

doi: 10.17586/2226-1494-2022-22-5-970-981

УДК 004.89

Прогнозирование летального исхода у пациентов с установленным диагнозом COVID-19

Игорь Николаевич Корсаков¹, Татьяна Леонидовна Каронова²✉,
Александра Олеговна Конради³, Аркадий Дмитриевич Рубин⁴,
Дмитрий Ильич Курапеев⁵, Алёна Тимуровна Черникова⁶,
Арина Алексеевна Михайлова⁷, Евгений Владимирович Шляхто⁸

^{1,2,3,4,5,6,7,8} Национальный медицинский исследовательский центр им. В.А. Алмазова, Санкт-Петербург, 197341, Российская Федерация

¹ korskakov_in@almazovcentre.ru, <https://orcid.org/0000-0003-2343-9641>

² karonova@mail.ru✉, <https://orcid.org/0000-0002-1547-0123>

³ Konradi_ao@almazovcenter.ru, <https://orcid.org/0000-0001-8169-7812>

⁴ rubin_ad@almazovcentre.ru, <https://orcid.org/0000-0001-5469-5635>

⁵ dkurapeev@gmail.com, <https://orcid.org/0000-0002-2190-1495>

⁶ arabicaa@gmail.com, <https://orcid.org/0000-0002-4878-6909>

⁷ armikhaylova@yandex.ru, <https://orcid.org/0000-0001-6066-3525>

⁸ e.shlyakhto@almazovcentre.ru, <https://orcid.org/0000-0003-2929-0980>

Аннотация

Предмет исследования. Появление нового коронавируса SARS-CoV-2 является основой развития пандемии COVID-19. Пандемия привела к резкому росту нагрузки на системы здравоохранения высокой летальности и существенным трудностям в организации медицинской помощи. Для прогнозирования течения заболевания и определения показаний к более агрессивному лечению предложено множество различных клинических и биологических маркеров. Однако не всегда клинико-лабораторная оценка состояния точно предсказывает развитие тяжелого течения болезни. В этом случае технологии, основанные на искусственном интеллекте, могут существенно влиять на оценку прогнозирования. Выполнен системный анализ факторов, влияющих на течение инфекционного заболевания у пациентов с диагностированным COVID-19, госпитализированных в стационар.

Метод. Предложенный подход основан на применении машинного обучения для прогнозирования летального исхода у пациентов с установленным диагнозом COVID-19. В основе подхода лежит анализ анамнеза пациента и клинические, лабораторные и инструментальные данные, полученные в первые 72 часа нахождения пациента в стационаре. **Основные результаты.** Алгоритм машинного обучения для прогнозирования летального исхода у пациентов с COVID-19 в течение 72 часов госпитализации продемонстрировал высокую чувствительность (0,816) и специфичность (0,865). **Практическая значимость.** Созданный алгоритм может помочь улучшить оказание медицинской помощи пациентам, снизить смертность и свести к минимуму нагрузку на врачей во время пандемии COVID-19. При ограничении ресурсов в период пандемии, включая аппараты искусственной вентиляции легких, точное прогнозирование состояния пациентов с тяжелым течением и возможным летальным исходом может дать важные рекомендации для разделения групп пациентов и рационального распределения лечебных ресурсов.

Ключевые слова

машинное обучение, математическая модель, классификация, метрики модели, ROC-анализ, риск-фактор, COVID-19

Благодарности

Исследование выполнено при финансовой поддержке Министерства науки и высшего образования Российской Федерации (Соглашение № 075-15-2022-301 от 20.04.2022).

Ссылка для цитирования: Корсаков И.Н., Каронова Т.Л., Конради А.О., Рубин А.Д., Курапеев Д.И., Черникова А.Т., Михайлова А.А., Шляхто Е.В. Прогнозирование летального исхода у пациентов с установленным диагнозом COVID-19 // Научно-технический вестник информационных технологий, механики и оптики. 2022. Т. 22, № 5. С. 970–981. doi: 10.17586/2226-1494-2022-22-5-970-981

© Корсаков И.Н., Каронова Т.Л., Конради А.О., Рубин А.Д., Курапеев Д.И., Черникова А.Т., Михайлова А.А., Шляхто Е.В., 2022

Prediction of fatal outcome in patients with confirm COVID-19

Igor N. Korsakov¹, Tatiana L. Karonova²✉, Alexandra O. Konradi³, Arkadii D. Rubin⁴,
Dmitry I. Kurapeev⁵, Alena T. Chernikova⁶, Arina A. Mikhaylova⁷, Evgeny V. Shlyakhto⁸

^{1,2,3,4,5,6,7,8} Almazov National Medical Research Centre, Saint Petersburg, 197341, Russian Federation

¹ korsakov_in@almazovcentre.ru, <https://orcid.org/0000-0003-2343-9641>

² karonova@mail.ru✉, <https://orcid.org/0000-0002-1547-0123>

³ Konradi_ao@almazovcenter.ru, <https://orcid.org/0000-0001-8169-7812>

⁴ rubin_ad@almazovcentre.ru, <https://orcid.org/0000-0001-5469-5635>

⁵ dkurapeev@gmail.com, <https://orcid.org/0000-0002-2190-1495>

⁶ arabicaa@gmail.com, <https://orcid.org/0000-0002-4878-6909>

⁷ armikhaylova@yandex.ru, <https://orcid.org/0000-0001-6066-3525>

⁸ e.shlyakhto@almazovcentre.ru, <https://orcid.org/0000-0003-2929-0980>

Abstract

SARS-CoV-2, the new coronavirus underlying the development of the COVID-19 pandemic, has led to a sharp increase in the burden on healthcare systems, high mortality and significant difficulties in organizing medical care. The aim of the study was to conduct a systematic analysis of factors affecting the course of infectious disease in patients with diagnosed COVID-19 hospitalized. In order to predict the course of the disease and determine the indications for more aggressive treatment, many different clinical and biological markers have been proposed, however, clinical and laboratory assessment of the condition is not always simple and can clearly predict the development of a severe course. Technologies based on artificial intelligence (AI) have played a significant role in predicting the development of the disease. One of the main requirements during a pandemic is an accurate prediction of the required resources and likely outcomes. In the present study, a machine learning (ML) approach is proposed to predict the fatal outcome in patients with an established diagnosis of COVID-19 based on the patient's medical history and clinical, laboratory and instrumental data obtained in the first 72 hours of the patient's stay in the hospital. A machine learning algorithm for predicting the lethal outcome in patients with COVID-19 during 72 hours of hospitalization demonstrated high sensitivity (0.816) and specificity (0.865). Given the serious concerns about limited resources, including ventilators, during the COVID-19 pandemic, accurately predicting patients who are likely to require artificial ventilation can help provide important recommendations regarding patient triage and resource allocation among hospitalized patients. In addition, early detection of such persons may allow for routine ventilation procedures, reducing some of the known risks associated with emergency intubation. Thus, this algorithm can help improve patient care, reduce patient mortality and minimize the burden on doctors during the COVID-19 pandemic.

Keywords

machine learning, mathematical model, classification, model metrics, ROC analysis, risk factor, COVID-19

Acknowledgements

This work was financially supported by the Ministry of Science and Higher Education of the Russian Federation (Agreement No. 075-15-2022-301).

For citation: Korsakov I.N., Karonova T.L., Konradi A.O., Rubin A.D., Kurapeev D.I., Chernikova A.T., Mikhaylova A.A., Shlyakhto E.V. Prediction of fatal outcome in patients with confirm COVID-19. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 5, pp. 970–981 (in Russian). doi: 10.17586/2226-1494-2022-22-5-970-981

Введение

Одно из основных требований во время пандемии — точное прогнозирование необходимых ресурсов и вероятных результатов лечения. Алгоритмы машинного обучения используются давно, в частности в работе [1] проведены исследования с целью применения моделей прогнозирования для уточнения потребности в медицинском кислороде. Разработана модель прогнозирования, которая определяет пациента с COVID-19, нуждающегося в кислороде. Авторы использовали информацию об уровне С-реактивного белка, наличии артериальной гипертензии, возрасте и параметрах клинического анализа крови в виде количества нейтрофилов и лимфоцитов у 221 пациента.

Для прогноза необходимости искусственной вентиляции легких в работе [2] использована группа из 1980 пациентов с COVID-19. Анализируемые параметры включали демографические данные, анамнез пациента, показатели жизнедеятельности в отделении

неотложной помощи и лабораторные параметры. Их результаты показали, что возраст и высокая температура тела были связаны с повышенным риском перевода больного на искусственную вентиляцию легких (ИВЛ). Так, подход машинного обучения был использован для прогнозирования ИВЛ у 197 пациентов с COVID-19, где входные данные включали 12 клинических характеристик, собранных в больницах США [3]. Их модель предсказала потребность в ИВЛ, применяя показатели и другие переменные, такие как артериальное давление и частота сердечных сокращений (ЧСС). В работе [4] произведен систематический обзор роли использования машинных методов во время пандемии.

В настоящей работе предложен подход машинного обучения для прогнозирования летального исхода у пациентов с установленным диагнозом COVID-19 на основе анамнеза пациента и клинических, лабораторных и инструментальных данных, полученных в первые 72 часа нахождения пациента в стационаре.

Одно из основных преимуществ полученных результатов работы по сравнению с ранее опубликованными — прогнозирование результата (летальный исход) в начальный момент поступления пациента в стационар путем оценки имеющихся симптомов и анамнеза пациента, не требуя длительного наблюдения за пациентом и дополнительной информации. Это может ускорить процесс планирования ресурсов, особенно во время пика заболевания, и избежать возникновения дефицита. Другое преимущество, имеющее важное значение с медицинской точки зрения — установление факторов риска, влияющих на смертность от COVID-19, используя данные, собранные в Национальном медицинском исследовательском центре (НМИЦ им. В.А. Алмазова) за период первых трех волн пандемии.

Полученные результаты могут помочь стационарам в прогнозировании потребности учреждения в необходимых лекарствах, оборудовании и эффективном управлении данными результатами.

Описание предложенного метода

Регистрация пациентов. В работе исследованы данные пациентов, госпитализированных в НМИЦ им. В.А. Алмазова в период с 13 мая 2020 года по конец августа 2021 года. Зарегистрированные пациенты включены в клиническое исследование при условии, что набор их жизненно важных показателей и лабораторных измерений выполнен в течение первых 72 часов после поступления в стационар, и, если подтвержден положительный результат на COVID-19, установленный с помощью полимеразной цепной реакции. В общей сложности проведен анализ данных 4071 пациента, госпитализированных в течение периода исследования (рис. 1), что соизмеримо с исследованиями в работе [5].

При поступлении пациента в отделение неотложной помощи или стационар автоматически начинается сбор данных о доступных жизненно важных показателях и лабораторных измерениях. Данные, полученные в первые 72 часа госпитализации, используются для создания алгоритма машинного обучения и расчета риска летального исхода. Все данные автоматически извлечены из электронной медицинской карты, не требуя вмешательства врача.

Отметим, что исследование представляет минимальный риск для людей, поскольку сбор данных был пассивным и не представлял угрозы для участвующих субъектов. Все данные о пациентах хранились в соответствии с Федеральным законом «О персональных данных».

Обработка данных. Рассмотрим подробнее типы данных пациентов при поступлении в стационар.

Первый тип данных: информация о состоянии и хронических заболеваниях пациента, которые внесены в электронную историю болезни (табл. 1).

Второй тип данных: жизненно важные показатели, такие как диастолическое артериальное давление (ДАД, мм рт. ст.), систолическое артериальное давление (САД, мм рт. ст.), ЧСС (1/мин), температура (°C), частота дыхания (ЧД, 1/мин), насыщение кислородом (SpO_2 , %), индекс массы тела (ИМТ, кг/м²), частота дыхательных движений (ЧДД, 1/мин), С-реактивный белок (СРБ, мг/л) (табл. 2).

Третий тип данных — данные лабораторных анализов. Некоторые анализы выполнялись только у части пациентов с хроническими заболеваниями и не вошли в конечную таблицу (табл. 2).

Проведенный корреляционный анализ выявил зависимости между признаками и их влияние на поставленный диагноз.

Практически не один из признаков не имеет явной зависимости с целевой функцией, что предполагает выполнение дальнейшего анализа переменных. С полной таблицей признаков можно ознакомиться по ссылке¹. В табл. 3 приведены основные категориальные признаки исследуемых пациентов для создания алгоритма.

Диагностика сопутствующих заболеваний, имевшихся во время пребывания пациента в стационаре, включая хронические заболевания бронхолегочной системы, хронические заболевания эндокринной системы и онкологические заболевания, оценена по МКБ-10.

В качестве целевой переменной, которая использована при обучении модели, применено понятие «Исход заболевания».

Из всех имеющихся признаков для уменьшения ошибок при ручном вводе данных отобрано 11 наиболее значимых, которые вошли в разработанный алгоритм. В качестве методов отбора признаков использованы алгоритмы Boruta и Portable Instant Mining Platform (PIMP), финальные данные отображены в табл. 4. По данным: выписано — 2858 пациентов, из них летальный исход — 196. Из данных случайным образом отобрано 5 % пациентов — 153 (140 выписаны, 13 летальный исход), которые не были использованы ни в обучение, ни в тестировании модели.

¹ Таблица корреляции признаков [Электронный ресурс]. Режим доступа: https://github.com/ikorsa/prediction_lethal_COVID-19/blob/main/Correlation.png (дата обращения: 18.09.2022).

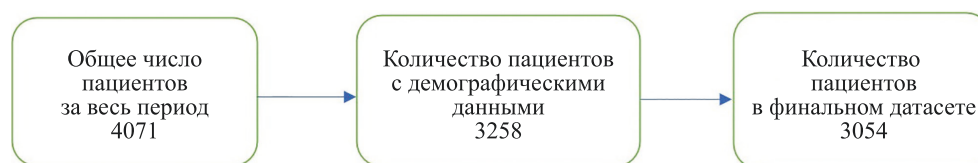


Рис. 1. Блок-схема включения пациентов

Fig. 1. Flowchart of patients in the study

Таблица 1. Пример данных пациента из истории болезни, которые были использованы при разработке алгоритма
 Table 1. An example of data from the patient's medical history used in the algorithm development

Поле	Заполненность, %
Регистрационный номер (ID пациента)	100
Возраст	100
Пол	100
Дата постановки диагноза	99,3
Дата госпитализации пациента	100
Время госпитализации пациента	99,9
Дата появления клинических симптомов	78,3
Диагноз основного заболевания по Международной классификации болезней (МКБ-10) при поступлении	99,3
Диагноз осложнения при поступлении	35,2
Подтвержден рентгенологически	100
Амбулаторное лечение	100
Вакцинация от гриппа	100
Вакцинация пневмококковой инфекции	100
Противовирусное лечение	100
Хронические заболевания бронхолегочной системы	82,4
Хронические заболевания сердечно-сосудистой системы	51,4
Хронические заболевания эндокринной системы	12,2
Онкологические заболевания	2,7
Болезнь, вызванная вирусом иммунодефицита человека	0,02
Туберкулез	0,07
Другие заболевания	22,7

Таблица 2. Значимые признаки, полученные в первые 72 часа госпитализации пациентов, для создания алгоритма машинного обучения

Table 2. Significant for the machine learning algorithm development features obtained in the first 72 hours of hospitalization

Признаки	Количество измерений	Среднее значение	Отклонение	Ошибка	Доверительный интервал (95 %)
Возраст, лет	3114	62	14	0,26	61,5
ИМТ, кг/м ²	3114	29,5	5,8	0,10	29,3
Ферритин, нг/л	3086	659	596	11	637,9
СРБ, мг/л	705	67,5	65,2	2,46	62,7
Протромбиновое время, с	2981	14	6,7	0,12	13,7
Гемоглобин, г/л	3113	133	18	0,32	133,1
Лейкоциты, 10 ⁹ 1/л	3113	7,8	8,9	0,16	7,5
Нейтрофилы, 10 ⁹ 1/л	2455	5,9	4,1	0,08	5,7
Лимфоциты, 10 ⁹ 1/л	3099	1,5	6,2	0,11	1,2
Соотношение нейтрофилы/лимфоциты	2455	6,4	7,7	0,16	6,1
Тромбоциты, 10 ⁹ 1/л	3113	221	99,9	1,79	217,6
Общий белок, г/л	1352	67,9	8,6	0,23	67,5
Креатинин, мкмоль/л	3110	89,1	52,6	0,94	87,3
Мочевина, ммоль/л	3078	6,4	4,2	0,08	6,3
ЧДД, 1/мин	3114	19	6	0,10	19,3
Пульсоксиметрия, %	3114	96	4	0,08	95,7
Процент поражения легочной ткани, %	3114	36	20	0,37	35,6
ЧСС, 1/мин	3114	78	15	0,27	77,2
САД, мм рт.ст.	3114	125	17	0,31	124,7
ДАД, мм рт.ст.	3114	76	11	0,20	75,9
Температура тела, °С	3114	37	0,7	0,01	36,9

Таблица 3. Значимые категориальные признаки, полученные в первые 72 часа госпитализации пациентов, для создания алгоритма машинного обучения

Table 3. Significant for the machine learning algorithm development categorical features obtained in the first 72 hours of hospitalization

Признаки	Значение	Описание	Количество	Доля, %
Пол	0	женский	1624	52,2
	1	мужской	1490	47,9
Исход заболевания	0	выписан	2915	93,6
	1	умер	199	6,4
Заболевания бронхолегочной системы	1	наличие	2591	83,2
	0	отсутствует	523	16,8
Сердечно-сосудистые заболевания	0	отсутствует	1570	50,4
	1	наличие	1544	49,6
Заболевания эндокринной системы	0	отсутствует	2736	87,9
	1	наличие	378	12,1
Онкологические заболевания	0	отсутствует	3054	98,1
	1	наличие	60	1,9
Сахарный диабет	0	отсутствует	2891	92,8
	1	наличие	223	7,2

Остальные данные разделены на два класса в соотношении 70/30 %: данные для обучения (2031 пациентов, из них с летальным исходом — 126) и данные для тестирования (870 пациентов, из них с летальным исходом — 57), которые вместе составили набор для финального обучения (табл. 4).

Модель машинного обучения. Модель создана с использованием библиотеки PyCaret, которая позволила проверить несколько алгоритмов за одно обращение [6]. Результаты обучения представлены в табл. 5.

В табл. 6 и 7 приведены значения метрик RF и CatBoost (модели с лучшей чувствительностью (Recall)) при K-Fold валидации.

В приведенных сравнительных результатах (табл. 6) видно, что у исследуемых моделей наблюдается небольшой разброс значений метрик Точность и Чувствительность для значений K-Fold валидации.

Полученные результаты свидетельствуют о том, что представленный датасет практически однороден и не имеет выбросов.

Для прогнозирования летального исхода одна из важных метрик — чувствительность метода, в связи с важностью не пропустить пациентов с большим риском летального исхода. Для каждого метода построены графики чувствительности (рис. 2).

Результаты

Известно, что идеальная модель обладает 100 % чувствительностью и специфичностью. Однако на практике добиться этого невозможно, более того, невозможно одновременно повысить и чувствительность, и специфичность модели. Компромисс можно найти с помощью порога отсечения, который нужен для при-

Таблица 4. Статистические данные

Table 4. Statistical knowledge

Признак	Количество	Среднее	Отклонение	Ошибка	Доверительный интервал (95 %)
Возраст, лет	3054	62	14	0,26	61,37
Процент поражения легочной ткани, %	3054	36	20	0,37	35,65
Тромбоциты, 10^9 /л	3053	222	100	1,82	217,99
СРБ, мг/л	682	67,5	63	2,41	62,74
Пульсоксиметрия, %	3054	96	4	0,08	95,66
ЧДД, 1/мин	3054	19	6	0,10	19,28
ЧСС, 1/мин	3054	78	15	0,28	77,21
Креатинин, мкмоль/л	3050	89	52,7	0,95	87,11
Мочевина, ммоль/л	3018	6,4	4,2	0,08	6,25
Общий белок, г/л	1314	68	8,7	0,24	67,50

Таблица 5. Результаты обучения

Table 5. Machine learning results

Модель	Точность	AUC	Чувствительность	Специфичность
Random Forest Classifier	0,862	0,908	0,816	0,865
SVM — Linear Kernel	0,691	0,000	0,761	0,688
K Neighbors Classifier	0,715	0,692	0,617	0,720
CatBoost Classifier	0,906	0,896	0,616	0,920
Extreme Gradient Boosting	0,927	0,904	0,504	0,948
Light Gradient Boosting Machine	0,929	0,905	0,449	0,953
Decision Tree Classifier	0,908	0,655	0,376	0,934

Примечание: AUC (Area Under Curve, ROC-площадь под кривой).

Таблица 6. Метрики моделей RF

Table 6. Metrics of RF Models

Выборка	Точность	AUC	Чувствительность	Специфичность
0	0,850	0,905	0,889	0,848
1	0,891	0,926	0,778	0,897
2	0,824	0,893	0,667	0,832
3	0,850	0,887	0,778	0,853
4	0,871	0,949	1,000	0,864
5	0,865	0,906	0,778	0,870
6	0,881	0,927	0,778	0,886
7	0,860	0,962	1,000	0,853
8	0,855	0,783	0,600	0,869
9	0,875	0,944	0,889	0,874
Среднее	0,862	0,908	0,816	0,865
Стандартное отклонение	0,018	0,048	0,124	0,018

Таблица 7. Метрики моделей CatBoost

Table 7. CatBoost Model Metrics

Выборка	Точность	AUC	Чувствительность	Специфичность
0	0,876	0,903	0,444	0,897
1	0,927	0,891	0,556	0,946
2	0,886	0,857	0,556	0,902
3	0,891	0,881	0,556	0,908
4	0,917	0,958	0,778	0,924
5	0,907	0,894	0,444	0,929
6	0,912	0,924	0,556	0,929
7	0,922	0,969	0,889	0,924
8	0,902	0,775	0,600	0,918
9	0,917	0,905	0,778	0,923
Среднее	0,906	0,896	0,616	0,920
Стандартное отклонение	0,016	0,051	0,142	0,014

Примечание: AUC (Area Under Curve, ROC-площадь под кривой).

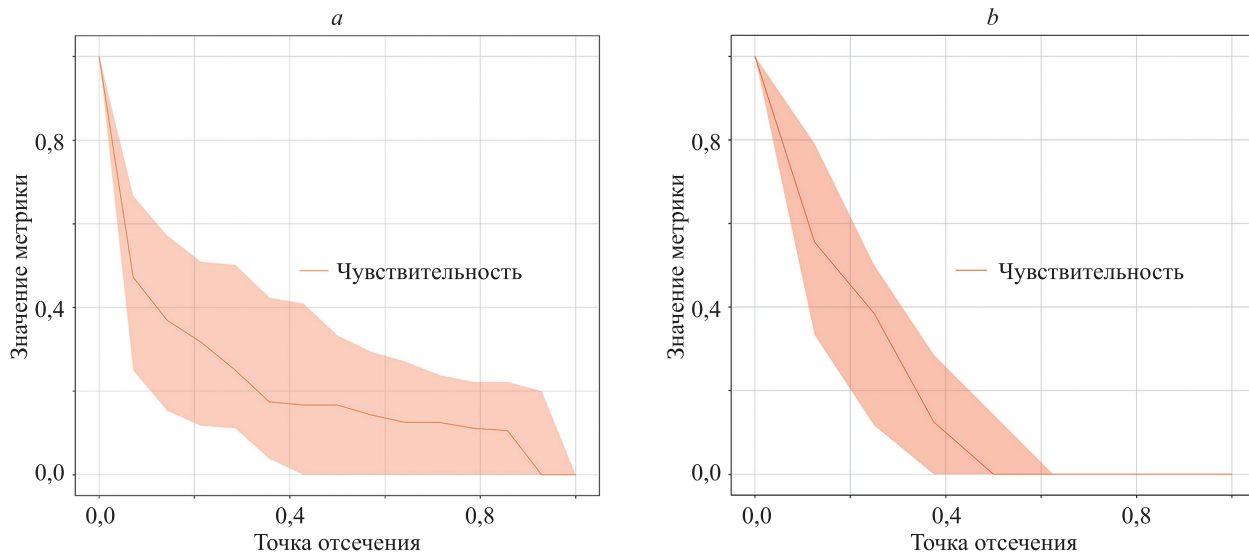


Рис. 2. Метрики методов для разных «точек отсечения» (threshold): Catboost (a) и Random Forest (b) Classifiers

Fig. 2. Catboost method metrics for different cut-off points (threshold) (a). Random Forest Classifier method metrics for different cut-off points (threshold) (b)

менения модели на практике: относить новые примеры к одному из двух классов. С большими допущениями можно считать, что чем больше показатель AUC, тем лучшей прогностической силой обладает модель. Отметим, что показатель AUC предназначен для сравнительного анализа нескольких моделей и не содержит никакой информации о чувствительности и специфичности модели.

Согласно экспертной шкале для значений AUC, по которой можно судить о качестве модели, значения: 0,6–0,7 свидетельствуют о среднем качестве модели, 0,7–0,8 — хорошем качестве, 0,8–0,9 — очень хорошем качестве и 0,9–1,0 — отличном качестве. Значения ниже 0,6 принято принимать за неудовлетворительное качество.

Выполним оценку модели прогнозирования летального исхода на двух наборах данных: тестовых — 870 пациентов и 5 % набора — 153 пациента, который не использовался при обучении модели. Дополнительно применим внешние данные из другого медицинского учреждения. Всего для обучения модели использован 2901 пациент, у которых был подтвержден диагноз COVID-19.

При использовании алгоритма Catboost Classifier установлено, что он обладает хорошей дискретизацией на всех типах данных (чувствительность до 99,2 %), кроме внешних, что свидетельствует о преимуществах его использования только на той базе, на которой он составлялся и его низкой прогностической способностью на внешних базах (чувствительность 37,5 %). Вместе с этим алгоритм Random Forest Classifier характеризуется хорошим балансом и подлежит дальнейшему исследованию в качестве алгоритма для модели калькулятора (для AUC 0,90 чувствительность 0,816 и специфичность 0,865) (рис. 3).

На практике значения Точности и Полноты гораздо более удобней рассчитывать с использованием матрицы неточностей рис. 4, которая часто используется

для описания эффективности модели классификации на наборе тестовых данных с известными истинными значениями.

В итоге по проведенным тестам наиболее точным алгоритмом для задачи прогнозирования летального исхода при COVID-19 выбран Random Forest Classifier.

Интерпретация результатов. В настоящее время у медицинских работников сохраняется недоверие к использованию прогностических моделей в медицине. В первую очередь, это связано с тем, что не всегда понятно, почему модель предложила то или другое решение. В общем считается, что математические модели, используемые в медицине, являются «черным ящиком». Вторая проблема — интерпретация самого результата прогноза. В частности, прогнозные модели выдают вероятность наступления того или иного события. Врачам

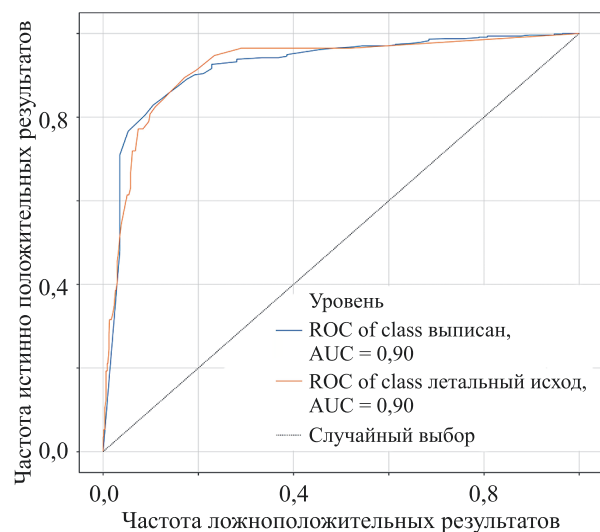


Рис. 3. График ROC-кривой

Fig. 3. ROC curve plot

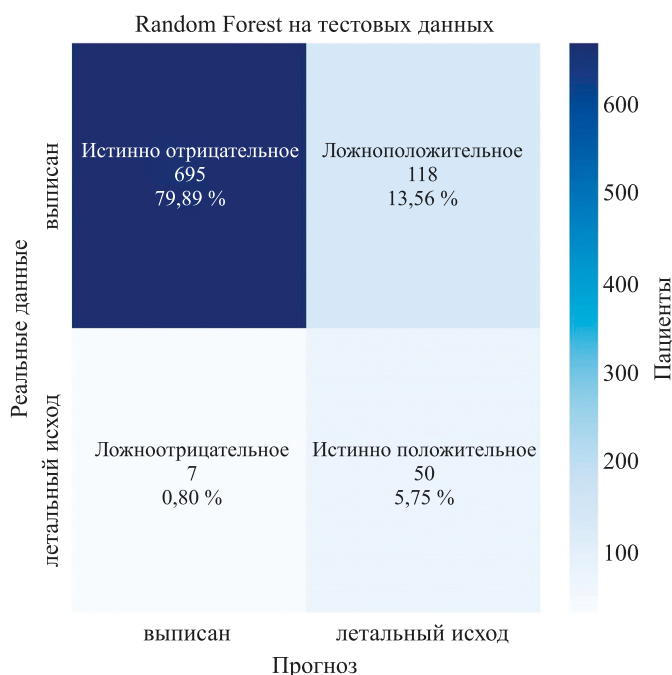


Рис. 4. Матрица неточностей

Fig. 4. Confusion matrix

же более привычно получать результат в виде категорий рисков: низкий, средний или высокий. Распределение результатов плотности вероятности модели приведено на рис. 5. На графике видно, что при вероятности от 0,2 до 0,5 результаты могут быть как с летальным исходом, так и не с летальным исходом, в то время как при вероятности от 0 до 0,2 можно говорить о низком риске, а при вероятности с 0,5 до 1,0 — высоком риске.

В последнее время стали появляться алгоритмы, которые позволяют интерпретировать вклад в конечный результат каждого признака. Такие алгоритмы интерпретируют не само предсказание, а одну из следующих величин:

- 1) разницу между двумя предсказаниями;
- 2) разницу между текущим и усредненным предсказаниями.

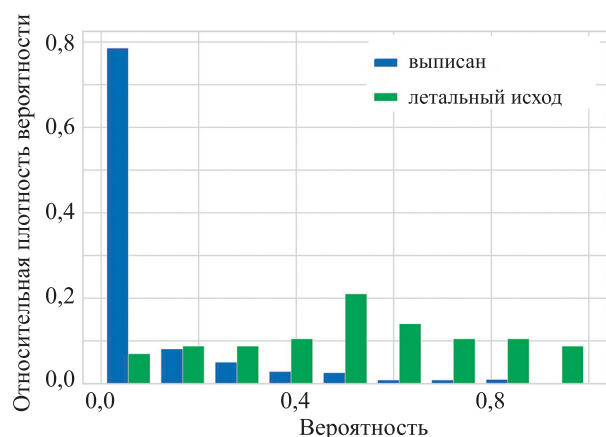


Рис. 5. Распределение вероятности летального исхода (нормированное)

Fig. 5. Probability distribution of lethal outcome (normalized)

Это означает, что алгоритм объясняет изменение в предсказании, вызванное изменением входных признаков (в первом случае) или появлением информации о входных признаках (во втором случае).

При этом возможно вычислить вклад каждого признака. Такой подход используется в методе SHAP (SHapley Additive explanation) [7] и [8].

Другой подход — LIME (Local Interpretable Model-Agnostic Explanations), суть которого в аппроксимации предсказания модели в окрестности тестового примера с более простой, легко интерпретируемой моделью, которая использует упрощенное представление. Например, если модель линейна, то каждому изменению сопоставляется некий вес. Если аппроксимация адекватна, то имеем наличие изменения, которое линейно влияет на предсказание модели. Отметим, что могут быть случаи, когда модель не сможет обучиться.

В данной работе для интерпретации результатов и оценки влияния признаков на результат модели использованы библиотеки SHAP и LIME.

Выбранный метод Random Forest Classifier показал следующий порядок значимости признаков: наиболее значимый — «Общий белок», наименее значимый — ЧСС (рис. 6).

График SHAP (рис. 6, б) позволяет оценить влияние признаков на прогноз модели, что является важной частью понимания врачом работы модели. Наиболее полно этот вопрос рассмотрен в работах [9, 10]. В первую очередь интерпретация модели помогает оценить ее качество. Если получить правила, которыми руководствуется модель при предсказании, то возможно оценить их правдоподобность.

Один из подходов для понимания работы модели — прогноз для конкретного пациента (рис. 7).

Как правило, при машинном обучении невозможно добиться полной интерпретируемости, но даже частичная интерпретация помогает оценить ее качество. Если мы выясним, как именно $f(x)$ преобразует входные признаки в выходной результат и какими правилами $E[f(x)]$ руководствуется при прогнозировании, то сможем оценить правдоподобность этих правил.

На рис. 7, а показан «Waterfall plot», объясняющий прогнозирование летального исхода у пациента №15 из представленного датасета. SHAP VALUES получены с помощью метода Tree SHAP. Схема читается снизу вверх, а признаки упорядочены по возрастанию их SHAP-значения (values). Например, SHAP VALUES +0,64 для признака «Общий белок» (имеющего значение 40,9) говорит о том, что это значение у данного пациента увеличивает величину предсказания модели, по сравнению с отсутствием признака «Общий белок» при произвольном наличии других признаков. Красный цвет свидетельствует об увеличении риска летального исхода при конкретных значениях показателей.

На рис. 7, б приведены данные пациента №12, также демонстрирующие вклад клинических и лабораторных признаков в риск развития летального исхода. Как видно, значения большинства признаков ассоциированы со снижением риска (обозначены синим цветом).

Библиотека PyCaret также позволяет построить график зависимости вклада признака в конечный ре-

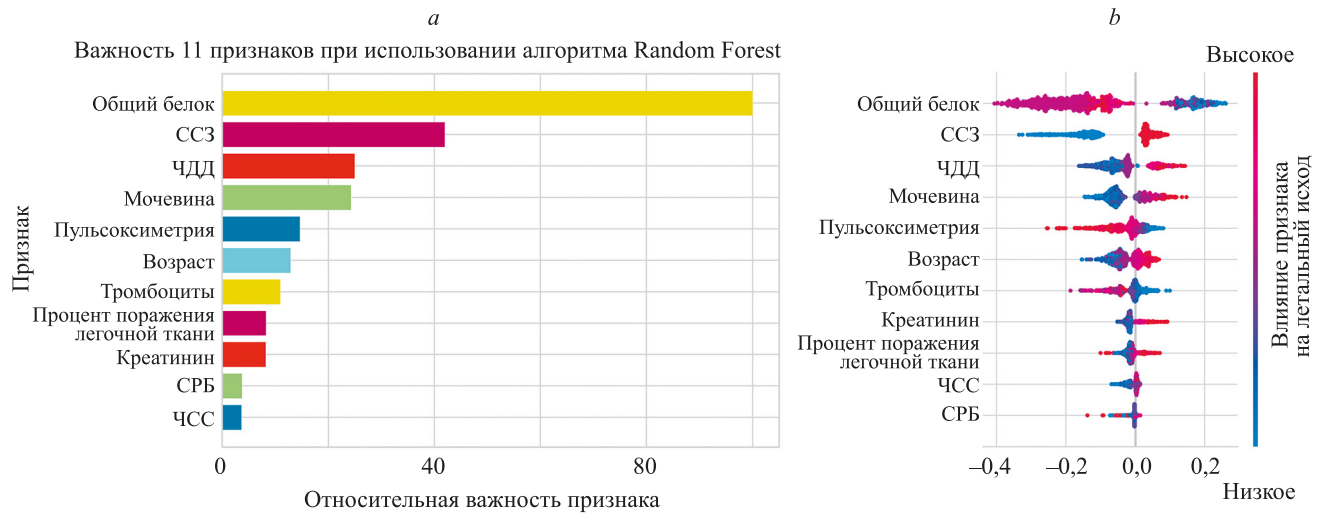


Рис. 6. Важность признаков (a), библиотека SHAP: влияние признаков на результат модели (b)

Fig. 6. Importance of features (a), SHAP Library: Influence of Features on Model Result (b)

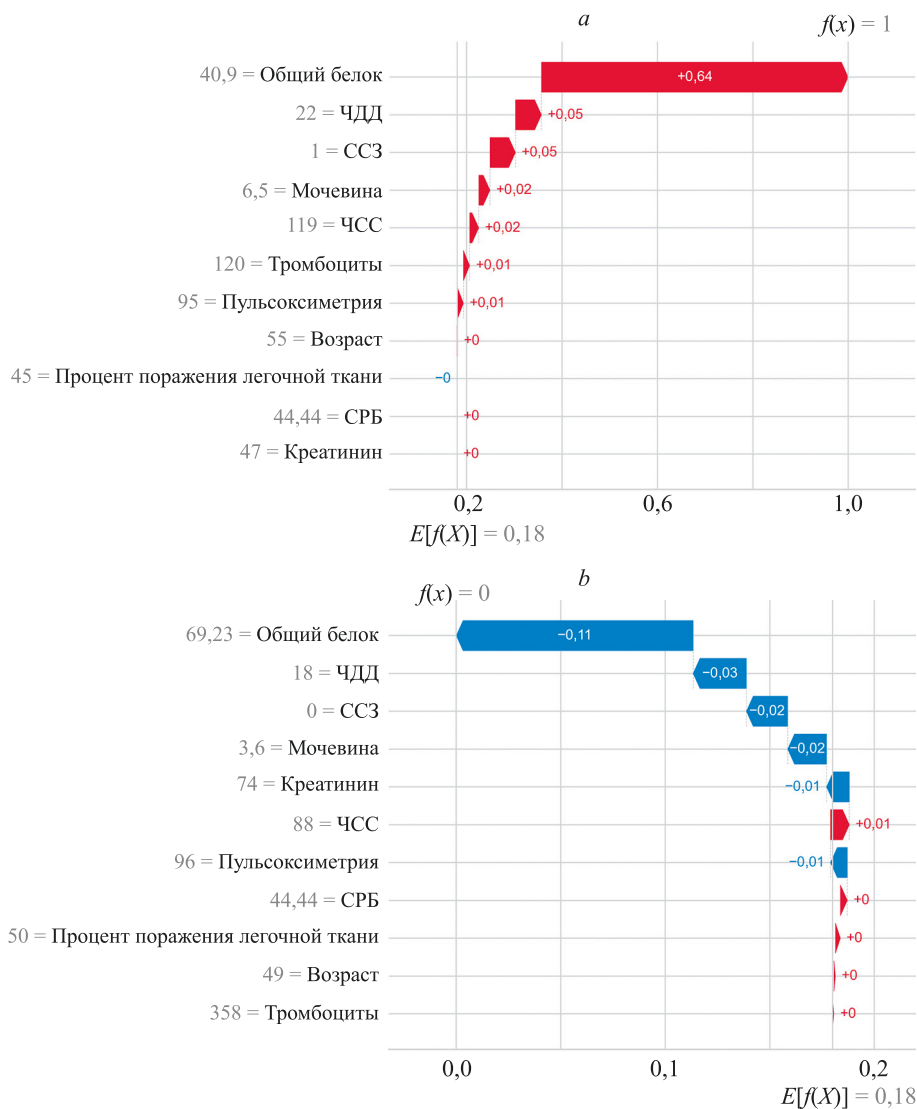


Рис. 7. Графики «Waterfall plot» (SHAP) для пациентов №15 (a) и №12 (b).

Размерности значений признаков представлены в табл. 2

Fig. 7. Waterfall plot (SHAP) charts for patients no.15 (a) and no.12 (b)

The dimensions of the feature values are presented in Table 2

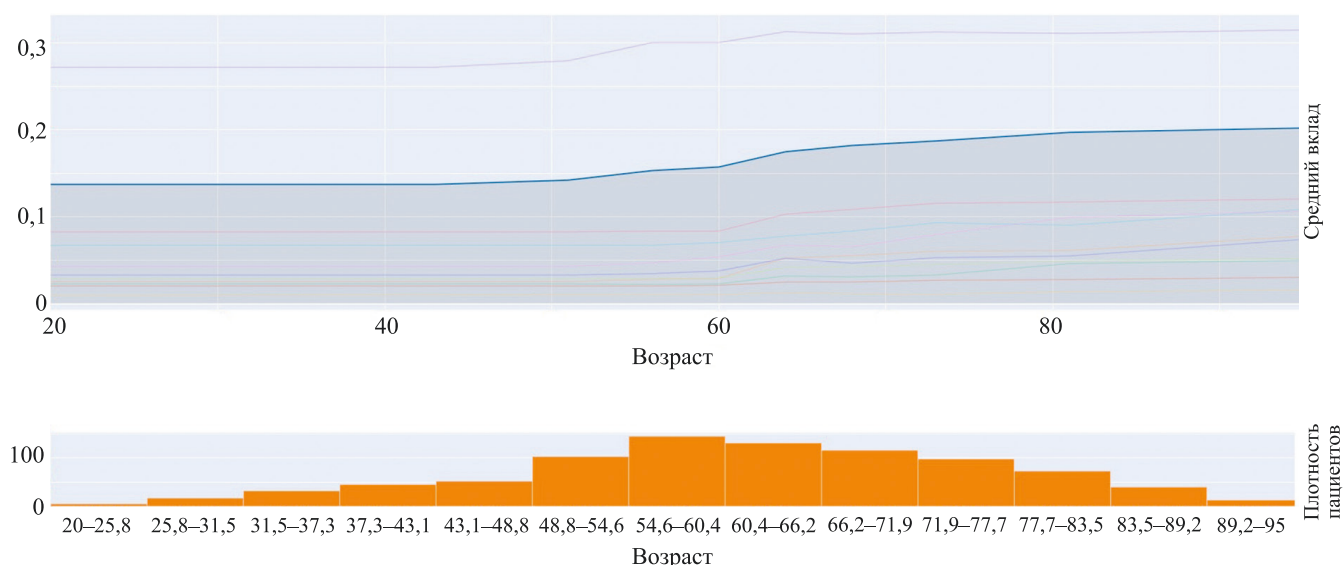


Рис. 8. Вклад признака «Возраст» в конечную модель и плотность распределения признака

Fig. 8. The contribution of the feature "Age" to the final model and the distribution density of the feature

зультат, на рис. 8 показана зависимость результата от признака «Возраст».

Обсуждение

Проведенное исследование построено на данных, полученных при работе инфекционного стационара НМИЦ им. В.А. Алмазова с пациентами с COVID-19, использованных для прогнозирования летального исхода. Разработанный алгоритм и калькулятор, построенный на его использовании, с высокой чувствительностью и специфичностью предсказывают вероятность летального исхода у пациентов с COVID-19, что отличает его от ранее предложенных алгоритмов, учитывающих ретроспективные данные [11,12].

Высокая чувствительность и специфичность, достигнутые алгоритмом, демонстрируют, что он способен точно различать пациентов с COVID-19 с высоким и низким рисками летального исхода уже в первые 72 часа госпитализации. Высокая чувствительность, в частности, предполагает, что алгоритм маловероятно даст ложноотрицательные данные, и что пациенты, нуждающиеся в повышенном внимании со стороны медперсонала, скорей всего не будут пропущены. Кроме того, повышение чувствительности алгоритма по сравнению с традиционной системой подсчета очков показало, что алгоритм способен обнаружить больше пациентов, которым требуется повышенное внимание; что должно способствовать улучшению маршрутизации пациентов и распределению ресурсов.

Отметим, что алгоритм также имеет большую чувствительность при более высокой специфичности по сравнению с NEWS2 [13]. Для базовой модели («NEWS2 + возраст») дискриминация была умеренной при внутренней валидации ($AUC = 0,764$; 95 % доверительный интервал: 0,737, 0,794).

Точное и ранее прогнозирование риска ухудшения состояния пациента может улучшить процедуры сорти-

ровки и маршрутизации пациентов и распределение ресурсов.

Заметим, что предложенная модель предсказала вероятность летального исхода, используя только доступные лабораторные исследования и данные жизненно важных признаков. Демографические данные для расчета риска не требовались, что отличает ее от уже имеющихся [14, 15]. Модель способна генерировать прогноз в отсутствие определенных входных данных. В реальной клинической практике не всегда всем больным требуется проведение полного спектра обследования, что при построении моделей нередко может влиять на результат. Пропущенные в базе данных значения не внесли существенного вклада и не искажали результаты при расчете риска. Таким образом, разработанная модель может быть использована для оценки прогноза даже в отсутствие некоторых признаков.

Отметим следующие выявленные ограничения представленной модели. Несмотря на то, что в выборку было включено достаточно большое количество пациентов, и использовалась внешняя валидация на больных из других медицинских центров, конечная общая выборка представлена 3054 больными, где летальный исход был представлен в 6,42 % случаев. Все это диктует необходимость проведения дальнейшей апробации разработанного калькулятора на внешних данных с включением показателей тяжелых больных. Также необходимо отметить, что на основании полученных результатов нельзя сделать выводы о производительности алгоритма в новых условиях при изменении штамма вируса и различий в течение заболевания. Дополнительно, в связи с принятым ограничением пациентов с подтвержденным диагнозом COVID-19, невозможно сделать вывод о производительности алгоритма для прогнозирования летального исхода в проспективных условиях у пациентов, не страдающих COVID-19, и у пациентов с подозрением на COVID-19, не подтвержденный в последствии.

Заключение

Алгоритм машинного обучения для прогнозирования летального исхода у пациентов с COVID-19 в течение 72 часов госпитализации продемонстрировал высокую чувствительность (0,816) и специфичность (0,865), и превзошел обычно используемую систему оценки раннего предупреждения заболевания. Алгоритм способен обнаруживать на 11 % больше пациентов, чем модифицированный показатель раннего предупреждения при одновременном снижении ложноположительных результатов. Учитывая серьезные опасения по поводу ограниченных ресурсов, включая аппараты искусствен-

ной вентиляции легких, во время пандемии COVID-19, точное прогнозирование пациентов, которым, вероятно, потребуется искусственная вентиляция легких, может помочь дать важные рекомендации в отношении сортировки пациентов и распределения ресурсов среди госпитализированных больных. Кроме того, раннее выявление таких лиц позволит проводить плановые процедуры вентиляции легких, снижая известные риски, связанные с экстренной интубацией. В результате реализации представленного алгоритма сможет помочь улучшить оказание медицинской помощи пациентам, снизить смертность больных и свести к минимуму нагрузку на врачей во время пандемии COVID-19.

Литература

1. Lee E.E., Hwang W., Song K.H., Jung J., Kang C.K., Kim J.H., Oh H.S., Kang Y.M., Lee E.B., Chin B.S., Song W., Kim N.J., Park J.K. Predication of oxygen requirement in COVID-19 patients using dynamic change of inflammatory markers: CRP, hypertension, age, neutrophil and lymphocyte (CHANeL) // *Scientific Reports*. 2021. V. 11. N 1. P. 13026. <https://doi.org/10.1038/s41598-021-92418-2>
2. Yu L., Halalau A., Dalal B., Abbas A.E., Ivascu F., Amin M., Nair G.B. Machine learning methods to predict mechanical ventilation and mortality in patients with COVID-19 // *PLoS ONE*. 2021. V. 16. N 4. P. e0249285. <https://doi.org/10.1371/journal.pone.0249285>
3. Burdick H., Lam C., Mataraso S., Siefkas A., Braden G., Dellinger R.P., McCoy A., Vincent J.L., Green-Saxena A., Barnes G., Hoffman J., Calvert J., Pellegrini E., Das R. Prediction of respiratory decompensation in Covid-19 patients using machine learning: The READY trial // *Computers in Biology and Medicine*. 2020. V. 124. P. 103949. <https://doi.org/10.1016/j.combiomed.2020.103949>
4. Syeda H.B., Syed M., Sexton K.W., Syed S., Begum S., Syed F., Prior F., Yu F. Role of machine learning techniques to tackle the COVID-19 crisis: systematic review // *JMIR Medical Informatics*. 2021. V. 9. N 1. P. e23811. <https://doi.org/10.2196/23811>
5. Jehi L., Ji X., Milinovich A., Erzurum S., Merlino A., Gordon S., Young J.B., Kattan M.W. Development and validation of a model for individualized prediction of hospitalization risk in 4,536 patients with COVID-19 // *PLoS ONE*. 2020. V. 15. N 8. P. e0237419. <https://doi.org/10.1371/journal.pone.0237419>
6. Moez Ali. PyCaret: An open source, low-code machine learning library in Python // *PyCaret version 2.1*. 2020 [Электронный ресурс]. URL: <https://www.pycaret.org> (дата обращения: 21.07.2022).
7. Lundberg S.M., Erion G., Chen H., DeGrave A., Prutkin J.M., Nair B., Katz R., Himmelfarb J., Bansal N., Lee S.-I. From local explanations to global understanding with explainable AI for trees // *Nature Machine Intelligence*. 2020. V. 2. N 1. P. 56–67. <https://doi.org/10.1038/s42256-019-0138-9>
8. Chen H., Lundberg S.M., Lee S.I. Explaining models by propagating shapley values of local components // *Studies in Computational Intelligence*. 2021. V. 914. P. 261–270. https://doi.org/10.1007/978-3-030-53352-6_24
9. Linardatos P., Papastefanopoulos V., Kotsiantis S. Explainable ai: A review of machine learning interpretability methods // *Entropy*. 2021. V. 23. N 1. P. 18. <https://doi.org/10.3390/e23010018>
10. Li X., Xiong H., Li X., Wu X., Zhang X., Liu J., Bian J., Dou D. Interpretable deep learning: interpretation, interpretability, trustworthiness, and beyond // *Knowledge and Information Systems*. 2022. in press. <https://doi.org/10.1007/s10115-022-01756-8>
11. Yang Y., Zhu X.F., Huang J., Chen C., Zheng Y., He W., Zhao L.H., Gao Q., Huang X.X., Fu L.J., Zhang Y., Chang Y.Q., Zhang H.J., Lu Z.J. Nomogram for prediction of fatal outcome in patients with severe COVID-19: a multicenter study // *Military Medical Research*. 2021. V. 8. N 1. P. 21. <https://doi.org/10.1186/s40779-021-00315-6>
12. Zarei J., Jamshidnezhad A., Haddadzadeh Shoushtari M., Mohammad Hadianfard A., Cheraghi M., Sheikhtaheri A. Machine learning models to predict in-hospital mortality among inpatients with COVID-19: Underestimation and overestimation bias analysis in

References

1. Lee E.E., Hwang W., Song K.H., Jung J., Kang C.K., Kim J.H., Oh H.S., Kang Y.M., Lee E.B., Chin B.S., Song W., Kim N.J., Park J.K. Predication of oxygen requirement in COVID-19 patients using dynamic change of inflammatory markers: CRP, hypertension, age, neutrophil and lymphocyte (CHANeL). *Scientific Reports*, 2021, vol. 11, no. 1, pp. 13026. <https://doi.org/10.1038/s41598-021-92418-2>
2. Yu L., Halalau A., Dalal B., Abbas A.E., Ivascu F., Amin M., Nair G.B. Machine learning methods to predict mechanical ventilation and mortality in patients with COVID-19. *PLoS One*, 2021, vol. 16, no. 4, pp. e0249285. <https://doi.org/10.1371/journal.pone.0249285>
3. Burdick H., Lam C., Mataraso S., Siefkas A., Braden G., Dellinger R.P., McCoy A., Vincent J.L., Green-Saxena A., Barnes G., Hoffman J., Calvert J., Pellegrini E., Das R. Prediction of respiratory decompensation in Covid-19 patients using machine learning: The READY trial. *Computers in Biology and Medicine*, 2020, vol. 124, pp. 103949. <https://doi.org/10.1016/j.combiomed.2020.103949>
4. Syeda H.B., Syed M., Sexton K.W., Syed S., Begum S., Syed F., Prior F., Yu F. Role of machine learning techniques to tackle the COVID-19 crisis: systematic review. *JMIR Medical Informatics*, 2021, vol. 9, no. 1, pp. e23811. <https://doi.org/10.2196/23811>
5. Jehi L., Ji X., Milinovich A., Erzurum S., Merlino A., Gordon S., Young J.B., Kattan M.W. Development and validation of a model for individualized prediction of hospitalization risk in 4,536 patients with COVID-19. *PLoS ONE*, 2020, vol. 15, no. 8, pp. e0237419. <https://doi.org/10.1371/journal.pone.0237419>
6. Moez Ali. PyCaret: An open source, low-code machine learning library in Python. *PyCaret version 2.1*. 2020. Available at: <https://www.pycaret.org> (accessed: 21.07.2022).
7. Lundberg S.M., Erion G., Chen H., DeGrave A., Prutkin J.M., Nair B., Katz R., Himmelfarb J., Bansal N., Lee S.-I. From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence*, 2020, vol. 2, no. 1, pp. 56–67. <https://doi.org/10.1038/s42256-019-0138-9>
8. Chen H., Lundberg S.M., Lee S.I. Explaining models by propagating shapley values of local components. *Studies in Computational Intelligence*, 2021, vol. 914, pp. 261–270. https://doi.org/10.1007/978-3-030-53352-6_24
9. Linardatos P., Papastefanopoulos V., Kotsiantis S. Explainable ai: A review of machine learning interpretability methods. *Entropy*, 2021, vol. 23, no. 1, pp. 18. <https://doi.org/10.3390/e23010018>
10. Li X., Xiong H., Li X., Wu X., Zhang X., Liu J., Bian J., Dou D. Interpretable deep learning: interpretation, interpretability, trustworthiness, and beyond. *Knowledge and Information Systems*, 2022, in press. <https://doi.org/10.1007/s10115-022-01756-8>
11. Yang Y., Zhu X.F., Huang J., Chen C., Zheng Y., He W., Zhao L.H., Gao Q., Huang X.X., Fu L.J., Zhang Y., Chang Y.Q., Zhang H.J., Lu Z.J. Nomogram for prediction of fatal outcome in patients with severe COVID-19: a multicenter study. *Military Medical Research*, 2021, vol. 8, no. 1, pp. 21. <https://doi.org/10.1186/s40779-021-00315-6>
12. Zarei J., Jamshidnezhad A., Haddadzadeh Shoushtari M., Mohammad Hadianfard A., Cheraghi M., Sheikhtaheri A. Machine learning models to predict in-hospital mortality among inpatients with COVID-19: Underestimation and overestimation bias analysis in subgroup populations. *Journal of Healthcare Engineering*, 2022, pp. 1644910. <https://doi.org/10.1155/2022/1644910>

- subgroup populations // Journal of Healthcare Engineering. 2022. P. 1644910. <https://doi.org/10.1155/2022/1644910>
13. Baker K.F., Hanrath A.T., van der Loeff I.S., Kay L.J., Back J., Duncan C.J. National early warning score 2 (NEWS2) to identify inpatient COVID-19 deterioration: a retrospective analysis // *Clinical Medicine, Journal of the Royal College of Physicians of London*. 2021. V. 21. N 2. P. 84–89. <https://doi.org/10.7861/clinmed.2020-0688>
 14. Kondapalli A.R., Koganti H., Challagundla S.K., Guntaka C., Biswas S. Machine learning predictions of COVID-19 second wave end-times in Indian states // *Indian Journal of Physics*. 2022. V. 96. N 8. P. 2547–2555. <https://doi.org/10.1007/s12648-021-02195-x>
 15. Mahdavi M., Choubdar H., Zabe E., Rieder M., Safavi-Naeini S., Jobbagy Z., Ghorbani A., Abedini A., Kiani A., Khanlarzadeh V., Lashgari R., Kamrani E. A machine learning based exploration of COVID-19 mortality risk // *Plos ONE*. 2021. V. 16. N 7. P. e0252384. <https://doi.org/10.1371/journal.pone.0252384>

Авторы

Корсаков Игорь Николаевич — кандидат физико-математических наук, специалист по информационному обеспечению, Национальный медицинский исследовательский центр им. В.А. Алмазова, Санкт-Петербург, 197341, Российская Федерация, [sc 57189603967](https://orcid.org/0000-0003-2343-9641), <https://orcid.org/0000-0003-2343-9641>, korsakov_in@almazovcentre.ru

Каронова Татьяна Леонидовна — доктор медицинских наук, доцент, главный научный сотрудник, заведующий научно-исследовательской лабораторией, профессор кафедры, Национальный медицинский исследовательский центр им. В.А. Алмазова, Санкт-Петербург, 197341, Российская Федерация, [sc 55812730000](https://orcid.org/0000-0002-1547-0123), <https://orcid.org/0000-0002-1547-0123>, karonova@mail.ru

Конради Александра Олеговна — доктор медицинских наук, академик РАН, заместитель генерального директора по научной работе, Национальный медицинский исследовательский центр им. В.А. Алмазова, Санкт-Петербург, 197341, Российская Федерация, [sc 7004144504](https://orcid.org/0000-0001-8169-7812), <https://orcid.org/0000-0001-8169-7812>, Konradi_ao@almazovcenter.ru

Рубин Аркадий Дмитриевич — доктор медицинских наук, доцент, директор Лечебно-реабилитационного комплекса, Национальный медицинский исследовательский центр им. В.А. Алмазова, Санкт-Петербург, 197341, Российская Федерация, <https://orcid.org/0000-0001-5469-5635>, rubin_ad@almazovcentre.ru

Курапеев Дмитрий Ильич — кандидат медицинских наук, заместитель генерального директора по информационным технологиям, Национальный медицинский исследовательский центр им. В.А. Алмазова, Санкт-Петербург, 197341, Российская Федерация, [sc 57225231263](https://orcid.org/0000-0002-2190-1495), <https://orcid.org/0000-0002-2190-1495>, dkurapeev@gmail.com

Черникова Алёна Тимуровна — младший научный сотрудник, Национальный медицинский исследовательский центр им. В.А. Алмазова, Санкт-Петербург, 197341, Российская Федерация, <https://orcid.org/0000-0002-4878-6909>, arabicaa@gmail.com

Михайлова Арина Алексеевна — ординатор кафедры, Национальный медицинский исследовательский центр им. В.А. Алмазова, Санкт-Петербург, 197341, Российская Федерация, <https://orcid.org/0000-0001-6066-3525>, armikhaylova@yandex.ru

Шляхто Евгений Владимирович — доктор медицинских наук, академик РАН, генеральный директор, Национальный медицинский исследовательский центр им. В.А. Алмазова, Санкт-Петербург, 197341, Российская Федерация, [sc 16317213100](https://orcid.org/0000-0003-2929-0980), <https://orcid.org/0000-0003-2929-0980>, e.shlyakhto@almazovcentre.ru

Authors

Igor N. Korsakov — PhD (Physics & Mathematics), IT Specialist, Almazov National Medical Research Centre, Saint Petersburg, 197341, Russian Federation, [sc 57189603967](https://orcid.org/0000-0003-2343-9641), <https://orcid.org/0000-0003-2343-9641>, korsakov_in@almazovcentre.ru

Tatiana L. Karonova — D. Sc. (Medicine), Associate Professor, Chief Researcher, Head of Laboratory, Chair Professor, Almazov National Medical Research Centre, Saint Petersburg, 197341, Russian Federation, [sc 55812730000](https://orcid.org/0000-0002-1547-0123), <https://orcid.org/0000-0002-1547-0123>, karonova@mail.ru

Alexandra O. Konradi — D. Sc. (Medicine), Academician of the RAS, Deputy Director General for Research, Almazov National Medical Research Centre, Saint Petersburg, 197341, Russian Federation, [sc 7004144504](https://orcid.org/0000-0001-8169-7812), <https://orcid.org/0000-0001-8169-7812>, Konradi_ao@almazovcenter.ru

Arkadii D. Rubin — D. Sc. (Medicine), Associate Professor, Medical Director of the Treatment and Rehabilitation Facility, Almazov National Medical Research Centre, Saint Petersburg, 197341, Russian Federation, <https://orcid.org/0000-0001-5469-5635>, rubin_ad@almazovcentre.ru

Dmitry I. Kurapeev — PhD (Medicine), Deputy CEO for Information Technology, Almazov National Medical Research Centre, Saint Petersburg, 197341, Russian Federation, [sc 57225231263](https://orcid.org/0000-0002-2190-1495), <https://orcid.org/0000-0002-2190-1495>, dkurapeev@gmail.com

Alena T. Chernikova — Junior Researcher, Almazov National Medical Research Centre, Saint Petersburg, 197341, Russian Federation, <https://orcid.org/0000-0002-4878-6909>, arabicaa@gmail.com

Arina A. Mikhaylova — Department Resident, Almazov National Medical Research Centre, Saint Petersburg, 197341, Russian Federation, <https://orcid.org/0000-0001-6066-3525>, armikhaylova@yandex.ru

Evgeny V. Shlyakhto — D. Sc. (Medicine), Academician of the RAS, Director General, Almazov National Medical Research Centre, Saint Petersburg, 197341, Russian Federation, [sc 16317213100](https://orcid.org/0000-0003-2929-0980), <https://orcid.org/0000-0003-2929-0980>, e.shlyakhto@almazovcentre.ru

Статья поступила в редакцию 18.07.2022
Одобрена после рецензирования 07.09.2022
Принята к печати 30.09.2022

Received 18.07.2022
Approved after reviewing 07.09.2022
Accepted 30.09.2022



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»