

И.Г. Перова¹, Е.В. Бодянский²¹ ХНУРЭ, г. Харьков, Украина, rikywenok@gmail.com;² ХНУРЭ, г. Харьков, Украина, yevgeniy.bodyanskiy@nure.ua

НЕЙРО-ФАЗЗИ СИСТЕМА ДЛЯ ЗАДАЧ ОБРАБОТКИ МЕДИЦИНСКИХ ДАННЫХ В СИТУАЦИЯХ МНОЖЕСТВА ДИАГНОЗОВ

В работе предложена структура и алгоритм обучения диагностирующей нейро-фаззи системы, предназначенной для обработки медицинских данных в онлайн режиме. Предлагаемая система может работать на малых объемах обучающей выборки в режиме реального времени, а также в ситуациях, когда у пациента имеются сопутствующие заболевания, то есть возможен вариант наличия одновременно нескольких диагнозов.

НЕЙРО-ФАЗЗИ СИСТЕМА, ФАЗЗИФИКАЦИЯ, ГИПЕРБОЛИЧЕСКИЙ ТАНГЕНС, АКТИВАЦИОННАЯ ФУНКЦИЯ, СИНАПТИЧЕСКИЙ ВЕС

Введение

Для решения широкого круга задач интеллектуального анализа данных (Data Mining), связанных прежде всего с диагностикой, классификацией, кластеризацией, распознаванием образов и т.п. сегодня все чаще используют методы вычислительного интеллекта (Computational intelligence) и, в первую очередь, мягкие вычисления (Soft Computing) и машинное обучение (Machine Learning) [1-8]. Здесь в качестве наиболее эффективных можно выделить нейро-фаззи-системы, благодаря их способности к обучению, включая самообучение, универсальным аппроксимирующим возможностям, лингвистической интерпретируемости и «прозрачности» получаемых результатов. Здесь наибольшее распространение получили ANFIS и TSK-системы различного порядка как аппроксиматоры и экстраполяторы и NEFCLASS [9] и его различные модификации, ориентированные на решение задач классификации — распознавания образов.

Вместе с тем существует достаточно широкий класс задач, где отмеченные системы неэффективны. Прежде всего, это задачи, где обучающая выборка коротка, данные на обработку подаются последовательно в виде потока [10], а собственно обучение системы должно происходить одновременно с анализом поступающей информации. Такая ситуация достаточно часто возникает в Medical Data Mining [11, 12] и осложняется тем, что обрабатываемые данные имеют нестационарный характер, а размерность входного пространства признаков может быть соизмерима с объемом обучающей выборки. Если же говорить о задаче диагностирования, то во-первых для редко встречающихся диагнозов выборка вообще может быть сверхмалой, а во-вторых, количество возможных диагнозов (особенно в случае массовых обследований) может изменяться по ходу процесса исследования. Также возможна ситуация, когда у пациента имеются сопутствующие заболевания и необходимо оценить уровень принадлежности к каждому из рассматриваемых диагнозов. Понятно, что традиционные диагностирующие нейро-фаззи-системы

типа NEFCLASS могут не справиться с возникающими проблемами.

1. Быстрая диагностирующая нейро-фаззи-система

Рассмотрим архитектуру диагностирующей нейро-фаззи-системы (DNFS), приведенную на рис.1 и состоящую из 6-ти последовательно соединенных слоев [13]. На входной слой системы подается $(nЧ1)$ -мерный вектор входных сигналов-признаков $x(k) = (x_1(k), x_2(k), \dots, x_n(k))^T \in R^n$, где $k = 1, 2, \dots$ — в данном случае текущее время. Первый скрытый слой системы содержит n^h функций принадлежности $\mu_{il}(x_i(k))$, $i = 1, 2, \dots, n$; $l = 1, 2, \dots, h$ и выполняет фаззификацию входного пространства признаков. При этом поскольку в рассматриваемой системе производится рассеянное разбиение этого пространства, то в качестве функции принадлежности используются стандартные колоколообразные конструкции с бесконечными носителями. Чаще всего это традиционные гауссианы или более экзотические функции, например, производные гиперболического тангенса.

Второй скрытый слой производит агрегирование уровней принадлежности, вычисленный в первом слое и состоит из h элементарных блоков умножения. Третий скрытый слой — это слой синаптических весов w_{jl} ($j = 1, 2, \dots, m$ — число возможных диагнозов, которое принимается исходя из сугубо эмпирических соображений), подлежащих уточнению в процессе обучения. Это наиболее ответственный слой DNFS, поскольку от точности и скорости обучения зависит эффективность системы в целом. Общее количество синаптических весов составляет m^h . Четвертый скрытый слой образован $m+1$ сумматорами также реализующими элементарные операции. В пятом скрытом слое, образованном m блоками деления, производится дефаззификация типа «центр тяжести». И наконец выходной (шестой) слой содержит m нелинейных активационных функций, при этом в задачах диагностики достаточно часто используются простейшие сигнум-функции, принимающие значение $+1$ в случае правильного диагноза и -1 — в противном

случае. В таком случае выходные сигналы DNFS $y_j(k)$ могут принимать только два значения ± 1 .

При подаче на вход системы вектора признаков $x(k) \in R^n$ на выходе первого скрытого слоя появляется hn значений $\mu_{li}(x_i(k))$, на выходе второго скрытого слоя – h сигналов $\prod_{i=1}^n \mu_{li}(x_i(k))$, на выходе третьего скрытого слоя – mh значений $w_{jl} \prod_{i=1}^n \mu_{li}(x_i(k))$, четвертого – $m+1$ сигналов: $\sum_{l=1}^h w_{jl} \prod_{i=1}^n \mu_{li}(x_i(k))$ и $\sum_{l=1}^h \prod_{i=1}^n \mu_{li}(x_i(k))$, пятого – $u_j(k) = \frac{\sum_{l=1}^h w_{jl} \prod_{i=1}^n \mu_{li}(x_i(k))}{\sum_{l=1}^h \prod_{i=1}^n \mu_{li}(x_i(k))} = \sum_{l=1}^h w_{jl} \frac{\prod_{i=1}^n \mu_{li}(x_i(k))}{\sum_{l=1}^h \prod_{i=1}^n \mu_{li}(x_i(k))} = \sum_{l=1}^h w_{jl} \phi_l(x(k)) = w_j^T \phi(x(k))$ и шестого – m диагностирующих сигналов $y_j(k) = \text{sign } u_j(k)$.

Таким образом рассмотренная система представляет собой модификацию нейро-фаззи-системы Ванга-Менделя [14], предназначенную для решения задач online диагностики-классификации.

Вводя в шестой слой вместо сигнум-функций активационные функции гиперболического тангенса принятые в искусственных нейронных сетях:

$$y_j(k) = \tanh \gamma u_j(k) = \frac{1 - e^{-2\gamma u_j(k)}}{1 + e^{-2\gamma u_j(k)}}$$

при этом с увеличением параметра γ функция $\tanh \gamma u_j$ как угодно близко приближается к $\text{sign } u_j$, не претерпевает при этом разрывов производной.

Вводя далее традиционный квадратичный критерий обучения

$$E_j(k) = \frac{1}{2} e_j^2(k) = \frac{1}{2} (d_j(k) - \tanh \gamma w_j^T \phi(x(k)))^2 = \frac{1}{2} (d_j(k) - \tanh \gamma u_j(k))^2$$

можно записать стандартное δ -правило обучения элементарного персептрона Розенблатта

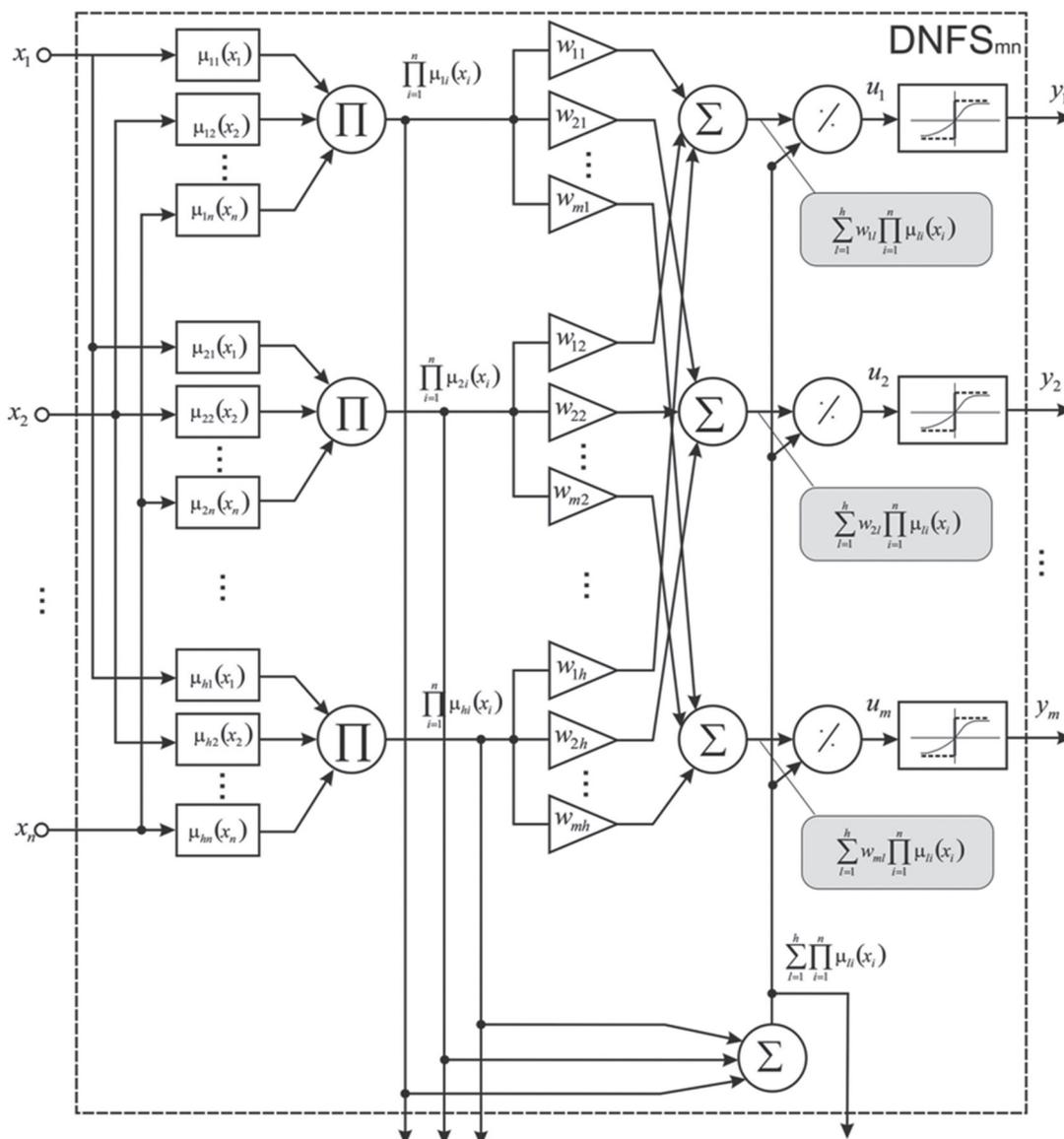


Рис. 1. Диагностирующая нейро-фаззи-система с n входами и m выходами DNFSmn

$$w_j(k+1) = w_j(k) + \eta(k)e_j(k)\gamma(1 - y_j^2(k))\phi(x(k)) = w_j(k) + \eta(k)\delta_j(k)\phi(x(k)), \quad (1)$$

где $\eta(k) > 0$ – параметр шага обучения, $\delta_j(k)$ – δ -ошибка обучения j -го вектора системы на k -м такте настройки.

Воспользовавшись идеями квазиньютоновского обучения [16] можно ввести в рассмотрение оптимизированный вариант (1) вида [17]:

$$w_j(k+1) = w_j(k) + \frac{\delta_j(k)\phi(x(k))}{\eta + \|\phi(x(k))\|^2}$$

или в матричной форме типа (2):

$$W(k+1) = W(k) + \frac{\delta(k)\phi(x(k))}{\eta + \|\phi(x(k))\|^2},$$

а при $\eta = 0$

$$W(k+1) = W(k) + \delta(k)\phi^+(x(k)),$$

где $\delta(k) = (\delta_1(k), \delta_2(k), \dots, \delta_m(k))^T$,

$$\begin{aligned} \delta_j(k) &= e_j(k)\gamma(1 - y_j^2(k)) = \\ &= (d_j(k) - \tanh \gamma u_j(k))\gamma(1 - (\tanh \gamma u_j(k))^2). \end{aligned}$$

Улучшить аппроксимирующие свойства рассматриваемой системы можно, вводя дополнительно контур обучения параметра γ , задающего форму (“крутизну”) активационной функции, хотя обычно этот параметр полагается постоянным. Для этого может использоваться подход, предложенный в [18], где для настройки предлагается использовать градиентную оптимизацию.

Записывая j -й выход системы в виде

$$y_j(k) = \tanh \gamma_j u_j(k) = \frac{1 - e^{-2\gamma_j u_j(k)}}{1 + e^{-2\gamma_j u_j(k)}},$$

а критерий обучения –

$$\begin{aligned} E_j(k) &= \frac{1}{2}(d_j(k) - \tanh \gamma_j w_j^T \phi(x(k)))^2 = \\ &= \frac{1}{2}(d_j(k) - \tanh \gamma_j u_j(k))^2, \end{aligned}$$

можно вместо традиционного δ -правила обучения (1) записать модифицированный алгоритм

$$\begin{cases} w_j(k+1) = w_j(k) + \eta_w(k)e_j(k)\gamma_j(k) \cdot \\ \cdot (1 - y_j^2(k))\phi(x(k)) = \\ = w_j(k) + \eta_w(k)\gamma_j(k)\delta_j(k)\phi(x(k)), \\ \gamma_j(k+1) = \gamma_j(k) + \eta_\gamma(k)e_j(k)(1 - y_j^2(k))u_j(k) = \\ = \gamma_j(k) + \eta_\gamma(k)\delta_j(k)u_j(k) = \\ = \gamma_j(k) + \eta_\gamma(k)\delta_j(k)w_j^T(k)\phi(x(k)). \end{cases} \quad (3)$$

Вводя далее составные векторы настраиваемых параметров $\tilde{w}_j^T(k) = (w_j^T(k), \gamma_j(k))$ и сигналов $\phi_j(x(k)) = (\gamma_j(k)\phi^T(x(k)), u_j(k))$, несложно ввести алгоритм обучения всех параметров выходных перцептронов вида

$$\tilde{w}_j^T(k+1) = w_j(k) + \eta(k)\delta_j(k)\phi_j(x(k))$$

или его оптимизированные варианты

$$\begin{cases} \tilde{w}_j^T(k+1) = \tilde{w}_j(k) + \frac{\delta_j(k)\phi_j(x(k))}{\eta + \|\phi_j(x(k))\|^2}, \\ \tilde{w}_j^T(k+1) = \tilde{w}_j^T(k) + \delta_j(k)\phi_j^+(x(k)) \end{cases} \quad (4)$$

Поскольку на каждом выходе системы значения $\gamma_j(k)$ в процессе обучения могут оказаться различными, отдельная настройка выходных перцептронов с помощью процедур (4) предпочтительнее матричной процедуры (2).

2. Быстрая диагностирующая нейро-фаззи-система в ситуации множества диагнозов

Достаточно часто может возникать ситуация, когда у пациента возможно наличие не одного, а сразу нескольких заболеваний, соответственно у него возможно наличие множества диагнозов. В этой связи целесообразно оценить не только факт наличия заболевания, но и нечеткий уровень принадлежности к каждому конкретному заболеванию.

Для этого в шестой выходной слой системы вместо сигмоидальных функций активации вводятся дополнительные элементы: функции активации типа «выпрямитель», используемые в нейроне Фукушимы [20, 21], сумматор и m блоков деления. При этом этот слой приобретает вид, приведенный на рис. 2.

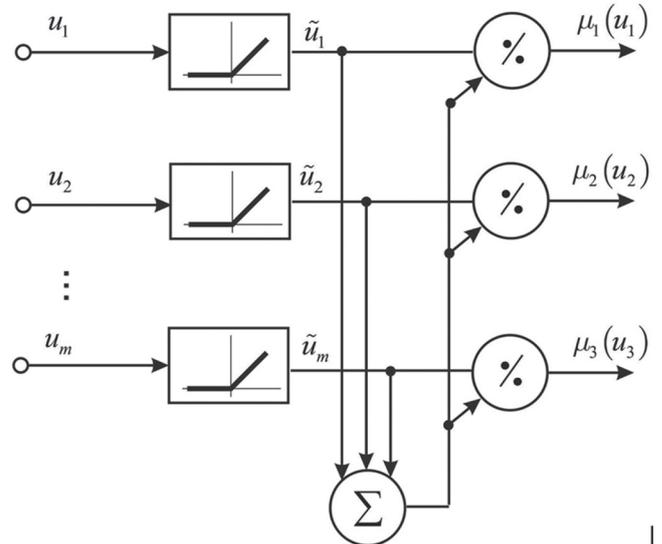


Рис. 2. Выходной слой системы

Элементы типа «выпрямитель» выделяют сигналы, определяющие сам факт наличия того или иного заболевания, а сумматор и делители вычисляют уровни принадлежности в виде

$$0 \leq \mu_j(u_j(k)) = \frac{\psi(\tilde{u}_j(k))}{\sum_{j=1}^m \psi(\tilde{u}_j(k))} = \frac{u_j(k)}{\sum_{j=1}^m u_j(k)} \leq 1,$$

где

$$\psi(u_j(k)) = \begin{cases} u_j(k), & \text{если } u_j(k) > 0 \\ 0, & \text{в противном случае.} \end{cases}$$

Необходимо заметить, что дополнительным преимуществом, обеспечиваемым такой конструкцией выходного слоя, является возможность использовать для обучения оптимальные Гауссовско-Ньютоновские процедуры типа рекуррентного метода наименьших квадратов в виде:

$$\begin{cases} w_j(k+1) = w_j(k) + \frac{P(k)(d_j(k) - w_j^T(k)\varphi(x(k)))}{1 + \varphi^T(x(k))P(k)\varphi(x(k))} \varphi(x(k)) \\ P(k+1) = P(k) - \frac{P(k)\varphi(x(k))\varphi^T(x(k))P(k)}{1 + \varphi^T(x(k))P(k)\varphi(x(k))} \end{cases} \quad (4)$$

или в матричном виде

$$\begin{cases} W(k+1) = W(k) + \frac{e(k)\varphi^T(x(k))P(k)}{1 + \varphi^T(x(k))P(k)\varphi(x(k))} \\ P(k+1) = P(k) - \frac{P(k)\varphi(x(k))\varphi^T(x(k))P(k)}{1 + \varphi^T(x(k))P(k)\varphi(x(k))} \end{cases}, \quad (5)$$

где

$$e(k) = (e_1(k), e_2(k), \dots, e_m(k))^T.$$

Важно отметить также, что использование оптимальных по быстродействию процедур (4) и (5) позволяет не только сократить время обработки информации, но и уменьшить объём обучающей выборки, что особенно важно в случае редко встречающихся диагнозов.

Выводы

В работе предложены архитектура и метод обучения диагностирующей нейро-фаззи-системы, предназначенной для решения широкого класса задач Data Stream Mining и, прежде всего, Medical Data Mining в online режиме в ситуациях, когда количество возможных диагнозов заранее неизвестно и может изменяться в процессе обучения-диагностирования, а также в тех ситуациях, когда пациент может иметь не один, а несколько диагнозов одновременно. Предлагаемая система проста в численной реализации и характеризуется высокой скоростью обучения, что позволяет использовать ее как на малых обучающих выборках, так и на больших массивах данных, поступающих в реальном времени.

Список литературы: 1. Jang J.-S.R. Neuro-Fuzzy and Soft Computing / J.-S.R. Jang, C.-T. Sun, E. Mizutani. — Prentice Hall, Upper Saddle River, NJ. — 1997. — 640 p. 2. Friedman J. The Elements of Statistical Learning. Data Mining, Inference and Prediction / Friedman J., Hastie T., Tibshirani R. — Berlin: Springer, 2003. — 552 p. 3. Osowski S. Sieci neuronowe do przetwarzania informacji. — Warszawa: Oficyna Wydawnicza Politechniki Warszawskiej, 2006. — 422s. 4. L. Ruthowski, “Computational Intelligence. Methods and Techniques”, Berlin-Heidelberg: Springer-Verlag, 2008. — 514 p. 5. Mumford C.L., Jain L.C. Computational Intelligence. Collaboration, Fusion and Emergence. — Berlin: Springer-Verlag, 2009. — 725p. 6. Kruse R., Borgelt C., Klawonn F., Moewes C., Steinbrecher M., Held P. Computational Intelligence. Methodological introduction. — Berlin: Springer-Verlag, 2013. — 488 p. 7. Kroll A. Computational Intelligence. Eine Einführung in Probleme, Methoden und technische Anwendungen. —

Muenchen: Oldenbourg Verlag, 2013. — 428 s. 8. K.-L. Du, M.N.S. Swami, Neural Networks and Statistical Learning, London: Springer-Verlag, 2014. — 824 p. 9. Nauck D., Kruse R. NEFCLASS - a neuro-fuzzy approach for the classification of data / Eds. K.M. George e.a. Proc. of the 1995 ACM Symposium on Applied Computing, “Applied Computing 1995” Nashville, 1995, pp. 461-465. 10. Bifet A. Adaptive Stream Mining. Pattern Learning and Mining from Evolving Data Streams. — Amsterdam. IOS Press, 2010. — 224 p. 11. Rizzo R. Computational Intelligence Methods for Bioinformatics and Biostatistics / Rizzo R. - In Lecture Notes in Bioinformatics (7th International Meeting, CIBIB 2010, Palermo, Italy, September 16-18, 2010). — Springer. — 2011. — 301 p. 12. Kountchev R. Advances in Intelligent Analysis of Medical Data and Decision Support Systems (Studies in Computational Intelligence) / Kountchev R. and etc - Springer. — 2013. — 246 p. 13. Бодянский Е.В., Винокурова Е.А., Мулеца П.П., Перова И.Г. Диагностирующая нейро-фаззи-система и ее адаптивное обучение в задачах интеллектуальной обработки данных медико-биологических исследований / Системные технологии. — 2014. — Вып. 2 (91). — с. 125-135. 14. Wang L.X. Fuzzy basis functions, universal approximation, and orthogonal least squares learning / L.X. Wang, J.M. Mendel // IEEE Trans. on Neural Network. — 1992. — 3. — P. 807-814. 15. Shynk J.J. Performance surfaces of a single-layer perceptron / Shynk J.J. // IEEE Trans. on Neural Networks. — 1990. — 1. — P. 268-274. 16. Shepherd A.J. Second-Order Methods for Neural Networks / A.J. Shepherd. — London: Springer-Verlag. — 1997. — 145 p. 17. Otto P., Bodyanskiy Ye., Kolodyazhniy V. A new learning algorithm for forecasting neural-fuzzy network // Integrated Computer-Aided Engineering. — 10. — №4. — Amsterdam : IOS Press, 2003. — P. 399-409. 18. Kruschke J.K., Movellan J.R. Benefits of gain: speeded learning and minimal layers backpropagation networks // IEEE Trans. on Systems, Man, and Cybernetics. — 1991. — 21. — P. 273-280. 19. Turuta O., Perova I., Deineko A. Evolving Flexible Neuro-Fuzzy System for Medical Diagnostics Tasks // International Journal of Computer Science and Mobile Computing, Vol. 4, Issue. 8, August 2015, P. 475-480. 20. Fukushima K. Cognitron: a self-organizing multilayer neural network // Biolog. Cybernetics, vol. 20, 1975, P. 121-136. 21. Fukushima K. Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position // Biolog. Cybernetics, vol. 36, 1980, P. 193-202.

Поступила в редколлегию 14.04.2015

УДК 615.471:616-071:004.032.26

Нейро-фаззі система для задач обробки медичних даних в ситуаціях великої кількості діагнозів. /І.Г. Перова, Є.В. Бодяньський // Біоніка інтелекту: наук.-техн. журнал. — 2015. — № 2 (85). — С. 86—89.

В статті розглянуті структура і алгоритм навчання діагностуючої нейро-фаззі системи для обробки медичних даних в онлайн режимі при наявності одного або більше діагнозів.

Л. 2. Бібліогр.: 21 найм.

UDK 615.471:616-071:004.032.26

Neuro-Fuzzy system for Medical Data Mining Tasks in Condition of Many Diagnosis/ I.G. Perova, Y.V. Bodyanskiy // Bionics of Intelligence: Sci. Mag. — 2015. — № 2 (85). — P. 86—89.

In article architecture and training method for diagnostic neuro-fuzzy-system for Medical Data Mining Tasks in online-mode and in condition of various diagnosis are proposed.

Fig. 2. Ref.: 21 items.