

ВНУТРЕННИЕ БОЛЕЗНИ

INTERNAL DISEASES

УДК 004.032.26

doi:10.21685/2072-3032-2022-4-1

Нейросетевая модель для ранней диагностики хронической сердечной недостаточности

В. И. Горбаченко¹, В. В. Потапов²,
О. К. Зенин³, И. С. Милтых⁴, Д. Н. Грибков⁵

^{1,3,4,5}Пензенский государственный университет, Пенза, Россия

²Донецкий национальный медицинский университет
имени М. Горького, Донецк, Россия

¹gorvi@mail.ru, ²x3x3x23@rambler.ru, ³zen.olegz@gmail.com,
⁴ilia@miltykh.com, ⁵rororol23@mail.ru

Аннотация. *Актуальность и цели.* По данным Всемирной организации здравоохранения, заболевания сердечно-сосудистой системы являются самой частой причиной инвалидизации и смертности. Широкая распространенность, а также тяжесть хронической сердечной недостаточности диктуют необходимость разработки новых методов ранней диагностики и контроля лечения. Ранняя диагностика уменьшает количество пациентов, которым требуется стационарное лечение, сокращает число дней нетрудоспособности, а также снижает вероятность неблагоприятных исходов. *Материалы и методы.* В состав биологических жидкостей человека входят поверхностно-активные вещества, которые способны изменять поверхностное натяжение, ускорять и замедлять процессы переноса вещества и газов. Их количество в биологических жидкостях человека может изменяться в связи с различными патологиями, что может являться ранним маркером заболевания. Для ранней диагностики хронической сердечной недостаточности предложено использовать методы тензиометрии. Для выбора объективного критерия диагностики патологического состояния применено нейросетевое моделирование. Данные для обучения сети получены с помощью тензиометра. *Результаты.* Применение нейронной сети позволяет с высокой точностью прогнозировать хроническую сердечную недостаточность по результатам тензиометрии крови. Реализована нейронная сеть, позволившая диагностировать хроническую сердечную недостаточность на ранних этапах с точностью 98 %. *Выводы.* Данные о тензиометрических параметрах пациентов с хронической сердечной недостаточностью и здоровых людей являются ключевым моментом в разработке новых диагностических подходов. Использование результатов анализа изменений тензиометрических показателей плазмы и сыворотки крови с последующей обработкой с помощью нейронной сети позволяет с высокой точностью проводить раннюю диагностику хронической сердечной недостаточности.

Ключевые слова: сердечная недостаточность, тензиометрия, обучение нейросетевых моделей, метрики качества нейросетевой модели

Финансирование: данное исследование реализовано в рамках «Ректорского гранта» для обучающихся Пензенского государственного университета.

Для цитирования: Горбаченко В. И., Потапов В. В., Зенин О. К., Милтых И. С., Грибков Д. Н. Нейросетевая модель для ранней диагностики хронической сердечной недостаточности // Известия высших учебных заведений. Поволжский регион. Медицинские науки. 2022. № 4. С. 5–15. doi:10.21685/2072-3032-2022-4-1

A neural network model for early diagnosis of chronic heart failure

V.I. Gorbachenko¹, V.V. Potapov², O.K. Zenin³, I.S. Milytkh⁴, D.N. Gribkov⁵

^{1,3,4,5}Penza State University, Penza, Russia

²Donetsk National Medical University named after M. Gorky, Donetsk, Russia

¹gorvi@mail.ru, ²x3x3x23@rambler.ru, ³zen.olegz@gmail.com,

⁴ilia@milytkh.com, ⁵rororol23@mail.ru

Abstract. *Background.* The WHO reports that cardiovascular diseases are the leading cause of disability and mortality. The high prevalence as well as the severity of chronic heart failure requires the development of new methods of early diagnosis and treatment monitoring. Early diagnosis reduces the number of patients that require inpatient treatment, reduces the number of days of disability, and also reduces the probability of adverse outcomes. *Materials and methods.* Human bodily fluids include surfactants that can change surface tension, accelerate and decelerate the transport of substances and gases. Their amount in human biological fluids may change due to various pathologies, which may be an early marker of the disease. Tensiometry methods have been suggested for early diagnosis of CHF. To select an objective criterion for diagnosing a pathological condition, neural network modeling was applied. Training data for the neural model was obtained using a tensiometer. *Results.* Neural network allows to diagnose chronic heart failure based on blood tensiometry results with high accuracy. The neural network allows to diagnose CHF at early stages with 98% accuracy. *Conclusions.* Tensiometric data of patients with CHF and healthy people are a key point in the development of new diagnostic approaches. Analysis of changes in tensiometric parameters of blood plasma and serum with subsequent neural network processing allows an early diagnosis of chronic heart failure with high accuracy.

Keywords: heart failure, tensiometry, neural network model training, neural network model quality metrics

Acknowledgements: the research was realized within the “Rector’s grant” for students of Penza State University.

For citation: Gorbachenko V.I., Potapov V.V., Zenin O.K., Milytkh I.S., Gribkov D.N. A neural network model for early diagnosis of chronic heart failure. *Izvestiya vysshikh uchebnykh zavedeniy. Povolzhskiy region. Meditsinskie nauki* = *University proceedings. Volga region. Medical sciences*. 2022;(4):5–15. (In Russ.). doi:10.21685/2072-3032-2022-4-1

Введение

Хроническая сердечная недостаточность (ХСН) – это хронический синдром, характеризующийся ухудшением качества и уменьшением продолжительности жизни. Неблагоприятный прогноз [1, 2] и широкое распространение среди лиц старше 65 лет [3, 4] делают ХСН одной из острейших проблем здравоохранения, которая влияет на экономику и качество жизни [5]. Поэтому одним из важнейших направлений современной кардиологии является разработка методов ранней диагностики данного состояния.

Для ранней диагностики ХСН применяется метод полимеразной цепной реакции с гибридационно-флуоресцентной детекцией продуктов амплификации центрифугированной пробы крови пациента [6]. Однако слож-

ность и дороговизна данной методики, а также необходимость обеспечения полной стерильности являются основными ее недостатками.

Другим методом ранней диагностики развития хронической сердечной недостаточности является определение тензио- и реометрических параметров сыворотки и плазмы крови. В состав биологических жидкостей человека входят поверхностно-активные вещества (ПАВ), которые способны изменять поверхностное натяжение, ускорять и замедлять процессы переноса вещества (например, субстратов метаболизма) и газов (например, кислорода и углекислого газа). Количество ПАВ в биологических жидкостях человека (например, в крови) может изменяться в связи с различной патологией. Однако для данного подхода существует проблема выбора объективного критерия диагностики патологического состояния, которая может быть решена путем применения методов нейросетевого моделирования [7].

Цель исследования: разработать нейросетевую модель для ранней диагностики ХСН на основе тензио- и реометрических показателей плазмы и сыворотки крови.

Материалы и методы

Определение динамического поверхностного натяжения производили с применением тензиометра PAT-1 (SINTERFACE Technologies, Берлин, Германия), используя метод «висящей капли». Исследовали динамическое и равновесное поверхностное натяжение, модуль вязкоупругости при частотах 0,1 и 0,01 Гц, а также фазовый угол. Так как динамическое поверхностное натяжение считается равновесным после 1500–2000 с, осцилляцию капли начинали после 2000 с. Определение параметров динамического поверхностного натяжения производили в нативных и депротеинизированных пробах [8–10].

Исходными данными для построения нейросетевой модели служили результаты тензиометрических исследований жидкой части крови (плазмы и сыворотки) здоровых добровольцев и людей, страдающих ХСН, представленные в виде таблиц Excel [11]. Данные также включали пол и возраст пациента, закодированные по рекомендациям Всемирной организации здравоохранения.

Для применения нейросетевой модели необходимо подготовить табличные обучающие данные [12]. Пропуски данных и выбросы отсутствуют. Числовые признаки были масштабированы. Качественные признаки кодировались с помощью One-Hot Encoding [13]. Проблема заключается в сложности получения большого объема обучающей выборки. Для решения этой проблемы применена аугментация данных, т.е. искусственное увеличение объема обучающей выборки. Аугментация изображений достаточно хорошо развита [14] и сводится в основном к изменениям размера и различным трансформациям исходных изображений. Аугментация табличных данных менее развита и сводится к генерации синтетических данных, статистически близких к исходным данным.

Известны следующие основные методы генерации синтетических данных [15, 16]:

- статистические методы;
- методы, основанные на машинном обучении.

Статистическими методами можно создать набор данных без использования реальных данных, если известно распределение реальных данных. Статистическими методами можно генерировать синтетические данные, определив статистическое распределение для имеющегося набора данных.

Методы, основанные на машинном обучении, позволяют генерировать наиболее реалистичные синтетические данные, включая изображения и тексты. Наибольшую популярность получили три метода машинного обучения генерирования синтетических данных.

Вариационный автокодировщик [17, 18], в отличие от простейшего автокодировщика, в процессе обучения формирует не значения элементов скрытого (латентного) слоя, а два вектора – вектор μ средних значений элементов скрытого слоя и вектор σ стандартных отклонений элементов. Вектор h входа декодера формируется из случайного вектора z , элементы которого распределены по нормальному закону с нулевым математическим ожиданием и единичным стандартным отклонением, по формуле:

$$h = z \odot \sigma + \mu, \quad (1)$$

где $\odot$ – операция поэлементного умножения.

Обучение вариационного автокодировщика производится методом обратного распространения ошибки. Функция потерь включает средний квадрат ошибки реконструкции входного вектора и регуляризатор, являющийся расстоянием Кульбака – Лейблера [19] между случайным скрытым вектором с параметрами μ , σ и случайного вектора, компоненты которого распределены по нормальному закону с нулевым математическим ожиданием и единичным стандартным отклонением. За счет члена регуляризации распределение скрытого представления приближается к стандартному нормальному распределению. Обученный вариационный автокодировщик позволяет из случайного вектора z формировать векторы, близкие к векторам обучающей выборки. Для этого используется только декодировщик, на вход которого подаются векторы, сформированные по (1).

Генеративно-состязательные сети [20] реализуют алгоритм машинного обучения без учителя, использующий две нейронных сети. Одна сеть (генератор) генерирует образцы, а другая сеть (дискриминатор) учится отличать правильные образцы от неправильных.

Диффузионные модели. Известно несколько генеративных моделей на основе диффузии [21–23]. В основе различных диффузионных моделей лежат схожие идеи. Диффузионные модели сначала искажают обучающие данные, постепенно добавляя гауссовский шум, стирая детали, пока данные не превратятся в чистый шум, а затем нейронная сеть обучается, чтобы восстановить обучающие данные.

Генеративно-состязательные сети и диффузионные модели широко используются для синтеза реалистических изображений, но очень сложны в обучении и требуют больших вычислительных затрат. Статистические методы требуют наличия достаточно большого объема данных для оценки законов распределения генерируемых признаков. Поэтому для синтеза табличных данных использован вариационный автокодировщик. Известно несколько свободно распространяемых библиотек генерации табличных данных

[24, 25]. Мы использовали библиотеку [25], основанную на вариационном автокодировщике.

В нашем случае признаки примеров, относящихся к пациентам с патологией и без патологии, распределены по различным законам. Поэтому для генерации синтетических данных выборка была разделена на две и генерация производилась отдельно для пациентов с патологией и без патологии. Исходные данные содержали 132 обучающих примера, из них 65 примеров с патологией, 67 примеров без патологии. На основе библиотеки *deep_tabular_augmentation* [25], использующей вариационный автокодировщик, было сгенерировано по 660 примеров, соответствующих пациентам с патологией и без патологии. Лучшие результаты получены для длины векторов латентного слоя, равной 3. Для исключения корреляционных связей между сгенерированными признаками, как рекомендуется в [25], использовалась генерация с добавлением шума. Для визуальной оценки качества генерации признаков использовались диаграммы рассеяния (*scatter plot*) реальных и сгенерированных данных (*fake data*). Примеры диаграмм рассеяния приведены на рис. 1.

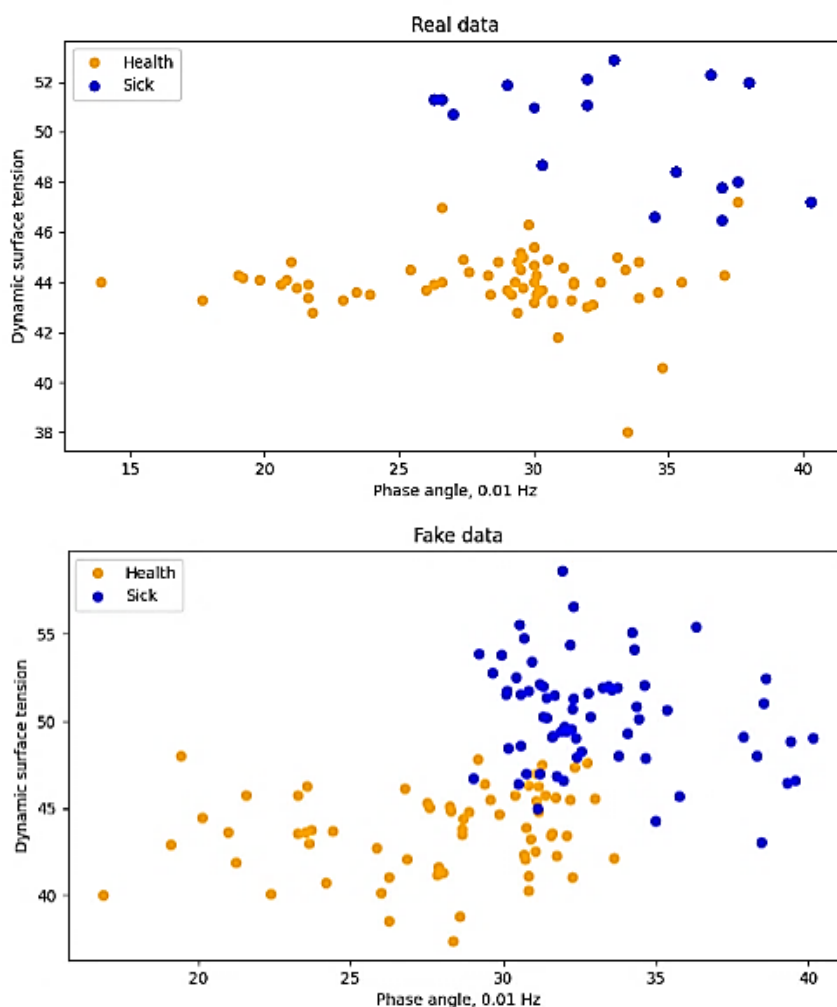


Рис. 1. Диаграммы рассеяния для фазового угла 0,1 Гц

Диаграммы рассеяния показывают, что средние значения и диапазоны изменения реальных и сгенерированных данных близки. Реальные и сгенерированные данные объединялись в единый набор данных, использовавшийся для обучения нейронной сети.

Результаты и обсуждение

Машинное обучение широко применяется в кардиологии для персонализированного прогнозирования и оценки сердечно-сосудистого риска. Разработаны нейросетевые модели определения прогностических фенотипов у пациентов с ХСН и сохраненной фракцией выброса [26]. В то же время существующие подходы нацелены прежде всего на определение риска обострения, а не на раннюю диагностику, поэтому в них не используются ранние маркеры развития ХСН.

Широко известно, что поверхностно-активные вещества в составе биологических жидкостей, находясь на границе раздела фаз, способны изменять межфазное (поверхностное) натяжение. Для различных заболеваний сердца, например ишемической болезни сердца, характерно увеличение поверхностного натяжения сыворотки крови за счет повышения концентрации фосфолипидов, фосфатидилхолина, креатинина, креатина, альбумина и других веществ [9, 10]. Однако существует сложность в выборе объективного диагностического критерия из-за ненормального распределения данных. Для выбора объективного критерия диагностики патологического состояния применено нейросетевое моделирование. Решаемая на нейронной сети задача представляет собой задачу двухклассовой классификации: наличие/отсутствие патологии. Для решения этой задачи реализована полносвязная нейронная сеть.

Эксперименты с нейронной сетью проводились с использованием библиотек TensorFlow и Keras. В обучающем наборе данных 80 % примеров выделялось для обучения сети, а 20 % – для тестирования качества обучения. В данных для обучения 80 % выделялось для тренировки сети, а 20 % – для проверки (валидации) сети на переобучение. Сеть обучалась одним из наиболее эффективных современных алгоритмов переменной скоростью обучения – алгоритмом Adam [27]. В качестве функции потерь применялась двоичная кросс-энтропия. Лучшие результаты получены для нейронной сети следующей конфигурации: количество слоев – 3, количество нейронов в каждом из скрытых слоев – 64, функция активации в скрытых слоях – гиперболический тангенс. В выходном слое использована функция активации softmax.

Графики изменения функции потерь и точности классификации тренировочных и валидационных данных для обучения этой сети показаны на рис. 2. Графики рис. 2 показывают, что переобучение отсутствует.

Были рассчитаны стандартные для задач классификации метрики качества [28]. Матрица ошибок, сформированная на тестовых данных, показывает высокую точность диагностики как наличия, так и отсутствия патологии. Стандартные метрики классификатора имеют следующие значения. Точность – число правильно классифицированных наблюдений, отнесенное к общему числу наблюдений: $R_{AC} = 0,97$. Прецизионность – точность предсказания положительного исхода: для случая патологии $R_{PR0} = 0,98$, для случая отсутствия патологии $R_{PR1} = 0,96$. Полнота – доля истинно положительных примеров, правильно определенных классификатором: для случая патологии

$S_{R0} = 0,96$, для случая отсутствия патологии $S_{R1} = 0,98$. F1-score – гармоническое среднее между прецизионностью и полнотой:

$$F_1 = 2 \frac{R_{PR} \cdot S_R}{R_{PR} + S_R} = 0,97.$$

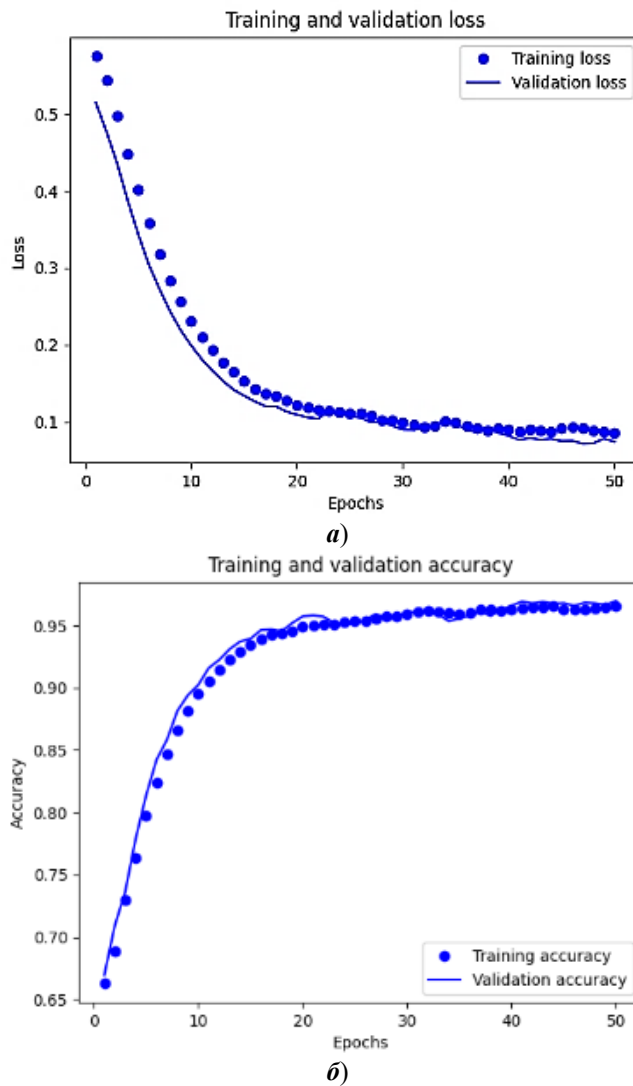


Рис. 2. Графики изменения функции потерь (а) и точности классификации (б) на тренировочных и валидационных данных

Площадь под ROC-кривой, равная $AUC = 0,97$, говорит о высокой прогностической силе модели.

Таким образом, применение нейронной сети позволяет с высокой точности прогнозировать хроническую сердечную недостаточность по результатам тензиометрии крови.

Заключение

Данные о тензиометрических параметрах пациентов с ХСН и здоровых людей являются ключевым моментом в разработке новых диагностических подходов. Использование результатов анализа изменений тензиометрических показателей плазмы и сыворотки крови с последующей обработкой с помощью нейронной сети позволяет с высокой точностью проводить раннюю диагностику хронической сердечной недостаточности. Дальнейшее направление исследований направлено на совершенствование нейросетевого подхода к диагностике хронической сердечной недостаточности и внедрению нейросетевой программы в виде приложения с удобным для врача интерфейсом.

Список литературы

1. Chang G., Zhang W., Zhang M., Ding G. Clinical Value of Circulating ZFAS1 and miR-590-3p in the Diagnosis and Prognosis of Chronic Heart Failure // *Cardiovascular Toxicology*. 2021. Vol. 21, № 11. P. 880–888.
2. Jessup M., Brozena S. Heart Failure // *New England Journal of Medicine*. 2003. Vol. 348, № 20. P. 2007–2018.
3. Christiansen M. N. [et al.]. Age-Specific Trends in Incidence, Mortality, and Comorbidities of Heart Failure in Denmark, 1995 to 2012 // *Circulation*. 2017. Vol. 135, № 13. P. 1214–1223.
4. Mosterd A., Hoes A. W. Clinical epidemiology of heart failure // *Heart*. 2007. Vol. 93, № 9. P. 1137–1146.
5. Berry C., Murdoch D. R., McMurray J. J. V. Economics of chronic heart failure // *European Journal of Heart Failure*. 2001. Vol. 3, № 3. P. 283–291.
6. Патент 2569458 Российская Федерация, МПК C12Q 1/68. Способ диагностики хронической сердечной недостаточности путем использования циркулирующих уровней микроРНК / Жиров И. В. ; заявитель и патентообладатель ФГБУ "РКНПК" МЗ РФ. № 2014148454/10 ; заявл. 02.12.2014 ; опубл. 27.11.2015, Бюл. № 33. 5 с.
7. Fauvel K. Fromont E., Masson V. XEM: An explainable-by-design ensemble method for multivariate time series classification // *Data Mining and Knowledge Discovery*. 2022. Vol. 36, № 3. P. 917–957.
8. Патент № 2745337 C1 Российская Федерация, МПК G01N 13/02, G01N 33/49. Способ ранней диагностики развития хронической сердечной недостаточности / Моисеева И. Я., Зенин О. К., Потапов В. В. [и др.] ; заявитель ФГБОУ ВО «Пензенский государственный университет». № 2020114334 ; заявл. 21.04.2020 ; опубл. 24.03.2021.
9. Моисеева И. Я., Зенин О. К., Потапов В. В. [и др.]. Новый способ ранней диагностики развития хронической сердечной недостаточности // *Известия высших учебных заведений. Поволжский регион. Медицинские науки*. 2021. № 3. С. 19–26.
10. Потапов В. В., Шраменко Е. К., Зенин О. К. Поверхностное натяжение и дилатационная реология крови у пациентов, оперированных на сердце в условиях искусственного кровообращения // *Коллоидный журнал*. 2020. Т. 82, № 1. С. 61–65.
11. Свидетельство о государственной регистрации базы данных № 2022621033. База данных результатов тензиометрии сыворотки крови здоровых добровольцев и больных с хронической сердечной недостаточностью / В. В. Потапов, О. Зенин, И. Я. Моисеева, И. С. Милтых. № 2021623094. заявл. 28.04.2022 ; зарег. 05.05.2022.
12. Чжен Э., Казари А. Машинное обучение. Конструирование признаков: принципы и техника для аналитиков. М. : Эксмо, 2022. 240 с.

13. Brownlee J. Why One-Hot Encode Data in Machine Learning? URL: <https://machinelearningmastery.com/why-one-hot-encode-data-in-machine-learning/> (дата обращения: 03.09.2022).
14. Shorten C., Khoshgoftaar T. M. A survey on Image Data Augmentation for Deep Learning // *Journal of Big Data*. 2019. Vol. 6, № 1. P. 1–60.
15. El Emam K., Mosquera L., Hoptroff R. Practical synthetic data generation: balancing privacy and the broad availability of data. Practical synthetic data generation. First Edition. Sebastopol, CA : O'Reilly Media, Inc, 2020. 151 p.
16. Synthetic Data Generation: 3 Key Techniques and Tips for Success. URL: <https://datagen.tech/guides/synthetic-data/synthetic-data-generation/> (дата обращения: 03.09.2022).
17. Kingma D. P., Welling M. Auto-Encoding Variational Bayes // *arXiv*. 2014.
18. Аггарвал Ч. Нейронные сети и глубокое обучение : учебный курс. СПб. : ООО «Диалектика», 2020. 752 с.
19. Бишоп К. Распознавание образов и машинное обучение. СПб. : ООО «Диалектика», 2020. 960 с.
20. Goodfellow I. J. [et al.]. Generative Adversarial Networks // *arXiv*. 2014.
21. Sohl-Dickstein J. [et al.]. Deep Unsupervised Learning using Nonequilibrium Thermodynamics. 2015.
22. Song Y., Ermon S. Generative Modeling by Estimating Gradients of the Data Distribution // *arXiv*. 2020.
23. Ho J., Jain A., Abbeel P. Denoising Diffusion Probabilistic Models // *arXiv*. 2020.
24. The Synthetic Data Vault. Put synthetic data to work! URL: <https://sdv.dev/> (дата обращения: 03.09.2022).
25. Schmidt L. Data Augmentation for Tabular Data. URL: <https://medium.com/analytics-vidhya/data-augmentation-for-tabular-data-f75c94398c3e> (дата обращения: 03.09.2022).
26. Kwon J., Lee Y., Lee Y., Lee S., Park J. An Algorithm Based on Deep Learning for Predicting In-Hospital Cardiac Arrest // *Journal of the American Heart Association*. 2018. Vol. 7, № 13. P. e008678.
27. Kingma D. P., Ba J. Adam: A Method for Stochastic Optimization // *arXiv*. 2017.
28. Брюс П., Брюс Э., Гедек П. Практическая статистика для специалистов Data Science. СПб. : БХВ-Петербург, 2021. 352 с.

References

1. Chang G., Zhang W., Zhang M., Ding G. Clinical Value of Circulating ZFAS1 and miR-590-3p in the Diagnosis and Prognosis of Chronic Heart Failure. *Cardiovascular Toxicology*. 2021;21(11):880–888.
2. Jessup M., Brozena S. Heart Failure. *New England Journal of Medicine*. 2003;348(20):2007–2018.
3. Christiansen M.N. et al. Age-Specific Trends in Incidence, Mortality, and Comorbidities of Heart Failure in Denmark, 1995 to 2012. *Circulation*. 2017;135(13):1214–1223.
4. Mosterd A., Hoes A.W. Clinical epidemiology of heart failure. *Heart*. 2007;93(9):1137–1146.
5. Berry C., Murdoch D.R., McMurray J.J.V. Economics of chronic heart failure. *European Journal of Heart Failure*. 2001;3(3):283–291.
6. Patent 2569458 Russian Federation, MPK C12Q 1/68. *Sposob diagnostiki khronicheskoy serdechnoy nedostatochnosti putem ispol'zovaniya tsirkuliruyushchikh urovney mikroRNK = Method for diagnosing chronic heart failure by using circulating miRNA levels*. Zhiron I.V.; applicant and patent holder FGBU "RKNPK" MZ RF. № 2014148454/10; appl. 02.12.2014; publ. 27.11.2015, Bull. № 33. 5 s. (In Russ.)
7. Fauvel K., Fromont E., Masson V. XEM: An explainable-by-design ensemble method for multivariate time series classification. *Data Mining and Knowledge Discovery*. 2022;36(3):917–957.

8. Patent № 2745337 C1 Russian Federation, MPK G01N 13/02, G01N 33/49. *Sposob ranney diagnostiki razvitiya khronicheskoy serdechnoy nedostatochnosti = A method for early diagnosis of the development of chronic heart failure*. Moiseeva I.Ya., Zenin O.K., Potapov V.V. et al.; applicant FGBOU VO «Penzen-skiy gosudarstvennyy universitet». № 2020114334; appl. 21.04.2020; publ. 24.03.2021. (In Russ.)
9. Moiseeva I.Ya., Zenin O.K., Potapov V.V. et al. A new method for early diagnosis of the development of chronic heart failure. *Izvestiya vysshikh uchebnykh zavedeniy. Povolzhskiy region. Meditsinskie nauki = University proceedings. Volga region. Medical sciences*. 2021;(3):19–26. (In Russ.)
10. Potapov V.V., Shramenko E.K., Zenin O.K. Surface tension and dilated blood rheology in patients operated on the heart under cardiopulmonary bypass. *Kolloidnyy zhurnal = Colloid journal*. 2020;82(1):61–65. (In Russ.)
11. *Svidetel'stvo o gosudarstvennoy registratsii bazy dannykh № 2022621033. Baza dannykh rezul'tatov tenziometrii syvorotki krovi zdorovykh dobrovol'tsev i bol'nykh s khronicheskoy serdechnoy nedostatochnost'yu = Certificate of state registration of the database No. 2022621033. Database of blood serum tensiometry results of healthy volunteers and patients with chronic heart failure*. V.V. Potapov, O. Zenin, I.Ya. Moiseeva, I.S. Milyukh. № 2021623094. appl. 28.04.2022; reg. 05.05.2022. (In Russ.)
12. Chzhen E., Kazari A. Mashinnoe obuchenie. *Konstruirovaniye priznakov: printsipy i tekhnika dlya analitikov = Feature engineering: principles and techniques for analysts*. Moscow: Eksmo, 2022:240. (In Russ.)
13. Brownlee J. *Why One-Hot Encode Data in Machine Learning?* Available at: <https://machinelearningmastery.com/why-one-hot-encode-data-in-machine-learning/> (accessed 03.09.2022).
14. Shorten C., Khoshgoftaar T.M. A survey on Image Data Augmentation for Deep Learning. *Journal of Big Data*. 2019;6(1):1–60.
15. El Emam K., Mosquera L., Hoptroff R. *Practical synthetic data generation: balancing privacy and the broad availability of data. Practical synthetic data generation. First Edition*. Sebastopol, CA: O'Reilly Media, Inc, 2020:151.
16. *Synthetic Data Generation: 3 Key Techniques and Tips for Success*. Available at: <https://datagen.tech/guides/synthetic-data/synthetic-data-generation/> (accessed 03.09.2022).
17. Kingma D.P., Welling M. Auto-Encoding Variational Bayes. *arXiv*. 2014.
18. Aggarwal Ch. *Neyronnye seti i glubokoe obuchenie: uchebnyy kurs = Neural networks and deep learning: textbook*. Saint Petersburg: OOO «Dialektika», 2020:752. (In Russ.)
19. Bishop K. *Raspoznavaniye obrazov i mashinnoe obuchenie = Pattern recognition and machine learning*. Saint Petersburg: OOO «Dialekti-ka», 2020:960. (In Russ.)
20. Goodfellow I.J. et al. Generative Adversarial Networks. *arXiv*. 2014.
21. Sohl-Dickstein J. et al. *Deep Unsupervised Learning using Nonequilibrium Thermodynamics*. 2015.
22. Song Y., Ermon S. Generative Modeling by Estimating Gradients of the Data Distribution. *arXiv*. 2020.
23. Ho J., Jain A., Abbeel P. Denoising Diffusion Probabilistic Models. *arXiv*. 2020.
24. *The Synthetic Data Vault. Put synthetic data to work!* Available at: <https://sdv.dev/> (accessed 03.09.2022).
25. Schmidt L. *Data Augmentation for Tabular Data*. Available at: <https://medium.com/analytics-vidhya/data-augmentation-for-tabular-data-f75c94398c3e> (accessed 03.09.2022).
26. Kwon J., Lee Y., Lee Y., Lee S., Park J. An Algorithm Based on Deep Learning for Predicting In-Hospital Cardiac Arrest. *Journal of the American Heart Association*. 2018;7(13):e008678.
27. Kingma D.P., Ba J. Adam: A Method for Stochastic Optimization. *arXiv*. 2017.
28. Bryus P., Bryus E., Gedek P. *Prakticheskaya statistika dlya spetsialistov Data Science = Practical statistics for specialists of the Data Science*. Saint Petersburg: BKhV-Peterburg, 2021:352. (In Russ.)

Информация об авторах / Information about the authors

Владимир Иванович Горбаченко

доктор технических наук, профессор,
заведующий кафедрой компьютерных
технологий, Пензенский
государственный университет
(Россия, г. Пенза, ул. Красная, 40)

E-mail: gorvi@mail.ru

Vladimir I. Gorbachenko

Doctor of engineering sciences, professor,
head of the sub-department of computer
technologies, Penza State University
(40 Krasnaya street, Penza, Russia)

Владимир Владимирович Потанов

кандидат медицинских наук, доцент,
ассистент кафедры анестезиологии,
Донецкий национальный медицинский
университет имени М. Горького
(г. Донецк, пр. Ильича, 16)

E-mail: x3x3x23@rambler.ru

Vladimir V. Potapov

Candidate of medical sciences, associate
professor, assistant of the sub-department
of anesthesiology, Donetsk National
Medical University named after
of M. Gorky (16 Ilyicha avenue, Donetsk)

Олег Константинович Зенин

доктор медицинских наук, профессор,
профессор кафедры анатомии человека,
Медицинский институт, Пензенский
государственный университет
(Россия, г. Пенза, ул. Красная, 40)

E-mail: zen.olegz@gmail.com

Oleg K. Zenin

Doctor of medical sciences, professor,
professor of the sub-department
of human anatomy, Medical Institute,
Penza State University (40 Krasnaya
street, Penza, Russia)

Илья Сергеевич Милтых

студент, Медицинский институт,
Пензенский государственный
университет (Россия, г. Пенза,
ул. Красная, 40)

E-mail: ilia@miltykh.com

Ilya S. Miltykh

Student, Medical Institute,
Penza State University
(40 Krasnaya street, Penza, Russia)

Дмитрий Николаевич Грибков

студент, Пензенский государственный
университет (Россия, г. Пенза,
ул. Красная, 40)

E-mail: rororol23@mail.ru

Dmitriy N. Gribkov

Student, Penza State University
(40 Krasnaya street, Penza, Russia)

Авторы заявляют об отсутствии конфликта интересов / The authors declare no conflicts of interests.

Поступила в редакцию / Received 27.09.2022

Поступила после рецензирования и доработки / Revised 15.10.2022

Принята к публикации / Accepted 11.11.2022