

DOI: <https://doi.org/10.36425/rehab678920>

EDN: JEMEV0

Возможности машинного обучения в прогнозировании ранних функциональных исходов у пациентов с инсультом

Е.М. Черных, Н.М. Хасанова, А.А. Карякин, З.Э. Джафарова, А.А. Ключас

Северный государственный медицинский университет, Архангельск, Россия

АННОТАЦИЯ

Обоснование. Задача создания универсального прогностического инструмента для пациентов с инсультом до конца не решена.

Цель исследования — на основе алгоритмов машинного обучения разработать модели, способные с первых суток от дебюта инсульта прогнозировать функциональный исход с оценкой по шкале инсульта Национального института здоровья США (NIHSS) и модифицированной шкале Рэнкина (mRs) при выписке из стационара на основании клинико-анамнестических предикторов.

Методы. На основе базы данных, созданной из 5225 записей о пациентах с инсультом, выписанных из неврологических отделений, разработаны модели на основе алгоритмов Artificial Neural Network (ANN) и Random Forest (RF). Использовали 21 переменную, включая демографические данные пациентов, исходные баллы по NIHSS и mRS, тип инсульта, время от начала заболевания до госпитализации, сведения о сопутствующих заболеваниях и выполненной экстренной реваскуляризации. Прогнозировали исходы с оценкой в баллах по шкалам NIHSS и mRs. Алгоритмы решали задачу классификации по множественному набору значений исхода: 26 классов для модели I (0–25 баллов по NIHSS) и 6 классов в для модели II (0–5 баллов по mRs). Качество моделей оценивали с помощью метода площади под кривой операционной характеристики приёмника (ROC-AUC). Вклад каждого предиктора в результат прогнозирования оценивали с помощью метода SHapley Additive exPlanations (SHAP).

Результаты. Определены результаты прогностической эффективности ANN по AUC-ROC: 0,771 для модели I и 0,844 для модели II. RF AUC-ROC составила 0,778 для модели I и 0,845 для модели II. Для дальнейшей работы выбран алгоритм RF ввиду его лучшей интерпретируемости. Наиболее значимыми признаками, влияющими на прогнозируемые исходы, стали исходные значения NIHSS и mRs, возраст пациента, время доезда до стационара от начала инсульта, тип инсульта. При проверке эффективности алгоритма RF на внешней валидационной выборке из 783 записей показатели ROC AUC составили 0,786 для модели I и 0,774 для модели II. Для практического применения создан калькулятор.

Заключение. Разработанные модели на основе алгоритма RF способны в течение первых суток от начала инсульта достоверно прогнозировать ранний функциональный исход с оценкой по NIHSS и mRs на основании клинико-анамнестических предикторов. Полученный инструмент полезен для формирования персонализированных терапевтических и реабилитационных стратегий в острый период инсульта. Универсальность использования моделей позволяет их применение в удалённых сельских медицинских учреждениях с дефицитом или отсутствием профильных специалистов и диагностического оборудования.

Ключевые слова: инсульт; прогнозирование; ранний функциональный исход; универсальная прогностическая модель; алгоритм Random Forest.

Как цитировать:

Черных Е.М., Хасанова Н.М., Карякин А.А., Джафарова З.Э., Ключас А.А. Возможности машинного обучения в прогнозировании ранних функциональных исходов у пациентов с инсультом // Физическая и реабилитационная медицина, медицинская реабилитация. 2025. Т. 7, № 2. С. 82–94. DOI: 10.36425/rehab678920 EDN: JEMEV0

DOI: <https://doi.org/10.36425/rehab678920>

EDN: JEMEVO

Machine Learning for Predicting Early Functional Outcomes in Patients with Stroke

Ekaterina M. Chernykh, Nina M. Khasanova, Aleksey A. Karyakin, Zohra E. Dzhafarova, Alexey A. Kliukas

Northern State Medical University, Arkhangelsk, Russia

ABSTRACT

BACKGROUND: Development of a universal predictive tool for patients with stroke remains a challenge.

AIM: The study aimed to develop machine learning–based models that could predict functional outcomes from the first day after stroke. The models were trained using clinical and anamnestic predictors, and functional outcomes were evaluated using the National Institutes of Health Stroke Scale (NIHSS) and the modified Rankin Scale (mRS) at hospital discharge.

METHODS: Models based on artificial neural network (ANN) and random forest (RF) algorithms were developed using a database created from 5,225 records of patients with stroke discharged from neurological departments. Twenty-one parameters were used, including patient demographics; baseline National Institutes of NIHSS and mRS scores; stroke type; time from stroke onset to hospitalization; comorbidities; and emergency revascularization. Outcomes were predicted using NIHSS and mRS scores. The algorithms solved the classification problem for multiple sets of outcome values. Model I had 26 classes (NIHSS score of 0–25), while model II had 6 classes (mRS score of 0–5). The quality of the models was evaluated using the area under the receiver operating characteristic curve (ROC–AUC). The contribution of each predictor was evaluated using the SHapley Additive exPlanations (SHAP).

RESULTS: The predictive value of the ANN was determined based on the AUC–ROC: 0.771 for model I and 0.844 for model II. The RF AUC–ROC was 0.778 for model I and 0.845 for model II. The RF algorithm was chosen for further work due to its better interpretability. The most significant features that influenced the predicted outcomes were baseline NIHSS and mRS scores, patient age, time from stroke onset to admission, and stroke type. When RF performance was tested on an external validation set of 783 records, ROC–AUC values were 0.786 for model I and 0.774 for model II. A calculator was developed for practical use.

CONCLUSION: The proposed RF-based models can reliably predict an early functional outcome within the first day of stroke onset, using NIHSS and mRS scores and clinical and anamnestic predictors. This tool can be used to develop personalized therapeutic and rehabilitation strategies in the acute phase of a stroke. These models are versatile to be used in rural and remote healthcare organizations that lack specialized staff and diagnostic equipment.

Keywords: stroke; prognosis; early functional outcome; universal prognostic model; Random Forest algorithm.

To cite this article:

Chernykh EM, Khasanova NM, Karyakin AA, Dzhafarova ZE, Kliukas AA. Machine Learning for Predicting Early Functional Outcomes in Patients with Stroke. *Physical and rehabilitation medicine, medical rehabilitation*. 2025;7(2):82–94. DOI: 10.36425/rehab678920 EDN: JEMEVO

Submitted: 26.04.2025

Accepted: 01.05.2025

Published online: 19.06.2025

Список сокращений

ANN (Artificial Neural Network) — вычислительная модель, вдохновлённая структурой и функциями биологических нейронных сетей

DT (Decision Trees) — деревья решений

mRS (Modified Rankin Scale) — модифицированная шкала Рэнкина для оценки степени инвалидности и зависимости в повседневной жизни лица, перенёвшего инсульт или другие неврологические нарушения

NIHSS (National Institutes of Health Stroke Scale) — шкала инсульта Национального института здоровья США для определения степени его тяжести

RBF (Radial Basis Function kernel) — ядро радиальной базисной функции

RF (Random Forest) — метод случайного леса; алгоритм машинного обучения, который состоит из множества отдельных независимых «решающих деревьев»

ROC-AUC (Receiver Operating Characteristic — Area Under Curve) — метрика оценки качества модели машинного обучения

SHAP (SHapley Additive exPlanations) — метод интерпретации машинного обучения, объясняющий вклад каждого признака в предсказание конкретного наблюдения

SVC (Support Vector Classifier) — метод опорных векторов

ОБОСНОВАНИЕ

Инструменты на основе искусственного интеллекта и машинного обучения востребованы на современном этапе исследовательской и клинической медицинской деятельности. Они активно используются в системах принятия врачебных решений, диагностике и прогнозировании заболеваний, в том числе в неврологии, что ранее казалось невозможным [1, 2].

Обращает на себя внимание максимальное количество разработок и моделей с использованием машинного обучения для социально значимых неврологических заболеваний — эпилепсии, болезни Альцгеймера и инсульта [2].

Последнее десятилетие в неврологии инсульта стало революционным. Появилась возможность использования принципиально новых методов для диагностики и лечения, в том числе с применением функционала искусственного интеллекта. За этот период методы машинного обучения использовали у пациентов с инсультом для диагностики, прогнозирования начальных признаков инсульта, оценки тяжести состояния, определения состава тромба при ишемическом инсульте, анализа отёка головного мозга, прогнозирования увеличения гематомы при геморрагическом инсульте и прогнозирования исхода заболевания [1, 3].

Внедрение алгоритмов машинного обучения в клиническую практику предполагает четыре этапа рабочего процесса: мероприятия на этапе первичной профилактики; лечение в остром периоде; прогнозирование исходов после постановки диагноза и вторичную профилактику. Первичная профилактика включает в себя персонализированное или стратифицированное прогнозирование рисков для пациентов и выявление проблем в оказании медицинской помощи, а интеграция искусственного интеллекта в лечение в острый период направлена на помощь врачам в постановке диагноза и принятии тактических решений [4].

Алгоритмы машинного обучения для прогнозирования при установленном диагнозе острого цереброваскулярного заболевания и алгоритмы вторичной профилактики демонстрируют прогнозируемые результаты, выявляющие пациентов, которые будут по-разному восстанавливаться в ходе лечения или, возможно, потребуют более тщательного наблюдения либо дополнительных лечебных мероприятий из-за повышенного риска рецидива. В совокупности алгоритмы машинного обучения должны помогать в принятии клинических решений на каждом этапе, предоставляя рекомендации и указывая на возможные упущения, проигнорированные при рутинном обследовании [5].

Прогнозирование клинических и функциональных исходов у пациентов с острым инсультом обычно основывается на многофакторной информации, такой как демографические, клинические, лабораторные и радиологические данные [6].

Модели прогнозирования на основе искусственного интеллекта могут давать объективные результаты после изучения входных данных и выполнения многочисленных расчётов. Прогностические модели машинного обучения, основанные на распознавании и сегментации признаков изображений, значительно упростили быструю диагностику инсульта, но прогноз исхода инсульта зависит также от множества других клинических и анамнестических факторов [7]. Стратегии прогнозирования исхода инсульта с использованием искусственного интеллекта подразумевают поиск оптимального набора входных предикторов для выполнения наиболее точного и достоверного прогноза, а также выбор алгоритма с наиболее сбалансированным сочетанием точности, интерпретируемости и простоты использования модели. По данным разных авторов, ключевыми предикторами для прогнозирования исходов инсульта являются возраст пациента, исходная тяжесть

неврологического дефицита, индивидуальный преморбидный фон, биохимические показатели, а также данные нейровизуализационных исследований [8–10], причём включение нескольких типов данных в одну модель искусственного интеллекта повышает производительность модели и улучшает точность прогнозирования [11].

Среди всего разнообразия алгоритмов наиболее эффективными в прогнозировании функциональных исходов инсульта зарекомендовали себя деревья решений (Decision Trees, DT) и метод случайного леса (Random Forest, RF), градиентный бустинг (Extreme Gradient Boosting, XGB), метод опорных векторов (Support Vector Machine, SVM) и искусственные нейронные сети (Artificial Neural Network, ANN) с максимальным достижением точности 95,5% в случае AdaBoost [7, 12, 13]. Предпочтение применения того или иного алгоритма зависит от специфики исходных данных и цели исследования, при этом ни один алгоритм не является абсолютно точным, а результаты варьируют в зависимости от используемых данных и методологии. И хотя разработки последних лет демонстрируют преимущество алгоритмов машинного обучения, в частности глубоких нейронных сетей (Deep Neural Networks, DNN), перед традиционными методами в прогнозировании долгосрочных результатов у пациентов с ишемическим инсультом, эффективность моделей машинного обучения по сравнению с регрессионными моделями, по данным разных исследователей, неоднозначна [12, 14]. Таким образом, создание надёжного прогностического инструмента для пациентов с инсультом остаётся сложной и не решённой до конца задачей.

Цель исследования — на основе алгоритмов машинного обучения разработать, обучить и апробировать модели, способные с первых суток от дебюта инсульта прогнозировать функциональный исход с оценкой по шкале инсульта Национального института здоровья США (National Institutes of Health Stroke Scale, NIHSS) и модифицированной шкале Ранкина (modified Rankin scale, mRs) при выписке из стационара на основании клинико-анамнестических предикторов.

МЕТОДЫ

Дизайн исследования

Выполнено историческое когортное исследование, проведённое среди пациентов с острым инсультом, выписанных из неврологических отделений Регионального сосудистого центра (РСЦ) ГБУЗ Архангельской области «Архангельская областная клиническая больница» (РСЦ АОКБ) и РСЦ ГБУЗ Архангельской области «Первая городская клиническая больница имени Е.Е. Волосевич» (РСЦ 1 ГКБ) в период с 1 января 2017 года по 31 декабря 2020 года.

Критерии соответствия

Критерии включения: все пациенты с инсультом, выписанные из РСЦ АОКБ и РСЦ 1 ГКБ с января 2017 по декабрь 2020 года.

Критерии исключения: пациенты с диагнозом аневризматической болезни головного мозга; субарахноидальное кровоизлияние; пациенты, в первичной документации которых отсутствовала оценка по NIHSS и mRs.

Условия проведения

Все пациенты, включённые в исследование, находились на лечении в РСЦ АОКБ и РСЦ 1 ГКБ.

Продолжительность исследования

Все пациенты были госпитализированы в РСЦ АОКБ и РСЦ 1 ГКБ в период с января 2017 года по декабрь 2020 года.

Описание медицинского вмешательства

Все пациенты получали стандартную терапию, соответствующую клиническим рекомендациям; реперфузионные методики применялись только у пациентов РСЦ 1 ГКБ; реабилитационные мероприятия осуществлялись силами специалистов мультидисциплинарной бригады. Дополнительных медицинских вмешательств медикаментозной и немедикаментозной природы в процессе исследования не проводилось.

Построение моделей машинного обучения

Предварительная обработка данных. Для обеспечения целостности и единообразия набора данных текстовое содержимое клинических документов было предварительно обработано, и на их основе создана компьютерная база данных с возможностью проведения статистического анализа, который проводился с использованием пакетов Microsoft Excel 2010 (США), Stata версии 18 (Stata Corp., США).

Выбор предикторов. Для разработки моделей машинного обучения изначально из базы данных была извлечена 21 переменная, включая демографические данные пациентов, исходные баллы по NIHSS и mRS, подтипы инсульта в соответствии с классификационной системой Trial of ORG 10472 in Acute Stroke Treatment (TOAST), время от начала заболевания до госпитализации, сведения о сопутствующих заболеваниях (табл. 1). Для отбора признаков использовались четыре специализированных метода (SelectKBest, SelectPercentile, Logistic Regression, Random Forest), а их эффективность оценивалась с помощью классификатора SVC с RBF-ядром. Важность финальных предикторов определена независимым методом SHAP, выявляющим вклад каждого признака в прогноз.

Определение цели прогнозирования. Целью прогнозирования являлось определение наиболее вероятного балла по шкалам NIHSS и mRs к моменту выписки пациента из РСЦ. Ввиду различия оценочных шкал по характеристикам функционального состояния разрабатывали две модели для прогнозирования исхода по каждой шкале с оценкой в баллах (модель I для NIHSS, модель II для mRs).

Таблица 1. Характеристики и количественная оценка переменных, используемых для разработки моделей машинного обучения**Table 1.** Characteristics and quantification of parameters used to develop machine learning models

Переменная	Характеристика	Кодировка	Метрика
1	Возраст (возрастные группы), лет	age	<ul style="list-style-type: none"> • младше 20 • 20–29 • 30–39 • 40–44 • 45–49 • 50–59 • 60–69 • 70–74 • 75–79 • 80–89 • 90 и старше
2	Пол	gender	мужчины/женщины
3	mRs при поступлении	rank_adm	балл
4	NIHSS при поступлении	nihss_adm	балл
5	Тип инсульта	stroke	<ul style="list-style-type: none"> • атеротромботический • кардиоэмболический • криптогенный • лакунарный • внутримозговое кровоизлияние • венозный инсульт • другой уточнённой этиологии
6	Время от начала заболевания до госпитализации	time	<ul style="list-style-type: none"> • до 6 ч • 6–24 ч • более 24 ч
7	Проведение мероприятий экстренной реваскуляризации головного мозга	revasc	да/нет
<i>Наличие сопутствующих заболеваний</i>			
8	артериальная гипертензия	hyper	да/нет
9	сахарный диабет	diabet	да/нет
10	фибрилляция предсердий	atrialf	да/нет
11	церебральный атеросклероз	cerebr	да/нет
12	ишемическая болезнь сердца	cardiac	да/нет
13	хроническая сердечная недостаточность	heartf	да/нет
14	ожирение	obes	да/нет
15	острый инфаркт миокарда в анамнезе (постинфарктный кардиосклероз)	pics	да/нет
16	стенокардия	angina	да/нет
17	хроническая болезнь почек	kidney	да/нет
18	онкозаболевание	onco	да/нет
19	количество сопутствующих заболеваний	quantity	абс.
20	mRs при выписке	rank_dis	балл
21	NIHSS при выписке	nihss_dis	балл

Примечание. mRS — модифицированная шкала Рэнкина; NIHSS — шкала инсульта Национального института здоровья США.

Note. mRS, Modified Rankin Scale; NIHSS, National Institutes of Health Stroke Scale.

Построение моделей прогнозирования. Были использованы два алгоритма машинного обучения — ANN и RF. ANN состоит из слоёв взаимосвязанных искусственных нейронов. Искусственный нейрон создан по образу

биологического нейрона, получает несколько входных данных, умноженных на весовые коэффициенты, и выдаёт сумму входных данных. Алгоритм RF — ансамблевый метод, состоящий из множества деревьев решений (DT),

включающих несколько условий, которые могут быть истинными или ложными, с использованием входных переменных. Для окончательной классификации используется сумма решений, принятых алгоритмом деревьев решений (DT), что повышает точность прогнозирования и уменьшает вероятность переобучения [15].

Модели машинного обучения были обучены с использованием отобранных ранее переменных в качестве входных данных для классификации пациентов с разными вариантами функциональных исходов при выписке. Для ANN использовали четыре скрытых слоя, для алгоритма RF — 500 деревьев решений (для каждой модели).

Вклад каждого предиктора в результат прогнозирования оценивали с помощью метода аддитивных объяснений Шепли (SHapley Additive exPlanations, SHAP) [16].

Анализ данных и построение моделей проводили с помощью языка программирования Python с открытым исходным кодом.

Обучение осуществлялось на стационарном компьютере со следующими техническими характеристиками: видеокарта Nvidia GeForce GTX 4090 (24 Gb видеопамяти), процессор Intel i3500 (8 ядер, 32 потока, базовая частота 3,6 GHz), оперативная память DDR4 (128 Gb памяти с частотой 3200 MHz).

Оценка моделей прогнозирования. Качество моделей оценивали с помощью метода площади под кривой операционной характеристики приёмника (Receiver Operating Characteristic – Area Under Curve, ROC-AUC).

Исходы исследования

Основной исход исследования: оценка достижения пациентом с инсультом определённого функционального состояния, оценённого в баллах по шкалам NIHSS и mRS, к моменту выписки из РСЦ.

Методы регистрации исходов

Информацию об исходах извлекали из историй болезни и базы данных выписных эпикризов.

Статистический анализ

Размер выборки предварительно не рассчитывался.

Предварительно были выполнены описательные статистики. Категориальные переменные представлены как абсолютные значения и процентные доли. Числовые характеристики записывали в виде среднего арифметического (среднее значение балла по NIHSS) и показателя медианы (Me для баллов по mRS). Проверка законов распределения значений количественных показателей выполнялась с использованием статистического критерия Колмогорова–Смирнова. Детальное описание статистического анализа представлено в предыдущей публикации [17]. Математический и статистический анализ результатов проводился с использованием пакетов Microsoft Excel 2010 (США) и Stata версии 18 (Stata Corp., США).

РЕЗУЛЬТАТЫ

Объекты (участники) исследования

В исследование были включены 5227 человек в возрасте от 18 до 98 лет в острый период инсульта: 3588 пациентов РСЦ 1 ГКБ и 1639 пациентов РСЦ АОКБ. После предварительной обработки данных количество записей для построения прогностической модели составило 5225. Записи изучаемой когорты после предварительной проверки и очистки были разделены случайным образом на две выборки: 4180 (80%) записей были определены в набор для обучения и кросс-валидации, остальные 1045 (20%) использовались в качестве набора для тестирования с целью оценки производительности обученных моделей.

Основные результаты исследования

Выбор предикторов. Выделенная из базы данных исходная 21 переменная была обработана пятью инструментами для отбора признаков в задачах машинного обучения (selectkbest, selectpercentile, logisticregression, RF и SVC с ядром RBF).

Наиболее эффективным алгоритмом для решения данной задачи оказался SVC с ядром RBF и предварительно удалёнными высококоррелированными признаками (средняя AUC 0,827 для модели I, средняя AUC 0,849 для модели II). Из дальнейших расчётов были удалены переменные atrialf, cardiac, heartf в связи с выявленной высокой корреляцией между категориальными признаками и целевыми переменными (балл по NIHSS и mRS при выписке).

В итоге для модели I (nihss_dis) были отобраны 11 признаков: «revasc», «time», «gender», «age», «stroke», «diabet», «pics», «kidney», «quantity», «nihss_adm», «rank_adm». Наиболее значимыми предикторами (в порядке убывания) стали «nihss_adm», «age», «time», «quantity», «stroke» (рис. 1).

Для модели II (rank_dis) были отобраны 12 признаков: «time», «gender», «age», «stroke», «hyper», «diabet», «pics», «angina», «kidney», «quantity», «nihss_adm», «rank_adm». Наиболее значимыми предикторами (в порядке убывания) стали «nihss_adm», «rank_adm», «time», «stroke», «age» (рис. 2).

Построение и обучение моделей прогнозирования

Для выполнения прогноза были выбраны алгоритмы ANN и RF для двух моделей с исходами nihss_dis и rank_dis. Подбор оптимальных параметров для построения выполняли по сетке.

Оба алгоритма решали задачу классификации по множественному набору значений исхода (26 классов для модели I, соответствующие баллам по NIHSS от 0 до 25, и 6 классов для модели II, соответствующие баллам по mRS от 0 до 5).

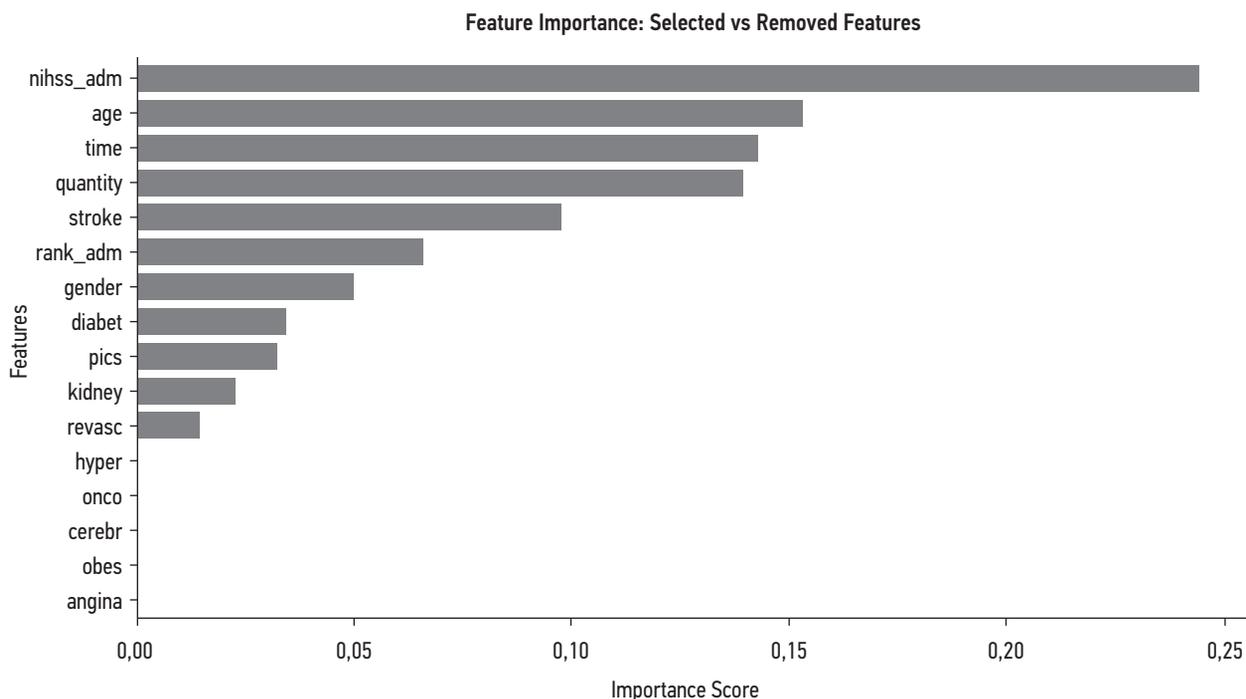


Рис. 1. Отобранные признаки для прогнозирования исхода nihss_dis.

Fig. 1. Features selected to predict nihss_dis.

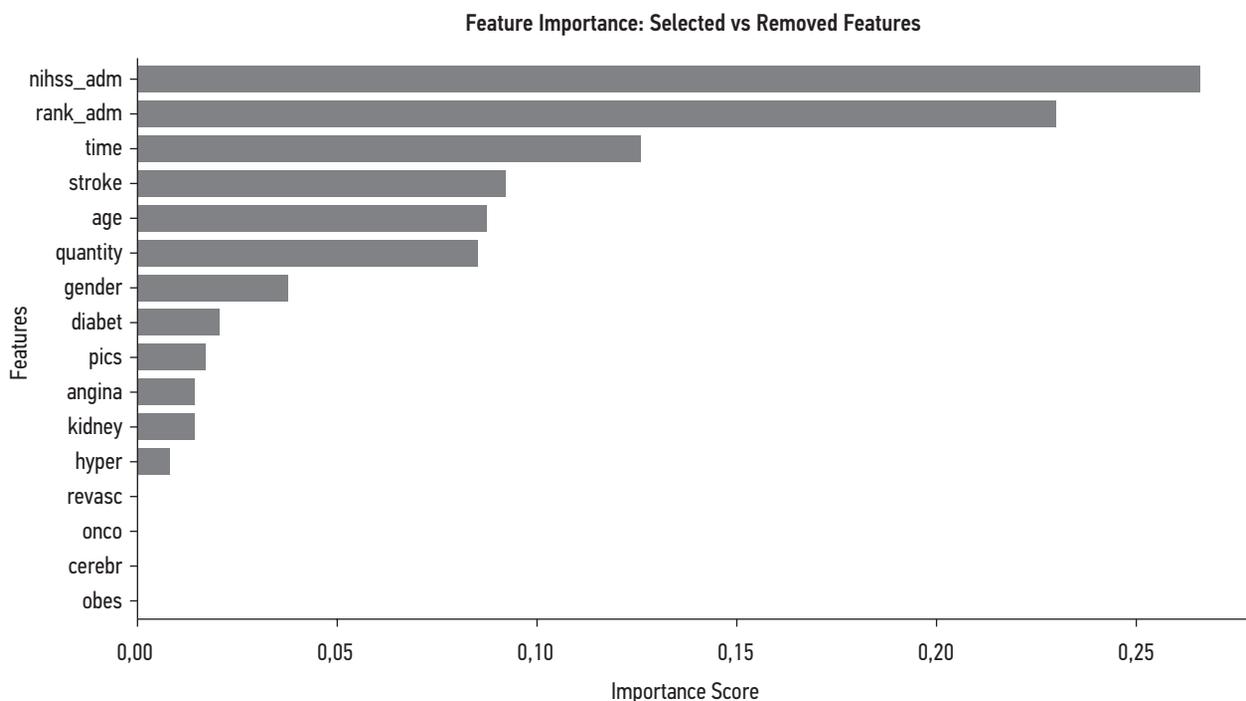


Рис. 2. Отобранные признаки для прогнозирования исхода rank_dis.

Fig. 2. Features selected to predict rank_dis.

Архитектура разрабатываемой ANN соответствовала многослойному перцептрону. Входной слой для каждой модели включал 11 нейронов для модели I и 12 нейронов для модели II, соответствующих количеству выделенных

предикторов. Четыре промежуточных слоя ANN содержали от 1024 до 64 нейронов. Выходной слой для каждой модели представлял собой один нейрон. После обучения построенных моделей получены результаты эффективности

по AUC-ROC: 0,771 для модели I и 0,844 для модели II. Процесс обучения ANN занял 68 минут. Показатели эффективности каждого алгоритма для моделей I и II в выборке для обучения представлены в табл. 2.

При формировании RF использовался обучающий набор с 11 и 12 ранее отобранными признаками для моделей I и II соответственно с построением 500 деревьев решений (DT) для каждой модели. Процесс обучения RF занял 12 минут.

Для дальнейшей работы было решено выбрать алгоритм RF ввиду его лучшей интерпретируемости

и сравнимых показателей эффективности относительно ANN.

Для проверки прогностической эффективности алгоритма RF был использован внешний набор записей 783 пациентов с инсультом, проходившими лечение в РСЦ 1 ГKB с 01.01.2022 по 31.12.2022. Показатели ROC-AUC при отработке алгоритма RF на валидационной выборке составили 0,786 для модели I и 0,774 для модели II (рис. 3, 4).

С целью определения вклада каждого фактора в принятие решений моделью использовали метод

Таблица 2. Показатели прогностической эффективности алгоритмов ANN и RF для моделей I и II

Table 2. Predictive performance parameters of ANN and RF algorithms for models I and II

Модель	Алгоритм, ROC-AUC	
	ANN	RF
I (для исхода nihss_dis)	0,771	0,778
II (для исхода rank_dis)	0,844	0,845

Примечание. ROC-AUC — метрика оценки качества модели машинного обучения; ANN — вычислительная модель, вдохновлённая структурой и функциями биологических нейронных сетей; RF — метод случайного леса.

Note. ROC-AUC, metric used to evaluate the quality of a machine learning model. ANN, computational model inspired by the structure and function of biological neural networks. RF, random forest method.

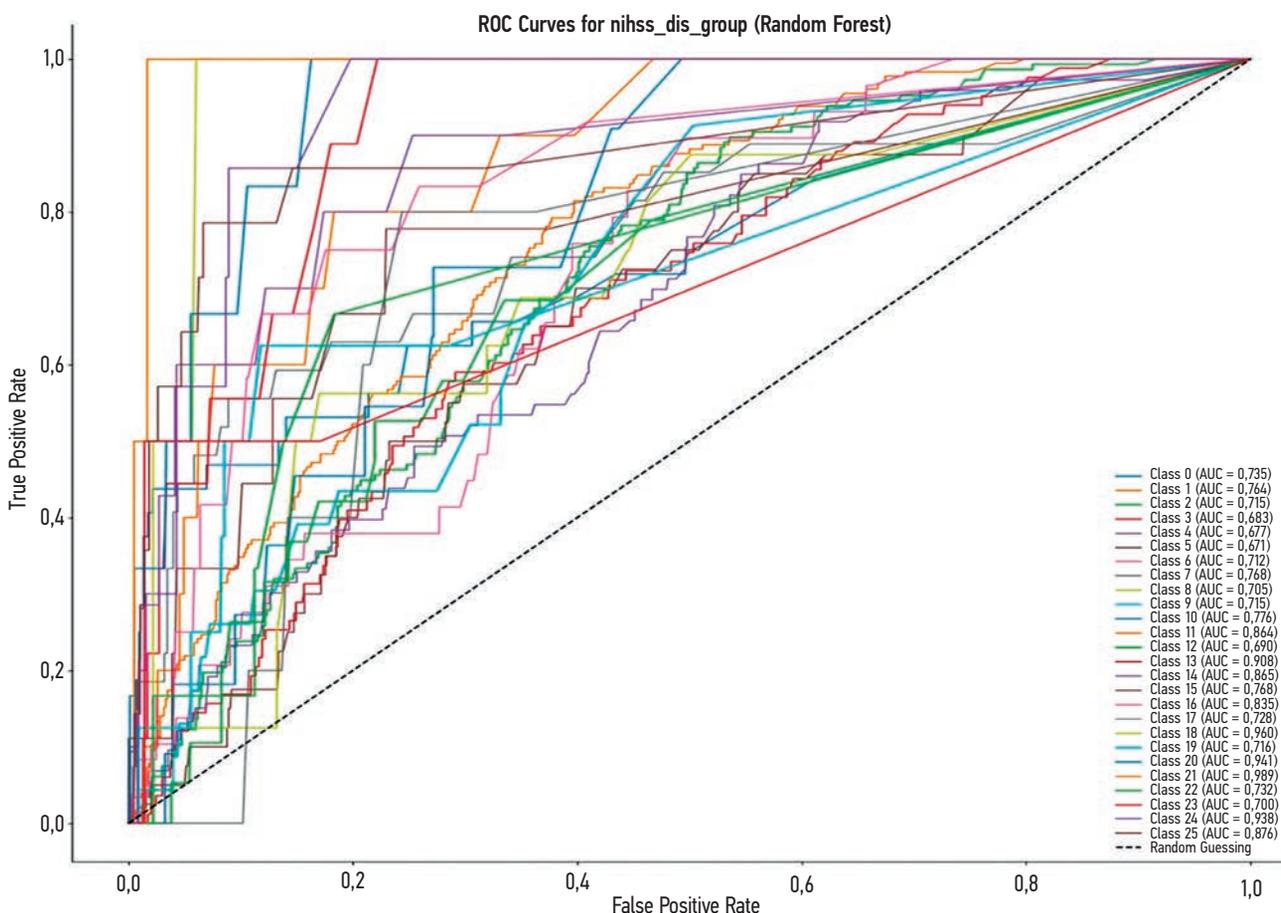


Рис. 3. ROC-AUC алгоритма RF для модели I на валидационной выборке. ROC-AUC — метрика оценки качества модели машинного обучения; RF — метод случайного леса.

Fig. 3. ROC-AUC of RF for model I on the validation set. ROC-AUC, metric used to evaluate the performance of a machine learning model. RF, random forest.

SHAP. Наиболее важными переменными в модели I были «nihss_adm», «rank_adm», «time», «age», «stroke» (в порядке убывания) (рис. 5). Наиболее значимыми признаками, влияющими на прогнозируемый исход в модели II, стали «rank_adm», «nihss_adm», «age», «time», «stroke» (в порядке убывания) (рис. 6). Среди факторов, отражающих наличие сопутствующих заболеваний, для обеих моделей наиболее значимый вклад внёс сахарный диабет (см. рис. 5, 6).

Для практического применения на основании разработанных моделей был создан калькулятор с возможностью использования в формате одностраничного сайта. На примере пациента из валидационной выборки продемонстрирована возможность реализации программы в клинической практике в виде online-калькулятора (рис. 7).

ОБСУЖДЕНИЕ

Резюме основного результата исследования

В результате выполненной работы разработаны и обучены модели на основе алгоритмов RF и ANN, способные с первых суток от начала инсульта прогнозировать функциональный исход с оценкой по NIHSS и mRS при выписке из стационара на основании клинико-анамнестических предикторов. Модели, обученные на основе алгоритма RF, апробированы путём внешней валидации, результаты разработанных моделей включены в структуру калькулятора для выполнения прогноза.

Обсуждение основного результата исследования

Точность прогнозирования функционального восстановления после инсульта важна для проведения своевременных и целенаправленных мероприятий, улучшения

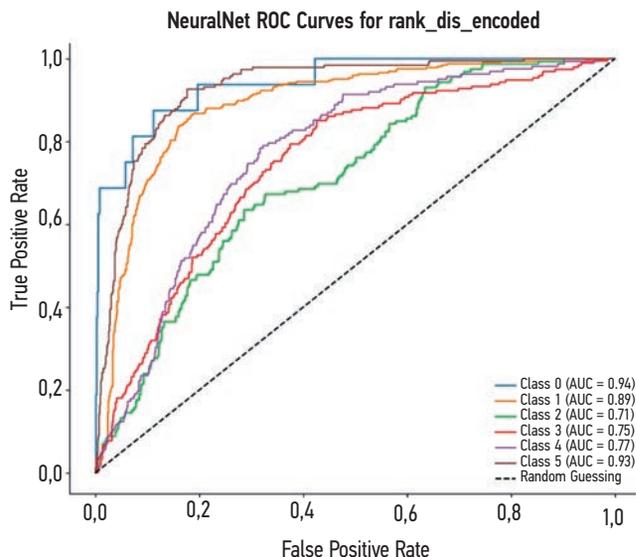


Рис. 4. ROC-AUC алгоритма RF для модели II на валидационной выборке. ROC-AUC — метрика оценки качества модели машинного обучения; RF — метод случайного леса.

Fig. 4. ROC-AUC of RF for model II on the validation set. ROC-AUC, metric used to evaluate the performance of a machine learning model. RF, random forest.

восстановления, распределения ресурсов, а также снижения экономических последствий постинсультной инвалидности [18]. И хотя прогностическая модель как объективный инструмент для принятия решений должна служить оптимизации реабилитационного процесса у пациентов с инсультом и снижению неопределённости в отношении прогнозируемых функций, отношение специалистов к возможностям выполнения такого прогноза неоднозначно. Исследование С. Кіæг и соавт., выполненное в 2018 году [19], показало, что только 35% специалистов

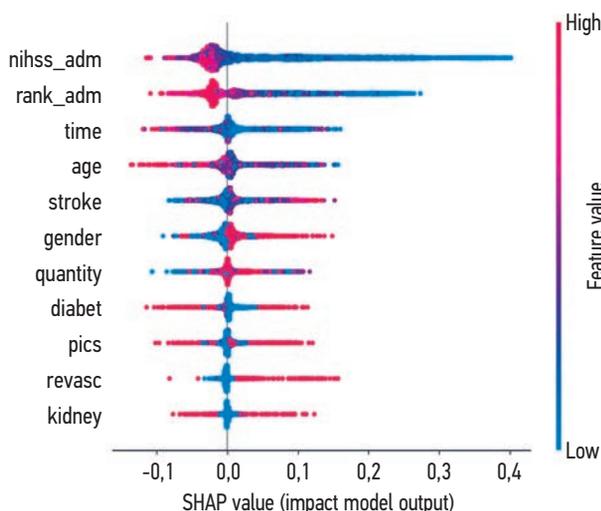


Рис. 5. Важность выбранных переменных для прогнозирования для модели I с исходом nihss_dis.

Fig. 5. Value of parameters selected to predict nihss_dis for model I.

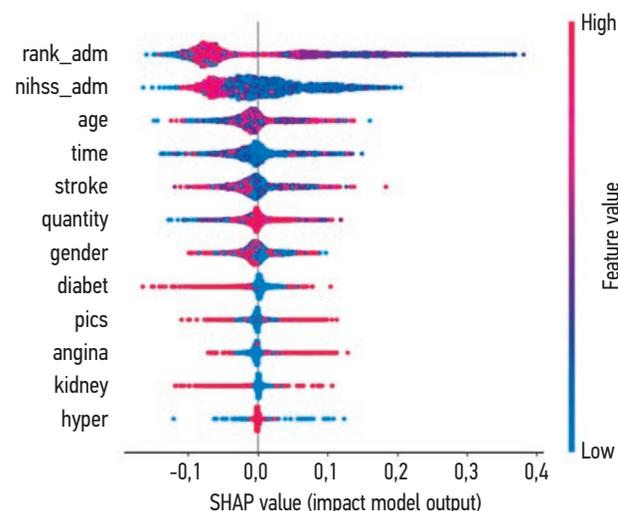


Рис. 6. Важность выбранных переменных для прогнозирования для модели II с исходом rank_dis.

Fig. 6. Value of parameters selected to predict rank_dis for model II.

Введите данные для прогнозирования исхода по шкалам National Institutes of Health Stroke Scale (NIHSS) и modified Rankin scale (mRS)

Время доезда от дебюта инсульта: 6-24 часа

Пол пациента: Мужской

Возраст пациента: 50-59 лет

Тип инсульта: Криптогенный

NIHSS при поступлении (балл): 9

mRS при поступлении (балл): 3

Артериальная гипертензия: НЕТ ДА

Сахарный диабет: НЕТ ДА

ОИМ в анамнезе (ПИКС): НЕТ ДА

Стенокардия: НЕТ ДА

ХБП: НЕТ ДА

Были ли проведены мероприятия экстренной реваascularизации головного мозга?: НЕТ ДА

Наиболее вероятный балл по mRS: 3 Наиболее вероятный балл по NIHSS: 1

Рис. 7. Пример использования калькулятора для прогнозирования раннего функционального исхода инсульта.

Fig. 7. An example of using a calculator to predict an early functional outcome of stroke.

по трудотерапии и физиотерапевтов были осведомлены о моделях прогнозирования, и только 9% использовали модели прогнозирования в своей клинической практике.

Вопрос клинического применения прогностических моделей наиболее значимо встает в условиях удаленности медицинского учреждения, оказывающего помощь пациентам с инсультом, дефицита профильных специалистов и диагностического оборудования, с которыми сталкиваются сельские больницы. Помимо того, что сельские пациенты с инсультом по сравнению с городскими реже получают тромболитическую терапию и имеют более высокий уровень смертности, они также находятся в менее выигрышной ситуации по возможностям лечения, наблюдения и реабилитации в постстационарный период [20, 21].

Активное внедрение телемедицинских технологий позволяет улучшать эту ситуацию, при этом выявляется необходимость получения простого и универсального прогностического инструмента, который позволит даже в отсутствие профильного специалиста спрогнозировать ранний функциональный исход у пациента с инсультом и принять решение о реабилитационной маршрутизации, спланировать транспортировку в реабилитационное отделение или использовать возможности телереабилитации на месте [22].

В исследованиях последних лет подчеркивается важность использования как клинических, так и нейровизуализационных данных для создания оптимальной модели прогнозирования исходов инсульта, так как интеграция данных об очаге поражения и других радиомических характеристик может значительно увеличивать точность прогнозов [6, 23]. Так, наиболее точным краткосрочным

прогностическим инструментом для пациентов с инсультом, получающих системную тромболитическую терапию в сельской местности, определена шкала для прогнозирования функционального исхода через 3 месяца у пациентов с инсультом MRI-DRAGON (Magnetic Resonance Imaging-DRAGON) [24]. Однако применение моделей с включением в качестве прогностических факторов данных нейровизуализации в сельских больницах часто ограничивается возможностью выполнения только скрининговой компьютерной томографии головного мозга. Именно поэтому важной характеристикой оптимальной прогностической модели является её универсальность с возможностью использования не только в специализированных сосудистых центрах, но и в условиях ограниченности ресурсов.

Традиционный подход к прогнозированию исходов инсульта предполагает использование классических статистических моделей, в частности логистической регрессии. Модели логистической регрессии выявляют и подтверждают прогностические переменные, их главное преимущество заключается в лёгкости реализации и интерпретации. Однако в последнее десятилетие предпринимались неоднократные попытки разработки и введения в клиническую практику алгоритмов машинного обучения, способных учитывать нелинейные и сложные взаимодействия между несколькими прогностическими переменными. Так, по результатам систематического обзора методов прогнозирования исходов инсульта на основе структурированных данных (с исключением анализа изображений и текста), W. Wang и соавт. [12] продемонстрировали, что в шести исследованиях модели машинного обучения превосходили сравниваемые регрессионные модели, а в пяти исследованиях существенной разницы между статистическими и моделями машинного обучения не обнаружено. Данные исследований групп S.A. Alaka [25] и J. Neo [26] также показали, что алгоритмы логистической регрессии и машинного обучения обладают сопоставимой точностью прогнозирования. А в работе G. Çelik и соавт. [27] при прогнозировании 10-дневной смертности пациентов с инсультом получены результаты, свидетельствующие о том, что логистическая регрессия уступала ANN в показателях точности на этапе обучения, но превосходила ANN на этапе тестирования.

При выборе оптимального алгоритма для решения задачи прогнозирования необходимо сравнивать показатели эффективности не только с классическими статистическими методами, но и с показателями других алгоритмов машинного обучения. Ранее выполненные исследования демонстрируют лучшую сравнительную эффективность в прогнозировании исхода инсульта на структурированных данных SVC в трёх исследованиях, ANN — в двух исследованиях, RF — в двух исследованиях, алгоритма логистической регрессии — в двух исследованиях [12]. Нами были обучены алгоритмы RF и ANN для прогнозирования ранних функциональных исходов у пациентов с инсультом по двум вариантам: с оценкой по NIHSS

и mRs. Оба алгоритма после выполнения внутренней валидации продемонстрировали сравнимые показатели прогностической эффективности в отношении обоих исходов: хорошей для модели I и очень хорошей для модели II (см. табл. 2). Полученные нами результаты показали более значимую прогностическую ценность алгоритма RF в сравнении с аналогичными моделями, прогнозирующими исход с оценкой mRs, в работе S.A. Alaka и соавт. (ROC-AUC 0,63, 0,67 и 0,70) [25]. Схожие с нашими показатели эффективности для моделей с исходом по mRS были получены в работе J. Neo, где ROC-AUC для ANN составила 0,888, а для RF — 0,857 [26]. В изученной литературе нами не обнаружено сведений для сравнения эффективности моделей машинного обучения по прогнозированию исхода с оценкой результата по NIHSS. Сравнение эффективности классификаторов RF и ANN было найдено в публикации N.F. Zahidi и соавт., где, прогнозируя состояние здоровья на основе данных о жизненно важных показателях, авторы выявили большую точность алгоритма RF относительно ANN [28].

Для дальнейшей работы по созданию прогностической модели нами было принято решение выбрать один алгоритм, которым стал RF. В пользу выбранного алгоритма сыграло затраченное на него время обучения и интерпретируемость модели. Так, для обучения RF, который использует способы для снижения риска переобучения и требуемого времени обучения, потребовалось 12 минут в сравнении с 68 минутами для ANN.

Клиническому применению моделей машинного обучения препятствует проблема интерпретируемости, связанная с «чёрным ящиком», который скрывает методы работы алгоритма от пользователя, и практикующим врачам сложно доверять прогнозам модели и учитывать их при принятии решений. Ансамблевые методы, к которым относится RF, в целом обеспечивают лучшую интерпретируемость по сравнению с ANN, поскольку они позволяют количественно оценить важность признаков и создают более прозрачные прогнозы [29]. В нашем исследовании для интерпретации разработанных моделей использовался метод SHAP. С помощью диаграммы `shap.plots.beeswarm` проиллюстрировано влияние каждого признака на прогнозируемый исход (см. рис. 5, 6). Важность вклада каждого фактора представлена в порядке убывания (сверху вниз), а цветовое решение указывает на зависимость исхода от большего (красный цвет) или меньшего показателя признака (синий цвет). С одной стороны, это демонстрирует значимость вклада выделенных факторов в результат прогноза, с другой — свидетельствует об интерпретируемости модели, которая выводит признаки, в большей степени влияющие на исход, коррелирующие с теми, которые были получены на стадии отбора (см. рис. 1, 2).

Наиболее важными переменными, с учётом двух проанализированных моделей, стали возраст пациента, а также исходные показатели по шкалам NIHSS и mRs,

отражающие тяжесть инсульта на момент госпитализации (см. рис. 5, 6). Эти данные подтверждаются результатами ранее проведённых исследований, например, S.A. Alaka и соавт. [25], прогнозировавшими функциональные исходы у пациентов с инсультом на 90-й день от начала заболевания. Обе наши модели продемонстрировали прогностическое значение такого фактора, как время, прошедшее от момента дебюта инсульта до поступления в специализированный неврологический стационар, при том что фактор проведения экстренных реваскуляризационных методик имел меньшее значение в модели I с исходом по NIHSS (см. рис. 5) и не учитывался в модели II с исходом по mRs (см. рис. 6). Из сопутствующей патологии наиболее значимым фактором, влияющим на исход, оказался сахарный диабет, хотя, по данным литературных источников, большее прогностическое значение имеет не само наличие сахарного диабета в анамнезе, а уровень глюкозы крови в острейший период [25].

Внешняя проверка — валидация (оценка точности прогнозов на данных из другого источника, которые использовались для разработки модели) — важна для оценки применимости модели в условиях, отличных от условий обучающей и тестирующей выборки [26]. Это неотъемлемая часть внедрения модели в рутинную клиническую практику, позволяющая оценить применимость прогнозов к новым данным. В нашем исследовании модель I, прогнозирующая исход по NIHSS, показала схожие результаты на обучающем (0,778) и валидационном (0,786) наборах данных (см. табл. 2 и рис. 4). Модель II, прогнозирующая исход по mRs, на валидационном наборе продемонстрировала меньшую эффективность (0,774), чем на обучающей выборке (0,845), что может указывать на наличие признаков переобучения.

В перспективе возможно улучшение прогностических характеристик моделей при условии их интеграции в систему электронной истории болезни и амбулаторной карты пациента и добавления характеристик, доступных в более поздние периоды времени.

Ограничения исследования

Данное исследование ограничено его ретроспективным характером, а также необходимостью валидации моделей на данных из других лечебных учреждений.

ЗАКЛЮЧЕНИЕ

Разработанные нами модели на основе алгоритма RF способны в течение первых 24 часов от начала инсульта достоверно прогнозировать ранний функциональный исход с оценкой по NIHSS и mRs на основании клинико-анамнестических предикторов. Полученный инструмент полезен для формирования персонализированных терапевтических и реабилитационных стратегий в острый период инсульта, определения реабилитационного

маршрута и проведения прогностических бесед с пациентами и их родственниками. Универсальность и простота использования моделей определяет их ценность и позволяет применение в удалённо расположенных сельских медицинских учреждениях с дефицитом или отсутствием опытных профильных специалистов и диагностического оборудования.

ДОПОЛНИТЕЛЬНАЯ ИНФОРМАЦИЯ

Вклад авторов. Н.М. Хасанова — разработка концепции, общее руководство, редактирование; Е.М. Черных — разработка концепции, сбор, обработка данных, анализ данных, написание текста рукописи; А.А. Карякин — разработка концепции, анализ данных, статистическая обработка, написание текста рукописи; З.Э. Джафарова — сбор, обработка данных, написание текста рукописи; А.А. Ключас — сбор, обработка данных. Все авторы одобрили рукопись (версию для публикации), а также согласились нести ответственность за все аспекты работы, гарантируя надлежащее рассмотрение и решение вопросов, связанных с точностью и добросовестностью любой её части.

Этическая экспертиза. На проведение исследования получено положительное этическое заключение локального этического комитета ФГБОУ ВО «Северный государственный медицинский университет» (протокол заседания № 09/10-2 от 26.10.2022).

Источники финансирования. Отсутствуют.

Раскрытие интересов. Авторы заявляют об отсутствии отношений, деятельности и интересов за последние три года, связанных с третьими лицами (коммерческими и некоммерческими), интересы которых могут быть затронуты содержанием статьи.

Оригинальность. При проведении исследования и создании настоящей работы авторы не использовали ранее опубликованные сведения (текст, иллюстрации, данные).

Доступ к данным. Редакционная политика в отношении совместного использования данных к настоящей работе неприменима.

Генеративный искусственный интеллект. При создании настоящей статьи технологии генеративного искусственного интеллекта не использовались.

ADDITIONAL INFORMATION

Author contributions. N.M. Khasanova: concept development, general management, editing; E.M. Chernykh: concept development, data collection, processing, data analysis, writing the manuscript; A.A. Karyakin: concept development, data analysis, statistical processing, writing the manuscript; Z.E. Dzhaфарova: collection, processing of data, writing of the manuscript; A.A. Klyukas: data collection and processing. Thereby, all authors provided approval of the version to be published and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Ethics approval. A positive ethical opinion was received for the study from the local ethics committee of the Federal State Budgetary Educational Institution of Higher Education "Northern State Medical University" (No. 09/10-2 dated October 26, 2022).

Funding sources. No funding.

Disclosure of interests. The authors have no relationships, activities or interests for the last three years related with for-profit or not-for-profit third parties whose interests may be affected by the content of the article.

Statement of originality. When conducting the research and creating this work, the authors did not use previously published information (text, illustrations, data).

Data availability statement. Editorial policy on data sharing not applicable to this work. Access to the data obtained in this study is closed due to confidential patient information.

Generative AI. Generative AI technologies were not used for this article creation.

СПИСОК ЛИТЕРАТУРЫ | REFERENCES

- Bösel J, Mathur R, Cheng L, et al. AI and neurology. *Neurol Res Pract*. 2025;7(1):11. doi: 10.1186/s42466-025-00367-2
- AbuAlrob MA, Mesraoua B. Harnessing artificial intelligence for the diagnosis and treatment of neurological emergencies: A comprehensive review of recent advances and future directions. *Front Neurol*. 2024;15:1485799. doi: 10.3389/fneur.2024.1485799
- Mainali S, Darsie ME, Smetana KS. Machine learning in action: Stroke diagnosis and outcome prediction. *Front Neurol*. 2021;12:734345. doi: 10.3389/fneur.2021.734345
- Abedi V, Kawamura Y, Li J, et al. Editorial: Machine learning in action: Stroke diagnosis and outcome prediction. *Front Neurol*. 2022;13:984467. doi: 10.3389/fneur.2022.984467
- Abedi V, Khan A, Chaudhary D, et al. Using artificial intelligence for improving stroke diagnosis in emergency departments: A practical framework. *Ther Adv Neurol Dis*. 2020;13:1756286420938962. doi: 10.1177/1756286420938962
- Rajashekar D, Hill MD, Demchuk AM, et al. Prediction of clinical outcomes in acute ischaemic stroke patients: A comparative study. *Front Neurol*. 2021;12:663899. doi: 10.3389/fneur.2021.663899
- Yang Y, Tang L, Deng Y, et al. The predictive performance of artificial intelligence on the outcome of stroke: A systematic review and meta-analysis. *Front Neurol*. 2023;17:1256592. doi: 10.3389/fnins.2023.1256592
- Qu S, Zhou M, Jiao S, et al. Optimizing acute stroke outcome prediction models: Comparison of generalized regression neural networks and logistic regressions. *PLoS One*. 2022;17(5):e0267747. doi: 10.1371/journal.pone.0267747
- Lee J, Park KM, Park S. Interpretable machine learning for prediction of clinical outcomes in acute ischemic stroke. *Front Neurol*. 2023;14:1234046. doi: 10.3389/fneur.2023.1234046
- Gaviria E, Eltayeb Hamid AH. Neuroimaging biomarkers for predicting stroke outcomes: A systematic review. *Health Sci Rep*. 2024;7(7):e2221. doi: 10.1002/hsr2.2221
- Pedersen M, Verspoor K, Jenkinson M, et al. Artificial intelligence for clinical decision support in neurology. *Brain Commun*. 2020;2(2):fcaa096. doi: 10.1093/braincomms/fcaa096
- Wang W, Kiik M, Peek N, et al. A systematic review of machine learning models for predicting outcomes of stroke with structured data. *PLoS One*. 2020;15(6):e0234722. doi: 10.1371/journal.pone.0234722
- Alyasein S, Alqaran R, Al-Aiad A. Using machine learning algorithms for predicting stroke disease-2024. In: 15th International Conference on Information and Communication Systems (ICICS), Irbid, Jordan; 2024. P. 1–6. doi: 10.1109/ICICS63486.2024.10638280
- Heo J, Yoon JG, Park H, et al. Machine learning-based model for prediction of outcomes in acute stroke. *Stroke*. 2019;50(5):1263–1265. doi: 10.1161/STROKEAHA.118.024293
- Daidone M, Ferrantelli S, Tuttolomondo A. Machine learning applications in stroke medicine: Advancements, challenges, and future perspectives. *Neural Reg Res*. 2024;19(4):769–773. doi: 10.4103/1673-5374.382228
- Jeon ET, Jung SJ, Yeo TY, et al. Predicting short-term outcomes in atrial-fibrillation-related stroke using machine learning. *Front Neurol*. 2023;14:1243700. doi: 10.3389/fneur.2023.1243700
- Chernykh EM, Khasanova NM, Grzhibovsky AM, Bogdanova AI. Functional predictor variables in the acute period of stroke in patients living in the Arctic Macroregion: Multivariate modeling findings. *Physical and rehabilitation medicine, medical rehabilitation*. 2024;6(4):336–347. doi: 10.36425/rehab637339 EDN: LNEYEF

18. Van der Groen O, Ghosh M, Norman R, et al. Point of view on outcome prediction models in post-stroke motor recovery. *Neurorehab Neural Rep.* 2024;38(5):386–398. doi: 10.1177/15459683241237975
19. Kiżer C, Lundquist CB, Brunner I. Knowledge and application of upper limb prediction models and attitude toward prognosis among physiotherapists and occupational therapists in the clinical stroke setting. *Top Stroke Rehabil.* 2021;28(2):135–141. doi: 10.1080/10749357.2020.1783915
20. Man S, Bruckman D, Uchino K, et al. Rural hospital performance in guideline-recommended ischemic stroke thrombolysis, secondary prevention, and outcomes. *Stroke.* 2024;55(10):2472–2481. doi: 10.1161/STROKEAHA.124.047071
21. Loccoh EC, Joynt Maddox KE, Wang Y, et al. Rural-urban disparities in outcomes of myocardial infarction, heart failure, and stroke in the United States. *J Am College Cardiol.* 2022;79(3):267–279. doi: 10.1016/j.jacc.2021.10.045
22. Ali F, Hamid U, Zaidat O, et al. Role of artificial intelligence in telestroke: An overview. *Front Neurol.* 2020;11:559322. doi: 10.3389/fneur.2020.559322
23. Zhou Y, Wu D, Yan S, et al. Feasibility of a clinical-radiomics model to predict the outcomes of acute ischemic stroke. *Korean J Radiol.* 2022;23(8):811–820. doi: 10.3348/kjr.2022.0160
24. Loggini A, Saleh Velez FG, Hornik J, et al. Validation of prognostic scales for functional outcome in ischemic stroke patients treated with intravenous thrombolysis in a rural setting. *Cerebrovasc Dis Extra.* 2024;14(1):148–157. doi: 10.1159/000541801
25. Alaka SA, Menon BK, Brobbey A, et al. Functional outcome prediction in ischemic stroke: a comparison of machine learning algorithms and regression models. *Front Neurol.* 2020;11:889. doi: 10.3389/fneur.2020.00889
26. Heo J, Yoon JG, Park H, et al. Machine learning-based model for prediction of outcomes in acute stroke. *Stroke.* 2019;50(5):1263–1265. doi: 10.1161/STROKEAHA.118.024293
27. Çelik G, Baykan ÖK, Kara Y, Tireli H. Predicting 10-day mortality in patients with strokes using neural networks and multivariate statistical methods. *J Stroke Cerebrovasc Dis.* 2014;23(6):1506–1512. doi: 10.1016/j.jstrokecerebrovasdis.2013.12.018
28. Zahidi NF, Tukiran Z, Tukiran I. Analyzing Classification Models: Random Forest vs. Neural Networks in Health Prediction. In: *2024 International Conference on Future Technologies for Smart Society (ICFTSS)*. Kuala Lumpur, Malaysia; 2024. P. 183–187. doi: 10.1109/ICFTSS61109.2024.10691342
29. Khairunnisa A, Notodiputro KA, Sartono B. A comparative study of random forest and double random forest models from view points of their interpretability. *Sci J Inform.* 2024;11(1). doi: 10.15294/sji.v11i1.48721

ОБ АВТОРАХ

*** Черных Екатерина Михайловна;**

адрес: Россия, 163069, Архангельск, пр. Троицкий, д. 51;
ORCID: 0000-0002-6523-7071;
eLibrary SPIN: 8296-2286;
e-mail: raduga0302@mail.ru

Хасанова Нина Минувалиевна, канд. мед. наук, доцент;

ORCID: 0000-0003-0729-3726;
eLibrary SPIN: 6834-6281;
e-mail: khasanovanina@rambler.ru

Карякин Алексей Андреевич, канд. техн. наук, доцент;

ORCID: 0000-0002-4458-8702;
eLibrary SPIN: 7296-3303;
e-mail: biophysica@yandex.ru

Джафарова Зохра Эйвазовна;

ORCID: 0009-0003-2429-9158;
e-mail: zohrajafarova@yandex.ru

Клюкас Алексей Артемович;

ORCID: 0009-0005-3428-187X;
e-mail: alexeythekly@gmail.com

AUTHORS' INFO

*** Ekaterina M. Chernykh;**

address: 51 Troitsky ave, Arkhangelsk, Russia, 163069;
ORCID: 0000-0002-6523-7071;
eLibrary SPIN: 8296-2286;
e-mail: raduga0302@mail.ru

Nina M. Khasanova, MD, Cand. Sci. (Medicine), Assistant Professor;

ORCID: 0000-0003-0729-3726;
eLibrary SPIN: 6834-6281;
e-mail: khasanovanina@rambler.ru

Aleksey A. Karyakin, Cand. Sci. (Engineering), Assistant Professor;

ORCID: 0000-0002-4458-8702;
eLibrary SPIN: 7296-3303;
e-mail: biophysica@yandex.ru

Zohra E. Dzhaifarova;

ORCID: 0009-0003-2429-9158;
e-mail: zohrajafarova@yandex.ru

Alexey A. Kliukas;

ORCID: 0009-0005-3428-187X;
e-mail: alexeythekly@gmail.com

* Автор, ответственный за переписку / Corresponding author