

УДК 004.032.26

Е. В. Бодянский, проф. д.т.н.

А. А. Дейнеко, канд. техн. наук

Я. В. Куценко, асп.

ЯДЕРНАЯ САМООРГАНИЗУЮЩАЯСЯ КАРТА НА ОСНОВЕ РАДИАЛЬНО-БАЗИСНОЙ НЕЙРОННОЙ СЕТИ

***Аннотация.** В статье предложена гибридная искусственная нейронная сеть, объединяющая в себе идеи ядерных систем и самообучения и построенная на основе радиально-базисной нейронной сети и самоорганизующейся карты. Предложенная система позволяет решать задачу on-line кластеризации в условиях, когда образующие исходными данными классы имеют произвольную форму.*

***Ключевые слова:** гибридная искусственная нейронная сеть, радиально-базисная нейронная сеть, ядерная система, самоорганизующаяся карта.*

Є. В. Бодянський, проф. д.т.н.

А. О. Дейнеко, канд. техн. наук

Я. В. Куценко, асп

ЯДЕРНА САМООРГАНІЗОВНА МАПА НА ОСНОВІ РАДІАЛЬНО-БАЗИСНОЇ НЕЙРОНОЇ МЕРЕЖІ

***Анотація.** У статті запропонована гібридна штучна нейронна мережа, яка об'єднує в собі ідеї ядерних систем і самонавчання та побудована на основі радіально-базисної нейронної мережі і самоорганізованої мапи. Запропонована система дозволяє вирішувати задачу on-line кластеризації в умовах, коли утворені вихідними даними класи мають довільну форму.*

***Ключові слова:** гібридна штучна нейронна мережа, радіально-базисна нейронна мережа, ядерна система, самоорганізовна мапа.*

Ye. Bodyanskiy D.Sci., Prof.

A. Deineko Ph.D.

Ya. Kutsenko Ph.D. student

KERNEL SELF-ORGANIZING MAP BASED ON RADIAL-BASIS NEURAL NETWORK

***Abstract.** A hybrid artificial neural network is proposed in the present paper. The network combines the principle of kernel systems and self-learning, and is based on radial basis neural networks and self-organizing maps. The proposed system allows to solve the problem of on-line clustering under the conditions, when the classes formed by the initial data have an arbitrary form.*

***Keywords:** hybrid artificial neural network, radial basis neural network, kernel system, self-organizing map.*

Введение: В настоящее время самообучающиеся системы вычислительного интеллекта [1, 2] и, прежде всего, искусственные нейронные сети, настраивающие свои параметры без учителя [3, 4], получили широкое распространение при решении различных задач интеллектуального анализа данных (Data Mining, Exploratory Data Analysis). Здесь особо следует отметить нейронные сети Т. Кохонена (самоорганизующиеся карты, SOM) [5], предназначенные для решения задач кластеризации больших массивов информации, благодаря своей вычислительной простоте, эффективности и возможности ра-

боты в on-line режиме путем последовательной обработки данных по мере их поступления.

Процесс настройки этих сетей реализуется в режиме самообучения на основе принципов «Победитель получает все» (WTA) или «Победитель получает больше» (WTM), при этом априорно предполагается, что исходная структура данных такова, что формируемые кластеры взаимно не пересекаются и имеют выпуклую форму, т.е. в процессе обучения нейросети могут быть построены разделяющие гиперплоскости, четко разграничивающие разные классы.

В случае пересекающихся классов могут быть использованы методы нечеткого (fuzzy) кластерного анализа [6, 7], в том числе нечеткие самоорганизующиеся карты [8-11], реализующие в той или иной форме метод нечетких С-средних (FCM) [12].

В случае невыпуклых классов ситуация представляется более сложной. Здесь следует отметить, что сети Кохонена фактически реализуют классический метод К-средних [13, 14], а потому клетки Г. Вороного, формируемые в процессе самообучения, ограничены гиперплоскостями. Здесь следует отметить, что сети Кохонена фактически реализуют классический метод К-средних [13, 14], а потому клетки Г. Вороного, формируемые в процессе самообучения, ограничены гиперплоскостями. Естественно, что существуют методы кластеризации для классов произвольной формы [15-19], однако реализующие их алгоритмы достаточно сложны с вычислительной точки зрения и никоим образом не предназначены для решения задач Dynamic Data Mining и Data Stream Mining [20], где информация должна обрабатываться в реальном времени.

На сегодня известны, так называемые, ядерные самоорганизующиеся карты (Kernel SOM) [21- 23], построенные с использованием ядер Дж. Мерсера [24] (обычно типа потенциальных функций) [25]) и основанные

на минимизации критерия эмпирического риска [26], лежащего в основе машин опорных векторов (SVM) [27, 28].

Нейронные сети-машины опорных векторов являются весьма эффективным аппаратом для решения широкого класса задач Data Mining, однако при этом они страдают от, так называемого, «проклятия размерности», поскольку количество нейронов в SVM-нейронной сети совпадает с числом векторов наблюдений в выборке данных, подлежащих кластеризации.

В связи с этим представляется целесообразным вместо традиционного SVM-подхода в ядерных системах использовать идеи, лежащие в основе радиально-базисных нейронных сетей [29-31] и связанные с теоремой Т.Кавера [32], утверждающей, что линейно неразделимая задача распознавания образов в исходном пространстве R^n может стать линейно разделимой в пространстве более высокой размерности $R^h (n+1 \leq h \leq N)$. Свойства такой нейронной сети полностью определяются радиально-базисными функциями φ , используемыми в качестве активационных и формирующими некоторый базис для входных векторов-образов.

Архитектура ядерной самоорганизующейся карты на основе радиально-базисной нейронной сети.

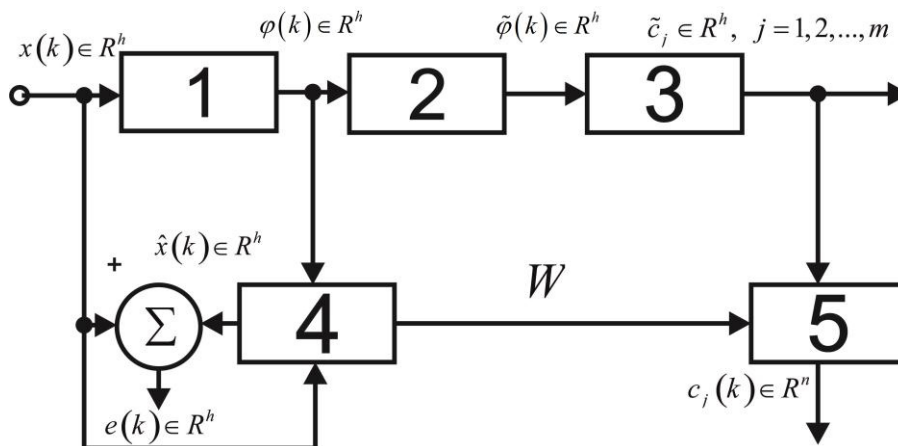


Рис. 1. Ядерная самоорганизующаяся карта на основе радиально-базисной нейронной сети

На рис. 1 приведена архитектура рассматриваемой ядерной самоорганизующейся карты на основе радиально-базисной нейронной сети.

Исходной информацией в данном случае является центрированная выборка (возможно растущая) векторов наблюдений $x(1), x(2), \dots, x(k), \dots, x(N), \dots$; $x(k) = (x_1(k), \dots, x_i(k), \dots, x_n(k))^T \in R^n$ таких, что $-1 \leq x_i(k) \leq 1$, $\frac{1}{N} \sum_{k=1}^N x_1(k) = 0$, которая должна быть разбита на m кластеров произвольной формы, при этом k здесь может быть как номером наблюдений, так и моментом текущего времени.

Векторы наблюдений $x(k)$ последовательно поступают на слой радиально-базисных функций 1, полностью совпадающий по структуре с первым скрытым слоем стандартной радиально-базисной сети и сформированный ядерными (колоколообразными, потенциальными) функциями активации $\varphi_1, \varphi_2, \dots, \varphi_1, \dots, \varphi_n$ ($n+1 \leq h \leq N$), с помощью которых производится повышение размерности исходного пространства входов. В качестве таких функций могут быть использованы традиционные гауссианы

$$\varphi_l = e^{-\frac{\|x-c_l\|^2}{2\sigma_l^2}}$$

(хотя, конечно могут применяться и другие колоколообразные функции), а в качестве их центров c_l в простейшем случае могут быть взяты h произвольно выбранных векторов наблюдений $c_l = x(l)$ (концепция «нейроны в точках данных» [33]). Таким образом, при подаче на вход системы вектора наблюдений $x(k)$ на выходе первого слоя формируется векторный сигнал $\varphi(k) = (\varphi_1(k), \dots, \varphi_l(k), \dots, \varphi_n(k))^T \in R^h$, где

$$\varphi_l(k) = e^{-\frac{\|x(k)-c_l\|^2}{2\sigma_l^2}}$$

Второй слой системы – слой нормализации 2 реализует элементарное преобразование

$$\tilde{\varphi}_l(k) = \frac{\varphi_l(k)}{\|\varphi_l(k)\|},$$

необходимое для эффективной работы третьего слоя – самоорганизующейся карты 3.

Именно в этом слое и решается задача кластеризации, т.е. разбиения последовательности образов $\tilde{\varphi}(1), \dots, \tilde{\varphi}(k), \dots$, на m кластеров с нахождением в процессе самообучения прототипов – центроидов классов $\tilde{c}_j^K \in R^h, j=1, 2, \dots, m$.

Основная проблема состоит в том, чтобы эффективно сформировать базис, образованный радиально-базисными функциями, в котором можно было бы произвести кластеризацию с помощью карты Кохонена. Для этого, прежде всего, необходимо оценить насколько удачно было выбрано количество h и центры c_l радиально-базисных функций первого слоя.

Для этого предназначен слой восстановления входного пространства 4, представляющий собой по сути выходной слой радиально-базисной сети с h входами и n выходами и содержащий nh настраиваемых синоптических весов и n сумматоров.

Таким образом, первый и четвертый слои образуют многовыходную радиально-базисную нейронную сеть, отличием которой от стандартной является то, что в качестве обучающего здесь используется входной сигнал, т.е. сеть работает в режиме ассоциации [3]. На выходе четвертого слоя формируется сигнал $\hat{x}(k) \in R^n$, являющийся оценкой входного сигнала $x(k)$. Качество восстановления оценивается на основе векторной ошибки

$$e(k) = x(k) - \hat{x}(k)$$

с помощью скалярного критерия

$$\bar{e} = \frac{1}{N} \sum_{k=1}^N \frac{\|x(k) - \hat{x}(k)\|}{\|x(k)\|}. \quad (1)$$

Если окажется, что значение \bar{e} превышает некоторый заданный порог e_T , принимается решение о том, что количество нейронов в первом слое должно быть увеличено. Этот процесс продолжается до обеспечения требуемого качества восстановления входного пространства. Результатом обучения четвертого слоя является $(n \times h)$ - матрица синаптических весов $W(N)$, полученная на основании N наблюдений.

Эта матрица является исходной информацией для пятого слоя восстановления прототипов кластеров в исходном пространстве R^n . При этом прототипы, сформированные самоорганизующейся картой в h -мерном пространстве, проецируются в исходное n -мерное пространство с помощью элементарного преобразования

$$c_j^K(k) = W(N)\tilde{c}_j^K(k) \quad \forall j = 1, 2, \dots, m.$$

Таким образом, рассматриваемая здесь система является по сути объединением двух нейросетей: эволюционирующей радиально-базисной нейронной сети (ERBFN) и самоорганизующейся карты Кохонена (SOM), которые параллельно настраивают свои синаптические веса в режиме самообучения, одновременно с этим решая задачу кластеризации данных, образующих классы произвольной формы.

Обучение ядерной самоорганизующейся карты на основе радиально-базисной нейронной сети. Обучение введенной системы может рассматриваться как две относительно независимые задачи: самообучение радиально-базисной подсистемы и самообучение собственно самоорганизующейся карты.

Задача слоя восстановления входного пространства состоит в нахождении $(n \times h)$ -матрицы синаптических весов $W = \{w_{ij}\}$ по выборке, содержащей N наблюдений путем минимизации критерия обучения

$$E^N = \sum_{k=1}^N E(k) = \frac{1}{2} \sum_{k=1}^N \|x(k) - W\varphi(k)\|^2 = \frac{1}{2} \sum_{k=1}^N \|e(k)\|^2. \quad (2)$$

Минимизация критерия (2) в пакетном варианте ведет к оценке наименьших квадратов вида

$$W(N) = \left(\sum_{k=1}^N x(k)\varphi^T(k) \right) \left(\sum_{k=1}^N \varphi(k)\varphi^T(k) \right)^{-1}$$

либо в рекуррентной форме [34] –

$$\begin{cases} W(k) = W(k-1) + \\ + \frac{(x(k) - W(k-1)\varphi(k))\varphi^T(k)P(k-1)}{1 + \varphi^T(k)P(k-1)\varphi(k)}, \\ P(k) = P(k-1) - \\ - \frac{P(k-1)\varphi(k)\varphi^T(k)P(k-1)}{1 + \varphi^T(k)P(k-1)\varphi(k)}. \end{cases} \quad (3)$$

Улучшить качество восстановления входного пространства можно, настраивая не только синаптические веса четвертого слоя, но и параметры центров c_i и ширины σ_i активационных функций первого слоя.

С учетом очевидных соотношений

$$\begin{cases} \hat{x}(k) = \sum_{i=1}^h w_{ii}(k-1)\varphi_i(k), \\ e_i^2(k) = (x_i(k) - \sum_{i=1}^h w_{ii}(k-1)\varphi_i(k))^2, \\ E(k) = \frac{1}{2} \sum_{i=1}^n e_i^2(k), \\ \nabla_{c_i} E(k) = e_i(k) w_{ii}(k-1) \varphi' \left(\frac{\|x(k) - c_i\|^2}{2\sigma_i^2} \right) \frac{x(k) - c_i}{\sigma_i^2}, \\ \frac{\partial E(k)}{\partial \sigma_i^{-2}} = -e_i(k) w_{ii}(k-1) \varphi' \left(\frac{\|x(k) - c_i\|^2}{2\sigma_i^2} \right) \frac{\|x(k) - c_i\|^2}{2}, \end{cases}$$

можно ввести рекуррентные градиентные алгоритмы обучения вида:

$$\left\{ \begin{array}{l} c_i(k) = c_i(k-1) - \eta_c(k) e_i(k) w_{ii}(k-1) \times \\ \times e^{-\frac{\|x(k) - c_i(k-1)\|^2}{2\sigma_i^2(k-1)}} \frac{x(k) - c_i(k-1)}{\sigma_i^2(k-1)}, \\ \sigma_i^{-2}(k) = \sigma_i^{-2}(k-1) + \eta_\sigma(k) e_i(k) w_{ii}(k-1) \times \\ \times e^{-\frac{\|x(k) - c_i(k-1)\|^2}{2\sigma_i^2(k-1)}} \frac{\|x(k) - c_i(k-1)\|^2}{2}, \end{array} \right. \quad (6)$$

где $\eta_c(k), \eta_\sigma(k)$ – скалярные параметры шага обучения.

Таким образом, соотношения (3), (4) есть алгоритм обучения всех параметров радиально-базисной нейронной сети, при этом процесс обучения (включая изменение числа радиально-базисных функций) продолжается до достижения требуемого значения критерия (1).

Обучение слоя 3 – собственно карты Кохонена состоит в разбиении последовательности векторов - образов $\tilde{\varphi}(k)$ на m кластеров, каждый из которых характеризуется собственным прототипом-центроидом $\tilde{c}_j^K(k) \in R^h, j=1,2,\dots,m$, непрерывно уточняемых при подаче очередного наблюдения $\tilde{\varphi}(k)$. Число же кластеров m полагается априорно заданным.

Как и любая другая процедура самообучения, процесс настройки начинается с инициализации синаптических весов сети, в качестве которых и выступают начальные значения прототипов $\tilde{c}_j^K(0)$. При этом эти значения в процессе обработки нормируются подобно входным образам, т.е.

$$\|\tilde{c}_j^K(k)\| = 1.$$

При подаче на вход третьего слоя сигнала $\tilde{\varphi}(k)$ в начале вычисляется m расстояний

$$D(\tilde{\varphi}(k), \tilde{c}_j^K(k-1)) = \|\tilde{\varphi}(k) - \tilde{c}_j^K(k-1)\| \quad \forall j = 1, 2, \dots, m, \quad (5)$$

при этом если в качестве расстояний используется евклидова метрика, то вместо (5) гораздо удобнее использовать меру подобия

$$SM(\tilde{\varphi}(k), \tilde{c}_j^K(k-1)) = \tilde{\varphi}^T(k) \tilde{c}_j^K(k-1) = \cos(\tilde{\varphi}(k), \tilde{c}_j^K(k-1)) = \cos \theta_j(k).$$

На основании (5) или (6) определяется нейрон-победитель, «ближайший» ко входному образу такой, что

$$D(\tilde{\varphi}(k), \tilde{c}_*^K(k-1)) = \min_j D(\tilde{\varphi}(k), \tilde{c}_j^K(k-1))$$

или

$$SM(\tilde{\varphi}(k), \tilde{c}_*^K(k-1)) = \max_j SM(\tilde{\varphi}(k), \tilde{c}_j^K(k-1)).$$

Далее настраиваются веса этого нейрона с помощью правила самообучения Кохонена [5] в форме

$$\tilde{c}_j^K(k) = \begin{cases} \frac{\tilde{c}_j^K(k-1) + \eta(k)(\tilde{\varphi}(k) - \tilde{c}_j^K(k-1))}{\|\tilde{c}_j^K(k-1) + \eta(k)(\tilde{\varphi}(k) - \tilde{c}_j^K(k-1))\|}, & \text{если } j\text{-й нейрон победил,} \\ \tilde{c}_j^K(k-1) & \text{в противном случае.} \end{cases} \quad (7)$$

Наличие знаменателя в первом соотношении (7) автоматически обеспечивает протекание процесса обучения на единичном гипершаре.

Процедура (7) реализует принцип «Победитель получает все» (WTA), при этом вектор синаптических весов нейрона-победителя $\tilde{c}_*^K(k-1)$ «подтягивается» ко входному образу $\tilde{\varphi}(k)$ на расстояние, определяемое шагом поиска $0 < \eta(k) < 1$. Регулирование шага поиска обычно производится, исходя из эмпирических соображений, а общая рекомендация сводится к тому, что шаг должен монотонно уменьшаться в процессе самообучения. Это условие может быть обеспечено при выборе шага согласно соотношению [35]

$$\eta(k) = r^{-1}(k), \\ r(k) = \alpha r(k-1) + \|\tilde{\varphi}(k)\|^2 = \alpha r(k-1) + 1, \quad 0 < \alpha \leq 1,$$

при этом при $\alpha = 1$ параметр шага $\eta(k) = \frac{1}{k}$ удовлетворяет условиям А. Дворецкого, а получаемые результаты совпадают с оценками К-средних.

Варьируя фактором забывания α , можно обеспечить широкий интервал изменения шага

$$\frac{1}{k} \leq \eta(k) \leq 1.$$

В результате предъявления самоорганизующейся карте N образов $\tilde{\varphi}(k)$ будет получено m прототипов-центроидов $\tilde{c}_j^k(N)$, которые далее из пространства повышенной размерности R^h могут быть спроецированы в исходное пространство R^n с помощью простого соотношения

$$c_j^k(N) = W(N)\tilde{c}_j^k(N) \quad \forall j = 1, 2, \dots, m.$$

Заключение. В статье предложена гибридная искусственная нейронная сеть, объединяющая в себе идеи ядерных систем и самообучения и построенная на основе эволюционирующей радиально-базисной нейронной сети и самоорганизующейся карты. Предложенная система позволяет решать задачу on-line кластеризации в условиях, когда образуемые исходными данными классы имеют произвольную форму. Введенная нейронная сеть проста в реализации и позволяет решать достаточно широкий класс задач динамического интеллектуального анализа данных (DDM) и интеллектуального анализа потоков данных (DSM).

Список использованной литературы

1. Rutkowski, L. (2008) Computational Intelligence. Methods and Techniques. Springer-Verlag. Berlin. Heidelberg. p. 514.
2. Kruse, R., Borgelt C., Klawonn F., Moewes C., Steinbrecher M., Held P. (2013) Computational Intelligence. A Methodological Introduction. Springer. Berlin. p. 488.

3. Haykin, S. (1999) Neural Networks. A Comprehensive Foundation. Upper Saddle River, N.J.: Prentice Hall, Inc. p. 842.
4. Du, K. - L., M. N. S. Swamy (2014) Neural Networks and Statistical Learning. Springer - Verlag. London. p. 824.
5. Kohonen T. (1995) Self-Organizing Maps. Springer-Verlag. Berlin. p.362.
6. Höppner F. Klawonn F., Kruse R. (1996) Fuzzy-Clusteranalyse. Verfahren für die Bilderkennung, Klassifikation und Datenanalyse. Vieweg. Braunschweig. s. 280.
7. Höppner, F., Klawonn F., Kruse R., Runkler T. (1999) Fuzzy Cluster Analysis: Methods for Classification, Data Analysis and Image Recognition. John Wiley & Sons. Chichester. p.289.
8. Tsao, E.C.-K., Bezdek J.C, Tsao, J.C., Pal N.R. (1994) Fuzzy Kohonen clustering networks. Pattern Recognition. 27. pp.757–764.
9. Pascual – Marqui R. D., Pascual – Montano A.D., Kochi K., Caroso J.M. (2001) Smoothly distributed fuzzy C-means: a new self-organizing map. Pattern Recognition. 34. pp. 2395–2402.
10. Bodyanskiy Ye., Kolodyazhniy V., Stephan A. (2002) Recursive fuzzy clustering algorithms. Proc. East West Fuzzy Coll. Zittau-Görlitz: HS. pp. 164-172.
11. Дейнеко А.О., Дейнеко Ж.В., Дейнеко О.П., Турута О.П., Бодянский Е.В. (2014) Метод комбінованого навчання-самонавчання нейро-фаззі систем. Системні технології. Регіональний міжвузівський збірник наукових праць. № 2(91). сс. 145 – 153.
12. Bezdek, J. - C. (1981) Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum Press. N.Y. p. 272.
13. Kaufman L., Rousseeuw P. (1989) Finding Groups in Data. John Wiley & Sons. N.Y. p.320.
14. Xu, R., Wunsch D.C. (2009) Clustering. IEEE Press Series on Computational Intelligence. Hoboken, NJ:John Wiley & Sons, Inc. p. 370.
15. Gan G., Ma Ch, Wu J (2007) Data Clustering: Theory, Algorithms and Applications. Philadelphia, Pennsylvania: SIAM. p. 455.
16. Jain A. K., Dubes R. C. (1988) Algorithms for Clustering Data. Englewood Cliffs, N.J.: Prentice-Hall. p. 318.

17. Han, J., M. Kamber (2006) *Data Mining: Concepts and Techniques*. Amsterdam: Morgan Kaufmann Publ. p. 754.

18. Abonyi, J., Feil B. (2007) *Cluster Analysis for Data Mining and System Identification*. Basel: Birkhauser. p. 303.

19. Olson, D. L., Dursun D. (2008) *Advanced Data Mining Techniques*. Berlin: Springer. p. 180.

20. Bifet, A. (2010) *Adaptive Stream Mining: Pattern Learning and Mining from Evolving Data Streams*. IOS Press. p. 224.

21. MacDonald D., Fyfe C. (2002) Clustering in data space and feature space. *ESANN'2002 Proc. European Symp. on Artificial Neural Networks*. Bruges (Belgium), 24-26 April. pp. 137-142.

22. Girolami, M. (2002) Mercer kernel-based clustering in feature space. *IEEE Trans. on Neural Networks*. 13. №3. pp. 780-784.

23. Camastra, F., Verri A (2005) A novel kernel method for clustering. *IEEE Trans. on Pattern Analysis and Machine Intelligence*. No. 5. pp. 801-805.

24. Schölkopf B., Smola A. (2002) *Learning with Kernels*. Cambridge, M.A.: MIT Press. – 2002.

25. Айзерман, М. А., Браверман Э. М., Розоноэр Л. И. (1970) Метод потенциальных функций в теории обучения машин. М.: Наука. с. 384.

26. Вапник В.Н., Червоненкис А.Я. (1974) *Теория распознавания образов (статистические проблемы обучения)*. М.: Наука. с. 416.

27. Cortes, C., Vapnik V. (1995) Support Vector Networks. *Machine Learning*. 20. pp. 273–297.

28. Suykens J.A.K., Gestel, T.V., Brabanter J.D., Moor B.D., Vandewalle J. (2002) *Least Squares Support Vector Machines*. Singapore: World Scientific. p. 294.

29. Moody J., Darken C.J. (1989) Fast learning in networks of locally-tuned processing units. *Neural Computation*. pp.281-294.

30. Park J., Sandberg I. W. (1991) Universal approximation using radial-basis-function networks. *Neural Computation*. 3. pp.246-257.

31. Poggio T., Girosi F., Memo A.I. (1994) A Theory of Networks for Approximation and Learning. № 1140, C.B.I.P. Paper № 31. – Massachusetts Institute of Technology. p. 63.

32. Cover T. M. (1965) Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition. *IEEE Trans. on Electronic Computers*. 14. pp. 326-334.

33. Zahirniak D., Chapman R., Rogers S., Suter B., Kabrisky M., Piaty V. (1990) Pattern recognition using radial basis function network. *Proc 6th Ann. Aerospace Application of Artificial Intelligence Conf.* – Dayton, OH. pp. 249-260.

34. Ljung, L. (1999) *System Identification: Theory for the Users*. N.Y.: Prentice-Hall.p. 519.

35. Бодянский Е.В., Плисс И.П., Соловьева Т.В. (1989) Адаптивное обобщенное прогнозирование многомерных случайных последовательностей. *Доклады АН УССР. А. №9*. с. 73-75.

Получено 6.10.2015.

References

1. Rutkowski, L. (2008) *Computational Intelligence. Methods and Techniques*. Springer-Verlag. Berlin. Heidelberg. p. 514.

2. Kruse, R., Borgelt C., Klawonn F., Moewes C., Steinbrecher M., Held P. (2013) *Computational Intelligence. A Methodological Introduction*. Springer. Berlin. p. 488.

3. Haykin, S. (1999) *Neural Networks. A Comprehensive Foundation*. Upper Saddle River, N.J.: Prentice Hall, Inc. p. 842.

4. Du, K. - L., M. N. S. Swamy (2014) *Neural Networks and Statistical Learning*. Springer - Verlag. London. p. 824.

5. Kohonen T. (1995) *Self-Organizing Maps*. Springer-Verlag. Berlin. p.362.

6. Höppner F. Klawonn F., Kruse R. (1996) *Fuzzy-Clusteranalyse. Verfahren für die Bilderkennung, Klassifikation und Datenanalyse*. Vieweg. Braunschweig. s. 280.

7. Höppner, F., Klawonn F., Kruse R., Runkler T. (1999) *Fuzzy Cluster Analysis: Methods for Classification, Data Analysis and Image Recognition*. John Wiley & Sons. Chichester. p.289.

8. Tsao, E.C.-K., Bezdek J.C, Tsao, J.C., Pal N.R. (1994) Fuzzy Kohonen clustering networks. *Pattern Recognition*. 27. pp.757–764.

9. Pascual – Marqui R. D., Pascual – Montano A.D., Kochi K., Caroso J.M. (2001)

Smoothly distributed fuzzy C-means: a new self-organizing map. *Pattern Recognition*. 34. pp. 2395–2402.

10. Bodyanskiy Ye., Kolodyazhnyi V., Stephan A. (2002) Recursive fuzzy clustering algorithms. *Proc. East West Fuzzy Coll. Zittau-Görlitz: HS*. pp. 164-172.

11. Dejneko A.O., Dejneko Zh.V., Dejneko O.P., Turuta O.P., Bodjanskij E.V. Metod kombinovanogo navchannja-samonavchannja nejro-fazzi sistem. *Sistemni tehnologii. [Adaptive method of combining learning-selflearning for neuro-fuzzy systems]* (2014) *Regional'nij mizhvuzivs'kij zbirnik naukovih prac'*. № 2(91). pp. 145 – 153. (In Ukrainian)

12. Bezdek, J. - C. (1981) *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press. N.Y. p. 272.

13. Kaufman L., Rousseeuw P. (1989) *Finding Groups in Data*. John Wiley & Sons. N.Y. p.320.

14. Xu, R., Wunsch D.C. (2009) *Clustering*. IEEE Press Series on Computational Intelligence. Hoboken, NJ: John Wiley & Sons, Inc. p. 370.

15. Gan G., Ma Ch, Wu J (2007) *Data Clustering: Theory, Algorithms and Applications*. Philadelphia, Pennsylvania: SIAM. p. 455.

16. Jain A. K., Dubes R. C. (1988) *Algorithms for Clustering Data*. Englewood Cliffs, N.J.: Prentice-Hall. p. 318.

17. Han, J., M. Kamber (2006) *Data Mining: Concepts and Techniques*. Amsterdam: Morgan Kaufmann Publ. p. 754.

18. Abonyi, J., Feil B. (2007) *Cluster Analysis for Data Mining and System Identification*. Basel: Birkhauser. p. 303.

19. Olson, D. L., Dursun D. (2008) *Advanced Data Mining Techniques*. Berlin: Springer. p. 180.

20. Bifet, A. (2010) *Adaptive Stream Mining: Pattern Learning and Mining from Evolving Data Streams*. IOS Press. p. 224.

21. MacDonald D., Fyfe C. (2002) Clustering in data space and feature space. *ESANN'2002 Proc. European Symp. on Artificial Neural Networks*. Bruges (Belgium), 24-26 April. pp. 137-142.

22. Girolami, M. (2002) Mercer kernel-based clustering in feature space. *IEEE Trans. on Neural Networks*. 13. №3. pp. 780-784.

23. Camastra, F., Verri A (2005) A novel kernel method for clustering. *IEEE Trans. on Pattern Analysis and Machine Intelligence*. No. 5. pp. 801-805.

24. Schölkopf B., Smola A. (2002) *Learning with Kernels*. Cambridge, M.A.: MIT Press. – 2002.

25. Ajzerman, M. A., Braverman Je. M., Rozonojer L. I. Metod potencial'nyh funkcij v teorii obuchenija mashin. [Theoretical foundation of potential functions method in pattern recognition learning] (1970) *M.: Nauka*. p. 384. (In Russian).

26. Vapnik V.N., Chervonenkis A.Ja. Teorija raspoznavanija obrazov (statisticheskie problemy obuchenija). [The Nature of Statistical Learning Theory.] (1974) *M.: Nauka*. p. 416. (In Russian).

27. Cortes, C., Vapnik V. (1995) Support Vector Networks. *Machine Learning*. 20. pp. 273–297.

28. Suykens J.A.K., Gestel, T.V., Brabanter J.D., Moor B.D., Vandewalle J. (2002) *Least Squares Support Vector Machines*. Singapore: World Scientific. p. 294.

29. Moody J., Darken C.J. (1989) Fast learning in networks of locally-tuned processing units. *Neural Computation*. pp.281-294.

30. Park J., Sandberg I. W. (1991) Universal approximation using radial-basis-function networks. *Neural Computation*. 3. pp.246-257.

31. Poggio T., Girosi F., Memo A.I. (1994) A Theory of Networks for Approximation and Learning. № 1140, C.B.I.P. Paper № 31. – Massachusetts Institute of Technology. p. 63.

32. Cover T. M. (1965) Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition. *IEEE Trans. on Electronic Computers*. 14. pp. 326-334.

33. Zahiriak D., Chapman R., Rogers S., Suter B., Kabrisky M., Piaty V. (1990) Pattern recognition using radial basis function network. *Proc 6th Ann. Aerospace Application of Artificial Intelligence Conf.* – Dayton, OH. pp. 249-260.

34. Ljung, L. (1999) System Identification: Theory for the Users. N.Y.: Prentice-Hall.p. 519.
35. Bodjanskij E.V., Pliss I.P., Solov'eva T.V. Adaptivnoe obobshhennoe prognozirovanie mnogomernyh sluchajnyh posledovatel'nostej. [Adaptive generalized forecasting of multivariable random sequences] (1989), *Doklady AN USSR. A.* №9. pp. 73-75. (In Russian).



Бодянский Евгений Владимирович
Руководитель ПНИЛ АСУ, д.т.н., проф., профессор кафедры искусственного интеллекта, Харьковский Национальный Университет Радиоэлектроники
Харьков, пр. Ленина, 14, ауд. 288
bodya@kture.kharkov.ua



Дейнеко Анастасия Александровна
Младший научный сотрудник, ПНИЛ АСУ, к.т.н.
Харьковский Национальный Университет Радиоэлектроники
Харьков, пр. Ленина, 14, ауд. 511
(050)753-60-50
Anastasiya.deineko@gmail.com



Куценко Яна Владимировна
Аспирантка кафедры искусственного интеллекта, Харьковский Национальный Университет Радиоэлектроники
Харьков, пр. Ленина, 14, ауд. 288
kutcenochka@gmail.com