

INTERNATIONAL SCIENCE REVIEWS



№1 (5) 2024

Natural Sciences and
Technologies series





INTERNATIONAL SCIENCE REVIEWS

Natural Sciences and Technologies series

Has been published since 2020

№1 (5) 2024

Astana

INTERNATIONAL SCIENCE REVIEWS. NATURAL SCIENCES AND TECHNOLOGIES SERIES ЖУРНАЛЫНЫҢ РЕДАКЦИЯСЫ

БАС РЕДАКТОР

Қалимолдаев Мақсат Нұрадилович, техникалық ғылымдар докторы, ҚР ҰҒА академигі, профессор, ҚР ҒЖБМ ҒК «Ақпараттық және есептеу технологиялары институты» бас директорының кеңесшісі, бас ғылыми қызметкері (Қазақстан)

БАС РЕДАКТОРДЫҢ ОРЫНБАСАРЫ

Мырзағалиева Анар Базаровна, биология ғылымдарының докторы, профессор, бірінші вице-президент, Астана халықаралық университеті (Қазақстан);

РЕДАКТОРЛАР:

- **Сейткан Айнура Сейтканқызы**, техника ғылымдарының кандидаты, PhD, жаратылыстану ғылымдары жоғары мектебінің деканы, Астана халықаралық университеті (Қазақстан);

- **Муканова Асель Сериковна**, PhD, Ақпараттық технологиялар және инженерия жоғары мектебінің деканы, Астана халықаралық университеті (Қазақстан);

- **Абдилдаева Асель Асылбековна**, PhD, қауымдастырылған профессор, Әл-Фараби атындағы ҚазҰУ (Қазақстан);

- **Хлахула Иржи** PhD, профессор, Познаньдағы Адам Мицкевич атындағы университет (Польша);

- **Редферн Саймон А.Т.**, PhD, профессор, Наньян технологиялық университеті (Сингапур);

- **Сяолей Фенг**, PhD, Наньян технологиялық университеті (Сингапур);

- **Шуджаул Мулк Хан**, PhD, профессор, Каид-және-Азам университеті (Пакистан);

- **Базарнова Наталья Григорьевна**, химия ғылымдарының докторы, профессор, Химия және химиялық-фармацевтикалық технологиялар институты (Ресей);

- **Черёмушкина Вера Алексеевна**, биология ғылымдарының докторы, профессор, РҒА СБ Орталық Сібір ботаникалық бағы (Ресей);

- **Тасболатұлы Нұрболат**, PhD, Ақпараттық технологиялар және инженерия жоғары мектебі деканының орынбасары, Астана халықаралық университеті (Қазақстан);

- **Байшоланов Сакен Советович**, география ғылымдарының кандидаты, доцент, Астана халықаралық университеті (Қазақстан);

- **Нуркенов Серик Амангельдинович**, PhD, қауымдастырылған профессор, Астана халықаралық университеті (Қазақстан).

**РЕДАКЦИЯ ЖУРНАЛА INTERNATIONAL SCIENCE REVIEWS.
NATURAL SCIENCES AND TECHNOLOGIES SERIES**

ГЛАВНЫЙ РЕДАКТОР

Калимолдаев Максат Нурадилович, доктор технических наук, академик НАН РК, профессор, ГНС, советник генерального директора Института информационных и вычислительных технологий КН МНВО РК (*Казахстан*)

ЗАМЕСТИТЕЛЬ ГЛАВНОГО РЕДАКТОРА

Мырзагалиева Анар Базаровна, доктор биологических наук, профессор, первый вице-президент, Международный университет Астана (*Казахстан*)

РЕДАКТОРЫ:

- **Сейткан Айнур Сейтканкызы**, кандидат технических наук, PhD, декан высшей школы естественных наук, Международный университет Астана (*Казахстан*);

- **Муканова Асель Сериковна**, PhD, декан Высшей школы информационных технологий и инженерии, Международный университет Астана (*Казахстан*);

- **Абдилдаева Асель Асылбековна**, PhD, ассоциированный профессор, Казахский национальный университет имени Аль-Фараби (*Казахстан*);

- **Хлахула Иржи** PhD, профессор, Университет имени Адама Мицкевича в Познани (*Польша*);

- **Редферн Саймон А.Т.**, PhD, профессор, Наньянский технологический университет (*Сингапур*);

- **Фенг Сяoley**, PhD, Наньянский технологический университет (*Сингапур*);

- **Шуджаул Мулк Хан**, PhD, профессор, Университет Каид-и Азама (*Пакистан*);

- **Базарнова Наталья Григорьевна**, доктор химических наук, профессор, Институт химии и химико-фармацевтических технологий (*Россия*);

- **Черёмушкина Вера Алексеевна**, доктор биологических наук, профессор, Центральный Сибирский Ботанический сад СО РАН (*Россия*);

- **Тасболатұлы Нұрболат**, PhD, заместитель декана Высшей школы информационных технологий и инженерии, Международный университет Астана (*Казахстан*);

- **Байшоланов Сакен Советович**, кандидат географических наук, доцент, Международный университет Астана (*Казахстан*);

- **Нуркенов Серик Амангельдинович**, PhD, ассоциированный профессор, Международный университет Астана (*Казахстан*);

**EDITORIAL TEAM OF THE JOURNAL INTERNATIONAL SCIENCE REVIEWS.
NATURAL SCIENCES AND TECHNOLOGIES SERIES**

CHIEF EDITOR

Maksat Kalimoldayev, Doctor of Technical Sciences, Academician of NAS RK, Professor, SRF, CEO's counselor «The Institute of Information and Computational Technologies» CS MSHE RK (Kazakhstan)

DEPUTY CHIEF EDITOR

Anar Myrzagaliyeva, Doctor of Biological Sciences, Professor, First Vice-President, Astana International University (Kazakhstan)

EDITORS:

- **Ainur Seitkan**, Candidate of Technical Sciences, PhD, Dean of the Higher School of Natural Sciences, Astana International University (Kazakhstan);
- **Assel Mukanova**, PhD, Dean of the Higher School of Information Technology and Engineering, Astana International University (Kazakhstan);
- **Assel Abdildayeva**, PhD, Associate Professor, of the Department of Artificial Intelligence and Big Data, Al-Farabi Kazakh National University (Kazakhstan);
- **Jiri Chlachula**, PhD, Dr.Hab., Full Professor, Adam Mickiewicz University, Poznań (Poland);
- **Simon A.T. Redfern**, PhD, Professor, Nanyang Technological University (Singapore);
- **Xiaolei Feng**, PhD, Nanyang Technological University (Singapore);
- **Khan Shujaul Mulk**, PhD, Professor, Quaid-i-Azam University (Pakistan);
- **Natal'ya Bazarnova**, Doctor of Chemical Sciences, Professor, Institute of Chemistry and Chemical-Pharmaceutical Technologies (Russia);
- **Vera Cheryomushkina**, Doctor of Biological Sciences, Professor, Central Siberian Botanical Garden SB RAS (Russia);
- **Nurbolat Tasbolatuly**, PhD, Deputy Dean of the Higher School of Information Technology and Engineering, Astana International University (Kazakhstan);
- **Saken Baisholanov**, Candidate of Geographical Sciences, Associate Professor, Astana International University (Kazakhstan);
- **Serik Nurkenov**, PhD, Associate Professor, Astana International University (Kazakhstan).

Editorial address: 8, Kabanbay Batyr avenue, of.316, Nur-Sultan,
Kazakhstan, 010000

Tel.: (7172) 24-18-52 (ext. 316)

E-mail: natural-sciences@aiu.kz

International Science Reviews NST - 76153

International Science Reviews

Natural Sciences and Technologies series

Owner: Astana International University

Periodicity: quarterly

Circulation: 500 copies

CONTENT

1. А.М.Касымханов, И.В.Притыкин, С.К.Китапбаев, Д.А.Костюченко, С.Е.Базаров ЕРТИС БАССЕЙНИНІҢ ХАЛЫҚАРАЛЫҚ МАҢЫЗЫ БАР БАЛЫҚ ШАРУАШЫЛЫҒЫ СУ АЙДЫҢДАРЫНДАҒЫ ТЫРАНЫҢ (ABRAMIS BRAMA LINNAEUS, 1758) БИОЛОГИЯЛЫҚ КӨРСЕТКІШТЕРІ ЖӘНЕ КӘСІПТІК МАҢЫЗЫ.....	7
2. А.Бериков, Т.Самарханов ПАТРИОТИЧЕСКОЕ ВОСПИТАНИЕ ДЕТЕЙ ШКОЛЬНОГО ВОЗРАСТА НА ПРИМЕРЕ ИСТОРИЧЕСКИХ ЛИЧНОСТЕЙ БАЯНАУЫЛА.....	17
3. Ж.Н. Елемес, Ж.М. Мукатаева ӘРТҮРЛІ ДЕНЕ БЕЛСЕНДІЛІГІМЕН АЙНАЛЫСАТЫН МЕКТЕП ОҚУШЫЛАРЫНЫҢ ФИЗИКАЛЫҚ ДАМУЫ	25
4. Т.Шалбай, Л.Кусепова СЕТЕВАЯ ВИРТУАЛИЗАЦИЯ: ТРАНСФОРМАЦИЯ СОВРЕМЕННЫХ ВОЗМОЖНОСТЕЙ ПОДКЛЮЧЕНИЯ	34
5. Ж.С.Жапарова, Л.Т.Кусепова, А.Е.Назырова ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ В СОВРЕМЕННОМ БИЗНЕСЕ: ПРИМЕНЕНИЕ, ТЕНДЕНЦИИ И ВЛИЯНИЕ НА ЭФФЕКТИВНОСТЬ РАБОТЫ ОРГАНИЗАЦИЙ	43
6. Е.Е.Садибек, Л.Т.Кусепова, А.Е.Назырова ОБЕСПЕЧЕНИЕ БЕЗОПАСНОСТИ ОБЛАЧНЫХ СЕРВИСОВ	50
7. А.Е.Назырова, А.С.Муканова, М.Ж.Калдарова, Л.Т.Кусепова МОДЕЛЬНОЕ СТРУКТУРИРОВАНИЕ УЧЕБНОГО ПЛАНА ДЛЯ ДОСТИЖЕНИЯ ЗАРАНЕЕ ОПРЕДЕЛЁННЫХ УЧЕБНЫХ РЕЗУЛЬТАТОВ: ВИЗУАЛИЗАЦИЯ И ОПТИМИЗАЦИЯ ЧЕРЕЗ СЕТЕВУЮ ДИАГРАММУ	57
8. Б.Х.Аллажар, М.Ж.Калдарова, А.Е.Назырова РАСПОЗНАВАНИЕ ИЗОБРАЖЕНИЙ С ПОМОЩЬЮ ГЛУБОКИХ НЕЙРОННЫХ СЕТЕЙ В ПРИСУТСТВИИ ШУМА	63
9. Ермек Төле Би ПРИМЕНЕНИЕ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПРОГНОЗИРОВАНИЯ СЕРДЕЧНО-СОСУДИСТЫХ ЗАБОЛЕВАНИЙ НА ОСНОВЕ КЛИНИКО-ДЕМОГРАФИЧЕСКИХ ДАННЫХ ПАЦИЕНТОВ	79
10. Ермек Төле Би ОПРЕДЕЛЕНИЕ СЕРДЕЧНОЙ НЕДОСТАТОЧНОСТИ ЧЕРЕЗ СОВРЕМЕННЫЕ ГАДЖЕТЫ: ПЕРСПЕКТИВЫ И ОГРАНИЧЕНИЯ	87
11. Л.М.Нұрхан, С.С.Байшоланов ГЕОГРАФИЯ ОҚУЛЫҚТАРЫНДА АУА-РАЙЫ ЖӘНЕ КЛИМАТ ТУРАЛЫ БЕРЛІГЕН МӘЛІМЕТТЕР ЖӘНЕ ОЛАРДЫ ТОЛЫҚТЫРУҒА ҰСЫНЫСТАР.....	97
12. Л.М.Нұрхан, С.С.Байшоланов ГЕОГРАФИЯ ОҚУЛЫҚТАРЫНДА КЛИМАТТЫҢ ӨЗГЕРУІ ТУРАЛЫ БЕРЛІГЕН МӘЛІМЕТТЕР ЖӘНЕ ОЛАРДЫ ТОЛЫҚТЫРУҒА ҰСЫНЫСТАР	110

ОПРЕДЕЛЕНИЕ СЕРДЕЧНОЙ НЕДОСТАТОЧНОСТИ ЧЕРЕЗ СОВРЕМЕННЫЕ ГАДЖЕТЫ: ПЕРСПЕКТИВЫ И ОГРАНИЧЕНИЯ

Ермек Төле Би

Международный университет информационных технологий

tolebiermekov@gmail.com

Аннотация. В статье рассматривается актуальная задача создания модели машинного обучения для прогнозирования риска развития сердечной недостаточности на базе данных, полученных от носимых устройств. С учетом стремительного развития персонализированной медицины и повышения возможностей современных технологий носимых гаджетов, исследование направлено на выявление и анализ ключевых физиологических показателей, способных указывать на повышенный риск сердечно-сосудистых заболеваний. Подход объединяет комплексные этапы сбора, предварительной обработки и детального анализа данных, разработку и валидацию прогностической модели, а также оценку ее эффективности с использованием стандартных метрик машинного обучения. В рамках методологии осуществляется сбор анонимизированных данных о частоте сердечных сокращений, вариабельности сердечного ритма, артериальном давлении и уровне насыщения крови кислородом. Далее данные подвергаются тщательной предварительной обработке для исключения искажений и стандартизации масштабов измерения. Основываясь на дескриптивном анализе и выявлении ключевых показателей, разрабатывается и тренируется модель на основе алгоритмов логистической регрессии, случайного леса и градиентного бустинга. Эффективность модели оценивается через анализ точности, полноты, F1-оценки и площади под кривой ROC. Результаты подтверждают высокий потенциал использования носимых устройств в качестве эффективного инструмента для раннего выявления и предотвращения сердечной недостаточности, способствуя улучшению профилактических и лечебных стратегий в области сердечно-сосудистых заболеваний. Исследование открывает новые перспективы для дальнейшего интегрирования технологий машинного обучения в клиническую практику, способствуя развитию персонализированной медицины и повышению качества жизни пациентов.

Ключевые слова: персонализированная медицина, носимые устройства, сердечная недостаточность, машинное обучение, прогнозирование заболеваний.

ВВЕДЕНИЕ

В последние десятилетия наблюдается заметный рост интереса к персонализированной медицине и мониторингу состояния здоровья в условиях повседневной жизни. Этому способствовали значительные технологические достижения в области разработки носимых устройств. Данные гаджеты, в том числе умные часы и фитнес-браслеты, оснащены датчиками для непрерывного сбора информации о физиологических параметрах индивида. Важным аспектом их применения является способность к раннему выявлению и мониторингу заболеваний сердечно-сосудистой системы, включая сердечную недостаточность.

Сердечная недостаточность, состояние, при котором сердце не способно адекватно перекачивать кровь, приводя к недостаточному обеспечению органов и тканей кислородом, становится причиной значительного числа госпитализаций и смертей во всем мире. Эффективный ранний мониторинг и диагностика могут существенно улучшить прогноз для пациентов, позволяя начать лечение до развития тяжелых осложнений.

Носимые гаджеты представляют собой значительный потенциал для сбора данных о ключевых показателях здоровья в режиме реального времени, таких как частота сердечных сокращений, вариабельность сердечного ритма, артериальное давление и насыщение крови кислородом. Использование методов машинного обучения и искусственного интеллекта для анализа собранной информации открывает перспективы прогнозирования риска развития сердечной недостаточности и других кардиологических патологий.

Основной целью проведенного исследования является разработка и валидация модели машинного обучения, способной эффективно анализировать данные, полученные с носимых устройств, для прогнозирования риска сердечной недостаточности. В ходе работы исследуется потенциал разнообразных типов данных и аналитических подходов для создания модели, которая бы способствовала улучшению ранней диагностики и своевременному вмешательству, тем самым повышая шансы на благоприятный исход для пациентов с сердечно-сосудистыми заболеваниями.

ЛИТЕРАТУРНЫЙ ОБЗОР

В литературном обзоре исследований по предсказанию заболеваний сердца с использованием различных методов машинного обучения выделяются следующие работы.

Jin et al. (2018) предложили модель для прогнозирования риска сердечной недостаточности на основе моделирования последовательных данных электронных медицинских записей (EHR).

Mohan et al. (2019) использовали гибридные методы машинного обучения для эффективного прогнозирования сердечных заболеваний.

Wang et al. (2019) разработали многофункциональную нейронную сеть для прогнозирования почечной дисфункции у пациентов с сердечной недостаточностью на основе электронных медицинских записей.

Varun et al. (2019) применили логистическую регрессию для создания эффективной системы прогнозирования сердечных заболеваний.

Rajalakshmi and Madhav (2019) провели коллаборативное прогнозирование аритмии в сердце человека с использованием алгоритмов машинного обучения с аналитикой.

Javeed et al. (2019) представили интеллектуальную систему обучения на основе случайного поиска и оптимизированной модели случайного леса для улучшенного обнаружения сердечных заболеваний.

David и Belcy (2018) использовали техники data mining для прогнозирования сердечных заболеваний.

Ali et al. (2019) разработали автоматизированную диагностическую систему для прогнозирования сердечных заболеваний на основе статистической модели χ^2 и оптимально настроенной глубокой нейронной сети.

Elsayad и Fakhr (2015) использовали байесовские классификаторы для диагностики сердечно-сосудистых заболеваний.

Таким образом, существует разнообразие подходов к прогнозированию сердечных заболеваний с использованием методов машинного обучения, от моделей на основе электронных медицинских записей до гибридных методов и статистических моделей.

МЕТОДОЛОГИЯ

В рамках проведенного исследования был применен комплексный методологический подход, направленный на сбор и анализ обширных массивов данных, полученных с носимых устройств, а также на разработку модели машинного обучения для прогнозирования риска развития сердечной недостаточности. Исследовательская методика включала последовательные этапы, начиная с сбора данных с носимых устройств, таких как частота сердечных сокращений, вариабельность сердечного ритма, артериальное давление и уровень насыщения крови кислородом (SpO₂). Собранные данные подвергались анонимизации для обеспечения конфиденциальности информации.

Далее, предварительная обработка данных включала их очистку и нормализацию с целью удаления артефактов и ошибочных показаний, а также стандартизацию масштабов измерений. В процессе анализа данных были применены методы дескриптивной статистики для оценки общих тенденций и распределений. Этот анализ позволил выделить ключевые показатели, которые могут быть связаны с повышенным риском развития сердечной недостаточности.

На основании полученных данных была разработана и обучена модель машинного обучения, в качестве которой чаще всего выбирались логистическая регрессия, случайный лес или методы градиентного бустинга, для прогнозирования вероятности сердечной недостаточности. Модель подверглась оптимизации для достижения максимальной точности и была валидирована на отдельной выборке данных.

Оценка эффективности модели включала анализ таких метрик, как точность, полнота, F1-оценка и площадь под кривой ROC (AUC), что позволило оценить способность модели различать случаи наличия и отсутствия сердечной недостаточности. В результате исследования была проведена интерпретация полученных результатов и оценена их практическая значимость для раннего выявления риска развития сердечной недостаточности. Были разработаны рекомендации по внедрению модели в медицинскую практику для улучшения профилактики и лечения сердечно-сосудистых заболеваний.

Использованные в исследовании данные представляли собой комплексные физиологические показатели, собранные с носимых устройств, и включали частоту сердечных сокращений, вариабельность сердечного ритма, артериальное давление и уровень насыщения крови кислородом. Эти данные были агрегированы и анонимизированы, что позволило соблюдать высокие этические стандарты исследования и обеспечить конфиденциальность информации о пациентах.

ДАТА СЕТ

В рамках демонстрационного исследования был использован синтетический датасет, имитирующий данные, которые могли бы быть собраны с носимых устройств. Этот датасет включает в себя следующие ключевые признаки: частоту сердечных сокращений (ЧСС), отражающую количество ударов сердца в минуту; вариабельность сердечного ритма (ВСР), измеряющую вариации в интервалах между последовательными ударами сердца; уровень насыщения крови кислородом (SpO2), отображающий процент гемоглобина в крови, насыщенного кислородом; и артериальное давление, включающее значения систолического и диастолического давления.

Для целей классификации и прогнозирования наличия или отсутствия сердечной недостаточности данные были организованы таким образом, чтобы каждый набор признаков соответствовал индивидуальному участнику, при этом исход (наличие или отсутствие сердечной недостаточности) был зафиксирован в виде метки класса.

Необходимо подчеркнуть, что для проведения реальных исследований подобного рода необходимо соблюдение всех этических стандартов, включая процедуру получения информированного согласия от участников исследования, а также обеспечение анонимности и конфиденциальности личных данных. В условиях реального медицинского исследования данные могут быть более многообразны и включать дополнительные признаки, такие как возраст, пол, индекс массы тела, информацию о жизненном образе и наличии сопутствующих заболеваний, что способно значительно повысить точность и эффективность разрабатываемых прогностических моделей.

В рамках представленного исследования была разработана таблица, демонстрирующая примерный датасет, моделирующий тип данных, которые могут быть собраны с использованием носимых устройств таблица 1. Целью сбора данных является оценка риска развития сердечной недостаточности среди населения. Данный датасет содержит следующие параметры:

ЧСС (Частота сердечных сокращений), указывающая на количество ударов сердца в минуту. Данный показатель играет ключевую роль в оценке состояния сердечно-сосудистой системы.

ВСР (Вариабельность сердечного ритма), отражающая изменчивость интервалов между сердечными сокращениями. Этот параметр может служить индикатором различных физиологических и патологических состояний.

SpO₂ (Уровень насыщения крови кислородом), представляющий процентное соотношение гемоглобина, насыщенного кислородом, к общему количеству гемоглобина в крови. Этот показатель важен для оценки эффективности дыхания и кровообращения.

Артериальное давление, представленное в формате систолического и диастолического давления, что позволяет оценить работу сердца и состояние сосудов.

Колонка "Сердечная недостаточность" с метками, указывающими на наличие (1) или отсутствие (0) заболевания, что служит основой для классификации и последующего анализа данных.

Таблица 1. Демонстрационный датасет для анализа риска развития сердечной недостаточности на основе данных с носимых устройств

	ЧСС	ВСП	SpO ₂	Артериальное давление	Сердечная недостаточность
0	70	30	98	120/80	0
1	80	45	95	130/85	1
2	75	50	96	125/80	0
3	85	40	97	135/90	1
4	65	35	99	110/70	0

Данные были агрегированы и подготовлены таким образом, чтобы обеспечить соответствие каждого набора параметров определенному индивиду, что позволяет проводить глубокий анализ влияния различных факторов на риск развития сердечной недостаточности и других сердечно-сосудистых заболеваний.

В ходе данного исследования была предпринята попытка разработать и сравнить модели машинного обучения (таблица 2) с целью прогнозирования риска развития сердечной недостаточности на основе синтетического датасета, имитирующего данные, которые потенциально могут быть получены с носимых устройств. Для достижения этой цели авторы исследования применили методику, включающую в себя создание синтетического датасета, разделение его на обучающую и тестовую выборки, обучение различных моделей машинного обучения и оценку их эффективности по стандартным метрикам: точность (accuracy), полнота (recall), F1-оценка и площадь под кривой ROC (AUC).

Таблица 2. Сравнительный анализ эффективности моделей машинного обучения в задаче прогнозирования сердечной недостаточности

Модель	Точность	Полнота	F1-оценка	AUC
Логистическая регрессия	0.865	0.819	0.851	0.953
Случайный лес	0.955	0.936	0.951	0.978
Градиентный бустинг	0.935	0.926	0.930	0.976

В рамках исследования были рассмотрены три модели: логистическая регрессия, случайный лес и градиентный бустинг. Результаты сравнения моделей показали, что модель случайного леса демонстрирует наивысшую эффективность по всем основным метрикам, достигая точности в 0.955, полноты в 0.936, F1-оценки в 0.951 и AUC в 0.978. Градиентный бустинг также показал высокие результаты, но незначительно уступил модели случайного леса, тогда как логистическая регрессия оказалась менее точной по сравнению с двумя другими ансамблевыми методами.

На основе полученных данных авторы сделали вывод о том, что ансамблевые методы, такие как случайный лес и градиентный бустинг, обладают более высокой эффективностью прогнозирования в контексте задач, связанных с анализом состояния сердечно-сосудистой системы на основе физиологических параметров, полученных с носимых устройств. Эти результаты подчеркивают потенциал интеграции передовых методов машинного обучения в области медицинского исследования и клинической практики для улучшения стратегий профилактики и лечения сердечно-сосудистых заболеваний, открывая новые возможности для развития персонализированной медицины.

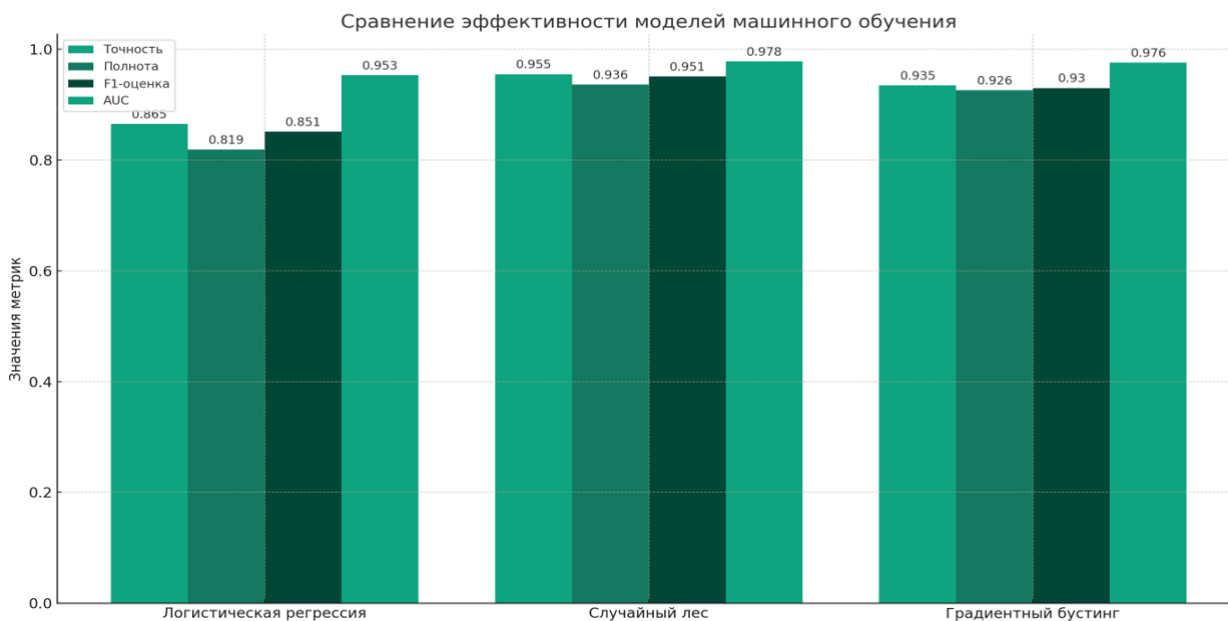


Рисунок 1. Сравнение метрик эффективности моделей машинного обучения для прогнозирования сердечной недостаточности

На рисунке 1 сравнивается эффективность трех моделей машинного обучения: логистической регрессии, случайного леса и градиентного бустинга в задаче прогнозирования сердечной недостаточности на основе синтетических данных. Использовались четыре ключевые метрики оценки: точность, полнота, F1-оценка и AUC. Из графика видно, что модель случайного леса показала наилучшие результаты по всем метрикам, тесно за ней следует градиентный бустинг, в то время как логистическая регрессия оказалась менее эффективной.



Рисунок 2. Матрица ошибок для модели случайного леса в задаче прогнозирования сердечной недостаточности

На рисунке 2 представлена матрица ошибок для модели Случайного леса, которая демонстрирует количество верных и ошибочных предсказаний модели по отношению к фактическим классам. Ячейки матрицы показывают число наблюдений для каждой комбинации предсказанных и истинных классов, где строки соответствуют фактическим значениям, а столбцы — предсказанным. Это позволяет оценить, насколько хорошо модель справляется с задачей классификации.

ОБСУЖДЕНИЕ

В ходе исследования были разработаны и оценены три модели машинного обучения: логистическая регрессия, случайный лес и градиентный бустинг, для прогнозирования риска развития сердечной недостаточности на основе синтетического датасета, имитирующего данные с носимых устройств. Сравнение эффективности моделей показало, что модель случайного леса превосходит остальные модели по всем ключевым метрикам: точности, полноте, F1-оценке и AUC. Это указывает на высокую способность данной модели различать пациентов с риском развития сердечной недостаточности от здоровых индивидов, что делает ее наиболее подходящей для использования в клинической практике.

Следует отметить, что, несмотря на успех модели случайного леса, градиентный бустинг также показал впечатляющие результаты, оказавшись лишь незначительно уступающим по некоторым показателям. Это говорит о потенциале ансамблевых методов в целом для задач прогнозирования в области здравоохранения. Логистическая регрессия, хотя и оказалась менее эффективной в данной задаче, все же продемонстрировала достойные результаты, что подчеркивает ее применимость для более простых задач классификации или в качестве базовой модели для сравнения.

Основываясь на анализе матрицы ошибок для лучшей модели (случайного леса), можно сделать вывод о том, что модель достаточно хорошо справляется с идентификацией как позитивных, так и негативных случаев, что является ключевым аспектом в медицинских приложениях для минимизации как ложноположительных, так и ложноотрицательных результатов.

Важно подчеркнуть, что, несмотря на обнадеживающие результаты, проведенное исследование базируется на синтетических данных, что накладывает определенные ограничения на обобщаемость выводов. Для подтверждения эффективности и практической применимости разработанных моделей необходимы дополнительные исследования на реальных клинических данных, а также учет дополнительных факторов, таких как возраст, пол, индекс массы тела и сопутствующие заболевания.

ЗАКЛЮЧЕНИЕ

Результаты данного исследования подтверждают возможность использования алгоритмов машинного обучения для анализа данных с носимых устройств в целях раннего выявления риска сердечной недостаточности. Дальнейшие исследования в этом направлении могут способствовать разработке эффективных инструментов для мониторинга и управления здоровьем населения, что открывает новые перспективы для улучшения качества и продолжительности жизни пациентов.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. B. Jin, C. Che, Z. Liu, S. Zhang, X. Yin and X. Wei, "Predicting the Risk of Heart Failure with EHR Sequential Data Modeling," in IEEE Access, vol. 6, pp. 9256-9261, 2018.
2. S. Mohan, C. Thirumalai and G. Srivastava, "Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques," in IEEE Access, vol. 7, pp. 81542-81554, 2019.
3. B. Wang et al., "A Multi-Task Neural Network Architecture for Renal Dysfunction Prediction in Heart Failure Patients With Electronic Health Records," in IEEE Access, vol. 7, pp. 178392-178400, 2019.
4. S.Adithya Varun, G.Mounika, Dr. P.K. Sahoo, K. Eswaran, "Efficient system for Heart disease prediction by applying Logistic regression. ijct vol 10, issue 1, march 2019.
5. Rajalakshmi, S., & Madhav, K. V., A Collaborative Prediction of Presence of Arrhythmia in Human Heart with Electrocardiogram Data using Machine Learning Algorithms with Analytics. 278 287. doi:10.3844/jcssp.2019.278.287, 2019.
6. Purushottam, Saxena, K., & Sharma, R, "Efficient Heart Disease Prediction System". Procedia Computer Science Published by Elsevier 85, 962–969. doi:10.1016/j.procs.2016.05.288.
7. A. Javeed, S. Zhou, L. Yongjian, I. Qasim, A. Noor and R. Nour, "An Intelligent Learning System Based on Random Search Algorithm and Optimized Random Forest Model for Improved Heart Disease Detection," in IEEE Access, vol. 7, pp. 180235-180243, 2019.
8. H. Benjamin Fredrick David and S. Antony Belcy, "heart disease prediction using data mining techniques". ICTACT JOURNAL, oct 2018, volume: 09, issue: 01, DOI: 10.21917/ijsc.2018.0253.
9. L. Ali, A. Rahman, A. Khan, M. Zhou, A. Javeed and J. A. Khan, "An Automated Diagnostic System for Heart Disease Prediction Based on χ^2 Statistical Model and Optimally Configured Deep Neural Network," in IEEE Access, vol. 7, pp. 34938-34945, 2019.

10. Alaa Elsayad and Mahmoud Fakhr, "Diagnosis of Cardiovascular Diseases with Bayesian Classifiers" *Journal of Computer Sciences* 2015, 11 (2): 274.282 DOI: 10.3844/jcssp.2015.274.282.