

Сети встречного распространения

Введение в сети встречного распространения

Возможности сети встречного распространения превосходят возможности однослойных сетей. Время же обучения по сравнению с обратным распространением может уменьшаться в сто раз. Встречное распространение не столь общо, как обратное распространение, но оно может давать решение в тех приложениях, где долгая обучающая процедура невозможна. Помимо преодоления ограничений других сетей встречное распространение обладает собственными интересными и полезными свойствами.

Во встречном распространении объединены два хорошо известных алгоритма:

- самоорганизующаяся карта Кохонена;
- звезда Гроссберга.

Объединение этих алгоритмов ведет к свойствам, которых нет ни у одного из них в отдельности.

Сети встречного распространения

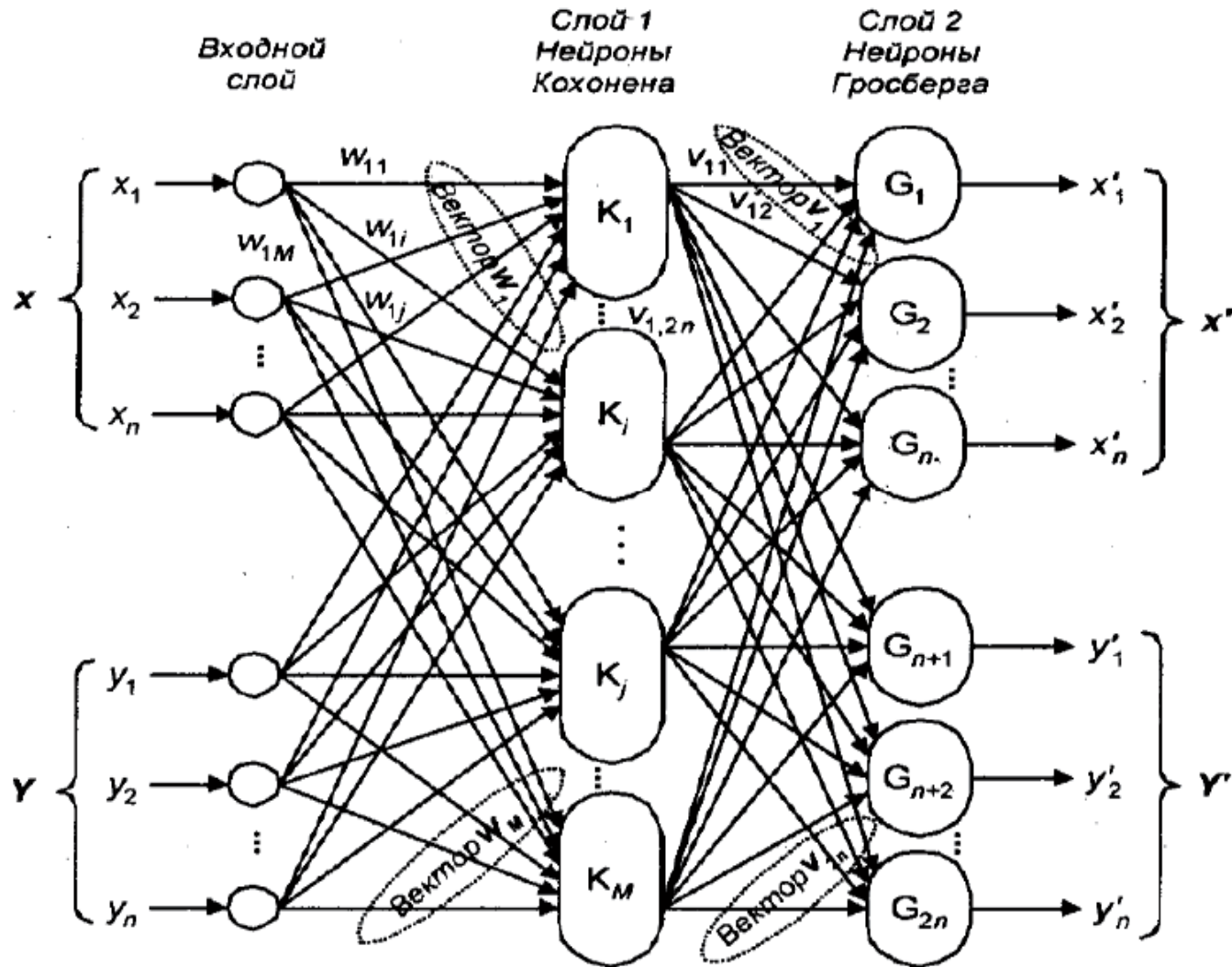
Сеть встречного распространения функционирует подобно столу справок, способному к обобщению. В процессе обучения входные векторы ассоциируются с соответствующими выходными векторами. Эти векторы могут быть двоичными, состоящими из нулей и единиц, или непрерывными. Когда сеть обучена, приложение входного вектора приводит к требуемому выходному вектору. Обобщающая способность сети позволяет получать правильный выход даже при приложении входного вектора, который является неполным или слегка неверным. Это позволяет использовать данную сеть для распознавания образов, восстановления образов и усиления сигналов.

Структура сети встречного распространения

На рисунке показана упрощенная версия прямого действия сети встречного распространения. На нем иллюстрируются ее функциональные свойства.

Сети встречного распространения

Сеть с встречным распространением без обратных связей.



Сети встречного распространения

Введение в сети встречного распространения

Нейроны слоя 0 (показаны кружками) служат лишь точками разветвления и не выполняют вычислений. Каждый нейрон слоя 0 соединен с каждым нейроном слоя 1 (называемого слоем Кохонена) отдельным весом w_{mn} . Эти веса в целом рассматриваются как матрица весов W . Аналогично, каждый нейрон в слое Кохонена (слое 1) соединен с каждым нейроном в слое Гроссберга (слое 2) весом v_{np} . Эти веса образуют матрицу весов V . Отличие этих сетей состоит в операциях, выполняемых нейронами Кохонена и Гроссберга.

Как и многие другие сети, встречное распространение функционирует в двух режимах: в нормальном режиме, при котором принимается входной вектор X и выдается выходной вектор Y , и в режиме обучения, при котором подается входной вектор и веса корректируются, чтобы дать требуемый выходной вектор.

Сети встречного распространения

Функционирование сети встречного распространения

Слой Кохонена. В своей простейшей форме слой Кохонена функционирует в духе «победитель забирает все», т. е. для данного входного вектора один и только один нейрон Кохонена выдает на выходе логическую единицу, все остальные выдают нуль. Нейроны Кохонена можно воспринимать как набор электрических лампочек, так что для любого входного вектора загорается одна из них.

Ассоциированное с каждым нейроном Кохонена множество весов соединяет его с каждым входом. Например, на рисунке нейрон Кохонена K_1 имеет веса $w_{11}, w_{21}, \dots, w_{m1}$, составляющие весовой вектор W_1 . Они соединяются через входной слой с входными сигналами x_1, x_2, \dots, x_m , составляющими входной вектор X . Подобно нейронам большинства сетей выход каждого нейрона Кохонена является просто суммой взвешенных входов.

Сети встречного распространения

Это может быть выражено следующим образом:

$$N_j = w_{1j}x_1 + w_{2j}x_2 + \dots + w_{mj}x_m$$

где N_j - это выход N нейрона Кохонена j , $N_j = \sum_i x_i w_{ij}$ или в векторной записи $\mathbf{N} = \mathbf{XW}$, где \mathbf{N} - вектор выходов N_j слоя Кохонена.

Нейрон Кохонена с максимальным значением N является «победителем». Его выход равен единице, у остальных он равен нулю.

Слой Гроссберга. Слой Гроссберга функционирует в сходной манере. Его выход U является взвешенной суммой выходов k_1, k_2, \dots, k_n слоя Кохонена, образующих вектор \mathbf{K} . Вектор соединяющих весов, обозначенный через \mathbf{V} , состоит из весов $V_{11}, V_{21}, \dots, V_{np}$.

Сети встречного распространения

Тогда выход Y_j каждого нейрона Гроссберга есть

$$Y_j = \sum_i k_i w_{ij}$$

где Y_j - выход j -го нейрона Гроссберга, или в векторной форме $\mathbf{Y} = \mathbf{KV}$, где \mathbf{Y} - выходной вектор слоя Гроссберга; \mathbf{K} - выходной вектор слоя Кохонена; \mathbf{V} - матрица весов слоя Гроссберга.

Если слой Кохонена функционирует таким образом, что лишь у одного нейрона величина выхода равна единице, а у остальных равна нулю, то лишь один элемент вектора \mathbf{K} отличен от нуля, и вычисления очень просты. Фактически каждый нейрон слоя Гроссберга лишь выдает величину веса, который связывает этот нейрон с единственным ненулевым нейроном Кохонена.

Обучение слоя Кохонена. Слой Кохонена классифицирует входные векторы в группы схожих векторов.

Сети встречного распространения

Это достигается с помощью такой подстройки весов слоя Кохонена, что близкие входные векторы активируют один и тот же нейрон данного слоя. Задачей слоя Гроссберга является получение требуемых выходов.

Обучение Кохонена является самообучением, протекающим без учителя. Поэтому трудно (и не нужно) предсказывать, какой именно нейрон Кохонена будет активироваться для заданного входного вектора. Необходимо лишь гарантировать, чтобы в результате обучения разделялись несхожие входные векторы.

Предварительная обработка входных векторов

Весьма желательно (хотя и не обязательно) нормализовать входные векторы перед тем, как предъявлять их сети. Это выполняется с помощью деления каждой компоненты входного вектора на длину вектора. Эта длина находится извлечением квадратного корня из суммы квадратов компонент вектора.

Сети встречного распространения

В алгебраической форме записи имеем

$$x'_i = \frac{x_i}{\sqrt{x_1^2 + x_2^2 + \dots + x_n^2}}$$

Это превращает входной вектор в единичный вектор с тем же самым направлением, т. е. в вектор единичной длины в n-мерном пространстве. Приведенное уравнение обобщает случай двух измерений, когда длина вектора равна гипотенузе прямоугольного треугольника, образованного его x и y компонентами, как это следует из теоремы Пифагора.

При обучении слоя Кохонена на вход подается входной вектор и вычисляются его скалярные произведения с векторами весов, связанными со всеми нейронами Кохонена. Нейрон с максимальным значением скалярного произведения объявляется «победителем» и его веса подстраиваются.

Сети встречного распространения

Так как скалярное произведение является мерой сходства между входным вектором и вектором весов, то процесс обучения состоит в выборе нейрона Кохонена с весовым вектором, наиболее близким к входному вектору, и дальнейшем приближении весового вектора к входному.

Процесс является процессом самообучения, выполняемым без учителя.

Сеть самоорганизуется таким образом, что данный нейрон Кохонена имеет максимальный выход для данного входного вектора.

Уравнение, описывающее процесс обучения имеет следующий вид:

$$w_n = w_c + \alpha(x - w_c),$$

где w_n - новое значение веса, соединяющего входную компоненту x с выигравшим нейроном;

w_c - предыдущее значение этого веса;

α - коэффициент скорости обучения, который может варьироваться в процессе обучения.

Сети встречного распространения

Каждый вес, связанный с выигравшим нейроном Кохонена, изменяется пропорционально разности между его величиной и величиной входа, к которому он присоединен. Направление изменения минимизирует разность между весом и его входом. Сначала находится вектор $x - w_c$. Для этого проводится отрезок из конца w_c в конец x . Затем этот вектор укорачивается умножением его на скалярную величину α , меньшую единицы, в результате чего получается вектор изменения δ . Окончательно новый весовой вектор w_n является отрезком, направленным из начала координат в конец вектора δ . Отсюда можно видеть, что эффект обучения состоит во вращении весового вектора в направлении входного вектора без существенного изменения его длины.

Сети встречного распространения

Переменная k является коэффициентом скорости обучения, который вначале обычно равен $0,7$ и может постепенно уменьшаться в процессе обучения. Это позволяет делать большие начальные шаги для быстрого грубого обучения и меньшие шаги при подходе к окончательной величине.

Если бы с каждым нейроном Кохонена ассоциировался один входной вектор, то слой Кохонена мог бы быть обучен с помощью одного вычисления на вес. Веса нейрона-победителя приравнивались бы к компонентам обучающего вектора ($\alpha = 1$). Как правило, обучающее множество включает много сходных между собой входных векторов, и сеть должна быть обучена активировать один и тот же нейрон Кохонена для каждого из них. В этом случае веса этого нейрона должны получаться усреднением входных векторов, которые должны его активировать.

Сети встречного распространения

Постепенное уменьшение величины α уменьшает воздействие каждого обучающего шага, так что окончательное значение будет средней величиной от входных векторов, на которых происходит обучение. Таким образом, веса, ассоциированные с нейроном, примут значение вблизи «центра» входных векторов, для которых данный нейрон является «победителем».

Выбор начальных значений весовых векторов. Всем весам сети перед началом обучения следует придать начальные значения. Общепринятой практикой является присваивание весам небольших случайных значений. При обучении слоя Кохонена случайно выбранные весовые векторы следует нормализовать. Окончательные значения весовых векторов после обучения совпадают с нормализованными входными векторами. Поэтому нормализация перед началом обучения приближает весовые векторы к их окончательным значениям, сокращая, таким образом, обучающий процесс.

Сети встречного распространения

Рандомизация весов слоя Кохонена может породить серьезные проблемы при обучении, так как в результате ее весовые векторы распределяются равномерно по поверхности гиперсферы. Из-за того, что входные векторы, как правило, распределены неравномерно и имеют тенденцию группироваться на относительно малой части поверхности гиперсферы, большинство весовых векторов будут так удалены от любого входного вектора, что они никогда не будут давать наилучшего соответствия. Эти нейроны Кохонена будут всегда иметь нулевой выход и окажутся бесполезными. Более того, оставшихся весов, дающих наилучшие соответствия, может оказаться слишком мало, чтобы разделить входные векторы на классы, которые расположены близко друг к другу на поверхности гиперсферы.

Сети встречного распространения

Допустим, что имеется несколько множеств входных векторов, все множества сходные, но должны быть разделены на различные классы. Сеть должна быть обучена активировать отдельный нейрон Кохонена для каждого класса. Если начальная плотность весовых векторов в окрестности обучающих векторов слишком мала, то может оказаться невозможным разделить сходные классы из-за того, что не будет достаточного количества весовых векторов в интересующей нас окрестности, чтобы приписать по одному из них каждому классу входных векторов. Наоборот, если несколько входных векторов получены незначительными изменениями из одного и того же образца и должны быть объединены в один класс, то они должны включать один и тот же нейрон Кохонена.

Сети встречного распространения

Если же плотность весовых векторов очень высока вблизи группы слегка различных входных векторов, то каждый входной вектор может активировать отдельный нейрон Кохонена. Это не является катастрофой, так как слой Гроссберга может отобразить различные нейроны Кохонена в один и тот же выход, но это расточительная трата нейронов Кохонена.

Наиболее желательное решение состоит в том, чтобы распределять весовые векторы в соответствии с плотностью входных векторов, которые должны быть разделены, помещая тем самым больше весовых векторов в окрестности большого числа входных векторов. На практике это невыполнимо, однако существует несколько методов приближенного достижения тех же целей.

Сети встречного распространения

Такими методами являются:

- метод выпуклой комбинации;
- метод добавления шума к входным векторам;
- метод случайного распределения начальных весов;
- метод «чувство справедливости».

1. Метод выпуклой комбинации, состоит в том, что все веса приравниваются одной и той же величине $w_i = \frac{1}{\sqrt{n}}$, где n - число входов.

Благодаря этому все весовые векторы совпадают и имеют единичную длину. Каждой же компоненте входа X придается значение $x_i = \alpha x_i + \frac{1 - \alpha}{\sqrt{n}}$.

Вначале α очень мало, вследствие чего все входные векторы имеют длину, близкую к $\frac{1}{\sqrt{n}}$, и почти совпадают с векторами весов.

Сети встречного распространения

В процессе обучения сети α постепенно возрастает, приближаясь к единице. Это позволяет разделять входные векторы и окончательно приписывать им их истинные значения. Весовые векторы отслеживают один или небольшую группу входных векторов и в конце обучения дают требуемую картину выходов. Метод выпуклой комбинации хорошо работает, но замедляет процесс обучения, так как весовые векторы подстраиваются к изменяющейся цели.

2. Добавление шума к входным векторам. При использовании этого метода входные вектора подвергаются случайным изменениям, схватывая в конце концов весовой вектор. Этот метод также работоспособен, но еще более медленен, чем метод выпуклой комбинации.

Сети встречного распространения

3. Случайное распределение начальных весов. Этот метод начинается со случайных весов, но на начальной стадии обучающего процесса подстраивает все веса, а не только связанные с выигравшим нейроном Кохонена. Тем самым весовые векторы перемещаются ближе к области входных векторов. В процессе обучения коррекция весов начинается производиться лишь для ближайших к победителю нейронов Кохонена. Этот радиус коррекции постепенно уменьшается, так что в конце концов корректируются только веса, связанные с выигравшим нейроном Кохонена.
4. «Чувство справедливости». Данный метод наделяет каждый нейрон Кохонена «Чувством справедливости». Если он становится победителем чаще своей законной доли времени (примерно $1/k$, где k - число нейронов Кохонена), он временно увеличивает свой порог, что уменьшает его шансы на выигрыш, давая тем самым возможность обучаться и другим нейронам.

Сети встречного распространения

Во многих приложениях точность результата существенно зависит от распределения весов. К сожалению, эффективность различных решений исчерпывающим образом не оценена и остается проблемой.

Метод интерполяции. До сих пор рассматривался алгоритм обучения, в котором для каждого входного вектора активировался лишь один нейрон Кохонена. Это называется методом аккредитации. Его точность ограничена, так как выход полностью является функцией лишь одного нейрона.

В методе интерполяции целая группа нейронов Кохонена, имеющих наибольшие выходы, может передавать свои выходные сигналы в слой Гроссберга. Число нейронов в такой группе должно выбираться в зависимости от задачи, убедительных данных относительно оптимального размера группы не имеется.

Сети встречного распространения

Метод интерполяции способен устанавливать более сложные соответствия и может давать более точные результаты. По-прежнему, однако, нет убедительных данных, позволяющих сравнить режимы интерполяции и аккредитации.

Статистические свойства обученной сети. Метод обучения Кохонена обладает полезной и интересной способностью извлекать статистические свойства из множества входных данных. Как показано Кохоненом, для полностью обученной сети вероятность того, что случайно выбранный входной вектор (в соответствии с функцией плотности вероятности входного множества) будет ближайшим к любому заданному весовому вектору, равна $1/k$, где k - число нейронов Кохонена. Это является оптимальным распределением весов на гиперсфере.

Сети встречного распространения

Обучение слоя Гроссберга. Слой Гроссберга обучается относительно просто. Входной вектор, являющийся выходом слоя Кохонена, подается на слой нейронов Гроссберга, и выходы слоя Гроссберга вычисляются, как при нормальном функционировании. Далее, каждый вес корректируется лишь в том случае, если он соединен с нейроном Кохонена, имеющим ненулевой выход. Величина коррекции веса пропорциональна разности между весом и требуемым выходом нейрона Гроссберга, с которым он соединен. В символьной записи

$$v_{ijn} = v_{ijc} + \beta(y_j - v_{ijc})k_i,$$

где k - выход i -го нейрона Кохонена (только для одного нейрона Кохонена он отличен от нуля);

y_j - j -ая компонента вектора желаемых выходов.

Сеть адаптивной резонансной теории (ART-сеть)

Определение сети ART

Мозг человека выполняет трудную задачу обработки непрерывного потока сенсорной информации, получаемой из окружающего мира. Из потока тривиальной информации он должен выделить жизненно важную информацию, обработать ее и, возможно, зарегистрировать в долговременной памяти. Понимание процесса человеческой памяти представляет собой серьезную проблему; новые образы запоминаются в такой форме, что ранее запомненные не модифицируются и не забываются, т.е. память остается пластичной и стабильной.

Традиционные искусственные нейронные сети оказались не в состоянии решить проблему стабильности-пластичности. Очень часто обучение новому образу уничтожает или изменяет результаты предшествующего обучения. В некоторых случаях это не существенно.

Сеть адаптивной резонансной теории (ART-сеть)

Если имеется только фиксированный набор обучающих векторов, они могут предъявляться при обучении циклически. В сетях с обратным распространением, например, обучающие векторы подаются на вход сети последовательно до тех пор, пока сеть не обучится всему входному набору. Если, однако, полностью обученная сеть должна запомнить новый обучающий вектор, он может изменить веса настолько, что потребуются полное переобучение сети. В реальной ситуации сеть будет подвергаться постоянно изменяющимся воздействиям; она может никогда не увидеть один и тот же обучающий вектор дважды. При таких обстоятельствах сеть часто не будет обучаться; она будет непрерывно изменять свои веса, не достигая удовлетворительных результатов.

Сеть адаптивной резонансной теории (АРТ-сеть)

В настоящее время известна сеть, в которой только четыре обучающих вектора, предъявляемых циклически, заставляют веса сети изменяться непрерывно, никогда не сходясь. Такая временная нестабильность явилась одним из главных факторов, заставивших Гроссберга и его сотрудников исследовать радикально отличные конфигурации. Адаптивная резонансная теория (АРТ) является одним из результатов исследования этой проблемы.

Сети и алгоритмы АРТ сохраняют пластичность, необходимую для изучения новых образов, в то же время предотвращая изменение ранее запомненных образов. Эта способность стимулировала большой интерес к АРТ, но многие исследователи нашли теорию трудной для понимания.

Сеть адаптивной резонансной теории (ART-сеть)

Сети ART (Adapt ve Resonance Theory) образуют класс различных нейронных сетей, предложенных Карпентером и Гроссбергом (Бостонский университет) в период 1987-1991 гг. Парадигма использует неконтролируемое обучение, анализирует значимые входные данные, выявляет возможные признаки и классифицирует образы в входном векторе.

Сеть ART состоит из двух взаимосвязанных слоев нейронов, расположенных между входным и выходным слоями. Каждый входной образ низшего слоя резонанса стимулирует ожидаемый образ на высшем слое, который пересылается к низшему слою, чтобы влиять на следующий вход. Это создает «резонанс» между низшим и высшим слоями для облегчения сетевой адаптации образов.

Сеть адаптивной резонансной теории (ART-сеть)

Сеть преимущественно используется в биологическом моделировании, тем не менее, существуют некоторые технические применения. Главным ограничением сетевой архитектуры является ее шумовая чувствительность. Даже небольшое количество шума во входном векторе путает обобщающие возможности наученной сети.

На практике данные, используемые для обучения или самообучения сети, часто нестабильны. Например, если на вход обычной нейронной сети с прямыми связями, обучаемую с помощью алгоритма с обратным распространением ошибки, подать образ такого класса, который не был представлен в обучающей последовательности или во множестве образов, подлежащих автоматической классификации или кластеризации. Здесь мы сталкиваемся с двумя противоречивыми требованиями или свойствами нейронной сети.

Сеть адаптивной резонансной теории (ART-сеть)

С одной стороны очень важно, чтобы она была способна выявлять (обнаруживать) образы новых классов, ранее не представленных сети. Это свойство пластичности. С другой же стороны изученные классы образов должны сохраняться - свойство устойчивости нейронных сетей. Эти два свойства - пластичности и стабильности в известной мере противоречивы - дилемма пластичности-стабильности. Сети ART и были разработаны для разрешения этой дилеммы, а именно: установление новых ассоциаций (классов) НС без забывания старых ассоциаций (классов).

Семейство ART-сетей включает:

- ART-1: для бинарных входных векторов, когда признаки распознаваемых образов принимают два значения 1 или 0;
- ART-2: расширение ART-1-сетей на непрерывные входные векторы;

Сеть адаптивной резонансной теории (ART-сеть)

- ART-2a: оптимальная версия ART-2-сетей, отличающаяся повышенной скоростью сходимости;
- ART-3: моделирование временных и химических процессов (биологических механизмов) на базе ART-2;
- ARTMAP: комбинация двух ART-сетей (например, ART-1 и ART-2);
- Fuzz(ART: гибридная сеть, объединяющая нечеткую логику (Fuzz(Log k) и ART сети.

Архитектура сети ART-1

Сеть ART-1 реализует алгоритм кластеризации, очень похожий на алгоритм «последовательного лидера». Следуя этому алгоритму первый входной сигнал считается образцом первого кластера. Следующий входной сигнал сравнивается с образцом первого кластера.

Сеть адаптивной резонансной теории (ART-сеть)

Говорят, что входной сигнал «направляется за лидером» и принадлежит первому кластеру, если расстояние до образца первого кластера меньше порога. В противном случае второй входной сигнал-образец второго кластера. Этот процесс повторяется для всех следующих входных сигналов. Таким образом, число кластеров растет с течением времени и зависит как от значения порога, так и от метрического свидетельства расстояния, используемое для сравнения входных сигналов и образцов классов.

Основная часть сети ART-1 похожа на сеть Хемминга. С помощью последовательных связей вычисляется соответствие входных сигналов и образцов кластеров. Максимальное значение соответствия усиливается с помощью латеральных связей выходных нейронов.

Сеть адаптивной резонансной теории (ART-сеть)

Сеть ART-1 отличается от сети Хемминга обратными связями от выходных нейронов к входным, кроме того есть возможность выключать выходной нейрон с максимальным значением соответствия и проводить тестирования соответствия входного сигнала и образцов кластеров, как того требует алгоритм «последовательного лидера».

Принцип работы ART-сетей сравнительно прост. При вводе значений признаков некоторого образа ART-1-сеть пытается сопоставить ему некоторый класс из числа уже изученных. Если такой класс удастся найти, то производится сравнительно небольшая модификация прототипа (стереотипа, типичного представителя) этого класса для того, чтобы он хорошо отображал и новый образ. Классификация образа на этом заканчивается. Если же такой класс найти не удастся, то образуется (вводится) новый класс.

Сеть адаптивной резонансной теории (ART-сеть)

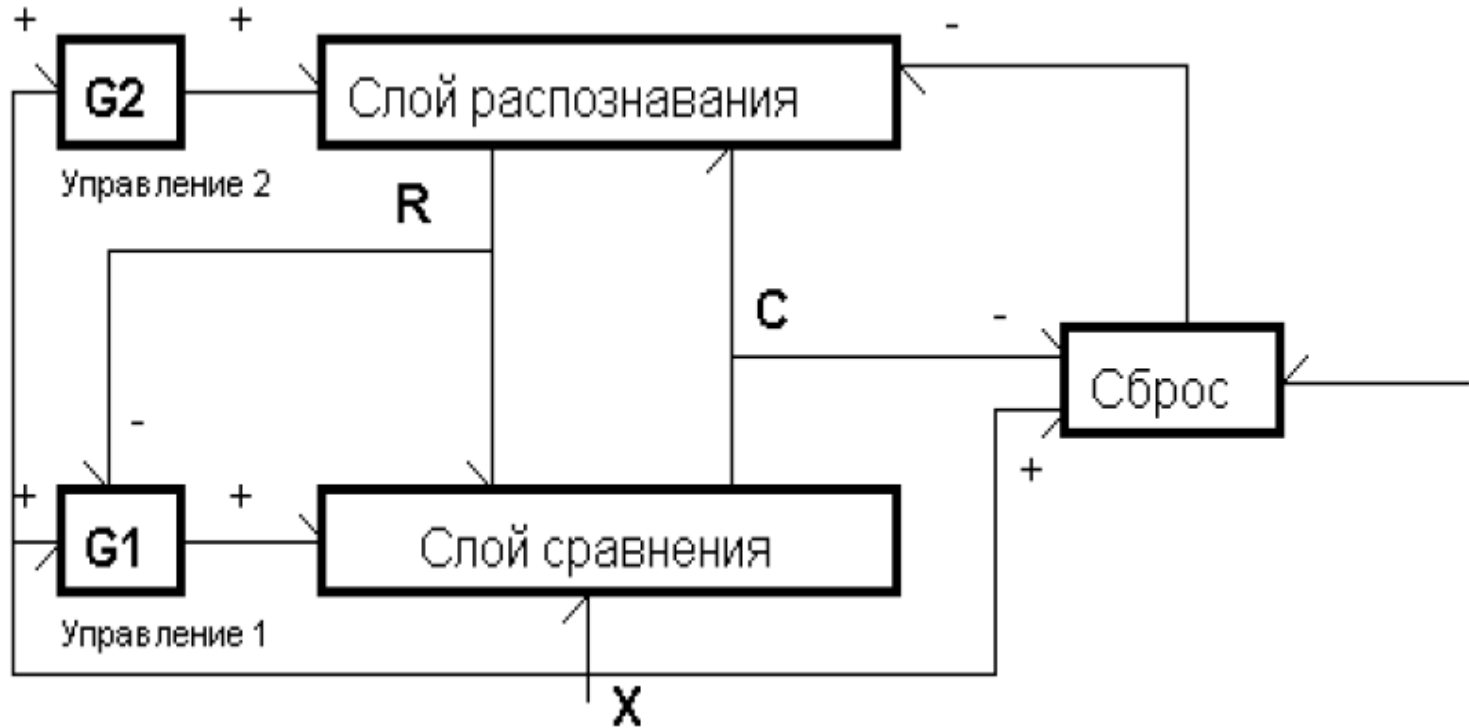
При этом предъявленный образ несколько модифицируется и используется затем в качестве прототипа (стереотипа, типичного представителя) для нового класса. При этом уже изученные классы не изменяются.

Сеть ART-1 состоит из пяти функциональных модулей (рисунок): двух слоев нейронов - слоя сравнения и слоя распознавания, и трех управляющих специализированных нейронов - сброса, управления 1 и управления 2.

Слой сравнения. Слой сравнения получает двоичный входной вектор X и первоначально пропускает его неизменным для формирования выходного вектора C . На более поздней фазе в распознающем слое вырабатывается двоичный вектор R , модифицирующий вектор C , как описано ниже.

Сеть адаптивной резонансной теории (ART-сеть)

Общая схема нейронной сети ART-1.



Каждый нейрон в слое сравнения получает три двоичных входа (0 или 1): (1) компонента x входного вектора X ; (2) сигнал обратной связи R - взвешенная сумма выходов распознающего слоя; (3) вход от G_1 (один и тот же сигнал подается на все нейроны этого слоя).

Сеть адаптивной резонансной теории (ART-сеть)

Чтобы получить на выходе нейрона единичное значение, как минимум два из трех его входов должны равняться единице; в противном случае его выход будет нулевым. Первоначально выходной сигнал G_1 Приемника 1 установлен в единицу, обеспечивая один из необходимых для возбуждения нейронов входов, а все компоненты вектора R установлены в 0; следовательно, в этот момент вектор S идентичен двоичному входному вектору X .

Слой распознавания. Слой распознавания осуществляет классификацию входных векторов. Каждый нейрон в слое распознавания имеет соответствующий вектор весов V_j . Только один нейрон с весовым вектором, наиболее соответствующим входному вектору, возбуждается; все остальные нейроны заторможены.

Сеть адаптивной резонансной теории (ART-сеть)

Как показано на рисунке, нейрон в распознающем слое имеет максимальную реакцию, если вектор C , являющийся выходом слоя сравнения, соответствует набору его весов, следовательно, веса представляют запомненный образ или экземпляр для категории входных векторов. Эти веса являются действительными числами, а не двоичными величинами. Двоичная версия этого образа также запоминается в соответствующем наборе весов слоя сравнения; этот набор состоит из весов связей, соединяющих определенные нейроны слоя распознавания, один вес на каждый нейрон слоя сравнения.

В процессе функционирования каждый нейрон слоя распознавания вычисляет свертку вектора собственных весов и входного вектора C . Нейрон, имеющий веса, наиболее близкие вектору C , будет иметь самый большой выход, тем самым выигрывая соревнование и одновременно затормаживая все остальные нейроны в слое.

Сеть адаптивной резонансной теории (ART-сеть)

Сброс. Модуль сброса измеряет величину сходства между векторами X и C . Если эта величина больше порогового значения, вырабатывается сигнал сброса возбужденного нейрона в слое распознавания.

В процессе функционирования модуль сброса вычисляет сходство как отношение количества единиц в векторе C к их количеству в векторе X . Если это отношение ниже значения параметра сходства, вырабатывается сигнал сброса.

Начальное значение нейрона управления 1 полагается равным единице: $G_1 = 1$. Входной двоичный вектор X поступает на слой сравнения, который первоначально пропускает его без изменения, при этом выходной вектор слоя сравнения $C = X$. Это достигается применением так называемого правила 2/3 для нейронов слоя сравнения.

Сеть адаптивной резонансной теории (ART-сеть)

Каждый из нейронов этого слоя имеет три двоичных входа - сигнал от соответствующей компоненты вектора X , сигнал от нейрона управления 1 и сигнал обратной связи из слоя распознавания P (который в начальный момент равен нулю).

Для активации нейрона в слое сравнения требуется, чтобы по крайней мере два из трех сигналов были равны единице, что и достигается в начальный момент входом от управления 1 и активными компонентами вектора X .

Выработанный слоем сравнения сигнал C поступает на входы нейронов слоя распознавания. Каждый нейрон слоя распознавания имеет вектор весов b_j - действительных чисел, при этом возбуждается только один нейрон этого слоя, вектор весов которого наиболее близок к C . Это может быть достигнуто, например, за счет механизма латерального торможения типа «Победитель забирает все».

Сеть адаптивной резонансной теории (ART-сеть)

Выход нейрона-победителя устанавливается равным единице, остальные нейроны полностью заторможены. Сигнал обратной связи от нейрона-победителя поступает обратно в слой сравнения через синаптические веса T . Вектор T , по существу, является носителем критических черт категории, определяемой выигравшим нейроном.

Выход нейрона управления 1 равен единице, только когда входной образ X имеет ненулевые компоненты, то есть этот нейрон выполняет функцию детекции факта поступления образа на вход. Однако, когда возникает ненулевой отклик нейронов слоя распознавания R , значение управления 1 обнуляется $G_1 = 0$.

Сигнал нейрона управления 2 также устанавливается на единицу при ненулевом векторе X . Задачей этого нейрона является погашение активности на слое распознавания, если в сеть не поступило никакой информации.

Сеть адаптивной резонансной теории (ART-сеть)

Итак, при генерации отклика R слоя распознавания выход $G_1 = 0$, и теперь нейроны слоя сравнения активируются сигналами образа X и отклика R . Правило двух третей приводит к активации только тех нейронов слоя сравнения, для которых X и R являются единичными. Таким образом, выход слоя сравнения S теперь уже не равен в точности X , а содержит лишь те компоненты X , которые соответствуют критическим чертам победившей категории. Этот механизм в теории ART получил название адаптивной фильтрации образа X .

Теперь задачей системы является установить, достаточен ли набор этих критических черт для окончательного отнесения образа X к категории нейрона-победителя. Эту функцию осуществляет нейрон сброса, который измеряет сходство между векторами X и S .

Сеть адаптивной резонансной теории (ART-сеть)

Выход нейрона сброса определяется отношением числа единичных компонент в векторе C к числу единичных компонент исходного образа X . Если это отношение ниже некоторого определенного уровня сходства, нейрон выдает сигнал сброса, означающий что уровень резонанса образа X с чертами предлагаемой категории не достаточен для положительного заключения о завершении классификации. Условием возникновения сигнала сброса является соотношение $\frac{\|C\|}{\|X\|} < \rho$, где $\rho < 1$ - параметр сходства.

Сигнал сброса выполняет полное торможение нейрона-победителя-неудачника, который не принимает в дальнейшем участия в работе сети.

Сеть адаптивной резонансной теории (ART-сеть)

Принцип работы ART-сети

В ART-1-сетях различают следующие пять фаз внутренней обработки информации.

1. Инициализация сети. В начале инициализируются обе весовые матрицы, а также параметр толерантности.
2. Распознавание. На этой фазе для входного вектора или для вектора сравнения, определенного на основе вектора входного вектора определяется наиболее близкий класс.
3. Сравнение. Производится сравнение ожидаемого вектора со входным вектором или вектором сравнения. При слишком малом совпадении векторов осуществляется повторное распознавание.
4. Поиск. Производится поиск альтернативного класса или же при необходимости открывается новый класс.

Сеть адаптивной резонансной теории (ART-сеть)

5. Адаптация весов. На этой стадии осуществляется модификация обеих весовых матриц.

Опишем последовательно события, происходящие в сети ART в процессе классификации.

Инициализация. Перед началом процесса обучения сети все весовые векторы V_j и T_j , а также параметр сходства (толерантности) ρ , должны быть установлены в начальные значения.

Веса векторов V_j все инициализируются в одинаковые малые значения. Эти значения должны удовлетворять условию

$$b_{ij} < \frac{L}{L-1+m} \quad \text{для всех } i, j,$$

где m - количество компонент входного вектора,

L - константа, большая 1 (обычно $L=2$).

Сеть адаптивной резонансной теории (ART-сеть)

Эта величина является критической; если она слишком большая, сеть может распределить все нейроны распознающего слоя одному входному вектору.

Параметр сходства ρ устанавливается в диапазоне от 0 до 1 в зависимости от требуемой степени сходства между запомненным образом и входным вектором. При высоких значениях сеть относит к одному классу только очень слабо отличающиеся образы. С другой стороны, малое значение заставляет сеть группировать образы, которые имеют слабое сходство между собой. Может оказаться желательной возможность изменять коэффициент сходства на протяжении процесса обучения, обеспечивая только грубую классификацию в начале процесса обучения, и затем постепенно увеличивая коэффициент сходства для выработки точной классификации в конце процесса обучения.

Сеть адаптивной резонансной теории (ART-сеть)

Фаза распознавания. Нулевые значения компонент входного вектора X устанавливают сигнал нейрона управления 2 в нуль, одновременно устанавливая в нуль выходы нейронов слоя распознавания. При возникновении ненулевых значений X , оба сигнала управления (G_1 и G_2) устанавливаются равными единице. При этом по правилу двух третей выходы нейронов слоя сравнения C в точности равны компонентам X . Вектор C поступает на входы нейронов слоя распознавания, которые в конкурентной борьбе определяют нейрон-победитель, описывающий предполагаемый результат классификации. В итоге выходной вектор R слоя распознавания содержит ровно одну единичную компоненту, остальные значения равны нулю. Ненулевой выход нейрона-победителя устанавливает в нуль сигнал управления 1 : $G_1=0$. По обратной связи нейрон-победитель посылает сигналы в слой сравнения, и начинается фаза сравнения.

Сеть адаптивной резонансной теории (ART-сеть)

Фаза сравнения. В слое сравнения вектор сигналов отклика слоя распознавания сравнивается с компонентами вектора X . Выход слоя сравнения C теперь содержит единичные компоненты только в тех позициях, в которых единицы имеются и у входного вектора X и у вектора обратной связи P . Если в результате сравнения векторов C и X не будет обнаружено значительных отличий, то нейрон сброса остается неактивным. Вектор C вновь вызовет возбуждение того же нейрона-победителя в слое распознавания, что и удачно завершит процесс классификации. В противном случае будет выработан сигнал сброса, который затормозит нейрон-победитель в слое распознавания, и начнется фаза поиска.

Фаза поиска. В результате действия тормозящего сигнала сброса все нейроны слоя распознавания получают нулевые выходы, и, следовательно, нейрон управления 1 примет единичное значение активности.

Сеть адаптивной резонансной теории (ART-сеть)

Снова выходной сигнал слоя сравнения S установится равным в точности X , как и в начале работы сети. Однако теперь в конкурентной борьбе в слое распознавания предыдущий нейрон-победитель не участвует, и будет найдена новая категория - кандидат. После чего опять повторяется фаза сравнения.

Итерационный процесс поиска завершается двумя возможными способами.

1. Найдется запомненная категория, сходство которой с входным вектором X будет достаточным для успешной классификации. После этого происходит обучающий цикл, в котором модифицируются веса b_i и t_i векторов B и T возбужденного нейрона, осуществившего классификацию.
2. В процессе поиска все запомненные категории окажутся проверенными, но ни одна из них не дала требуемого сходства.

Сеть адаптивной резонансной теории (ART-сеть)

В этом случае входной образ X объявляется новым для нейронной сети, и ему выделяется новый нейрон в слое распознавания. Весовые вектора этого нейрона V и T устанавливаются равными вектору X .

Важно понимать, почему требуется фаза поиска и окончательный результат классификации не возникает с первой попытки. Обучение и функционирование сети ART происходит одновременно. Нейрон-победитель определяет в пространстве входных векторов ближайший к заданному входному образу вектор памяти, и если бы все черты исходного вектора были критическими, это и было бы верной классификацией. Однако множество критических черт стабилизируется лишь после относительно длительного обучения. На данной фазе обучения лишь некоторые компоненты входного вектора принадлежат актуальному множеству критических черт, поэтому может найтись другой нейрон-классификатор, который на множестве критических черт окажется ближе к исходному образу.

Сеть адаптивной резонансной теории (ART-сеть)

Он и определяется в результате поиска. Отметим, что после относительной стабилизации процесса обучения классификация выполняется без фазы поиска. В этом случае говорят, что формируется прямой доступ к памяти. Возникновение в процессе обучения прямого доступа доказывается в теории ART.

Обучение сети ART. Обучение представляет собой процесс, в котором набор входных векторов подается последовательно на вход сети и веса сети изменяются при этом таким образом, чтобы сходные векторы активизировали соответствующие нейроны. Заметим, что это - неуправляемое обучение, нет учителя и нет целевого вектора, определяющего требуемый ответ.

Различают два вида обучения: медленное и быстрое.

Сеть адаптивной резонансной теории (ART-сеть)

При медленном обучении входной вектор предъявляется настолько кратко временно, что веса сети не имеют достаточного времени для достижения своих асимптотических значений в результате одного предъявления.

В этом случае значения весов будут определяться скорее статистическими характеристиками входных векторов, чем характеристиками какого-то одного входного вектора.

Динамика сети в процессе медленного обучения описывается дифференциальными уравнениями.

Быстрое обучение является специальным случаем медленного обучения, когда входной вектор прикладывается на достаточно длительный промежуток времени, чтобы позволить весам приблизиться к их окончательным значениям. В этом случае процесс обучения описывается только алгебраическими выражениями.

Сеть адаптивной резонансной теории (ART-сеть)

Кроме того, компоненты весовых векторов T_j принимают двоичные значения, в отличие от непрерывного диапазона значений, требуемого в случае быстрого обучения. Будем рассматривать только быстрое обучение. Процесс обучения происходит без учителя, на основе самоорганизации. Обучение производится для весов нейрона-победителя в случае как успешной, так и неуспешной классификации. При этом веса вектора B стремятся к нормализованной величине компонент вектора C :

$$b_i = \frac{Lc_i}{L-1 + \sum_k c_k}$$

При этом роль нормализации компонент крайне важна. Вектора с большим числом единиц приводят к небольшим значениям весов B , и наоборот. Таким образом, произведение

$$BC = \sum_i b_i c_i$$

оказывается масштабированным.

Сеть адаптивной резонансной теории (ART-сеть)

Операция масштабирования приводит к тому, что возможно правильное различение векторов, даже если один является подмножеством другого.

Пример. Пусть нейрон X_1 соответствует образу (100000), а нейрон X_2 - образу (111100). Эти образы являются, очевидно, различными. При обучении без нормализации при поступлении в сеть первого образа, он даст одинаковые скалярные произведения, равные 1, как с весами нейрона X_1 , так и X_2 . Нейрон X_2 в присутствии небольших шумовых отклонений в значениях весов, может выиграть конкуренцию. При этом веса его вектора T становятся равными (100000), и образ (111100) будет безвозвратно «забыт» сетью.

При применении нормализации исходные скалярные произведения будут равны 1 для нейрона X_1 , и значению $2/5$ для нейрона X_2 (при $L=2$). Тем самым, нейрон X_1 выигрывает конкурентное соревнование.

Сеть адаптивной резонансной теории (ART-сеть)

Рассмотрим основные теоремы теории ART.

1. После стабилизации процесса обучения предъявление одного из обучающих векторов (или вектора с существенными характеристиками категории) будет активизировать требуемый нейрон слоя распознавания без поиска. Эта характеристика «прямого доступа» определяет быстрый доступ к предварительно изученным образам.
2. Процесс поиска является устойчивым. После определения выигравшего нейрона в сети не будет возбуждений других нейронов в результате изменения векторов выхода слоя сравнения C ; только сигнал сброса может вызвать такие изменения.
3. Процесс обучения является устойчивым. Обучение не будет вызывать переключения с одного возбужденного нейрона слоя распознавания на другой.

Сеть адаптивной резонансной теории (АРТ-сеть)

4. Процесс обучения конечен. Любая последовательность произвольных входных векторов будет производить стабильный набор весов после конечного количества обучающих серий; повторяющиеся последовательности обучающих векторов не будут приводить к циклическому изменению весов.

Нерешенные проблемы и недостатки АРТ-1

Сети АРТ, при всех их замечательных свойствах, имеют ряд недостатков. Одним из них является большое количество синаптических связей в сети в расчете на единицу запоминаемой информации. При этом многие из весов оказываются после обучения нулевыми. Эту особенность следует учитывать при аппаратных реализациях.

Сеть АРТ-1 приспособлена к работе только с битовыми векторами. Это неудобство преодолевается в сетях АРТ-2 и АРТ-3.

Сеть адаптивной резонансной теории (АРТ-сеть)

Однако в этих архитектурах, равно как и в АРТ-1, сохраняется главный недостаток АРТ - локализованность памяти. Память нейронной сети АРТ не является распределенной, некоторой заданной категории отвечает вполне конкретный нейрон слоя распознавания. При его разрушении теряется память обо всей категории. Эта особенность не позволяет говорить о сетях адаптивной резонансной теории, как о прямых моделях биологических нейронных сетей. Память АРТ сетей является распределенной, а не локализованной. Основная особенность АРТ-1-сетей состоит в способности к формированию новых классов в процессе обучения или иначе в разрешении дилеммы стабильности- пластичности.

Разновидности АРТ-сетей

Приведем краткий обзор сетей АРТ-2, АРТ-2а, АРТ-3, АРТМАР и FUZZY-АРТ.

Сеть адаптивной резонансной теории (АРТ-сеть)

Сети АРТ-2. Основной отличительной чертой нейронной сети АРТ-2 является возможность работы с аналоговыми векторами и сигналами. По сравнению с АРТ-1 в архитектуре сети сделаны некоторые изменения, позволяющие отдельным подсистемам функционировать асинхронно, что принципиально для аппаратных реализаций. Важным отличием аналоговых сигналов от битовых является принципиальная возможность аналоговых векторов быть сколь угодно близкими друг к другу (в то время как пространство битовых векторов дискретно). Это накладывает дополнительные требования на функционирование нейронов слоя сравнения - требуется более тонкий и чувствительный механизм для выделения областей резонанса. Общим решением здесь является переход к многослойной архитектуре, с все более точной настройкой при переходе от слоя к слою, что и применено в АРТ-2. Функционирование слоя распознавания принципиально не изменяется.

Сеть адаптивной резонансной теории (АРТ-сеть)

Сети АРТ-2 успешно применяются для распознавания движущихся изображений. Поскольку нейросистемы АРТ не содержат механизма инвариантного распознавания, то в сочетании с ними применяются специализированные (часто не нейросетевые) системы инвариантного представления образов, например двумерное преобразование Фурье, или более сложные алгоритмы.

Сети АРТ-3. Следующим шагом в развитии АРТ явилась сеть АРТ-3. Особенности обучения нейронов сетей АРТ-1 и АРТ-2 не позволяют использовать эти сети, как элементы более крупных иерархических нейросистем, в частности, компоновать из них многослойные сети. Это затрудняет представление в АРТ иерархически организованной информации, что характерно для систем восприятия человека и животных.

Сеть адаптивной резонансной теории (ART-сеть)

Эти проблемы решены в сети ART-3, которая выступает как многослойная архитектура. При переходе от слоя к слою происходит контрастирование входных образов и запоминание их в виде все более общих категорий. При этом основной задачей каждого отдельного слоя является сжатие входящей информации.

Образ входит в адаптирующийся резонанс между некоторой парой слоев, в дальнейшем этот резонанс распространяется на следующие слои иерархии. В ART-1 и ART-2 недостаточный уровень резонанса приводил к генерации сигнала сброса, что приводило к полному торможению слоя распознавания. В случае многослойной сети ART-3 это недопустимо, так как это разрывает поток информации. Поэтому в ART-3 введен специальный механизм зависимости активности синапсов обратных связей от времени, аналогичный рефрактерному торможению биологического нейрона после передачи возбуждения.

Сеть адаптивной резонансной теории (ART-сеть)

Поэтому вместо полного сброса сигнала происходит торможение синаптических сигналов обратной связи, и слой сравнения получает исходное состояние возбуждения для выполнения фазы поиска нового резонанса.

ARTMAP. Они объединяют элементы обучения и самообучения (или обучения с поощрением и без поощрения, Supervised and Unsupervised Training). Для этого обычно формируется комбинация из двух ART-сетей.

FUZZY-ART-сети (нечеткие ART-сети). Они представляют собой прямое расширение ART-1-сетей средствами нечеткой логики.

В них применяются следующие операторы:

- для определения класса образов в слое распознавания;
- для расчета степени сходства (Reset-критерий);
- для адаптации весов.

Сеть адаптивной резонансной теории (ART-сеть)

Таким образом, сети ART являются интересным и перспективным видом нейронных систем. Они способны решить дилемму стабильности-пластичности, а их архитектура сконструирована по принципу биологического подобия. Однако они могут оказаться не в состоянии моделировать распределенную память, которую многие рассматривают как важную характеристику функций мозга. Слабым местом ART является то, что разрушение отдельных элементов ART разрушает всю память. Память мозга, напротив, распределена по веществу мозга, запомненные образы могут часто пережить значительные физические повреждения мозга без полной их потери.

Использованные материалы

В презентации использованы материалы, предоставленные:
Бутырским Евгением Юрьевичем, доктором физико-математических наук, профессором кафедры теории управления СПбГУ;
Гришкиным Валерием Михайловичем, кандидат технических наук, доцентом кафедры компьютерного моделирования и многопроцессорных систем.