

УДК 519.95

МОДЕЛЬ ИСКУССТВЕННОЙ НЕЙРОННОЙ СЕТИ (ИНС) С РЕАЛИЗАЦИЕЙ МОДУЛЬНОГО ПРИНЦИПА ОБУЧЕНИЯ

© А.А. Арзамасцев, В.П. Рыков

Ключевые слова: искусственные нейронные сети; модульный принцип обучения.

Рассматривается новый подход к обучению искусственных нейронных сетей (модульный принцип); описывается программный комплекс, предназначенный для моделирования ИНС, поддерживающий модульный принцип обучения.

ВВЕДЕНИЕ

На сегодняшний день *искусственные нейронные сети* (ИНС) представляют собой мощный инструмент моделирования, хорошо зарекомендовавший себя в научных исследованиях и практических приложениях. Нейронные сети широко используются в системах искусственного интеллекта, прогнозировании, распознавании образов при принятии решений и т. д. [1–4].

Однако для существующих программ моделирования ИНС (MATLAB – The MathWorks, Statistika – StatSoft, NNC – разработка вычислительного центра РАН и др.) характерны следующие недостатки:

- возможность осуществлять обучение только всей сети, в то время как при модульном подходе в моделировании было бы желательно осуществлять обучение отдельных ее частей без изменения настроек уже отлаженных модулей; отсутствие данной возможности не позволяет использовать модульный принцип моделирования и существенно увеличивает время обучения ИНС-модели;

- недостаточно развитые принципы структурной организации ИНС и «слабый» конструктор сети; данное обстоятельство приводит к существенной доле ручного труда и значительным временным затратам при построении сложных структур ИНС-моделей.

Поэтому идея разработки программного комплекса, предназначенного для моделирования ИНС, поддерживающего возможность обучения сети по частям (модульный принцип) и обладающего удобным конструктором для легкого и быстрого построения структуры сети (отдельными блоками), является актуальной.

Цель данной работы – разработать программный комплекс, предназначенный для моделирования искусственных нейронных сетей, отличием которого является модульный принцип обучения сети, развитый графический интерфейс и возможность ручного ввода весовых коэффициентов.

ПРЕИМУЩЕСТВА МОДУЛЬНОГО ПОДХОДА

Модульный принцип – процесс обучения нейронной сети отдельными частями (модулями). В искусствен-

ных нейронных сетях такой подход представляет собой поиск решения не во всем пространстве весовых коэффициентов (например, размерности n), а лишь в некоторой его части (в пространстве $n-k$, где k – число неизменяемых коэффициентов).

Использование модульного принципа обучения в ИНС предполагает *ряд весомых преимуществ*:

- высокая скорость обучения сети;
- легкость моделирования тех объектов, у которых уже известно устройство некоторых подсистем;
- более удобный метод подбора структуры нейронной сети;
- контроль над обучением.

Покажем преимущества модульного подхода к обучению на примере двух наиболее широко используемых нейронных сетей.

Преимущества модульного подхода при использовании полносвязного двухслойного персептрона.

Будем считать, что количество нейронов в скрытых слоях идентично числу нейронов входного слоя – n . Кроме того, для удобства деления нейронной сети на части будем считать, что n четно.

Данную нейронную сеть будем делить на две равные части (рис. 1).

Формула для расчета количества весовых коэффициентов в данном случае имеет вид:

$$N_{\text{кол}} = n \cdot L_1 + L_1 \cdot L_2 + L_2. \quad (1)$$

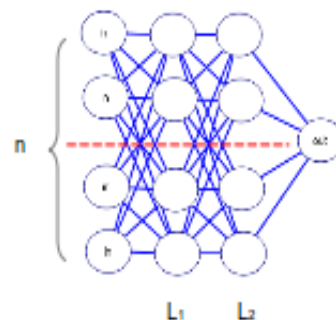


Рис. 1. Макет нейронной сети, разделенной на две части

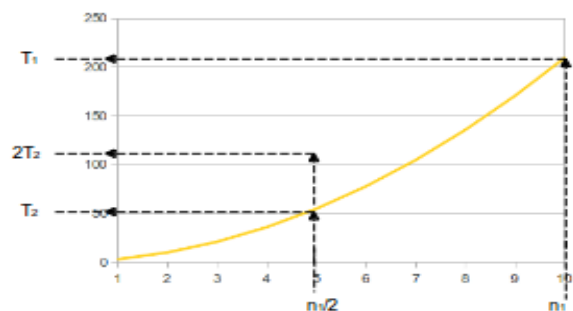


Рис. 2. График зависимости времени обучения от количества нейронов

Так как по условию:

$$n = L_1 = L_2, \tag{2}$$

тогда формулу (1) можно переписать в виде:

$$N_{\text{кол}} = n^2 + n^2 + n = 2n^2 + 1. \tag{3}$$

Приведенный график (рис. 2) показывает, что время, необходимое на обучение нейронной сети полностью (количество нейронов равно n_1), существенно больше, чем время последовательного обучения двух половин сети:

$$2 \cdot T_2 < T_1. \tag{4}$$

Преимущества модульного подхода при использовании ИНС, построенной по принципу А.Н. Колмогорова.

Теорема (Колмогорова): каждая непрерывная функция n переменных, заданная на единичном кубе n -мерного пространства, представима в виде [5]:

$$f(x_1, x_2, \dots, x_n) = \sum_{q=1}^{2n+1} h_q \left[\sum_{p=1}^n \varphi_p(x_p) \right], \tag{5}$$

где функции φ непрерывны, а функции h , кроме того, еще и стандартны, т. е. не зависят от выбора функции f .

В частности, каждая непрерывная функция двух переменных x и y представима в виде:

$$f(x, y) = \sum_{q=1}^5 h_q \left[\varphi_q(x) + \varphi_q(y) \right]. \tag{6}$$

Схема ИНС-модели, соответствующая теореме Колмогорова, представлена на рис. 3.

Аналогично предыдущему случаю будем делить сеть на две равные части. Формула для расчета количества весовых коэффициентов в рассматриваемом случае имеет вид:

$$N_{\text{кол}} = n \cdot (2n + 1) + n \cdot (2n + 1) + (2n + 1), \tag{7}$$

$$N_{\text{кол}} = 4n^2 + 4n + 1.$$

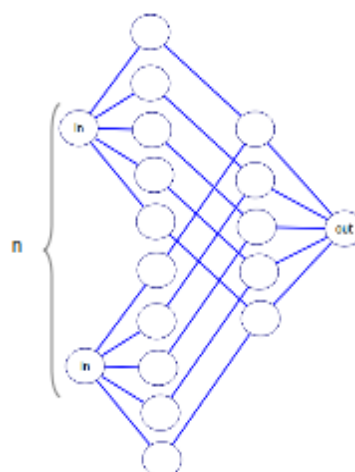


Рис. 3. Нейронная сеть, построенная по теореме Колмогорова

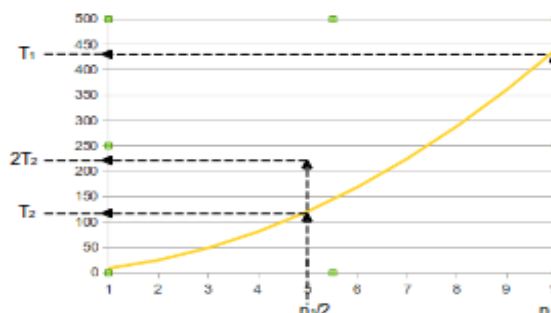


Рис. 4. График зависимости времени обучения от количества нейронов

Как и в предыдущем примере, зависимость времени обучения ИНС-модели от числа нейронов в сети является параболой второй степени.

Из приведенного графика (рис. 4) следует, что аналогично предыдущему примеру время, необходимое на обучение нейронной сети полностью (количество нейронов равно n_1), существенно больше, чем время последовательного обучения сети частями:

$$2 \cdot T_2 < T_1. \tag{8}$$

Таким образом, из приведенных примеров следует, что разработка программного комплекса, поддерживающего модульный принцип обучения сети, имеет преимущество.

**ПРОГРАММНЫЙ КОМПЛЕКС
ДЛЯ МОДЕЛИРОВАНИЯ ИНС С РЕАЛИЗАЦИЕЙ
МОДУЛЬНОГО ПРИНЦИПА ОБУЧЕНИЯ**

Программа разработана на языке программирования C++ в визуальной среде Qt Creator. Данный выбор был обусловлен некоторыми преимуществами программных средств: кросс-платформенность, бесплатность. Программный комплекс предназначен для за-

пуска на персональном компьютере под управлением ОС GNU/Linux или MS Windows. Программный комплекс защищен свидетельством о государственной регистрации программы для ЭВМ № 2012618141 [7].

Среди *основных возможностей программы* следует выделить:

- конструирование «больших» нейронных сетей (количество нейронов до 500) прямого распространения с любым количеством входов и выходов;
- возможность быстрого построения нейронной сети путем перемещения, копирования или удаления отдельных ее блоков;
- поддержка модульного принципа обучения;
- возможность ввода весовых коэффициентов вручную;
- поддержка основных алгоритмов обучения сети.

Графический интерфейс программы (рис. 5) представляет собой систему вкладок, каждая из которых отвечает за определенную функцию:

- 1) *сеть* – конструирование нейронной сети;
- 2) *выборка* – работа с данными для обучения сети;
- 3) *параметры* – алгоритм обучения сети;
- 4) *обучение* – остановка/начало обучения;
- 5) *расчет* – работа с обученной сетью.

ВЫЧИСЛИТЕЛЬНЫЕ ЭКСПЕРИМЕНТЫ

Работоспособность программного комплекса продемонстрируем с помощью двух стандартных для нейронных сетей примеров.

Восстановление внутренней структуры объекта мультипликативного типа по обучающей выборке с помощью ИНС-модели.

Цель данного эксперимента – «обучить» нейронную сеть операции умножения. Необходимо задать мультипликативную функцию вида:

$$Y = X_1 \cdot X_2. \quad (9)$$

Входные значения для обучающей выборки – X_1 и X_2 были сгенерированы случайным образом в интервале $[-10; 10]$ при помощи программы Microsoft Excel.

При попытке использования нейронной структуры с одним скрытым слоем, содержащим единственный нейрон, результат работы при любой активационной функции не соответствовал реальным значениям. Поэтому для получения результатов, наиболее близких к реальным, необходимо выразить операцию умножения через операции сложения и возведения в степень, как это показано в формуле (10):

$$(x_1 + x_2)^2 = x_1^2 + 2 \cdot x_1 \cdot x_2 + x_2^2, \\ x_1 \cdot x_2 = \frac{(x_1 + x_2)^2 - x_1^2 - x_2^2}{2}. \quad (10)$$

Структура нейронной сети, соответствующая формуле (10), имеет вид, представленный на рис. 6.

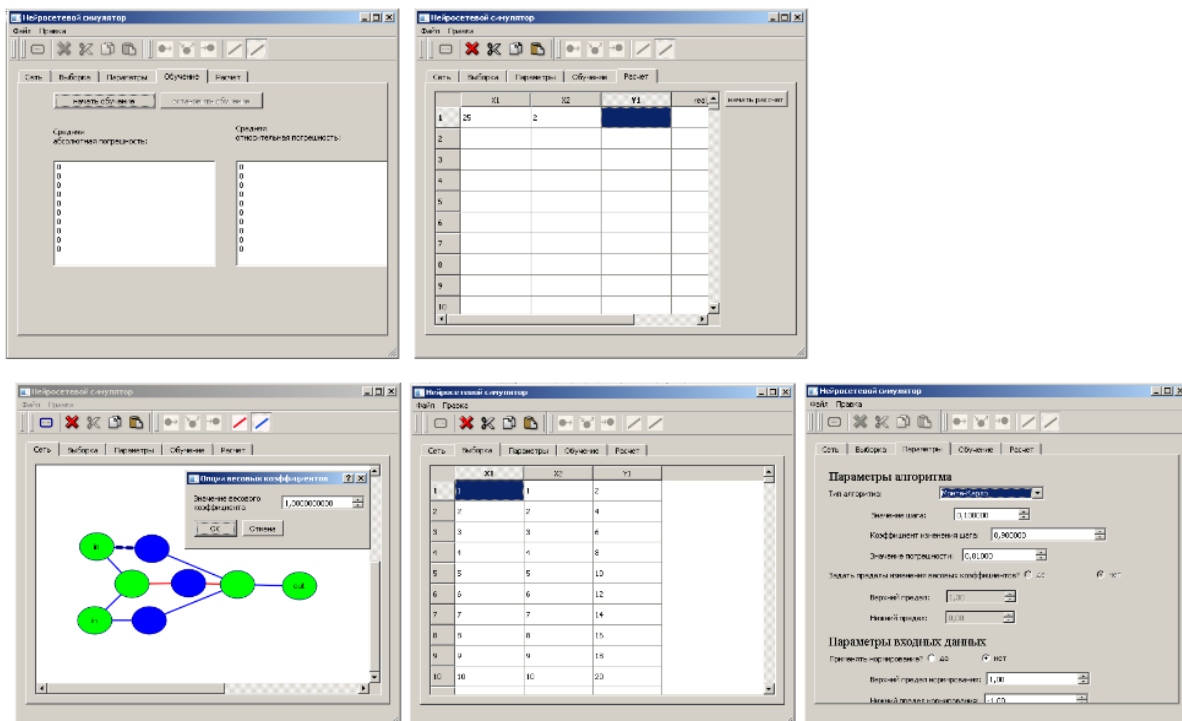


Рис. 5. Интерфейс программы

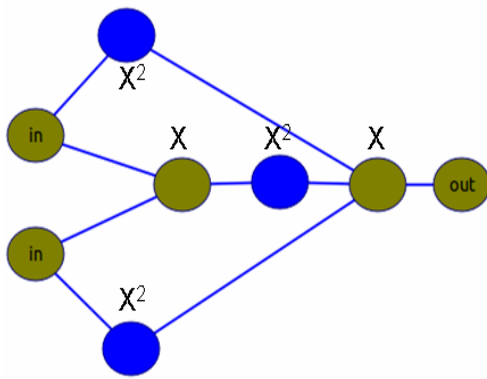


Рис. 6. Структура нейронной сети для операции умножения

	X1	X2	Y1	real_Y1
1	4	-7	-28	-28.06920016
2	-1	-10	10	9.905131799
3	10	1	10	9.976102052
4	6	-2	-12	-12.02092614
5	-6	8	-48	-48.10103897
6	5	9	45	44.93679828
7	-2	4	-8	-8.021611775
8	3	-5	-15	-15.03594166
9	3	0	0	-0.0025613...
10	0	10	0	-0.1001461...
11	-8	-9	72	71.94071812
12	10	2	20	19.97866021
13	7	-2	-14	-14.02573832
14	4	4	16	15.98832323
15	6	-3	-18	-18.02927097
16	-8	-5	40	39.97899984
17	3	7	21	20.9600484
18	9	-7	-63	-63.10716741
19	6	-8	-48	-48.10103897

Рис. 7. Результаты вычислений

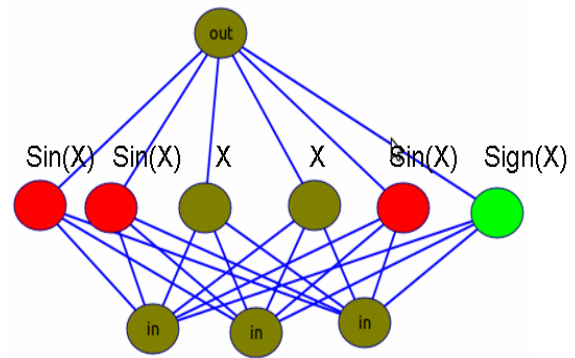


Рис. 8. Структура нейронной сети для прогнозирования

Для обучения созданной сети был выбран градиентный метод с заданным шагом, равным 0,01. Значение желаемой погрешности также равняется 0,01. Сравнение значений, полученных в результате работы нейронной сети, со значениями реального объекта представлено на рис. 7.

В данном эксперименте максимальная погрешность между реальным и вычисленным значениями составила 0,2 %.

Прогнозирование курса доллара по обучающей выборке с помощью ИНС-модели. Финансовое прогнозирование является одной из основных задач, решаемых при помощи нейронных сетей, поэтому в качестве второго эксперимента для проверки работоспособности программного комплекса было выбрано прогнозирование валютного курса американского доллара по отношению к рублю (использование выбранной валюты не имеет принципиального значения). Исходными данными для эксперимента послужили изменения курса американского доллара в период с 1 ноября 2011 г. по 29 февраля 2012 г.

Нейронная сеть для данного эксперимента, подобранная конструктивным методом [6], представлена на рис. 8.

Для обучения созданной сети был выбран градиентный метод с заданным шагом, равным 0,01. Значение желаемой погрешности также равняется 0,01. График на рис. 9 показывает сравнение реальных и вычисленных значений курса доллара за обучающий период

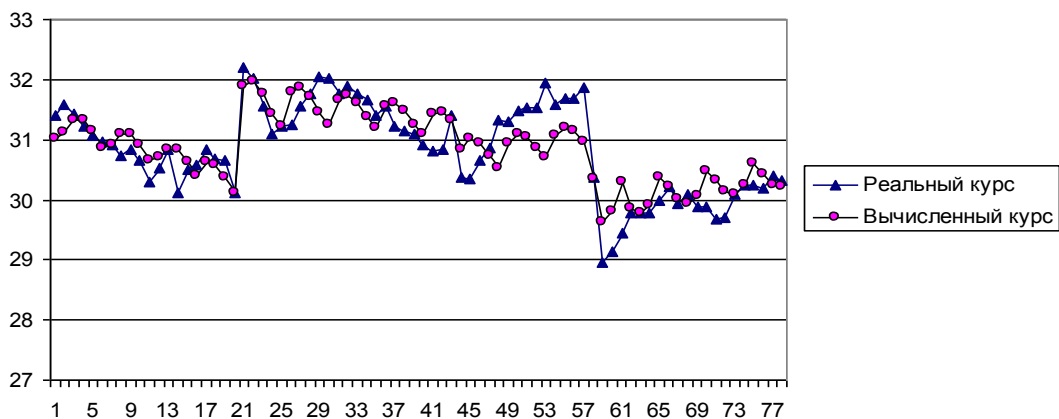


Рис. 9. Результат обучения нейронной сети

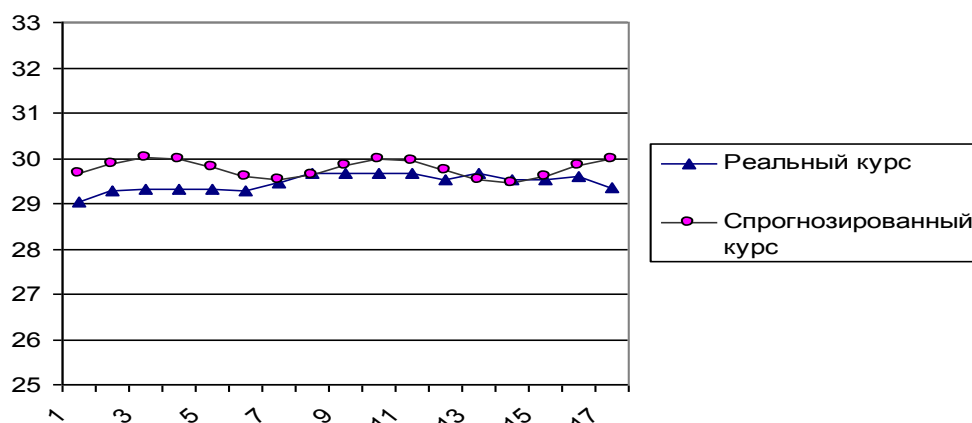


Рис. 10. Прогноз курса доллара

Дата		Курс	Прогно:	
17	3	12	29.36	29.99271074
16	3	12	29.58	29.83223746
15	3	12	29.51	29.59480931
14	3	12	29.51	29.45933896
13	3	12	29.67	29.52311391
12	3	12	29.54	29.73344216
11	3	12	29.66	29.93170654
10	3	12	29.66	29.97413035
9	3	12	29.66	29.83811515
8	3	12	29.66	29.63584561
7	3	12	29.45	29.52713036
6	3	12	29.29	29.59833504
5	3	12	29.3	29.79838953
4	3	12	29.3	29.97906025
3	3	12	29.3	30.00822769
2	3	12	29.29	29.87041993
1	3	12	29.03	29.67876597

Рис. 11. Результаты прогноза

(по оси абсцисс – индекс примера из обучающей выборки, по оси ординат – значение курса доллара).

Попытка прогнозирования была осуществлена на 17 дней вперед, а именно на период с 1 по 17 марта 2012 г. Результаты прогноза, произведенного программой, представлены на рис. 10.

Числовые значения курса американского доллара по отношению к рублю, вычисленные программой, оказались достаточно близкими к реальным значениям (рис. 11).

В данном эксперименте максимальная погрешность составила 2,4 %.

Таким образом, проведенные вычислительные эксперименты показали, что программный комплекс является пригодным для моделирования различных объектов и может применяться для решения практических задач.

ВЫВОДЫ

Разработан программный комплекс, предназначенный для моделирования искусственных нейронных сетей, отличием которого является модульный принцип обучения, развитый графический интерфейс и возможность ручного ввода весовых коэффициентов.

Реализованные возможности:

- реализован удобный графический интерфейс, предоставляющий легкую работу с нейронной сетью и обучающей выборкой;
- возможность сохранения разработанных данных в файл проекта;
- удобная загрузка как целых проектов, так и отдельных файлов;
- возможность «быстрого» построения нейронной сети прямого распространения (до 500 нейронов) путем перемещения, удаления или копирования отдельных ее блоков;
- реализованы наиболее распространенные алгоритмы обучения сети с поддержкой модульного принципа и возможностью параметризации каждого из них: простой градиентный метод, метод Монте-Карло, метод покоординатного спуска, а также метод простого сканирования;
- проведены вычислительные эксперименты, доказывающие, что программный комплекс является пригодным для моделирования различных объектов и может применяться для решения практических задач.

ЛИТЕРАТУРА

1. Арзамасцев А.А. Математическое и компьютерное моделирование: учебное пособие. Тамбов: ИМФИ ТГУ им. Г.Р. Державина, 2010.
2. Лекция № 5. Задача прогнозирования временного ряда. URL: <http://i-intellect.ru/lectures-of-neural-networks/5.html>. Загл. с экрана.
3. Аналитические технологии для прогнозирования и анализа данных. URL: <http://www.neuroproject.ru/tutorial.php>. Загл. с экрана.
4. Осовский С. Нейронные сети для обработки информации / пер. с польского И.Д. Рудинского. М.: Финансы и статистика, 2002. 344 с.

5. Колмогоров А.Н. О представлении непрерывных функций нескольких переменных суперпозициями непрерывных функций меньшего числа переменных // Докл. АН СССР. 1956. Т. 108. № 2. С. 179-182.
6. Арзамасцев А.А., Зенкова Н.А. Искусственный интеллект и распознавание образов: учебное пособие. Тамбов: ИМФИ ТГУ им. Г.Р. Державина, 2010.
7. Арзамасцев А.А., Рыков В.П., Крючин О.В. Симулятор искусственной нейронной сети с реализацией модульного принципа обучения. Свидетельство о государственной регистрации программы для ЭВМ № 2012618141 от 7.09.2012 г.

Поступила в редакцию 10 сентября 2012 г.

Arzamastsev A.A., Rykov V.P. ARTIFICIAL NEURAL NETWORK (ANN) WITH IMPLEMENTATION OF MODULAR PRINCIPLE OF TRAINING

A new approach to teaching of artificial neural networks (modular principle) is considered; a software package designed to model the INS, which supports modular training, is described.

Key words: artificial neural network; modular principle of training.