



ровании новых задач интеллектуального управления данный аппарат на основе заданных аспектов и предикатов позволит производить адаптацию общей компьютерной модели системы к их автоматизированному решению с минимальными временными затратами на выполнение соответствующих вычислительных экспериментов.

ЛИТЕРАТУРА

1. Оптимальное управление природно-экономическими системами / под ред. В.И. Гурмана, А.И. Москаленко. – М.: Наука, 1980.
2. Бурков В.Н., Новиков Д.А., Щепкин А.В. Модели и механизмы управления эколого-экономическими системами // Проблемы управления. – 2009. – № 1. – С. 2-7.
3. Малинецкий Г.Г., Маненков С.К., Митин Н.А., Шишов В.В. Когнитивный вызов и информационные технологии // Препринты ИПМ им. М.В.Келдыша. – 2010. – № 46. URL: http://library.keldysh.ru/preprint.asp?id=2010_46.
4. Дмитриев В.М., Ганджа Т.В., Затик О.С. Компьютерная модель эколого-экономической системы нефтегазодобывающего региона для формирования экологических программ // Информатика и системы управления. – 2011. – № 4(30). – С 27-39.
5. Дмитриев В.М., Ганджа Т.В., Коротина Т.Ю. Генератор моделей компонентов с энергетическими связями физически неоднородных цепей на базе интерактивной математической панели // Доклады Томского государственного университета систем управления и радиоэлектроники. – 2009. – № 2 (20). – С.94-99.

Статья представлена к публикации членом редколлегии А.А. Шелупановым.

E-mail:

Дмитриев Вячеслав Михайлович – ofo45@mail.ru;

Ганджа Тарас Викторович – gandgatv@gmail.com;

Затик Ольга Сергеевна – olga_sur@mail.ru.

УДК 004.93

© 2014 г. С.А. Субботин, канд. техн. наук
(Запорожский национальный технический университет, Украина)

ОЦЕНКА СВОЙСТВ НЕЙРОМОДЕЛЕЙ В ЗАДАЧАХ ДИАГНОСТИРОВАНИЯ И АВТОМАТИЧЕСКОЙ КЛАССИФИКАЦИИ

Предложен комплекс показателей, позволяющих оценивать свойства нейронных сетей. Применение предложенных показателей на практике позволяет автоматизировать процесс построения нейросетевых моделей для решения задач диагностики и автоматической классификации.

Ключевые слова: нейронная сеть, пластичность, однородность, параллелизм, иерархичность, предорганизация, функциональная блочность, обобщение, обучаемость, интерпретабельность, надежность, распределенность памяти, вариативность, эффективность.

Введение

Для построения диагностических и распознающих моделей по прецедентам широкое распространение на практике получили нейромодели, представляющие собой массивованно-параллельные соединения простых, обычно адаптивных элементов (нейронов), имеющих непустое множество входов, вместе с их иерархической организацией, обладающее свойствами нелинейности (целое суммы частей), нелокальности (дальнодействия), нестационарности (обучаемости) и обратной связи (исключение бесконечного регресса) [1].

Наиболее значимыми классами сетевых вычислительных структур являются искусственные нейронные [1, 2] и нейро-нечеткие сети [3]. В настоящее время разработан широкий спектр моделей и методов построения нейронных и нейро-нечетких сетей [1 – 3]. Такое разнообразие существующих моделей приводит к необходимости решения задачи автоматизации анализа свойств и сравнения сетевых вычислительных моделей.

Целью данной работы является создание математического обеспечения для автоматизации анализа свойств и сравнения сетевых вычислительных моделей.

Постановка задачи

Пусть задана обучающая выборка $\langle x, y \rangle$, где $x = \{x^s\}$, $x^s = \{x_j^s\}$, $y = \{y^s\}$; $s = 1, 2, \dots, S$, $j = 1, 2, \dots, N$, x^s – s -й экземпляр выборки; x_j^s – значение j -го признака s -го экземпляра; S – число экземпляров выборки; N – число признаков; y^s – значение выходного признака, сопоставленное s -му экземпляру выборки.

Тогда задача построения диагностической (распознающей) модели $y = f(w, x)$ по выборке прецедентов $\langle x, y \rangle$ (задача обучения с учителем) заключается в идентификации такой структуры модели f и значений ее параметров w , при которых достигается экстремум заданной целевой функции обучения F .

При использовании нейронных сетей в качестве базиса для построения диагностических и распознающих моделей необходимо обеспечить выбор наилучшей из множества возможных моделей, что требует выделения наиболее значимых их характеристик.

Поскольку наиболее широко используемым типом сетей являются слоистые сети прямого распространения, будем характеризовать сети кортежем $\langle f, w \rangle$, где $f = \langle M, \{N_\eta\}, \{\langle \varphi^{(\eta,i)}, \psi^{(\eta,i)} \rangle\} \rangle$, $w = \{w_j^{(\eta,i)}\}$; M – число слоев сети; N_η – число нейронов в η -м слое сети; $w_j^{(\eta,i)}$ – весовой коэффициент j -го входа i -го нейрона η -го слоя; $\varphi^{(\eta,i)}$ – дискриминантная функция i -го нейрона η -го слоя $\psi^{(\eta,i)}$ – функция активации i -го нейрона η -го слоя. Также обозначим N_n – число нейронов в сети; $\varphi(i)$ и $\psi(i)$ – соответственно дискриминантная и активационная функции i -го нейрона (при сплошной нумерации нейронов сети); $w_{i,j}$ – вес связи между i -м и j -м нейронами сети (если связь отсутствует, примем $w_{i,j} = 0$), где $i, j = 1, 2, \dots, N_n$.

Для нейронных сетей чаще всего на практике структура модели f задается пользователем либо подбирается путем эволюционной оптимизации, а значения па-

раметров w настраиваются в основном на основе градиентных и реже – эволюционных методов. Для автоматизации сравнительного анализа нейромоделей необходимо разработать набор показателей, а также методы их расчета, отражающие важнейшие свойства нейромоделей. Целесообразно при этом обеспечить такой масштаб шкалы критериев, при котором они будут легко объединяться друг с другом.

Анализ литературы

Наиболее широко используемыми критериями сравнения нейросетевых моделей, используемых для задания целевых функций F , являются точность, характеризующая ошибкой обучения и ошибкой распознавания тестовых данных, время и количество итераций, затраченных на процесс обучения, время распознавания тестовых данных, число слоев и нейронов сети [1, 2]. В [2] предложен ряд способов для оценки сложности нейросетей (НС). Однако такой набор показателей отражает лишь небольшую часть свойств нейромоделей. Поэтому автором в [4, 5] был предложен комплекс критериев, характеризовавших сложность, избыточность памяти, логическую прозрачность и качество аппроксимации нейросетевых моделей.

Однако такие важные свойства НС как иерархичность [6], пластичность [7], однородность [2], параллелизм вычислений [1, 2], предорганизация в обучении [6], функциональная блочность [6], обучаемость [1, 7], интерпретируемость [2, 4, 5, 8], надежность [9 – 11], обобщение [7, 12], распределенность памяти [1, 2], вариативность моделей [2, 13], эффективность до настоящего времени либо не имели аналитических формул для количественного выражения, либо оценивались показателями с различными диапазонами значений. Поэтому в настоящей работе именно данные свойства выбраны для анализа.

Свойства структуры нейромодели

К свойствам архитектуры нейромодели следует отнести пластичность, однородность, параллелизм, иерархичность, предорганизацию в обучении и функциональную блочность.

Пластичность обуславливает сложность поведения НС, которое рассматривается как результат взаимодействия многих элементов. Каждый из них ограничивает действие других и сам ограничивается другими на пути к формированию глобального наблюдаемого поведения. Различают нейронную пластичность (в качестве пластических элементов рассматриваются нейроны), а также синаптическую пластичность (модификация силы синаптической связи между нейронами) [7].

Относительный коэффициент нейронной пластичности зададим как:

$$I_{np} = \frac{\sum_{i=1}^{N_{н.}} g_{np}(\varphi(i))g_{np}(\psi(i))}{N_{н.}^{\max} g_{np}^{\max}(\varphi)g_{np}^{\max}(\psi)},$$

где $N_{н.}^{\max}$ – максимально возможное количество нейронов; $g_{np}(element)$ – харак-

характеристика пластичности функционального элемента *element* нейрона (будем полагать $\mathcal{G}_{np}(element)$ равным числу возможных состояний элемента *element* для дискриминантной функции *i*-го нейрона $\mathcal{G}_{np}(\varphi(i))$, равна числу возможных видов дискриминантных функций, которые могут быть заданы для данного нейрона; для функции активации *i*-го нейрона $\mathcal{G}_{np}(\psi(i))$ будет равняться числу возможных видов функций активации, которые могут быть заданы для данного нейрона); $\mathcal{G}_{np}^{\max}(\varphi)$ и $\mathcal{G}_{np}^{\max}(\psi)$ – соответственно максимальные характеристики пластичности дискриминантных и активационных функций среди всех нейронов сети. Относительный коэффициент нейронной пластичности сети будет принимать значения в диапазоне от 0 до 1, и чем больше его значение, тем выше уровень нейронной пластичности сети.

Относительный коэффициент синаптической пластичности сети определим как:

$$I_{sp} = \frac{\sum_{i=1}^{N_n} \sum_{j=1}^{N_n} \mathcal{G}_{sp}(i, j)}{N_n \cdot 2 \cdot \text{round}\left(\frac{w^{\max} - w^{\min}}{\Delta w}\right)},$$

где w^{\max} и w^{\min} – соответственно максимально и минимально возможные значения весов сети; Δw – минимально возможное изменение веса с учетом разрядности вычислительной сетки ЭВМ; round – функция округления к ближайшему целому числу; $\mathcal{G}_{sp}(i, j)$ – характеристика пластичности связи от *i*-го нейрона к *j*-му ($\mathcal{G}_{sp}(i, j) = 0$, если связь не может существовать либо связь жестко задана и ее вес не может изменяться; в остальных случаях:

$$\mathcal{G}_{sp}(i, j) = \text{round}\left(\frac{w_{i,j}^{\max} - w_{i,j}^{\min}}{\Delta w_{i,j}}\right),$$

здесь $w_{i,j}^{\max}$ и $w_{i,j}^{\min}$ – соответственно максимально и минимально возможные значения веса связи от *i*-го нейрона к *j*-му; $\Delta w_{i,j}$ – минимально возможное изменение веса с учетом размера разрядной сетки ЭВМ. Относительный коэффициент синаптической пластичности сети будет принимать значения в диапазоне от 0 до 1, и чем больше его значение, тем выше уровень синаптической пластичности сети при фиксированном количестве нейронов.

Относительный коэффициент пластичности сети определим на основе введенных относительных коэффициентов нейронной и синаптической пластичностей: $I_{pl} = I_{np} I_{sp}$. Относительный коэффициент пластичности сети будет принимать значения в диапазоне от 0 до 1, и чем больше его значение, тем выше уровень пластичности сети, следовательно, сеть обладает большими адаптационными способностями.

Однородность нейроэлементов сети заключается в том, что нейросети строятся из множества простых унифицированных типовых элементов (нейронов),

выполняющих элементарные действия (умножение, сложение, вычисление простейшей нелинейной функции) и соединенных между собой различными связями [2].

Однородность дискриминантных функций будем оценивать в виде:

$$I_{hd} = \frac{2 \sum_{i=1}^{N_H} \sum_{j=i}^{N_H} \{1 | \varphi(i) \equiv \varphi(j)\}}{N_H(N_H - 1)}.$$

Однородность функций активации предлагается определять по формуле:

$$I_{ha} = \frac{2 \sum_{i=1}^{N_H} \sum_{j=i}^{N_H} \{1 | \psi(i) \equiv \psi(j)\}}{N_H(N_H - 1)}.$$

Однородность нейронов сети определим как:

$$I_{hn} = \frac{2 \sum_{i=1}^{N_H} \sum_{j=i}^{N_H} \{1 | (\psi(i) \equiv \psi(j)) \wedge (\varphi(i) \equiv \varphi(j))\}}{N_H(N_H - 1)}.$$

Показатели однородности будут изменяться от 0 до 1, и чем больше их значения, тем более однородны соответствующие элементы сети.

Параллельная архитектура и распределенность вычислений – обработка информации в НС выполняется несколькими процессорными элементами, в то время как в обычных компьютерах, имеющих только один центральный процессор, информация обрабатывается последовательно, по шагам [1, 2].

Параллелизм вычислений сети прямого распространения без латеральных связей будем характеризовать с помощью критерия:

$$I_{par} = \frac{\sum_{\eta=1}^M \frac{1}{\eta} \sum_{i=1}^{N_\eta} (N_{\eta-1} + 3)}{N_w + 2N_H}.$$

Иерархическая организация структуры, целостность и дробимость элементов НС: использование иерархических структур позволяет строить модели сложных объектов из более простых; работа иерархической структуры требует, чтобы информационный элемент в каждом иерархическом уровне вел себя как единое целое, но при переходе с уровня на уровень допускал дробление, причем при переходе с верхнего иерархического уровня на нижний это дробление соответствует выделению составляющих его элементов, а при переходе с нижнего уровня на верхний оно соответствует включению определенной части этого элемента в более сложный объект [6].

Определим иерархичность сети как:

$$I_h = \frac{\sum_{\eta=1}^M \eta N_\eta}{MN_H}.$$

Предорганизация в обучении – свойство сети, отражающее ее пригодность

для решения задачи в начале процесса структурно-параметрической идентификации [6]. Данное свойство будем характеризовать набором показателей.

Относительный показатель предорганизации в обучении определим как:

$$I_{po} = \frac{1}{2N_w} \sum_{\eta=1}^M \sum_{i=1}^{N_{\eta}} \sum_{j=1}^{N_{\eta-1}} \mathfrak{G}_{po}(w_j^{(\eta,i)}) + \frac{1}{4N_H} \sum_{\eta=1}^M \sum_{i=1}^{N_{\eta}} \mathfrak{G}_{po}(\varphi^{(\eta,i)}) + \frac{1}{4N_H} \sum_{\eta=1}^M \sum_{i=1}^{N_{\eta}} \mathfrak{G}_{po}(\psi^{(\eta,i)}),$$

где M – количество слоев в сети; N_H – количество нейронов в сети; \mathfrak{G}_{po} – эвристическая функция, показывающая степень неслучайности задания значения параметра сети в предорганизации ее обучения:

$$\mathfrak{G}_{po}(element) = \begin{cases} 0, & element = rand; \\ 1, & element \neq rand, \end{cases}$$

либо

$$\mathfrak{G}_{po}(element) = 1 - \frac{\max(rand) - \min(rand)}{\max(element) - \min(element)},$$

где $element$ – условное обозначение параметра функции; $rand$ – случайная компонента элемента.

Относительный показатель автоматизации предорганизации в обучении зададим как:

$$I_{apo} = \frac{\mathfrak{G}_L + \mathfrak{G}_H}{4} + \frac{1}{4N_w} \sum_{\eta=1}^M \sum_{i=1}^{N_{\eta}} \sum_{j=1}^{N_{\eta-1}} (\mathfrak{G}_a(w_j^{(\eta,i)}) \mathfrak{G}_{po}(w_j^{(\eta,i)})) + \\ + \frac{1}{8N_H} \sum_{\eta=1}^M \sum_{i=1}^{N_{\eta}} (\mathfrak{G}_a(w_j^{(\eta,i)}) \mathfrak{G}_{po}(\varphi^{(\eta,i)})) + \frac{1}{8N_H} \sum_{\eta=1}^M \sum_{i=1}^{N_{\eta}} (\mathfrak{G}_a(w_j^{(\eta,i)}) \mathfrak{G}_{po}(\psi^{(\eta,i)})),$$

где \mathfrak{G}_L – эвристическая мера, показывающая степень автоматизации задания числа слоев в предорганизации обучения сети ($\mathfrak{G}_L = 1$, если число слоев определяется автоматически; в противном случае принимают $\mathfrak{G}_L = 0$); \mathfrak{G}_H – эвристическая мера, показывающая степень автоматизации задания числа нейронов в предорганизации обучения сети ($\mathfrak{G}_H = 1$, если число нейронов определяется автоматически; в противном случае принимают $\mathfrak{G}_H = 0$); $\mathfrak{G}_a(element)$ – эвристическая функция, показывающая степень автоматизации задания значения параметра $element$ сети в предорганизации ее обучения ($\mathfrak{G}_a(element) = 0$, если значение параметра $element$ задает пользователь, в противном случае – $\mathfrak{G}_a(element) = 1$).

Функциональная блочность – свойство построения архитектуры сети из стандартизированных функциональных блоков, внутренняя организация и свойства которых определены заранее и известны разработчику [6].

Поскольку все нейронные и нейро-нечеткие сети по своей природе состоят из нейронов, обладающих однотипностью выполняемых операций, относительную функциональную блочность будем рассматривать на внешнем уровне по отношению к уровню нейронов:

$$I_{fb} = 1 - \frac{2 \sum_{i=1}^{N_H} \sum_{j=i+1}^{N_H} \mathfrak{G}_{fb}(i, j)}{N_H(N_H - 1)},$$

где $\vartheta_{fb}(i, j)$ – показатель принадлежности i -го и j -го нейронов сети к одному и тому же функциональному блоку ($\vartheta_{fb}(i, j)=1$, если нейроны принадлежат к одному и тому же блоку, $\vartheta_{fb}(i, j)=0$ – в противном случае).

Если уровень функциональной блочности равен нулю, это означает, что разработчику сети понятны только преобразования на уровне конкретных нейронов, однако нет таких групп нейронов, функции которых были бы ему понятны. Если же уровень функциональной блочности равен единице, то все нейроны сети объединены в систему блоков, функционирование которых понятно на более высоком уровне, чем уровень конкретного нейрона.

Свойства параметров нейромодели

К свойствам настраиваемых параметров нейромодели можно отнести обобщение, обучаемость, интерпретабельность, надежность, распределенность памяти, вариативность и эффективность нейромоделей.

Обобщение – способность сети интегрировать частные данные для определения закономерностей и пролонгации результатов, что позволяет после обучения на одних данных применять полученные знания для других данных [7, 12].

Коэффициент обобщения обученной сети определим как:

$$I_{gen} = \frac{NS}{N_w N_n} I_h \exp(-(E_{об.} - E_{тест.})^2),$$

где $E_{об.}$ и $E_{тест.}$ – соответственно ошибка сети для обучающей и тестовой выборок.

Если коэффициент обобщения будет существенно больше единицы, то сеть проявляет большие обобщающие способности, если же коэффициент обобщения существенно меньше единицы, то сеть не проявляет обобщающих свойств.

Обучаемость – свойство нейросетей совершенствовать свою работу (обучаться или адаптироваться), используя примеры для настройки на решение определенной задачи [1, 7];

Определим показатель обучаемости сети как:

$$I_{lr} = \frac{I_{pl}(1-L)L_{net}}{NS},$$

где L_{net} – константа Липшица (сложность) нейронной сети net [2]; L – модифицированная константа Липшица для обучающей выборки:

для задач с вещественной выходной переменной

$$L(x, y) = \sqrt{\max_{\substack{s=1,2,\dots,S; \\ g=s+1,\dots,S}} \left\{ \frac{v_y(y^s - y^g)^2}{1 + \sum_{i=1}^N v_i(x_i^s - x_i^g)^2} \right\}}, v_i = \left(\frac{1}{\max_{s=1,2,\dots,S} \{x_i^s\} - \min_{s=1,2,\dots,S} \{x_i^s\}} \right)^2,$$

$$v_y = \left(\frac{1}{\max_{s=1,2,\dots,S} \{y^s\} - \min_{s=1,2,\dots,S} \{y^s\}} \right)^2;$$

для задач с дискретной выходной переменной

$$L(x, y) = \left(1 + \min_{\substack{s=1,2,\dots,S; \\ g=s+1,\dots,S}} \left\{ \sum_{i=1}^N v_i (x_i^s - x_i^g)^2 \right\} \right)^{-0,5}.$$

Чем больше будет значение показателя обучаемости, тем большими возможностями обладает нейронная сеть net для решения задачи аппроксимации таблично заданной зависимости $y = f(x)$.

Интерпретабельность (логическая прозрачность) – свойство сети быть понятной для восприятия и анализа человеком [2, 8]. Очевидно, что сеть является более интерпретабельной, если она иерархичная, функционально блочная, а ее веса контрастны (близки по значению к нулю или единице), причем среднее число входов нейрона не превышает 5-7 (это число обусловлено особенностями человеческой психики).

Наряду с ранее предложенными в [4, 5] характеристиками логической прозрачности, можно эвристически определить интерпретабельность через иерархичность и функциональную блочность:

$$I_{\text{interp.}} = \frac{I_{fb} I_h N_h}{\sum_{\eta=1}^M \sum_{i=1}^{N_\eta} \sum_{j=1}^{N_{\eta-1}} \mathfrak{G}_i(w_j^{(\eta,i)})}, \quad \mathfrak{G}_i(w_j^{(\eta,i)}) = \begin{cases} 1 - |\sin(\pi w_j^{(\eta,i)})|, & 0 \leq w_j^{(\eta,i)} \leq 1; \\ \frac{1}{w_j^{(\eta,i)}}, & w_j^{(\eta,i)} > 1; \\ \frac{1}{|w_j^{(\eta,i)} - 1|}, & w_j^{(\eta,i)} < 0. \end{cases}$$

Надежность и устойчивость сети к отказам отдельных элементов, ее составляющих, проявляется в том, что отказ одного или нескольких нейроэлементов сети не приводит к отказу всей НС и не может существенно влиять на работу сети в целом [9 – 11].

Определим показатели надежности весов сети относительно обучающей выборки как:

$$I_{wre}^{avg} = 1 - \frac{\sum_{\eta=1}^M \sum_{i=1}^{N_\eta} \sum_{j=1}^{N_{\eta-1}} |E(w) - E(w|_{w_j^{(\eta,i)}=0})|}{N_w \sum_{i=1}^{N_M} |\psi_{\max}^{(M,i)} - \psi_{\min}^{(M,i)}|}, \quad I_{wre}^{\max} = 1 - \frac{\max_{\substack{\eta=1,\dots,M; \\ i=1,\dots,N_\eta; \\ j=1,\dots,N_{\eta-1}}} |E(w) - E(w|_{w_j^{(\eta,i)}=0})|}{\max_{i=1,\dots,N_M} |\psi_{\max}^{(M,i)} - \psi_{\min}^{(M,i)}|},$$

где $E(w) = \sum_{s=1}^S \sum_{i=1}^{N_M} (y_i^s - \psi_s^{(M,i)})^2$ – ошибка сети при наборе весовых коэффициентов

w ; $\psi_s^{(M,i)}$ – значение на i -м выходе сети для s -го экземпляра.

Определим показатели надежности нейронов сети относительно обучающей выборки как:

$$I_{\text{н.ге}}^{\text{avg}} = 1 - \frac{\sum_{\eta=1}^M \sum_{i=1}^{N_{\eta}} \sum_{j=1}^{N_{\eta-1}} \left| E(w) - E(w |_{\forall p=1,2,\dots,N_{\eta+1}: w_i^{(\eta+1,p)}=0}) \right|}{N_w \sum_{i=1}^{N_M} \left| \psi_{\text{max}}^{(M,i)} - \psi_{\text{min}}^{(M,i)} \right|},$$

$$I_{\text{н.ге}}^{\text{max}} = 1 - \frac{\max_{\substack{\eta=1,\dots,M; \\ i=1,\dots,N_{\eta}; \\ j=1,\dots,N_{\eta-1}}} \left| E(w) - E(w |_{\forall p=1,2,\dots,N_{\eta+1}: w_i^{(\eta+1,p)}=0}) \right|}{\max_{i=1,\dots,N_M} \left| \psi_{\text{max}}^{(M,i)} - \psi_{\text{min}}^{(M,i)} \right|}.$$

Распределенность памяти – информация в сети хранится по многим адресам, распределенным образом, так что каждый элемент данных представляется шаблоном активности, распределенным по многим вычислительным элементам, и каждый вычислительный элемент участвует в представлении многих различных элементов данных [1, 2];

Распределенность памяти сети представим как усредненную долю весов, оказывающих существенное влияние на выход сети, в общем числе весов для обучающей выборки:

$$I_{\text{md}} = \frac{\sum_{s=1}^S \sum_{\eta=1}^M \sum_{i=1}^{N_{\eta}} \sum_{j=1}^{N_{\eta-1}} (1 - e^{-(E(w, x^s) - E(w, x^s |_{w_j^{(\eta,i)}=0}))^2})}{SN_w}.$$

Вариативность моделей аппроксимируемой зависимости по одной и той же таблице данных сетью заключается в том, что при неизменной структуре, но разных начальных значениях весов либо разных методах их настройки после обучения в общем случае может получиться несколько различных нейромоделей [2, 13].

Определим абсолютный показатель вариативности моделей для сети:

$$I_v = \prod_{\eta=1}^M \prod_{i=1}^{N_{\eta}} \left(\mathcal{G}_v(\varphi^{(\eta,i)}) \mathcal{G}_v(\psi^{(\eta,i)}) \prod_{j=1}^{N_{\eta-1}} \mathcal{G}_v(w_j^{(\eta,i)}) \right),$$

где $\mathcal{G}_v(\varphi^{(\eta,i)})$, $\mathcal{G}_v(\psi^{(\eta,i)})$, $\mathcal{G}_v(w_j^{(\eta,i)})$ – соответственно показатели вариативности типов дискриминантных и активационных функций нейронов, а также значений весов в процессе обучения сети (если соответствующий элемент сети в процессе обучения может принимать только одно фиксированное состояние, то полагают $\mathcal{G}_v(\text{element}) = 1$). Вариативность дискриминантных и активационных функций определяется архитектурой сети и методом обучения, для весов вариативность предлагается оценивать по формуле $\mathcal{G}_v(w_j^{(\eta,i)}) = 1$, если вес является неизменным в

процессе обучения; $\mathcal{G}_v(w_j^{(\eta,i)}) = \text{round} \left(\frac{w_{j \text{max}}^{(\eta,i)} - w_{j \text{min}}^{(\eta,i)}}{\Delta w_j^{(\eta,i)}} \right)$, если вес может изменяться.

Здесь $w_{j \text{max}}^{(\eta,i)}$, $w_{j \text{min}}^{(\eta,i)}$ – соответственно максимально и минимально возможные зна-

чения веса $w_j^{(\eta,i)}$; $\Delta w_j^{(\eta,i)}$ – минимально возможное изменение веса с учетом размера разрядной сетки ЭВМ.

Эффективность (качество) решения задач НС определяется точностью (ошибкой) решения задачи для обучающих и тестовых данных, простотой, интерпретабельностью и быстродействием полученной нейросетевой модели, а также затратами на построение нейросетевой модели (требования к аппаратным средствам, итерационность и затраты времени метода обучения).

Обобщенный показатель эффективности нейромодели на основе предложенных показателей определим как

$$I_{ef} = I_{pl} I_{par} I_h I_{fb} I_{gen} I_{lr} I_{interp.} (I_{wre}^{avg} + I_{н.}^{avg}) \left(\frac{1 + 0,5(I_{po} + I_{apo})}{6} \right) e^{-E}.$$

Обобщенный показатель эффективности может быть использован как для сравнения нейромоделей и методов их синтеза, так и для оптимизации процесса построения нейромоделей.

Эксперименты и результаты

Предложенный комплекс показателей был программно реализован как дополнение к разработанной автором компьютерной программе "Автоматизированная система синтеза нейросетевых и нейро-нечетких моделей для неразрушающей диагностики и классификации образов по признакам" (свидетельство о регистрации авторского права на произведение № 35431, выданное 21.10.2010 г. Государственным департаментом интеллектуальной собственности Украины) на языке макросов, использующем базовые средства ядра сред математических пакетов Scilab и Matlab.

Для проверки практической применимости предложенного комплекса показателей проведено его экспериментальное исследование при решении набора задач технического и биомедицинского диагностирования и автоматической классификации образов по признакам [14], характеристики которых представлены в табл. 1.

Здесь типы задач обозначены: "К" – классификация, "О" – оценивание.

Результаты проведенных экспериментов показали, что предложенные показатели позволяют количественно выразить основные свойства нейромоделей.

В частности, приведенный в табл. 2 фрагмент полученных результатов экспериментов на примере задачи распознавания ирисов Фишера с помощью многослойных нейронных сетей прямого распространения (модели 1 – 3) и радиально-базисных сетей (модель 4) с различным числом нейронов демонстрирует работоспособность предложенных показателей для сравнения получаемых нейромоделей.

Результаты проведенных экспериментов позволяют рекомендовать предложенный комплекс показателей для использования на практике.

На рис. 1 – 3 представлены экспериментально полученные зависимости между предложенными показателями на наборе решавшихся задач различной размерности и сложности.

Таблица 1

Задачи	Число признаков		Число экземпляров	Тип задачи
	входных	выходных		
Диагностирование лопаток авиадвигателей	10240	1	32	К
Моделирование зависимости частоты собственных колебаний от геометрических параметров пера лопаток	13	1	65	О
Прогнозирование повышения поверхностной прочности деталей газотурбинных авиадвигателей после упрочняющей обработки	16	1	57	О
Моделирование взаимосвязей параметров испытаний авиадвигателей в процессе проливки сопловых аппаратов	6	9	491	О
Первичное дифференциальное диагностирование острых воспалений мочевыделительной системы	6	2	120	К
Диагностирование рака молочной железы	10	1	569	К
Моделирование суммарного показателя качества жизни (СПКЖ) больных хроническим обструктивным бронхитом	95	1	117	О
Диагностирование хронического обструктивного бронхита	28	1	205	К
Диагностирование сердечной аритмии	279	1	452	К
Моделирование зависимости состояния здоровья детского населения от загрязнения окружающей среды	43	1	954	О
Автоматическая классификация ирисов Фишера	4	1	150	К
Автоматическая классификация автотранспортных средств	26	1	1062	К
Автоматическая классификация сельскохоз. растений	55	1	248	К

Таблица 2

Показатель	Нейромодель 1	Нейромодель 2	Нейромодель 3	Нейромодель 4
E	0,0015774	0,0015774	0,0016285	$3,5754 \cdot 10^{-24}$
N_n	3	3	5	151
I_{par}	0,86842	0,86842	0,71264	0,93641
I_{ha}	1	1	1	0,98675
I_{hd}	1	1	1	0,98675
I_{hn}	1	1	1	0,98675
I_{np}	0,1	0,1	0,1	0,1
I_{sp}	0,35135	0,35135	0,31148	0,03847
I_{pl}	0,035135	0,035135	0,031148	0,003847
I_h	0,44444	0,44444	0,3	0,33554
I_v	26611	26611	38893	1844347
I_{fb}	0	0	0	0,013245
$I_{interp.}$	0,048509	0,056079	0,036493	0,055908
I_{po}	0,5	0,5	0,5	0,5
I_{apo}	0,5	0,5	0,5	0,5
I_{md}	0,029729	0,063077	0,038288	0,62625
I_{wre}^{avg}	0,079283	0,12968	0,091169	1
I_{wre}^{max}	0,34076	0,34076	0,34076	1
$I_{h.re}^{avg}$	0,036933	0,073035	0,083012	1
$I_{h.re}^{max}$	0,15492	0,34076	0,34076	1
L_{net}	$1,7043 \cdot 10^8$	$4,8703 \cdot 10^8$	38618	$5,1333 \cdot 10^5$
I_{lr}	4727,5	13509	0,94961	1,559
I_{gen}	14,975	14,975	5,3164	$2,9115 \cdot 10^{-42}$
I_{ef}	0	0	0	$7,81472 \cdot 10^{-51}$

Как следует из рис. 1, с увеличением значения I_{apo} уменьшается значение I_{hn} и увеличивается значение $I_{interp.}$, а с увеличением значения $\ln(I_{gen})$ увеличиваются значения $I_{н.ге}^{avg}$ и $\ln(I_{ef})$; с увеличением значения I_{hn} уменьшаются значения I_{fb} и $I_{interp.}$. Как видно из рис. 2, с ростом значения $\ln(I_{md})$ возрастают значения $\ln(I_{ef})$ и $\ln(I_{н.ге}^{avg})$; с ростом значения $\ln(I_{par})$ возрастают значения $\ln(I_{ef})$ и $\ln(I_{н.ге}^{avg})$ и уменьшается значение $\ln(I_v)$; с ростом значения $\ln(I_{md})$ возрастает значение $\ln(I_{wre}^{avg})$. Из рис. 3 видно, что с увеличением значения $\ln(I_{pl})$ возрастает значение $\ln(I_{wre}^{avg})$, а с ростом значения I_h наблюдается рост значения $I_{interp.}$. Использование полученных зависимостей позволяет упростить расчет пользовательских критериев качества нейромоделей посредством исключения качественно эквивалентных составляющих.

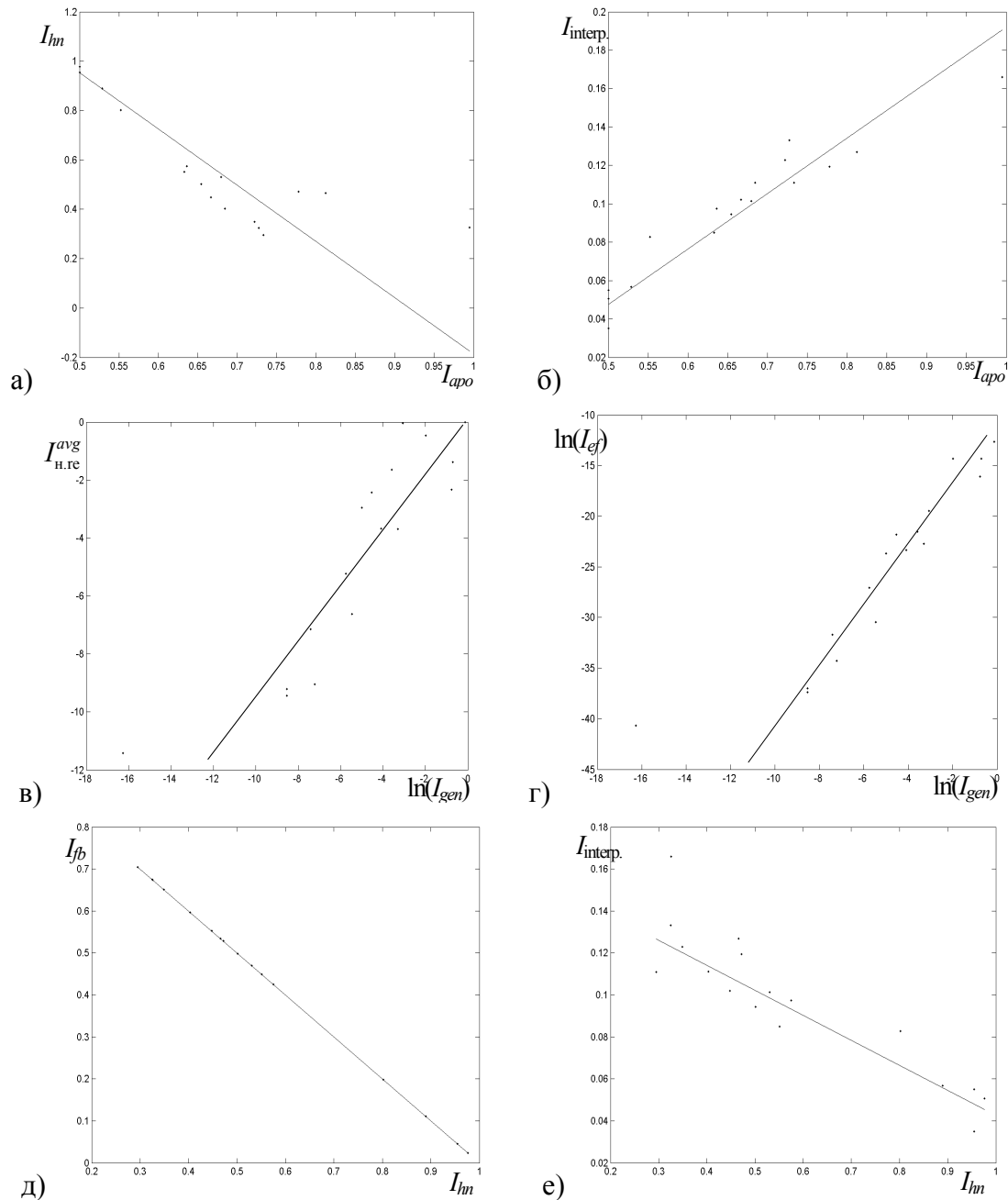


Рис. 1. Графики зависимостей: а) I_{hn} от I_{apo} ; б) $I_{interp.}$ от I_{apo} ; в) $I_{н.ге}^{avg}$ от $\ln(I_{gen})$; г) $\ln(I_{ef})$ от $\ln(I_{gen})$; д) I_{fb} от I_{hn} ; е) $I_{interp.}$ от I_{hn} .

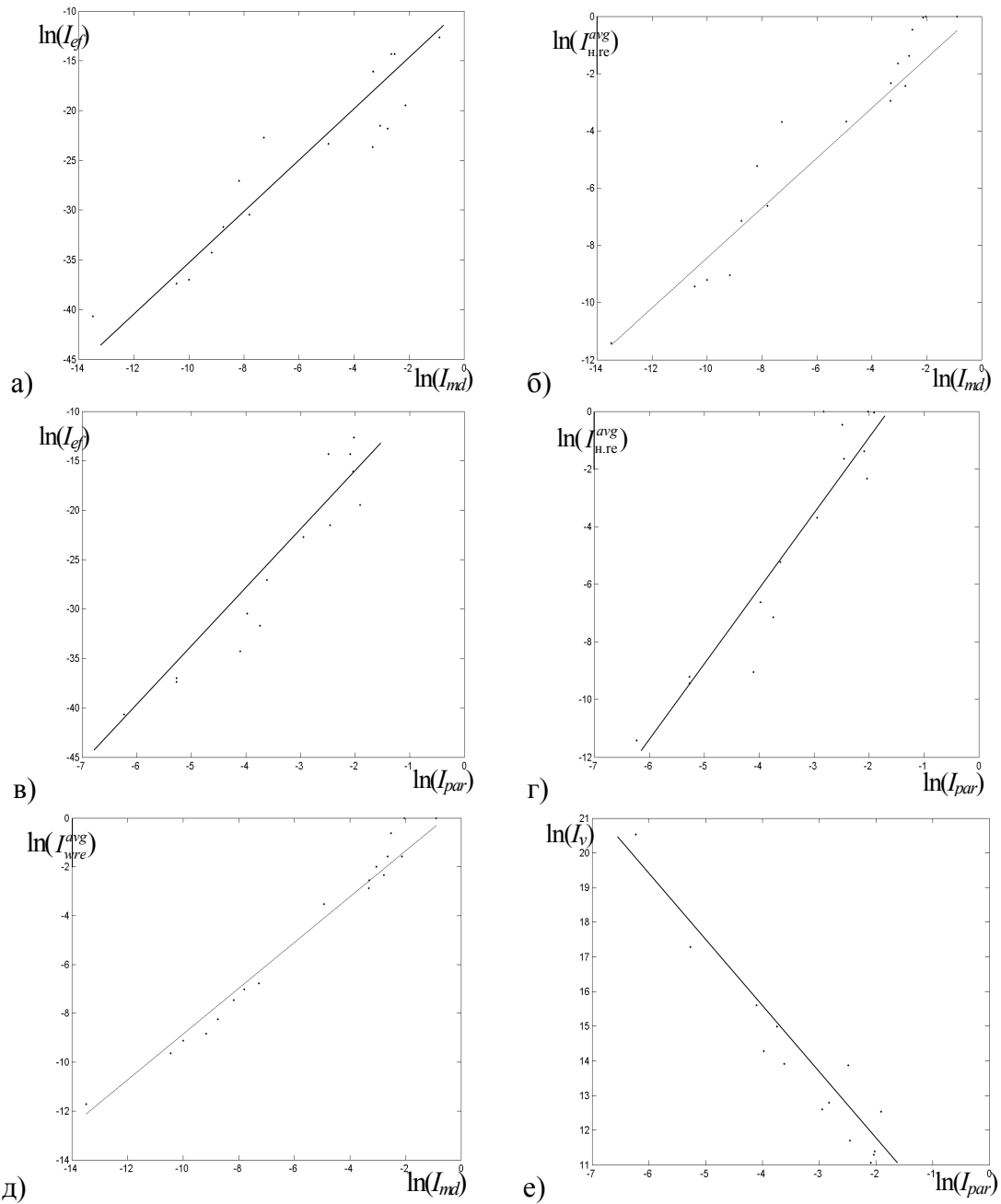


Рис. 2. Графики зависимостей: а) $\ln(I_{ef})$ от $\ln(I_{md})$; б) $\ln(I_{h.re}^{avg})$ от $\ln(I_{md})$; в) $\ln(I_{ef})$ от $\ln(I_{par})$; г) $\ln(I_{h.re}^{avg})$ от $\ln(I_{par})$; д) $\ln(I_{wre}^{avg})$ от $\ln(I_{md})$; е) $\ln(I_v)$ от $\ln(I_{par})$.

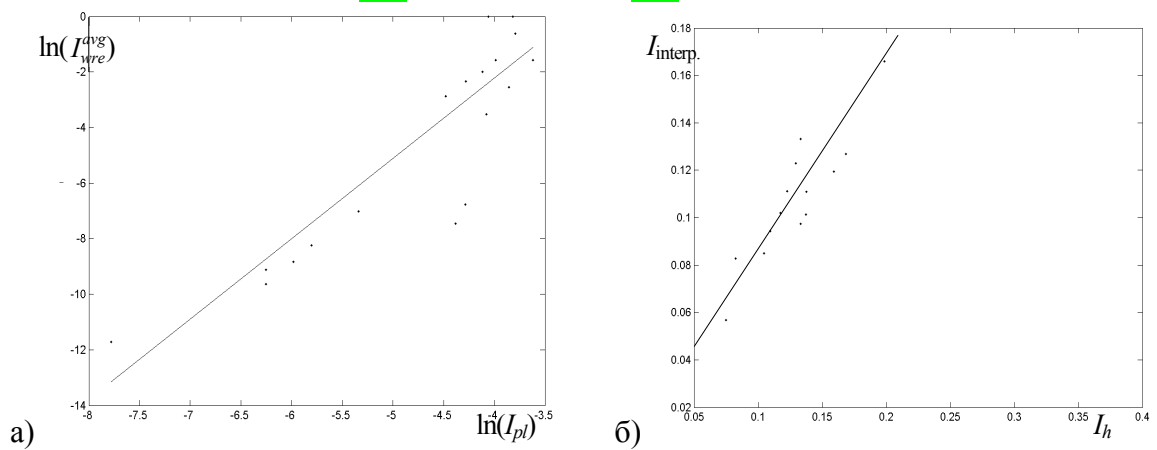


Рис. 3. Графики зависимостей: а) $\ln(I_{wre}^{avg})$ от $\ln(I_{pl})$; б) I_{interp} от I_h .

Заклучение

Предложен набор показателей, характеризующих свойства нейронных сетей, позволяющий автоматизировать решение задачи анализа свойств и сравнения нейросетевых моделей при решении задач диагностики и распознавания образов.

Результаты проведенных экспериментов подтвердили работоспособность предложенных показателей, что позволяет использовать их на практике при решении задач диагностирования и распознавания образов.

Работа выполнена как часть госбюджетной темы Запорожского национального технического университета "Интеллектуальные информационные технологии автоматизации проектирования, моделирования, управления и диагностирования производственных процессов и систем".

ЛИТЕРАТУРА

1. Терехов В.А., Ефимов Д.В., Тюкин И.Ю., Антонов В. Н. Нейросетевые системы управления. – СПб.: Изд-во СПбГУ, 1999.
2. Миркес Е. М. Нейроинформатика: Учебное пособие для студентов. – Красноярск: ИПЦ КГТУ, 2002.
3. Субботин С.О., Олійник А.О., Олійник О.О. Неітеративні, еволюційні та мультиагентні методи синтезу нечіткологічних і нейромережних моделей. – Запоріжжя: ЗНТУ, 2009.
4. Субботин С.А. Методика и критерии сравнения моделей и алгоритмов синтеза искусственных нейронных сетей // Радиоелектроніка. Інформатика. Управління. – 2003. – № 2. – С. 109-114.
5. Субботин С.А. О сравнении нейросетевых моделей // Нейроинформатика и ее приложения: Материалы XI Всероссийского семинара. – Красноярск: ИВМ СО РАН, 2003. – С. 152-153.
6. Нейрокомпьютер как основа мыслящих ЭВМ. – М.: Наука, 1993.
7. Головкин В.А. Нейронные сети: обучение, организация и применение: Учебное пособие для вузов / общ. ред. А. И. Галушкина. – М.: ИПРЖР, 2001.
8. Wang Ed. J. Encyclopedia of Data Warehousing and Mining. – Hershey: Idea Group, 2006.
9. Van Gorp J. Nonlinear identification with neural networks and fuzzy logic : proefschrift ... doctor. – Brussel: Vrije Universiteit Brussel, 2000.
10. Гаскаров Д. В., Голинкевич Т.А., Мозгалевский А.В. Прогнозирование технического состояния и надежности радиоэлектронной аппаратуры. – М.: Советское радио, 1974.
11. Hoekstra A., Tholen S.A., Duijn R.P.W. Estimating the reliability of neural network classifications // Artificial Neural Networks: International Conference ICANN 96, 16–19 July, Bochum : Proceedings / [Eds.: C. von der Malsburg, W. von Seelen, J. C. Vorbrüggen, B. Sendhoff]. – Berlin: Springer, 1996. – P. 53-58. – (Lecture Notes in Computer Science. vol. 1112).
12. Wang W., Gelder P. Van, Vrijling J.K. Some issues about the generalization of neural networks for time series prediction // Artificial Neural Networks: Formal Models and Their Applications : 15th International Conference ICANN 2005, 11-15 September 2005, Warsaw : Proceedings / [Eds.: W. Duch et al.]. – Berlin: Springer, 2005. – Part III. – P. 559–564.
13. Basalyga G., Salinas E. When response variability increases neural network robustness to synaptic noise // Neural Computation. – 2006. –Vol. 18, № 6. – P. 1349-1379.
14. Субботин С.А., Олейник Ан. А., Гофман Е.А. и др. Интеллектуальные информационные технологии проектирования автоматизированных систем диагностирования и распознавания образов. – Харьков: Компания СМІТ, 2012.

Статья представлена к публикации членом редколлегии О.С. Амосовым.

E-mail:

Субботин Сергей Александрович – subbotin@zntu.edu.ua.