

Адаптивная резонансная теория: нейронные сети ART1 и Fuzzy ART.

Один из основных исследовательских интересов в области нейронных сетей заключается в разработке моделей функций мозга. Прогресс в этом направлении зависит от результатов исследований, полученных из областей нейрофизиологии, психологии и математики. Большой интерес представляет моделирование того, как мозг обрабатывает меняющийся с течением времени поток сенсорной информации. С одной стороны человеческая память запоминает новую информацию, с другой – сохраняет уже накопленную. Ключевая проблема многих современных нейронных сетей заключается в том, что они не могут быстро изучать новую информацию в динамической среде, не забывая старую. Предъявление всего одного нового образа часто уничтожает или изменяет результаты предшествующего обучения. Если сеть работает в реальном времени, например, обрабатывает сенсорную информацию, то обучающее множество может все время меняться. Для большинства моделей нейронных сетей это приводит к отсутствию обучения. Адаптивная резонансная теория (ART) решает дилемму стабильности и пластичности при запоминании образов. А именно, при предъявлении нового образа, восприятие системы должно быть пластично или адаптировано к новой информации, и при этом стабильно, то есть не разрушать память о старых образах.

Адаптивная резонансная теория, как теория когнитивной обработки информации, была введена Stephen Grossberg [1]. Из этой работы появились целые классы сетей, используемых для распознавания, классификации и запоминания различных объектов. ART1 – самая первая модель нейронной сети, которая была предложена Gail Carpenter и Stephen Grossberg [2]. Она представляет собой классификатор двоичных входных данных и обучается без учителя. Сеть ART2 [3] классифицирует как двоичные, так и аналоговые входные данные (компоненты векторов действительные числа), но механизм обучения тот же, что и у ART1. Нечеткая нейронная сеть Fuzzy ART [5], как и ART2, является модификацией сети ART1, позволяя обрабатывать как двоичные, так и аналоговые входные данные с помощью включения операторов из теории нечетких множеств. Она является наиболее широко используемой моделью ART. Модель ART3 [4] основана на ART2 и близка к синаптическим процессам реальных нейронов. Нейронная сеть ARTMAP [6] комбинирует классификаторы ART1 и ART2 для обучения с учителем. Существуют много вариантов сетей ART для обучения с учителем и без, например, Gaussian ART, Hypersphere ART, Gaussian ARTMAP, Hypersphere ARTMAP, Fuzzy ARTMAP, ToroART и так далее. В этой лекции мы рассмотрим нейронные сети ART1 и Fuzzy ART.

1 Нейронная сеть ART1.

Нейронная сеть ART1 работает только с двоичными входными данными. Согласно работе [2], общая схема ART состоит из двух типов памяти: кратковременной (STM) и долговременной (LTM), которые связаны между собой. Текущий входной образ помещается в кратковременную память и сравнивается с хранящимися в долговременной памяти. Если содержимое STM похоже на образ из LTM, то возникает состояние резонанса - т.е. система "узнаёт" входной образ. В противном случае текущий входной образ считается "новым". Таким образом реализуется детектор новизны образов. Кратковременная и долговременная памяти описываются дифференциальными уравнениями, которые заменяются алгебраическими уравнениями, соответствующие стационарному состоянию системы.

В отличие от большинства существующих архитектур нейронных сетей, ART-сети не предусматривают строгого деления жизненного цикла на стадии обучения и практического использования. Они продолжают обучаться на протяжении всего времени их практического использования.

Пусть $x = (x_1, \dots, x_N)$ – двоичный входной вектор. Сеть состоит из двух слоев нейронов: входного или нижнего слоя, в котором число нейронов фиксировано и равно размерности входного вектора, и выходного или верхнего слоя с переменным количеством нейронов, где каждому нейрону соответствует один класс объектов или кластеров. Связь между нейронами, принадлежащими различным слоям, осуществляется по принципу каждый с каждым. Всем связям поставлен в соответствие некоторый вес. Используется два набора весов между слоями. Пусть b_{ij}^t – восходящий вес связи от входного i -го нейрона к выходному j -му нейрону в момент времени t , а c_{ji}^t – нисходящий вес связи от выходного j -го нейрона к входному i -му нейрону в момент времени t . На рисунке 1 показана структура нейронной сети ART1.

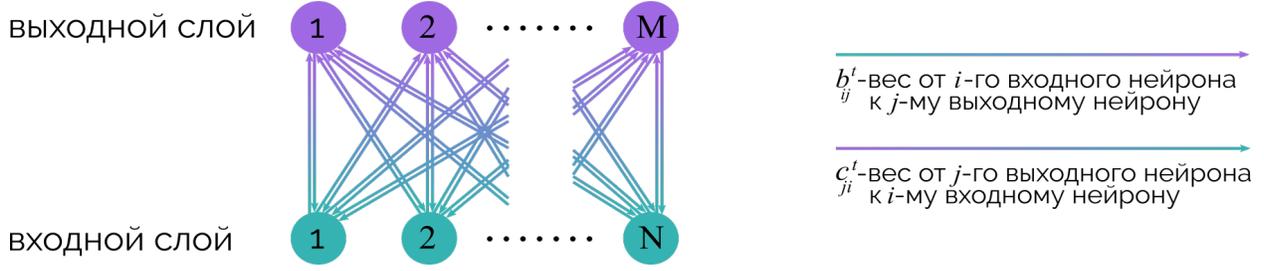


Рис. 1: Нейронная сеть ART1.

Введем следующие обозначения. Для любых $u = (u_1, \dots, u_N)$ и $v = (v_1, \dots, v_N)$

$$\|u\| = \sum_{i=1}^N |u_i| \text{ — норма в пространстве } L_1,$$

$$\langle u, v \rangle = \sum_{i=1}^N u_i v_i \text{ — скалярное произведение в пространстве } L_1.$$

На первоначальном этапе в выходном слое только один нейрон, то есть $M = 1$. Его веса инициализируются следующим образом:

$$b_{i1}^0 = \frac{L}{L-1+N}, \quad c_{1i}^0 = 1, \quad 1 \leq i \leq N, \tag{1}$$

где $L > 1$ — некоторая константа (чаще всего $L = 2$). Важнейшим параметром в алгоритме является параметр схожести r , $0 \leq r \leq 1$, который показывает насколько входной вектор должен совпадет с одним из классов или образов. Этот параметр также устанавливается на первоначальном этапе. Чем он меньше, тем менее похожие векторы будут отнесены к одному классу. Близкое к единице значение параметра требует почти полного совпадения.

Алгоритм включает следующие этапы:

1. при предъявлении сети нового входного вектора x для каждого нейрона выходного слоя вычисляется функция состояния

$$y_j = \langle b_j^t, x \rangle = \sum_{i=1}^N b_{ij}^t x_i, \quad 1 \leq j \leq M,$$

где M — включает $M - 1$ зафиксированных нейронов (то есть те, которые выбирались хотя бы один раз для представления входного вектора) и один незафиксированный (то есть тот, у которого веса инициализируются по формулам (1)). На первоначальном этапе зафиксированных нейронов нет. Один незафиксированный нейрон гарантирует выполнение условия резонанса (2) на следующем шаге. Из этих M нейронов выбирается нейрон с наибольшим значением $y_k = \max_j y_j$, который считается "победителем".

Если значение y_k принимают сразу несколько нейронов, то выбирается нейрон с наименьшим индексом.

2. выбранный k -ый нейрон из выходного слоя и входной вектор сравниваются следующим образом:

$$\frac{\langle c_k^t, x \rangle}{\|x\|} \geq r, \quad \text{где } c_k^t = (c_{k1}^t, \dots, c_{kN}^t). \tag{2}$$

Если условие (2) выполнено для зафиксированного нейрона k , $1 \leq k \leq M - 1$, то система пришла в состояние резонанса и входной вектор успешно классифицирован. Номер класса соответствует номеру нейрона-победителя k и переходим к следующему шагу 3. Если условие (2) выполнено для незафиксированного нейрона M , то это означает, что система обнаружила новый образ, который не похож ни на один, из хранящихся в памяти ЛТМ, и этот образ будет представлять незафиксированный M -ый нейрон. В этом случае увеличиваем M на единицу и резервируем новый нейрон выходного слоя для нового образа (или класса) с весами $c_{M+1,i}^{t+1}$, $b_{i,M+1}^{t+1}$, равными правой части в формулах (1), и переходим на шаг 4;

3. для k -го нейрона пересчитаем весовые коэффициенты по формулам:

$$c_{ki}^{t+1} = c_{ki}^t x_i, \quad b_{ik}^{t+1} = \frac{L c_{ki}^t x_i}{L - 1 + \langle c_k^t, x \rangle}, \quad 1 \leq i \leq N; \quad (3)$$

4. веса M -го незафиксированного нейрона пересчитываются по следующим формулам:

$$c_{Mi}^{t+1} = x_i, \quad b_{iM}^{t+1} = \frac{L x_i}{L - 1 + \|x\|},$$

которые следуют из (3) для $c_{ki}^t = 1, 1 \leq i \leq N$.

Заметим, если значение параметра L взять достаточно большое, то функция состояния для незафиксированного нейрона будет больше, чем для зафиксированных. Поэтому каждый раз при предъявлении сети нового входного вектора будет создаваться новый класс, что приведет к неэффективной классификации.

Сеть ART решает дилемму стабильности-пластичности и позволяет быстро запоминать новые образы без утраты старых. Однако основными недостатками адаптивной резонансной нейронной сети считается чувствительность к векторам, содержащим шум, а также к порядку предъявления входных векторов, так как любой идентифицированный образ подстраивает веса нейрона соответствующего класса под новую информацию.

2 Нечеткая нейронная сеть Fuzzy ART.

Важным отличием аналоговых данных от двоичных является принципиальная возможность аналоговых быть сколь угодно близкими друг к другу (в то время как пространство двоичных данных дискретно). Это накладывает дополнительные требования на функционирование нейронов выходного слоя или слоя сравнения – требуется более тонкий и чувствительный механизм для выделения областей резонанса. Нечеткая нейронная сеть Fuzzy ART расширяет возможности сети ART1, позволяя обрабатывать аналоговые входные данные с помощью включения операторов из теории нечетких множеств [12].

Пусть $x = (x_1, \dots, x_N)$ – входной вектор, где каждая компонента – действительное число из отрезка $[0, 1]$. Сеть состоит из двух слоев нейронов: входного, где число нейронов фиксировано, и выходного с переменным количеством нейронов, где каждому нейрону соответствует один класс объектов или кластеров. Связь между нейронами, принадлежащими различным слоям, осуществляется по принципу каждый с каждым.

В нечеткой нейронной сети Fuzzy ART возникает проблема размножения категорий [9]. А именно с течением времени значения компонент весового вектора уменьшаются до такой степени, что кластер больше не является репрезентативным для своей категории и, следовательно, необходимо сгенерировать набор новых кластеров, которые будут соответствовать входным данным с более высокими значениями компонент весового вектора. Чтобы предотвратить это, входной вектор либо нормализуется

$$I = \frac{x}{\|x\|},$$

либо производится дополнительное кодирование

$$I = (x, x^c) = (x_1, \dots, x_N, x_1^c, \dots, x_N^c), \quad (4)$$

где $x^c \equiv 1 - x$ дополнение к вектору x . Каждый класс j соответствует вектору адаптивных весов $w_j^t = (w_{j1}^t, w_{j2}^t, \dots, w_{j2N}^t)$, $j = 1, \dots, M$, которые включают восходящие и нисходящие веса. Число категорий M может быть произвольно большим.

Для любых $u = (u_1, \dots, u_N)$ и $v = (v_1, \dots, v_N)$ введем оператор пересечения \wedge из теории нечетких множеств [12] (также можно посмотреть в первой части лекций по Теории возможностей):

$$(u \wedge v)_i \equiv \min(u_i, v_i), \quad (5)$$

$$\|u \wedge v\| = \sum_{i=1}^N \min(u_i, v_i).$$

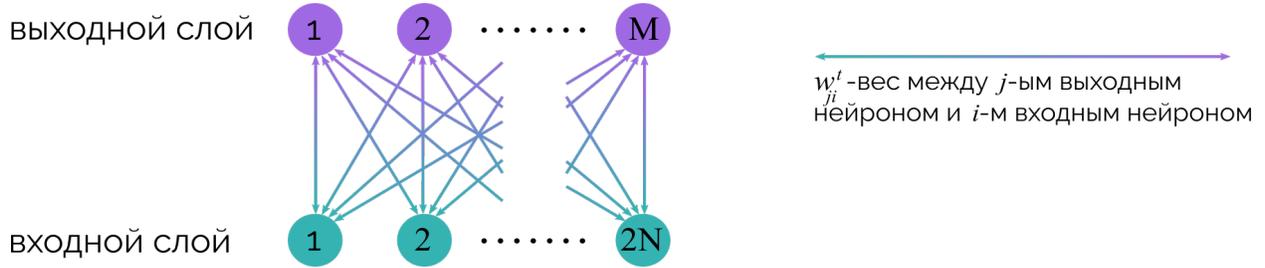


Рис. 2: Нечеткая нейронная сеть Fuzzy ART.

Нечеткая нейронная сеть Fuzzy ART определяется следующими параметрами: параметром $\alpha \in (0, +\infty)$, обучающим параметром $\beta \in [0, 1]$ и параметром похожести $r \in [0, 1]$. Параметр α обычно выбирается маленьким. На начальном этапе в выходном слое только один незафиксированный нейрон с единичными весами:

$$w_{1i}^0 = 1, 1 \leq i \leq 2N.$$

Алгоритм включает следующие этапы:

1. при предъявлении сети нового входного вектора $I = (x, x^c)$ для каждого нейрона выходного слоя вычисляется функция состояния

$$y_j = \frac{\|I \wedge w_j^t\|}{\alpha + \|w_j^t\|}, j = 1, \dots, M, \quad (6)$$

где M включает $M - 1$ зафиксированных нейронов и один незафиксированный. На начальном этапе зафиксированных нейронов нет. Один незафиксированный нейрон гарантирует выполнение условие резонанса (7) на следующем шаге. Из этих M нейронов выбирается нейрон с наибольшим значением $y_k = \max_j y_j$, который считается "победителем". Если значение y_k принимают сразу несколько нейронов, то выбирается нейрон с наименьшим индексом;

2. для выбранного k -го нейрона проверяется условие резонанса:

$$\frac{\|I \wedge w_k^t\|}{\|I\|} \geq r. \quad (7)$$

Если условие (7) выполнено для зафиксированного нейрона k , $1 \leq k \leq M - 1$, то система пришла в состояние резонанса и входной вектор считается похожим на k -ый нейрон (или класс). Номер класса соответствует номеру нейрона-победителя k . После этого переходим к следующему шагу 3. Если условие (7) выполняется для незафиксированного нейрона M , то это означает, что система обнаружила новый образ. В этом случае увеличиваем M на единицу и резервируем новый нейрон выходного слоя для нового образа (или класса) с единичными весами $w_{M+1,i}^{t+1} = 1, 1 \leq i \leq 2N$ и переходим к шагу 4;

3. для k -го нейрона пересчитываем весовой вектор:

$$w_k^{t+1} = \beta(I \wedge w_k^t) + (1 - \beta)w_k^t, 1 \leq k \leq M - 1; \quad (8)$$

4. веса M -го незафиксированного нейрона пересчитываются согласно формуле:

$$w_M^{t+1} = I.$$

В формуле (8) параметр β определяет скорость обучения сети. Для быстрого обучения $\beta = 1$, для медленного $0 < \beta < 1$. Значения компонент весового вектора w_j^t могут только уменьшаться в течении обучения. А классы, наоборот, расширяются в течении обучения.

Обсудим параметр α . Вектор a называется нечетким подмножеством множества b [12], если

$$a \wedge b = a.$$

При $\alpha \rightarrow 0$ функция состояния y_j в (6) отражает степень, с которой весовой вектор w_j^t является нечетким подмножеством входного вектора I . Если для нейрона k

$$y_k = \frac{\|I \wedge w_k^t\|}{\|w_k^t\|} = 1,$$

тогда w_k^t является нечетким подмножеством I и k -ый нейрон оказался нейроном-победителем. Чтобы для этого нейрона выполнялось условие резонанса

$$\frac{\|I \wedge w_k^t\|}{\|I\|} = \frac{\|w_k^t\|}{\|I\|} \geq r,$$

необходимо чтобы $\|w_k^t\|$ быть максимальным. Тогда согласно (8), $w_k^{t+1} = w_k^t$ и значения весов сохраняются. Если условие резонанса не выполняется для выбранного нейрона k с максимальным $\|w_k^t\|$, то оно не будет выполняться и для всех остальных нейронов j , для которых w_j – нечеткое подмножество I .

При достаточно большом значении параметра α функция состояния y_j в (6) для незафиксированного нейрона всегда будет больше функции состояния для зафиксированных нейронов, так как для незафиксированного нейрона в силу нормализации I

$$\frac{\|I \wedge w_j^t\|}{\alpha + \|w_j^t\|} = \frac{\|I\|}{\alpha + \|w_j^t\|} = \frac{N}{\alpha + 2N}.$$

Поэтому каждый раз при предъявлении нового входного вектора нейронная сеть всегда будет выбирать незафиксированный нейрон. Таким образом, каждый входной вектор будет формировать собственный класс в выходном слое и нейронная сеть не будет выполнять полезную классификацию. Однако, как упомянуто в [7], начальные значения весовых векторов могут быть взяты достаточно большими, чтобы решить эту проблему и алгоритм нейронной сети Fuzzy ART для больших α также может быть полезен, как и для маленьких α .

Дополнительный свойства нечеткой нейронной сети Fuzzy ART можно посмотреть в работе J. Huang [8]. Далее мы покажем, почему возникает размножение категорий без дополнительного кодирования (4).

2.1 Геометрия Fuzzy ART без дополнительного кодирования.

Рассмотрим алгоритм Fuzzy ART без дополнительного кодирования входящего вектора x . Пусть $N = 2$ и $\beta = 1$. Для заданного вектора x формула (6) определяет многоугольник

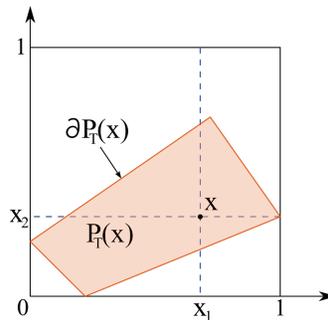
$$y_j(x) = \frac{\|x \wedge w_j\|}{\alpha + \|w_j\|}, \alpha > 0, \quad (9)$$

где для простоты мы опустили верхний индекс t у весового вектора w_j^t . Для любого $T \in [0, 1]$ определим область

$$P_T(x) = \left\{ w \in [0, 1] \times [0, 1] : \frac{\|x \wedge w\|}{\alpha + \|w\|} \geq T \right\},$$

граница которой

$$\partial P_T(x) = \left\{ w \in [0, 1] \times [0, 1] : \frac{\|x \wedge w\|}{\alpha + \|w\|} = T \right\}.$$



Уменьшение значений T порождает семейство более крупных многоугольников $P_T(x)$, причем большее семейство будет содержать меньшее, и все они будут содержать точку x . Чтобы вектор $w_j \in P_T(x)$, необходимо чтобы $y_j(x) \geq T$. Таким образом, чтобы найти нейрон k , для которого $y_j(x)$, $j = k$, максимально, наибольшее T должно быть таким, чтобы

$$x \in P_T(x), w_k \in \partial P_T(x) \text{ и } w_j \notin P_S(x), j = 1, \dots, M, \text{ для любого } S > T.$$

Для этого T уменьшается и, соответственно, многоугольник $P_T(x)$ расширяется до тех пор пока его граница $\partial P_T(x)$ не пересечет первый вектор w_j , $j = 1, \dots, M$ для некоторого T . Тогда $k = j$ и $y_k(x) = T$.

После того, как мы выбрали нейрон k , проверяем условие резонанса:

$$\|x \wedge w_k\| \geq r\|x\|. \quad (10)$$

Вектор x делит единичный квадрат на 4 области.

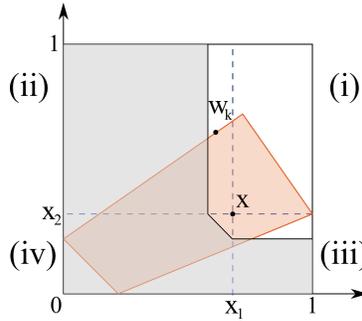


Рис. 3: Входящий вектор x делит единичный квадрат на 4 области. Серым цветом закрашена область, где условие (7) не выполняется для нейрона k . Соответственно, остальная часть – это область резонанса.

Если w_k из области (i), то $\|x \wedge w_k\| = \|x\|$ и условие (7) выполняется для любого $r \in [0, 1]$. Если w_k из области (iv), то $\|x \wedge w_k\| = \|w_k\|$ и нейрон k блокируется, если $\|w_k\| < r\|x\|$. Граница области, где условие резонанса не выполнено, в (iv) определяется прямой линией $\{w : \|w\| = r\|x\|\}$, а в областях (ii) и (iii) соответственно вертикальной и горизонтальной прямыми линиями. На рисунке 3 серым цветом закрашена вся область, где условие резонанса не выполнено.

После того, как нейрон k выбран в резонансной области, $x \wedge w_k$ становится новым весовым вектором для категории k , т.е. $w_k^{t+1} = x \wedge w_k^t$, так как $\beta = 1$. На рисунке 4 показано, как w_k проецируется на область (iv), а именно на закрашенный треугольник.

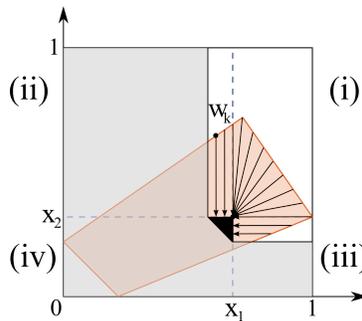


Рис. 4: Выбранный из резонансной области весовой вектор w_k проецируется на закрашенный треугольник в течении обучения.

В течении обучения w_k стягивается к началу координат. Однако, когда вектор w_k стремится к началу координат, он покидает область резонанса для большинства входящих векторов x . Чтобы удовлетворять условию резонанса, будущие входящие векторы вынуждены перетаскивать остальные весовые векторы к началу координат или создавать новые нейроны (классы) в выходном слое. На следующем рисунке для двух разных

весовых векторов w_k и w'_k показаны множества входящих векторов x , для которых возникает резонанс, если выбраны классы k и k' соответственно. При стремлении $\|w_k\|$ к нулю множество входящих векторов x , для которых возникает резонанс, также будет стремиться к нулю. Это приводит к тому, что кластер больше не является репрезентативным для своей категории и требуется сгенерировать новый набор кластеров.

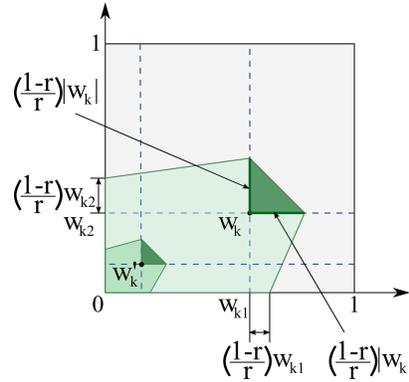


Рис. 5: Для заданного класса k множество входящих векторов x , для которых возникает резонанс, закрашен светло-зеленым цветом, включая темно-зеленый треугольник. Темно-зеленый треугольник – это та область, где веса не меняются в течении обучения.

Если входящий вектор x попадает в темно-зеленый треугольник, то вес w_k не меняется в течении обучения, так как $w_k^{t+1} = x \wedge w_k^t = w_k^t$.

2.2 Геометрия нечеткой ART сети с дополнительным кодированием.

Как и в предыдущем параграфе предположим размерность $N = 2$, а параметр $\beta = 1$. Тогда для входящего двумерного вектора $x = (x_1, x_2)$ после дополнительного кодирования получаем вектор

$$I = (x_1, x_2, 1 - x_1, 1 - x_2). \quad (11)$$

Следуя (11), весовой вектор w_j представим в виде:

$$w_j = (u_j, v_j^c), \quad (12)$$

где u_j, v_j – векторы размерности 2. Верхний индекс t мы опустили для простоты. Каждый нейрон или класс j можно геометрически представить, как четырехугольник R_j . Пусть u_j – определяет левый нижний угол прямоугольника R_j , а v_j – верхний правый угол (Рис. 6). Под размером R_j будем понимать длину его диагонали, а именно $\|R_j\| = \|v_j - u_j\|$, которая также равна длине вектора суммы длины и высоты R_j .

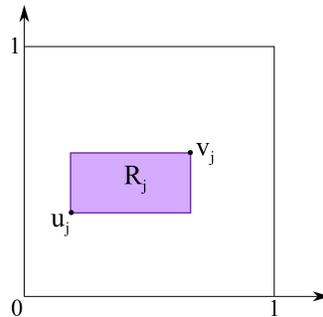


Рис. 6: Каждый нейрон имеет геометрическую интерпретацию четырехугольника с углами (u_j, v_j) .

Помимо оператора (5), введем оператор \vee из теории нечетких множеств. Для любых $u = (u_1, \dots, u_N)$ и $v = (v_1, \dots, v_N)$

$$(u \vee v)_i \equiv \max(u_i, v_i).$$

Предположим, что входящий вектор I активировал победивший нейрон k , то есть возникает резонанс и выполнено (7). Тогда веса меняются согласно формуле:

$$w_k^{new} = I \wedge w_k = (x, x^c) \wedge (u_k, v_k^c) = (x \wedge u_k, x^c \wedge v_k^c) = (x \wedge u_k, (x \vee v_k)^c) = (u_k^{new}, (v_k^{new})^c).$$

Откуда следует, что нижний левый и верхний правый углы обновленного четырехугольника R_k^{new} меняются согласно соотношениям:

$$u_k^{new} = x \wedge u_k, v_k^{new} = x \vee v_k. \quad (13)$$

Заметим, в течении обучения норма весового вектора $\|w_k\| = \|(u_k, v_k^c)\|$ уменьшается. Если $x \in R_k$, тогда согласно (13) прямоугольник R_k не меняется. Если $x \notin R_k$, тогда R_k увеличивается до минимального прямоугольника, содержащего x и R_k . Обозначим его $R_k \oplus x$.

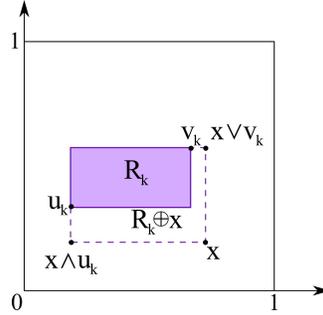


Рис. 7: В течении быстрого обучения R_k расширяется до $R_k \oplus x$ – минимального четырехугольника, содержащего x и R_k .

Покажем какой максимальный размер может достигать прямоугольник R_k , сохраняя условие резонанса (7). Из (13) $\|R_k \oplus x\| = \|(x \vee v_k) - (x \wedge u_k)\|$. Так как над входящем вектором x проведено дополнительное кодирование, то

$$\|I\| = \|(x, x^c)\| = \sum_{i=1}^2 x_i + \sum_{i=1}^2 (1 - x_i) = 2,$$

а условие резонанса (7) запишется в виде:

$$\|I \wedge w_k\| \geq 2r. \quad (14)$$

Однако

$$\|I \wedge w_k\| = \|((x \wedge u_k), (x \vee v_k)^c)\| = \|x \wedge u_k\| + 2 - \|x \vee v_k\| = 2 - \|R_k \oplus x\|.$$

Учитывая (14), $\|R_k \oplus x\| \leq 2(1 - r)$. Если параметр r близок к 1, тогда все R_j маленькие. Если r близок к нулю, то некоторые R_j могут расти, заполняя почти весь квадрат $[0, 1] \times [0, 1]$.

Мы показали, что для произвольных входящих аналоговых векторов с дополнительным кодированием нечеткая нейронная сеть Fuzzy ART в случае $\beta = 1$ формирует четырехугольные классы R_j , которые во время обучения растут до максимального размера $\|R_j\| \leq 2(1 - r)$, при этом $\|w_j\|$ убывает. Для случая размерности N получен аналогичный результат [5].

Нечеткая нейронная сеть Fuzzy ART расширяет возможности нейронной сети ART1, классифицируя аналоговые входные сигналы. Но зависимость классификации от порядка представления входных векторов и чувствительность к шуму остаются. Для уменьшения влияния шума были разработаны нейронные сети Gaussian ART, Gaussian ARTMAP [11], а также ТороART [10].

Литература

- [1] S. Grossberg, Adaptive pattern classification and universal recoding. II. Feedback, expectation, olfaction, and illusions, *Bioi. Cybemet.* 23, 1976, 187-202.
- [2] G. A. Carpenter, S. Grossberg, A massively parallel architecture for a self-organizing neural pattern recognition machine, *Computer Vision, Graphics, and Image Processing*, Vol. 37, Issue 1, 1987, 54-115.

- [3] G. A. Carpenter, S. Grossberg, ART 2: self-organization of stable category recognition codes for analog input patterns, *Applied Optics*, Vol. 26, Issue 23, 1987, 4919-4930.
- [4] G. A. Carpenter, S. Grossberg, ART 3: Hierarchical search using chemical transmitters in self-organizing pattern recognition architectures, *Neural Networks*, Vol. 3, Issue 2, 1990, 129-152.
- [5] G. A. Carpenter, S. Grossberg, D.B. Rosen, Fuzzy ART: Fast stable learning and categorization of analog patterns by an adaptive resonance system, *Neural Networks*, Vol. 4, Issue 6, 1991, 759-771.
- [6] G. A. Carpenter, S. Grossberg, J.H. Reynolds, ARTMAP: supervised real-time learning and classification of nonstationary data by a self-organizing neural network, *Neural Networks*, Vol. 4, Issue 5, 1991, 565-588.
- [7] M. Georgiopoulos, I. Dagher, G. L. Heileman, G. Bebis, Properties of learning of a Fuzzy ART Variant, *Neural Networks*, Vol. 12, Issue 6, 1999, 837-850.
- [8] J. Huang, M. Georgiopoulos, G.L. Heileman, Fuzzy ART properties. *Neural Networks*, Vol. 8, Issue 2, 1995, 203-213.
- [9] B. Moore, ART1 and pattern clustering. In proceedings of the 1988 Connectionist Models Summer School edited by Touretzky, D., Hinton, G., Sejnowski, T., Published by Morgan Kaufmann, San Mateo, CA, 1989.
- [10] M. Tscherepanow, TopoART: A Topology Learning Hierarchical ART Network, In Proceedings of the International Conference on Artificial Neural Networks (ICANN), Part III, 2010, 157-167.
- [11] J.R. Williamson, Gaussian ARTMAP: A Neural Network for Fast Incremental Learning of Noisy Multidimensional Maps. *Neural Networks*, Vol. 9, 1996, 881 - 897.
- [12] L.A. Zadeh, Fuzzy sets, *Information and Control*, V.8, 1965.