

УДК 004.8

DOI: [10.26102/2310-6018/2022.36.1.004](https://doi.org/10.26102/2310-6018/2022.36.1.004)

Анализ влияния пандемии COVID-19 на развитие человеческого капитала региона с помощью алгоритмов машинного обучения

И.Л. Каширина[✉], Т.В. Азарнова, Ю.В. Бондаренко

*Воронежский государственный университет, Воронеж, Российская Федерация
kash.irina@mail.ru[✉]*

Резюме. Пандемия коронавирусной инфекции COVID-19 оказала большое влияние на формирование и развитие человеческого капитала за счет своего негативного воздействия на образование и здоровье населения. Это заболевание уже унесло сотни тысяч жизней и вызвало длительные нарушения здоровья людей, а также лишило многих доступа к качественному образованию. В связи с этим в период пандемии COVID-19 большое значение приобретает разработка современных и точных методов анализа, моделирования и прогнозирования динамики распространения этого заболевания, позволяющих выявить факторы, оказывающие существенное влияние на процесс распространения инфекции. В статье обсуждаются этапы построения моделей машинного обучения для проведения предикативного анализа заболеваемости COVID-19, позволяющего исследовать динамику распространения данного вируса на региональном уровне, выявить влияние различных факторов на степень тяжести, длительность протекания болезни и впоследствии разработать своевременные сценарии управления человеческим капиталом региона с целью снижения негативного воздействия пандемии. Для разработки методов использовался большой массив деперсонифицированных данных по распространению COVID-19 в Воронежской области, предоставленных Воронежским областным клиническим консультативно-диагностическим центром (ВОККДЦ). В статье представлены результаты разведочного анализа имеющихся данных, выявлены дополнительные признаки, которые могут быть использованы для построения моделей машинного обучения, разработаны методы интерактивной визуализации и прогнозирования динамики COVID-19.

Ключевые слова: человеческий капитал, COVID-19, машинное обучение, прогнозирование динамики, разведочный анализ данных.

Благодарности: Исследование выполнено при поддержке РФФИ, проект 19-29-07400.

Для цитирования: Каширина И.Л., Азарнова Т.В., Бондаренко Ю.В. Анализ влияния пандемии COVID-19 на развитие человеческого капитала региона с помощью алгоритмов машинного обучения. *Моделирование, оптимизация и информационные технологии*. 2022;10(1). Доступно по: <https://moitvvt.ru/ru/journal/pdf?id=1137> DOI: 10.26102/2310-6018/2022.36.1.004

Analysis of the COVID-19 pandemic impact on the development of human capital in the region using machine learning algorithms

I.L. Kashirina[✉], T.V. Azarnova, Yu.V. Bondarenko

*Voronezh State University, Voronezh, Russian Federation
kash.irina@mail.ru[✉]*

Abstract: The COVID-19 pandemic has had a major impact on the formation and development of human capital through its negative effect on education and public health. This disease has already claimed hundreds of thousands of lives, caused long-term health problems and deprived many of them of access to quality education. In this regard, during the COVID-19 pandemic, it is of great importance to design modern and accurate methods for analyzing, modeling and predicting the dynamics of the spread of this

disease, which enable to identify factors that significantly affect the spread of the infection. The article discusses the stages of constructing machine learning models for conducting a predicative analysis of the COVID-19 incidence, which makes it possible to study the dynamics of the spread of this virus at the regional level, identify the influence of various factors on the severity, the duration of the disease, and subsequently create timely scenarios for managing the human capital of the region in order to reduce the negative impact of the pandemic. To devise the methods, a large array of depersonalized data on the spread of COVID-19 in the Voronezh region, provided by the Voronezh Regional Clinical Consultative and Diagnostic Center, was used. The article presents the results of an exploratory analysis of the available data, highlights additional features that can be employed to build machine learning models and develops methods for interactive visualization and forecasting of COVID-19 dynamics.

Keywords: human capital, COVID-19, machine learning, trend forecasting, exploratory data analysis.

Acknowledgments: The research was supported by the Russian Foundation for Basic Research (RFBR), project 19-29-07400.

For citation: Kashirina I.L., Azarnova T.V., Bondarenko Yu.V. Analysis of the COVID-19 pandemic impact on the development of human capital in the region using machine learning algorithms.

Modeling, Optimization and Information Technology. 2022;10(1). Available from:

<https://moitvvt.ru/ru/journal/pdf?id=1137> DOI: 10.26102/2310-6018/2022.36.1.004 (In Russ).

Введение

Во время пандемии COVID-2019 на рынке труда произошли глобальные и существенные изменения, что привело к переоценке человеческого капитала. Пандемия COVID-2019, которая охватила весь мир, практически остановила функционирование многих сфер деятельности. В условиях пандемии роль и функции человеческого капитала кардинально изменились в связи с изменившимися условиями. Доклад группы Всемирного Банка “COVID-19 и человеческий капитал”, выпущенный в конце 2020 года, посвящен оценке влияния пандемии COVID-19 на показатели состояния здоровья и образования населения [1]. В докладе отмечается, что в условиях пандемии значительно возросли требования к обоснованности стратегических решений, принимаемых в области управления человеческим капиталом. Принимаемые решения должны учитывать достоверные прогнозы развития эпидемиологической ситуации в отдельных регионах и стране в целом. При этом большой популярностью при моделировании эпидемиологических процессов в последнее время пользуются алгоритмы машинного обучения, которые сами способны находить закономерности и аппроксимировать зависимости, опираясь на имеющуюся базу наблюдений [2-4]. Для обучения таких моделей и разработки точных прогнозных инструментов необходимо применение деперсонифицированных баз медицинских данных, а также выявление дополнительных факторов, которые оказывают влияние на развитие эпидемиологического процесса [5].

Таким образом, в настоящее время существует необходимость в разработке с помощью инструментов машинного обучения специальных интеллектуальных подходов к исследованию различных аспектов распространения эпидемии COVID-19 в разрезе их влияния на человеческий капитал на региональном уровне.

Процесс разработки подобных подходов можно разбить на следующие этапы.

1. Разведочный анализ данных, включающий выявление и исследование признаков, оказывающих влияние на распространение COVID-19.

2. Разработка инструментов краткосрочного и долгосрочного прогнозирования динамики волн распространения эпидемии с применением гибридных технологий моделирования нелинейных зависимостей.

3. Разработка новых инструментов визуализации, позволяющих отображать в режиме ежедневного обновления данные по заболеваемости COVID-19 с интерактивным отображением на масштабируемой гугл-карте региона.

4. Построение модели профиля региона, отражающей наиболее существенные с точки зрения распространения COVID 19 показатели функционирования региона, и разработка инструментов предсказания зонального уровня массовости и тяжести протекания эпидемиологического процесса с учетом анализа траектории динамики развития заболевания в России в целом и эффектов кросс корреляции с другими регионами.

5. Разработка нейросетевых инструментов прогнозирования динамики госпитализаций, позволяющих учитывать ретроспективную динамику балансовых соотношений количества заболевших по районам области.

6. Разработка методов сценарного анализа развития ситуации в регионе, позволяющих оценить вероятность различных сценариев развития эпидемии, с учетом внешних и внутренних факторов регулирования ситуации.

В статье описана реализация 1, 2 и 3 этапов данного исследования. Целью исследования является построение моделей машинного обучения для проведения предикативного анализа заболеваемости COVID 19, позволяющих моделировать динамику распространения данного вируса на региональном уровне, выявлять влияние различных факторов на количество госпитализаций и впоследствии разработать своевременные сценарии управления человеческим капиталом региона с целью снижения негативного воздействия пандемии.

Описание исходных данных

В качестве исходных данных для построения моделей машинного обучения использовался набор деперсонифицированных данных, предоставленных Воронежским областным клиническим консультативно-диагностическим центром (ВОККДЦ), включающий данные обо всех ПЦР тестах на COVID-19, которые были проведены в Воронежской области в период с марта 2020 года по декабрь 2021 года.

id	Пол	Дата забора	Результат	Район	Проводящая исследование МО	Направлен	Стационар тяжёлый	Осложнения после ковид	Прошел первый этап УД	Вакцина более двух недель назад	Возраст
a8961435	Женский	2021-07-15	+	Павловский район	БУЗ ВО "ВГКП № 7"	Амбулаторно	Нет	Нет	Нет	Нет	46
bfc6eb1	Женский	2021-08-10	-	Семилукский район	АУЗ ВО "ВОККДЦ"	Амбулаторно	Нет	Нет	Нет	Нет	71
e37a4c7d	Женский	2021-08-11	+	городской округ Воронеж	БУЗ ВО "ОДКБ №2"	Стационарно	Нет	Да	Нет	Нет	61
184974b5	Мужской	2021-07-31	+	городской округ Воронеж	БУЗ ВО "ВГКП № 7"	Амбулаторно	Нет	Нет	Нет	Нет	42
4c71edb3	Женский	2021-07-19	-	городской округ Воронеж	БУЗ ВО "ОДКБ №2"	Амбулаторно	Нет	Нет	Нет	Нет	9
...

Рисунок 1– Фрагмент исходных данных
Figure 1 – Fragment of initial data

Датасет содержит следующие показатели (Рисунок 1): уникальный id пациента; пол; возраст; дата забора теста; результат теста (положительный или отрицательный); район Воронежской области, в котором проживает пациент; медицинская организация, которая проводила тестирование; тест сдан амбулаторно или в стационаре; был ли тест сдан в одном из стационаров, в которые направляются преимущественно пациенты с тяжелыми случаями заболевания; зарегистрированы ли у данного пациента осложнения

после перенесенного заболевания COVID-19, состоит ли пациент на диспансерном учете для реабилитации после COVID-19, был ли пациент привит вторым компонентом вакцины более 2 недель назад. База данных постоянно пополняется. На 1 декабря 2021 года она содержала более 1.6 миллиона записей, включающих сведения о результатах ПЦР тестирования 665 тысяч уникальных пациентов.

Разведочный анализ данных

Исследуем распределения значений показателей, представленных в исходном наборе данных. Столбец “Результат” содержит положительные значения у 361852 уникальных пациентов, сдававших тест ПЦР.

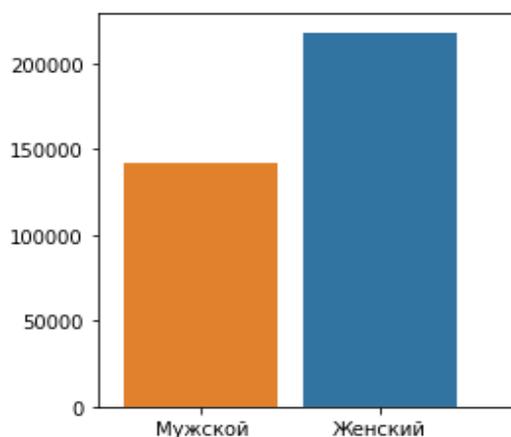


Рисунок 2 – Соотношение мужчин и женщин с ПЦР+
Figure 2 – Ratio of men and women with PCR+

Согласно значениям в столбце “Пол”, исходный датасет содержит 60 % женщин и 40 % мужчин, то есть женщины в Воронежской области сдают ПЦР тесты в полтора раза чаще, чем мужчины. Если рассмотреть только случаи положительных тестов, то среди них пропорции такие же: 60 % женщин и 40 % мужчин. Исходя из этого можно сделать предположение, что заболеваемость COVID-19 не зависит от пола.

Средний возраст пациента в исходном датасете составляет 44 года, среди пациентов с ПЦР+ 47 лет. На Рисунке 3 приведено распределение по возрасту для пациентов с ПЦР+ (отдельно для женщин и мужчин).

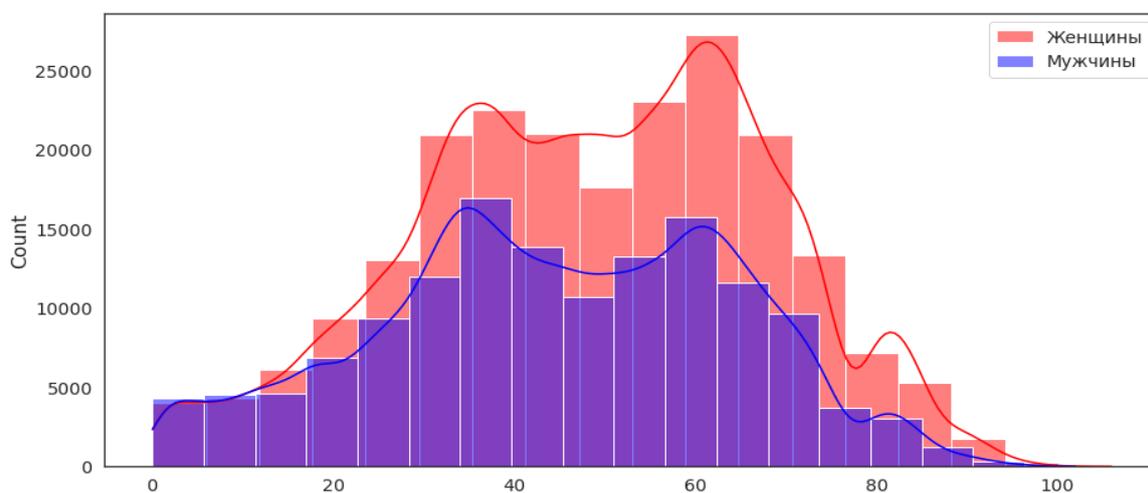


Рисунок 3– Распределение по возрасту пациентов с ПЦР+
Figure 3 – Age distribution of patients with PCR+

Как видно, распределения являются бимодальными – первый пик приходится на возраст 35-40 лет, второй – 60-65 лет. До апреля 2021 года модальный возраст в 35-40 лет на графике отсутствовал, возможно, его появление связано со штаммом дельта, которым болели и более молодые люди.

На Рисунке 4 представлен сравнительный график заболеваемости на 1 декабря 2021 года в районах Воронежской области за весь период пандемии (общее число пациентов района, имевших положительный тест ПЦР в пересчете на тысячу жителей этого района). Видно, что число ПЦР+ пациентов в разных районах существенно отличается (что, возможно, связано с доступностью тестирования, в числе прочих факторов).

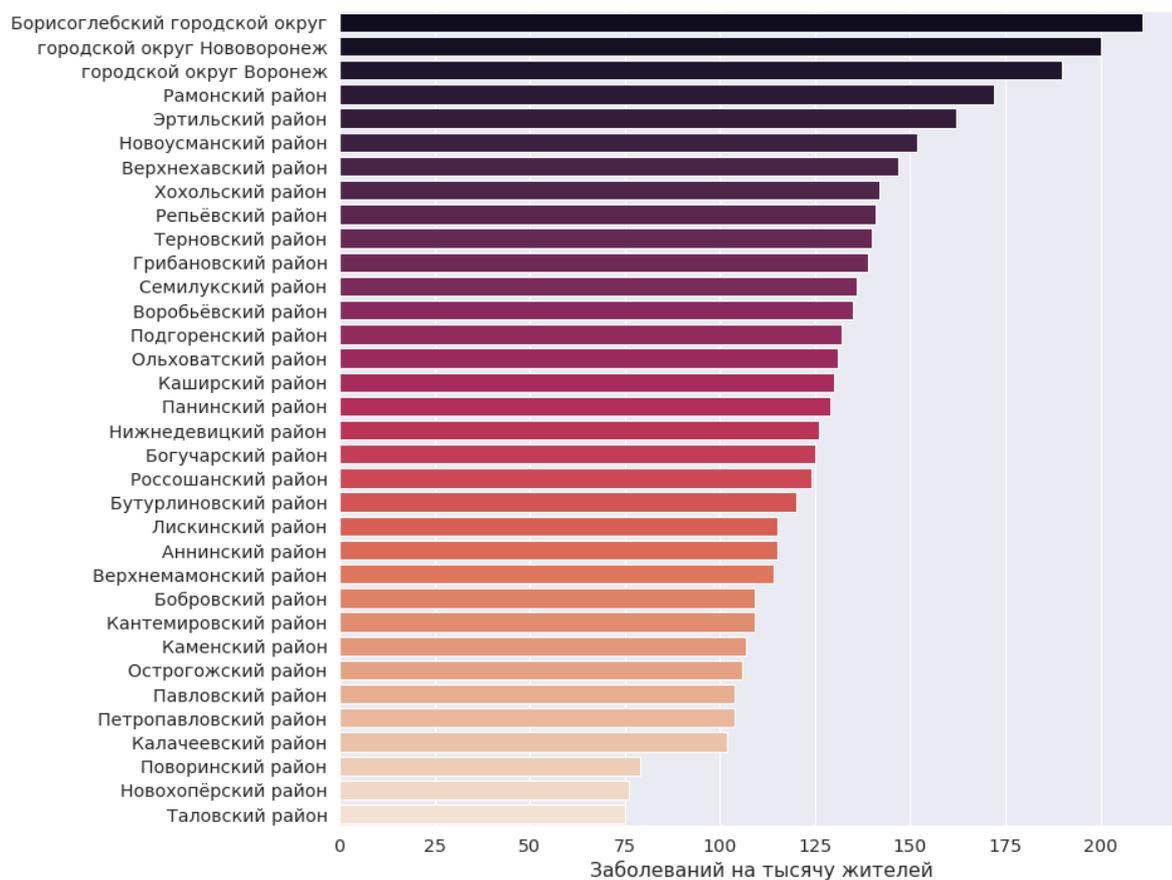


Рисунок 4– Сравнение районов по заболеваемости.
Figure 4– Comparison of regions by incidence.

Динамика активных случаев заболевания в целом по Воронежской области с начала пандемии изображена на Рисунке 5. Следует отметить, что заболевание считалось активным в период от первого положительного до первого отрицательного теста у данного пациента. В случаях, когда повторные отрицательные тесты у пациента отсутствовали, заболевание считалось активным в течение двух недель после появления первого положительного теста.

В процессе исследования динамики активных заболеваний было обнаружено, что в датасете есть ряд пациентов, у которых на протяжении длительного времени (2-3 месяца подряд) с промежутками менее двух недель выявлялись положительные тесты (так называемый синдром лонг-ковида). В связи с чем дополнительно был проведен анализ продолжительности заболевания (исследовалось распределение времени между

первым положительным и первым отрицательным после него тестом). Средняя продолжительность заболевания составила 15 дней (со средним квадратическим отклонением в 6 дней), модальная – 14 дней.

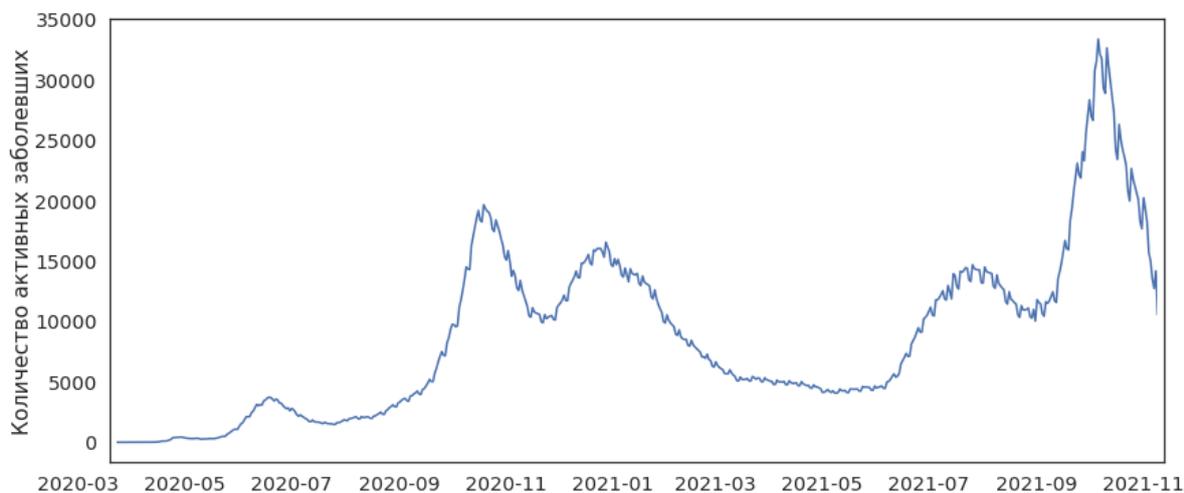


Рисунок 5 – Динамика активных случаев заболевания.

Figure 5 – Dynamics of active cases of the disease.

Дополнительно были исследованы случаи повторных заболеваний, чтобы понять, как учитывать в моделях прогнозирования динамики заболеваемости случаи повторных заболеваний. Заболевание считалось повторным, если у одного и того же пациента через 3 и более месяцев повторно регистрировался положительный тест. Таких случаев в выборке пока достаточно мало – для повторно заболевших (среди всех пациентов с ПЦР+) составляет чуть более 1 процента (1.17 %). На Рисунке 7 приводится гистограмма распределения периода между повторными заболеваниями. Среднее значение этого периода составляет 282 дня со средним квадратическим отклонением в 129 дней.

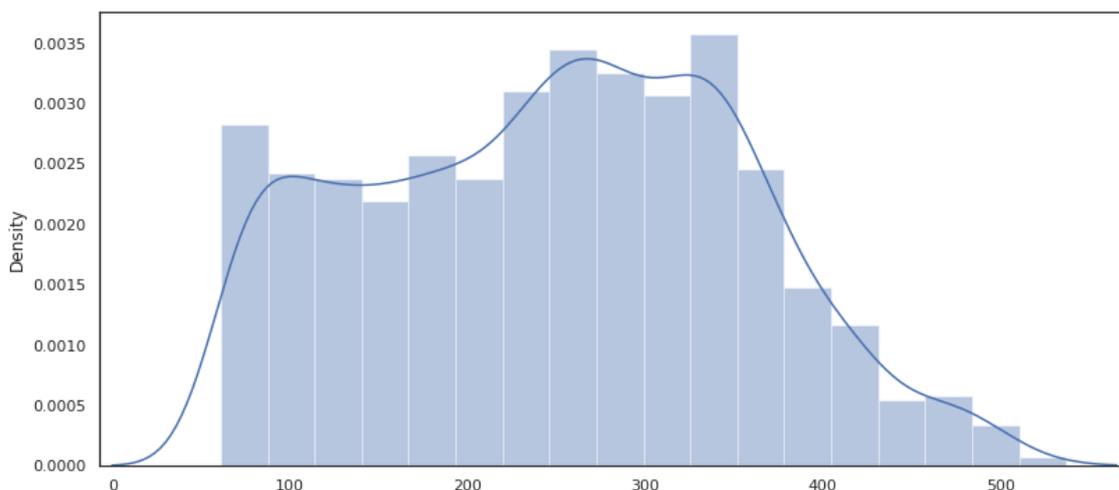


Рисунок 6 – Распределение периода между повторными заболеваниями.

Figure 6 – Distribution of the period between recurrent diseases.

Это распределение не отвечает гипотезе о нормальности, среди повторно заболевших достаточно высока доля тех, кто вновь заразился через 3-4 месяца, и эта доля растет.

Для выявления дополнительных признаков, оказывающих влияние на распространение пандемии, был проведен корреляционный анализ взаимосвязи между случаями COVID-19 и некоторыми внешними факторами. Для оценки корреляции были рассмотрены такие характеристики, как число новых случаев ПЦР+ в определенный день в Воронежской области, скользящие средние для температуры, количество осадков и новых случаев ПЦР+ за предшествующую неделю, число активных случаев в предыдущий день, а также количество запросов по поисковым словам “Лечение коронавируса” (запрос 1), “Covid” (запрос 2) и “Вызвать скорую” (запрос 3) за предшествующую неделю.

Статистика запросов была взята с сайта <https://wordstat.yandex.ru>. История запросов Yandex – это один из информационных ресурсов, который содержит полезную информацию для прогнозирования и оценки эпидемии COVID-19. На Рисунке 7 представлен фрагмент выдачи истории запросов по слову “Covid” в Воронежской области за период с марта 2020 по январь 2022 года. Во время пандемии COVID-19 было опубликовано несколько исследований с использованием веб-данных, которые показали, что данные выдачи поисковых систем могут быть полезны для прогнозирования дальнейшего развития эпидемии [5, 6, 7]. Для поиска новых методов и подходов к эпиднадзору за заболеваниями крайне важно использовать данные Интернета, которые обновляются в режиме реального времени. При анализе связи заболеваемости с числом поисковых запросов был обнаружен интересный факт – пики в запросах, связанных с COVID, происходили за 3-5 дней до того, как достигался пик по новым заболевшим в базе ПЦР-тестов. Это может быть связано с тем, что люди, которые чувствуют себя плохо, сначала ищут информацию в сети интернет, прежде чем обратиться к врачу и сдадут ПЦР-тест.



Рисунок 7– История показов по фразе COVID в Воронежской области.
Figure 7 – History of impressions for the phrase COVID in the Voronezh region.

Метеорологические данные анализа были использованы потому, что в предыдущих исследованиях было отмечено, что они тоже могут оказывать влияние на заболеваемость [8].

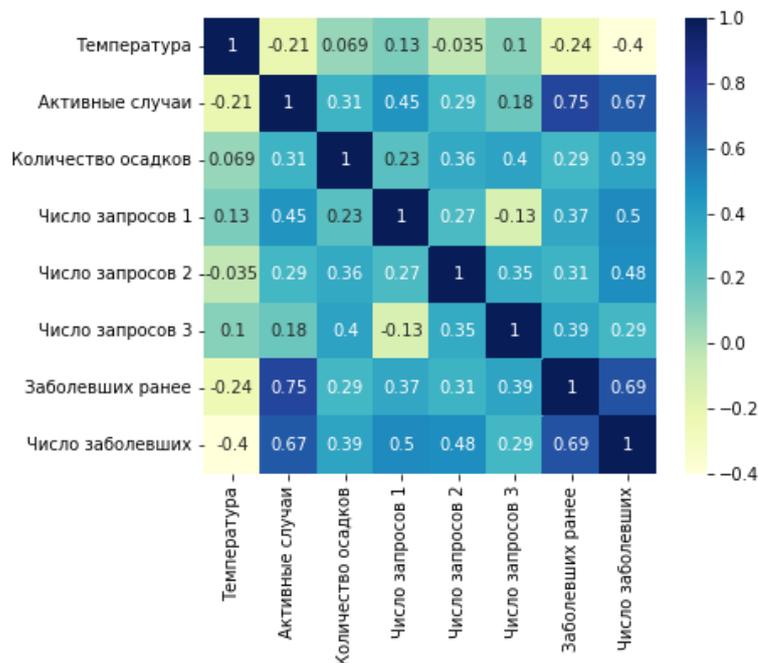


Рисунок 8 – Корреляция между новыми случаями COVID и другими факторами
Figure 8 – Correlation between new COVID cases and other factors

Матрица корреляции, приведенная на Рисунке 8, показывает степень линейной зависимости перечисленных признаков между собой. Видно, что самая сильная корреляция числа заболевших в определенный день наблюдается с числом заболевших ранее и числом активных случаев. Однако все перечисленные показатели оказались значимы, включая температуру и количество осадков.

Прогнозирование новых случаев

Для построения базовой (baseline) модели прогнозирования динамики распространения COVID-19 в Воронежском регионе использовалась библиотека Prophet. Prophet – это программное обеспечение с открытым исходным кодом, разработанное научной группой Facebook по обработке данных, реализованное в R и Python. Prophet содержит процедуру прогнозирования данных временных рядов на основе аддитивной модели, в которой нелинейные тенденции соответствуют годовой, еженедельной и ежедневной сезонности, а также учитывает эффекты праздников [10]. Prophet устойчива к изменениям тренда и, как правило, хорошо справляется с выбросами в данных. Прогнозное значение (в момент времени t) $y(t)$ определяется следующим образом:

$$y(t) = g(t) + s(t) + h(t) + \varepsilon, \quad \text{где}$$

$g(t)$ – неперіодические изменения в данных временного ряда, которые моделируются с использованием функции логистического роста;

$s(t)$ – периодические изменения в данных (например, еженедельная / годовая сезонность);

$h(t)$ – праздничные эффекты с нерегулярным расписанием;

ε – случайные изменения, не учитываемые моделью.

Модель логистического роста, используемая при определении тренда, соответствует логистической функции активации нейронной сети и имеет следующий вид:

$$g(t) = C / (1 + \exp(-k(t - m))) ,$$

где параметр C указывает на силу роста, k указывает скорость роста, а m -параметр смещения. Параметры C , k и m настраиваются по обучающей выборке.

Prophet использует ряды Фурье для прогнозирования эффектов сезонности:

$$s(t) = \sum_{n=1}^N \left(a_n \cos \frac{2n\pi t}{P} + b_n \sin \frac{2n\pi t}{P} \right),$$

где P – период, а параметры N , a_1, a_2, \dots, a_N и b_1, b_2, \dots, b_N оцениваются по обучающей выборке.

На Рисунке 9 синим цветом изображены реальные данные динамики новых случаев, красным цветом – прогноз, полученный с помощью модели Prophet. Данные до 1 июля 2021 года использовались для обучения модели (на графике можно увидеть построенную модель аппроксимацию обучающей выборки), последующие дни – для прогнозирования. На графике бирюзовым цветом изображен доверительный интервал полученного прогноза. В качестве метрик точности прогноза использовалась средняя абсолютная ошибка MAE и средняя процентная ошибка MAPE:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \tilde{y}_i|, \quad MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \tilde{y}_i|}{y_i} * 100\%$$

Видно, что с увеличением интервала прогнозирования, точность прогноза падает. Средняя абсолютная ошибка, рассчитанная для прогноза динамики на 15 дней вперед, составила MAE= 50.8. Средняя процентная ошибка для этого же периода составила MAPE = 6 %, то есть прогнозное значение в среднем на 6 % отклонялось от реального.

Библиотека Prophet строит прогнозы, опираясь на предшествующую динамику временного ряда. Однако, она не предоставляет возможности учитывать дополнительные признаки (такие, как метеорологические факторы или количество поисковых запросов).

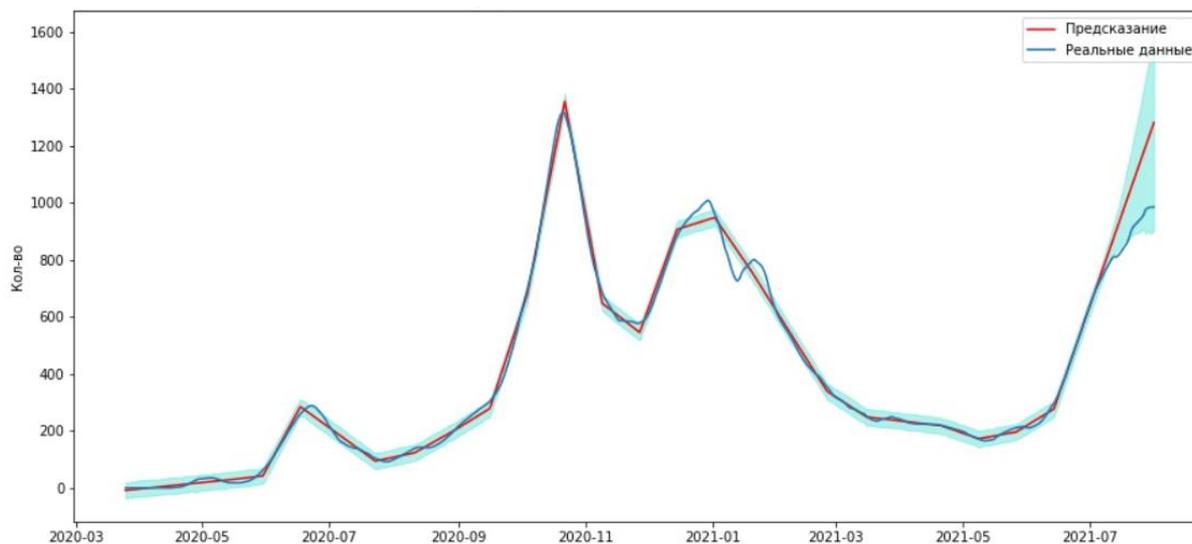


Рисунок 9– Прогнозирование динамики новых случаев COVID-19 в Воронежском регионе
Figure 9 – Forecasting the dynamics of new cases of COVID-19 in the Voronezh region

Поэтому построенная с использованием этой библиотеки модель прогнозирования будет использоваться в качестве базовой (baseline), то есть как отправная точка для последующего повышения точности, а в качестве основного подхода разрабатывается

гибридная модель глубокого обучения, которая включает рекуррентные сети долгой краткосрочной памяти (LSTM), позволяющие на входе модели использовать несколько временных рядов, при этом признаки из временных рядов будут автоматически извлекаться с помощью сверточной нейронной сети (CNN) [10].

Визуализация заболеваемости на масштабируемой гугл-карте региона

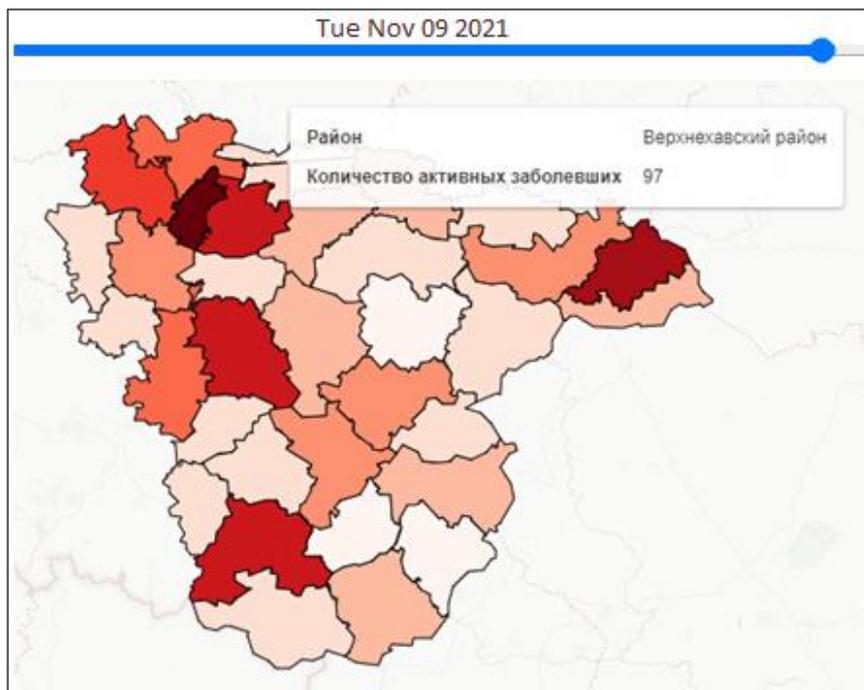


Рисунок 10– Визуализация активных случаев заболевания
 Figure 10 – Visualization of active cases

В рамках данного исследования разработан новый инструмент визуализации, позволяющий отображать данные по заболевшим COVID-19 (активные случаи), а также новые случаи заболевания с интерактивным отображением на масштабируемой гугл-карте области.

Инструмент поможет наглядно визуализировать различия между районами, которые являются горячими точками COVID-19. С помощью специализированного плагина Choropleth (реализованного в библиотеке python geopandas для анализа и визуализации географических данных) можно детально рассмотреть, как меняется ситуация в различных районах города и области на протяжении всего времени пандемии (изменять дату можно с помощью встроенного ползунка времени Time Slider). На Рисунке 10 в качестве примера представлена визуализация активных случаев заболевания на 9 ноября 2021 года, всплывающие подсказки содержат название района и число заболевших.

Заключение

В предлагаемом исследовании с использованием инструментов визуализации проанализированы распределения основных признаков, влияющих на распространение пандемии COVID-19, исследованы закономерности повторных случаев заболевания, построена начальная модель прогнозирования динамики распространения COVID-19.

Ошибка построенной модели составила 6 % по метрике MAPE, что несколько превышает желательный 5 % интервал для погрешности модели по данной метрике.

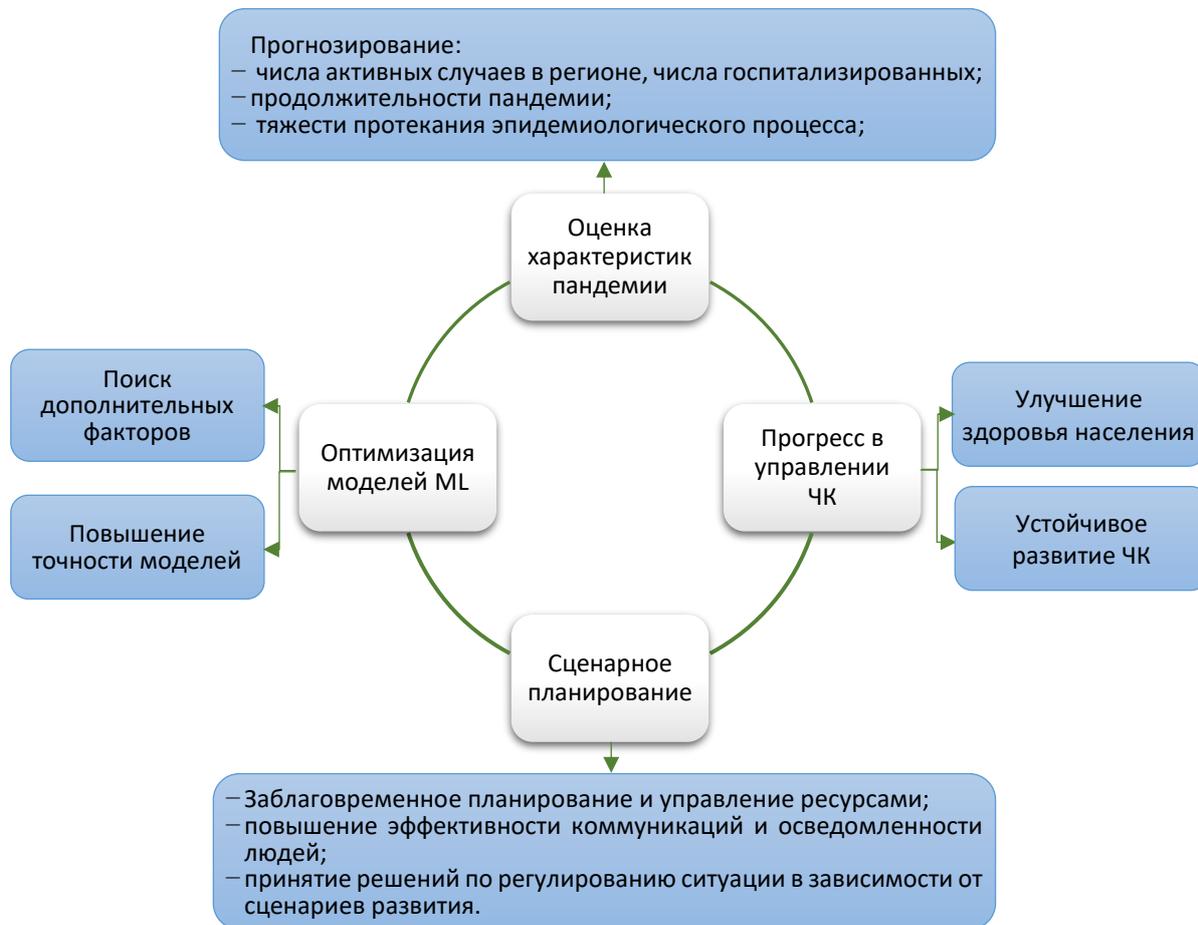


Рисунок 11 – Использование моделей машинного обучения (ML) для управления человеческим капиталом (ЧК) региона в условиях пандемии Covid 19

Figure 11 – Application of machine learning models to manage the human capital of the region in the context of the Covid 19 pandemic

Для повышения точности прогнозирования ведется разработка гибридной нейросетевой модели глубокого обучения CNN-LSTM, которая будет использовать дополнительные признаки, описанные в данном исследовании.

Построенные модели могут быть использованы для принятия своевременных управляющих решений, направленных на снижение негативных последствий пандемии на человеческий капитал. Практическое значение моделей прогнозирования случаев инфицирования COVID-19 для управления развитием человеческого капитала включает в себя оценку характеристик пандемии для данного прогнозируемого периода и местности, сценарное планирование в секторе управления регионом, оптимизацию моделей и оценку корреляции различных факторов с динамикой заболеваемости (Рисунок 11).

СПИСОК ИСТОЧНИКОВ

1. World Bank. Europe and Central Asia Economic Update, Fall 2020: *COVID-19 and Human Capital*. Washington, DC: World Bank, 2020. <https://openknowledge.worldbank.org/handle/10986/34518>.
2. Yadaw A.S. Clinical features of COVID-19 mortality: development and validation of a

- clinical prediction model. *Lancet Digit Health*. 2020;2:516–525.
- Anastassopoulou C. Data-based analysis, modelling and forecasting of the COVID-19 outbreak. *PLOS ONE*. 2020;15:1–21.
 - Chaurasia V. Application of machine learning time series analysis for prediction covid-19 pandemic. *Research on Biomedical Engineering*. 2020:1–13.
 - Kashirina I., Bondarenko Y., Azarnova T. Analysis and forecasting of the market of educational services of the region. *Proceedings – 2021 1st International Conference on Technology Enhanced Learning in Higher Education, TELE 2021*. Lipetsk; 2021. P. 30–34.
 - Ayyoubzadeh S., Ayyoubzadeh S., Zahedi H., Ahmadi M., Niakan Kalhori S. Predicting COVID-19 Incidence Through Analysis of Google Trends Data in Iran: Data Mining and Deep Learning Pilot Study. *JMIR Public Health Surveill*. 2020;6(2):e18828.
 - Prediction of COVID-19 Outbreaks Using Google Trends in India: A Retrospective Analysis Healthc Inform Res*. 2020;26(3):175–184.
 - Mavragani A., Gkillas K. COVID-19 predictability in the United States using Google Trends time series. *Sci Rep*. 2020;10:20693.
 - Фирюлина М.А., Каширина И.Л. Прогнозирование развития инфаркта миокарда на основании сезонных и метеорологических факторов. *Вестник Воронежского института высоких технологий*. 2021;2(37):19–24.
 - Devaraj J., Elavarasan R.M., Pugazhendhi R., Shafiullah G., Ganesan S., Jeysree A.K., Khan I.A., Hossain E. Forecasting of COVID-19 cases using deep learning models: Is it reliable and practically significant? *Results Phys*. 2021;21:e103817.
 - Mahanty M., Swathi K., Teja K.S., Kumar P.H., Sravani A. Forecasting the spread of COVID-19 pandemic with Prophet. *Revue d'Intelligence Artificielle*. 2021;35(2):115–122.
 - Ketu S., Mishra P.K. India perspective: CNN-LSTM hybrid deep learning model-based COVID-19 prediction and current status of medical resource availability. *Soft Comput*. 2022;26:645–664.
 - Яковенко Н.В., Каширина И.Л., Щепина И.Н. и др. *Человеческий капитал как драйвер развития цифровой экономики региона: теоретические и инструментальные основы исследования*. Воронеж: Цифровая полиграфия; 2020. 234 с.

REFERENCES

- World Bank. Europe and Central Asia Economic Update, Fall 2020: *COVID-19 and Human Capital*. Washington, DC: World Bank, 2020. <https://openknowledge.worldbank.org/handle/10986/34518>.
- Yadaw A.S. Clinical features of COVID-19 mortality: development and validation of a clinical prediction model. *Lancet Digit Health*. 2020;2:516–525.
- Anastassopoulou C. Data-based analysis, modelling and forecasting of the COVID-19 outbreak. *PLOS ONE*. 2020;15:1–21.
- Chaurasia V. Application of machine learning time series analysis for prediction covid-19 pandemic. *Research on Biomedical Engineering*. 2020:1–13.
- Kashirina I., Bondarenko Y., Azarnova T. Analysis and forecasting of the market of educational services of the region. *Proceedings – 2021 1st International Conference on Technology Enhanced Learning in Higher Education, TELE 2021*. Lipetsk; 2021. P. 30–34.
- Ayyoubzadeh S., Ayyoubzadeh S., Zahedi H., Ahmadi M., Niakan Kalhori S. Predicting COVID-19 Incidence Through Analysis of Google Trends Data in Iran: Data Mining and Deep Learning Pilot Study. *JMIR Public Health Surveill*. 2020;6(2):e18828.

7. Prediction of COVID-19 Outbreaks Using Google Trends in India: A Retrospective Analysis *Health Inform Res.* 2020;26(3):175–184.
8. Mavragani A., Gkillas K. COVID-19 predictability in the United States using Google Trends time series. *Sci Rep.* 2020;10:20693.
9. Firyulina, M. A., Kashirina I. L. Forecasting the development of myocardial infarction based on seasonal and meteorological factors. *Vestnik Voronezhskogo instituta vysokikh tekhnologii = Bulletin of the Voronezh Institute of High Technologies.* 2021;2(37):19–24. (In Russ.)
10. Devaraj J., Elavarasan R.M., Pugazhendhi R., Shafiullah G., Ganesan S., Jeysree A.K., Khan I.A., Hossain E. Forecasting of COVID-19 cases using deep learning models: Is it reliable and practically significant? *Results Phys.* 2021;21:e103817.
11. Mahanty M., Swathi K., Teja K.S., Kumar P.H., Sravani A. Forecasting the spread of COVID-19 pandemic with Prophet. *Revue d'Intelligence Artificielle.* 2021;35(2):115–122.
12. Ketu S., Mishra P.K. India perspective: CNN-LSTM hybrid deep learning model-based COVID-19 prediction and current status of medical resource availability. *Soft Comput.* 2022;26:645–664.
13. Yakovenko N.V., Kashirina I.L., Shchepina I.N. and others. *Human capital as a driver for the development of the digital economy of the region: theoretical and instrumental foundations of the study.* Voronezh: Digital printing; 2020. 234 p. (In Russ.)

ИНФОРМАЦИЯ ОБ АВТОРАХ / INFORMATION ABOUT THE AUTHORS

Каширина Ирина Леонидовна, д.т.н., профессор, кафедра математических методов исследования операций, Воронежский государственный университет, Воронеж, Российская Федерация.
e-mail: kash.irina@mail.ru
ORCID: [0000-0002-8664-9817](https://orcid.org/0000-0002-8664-9817)

Irina Leonidovna Kashirina, Doctor of Technical Sciences, Professor, Mathematical Methods of Operations Research Department, Voronezh state university, Voronezh, Russian Federation.

Азарнова Татьяна Васильевна, д.т.н., заведующий кафедрой математических методов исследования операций, Воронежский государственный университет, Воронеж, Российская Федерация.
e-mail: ivdas92@mail.ru
ORCID: [0000-0001-6342-9355](https://orcid.org/0000-0001-6342-9355)

Azarnova Tatyana Vasilievna, Doctor of Technical Sciences, Head of the Department of Mathematical Methods of Operations Research, Voronezh State University, Voronezh, Russian Federation.

Бондаренко Юлия Валентиновна, д.т.н., профессор, кафедра математических методов исследования операций, Воронежский государственный университет, Воронеж, Российская Федерация.
e-mail: bond.julia@mail.ru
ORCID: [0000-0002-5442-9123](https://orcid.org/0000-0002-5442-9123)

Bondarenko Yulia Valentinovna, Doctor of Technical Sciences, Professor, Department of Mathematical Methods of Operations Research, Voronezh State University, Voronezh, Russian Federation.

Статья поступила в редакцию 26.01.2022; одобрена после рецензирования 15.02.2022; принята к публикации 22.02.2022.

The article was submitted 26.01.2022; approved after reviewing 15.02.2022; accepted for publication 22.02.2022.