



УДК 681.327.12.001.362

© 2005 г. **Н.С. Безруков,**
Е.Л. Еремин, д-р техн. наук
(Амурский государственный университет, Благовещенск)

ПОСТРОЕНИЕ И МОДЕЛИРОВАНИЕ АДАПТИВНОЙ НЕЙРО-НЕЧЕТКОЙ СИСТЕМЫ В ЗАДАЧЕ МЕДИЦИНСКОЙ ДИАГНОСТИКИ

В работе предлагается один из подходов к построению медицинских систем диагностики на основе аппарата нечеткой логики и искусственных нейронных сетей. Функционирование разработанной имитационной модели показано на примере диагностики онкологического заболевания.

Введение

К концу XX в. дисбаланс между объемом информации и возможностями клинициста-аналитика по ее переработке значительно возрос, причем эта тенденция сохраняется и в настоящее время. Происходящее имеет место в силу того, что, наряду с резким увеличением потоков информации и традиционно невысокой скоростью ее обработки, методы и способы анализа информации развиваются достаточно медленно. Об объеме медицинских данных, складированных на разных типах носителей, может свидетельствовать тот факт, что только научные учреждения за один день записывают и фиксируют информации примерно на 1 терабайт (данные аналитического отдела американской компании GTE). В такой ситуации проблема анализа информации оказывается поистине узким местом, в частном – ограничивающим оперативность медицинской помощи, в общем – сдерживающим развитие медицинской науки [1].

Современные методы интеллектуального анализа данных, способы автоматизированного извлечения знаний непосредственно из исходных данных отчасти способны снять остроту указанной проблемы. Тем более, что такие методы уже успешно используются, например, в экономике при анализе рынка, в маркетинге, прогнозе фондовых котировок и других бизнес-приложениях [2]. Однако в области медицины методы интеллектуальной поддержки принятия решений, основанные на компьютерной обработке и анализе данных, еще не имеют достаточно широкого применения, хотя автоматизированный анализ информации уже позволяет решать ряд ме-

дицинских задач, в частности связанных со стратегией лечения больного (относительно цены и качества), равномерной загрузкой объектов медицинского обслуживания (поликлиник, лабораторий, диагностических центров), диагностированием заболеваний и т.д.

Некоторые способы диагностики на основе нечеткой логики

В работах [3, 4], в рамках теории нечетких множеств, предложен способ формирования диагностирующих моделей, где разработчику системы диагностики предлагается вести диалог с экспертом, в ходе которого он определяет функции принадлежности для входов и выходов и правила их взаимосвязи на основе «частотного» метода или метода парных сравнений степеней принадлежности. Иной подход, но уже без диалога с экспертом рассмотрен в работе [1], где для построения модели диагностики ракового заболевания используется метод «деревья решений» и регрессионных моделей. Уместно отметить то обстоятельство, что модель «деревья решений» в некоторых выборках позволяет правильно определить диагноз, используя не все диагностические признаки заболевания. Это важно потому, что на практике для диагностирования какого-либо заболевания врач тоже не требует от пациента собирать сразу все возможные анализы (признаки болезни). Действительно, врач снимает признаки последовательно и группами, т.е. когда он не способен диагностировать на первой группе, то он просит пациента предоставить еще какие-либо анализы, собирая тем самым полную картину о пациенте до тех пор, пока не поставит диагноз.

Но какие информационные факторы следует объединять в единую группу при отсутствии врача? Одно из решений подобной задачи может быть получено с помощью методов разделения независимых компонент BSS (blind signal separation) [5] и анализа независимых компонент ICA (independent component analysis) [6]. Метод ICA является следствием BSS и имеет различные сферы применения, в том числе используется и при анализе электроэнцефалограмм мозга. Оба метода объемны, но их основой является анализ энтропии и разброса данных (гауссово распределение). Достаточно простой алгоритм этого класса, функционирование которого связано с оценкой энтропии – box-counting (подсчет ячеек), рассмотрен в [2].

Кроме того, при диагностировании болезни у одних пациентов для точного диагноза достаточно небольшого числа признаков, а другим требуется дополнительное обследование. Зачастую это обусловлено тем, что между здоровыми и больными иногда нет четкой разделяющей границы.

В таких случаях для построения модели диагностики целесообразно применять гибридные сети [7]. В качестве таковой можно воспользоваться известной сетью, реализованной в программной среде Matlab – это так называемая адаптивная нечеткая нейронная система вывода ANFIS (Adaptive Neuro-Fuzzy Inference System). В ANFIS выводы делаются на основе аппарата нечеткой логики, а соответствующие функции принадлежности само-

настраиваются с помощью алгоритма обратного распространения ошибки [8], в частности при обучении, выявлении эффекта переобучения и определении количества настраиваемых коэффициентов нейронной сети. При этом характерно, что для сокращения сроков обучения в нейронную сеть невозможно ввести какую-либо априорную информацию, – например, знания эксперта.

С другой стороны, эффект переобучения нейронной сети, когда сеть заучивает выборку вместо того, чтобы ее обобщать, как показано в работе [2], можно уменьшить. Этого можно достичь за счет искусственного увеличения обучающей выборки или путем выбора оптимального количества настраиваемых коэффициентов нейронной сети (условие минимума эффекта переобучения) по формуле

$$d \approx \frac{W^2}{P}, \quad (1)$$

где d – размерность векторного входа нейронной сети, $d \geq 2$; P – число групп данных или размер обучающей выборки для настройки нейронной сети; W – число весов (степеней свободы) в нейронной сети.

Постановка задачи

Пусть вектор-строка $x = x[n]$ сформирована из l диагностических признаков соответствующего заболевания и записана следующим образом:

$$x[n] = (x_1[n], x_2[n], \dots, x_l[n]), \quad n = 1, 2, \dots, P, \quad x = x[n] \in R^l,$$

где R^l – известное множество числовых значений всех l -признаков.

Требуется построить и обучить диагностическую систему:

используя сходящуюся древовидную структуру, узлами которой являются однотипные гибридные сети $G(\mathbf{u})$, с векторными входами $\mathbf{u} = \mathbf{u}[n]$, составленными из некоторых диагностических и внутренних (промежуточных) признаков;

определяя в ходе обучения априорно неизвестные параметры гибридных сетей, с помощью алгоритма обратного распространения ошибки и в условиях выполнения требования о минимуме эффекта переобучения гибридной сети

$$P \approx \frac{(8d + 2^d(d+1))^2}{d} \quad (2)$$

таким образом, чтобы по заданному математическому описанию гибридной сети финального слоя древовидной структуры

$$y = y[n] = \text{sign}(G_{fin}(\mathbf{u}) - N) + N, \quad N = \frac{y_{\min} + y_{\max}}{2}, \quad (3)$$

где y – выход (диагноз), принимающий одно из значений y_{\min} (здоров) или y_{\max} (болен), система диагностики обеспечивала бы вычисление диагноза соответствующего заболевания.

Разработка диагностической системы

Диагностическая система имеет сходящуюся древовидную структуру, состоящую из входного (нулевого) слоя с диагностическими l -признаками и последующими k -слоями, каждый из которых имеет m -узлов. Количество слоев и узлов в каждом слое определяется по следующему алгоритму до момента, пока $t_k \leq 0$ и $m_k = 1$:

$$\left\{ \begin{array}{l} m_1 = \frac{m_0 - t_1}{d}, \\ m_2 = \frac{m_1 + t_1 - t_2}{d}, \\ m_3 = \frac{m_2 + t_2 - t_3}{d}, \\ \vdots \\ m_k = \frac{m_{k-1} + t_{k-1} - t_k}{d}, \end{array} \right. \quad (4)$$

$$M = [m_1, m_2, \dots, m_k], \quad m_0 = l.$$

где k – количество слоев в диагностической системе; m_k – число, округленное до меньшего целого от деления (m_k / d) и определяющее количество узлов в k -слое; $t_k = m_{k-1} - dm_k$ – остаток от деления, определяющий количество незадействованных признаков в k -слое, переходящих на следующий слой.

На каждый узел слоя должны последовательно подаваться группы диагностических и(или) внутренних признаков:

$$X_{i,j} = (x_1, x_2, \dots, x_d)_{i,j}, \quad i = \overline{1, 2, \dots, k}, \quad j = \overline{1, 2, \dots, m_i}. \quad (5)$$

При формировании групп используются следующие правила:

1) признаки объединяются в группу, если они «не предсказывают» друг друга, т.е. когда они имеют наименьшие значения кросс-энтропии, которую вычисляют по формуле

$$I(x, y) = \log_{10} \frac{N_{X_1} \cdot N_{X_2} \cdot \mathbf{K} \cdot N_{X_d}}{N_{X_1, X_2, \mathbf{L}, X_d}}, \quad (6)$$

где N_X – количество промежутков с наличием X_1 ; N_d – количество промежутков с наличием X_d ; $N_{X_1, X_2, \mathbf{L}, X_d}$ – количество клеток, в которых содержатся точки с координатами (X_1, X_2, \dots, X_d) ;

2) каждый признак (диагностический или внутренний) используется только в одной группе и только один раз;

3) признаки (диагностические или внутренние), не вошедшие в группу в верхнем слое, участвуют в соревновании на образования группы на нижних слоях.

В качестве типового узла в диагностической системе выступает гиб-

формируются группы (входные векторы) для узлов, образуемых на последующем $i+1$ -ом слое.

Иллюстративный пример

Построим, обучим и оценим работоспособность диагностической системы на примере рака груди, используя данные – исходную выборку на 683 пациентов доктора Вольберга [10, 11]. Подобная система диагностики была рассмотрена в работе [1] и реализована с помощью программы Poly-Analyst 4.6. По данным цитологического обследования (табл. 1), включающего 9 диагностических признаков ($x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9$), врач определял, больна пациентка или нет, т.е. ставил диагноз – y . Каждый из признаков менялся в диапазоне от 1 до 10, а диагноз был закодирован числами $y_{\min} = 2$ и $y_{\max} = 4$ соответственно для здоровых и больных пациентов.

Таблица 1

Фактор, или признак	Обозначение в таблице данных	Показатель	Диапазон изменения
x_1	clumpThickness	Толщина агрегации клеток	1 – 10
x_2	cellSize	Однородность размера клеток	1 – 10
x_3	cellShape	Однородность формы клеток	1 – 10
x_4	arginalAdhesion	Поверхностная адгезия	1 – 10
x_5	epithelialSize	Размер единичной эпителиальной клетки	1 – 10
x_6	bareNuclei	Обнаженные ядра	1 – 10
x_7	blandChromatin	Хроматин	1 – 10
x_8	normalNucleoli	Нормальные ядрышки	1 – 10
x_9	Mitoses	Митоз	1 – 10
y	Class	Диагноз (здоров или болен)	2 или 4

На рис. 1 изображены гистограммы признаков относительно диагноза – болен (светлые полосы) или здоров (темные полосы) человек и где для наглядности гистограмм по оси ординат отсечен ряд значений, хотя фактически некоторые из них доходят до 300. В частности, из гистограммы для седьмого диагностического признака (blandChromatin) видно, что при значениях от 1 до 2 пациент с высокой степенью уверенности здоров, а при значениях от 7 до 10 – скорее болен, чем здоров. Схожая зависимость наблюдается и по другим признакам. Достаточно очевидно, что совместная оценка признаков должна уменьшать ошибку диагноза.

С целью объективной оценки качества разрабатываемой системы диагностики, следуя [2, 3], разобьем исходную выборку на две – обучающую (480 пациентов) и проверочную (203 пациента) выборки.

Для диагностических признаков обучающей выборки

$$x^T = (x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9), \quad (10)$$

построение диагностической системы будем осуществлять согласно (3),

обеспечивая выполнение условия минимума эффекта переобучения
 $P = 480 \Rightarrow d = 2.$ (11)

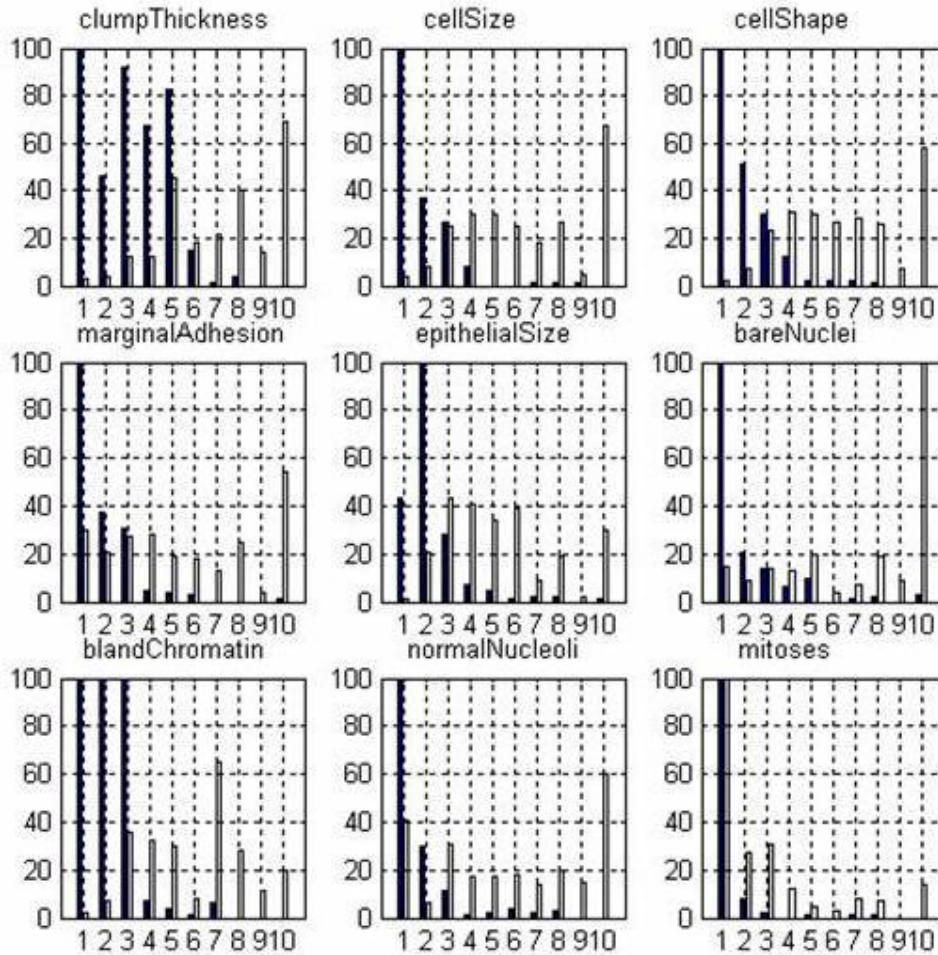


Рис. 1. Гистограммы диагностических признаков.

Структура диагностической системы состоит из 4 слоев, на вход первого поступает вектор из 9 диагностических признаков. Количество узлов в каждом слое согласно (4) определяется соотношениями

$$\begin{cases} m_1 = \frac{9-1}{2} = 4, \\ m_2 = \frac{4+1-1}{2} = 2, \\ m_3 = \frac{2+1-1}{2} = 1, \\ m_4 = \frac{1+1-0}{2} = 1, \end{cases} \quad (12)$$

т.е. $M = [4, 2, 1, 1]$.

Один из вариантов заполнения ячеек значениями кросс-энтропии, вычисленными по формуле (6), на примере диагностических признаков x_1 и x_9 , при количестве $N_{x_1} = 10$, $N_{x_9} = 9$, $N_{x_1, x_9} = 53$ – соответственно не нуле-

вых строк, столбцов и ячеек таблицы, представлен в табл. 2.

Таблица 2

		Значения фактора x_9									
		1	2	3	4	5	6	7	8	9	10
Значения фактора x_1	1	93	1	0	0	0	0	1	1	0	0
	2	31	0	0	1	0	0	0	0	0	0
	3	74	1	2	1	0	0	1	1	0	0
	4	52	1	0	0	0	0	0	0	0	0
	5	70	4	6	1	2	0	0	0	0	1
	6	15	4	3	0	0	0	2	0	0	1
	7	6	3	4	3	0	0	1	0	0	0
	8	22	5	1	2	0	1	2	2	0	2
	9	5	1	1	1	1	1	0	0	0	1
	10	20	7	7	3	1	0	1	3	0	3

Результаты вычисления по формуле (6) значений кросс-энтропии для всех диагностических признаков по отношению друг к другу, приведены в табл. 3. Очевидно, что наибольшее значение кросс-энтропия принимает у одноименных признаков.

Таблица 3

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
x_1	1.0000	0.1612	0.1367	0.1308	0.1739	0.1549	0.1367	0.1427	0.2300
x_2	0.1612	1.0000	0.2076	0.1427	0.2147	0.1612	0.1871	0.1367	0.2382
x_3	0.1367	0.2076	1.0000	0.1192	0.2076	0.1675	0.1739	0.1308	0.2640
x_4	0.1308	0.1427	0.1192	1.0000	0.1871	0.2218	0.1805	0.1192	0.2382
x_5	0.1739	0.2147	0.2076	0.1871	1.0000	0.2076	0.1871	0.1739	0.2821
x_6	0.1549	0.1612	0.1675	0.2218	0.2076	1.0000	0.1871	0.1612	0.2821
x_7	0.1367	0.1871	0.1739	0.1805	0.1871	0.1871	1.0000	0.1938	0.2218
x_8	0.1427	0.1367	0.1308	0.1192	0.1739	0.1612	0.1938	1.0000	0.2640
x_9	0.2300	0.2382	0.2640	0.2382	0.2821	0.2821	0.2218	0.2640	0.9542

Последовательно выбирая наименьшие значения кросс-энтропии – так, чтобы каждый из признаков был задействован только один раз (в табл. 3 такие значения выделены жирным курсивом), получаем для первого слоя следующее разбиение на группы:

$$(x_4, x_3)_{1,1}, (x_7, x_1)_{1,2}, (x_8, x_2)_{1,3}, (x_6, x_5)_{1,4}, \quad (13)$$

где признак x_9 не вошел ни в одну из групп.

Используя для реализации разрабатываемой системы диагностики программную среду Matlab [7], проведем обучение ее гибридных сетей, обучая каждую из сетей на 20 циклах.

Структура гибридной нейронной сети с двумя входами и одним выходом $G(\Theta_{1,1}, \Theta_{1,2}, \Theta_{2,1}, \Theta_{2,2}, a_{0,1}, \dots, a_{2,4}, x_1, x_2)$ однотипна для всех узлов и описывается согласно (7) следующими правилами:

$$\begin{aligned} \Pi_1: \quad & \text{если } x_1 \text{ есть } A_{1,1}(\Theta_{1,1}, x_1) \text{ и } x_2 \text{ есть } A_{2,1}(\Theta_{2,1}, x_2), \\ & \text{то } z_1 = a_{0,1} + a_{1,1}x_1 + a_{2,1}x_2; \end{aligned}$$

- П₂: если x_1 есть $A_{1,1}(\Theta_{1,1}, x_1)$ и x_2 есть $A_{2,1}(\Theta_{2,1}, x_2)$,
то $z_2 = a_{0,2} + a_{1,2}x_1 + a_{2,2}x_2$;
- П₃: если x_1 есть $A_{1,1}(\Theta_{1,1}, x_1)$ и x_2 есть $A_{2,1}(\Theta_{2,1}, x_2)$,
то $z_3 = a_{0,3} + a_{1,3}x_1 + a_{2,3}x_2$;
- П₄: если x_1 есть $A_{1,1}(\Theta_{1,1}, x_1)$ и x_2 есть $A_{2,1}(\Theta_{2,1}, x_2)$,
то $z_4 = a_{0,4} + a_{1,4}x_1 + a_{2,4}x_2$.

Структурная схема гибридной сети с двумя входами и одним выходом изображена на рис. 2, где буквой T обозначен блок вычисления так называемой T -нормы.

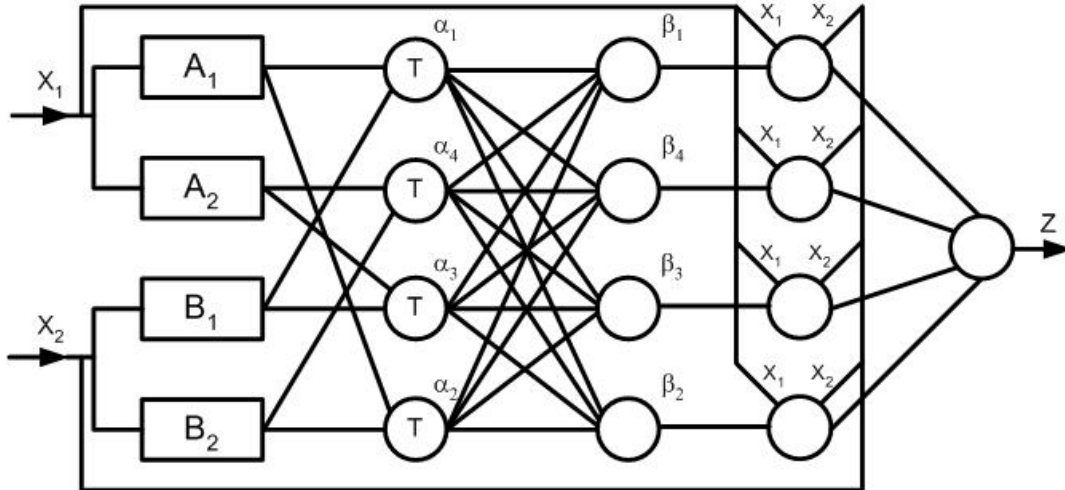


Рис. 2. Структурная схема гибридной сети с двумя входами и одним выходом.

Четкое значение выходной переменной гибридной сети, следуя (8), определяется в виде

$$Z = \frac{a_1 z_1 + a_2 z_2 + a_3 z_3 + a_4 z_4}{a_1 + a_2 + a_3 + a_4} = \sum_{j=1}^4 b_j z_j.$$

Соединяя выходы первого слоя по правилу наименьшего значения кросс-энтропии (рис. 3), получаем группы для второго слоя диагностической системы $(y_{1,1}, y_{1,2})_{2,1}$, $(y_{1,3}, y_{1,4})_{2,2}$. Следует отметить, что диагностический признак x_9 последовательно участвовал в соревновании на вхождение в группу каждого из слоев, но в силу большого значения своей кросс-энтропии (см. рис. 1) он эти соревнования проигрывал и поэтому смог участвовать только в образовании последней группы – 4-го слоя.

Диагноз на выходе системы вычислялся согласно формуле (3) при значении $N = 3$.

Завершив построение и обучение диагностической системы, проведем оценку качества ее функционирования, вычисляя ошибки по каждому из ее узлов. Ошибкой системы будем считать неверно поставленный диагноз относительно «истинного» значения, т.е. известного врачебного диагноза. Разброс значений ошибок по каждому слою системы представлен на рис. 4. Характер изменения ошибок на проверочных (пунктирная линия) и обучающих (сплошная линия) данных представлен на рис. 4.

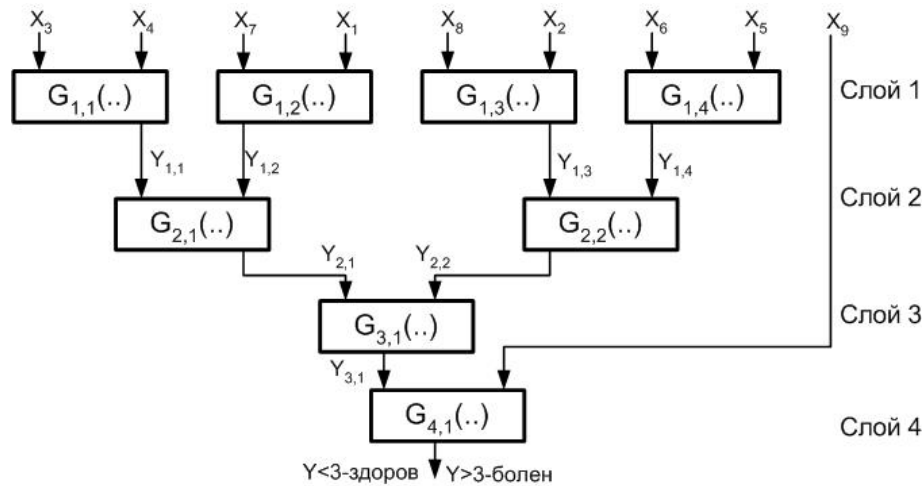


Рис. 3. Блок-схема системы диагностики заболевания рака груди.

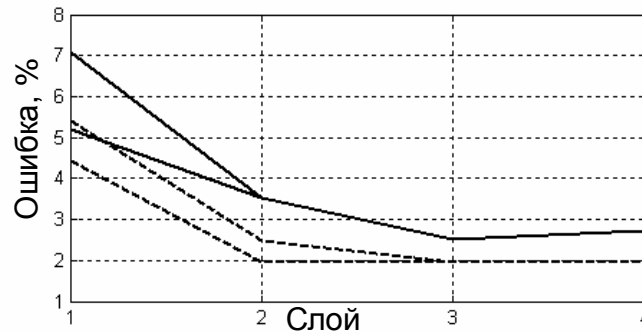


Рис. 4. Ошибки системы диагностики по слоям на проверочных и обучающих данных.

Из графиков на рис. 4 следует, что для рассматриваемого примера ошибка диагностики заметно убывает при переходе от 1-го слоя к 3-му, а последний – 4-й слой практически не улучшает качества диагностики, что позволяет структуру системы диагностики упростить и построить ее как 3-слойную. Последнее, вообще говоря, было понятно и раньше, когда отмечалось, что диагностический признак x_9 , участвующий в образовании 4-го слоя, имеет большое значение кросс-энтропии, т.е. фактически является малоинформативным признаком.

Заключение

Ошибки на выходе имитационной модели разработанной системы диагностики составили при тренировочных данных – 2,71%, а при проверочных – 1,97%. Поскольку значения ошибок не только соизмеримы, но и достаточно малы, то разработанную систему можно считать вполне работоспособной. Отметим также, что полученные значения ошибок оказались меньше, чем у аналогичной диагностической системы реализованной с помощью программы PolyAnalyst 4.6 [1], где ошибка диагноза составила около 3÷5 %.



ЛИТЕРАТУРА

1. *Арсеньев С.* Извлечение знаний из медицинских баз данных, <http://www.megaputer.ru>.
2. *Ежов А.А., Шумский С.А.* Нейрокомпьютинг и его применение в экономике и бизнесе, <http://www.neuroproject.ru>.
3. *Королева С.А.* Разработка методов и средств контроля и прогнозирования состояния здоровья в задачах профессионального отбора на основе нечеткой логики принятия решений / Дис. ...канд. техн. наук. Курск, 2005.
4. *Уварова А.Н.* Разработка методов и средств комплексной диагностики шизофрении на основе нечеткой логики принятия решений / Дис. ...канд. техн. наук. Курск, 2004.
5. *Cardoso J.-F.* Blind signal separation: statistical principles. CNRS and ENST, <http://www.tsi.enst.fr/~cardoso.html>.
6. *Huvarinen A., Oja E.* Independent Component Analysis: A Tutorial. – Laboratory of Computers and Information Science, <http://www.cis.hut.fi/projects/ica>.
7. *Круглов В.В., Дли М.И., Голунов Р.Ю.* Нечеткая логика и искусственные нейронные сети, <http://www.neuroproject.ru>.
8. *Рутковская Д., Пилиньский В., Рутковский Л.* Нейронные сети, генетические алгоритмы и нечеткие системы: Пер. с польского И.Д. Рудинского. М: Горячая линия – Телеком, 2004.
9. *Tagagi T., Sugeno M.* Fuzzy Identification of Systems and Its Applications to Modeling and Control //IEEE Transactions on Systems, Man and Cybernetics. 1985. Vol. 15. P.116-132.
10. <http://sci2s.ugr.es/keel-dataset/problemas/clasificacion/wisconsin.php#dataset>.
11. <http://www.wisc.edu/wolberg/breast.html>.

УДК 531/534:57+612.7

© 2005 г. **А.В. Бушманов**, канд. техн. наук,
(Амурский государственный университет, Благовещенск)

ПРОЧНОСТНОЙ АНАЛИЗ ЧРЕСКОСТНЫХ ФИКСИРУЮЩИХ УСТРОЙСТВ

Рассматривается задача определения прочности конструкций для чрескостного остеосинтеза в зависимости от расположения фиксирующих устройств на кости.

Введение

В последние годы в травматологии формируется концепция функционального лечения переломов, сущность которой заключается в репозиции и удержании отломков с сохранением движений в смежных суставах.