

УДК 519.95

РАЗРАБОТКА ЭКСПЕРТНОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ С РАЗВИВАЮЩИМСЯ ИНТЕЛЛЕКТУАЛЬНЫМ ЯДРОМ НА БАЗЕ ИНС-МОДЕЛЕЙ

© А.А. Арзамасцев, Н.А. Зенкова, А.В. Неудахин

Ключевые слова: автоматизированная экспертная система, развивающееся интеллектуальное ядро, нейросетевые модели, математическая формализация задач накопления информации.

Статья посвящена разработке автоматизированной технологии построения экспертных систем, сочетающих в себе функциональные преимущества развивающегося интеллектуального ядра и сетевых технологий взаимодействия. Приведены результаты математической формализации задач накопления информации и построения на ее основе развивающегося интеллектуального ядра.

Необходимость в экспертных системах (ЭС) как в интеллектуальных помощниках, способных решать трудноформализуемые задачи, существует во многих сферах человеческой деятельности: промышленности, научных исследованиях, социальной и экономической сферах.

Ранее нами была разработана автоматизированная технология построения ЭС с интеллектуальным ядром на основе нейросетевых моделей [1–2]. Концептуальная модель таких ЭС приводится на рис. 1. Основными отличительными компонентами данной модели являются веб-интерфейс, база данных, используемая для долговременного хранения технической информации, а также интеллектуальное ядро, включающее развивающуюся систему представления знаний экспертов, основанную на нейросетевых моделях.

Для разработки автоматизированной технологии построения ЭС, соответствующих разработанной концептуальной модели, была проведена математическая формализация задач накопления информации и построения на ее основе развивающегося интеллектуального ядра [3].

Формализация проблемы. Пусть имеется распределенная система получения исходной информации, включающая N терминалов, каждый из которых предназначен для ввода n независимых параметров, имеющих m уровней каждый. Работа терминалов организована таким образом, что каждый из них работает лишь часть времени так, что интервалы их простоя являются случайными величинами, заданными плотностями распределений $p_1(\tau), p_2(\tau), \dots, p_N(\tau)$. Время ввода информации не зависит от номера терминала, но линейно зависит от n , т. е. $\tau_{\text{input}} = k_1 n$, где k_1 – некоторый коэффициент пропорциональности. Будем считать, что очередная порция информации добавляется к базе данных в виде одной записи, когда на i -м терминале ($i = 1, \dots, N$) каждому независимому параметру x_j ($j = 1, \dots, n$) присвоен соответствующий уровень l_k ($k = 1, \dots, m$).

Будем считать, что передача информации с любого терминала в базу данных осуществляется мгновенно, т. к. это время на несколько порядков меньше времени, затрачиваемого пользователем на ввод информации.

Тогда среднее время одного цикла работы i -го терминала (ввод + ожидание следующего ввода информации), соответствующее добавлению в базу данных одной записи, может быть вычислено как:

$$\bar{\tau}_i = k_1 n + \int_0^{t_i^{\max}} \tau p_i(\tau) d\tau.$$

Поэтому за период времени T (например $T = 24$ ч) на i -м терминале могут быть осуществлены r_i циклов ввода, соответствующие добавлению в базу данных r_i записей:

$$r_i = \frac{T}{t_i^{\max}} \cdot \left(k_1 n + \int_0^{t_i^{\max}} \tau p_i(\tau) d\tau \right).$$

со всех терминалов за период T , составит:

$$R = \sum_{i=1}^N r_i = T \sum_{i=1}^N \frac{1}{k_1 n + \int_0^{t_i^{\max}} \tau p_i(\tau) d\tau}.$$

Выполним оценку периода первичного накопления данных в БД – t_1 , предшествующего первому обучению интеллектуального ядра (ИНС-модели).

В соответствии с теоремами А.Н. Колмогорова [4–5] о представимости функций нескольких переменных с помощью суперпозиций и сумм функций одного переменного, можно утверждать, что каждая непрерывная функция n переменных, заданная на единичном кубе

n -мерного пространства, представима в виде: $f(x_1, x_2, \dots, x_n) = \sum_{q=1}^{2n+1} h_q \left[\sum_{p=1}^n \varphi_q^p(x_p) \right]$, где

функции $h_q(u)$ непрерывны, а функции $\varphi_q^p(x_p)$, кроме того, еще и стандартны, т. е. не зависят от выбора функции f .

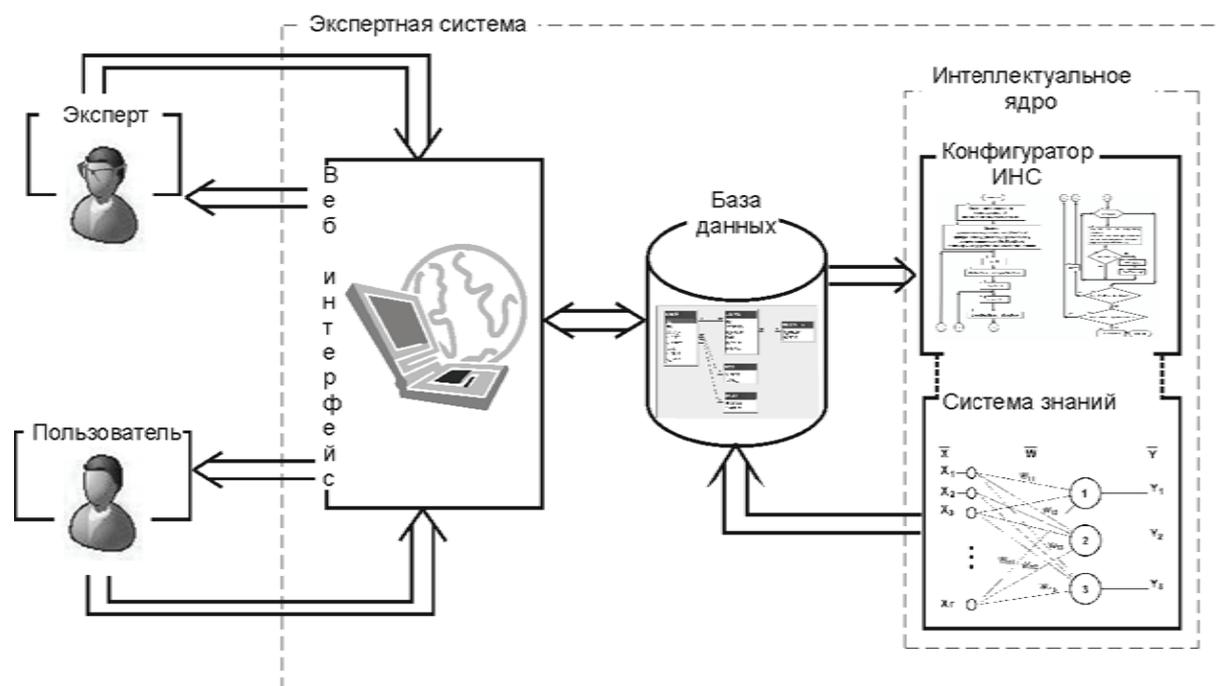


Рис. 1. Концептуальная модель веб-ориентированных нейросетевых ЭС

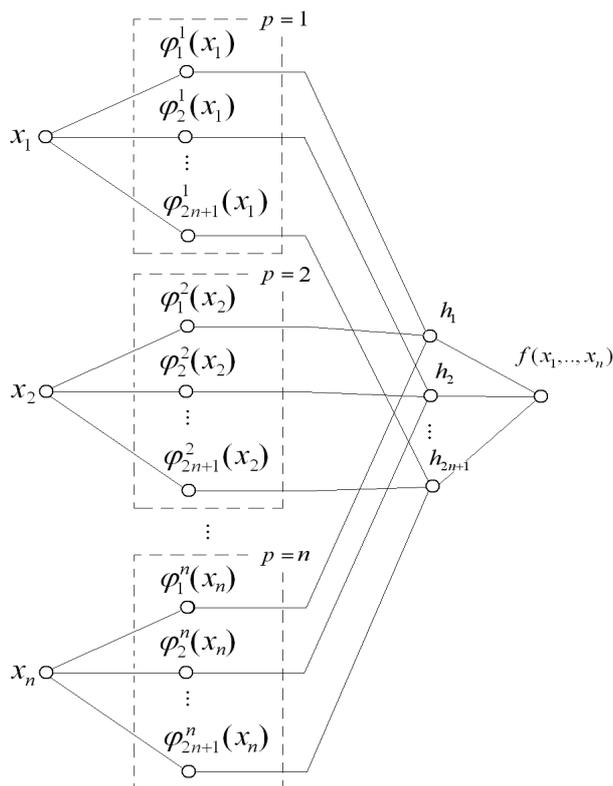


Рис. 2. Представление функции n переменных в виде ИНС-модели

Фактически это означает, что «минимальная» структура ИНС-модели, позволяющая аппроксимировать функцию n переменных, должна иметь следующий вид: входной слой, состоящий из n нейронов, первый скрытый слой, включающий $(2n+1) \cdot n$ функциональных нейронов, второй скрытый слой, состоящий из $(2n+1)$ функциональных нейронов, и выходной, суммирующий нейроны (рис. 2). Указанная нейронная сеть не является полностью связанной и имеет $(2n+1)n + (2n+1)n + 2n+1 = (2n+1)^2$ связей (степеней свободы при обучении ИНС-модели).

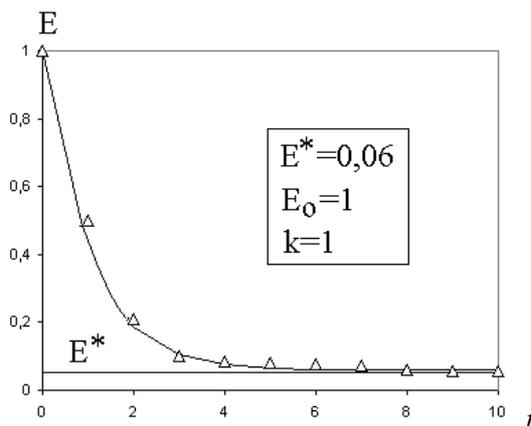


Рис. 3. Корреляция эмпирических данных по среднеквадратичной погрешности обучения сети – E (треугольники) от времени обучения с результатами расчета по уравнению (6): на графике показаны приведенные безразмерные значения

Необходимо заметить, что теоремы А.Н. Колмогорова не несут никакой информации о типе нелинейности функций φ и h . Так, в интеллектуальном ядре экспертной системы, базируемом на ИНС-модели, мы используем активационную функцию $f(s) = \frac{1}{1 + e^{-\alpha s}}$ и полностью

занную сеть – многослойный персептрон. Для этого случая число степеней свободы ИНС-модели будет $(2n + 1) \cdot (3n^2 + n + 1)$. Учитывая, что число записей в БД должно быть не меньше, чем число степеней свободы, получим оценку для первичного периода накопления данных:

$$t_1 \geq K \cdot \left(\sum_{i=1}^N \frac{1}{k_1 n + \int_0^{t_i^{\max}} p_i(\tau) d\tau} \right)^{-1}, \quad (1)$$

где $K = (2n + 1)^2$ для ИНС-модели, построенной в соответствии с теоремой А.Н. Колмогорова, и $K = (2n + 1)(3n^2 + n + 1)$ для полносвязанной ИНС-модели.

Процесс обучения интеллектуального ядра ЭС заключается в минимизации функционала вида:

$$E\{\mathbf{w}[S(n)]\} = \|\Delta\| = \|\mathbf{Y}^{tabl} - \mathbf{Y}^{net}(\mathbf{w})\| = \sqrt{\sum_{i=1}^R \sum_{j=1}^n (Y_{ij}^{tabl} - Y_{ij}^{net}(\mathbf{w}))^2} \rightarrow \min_{\mathbf{w}}. \quad (2)$$

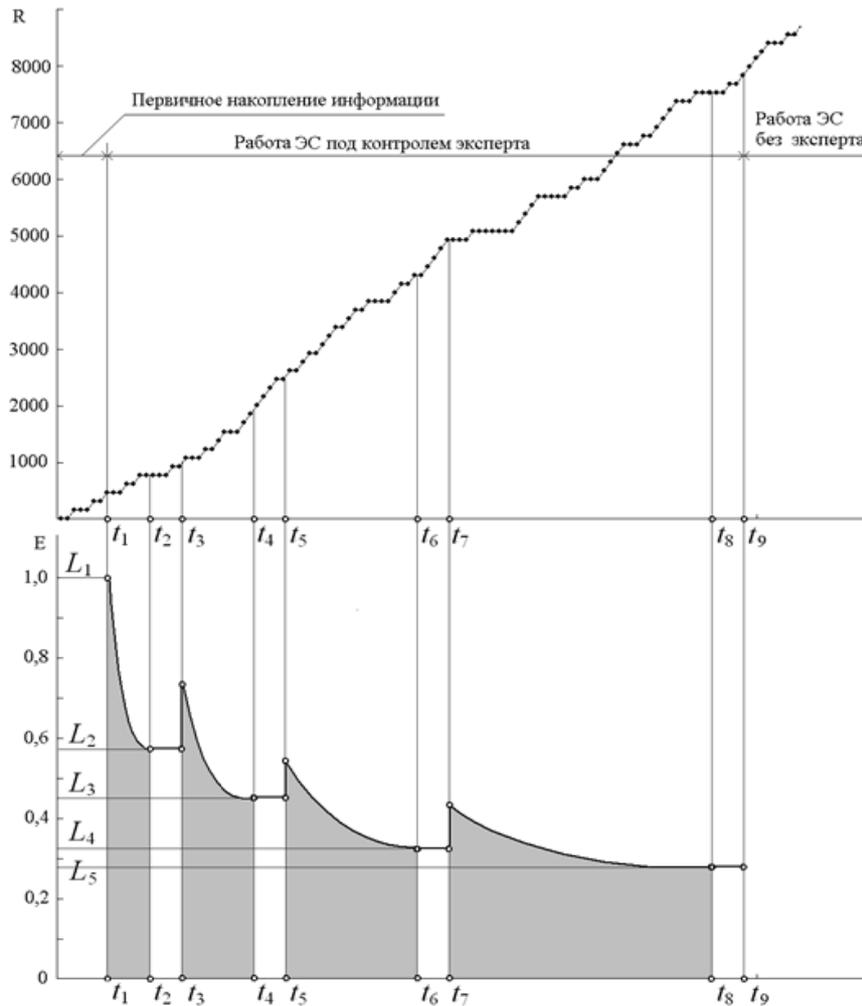


Рис. 4. График накопления первичной информации (верхняя часть рисунка): R – число записей в базе данных, полученных с терминалов; процесс обучения ИНС-модели, составляющей интеллектуальное ядро ЭС; E – приведенная среднеквадратическая ошибка ИНС-модели; по оси абсцисс – безразмерное время

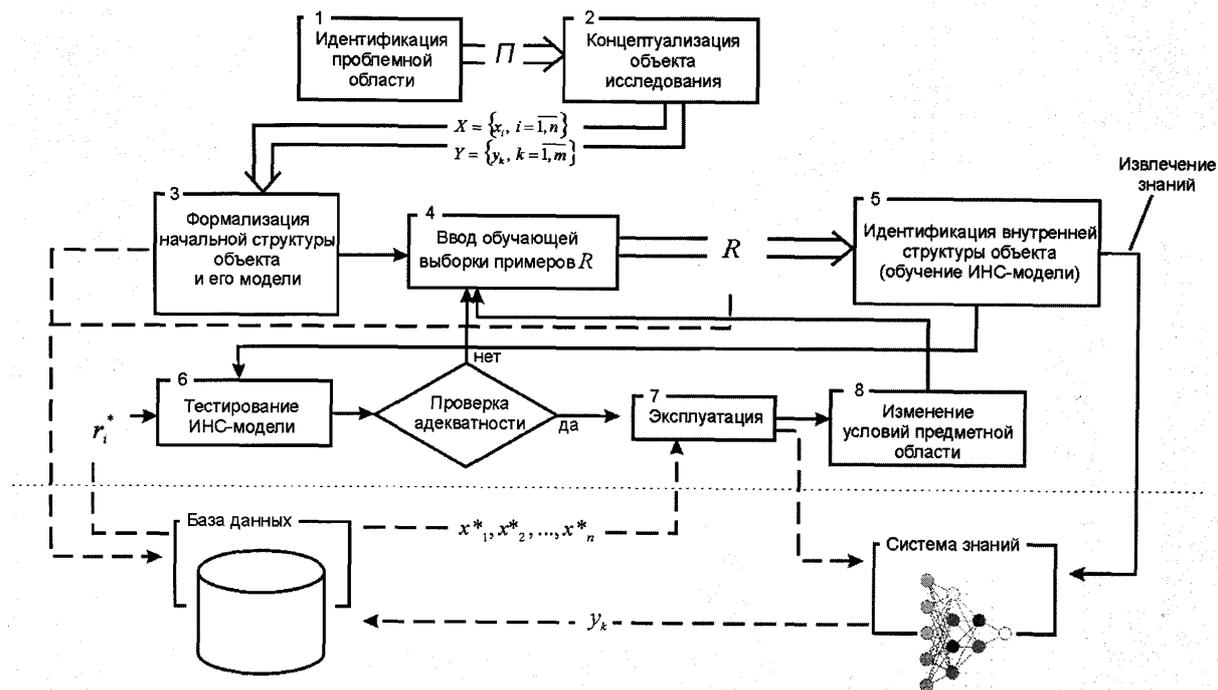


Рис. 5. Функциональная модель жизненного цикла построения ЭС, обладающих интеллектуальным ядром на основе ИНС-моделей и веб-интерфейсом

Вектор w определяется структурой сети S , которая, в свою очередь, связана с числом ее входов n так, что имеется однозначное соответствие n и $w(S)$. Поэтому вектор w^* , зависящий от структуры сети и минимизирующий функционал (2), соответствует оптимальной ИНС-модели интеллектуального ядра и может быть определен как:

$$w^* = \arg \min_{w \in \Omega} F(w) . \quad (3)$$

Минимизацию функционала (2) осуществляем двумя методами: градиентным, так, что последующие значения вектора w вычисляются по формуле $w^{t+1} = w^t - h(t) \text{grad}(E(w^t))$, и методом сканирования. Для указанных методов оценки времени обучения могут быть получены следующим образом: $\Theta_{grad} = k_2 n R$ и $\Theta_{skan} = k_3 R \alpha^N$. Здесь k_2 и k_3 – коэффициенты, зависящие от технических характеристик используемого оборудования, α – число разбиений диапазона изменения переменной.

Процесс обучения ИНС-модели, составляющей интеллектуальное ядро ЭС, зависит от количества и качества данных в обучающей выборке и может быть описан следующим дифференциальным уравнением:

$$\frac{dE(t)}{dt} = k_4 [E^* - E(t)],$$

с начальным условием:

$$E(0) = E_0 , \quad (4)$$

где $E(t)$, E_0 и E^* – значения приведенной среднеквадратичной погрешности, ее начального значения и уровня, на котором она будет зафиксирована в конце цикла обучения ИНС-модели; k_4 – параметр удельной скорости обучения, зависящий от n , объема обучающей выборки и используемого метода минимизации невязки (2); t – безразмерное время.

Разделяя переменные в уравнении (4) и интегрируя, получим:

$$\int \frac{dE(t)}{k(E^* - E(t))} = \int dt$$

или

$$-\frac{1}{k} \ln[k(E^* - E(t))] = t + c. \quad (5)$$

Определив постоянную интегрирования из начальных условий (4), получим решение в виде:

$$E(t) = E^*(1 - e^{-kt}) + E_0 e^{-kt}. \quad (6)$$

Предложенная математическая формализация на практике позволяет решить следующие задачи: оценить первичный период накопления информации, необходимой для построения ИНС-модели исследуемого объекта; определить начальную структуру ИНС-модели на основе входных характеристик объекта исследования; сделать оценку количества циклов обучения интеллектуального ядра в присутствии эксперта.

На основе полученных результатов была разработана функциональная модель (рис. 5) жизненного цикла процесса построения ЭС с интеллектуальным ядром на основе нейросетевых моделей и поддерживающих механизмы удаленного доступа.

Практическая реализация. Информационная технология построения ЭС на основе нейросетевой модели и с распределенным вводом данных, соответствующая функциональной модели ЭС (рис. 5), заключается в последовательном выполнении следующих шагов:

- изучение объекта, для которого разрабатывается ЭС; выделение векторов входных характеристик и возможных выходных состояний данного объекта;
- накопление первичной информации об объекте в БД посредством распределенной сетевой системы сбора данных в течение периода, определяемого аналитической моделью (1);
- формирование первичной структуры ИНС-модели в соответствии с теоремой А.Н. Колмогорова и рис. 2;
- планирование поэтапного процесса обучения развивающейся ИНС-модели по предложенному уравнению (6); обучение ИНС-модели согласно предложенным аналитическим моделям;
- совершенствование ИНС-модели на основе вновь поступающих эмпирических данных под контролем эксперта, который определяет законченность формирования системы представления знаний;
- эксплуатация ЭС (без эксперта).

На основе сравнения разработанной технологии с существующими аналогами выделены ее функциональные преимущества, заключающиеся в следующем: универсальность – применительно к различным предметным областям (без внесения конструктивных изменений в программное обеспечение); многопользовательский доступ к одному объекту исследования; автоматизация процесса проектирования базы знаний за счет конструктивных алгоритмов построения ИНС-модели; поддержка сетевых технологий взаимодействия.

Выявлены условия, при которых может быть осуществлено применение разработанной технологии, такие как: возможность формализовать начальную структуру объекта и закодировать примеры его диагностики в числовом виде; наличие эксперта, обладающего высоким профессиональным уровнем (выше уровня рядового специалиста), способного адекватно оцени-

вать параметрические данные объекта; достаточная квалификация пользователей, имеющих опыт использования сервисов, предоставляемых консорциумом WWW.

Осуществлена также компьютерная реализация информационной технологии на основе разработки специализированного программного комплекса.

Организация программного комплекса, который представляет собой веб-ориентированную интерактивную систему, состоящую из двух подсистем: информационной и интеллектуальной – представлена на рис. 6.

Основным назначением информационной подсистемы являются накопление, хранение, визуализация технической информации, а также обеспечение веб-ориентированного интерфейса. Компонентами подсистемы являются программно-логическое ядро, реализованное с помощью PHP-скриптов, и серверная система управления базами данных MySQL, с помощью которой была разработана структура связанных между собой таблиц, каждая из которых хранит определенную техническую информацию.

Интеллектуальная подсистема предназначена для реализации процессов построения, обучения, модификации, эксплуатации ИНС-моделей, являющихся самостоятельными системами представления знаний в ЭС. Подсистема включает 4 основных компонента: блок мониторинга данных, который обеспечивает передачу входных и выходных данных, отслеживает поступление новой информации в базу данных; преобразователь, осуществляющий подготовку обучающей выборки для сети; конструктор, обеспечивающий автоматизированное построение и обучение интеллектуальной модели (ИНС) исследуемого объекта.

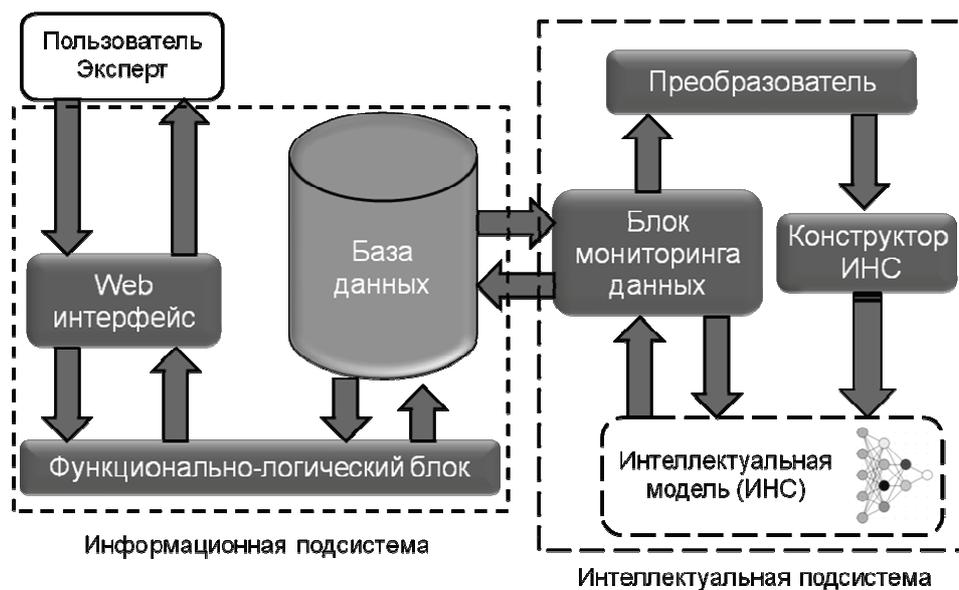


Рис. 6. Структурная схема организации программного комплекса

Для апробации разработанного программного комплекса были построены три экспертные системы в двух предметных областях: медицинской и социологической.

Медицинская ЭС решала проблему общеклинического анализа крови по десяти показателям (уровень гемоглобина (г/л); скорость оседания эритроцитов (мм/ч); лейкоциты (10⁹/л); эритроциты (10¹²/л); цветовой показатель; лейкоцитарная формула (5 составляющих)). Для построения модели знаний ЭС потребовалась обучающая выборка объемом в 350 примеров, при которой уровень погрешности ИНС модели достиг допустимого значения и стабилизировался, следовательно при этом ЭС стала пригодна для использования в режиме эксплуатации. Входными данными для ЭС были значения показателей крови, по рассчитанному ИНС-моделью выходному значению определялось состояние здоровья пациента (выходы: 1 – здоров, 2 – стационарное лечение, 3 – амбулаторное лечение, 4 – экстренная госпитализация (табл. 1)).

Таким образом, получена структура ИНС-модели (рис. 7), включающая 10 входных нейронов, 210 нейронов первого скрытого слоя, 21 нейрон второго скрытого слоя, один выходной суммирующий нейрон (в табл. 2 представлен фрагмент массива значений весовых коэффициентов синаптических связей нейронов, полученного в результате обучения ИНС-модели).

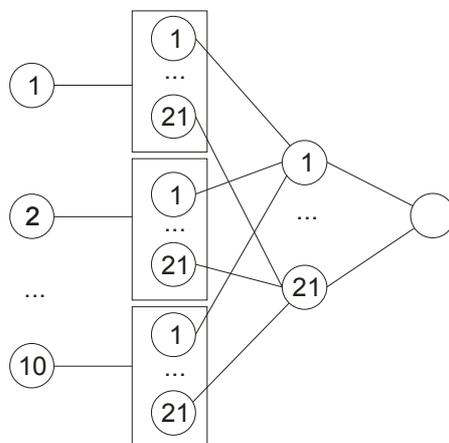


Рис. 7. Структурная схема ИНС-модели медицинской ЭС

Таблица 1

Результаты тестирования медицинской ЭС

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	y
1	134	5	5,4	4,19	0,96	0	3	60	30	8	3
2	111	5	5,9	3,82	0,87	0	5	53	36	6	2
...
50	137	6	3,2	4,28	0,96	1	1	53	30	9	1

Таблица 2

Коэффициенты синаптических связей нейронов входного и первого скрытого слоев
ИНС-модели медицинской ЭС

Номера нейронов	1	2	3	4	5	6	7	8	9	10	11	12	13	14	...	210
1	0,074	0,052	0,099	0,063	0,076	0,080	0,090	0,034	0,033	0,064	0,071	0,050	0,089	0,078	...	0,040
2	0,047	0,009	0,039	0,053	0,079	0,036	0,070	0,099	0,068	0,019	0,096	0,090	0,011	0,011	...	0,038
...
10	0,013	0,001	0,057	0,006	0,018	0,027	0,076	0,087	0,008	0,017	0,086	0,054	0,074	0,079	...	0,012

При тестировании ЭС на данных, являющихся результатами общеклинического анализа крови и не участвовавших в процессе построения ИНС-модели, правильно было распознано 70 % введенных данных, для 25 % была дана близкая классификация, в 5 % случаев система ошиблась. Указанные показатели позволяют сделать вывод о применимости разработанной ЭС в медицинских учреждениях в качестве интеллектуального помощника при проведении общеклинического анализа крови начинающими специалистами.

Результаты этого примера применения и других (для социологических исследований), позволяют сделать вывод о работоспособности программного комплекса, являющегося практической реализацией разработанной автоматизированной технологии.

ЛИТЕРАТУРА

1. Арзамасцев А.А., Зенкова Н.А., Неудахин А.В. Автоматизированная технология построения экспертных систем с интеллектуальным ядром на основе ИНС-моделей // Открытое образование. 2008. № 3 (68). С. 35-39.
2. Арзамасцев А.А., Зенкова Н.А., Неудахин А.В. Технология построения медицинской экспертной системы на основе аппарата искусственных нейронных сетей // Информационные технологии. 2009. № 8. С. 60-63.
3. Арзамасцев А.А., Зенкова Н.А., Неудахин А.В. Формализация проблемы разработки экспертной информационной системы с развивающимся интеллектуальным ядром на базе ИНС-моделей // Вестн. Тамб. ун-та. Сер. Естеств. и техн. науки. Тамбов, 2010. Т. 15. Вып. 1. С. 287-290.
4. 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.
5. Колмогоров А.Н. О представлении непрерывных функций нескольких переменных в виде суперпозиции непрерывных функций одного переменного // Докл. АН СССР. 1957. Т. 114. № 5. С. 953-956. (Цит. по книге: Горбань А.Н., Россиев Д.А. Нейронные сети на персональном компьютере. Новосибирск: Наука, 1996).

Поступила в редакцию 1 сентября 2010 г.

Arzamastsev A.A., Zenkova N.A., Neudahin A.V. Development of of expert informational system with developing intellectual kernel on the basis of ANN-models

Article is devoted to working out of the automated technology of construction of the expert systems combining functional advantages of the developing intellectual kernel and network technologies of interaction. Results of mathematical formalization of accumulation problems of information and construction on its basis the developing intellectual kernel are given.

Key words: automated expert system; developing intellectual kernel; neuro-net models; mathematical formalization of problems of accumulation of information.