

УДК 517.958

ИСПОЛЬЗОВАНИЕ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ РЕШЕНИЯ ЗАДАЧ КЛАССИФИКАЦИИ НА ПРИМЕРЕ МОДЕЛИРОВАНИЯ МЕДИЦИНСКОГО ОБЪЕКТА

© О.В. Крючин, Н.А. Зенкова

Ключевые слова: искусственные нейронные сети; адаптивно-резонансная теория; ИНС-модель; нейросетевой классификатор; моделирование медицинского объекта; общеклиническое исследование крови.

В работе описываются искусственные нейронные сети, структуры которых базируются на адаптивно-резонансной теории. Также предлагается применение данного типа искусственных нейронных сетей для решения задач классификации на примере моделирования медицинского объекта.

ВВЕДЕНИЕ

В настоящее время искусственные нейронные сети (ИНС) нашли широкое применение в различных сферах человеческой деятельности. Наиболее известным типом ИНС является многослойный перцептрон (multi-layer perceptron – MLP), который позволяет решать практически любые виды задач. Однако этот тип ИНС имеет ряд недостатков при решении некоторых типов задач, например, классификации. Так, очень часто обучение новому образцу уничтожает или изменяет результаты предшествующего обучения. В некоторых случаях это несущественно, например, если имеется только фиксированный набор обучающих векторов, то они могут предъявляться при обучении циклически. В сетях с обратным распространением обучающие векторы подаются на вход сети последовательно до тех пор, пока сеть не обучится всему входному набору. Однако если полностью обученная сеть должна запомнить новый обучающий вектор, то он может изменить веса настолько, что потребуются полное переобучение сети [1, 2].

В реальной ситуации сеть будет подвергаться постоянно изменяющимся воздействиям и может никогда не увидеть один и тот же обучающий вектор дважды. При таких обстоятельствах сеть часто не будет обучаться, а будет непрерывно изменять свои веса, не достигая удовлетворительных результатов. Кроме того, существуют ИНС, в которых только четыре обучающих вектора, предъявляемых циклически, заставляют веса сети изменяться непрерывно, никогда не сходясь [1].

Данные обстоятельства послужили одной из причин для разработки Гроссбергом в 1969 г. адаптивной резонансной теории (АРТ) и построения на базе данной теории искусственных нейронных сетей. Разработанные АРТ-структуры ИНС предназначены для решения задач классификации.

Целью данной работы является использование нейросетевого классификатора для моделирования медицинского объекта, заключающегося в определении зависимости наиболее подходящих условий лечения от результатов общеклинического исследования крови, и

разработка медицинской экспертной системы, базирующейся на результатах моделирования.

Ранее нами была разработана ИНС-модель медицинского объекта с использованием многослойного перцептрона, однако адекватность такой модели оказалась недостаточно удовлетворительной для практического применения [3].

Искусственные нейронные сети, базирующиеся на адаптивно-резонансной теории. В настоящее время существует несколько типов АРТ-структур. Такая сеть состоит из двух слоев – входного слоя сравнения (имеющего L нейронов), на который подаются образы – наборы входных данных и выходного слоя распознавания (имеющего P нейронов), определяющего количество классов – групп образов. Каждый нейрон входного слоя соединен с каждым нейроном выходного слоя восходящими синаптическими связями ($\vec{W}$), а каждый нейрон выходного слоя соединен с каждым нейроном входного слоя нисходящими связями ($\vec{\bar{W}}$). Пример АРТ-структуры представлен на рис. 1 (темным цветом отображены нейроны входного слоя сравнения, а светлым – выходного слоя распознавания) [4, 5].

Особенность структуры АРТ-2 заключается в том, что входной слой состоит из 6 подслоев, каждый из которых содержит L нейронов. Наличие этих подслоев обусловлено задачей нормализации входных значений

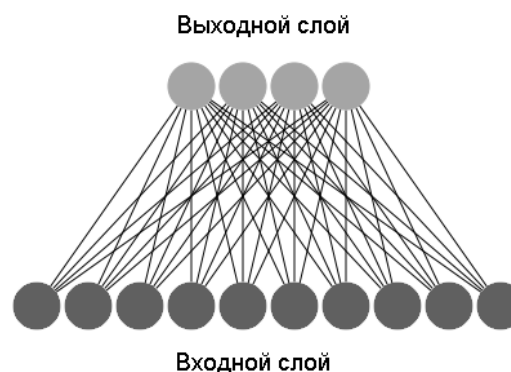


Рис. 1. Структура сети АРТ

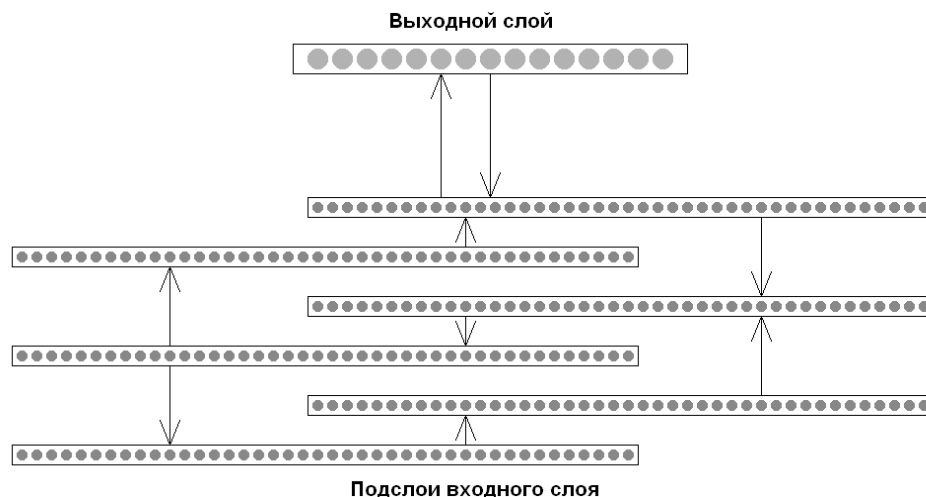


Рис. 2. Подробная структура сети АРТ-2

(благодаря им сеть АРТ-2, в отличие от АРТ-1, может работать с образами, содержащими не только бинарные, но и вещественные числа). Структура этих подслоев изображена на рис. 2 ($L = 57$, $P = 15$).

Входными данными сети являются последовательно подаваемые вещественные вектора произвольной размерности L – $\vec{x} = (x_0, x_1, \dots, x_{L-1})$.

Параметрами сети являются:

- порог близости c_p – действительное число из интервала $(0; 1)$, при этом чем ближе оно к 1, тем требование близости является более строгим;
- c_a , c_b , c_c – произвольные положительные числовые коэффициенты, выполняющие нормализующую функцию;
- δ_ε – малое действительное положительное число, применяемое для предотвращения деления на ноль;
- δ_θ – малое действительное положительное число, служащее критерием отсека шума с использованием следующего соотношения:

$$f(x) = \begin{cases} 0, & 0 \leq |x| \leq \delta_\theta, \\ x, & |x| > \delta_\theta. \end{cases}$$

Алгоритм обучения. В начале работы сеть содержит по L нейронов в каждом подслое входного слоя и один нейрон выходного слоя ($P = 1$). Веса связей инициализируются следующими значениями:

$$w_{0,i} = \frac{1}{1+L}, \quad \hat{w}_{i,0} = 0, \quad i = 0..L-1. \quad (1)$$

Состояния всех нейронов в подслоях входного слоя являются нулевыми.

При поступлении очередного вектора $\vec{x}_k = (x_{k,0}, x_{k,1}, \dots, x_{k,L-1})$ выполняются следующие операции:

- все нейроны выходного слоя становятся активными;

– последовательно вычисляются состояния нейронов в подслоях входного слоя:

$$\hat{y}_{0,i} = x_{k,i} + c_a \hat{y}_{2,i}, \quad i = 0..L-1, \quad (2)$$

$$\hat{y}_{1,i} = \frac{\hat{y}_{0,i}}{\delta_\varepsilon + \sum_{j=0}^{L-1} \hat{y}_{0,j}}, \quad i = 0..L-1, \quad (3)$$

$$\hat{y}_{2,i} = f(\hat{y}_{1,i}) + c_b f(\hat{y}_{5,i}), \quad i = 0..L-1, \quad (4)$$

$$\hat{y}_{3,i} = \frac{\hat{y}_{2,i}}{\delta_\varepsilon + \sum_{j=0}^{L-1} \hat{y}_{2,j}}, \quad i = 0..L-1, \quad (5)$$

$$\hat{y}_{4,i} = \hat{y}_{3,i}, \quad i = 0..L-1, \quad (6)$$

$$\hat{y}_{5,i} = \frac{\hat{y}_{4,i}}{\delta_\varepsilon + \sum_{j=0}^{L-1} \hat{y}_{4,j}}, \quad i = 0..L-1. \quad (7)$$

Здесь $\vec{x}$ – вектор входных сигналов, $\vec{\hat{y}}_k$ – значения k -го подслоя входного слоя.

Вычисляются значения нейронов выходного слоя

$$y_{1,i} = \sum_{j=0}^{L-1} w_{i,j} \hat{y}_{4,j}, \quad i = 0..P-1. \quad (8)$$

Если существуют активные нейроны, то из них выбирается «нейрон-победитель» (нейрон, имеющий максимальное значение), после чего происходит пересчет значений нейронов входного слоя:

$$\hat{y}_{4,i} = \hat{y}_{3,i} + y_{1,l} \hat{w}_{l,i}, \quad i = 0..L-1, \quad (9)$$

где l – индекс «нейрона-победителя».

Затем «нейрон-победитель» проходит проверку на близость:

$$\frac{c_p}{\delta_\varepsilon + r} \leq 1, \quad (10)$$

$$\text{где } r = \sum_{i=0}^{L-1} \left(\frac{\hat{y}_{3,i} + c_c \hat{y}_{4,i}}{\delta_\varepsilon + \sum_{j=0}^{L-1} \hat{y}_{3,i} + c_c \sum_{j=0}^{L-1} \hat{y}_{4,i}} \right).$$

При выполнении условия (10) весовые коэффициенты «нейрона-победителя» корректируются:

$$w_{i,i} = \hat{w}_{i,i} = \frac{\hat{y}_{4,i}}{1 - y_{1,i}}, \quad i = 0..L-1, \quad (11)$$

в противном случае «нейрон-победитель» становится неактивным и из оставшихся активных нейронов выбирается новый «победитель». Если активных нейронов больше нет, то генерируется новый выходной нейрон (появляется новый класс) со следующими связями:

$$w_{p-1,i} = \frac{\hat{y}_{4,i}}{1 + \sum_{j=0}^{L-1} \hat{y}_{4,i}} \quad i = 0..L-1, \quad (12)$$

$$\hat{w}_{i,p-1} = \hat{y}_{4,i}, \quad i = 0..L-1. \quad (13)$$

Результатом работы сети является либо номер «нейрона-победителя» (i), прошедшего проверку на близость, т. е. факт распознавания вектора и его отнесения к уже известному классу (кластеру) объектов, либо сообщение об отсутствии кластера для данного вектора и создании такого кластера [6, 7].

Объект моделирования и формирование обучающей выборки. В качестве данных для моделирования были использованы данные центральной реабилитационной больницы г. Рассказова. Данные представляют собой результаты общеклинического исследования анализа крови пациентов.

Для обучения сети были использованы результаты анализа крови четырехсот пациентов. Возраст пациентов строго ограничен – старше 18 лет, т. к. у людей младшего возраста наблюдается изменение в нормах по показателям, входящим в состав результата анализа крови.

Общеклиническое исследование крови, являясь одним из важнейших диагностических методов, тонко отражает реакцию кроветворных органов на воздействие на организм различных физиологических и патологических факторов. Во многих случаях оно играет большую роль в постановке диагноза, а при заболеваниях системы кроветворения ему отводится ведущая роль. В понятие «общеклиническое исследование крови» входят определение концентрации гемоглобина, подсчет количества эритроцитов, цветового показателя, лейкоцитов, скорости оседания эритроцитов (СОЭ) и лейкоцитарной формулы. В необходимых случаях дополнительно определяют время свертывания крови, длительность кровотечения, количество ретикулоцитов и тромбоцитов. В настоящее время большинство показателей выполняют на автоматических гематологиче-

ских анализаторах, которые в состоянии одновременно определять от 5 до 24 параметров. Из них основными являются концентрация гемоглобина, гематокрит, количество эритроцитов, средний объем эритроцита, средняя концентрация гемоглобина в эритроците, среднее содержание гемоглобина в эритроците, полуширина распределения эритроцитов по размерам, количество тромбоцитов, средний объем тромбоцита, количество лейкоцитов.

Входами обучающей выборки являлись десять показателей общего анализа крови, включающие:

- гемоглобин (г/л);
- скорость оседания эритроцитов (мм/ч);
- лейкоциты ($10^9/\text{л}$);
- эритроциты ($10^{12}/\text{л}$);
- цветовой показатель;
- лейкоцитарная формула:
 - базофилы;
 - нейтрофилы палочкоядерные;
 - нейтрофилы сегментоядерные;
 - лимфоциты;
 - моноциты.

Значения каждого показателя были пронормированы в диапазоне $[-1; 1]$ в связи с тем, что значения этих показателей находились в различных диапазонах [3].

Построение ИНС-модели. Для построения ИНС-модели было использовано 400 строк обучающей выборки. Выборка содержала 4 класса, соответствующие определенному состоянию пациента и соответствующим условиям его лечения – госпитализация, стационарное лечение, амбулаторное лечение и отсутствие необходимости лечения. После анализа выборки были исключены 2 противоречивых образа (идентичные образы содержались в разных классах). После проведения обучения было получено 7 классов, которые полностью укладывались в исходные (образы, находящиеся в общем полученном и разных исходных классах, отсутствовали), что позволяет сделать вывод о достаточной адекватности построенной модели.

Построенная ИНС-модель была проверена на адекватность. Исходная выборка, состоящая из четырехсот результатов общего анализа крови пациентов, была разбита на две части. Первая часть – «выборка для обучения» – представляет собой выборку, предназначенную для обучения ИНС. Она состоит из трехсот результатов анализа крови, выбранных таким образом, чтобы каждого состояния в ней было ровно 3/4 от состояний исходной выборки. Вторая часть представляет собой т. н. «тестовое множество». Оно предназначено для оценки адекватности ИНС-модели. Она состоит из ста результатов анализа крови, выбранных таким образом, чтобы каждого состояния в ней было ровно 1/4 от

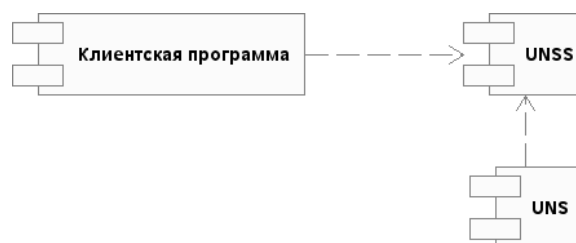


Рис. 3. Архитектура экспертной системы



Рис. 4. Интерфейс клиентской программы ЭС

состояний исходной выборки. Ошибка на тестовом множестве не превысила 10 %, что подтверждает адекватность модели.

Программная реализация. Описанная выше модель была реализована в виде экспертной системы (ЭС), разработанной на языке C++ (интерфейс пользователя разработан с помощью библиотеки Qt4).

На рис. 3 изображена структура данной ЭС, состоящей из трех компонентов (все эти компоненты разработаны на кафедре компьютерного и математического моделирования Тамбовского государственного университета им. Г.Р. Державина):

- нейросетевой симулятор *UNS (Universal Neural Network's Simulator)*, разработанный на кафедре компьютерного и математического моделирования [8];
- сервер симулятора *QUNSS (Universal Neural Network's Simulator's Server)*, также разработанный на кафедре компьютерного и математического моделирования;
- компонент взаимодействия с пользователем (интерфейс программы представлен на рис. 4).

Эти компоненты могут располагаться как сосредоточенно (все компоненты на одном компьютере), так и распределенно (на нескольких компьютерах). Первые два компонента предназначены для работы под управлением операционной системы *GNU/Linux*, а компонент взаимодействия с пользователем может работать как под управлением *GNU/Linux*, так и *MS Windows*.

На рис. 4 представлен интерфейс клиентской программы. В правой части окна программы находятся поля для ввода адреса сервера симуляторов – *UNSS* – ip-адрес (хост) и порт, а также поле для ввода имени файла, хранящего конфигурацию ИНС. В левой части находятся поля для ввода значений показателей, являющихся результатами общеклинического исследования крови. Для получения наиболее подходящего условия лечения необходимо нажать кнопку «Вычислить», после чего оно будет выведено в поле «Заклучение».

ВЫВОДЫ

Таким образом, была построена ИНС-модель медицинского объекта с использованием нейросетевого

классификатора. На основе данной модели возможно построение программы, реализующей медицинскую экспертную систему. Проведенные в работе исследования показали, что достигнутая адекватность является приемлемой для практического использования.

ЛИТЕРАТУРА

1. Круглов В.В., Борисов В.В. Искусственные нейронные сети. М.: Горячая линия – Телеком, 2001.
2. Уоссермен Ф. Нейрокомпьютерная техника. М.: Мир, 1992.
3. Зенкова Н.А., Сергеева М.С. Экспертная система для медицинского объекта на основе искусственных нейронных сетей // Вестник Тамбовского университета. Серия Естественные и технические науки. Тамбов, 2009. Т. 14. Вып. 1. С. 246-247.
4. Hartley K. Seeing the need for ART // Sci. News. 1987. V. 132. № 14.
5. Carpenter G.A., Grossberg S. Category learning and adaptive pattern recognition: a neural network model // Proceedings, Third Army Conference Applied Mathematics on and Computing. 1985. ARO Report 86-1. P. 37-56.
6. Carpenter G.A., Grossberg S. Invariant pattern recognition and recall by an attentive self-organizing ART architecture in a nonstationary world // Proceedings first international conference on neural networks. San Diego (IEEE, New York), 1987.
7. Carpenter G.A., Grossberg S. ART 2: Self-organization of stable category recognition codes for analog input patterns // Applied Optics. 1987. V. 26 (23). P. 4919-4930.
8. Крючин О.В., Арзамасцев А.А., Королев А.Н., Горбачев С.И., Семенов Н.О. Универсальный симулятор, базирующийся на технологии искусственных нейронных сетей, способный работать на параллельных машинах // Вестник Тамбовского университета. Серия Естественные и технические науки. Тамбов, 2008. Т. 13. Вып. 5. С. 372-375.

Поступила в редакцию 18 апреля 2011 г.

Kryuchin O.V., Zenkova N.A. USE OF ARTIFICIAL NEURAL NETWORKS FOR CLASSIFICATION PROBLEMS SOLUTION ON EXAMPLE OF MEDICAL OBJECT MODELING

This article describes artificial neural networks which structure based on the adaptive-resonance theory. The usage of this type of artificial neural networks for the solving problems of the classification is presented. The example of this problem which is the simulation of the medical object is discussed.

Key words: artificial neural networks; adaptive-resonance theory; ANN-model; neural network classifier; simulation of medical object; blood analysis.