

ТРЕХСЛОЙНЫЙ ПЕРЦЕПТРОН, СПОСОБНЫЙ ДООБУЧАТЬСЯ

В.Д. Дмитриенко¹, А.Ю. Заковоротный², В.А. Бречко³

^{1,2,3} *Национальный технический университет "Харьковский политехнический институт", г. Харьков, Украина*

THREE-LAYERED PERCEPTRON THAT IS ABLE TO LEARN

V.D. Dmitrienko¹, A.Y. Zakovorotnyi², V.A. Brechko³

^{1,2,3} *National Technical University "KhPI", Kharkov, Ukrainian
arcade@i.ua*

Впервые предложена нейросетевая архитектура и алгоритмы функционирования трехслойного перцептрона, способного дообучаться без потерь запомненной ранее информации. Новая нейронная сеть может стать альтернативой дискретным сетям адаптивной резонансной теории.

Ключевые слова: *перцептрон, способный дообучаться, нейронная сеть, сети адаптивной резонансной теории.*

Neural network architecture and algorithms of a three-layered perceptron that is able to learn without loss of previously stored information were first proposed. The new neural network can become an alternative to discrete networks of adaptive resonance theory.

Keywords: *perceptron that is able to learn, neural network, the network of adaptive resonance theory.*

Постановка проблемы и анализ литературы

Системы распознавания, классификации и управления на основе искусственных нейронных сетей широко применяются для решения самых различных задач. Успешное применение нейронных сетей при решении разнообразных задач обусловлено их универсальностью, возможностью замены различных традиционных проблем, возникающих при решении конкретных прикладных задач, обучением нейронных сетей на примерах, возможностью самообучения и разработки нейронных сетей под решаемую задачу, и т.д. [1 – 7].

При наличии достаточно полной информации об объекте исследования, возможно использование целого ряда различных видов нейронных сетей, подходящих для создания распознающих, классифицирующих и управляющих систем. Однако при разработке систем для реальных объектов зачастую возникают ситуации, когда информация об объекте далеко не полная или будет уточняться в процессе функционирования объекта. Это обстоятельство резко сокращает круг нейрон-

ных сетей, которые целесообразно использовать в таких ситуациях, поскольку обучение новому образу в общем случае приведет к полному переобучению нейронной сети. Также ограничение на применяемые сети накладывает и нечувствительность сетей к новым данным во входной информации. Все вышеперечисленное привело к разработке принципиально новых конфигураций нейронных сетей, использующих в своей основе адаптивную резонансную теорию [8 – 12] и решающих проблему стабильности – пластичности, то есть проблему восприятия и запоминания новой информации при сохранении уже имеющейся.

Нейронные сети адаптивной резонансной теории относят входное изображение к одному из известных классов изображений, если оно в достаточной степени похоже на прототип этого класса. Степень сходства с прототипом класса при этом определяется по заданному параметру сходства. Если найденный прототип соответствует входному изображению с заданной точностью, то он модифицируется, чтобы стать более похожим на предъявленное изображение. Если входное изображение сети адаптивной резонансной теории (АРТ) не похоже в достаточной степени ни на одно из изображений, хранящихся в памяти нейронной сети (в весах связей), то на его основе создается новый класс. Это возможно благодаря наличию в сети избыточных нейронов, которые не используются до тех пор, пока в этом нет необходимости (если избыточных нейронов уже нет и входное изображение не относится ни к одному из известных классов, то оно не вызывает реакции сети). Таким образом, нейронные сети АРТ могут запоминать новую информацию без искажения запомненной ранее информации или переобучения сети.

Указанное достоинство сетей АРТ достигается за счет введения в архитектуру нейронных сетей избыточных нейронов, которые используются только для дообучения сети. Введение избыточных нейронов может позволить решить проблему дообучения или проблему стабильности – пластичности и других нейронных сетей.

Целью статьи является разработка стабильно-пластичных нейронных сетей на основе перцептрона.

Основной раздел

Стабильность дискретных нейронных сетей АРТ-1 [10] обеспечивается наличием дополнительных нейронов, на которые и записывается новая информация. Подобные дополнительные нейроны в скрытом A -слое и выходном Y -слое нейронов может иметь и трехслойный перцептрон (*рис. 1*). Будем использовать эту сеть для запоминания и распознавания только k различных классов изображений, каждый из которых задается сигналом близким к единице на выходе одного Y -нейрона и сигналами близкими к -1 на выходах других Y -нейронов (при би-

полярном кодировании информации). При подаче на входы X -нейронов изображения, относящегося к одному из классов изображений, хранящихся в памяти сети, суммарный выходной сигнал нейронов будет близок к $(k-1) \cdot (-1) + 1 = -(k-2)$. Если входное изображение не относится ни к одному из классов, хранящихся в памяти нейронной сети, то суммарный выходной сигнал будет отличаться от $-(k-2)$.

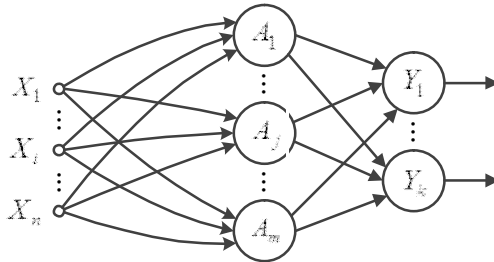


Рис. 1. Трехслойный перцептрон

Архитектуру новой сети условно можно разбить на $(N+1)$ однотипных блоков (рис. 2), каждый из которых хранит в весах своих связей множество изображений $M_l = \{S^{l1}, S^{l2}, \dots, S^{ln_l}\}$, $l = 0, 1, \dots, N$.

Любой l -й ($l = 0, 1, \dots, N$) блок содержит k_l A -нейронов скрытого слоя: $A_1^l, A_2^l, \dots, A_{k_l}^l$, n_l Y -нейронов распознающего или выходного слоя: $Y_1^l, Y_2^l, \dots, Y_{n_l}^l$, и столько же D -нейронов ($D_1^l, D_2^l, \dots, D_{n_l}^l$). D -нейроны имеют функцию активации вида

$$U_{\text{вых } D} = \begin{cases} 1, & \text{если } (1-\delta) \leq U_{\text{вх } D} < 1, \\ 0, & \text{если } (-1+\delta) < U_{\text{вх } D} < 1-\delta, \\ -1, & \text{если } -1 < U_{\text{вх } D} \leq -1+\delta, \end{cases} \quad (1)$$

где $U_{\text{вых } D}$, $U_{\text{вх } D}$ – соответственно выходной и входной сигнал D -нейрона; δ – максимально допустимая ошибка приближения сигналов "1" и "-1" с помощью нейронов с биполярной сигмоидальной функцией активации.

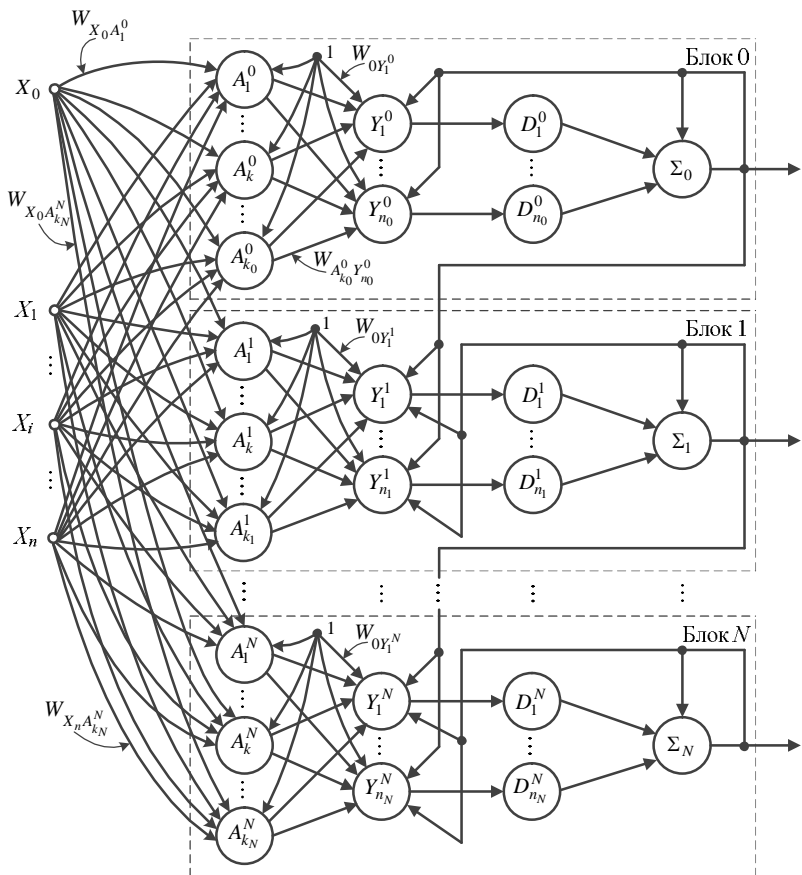


Рис. 2. Архитектура трехслойного перцептрона, который может дообучаться N раз

Каждый блок нейронной сети содержит также управляющий Σ -нейрон, с помощью которого определяется принадлежность входного вектора к соответствующему множеству изображений. Входы A -нейронов всех $(N + 1)$ блоков сети соединены с выходами сенсорных нейронов X_1, X_2, \dots, X_n .

Нейроны $A_1^0, \dots, A_{k_0}^0, Y_1^0, \dots, Y_{n_0}^0$ нулевого блока используются для запоминания исходной информации в виде множества изображений M_0 . Блок 1 и последующие блоки нейронной сети используются для обучения перцептрона по мере накопления новой информации и последовательного запоминания множеств изображений M_1, M_2, \dots, M_N . В каждом блоке выходные сигналы Y -нейронов поступают на входы

D -нейронов, имеющих функцию активации вида (1) и преобразующих непрерывные сигналы Y -нейронов в дискретные сигналы: 1, 0 и -1 .

В каждом блоке нейронов сигналы D -нейронов поступают на входы управляющего нейрона, имеющего функцию активации вида

$$U_{\text{вых } \Sigma_d} = \begin{cases} 0, & \text{если } U_{\text{вх } \Sigma_d} = \sum_{k=1}^{n_d} U_{\text{вых } D_k^d} = -(n_d - 2), \\ 1, & \text{если } U_{\text{вх } \Sigma_d} \neq -(n_d - 2), \quad d = 0, 1, \dots, N, \end{cases} \quad (2)$$

где $U_{\text{вых } \Sigma_d}$, $U_{\text{вх } \Sigma_d}$ – соответственно выходной и входной сигнал нейрона Σ_d , $d = 0, 1, \dots, N$; $U_{\text{вых } D_k^d}$ – выходной сигнал нейрона D_k^d ($d = 0, 1, \dots, N$; $k = 1, \dots, n_d$); n_d – число Y и D нейронов в блоке d ($d = 0, 1, \dots, N$).

Если $U_{\text{вых } \Sigma_d} = 0$, то это является признаком, что входное изображение принадлежит множеству M_d , а его конкретный класс определяется комбинацией выходных сигналов D -нейронов. Если $U_{\text{вых } \Sigma_d} = 1$, то это означает, что входное изображение не принадлежит множеству M_d . В этом случае выходной сигнал $U_{\text{вых } \Sigma_d}$ по цепи обратной связи фиксирует единичный выходной сигнал нейрона Σ_d , блокирует нейроны $Y_1^d, \dots, Y_{n_d}^d$ и подключает к распознаванию нейроны группы $(d + 1)$.

Рассмотрим алгоритмы функционирования нейронной сети.

Алгоритм обучения перцептрона, способного дообучаться

Шаг 1. Задаются:

- множество M_0 из n_0 обучающих пар (входное изображение – заданный выходной вектор сети);
- максимальное число N возможных дообучений нейронной сети;
- начальные веса связей нейронной сети;
- нулевые выходные сигналы нейронов блока 0 нейронной сети;
- начальное значение параметра l , с помощью которого подсчитывается число дообучений сети, $l = 0$;
- начальное значение параметра r , с помощью которого подсчитывается число нераспознанных входных изображений;

Блокируется работа Y -нейронов всех блоков сети кроме нулевого.

Шаг 2. Одним из алгоритмов метода обратного распространения ошибки блок 0 обучается распознаванию всех изображений из множества M_0 . Увеличивается значение параметра l на единицу.

Шаг 3. Нейронная сеть переводится в режим распознавания или классификации входной информации, задается начальное значение параметра $r = 0$. Формируется множество $M_{l+1} = \{S_1^{(l+1)}, \dots, S_{n_{(l+1)}}^{(l+1)}\}$ входных изображений, информация о которых отсутствует в памяти нейронной сети и которые являются представителями новых классов изображений. При записи в множество M_{l+1} нового изображения $S_k^{(l+1)}$ параметр r увеличивается на единицу и проверяется условие необходимости дообучения нейронной сети (например, по величине параметра r и числу дообучений нейронной сети). Если условие не выполняется, то сеть продолжает функционировать в заданном режиме, иначе – переход на следующий шаг алгоритма.

Шаг 4. Одним из алгоритмов метода обратного распространения ошибки блок $(l+1)$ обучается распознаванию всех изображений из множества M_{l+1} . Число l увеличивается на единицу: $l = l+1$. Проверяется условие окончания работы алгоритма, если оно не выполняется, то переход на шаг 3 алгоритма, иначе – переход на последний шаг алгоритма.

Шаг 5. Останов.

Возможна архитектура сети и алгоритм обучения, когда число групп нейронов и число нейронов в каждой группе не задается заранее, а формируется в процессе обучения.

Алгоритм функционирования перцептрона, способного дообучаться в режиме распознавания

Шаг 1. Иницируются веса связей и обнуляются выходные сигналы всех нейронов сети. Задается число L ($1 < L < N$) дообучений нейронной сети. Задается число l ($l = 0$), с помощью которого подсчитывается число блоков, которые могут быть использованы для распознавания входного изображения. Блокируется работа Y -нейронов всех блоков кроме нулевого.

Шаг 2. На вход сети подается некоторое изображение $S^* = (S_1^*, S_2^*, \dots, S_n^*)$.

Шаг 3. Определяются входные и выходные сигналы нейронов X -слоя:

$$U_{\text{вх}X_i} = S_i^*, \quad U_{\text{вых}X_i} = U_{\text{вх}X_i}, \quad i = 1, \dots, n.$$

Шаг 4. Определяются входные и выходные сигналы нейронов скрытого слоя нулевого блока:

$$U_{\text{вх } A_k^0} = 1 \cdot W_{X_0 A_k^0} + \sum_{i=1}^n U_{\text{вых } X_i} W_{X_i A_k^0}, \quad k = 1, \dots, k_0;$$

$$U_{\text{вых } A_k^0} = f_{A_k^0}(U_{\text{вх } A_k^0}), \quad k = 1, \dots, k_0,$$

где $U_{\text{вх } A_k^0}$, $U_{\text{вых } A_k^0}$ ($k = 1, \dots, k_0$) – соответственно входные и выходные сигналы нейронов скрытого A^0 -слоя; $W_{X_0 A_k^0}$ ($k = 1, \dots, k_0$) – веса связей для сигнала смещения на нейроны $A_1^0, \dots, A_{k_0}^0$; $W_{X_i A_k^0}$ ($i = 1, \dots, n$; $k = 1, \dots, k_0$) – веса связей от нейронов входного слоя к нейронам скрытого слоя; $f_{A_k^0}$ – функция активации нейронов A^0 -слоя.

Шаг 5. Определяются входные и выходные сигналы распознающих нейронов нулевого блока:

$$U_{\text{вх } Y_p^0} = 1 \cdot W_{0 Y_p^0} + \sum_{k=1}^{k_0} U_{\text{вых } A_k^0} W_{A_k^0 Y_p^0}, \quad p = 1, \dots, n_0;$$

$$U_{\text{вых } Y_p^0} = f_{Y_p^0}(U_{\text{вх } Y_p^0}), \quad p = 1, \dots, n_0,$$

где $U_{\text{вх } Y_p^0}$, $U_{\text{вых } Y_p^0}$, $p = 1, \dots, n_0$ – соответственно входные и выходные сигналы нейронов Y -слоя; $W_{0 Y_p^0}$, $p = 1, \dots, n_0$ – веса связей для сигнала смещения на нейроны $Y_1^0, \dots, Y_{n_0}^0$; $W_{A_k^0 Y_p^0}$, $k = 1, \dots, k_0$, $p = 1, \dots, n_0$ – веса связей от элементов A -слоя к нейронам Y -слоя; $f_{Y_p^0}$ – функция активации нейронов Y -слоя.

Шаг 6. Определяются входные и выходные сигналы D -нейронов нулевого блока:

$$U_{\text{вх } D_p^0} = U_{\text{вых } Y_p^0}, \quad p = 1, \dots, n_0; \quad U_{\text{вых } D_p^0} = \begin{cases} 1, & \text{если } (1 - \delta) \leq U_{\text{вх } D_p^0} < 1, \\ 0, & \text{если } (-1 + \delta) < U_{\text{вх } D_p^0} < 1 - \delta, \\ -1, & \text{если } -1 < U_{\text{вх } D_p^0} \leq -1 + \delta, \end{cases}$$

где $U_{\text{вх } D_p^0}$, $U_{\text{вых } D_p^0}$, $p = 1, \dots, n_0$ – соответственно входные и выходные сигналы нейронов D -слоя нулевого блока; δ – максимально допусти-

мая ошибка приближения сигналов "1" и "-1" с помощью нейронов с биполярной функцией активации.

Шаг 7. Определяются входные и выходные сигналы управляющего нейрона Σ_0 :

$$U_{\text{вых } \Sigma_0} = \begin{cases} 0, & \text{если } U_{\text{вх } \Sigma_0} = \sum_{p=1}^{n_0} U_{\text{вых } D_p^0} = -(n-2), \\ 1, & \text{если } U_{\text{вх } \Sigma_0} = \sum_{p=1}^{n_0} U_{\text{вых } D_p^0} \neq -(n-2). \end{cases}$$

Если $U_{\text{вых } \Sigma_0} = 0$, то с помощью выходных сигналов D^0 -нейронов определяется класс, к которому относится входное изображение. Затем осуществляется переход на шаг 1 алгоритма. Если $U_{\text{вых } \Sigma_0} = 1$, то – входное изображение не принадлежит множеству M_0 . В этом случае по цепи обратной связи фиксируется единичный выходной сигнал нейрона Σ_0 . Этот сигнал блокирует нейроны $Y_1^0, \dots, Y_{n_0}^0$ и подключает к распознаванию нейроны первой группы.

Шаг 8. Число l увеличивается на единицу: $l = l + 1$. Определяются входные и выходные сигналы нейронов слоев A , Y , D блока l :

$$\begin{aligned} U_{\text{вх } A_k^l} &= 1 \cdot W_{X_0 A_k^l} + \sum_{i=1}^{n_l} U_{\text{вых } X_i} W_{X_i A_k^l}, \quad k = 1, \dots, k_l; \\ U_{\text{вых } A_k^l} &= f_{A_k^l}(U_{\text{вх } A_k^l}), \quad k = 1, \dots, k_l; \\ U_{\text{вх } Y_p^l} &= 1 \cdot W_{0Y_p^l} + \sum_{k=1}^{k_l} U_{\text{вых } A_k^l} W_{A_k^l Y_p^l}, \quad p = 1, \dots, n_l; \\ U_{\text{вых } Y_p^l} &= f_{Y_p^l}(U_{\text{вх } Y_p^l}), \quad p = 1, \dots, n_l; \\ U_{\text{вх } D_p^l} &= U_{\text{вых } Y_p^l}, \quad p = 1, \dots, n_l; \quad U_{\text{вых } D_p^l} = \begin{cases} 1, & \text{если } (1-\delta) \leq U_{\text{вх } D_p^l} < 1, \\ 0, & \text{если } (-1+\delta) < U_{\text{вх } D_p^l} < 1-\delta, \\ -1, & \text{если } -1 < U_{\text{вх } D_p^l} \leq -1+\delta, \end{cases} \end{aligned}$$

где $U_{\text{вх } A_k^l}$, $U_{\text{вых } A_k^l}$ ($k = 1, \dots, k_l$), $U_{\text{вх } Y_p^l}$, $U_{\text{вых } Y_p^l}$, $U_{\text{вх } D_p^l}$, $U_{\text{вых } D_p^l}$ ($p = 1, \dots, n_l$) – соответственно входные и выходные сигналы нейронов A^l -слоя, Y^l -слоя и D^l -слоя блока l нейронной сети; $W_{X_0 A_k^l}$ ($k = 1, \dots, k_l$), $W_{0Y_p^l}$ ($p = 1, \dots, n_l$) – веса связей для сигналов смещения соответствен-

но на нейроны A^l и Y^l слоев; $W_{X_i A_k^l}$, $W_{A_k^l Y_p^l}$ ($i=1, \dots, n$; $k=1, \dots, k_l$; $p=1, \dots, n_l$) – соответственно веса связей от нейронов входного слоя к нейронам скрытого A^l -слоя и от нейронов A^l -слоя к нейронам Y^l -слоя.

Шаг 9. Определяются входной $U_{\text{вх } \Sigma_l}$ и выходной $U_{\text{вых } \Sigma_l}$ сигналы управляющего нейрона Σ_l :

$$U_{\text{вх } \Sigma_l} = \sum_{p=1}^{n_l} U_{\text{вых } D_p^l}; \quad U_{\text{вых } \Sigma_l} = \begin{cases} 0, & \text{если } U_{\text{вх } \Sigma_l} = -(n_l - 2), \\ 1, & \text{если } U_{\text{вх } \Sigma_l} \neq -(n_l - 2). \end{cases}$$

Если $U_{\text{вых } \Sigma_l} = 0$, то с помощью выходных сигналов D^l -нейронов определяется класс, к которому относится входное изображение. Затем осуществляется переход на шаг 1 алгоритма.

Если $U_{\text{вых } \Sigma_l} = 1$ и $l = L$, то – переход на шаг 1 алгоритма, поскольку входное изображение не принадлежит множеству M_l и в памяти сети отсутствует класс изображений, к которому может относиться входное изображение. Оно воспринимается как представитель нового класса изображений, который необходимо запомнить на $(L + 1)$ -м дообучении нейронной сети в $(L + 1)$ -м блоке.

Если $U_{\text{вых } \Sigma_l} = 1$ и $l < L$, то входное изображение не принадлежит множеству M_l , но может принадлежать множеству M_{l+1} . Сигнал с выхода нейрона Σ_l по цепи обратной связи фиксирует свой единичный выходной сигнал, блокирует работу нейронов $Y_1^l, \dots, Y_{n_l}^l$ и подключает к распознаванию нейроны группы $(l + 1)$.

Шаг 10. Проверяются условия прекращения работы алгоритма. Если они выполняются, то переход на шаг 11 алгоритма, в противном случае – переход на шаг 8 алгоритма.

Шаг 11. Останов.

Полученные результаты несложно обобщить и на перцептроны, имеющие более трех слоев нейронов.

Выводы

Впервые разработана нейросетевая архитектура и алгоритмы функционирования трехслойного перцептрона, способного дообучаться в процессе своего функционирования.

Библиографический список

1. **Suzuki K.** Artificial Neural Networks: Architectures and Applications / K. Suzuki. – Publisher: InTech, 2013. – 256 p.

2. **Bianchini M.** Handbook on Neural Information Processing (Intelligent Systems Reference Library) / M. Bianchini. – Publisher: Springer, 2013. – 499 p.
3. **Cirrincione M.** Power Converters and AC Electrical Drives with Linear Neural Networks (Energy, Power Electronics, and Machines) / M. Cirrincione, M. Pucci, G. Vitale. – Publisher: CRC Press, 2012. – 631 p.
4. **Галушкин А.И.** Нейронные сети. Основы теории / А.И. Галушкин. – М.: Горячая линия – Телеком, 2012. – 496 с.
5. **Капля В.И.** Системы искусственного интеллекта: учебное пособие. – Волгоград: ИУНЛ ВолгГТУ. – 2011. – 97 с.
6. **Russell S.** Artificial Intelligence: A Modern Approach, Third Edition / S. Russell, P. Norvig. – Publisher: Prentice Hall, 2010. – 1152 p.
7. **Девятков В. В.** Системы искусственного интеллекта / Гл. ред. И. Б. Фёдоров. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2001. – 352 с.
8. **Carpenter G.A.** Massively parallel architecture for self-organising neural pattern recognition machine / G.A. Carpenter, S.A. Grossberg // Computing, Vision, Graphics and Image Processing. – 1987. – Vol. 37. – P. 54 – 115.
9. **Grossberg S.** Competitive learning: From interactive activation to adaptive resonance / S. Grossberg // Cognitive Science. – 1987. – Vol. 11. – P. 23 – 63.
10. **Fausett L.** Fundamentals of Neural Networks. Architectures, Algorithms and Applications / L. Fausett. – New Jersey: Prentice Hall Int., Inc., 1994. – 461 p.
11. **Дмитриенко В.Д.** Нейросетевое устройство направленных ассоциаций / В.Д. Дмитриенко, А.Ю. Заковоротный, Хавина И.П. // Научные ведомости. Серия: История, политология, экономика, информатика. – Белгород: БГУ, 2010. – № 7(78). Вып. 14/1. – С. 110–119.
12. **Дмитриенко В.Д.** Ассоциативная нейронная сеть АРТ / В.Д. Дмитриенко, А.Ю. Заковоротный, В.А. Бречко // Сборник трудов Международной молодежной конференции "Прикладная математика, управление и информатика". – Белгород: ИД "Белгород", 2012. – Т. 1. – С. 115–118.