



<http://dx.doi.org/10.35596/1729-7648-2025-23-5-99-104>

УДК 004.032.26

## АДАПТАЦИЯ НЕЙРОСЕТЕВОГО КЛАССИФИКАТОРА К НЕУСТОЙЧИВЫМ ВХОДНЫМ ДАННЫМ

В. В. МАЦКЕВИЧ, Ц. ГО

*Белорусский государственный университет (Минск, Республика Беларусь)*

**Аннотация.** Сегодня задачи классификации решаются, как правило, с применением нейронных сетей. В то же время при выборе архитектуры сети особое внимание уделяется сжимающим слоям, а архитектура многослойного персептрона выбирается интуитивно, хотя качество решения во многом зависит от вида разделяющей поверхности. В статье предлагается алгоритм проверки эффективности архитектуры многослойного персептрона на основе анализа свойств входных данных. Алгоритм основан на специальном, разработанном авторами, числовом показателе эффективности архитектуры нейронной сети – коэффициенте перекоса. Коэффициент рассчитывается на основе матрицы несоответствий, что не требует большого объема вычислений. Экспериментально показано, что выбор правильной архитектуры классификатора может повысить качество решения на 50 %.

**Ключевые слова:** нейронные сети, многослойный персептрон, архитектура, классификация, обучение, входные данные, сжатие, оптимизационная задача, пространство решений, метод отжига, метод градиентного спуска.

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

**Для цитирования.** Мацкевич, В. В. Адаптация нейросетевого классификатора к неустойчивым входным данным / В. В. Мацкевич, Ц. Го // Доклады БГУИР. 2025. Т. 23, № 5. С. 99–104. <http://dx.doi.org/10.35596/1729-7648-2025-23-5-99-104>.

## NEURAL NETWORK CLASSIFIER ADAPTATION TO UNSTABLE INPUT DATA

VADIM V. MATSKEVICH, JIRAN GUO

*Belarusian State University (Minsk, Republic of Belarus)*

**Abstract.** Today, classification problems are typically solved using neural networks. When choosing a network architecture, particular attention is paid to compression layers, while the multilayer perceptron architecture is chosen intuitively, although the quality of the solution largely depends on the type of separating surface. This article proposes an algorithm for testing the effectiveness of a multilayer perceptron architecture based on an analysis of input data properties. The algorithm is based on a special numerical metric for assessing the effectiveness of a neural network architecture, the skew coefficient, developed by the authors. The coefficient is calculated based on the confusion matrix, which does not require a large amount of calculations. Experiments have shown that choosing the right classifier architecture can improve solution quality by 50 %.

**Keywords:** neural networks, multilayer perceptron, architecture, classification, training, input data, compression, optimization problem, solution space, annealing method, gradient descent method.

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

**For citation.** Matskevich V. V., Guo J. (2025) Neural Network Classifier Adaptation to Unstable Input Data. *Doklady BGUIR*. 23 (5), 99–104. <http://dx.doi.org/10.35596/1729-7648-2025-23-5-99-104> (in Russian).

## Введение

В настоящее время нейронные сети широко применяются для решения различных прикладных задач [1]. По мере роста мощностей вычислительных устройств расширяется класс решаемых задач, однако львиную их долю по-прежнему занимают задачи классификации. Несмотря на различие используемых для их решения архитектур сетей, в конце каждой из них расположен многослойный персептрон, который и выполняет роль классификатора [2, 3]. Его значимость на практике занижается, поскольку повышенное внимание при решении подобных задач уделяется проектировке архитектуры сжимающих слоев и организации процесса обучения громоздкой сети. Однако, как известно, архитектура нейронной сети, в том числе и классификатора, играет решающую роль в качестве конечного решения [4]. На практике же архитектуру многослойного персептрона выбирают интуитивно по размерам сжатых входных образов и количеству классов входных данных. При этом архитектура персептрона остается неизменной на протяжении всего решения задачи [5]. Однако входные данные большой размерности содержат множество различных мелких объектов, которые могут изменить свойства разделяющей поверхности классов [6], и, кроме того, – разнообразные шумы, способные сохраниться при сжатии, тем самым повышая требования к архитектуре классификатора.

В статье предлагается алгоритм проверки эффективности архитектуры нейросетевого классификатора на основе анализа свойств входных данных. Алгоритм не зависит от типа и размерности входных данных, а также инвариантен к количеству классов, архитектуре сжимающих слоев и алгоритмам предварительной обработки данных. При этом он не требует большого объема вычислений. Данные свойства алгоритма делают его универсальным для оценки архитектуры многослойного персептрона. Эффективность представленного в статье алгоритма проверена экспериментально на примере решения прикладной задачи.

## Многомерная матрица несоответствий и пространство решений

Обучение нейронных сетей является задачей условной оптимизации, пространство решений которой обладает рядом особенностей:

- оптимальное решение в этом пространстве очень неустойчивое, иными словами, незначительное изменение значения любого из параметров нейронной сети приводит к совершенно иному результату;
- пространство имеет большую размерность, так как современные нейронные сети содержат большое количество настраиваемых параметров;
- пространство обладает свойством своеобразной неоднородности. Известно, что результатом работы сети является вектор значений нейронов выходного слоя. Для разных входных данных должны быть различные выходные значения. Учитывая специфику функционирования нейронной сети, можно показать, что оптимальные значения параметров будут находиться в достаточно узких диапазонах, а само оптимальное решение в силу особенности пространства – в некотором подпространстве малой мощности.

Входное значение в каждом нейроне – это сумма произведений весов и выходных значений предыдущего слоя. Поскольку веса при обработке данных зафиксированы, а входные данные имеют (если задачей не оговорено иное) многомерное равномерное распределение, каждое слагаемое является равномерно распределенной случайной величиной на некотором отрезке. По мере роста числа слагаемых (размерности входных данных) из закона больших чисел следует, что сумма сходится к нормальному распределению. А нормальное распределение, как известно, имеет маленькую дисперсию. Поэтому при разных значениях входных данных значения сумм будет трудно различить. Это приводит к необходимости использовать веса с большим соотношением. Из чего следует, что достаточно их расположить в узком интервале значений с нулем в центре. Как правило, при обучении нейронных сетей в задачах классификации точка оптимального решения является неустойчивой, поскольку в ее окрестности много точек с различными значениями целевой функции. Отсюда вывод – для любого параметра существует относительно небольшой отрезок значений, вероятность нахождения оптимального решения в котором существенно выше, чем на оставшемся диапазоне его изменений. Для формализации данного факта рассмотрим следующие определения.

Пространство решений множества  $\Omega$  назовем неоднородным, если существует его разбиение на подмножества  $\Omega_1, \Omega_2$  такие, что:

$$\begin{cases} \Omega = \Omega_1 \cup \Omega_2, \Omega_1 \cap \Omega_2 = \emptyset; \\ \forall \Omega_3, \Omega_4 : |\Omega_3| = |\Omega_4|, \Omega_3 \subseteq \Omega_1, \Omega_4 \subseteq \Omega_2; \\ P(x^* \in \Omega_3) \gg P(x^* \in \Omega_4), \end{cases}$$

где  $x^*$  – оптимальное решение.

Отношение вероятности нахождения оптимального решения в некотором подмножестве множества  $\Omega$  к мощности подмножества назовем плотностью этого подмножества.

Из введенных определений следует вполне очевидное утверждение, что если множество  $\Omega$  обладает свойством неоднородности, то в нем существуют подмножества с заметно различающимися плотностями. То есть пространство решений  $\Omega$  в задаче обучения нейронных сетей обладает свойством своеобразной неоднородности.

*Замечание.* Как было показано в [7], учет данного свойства позволяет существенно ускорить обучение нейронных сетей.

Теперь опишем пространство решений для задачи классификации. Пусть для ее решения используется нейросетевой классификатор, а именно – многослойный персептрон. Пусть в исходной задаче задано  $n$  классов входных данных. Классификацию будем производить стандартным образом: какой нейрон выходного слоя обладает наибольшим выходным значением, такому классу и принадлежит обрабатываемый объект.

Пусть задана квадратная матрица  $n \times n$ , где  $a_{ij}$  – количество раз, когда  $i$ -й класс был выбран  $j$ -м нейроном выходного слоя. Эта матрица также известна как многомерная матрица несоответствий [8]. Она применяется для оценки качества конечного решения. Сумма элементов главной диагонали определяют количество правильных ответов сети. Сумма всех остальных элементов матрицы определяет количество неверных классификаций. Данная матрица также показывает, для каких классов были допущены ошибки, сколько раз и какой нейрон сети был выбран сетью для конкретного класса ошибочно.

Воспользуемся данными свойствами матрицы и покажем, что ее можно использовать для численной оценки эффективности архитектуры нейронной сети. Сумма элементов одного столбца матрицы определяет количество «срабатываний» одного нейрона выходного слоя сети для различных входных данных. Воспользуемся последним фактом и на его основе построим алгоритм проверки эффективности архитектуры многослойного персептрона. Для этого введем следующее определение.

*Определение.* Коэффициентом перекоса  $q$  многомерной матрицы несоответствий  $A$  называется величина, определяющая отношение наибольшей суммы элементов столбца матрицы к сумме всех элементов:

$$q = \frac{\max_{j=1, n} \sum_{i=1}^n a_{ij}}{\sum_{i=1}^n \sum_{j=1}^n a_{ij}}. \quad (1)$$

Величина  $q$  для задачи из  $n$  классов может варьироваться от  $1/n$  до 1. Чем меньше значение этого коэффициента, тем, очевидно, лучше архитектура нейронной сети. Значение  $q = 1$  указывает на то, что нейронная сеть полностью «невосприимчива» к входным данным из различных классов.

### Алгоритм проверки эффективности архитектуры многослойного персептрона

На основе многомерной матрицы несоответствий предлагается эвристический алгоритм проверки эффективности архитектуры, который представляет собой итерационный процесс с проверкой критерия оптимальности в конце каждой итерации.

Шаг 0. Определение архитектуры многослойного персептрона. На данном шаге, исходя из задачи, эксперт определяет количество слоев сети, количество нейронов в каждом из слоев и функ-

ции активации нейронов. Также устанавливается счетчик числа замеров  $t$ , который принимается равным нулю.

Шаг 1. Генерация начального приближения. Инициализируются параметры обучаемого многослойного персептрона. Например, веса сети могут задаваться реализациями равномерно распределенных случайных величин на отрезке с центром в нуле, а все остальные параметры зануляются.

Шаг 2. Построение многомерной матрицы несоответствий. Многослойный персептрон обрабатывает входные данные из обучающей выборки, а в матрицу записывается статистика откликов сети с учетом меток классов.

Шаг 3. Вычисление коэффициента перекоса  $q$  по формуле (1).

Шаг 4. Повторение оценки. Счетчик числа замеров увеличивается на единицу. Если счетчик равен  $T_2$  – переход на шаг 5. В противном случае – переход на шаг 1.

Шаг 5. Проверка критерия оптимальности. Если среднее значение  $q$  меньше заданного порога  $T_1$ , то алгоритм завершается, в противном случае – переход на шаг 0.

*Замечание 1.* Чем больше классов входных данных в задаче, тем ниже рекомендуется выставить значение параметра  $T_1$ . Это можно объяснить тем, что вероятность «неудачной» генерации значений параметров при росте числа классов снижается. Рекомендуемое значение параметра для двух классов 0,9.

*Замечание 2.* Величину  $T_2$  рекомендуется выставить 20. Большее значение  $T_2$  существенно замедляет алгоритм, а меньшее приводит к большой погрешности оценки. «Неудачные» инициализации с большой долей вероятности указывают на проблемы с архитектурой сети.

*Замечание 3.* В случае использования большой обучающей выборки в целях экономии вычислительных ресурсов алгоритм допускает использование фрагмента обучающей выборки. Однако фрагмент должен содержать данные всех возможных классов для оценки эффективности архитектуры.

### Методика проведения эксперимента

В рамках экспериментов рассмотрим прикладную задачу детектирования внепогодных изменений поверхности Земли по результатам анализа спутниковых снимков [9]. Разобьем спутниковые снимки на блоки  $20 \times 20$  пикселей и решим задачу независимо для каждого блока. С учетом масштаба снимков блоки указанного размера вмещают наименьший возможный объект исследования. Для простоты будем решать данную задачу с использованием простого многослойного персептрона. Поскольку по условию задачи непогодные изменения определяются на основе пары снимков одной и той же местности в разные моменты времени, на вход нейронной сети поступают два цветных изображения разрешением  $20 \times 20$  пикселей. Следовательно, входной слой будет содержать 2400 нейронов, а выходной – всегда два с функцией активации софтмакс (один нейрон указывает на наличие непогодных изменений, второй – на их отсутствие). Частично зафиксируем архитектуру, определив, что в обучаемой нейронной сети один скрытый слой.

Теперь начнем перебирать различные архитектуры многослойного персептрона и проверять их эффективность. Будем вычислять средний коэффициент перекоса для каждой архитектуры и обучать нейронные сети с различными значениями этого перекоса для всех рассматриваемых архитектур. В экспериментах проверялись различные виды ошибок, допущенных при проектировании архитектуры сети, и их влияние на конечный результат и коэффициент перекоса.

Количество нейронов в скрытом слое и их функция активации в экспериментах варьировались. Для оценки качества решения использовались относительные величины. За единицу брали нейронную сеть с наилучшей обобщающей способностью. Использовали алгоритм обучения на основе больцмановского отжига [7], реализованный в рамках программного комплекса для обучения глубоких нейронных сетей [10].

В целях экономии времени без существенной потери точности результатов количество итераций было ограничено величиной 0,5 млн. В данной задаче это никак не повлияло на качество решения, поскольку дальнейшие итерации не приводили к существенному улучшению решения. Результаты экспериментов и исследуемые архитектуры отображены в табл. 1.

**Таблица 1.** Оценка эффективности архитектуры нейронной сети  
**Table 1.** Neural network architecture efficiency evaluation

Архитектура нейронной сети	Средний коэффициент перекоса	Качество конечного решения
16, биполярная сигмоидная	0,918	0,865
16, сигмоидная	0,906	0,867
16, ReLU	0,976	0,663
32, биполярная сигмоидная	0,886	0,897
32, сигмоидная	0,862	0,944
32, ReLU	0,920	0,795
64, биполярная сигмоидная	0,853	0,952
64, сигмоидная	0,785	1,000
64, ReLU	0,891	0,883

*Источник:* собственная разработка.

### Результаты исследований и их обсуждение

По результатам экспериментов можно отметить, что чем меньше среднее значение коэффициента перекоса, тем выше эффективность архитектуры нейронной сети. Также из табл. 1 видно, что если среднее значение коэффициента перекоса превышает 0,9, то архитектура крайне неэффективна.

Данный алгоритм проверяет не только эффективность архитектуры нейронной сети, но и алгоритм генерации начального приближения в задаче обучения. Если обучение производилось с начального приближения, для которого  $q = 1$  (т. е. располагалось за пределами области повышенной плотности), то качество решения уменьшалось более чем на 10 %. Если при проверке эффективности среднее значение коэффициента  $q$  для различных архитектур было близко или равно единице, то это означало, что начальные параметры нейронной сети задавались неверно. Отдельно стоит отметить что разница в качестве решения между наилучшей архитектурой и наихудшей превышает 50 %, что показывает, насколько важен выбор правильной архитектуры нейронной сети. Выбор эффективной архитектуры многослойного персептрона без его обучения позволяет не только повысить качество решения, но и существенно ускорить разработку систем на их основе, поскольку в этом случае экономится время, которое могло быть потрачено на обучение неэффективных архитектур.

### Заключение

1. Проведены исследования пространства решения для оптимизационной задачи обучения нейронных сетей. Показано, что решение обладает свойством своеобразной неоднородности, т. е. существует область пространства малого размера, вероятность расположения оптимального решения в которой гораздо выше, чем в остальной части пространства.

2. Разработан алгоритм проверки эффективности архитектуры многослойного персептрона. Экспериментально установлено, что алгоритм позволяет проверить не только эффективность архитектуры нейронной сети, но и эффективность алгоритма генерации начального приближения.

3. С помощью предложенного алгоритма можно численно сравнивать эффективность архитектур с помощью усредненного значения коэффициента перекоса. Алгоритм не требует больших вычислительных ресурсов, что позволяет ускорить обучение нейронных сетей путем быстрого выбора эффективной архитектуры, а также повысить качество конечного решения и эффективность алгоритма генерации начального приближения.

### Список литературы / References

1. Zhuang H., Lin Zh., Yang Y., Toh K.-A. (2025) An Analytic Formulation of Convolutional Neural Network Learning for Pattern Recognition. *Information Sciences. Elsevier*. 686. <https://doi.org/10.1016/j.ins.2024.121317>.

2. Li Z., Yu Y., Yin Ch., Shi Y. (2025) AdaNet: A Competitive Adaptive Convolutional Neural Network for Spectral Information Identification. *Pattern Recognition. Elsevier*. 163. <https://doi.org/10.1016/j.pat-cog.2025.111472>.
3. Talebzadeh H., Talebzadeh M., Satarpour M., Jalali F., Farhadi B., Vahdatpour M. S. (2024) Enhancing Breast Cancer Diagnosis Accuracy Through Genetic Algorithm-Optimized Multilayer Perceptron. *Multiscale and Multidisciplinary Modeling, Experiments and Design*. 7 (4), 4433–4449. <https://doi.org/10.1007/s41939-024-00487-3>.
4. Gao Y., Hu Z., Chen W. A., Liu M., Ruan Y. (2025) A Revolutionary Neural Network Architecture with Interpretability and Flexibility Based on Kolmogorov-Arnold for Solar Radiation and Temperature Forecasting. *Applied Energy. Elsevier*. 378 (A). <https://doi.org/10.1016/j.apenergy.2024.124844>.
5. Du J., Zeng J., Chen Ch., Ni M., Guo Ch., Zhang Sh., et al. (2025) Acoustic Emission Monitoring for Damage Diagnosis in Composite Laminates Based on Deep Learning with Attention Mechanism. *Mechanical Systems and Signal Processing. Elsevier*. 222. <https://doi.org/10.1016/j.ymsp.2024.111770>.
6. Altalhan M., Algarni A., Alouane M. T. H. (2025) Imbalanced Data Problem in Machine Learning: A Review. *IEEE Access*. 13, 13686–13699. <http://doi.org/10.1109/ACCESS.2025.3531662>.
7. Krasnoproshin V. V., Matskevich V. V. (2024) Random Search in Neural Networks Training. *Pattern Recognition and Image Analysis*. 34 (2), 309–316. <http://dx.doi.org/10.1134/S105466182470010X>.
8. Heydarian M., Doyle T. E., Samavi R. (2022) MLCM: Multi-Label Confusion Matrix. *IEEE Access*. 10, 19083–19095. <http://dx.doi.org/10.1109/ACCESS.2022.3151048>.
9. Zhou X., Bu Q., Matskevich V., Nedzved A. (2024) Landscape's Non-Natural Changes Detection System by Satellites Images Based on Local Areas. *Pattern Recognition and Image Analysis*. 34 (2), 365–378. <http://dx.doi.org/10.1134/S1054661824700159>.
10. Krasnoproshin V. V., Matskevich V. V. (2023) Neural Network Software Technology Trainable on the Random Search Principles. *Open Semantic Technologies for Intelligent Systems, Research Papers Collection*. 7, 133–140.

Поступила 31.03.2025

Received: 31 March 2025

Принята в печать 27.06.2025

Accepted: 27 June 2025

#### Вклад авторов / Authors' contribution

Авторы внесли равный вклад в написание статьи / The authors contributed equally to the writing of the article.

#### Сведения об авторах

**Мацкевич В. В.**, канд. техн. наук, доц. каф. информационных систем управления, Белорусский государственный университет

**Го Цзижань**, асп. каф. информационных систем управления, Белорусский государственный университет

#### Адрес для корреспонденции

220030, Республика Беларусь,  
Минск, просп. Независимости, 4  
Белорусский государственный университет  
Тел.: +375 29 125-49-07  
E-mail: matskevich1997@gmail.com  
Мацкевич Вадим Владимирович

#### Information about the authors

**Matskevich V. V.**, Cand. Sci. (Tech.), Associate Professor at Information Management Systems Department, Belarusian State University

**Guo J.**, Postgraduate at Information Management Systems Department, Belarusian State University

#### Address for correspondence

220030, Republic of Belarus,  
Minsk, Nezavisimosti Ave., 4  
Belarusian State University  
Тел.: +375 29 125-49-07  
E-mail: matskevich1997@gmail.com  
Matskevich Vadim Vladimirovich