

<https://doi.org/10.29188/2712-9217-2023-9-4-36-44>

Анализ влияния факторов окружающей среды на здоровье человека с применением цифровых решений

Мнение эксперта

И.А. Шадеркин¹, А.П. Дьяченко², Е.А. Чулюкова³, Т.В. Пшеничный³, Л.В. Ковека⁴, В.Е. Храмцова⁴, Н.С. Гугнявых⁵, А.И. Кузьмина⁶

¹ Институт цифровой медицины Первого Московского государственного медицинского университета им. И.М. Сеченова (Сеченовский Университет); д. 1, стр. 2, Абрикосовский пер., Москва, 119435, Россия

² ФГБОУ ВО «Волгоградский государственный медицинский университет» МЗ РФ; д. 1, площадь Павших Борцов, Волгоград, 400066, Россия

³ Московский государственный университет им. М. В. Ломоносова; дом 1, Ленинские горы, Москва, 119234, Россия

⁴ ФГБОУ ВО «Южно-Уральский государственный медицинский университет» МЗ РФ; дом 64, ул. Воровского, Челябинск, 454092, Россия

⁵ Санкт-Петербургский государственный университет; д. 7/9, Университетская набережная, Санкт-Петербург, 199034, Россия

⁶ Санкт-Петербургский государственный университет ветеринарной медицины; дом 5, ул. Черниговская, Санкт-Петербург, 196084, Россия

Контакт: Шадеркин Игорь Аркадьевич, info@uroweb.ru

Аннотация:

Введение. В статье рассматриваются возможности применения цифровых решений анализа влияния факторов окружающей среды на здоровье человека.

Целью является предложение цифровой модели для гигиенической оценки комплексного влияния факторов окружающей среды на здоровье человека.

Решаемая проблема: в настоящее время информация по оценке влияния многочисленных факторов окружающей среды на здоровье человека характеризуется высокой рассеянностью, не позволяя учитывать их в комплексе. Использование цифровых решений позволит проводить комплексную гигиеническую оценку этих факторов. Новизна работы заключается в применении цифровых технологий при анализе и оценке исследуемых факторов окружающей среды.

Материалы и методы. В процессе работы использованы научные и обзорные статьи, отчёты и доклады государственных органов, общие методы научного исследования: анализ, сравнение, обобщение, синтез, системный подход, структурный анализ; статистический метод; гигиенический анализ.

Результаты. Для комплексной оценки факторов окружающей среды на здоровье человека нами предложено применить модель случайного леса машинного обучения на наборе данных факторов окружающей среды. На основе посчитанных моделью весов можно сделать выводы о том, какие факторы наиболее существенно влияют на возникновения тех или иных заболеваний.

Выводы. Результаты исследования перспективны для внедрения в практику оценки качества окружающей среды исследуемой области. Их можно использовать для разработки методов и технологий для предотвращения и сокращения негативного воздействия этих факторов на здоровье людей.

Ключевые слова: факторы окружающей среды; здоровье человека; цифровые решения; количественная оценка; корреляция.

Для цитирования: Шадеркин И.А., Дьяченко А.П., Чулюкова Е.А., Пшеничный Т.В., Ковека Л.В., Храмцова В.Е., Гугнявых Н.С., Кузьмина А.И. Анализ влияния факторов окружающей среды на здоровье человека с применением цифровых решений. Российский журнал телемедицины и электронного здравоохранения 2023;9(4):36-44; <https://doi.org/10.29188/2712-9217-2023-9-4-36-44>

Analysis of the impact of environmental factors on human health using digital solutions

Expert opinion

<https://doi.org/10.29188/2712-9217-2023-9-4-36-44>**I.A. Shaderkin¹, A.P. Dyachenko², E.A. Chulyukova³, T.V. Pshenichny³, L.V. Koveka⁴, V.E. Khramtsova⁴, N.S. Gugnyavykh⁵, A.I. Kuzmina⁶**¹ Institute of Digital Medicine of the First Moscow State Medical University them Sechenov (Sechenov University), Abrikosovskiy per., 1, bldg. 2, Moscow, 119435, Russia² Federal State Budgetary Educational Institution of Higher Education «Volgograd State Medical University» of the Ministry of Health of the Russian Federation; 1 Fallen Fighters Square, Volgograd, 400066, Russia³ Lomonosov Moscow State University; 1, Leninskie Gory, Moscow, 119234, Russia⁴ Federal State Budgetary Educational Institution of Higher Education «South-Ural State Medical University» of the Ministry of Healthcare of the Russian Federation; 64, Vorovskogo str., Chelyabinsk, 454092, Russia⁵ St. Petersburg State University; 7/9, Universitetskaya embankment, St. Petersburg, 199034, Russia⁶ St. Petersburg State University of Veterinary Medicine; 5, Chernigovskaya, St. Petersburg, 196084, Russia**Contact:** Igor A. Shaderkin, info@uroweb.ru**Annotation:****Introduction.** The article discusses the possibilities of using digital solutions to analyze the impact of environmental factors on human health.**The aim** is to propose a digital model for the hygienic assessment of the complex impact of environmental factors on human health.

The problem to be solved: currently, information on the assessment of the impact of numerous environmental factors on human health is characterized by high dispersion, not allowing them to be taken into account in a complex. The use of digital solutions will allow for a comprehensive hygienic assessment of these factors. The novelty of the work lies in the use of digital technologies in the analysis and assessment of the studied environmental factors.

Materials and methods. In the process of work, scientific and review articles, reports and reports of government agencies, general methods of scientific research are used: analysis, comparison, generalization, synthesis, systematic approach, structural analysis; statistical method; hygienic analysis.**Results.** For a comprehensive assessment of environmental factors on human health, we proposed to apply a random forest machine learning model on a data set of environmental factors. Based on the scales calculated by the model, it is possible to draw conclusions about which factors most significantly affect the occurrence of certain diseases.**Conclusions.** The results of the study are promising for implementation in the practice of assessing the environmental quality of the study area. They can be used to develop methods and technologies to prevent and reduce the negative impact of these factors on human health.**Key words:** environmental factors; human health; digital solutions; quantitative assessment; correlation.**For citation:** Shaderkin I.A., Dyachenko A.P., Chulyukova E.A., Pshenichny T.V., Koveka L.V., Khramtsova V.E., Gugnyavykh N.S., Kuzmina A.I. Analysis of the impact of environmental factors on human health using digital solutions. Russian Journal of Telemedicine and E-Health 2023;9(4):36-44; <https://doi.org/10.29188/2712-9217-2023-9-4-36-44>**■ ВВЕДЕНИЕ**

В настоящее время во всем мире проблема профилактики неблагоприятного воздействия факторов окружающей среды на здоровье человека привлекает внимание многих ученых – врачей, экологов, биологов, физиков. Это вызвано постоянным ростом числа факторов, оказывающих негативное воздействие на здоровье человека. По данным Всемирной организации здравоохранения (ВОЗ) около 1,4 млн. случаев преждевременной смерти могут быть связаны с экологическими факторами, причем 569 тыс. из них связаны с загрязнением окружающего воздуха, а 154 тыс. – с загрязнением воздуха внутри помещений [1]. Однако информация по оценке влияния многочисленных факторов среды помещения и окружающей среды на здоровье человека характеризуется рассеянностью, не позволяя учитывать все факторы в комплексе. Цифровые решения

могут быть эффективным инструментом для анализа этих факторов и разработки научного обоснования влияния их на здоровье человека [2].

Анализ комплексной оценки влияния факторов окружающей среды на здоровье человека можно провести методом корреляции показателей здоровья населения по классам основных нозологий на определенной территории с показателями факторов окружающей среды на этой территории. Для этого возможно использование информационных технологий анализа данных, установления причинно-следственных связей изменения показателей здоровья населения с показателями факторов окружающей среды, оценки рисков их воздействия на здоровье. Полученные результаты могут использоваться для разработки профилактических мероприятий, направленных на уменьшение отрицательного влияния факторов среды на здоровье человека. ►►

В статье собран и обработан материал по формированию модели практического применения результатов цифровых решений анализа оценки влияния факторов окружающей среды на здоровье населения.

■ МАТЕРИАЛЫ И МЕТОДЫ

Авторами был проведен анализ научных статей по теме исследования в общедоступных электронных базах Medline, PubMed, Google Scholar, Elibrary, по указанным ключевым словам, отчетам и докладом государственных органов, показателям Росстата. Используются методы: структурный анализ; корреляционный анализ, метод ансамблевого машинного обучения, в частности дерева решения и алгоритм случайного леса.

Цель работы: предложить алгоритм прогнозирования воздействия комплекса факторов окружающей среды.

■ РЕЗУЛЬТАТЫ

В настоящее время в научной литературе предлагаются разные подходы оценки влияния факторов окружающей среды на здоровье населения [3-5]. В большинстве случаев они основываются на вероятностном математико-статистическом анализе воздействия вредного фактора на здоровье населения, с помощью которого можно количественно оценить уровень связи между факторами риска и критериями общественного здоровья и построить вероятностную математическую модель воздействия вредного фактора на здоровье населения [3, 4]. Количественная оценка риска может осуществляться с помощью корреляционного анализа. По коэффициентам корреляций можно судить о связи между загрязнением среды и состоянием здоровья населения [3].

Однако рост уровня заболеваемости населения, связанный с воздействием вредных факторов окружающей среды, может не носить линейный характер, так как пороги чувствительности и реакция защитных сил организма индивидуальны. Поэтому на начальном этапе воздействие вредных факторов может не отражаться на статистике заболеваемости или отражаться незначительно. При возрастающем воздействии в дальнейшем происходит срыв механизмов адап-

тации, и заболеваемость возрастает. Поэтому степень вредных воздействий факторов можно установить только на основании закона больших чисел – количественные закономерности массовых явлений проявляются лишь в достаточно большом их числе. Для этих целей необходимо выполнить анализ больших по численности групп населения. В небольших группах выводы могут быть ошибочные, так как большую роль будут играть случайные факты и индивидуальные особенности организма человека. При анализе большого количества событий случайности сглаживаются, и погрешности результатов снижаются [6].

Клепиковым О.В. и Студеникиной Е.М. предложены методы оценки влияния загрязнения атмосферного воздуха на здоровье населения при помощи анализа больших чисел. За величину, характеризующую загрязненность атмосферы принимается комплексный индекс загрязнения атмосферы (ИЗА), учитывающий несколько приоритетных загрязнителей (как правило – пять) [1].

По значению коэффициента парной корреляции можно судить о тесноте взаимосвязи между изучаемыми показателями. Построение регрессионных моделей в таких исследованиях позволяет оценить направленность, силу, вид связи, прогнозировать значения. Моделировать процессы можно как по отдельным компонентам, так и в комплексе [3].

При оценке взаимосвязи здоровья населения с загрязнением окружающей среды часто используется линейное программирование. Для прогноза периодических процессов по известному спектру частот используется Фурье-анализ. Методы моделирования и прогнозирования временных рядов позволяют выявить тенденции изменения фактических значений параметра во времени и прогнозировать его будущие значения [3].

Коротковым П.А., Трубяновым А.Б. и соавт. для анализа статистической связи показателей загрязнения окружающей среды с показателями экологически зависимой заболеваемости предложена методология исследования, основанная на подходах к корреляционно-регрессионному анализу панельных данных. Для этого рассчитываются коэффициенты корреляции Пирсона и ранговые коэффициенты корреляции Спирмена, затем строятся модели регрессии для панельных данных: модель с фиксированными эффектами и модель со случайными эффектами. Источни-

ками панельных данных являются показатели подразделений Росстата, Роспотребнадзора и Минздрава: 6 основных показателей загрязнения воды и атмосферы и 7 приоритетных показателей заболеваемости населения за определенный период (несколько лет) [4].

В рассматриваемых выше методиках данные о заболеваемости населения принимаются по статистическим данным. Однако для более углубленного анализа влияния негативных факторов окружающей среды на здоровье необходимо изучение показателей здоровья конкретной местности. Для этих целей можно воспользоваться единой интегрированной платформой, на которую будут поступать от медицинских учреждений все данные о здоровье пациентов. Благодаря таким данным можно проводить детальный анализ влияния факторов среды на здоровье, учитывающий и наличие загрязняющих веществ в организме пациентов, и их ответную реакцию на изменение факторов среды. Хранение больших данных можно обеспечить в облачных хранилищах.

На основе изученных материалов и методик анализа больших данных в данной работе мы в качестве примеров выполнили:

1) корреляционный анализ и прогноз показателей загрязнения атмосферного воздуха и заболеваемости детей от 0 до 14 лет по болезням органов дыхания по Волгоградской области за пять лет – с 2018 по 2022 гг.;

2) алгоритм прогнозирования воздействия комплекса факторов окружающей среды.

Корреляционный анализ и прогноз показателей загрязнения атмосферного воздуха и заболеваемости детей от 0 до 14 лет по болезням органов дыхания по Волгоградской области за пять лет – с 2018 по 2022 гг.

Исследованием комплексной гигиенической оценки факторов окружающей среды, проведен-

ным Клепиковым О.В. и др. на основе города Воронежа, установлено, что приоритетным фактором неблагоприятного воздействия на здоровье населения является атмосферный воздух [7].

Индекс загрязнения атмосферного воздуха (ИЗА) по пяти приоритетным загрязнителям принят по статистическим данным, приведенным в Докладе Комитета природных ресурсов, лесного хозяйства и экологии Волгоградской области о состоянии окружающей среды Волгоградской области в 2022 году [8]. Влиянию вредных факторов окружающей среды наиболее подвержены дети. Важно, что дети более привязаны к исследуемой территории, так как в основном они живут и учатся на одном территориальном участке, на их здоровье не отражаются профессиональные факторы. Все это уменьшает величину случайных событий при анализе данных, поэтому предлагается проводить анализ статистических данных о заболеваемости детей по основным нозологиям. Показатели заболеваемости детей приняты по данным Росстата и Роспотребнадзора по Волгоградской области [9-14]. В заболеваниях органов дыхания учитываются: бронхиты, пневмонии, бронхиальная астма, хроническая обструктивная болезнь легких, бронхоэктатическая болезнь, плеврит, эмфизема легких. По данным проведенных исследований прослеживается взаимосвязь этих заболеваний с такими загрязнителями воздуха: взвешенные вещества, диоксид азота, бензапирен [15, 16]. В таблице 1 приведены показатели ИЗА и заболеваемости, на основании которых произведены расчеты.

Значение коэффициента корреляции (R) может быть: $-1 < R < +1$, что свидетельствует:

- R близко к 0 – свидетельствует об отсутствии корреляционной связи между концентрациями вредных веществ и уровнем заболеваемости;
- R близко к 1 – существует положительная связь; ►►

Таблица 1. Показатели ИЗА и болезней органов дыхания по Волгоградской области за 2018–2022 гг.
Table 1. Indicators of IZA and respiratory diseases in the Volgograd region for 2018–2022

Годы	Болезни органов дыхания (Волгоград) (на 100 тыс. детского населения)	ИЗА (Волгоград)
2018	115191,80	4,30
2019	113388,70	3,20
2020	105602,60	3,50
2021	115278,10	10,60
2022	122320,00	10,60

- R близко к -1 – существует отрицательная связь.

Анализ с применением электронных таблиц Microsoft Excel приведен на гистограмме (рис. 1). Коэффициент корреляции, рассчитанный по формуле в электронной таблице Microsoft Excel равен 0,70, что свидетельствует о наличии положительной зависимости.

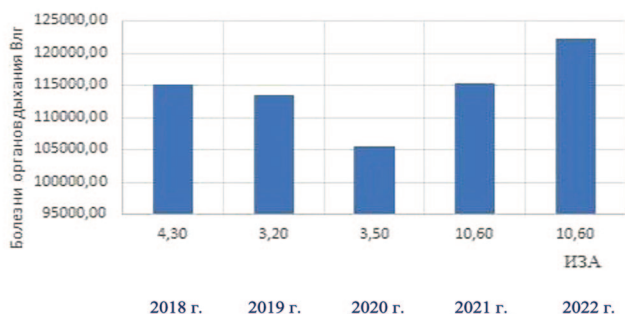


Рис. 1 Вероятностная взаимосвязь между уровнем загрязнения атмосферного воздуха и уровнем заболеваемости органов дыхания по Волгоградской области за 2018 – 2022 гг. (расчет по формуле корреляции в электронной таблице Microsoft Excel) (выполнено авторами)
Fig.1 Probabilistic relationship between the level of atmospheric air pollution and the level of respiratory morbidity in the Volgograd region for 2018 – 2022. (calculation using the correlation formula in a Microsoft Excel spreadsheet) (completed by the authors)

На основе данных таблицы 1 с помощью электронной таблицы Microsoft Excel был сделан график и построена линия тренда (аппроксимирующая функция $y(x)$, которая с максимальной степенью близости приближается к опытной зависимости) (рис. 2).

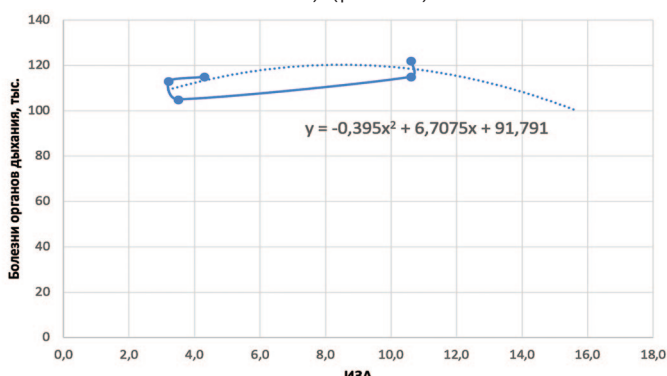


Рис.2. График зависимости между уровнем загрязнения атмосферного воздуха и уровнем заболеваемости органов дыхания и линия тренда (выполнено авторами)
Fig.2. Graph of the relationship between the level of atmospheric air pollution and the level of respiratory diseases and the trend line (completed by the authors)

Как видно на рисунке, зависимость не является линейной.

Прогноз воздействия загрязнения атмосферного воздуха на болезни органов дыхания выполнен в облачном сервисе Google Collab. В качестве исследования была выбрана задача прогнозирования возникновения заболевания одной из групп у человека при известных числовых показателях ИЗА, взятых из официаль-

ных источников по Волгоградской области, представленных в таблице 1.

Предлагается применить метод ансамблевого машинного обучения в виде алгоритма случайного леса и дерева решений на примере набора данных с помощью языка программирования Python, где в качестве признаков будут использоваться числовые показатели индекса загрязнения атмосферы, а в качестве целевой переменной — группы болезней органов дыхания. Random forest («Случайный лес») – алгоритм машинного обучения, заключающийся в использовании комитета (ансамбля) решающих деревьев [17-19]. Ниже приведена ссылка, в которой мы реализовали метод алгоритма случайного леса.

<https://colab.research.google.com/drive/1p4ozfSrOPI2AzUONwIVSzmYzTvCtiDw?usp=sharing>

Приводим пример написания кода:

- прогноз обучающей модели (рис. 3):

```
#прогноз
y_pred = classifier.predict(x_test)
y_pred

array(['10,6'], dtype=object)

#проверка прогноза
from sklearn.metrics import classification_report, confusion_matrix
print(confusion_matrix(y_test, y_pred))
print(classification_report(y_test, y_pred))

[[1]]
```

	precision	recall	f1-score	support
10,6	1.00	1.00	1.00	1
accuracy			1.00	1
macro avg	1.00	1.00	1.00	1
weighted avg	1.00	1.00	1.00	1

Рис. 3. Прогноз обучающей модели (выполнено авторами)
Fig. 3. The forecast of the training model (completed by the authors)

- построение дерева решений (рис. 4):

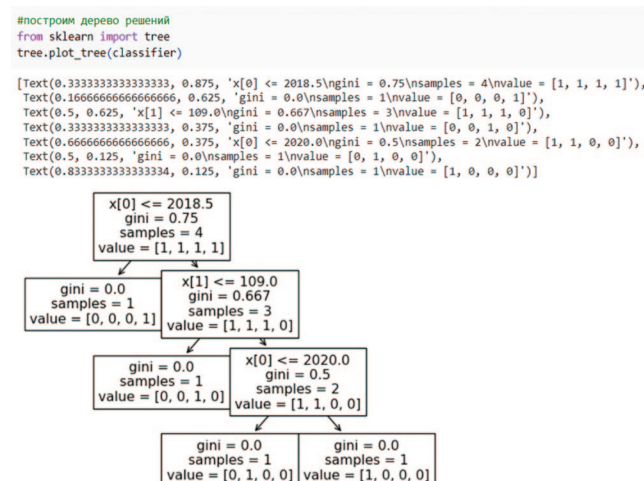


Рис. 4. Построение дерева решений (выполнено авторами)
Fig. 4. Building a decision tree (completed by the authors)

- проверка прогноза (рис. 5):

```
#прогноз
y_pred = classifier.predict(x_test)
y_pred

array(['10,6'], dtype=object)

#проверка прогноза
from sklearn.metrics import classification_report, confusion_matrix
print(confusion_matrix(y_test, y_pred))
print(classification_report(y_test, y_pred))

[[1]]
      precision    recall  f1-score   support

   10,6         1.00      1.00      1.00         1

 accuracy         1.00
 macro avg         1.00
 weighted avg       1.00
```

Рис. 5. Проверка прогноза (выполнено авторами)
Fig. 5. Checking the forecast (completed by the authors)

Вывод диаграммы вероятностной зависимости между уровнем загрязнения атмосферного воздуха и уровнем заболеваемости органов дыхания показывает наличие зависимости, однако она не является линейной (рис. 6):

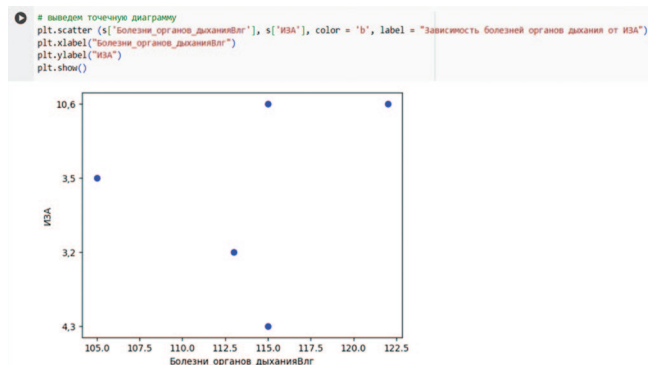


Рис. 5. Проверка прогноза (выполнено авторами)
Fig. 5. Checking the forecast (completed by the authors)

Алгоритм прогнозирования воздействия комплекса факторов окружающей среды

В качестве исследования была выбрана задача прогнозирования возникновения заболевания одной из групп у человека при известных числовых показателях воздействия окружающей среды.

Предлагается применить метод ансамблевого машинного обучения в виде алгоритма случайного леса и дерева решений на примере набора данных с помощью языка программирования Python, где в качестве признаков будут использоваться числовые показатели воздействия различных факторов на организм человека, а в качестве целевой переменной – группы болезней, которые могут возникнуть [20].

Дерево решений – это структура данных, которая представляет собой древовидную модель принятия решений в виде логических правил. Каждый узел дерева представляет собой признак, по которому происходит разделение данных, а каждое ребро – условие (например, «больше 30 лет»). Листовые узлы дерева содержат предсказание (например, класс объекта).

Процесс построения дерева решений включает следующие этапы:

1. Выбор признака для разделения дерева на узлы:

На каждом узле дерева выбирается признак, который наилучшим образом разделяет данные на подгруппы. Этот выбор осуществляется на основе критерия информативности.

2. Разделение данных:

После выбора признака происходит разделение данных на две или более подгруппы в зависимости от значения выбранного признака.

3. Построение поддеревьев:

Для каждой подгруппы данных происходит рекурсивное построение поддерева, повторяя процесс выбора признака и разделения данных.

4. Остановка построения дерева:

Построение дерева может быть остановлено при достижении определенного критерия, например, максимальной глубины дерева, минимального числа объектов в листе или при отсутствии улучшения качества разделения.

Деревья решений могут быть склонны к переобучению, поэтому модель случайного леса использует несколько деревьев для уменьшения этого эффекта. Каждое дерево обучается на случайной подвыборке данных и случайном подмножестве признаков, что способствует разнообразию деревьев и повышению качества модели.

В качестве параметров были выбраны следующие факторы: уровни взвешенных веществ, оксида углерода, диоксида азота, диоксида серы и оксида азота в атмосферном воздухе, нитратов, ртути, бензола, марганца, свинца в почве, а также уровни железа, меди, цинка и фенола в водопроводной воде, уровень радиоактивной нагрузки.

В качестве предсказываемых классов рассматриваются заболевания органов дыхания, заболевания печени и почек, онкологические заболевания, сердечно-сосудистые заболевания, отсутствие заболеваний. Ниже приведена ►►

реализация алгоритма случайного леса:

Загрузка необходимых библиотек:

```
• from sklearn.ensemble import RandomForestClassifier
• from sklearn.model_selection import train_test_split
• from sklearn.metrics import accuracy_score
```

Далее разделяем данные на обучающий и тестовый наборы, здесь X – параметры воздействия окружающей среды, Y – величина, принадлежащая к одному из классов: заболевание органов дыхания, заболевание печени и почек, онкологическое заболевание, сердечно-сосудистое заболевание, отсутствие заболеваний.

```
X_train, X_test, y_train, y_test = train_test_split(X,
y, test_size=0.2, random_state=42)
```

Инициализация и обучение модели Random Forest:

```
clf = RandomForestClassifier()
clf.fit(X_train, y_train)
```

Предсказание на тестовом наборе:

```
y_pred = clf.predict(X_test)
```

Оценка качества модели:

```
accuracy = accuracy_score(y_test, y_pred)
print("Accuracy:", accuracy)
```

На основе посчитанных моделью весов можно сделать выводы о том, какие факторы наиболее существенно влияют на возникновение тех или иных заболеваний. Получить веса можно с помощью приведенного ниже кода:

```
importances = clf.feature_importances_
for class_idx, class_name in enumerate(
clf.classes_):
print(f"Class: {class_name}")
for idx, importance in enumerate(importances[class_idx]):
print(f"Feature {idx}: Importance {importance}")
print("\n")
```

■ ОБСУЖДЕНИЕ

Изучение собранных материалов и проведенный анализ на примере Волгоградской области показал, что имеется вероятностная зависимость между уровнем загрязнения атмосферного воздуха и уровнем заболеваемости органов дыхания среди детского населения, однако она не является линейной.

По расчетам в электронной таблице Microsoft Excel значение коэффициента корреля-

ции равно 0,7 что свидетельствует о наличии положительной зависимости между уровнем загрязнения атмосферного воздуха и уровнем заболеваемости органов дыхания, не являющейся линейной (рис. 2). В облачном сервисе Google Collab с помощью метода лесов можно сделать прогноз и его оценку, а также построить точечную диаграмму зависимостей. Мы видим, что в обеих цифровых технологиях существует положительная связь между факторами загрязнения атмосферного воздуха и болезнями органов дыхания.

Для прогнозирования воздействия комплекса факторов окружающей среды на здоровье населения предлагается применить метод ансамблевого машинного обучения в виде алгоритма случайного леса и дерева решений на примере набора данных с помощью языка программирования Python, где в качестве признаков будут использоваться числовые показатели воздействия различных факторов на организм человека, а в качестве целевой переменной – группы болезней, которые могут возникнуть.

Для обучения модели возможно применение баз данных на основе мониторинга факторов прошлых лет. Подойдут данные с разных инфраструктурных территорий страны и рассмотрением различных показателей.

Уже к обученной модели можно подключать как большие объёмы, такие как анализ состава воздуха по городу, так и локальные, собранные непосредственно в исследуемом помещении.

Существуют обновляемые в реальном времени базы данных, которые позволяют проводить операции ввода/изменения данных и получать актуальные результаты немедленно. Примерами таких баз данных являются Apache Kafka, Amazon DynamoDB, Google Firebase Realtime Database и другие. Эти базы данных позволяют мгновенно обновлять информацию и реагировать на изменения в реальном времени. Именно одну из таких можно использовать для прогнозирования и мониторинга зависимости.

В качестве сбора данных можно использовать различные датчики и анализаторы, обслуживание которых будет быстрым и простым. Для удобства и упрощения на каждый фактор можно завести отдельный датчик, передающий информацию в обновляемую базу данных.

ВЫВОДЫ

1. Для исследования зависимостей уровня заболеваемости от факторов окружающей среды можно использовать различные виды цифровых моделей.

2. Степень воздействия вредных факторов окружающей среды можно установить только на основании закона больших чисел, так как количественные закономерности массовых явлений проявляются лишь в достаточно большом их числе [20–26].

3. Корреляционный анализ показателей загрязнения атмосферного воздуха и заболеваемости детей от 0 до 14 лет по болезням органов дыхания по Волгоградской области за пять лет, проведенный по двум цифровым технологиям, показывает наличие положительной связи

между факторами загрязнения атмосферного воздуха и болезнями органов дыхания. Однако эта зависимость не является линейной.

4. Для комплексной оценки факторов окружающей среды на здоровье человека нами предлагается применить модель случайного леса машинного обучения на наборе данных факторов окружающей среды. На основе посчитанных моделью весов можно сделать выводы о том, какие факторы наиболее существенно влияют на возникновения тех или иных заболеваний.

5. Преимущество цифровой платформы сбора и обработки данных перед отдельными модулями в том, что она позволяет осуществить в одном месте все показатели региона, выполнить анализ данных, автоматически сообщать сигналы, где возникают риски негативного воздействия на здоровье. ▄

ЛИТЕРАТУРА

1. Atmospheric air pollution. The World Health Organization 2022. [Electronic resource]. URL: [https://www.who.int/ru/news-room/factsheets/detail/ambient-\(out-door\)-air-quality-and-health](https://www.who.int/ru/news-room/factsheets/detail/ambient-(out-door)-air-quality-and-health)
2. Лебедев Г.С., Шадеркин И.А., Лебедева Н.А. Модифицируемые факторы среды помещения: влияние на здоровье человека и цифровой мониторинг. *Российский журнал телемедицины и электронного здравоохранения* 2023;9(1):21–48. [Lebedev G.S., Shaderkin I.A., Lebedeva N.A. Modifiable Indoor Environmental Factors: Impact on Human Health and Digital Monitoring. *Rossiyskiy zhurnal telemeditsiny i elektronnoogo zdravookhraneniya = Russian Journal of Telemedicine and E-Health* 2023;9(1):21–48. (In Russian)]. <https://doi.org/10.29188/2712-9217-2023-9-1-21-48>
3. Куролап С.А., Клепиков О.В. Экологические риски территорий интенсивного техногенного освоения: монография. *Цифровая полиграфия* 2019:191. [Kurolap S.A., Klepikov O.V. Ecological risks of territories of intensive technogenic development: monograph. *Digital Printing* 2019:191. (In Russian)].
4. Гребенева О.В., Сакиев К.З., Отарбаева М.Б., Жанбасинова Н.М. Построение моделей, отражающих влияние окружающей среды на состояние здоровья населения, в программе Statistica. *Гигиена труда и медицинская экология* 2017;54(1):54–89. [Grebeneva O.V., Sakiev K.Z., Otarbaeva M.B., Zhanbasinova N.M. Building models reflecting the impact of the environment on the health of the population in the Statistica program. *Gigiyena truda i meditsinskaya ekologiya = Occupational hygiene and medical ecology* 2017;54(1):54–89. (In Russian)].
5. Коротков П.А., Трубянов А.Б., Авдеева А.А., Гисмиева А.И. Статистический анализ влияния загрязнения среды обитания на заболеваемость населения в Республике Марий Эл. *Статистика и Экономика* 2020;17(3):58–66. [Korotkov P.A., Trubyanov A.B., Avdeeva A.A., Gismieva A.I. Statistical Analysis of Environmental Pollution Impact on Population Morbidity in the Republic of Mari El. *Statistika i Ekonomika = Statistics and Economics* 2020;17(3):58–66. (In Russian)]. <https://doi.org/10.21686/2500-3925-2020-3-58-66>
6. Нифонтов Н.С., Маслова Е.Ю. Закон больших чисел и теорема Чебышева. *Вестник Академии знаний* 2017;20(1):23–5. [Nifontov N.S., Maslova E.Yu. The law of large numbers and Chebyshev's theorem. *Vestnik Akademii znaniy = Bulletin of the Academy of Knowledge* 2017;20(1):23–5. (In Russian)].
7. Абдуганиева Ш.Х., Никонорова М.Л. Цифровые решения в медицине. *Крымский журнал экспериментальной и клинической медицины* 2022;12(2):73–85. [Abduganieva Sh.Kh., Nikonorova M.L. Digital solutions in medicine. *Krymskiy zhurnal eksperimental'noy i klinicheskoy meditsiny = Crimean Journal of Experimental and Clinical Medicine* 2022;12(2):73–85. (In Russian)]. <https://doi.org/10.37279/2224-6444-2022-12-2-73-85>
8. Доклад «О состоянии окружающей среды Волгоградской области в 2022 году». Комитет природных ресурсов, лесного хозяйства и экологии Волгоградской области. ТЕМПОРА 2023:300. [Электронный ресурс]. [Report «On the state of the environment of the Volgograd region in 2022». Committee of Natural Resources, Forestry and Ecology of the Volgograd region. ТЕМПОРА 2023:300. [Electronic resource]. (In Russian)]. URL: https://oblkomprirroda.volgograd.ru/upload/iblock/5d6/Ekologiya_2023-_1_.pdf
9. Государственный доклад «О состоянии санитарно-эпидемиологического благополучия населения в Волгоградской области в 2022 году». [Электронный ресурс]. [State report «On the state of sanitary and epidemiological welfare of the population in the Volgograd region in 2022» [Electronic resource]. (In Russian)]. URL:https://34.rospotrebnadzor.ru/upload/iblock/14a/230523_34_ГД_2022_28.02.23.pdf
10. Государственный доклад «О состоянии санитарно-эпидемиологического благополучия населения в Волгоградской области в 2021 году». [Электронный ресурс]. [State report «On the state of sanitary and epidemiological welfare of the population in the Volgograd region in 2021» [Electronic resource]. (In Russian)]. URL: https://34.rospotrebnadzor.ru/upload/iblock/ecc/34_ГД_2021.pdf
11. Государственный доклад «О состоянии санитарно-эпидемиологического благополучия населения в Волгоградской области в 2020 году». [Электронный ресурс]. [State report «On the state of sanitary and epidemiological welfare of the population in the Volgograd region in 2020» [Electronic resource]. (In Russian)]. URL: <https://34.rospotrebnadzor.ru/content/282/11816/>
12. Государственный доклад «О состоянии санитарно-эпидемиологического благополучия населения в Волгоградской области в 2019 году» [Электронный ресурс]. [State report «On the state of sanitary and epidemiological welfare of the population in the Volgograd region in 2019» [Electronic resource]. (In Russian)]. URL: https://34.rospotrebnadzor.ru/upload/iblock/72e/34_ГД_2019.pdf
13. Медико-санитарный атлас Волгоградской области по данным СГМ за 2022 год. Часть 1. [Электронный ресурс]. [The medical and sanitary atlas of the Volgograd region according to the SGM data for 2022. Part 1. [Electronic resource]. (In Russian)]. URL: <https://34.rospotrebnadzor.ru/content/204/14166/>
14. Статистические издания. Здравоохранение в России [Электронный ресурс]. [Statistical publications. Healthcare in Russia [Electronic resource]. (In Russian)]. URL: <https://rosstat.gov.ru/folder/210/document/13218>
15. Клепиков О.В., Самойлов А.С., Ушаков И.Б., Попов В.И., Куролап С.А. Комплексная оценка состояния окружающей среды промышленного города. *Гигиена и санитария* 2018;97(8):686–92. [Klepikov O.V., Samoilov A.S., Ushakov I.B., Popov V.I., Kurolap S.A. Comprehensive assessment of the state of the environment of an industrial city. *Gigiyena i sanitariya = Hygiene and sanitation* 2018;97(8):686–92. (In Russian)].
16. Лазарева Н.В. Комплексная оценка состояния репродуктивного и соматического здоровья населения от воздействия факторов риска окружающей среды. *Медико-фармацевтический журнал «Пulse»* 2015;17(4):277–82. [Lazareva N.V. Comprehensive assessment of the state of reproductive and somatic health of the population from the effects of environmental risk factors. *Mediko-farmatsevticheskiy zhurnal «Pul's» = Medical and pharmaceutical journal «Pulse»* 2015;17(4):277–82. (In Russian)].
17. Лимановская О.В., Алферьева Т.И. Основы машинного обучения : учебное пособие. Издательство Уральского университета 2020:88. [Limanovskaya O.V., Alfereva T.I. Fundamentals of Machine Learning: Tutorial. Ural University Publishing House 2020:88.].
18. Картиев С.Б., Курейчик В.М. Алгоритм классификации, основанный на принципах случайного леса, для решения задачи прогнозирования. Программные продукты и системы 2016;30:11–5. [Kartiev S.B., Kureichik V.M. A classification algorithm

ЛИТЕРАТУРА

based on random forest principles to solve the forecasting problem. *Programmnyye produkty i sistemy = Software products and systems* 2016;30:11-5. (In Russian)].

19. Баширов А.Н., Воронов В.И. Прогнозирование вероятности возникновения бронхиальной астмы у детей с применением алгоритма случайного леса. *Современные наукоемкие технологии* 2019;12(2):249-55. [Bashirov A.N., Voronov V.I. Forecasting the probability of bronchial asthma in children using the random forest algorithm. *Sovremennyye naukoemkiye tekhnologii = Modern science-intensive technologies* 2019;12(2):249-55. (In Russian)]. <https://doi.org/10.17513/snt.37867>

20. Guido S., Mueller A. Introduction to Machine Learning with Python. O'Reilly Media Inc. 2016. [Electronic resource]. URL: https://inprogammer.com/wp-content/uploads/2023/02/Andreas-C.-Muller-Sarah-Guido-Introduction-to-Machine-Learning-with-Python_-A-Guide-for-Data-Scientists.pdf

21. «РД 52.04.186-89. Руководство по контролю загрязнения атмосферы». [Электронный ресурс]. [«RD 52.04.186-89. Guidelines for the control of atmospheric pollution». [Electronic resource]. (In Russian)]. URL: <https://www.consultant.ru/cons/cgi/online.cgi?from=10899-0&req=doc&rnd=GAJmA&base=ESU&n=24758#OV9YHCUR9NCURxO9>

22. Игнатьева Л.П., Потапова М.О., Чирцова М.В. Эколого-гигиенические критерии оценки загрязнения атмосферного воздуха: учебное пособие для студентов. ИГМУ 2022:79. [Ignatieva L.P., Potapova M.O., Chirtsova M.V. Ecological and hygienic criteria for assessing atmospheric air pollution: a textbook for students. IGМУ 2022:79].

23. Богданова А.М., Тымченко С.Л., Евстафьева Е.В. Алгоритм анализа данных медицинской статистики с оценкой экологически обусловленной заболеваемости в регионе. *Вычислительная биология и искусственный интеллект для*

персонализированной медицины 2022. [Bogdanova A.M., Tymchenko S.L., Evstafyeva E.V. Algorithm for analyzing medical statistics data with an assessment of ecologically determined morbidity in the region. *Vychislitel'naya biologiya i iskusstvennyy intellekt dlya personalizirovannoy meditsiny = Computational biology and artificial intelligence for personalized medicine* 2022. (In Russian)]. <http://dx.doi.org/10.14341/CBAI-2022-9>

24. Галеева Э.М., Галимова Р.Г., Теплова Д.С. К вопросу о комплексной оценке состояния окружающей среды в г. Уфа. *Российский журнал прикладной экологии* 2018;13(1):47-51. [Galeeva E.M., Galimova R.G., Teplova D.S. On the issue of a comprehensive assessment of the state of the environment in Ufa. *Rossiyskiy zhurnal prikladnoy ekologii = Russian Journal of Applied Ecology* 2018;13(1):47-51. (In Russian)].

25. Табаров С.Ф. Модель системы мониторинга и анализа состояния окружающей среды на примере субъектов РФ. *Вестник Воронежского государственного аграрного университета* 2023;8(2):207-26. [Tabarov S.F. Model of the environmental monitoring and analysis system on the example of the subjects of the Russian Federation. *Vestnik Voronezhskogo gosudarstvennogo agrarnogo universiteta = Bulletin of the Voronezh State Agrarian University* 2023;8(2):207-26. (In Russian)].

26. Табаров С.Ф. Методология отбора экологических показателей для оценки качества окружающей среды в регионах РФ. *Геополитика и экодинамика регионов* 2022;8(4):13-21. [Tabarov S.F. Methodology of selecting environmental indicators for assessing environmental quality in the regions of the Russian Federation. *Geopolitika i ekodinamika regionov = Geopolitics and ecodynamics of regions* 2022;8(4):13-21. (In Russian)].

Сведения об авторах:

Шадеркин И.А. – к.м.н., заведующий лабораторией электронного здравоохранения Института цифровой медицины Первого Московского государственного медицинского университета им. И.М. Сеченова (Сеченовский университет); Москва, Россия; PИNЦ Author ID 695560, <https://orcid.org/0000-0001-8669-2674>

Дьяченко А.П. – студентка Волгоградского государственного медицинского университета; Волгоград, Россия

Чулюкова Е.А. – студентка Московского государственного университета им. М.В. Ломоносова; Москва, Россия

Пшеничный Т.В. – студент Московского государственного университета им. М.В. Ломоносова; Москва, Россия

Ковека Л.В. – студентка Южно-Уральского государственного медицинского университета; Челябинск, Россия

Храмцова В.Е. – студентка Южно-Уральского государственного медицинского университета; Челябинск, Россия

Гугнявых Н.С. – студент Санкт-Петербургского государственного университета; Санкт-Петербург, Россия

Кузьмина А.И. – студентка Санкт-Петербургского государственного университета ветеринарной медицины; Санкт-Петербург, Россия

Вклад авторов:

Шадеркин И.А. – определение научного интереса, дизайн исследования, 30%
 Дьяченко А.П. – литературный обзор, статистические данные, 10%
 Чулюкова Е.А. – написание текста, математическая обработка, 10%
 Пшеничный Т.В. – написание текста, математическая обработка, 10%
 Ковека Л.В. – литературный обзор, статистические данные, 10%
 Храмцова В.Е. – литературный обзор, математическая обработка, 10%
 Гугнявых Н.С. – написание текста, статистические данные, 10%
 Кузьмина А.И. – написание текста, математическая обработка, 10%

Конфликт интересов: Автор заявляет об отсутствии конфликта интересов.

Финансирование: Исследование проведено без спонсорской поддержки.

Статья поступила: 11.11.23

Результат рецензирования: 18.12.23

Принята к публикации: 20.12.23

Information about authors:

Shaderkin I.A. – PhD, Head of the Laboratory of Electronic Health, Institute of Digital Medicine, Sechenov University; Moscow, Russia; info@uroweb.ru, RCSI Author ID 695560; <https://orcid.org/0000-0001-8669-2674>

Dyachenko A.P. – student of Volgograd State Medical University; Volgograd, Russia

Chulyukova E.A. – student of Lomonosov Moscow State University; Moscow, Russia

Pshenichny T.V. – student of Lomonosov Moscow State University; Moscow, Russia

Koveka L.V. – student of the South Ural State Medical University; Chelyabinsk, Russia

Khramtsova V.E. – student of the South Ural State Medical University; Chelyabinsk, Russia

Gugnyavykh N.S. – student of St. Petersburg State University; St. Petersburg, Russia

Kuzmina A.I. – student of St. Petersburg State University of Veterinary Medicine; St. Petersburg, Russia

Authors Contribution:

Shaderkin I.A. – identification of scientific interest, research design, 30%
 Dyachenko A.P. – literature review, statistical data, 10%
 Chulyukova E.A. – text writing, mathematical processing, 10%
 Wheat T.V. – text writing, mathematical processing, 10%
 Koveka L.V. – literature review, statistical data, 10%
 Khramtsova V.E. – literature review, mathematical processing, 10%
 Gugnyavykh N.S. – text writing, statistical data, 10%
 Kuzmina A.I. – text writing, mathematical processing, 10%

Conflict of interest. The author declare no conflict of interest.

Financing. The study was performed without external funding.

Received: 11.11.23

Review result: 18.12.23

Accepted for publication: 20.12.23