УДК 004.8

А.Т. Нгуен, А.М. Кориков

Нейронные сети с полиномиальными кусочно-непрерывными функциями активации для поиска закономерностей в данных

Исследовано применение моделей нейронных сетей, использующих полиномиальные кусочно-непрерывные функции активации (ПКНФА) нейронов, для решения задач обнаружения закономерностей в наборах данных. Разработана структура многослойной однонаправленной нейронной сети для четырех видов ПКНФА с использованием метода скользящего окна для прогнозирования временных рядов и определения амплитуды детерминированного полезного сигнала на фоне белого шума. В процессе обучения нейронной сети используется алгоритм Левенберга–Марквардта обратного распространения. По результатам тестирования проведено сравнение четырех нейронных сетей с различными видами ПКНФА с известными сетями. На основе сравнительного анализа разработаны рекомендации по использованию нейронных сетей с ПКНФА для обнаружения закономерностей в наборах данных.

Ключевые слова: нечеткие функции активации, нечеткие нейронные сети, обучение нейронных сетей, прогнозирование, определение параметров сигнала, поиск закономерностей в данных.

doi: 10.21293/1818-0442-2019-22-1-71-76

Проблема поиска закономерностей, скрытых в экспериментальных данных, является актуальной для многих разделов науки и техники. Во многих прикладных задачах наборы экспериментальных данных представляют собой последовательности наблюдений (данных), упорядоченные по времени. Такие последовательности обычно называют временными рядами [1]. При анализе временных рядов в основном используются методы математической статистики [1, 2], а в последние годы все шире применяются нейронные сети (см., например, [3]).

В наших работах [4–6] исследуются модели нейронных сетей (НС), использующих импульсные нечеткие функции активации (ФА) нейронов НС. В [4–6] используется четыре вида полиномиальных кусочно-непрерывных функций активации (ПКНФА). Цитируемые работы доступны практически всем читателям нашего журнала, поэтому далее ограни-

чимся только представлением внешнего вида этих ПКНФА на рис. 1.

Особенности построения ПКНФА, представленных на рис. 1, их настройка и применение изложены в [4–6]. Модель нейронной сети, использующую полиномиальную кусочно-непрерывную функцию активации, представленную на рис. 1, *а*, обозначим для краткости через НС-*а*. Соответственно через НС-*б*, НС-*в*, НС-*г* обозначим НС, использующие ПКНФА, представленные на рис. 1, *б–г*.

Исследуем эффективность решения задачи поиска закономерностей, скрытых во временных рядах, на основе новых моделей нейронных сетей НС-*а*, НС-*б*, НС-*в*, НС-*г*. Для сравнительного анализа эффективности НС используем модель НС с известной (сигмоидной) ФА. Данную пятую модель НС обозначим как НС-*сигм*.

Рис. 1. Внешний вид ПКНФА

Модели нейронной сети на основе скользящего окна для прогнозирования временных рядов

Метод скользящего окна широко применяется для прогнозирования временных рядов [7, 8]. Размер

окна и сегмента увеличивается до достижения наименьшей погрешности [7]. После выбора первого сегмента следующий сегмент выбирается [8] с конца первого сегмента. Процесс повторяется до тех пор, пока все данные временных рядов не будут сегмен-

тированы. Начальный размер скользящего окна выбирается опытным путем, а затем определяется по набору данных. Структура многослойной однона-

правленной сети (МОНС) состоит из входного слоя, одного скрытого слоя, выходного слоя и представлена на рис. 2.

Рис. 2. МОНС для реализации метода скользящего окна

Математическая модель МОНС может быть представлена в следующем виде:

$$\hat{X}_t = w_0 + \sum_{j=1}^Q w_j \cdot f \left(w_{0j} + \sum_{i=1}^P w_{ij} \cdot X_{t-i} \right), \quad (1)$$

где P и Q – количество входных и скрытых нейронов соответственно, $w_j, j=0, 1, \dots, Q$ – элементы весовых векторов $\mathbf{w}$ от скрытого слоя до выхода, $w_{ij}, i=1, 2, \dots, P$ & $j=1, 2, \dots, Q$ – элементы весовых векторов $\mathbf{w}$ от входа к скрытым нейронам. Для согласованности сети и сокращения времени вычисления входные данные нормализуются до начала обработки:

$$X_{\text{норм}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}}.$$

В наборе данных 70% данных используется для обучения НС и 30% – для тестирования. Для обучения выбран алгоритм Левенберга–Марквардта обратного распространения (ЛМОР) [9]. Обновление весовых векторов $\mathbf{w}$ выполняется по формуле [10]:

$$\mathbf{W}_{k+1} = \mathbf{W}_k - (\mathbf{J}^T \mathbf{J} + \mu \mathbf{I})^{-1} \mathbf{J}^T \mathbf{e}, \quad (2)$$

где $\mathbf{W}$ – обобщенный весовой вектор, содержащий элементы всех весовых векторов $\mathbf{w}$; $\mathbf{I}$ – единичная матрица; μ – параметр шага обучения, значение которого уменьшается после каждого успешного шага; $\mathbf{J}$ – матрица Якобиана, содержащая первые

производные сетевых ошибок относительно весов; $\mathbf{e}$ – вектор сетевых ошибок во всех образцах обучения. Для оценки матрицы Якобиана используется стандартный алгоритм обратного распространения для аппроксимации матрицы Гессияна [11]. В алгоритме (2) при вычислении матрицы Якобиана используется среднеквадратичная ошибка (СКО) δ для сетевого обучения:

$$\delta = \frac{\sum_{i=1}^n (\hat{Z}_i - Z_i)^2}{n}, \quad (3)$$

где Z_i обозначает i -е значение из выборки данных; $\hat{Z}_i$ – оценка, полученная сетью для значения Z_i ; n – количество выборок данных для сетевого обучения.

На основе МОНС, представленной на рис. 2, созданы пять моделей НС: четыре модели НС (НС-а, НС-б, НС-в, НС-г) с ПКНФА и модель НС с сигмоидной ФА – НС-сигм.

Модели данных

Исследуется последовательность случайных величин (случайный процесс) на некотором интервале времени T . Значение процесса z_t в каждый момент времени t является случайной величиной. Такие процессы называются временными рядами. Модель временного ряда можно представить последовательностью наблюдений:

$$z_t = y_t + \varepsilon_t, \quad (4)$$

где y_t – регулярная (систематическая) составляющая исследуемого временного ряда, ε_t – аддитивная случайная помеха. В свою очередь в модели (4) регулярную составляющую можно представить различными моделями сигналов: вероятностный (стохастический) полезный сигнал (тренд курса обмена валют на финансовых рынках и др.), детерминированный полезный сигнал (сигналы в радиотехнике, радионавигации, связи и т.п.). Условно назовем модель (4) со стохастическим полезным сигналом –

первая модель данных, а модель (4) с детерминированным полезным сигналом – вторая модель данных.

Временные ряды, соответствующие первой модели данных, представлены на рис. 3. На этом рисунке представлено три набора данных (три временных ряда): набор данных 1 – ежемесячный австралийский импорт из Японии от июля 1965 г. до октября 1993 г. [12], набор данных 2 – обменный курс австралийского доллара (доллар Австралии за 1 доллар США) от июля 1969 г. до августа 1995 г. [13] и набор данных 3 – экспорт квартальных австралийских национальных счетов от сентября 1959 г. до июня 1995 г. [14].

Рис. 3. Наборы данных (первая модель)

В качестве второй модели данных используем аддитивную смесь полезного сигнала y и белого шума ε . Детерминированный полезный сигнал представим в виде синусоиды:

$$y = x \times \sin(2\pi ft), \quad (5)$$

где x – амплитуда полезного сигнала; f – частота полезного сигнала. Белый шум ε имеет нормальный закон распределения $N(0, \sigma^2)$, т.е. имеет нулевое математическое ожидание и дисперсию σ^2 .

Моделирование нейронных сетей

Для каждого набора данных проведено по 5 экспериментов, в которых последовательно исполь-

зуется пять моделей НС: модель НС с сигмоидной ФА – НС-*сигм* и четыре модели НС (НС-*а*, НС-*б*, НС-*в*, НС-*г*) с новыми ПКНФА. В этих экспериментах используются одинаковые структуры для каждого набора данных, т.е. для каждого набора данных структуры НС имеют одинаковое количество нейронов (10 нейронов) в скрытом слое, размеры скользящего окна также не отличаются.

В табл. 1 приведены СКО, вычисленные по формуле (3), для определения лучшей модели НС для прогнозирования временных рядов с первой моделью данных.

Таблица 1

Данные	СКО для результатов экспериментов с первой моделью данных				
	Тип НС				
	НС- <i>сигм</i>	НС- <i>а</i>	НС- <i>б</i>	НС- <i>в</i>	НС- <i>г</i>
Набор данных 1	23×10^{-4}	25×10^{-4}	25×10^{-4}	23×10^{-4}	24×10^{-4}
Набор данных 2	18×10^{-4}	17×10^{-4}	16×10^{-4}	16×10^{-4}	21×10^{-4}
Набор данных 3	$2,74 \times 10^{-4}$	$7,01 \times 10^{-4}$	$2,33 \times 10^{-4}$	$2,89 \times 10^{-4}$	$4,38 \times 10^{-4}$

Из результатов тестирования НС следует, что разрабатываемые нами модели нечетких нейронов и нечетких НС второго типа [4–6] вполне успешно решают задачи прогнозирования временных рядов со стохастическим полезным сигналом (первая модель данных). СКО для прогнозируемых данных в НС с ПKNФА вида a , b , v , z (см. рис. 1) незначительно отличаются от СКО для прогнозируемых данных в обычной НС (НС-сигм). Эксперименты показывают, что модель НС с ФА вида b (НС- b) наиболее эффективна. Модели нейронной сети с ФА вида v (НС- v) также показывают хорошие результаты, но они незначительно отличаются от обычной НС. В наборах данных на рис. 3 нечеткие НС прогнозируют тренды, которые можно интерпретировать как вероятностные (стохастические) полезные сигналы на фоне случайных помех. Разрабатываемые нечеткие НС второго типа (особенно НС- b с ФА вида b) эффективно решают подобные задачи.

В экспериментах со второй моделью данных (5) выбираем амплитуды из интервала (1; 1,5); дискретное время t – это 500 эквидистантных значений от 0 до 0,5 с; задаем $f = 60$ Гц. Затем добавляем гауссовский белый шум (отношение сигнал/шум = 0,05). Отношение сигнал/шум – безразмерная величина, равная отношению мощности полезного сигнала к мощности шума [15. С. 51]. Чем больше это отношение, тем меньше шум влияет на характеристики сигнала. Таким образом, на вход НС поступает временной ряд, представляющий собой упорядоченную последовательность значений, характер которой меняется во времени t . Задачей НС является прогноз значения x – амплитуды полезного сигнала на входе НС. Подобные задачи возникают во многих приложениях НС. На рис. 4 показан чистый оригинальный сигнал – синусоида (5) с амплитудой 1,5 и этот же сигнал после добавления к нему гауссовского белого шума с отношением сигнал/шум = 0,05.

Рис. 4. Сигнал для экспериментов со второй моделью данных

Для тестирования НС со второй моделью данных выбираем 5 значений амплитуды в интервале от 1,1 до 1,5 и для каждого значения амплитуды создаем 20 зашумленных сигналов. Так же, как и с первой моделью данных, проводится 5 экспериментов,

в которых используется обычная НС с сигмоидной ФА для определения значения амплитуды. В табл. 2 приведены СКО, вычисленные по формуле (3), по выбору лучшей модели НС для определения амплитуды полезного сигнала на фоне белого шума.

Таблица 2

СКО для результатов экспериментов со второй моделью данных

Тип НС	Амплитуда сигнала				
	1,1	1,2	1,3	1,4	1,5
НС- a	230×10^{-4}	43×10^{-4}	230×10^{-4}	53×10^{-4}	157×10^{-4}
НС- b	129×10^{-4}	28×10^{-4}	41×10^{-4}	82×10^{-4}	339×10^{-4}
НС- v	88×10^{-4}	16×10^{-4}	$4,96 \times 10^{-4}$	$1,93 \times 10^{-4}$	$4,23 \times 10^{-4}$
НС- z	164×10^{-4}	25×10^{-4}	$2,10 \times 10^{-4}$	100×10^{-4}	217×10^{-4}
НС-сигм	278×10^{-4}	74×10^{-4}	34×10^{-4}	63×10^{-4}	315×10^{-4}

Из результатов тестирования НС следует, что разрабатываемые нами модели нечетких нейронов и нечетких НС второго типа вполне успешно решают задачу определения амплитуды детерминированного полезного сигнала на фоне белого шума. СКО в НС

с ФА вида a , b , v , z (см. рис. 1) в большинстве экспериментов показывают более высокую точность определения параметра детерминированного полезного сигнала. Эксперименты показывают, что модель НС с ФА вида v (НС- v) наиболее эффективна. Моде-

ли нейронной сети с ФА вида a , b , c также показывают хорошие результаты, но они незначительно отличаются от НС-*сигм*.

Заключение

1. Разработана структура НС с использованием метода скользящего окна для прогноза временных рядов с ФА вида a , b , c (см. рис. 1). Процесс обучения реализуется с помощью алгоритма Левенберга–Марквардта обратного распространения. Для обучения и тестирования НС используются реальные наборы данных и модели в виде аддитивной смеси детерминированного полезного сигнала и белого шума.

2. Среднеквадратичная ошибка (СКО) используется для определения лучшей модели НС. В экспериментах с прогнозированием реальных временных рядов (см. рис. 3), СКО НС с ФА вида a , b , c незначительно отличаются от СКО обычной НС. Эксперименты показывают, что модель НС с ФА вида b , c (см. рис. 1) наиболее эффективны. Тем не менее тестирование разработанных НС на смеси синусоидальных сигналов и гауссовского белого шума с отношением сигнал/шум = 0,05 и их сравнительный анализ с известной НС (ФА-*сигмоид*) показали более лучшие результаты по эффективности и точности определения параметра детерминированного полезного сигнала, особенно НС- b (НС с ФА вида b).

Литература

1. Андерсон Т.В. Статистический анализ временных рядов. – М.: Мир, 1976. – 757с.
2. Обнаружение закономерностей в массивах экспериментальных данных / Н.Г. Загоруйко, И.А. Борисова, О.А. Кутненко, Д.А. Леванов // Вычислительные технологии. – 2013. – Т.18, спец. вып. – С. 12–20.
3. Ефремова Е.А. Алгоритм выбора даты размещения субфедерального облигационного займа на основе аппарата нейронных сетей / Е.А. Ефремова // Доклады ТУСУР. – 2009. – № 1(19), ч. 1. – С. 144–149.
4. Nguyen A.T. Neural network model with fuzzy activation functions for time series predictions / A.T. Nguyen, A.M. Korikov // Доклады ТУСУР. – 2016. – Т. 19, № 4. – С. 50–52.
5. Кориков А.М. Нейро-нечеткая классификация объектов и их состояний / А.М. Кориков, А.Т. Нгуен // Научный вестник НГТУ. – 2018. – № 3 (72). – С. 73–86.
6. Nguyen A.T. Models of neural networks with fuzzy activation functions / A.T. Nguyen, A.M. Korikov // IOP Conference Series: Materials Science and Engineering. – 2017. – Vol. 177. – URL: <http://iopscience.iop.org/article/10.1088/1757-899X/177/1/012031> (дата обращения: 01.03.2018).
7. Mozaffari L. Vehicle speed prediction via a sliding-window time series analysis and an evolutionary least learning machine: A case study on San Francisco urban roads / L. Mozaffari, A. Moxaffari, N.L. Azad // Engineering Science and Technology, An International Journal. – 2015. – Vol. 18. – P. 150–162.
8. Vafaeipour M. Application of sliding window technique for prediction of wind velocity time series / M. Vafaeipour, O. Rahbari, M.A. Rosen, F. Fazelpour, P. Ansarirad // International Journal of Energy and Environmental Engineering (Springer). – 2014. – Vol. 5. – P. 105–111.

9. Levenberg-Marquardt backpropagation training of multilayer neural networks for state estimation of a safety critical cyber-physical system / L.V. Chen, Yan Xing, Junzhi Zhang et al. // IEEE Transactions On Industrial Informatics. – 2018. – Vol. 14, No. 8. – P. 3436–3446.

10. Nguyen A.T. Application of Artificial Neural Networks to Predict Dynamic Responses of Wing Structures due to Atmospheric Turbulence / A.T. Nguyen, J.H. Han, A.T. Nguyen // International Journal of Aeronautical and Space Sciences. – 2017. – Vol. 18 (3). – P. 474–484.

11. Hagan M.T. Training feed forward networks with the Marquardt algorithm / M.T. Hagan, M.B. Menhaj // IEEE Trans. Neural Netw. – 1994. – Vol. 5. – P. 989–993.

12. Monthly Australian imports from Japan: thousands of dollars. Jul 65 – Oct 93. – URL: <https://datamarket.com/data/set/22qx/monthly-australian-imports-from-japan-thousands-of-dollars-jul-65-oct-93#!ds=22qx&display=line> (дата обращения 01.03.2018).

13. Exchange rate of Australian dollar: \$A for 1 US dollar. Monthly average: Jul 1969 – Aug 1995. – URL: <https://datamarket.com/data/set/22wv/exchange-rate-of-australian-dollar-a-for-1-us-dollar-monthly-average-jul-1969-aug-1995#!ds=22wv&display=line> (дата обращения: 01.03.2018).

14. Quarterly Australian national accounts exports: millions of dollars at 1989/90 prices. Sep. 59 – Jun. 95. – URL: <https://datamarket.com/data/set/22my/quarterly-australian-national-accounts-exports-millions-of-dollars-at-198990-prices-sep-59-jun-95#!ds=22my&display=line> (дата обращения: 01.03.2018).

15. Хоровиц П. Искусство схемотехники: в 3 т. – Т. 2. Пер. с англ. / П. Хоровиц, У. Хилл. – 4-е изд., перераб. и доп. – М.: Мир, 1993. – 371 с.

Нгуен Ань Ту

Аспирант отд. автоматизации и робототехники
Национального исследовательского
Томского политехнического университета (НИ ТПУ)
Ленина пр-т, 30, г. Томск, Россия, 634050
Тел.: +7-952-157-48-78
Эл. почта: nguyenanhtu@tpu.ru

Кориков Анатолий Михайлович

Д-р техн. наук, профессор, зав. каф. АСУ ТУСУРа,
вед. науч. сотр. Томского ф-ла
Института вычислительных технологий СО РАН
Ленина пр-т, 40, г. Томск, Россия, 634050
Тел.: +7 (382-2) 41-42-79
Эл. почта: korikov@asu.tusur.ru

Nguyen A.T., Korikov A.M.

Neural networks with polynomial piecewise-continuous activation functions for the problems of finding data patterns

This paper investigates the application of neural network models using polynomial piecewise-continuous activation function (PPCAF) to solve the problems of finding patterns in data sets. The authors apply multi-layered feed forward neural networks (NNs) with four types of PPCAF and use the sliding window method to predict a time series and determine the amplitude of a given signal on the white noise background. The training process of the multi-layered feed forward NN is conducted with the Levenberg-Marquardt back-propagation

algorithm. From the testing results, a comparison is made between a known multi-layered feed forward NN and four multi-layered feed forward NNs using various types of PPCAF. Based on the result of the comparison analysis, recommendation is made for the use of multi-layered feed forward NNs with PPCAF for the problems of finding patterns in data sets.

Keywords: fuzzy activation functions, fuzzy neural networks, neural network training, prediction, determining signal parameters, finding data patterns.

doi: 10.21293/1818-0442-2019-22-1-71-76

References

1. Anderson T.W. The statistical analysis of time series. Publ: Mir, 1976. – 757p.
2. Zagoruiko N.G., Borisova I.A., Kutnenko O.A., Levvanov D.A. Obnaruzhenie zakonomernostei v massivakh eksperimentalnykh dannikh [Detection of patterns in experimental data arrays]. *Computational Technologies*, 2013, vol. 18, Special ed., pp. 12–20 (in Russ.).
3. Efremova E.A. Algoritmy izbora dati razmesheniya subfederalnogo obligatsionnogo zaima na osnove apparata neironnykh setei [The algorithm choice date of bonded loan placing by using neural networks mathematical apparatus]. *Proceedings of TUSUR University*, 2009, vol. 19, no. 1, pp. 144–149 (in Russ.).
4. Melnik I.A., Erofeev L.Ya. Physical and geochemical model of low resistivity reservoir and its practical application. *Geology, geophysics and development of oil and gas fields*, 2014, no. 3, pp. 46–50 (in Russ.).
5. Korikov A.M., Nguyen A.T. Neuro-nechetkaya klassifikatsiya ob'ektov i ikh sostoyanii [A neuro-fuzzy classification of objects and their states]. *Science Bulletin of the Novosibirsk State Technical University*, 2018, vol. 72, no. 3, pp. 73–86 (in Russ.).
6. Nguyen A.T., Korikov A.M. Models of neural networks with fuzzy activation functions. *IOP Conference Series: Materials Science and Engineering*, 2017, vol. 177. Available at: <http://iopscience.iop.org/article/10.1088/1757-899X/177/1/012031> (accessed: March 01, 2018).
7. Mozaffari L. Vehicle speed prediction via a sliding-window time series analysis and an evolutionary least learning machine: A case study on San Francisco urban roads. *Engineering Science and Technology, An International Journal*, 2015, vol. 18, pp. 150–162.
8. Vafaeipour M., Rahbari O., Rosoen M.A., Fazelpour F., Ansarirad P. Application of sliding window technique for prediction of wind velocity time series. *International Journal of Energy and Environmental Engineering (Springer)*, 2014, vol. 5, pp. 105–111.
9. Chen L.V., Yan Xing, Junzhi Zhang et al. Levenberg-Marquardt backpropagation training of multilayer neural networks for state estimation of a safety critical cyber-physical system. *IEEE Transactions On Industrial Informatics*, 2018, vol. 14, no. 8, pp. 3436–3446.
10. Nguyen A.T., Han J.H., Nguyen A.T. Application of Artificial Neural Networks to Predict Dynamic Responses of Wing Structures due to Atmospheric Turbulence. *International Journal of Aeronautical and Space Sciences*, 2017, vol. 18, no. 3, pp. 474–484.
11. Hagan M.T., Menhaj M.B. Training feed forward networks with the Marquardt algorithm. *IEEE Trans. Neural Netw.*, 1994, vol. 5, pp. 989–993.
12. Monthly Australian imports from Japan: thousands of dollars. Jul 65 – Oct 93. Available at: <https://datamarket.com/data/set/22qx/monthly-australian-imports-from-japan-thousands-of-dollars-jul-65-oct-93#!ds=22qx&display=line> (accessed: 01 March 2018).
13. Exchange rate of Australian dollar: \$A for 1 US dollar. Monthly average: Jul 1969 – Aug 1995. Available at: <https://datamarket.com/data/set/22wv/exchangerate-of-australian-dollar-a-for-1-us-dollar-monthly-average-jul-1969-aug-1995#!ds=22wv&display=line> (accessed: 01 March 2018).
14. Quarterly Australian national accounts exports: millions of dollars at 1989/90 prices. Sep 59 – Jun 95. Available at: <https://datamarket.com/data/set/22my/quarterly-australian-national-accounts-exports-millions-of-dollars-at-198990-prices-sep-59-jun-95#!ds=22my&display=line> (accessed: March 01, 2018).
15. Horowitz P., Hill W. The Art of Electronics. M.: Mir, 1993, vol. 3, 371 p.

Anh T. Nguyen

PhD student, Department of Automation and Robotics,
Tomsk Polytechnic University
30, Lenin pr., Tomsk, Russia, 634050
Phone: +7-952-157-48-78
Email: nguyenanhtu@tpu.ru

Anatoly M. Korikov

Doctor of Engineering, Professor,
Head of the Department of Automated Control Systems
Tomsk State University of Control Systems
and Radioelectronics (TUSUR)
40, Lenin pr., Tomsk, Russia, 634050;
Leading researcher at Tomsk Branch of the Institute
of Computing Technologies SB RAS
Phone: +7 (382-2) 41-42-79
Email: korikov@asu.tusur.ru