

Кожемякин, Л. В. О качестве обучения искусственных нейронных сетей без предобработки исходных данных в условиях их ограниченного набора / Л. В. Кожемякин, А. О. Алексеев // Прикладная математика и вопросы управления. – 2023. – № 3. – С. 67–83. DOI 10.15593/2499-9873/2023.3.05

Библиографическое описание согласно ГОСТ Р 7.0.100–2018

Кожемякин, Л. В. О качестве обучения искусственных нейронных сетей без предобработки данных в условиях их ограниченного исходного набора / Л. В. Кожемякин, А. О. Алексеев. – Текст : непосредственный. – DOI 10.15593/2499-9873/2023.3.05 // Прикладная математика и вопросы управления / Applied Mathematics and Control Sciences. – 2023. – № 3. – С. 67–83.



ПРИКЛАДНАЯ МАТЕМАТИКА
И ВОПРОСЫ УПРАВЛЕНИЯ

№ 3, 2023

<https://ered.pstu.ru/index.php/amcs>



Научная статья

DOI: 10.15593/2499-9873/2023.3.05

УДК 51-76



О качестве обучения искусственных нейронных сетей без предобработки исходных данных в условиях их ограниченного набора

Л.В. Кожемякин^{1,2}, А.О. Алексеев^{1,2}

¹Пермский национальный исследовательский политехнический университет, Пермь, Российская Федерация

²Пермский центр поддержки принятия решений, Пермь, Российская Федерация

О СТАТЬЕ

Получена: 28 августа 2023
Одобрена: 10 сентября 2023
Принята к публикации:
16 сентября 2023

Финансирование

Исследование проведено при финансовой поддержке Минобрнауки России в рамках реализации программы деятельности научно-образовательного центра мирового уровня "Рациональное недропользование"

Конфликт интересов

Авторы заявляют об отсутствии конфликта интересов.

Вклад авторов

равноценен.

Ключевые слова:

искусственные нейронные сети, ограниченные данные, медицинские данные, болезнь Альцгеймера, предобработка данных, качество нейронных сетей

АННОТАЦИЯ

Рассматривается задача обучения искусственных нейронных сетей (ИНС) для диагностики болезни Альцгеймера на основе магнитной восприимчивости 10 вен головного мозга в условиях ограниченных исходных данных (81 пациент). В недавнем исследовании после сокращения числа исследуемых вен до 4, а также определения числа синаптических связей, сопоставимого с нейронной сетью, основанной на корне принятия решений, были успешно обучены полносвязные ИНС с одним и двумя скрытыми слоями. Известные результаты породили гипотезу, что обучение ИНС на основе тех же самых данных без предобработки также позволит получить качественные нейронные сети.

Для оценки качества ИНС использованы следующие показатели: средние квадратические ошибки обучения и обобщения (тестирования), коэффициенты детерминации (R^2) на множествах значений, используемых при обучении и при обобщении (тестировании), а также число обученных ИНС, удовлетворяющих определенному значению использованных выше показателей.

Для графического представления результатов многократных экспериментов по обучению и тестированию ИНС предложен новый график, позволяющий наглядно показать качество ИНС. Так, в работе демонстрируются показатели качества 800 ИНС, обученных без предварительной обработки данных, среди которых было семь нейронных сетей с одним скрытым слоем и девять нейронных сетей с двумя скрытыми слоями, для каждой из которых проводилось по 50 экспериментов. Для сравнения приводятся показатели качества 2400 ИНС, обученных с предварительной обработкой данных с такой же архитектурой: семь нейронных сетей с одним скрытым слоем и 9 нейронных сетей с двумя скрытыми слоями, для каждой из которых проводилось по 150 экспериментов.

© ПНИПУ

© Кожемякин Леонид Валерьевич – старший преподаватель кафедры «Экономика и финансы»; ведущий специалист, и.о. начальника отдела разработки и сопровождения программного обеспечения, e-mail: lvkozheymyakin@pstu.ru, ORCID: 0000-0001-5265-5787.

Алексеев Александр Олегович – кандидат экономических наук, доцент, начальник управления организации научных исследований, директор, e-mail: aoalekseev@pstu.ru, ORCID: 0000-0001-5033-6694.



Эта статья доступна в соответствии с условиями лицензии Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

Perm Polytech Style: Kozhemyakin L.V., Alekseev A.O. On the quality of training artificial neural networks without data preprocessing under conditions of their limited initial set. *Applied Mathematics and Control Sciences*. 2023, no. 3, pp. 67–83. DOI: 10.15593/2499-9873/2023.3.05

MDPI and ACS Style: Kozhemyakin, L.V.; Alekseev A.O. On the quality of training artificial neural networks without data preprocessing under conditions of their limited initial set. *Appl. Math. Control Sci.* **2023**, *3*, 67–83. <https://doi.org/10.15593/2499-9873/2023.3.05>

Chicago/Turabian Style: Kozhemyakin, Leonid V., and Aleksandr O. Alekseev. 2023. “On the quality of training artificial neural networks without data preprocessing under conditions of their limited initial set”. *Appl. Math. Control Sci.* no. 3: 67–83. <https://doi.org/10.15593/2499-9873/2023.3.05>



APPLIED MATHEMATICS
AND CONTROL SCIENCES
№ 3, 2023
<https://ered.pstu.ru/index.php/amcs>



Article

DOI: 10.15593/2499-9873/2023.3.05

UDC 51-76



On the quality of training artificial neural networks without initial data preprocessing under conditions of their limited set

L.V. Kozhemyakin^{1,2}, A.O. Alekseev^{1,2}

¹Perm National Research Polytechnic University, Perm, Russian Federation

²Perm Decision Making Support Center, Perm, Russian Federation

ARTICLE INFO

Received: 28 August 2023
Approved: 10 September 2023
Accepted for publication:
18 September 2023

Funding

This research was carried out with the financial support of the Ministry of Science and Higher Education of the Russian Federation in the framework of the program of activities of the Perm Scientific and Educational Center “Rational Subsoil Use”.

Conflicts of Interest

The authors declare no conflict of interest.

Author Contributions

equalent.

Keywords:

artificial neural networks, limited data, medical data, Alzheimer's disease, data preprocessing, quality of neural networks

ABSTRACT

The task of training artificial neural networks (ANNs) for the diagnosis of Alzheimer's disease based on the magnetic susceptibility of 10 brain veins under conditions of limited initial data (81 patients) is considered. In a recent study, fully-connected ANNs with one and two hidden layers were successfully trained in a recent study after reducing the number of veins to 4 and determining the number of synaptic connections comparable to a neural network based on a decision root. The known results gave rise to the hypothesis that training ANNs on the same data without preprocessing would also yield high-quality neural networks.

To assess the quality of ANNs, the following indicators were used: mean square errors of training and generalization (testing), coefficients of determination (R^2) on the sets of values used in training and generalization (testing), as well as the number of trained ANNs satisfying a certain value of the above indicators.

For graphical representation of the results of repeated experiments on training and testing of ANNs, a new graph is proposed to visualize the quality of ANNs. Thus, the paper demonstrates the quality indicators of 800 ANNs trained without data preprocessing, among which there were 7 neural networks with one hidden layer and 9 neural networks with two hidden layers, for each of which 50 experiments were conducted. For comparison, we present the quality indices of 2400 ANNs trained with preprocessing.

© PNRPU

© **Leonid V. Kozhemyakin** – Senior Lecturer Department of Economics and Finance, Leading Specialist and Acting Head of the Department of Software Development and Maintenance, e-mail: lvkozhemyakin@pstu.ru, ORCID: 0000-0001-5265-5787.

Aleksander O. Alekseev – CSc of Economical Sciences, Associate Professor, Head of the Administrative Directorate for Organization of Scientific Research, e-mail: aoalekseev@pstu.ru, ORCID: 0000-0001-5033-6694.



This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

Введение

Искусственные нейронные сети (ИНС) продолжают оставаться одними из наиболее популярных и эффективных средств моделирования и анализа сложных систем. Однако для достижения высокого уровня точности ИНС, несмотря на различные варианты их структур, от простого однослойного персептрона до сложной глубокой рекуррентной сети, необходим достаточно большой набор исходных данных для обучения, валидации и тестирования. На практике не всегда удается сформировать большой объем данных.

Последнее обстоятельство наиболее часто встречается в медицинских исследованиях. Ограниченность исходных медицинских данных частично объясняется этическими требованиями, так как пациенты или их законные представители должны дать согласие на использование их медицинских данных для научных целей и публикации в открытой печати. Например, в работе [1] исследователям потребовалось пять лет на сбор данных о 81 пациенте, включая 59 больных с клинически подтвержденной болезнью Альцгеймера и 22 пожилых пациентах, не имеющих проблем с когнитивными способностями.

Несмотря на сложность и высокую трудоемкость получения исходных медицинских данных, ИНС активно используются в этой области [2; 3], в частности для диагностики сердечно-сосудистых заболеваний [4–7], дисбактериоза кишечника [8], онкологических заболеваний и причин возникновения раковых опухолей [9–11], при моделировании прогноза исхода заболевания и вероятности рецидива опухоли [11–13]. В [14] предлагается использовать сверточные нейронные сети для автоматической сегментации отека у пациентов с внутримозговым кровоизлиянием. В [15] модели глубокого обучения используются для анализа осложнений отека мозга, вызванного лучевой терапией, у пациентов с внутричерепной опухолью.

В [16] исследовался вопрос об эффективности предобработки данных для обучения ИНС, которая могла бы применяться при диагностике болезни Альцгеймера. Под предобработкой подразумевался выбор наиболее существенных переменных и идентификация корня принятия решений (КПР), который, в свою очередь, представим в нейросетевой интерпретации. С помощью методов системно-когнитивного анализа (СКА) из 10 вен головного мозга было выбрано четыре, которые использовались в качестве входных сигналов для обучения нейронных сетей. В результате исследования была получена серия нейронных сетей, обладающих высокой степенью достоверности. Нейронные сети отличались друг от друга количеством скрытых слоев и количеством нейронов на скрытых слоях.

В ходе обсуждения этих результатов на научных семинарах и конференциях была высказана гипотеза о возможности обнаружения количественной связи между исследуемыми показателями без какой-либо предварительной обработки данных.

Это обстоятельство определило цель настоящей работы – оценить качество ИНС, обученных на основе исходных данных всех 10 вен головного мозга и установить связь между болезнью Альцгеймера и величиной магнитной восприимчивости (MSV) вен головного мозга (на примере 81 пациента) при сохранении конфигурации скрытых слоев ИНС.

Стоит отметить, что в работе [16] представлены только лучшие результаты обучения ИНС из многократно проведенных экспериментов с использованием нейронных сетей как с одним, так и с двумя скрытыми слоями. Поэтому в настоящей работе демонстрируются показатели качества всех 2400 ИНС, обученных с предварительной обработкой данных, среди которых было семь нейронных сетей с одним скрытым слоем и девять нейронных сетей с двумя скрытыми слоями, для каждой из которых проводилось по 150 экспериментов. Для достижения цели настоящей работы и сравнения результатов приводятся показатели качества 800 ИНС, обученных без предварительной обработки данных, среди кото-

рых было также семь ИНС с одним скрытым слоем и девять ИНС с двумя скрытыми слоями, для каждой из которых проводилось по 50 экспериментов.

Для оценки качества ИНС использованы следующие показатели: средние квадратические ошибки обучения и обобщения (тестирования), коэффициенты детерминации (R^2) на множествах значений, используемых при обучении и при обобщении (тестировании), а также число обученных ИНС, удовлетворяющих определенному значению использованных выше показателей. Для средней квадратической ошибки обучения и обобщения (тестирования) критерием качества выбрано 20 %, для $R^2 - 0,7$.

Данные и методы

Исходными медицинскими данными стали сведения о 81 пациенте, приведенные в приложении к работе [1]. Пациенты были обследованы с помощью мини-тестирования психического состояния (MMSE), Монреальского когнитивного теста (MoCA), задачи рисования часов (CDT) и шкалы активности повседневной жизни (ADL). Болезнь Альцгеймера была клинически диагностирована у 59 человек (21 мужчина и 38 женщин), включенные в группу AD (пациенты, имеющие болезнь Альцгеймера). 22 пациента (12 мужчин, 10 женщин), не имеющие проблем с когнитивными способностями, были включены в контрольную группу (группа CON).

В табл. 1 приведен фрагмент исходного набора данных, в котором содержатся значения MSV базальных вен левого и правого полушария головного мозга (L_BV и R_BV соответственно), внутренних мозговых вен левого и правого полушария головного мозга (L_ICV, R_ICV соответственно), вен таламуса левого и правого полушария головного мозга (L_TV, R_TV соответственно), перегородочных вен левого и правого полушария головного мозга (L_SV, R_SV соответственно), вен зубчатого ядра левого и правого полушария головного мозга (L_DNV, R_DNV).

Таблица 1

Фрагмент исходных данных MSV головного мозга¹

Субъект (пациент)	Группа	L_BV	R_BV	L_ICV	R_ICV	L_TV	R_TV	L_SV	R_SV	L_DNV	R_DNV
sub001	1	279	288	255	263	140	138	131	131	165	185
sub002	1	274	247	223	243	239	262	190	222	204	102
...											
sub080	0	249	232	221	232	154	142	165	151	161	143
sub081	0	300	259	295	299	216	197	163	131	162	129

Примечание: ¹ – в этой таблице в столбце «Группа» 1 обозначает AD-группу, 0 – CON-группу.

В текущей работе предложено оценить качество ИНС, обученных на основе всех исходных данных о 10 венах головного мозга (L_BV, R_BV, L_ICV, R_ICV, L_TV, R_TV, L_SV, R_SV, L_DNV, R_DNV) при сохранении конфигурации скрытых слоев ИНС и алгоритма обучения таких сетей, как и в работе [16].

Для построения архитектуры ИНС, обучения и тестирования был выбран «Нейросимулятор 5.0» [17]. Хочется отдельно отметить данный нейропакет ввиду его доступного для понимания пользовательского интерфейса и при этом достаточно хорошего функционала. Данный пакет позволяет без навыков программирования и знаний специализированных математических пакетов обработки информации (MATLAB, Maple, STATIS-

ТІСА и др.) провести нейросетевое моделирование с выбором архитектуры ИНС (рис. 1) и алгоритма обучения.

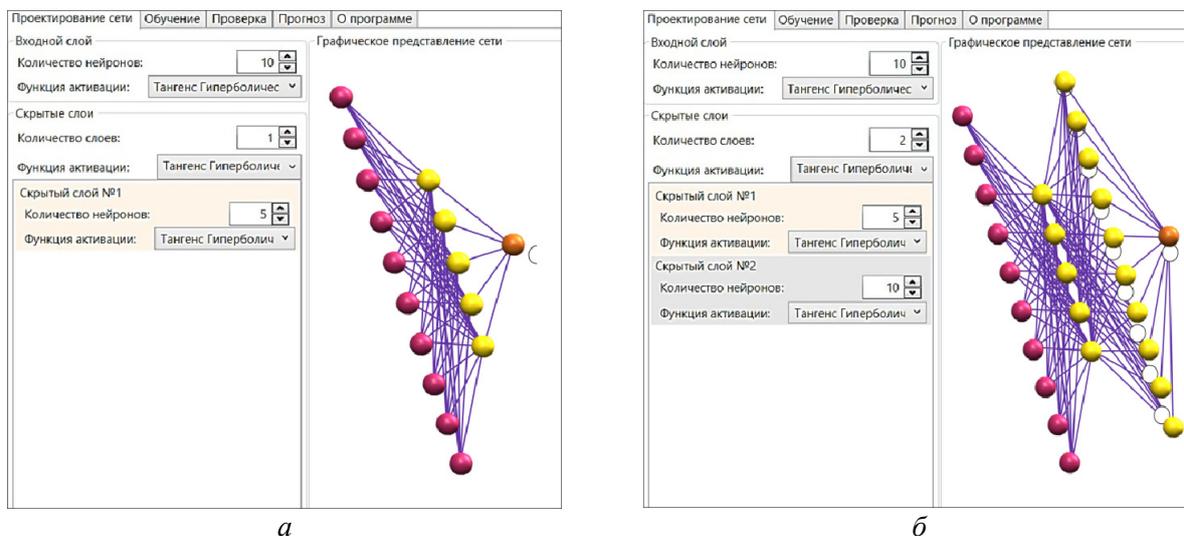


Рис. 1. Пример графического представления выбора структуры полносвязных ИНС, реализуемых в нейропакете «Нейросимулятор 5.0» с 10 входными сигналами и одним выходным сигналом: *a* – ИНС с одним скрытым слоем и 5 нейронами на скрытом слое, *б* – ИНС с двумя скрытыми слоями и 5 нейронами на первом и 10 нейронами на втором скрытых слоях

Для ИНС с одним скрытым слоем и 5 нейронами на скрытом слое (см. рис. 1, *a*) общее число нейронов равняется 16, число связей определено на уровне 55 (+6 смещений), для ИНС с двумя скрытыми слоями и 5 нейронами на первом и 10 нейронами на втором скрытых слоях общее число нейронов на всех слоях равняется 25, число связей определено на уровне 110 (+16 смещений) (см. рис. 1, *б*).

Однако «Нейросимулятор 5.0» не позволяет обеспечить неполносвязность в узлах ИНС, что является важным критерием при использовании ИНС класса Decision's-Root-based Neural Network [16]. Поэтому как отмечено в [16], ООО «Пермский центр поддержки принятия решений» разрабатывает компьютерное программное обеспечение типа SaaS, для анализа данных, в частности, для возможности нейросетевого моделирования сетей класса Decision's-Root-based Neural Network.

Строгой теории выбора оптимального количества скрытых слоев и нейронов в скрытых слоях пока не существует. Поэтому выбор структур ИНС, что в работе [16], что в текущем исследовании, не противоречит тем подходам, в результате которых чаще всего используют перцептроны, имеющие один или два скрытых слоя, причем количество нейронов в скрытых слоях обычно колеблется от $N_x/2$ до $3N_x$ [18, с. 137].

Все множество исходных данных о 81 пациенте было разбито на тренировочное (65 примеров) и тестовое (16 примеров) множества в соотношении 80 : 20 %, что является приемлемым согласно [18, с. 143].

В качестве подбора и корректировки синаптических весов $w_{ij}(t)$ нейросетевой модели (цель обучения ИНС) был выбран алгоритм упругого распространения *RPROP* – *Resilient back PROPagation*, учитывающий только знак соответствующей составляющей градиента:

$$\Delta w_{ij}(t) = -\eta(t) \operatorname{sign} \left(\frac{\partial \varepsilon(t)}{\partial w_{ij}} \right),$$

За активационную функцию нейронов скрытых слоев, входного слоя и выходного слоя был выбран тангенс гиперболический.

Результаты и их обсуждение

Результаты обучения ИНС без предварительной обработки исходных данных

Поскольку на первом шаге обучения синаптические коэффициенты генерируются случайным образом и корректируются в процессе обучения для повышения точности ИНС, качество обучения ИНС следует оценивать не только по лучшей ИНС, а смотреть на все попытки обучить ИНС в целом. Проведя серию из 800 экспериментов в попытке обучить без предварительной обработки данных ИНС на основе всех 10 вен головного мозга в условиях ограниченных данных (65 примеров), обнаружено, что ни одна из обученных на базе 10 вен головного мозга ИНС при тестировании (16 примеров) не соответствует критериям качества обучения ИНС. Ниже (табл. 2) представлены результаты обучения и тестирования ИНС с одним скрытым слоем, где использованы следующие обозначения: в столбце ИНС первое число означает число скрытых слоев, второе число – нейронов на скрытом слое.

Таблица 2

Сводная таблица результатов обучения полносвязных нейронных сетей с одним скрытым слоем и десятью входными сигналами

ИНС	Кол-во ИНС	Средняя кв. ошибка обучения (левая ось)	Средняя кв. ошибка обобщения (правая ось)	Кол-во ИНС, R^2 Train > 0,7	Кол-во ИНС, R^2 Test > 0,7	Кол-во ИНС, средн. кв. отн. Train (%) < 20 %	Кол-во ИНС, средн. кв. отн. Test (%) < 20 %
1-5	50	0,910	12,185	50	0	45	0
1-10	50	0,460	7,497	50	0	49	0
1-15	50	0,354	7,624	50	0	49	0
1-20	50	0,314	10,545	50	0	50	0
1-25	50	0,300	9,901	50	0	49	0
1-27	50	0,237	10,341	50	0	50	0
1-30	50	0,033	16,103	50	0	45	0

На рис. 2 видно, что при увеличении количества нейронов на единственном скрытом слое ИНС обучалась лучше, при этом изменение средней квадратической ошибки обобщения (тестирования) ведет себя немонотонно.

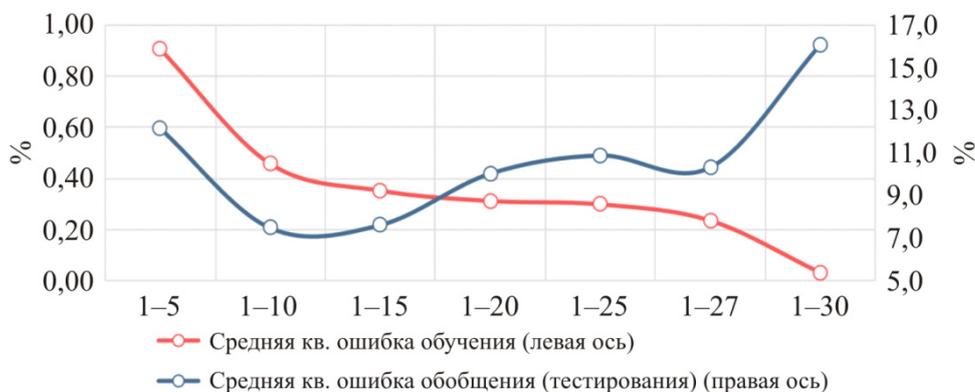


Рис. 2. Средние квадратические ошибки обучения и обобщения (тестирования) полносвязных ИНС с одним скрытым слоем с различным числом нейронов на скрытом слое

Для наглядного представления результатов многократных экспериментов по обучению и тестированию ИНС предложен новый вид столбчатой диаграммы, позволяющей графически показать качество обучения ИНС. С одной стороны, предлагаемая диаграмма одновременно демонстрирует два показателя: ось значений средней квадратической ошибки обобщения (тестирования) соответствует левой оси, ось значений коэффициента детерминации (R^2), используемых при обобщении (тестировании), соответствует правой оси. С другой стороны, диаграмма имеет цветовое форматирование: столбцы, соответствующие ИНС, удовлетворяющим критериям качества, выделяются зеленым цветом, не соответствующие – красным. Таким образом, аналитик данных сразу может видеть, насколько легко получить качественную ИНС при выбранной конфигурации.

Предлагаемые столбчатые диаграммы для ИНС с одним скрытым слоем и с различным числом нейронов на скрытых слоях приведены ниже (рис. 3).

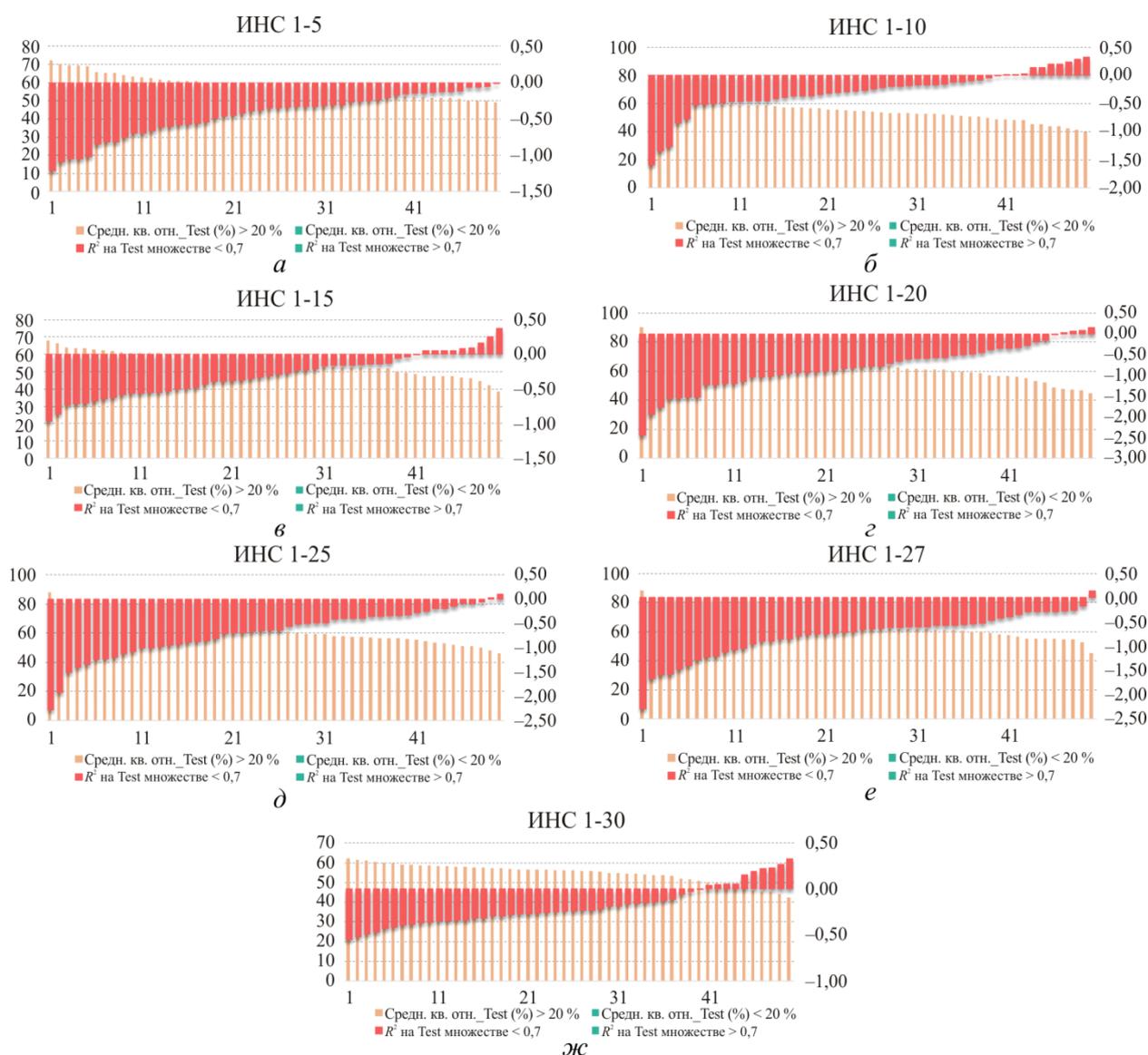


Рис. 3. Графическое представление (*a* – *ж*) результатов нейросетевого моделирования полностью связанных ИНС с одним скрытым слоем с различным числом нейронов на скрытом слое

Из рис. 3 видно, что ни одна из 350 ИНС с одним скрытым слоем не удовлетворяет заданным критериям качества. Более того, с увеличением нейронов ошибка тестирования у

ИНС 1–25, 1–27, 1–30 увеличивается, что видно из рис. 2. Здесь начинает проявляться эффект персонализации нейронов к исходным примерам, поскольку становится сопоставимо количество нейронов на скрытом слое и количество примеров для обучения.

Результаты нейросетевого моделирования второй группы ИНС с двумя скрытыми слоями и разным количеством нейронов представлены в табл. 3. В столбце ИНС и далее по тексту используются следующие обозначения: первое число определяет число скрытых слоев, второе число определяет количество нейронов на первом слое, второе число количество нейронов на втором слое. Всего было проведено девять серий экспериментов, в каждом из которых также осуществлялось по 50 попыток обучить ИНС, в которых шаг между количеством нейронами на скрытых слоях разных конфигураций ИНС был выбран на уровне пяти.

Таблица 3

Сводная таблица результатов обучения полносвязных нейронных сетей с двумя скрытыми слоями и десятью входными сигналами

ИНС	Кол-во ИНС	Средняя кв. ошибка обучения (левая ось)	Средняя кв. ошибка обобщения (правая ось)	Кол-во ИНС, R^2 Train > 0,7	Кол-во ИНС, R^2 Test > 0,7	Кол-во ИНС, средн. кв. отн. Train (%) < 20 %	Кол-во ИНС, средн. кв. отн. Test (%) < 20 %
2–5–5	50	0,087	11,852	50	0	50	0
2–5–10	50	0,133	8,796	50	0	50	0
2–5–15	50	0,148	8,360	50	0	50	0
2–10–5	50	0,013	11,580	50	0	50	0
2–10–10	50	0,002	8,572	50	0	50	0
2–10–15	50	0,003	8,165	50	0	50	0
2–15–5	50	0,007	11,230	50	0	50	0
2–15–10	50	0,005	8,219	50	0	50	0
2–15–15	50	0,004	7,811	50	0	50	0

На рисунках ниже представлены зависимости средней квадратической ошибки обучения (рис. 4, а) и обобщения (тестирования) (рис. 4, б) ИНС с двумя скрытыми слоями с различным числом нейронов на скрытом слое. Графики представляют собой зависимости изменения средней квадратической ошибки при фиксированном числе нейронов на первом скрытом слое от числа нейронов на втором скрытом слое. Например, график 2–15–N, соответствующий значению по оси абсцисс 2–N–10, соответствует структуре ИНС 2–15–10.



Рис. 4. Средние квадратические ошибки обучения и обобщения (тестирования) полносвязных ИНС с двумя скрытыми слоями в зависимости от числа нейронов на первом скрытом слое

На рис. 5 представлены столбчатые диаграммы для наглядного представления результатов 450 попыток обучить ИНС с двумя скрытыми слоями.



Рис. 5. Графическое представление (а – к) результатов нейросетевого моделирования полносвязных ИНС с двумя скрытыми слоями с различным числом нейронов на скрытых слоях

Из рисунка (см. рис. 4, б) видно, что при увеличении числа нейронов на втором скрытом слое ошибка обобщения (тестирования) снижается практически одинаково при любом числе нейронов на первом слое. При этом на рисунке (см. рис. 4, а) видно, что при недостаточном числе нейронов на первом скрытом слое ошибка может возрастать даже при обучении.

Из рис. 5 видно, что ни одна из 450 ИНС с двумя слоями не удовлетворяет заданным критериям качества, несмотря на их довольно «мягкие» значения.

Таким образом, всего из 800 ИНС, обученных без предварительной обработки данных, не удалось получить ни одной ИНС, у которой среднеквадратическая ошибка была бы меньше 20 %, а коэффициент детерминации был бы больше 70 %. Следовательно, высказанная в ходе научных семинаров гипотеза о возможности обучения качественной нейронной сети на основе всех исходных данных без их предварительной обработки не подтвердилась. С другой стороны, напомним: в работе [16] представлены результаты успешного обучения нейронных сетей на этих же исходных данных, которые были предварительно обработаны путем отбора наиболее значимых показателей и определения числа синаптических связей в нейронной сети, основанной на корне принятия решений, который идентифицировался после интервального кодирования исходных данных, что позволило определить конфигурацию полносвязных ИНС с сопоставимым числом синаптических коэффициентов.

Нельзя не забыть и тот факт, что в [16] представлены только лучшие результаты обучения, но сохраняется вопрос, насколько легко такие ИНС получить? Поскольку в настоящей работе представлена новая диаграмма, позволяющая наглядно показать результаты многократных обучений, то стоит привести все результаты, а не только лучшие. Это позволит сравнить результаты ИНС и оценить качество получаемых моделей в зависимости от того, была ли проведена предобработка данных или нет.

Результаты обучения ИНС с предварительной обработкой исходных данных

Ниже демонстрируются показатели качества 2450 обученных ИНС, среди которых было семь нейронных сетей с одним скрытым слоем (табл. 4) и девять нейронных сетей с двумя скрытыми слоями (табл. 5), для каждой из которых проводилось по 150 экспериментов.

Таблица 4

Сводная таблица результатов обучения полносвязных нейронных сетей с одним скрытым слоем с четырьмя входными сигналами

ИНС	Кол-во ИНС	Средняя кв. ошибка обучения (левая ось)	Средняя кв. ошибка обобщения (правая ось)	Кол-во ИНС, R^2 Train >0,7	Кол-во ИНС, R^2 Test >0,7	Кол-во ИНС, средн. кв. отн. Train (%) < 20 %	Кол-во ИНС, средн. кв. отн. Test (%) < 20 %
1–5	150	5,399	23,998	18	16	9	8
1–10	150	1,904	21,541	149	13	148	5
1–15	150	1,205	17,682	150	32	150	4
1–20	150	1,001	15,158	150	40	150	14
1–25	150	0,880	16,491	150	33	150	8
1–27	150	0,310	17,626	150	26	150	6
1–30	150	0,388	18,079	150	24	150	1

Из табл. 4 видно, что для каждой конфигурации одного скрытого слоя находятся ИНС, удовлетворяющие заданным в начале исследования критериям качества. Вероятность качественно обучить ИНС с одним скрытым слое выше при 20 нейронах, где также достигается минимальное значение средней квадратической ошибки при обобщении (тестировании), что можно заметить на рис. 6. Там же видно, что средняя квадратическая ошибка обучения снижается по мере роста числа нейронов, что свидетельствует уже о переобучении.

Из табл. 5 видно, что для каждой конфигурации двух скрытых слоев также находятся ИНС, удовлетворяющие заданным критериям качества. Вероятность качественно обучить ИНС с двумя скрытыми слоями выше также при 20 нейронах по количеству ИНС удовлетворяющих требованию; по коэффициенту детерминации лидирует ИНС 2–10–10, незна-

чительно ей уступает ИНС 2–5–15, которая, в свою очередь, занимает одну из лидирующих позиций по числу ИНС, удовлетворяющих критерию по среднеквадратической ошибке при обобщении (тестировании), среднее значение которой также одно из самых малых.

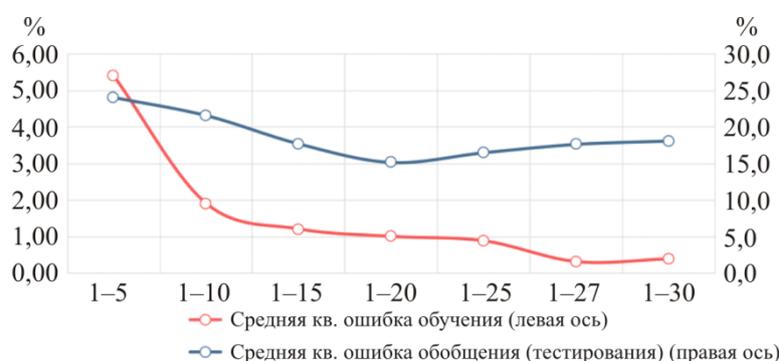


Рис. 6. Средние квадратические ошибки обучения и обобщения (тестирования) полносвязных ИНС с одним скрытым слоем и четырьмя входными сигналами в зависимости от числа нейронов на скрытом слое

Таблица 5

Сводная таблица результатов обучения полносвязных нейронных сетей с двумя скрытыми слоями с четырьмя входными сигналами

ИНС	Кол-во ИНС	Средняя кв. ошибка обучения (левая ось)	Средняя кв. ошибка обобщения (правая ось)	Кол-во ИНС, R^2 Train > 0,7	Кол-во ИНС, R^2 Test > 0,7	Кол-во ИНС, средн. кв. отн. кв. отн. Train (%) < 20 %	Кол-во ИНС, средн. кв. отн. кв. отн. Test (%) < 20 %
2–5–5	150	1,355	22,081	18	18	9	9
2–5–10	150	0,716	24,733	150	14	150	7
2–5–15	150	0,509	23,325	150	17	150	5
2–10–5	150	0,543	23,421	149	21	148	11
2–10–10	150	0,132	17,886	150	26	150	8
2–10–15	150	0,060	20,492	149	21	145	7
2–15–5	150	0,173	17,049	150	24	150	11
2–15–10	150	0,057	17,221	150	16	150	6
2–15–15	150	0,020	16,649	150	19	150	5

Сравнивая данные табл. 4 и 5, видим, что вероятность получить качественно обученную ИНС с одним скрытым слоем выше, нежели обучать ИНС с двумя скрытыми слоями. Это может свидетельствовать о качественном отборе наиболее значимых показателей.

К аналогичным выводам можно прийти, глядя на диаграммы ниже, где приведены результаты обучения ИНС с одним скрытым слоем и четырьмя входными сигналами (рис. 7), а на следующем рисунке (рис. 8) приведены девять ИНС с двумя скрытыми слоями и также четырьмя входными сигналами.

Из рис. 7 видно, что зеленые столбцы чаще встречаются на четвертой диаграмме (см. рис. 7, г), что согласуется с выводом выше о наибольшем качестве обучения ИНС при 20 нейронах на одном скрытом слое. Это свидетельствует о наглядности предложенной диаграммы, особенно когда они сопоставляются вместе.

На рис. 8 визуальнo выделяются две диаграммы, соответствующие ИНС 2–10–10 (см. рис. 8, д) и ИНС 2–5–15 (см. рис. 8, ж), среди которых преимущество ИНС 2–5–15 перед ИНС 2–10–10 еле заметно, но все же видно. Это еще раз подтверждает наглядность предложенной диаграммы.

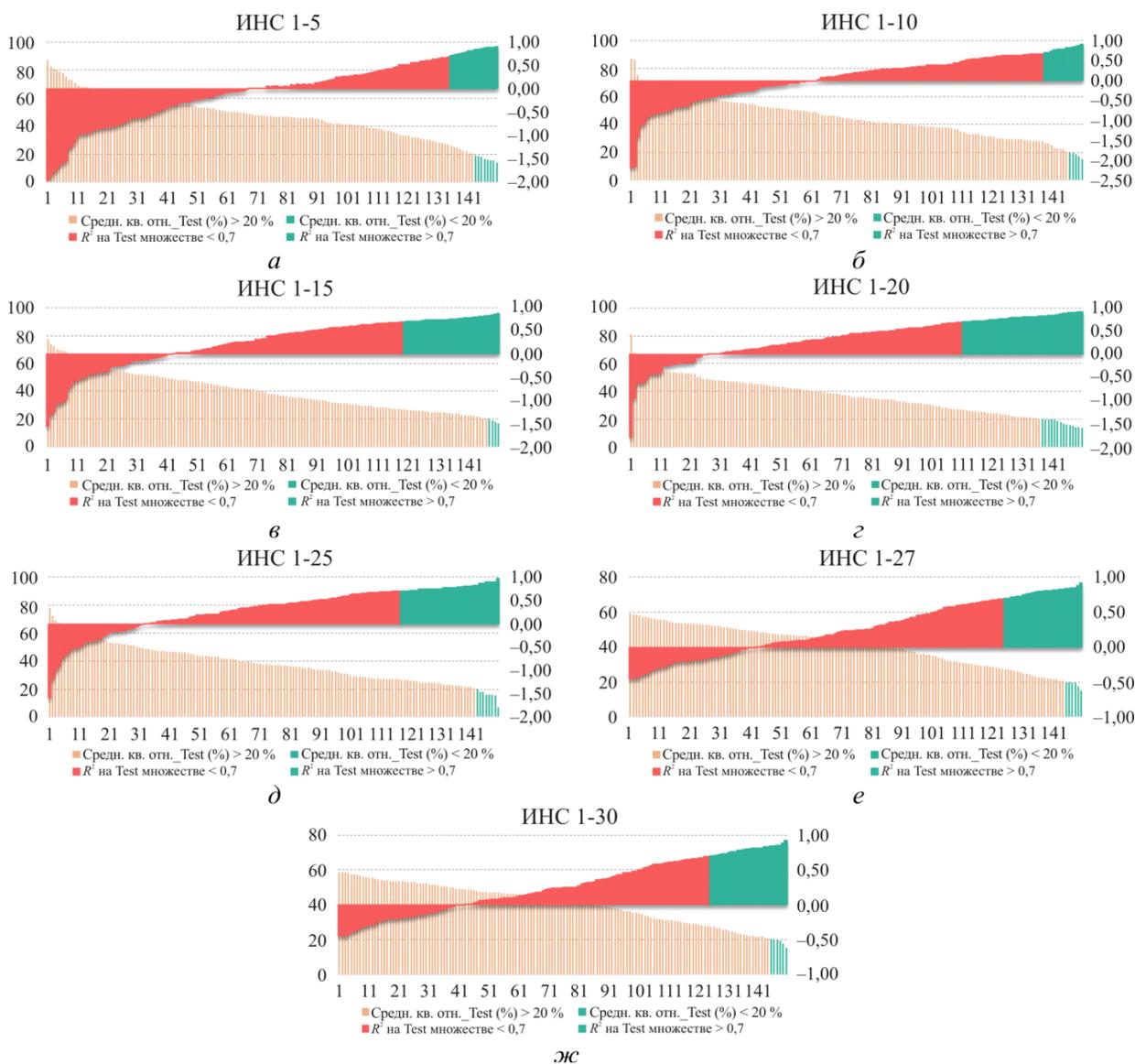


Рис. 7. Графическое представление (а – ж) результатов нейросетевого моделирования полносвязных ИНС с одним скрытым слоем в зависимости от числа нейронов на скрытом слое

Заключение

Несмотря на то, что с увеличением количества нейронов на скрытых слоях, как для ИНС с одним скрытым слоем, так и для ИНС с двумя скрытыми слоями, модель обучается лучше, что также находит подтверждение в других исследованиях, например, в [18, с. 134], говорить о точности и адекватности ИНС, обученных без предварительной обработки исходных данных, невозможно, так как средняя квадратическая ошибка обобщения (тестирования) превышает 20 %, а R^2 на тестирующем множестве не достигает уровня 0,7, и даже наоборот: в некоторых случаях увеличение количества нейронов худшим образом сказывается на средней квадратической ошибке обобщения (тестирования).

Исходя из проведенных 3200 экспериментов (2400 с предобработкой данных и 800 без предобработки), можно заключить, что экспериментально доказано, что гипотеза о наличии возможности обнаружения количественной связи между исследуемыми показателями без какой-либо предварительной обработки данных оказалась несостоятельной. Ни одна из обученных ИНС без предварительной обработки исходных данных не соответствовала установленным критериям качества.

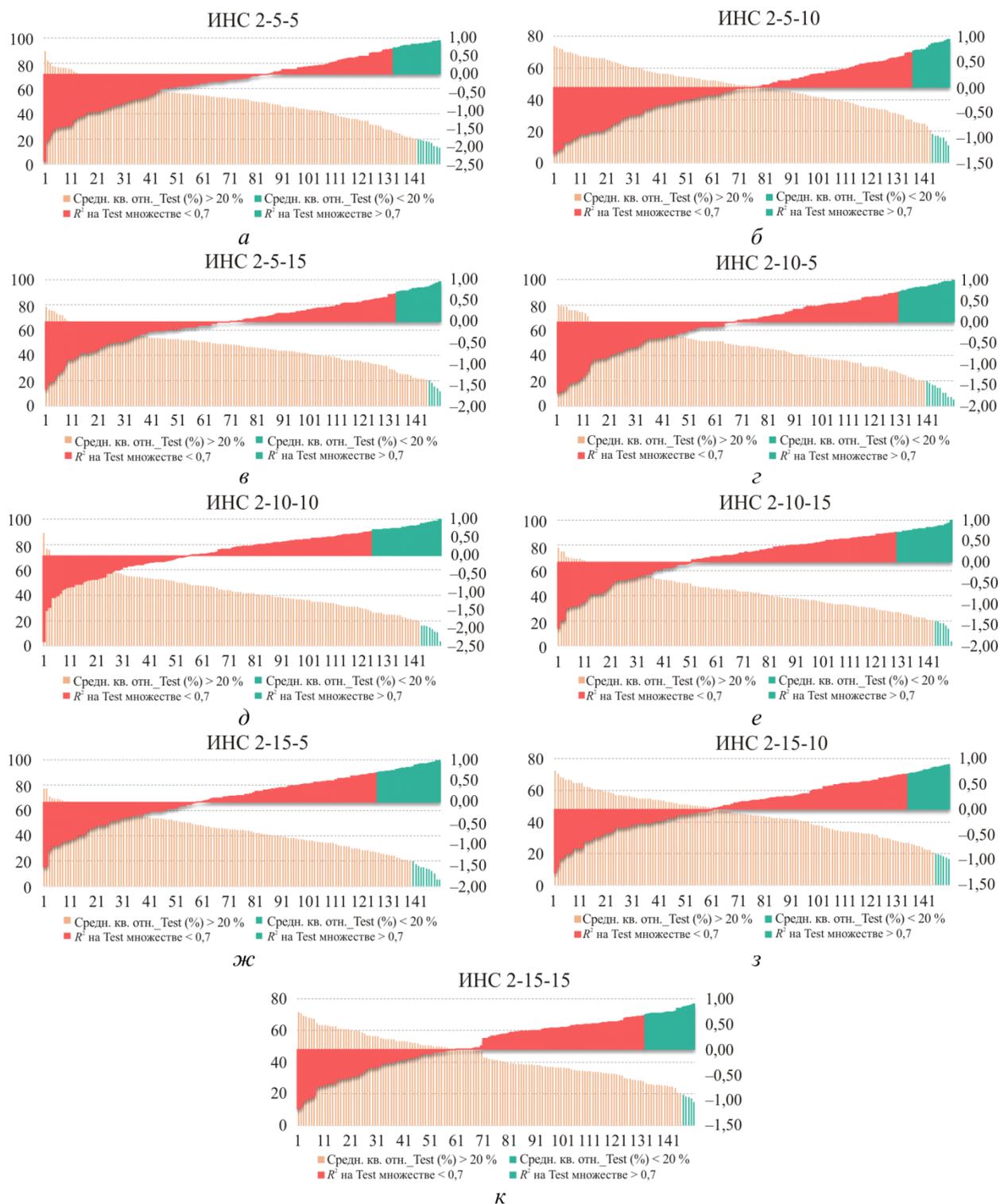


Рис. 8. Графическое представление (а – к) результатов нейросетевого моделирования полносвязных ИНС с двумя скрытыми слоями с различным числом нейронов на скрытых слоях

При всех конфигурациях ИНС с четырьмя входными сигналами (см. табл. 4, 5) удается обучить ИНС с заданными критериям качества, что в графическом представлении отображается зелеными столбцами на диаграммах (см. рис. 7, 8), что подтверждает необходимость предварительной обработки данных и сокращение числа учитываемых переменных.

Для достижения высокой точности результатов нейросетевого моделирования необходимо экспериментально подбирать конфигурацию нейронной сети. В эти эксперименты

дополнительно стоит включать такие этапы, как контроль переобучения сети (например, применение методов регуляризации весов или кросс-валидации), выбор алгоритма обучения (например, с применением адаптивного параметра обучения – скорости обучения для каждого веса или без адаптивного параметра), нейросетевую фильтрацию [19] для исключения выбросов в статистической информации, и многие другие тонкости, что является очень трудоемким исследованием.

Необходимость в предварительной обработке исходного набора данных имеет свое подтверждение. Методы и алгоритмы предобработки данных успешно себя показали в работе [16], когда были получены ИНС приемлемого уровня в тех же экспериментальных условиях, что и ИНС, представленные в текущей работе, но обученные сразу на исходном наборе данных без какой-либо предобработки данных.

Список литературы

1. Correlation between cerebral venous oxygen level and cognitive status in patients with Alzheimer's disease using quantitative susceptibility mapping / Y. Liu, J. Dong, Q. Song, N. Zhang, W. Wang, B. Gao, S. Tian, C. Dong, Z. Liang, L. Xie, Y. Miao / *Frontiers in Neuroscience*. – 2021. – Vol. 14. – Art. 570848. DOI 10.3389/fnins.2020.570848

2. Capabilities of neural network technologies for extracting new medical knowledge and enhancing precise decision making for patients / L.N. Yasnitsky, A.A. Dumler, N.A. Uteva, F.M. Cherepanov, V.L. Yasnitsky / *Expert Review of Precision Medicine and Drug Development*. – 2022. – Vol. 7, № 1. – P. 70–78.

3. Yasnitsky L.N. Artificial intelligence and medicine: history, current state, and forecasts for the future // *Current Hypertension Reviews*. 2020. – Vol. 16, № 3. – P. 210–215. DOI 10.2174/1573402116666200714150953

4. Prototyping neural networks to evaluate the risk of adverse cardiovascular outcomes in the population / L.A. Bogdanov, E.A. Komosky, V.V. Voronkova, D.E. Tolstosheev, G.V. Martsenyuk, A.S. Agienko, E.V. Indukaeva, A.G. Kutikhin, D.P. Tsygankova / *Fundamental and Clinical Medicine*. – 2021. – Vol. 6, № 4. – P. 67–81.

5. Kilic A. Artificial intelligence and machine learning in cardiovascular health care // *The Annals of Thoracic Surgery*. – 2020. – Vol. 109, № 5. – P. 1323–1329. DOI 10.1016/j.athoracsur.2019.09.042

6. Ясницкий Л.Н., Думлер А.А., Черепанов Ф.М. Возможности применения методов искусственного интеллекта для диагностики и прогнозирования сердечно-сосудистых заболеваний // *Искусственный интеллект: философия, методология, инновации. Материалы Пятой Всероссийской конференции студентов, аспирантов и молодых ученых / Московский государственный институт радиотехники, электроники и автоматики (технический университет), Москва, 9–11 ноября 2011 г.* – М.: МГТУ МИРЭА, 2011. – С. 146–148.

7. Хливненко Л.В. Прогнозирование исходов мерцательной аритмии с помощью искусственной нейронной сети // *Информатика: проблемы, методология, технологии: материалы 7-й Междунар. науч.-метод. конф. / Воронежский государственный университет.* – Воронеж, 2006. – С. 467–471.

8. Солодуха Т.В. Применение нейронных сетей для исследования дисбактериоза [Электронный ресурс]. – URL: <http://masters.donntu.edu.ua/publ2002/kita/solodukha.pdf> (дата обращения: 10.07.2023).

9. Multispectral imaging and artificial neural network: mimicking the management decision of the clinician facing pigmented skin lesions / M. Carrara, A. Bono, C. Bartoli, A. Colombo, M. Lualdi, D. Moglia, N. Santoro, E. Tolomio, S. Tomatis, G. Tragni, M. Santinami, R. Marchesini / *Physics in Medicine & Biology*. – 2007. – Vol. 52, № 9. – P. 2599–2613.
10. Predicting continuous values of prognostic markers in breast cancer from microarray gene expression profiles / S.K. Gruvberger-Saal, P. Edén, M. Ringnér, B. Baldetorp, G. Chebil, A. Borg, M. Fernö, C. Peterson, P. S. Meltzer / *Molecular Cancer Therapeutics*. – 2004. – Vol. 3, № 2. – P. 161–168.
11. Combined neural network and decision trees model for prognosis of breast cancer relapse / J.M. Jerez-Aragonés, J.A. Gómez-Ruiz, G. Ramos-Jiménez, J. Muñoz-Pérez, E.A. AlbaConejo / *Artificial intelligence in medicine*. – 2003. – Vol. 27, № 1. – P. 45–63.
12. Artificial neural networks applied to survival prediction in breast cancer / J. Lundin, H.B. Burke, S. Toikkanen, L. Pylkkänen, H. Joensuu // *Oncology*. – 1999. – Vol. 57, № 4. – P. 281–286.
13. Hybrid 3D/2D convolutional neural network for hemorrhage evaluation on head CT / P.D. Chang, E. Kuoy, J. Grinband, B.D. Weinberg, M. Thompson, R. Homo, J. Chen, H. Abcede, M. Shafie, L. Sugrue, C.G. Filippi, M.Y. Su, W.Yu, C. Hess, D. Chow // *AJNR Am. J. Neuroradiol.* – 2018. – Vol. 9. – P. 1609–1616. DOI 10.3174/ajnr. A5742
14. Using deep learning models to analyze the cerebral edema complication caused by radiotherapy in patients with intracranial tumor / P.J. Chao, L. Chang, C.L. Kang, C.H. Lin, C.S. Shieh, J.M. Wu, C.D. Tseng, I.H. Tsai, H.C. Hsu, Y.J. Huang, T.F. Lee // *Scientific Reports*. – 2022. – Vol. 12, № 1. – Art. 1555.
15. End-To-End Alzheimer's Disease Diagnosis and Biomarker Identification / S. Esmailzadeh, D.I. Belivanis, K.M. Pohl, E. Adeli // *Machine Learning Medical Imaging*. – 2018. – Vol. 11046. – P. 337–345.
16. Data Preprocessing and Neural Network Architecture Selection Algorithms in Cases of Limited Training Sets – On an Example of Diagnosing Alzheimer's Disease / A.A. Alekseev, L.V. Kozhemyakin, V.N. Nikitin, J.A. Bolshakova // *Algorithms*. – 2023. – Vol. 16. – Art. 219. DOI 10.3390/a16050219.
17. Черепанов Ф. М., Ясницкий Л.Н. Нейросимулятор 5.0. Свидетельство о государственной регистрации программы для ЭВМ № 2014618208. Заявка Роспатент № 2014614649. Зарегистрировано в Реестре программ для ЭВМ 12 августа 2014 г.
18. Ясницкий Л.Н. Искусственный интеллект. Элективный курс: учебное пособие. – М.: БИНОМ. Лаборатория знаний, 2011. – 197 с.
19. Черепанов Ф.М., Ясницкий Л.Н. Нейросетевой фильтр для исключения выбросов в статистической информации // *Вестник Пермского университета. Серия: Математика. Механика. Информатика*. – 2008. – № 4. – С. 151–155.

References

1. Liu Y., Dong J., Song Q., Zhang N., Wang W., Gao B., Tian S., Dong C., Liang Z., Xie L., Miao Y. Correlation between cerebral venous oxygen level and cognitive status in patients with Alzheimer's disease using quantitative susceptibility mapping. *Frontiers in Neuroscience*, 2021, vol. 14, art. 570848. DOI: 10.3389/fnins.2020.570848.
2. Yasnitsky L.N., Dumler A.A., Uteva N.A., Cherepanov F.M., Yasnitsky V.L. Capabilities of neural network technologies for extracting new medical knowledge and enhancing precise de-

cision making for patients. *Expert Review of Precision Medicine and Drug Development*, 2022, vol. 7, no. 1, pp. 70–78.

3. Yasnitsky L.N., Artificial intelligence and medicine: history, current state, and forecasts for the future. *Current Hypertension Reviews*, 2020, vol. 16, no. 3, pp. 210–215. DOI: 10.2174/1573402116666200714150953.

4. Bogdanov L.A., Komossky E.A., Voronkova V.V., Tolstosheev D.E., Martsenyuk G.V., Agienko A.S., Indukaeva E.V., Kutikhin A.G., Tsygankova D.P. Prototyping neural networks to evaluate the risk of adverse cardiovascular outcomes in the population. *Fundamental and Clinical Medicine*, 2021, vol. 6, no. 4, pp. 67–81.

5. Kilic A. Artificial intelligence and machine learning in cardiovascular health care. *The Annals of Thoracic Surgery*, 2020, vol. 109, no. 5, pp. 1323–1329. DOI: 10.1016/j.athoracsur.2019.09.042

6. Yasnitsky L.N., Dumler A. A., Cherepanov F. M. Vozmozhnosti primeneniia metodov iskusstvennogo intellekta dlia diagnostiki i prognozirovaniia serdechno-sosudistykh zabolevanii [The possibilities of using artificial intelligence methods for the diagnosis and prediction of cardiovascular diseases]. *Iskusstvennyi intellekt: filozofia, metodologiya, innovatsii. Materialy Piatoi Vserossiiskoi konferentsii studentov, aspirantov i molodykh uchenykh, 9-11 November 2011*, Moscow, Moscow State Institute of Radio Engineering, Electronics and Automation (Technical University), 2011, pp. 146 – 148.

7. Khlivnenko L.V. Prognozirovanie iskhodov mertsatel'noi aritmii s pomoshch'iu iskusstvennoi neuronnoi seti [Prediction of atrial fibrillation outcomes using an artificial neural network]. *Informatika: problemy, metodologiya, tekhnologii: materialy sed'moi mezhdunarodnoi nauchno-metodologicheskoi konferentsii, 9-10 February 2006*, Voronezh, Voronezh State University, 2006, pp. 467–471.

8. Solodukha T.V., Malt T.V. Application of neural networks for the study of dysbiosis, available at: <http://masters.donntu.edu.ua/publ2002/kita/solodukha.pdf>.

9. Carrara M., Bono A., Bartoli C., Colombo A., Lualdi M., Moglia D., Santoro N., Tolomio E., Tomatis S., Tragni G., Santinami M., Marchesini R. Multispectral imaging and artificial neural network: mimicking the management decision of the clinician facing pigmented skin lesions. *Physics in Medicine & Biology*, 2007, vol. 52, no. 9, pp. 2599–2613.

10. Gruvberger-Saal S.K., Edén P., Ringnér M., Baldetorp B., Chebil G., Borg A., Fernö M., Peterson C., Meltzer P.S. Predicting continuous values of prognostic markers in breast cancer from microarray gene expression profiles. *Molecular Cancer Therapeutics*, 2004, vol. 3, no. 2, pp. 161–168.

11. Jerez-Aragónés J.M., Gómez-Ruiz J.A., Ramos-Jiménez G., Muñoz-Pérez J., Alba-Conejo E.A. Combined neural network and decision trees model for prognosis of breast cancer relapse. *Artificial intelligence in medicine*, 2003, vol. 27, no. 1, pp. 45-63.

12. Lundin J., Burke H.B., Toikkanen S., Pylkkänen L., Joensuu H. Artificial neural networks applied to survival prediction in breast cancer. *Oncology*, 1999, vol. 57, no. 4, pp. 281-286.

13. Chang P.D., Kuoy E., Grinband J., Weinberg B.D., Thompson M., Homo R., Chen J., Abcede H., Shafie M., Sugrue L., Filippi C.G., Su M.Y., Yu W., Hess C., Chow D. Hybrid 3D/2D convolutional neural network for hemorrhage evaluation on head CT. *AJNR Am. J. Neuroradiol*, 2018, vol. 9, pp. 1609–1616. DOI: 10.3174/ajnr. A5742.

14. Chao P.J., Chang L., Kang C.L., Lin C.H., Shieh C.S., Wu J.M., Tseng C.D., Tsai I.H., Hsu H.C., Huang Y.J., Lee T.F. Using deep learning models to analyze the cerebral edema com-

plication caused by radiotherapy in patients with intracranial tumor. *Scientific Reports*, 2022, vol. 12, no. 1, art. 1555.

15. Esmailzadeh S., Belivanis D.I., Pohl K.M., Adeli E. End-To-End Alzheimer's Disease Diagnosis and Biomarker Identification. *Machine Learning Medical Imaging*, 2018, vol. 11046, pp. 337-345.

16. Alekseev A.A., Kozhemyakin L.V., Nikitin V.N., Bolshakova J.A. Data Preprocessing and Neural Network Architecture Selection Algorithms in Cases of Limited Training Sets – On an Example of Diagnosing Alzheimer’s Disease. *Algorithms*, 2023, vol. 16, art. 219. DOI: 10.3390/a16050219.

17. Cherepanov, F.M.; Yasnitsky, L.N. “Neurosimulator 5.0” [Computer Program RU 2014618208, 12 July 2014]. *Programmy Dlia Evm. Bazy Danykh. Topologii Integral’Nykh Mikroskhem* [Computer Programs. Database. Topologies of Integrated Circuits].

18. Yasnitsky L.N. *Iskusstvennyj intellekt. Jelevtivnyj kurs : uchebnoe posobie* [Artificial intelligence. Elective course : study guide]. Moscow, BINOM. Laboratorija znaniy, 2011, 197 p.

19. Cherepanov F. M., Yasnitskiy L. N. Nejrsetevoj fil'tr dlja iskljuchenija vybrosov v statisticheskoj informacii [Neural network filter to exclude outliers in statistical information]. *Bulletin of the Perm university: Mathematics. Mechanics. Information science*, 2008, no. 4, pp. 151–155.