Жетимекова Г.Ж.

РАЗРАБОТКА МОДЕЛИ РАСПОЗНАЮЩЕЙ НЕЙРОСЕТИ БЕЗ ОБУЧЕНИЯ

G.Zh. Zhetimekova

DEVELOPMENT OF A MODEL RECOGNIZING NEURAL NETWORKS WITHOUT TRAINING

УДК: 681.5

В статье рассматривается гибридная модель распознающей нейросетей без обучения на основе принципа покомпонентного анализа (PCA) и самоорганизующей карты Кохенена. Эта сеть применяется для распознавания и классификации как статичных образов, так и образов, подвергшихся определенным трансформациям.

Ключевые слова: *кластеризация, классификация, нечеткая логика, нечеткая система, анализ данных, нейросеть.*

The article deals with the hybrid model of the recognizing neural networks without training based on the principle of component-based analysis (PCA) and the self-organizing map of Cohenen. This network is used to recognize and classify both static images and images that underwent certain transformations.

Key words: *clustering, classification, fuzzy logic, fuzzy system, data analysis, neural network.*

Нейронные сети представляют собой мощный инструмент в искусственном интеллекте. Область их применения очень широкая, начиная от распознавания образов, заканчивая теорией игр и моделирования человеческого мозга. Существуют множество видов нейронной сетей: персептронные, обратного распространения, карты Кохенена, и другие. Виды нейронных сетей зависят от их природы, где существуют распознавание с обучением т.е. с учителем и распознавание без обучения. Данная статья посвящена распознаванию без обучения с помощью нейронных сетей. Сеть должна обнаружить для себя образцы, признаки, регулярность, корреляции или категории во входных данных.

Тип образца, в котором сеть без обучения обнаруживает во входных данных, зависит от архитектуры сети. Архитектура сети без обучения довольно проста, но осложнения и тонкость приходят главным образом от правил обучения. В этом случае, в отличие от сетей с обучением и с учителем, процесс распознавания совмещен с процессом обучения. Некоторые из известных сетей состоят лишь из единственного слоя, а известной многослойной сетью является сеть с прямой связью адаптивной теории резонанса (АТР).

Обычно используемый метод статистик для анализа данных использует принцип покомпонентного анализа (PCA). Извлекаемые признаки согласно принципу покомпонентного анализа неизменны. Это становится полезным (PCA) для извлечения собственных векторов каждого образца для сокращения размерности в описании исходных образов. Sanger и Oja оба разработали сети с прямой связью с одним уровнем, извлекающие первые M основных компонент входных данных с помощью модификации правила плоскости Hebbian.

В качестве примера используются встречающиеся формы цифр, для того чтобы их классифицировать (распознать). В противовес к подходам, основанным на нейронной сети, существуют другие методы, такие как методы типа принцип покомпонентного анализа (PCA) и типа пар линейной классификации. Для того чтобы оценить результаты с результатами из нашей нейронной сети, мы также осуществили PCA в двух направлениях: PCA был применен как для строк исходного образца, так и PCA был применен для столбцов для уменьшения размерности образца. Новым образцом, при вдвое уменьшенной размерности, стал сокращенный набор признаков входного образца для распознавания.

Распознавание образов, в известном смысле, независимо от позиции, размера, яркости и ориентации в визуальном пространстве было целью обширного текущего исследования. Для достижения максимальной утилизации и гибкости, используемые методы должны быть нечувствительны к небольшим изменениям форм и обеспечивать лучшее выполнение при повторных испытаниях. Форма граничных сегментов может быть описана количественно с использованием моментов. Другие методы включают кодирование цепочки, используемое Freeman, Фурье, многоугольное приближение, предложенное Pavlidis, и синтаксическое распознавание образов.

Моментные инварианты стали классическим инструментом для объектного распознавания в течение прошедших 30 лет. Существуют различные подходы к теоретическому образованию инвариантов вращения на основе моментов. Hu использовал результаты теории алгебраических инвариантов и получил семь известных инвариантов вращения двумерных объектов. Teague и Wallin предложили использовать моменты Zernike, а Wong использовал сложные одночлены, которые также происходят из теории алгебраических инвариантов, Ли применял преобразования Фурье-Меллина.

Преобразования Фурье-Меллина были введены для регистрации образов с нарушениями границ вследствие трансляции (т.е. сдвиг, перенос), вращения и масштабирования. Этот метод использует преобразования Фурье для трансляции изображений. Затем лог-полярное преобразование применяется к спектру величин, а вращение и масштабирование возвращается с использованием корреляции в лог-полярном пространстве. В этом методе используется тот факт, что, оперируя спектром величин изображений, можно избежать различий перевода, поскольку спектр образов и его оттранслированная счетная часть идентичны, различен только их фазовый спектр. Кроме того, лог-полярное преобразование является причиной вращения и масштабирования, которые образуют трансляцию и посредством которых может быть применена корреляция стадии для возвращения угла вращения и коэффициента масштабирования между парой входных образов.

Для достижения цели строится гибридная модель распознающей нейросетей без «обучения» на основе принципа покомпонентного анализа (PCA) и самоорганизующей карты Кохенена (рис. 1), используемая как для инвариантных (неизменяемых) так и для вариантных (изменяемых) образов.

На основе гибридной модели удается построить распознающий алгоритм, включающий два компонента:

1. Алгоритм без обучения Sanger Правило (PCA).

2. Распознавание образов с использованием сети Кохенена.

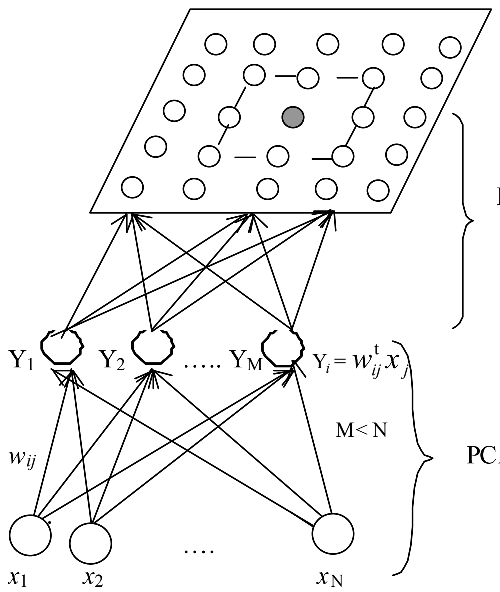

Рис. 1. Гибридная модифицированная модель распознающей нейросети.

В теории связей PCA известен как Karhunen-Loeve преобразование. Суть его состоит в том, чтобы найти набор *M* ортогональных векторов в пространстве данных, которые вычисляются в максимально возможной степени дисперсии данных. Проектирование данных от их первоначального *N*-мерного пространства на *M*-мерное подпространство, заполненное данными векторами, выполняет сокращение размерности, в общем случае $M<N$, что часто сохраняет большинство встроенной информации в данных.

Алгоритм состоит из следующих шагов [Press et. al., 1992; Bishop, 1996]:

1) Вычислить среднее из всех входов в данные, и затем его вычесть:

$$\mu = E(x) = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad (1)$$

2) Вычислить матрицу ковариации и ее собственные векторы λ и собственные значения w

$$C = E((x-\mu)(x-\mu)^t). \qquad (2)$$

$$Cw = \lambda w. \qquad (3)$$

$$\lambda = w^t C w \qquad (4)$$

3) Сохранить собственные векторы, соответствующие M наибольших собственных значений $\lambda^1 \geq \lambda^2 \geq \ \dots \ \geq \lambda^n$.

4) Спроектировать входные векторы на собственные векторы.

Нейронные сети могут быть использованы для изменения размерности. Двухуровневый perceptron с линейными устройствами вывода (количество скрытых модулей равно M, M<d) обучен отображать входные векторы на себя при помощи минимизации ошибки суммы квадратов и способен выполнить линейный анализ основного компонента. Если два дополнительных нелинейных скрытые уровни позволено поместить в сеть, то сеть может быть сделана так, чтобы выполнить нелинейный принцип покомпонентного анализа. В общем случае может быть показано, что k-ое направление основного компонента проходит вдоль направления собственного вектора, принадлежащего k-ому наибольшему собственному значению матрицы полной ковариации входного вектора. Принцип покомпонентного анализа (PCA) использует Hebbian сеть с двумя уровнями для того, чтобы вычислить собственные векторы и собственные значения входных образцов. Фактическое количество собственных значений определяется путем сравнения первых k наибольших собственных значений со среднеквадратичным отклонением вводов.

Выбор или извлечение признака является процессом выбора карты формы y= f(x), при котором выборка $x(x_1, x_2, \dots x_n)$ в n-мерном измеренном пространстве $\Re^n$ преобразована в точку $y(y_1, y_2, \dots, y_q)$ в q-мерном (q<n) $\Re^q$ пространстве признака. Основная цель данной задачи состоит в том, чтобы сохранить оптимальные существенные характеристики, необходимые для процесса распознавания и уменьшения размерности пространства так, чтобы была возможна эффективная классификация. Самый общий метод уменьшения размерности входного пространства – принцип покомпонентного анализа (PCA).

Существует схожее соответствие между поведением самоорганизующихся нейронных сетей и статистического метода принципа покомпонентного анализа. Фактически, самоорганизация, полностью связанная с двухуровневой сетью с i входами, и j выходами (i> j) использоваться для извлечения первых j основных компонентов из входного вектора, таким образом, сокращая размер входного вектора на j-i элементов. При j=1, такая сеть действует как максимум собственного фильтра (eigenfilter), извлекающий первый основной компонент из входного вектора.

Принцип покомпонентного анализа (PCA), использующий Hebbian двухуровневые сети показан на рисунке 2. Узлы входа и выхода используют линейные функции активации и полностью взаимосвязаны соединениями Hebbian. Алгоритм неконтролируемого Hebbian обучения извлекает первые j основных компонентов из входного вектора x, используя вектор веса w. Окончательный выходной вектор y, может затем использоваться в качестве входного вектора для регулярной сети с прямой связью.

Сеть без обучения по правило Hebbian с помощью регулирования весов определена в соответствии с уравнением (5)

$$\Delta w_i = \eta Y x_i \quad (5)$$

где η управляет скоростью обучения.

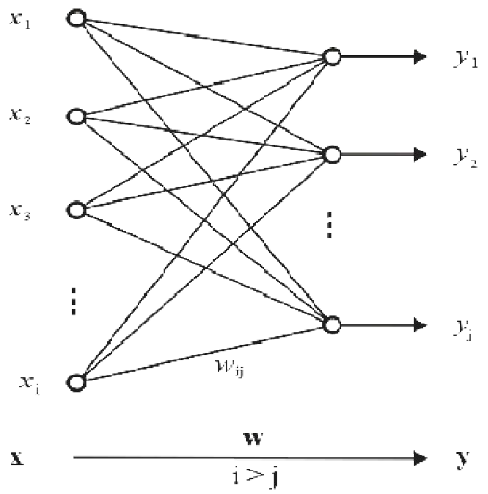

Рис. 2. Hebbian сеть с 2-х уровнями.

Предложено множество подходов для изменения правила Hebbian.

Правило Oja.

Данное правило может предотвратить расходимость обучения Hebiann, сдерживая рост вектора веса w. Для этого существуют несколько способов: простая нормализация всех весов $w_i' = \alpha w_i$ после каждого обновления, выбирая α так, чтобы $\left| w' \right|=1$. Но Oja [1982] предложил более используемый подход. Непосредственно модифицируя правило Hebbian, он показал, что, возможно, заставить весовой вектор приблизиться к постоянной длине без необходимости проведения всякой нормализации вручную.

Правило Oja соответствует добавлению для уменьшения веса, пропорционального Y^2 к Hebbian правилу. По правилу Oja находят модуль весового вектора, который максимизирует среднеквадратичный выход.

$$\Delta w_i = \eta Y (x_i - Y w_i) \quad (6)$$

и *M* согласно правилу Oja равно:

$$\Delta w_{ij} = \eta Y_i \, (x_j - \sum_{k=1}^{n} Y_k \, w_i) \quad (7)$$

Заметим, что это напоминает обучение по правилу обратной дельта; Δ w зависит от разности между фактическим входом и выхода по обратному распространению.

Обучение по правило Sanger применяются в следующем виде:

$$\Delta w_{ij} = \eta Y_i \, (x_j - \sum_{k=1}^{i} Y_k \, w_i) \quad (8)$$

Сеть для осуществления неконтролируемого обучения по правило Sanger показано на рисунке 3. На данном рисунке показана только одна входная строка x_1.

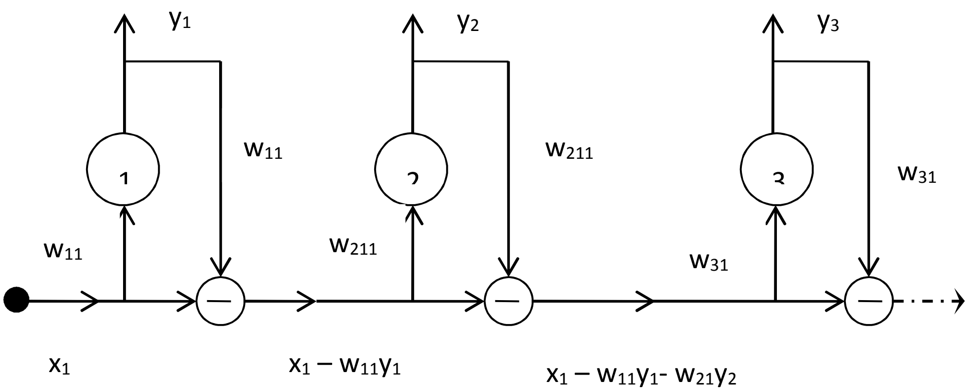

Рис. 3. Сеть для осуществления неконтролируемого обучения по правило Sanger.

Взвешенный выход y_i из модуля i вычтен из входа прежде, чем оно достигнет модуля i+1. Отметим, что каждый вес появляется дважды; оба значения должны сохраниться равными.

Для правила Sanger весовые векторы становятся точно первыми *M* направлениями главных компонентов. Правило Sanger обычно более полезно на практике, так как согласно данному правилу, отдельно извлекаются главные компоненты так, что дает восстанавливаемый результат на заданном наборе данных, в случае, если собственные значения не вырождены.

Правило Sanger, то есть изучение по модифицированному правилу Hebbian для вычисления весов сети, которая будет представлена собственными векторами и выходом, представляющий собой собственные значения. Алгоритм PCA осуществлен с помощью следующих шагов:

1. Заполнить начальные веса сети случайными числами между (0,1). Установить количество выходов m, который представляют собой собственные значения и изучение скоростью η. Входной вектор должен быть нормализован в среднем к нулю. Вычислить среднее из всех входных образцов, и вычесть полученное значение из каждого входа так, чтобы получить входы с нулевым-средним и таким образом, вычислить среднеквадратичное отклонение δ всех образцов.

2. Входной вектор передать вперед по сети согласно уравнению 1, чтобы вычислить выход

$$y_j(n)=\sum_{i=1}^{m} w_{ij}(n)\,x_i(n), \quad j=1,2,3,\ldots,m. \tag{9}$$

3. Синаптический вес w_{ij} (n) адаптировать в соответствии с обобщенной формой обучения Hebbian уравнением (6).

$$\Delta w_{ij}(n)=\eta\left[y_j(n)x_i(n)-\ y_j(n)\sum_{i=1}^{j} w_{ij}(n)x_i(n)\right]$$

$i = 1,2,\ldots, m; j = 1,2,\ldots, l.$ (10)

Веса модифицируются уравнением 11

$$w_{ij}(n+1)= w_{ij}(n)+\Delta w_{ij}(n) \tag{11}$$

4. Мы выполняем итерации по всем входным образцам, используя измененные веса для нового цикла n.

5. Выбирем набор из выхода, сравнивая все выходы с δ s.t. $A =\{ y_j \mid y_j > \delta \}$, обновляем m.

6. Условие повторения заключается в сравнении среднего из частей дельты весов Δ(n) с выбранной ошибкой ε. Если Δ(n) ≤ ε , то остановить процесс, иначе идти на шаг 2.

В общем случае, данный алгоритм сходится очень быстро. При выполнении необходимо отбросить исключение, когда отдельный вес растет выше соответствующего верхнего предела. Восстановить входы уменьшенной размерностью m<l из уравнения (11). Уменьшенные размерные образцы из PCA используются для входного уровня в сети Кохенена.

$$X_{new}=W^t\cdot X_{old} \tag{12}$$

В целях построение гибридной модифицированной модели распознавание образов выполняется с использованием сети Кохенена.

Один из типов конкурентоспособного самопроизвольного обучения является карта признаков Кохенена или самоорганизующуюся карта (СОК). В неконтролируемом обучении нет никаких ожидаемых выходов, представленных в нейронной сети, как в контролируемом алгоритме обучения с учителем типа обратного распространения. Самоорганизующая карта Кохенена (СОК) – система нейронной сети, разработанная Tuevo Кохенена, часто используется для классифицирования входов на различные категории. Нейронная сеть Кохенена работает отличным образом в отличие от нейронной сети с прямой связью. Нейронная сеть Кохенена содержит только слой входа и выхода нейронов, при этом нет ни одного скрытого слоя как показано на рисунке 4.

Подобно любой другой нейронной сети, использование карты Кохенена состоит из двух шагов: шаг обучения и шаг тестирования. При обучении, входные примеры последовательно используются в качестве входа в нейронной сети, каждый раз, когда изменяются веса связей. Входные данные используются до тех пор, пока не достигается сходимость нейронной сети. При тестировании, веса не изменяются, и выход нейронной сети используется в качестве ответа нейронной сети для заданных входных данных. Как показано на рис.4 карта Кохенена сформирована двумя слоями: входным слоем и слем выхода.

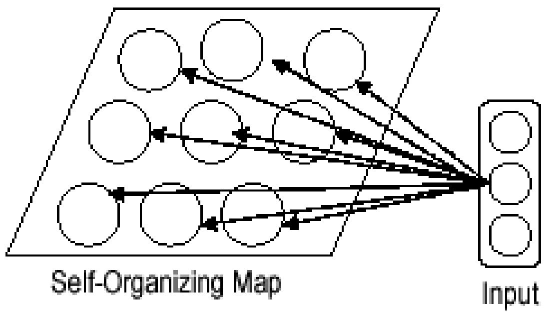

Рис. 4. Карты Кохенена.

Каждый нейрон слоя выхода связан с каждым нейроном входного слоя. Каждый нейрон определяет свой выход согласно взвешенной сумме в уравнении (13)

$$out_j=\sum_{i=1}^{n} w_{ij}\,x_i \tag{13}$$

Веса и входы обычно нормализованы, то есть величина веса и входных векторов установлена равной единице. Нейрон с наибольшим выходом – «победитель». Этот нейрон имеет заключительный выход 1, и все другие нейроны имеют выход нули.

Мексиканская функция «шляпы» показывает отношения между силой связи и расстоянием до нейрона победителя. В результате выполнения данной функции должна установиться конкурентоспособная среда для обучения. Только победившие нейроны и соседи участвуют в изучении для заданного входного образца как показано на рисунках 3 и 5 в области '+'. Другие нейроны не участвуют.

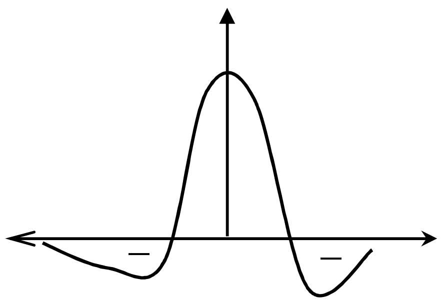

Рис. 5. Мексиканская функция НАТ.

Процесс обучения для карты Кохенена является прямым процессом. Для каждого набора обучения устанавливается один нейрон на выходном слое который, "победит" с минимальным расстоянием между его весами и входным вектором.

$$d_j = \sum_{i=1}^{n} (w_{ij} - x_i)^2 \qquad (14)$$

Размер окрестностей используется для моделирования эффекта мексиканской функции шляпы. Те нейроны, которые находятся в пределах расстояния, указанного размером окрестностей, участвуют в обучении и обновлениях векторов веса уравнением (15):

$$w_{ij}(t+1) = w_{ij}(t) + \eta(x_i - w_{ij}(t)) \qquad (15)$$

Нейроны, которые находятся вне данного расстояния размера окрестности, не участвуют в обучении. Обычно, размер окрестности запускается в качестве начального значения. И далее например 50% карты берется в качестве выходов и уменьшается в такой же пропорции в течении цикла обучения.

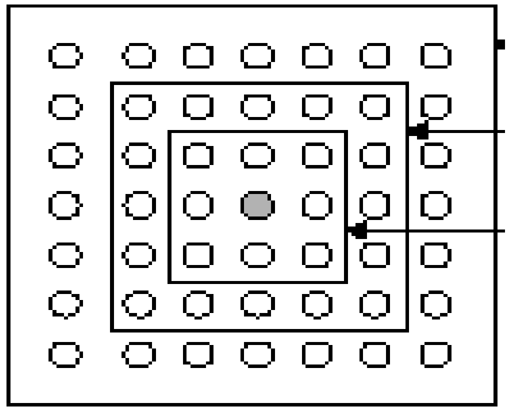

Рис. 6. Соседние окна победившего узла.

Коэффициент скорости обучения η - значение между 0 и 1, который дает скорость сходимости. Фиксированный, победивший нейрон будет иметь скорректированный вес такой, что он будет реагировать более строго на входе на следующий раз. Поскольку различные нейроны побеждают для различных образцов, то будет увеличена их способность распознать специфический образец.

Литература:

1. Hu M.K. Visual pattern recognition by moment invariants [Text] / M.K. Hu // IRE Trans. Inform. - 1962. - T 8. - P. 179-187.
2. Li Y. Reforming the theory of invariant moments for pattern recognition [Text] / Y. Li. - Pattern Recognition. - 1992. - Vol. 25. - P. 723-730.
3. Wong W.H. Generation of moment invariants and their uses for character recognition [Text] / W.H. Wong, W.C. Siu, K.M. Lam // Pattern Recognition Lett. - 1995. - Vol. 16. - P. 115-123.
4. Teague M.R. Image analysis via the general theory of moments [Text] / M.R. Teague. - J. Opt. Soc. Am. - 1980.- Vol. 70. - P. 920-930.
5. Wallin A. Complete sets of complex Zernike moment invariants and the role of the pseudoinvariants [Text] / A. Wallin, O. Kubler // IEEE Trans. Pattern Anal, Mach. Intell. - 1995. - Vol. 17. - P. 1106-1110.
6. Dusembaev A. One Approach for fuzzy classification using fuzzy clustering analysis [Text] / A. Dusembaev, W. Khedr // Doklady National Academy of Sciences Republic of Kazakhstan. - Astana. - 2004. - P. 112-118.

Рецензент: к.ф.-м.н., доцент Омаров А.М.