

**THE DIGITAL CALCULATIONS METHOD OF THE NEURAL NETWORK
ACTIVATION LEVEL
СПОСОБ МАТЕМАТИЧЕСКОГО РАСЧЕТА УРОВНЯ АКТИВАЦИИ
НЕЙРОННОЙ СЕТИ – ССМ**

Sergey Kozlov
Department of Decision Support Systems,
Riga Technical University,
1 Kalkju Street, LV – 1658, Latvia
mob. +3716476183, tel. +3715624717,
e-mail: tac@inbox.lv

Abstract. *The paper is devoted to the analysis of three-layer neural networks, learned by CC4 method, which was proposed by Subhash C. Kak. The radius of generalization is being investigated. It is found that the activation of a disclosed layer's element is mathematically dependent on the radius of generalization. The determined correlations weakly depend on the sizing and the amount of input data. This allows the modeling time of neural networks operations on the computer to be significantly reduced. Thus, the modeling time does not substantially depend on the size of a task.*

Keywords: pattern recognition, neural networks, generalization

1. Введение

Данная работа нацелена на исследование природы обобщения, которое наиболее полезно при отделении шумовых помех при распознавании образов. Один из методов обобщения, с одной стороны, встроен в алгоритм СС4, который используется для обучения нейронной сети, а с другой стороны, использует алгоритм работы самой нейронной сети.

Таким образом, сделана попытка построить алгоритм обобщения, который не базируется на других алгоритмах и понятиях. Такое разделение понятий на “Модель обобщения”, “Алгоритм обучения”, “Структуру и работу нейронной сети” обеспечивает, по мнению автора, во-первых, более точное понятие каждой сущности, и во-вторых, что наиболее полезно в техническом смысле, возможность увеличить скорость моделирования нейронной сети на компьютере. Это означает, что при моделировании можно заменить параллельные процессы на их имитацию – но с гораздо меньшими потерями времени. Следовательно, нет необходимости использовать ресурсы для организации параллельных процессов, и можно наиболее экономично использовать имитацию, причем без существенных потерь времени.

В данной работе сделан шаг к осуществлению такой концепции, и основное внимание уделено расчету активации скрытого элемента нейронной сети. Расчет активации скрытого элемента происходит очень часто, поэтому хотелось бы уменьшить время расчета, и к тому же сократить количество таких расчетов. В данной работе рассмотрен метод ССМ, с помощью которого можно свести расчет активации элемента к выбору значения из массива, который создан заранее и постоянно хранится в памяти. Вторая цель – сокращение количества расчетов, достигается за счет того, что в методе ССМ можно оперировать не двоичными цифрами, а десятичными. Метод ССМ основан на понятии “Модели обобщения”, заложенной в алгоритме СС4. Поэтому мы переходим к рассмотрению алгоритма СС4 и рассмотрению сущности понятия “Модели обобщения”. А затем рассмотрим собственно метод ССМ.

2. Метод СС4

Метод СС4 использует сетевую архитектуру, состоящую из трех уровней двоичных нейронов. Уровни входа и выхода полностью связаны. Число входных нейронов равно длине входного вектора плюс один дополнительный нейрон, являющийся нейроном смещения, который имеет постоянное значение 1. Число скрытых нейронов равно числу обучающих выборок. Каждый скрытый нейрон соответствует одной обучающей выборке. Рис.1 показывает структуру общей сети СС4.

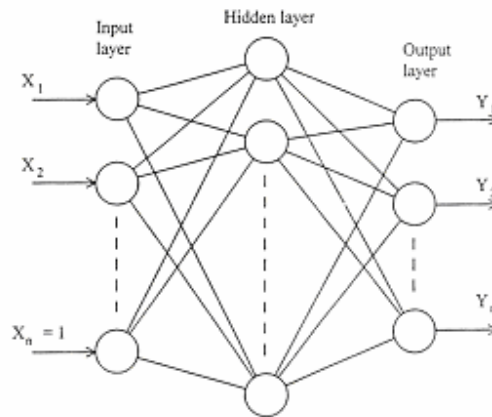


Рис. 1. Архитектура нейронной сети для метода СС4

Функция активации - двоичная ступенчатая функция, чей выход равняется 0, если сумма всех взвешенных входов не превышает порог; в противном случае выход равняется 1. Нейрон смещения был добавлен к входному уровню, чтобы значения порогов для всех нейронов были бы равны 0.

Веса входа и выхода определяются на основе исследования обучающих выборок. Для каждого вектора обучения, поданного на вход сети, если входной нейрон принимает значение 1, вес связи к скрытому нейрону, соответствующий этому вектору обучения, устанавливается равным 1. В противном случае он принимается равным (-1). Нейрон смещения обрабатывается иначе. Если s - сумма входов в обучающем векторе, исключая вход смещения, и желательный радиус обобщения - r , то вес связи между нейроном смещения и скрытым нейроном, соответствующий этому вектору обучения, равен $(r-s + 1)$. Таким образом, для любого обучающего вектора длины n , который включает смещение, веса входного уровня назначаются согласно следующему уравнению:

$$w_i[j] = \begin{cases} 1, & \text{if } x_i[j] = 1 \\ -1, & \text{if } x_i[j] = 0 \\ r - s + 1, & \text{if } j = n \end{cases}$$

Точно так же в выходном уровне, если обучающий вектор устанавливает значение 1 в выходном нейроне, то вес от скрытого нейрона до соответствующего выходного нейрона устанавливается равным 1. В противном случае устанавливается значение (-1).

Обучающий алгоритм СС4 может быть формально описан следующим образом:

```

for each training vector  $x_i$  [ $n$ ] do //  $n$  = длина вектора, включая смещение
     $s_i$  = no. of 1's in  $x_i$  [ $1:n-1$ ];
    for  $j$  = 1 to  $n - 1$  do
        if  $x_i$  [ $j$ ] = 1 then
             $w_i$  [ $j$ ] = 1; // Входной вес
        else
    
```

```

        wi [j] = -1;
    end
end
wi [n] = r - si + 1;           // r = радиус обобщения
for k = 1 to m do               // m = длина выходного вектора
    if yi [k] = 1 then
        ui [k] = 1;           // Выходной вес
    else
        ui [k] = -1;
    end
end
end
end

```

Причина для предписания весов таким образом может быть объяснена первым предположением, состоящим в том, что $r = 0$. Когда обучающийся вектор подан на вход сети, скрытый нейрон, соответствующий этой обучающей выборке, принимает значение 1 и зависит от значений входных нейронов, которые представлены с помощью 1, и принимает значение 0 для тех входных нейронов, которые представлены значением 0. В то же самое время скрытый слой получает сигнал $(-s + 1)$ от нейрона смещения. Таким образом, значения входов, полученных скрытыми нейронами, изменяются от $(-s + 1)$ до 1. В связи с тем, что скрытый нейрон использует двоичную функцию активации, а порог был принят равным нулю, скрытый нейрон будет активен. Все другие нейроны скрытого слоя принимают значение ноль или вход сети является негативным вследствие несоответствия в значениях входа (+1) и веса (+1). Таким образом, только один скрытый нейрон будет активен для каждого вектора обучающей выборки, поданного на вход сети. Во внешнем уровне способ, которым веса выхода были назначены, гарантирует, что единичный вход, активизирующий скрытый нейрон, осуществляет связь с обучающим вектором. При этом считают выход положительно коррелированным с обучающим вектором, а все другие нейроны скрытого слоя считаются отрицательно коррелированными с обучающим вектором.

Если тестовый вектор, который не был в наборе обучения, теперь подан на вход сети, все скрытые нейроны примут значение ноль, что соответствует негативному выходу сети из-за отрицательной корреляции, и ни один из выходных нейронов не будет активен. Отсюда следует, что, когда $r=0$, сеть узнает, что обучающая выборка воспроизводится корректно, но вывод не будет осуществляться.

Рассмотрим случай $r > 0$. Когда тестовый вектор подан на вход, сеть его блокирует, если данный вектор в значительной степени отличается от любого из обучающих векторов, которые сеть уже узнала и сохранила [1].

3. Понятие “Модель обобщения”

В данном разделе будет объяснено понятие “Модель обобщения”. Данное понятие достаточно абстрактно, и поэтому будет представлена конкретная модель. “Модель обобщения” определяется двумя параметрами:

1. Метод обучения нейронной сети;
2. Размерность задачи (или, что то же самое, структура нейронной сети).

В данной работе используется метод обучения СС4. Размерность задачи определяется количеством входных и выходных данных. Рассмотрим только входные данные, так как выходные - описываются аналогично. Определим входные данные:

1. Размерность данных: 1 байт (числа от 0 до 255);

2. Количество данных 256.

Как было описано выше, в методе СС4 “Модель обобщения” выбирается с помощью k – коэффициента обобщения. Причем значение k при нашей размерности находится в пределах от 1 до 8. “Модель обобщения” для каждого k можно описать бинарной матрицей. Такие матрицы представлены на рис. 2 – 9, где 1 – черная точка, а 0 – белая точка. Такая бинарная матрица описывает подобие входных элементов (ось y) с элементами в памяти (ось x), т.е. такими элементами, которыми была обучена нейронная сеть в процессе обучения методом СС4.

Для нашего примера такой вид модели (рис. 2 - 9) строго определяется самим методом обучения – СС4. Метод обучения СС4 в режиме тестирования (или при распознавании) динамически создает ту или иную часть “Модели обобщения”. Выбор этой части заложен в алгоритм работы нейронной сети, а если конкретнее, в расчет активации элемента скрытого слоя.

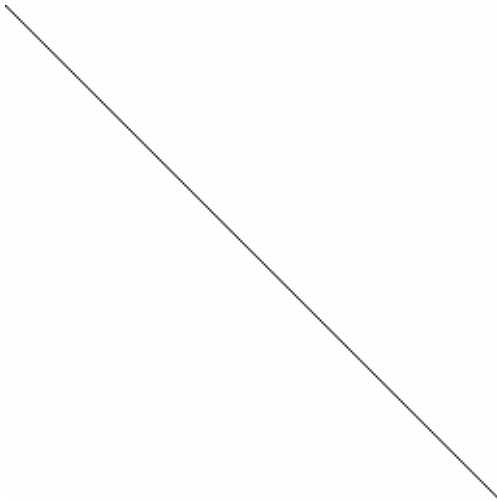


Рис. 2. “Модель обобщения” при $k=0$

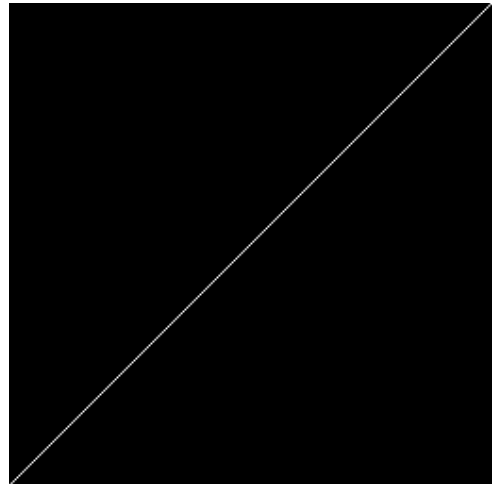


Рис. 3. “Модель обобщения” при $k=7$

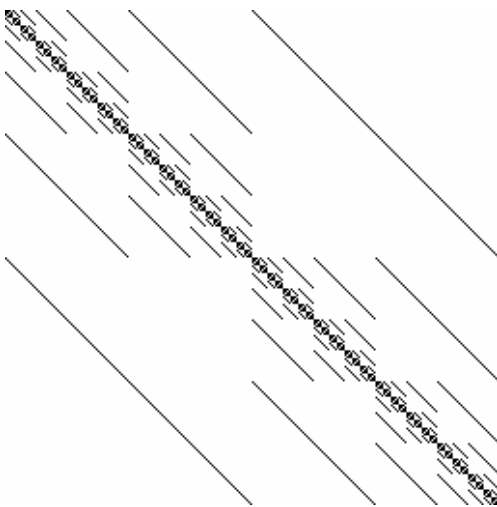


Рис. 4. “Модель обобщения” при $k=1$

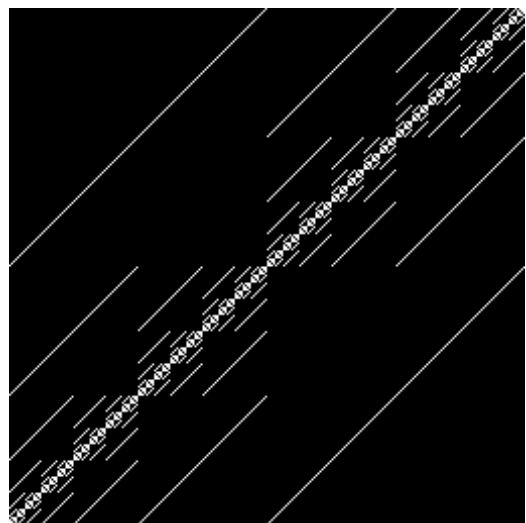


Рис. 5. “Модель обобщения” при $k=6$

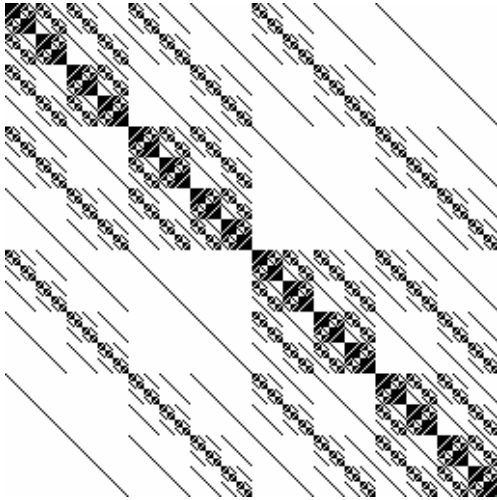


Рис. 6. “Модель обобщения” при $k=2$

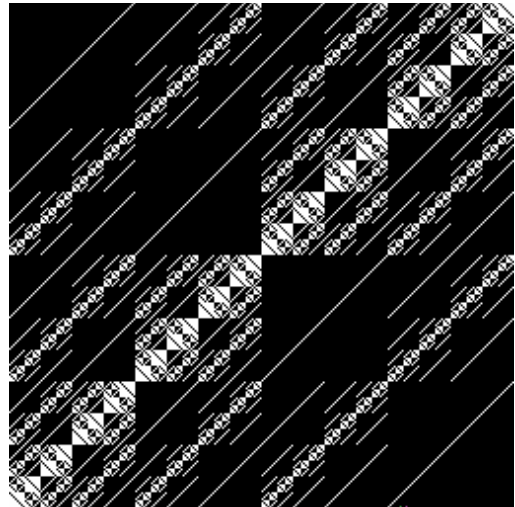


Рис. 7. “Модель обобщения” при $k=5$

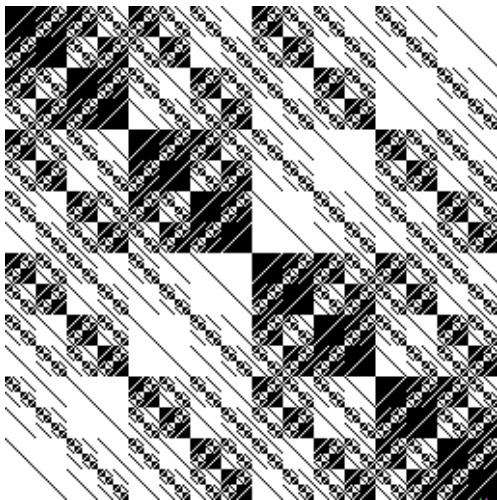


Рис. 8. “Модель обобщения” при $k=3$

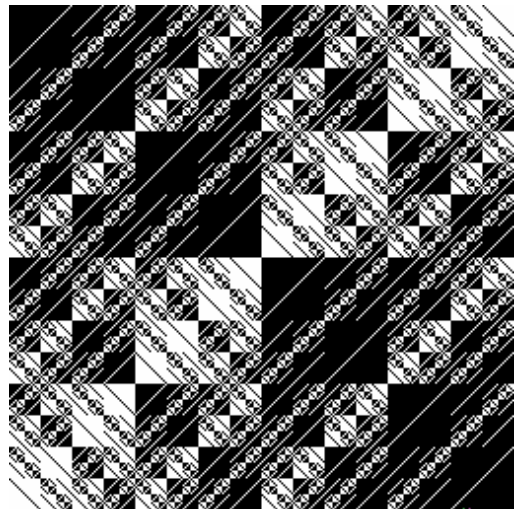


Рис. 9. “Модель обобщения” при $k=4$

Таким образом, построить бинарную матрицу, отражающую определенную “Модель обобщения” для i -го коэффициента обобщения можно, последовательно подавая на вход все варианты возможных входных данных (ось y) и соответствующую активацию каждого из элементов (ось x), т.е. в матрице отражается активация N -го элемента (1 – активен, 0 – неактивен) для M -ой входной последовательности.

4. Выбор подмоделей обобщения

Не вызывает сомнения тот факт, что строить полные “Модели обобщения” для расчета активации элемента неэффективно, так как нужно рассчитать все возможные варианты реакции элементов на стимул, даже тех, которые могут не быть в тестовой последовательности стимулов.

Но в любой “Модели обобщения” присутствует определенная симметрия и подобие различных частей. Поэтому имеет смысл построить набор достаточно малых по размерности “Моделей обобщения” и в конкретном случае применить ту или иную модель. Покажем это на примере. Пусть наша задача имеет следующую размерность:

1. Размерность данных – 2 байта (числа от 0 до 65535);
2. Количество данных – 65536,

т.е. теперь наша задача имеет в 2 раза большую размерность, чем было принято в предыдущем разделе. Но именно для этой задачи, мы сможем использовать 4 “Модели обобщения” (рис. 2 - 5) из 8 построенных ранее (рис. 2 - 9) для задачи, чья размерность была в 2 раза меньше.

Остается только определить, когда использовать ту или иную модель, что и будет описано в следующем разделе.

5. Стратегия выбора подмоделей обобщения

Ниже представлена стратегия выбора “Модели обобщения”. Данная стратегия может быть описана двумерной матрицей (или массивом) с размерностью $k_{\max} \times S_M$, где $k_{\max} = N/2$ - максимальный коэффициент обобщения (для нашего примера - 4), а N – количество входов нейронной сети, которая нужна для получения “Модели обобщения”, $S_M = 2^{k_{\max}}$ - размер модели (для нашего примера – 16). Элементы данной матрицы указывают на номер “Модели обобщения”. Под номером “Модели обобщения” понимается то, для какой степени обобщения k пригодна данная “Модель обобщения”. Заполняется данная матрица, исходя из следующего алгоритма:

```

for i=1 to kmax
    Model[i].Value[1] = kmax-i+1
    Model[i].Value[2] = kmax-i+2
end
for j=3 to Sm
    Model[1].Value[j] = kmax+1
end
for n=1 to N-1
    for i=2 to kmax
        for j=1+2^n to 2^(n+1)
            Model[i].Value[j] = Model[i-1].Value[j-2^n]
        end
    end
end

```

Для примера, приведенного в предыдущей главе, матрица, описывающая стратегию выбора “Модели обобщения”, показана на рис.10. Причем номер 5 соответствует “пустой” “Модели обобщения”.

4	5	5	5	5	5	5	5	5	5	5	5	5	5	5
3	4	4	5	5	5	4	5	5	5	5	5	5	5	5
2	3	3	4	4	5	3	4	4	5	4	5	5	5	5
1	2	3	3	4	4	2	3	3	4	3	4	4	4	5

Рис.10. Матрица, описывающая стратегию выбора “Модели обобщения”

6. Расчет активации методом ССМ

Чтобы рассчитать активность скрытого элемента нейронной сети, нужно знать следующие три величины:

1. k - текущий коэффициент обобщения;
2. x - число, которому обучен данный элемент (десятичное представление всех весов данного элемента);
3. y - число, поданное на вход (десятичное представление всех сигналов, поданных на вход).

Далее, исходя из этих величин, нужно рассчитать две величины по следующим формулам:

$$ModelNumber = ABS(INT(\frac{x-y}{S_M})) + 1;$$

$$ElementNumber = ABS(MOD(\frac{x-y}{S_M})) + 1,$$

где $ABS(x)$ - абсолютное значение числа;

$INT(x)$ - целая часть числа;

$MOD(x)$ - остаток от деления.

Используя эти значения, нужно обратиться вначале к массиву, описывающему стратегию выбора подмоделей обобщения, следующим образом:

$$Model = ModelArray[k + 1].Value[ModelNumber].$$

Получив номер модели, остается только обратиться к соответствующей модели и узнать, активен ли элемент скрытого слоя нейронной сети:

$$Aktiv = ElementArray[Model].Value[ElementNumber].$$

7. Заключение

В данной статье описан теоретический подход, который позволяет существенно сокращать время работы нейронной сети с помощью коэффициента обобщения. На момент написания статьи, этот метод только начал использоваться, в связи с чем не представляется возможным привести численные показатели эффективности метода ССМ. Однако, общая концепция позволяет утверждать, что при рекурсивном использовании этого метода становится возможным решить проблемы обобщения и распознавания в приемлемые сроки и использовать ресурсы более эффективно.

Kopsavilkums

Darba mērķis ir izpētīt vispārināšanu, kas visvairāk noder trokšņa (traucējumu) atdalīšanai veicot tēlu atpazīšanu. Viena no vispārināšanas metodēm ir no vienas puses iebūvēta algoritma CC4, kurš tiek izmantots neironu tīkla apmācībā, bet no otras puses, izmanto paša neironu tīkla darbības algoritmu.

Tādējādi, tiek mēģināts izveidot vispārināšanas algoritmu, kurš nebāzējas uz citiem algoritmiem un jēdzieniem. Jēdzienu sadalījums uz "Vispārināšanas modelis", "Apmācības algoritms", "Neironu tīkla struktūra un darbība" pēc autora domām dod, pirmkārt, katra jēdziena precīzāku izpratni un, otrkārt, kas ir svarīgi no tehniskā viedokļa, iespēju palielināt neironu tīkla modelēšanas ātrumu uz datora. Tas nozīmē, ka veicot modelēšanu, paralēlos procesus var aizvietot ar to imitāciju, pie tam ar daudz mazākiem laika zaudējumiem. Līdz ar to, nav nepieciešams izmantot resursus paralēlu procesu organizācijai un iespējams ekonomiskāk izmantot imitāciju, pie tam bez būtiskiem laika zaudējumiem.

Šajā darbā tiek sperts solis lai realizētu šādu koncepciju, un galvenā uzmanība tiek pievērsta neironu tīkla slēptā elementa aktivācijas izskaitļošanai. Slēptā elementa aktivācijas izskaitļošana tiek veikta ļoti bieži, tādēļ gribētos samazināt tā izskaitļošanas laiku, kā arī samazināt tādu aprēķinu apjomu. Šajā darbā tiek aplūkota metode ССМ, ar

kuras palīdzību var elementa aktivācijas izskaitļošanu vienkāršot līdz vērtības izvēlei no masīva, kurš radīts jau iepriekš un pastāvīgi glabājas atmiņā. Otrs mērķis - aprēķinu apjoma samazināšana tiek panākta ar to, ka metodē CCM var operēt nevis ar bināriem, bet gan ar decimāliem skaitļiem. Metode CCM balstās uz "Vispārināšanas modeļa" jēdzienu, kas iekļauts algoritmā CC4.

Bibliography

1. Subhash C. Kak. (1998). On generalization by neural networks, Information Sciences 111, P.293-302.
2. Fausett L. (1994). Fundamentals of Neural Networks: Architectures, Algorithms and Applications, Prentice Hall International Inc.
3. Rojas R. (1996). Neural Networks. A Systematic Introduction. Berlin, Springer – Verlag.