

УДК 54-145+519.254

## НЕКОТОРЫЕ АСПЕКТЫ ПАРАМЕТРИЗАЦИИ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ В ЗАДАЧАХ КАЧЕСТВЕННОГО ХИМИЧЕСКОГО АНАЛИЗА

© 2010 Я. Н. Краснянчин, А. В. Пантелеймонов, Ю. В. Холин

Изучены закономерности функционирования различных алгоритмов искусственных нейронных сетей при изменении ряда управляющих параметров. Предложена процедура формирования обучающей выборки, используемая для настройки нейронной сети. Показана удовлетворительная надежность классификации наборов тестовых данных при наличии в характеристиках исследуемых объектов грубых промахов с помощью аппарата нейронных сетей.

**Ключевые слова:** искусственная нейронная сеть, обучающая выборка, управляющий параметр, классификация, качественный анализ.

### ВВЕДЕНИЕ

Современный качественный химический анализ решает задачи идентификации и обнаружения аналитов, а также дискриминации образцов [1, 2]. Согласно [3, С. 22], идентификация "представляет собой отождествление аналита с известным химическим веществом или отнесение аналитического сигнала к известному веществу (соединению, группе соединений)". Дискриминация означает отнесение исследуемого образца к определенной классификационной единице (например, отнесение напитка к чаю, минеральной воде, соку и т.д. [2]). Таким образом, и идентификацию, и дискриминацию можно рассматривать как классификацию анализируемых образцов по набору их характеристик<sup>1</sup>.

Современный качественный анализ немислим без применения инструментальных методов (хроматография, различные спектроскопические методы и др.) [6]. Результаты применения этих методов приводят к получению многомерных массивов экспериментальных данных, к обработке которых зачастую необходимо применить современные методы анализа многомерных данных (методы хемометрии) – алгоритмы распознавания образов, классификации, дискриминантного анализа, искусственных нейронных сетей и т.д.

Для решения классификационных задач весьма часто привлекают методы управляемого обучения. Реализация этих методов предусматривает наличие обучающих (множество объектов с известной "классовой" принадлежностью) и тестовых выборок (множество исследуемых объектов, для которых "классовая" принадлежность неизвестна). Алгоритмы искусственных нейронных сетей стали мощным средством решения классификационных задач. Аппарат искусственных нейронных сетей позволяет работать с данными различной структуры и сложности [7–9]. Следует отметить, что для успешного применения алгоритмов нейронных сетей пользователю необходимо обладать более скромным уровнем вычислительных навыков, чем для использования традиционных методов статистического анализа.

Вопрос о параметрах сети, необходимых и достаточных для решения того или иного рода задач, представляет собой целое направление нейрокомпьютерной науки. Так как проблема синтеза искусственной нейронной сети напрямую зависит от решаемой задачи, выработать общие подробные рекомендации трудно.

В данной работе поставлены следующие цели: изучить закономерности функционирования алгоритмов нейронных сетей при различных значениях управляющих параметров (шаг обучения, среднеквадратичная ошибка, параметр функции активации, число нейронов) при классификации тестовых данных, определить оптимальный состав обучающей выборки, а также оценить робастность результатов предоставляемых реализованными алгоритмами нейронных сетей.

<sup>1</sup> Ср. с определениями: "Качественный анализ – анализ, в котором вещества идентифицируют или классифицируют на основе их химических или физических свойств, таких как химическая реакционная способность, растворимость, молекулярный вес, температура плавления, испускание или поглощение излучения, масс-спектры, периоды полураспада ядер и т.д." [4], "Qualitative analysis: classification of objects against specified criteria to meet agreed requirements" [5].

Для апробации алгоритмов классификации использовали различные наборы тестовых данных: данные с двуиерархической и дугообразной структурой [10]; 150 образцов цветков трех сортов ириса [11]; 178 образцов итальянских вин [12].

В работе использовали двухслойные нейронные сети (кроме нейронной сети встречного распространения, которая является трехслойной) (см. приложение А), а достаточное число нейронов определяли на основе оценки качества работы нейронной сети.

### ОЦЕНКА НАДЕЖНОСТИ И УСТОЙЧИВОСТИ АЛГОРИТМОВ НЕЙРОННЫХ СЕТЕЙ

Полученные результаты работы алгоритмов сетей сравнивали с результатами дискриминантного анализа (аналогом нейронной сети).

Имеющиеся наборы данных разделяли случайным образом на обучающую и тестовую выборки при различном их соотношении. Коэффициент  $T$  (%) указывает, какая доля общего числа образцов находится в обучающей выборке:

$$T = \frac{t}{M} \times 100, \% \quad (1)$$

где  $t$  – число образцов в обучающей выборке,  $M$  – общее число образцов.

Важно отметить, что вся информация, которая используется для настройки и обучения сети, содержится в наборе эталонов. Поэтому качество обучения сети, как и качество классификации анализируемых объектов (тестовой выборки), напрямую зависит от числа образцов в обучающей выборке, а также от того, насколько полно эти образцы описывают данную задачу.

О надежности алгоритмов судили по критериям качества работы нейросетей: доле неправильно классифицированных объектов тестовой выборки ( $P$ ) (при безошибочном обучении сети):

$$P = \frac{n}{N} \times 100, \% , \quad (2)$$

где  $n$  – число неправильно классифицированных объектов в тестовой выборке,  $N$  – общее число объектов в тестовой выборке; и средней квадратичной ошибке ( $mse$ , mean squared error)<sup>2</sup>:

$$mse = \frac{\sum_{i=1}^N (t_i - y_i)^2}{N} , \quad (3)$$

где  $t_i$  – целевой вектор,  $y_i$  – выходы сети.

Использование критерия  $mse$  возможно лишь в случае модельных данных, когда известны желаемые выходы сети для тестовой выборки. В случае классификации реальных данных применение критерия  $mse$  является невозможным.

Химические данные зачастую содержат в себе шум – нежелательную составляющую, природа которой может быть различной. Таким образом, для решения классификационных задач необходимы устойчивые к таким составляющим алгоритмы.

Для проверки устойчивости реализованных моделей применяли модель "грубых промахов". В исходные данные характеристик анализируемых объектов (тестовой выборки) вносили погрешности ( $\varepsilon$ ), рассчитываемые по формуле

$$\varepsilon = [(100 - q) \cdot \varepsilon_{Gauss}(0, \sigma) + q \cdot \varepsilon_{Laplas}(0, \sigma)] / 100 , \quad (4)$$

где  $q$ , % – интенсивность "грубых промахов";  $\varepsilon_{Gauss}$  – случайная величина, распределенная по закону Гаусса с нулевым средним и стандартным отклонением  $\sigma$ ;  $\varepsilon_{Laplas}$  – случайная величина, подчиняющаяся распределению Лапласа с нулевым средним и стандартным отклонением

$$\sigma = \frac{\sqrt{2}}{\alpha} :$$

<sup>2</sup> Ср. с обозначением средней квадратичной ошибки в (A.2).

$$p(\varepsilon) = \frac{\alpha}{2} \exp\{-\alpha|\varepsilon - \beta|\}, \quad \varepsilon \in (-\infty, \infty), \quad (5)$$

где  $p(\varepsilon)$  – плотность распределения,  $\alpha > 0$  – параметр масштаба,  $-\infty < \beta < +\infty$  – параметр сдвига.

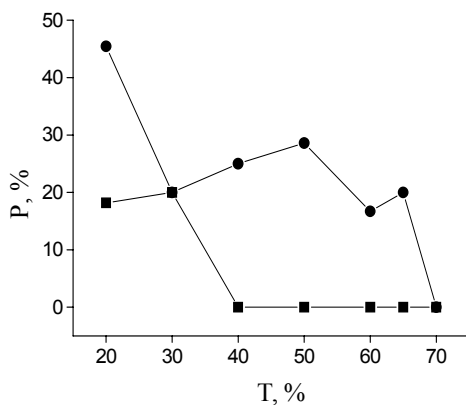
Поскольку коэффициент эксцесса распределения Лапласа ( $\gamma_2 = 3$ ) больше, чем нормального ( $\gamma_2 = 0$ ), с ростом  $q$  возрастает вероятность появления среди  $\varepsilon$  "грубых промахов" – погрешностей, более чем в два-три раза превышающих стандартные отклонения  $\sigma$  [13].

В исходные данные характеристик классифицируемых объектов вносили погрешности согласно (4) при  $q = 0, 10, \dots, 100\%$ .

Для определения оптимального числа образцов в обучающей выборке применяли искусственные нейронные сети, для реализации которых не требуется определение числа нейронов скрытого слоя: вероятностную нейронную сеть (PNN) и сеть радиального основания (RBN). Найденный оптимальный объем обучающей выборки применяли для обучения двунаправленной сети, сети прямой передачи сигнала и нейросети с линиями задержки на входах. Таким образом, изучаемые нейросети находятся в равных условиях, и можно судить о том, какая из предложенных архитектур нейронных сетей более надежна для классификации конкретной тестовой выборки.

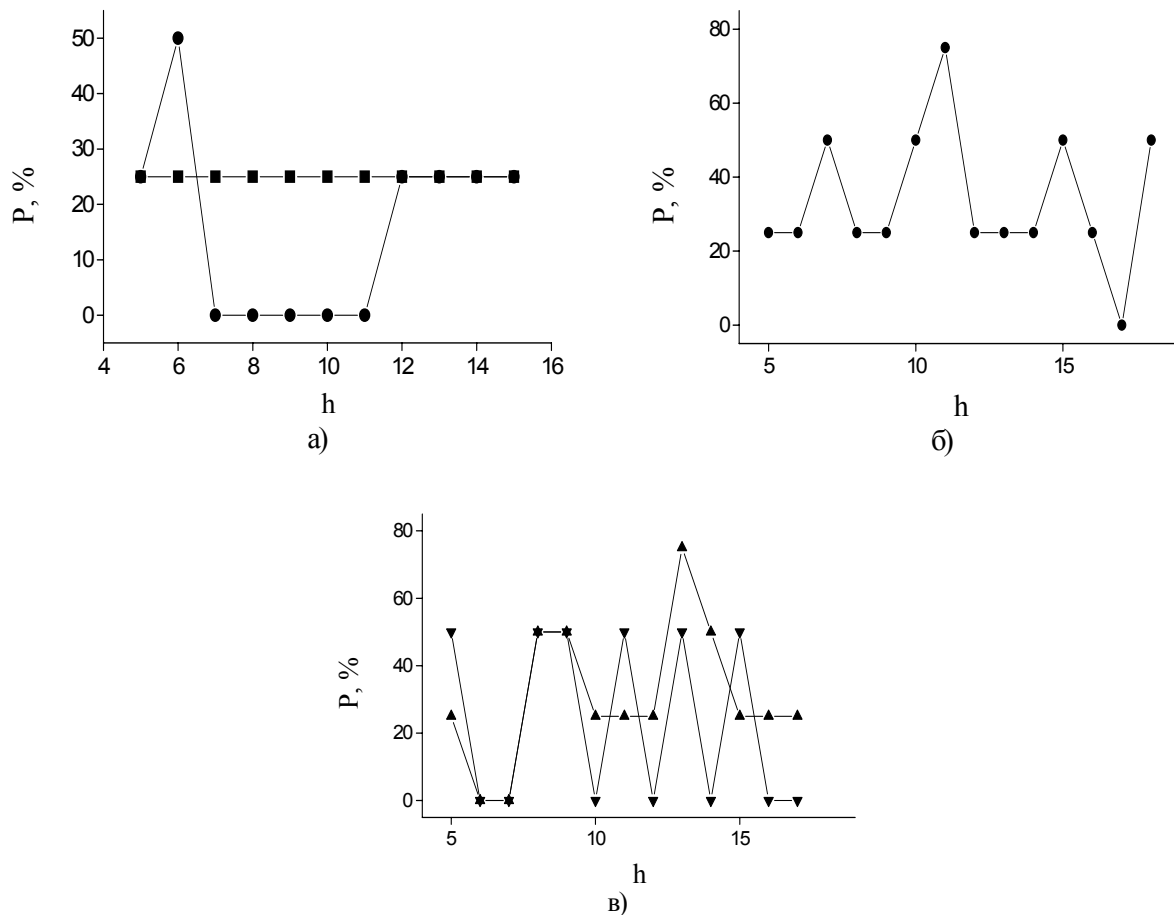
### ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

**Классификация данных с дугообразной структурой.** При классификации данных с дугообразной структурой нейронная сеть радиального основания ( $\delta = 1$ ) правильно классифицирует

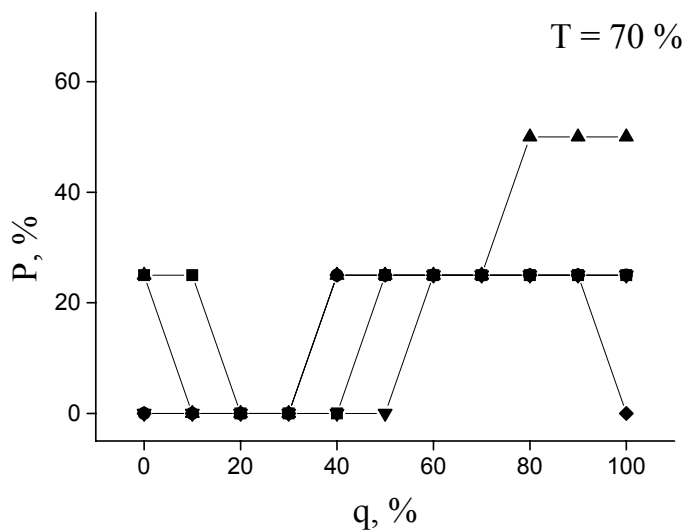


**Рис. 1.** Зависимость значений долей неправильно классифицированных образцов тестовой выборки данных с дугообразной структурой от объема обучающей выборки: —■— RBN, —●— PNN.

образцы в достаточно широком диапазоне значения коэффициента  $T$  ( $40 \leq T \leq 70$ ), при этом вероятностная нейронная сеть ( $\delta = 0.1$ ) способна правильно относить образцы в соответствующие группы только при  $T = 70\%$  (рис. 1). Оптимальные параметры архитектуры LVQ-сети следующие: число нейронов  $7 \leq h \leq 11$  при шаге обучения  $\eta = 0.03$  (рис. 2а). Динамическая сеть обучается безошибочно при числе нейронов  $5 \leq h \leq 18$ , что можно объяснить наличием линий задержки на входах, но правильное отнесение по классам образцов тестовой выборки достигается только при  $h = 17$  (рис. 2б). Для нейросети прямой передачи сигнала оптимальное число нейронов определялось при одновременной безошибочной классификации образцов тестовой и обучающей выборки. Данное условие выполняется при  $h = 6$  и  $h = 7$  (рис. 2в). Наиболее адаптивными алгоритмами для классификации данных с дугообразной структурой являются нейронная сеть прямой передачи сигнала и нейронная сеть с временными задержками: FFTDN безошибочно классифицирует образцы тестовой выборки при  $q = 50\%$ , а FFN при  $q = 100\%$ ; наименее адаптивной – двунаправленная нейронная сеть (рис. 3). Следует заметить, что дискриминантный анализ не отображает реальной структуры дугообразных данных.



**Рис. 2.** Зависимость значений долей неправильно классифицированных образцов тестовой выборки данных с дугообразной структурой от числа нейронов ( $T = 70\%$ ): а) при различных значениях коэффициента обучения для LVQN:  $\blacksquare$   $\eta = 0.01, 0.1, 0.05, 0.04$ ;  $\bullet$   $\eta = 0.03$ ; б)  $\bullet$  FFTDN; в)  $\blacktriangle$  FFN (тестовая выборка),  $\blacktriangledown$  FFN (обучающая выборка).

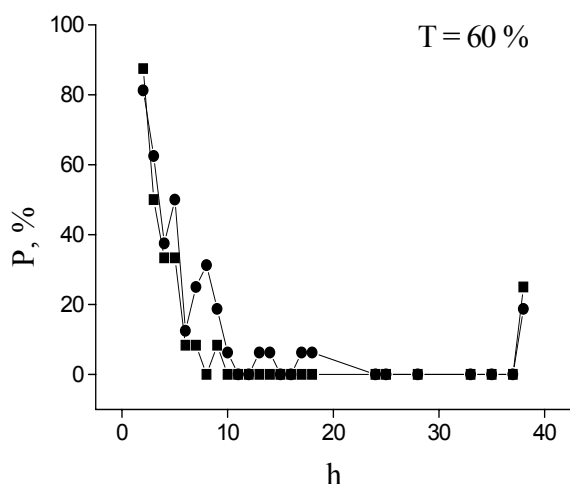


**Рис. 3.** Зависимость значений долей неправильно классифицированных образцов тестовой выборки данных с дугообразной структурой от доли грубых промахов, внесенных в исходные данные характеристик:  $\blacksquare$  PNN,  $\bullet$  RBN,  $\blacktriangle$  LVQN,  $\blacktriangledown$  FFTDN,  $\blacklozenge$  FFN.

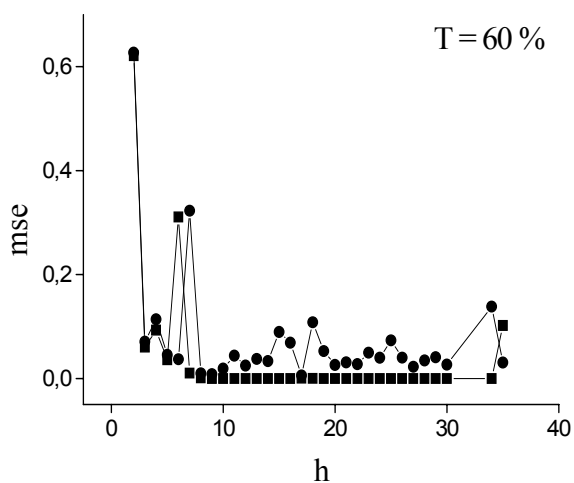
**Классификация данных с двуиерархической структурой.** В случае данных с двуиерархической структурой обучающая выборка формировалась таким образом, чтобы минимум два-три представителя всех девяти классов входили в ее состав для более полного описания поставленной задачи ( $T = 60\%$ ). В табл. 1 приведены значения управляющих параметров работы нейронных сетей для решения данной задачи. При обучении *LVQ*-сети добились минимального значения ошибки решения равного 5.4%, что соответствует 13 образцам. Такая большая ошибка при обучении ведет к неадекватному отнесению образцов тестовой выборки, поэтому исследование устойчивости указанного типа нейросети является нецелесообразным. На примере зависимости работы сети прямого распространения сигнала и сети с временными задержками от числа нейронов показано, что для оценки надежности алгоритмов можно использовать не только значение доли неправильно классифицированных объектов, но и значение средней квадратичной ошибки работы сети (рис. 4, 5). Сравнение долей неправильно классифицированных образцов тестовой выборки данных с двуиерархической структурой, полученных при использовании алгоритмов искусственных нейронных сетей и в результате применения дискриминантного анализа, указывает на то, что вероятностная нейронная сеть и сеть радиального основания наименее устойчивы к наличию в исходных данных "грубых промахов". Это можно объяснить отсутствием процедуры обучения. Предпочтительнее использовать нейронную сеть прямого распространения сигнала и динамическую сеть (рис. 6).

**Таблица 1.** Значения управляющих параметров реализации алгоритмов нейросетей для классификации данных с двуиерархической структурой

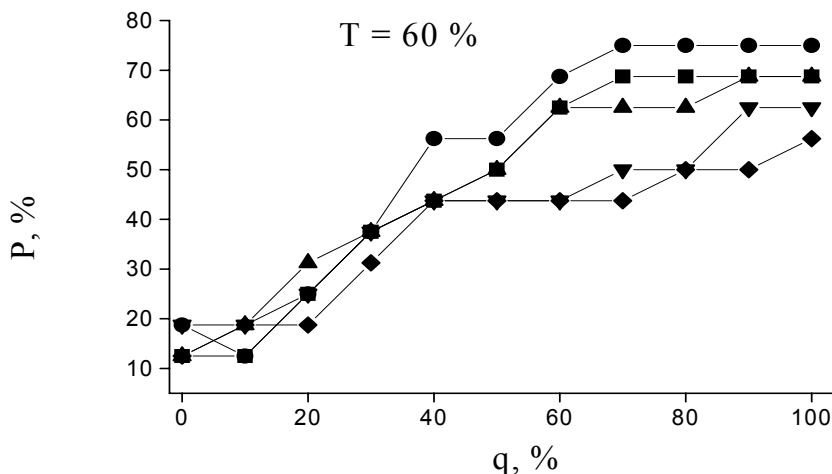
Нейронная сеть	Параметры
RBN	$goal = 0.3; \delta = 1.0$
PNN	$\delta = 0.1$
FFTDN	$9 \leq h \leq 14$
FFN	$h = 11, 12, 15, 16 - 24$
LVQN	$h = 38, \eta = 0.05$



**Рис. 4.** Зависимость значений долей неправильно классифицированных образцов тестовой выборки данных с двуиерархической структурой от числа нейронов для FFN: —■— обучающая выборка, —●— тестовая выборка.



**Рис. 5.** Зависимость значений средней квадратичной ошибки классификации образцов тестовой выборки данных с двуиерархической структурой от числа нейронов для FFTDN: —■— обучающая выборка, —●— тестовая выборка.



**Рис. 6.** Зависимость значений долей неправильно классифицированных образцов тестовой выборки данных с двумерной структурой от доли грубых промахов, внесенных в исходные данные характеристики: —■— PNN, —●— RBN, —▲— DA, —▼— FFTDN, —◆— FFN.

**Идентификация образцов цветков трех сортов ириса.** Безошибочная идентификация образцов ириса при использовании вероятностной нейронной сети ( $\delta = 0.1$ ), сети радиального основания ( $goal = 0.8$ ;  $\delta = 0.5$ ) и дискриминантного анализа наблюдается при  $T = 70\%$  (рис. 7), при этом при  $T = 68, 69, 71, 75\%$  ошибки в обучении сетей составляли от 2 до 8%. Для описания обучающей и тестовой выборок использовали следующие параметры: стандартное отклонение ( $s_i$ ) и размах ( $r_i$ ) каждой из характеристик исследуемых образцов:

$$s_i = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x}_i)^2}, \quad (6)$$

$$r_i = x_i^{\max} - x_i^{\min}, \quad (7)$$

где  $x_i$  – вектор значений характеристики объекта,  $\bar{x}_i$  – среднее значение характеристики,  $N$  – число образцов.

Оказалось, что, руководствуясь значениями указанных параметров, можно сформировать оптимальную выборку для обучения нейросетей, обеспечивающую адекватную последующую классификацию анализируемых образцов. Для этого достаточным является выполнение следующего условия: разница между значениями стандартного отклонения и размаха каждого из свойств исследуемых образцов в обучающей и тестовой выборках должна быть минимальной. В табл. 2 приведены значения суммы модуля разности стандартного отклонения и размаха четырех свойств образцов ириса в обучающей и тестовой выборках при их различных соотношениях:

$$S = \sum_{i=1}^M (|s_i^{обуч} - s_i^{тест}|), \quad (8)$$

$$R = \sum_{i=1}^M (|r_i^{обуч} - r_i^{тест}|), \quad (9)$$

где  $M$  – число характеристик, описывающих исследуемые объекты.

Действительно, при  $T = 70\%$  значения критериев  $S$  и  $R$  принимают минимальные значения из представленных в табл. 2 возможных вариантов.

Безошибочная идентификация образцов ириса при использовании  $FF$ - и  $FFTD$ -сетей наблюдается при  $h = 7; 19$  и  $h = 15$ , соответственно (рис. 8, 9). При обучении двунаправленной нейросети добились минимального значения ошибки сети равного 2.86% (3 образца) при следующих установочных параметрах:  $h = 40$  при  $\eta = 0.5$ .

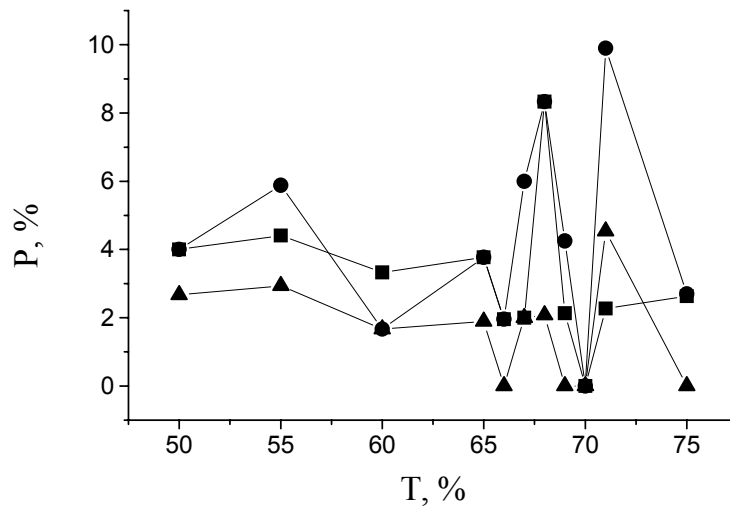


Рис. 7. Зависимость значений долей неправильно идентифицированных образцов ириса тестовой выборки от объема обучающей выборки: —■— PNN, —●— RBN, —▲— DA.

Таблица 2. Значения  $S$  и  $R$  при различном соотношении обучающей и тестовой выборок для идентификации образцов ириса

Параметр	$T, \%$					
	67	68	69	70	71	75
$S$	0.47	9.05	3.48	0.09	14.30	11.20
$R$	0.9	1.5	3.2	0.4	0.6	2.3

Наиболее устойчивыми к наличию "грубых промахов" в исходных характеристиках образцов ириса оказались двунаправленная сеть (несмотря на наличие ошибки при обучении) и сеть прямой передачи сигнала (рис. 10).

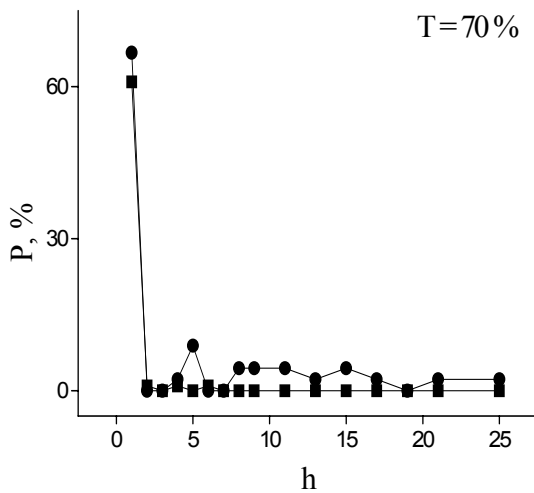


Рис. 8. Зависимость долей неправильно идентифицированных образцов ириса от числа нейронов для FFN: —■— обучающая выборка, —●— тестовая выборка.

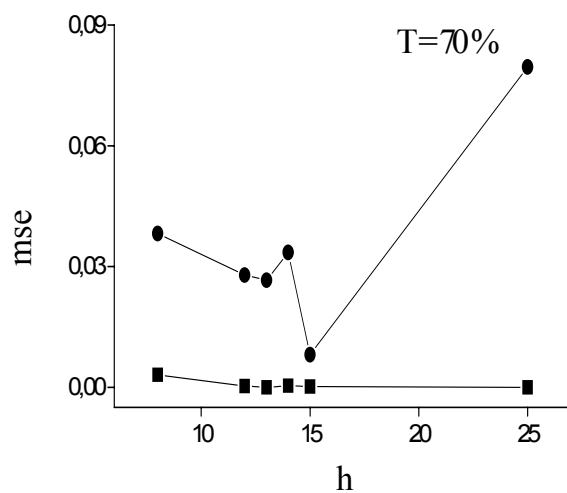
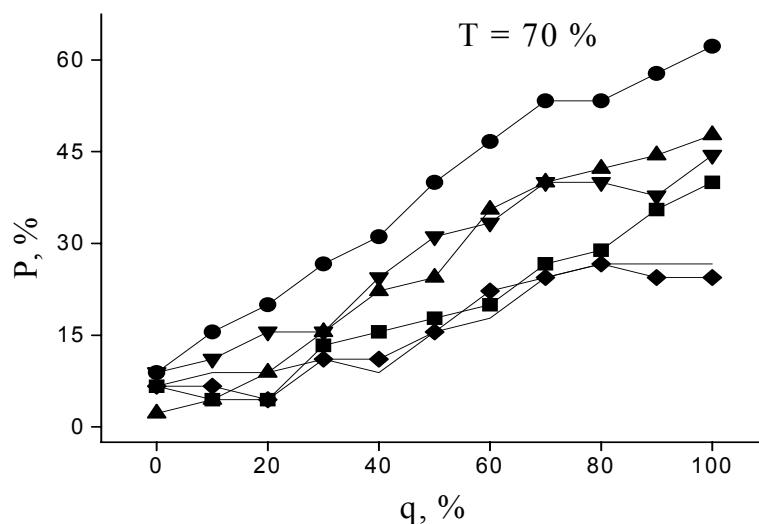


Рис. 9. Зависимость значений средней квадратичной ошибки идентификации образцов ириса от числа нейронов для FFTDN: —■— обучающая выборка, —●— тестовая выборка.



**Рис. 10.** Зависимость долей неправильно идентифицированных образцов ириса тестовой выборки от доли грубых промахов, внесенных в исходные данные характеристик: —■— PNN, —●— RBN, —▲— DA, —▼— FFTDN, —◆— FFN, —★— LVQN.

**Идентификация образцов вин.** Для адекватной идентификации образцов вин все реализованные алгоритмы нейронных сетей, кроме сети с временными задержками требуют предварительной нормализации данных путем приведения к нулевому среднему и единичной дисперсии. Это связано с большим разбросом значений характеристик, описывающих образцы вин (от долей единицы до нескольких сотен). Нормировка позволяет уравнивать вклад в модель от различных характеристик.

Для обоснования выбора значения  $T = 87\%$  в табл. 3 приведены значения суммы модуля разницы стандартных отклонений и размаха 13 свойств образцов вин (после процедуры нормализации) обучающей и тестовой выборок при  $T = 85, 86, 87, 88 \%$ .

В табл. 4 приведены значения управляющих параметров работы нейросетей для идентификации образцов вин.

**Таблица 3.** Значения  $S$  и  $R$  при различном соотношении обучающей и тестовой выборок для идентификации образцов вин

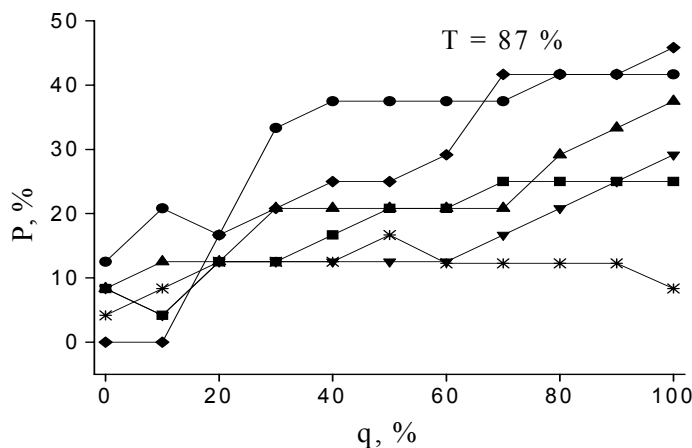
Параметр	$T, \%$			
	85	86	87	88
$S$	0.04	0.05	0.02	0.05
$R$	0.25	0.43	0.13	0.59

**Таблица 4.** Значения управляющих параметров реализации алгоритмов нейросетей для идентификации образцов вин

Нейронная сеть	Параметры
RBN	$goal = 0; \delta = 2.2$
PNN	$\delta = 0.1$
FFTDN	$h = 14$
FFN	$h = 14$
LVQN	$h = 40, \eta = 0.1$

При обучении LVQ-сети добились минимального значения ошибки равного 1.33 % (2 образца), и, не взирая на это, двунаправленная нейросеть обладает максимальной устойчивостью к наличию "грубых промахов" в исходных данных характеристик образцов вин (рис. 11).





**Рис. 11.** Зависимость значений долей неправильно идентифицированных образцов вин тестовой выборки от доли грубых промахов, внесенных в исходные данные характеристик: —■— PNN, —●— RBN, —▲— DA, —▼— FFTDN, —◆— FFN, —★— LVQN.

### ЗАКЛЮЧЕНИЕ

Обоснована процедура выбора параметров искусственных нейронных сетей различной архитектуры для решения задач классификации. Параметризация искусственных нейронных сетей зависит от сложности поставленной задачи. Приведенные результаты позволяют сделать следующие выводы: для реализации вероятностной сети достаточно задавать значение параметра отклонения гауссовой функции активации 0.1; для сети прямой передачи сигнала и динамической нейросети оптимальной функцией активации является сигмоидальная; наиболее устойчивыми к наличию в исходных данных грубых промахов являются нейронная сеть прямой передачи сигнала и двунаправленная нейросеть.

Сформулированные правила позволят упростить процедуру подбора параметров искусственных нейронных сетей для решения задач качественного химического анализа.

Авторы выражают благодарность репозитарию Center for Machine Learning and Intelligent Systems (School of Information and Computer Science, University of California, Irvine, CA) за предоставление доступа к массивам данных о характеристиках цветов ириса и образцов вин.

### ЛИТЕРАТУРА

1. Cardenas S., Valcarcel M. Analytical features in qualitative analysis // Trends Anal. Chem. – 2005. – V. 24, № 6. – P. 477–487.
2. Vlasov Yu., Legin A., Rudnitskaya A., Di Natale C., D'Amico A. Nonspecific sensor arrays ("electronic tongue") for chemical analysis of liquids // Pure Appl. Chem. – 2005. – V. 77, № 11. – P. 1965–1983.
3. Мильман Б.Л. Введение в химическую идентификацию. – СПб.: ВВМ, 2008. – 180 с.
4. IUPAC. Compendium of Chemical Terminology, 2<sup>nd</sup> ed. (the «Gold Book»). Compiled by A.D. McNaught, A. Wilkinson. Blackwell Scientific Publications, Oxford (1997); <http://goldbook.iupac.org/Q04973.html>.
5. Hardcastle W.A. (Editor). Qualitative analysis: a guide to best practice. – Cambridge: Royal Soc. of Chemistry, 1998. – 24 p.
6. Вершинин В.И., Дерендяев Б.Г., Лебедев К.С. Компьютерная идентификация органических соединений. – М.: Академкнига, 2003. – 197 с.
7. Balabin R.M., Safieva R.Z., Lomakina E.I. Gasoline classification using near infrared (NIR) spectroscopy data: comparison of multivariate techniques // Anal. Chim. Acta. – 2010. – V. 671. – P. 27–35.
8. Daszykowski M., Walczak B., Massart D.L. Density-based clustering for exploration of analytical data // Anal. Bioanal. Chem. – 2004. – V. 380. – P. 370–372.

9. Balabin R.M., Safieva R.Z., Lomakina E.I. Wavelet neural network (WNN) approach for calibration model building based on gasoline near infrared (NIR) spectra // Chemometr. and Intell. Lab. Systems. – 2008. – V. 93. – P. 58–62.
10. Коняев Д.С. Методы анализа данных и химической информатики в исследовании комплексообразования в растворах и на поверхности химически модифицированных кремнезёмов. Дисс. ... канд. хим. наук. – Х.: – 1999. – С. 139–141.
11. Iris Data Set (1988). UCI Machine Learning Repository [<http://archive.ics.uci.edu/ml/datasets/Iris>].
12. Wine Data Set (1991). UCI Machine Learning Repository [<http://archive.ics.uci.edu/ml/datasets/wine>].
13. Вучков И., Бояджиева Л., Солаков Е. Прикладной линейный регрессионный анализ. Пер. с болг. – М.: Финансы и статистика, 1987. – 239 с.
14. Дьяконов В.П., Круглов В.В. MATLAB 6.5 SP1/7/7 SP1/7 SP2+Simulink 5/6. Инструменты искусственного интеллекта и биоинформатики. Серия "Библиотека профессионала". – М.: СОЛОН-ПРЕСС, 2006. – 456 с.
15. Медведев В.С., Потемкин В.Г. Нейронные сети. MATLAB 6 / Под общ. ред. В.Г. Потемкина. – М.: ДИАЛОГ-МИФИ, 2002. – 496 с.
16. Заенцев И.В. Нейронные сети: основные модели. Учебное пособие к курсу "Нейронные сети". – Воронеж: Воронежский гос. ун-т, 1999. – 76 с.

## Приложение А

### Краткая характеристика нейронных сетей

К настоящему времени предложено и изучено большое количество нейросетевых архитектур. Каждый алгоритм имеет свои достоинства и недостатки, особенности реализации и последующего обучения. Вопрос создания нейронной сети решается в два этапа:

- 1) выбор типа сети – определение числа нейронов, вида активационной функции;

**Таблица А.1.** Нейронные сети и соответствующие характеристики

Название	Описание архитектуры	Управляющие параметры реализации сети
Сеть с радиальными базисными функциями (Radial Basic Function Network, RBN)	Двухслойная сеть без обратных связей, которая содержит скрытый слой радиально-симметричных скрытых нейронов, каждый из которых предназначен для хранения отдельного эталонного вектора. Выходной слой является линейным.	Значение допустимой среднеквадратичной ошибки работы сети ( <i>goal</i> ); степень сглаживания или отклонение гауссовой функции активации ( $\delta$ ).
Вероятностная нейронная сеть (Probabilistic Neural Network, PNN)	Архитектура базируется на архитектуре радиальной базисной сети, но в качестве второго слоя используют конкурирующий слой, который подсчитывает вероятность принадлежности входного вектора к тому или иному классу.	Степень сглаживания ( $\delta$ ).
Двунаправленная нейронная сеть (Learning Vector Quantization Network, LVQN)	Нейронная сеть встречного распространения содержит два слоя с последовательными связями. Первый слой – Кохонена (конкурирующий), второй – Гроссберга (линейный).	Число нейронов скрытого слоя ( <i>h</i> ), коэффициент обучения ( $\eta$ ).
Сеть прямой передачи сигнала (Feed Forward Network, FFN)	"Классическая" сеть с обучением по методу обратного распространения ошибки. Все связи направлены строго от входных нейронов к выходным. Выходной слой является линейным.	Число нейронов скрытого слоя ( <i>h</i> ), вид функции активации скрытого слоя.
Сеть с временной задержкой или динамическая сеть (Time Delay Feed Forward Neural Network, FFTDN)	Двухслойная сеть прямой передачи сигнала с линиями задержки, использующая метод обратного распространения ошибки. Выходной слой является линейным.	Число скрытых нейронов ( <i>h</i> ), вид функции активации скрытого слоя.

2) обучение сети – подбор таких значений ее весов, чтобы сеть работала нужным образом. Обучить нейросеть – значит сообщить ей, чего мы от нее добиваемся [14].

Обучение осуществляется путем последовательного предъявления входных векторов с одновременной подстройкой весов в соответствии с определенной процедурой. Обучение с учителем предполагает, что для каждого входного вектора существует целевой вектор, представляющий собой требуемый выход сети. В данном случае имеется в виду принадлежность к классам образцов обучающей выборки.

Ряд алгоритмов, принципиально отличающихся по архитектуре, реализован в MATLAB 6.5 [14, 15]. В табл. А.1 представлены типы рассмотренных в данной работе нейронных сетей, а также их основные характеристики.

Для понимания влияния управляющих параметров на поведение нейросети необходимо обсудить некоторые аспекты реализации представленных в табл. А.1 алгоритмов.

**Радиально-базисная сеть.** Обучение *RB*-сети подразумевает предварительное проведение кластеризации для определения центров классов (эталонов). Это могут быть случайно выбранные векторы из обучающей выборки или векторы, выбранные в результате функционирования алгоритма *k*-средних. После определения расположения центров необходимо задать гауссову функцию активации:

$$f_i = \exp \left[ - \frac{\sum_{j=1}^N (x_j - c_{ij})^2}{\delta_i^2} \right], \quad i = 1, \dots, L, \quad (\text{A.1})$$

где  $X = (x_1, \dots, x_N)$  – массив характеристик исследуемых объектов,  $C_i = (c_{i1}, \dots, c_{iN})$  – весовой вектор *i*-го эталонного нейрона скрытого слоя (центры кластеров), *L* – число эталонов,  $\delta_i$  – отклонение (сглаживающий фактор) гауссовой функции.

Величина отклонения определяет остроту пика гауссовой функции. Если эти функции выбраны слишком острыми, то сеть потеряет способность к обобщению; если же гауссовы функции чересчур широкие, то сеть не будет воспринимать мелкие детали (явление пере/недообучения). Как правило, отклонение выбирается таким образом, чтобы область значений каждой гауссовой функции захватывала несколько соседних центров; зачастую этому параметру присваивают значение 1.

Средняя квадратичная ошибка работы *RB*-сети подразумевает под собой ошибку работы сети в результате обучения:

$$goal = \frac{\sum_{i=1}^Q (t_i - y_i)^2}{Q}, \quad (\text{A.2})$$

где *Q* – число образцов в обучающей выборке.

Стараются обеспечить как можно меньшее значение среднеквадратичной ошибки, в идеальном случае – 0.

**Двухнаправленная нейронная сеть.** Обучение *LVQ*-сети выполняется комбинированным методом. Обучение слоя нейронов Кохонена заключается в подстройке векторов весов выбранного (выигравшего) нейрона согласно формуле коррекции весов:

$$w_n = w_{n-1} + \eta(x_i - w_{n-1}), \quad (\text{A.3})$$

где *n* – итерация,  $w_n$  – новое значение вектора весов, соединяющего входной компонент  $x_i$  с выигравшим нейроном,  $w_{n-1}$  – значение этого вектора на предыдущей итерации,  $\eta$  – коэффициент скорости обучения или шаг обучения.

Слой Гроссберга обучается с учителем, алгоритм использует заданные желаемые выходы (целевой вектор). При этом подстраиваются только те веса нейронов данного слоя, которые соединены с выигравшим нейроном слоя Кохонена:

$$w_n = w_{n-1} + \eta(t_i - w_{n-1}). \quad (\text{A.4})$$

**Сеть прямой передачи сигнала, динамическая нейросеть.** Для промежуточного слоя нейросети прямой передачи сигнала и динамической сети использовали сигмоидальную функцию активации для всех тестовых данных:

$$f(s) = \frac{e^s - e^{-s}}{e^s + e^{-s}}, \quad (\text{A.5})$$

$$s = \sum_{i=1}^n w_i x_i + b, \quad (\text{A.6})$$

где  $b$  – значение смещения нейрона.

Для выбора числа нейронов и слоев в сети нет строго определенной процедуры. Чем больше количество нейронов и слоев, тем шире возможности сети, но тем медленнее она обучается и работает. Если в сети слишком мало нейронов или слоев, то сеть не обучится, и ошибка при ее работе останется большой; если нейронов или слоев слишком много – сеть будет неспособна к обобщению результатов. Подбор необходимого числа нейронов зависит от многих факторов, в числе которых размерность задачи, объем обучающих данных и пространственная структура активационной функции [16].

*Поступила в редакцию 5 августа 2010 г.*

Я. М. Краснянчин, А. В. Пантелеймонов, Ю. В. Холин. Деякі аспекти параметризації штучних нейронних мереж у задачах якісного хімічного аналізу.

Вивчені закономірності функціонування різних алгоритмів штучних нейронних мереж при зміні ряду управляючих параметрів. Запропонована процедура формування навчальної вибірки, що використовується для налаштування нейронної мережі. Показано задовільну надійність класифікації наборів тестових даних за наявності в характеристиках досліджуваних об'єктів грубих промахів за допомогою апарату нейронних мереж.

**Ключові слова:** штучна нейронна мережа, навчальна вибірка, управляючий параметр, класифікація, якісний аналіз.

Ya. N. Krasnianchyn, A. V. Panteleimonov, Yu. V. Kholin. Parameterization of artificial neural networks in solving tasks of qualitative chemical analysis.

Regularities of functioning of different algorithms of artificial neural networks were studied via variation of command parameters. Procedure of forming the learning sample for tuning neural networks is offered. The adequacy of algorithms under study has been demonstrated for classification of test samples. Sufficient reliability of algorithms was also demonstrated for data sets with gross errors.

**Key words:** artificial neural network, learning sample, command parameter, classification, quality analysis.

Kharkov University Bulletin. 2010. № 932. Chemical Series. Issue 19(42).