

## Прогнозирование рецидива диффузного токсического зоба на основе клинических данных и сравнение методов машинного обучения и статистических моделей

Е. М. Дарий<sup>1</sup>✉, А. А. Мелдо<sup>1</sup>, С. В. Дора<sup>2</sup>, Ю. Ш. Халимов<sup>2</sup>

<sup>1</sup> Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В. И. Ульянова (Ленина), Санкт-Петербург, Россия

<sup>2</sup> Первый Санкт-Петербургский государственный медицинский университет им. акад. И. П. Павлова, Санкт-Петербург, Россия

✉ [evgeniadaryi@gmail.com](mailto:evgeniadaryi@gmail.com)

**Аннотация.** Проводится анализ моделей прогнозирования клинических событий на примере рецидива диффузного токсического зоба (ДТЗ) с применением методов машинного обучения и статистических подходов. Исследование охватывает данные 185 пациентов. Для определения наиболее значимых предикторов рецидива ДТЗ применялись алгоритмы случайного леса и экстремального градиентного бустинга, которые показали более высокую предсказательную способность (AUC 0.82–0.83) по сравнению с традиционными статистическими моделями. Данное исследование открывает новые перспективы для развития персонализированной медицины, в частности подходов к лечению пациентов с ДТЗ.

**Ключевые слова:** диффузный токсический зоб, машинное обучение, прогнозирование клинических событий, прогноз рецидива, тиреотоксикоз, факторы риска

**Для цитирования:** Прогнозирование рецидива диффузного токсического зоба на основе клинических данных и сравнение методов машинного обучения и статистических моделей / Е. М. Дарий, А. А. Мелдо, С. В. Дора, Ю. Ш. Халимов // Изв. СПбГЭТУ «ЛЭТИ». 2026. Т. 19, № 6. С. 63–71. doi: 10.32603/2071-8985-2026-19-6-63-71.

**Конфликт интересов.** Авторы заявляют об отсутствии конфликта интересов.

Original article

## Prediction of Recurrence of Diffuse Toxic Goiter Based on Clinical Data and Comparison of Machine Learning Methods and Statistical Models

E. M. Dariy<sup>1</sup>✉, A. A. Meldo<sup>1</sup>, S. V. Dora<sup>2</sup>, Yu. Sh. Khalimov<sup>2</sup>

<sup>1</sup> Saint Petersburg Electrotechnical University, Saint Petersburg, Russia

<sup>2</sup> Pavlov First Saint Petersburg State Medical University, Saint Petersburg, Russia

✉ [evgeniadaryi@gmail.com](mailto:evgeniadaryi@gmail.com)

**Abstract.** This paper analyzes models for predicting clinical events using the example of recurrence of diffuse toxic goiter (DTG), employing machine learning methods and statistical approaches. The study includes data from 185 patients. To identify the most significant predictors of DTG recurrence, random forest and extreme gradient boosting algorithms were applied, demonstrating higher predictive performance (AUC 0.82–0.83) compared to traditional statistical models. The results of this study open new prospects for the development of personalized medicine, particularly in approaches to the treatment of patients with DTG.

**Keywords:** diffuse toxic goiter, machine learning, clinical event prediction, recurrence prediction, thyrotoxicosis, risk factors

**For citation:** Prediction of Recurrence of Diffuse Toxic Goiter Based on Clinical Data and Comparison of Machine Learning Methods and Statistical Models / E. M. Dariy, A. A. Meldo, S. V. Dora, Yu. Sh. Khalimov // LETI Transactions on Electrical Engineering & Computer Science. 2026. Vol. 19, no. 6. P. 63–71. doi: 10.32603/2071-8985-2026-19-6-63-71.

---

**Conflict of interests.** The authors declare no conflicts of interests.

**Введение.** Диффузный токсический зоб (болезнь Грейвса, базедова) – заболевание щитовидной железы, основным проявлением которого служит специфическая триада симптомов: зоб (увеличение объема щитовидной железы более 18 мл у женщин и 23 мл у мужчин), тахикардия (учащенное сердцебиение) и эндокринная офтальмопатия (экзофтальм, «пучеглазие»). Развитие этого заболевания связано с выработкой иммунной системой специфических антител к клеткам щитовидной железы. Эти антитела стимулируют ткань щитовидной железы к росту и повышенной выработке гормонов щитовидной железы (тироксина и трийодтиронина), а в жировой клетчатке вокруг глазных яблок вызывают отек. Повышенное количество гормонов щитовидной железы носит название гипертиреоз, а если оно приводит к появлению специфических симптомов – тиреотоксикоз. Лечение диффузного токсического зоба многообразно, однако основная его цель – эффективное устранение тиреотоксикоза [1]. Несмотря на существующие методы лечения, ДТЗ склонен к повторному развитию [2]. Знание о взаимосвязях между многочисленными клиническими параметрами и риском рецидива может помочь в определении персонализированной лечебной тактики, улучшении качества медицинской помощи. В настоящее время в клинических рекомендациях, утвержденных Минздравом РФ [<https://cr.minzdrav.gov.ru/>], нет сведений о факторах, влияющих на течение заболевания. В научной литературе в решении задачи прогноза рецидива ДТЗ приведены примеры применения медицинской статистики (методы Шапиро–Уилка, Уилкоксона, Спирмена и т. п.) [3], [4].

В последние годы инструменты машинного обучения (Machine Learning, ML) находят все более широкое применение в медицинских исследованиях, в том числе для диагностики, прогнозирования и поддержки принятия клинических решений [5], [6]. Для решения этих задач используются различные алгоритмы ML, включая случайный лес (Random Forest), градиентный бустинг (Gradient Boosting) и нейронные сети (Artificial Neural Networks) [5], [7], [8]. Эти методы позволяют анализировать сложные многомерные данные, выяв-

лять значимые клинические предикторы и создавать прогностические модели с высокой точностью [9]. Статья посвящена применению методов машинного обучения для прогнозирования рецидива диффузного токсического зоба с целью оптимизации лечебной тактики.

**Материалы и методы.** Исходный набор данных содержал информацию о 185 пациентах с установленным диагнозом ДТЗ и включал 86 клинико-анамнестических параметров, содержащих информацию о демографических характеристиках, лабораторных показателях, данных ультразвукового исследования, особенностях медикаментозной терапии и течения заболевания. Целевой переменной служило наличие рецидива заболевания.

На этапе подготовки данных был выполнен строгий отбор признаков: исключены переменные с более 30 % пропущенных значений, что позволило сформировать репрезентативный набор из 22 наиболее информативных параметров. В окончательный набор вошли демографические показатели (год рождения, пол), характеристики заболевания (год начала, возраст манифестации, период начала), анамнестические данные (наследственность, курение), параметры щитовидной железы по данным УЗИ (объем, наличие и размер узлов), гормональные показатели в динамике (тиреотропный гормон (ТТГ), тироксин свободный (свободный Т4), трийодтиронин свободный (свободный Т3), их соотношение), наличие осложнений (экзофтальм, тиреотоксическая кардиомиопатия) и тип проводимой лекарственной терапии.

Для обеспечения качества данных выполнено заполнение пропущенных значений с использованием средних величин для нормально распределенных признаков и медианы для признаков с асимметричным распределением, пропуски категориальных переменных заполнялись модой. Для корректной работы алгоритмов машинного обучения и повышения их предсказательной способности исходные данные были предварительно нормализованы.

Для анализа применялись как традиционные статистические методы (логистическая регрессия, линейный и квадратичный дискриминантный анализ, наивный байесовский классификатор),

так и алгоритмы машинного обучения (метод  $k$ -ближайших соседей, машина опорных векторов, дерево решений, случайный лес, экстремальный градиентный бустинг, многослойный перцептрон). При выборе данного набора моделей авторы ориентировались на существующие памятки и методические рекомендации по выбору алгоритмов машинного обучения в зависимости от объема данных, их структуры и требований к интерпретируемости и качеству прогнозирования [10], [11].

Значимость признаков определена на основе случайного леса (рис. 1, *a*) и экстремального градиентного бустинга (рис. 1, *b*). Этот выбор обусловлен двумя критериями: высокой предсказательной силой и встроенной функциональностью оценки важности признаков, которая позволяет интерпретировать вклад каждого признака в итоговый прогноз.

Случайный лес (Random Forest) относится к числу широко применяемых ансамблевых алгоритмов машинного обучения и основан на комбинировании большого числа решающих деревьев. Данный метод реализует подход бэггинга (bootstrap aggregating), предполагающий многократное формирование обучающих подвыборок случайной выборкой объектов с возвращением. Основная идея алгоритма заключается в построении совокупности моделей одного типа и последующем объединении их предсказаний с целью повышения обобщающей способности и устойчивости результатов.

Архитектура случайного леса предполагает независимое обучение каждого дерева решений. Для построения отдельного дерева используется собственная обучающая выборка, сформированная методом бутстрэпа, в результате чего одни объекты могут присутствовать в выборке несколько раз, тогда как другие могут быть полностью исключены. Кроме того, на каждом шаге разбиения узла рассматривается не полный набор признаков, а лишь его случайно выбранное подмножество. Такая двойная случайность – по объектам и по признакам – существенно снижает корреляцию между отдельными деревьями ансамбля, что, в свою очередь, уменьшает риск переобучения и повышает устойчивость модели к шуму в данных.

В задачах классификации итоговое решение в случайном лесе формируется на основе агрегирования предсказаний всех деревьев. Каждое дерево «голосует» за определенный класс, а окончательный результат определяется как класс, получивший наибольшее количество голосов. Данный

механизм коллективного принятия решений позволяет компенсировать ошибки отдельных деревьев и обеспечивает более стабильные и точные прогнозы по сравнению с одиночными моделями.

Алгоритм XGBoost (экстремальный градиентный бустинг) представляет собой высокоэффективную реализацию градиентного бустинга над деревьями решений. В отличие от методов бэггинга, XGBoost использует последовательный подход к построению ансамбля, при котором каждое последующее дерево направлено на исправление ошибок, допущенных предыдущими моделями. Таким образом, процесс обучения носит итеративный характер, а итоговая модель формируется как аддитивная композиция базовых обучаемых алгоритмов.

Ключевая особенность XGBoost – использование градиентных методов оптимизации для минимизации заданной функции потерь. На каждом шаге обучения новые деревья строятся с учетом градиента и, при необходимости, гессии функции ошибки, что позволяет более точно аппроксимировать сложные зависимости в данных. Дополнительно в XGBoost реализованы механизмы регуляризации, ограничивающие сложность модели и способствующие снижению переобучения, что делает данный алгоритм особенно эффективным при работе с высокоразмерными и неоднородными наборами данных.

Оценка значимости признаков в ансамблевых моделях – также важный этап анализа. В случайном лесе важность признака определяется на основе его вклада в улучшение качества разбиений в узлах деревьев. Каждый раз, когда признак используется для разделения данных, фиксируется снижение выбранного критерия неоднородности (например, индекса Джини или энтропии). Итоговая значимость признака рассчитывается как сумма таких улучшений по всем узлам и всем деревьям ансамбля, после чего выполняется усреднение. В результате более информативные признаки, обеспечивающие существенное повышение качества разбиений, получают более высокие значения важности.

В XGBoost оценка важности признаков реализована с применением нескольких дополняющих друг друга метрик. Наиболее информативна метрика gain (прирост), отражающая вклад признака в уменьшение функции потерь при выполнении разбиений. Чем больше суммарный прирост качества, тем более значимым считается признак. Дополнительно используется метрика

weight (вес), показывающая частоту использования признака в узлах всех деревьев ансамбля, а также метрика cover (покрытие), характеризующая объем данных, затрагиваемых разбиениями по данному признаку. Совместный анализ этих показателей позволяет получить более детальное представление о роли признаков в процессе формирования прогноза.

Результаты оценки значимости признаков были получены на основе серии из 120 вычислительных экспериментов, после чего выполнено усреднение значений для повышения устойчивости выводов. Важности признаков обоих методов отображены на рис. 1: по горизонтали – средняя важность признаков, а по вертикали – признаки в порядке уменьшения их значимости для моделей.

В качестве основной метрики для оценки предсказательной способности построенных моделей использовалась площадь под ROC-кривой –

AUC (Area Under the Receiver Operating Characteristic Curve), которая широко применяется в задачах бинарной классификации, в том числе в медицинской диагностике [10]. ROC-кривая представляет собой графическое отображение зависимости между чувствительностью и специфичностью модели при различных значениях порога классификации. В отличие от таких показателей, как точность и полнота, AUC позволяет оценивать качество модели независимо от выбранного порога принятия решения, что делает ее удобным инструментом для сравнения различных алгоритмов машинного обучения [12]. Кроме того, AUC обладает статистической непротиворечивостью и более избирательна по сравнению с точностью [13].

Дополнительно для оценки устойчивости и воспроизводимости результатов использовалась метрика std, представляющая собой стандартное отклонение значений AUC, полученных в серии

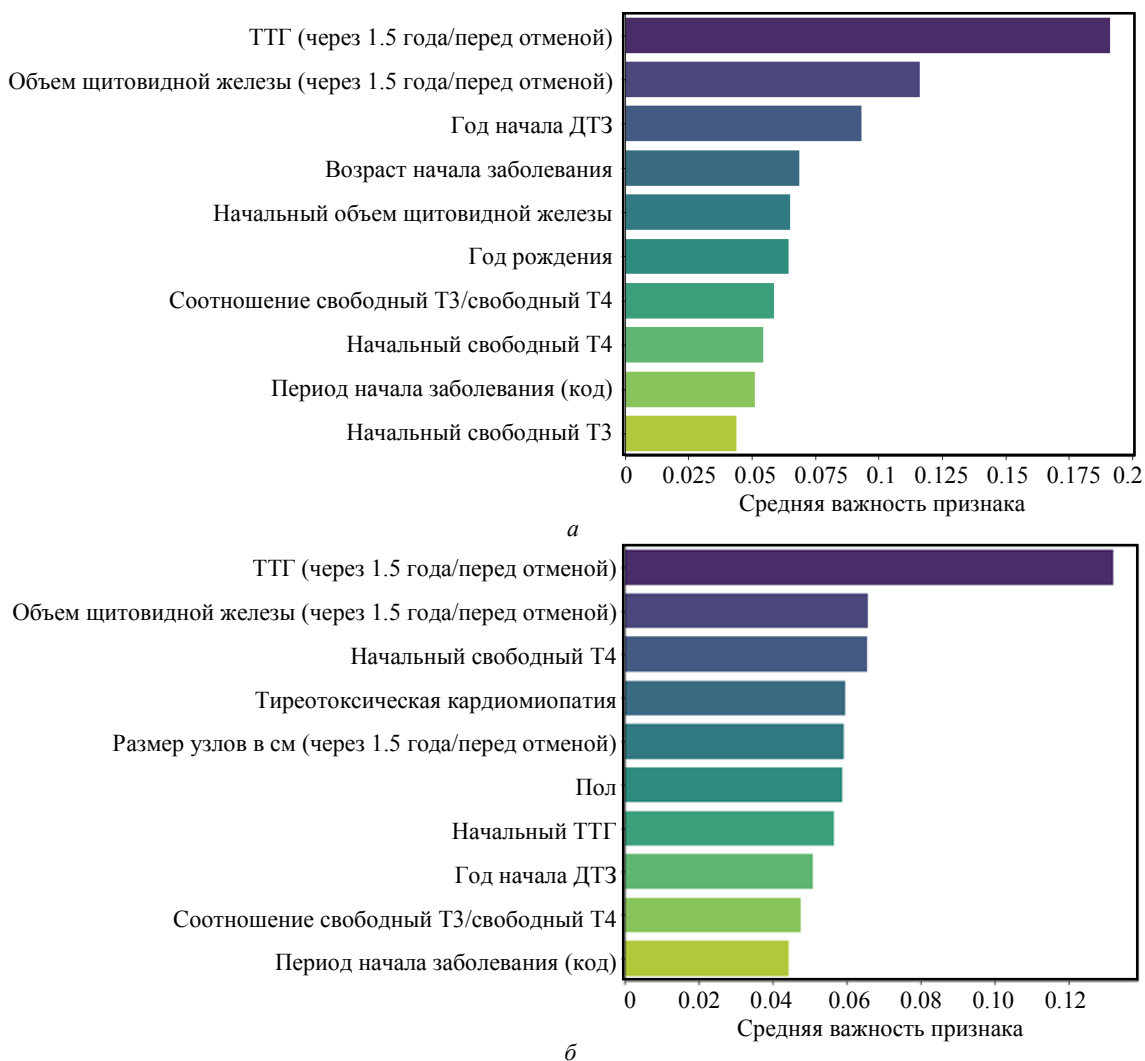


Рис. 1. Наиболее важные признаки: а – случайный лес; б – экстремальный градиентный бустинг  
Fig. 1. Most important features: a – random forest; б – extreme gradient boosting

экспериментов. Данная метрика позволяет количественно оценить степень вариативности результатов и чувствительность модели к изменениям обучающих данных. Низкие значения стандартного отклонения свидетельствуют о большей стабильности алгоритма и надежности полученных оценок качества [14].

Алгоритмы ML включают в себя ряд гиперпараметров, значения которых задаются до начала процесса обучения. В отличие от параметров, узнаваемых непосредственно в ходе обучения, гиперпараметры определяют структуру алгоритма ML, особенности процесса обучения и оказывают влияние на его обобщающую способность. В связи с ограниченным размером исходного набора данных в данной статье использовались значения гиперпараметров, заданные по умолчанию в применяемых программных библиотеках, т. е. формируемые на основе рекомендаций или эмпирических оценок [15].

Для обеспечения статистической надежности и минимизации дисперсии оценки прогностической способности моделей была применена процедура повторной стратифицированной  $k$ -блочной кросс-валидации. Учитывая ограниченный объем исходной выборки, методология включала 30 независимых повторений случайного разбиения с последующей 4-блочной валидацией, что в совокупности составило 120 оценок для каждой модели. При каждом повторении осуществлялось стратифицированное разделение на обучающую (75 % данных) и тестовую (25 % данных) подвыборки с сохранением распределения целевого признака (наличие рецидива), что обеспечивало

репрезентативность получаемых оценок. Такое соотношение 75/25 % при 4-блочном разделении позволило максимизировать объем обучающих данных при сохранении достаточного размера тестовой выборки для валидации. Для каждого разбиения выполнялось полное переобучение моделей, а итоговые метрики качества вычислялись как среднее арифметическое по всем 120 итерациям, что позволило получить надежные оценки эффективности алгоритмов, устойчивые к случайным колебаниям, связанным с малым объемом исходных данных. Арифметические средние полученных результатов для каждой модели представлены в таблице.

**Обсуждение результатов.** Проведенное исследование позволило выявить ключевые факторы, влияющие на риск рецидива ДТЗ, а также оценить сравнительную эффективность различных прогностических подходов. Наибольшую предсказательную способность продемонстрировали ансамблевые методы машинного обучения, что подтверждается результатами экспериментов (рис. 2, табл.): средние значения AUC составили 0.829 для экстремального градиентного бустинга и 0.821 для случайного леса. Остальные алгоритмы машинного обучения (метод опорных векторов, дерево решений, многослойный перцептрон и метод  $k$ -ближайших соседей) оказались менее эффективными: их показатели AUC находились в диапазоне от 0.648 до 0.738. Традиционные статистические методы (логистическая регрессия, линейный и квадратичный дискриминантный анализ, наивный байесовский классификатор) в целом продемонстрировали более низкие результаты со средним AUC 0.673 (где наименьший по-

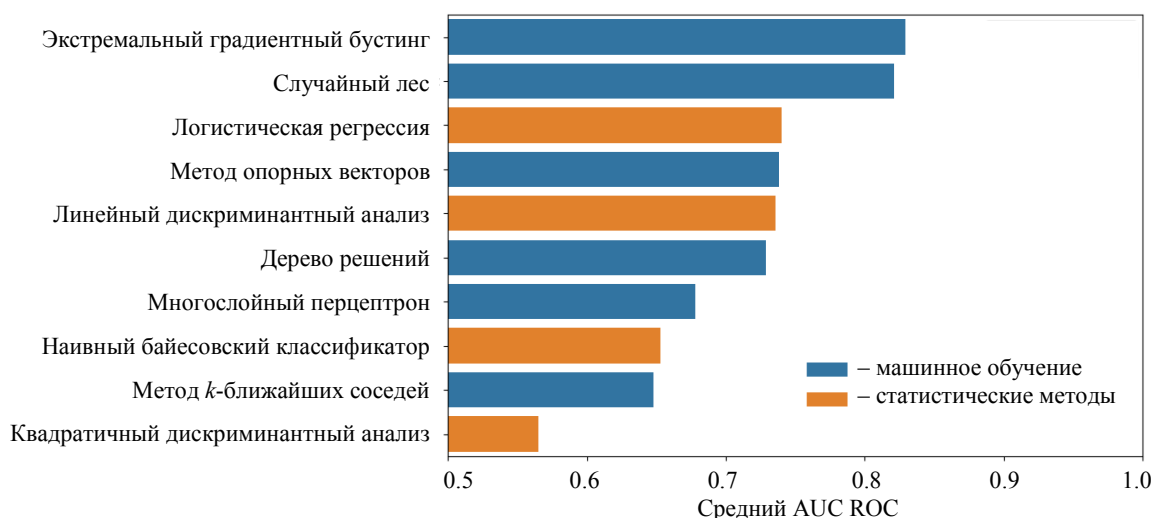


Рис. 2. Сравнение моделей по средним значениям AUC

Fig. 2. Comparison of models by mean AUC values

Средние значения AUC  
 Average AUC values

Модель	Средний AUC	std
Экстремальный градиентный бустинг	0.830	0.035
Случайный лес	0.821	0.046
Логистическая регрессия	0.740	0.051
Метод опорных векторов	0.738	0.063
Линейный дискриминантный анализ	0.736	0.026
Дерево решений	0.729	0.082
Многослойный перцептрон	0.678	0.054
Наивный байесовский классификатор	0.653	0.039
Метод $k$ -ближайших соседей	0.648	0.057
Квадратичный дискриминантный анализ	0.565	0.073
Статистические методы: средний AUC = 0.673		
Методы машинного обучения: средний AUC = 0.741		

казатель 0.565 зафиксирован у квадратичного дискриминантного анализа), тогда как среднее значение AUC для всей группы методов машинного обучения составило 0.741. Для подтверждения статистической значимости различий между группами методов был применен  $t$ -критерий Стьюдента (Student's  $t$ -test), который позволяет сравнивать средние значения показателей при двух выборках с нормальным распределением [16]. Результаты расчета продемонстрировали, что различия между традиционными статистическими методами и алгоритмами машинного обучения статистически значимы ( $p < 0.001$ , где  $p$  – достигнутый уровень статистической значимости, от англ.  $p$ -value), что подтверждает более высокую предсказательную способность ансамблевых моделей. Преимущество ансамблевых методов обусловлено их способностью автоматически учитывать нелинейные зависимости и взаимодействия между признаками, что ограничено реализуется в традиционных статистических моделях и особенно важно при анализе сложных медицинских данных [17], [18].

Следует отметить, что при интерпретации важности признаков в ансамблевых алгоритмах необходимо учитывать методологические особенности: непрерывные переменные (уровень ТТГ, объем щитовидной железы и т. п.) могут получать завышенные оценки важности по сравнению с бинарными признаками (например, наличие наследственности, экзофтальма) в силу большего количества возможных точек разбиения в деревьях решений.

Несмотря на это ограничение, анализ важности признаков выявил устойчивые предикторы рецидива ДТЗ. У модели случайного леса наиболее значимыми факторами оказались: уровень ТТГ через 1.5 г. после начала лечения/перед отменой терапии, объем щитовидной железы через 1.5 г. после начала лечения/перед отменой тера-

пии и календарный год с начала заболевания (рис. 1, а). У модели XGBoost также лидирующие позиции заняли уровень ТТГ и объем щитовидной железы через 1.5 г. после начала лечения/перед отменой, однако существенную роль приобрели начальный уровень свободного Т4, наличие тиреотоксической кардиомиопатии, размер узлов через 1.5 г. после начала лечения/перед отменой и пол (рис. 1, б). Можно заметить, что обе модели согласованно выделили важность ТТГ и объема щитовидной железы через 1.5 г. лечения, что подчеркивает значение контроля этих показателей в процессе терапии.

Важно отметить, что выявленные моделью ассоциации представляют собой лишь статистические закономерности и не могут трактоваться как прямые причинно-следственные связи. Тем не менее, они предоставляют ценную информацию для формирования гипотез о факторах, влияющих на течение ДТЗ, и открывают новые возможности для дальнейших исследований.

**Заключение.** Проведенное исследование демонстрирует перспективность применения методов машинного обучения для прогнозирования клинических событий в ходе заболеваний. Ансамблевые алгоритмы показали существенно более высокую предсказательную способность по сравнению с традиционными статистическими методами, что обусловлено их способностью выявлять сложные нелинейные взаимосвязи между многочисленными клиническими параметрами. Персонализированная оценка риска рецидива открывает новые возможности для оптимизации лечебной тактики и улучшения результата терапии пациентов с ДТЗ.

Полученные результаты обосновывают целесообразность интеграции в клиническую практику для разработки систем поддержки врачебных решений программных инструментов на основе методов машинного обучения. Перспективными

направлениями дальнейших исследований являются валидация моделей на новых наборах клинических данных, изучение влияния установленных важных признаков на течение ДТЗ, а также

разработка персонализированных алгоритмов терапии на основе индивидуального прогноза риска рецидива.

### Список литературы

1. Диффузный токсический зоб (Базедова болезнь). Официальный сайт ФГБУ «НМИЦ ТПМ» Минздрава России. Клиника высоких медицинских технологий им. Н. И. Пирогова СПбГУ. URL: <https://www.gosmed.ru/lechebnaya-deyatelnost/spravochnik-zabolevaniy/endokrino-logiya-bolezny/diffuznyy-toksicheskiy-zob/> (дата обращения: 18.01.2026).
2. Почему происходит рецидив тиреотоксикоза и как его избежать. Северо-западный центр эндокринологии и эндокринной хирургии. URL: [https://endoinfo.ru/theory\\_pacients/zabolevaniya-shchitovidnoy-zhelezy/pochemu-proiskhodit-retsivid-tireotoksikoza-i-kak-ego-izbezhat.html](https://endoinfo.ru/theory_pacients/zabolevaniya-shchitovidnoy-zhelezy/pochemu-proiskhodit-retsivid-tireotoksikoza-i-kak-ego-izbezhat.html) (дата обращения: 18.01.2026).
3. Макаров И. В., Галкин Р. А., Андреев М. М. Отдаленные результаты и их прогнозирование в хирургическом лечении диффузного токсического зоба // Эндокринная хирургия. 2013. № 2. С. 45–51.
4. Прогностическая модель вероятности ремиссии заболевания у больных диффузным токсическим зобом / А. Р. Волкова, С. В. Дора, Г. Г. Алламова, И. М. Абрамова, С. М. Панчоян // Endocrinology. 2020. Т. 19, № 2. С. 40–44. doi: 10.31550/1727-2378-2020-19-2-40-44.
5. Habehh H., Gohel S. Machine learning in healthcare // Curr. Genomics. 2021. Vol. 22, no. 4. P. 291–300. doi: 10.2174/1389202922666210705124359. URL: <https://pmc.ncbi.nlm.nih.gov/articles/PMC8822225/> (дата обращения: 18.01.2026).
6. Automated machine learning in medical research: A systematic literature mapping study / G. A. de Castro, L. G. Barioto, H. Cao Yu, R. M. Silva, H. M. Caseli, J. A. Machado-Neto, R. Cerri, A. Villavicencio, T. A. Almeida // Artificial Intelligence in Medicine. 2026. Vol. 171. P. 103302. doi: 10.1016/j.artmed.2025.103302.
7. Yıldız A. Ya., Kalaycı A. Gradient boosting decision trees on medical diagnosis over tabular data // 2025 IEEE Intern. Conf. on AI and Data Analytics (ICAD). Medford, Massachusetts, USA: IEEE, 2025. doi: 10.1109/ICAD65464.2025.11114069.
8. Методы машинного обучения как инструмент диагностических и прогностических исследований при ишемической болезни сердца / Б. И. Гельцер, М. М. Циванюк, К. И. Шахгельдян, В. Ю. Рублев // Российский кардиологический журн. 2020. Т. 25, № 12. С. 164–171. doi: 10.15829/1560-4071-2020-3999.
9. A Gradient boosting machine learning model for predicting early mortality in the emergency department triage: Devising a nine-point triage score / M. Klug, Y. Barash, S. Bechler, Y. S. Resheff, T. Tron, A. Ironi, Sh. Soffer, E. Zimlichman, E. Klang // J. Gen. Intern. Med. 2019. Vol. 35, no. 1. P. 220–227. doi: 10.1007/s11606-019-05512-7.
10. Domingos P. A few useful things to know about machine learning // Communications of the ACM. 2012. Vol. 55, no. 10. P. 78–87. doi: 10.1145/2347736.2347755.
11. Machine learning algorithm cheat sheet for azure machine learning designer. URL: <https://learn.microsoft.com/en-us/azure/machine-learning/algorithm-cheat-sheet?view=azureml-api-1> (дата обращения: 18.01.2026).
12. Huang J., Ling C. X. Using AUC and accuracy in evaluating learning algorithms // IEEE Trans. on Knowledge and Data Engin. 2005. Vol. 17, no. 3. P. 299–310. doi: 10.1109/TKDE.2005.50.
13. Ling Ch. X., Huang J., Zhang H. AUC: a statistically consistent and more discriminating measure than accuracy // Proc. of the 18<sup>th</sup> Intern. Joint Conf. on Artificial Intelligence (IJCAI 2003). Acapulco, Mexico: IJCAI, Inc., 2003. P. 519–526.
14. Kuhn M., Johnson K. Applied predictive modeling. New York: Springer, 2013. 600 p. doi: 10.1007/978-1-4614-6849-3.
15. Scikit-learn: Machine learning in Python / F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, É. Duchesnay // J. of Machine Learning Res. 2011. Vol. 12. P. 2825–2830.
16. Ruxton G. D. The unequal variance t-test is an underused alternative to Student's t-test and the Mann-Whitney U test // Behavioral Ecol. 2006. Vol. 17, no. 4. P. 688–690. doi: 10.1093/beheco/ark016.
17. Breiman L. Random Forests // Machine learning. 2001. Vol. 45, no. 1. P. 5–32. doi: 10.1023/A:1010933404324.
18. Hastie T., Tibshirani R., Friedman J. The Elements of statistical learning: Data mining, inference, and prediction. 2<sup>nd</sup> ed. New York: Springer, 2009. 745 p. doi: 10.1007/978-0-387-84858-7.

### Информация об авторах

**Дарий Евгения Марьяновна** – студентка бакалавриата кафедры вычислительной техники СПбГЭТУ «ЛЭТИ».  
E-mail: [evgeniadaryi@gmail.com](mailto:evgeniadaryi@gmail.com)

**Мелдо Анна Александровна** – д-р мед. наук, профессор кафедры вычислительной техники СПбГЭТУ «ЛЭТИ».  
E-mail: [anna.meldo@yandex.ru](mailto:anna.meldo@yandex.ru)  
<https://orcid.org/0000-0002-4906-9901>

**Дора Светлана Владимировна** – д-р мед. наук, профессор кафедры терапии факультетской с курсом эндокринологии и кардиологии с клиникой имени академика Г. Ф. Ланга, зам. главного врача клиники по медицинской части ФГБОУ ВО ПСПбГМУ им. И. П. Павлова Минздрава России, ул. Льва Толстого, 6–8, Санкт-Петербург, 197022, Россия.

E-mail: doras2001@mail.ru

<https://orcid.org/0000-0002-8249-6075>

**Халимов Юрий Шавкатович** – д-р мед. наук, профессор, зав. кафедрой терапии факультетской с курсом эндокринологии и кардиологии с клиникой имени академика Г. Ф. Ланга, проректор по лечебной работе ФГБОУ ВО ПСПбГМУ им. И. П. Павлова Минздрава России. Заслуженный врач РФ, Главный эндокринолог Комитета по здравоохранению Правительства С.-Петербурга. ул. Льва Толстого, 6–8, Санкт-Петербург, 197022, Россия.

E-mail: yushkha@mail.ru

<https://orcid.org/0000-0002-7755-7275>

#### **Вклад авторов:**

Дарий Е. М. – изучение литературы, предобработка данных, обучение моделей машинного обучения, подготовка текста.

Мелдо А. А. – руководство научной работой, разработка концепции, редактирование текста.

Дора С. В. – сбор материала, формирование медицинской базы данных, предоставление материала для исследования.

Халимов Ю. Ш. – организация сбора материала, формулировка целевых направлений исследования.

#### **References**

1. Diffuznyj toksicheskiy zob (Bazedova bolezn'). Oficial'nyj sayt FGBU «NMIC TPM» Minzdrava Rossii. Klinika vysokih medicinskih tehnologij im. N. I. Pirogova SPbGU. URL: <https://www.gosmed.ru/lechebnaya-deyatelnost/spravochnik-zabolevaniy/endokrino-logiya-bolezny/diffuznyy-toksicheskiy-zob/> (data obrashhenija: 18.01.2026). (In Russ.)
2. Pochemu proishodit recidiv tireotoksikoza i kak ego izbezhat'. Severo-zapadnyj centr jendokrinologii i jendokrinnoj hirurgii. URL: [https://endoinfo.ru/theory\\_pacients/zabolevaniya-shchitovidnoy-zhelezy/pochemu-proiskhodit-retsiv-tireotoksikoza-i-kak-ego-izbezhat.html](https://endoinfo.ru/theory_pacients/zabolevaniya-shchitovidnoy-zhelezy/pochemu-proiskhodit-retsiv-tireotoksikoza-i-kak-ego-izbezhat.html) (data obrashhenija: 18.01.2026). (In Russ.)
3. Makarov I. V., Galkin R. A., Andreev M. M. Otdalennye rezultaty i ih prognozirovanie v hirurgicheskom lechenii diffuznogo toksicheskogo zoba // Jendokrinnaia hirurgija. 2013. № 2. S. 45–51. (In Russ.)
4. Prognosticheskaja model' verojatnosti remissii zabolevaniya u bol'nyh diffuznym toksicheskim zobom / A. R. Volkova, S. V. Dora, G. G. Allamova, I. M. Abramova, S. M. Panchojan // Endocrinology. 2020. T. 19, № 2. S. 40–44. doi: 10.31550/1727-2378-2020-19-2-40-44. (In Russ.)
5. Habehh H., Gohel S. Machine Learning in Healthcare // Curr. Genomics. 2021. Vol. 22, no. 4. P. 291–300. doi: 10.2174/1389202922666210705124359. URL: <https://pmc.ncbi.nlm.nih.gov/articles/PMC8822225/> (data obrashhenija: 18.01.2026).
6. Automated machine learning in medical research: A systematic literature mapping study / G. A. de Castro, L. G. Barioto, H. Cao Yu, R. M. Silva, H. M. Caseli, J. A. Machado-Neto, R. Cerri, A. Villavicencio, T. A. Almeida // Artificial Intelligence in Medicine. 2026. Vol. 171. P. 103302. doi: 10.1016/j.artmed.2025.103302.
7. Yildiz A. Ya., Kalayci A. Gradient boosting decision trees on medical diagnosis over tabular data // 2025 IEEE Intern. Conf. on AI and Data Analytics (ICAD). Medford, Massachusetts, USA: IEEE, 2025. doi: 10.1109/ICAD65464.2025.11114069.
8. Metody mashinnogo obuchenija kak instrument diagnosticheskikh i prognosticheskikh issledovanij pri ishemichejskoj bolezni serdca / B. I. Gel'cer, M. M. Civanjuk, K. I. Shahgel'djan, V. Ju. Rublev // Rossijskij kardiologicheskij zhurn. 2020. T. 25, № 12. S. 164–171. doi: 10.15829/1560-4071-2020-3999. (In Russ.)
9. A gradient boosting machine learning model for predicting early mortality in the emergency department triage: Devising a nine-point triage score / M. Klug, Y. Barash, S. Bechler, Y. S. Resheff, T. Tron, A. Ironi, Sh. Soffer, E. Zimlichman, E. Klang // J. Gen. Intern. Med. 2019. Vol. 35, no. 1. P. 220–227. doi: 10.1007/s11606-019-05512-7.
10. Domingos P. A few useful things to know about machine learning // Communications of the ACM. 2012. Vol. 55, no. 10. P. 78–87. doi: 10.1145/2347736.2347755.
11. Machine learning algorithm cheat sheet for azure machine learning designer. URL: <https://learn.microsoft.com/en-us/azure/machine-learning/algorithm-cheat-sheet?view=azureml-api-1> (data obrashhenija: 18.01.2026).
12. Huang J., Ling C. X. Using AUC and accuracy in evaluating learning algorithms // IEEE Trans. on Knowledge and Data Engin. 2005. Vol. 17, no. 3. P. 299–310. doi: 10.1109/TKDE.2005.50.
13. Ling Ch. X., Huang J., Zhang H. AUC: a statistically consistent and more discriminating measure than accuracy // Proc. of the 18<sup>th</sup> Intern. Joint Conf. on Artificial

Intelligence (IJCAI 2003). Acapulco, Mexico: IJCAI, Inc., 2003. P. 519–526.

14. Kuhn M., Johnson K. Applied predictive modeling. New York: Springer, 2013. 600 p. doi: 10.1007/978-1-4614-6849-3.

15. Scikit-learn: Machine learning in Python / F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, É. Duchesnay // J. of Machine Learning Res. 2011. Vol. 12. P. 2825–2830.

16. Ruxton G. D. The unequal variance t-test is an underused alternative to Student's t-test and the Mann-Whitney U test // Behavioral Ecol. 2006. Vol. 17, no. 4. P. 688–690. doi: 10.1093/beheco/ark016.

17. Breiman L. Random forests // Machine Learning. 2001. Vol. 45, no. 1. P. 5–32. doi: 10.1023/A:1010933404324.

18. Hastie T., Tibshirani R., Friedman J. The Elements of statistical learning: Data mining, inference, and prediction. 2<sup>nd</sup> ed. New York: Springer, 2009. 745 p. doi: 10.1007/978-0-387-84858-7.

---

#### Information about the authors

**Evgeniya M. Dariy** – Bachelor's student of the Department of Computer Technology, Saint Petersburg Electrotechnical University.

E-mail: [evgeniadaryi@gmail.com](mailto:evgeniadaryi@gmail.com)

**Anna A. Meldo** – Dr Sci. (Med.), Professor of the Department of Computer Engineering, Saint Petersburg Electrotechnical University.

E-mail: [anna.meldo@yandex.ru](mailto:anna.meldo@yandex.ru)

<https://orcid.org/0000-0002-4906-9901>

**Svetlana V. Dora** – Dr Sci. (Med.), Professor, Department of Faculty Therapy with the Course of Endocrinology and Cardiology and the Clinic named after Academician G. F. Lang, Deputy Chief Medical Officer of the Clinic, Pavlov First Saint Petersburg State Medical University. 6–8 Lev Tolstoy St., Saint Petersburg, 197022, Russia.

E-mail: [doras2001@mail.ru](mailto:doras2001@mail.ru)

<https://orcid.org/0000-0002-8249-6075>

**Yuriy Sh. Khalimov** – Dr Sci. (Med.), Professor, Head of the Department of Faculty Therapy with the Course of Endocrinology and Cardiology and the Clinic named after Academician G. F. Lang, Vice-Rector for Clinical Work, Pavlov First Saint Petersburg State Medical University. Honored Doctor of the Russian Federation, Chief Endocrinologist of the Health Committee of the Government of St. Petersburg. 6–8 Lev Tolstoy St., Saint Petersburg, 197022, Russia.

E-mail: [yushkha@mail.ru](mailto:yushkha@mail.ru)

<https://orcid.org/0000-0002-7755-7275>

#### Author contribution statement:

Dariy E. M. – literature review, data preprocessing, training of machine learning models, text preparation.

Meldo A. A. – supervision of scientific work, conceptualization, text editing.

Dora S. V. – collection of material, formation of medical database, provision of material for research.

Khalimov Yu. Sh. – organization of material collection, formulation of target areas of research.

Статья поступила в редакцию 10.02.2026; принята к публикации после рецензирования 20.04.2026; опубликована онлайн 29.06.2026.

Submitted 10.02.2026; accepted 20.04.2026; published online 29.06.2026.

---