

УДК 519.24

## О СИНТЕЗЕ НЕЙРОННЫХ СЕТЕЙ\*

Д.О. РОМАННИКОВ

630073, РФ, г. Новосибирск, пр. Карла Маркса, 20, Новосибирский государственный технический университет, кандидат технических наук, доцент кафедры автоматизи. E-mail: dmitry.romannikov@gmail.com

Применение нейронных сетей для решения задач различной направленности, к которым можно отнести задачи классификации (в частности, задачи распознавания образов на изображениях), задачи сегментирования, прогнозирования и динамического программирования и другие, приобрело достаточно большую популярность в настоящее время. Но стоит отметить, что предлагаемые решения всех перечисленных задач основываются на применении методов обучения нейронной сети с использованием градиентного спуска или его модификаций, что влечет за собой набор недостатков: неопределенность в задании начальных значений, отсутствие обоснованных рекомендаций по выбору параметров алгоритмов обучения (шага обучения и других) и другие. В статье рассматривается метод построения нейронной сети, основанный на последовательном добавлении нейронов для соединения новых «ближайших» точек из обучаемого множества при помощи функции активации выпрямителя (*relu*), а в нейроне выходного слоя – линейной функции. В статье этот способ применяется на примере аппроксимации функции параболы. В результате получена нейронная сеть, аппроксимирующая функцию параболы. Структура нейронной сети представляет собой однослойный персептрон с множеством нейронов в скрытом слое с одним нейроном во входном и выходном слоях. Предлагаемый в статье метод позволяет аппроксимировать не только функцию параболы, но и любую другую функцию одной переменной с заданной точностью. Приведены графики сравнения оригинальной функции и аппроксимированной с целью демонстрации полученных результатов, также приведена часть структуры полученной нейронной сети.

**Ключевые слова:** нейронные сети, сети Петри, искусственный интеллект, преобразование, функция активации, keras, регуляризация, обучение

DOI: 10.17212/2307-6879-2018-1-104-111

---

\* Статья получена 11 декабря 2017 г.

## ВВЕДЕНИЕ

Методы обучения нейронных сетей [1–9], которые используются для инициализации начальных значений нейронов и непосредственного получения их окончательных значений, на сегодняшний день имеют некоторые общие недостатки, которые если не ограничивают их использование, то приводят к необходимости учета некоторых проблем на практике. К таким проблемам можно отнести отсутствие однозначных и обоснованных рекомендаций по выбору начальных значений нейронов, шагов обучения, времени (или критерия останова) процесса обучения нейронной сети.

В статье исследуется итеративный процесс построения нейронной сети, в котором добавление нового нейрона основывается на соединении ближайшей точки с новой точкой в множестве входных значений.

## 1. ОПИСАНИЕ ЗАДАЧИ

Поставленной в работе задачей является разработка метода синтеза нейронной сети для задачи аппроксимации функции параболы (рис. 1).

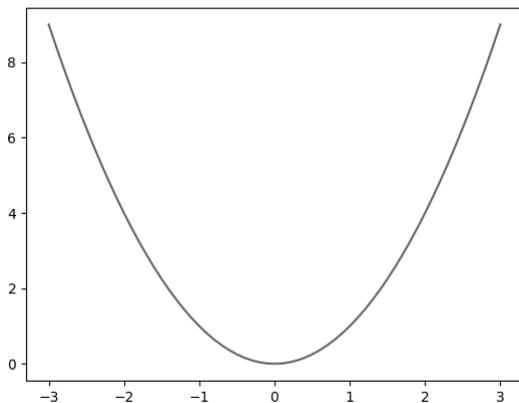


Рис. 1. График аппроксимируемой функции параболы  $f(x) = x^2$

Входными данными для этой задачи является набор данных для обучения со значениями функции в диапазоне от  $[-3, 3]$  с шагом, равным единице.

## 2. РЕШЕНИЕ

Основной идеей предлагаемого решения является итеративное построение нейронной сети с одним скрытым слоем (персептрона) и одним нейроном в выходном слое. Причем нейроны в скрытом слое будут использовать нелинейную функцию выпрямителя (*relu*), график которой приведен на рис. 2, а в нейроне выходного слоя – линейную функцию. Рассмотрим метод предлагаемого решения подробнее на примере аппроксимации параболической функции (см. рис. 1).

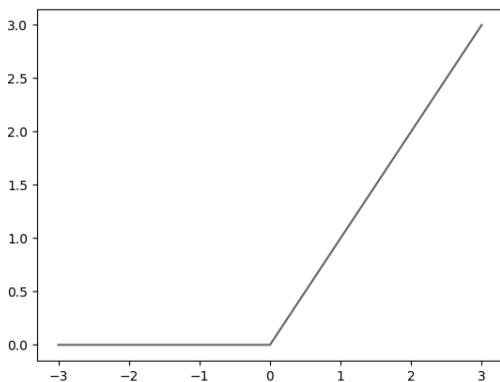


Рис. 2. График нелинейной функции нейрона скрытого слоя  $relu: f(x) = \{0, x \leq 0; x, x > 0\}$

Аппроксимация функции выполняется по шагам. Первым шагом является построение прямой между первой парой входных и выходных данных, для исследуемого примера этими данными будут пары точек  $(x_1, y_1)$ ,  $(x_2, y_2)$ :  $(-3, 9)$ ,  $(-2, 4)$ . Построение прямой  $y = ax + b$  между двумя точками выполняется из соотношений:  $a = (y_1 - y_2) / (x_1 - x_2)$ ,  $b = y_1 - (y_1 - y_2) / (x_1 - x_2) \cdot x_1$ . Для рассматриваемых пар точек прямая будет иметь вид  $y = -5x - 6$ . При этом нейронная сеть будет состоять из одного скрытого нейрона и одного нейрона в выходном слое, где нейрон скрытого слоя можно описать функцией  $relu(-5x - 6)$ . Важно отметить, что график вышеприведенной функции имеет ненулевые значения в диапазоне  $(-\infty, -6/5)$ . Второй шаг построения нейронной сети связан с обработкой следующей точки входных данных, которая является ближайшей к точкам  $(-2, 4)$ ,  $(-1, 1)$ . Тогда диапазон  $x \in (-2, -1]$ , можно

разделить на две части:  $x \in (-2, -6/5] \cup (-6/5, -1]$ . Такое разделение необходимо, так как в первой части диапазона для описания результирующей функции необходимо учитывать влияние функции  $relu(-5x - 6)$ , а во второй – нет. Тогда результирующая функция для отрезка  $(-2, -6/5]$  будет иметь вид  $f_2(x) = f_1(x) + ax + b$ , где  $f_1(x)$  – функция, посчитанная на предыдущем шаге, а  $ax + b$  – неизвестная часть, которую следует определить. И стоит заметить, что функция  $f_2(x)$  будет являться комбинацией линейной функций. Параметры  $a, b$  можно определить из вышеприведенных соотношений с поправкой на смещения:  $a = 2, b = 4$ . Далее логика вычислений повторяется: функция  $y = relu(2x + 4)$  имеет ненулевые значения, начиная от точки  $x = -2$ , а вычисление второй части диапазона  $(-6/5, -1]$  выполняется аналогичным образом, где  $f_3(x) = 2x + 4 + ax + b$ , тогда  $a = -5, b = -6$ . Важно отметить, что так как значение функции вычитается от остальных, то необходимо взять эту функцию со знаком минус:  $-relu(5x + 6)$ . Далее процесс определения остальных слагаемых функции выполняется аналогично, и в итоге получается нейронная сеть описываемая выражением  $relu(-5x - 6) + relu(2x + 4) - relu(5x + 6) + + relu(2x + 2) + relu(2x) + relu(2x - 2) + relu(2x - 4)$ , часть структуры которой приведена на рис 3, а на рис. 4 показаны графики оригинальной и аппроксимированной функций.

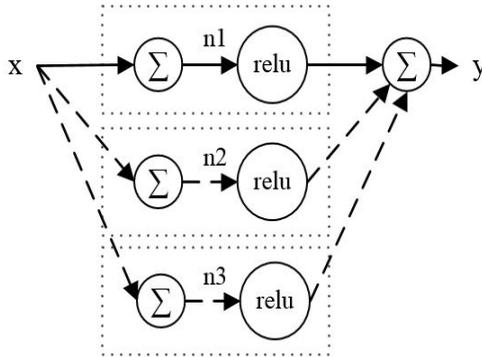


Рис. 3. Часть структуры синтезируемой нейронной сети

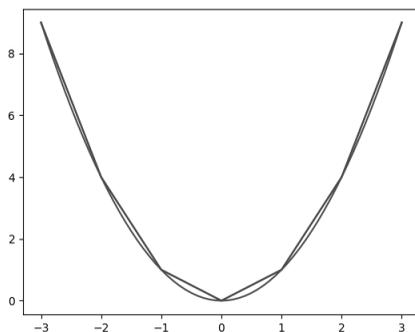


Рис. 4. Графики функции параболы и аппроксимированной по семи точкам функции параболы

Из рис. 4 видно, что выходное значение синтезируемой нейронной сети позволяет достаточно точно повторить график аппроксимированной функции.

## ЗАКЛЮЧЕНИЕ

В статье исследуется процесс аппроксимации функции параболы при помощи нейронной сети. Новизной предлагаемого способа аппроксимации является процесс синтеза нейронной сети, основанный на последовательном дополнении нейронов в структуру однослойного персептрона. В данном процессе использовались различные вариации функции выпрямителя (*relu*), однако могут использоваться и другие функции активации нейрона.

Стоит отметить, что вопрос разработки алгоритма аппроксимации функции нескольких переменных, который соответствует синтезу нейронной сети с несколькими входами, является вопросом дальнейших исследований.

## СПИСОК ЛИТЕРАТУРЫ

1. *Bishop C.* Pattern recognition and machine learning. – New York: Springer, 2007. – 738 p. – (Information science and statistics).
2. *Richert W., Coelho L.* Building machine learning systems with Python. – Birmingham: Packt Publ., 2013. – 290 p.

3. *Hastie T., Tibshirani R., Friedman J.* The elements of statistical learning: data mining, inference, and prediction. – 2<sup>nd</sup> ed. – New York: Springer, 2013. – 745 p. – (Springer series in statistics).
4. *Lantz B.* Machine learning with R. – Birmingham: Packt Publ., 2013. – 396 p.
5. *Mohri M., Rostamizadeh A., Talwalkar A.* Foundations of machine learning. – Cambridge, MA: MIT Press, 2012. – 432 p. – (Adaptive computation and machine learning series).
6. *Conway D., White J.M.* Machine learning for hackers. – Sebastopol, CA: O'Reilly, 2012. – 324 p.
7. Welcome to the Deep Learning tutorial [Electronic resource]. – URL: <http://deeplearning.stanford.edu/tutorial/> (accessed: 25.04.2018).
8. *Haykin S.* Neural networks: a comprehensive foundation. – New York: MacMillan Publ., 1994. – 1104 p.
9. *Романников Д.О.* О преобразовании сети Петри в нейронную сеть // Сборник научных трудов НГТУ. – 2016. – № 4 (86). – С. 98–103.
10. *Goodfellow I., Bengio Y., Courville A.* Deep learning. – Cambridge, MA: MIT Press, 2016.
11. *Krizhevsky A., Sutskever I., Hinton G.E.* ImageNet classification with deep convolutional neural networks // Proceedings of the 25th International Conference on Neural Information Processing Systems: NIPS 12. – Lake Tahoe, Nevada, 2012. – P. 1097–1105.
12. *Graves A., Mohamed A., Hinton G.* Speech recognition with deep recurrent neural networks // Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2013. – Vancouver, Canada, 2013.
13. *Deng L., Hinton G.E., Kingsbury B.* New types of deep neural network learning for speech recognition and related applications: an overview // Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2013. – Vancouver, Canada, 2013.

**Романников Дмитрий Олегович**, кандидат технических наук, доцент кафедры автоматике Новосибирского государственного технического университета. Основное направление научных исследований – нейронные сети, сети Петри. Имеет более 60 публикаций. E-mail: [dmitry.romannikov@gmail.com](mailto:dmitry.romannikov@gmail.com)

## About of synthesis of neural networks \*

### D.O. Romannikov

*Novosibirsk State Technical University, 20 Karl Marks Avenue, Novosibirsk, 630073, Russian Federation, candidate of Technical Sciences, associate professor of the automation department. E-mail: dmitry.romannikov@gmail.com*

The use of neural networks for solving problems of various orientations, to which classification problems (in particular, image recognition problems on images), tasks of segmentation, forecasting and dynamic programming, etc., have become quite popular at the present time. But, it is worth noting that the proposed solutions to all these problems are based on the use of neural network training methods using gradient descent or its modifications, which entails a set of shortcomings: uncertainty in setting initial values, lack of well-founded recommendations on the choice of parameters of algorithms training (step of training and others) and others. The method of constructing a neural network is based on the successive addition of neurons for the connection of new "nearest" points from the learning set using the function of activating the rectifier (relu), and in the neuron of the output layer it is a linear function. In this paper, this method is applied to the approximation of a parabola function. As a result, a neural network approximating the parabola function is obtained. The structure of a neural network is a single-layer perceptron with a set of neurons in a hidden layer with one neuron in the input and output layers. The method proposed in the article allows us to approximate not only the parabola function, but also any other function of one variable with a given accuracy. The graphs of comparison of the original function and approximated for the purpose of demonstrating the obtained results are given, as well as a part of the structure of the obtained neural network.

**Keywords:** neural networks, Petri nets, artificial intelligence, transformation, activation function, keras, regularization, training

DOI: 10.17212/2307-6879-2018-1-104-111

## REFERENCES

1. Bishop C. *Pattern recognition and machine learning. Information science and statistics*. New York, Springer, 2007. 738 p.
2. Richert W., Coelho L. *Building machine learning systems with Python*. Birmingham, Packt Publ., 2013. 290 p.
3. Hastie T., Tibshirani R., Friedman J. *The elements of statistical learning: data mining, inference, and prediction*. 2<sup>nd</sup> ed. *Springer series in statistics*. New York, Springer, 2013. 745 p.

---

\* Received 11 December 2017.

4. Lantz B. *Machine learning with R*. Birmingham, Packt Publ., 2013. 396 p.
5. Mohri M., Rostamizadeh A., Talwalkar A. *Foundations of machine learning. Adaptive computation and machine learning series*. Cambridge, MA, MIT Press, 2012. 432 p.
6. Conway D., White J.M. *Machine learning for hackers*. Sebastopol, CA, O'Reilly, 2012. 324 p.
7. *Welcome to the Deep Learning tutorial*. Available at: <http://deeplearning.stanford.edu/tutorial/> (accessed 25.04.2018).
8. Haykin S. *Neural networks: a comprehensive foundation*. New York, MacMillan Publ., 1994. 1104 p.
9. Romannikov D.O. O preobrazovanii seti Petri v neironnyuy set' [On the transformation of Petri nets in neural network]. *Sbornik nauchnykh trudov Novosibirskogo gosudarstvennogo tekhnicheskogo universiteta – Transaction of scientific papers of the Novosibirsk state technical university*, 2016, no. 4 (86), pp. 98–103.
10. Goodfellow I., Bengio Y., Courville A. *Deep learning*. Cambridge, MA, MIT Press, 2016.
11. Krizhevsky A., Sutskever I., Hinton G.E. ImageNet classification with deep convolutional neural networks. *Proceedings of the 25th International Conference on Neural Information Processing Systems: NIPS 12*, Lake Tahoe, Nevada, 2012, pp. 1097–1105.
12. Graves A., Mohamed A., Hinton G. Speech recognition with deep recurrent neural networks. *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2013*, Vancouver, Canada, 2013.
13. Deng L., Hinton G.E., Kingsbury B. New types of deep neural network learning for speech recognition and related applications: an overview. *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2013*, Vancouver, Canada, 2013.

Для цитирования:

Романников Д.О. О синтезе нейронных сетей // Сборник научных трудов НГТУ. – 2018. – № 1 (91). – С. 104–111. – doi: 10.17212/2307-6879-2018-1-104-111.

For citation:

Romannikov D.O. O sinteze neironnykh setei [About of synthesis of neural networks]. *Sbornik nauchnykh trudov Novosibirskogo gosudarstvennogo tekhnicheskogo universiteta – Transaction of scientific papers of the Novosibirsk state technical university*, 2018, no. 1 (91), pp. 104–111. doi: 10.17212/2307-6879-2018-1-104-111.