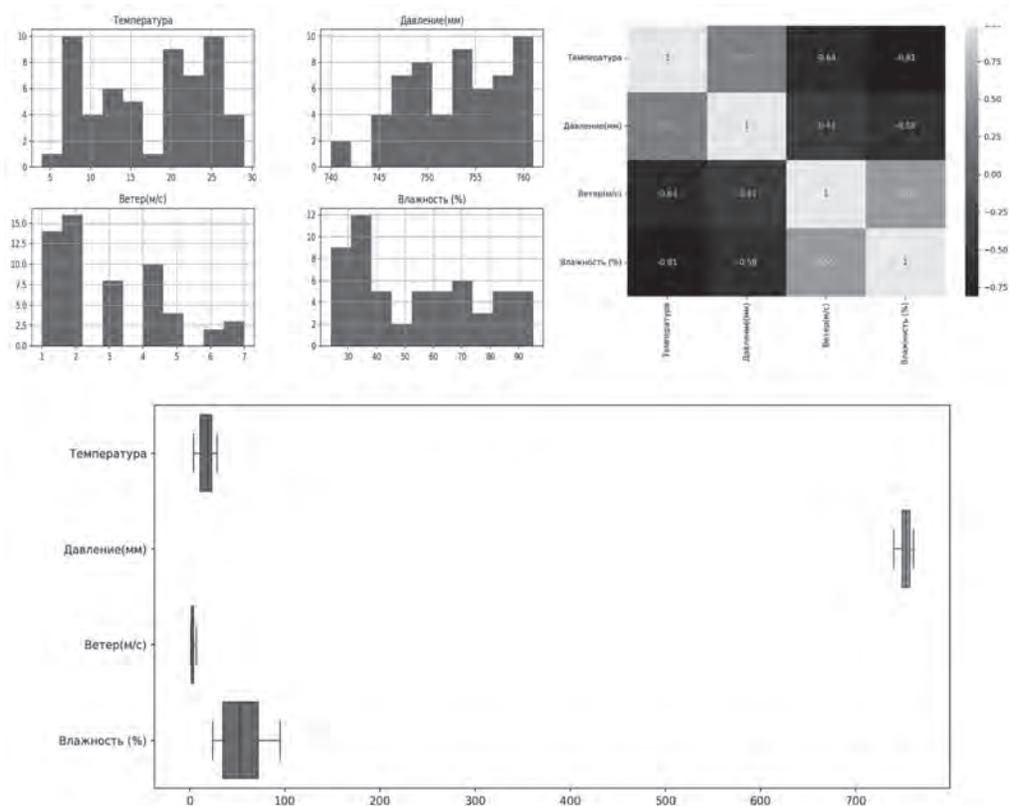


Рисунок 3. Визуализация данных

Прогнозирование метеорологических данных с помощью нейронной сети

Для прогнозирования переменной «Температура» данные были разделены на обучающую и тестовую выборки в соотношении 80 % на 20 %, где 80 % – (training set) – используется исключительно для обучения модели. Тестовая выборка (testing set) полностью независима от обучающей выборки и предназначена для оценки качества работы модели. Без ее применения нельзя гарантировать несмещенность оценок модели [11, с. 7].

Далее была выполнена стандартизация модели. Преобразование данных в набор, имеющий определенные статистические характеристики, но неопределенные минимальные и максимальные значения, называется стандартизацией данных [12, с. 87]. После этого была разработана архитектура нейронной сети с одним входным слоем (64 нейрона), одним скрытым слоем (32 нейрона) и одним выходным слоем, который прогнозирует температуру. Далее разработанная нейронная модель обучалась на тренировочных данных в течение 100 эпох с размером пакета 32. После этого оценка проводилась на тестовых данных. Здесь вычисляются метрики производительности: средняя абсолютная ошибка

Прогнозирование метеорологических показателей с использованием нейросетевых...

(MAE) и коэффициент детерминации (R^2), которые могут быть использованы для оценки качества прогноза модели. В данном случае ошибка $R^2 = 0,77$; MAE = 3,08, что считается хорошим результатом (см. Рисунок 4).

```
MAE: 3.086972916548428
R^2: 0.7733358995710944

Предсказанная температура воздуха на 28.10.24: 23 °
PS C:\Users\user> █
```

Рисунок 4. Вывод предсказанной погоды

В целом представленная модель – это полноценная система для обработки данных погоды, предсказания температуры и выполнения прогноза на основе новых входных данных. Она обеспечивает анализ, обработку, обучение и оценку системы, создавая основу для использования в реальных задачах предсказания.

Заключение

Таким образом, в статье показано применение нейронных сетей с использованием языка программирования Python и различных библиотек, в частности для прогнозирования метеорологических показателей. В статье также представлен предварительный анализ данных, выявлены ключевые взаимосвязи и построена модель нейронной сети, которая показала высокую точность предсказаний ($R^2 = 0,77$; MAE = 3,08). Результаты подтверждают эффективность нейронных сетей в этой области и их потенциал для точного прогнозирования.

Литература

1. Раскатова М.В., Терелецкова Е.Е., Сало А.А. Нейросети и перспективы их использования в экономико-хозяйственной деятельности // Цифровая трансформация: тенденции и перспективы : II Международная научно-практическая конференция, Москва, 20 декабря 2023 г. М. : Мир науки, 2023. С. 397–404. EDN UHTNLG
2. Хайкин С. Нейронные сети: полный курс. 2-е изд. / Пер. с англ. Н.Н. Кузусль, А.Ю. Шелестова. М. : Вильямс, 2006. 1104 с. ISBN 5-8459-0890-6.
3. Гафаров Ф.М., Галимянов А.Ф. Искусственные нейронные сети и их приложения: учеб. пособие. Казань : Изд-во Казанского университета, 2018. 121 с.
4. Головкин В.А., Краснопрошин В.В. Нейросетевые технологии обработки данных : учеб. пособие. Минск : БГУ, 2017. 263 с. ISBN 978-985-566-467-4. EDN GLVGIE.
5. Сысоев Д.В., Курипта О.В., Проскурин Д.К. Введение в теорию искусственного интеллекта : учеб. пособие. Воронеж : Воронежский государственный архитектурно-строительный университет, 2014. 170 с. ISBN 978-5-89040-498-5. EDN UGOVAX.
6. Марченко А.А. Python: большая книга примеров : учеб. пособие. М. : Изд-во Московского университета, 2023. 361 с. ISBN 978-5-19-011853-7.
7. Буйначев С.К., Боклаг Н.Ю. Основы программирования на языке Python : учеб. пособие. Екатеринбург : Изд-во Урал. ун-та, 2014. 92 с. ISBN 978-5-7996-1198-9. EDN HIMEBZ.
8. Маккинни У. Python и анализ данных: Первичная обработка данных с применением pandas, NumPy и Jupiter : учеб. пособие. 3-е изд. / Пер. с англ. А.А. Слинкина. М. : МК Пресс, 2023. 536 с. ISBN 978-5-93700-174-0.

9. Лемешевский С.В. Графическая визуализация данных : учеб. пособие. Минск : Институт математики НАН Беларуси, 2020. 25 с.
10. Румянцев М.И. Корреляционный анализ: методические указания. Магнитогорск : Магнитогорский государственный технический университет им. Г.И. Носова, 2017. 12 с.
11. Кугаевских А.В., Муромцев Д.И., Кирсанова О.В. Классические методы машинного обучения : учеб. пособие. СПб. : Университет ИТМО, 2022. EDN MJQVRZ.
12. Старовойтов В.В., Голуб Ю.И. Нормализация данных в машинном обучении // Информатика. 2021. Т. 18. № 3. С. 83–96. EDN JKANKM. DOI: 10.37661/1816-0301-2021-18-3-83-96

References

1. Raskatova M.V., Tereletsikova E.E., Salo A.A. (2023) Neural networks and prospects for their use in economic and economic activities. In: Ketoeva N.L., Zargaryan M.T. (Eds) *Tsifrovaya transformatsiya: tendentsii i perspektivy* [Digital transformation: Trends and prospects] : II International Scientific and Practical Conference. Moscow, December 20, 2023. Moscow : Mir Nauki Publ. Pp. 397–404. (In Russian).
2. Haykin S. (1998) *Neural Networks: A Comprehensive Foundation*. Prentice Hall, NJ : Upper Saddle River. 842 p. ISBN 0-13-273350-1. (Russian edition: transl. by N.N. Kussul and A.Yu. Shelestov. Moscow : Williams Publ., 2006. 1104 p.).
3. Gafarov F.M., Galimyanov A.F. (2018) *Iskusstvennye neironnye seti i ikh prilozheniya* [Artificial neural networks and their applications] : Training Manual. Kazan : Kazan University Publ. 121 p. (In Russian).
4. Golovko V.A., Krasnoprosin V.V. (2017) *Neirosetevye tekhnologii obrabotki dannykh* [Neural network technologies for data processing] : Training Manual. Minsk : Belarusian State University Publ. 263 p. ISBN 978-985-566-467-4. (In Russian).
5. Sysoev D.V., Kuripta O.V., Proskurin D.K. (2014) *Vvedeniye v teoriyu iskusstvennogo intellekta* [Introduction to Artificial Intelligence Theory] : Training Manual. Voronezh : Voronezh state agrarian university Publ. 170 p. ISBN 978-5-89040-498-5. (In Russian).
6. Marchenko A.L. (2023) *Python: bol'shaya kniga primerov* [Python: A big book of examples] : Training Manual. Moscow : Moscow University Publ. 361 p. ISBN 978-5-19-011853-7. (In Russian).
7. Buinachev S.K., Boklag N.Yu. (2014) *Osnovy programmirovaniya na yazyke Python* [Python Programming Basics] : Training Manual. Ekaterinburg : Ural Federal University Publ. 92 p. ISBN 978-5-7996-1198-9. (In Russian).
8. McKinney W. (2022) *Python for Data Analysis: Data Wrangling with pandas, NumPy, and Jupyter*. 3rd edition: O'Reilly Media. 582 p. ISBN 109810398X. (Russian edition: transl. by A.A. Slinkin, Moscow : MK-press, 2022. 536 p.)
9. Lemeshevsky S.V. (2020) *Graficheskaya vizualizatsiya dannykh* [Graphical data visualization] : Training Manual. Minsk : Institute of Mathematics of the National Academy of Sciences of Belarus. 25 p. (In Russian).
10. Rumyantsev M.I. (2017) *Korrelyatsionnyi analiz* [Correlation analysis] : Guidance. Magnitogorsk : Nosov Magnitogorsk State Technical University Publ. 12 p. (In Russian).
11. Kugaevskikh A.V., Muromtsev D.I., Kirsanova O.V. (2022) *Klassicheskie metody mashinnogo obucheniya* [Classic machine learning methods] : Training Manual. St. Petersburg : Saint Petersburg State University of Information Technologies, Mechanics and Optics Publ. 53 p. (In Russian).
12. Starovoitov V.V., Golub Yu.I. (2021) Data normalization in machine learning. *Informatics*. Vol. 18. No. 3. Pp. 83–96. DOI: 10.37661/1816-0301-2021-18-3-83-96 (In Russian).