

*СОВРЕМЕННЫЕ
ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ*

УДК 519.24

**СИНТЕЗ НЕЙРОННЫХ СЕТЕЙ С НЕСКОЛЬКИМИ
ПЕРЕМЕННЫМИ***

А.А. ВОЕВОДА¹, Д.О. РОМАННИКОВ²

¹ 630073, РФ, г. Новосибирск, пр. Карла Маркса, 20, Новосибирский государственный технический университет, доктор технических наук, профессор кафедры автоматики. E-mail: ucit@ucit.ru

² 630073, РФ, г. Новосибирск, пр. Карла Маркса, 20, Новосибирский государственный технический университет, кандидат технических наук, доцент кафедры автоматики. 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 исследуется процесс аппроксимации функций двух переменных при помощи алгоритма построения нейронной сети, которая не требует обучения. Основная идея алгоритма заключается в разделении всей поверхности входных координат на проекции поверхностей, через которые проходит плоскость функции активации нейрона (в статье использовалась функция выпрямителя – *relu*), и дальнейшем расчете выходного значения, основанного на следующих этапах: 1) определение принадлежности входной переменной одному из диапазонов, т. е. принадлежности одной из проекций участков плоскостей точек из обучающего множества; 2) исходя из определенной проекции формируется выходное значение при помощи нейрона и функции активации задающей поверхности. Предлагаемый способ синтеза нейронной сети позволяет определить достаточно малое количество нейронов, которое требуется для построения нейронной сети, а также их значения. Последнее особенно важно, так как позволяет существенно ускорить процесс обучения, а точнее, заменить его процессом расчета значений обучающих параметров. Важной особенностью моделей машинного обучения является их поведение в точках за пределами обучающей выборки.

* Статья получена 20 октября 2017 г.

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

DOI: 10.17212/2307-6879-2018-1-86-94

ВВЕДЕНИЕ

На сегодняшний день в процессе разработки систем, основанных на использовании нейронных сетей [1–9], самым затратным является процесс обучения. С другой стороны, в работах [10, 11] обсуждаются способы построения нейронных систем в таком виде, что они не требуют обучения, а основаны на расчете коэффициентов нейронной сети. Так, в частности, в [10] рассматривается пример построения нейронной сети для решения логико-арифметических задач, в [11] – пример сортировки массива чисел с использованием нейронных сетей и другие.

В статье исследуется процесс аппроксимации функций двух переменных при помощи алгоритма построения нейронной сети, который основан на разделении всей поверхности входных переменных на части, определении принадлежности входной точки одной из частей и формировании выходного сигнала при помощи функции активации нейрона.

ПОСТАНОВКА ЗАДАЧИ И ЕЕ РЕШЕНИЕ

Поставленной в работе задачей является разработка метода синтеза нейронной сети для задачи аппроксимации функции нескольких переменных на примере функций параболоиды, гиперболического параболоида и других функций. Входными данными для этой задачи является набор данных для обучения со значениями функции.

В отличие от синтеза функции одной переменной, при синтезе функции нескольких переменных достаточно сложно выполнять построение плоскостей (взятых с суммы входных переменных перед функцией активации) так, чтобы плоскости, проходящие через новые точки, не меняли значения уже построенных плоскостей. Таким образом, основной идеей предлагаемого решения является итеративное построение нейронной сети с двумя скрытыми слоями и одним нейроном в выходном слое. Причем нейроны в первом скрытом слое будут определять принадлежность входной точки к части проекции на плоскость входных переменных, а нейроны во втором слое нужны для

формирования значения функции на плоскости, соединяющей ближайšie три точки в пространстве. Целью такого построения является получение нейронной сети с уже «обученными» коэффициентами.

Решение вышеприведенной задачи прямо вытекает из ее постановки:

1. Первым шагом алгоритма необходимо выбрать любую точку из входного множества и две ближайšie к ней.

2. Построение определителя принадлежности точки к части поверхности выполняется при помощи векторного произведения всех пар сторон треугольника (образованного от трех точек для проведения поверхности). Если все знаки векторного произведения пар сторон одного знака или нулевые, то точка принадлежит части поверхности. Согласно [8, 10] такую функцию можно построить при помощи нейронов. Выходным значением данной функции является либо ноль, либо единица, что соответствует принадлежности точки части плоскости.

3. Формирование значения функции в вычисляемой точке выполняется при помощи суммы сигналов нейронов с функцией активации выпрямителя (*relu*). Вычисление обучаемых коэффициентов нейронной сети для данной функции может быть выполнено при помощи решения системы линейных уравнений с использованием уравнения плоскости, проходящий через три точки.

Одной из особенностей данной схемы построения нейронной сети является то, что в случае, когда точка лежит на стороне проекции треугольника, в данной схеме она определяется двумя нейронами определителя, что ведет к двойному суммированию в выходном нейроне. Для устранения данной проблемы необходимо либо учитывать пересечение со стороной только для одного треугольника, либо итоговую сумму делить на количество срабатываемых определителей.

Используя вышеприведенный алгоритм, можно получить структуру нейронной сети, приведенную на рис. 1. При помощи этой структуры и вышеприведенного алгоритма были получены функции параболоиды, гиперболического параболоида и другие (рис. 2).

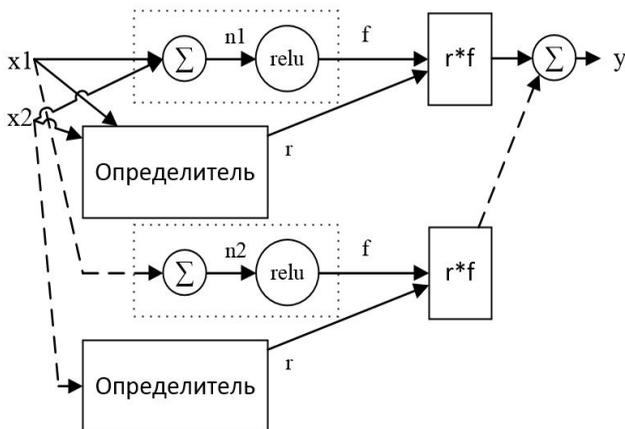


Рис. 1. Структура нейронной сети для формирования значений функции двух переменных

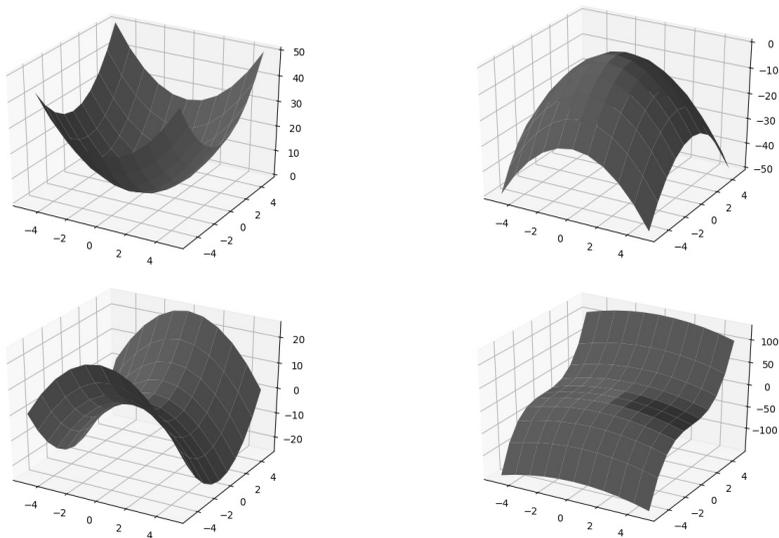


Рис. 2. Графики поверхностей, построенных при помощи структуры нейронной сети, изображенной на рис. 1. Приведены поверхности параболоида ($x_1^2 + x_2^2 = y$), отрицательного гиперолоида ($-x_1^2 - x_2^2 = y$), гиперболического параболоида ($x_1^2 - x_2^2 = y$) и функции третьего порядка ($x_1^3 - x_2^2 + 1 = y$)

Из рис. 2 видно, что выходное значение синтезируемой нейронной сети позволяет достаточно точно описать заданные поверхности.

ЗАКЛЮЧЕНИЕ

В статье исследуется процесс аппроксимации функций двух переменных при помощи алгоритма построения нейронной сети, которая не требует обучения. Основная идея алгоритма заключается в разделении всей поверхности входных координат на проекции поверхностей, через которые проходит плоскость функции активации нейрона (в статье использовалась функция выпрямителя – *relu*), и дальнейшем расчете выходного значения, основанного на следующих этапах: 1) определение принадлежности входной переменной одному из диапазонов, т. е. принадлежности одной из проекций участков плоскостей точек из обучающего множества; 2) исходя из определенной проекции формируется выходное значение при помощи нейрона и функции активации задающей поверхности.

Предлагаемый способ синтеза нейронной сети позволяет определить достаточно малое количество нейронов, которые требуются для построения нейронной сети, а также их значения. Последнее особенно важно, так как позволяет существенно ускорить процесс обучения, а точнее, заменить его процессом расчета значений обучающих параметров. Важной особенностью моделей машинного обучения является их поведение в точках за пределами обучающей выборки.

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. – 2nd 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. *Воевода А.А., Романников Д.О.* Синтез нейронной сети для решения логико-арифметических задач // Труды СПИИРАН. – 2017. – Вып. 54. – С. 205–223.
11. *Voevoda A.A., Romannikov D.O.* A binary array asynchronous sorting algorithm with using Petri nets // Journal of Physics: Conference Series. – 2017. – Vol. 803, N 1. – P. 012178.
12. *Goodfellow I., Bengio Y., Courville A.* Deep learning. – Cambridge, MA: MIT Press, 2016.
13. *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.
14. *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.
15. *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.

Воевода Александр Александрович, доктор технических наук, профессор кафедры автоматике Новосибирского государственного технического университета. Основное направление научных исследований – управление многоканальными объектами. Имеет более 300 публикаций. E-mail: ucit@ucit.ru

Романников Дмитрий Олегович, кандидат технических наук, доцент кафедры автоматике Новосибирского государственного технического универси-

тета. Основное направление научных исследований – нейронные сети, сети Петри. Имеет более 60 публикаций. E-mail: dmitry.romannikov@gmail.com

Synthesis of neural networks with several variables*

A.A. Voevoda¹, D.O. Romannikov²

¹Novosibirsk State Technical University, 20 Karl Marks Avenue, Novosibirsk, 630073, Russian Federation, doctor of Technical Sciences, professor of the automation department. E-mail: ucit@ucit.ru

²Novosibirsk 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 article investigates the process of approximation of functions of two variables using the neural network construction algorithm, which does not require training. The main idea of the algorithm is the separation of the entire surface of the input coordinates into the projection of the surfaces through which the plane of the neuron activation function passes (in the article the rectifier function was used) and further calculation of the output value based on the steps: 1) determining the belonging of the input variable to one of the ranges i.e. the belonging of one of the projections of sections of the planes of points from the learning set; 2) based on a certain projection, the output value is formed using a neuron and the activation function of the master surface. The proposed method of synthesis of a neural network allows us to determine a sufficiently small number of neurons, which are required for the construction of a neural network, as well as their values. The latter is especially important, because allows significantly accelerate the learning process, or rather, replace it with the process of calculating the values of training parameters. An important feature of machine learning models is their behavior at points outside the training sample.

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

DOI: 10.17212/2307-6879-2018-1-86-94

*Received 20 October 2017.

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*. 2nd ed. *Springer series in statistics*. New York, Springer, 2013. 745 p.
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 neironnuyu 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. Voevoda A.A., Romannikov D.O. Sintez neironnoi seti dlya resheniya logiko-arifmeticheskikh zadach [Synthesis of neural network for solving logical-arithmetic problems]. *Trudy SPIIRAN – SPIIRAS proceedings*, 2017, vol. 54, no. 5, pp. 205–223.
11. Voevoda A.A., Romannikov D.O. A binary array asynchronous sorting algorithm with using Petri nets. *Journal of Physics: Conference Series*, 2017, vol. 803, no. 1, p. 012178.
12. Goodfellow I., Bengio Y., Courville A. *Deep learning*. Cambridge, MA, MIT Press, 2016.
13. 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.
14. 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.

15. 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:

Воевода А.А., Романников Д.О. Синтез нейронных сетей с несколькими переменными // Сборник научных трудов НГТУ. – 2018. – № 1 (91). – С. 86–94. – doi: 10.17212/2307-6879-2018-1-86-94.

For citation:

Voevoda A.A., Romannikov D.O. Sintez neironnykh setei s neskol'kimi peremennymi [Synthesis of neural networks with several variables]. *Sbornik nauchnykh trudov Novosibirskogo gosudarstvennogo tekhnicheskogo universiteta – Transaction of scientific papers of the Novosibirsk state technical university*, 2018, no. 1 (91), pp. 86–94. doi: 10.17212/2307-6879-2018-1-86-94.