

ИСКУССТВЕННАЯ НЕЙРОННАЯ СЕТЬ ХОПФИЛДА

Персептрон не содержит обратных связей и относится к классу сетей, в которых сигналы распространяются в одном направлении от входов к выходам. Каждый нейрон передаёт сигналы другим нейронам только один раз. Процесс функционирования такой сети происходит без итераций и он всегда устойчивый.

В общем случае может быть построена нейронная сеть, содержащая разнообразные обратные связи, по которым сигналы с выходов нейронов передаются на их входы. Возникает итерационный процесс работы искусственной нейронной сети. Сети с обратными связями обладают более широким спектром возможностей, чем сети без обратных связей, но характеризуются и определёнными проблемами. Наличие обратных связей может приводить к возникновению неустойчивых процессов. Поэтому в такой сети могут отсутствовать стационарные состояния. Примером искусственной нейронной сети с обратными связями является сеть Хопфилда, предложенная в 1982 году американским учёным Джоном Джоозефом Хопфилдом.

Нейронная сеть Хопфилда — это однослойная нейронная сеть с обратными связями и симметричной матрицей связей. Особенность такой сети состоит в том, что она является динамической системой и результатом её работы является одно из положений равновесия. Оно достигается не мгновенно. Изменения в сети Хопфилда продолжаются до момента достижения равновесия. После этого момента каждое последующее состояние сети совпадает с её предыдущим состоянием.

Нейронную сеть Хопфилда можно определить как дискретно действующую динамическую систему с обратными связями, у которой выходные сигналы, полученные при выполнении прямой операции на одном шаге работы сети, служит основой для совершения такой же операции на следующем шаге работы сети.

Задача, решаемая сетью Хопфилда, характерна для *ассоциативной памяти*. Формулировка этой задачи выглядит следующим образом. Известны некоторые наборы двоичных входных воздействий, которые можно рассматривать как двоичные векторы - *образы*. Ими могут быть наборы двоичных сигналов, характеризующих изображения, звуки, состояние объектов или характеристики процессов. Такие наборы воздействий считаются *образцовыми*. Получив на вход произвольный двоичный вектор, предварительно обученная сеть Хопфилда должна сформировать на своём выходе тот образцовый вектор, с которым ассоциируется входной вектор, т.е. тот, на который больше всего похож входной вектор. Начальное состояние сети является *входным образом*, а состояние, достигнутое при равновесии, представляет собой *выходной образ*. Пример сети Хопфилда приведён на рис. 2.21.

Входы нейронов подсоединены к их выходам через цепи обратных связей, имеющие определённые, заданные при обучении сети, *коэффициенты веса*. При этом выходы нейронов подсоединяются ко входам всех нейронов, за исключением своих собственных входов.

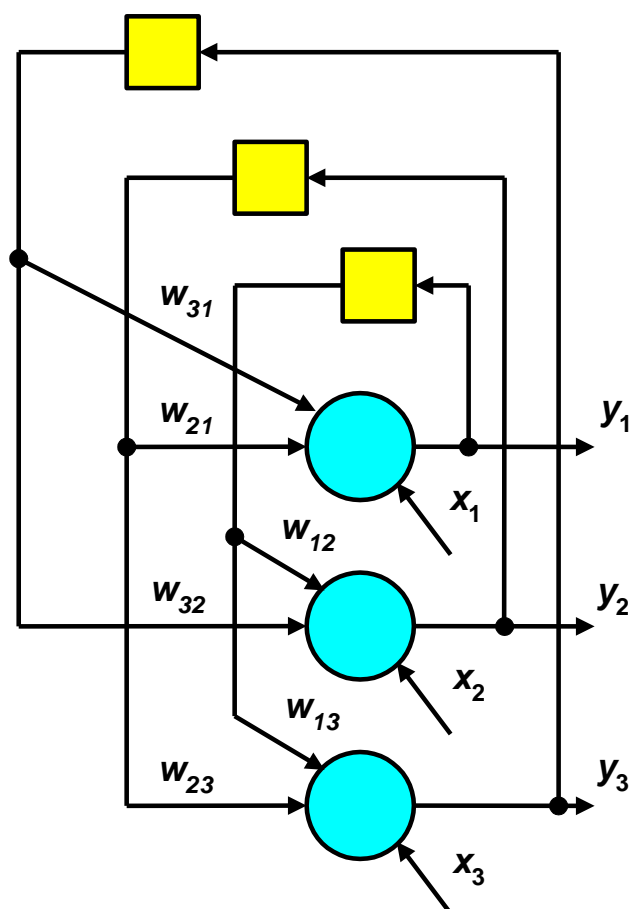


Рисунок 2.21 Пример структуры сети Хопфилда (входы показаны условно)

Сети, в которых имеются обратные связи, представляют собой **рекуррентные сети**. Каждая выполняемая сетью операция формирования очередного набора выходных сигналов на основании её исходного состояния, называется **итерацией**. Сеть представляет собой динамическую систему и поэтому происходящие в ней процессы могут быть устойчивыми или неустойчивыми. Устойчивый процесс сходится к некоторому фиксированному состоянию, называемому неподвижной точкой (аттрактором). При этом исходное состояние сети порождает точно такое же состояние в результате работы сети. Другими словами, неподвижная точка на входе приводит к образованию на выходе той же самой точки. Сеть Хопфилда может иметь несколько аттракто-

ров. Для каждого из них существует определённая область притяжения. Она представляет собой множество входных векторов состояния сети, приводящих к аттрактору в процессе итераций, совершаемых сетью как динамической системой. Поэтому можно говорить о том, что входные векторы, принадлежащие области притяжения аттрактора, *ассоциированы* с ним, а совокупность аттракторов образуют *память* сети Хопфилда.

Слово «*ассоциация*» происходит от латинского слова *associare*, которое означает соединять, объединять. Ассоциация – это связь между внешними воздействиями в виде образов объектов и явлений и возникающими в системе управления образами других объектов и явлений. Особенность ассоциативной памяти состоит в том, что *на основе ассоциации* она способна восстанавливать информацию, соответствующую аттракторам, имея на входе неполные или искажённые данные.

Таким образом, особенность сети Хопфилда состоит в том, что она обладает памятью и применяется для решения задач распознавания и восстановления образов. Сеть способна запомнить эталонные образы. Получив на вход некоторый неизвестный, несколько *искажённый* или *неполный* входной образ, сеть Хопфилда может извлечь из памяти тот эталонный образ, который окажется наиболее близким к входному образу.

Каждый образ кодируется в виде некоторого N – мерного вектора, где N – число нейронов в сети. Размерность входных и выходных векторов также равна N . Компоненты входного вектора имеют значения $+1$ или -1 . Например, если речь идёт о графическом образе, состоящем из чёрных и белых областей, то белому участку может соответствовать $+1$, а чёрному участку значение -1 .

Работа сети Хопфилда происходит следующим образом. Сначала выходам нейронов присваиваются значения компонентов вектора, описывающего входной образ. Фактически такая сеть не имеет входов в их привычном виде. Ввод данных о входном воздействии осуществляется *непосредственной установкой значений аксонов*.

Сеть ведёт себя как динамическая система и через некоторое количество циклов в результате процесса автоматической подстройки состояний нейронов сеть может сформировать на выходе вектор, соответствующий эталонному образу. Однако ёмкость сети Хопфилда ограничена, и в некоторых случаях сети не удаётся найти эталонный вектор и поэтому она выдаёт несуществующий образ.

Это происходит, если сделана попытка запомнить большое количество эталонных образов при недостаточном числе нейронов либо при наличии эталонных образов, слабо отличающихся друг от друга. В [22] отмечается, что количество M эталонных образов, запомненных сетью Хопфилда, состоящей из N нейронов, не может превышать N . Кроме того, число ошибочных реакций сети Хопфилда существенно возрастает, если предпринимается попытка запоминания количества M образов, превышающих $(0,138...0,15)N$ [18, 22, 49].

Если предположить, что допустимая погрешность формирования выходного образа не должна превышать 1% [49], то приблизительную оценку размера ограниченной ёмкости сети Хопфилда можно получить, воспользовавшись выражением

$$M \approx 0,14N.$$

В частности, сеть, изображённую на рис. 2.21 и имеющую три нейрона, ничему обучить нельзя. Реальная сеть Хопфилда должна иметь достаточно большое число нейронов, достигающее до нескольких сотен, а иногда и тысяч. Для повышения качества работы сети Хопфилда целесообразно увеличивать количество входящих в неё нейронов. Приближённая оценка количества N нейронов, необходимых для распознавания M образов, может быть выполнена по формуле

$$N \approx 15M.$$

Например, если мы хотим иметь возможность распознавать **2** образа, то сеть должна состоять, по крайней мере, из **30** нейронов. При необходимости распознавать **10** образов количество нейронов такой сети должно быть не менее **150**. А при создании сети, способной различать **50** образов потребуется не менее **750** нейронов.

Попытка записи большего числа образов, чем даёт записанная выше формула, приводит к потере возможности распознавания образов нейронной сетью. И хотя итерационный процесс может завершиться достижением устойчивого состояния, оно может оказаться ложным аттрактором (*химерой*) и не дать правильного ответа. Как правило, химеры оказываются составленными из фрагментов различных эталонных образов. В этом состоит одна из проблем и недостатков сети Хопфилда.

В сети Хопфилда должно выполняться условие симметричности связей. Это означает, что вес w_{ij} связи входа j нейрона с выходом i -го нейрона имеет то же значение, что и вес w_{ji} связи входа i -го нейрона с выходом j -го нейрона. Кроме того, связи входов нейронов с их собственными выходами отсут-

ствуют. Если все коэффициенты связей в сети Хопфилда свести в матрицу W , то её диагональные элементы будут нулевыми. Эти условия являются необходимыми для обеспечения устойчивости процесса функционирования сети.

2.9 Обучение сети Хопфилда

При обучении нейронной сети Хопфилда производится сохранение информации об образах-образцах в памяти сети. Алгоритм обучения такой сети значительно отличается от алгоритмов обучения других сетей, например, от алгоритма обучения персептрона по методу обратного распространения ошибки.

Обучением сети Хопфилда называется расчёт коэффициентов $N \times N$ - матрицы $W = \{w_{ij}\}$ коэффициентов обратных связей. Расчёт значений этих коэффициентов осуществляется однократно, после чего сеть будет способна распознавать и извлекать из памяти запомненные образы.

В случае запоминания одного входного образца

$$X = (x_1, x_2, \dots, x_m, \dots, x_N)^T$$

процедура обучения сети Хопфилда сводится к расчёту коэффициентов матрицы W по формуле

$$W = XX^T,$$

где X^T – транспонированный вектор X , который получается из вектора X заменой строк на столбцы. Таким образом, вектор X^T имеет вид

$$X^T = (x_1, x_2, \dots, x_m, \dots, x_N),$$

а матрица выглядит следующим образом:

$$W = \begin{bmatrix} x_1^2 & x_1x_2 & \dots & x_1x_N \\ x_1x_2 & x_2^2 & \dots & x_2x_N \\ \dots & \dots & \dots & \dots \\ x_1x_N & x_2x_N & \dots & x_N^2 \end{bmatrix}.$$

Важно отметить, что для устойчивости процессов работы сети необходимо **обнулить диагональные элементы матрицы W** , полученной в результате вычисления по приведённой выше формуле. В результате этого матрица W приобретает вид

$$W = \begin{bmatrix} 0 & x_1x_2 & \dots & x_1x_N \\ x_1x_2 & 0 & \dots & x_2x_N \\ \dots & \dots & \dots & \dots \\ x_1x_N & x_2x_N & \dots & 0 \end{bmatrix}.$$

Если требуется запомнить M образов-образцов X_1, X_2, \dots, X_M , то, найдя по вышеприведённой формуле матрицы для каждого из этих образов, результирующая формула коэффициентов обратных связей вычисляется следующим образом:

$$W = W_1 + W_2 + \dots + W_M.$$

Уравнения в развёрнутом виде, по которым можно рассчитать значения компонентов матрицы для обучения сети по M образцовым векторам, имеют следующий вид:

$$w_{ij} = \begin{cases} \frac{1}{M} \sum_{k=1}^M x_i^{(k)} x_j^{(k)} & \text{при } i \neq j \\ 0 & \text{при } i = j \end{cases},$$

где i и j – номера выходов и входов нейронов, $x_i^{(k)}$ и $x_j^{(k)}$ – i -ый и j -ый элементы вектора k -ого образца.

2.10 Функционирование обученной сети Хопфилда. Распознавание образов

Изменение состояния каждого нейрона в сети Хопфилда происходит по тому же правилу, что и для других искусственных нейронов МакКаллока и Питтса. Поступающие на входы нейронов сигналы с выходов других нейронов умножаются на весовые коэффициенты и суммируются. Затем сумма сравнивается с опорным сигналом и на основании активационной функции формируется выходной сигнал нейрона. Процесс поиска результата продолжается до тех пор, пока изменения выходного вектора не прекращаются. Возникающие при этом стационарные состояния могут интерпретироваться как образы, извлечённые из памяти нейронной сети.

Алгоритм функционирования обученной сети Хопфилда выглядит следующим образом.

1. По аналогии с персептроном говорят, что в начальный момент времени (на начальном нулевом $l=0$ шаге работы сети) на входы сети Хопфилда, имеющей N нейронов, подается неизвестный бинарный N -мерный вектор

$$X = (x_1, x_2, \dots, x_m, \dots, x_N).$$

На самом деле значения компонентов этого вектора присваиваются выходам нейронов. Поэтому обозначение на схеме сети входов носит условный характер. Таким образом, в начальный момент времени (при $l=0$) вектор

$$Y(l) = [y_1(l), y_2(l), \dots, y_m(l), \dots, y_N(l)],$$

характеризующий состояние сети, имеющей N нейронов, описывается следующим образом:

$$y_m(l) = x_m, \quad m = 1, \dots, N, \quad l = 0.$$

2. Рассчитываются новые состояния всех нейронов с номерами от 1 до N по формуле

$$S_n(l+1) = \sum_{m=1}^N w_{nm} y_m(l),$$

где w_{nm} – компонент $N \times N$ - матрицы W весовых коэффициентов; значения индексов n и m изменяется в диапазоне от 1 до N . Отметим, что значения элементов, стоящих на главной диагонали матрицы W всегда равны нулю.

После этого номер шага l увеличивается на единицу и определяются новые значения переменных, описывающих состояния нейронов, по формуле

$$y_m(l+1) = f[S_m(l+1)]$$

с учётом функции активации

$$f(u) = \text{sign}(u),$$

причём m изменяется в диапазоне от 1 до N .

Использование приведённой выше функции активации $f(u)$ имеет особенность. Если аргумент функции активации $S_m(l+1)$ m -го нейрона оказывается отрицательным, то переменная $y_m(l+1)$ приобретает значение -1. Если величина $S_m(l+1)$ положительна, то переменной $y_m(l+1)$ присваивается значение +1. Если же $S_m(l+1) = 0$, то выход m -го нейрона не изменяется, и он сохраняет предыдущее значение.

3. После каждого шага расчёта новых значений выходных сигналов сети производится проверка условия окончания процесса. изменения выходов сети за последнюю итерацию. Если произошли изменения значений этих выходных сигналов по сравнению со значениями на предыдущем шаге работы сети, то выполняется переход к пункту 2 для продолжения поиска решения. В противном случае считается, что решение найдено, а сформировавшийся на выходе сети вектор представляет собой образец, наиболее точно соответствующий входному вектору.

Различают **синхронный** и **асинхронный** режимы работы сети Хопфилда. При синхронной работе происходит одновременное изменение состояния всех нейронов. В этом режиме возможно бесконечное переключение между двумя разными состояниями. Фактически это означает, что сеть никогда не даст окончательный установившийся результат своей работы.

При асинхронной работе в данный момент времени происходит изменение состояния только одного нейрона. В этом случае процесс устойчив и изменения выходных сигналов со временем обязательно заканчиваются.

Сеть Хопфилда может работать в режиме фильтрации и восстанавливать образы на основе предъявленного сети частично искажённого образа. Такая возможность оказывается полезной при решении задачи распознавания образов. В [48] приведён пример, показывающий как нейронная сеть, состоящая из ста нейронов, восстанавливает изображение из его частично искажённого образа (рис. 2.22). При обучении в матрицу связей этой сети Хопфилда были записаны сведения о наборе чёрно-белых картинок. При этом чёрный цвет кодировался числом -1 , а белый цвет – числом $+1$. Если искажение входного вектора по сравнению с эталоном относительно невелико, то динамические процессы, происходящие в сети, завершаются формированием образа, совпадающего с эталоном. В данном случае был найден правильный образ, показанный на рисунке 2.22.

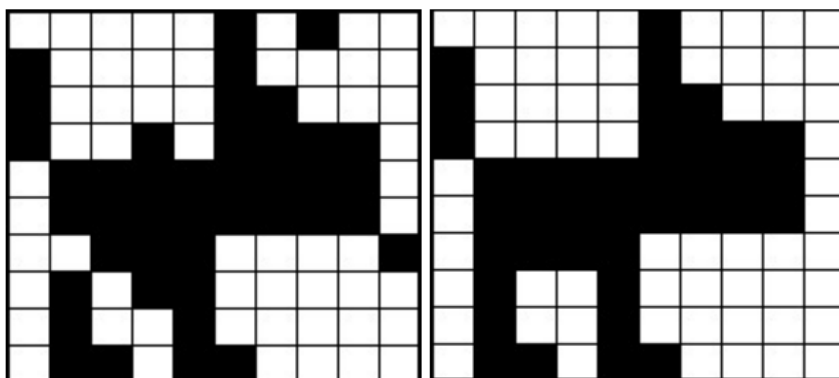


Рисунок 2.22 Восстановление изображения сетью Хопфилда (слева искажённый образ; справа восстановленный образ)

Но число распознаваемых образов не слишком велико в силу ограниченности памяти в сетях Хопфилда. Попытка записи информации о большем числе образов, чем может воспринять нейронная сеть, приводит к тому, что сеть перестаёт распознавать образы. Более того, факт достижения сетью устойчивого состояния не гарантирует, что сеть нашла правильный образ. В некоторых случаях, особенно при предъявлении сети сильно искажённых образов, сеть может сойтись к ложным аттракторам, называемым химерами. Они представляют собой комбинации фрагментов различных образов-эталонов.

Успешность распознавания образов существенно зависит от степени искажения предъявленного образа, вызванного действием шумов, по сравнению с эталонным образом. Ряд исследований показывают, что сеть Хопфилда успешно распознаёт и восстанавливает образ, если предъявленный образ искажён не более чем на 30 ... 40 %. Если представленные для распознавания образы зашумлены на 45 ... 60 %, то они распознаются нестабильно. Может возникать ошибочная реакция, в результате которой на выходе нейронной сети появляется образ, соответствующий совершенно другому эталону. А при степени искажения входного образа на 60% и более на выходе сети появляется негатив эталонного образа или его искаженный вариант. Весьма опасной ситуацией является попытка обучить сеть Хопфилда ряду похожих образов-эталонов при недостаточной памяти сети. Тогда возникает так называемый эффект перекрёстных ассоциаций, и сеть уже не может распознать ни один образ из числа похожих. Улучшение свойств сети Хопфилда достигается путём увеличения числа нейронов, ведущего к росту объёма её памяти.