

ФЕДЕРАЛЬНОЕ АГЕНТСТВО СВЯЗИ
Федеральное государственное образовательное бюджетное учреждение
высшего профессионального образования
«Поволжский Государственный Университет Телекоммуникации и Информатики»

Кафедра теоретических основ радиотехники и связи

Методическая разработка по контрольной работе
«Основы нейронных сетей»
по дисциплине
«Нейрокомпьютерные технологии»
для студентов
заочного факультета
дистанционной формы обучения

Самара 2013г

Данная методическая разработка содержит указания по контрольной работе, посвященной принципам функционирования и обучения нейронных сетей. Рассмотрены модели нейрона, функции активации, архитектура нейронных сетей, однослойные и многослойные персептроны, радиальные сети, а также рекуррентные сети (Хопфилда и Хемминга). Даны пояснения для выполнения контрольной работы, приведены примеры расчётов.

Данные методические указания предназначены для студентов заочной и дистанционной форм обучения.

Составитель: ст. преп. Харитонова А.А.

Редактор: д.т.н., проф. Горячкин О.В.

Рецензент: д.т.н., проф. Тарасов В.Н.

Содержание

Краткая теория

1. Основы нейронных систем
2. Однослойный персептрон. Линейное разделение классов. Многослойный персептрон. Нелинейное разделение классов.
3. Радиальная нейронная сеть. Нелинейное разделение двух классов.
4. Рекуррентные сети как ассоциативные запоминающие устройства.

Примеры расчёта

Задания на контрольную работу

Вопросы по контрольной работе для самостоятельной проверки

Литература

КРАТКАЯ ТЕОРИЯ

1 Основы нейронных систем

Нейрокомпьютеры (НК) – это ЭВМ нового поколения, центральный процессор которой построен на основе нейронной сети и реализует нейросетевые алгоритмы. Отличие нейрокомпьютеров от реально действующих классических вычислительных систем параллельного типа заключается в том, что для решения задач они используют не заранее разработанные алгоритмы, а специальным образом подобранные примеры, на которых учатся. Появление НК обусловлено следующими причинами:

- развитие элементной базы, позволяющее на одной плате реализовывать персональный компьютер – полнофункциональный компьютер (модель нейрона),
- необходимость решения важных практических задач, поставленных действительностью.

Сегодня большинство нейросетевых продуктов представлено в виде моделирующего программного обеспечения. Ведущие фирмы разрабатывают также и специализированные нейрочипы или нейроплаты в виде приставок к персональным ЭВМ. При сочетании нейрочипов и программного обеспечения производительность ЭВМ возрастает в сотни, а то и в тысячи раз. Однако эти приставки представляют собой классические процессоры (универсальные или специализированные). При таком подходе реализация суперЭВМ с «истинным параллелизмом», когда каждый нейрон работает в соответствии с логикой работы нейронной сети, представляется проблематичным[1].

Основные проблемы создания нейрокомпьютеров – разработка сверхпараллельных нейросетевых алгоритмов формализуемых задач и их распараллеливание в соответствии с архитектурой коммутационной системы, а также разработка новых методик решения неформализуемых задач. При этом рассматриваются три вида моделей нейронных сетей: физические, математические, технологические [1].

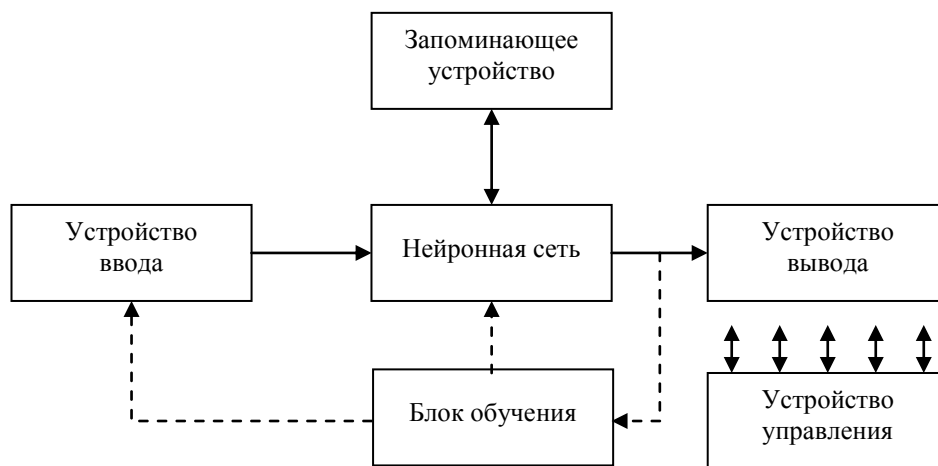


Рисунок 1 – Структурная схема абстрактного нейрокомпьютера

На рисунке 1 представлена структурная схема абстрактного нейрокомпьютера. Эту схему можно назвать обобщенной.

Основным операционным блоком НК, его процессором, является искусственная нейронная сеть. В первом приближении сеть представляет собой совокупность простейших модулей, называемых формальными нейронами, соединенными каналами передачи информации. Количественная характеристика каждого канала определяется решаемой задачей.

Нейронная сеть не производит вычислений, как это делает АЛУ (арифметико-логическое устройство) машин фон Неймана. Она трансформирует входной сигнал (входной образ) в выходной в соответствии со своей топологией и значениями коэффициентов межнейронной связи.

В запоминающем устройстве НК храниться не программа решения задачи, а программа изменения коэффициентов связи между нейронами.

Устройства ввода и вывода информации выполняют те же функции, что и в машине фон Неймана. Устройство управления служит для синхронизации работы всех структурных блоков нейрокомпьютера при решении конкретной задачи [1].

В работе нейрокомпьютера выделяют два главных режима работы – обучения и рабочий. Для того чтобы нейрокомпьютер решал требуемую задачу, его нейронная сеть должна пройти обучение на эту задачу. Суть режима обучения заключается в настройке коэффициентов межнейронных связей на совокупность входных образов этой задачи. С математической точки зрения процесс обучения

представляет собой решение задачи оптимизации, целью которой является минимизация функции ошибки на данном множестве примеров путем выборки коэффициентов межнейронных связей [1].

Биологические основы функционирования нейрона

Искусственные нейронные сети представляют собой попытку использовать процессы, происходящие в нервной системе, для выработки новых технологических решений.

Нервная клетка, сокращенно называемая нейроном, является основой нервной системы. Считается, что человеческий мозг содержит около 10^{11} нейронов, каждый из которых выполняет относительно примитивные функции суммирования весовых коэффициентов входных сигналов и сравнения полученной суммы с пороговым значением. Каждый нейрон имеет свои веса и свое пороговое значение. Они определяются местонахождением нейрона и решаемой им задачей и могут интерпретироваться аналогично содержимому локальной памяти процессора [1,3].

Громадное количество нейронов и межнейронных связей (до 1000 входов в каждый нейрон) приводит к тому, что ошибка в срабатывании отдельного нейрона остается незаметной или несущественной. Другая важная особенность нервных систем – высокая скорость их функционирования, несмотря на относительно длительный цикл срабатывания каждой отдельной клетки, измеряемый в миллисекундах. Она достигается благодаря параллельной обработке информации в мозге огромным количеством нейронов, соединенных многочисленными межнейронными связями [3].

Модель нейрона

Нейрон представляет собой единицу обработки информации в нейронной сети. На блок-схеме рисунок 2 показана модель нейрона, лежащего в основе искусственных нейронных сетей. В этой модели можно выделить три основных элемента:

1. Набор синапсов, каждый из которых характеризуется своим весом (силой).

2. Сумматор складывает входные сигналы, взвешенные относительно соответствующих синапсов нейрона. Эту операцию можно описать как линейную комбинацию.

3. Функция активации ограничивает амплитуду выходного сигнала нейрона. Эта функция также называется функцией сжатия. Обычно нормированный диапазон амплитуд выхода нейрона лежит в интервале $[0, 1]$ или $[-1, 1]$ [2].

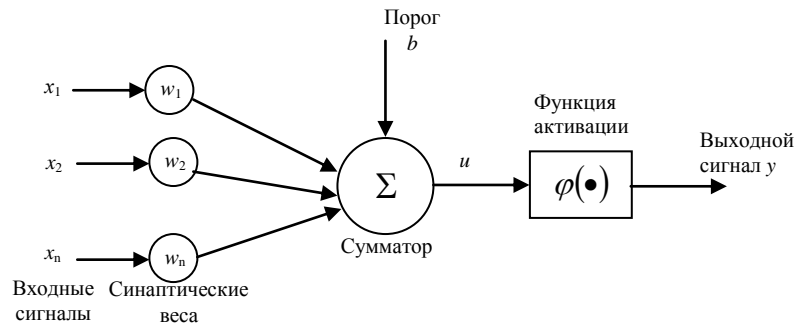


Рисунок 2 – Нелинейная модель нейрона

Пороговый элемент позволяет увеличить или уменьшить входной сигнал, подаваемый на функцию активации.

В математическом представлении функционирование нейрона можно описать следующей парой выражений

$$u = \sum_{j=1}^n w_j x_j, \quad (1)$$

$$y = \varphi(u + b) \quad (2)$$

где $\mathbf{X} = [x_1, x_2, \dots, x_n]$ – входные сигналы; $\mathbf{W} = [w_1, w_2, \dots, w_n]$ – синаптические веса нейрона; u – линейная комбинация входных воздействий; b – порог; $\varphi(\bullet)$ – функция активации; y – выходной сигнал нейрона.

Использование порога b обеспечивает эффект аффинного преобразования выхода линейного сумматора u . Преобразование плоскости называется аффинным, если оно взаимно однозначно (разные точки переводит в разные, и в каждую точку переходит какая-то точка) и образом любой прямой является прямая. Частным случаем аффинных преобразований являются просто движения. Движения — это параллельные переносы, повороты, различные симметрии и их комбинации.

Учитывая сказанное, имеем:

$$u = \sum_{j=0}^n w_j x_j, \quad (3)$$

$$y = \varphi(u) \quad (4)$$

В приведенной формуле подразумевается, что имеющий длину n вектор \mathbf{X} дополнен нулевым членом $x_0 = 1$ $\mathbf{X} = [x_0, x_1, x_2, \dots, x_n]$, тогда $\mathbf{W} = [w_0 = b, w_1, w_2, \dots, w_n]$ [2, 3].

Функции активации

Функции активации определяют выходной сигнал нейрона в зависимости от индуцированного локального поля (линейной комбинации входных воздействий). Можно выделить три основных типа функций активации.

1. Функция единичного скачка (пороговая функция или функция Хэвисайда).

$$y = \varphi(u) = \begin{cases} 1, & \text{если } u \geq 0; \\ 0, & \text{если } u < 0 \end{cases} \quad (5)$$

где u – это индуцированное локальное поле нейрона.

Эту модель в литературе называют моделью Мак-Каллока-Питца. Выходной сигнал нейрона принимает значение 1, если индуцированное локальное поле этого нейрона не отрицательно, и 0 – в противном случае [2].

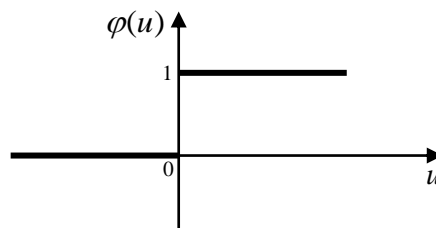


Рисунок 3 – Вид функции единичного скачка

2. Кусочно-линейная функция. Она описывается следующим выражением:

$$\varphi(u) = \begin{cases} 1, & u \geq +A \\ \frac{u}{2A} + \frac{1}{2}, & +A > u > -A \\ 0, & u \leq -A \end{cases} \quad (6)$$

где коэффициент усиления в линейной области оператора предполагается равным единице. Эту функцию активации можно рассматривать как аппроксимацию нелинейного усилителя.

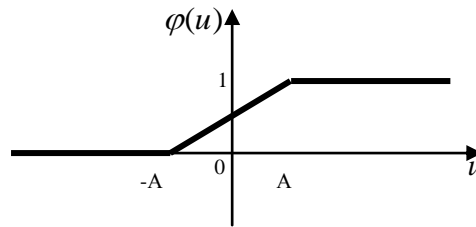


Рисунок 4 – Вид кусочно-линейной функции активации

Если линейная область оператора не достигает порога насыщения, он превращается в линейный сумматор.

Если коэффициент усиления линейной области принять бесконечно большим, то кусочно-линейная функция вырождается в пороговую (функцию единичного скачка).

3. Сигмоидальная функция. Ее график похож на букву S. Эта функция является самой распространенной. Это быстро возрастающая функция, которая поддерживает баланс между линейным и нелинейным поведением. Примером сигмоидальной функции может служить логистическая функция, задаваемая следующим выражением:

$$\varphi(u) = \frac{1}{1 + \exp(-\alpha u)}, \quad (7)$$

где α – параметр наклона сигмоидальной функции. Изменяя этот параметр, можно получить функции с различной крутизной. В пределе, когда $\alpha = \infty$, сигмоидальная функция вырождается в пороговую (функцию единичного скачка). Сигмоидальная функция является дифференцируемой, в то время как пороговая – нет.

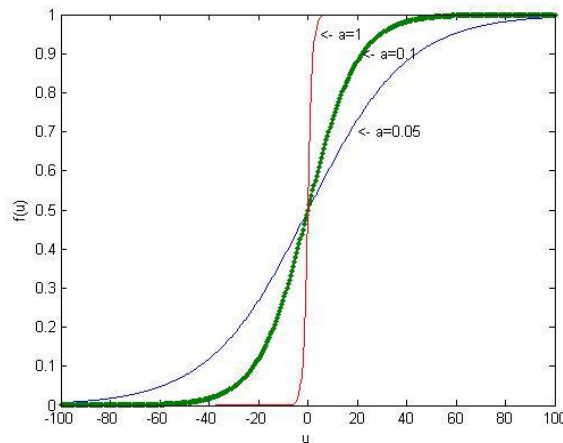


Рисунок 5 Вид сигмоидальной функции активации (унополярной)

Область значения приведенных функций активации лежит в интервале $[0,1]$, однако иногда требуется функция активации с областью значений в интервале $[-1,1]$. В этом случае функция активации должна быть симметричной относительно начала координат и являться нечетной функцией индуцированного локального поля. Эта функция обычно называется сигнум (биполярной сигмоидальной функцией). В данном случае сигмоидальная функция будет иметь форму гиперболического тангенса:

$$\varphi(u) = \text{th}(\pm \alpha u) \quad (8)$$

Такой вид сигмоидальной функции обеспечивает ряд преимуществ, в частности при обучении нейронных сетей.

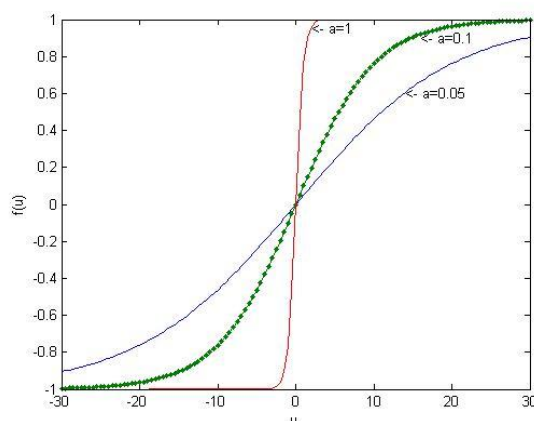


Рис. 5. Вид биполярной сигмоидальной функции активации

4. Степенные функции активации. Функции этого вида применяются при решении задач аппроксимации функций одной или нескольких переменных, для

реализации степенных рядов, для реализации сложных вычислительных алгоритмов. Математическая модель функции имеет вид:

$$\varphi(u) = u^k, \quad (1.9)$$

где $k = 2, 3, 4, \dots$ - показатель степени.

В настоящее время широко используют функции второй и третьей степени. Лишь степенные функции с нечетными показателями степени отвечают требованиям, предъявляемым к функциям активации – монотонны, дифференцируемы, непрерывны [1].

Перспективным направлением в теории нейронных сетей является применение функции активации с переменной «крутизной».

Приведенная модель нейрона является детерминистической, т.е. преобразование входного сигнала в выходной точно определено для всех значений входного сигнала. Однако в некоторых приложениях лучше использовать стохастические нейросетевые модели, в которых функция активации имеет вероятностную интерпретацию. В таких моделях нейрон может находиться в состояний: $+1$ или -1 . Решение о переключении состояния нейрона принимается с учетом вероятности этого события. Обозначим состояние нейрона символом z , а вероятность активации нейрона – функцией $P(u)$, где u - индуцированное локальное поле нейрона [3].

$$z = \begin{cases} +1, & \text{с вероятностью } P(u); \\ -1, & \text{с вероятностью } 1 - P(u). \end{cases} \quad (1.10)$$

Вероятность $P(u)$ описывается сигмоидальной функцией следующего вида:

$$P(u) = \frac{1}{1 + \exp(-u/T)}. \quad (1.11)$$

Архитектура сетей

Нейронная сеть – это ориентированный граф, состоящий из узлов, соединенных синаптическими и активационными связями, которые характеризуются следующими четырьмя свойствами.

1. Каждый нейрон представляется множеством линейных синаптических связей и, внешним порогом и функцией активации.

2. Синаптические связи входных сигналов определяются для взвешивания соответствующих входных сигналов.

3. Взвешенная сумма входных сигналов определяет индуцированное локальное поле каждого конкретного нейрона.

4. Активационные связи (функции) модифицируют индуцированное локальное поле нейрона, создавая выходной сигнал [2].

В общем случае выделяют три основных класса нейросетевых архитектур: 1. однослойные сети прямого распространения; 2. многослойные сети прямого распространения; 3. рекуррентные сети.

Однослойные сети прямого распространения – это сети, состоящие из входного слоя, состоящего из узлов источника (при подсчете числа слоев во внимание не принимаются) и выходного слоя, который состоит из нейронов (вычислительных узлов), в которые подается информация от входного слоя.

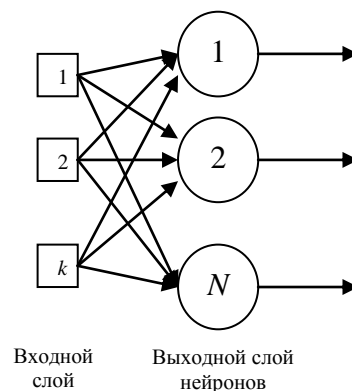


Рис. 6 Сеть прямого распространения с одним слоем нейронов

Многослойные сети прямого распространения – это сети, имеющие один или несколько скрытых слоев, узлы которых называются скрытыми нейронами (элементами). Если каждый выходной сигнал i -того слоя подается на вход всех нейронов $i+1$ -го, то говорят что сеть полносвязная. Если некоторые из синаптических связей отсутствуют, такая сеть называется неполносвязной.

Многослойные сети прямого распространения благодаря скрытым слоям могут вычислять статистики высокого порядка. Она позволяет выделять глобальные свойства данных с помощью локальных соединений за счет наличия дополнительных синаптических связей и повышать уровень взаимодействия нейронов.

В общем случае сеть прямого распространения с m входами, h_1 нейронами первого скрытого слоя, h_2 нейронами второго скрытого слоя и q нейронами выходного слоя называются сетью $m-h_1-h_2-q$ [2].

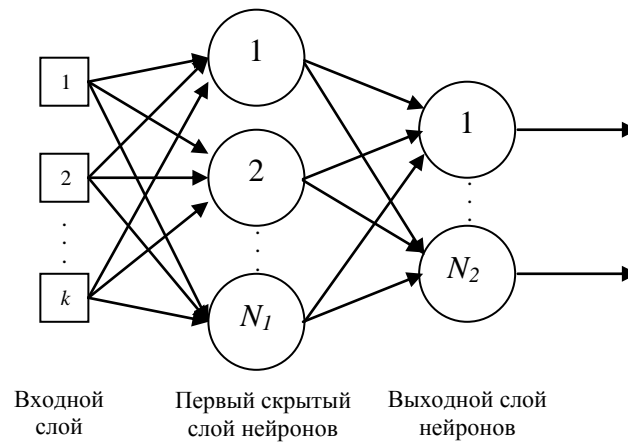


Рис. 7. Полносвязная сеть прямого распространения с одним скрытым и одним выходным слоем

Рекуррентные сети отличаются от сети прямого распространения наличием по крайней мере одной обратной связи.

Понятие обратной связи характерно для динамических систем, в которых выходной сигнал некоторого элемента системы оказывает влияние на входной сигнал этого элемента. Обратная связь присутствует в нервной системе любого животного.

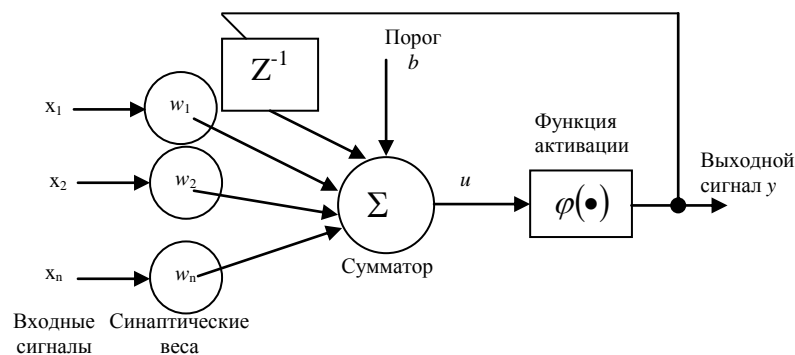


Рис. 8. Нейрон с обратной связью (рекуррентный нейрон)

$$y(n) = \sum_{l=0}^{\infty} w^{l+1} x(n-l) \quad (1.12)$$

$x(n-l)$ - это входной сигнал, задержанный на l единиц дискретизации, n - итерация (такт времени). Динамика реакции системы определяется весом w . Если $|w| < 1$,

сигнал $y(n)$ экспоненциально сходится и система устойчива, а если нет, то сигнал $y(n)$ расходится и система неустойчива [2].

Наличие обратных связей оказывает непосредственное влияние на способность таких сетей к обучению и на их производительность. Наличие элементов единичной задержки приводит к нелинейному динамическому поведению, если в сети содержатся нелинейные нейроны [2].

2. Однослойный персептрон. Линейное разделение классов.

Многослойный персептрон. Нелинейное разделение классов

Персептрон представляет собой простейшую форму нейронной сети, предназначенную для классификации линейно-разделимых сигналов. Персептрон состоит из одного нейрона с настраиваемыми синаптическими весами и порогом. Первый алгоритм настройки свободных параметров для такой сети был создан Розенблаттом для персептронной модели мозга. Розенблатт доказал, что если образы (векторы), используемые для обучения персептрона, выбраны из двух линейно-разделимых классов, то алгоритм персептрона сходится и формирует поверхность решений в форме гиперплоскости, разделяющей эти два класса. Персептрон, построенный только на одном нейроне, ограничен выполнением задачи разделения только двух классов (гипотез). Увеличивая размерность выходного слоя персептрона и включая в него несколько нейронов, можно решать задачи классификации на большее число классов.

Единичный нейрон также составляет основу адаптивного фильтра – функционального блока, являющегося основным для предметной области обработки сигналов.

Для построения линейного решающего правила необходимо выполнить несколько шагов:

1. Получение начального значения вектора весов на основании разделения центров масс для дальнейшего обучения персептрона.
2. Коррекция вектора весов персептрона по одному из алгоритмов обучения персептрона [4].

Линейное решающее правило,

построенное на основании разделения центров масс

Линейное разделение классов состоит в построении решающего правила, т.е. такого вектора $\mathbf{W} = [w_0, w_1, w_2, \dots, w_n]$, где w_0 – порог.

Разделение центров масс – простейший способ построения решающего правила. Суть этого способа заключается в вычислении вектора весов персептрона по следующей формуле

$$\mathbf{W} = \left(\sum_{i=1}^n \mathbf{X}^i - \sum_{j=1}^m \mathbf{X}^j \right) / (n + m), \quad (2.1.)$$

где \mathbf{X}^i , $i = 1, \dots, n$ вектора из множества первого класса, а \mathbf{X}^j , $j = 1, \dots, m$ - ко второму классу.

Линейные решающие правила, построенные на основании разделения центров масс, могут ошибаться на примерах из обучающей выборки. Однако этот метод полезен как средство определения начального значения вектора весов для алгоритма обучения персептрона.[4]

Алгоритм обучения персептрона по отдельным примерам

1. При изначально заданных значениях весов w_i $i = 1, \dots, N$ на вход нейрона подается обучающий вектор \mathbf{X} и рассчитывается значение выходного сигнала u , зная результат d . По результатам сравнения u с d уточняются значения весов.

2. Если $u=d$, то w_i не изменяются.

3. Если $u=0$, а $d=1$, то значение весов уточняется по формуле

$$w_i(t+1) = w_i(t) + \gamma x_i, \gamma \in (0,1) \quad (2.2.)$$

где γ - коэффициент (скорость) обучения, t - номер предыдущего цикла.

4. Если $u=1$, а $d=0$, то значение весов уточняется по формуле

$$w_i(t+1) = w_i(t) - \gamma x_i, \gamma \in (0,1) \quad (2.3.)$$

По завершении уточнения весовых коэффициентов представляются очередной обучающий вектор \mathbf{X} и связанное с ним ожидаемое значение d , и значение весов уточняется заново. Этот процесс многократно повторяется по всей обучающей выборке, пока не будут ликвидированы различия между всеми значениями u и соответствующим им ожидаемыми значениями d [4].

В итоге должна получиться n -мерная плоскость

$$w_0 + w_1 x_1 + w_2 x_2 + \dots + w_n x_n = 0 \quad (2.4.)$$

Многослойный персептрон. Нелинейное разделение классов

Многослойный персептрон успешно применяется для решения разнообразных сложных задач.

Многослойные персептроны имеют три отличительные признака:

1. Каждый нейрон сети имеет нелинейную функцию активации. Данная нелинейная функция является гладкой (т.е. дифференцируемой).
2. Сеть содержит один или несколько слоев скрытых нейронов, не являющихся частью входа и выхода сети. Эти нейроны позволяют сети обучаться решению сложных задач, последовательно извлекая наиболее важные признаки из входного образа (вектора).
3. Сеть обладает высокой степенью связности, реализуемой посредством синаптических соединений.

Благодаря этим свойствам и способностью обучаться на собственном опыте многослойного персептрона обеспечивает вычислительную мощь. Но эти качества являются причиной неполноты современных знаний о поведении этого рода сетей. Во-первых, распределенная форма нелинейности и высокая связность сети существенно усложняют теоретический анализ многослойного персептрона. Во-вторых, наличие скрытых нейронов делает процесс обучения более трудоемким для визуализации. Именно в процессе обучения необходимо определить, какие признаки входного сигнала следует представлять скрытыми слоями нейронной сети.

Обучение методом обратного распространения ошибки предполагает два прохода по всем слоям сети: прямого и обратного. При прямом проходе входной вектор подается на вход сети, после чего распространяется по сети от слоя к слою. В результате генерируется набор выходных сигналов (реакция сети на данный входной образ). Во время прямого прохода все синаптические веса фиксированы. Во время обратного прохода все синаптические веса настраиваются в соответствии с правилом коррекции ошибок, а именно: фактический выход сети вычитается из желаемого отклика, в результате чего формируется сигнал ошибки. Этот сигнал распространяется по сети в обратном направлении синаптических связей.

Синаптические связи настраиваются с целью максимального приближения выходного сигнала сети к желаемому в статистическом смысле.

Алгоритм обратного распространения ошибки реализует вычислительно эффективный метод обучения многослойного персептрона [4].

Выделение выпуклых областей

Серьезное ограничение разделяющих поверхностей однослойными сетями можно преодолеть, добавив дополнительные слои. Например, двухслойные сети, получаемые каскадным соединением однослойных сетей, способны выполнять более общие классификации, отделяя точки, содержащиеся в выпуклых ограниченных и неограниченных областях.

Область выпуклая, если для каждой двух ее точек соединяющий их отрезок целиком лежит в области. Область ограничена, если ее можно заключить в некоторый шар.

Включением достаточного числа нейронов во входной слой может быть образован выпуклый многоугольник (многогранник) желаемой формы.

Например, для выделения треугольной области в двумерном пространстве в первом слое потребуется три нейрона с весовыми коэффициентами $\mathbf{w}^{(1)} = [w_0^{(1)}, w_1^{(1)}, w_2^{(1)}]$, $\mathbf{w}^{(2)} = [w_0^{(2)}, w_1^{(2)}, w_2^{(2)}]$, $\mathbf{w}^{(3)} = [w_0^{(3)}, w_1^{(3)}, w_2^{(3)}]$, где верхний индекс указывает на номер нейрона, а нижний на весовой коэффициент этого нейрона. В итоге получим уравнения $w_0^{(1)} + x_1 \cdot w_1^{(1)} + x_2 \cdot w_2^{(1)} = 0$, $w_0^{(2)} + x_1 \cdot w_1^{(2)} + x_2 \cdot w_2^{(2)} = 0$, $w_0^{(3)} + x_1 \cdot w_1^{(3)} + x_2 \cdot w_2^{(3)} = 0$. Нейроны второго слоя должны выполнять функцию И.

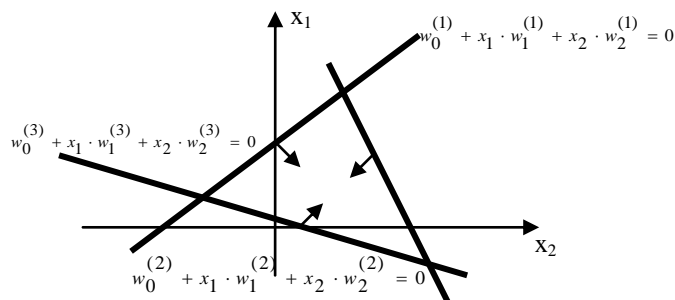


Рис. 2.1. Гиперплоскости, выделяющие на плоскости выпуклую (треугольную) область

Выделение невыпуклых областей

Для выделения невыпуклых областей используют трехслойные нейронные сети. Нейроны второго слоя не ограничены функцией И, а нейроны третьего слоя принимают в качестве входа набор выпуклых многогранников (многоугольников), и их логические комбинации могут быть невыпуклыми.

Например, выделение невыпуклой области, представленной в виде объединения трех треугольных областей. Пять нейронов первого слоя реализуют разделяющие гиперплоскости: 1-ая $w_0^{(1)} + x_1 \cdot w_1^{(1)} + x_2 \cdot w_2^{(1)} = 0$, 2-ая $w_0^{(2)} + x_1 \cdot w_1^{(2)} + x_2 \cdot w_2^{(2)} = 0$, ... 5-ая $w_0^{(5)} + x_1 \cdot w_1^{(5)} + x_2 \cdot w_2^{(5)} = 0$, два нейрона второго слоя реализуют трехвходовые функции И, нейрон третьего слоя реализует функцию ИЛИ.

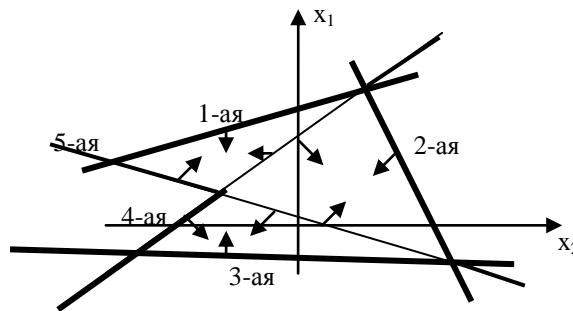


Рис. 2.2. Гиперплоскости, выделяющие на плоскости невыпуклую область

3. Радиальная нейронная сеть. Нелинейное разделение двух классов

Сети радиального типа представляют собой естественное дополнение сигмоидальных сетей. Сигмоидальный нейрон представляется в многомерном пространстве гиперплоскостью, разделяющей это пространство на две категории (класса).

Радиальный нейрон представляет собой гиперсферу, которая осуществляет шаровое разделение пространства вокруг центральной точки. Именно с этой точки зрения он является естественным дополнением сигмоидального нейрона, поскольку в случае круговой симметрии данных позволяет заметно уменьшить количество нейронов, необходимых для разделения различных классов. Структура радиальной сети включает в себя: входной слой, на который подаются сигналы, описываемые входным вектором x , скрытый слой с нейронами радиального типа и выходной слой, состоящий, как правило, из одного или нескольких линейных нейронов. Функция выходного нейрона сводится исключительно к взвешенному суммированию сигналов, генерируемых скрытыми нейронами.



Рис. 3.1. Иллюстрация способов разделения пространства данных: а) сигмоидальным нейроном; б) радиальным нейроном

Радиальные нейронные сети образуют семейство сетей, в которых нейроны реализуют функции, радиально изменяющиеся вокруг выбранного центра и принимающие ненулевые значения только в окрестности этого центра. Подобные функции, определяемые в виде $\varphi(x) = \varphi(\|x - c\|)$, будем называть радиальными базисными функциями.

Примерами радиальных базисных функций могут служить:

1. Мультиквадратичная функция: $\varphi(z) = (z^2 + a^2)^{1/2}$, где $a > 0$ и $z \in \mathbb{R}$
2. Обратная мультиквадратичная функция: $\varphi(z) = \frac{1}{(z^2 + a^2)^{1/2}}$, где $a > 0$ и $z \in \mathbb{R}$

3. Функция Гаусса: $\varphi(z) = \exp\left(-\frac{z^2}{2\sigma^2}\right)$, где $\sigma > 0$ и $z \in \mathbb{R}$

4. Сплайн тонких пластин (thin-plate spline): $\varphi(z) = z^2 \ln(z)$

В таких сетях роль нейрона заключается в отображении радиального пространства вокруг одиночной заданной точки (центра) либо вокруг группы таких точек, образующих кластер. Суперпозиция сигналов, поступающих от всех таких нейронов, которая выполняется выходным нейроном, позволяет получить отображение всего многомерного пространства.

Математическую основу функционирования радиальных нейронных сетей составляет теорема Т. Ковера о распознавании образов, в соответствии с которой нелинейные проекции образов в некоторое многомерное пространство могут быть линейно разделены с большей вероятностью, чем при их проекции в пространство меньшей размерности.

Если вектор радиальных функций в N -мерном входном пространстве обозначить $\varphi(x)$, то это пространство является нелинейно φ -разделимым на два пространственных класса X^+ и X^- тогда, когда существует такой вектор весов \mathbf{W} , что

$$\mathbf{W}^T \varphi(x) > 0, x \in X^+, \quad (3.1a)$$

$$\mathbf{W}^T \varphi(x) < 0, x \in X^-. \quad (3.1b)$$

Граница между этими классами определяется уравнением $\mathbf{W}^T \varphi(x) = 0$.

Доказано, что каждое множество образов, случайным образом размещенных в многомерном пространстве, является φ -разделяемым с вероятностью 1 при условии соответственно большой размерности этого пространства.

Для моделирования любой функции достаточно взять определенное число радиальных элементов. Обучение промежуточного слоя осуществляется по алгоритму объективной классификации.

Выход сети формируется в результате комбинации выходов скрытых радиальных элементов. Оказывается, что достаточно взять их линейную комбинацию. Поэтому эти сети имеют выходной слой, состоящий из элементов с линейными функциями активации, что позволяет не проводить обучение [3].

Топология радиальной базисной сети (*RBF – Radial Basis Function*) представлена на рис. 3.2.

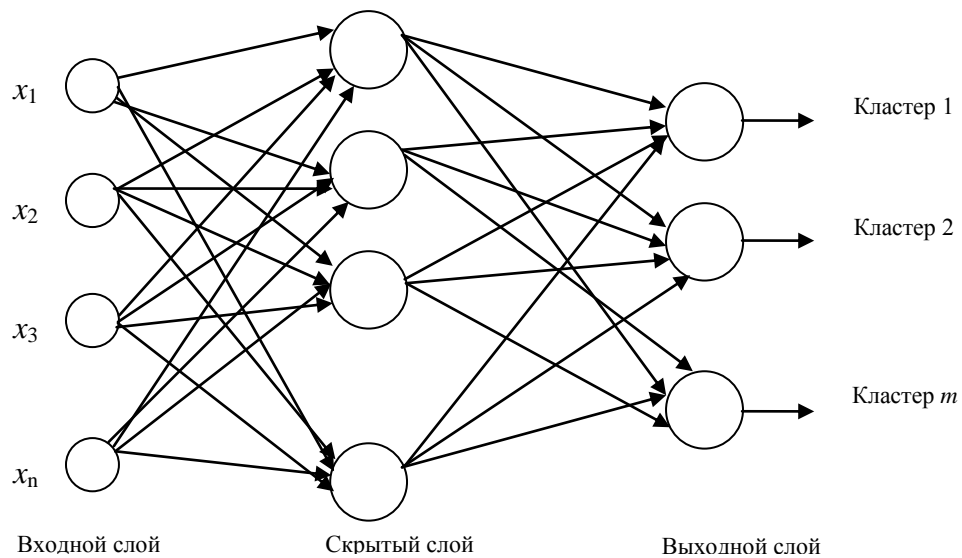


Рис. 3.2. Топология радиальной сети

Радиальные сети перед многослойным персептроном имеет ряд преимуществ:

- они моделируют произвольную нелинейную функцию с помощью всего одного скрытого слоя, поэтому избавляют конструктора сети от необходимости решать вопросы о числе слоев;
 - радиальные сети обучаются на порядок быстрее, чем многослойные персептроны.
- Однако в промежуточном слое для радиальных элементов необходимо определить положения их центров и величины отклонений [3].

Обучение радиальных сетей

Использование в разложении p базисных функций (p - это количество обучающих выборок) недопустимо как с теоретической, так и с практической точек зрения, так как обычно количество этих выборок очень велико, и в результате вычислительная сложность обучающего алгоритма становится чрезвычайной.

Если ограничиться L базисными функциями, то аппроксимирующее решение можно представить в виде

$$F(x) = \sum_{i=1}^L w_i \varphi(\|x - c_i\|) \quad (3.2)$$

где $L < p$, c_i ($i = 1, 2, \dots, L$) - множество центров, которые необходимо определить. В особом случае, если принять $L = p$, то можно получить точное решение $c_i = x_i$.

Задача аппроксимации состоит в подборе соответствующего количества радиальных функций и их параметров (центров и ширины, например), а также в таком подборе весов w_i , чтобы решение уравнения (3.2) было наиболее близким к точному.

Основной проблемой является подбор количества базисных функций, каждой из которых соответствует один скрытый нейрон. Подбор необходимого и достаточного количества нейронов зависит от многих факторов, в числе которых размерность задачи, объем обучающих данных и пространственная структура аппроксимируемой функции. Как правило, количество базисных функций L составляет определенную долю от объема обучающих данных p , причем фактическая величина этой доли зависит от размерности вектора x и от разброса ожидаемых значений d_i , соответствующих входным векторам x_i , для $i = 1, 2, \dots, p$.

Обучение сети начинается при каком-либо изначальном количестве нейронов, а впоследствии контролируется как степень уменьшения среднеквадратичной погрешности, так и изменение значений подбираемых параметров сети. Если среднее изменение значений весов после определенного числа обучающих циклов слишком мало $\sum_i (\Delta w_i) < \xi$, добавляются две базисные функции (2 нейрона) с центрами, соответствующими наибольшей и наименьшей погрешности адаптации, после чего обучение расширенной таким образом структуры продолжается. Одновременно контролируются абсолютные значения весов w_i всех отдельных нейронов. Если они меньше установленного вначале порога δ , соответствующие им нейроны подлежат удалению из сети. Как добавление нейронов, так и их удаление начинается после выполнения определенного количества обучающих циклов и может происходить в течение

всего процесса обучения вплоть до достижения требуемой точности отображения [3].

Гибридный алгоритм обучения радиальных сетей

В гибридном алгоритме процесс обучения разделяется на два этапа:

1. подбор линейных параметров сети (веса выходного слоя) при использовании метода псевдоинверсии;
2. адаптация нелинейных параметров радиальных функций (центра c_i и ширины σ_i этих функций).

Оба этапа тесно переплетаются. При фиксации конкретных значений центров и ширины радиальных функций (в первый момент это будут начальные значения) за один шаг подбираются величины линейных весов выходного слоя. Такая фиксация параметров радиальных функций позволяет определить значения самих функций $\varphi_i(x_k)$ для $i = 1, 2, \dots, L$ и $k = 1, 2, \dots, p$, где i - это номер радиальной функции, а k - номер очередной обучающей пары (x_k, d_k) . Функционирование сети при наличии L радиальных функций и p обучающих пар можно записать как систему уравнений

$$\begin{bmatrix} 1 & \varphi_1(x_1) & \varphi_2(x_1) & \dots & \varphi_L(x_1) \\ 1 & \varphi_1(x_2) & \varphi_2(x_2) & \dots & \varphi_L(x_2) \\ \dots & \dots & \dots & \dots & \dots \\ 1 & \varphi_1(x_p) & \varphi_2(x_p) & \dots & \varphi_L(x_p) \end{bmatrix} \cdot \begin{bmatrix} w_0 \\ w_1 \\ \dots \\ w_L \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ \dots \\ d_p \end{bmatrix} \quad (3.3)$$

На втором этапе при зафиксированных значениях выходных весов возбуждающие сигналы пропускаются по сети до выходного слоя, что позволяет рассчитать величину погрешности для последовательности векторов x . Далее происходит возврат к скрытому слою (обратное распространение). По величине погрешности определяется вектор градиента целевой функции относительно конкретных центров c_{ij} и ширины σ_{ij} [3].

Для последующего изложения предположим, что используется модель сети типа гиперрадиальной базисной функции с диагональной формой масштабируемой матрицы. Каждая радиальная функция определяется в общем виде как

$$\varphi_i(x_k) = \exp \left(-\frac{1}{2} \sum_{j=1}^N \frac{(x_{jk} - c_{ij})^2}{\sigma_{ij}^2} \right) \quad (3.4).$$

При существовании p обучающих пар целевую функцию можно задать в виде

$$E = \frac{1}{2} \sum_{k=1}^p [y_k - d_k]^2 = \frac{1}{2} \sum_{k=1}^p \left[\sum_{i=0}^K w_i \phi(x_k) - d_k \right]^2 \quad (3.5).$$

В результате дифференцирования этой функции получаем:

$$\frac{\partial E}{\partial c_{ij}} = \sum_{k=1}^p \left[(y_k - d_k) w_i \exp \left(-\frac{1}{2} \sum_{j=1}^N \frac{(x_{jk} - c_{ij})^2}{\sigma_{ij}^2} \right) \frac{(x_{jk} - c_{ij})}{\sigma_{ij}^2} \right] \quad (3.6)$$

$$\frac{\partial E}{\partial \sigma_{ij}} = \sum_{k=1}^p \left[(y_k - d_k) w_i \exp \left(-\frac{1}{2} \sum_{j=1}^N \frac{(x_{jk} - c_{ij})^2}{\sigma_{ij}^2} \right) \frac{(x_{jk} - c_{ij})^2}{\sigma_{ij}^3} \right] \quad (3.7)$$

Применение градиентного метода наискорейшего спуска позволяет провести уточнение центров и ширины радиальных функций согласно формулам:

$$c_{ij}(n+1) = c_{ij}(n) - \eta \frac{\partial E}{\partial c_{ij}}, \quad (3.8)$$

$$\sigma_{ij}(n+1) = \sigma_{ij}(n) - \eta \frac{\partial E}{\partial \sigma_{ij}} \quad (3.9)$$

Уточнение нелинейных параметров радиальной функции завершается очередной цикл обучения. Многократное повторение обоих этапов ведет к полному и быстрому обучению сети, особенно когда начальные значения параметров радиальной сети близки к оптимальным.[3]

4. Рекуррентные сети как ассоциативные запоминающие устройства

Главная особенность сетей с обратной связью между различными слоями нейронов (рекуррентные сети) это динамические зависимости на каждом этапе функционирования. Изменение состояния одного нейрона отражается на всей сети вследствие обратной связи типа «один ко многим». В сети возникает некоторый переходной процесс, который завершается формированием нового устойчивого состояния, отличающегося в общем случае от предыдущего. Если функцию активации нейрона обозначить $f(u)$, где u - взвешенная сумма его возбуждений, то состояние нейрона можно определить выходным сигналом

$$y_i = f(u_i) = f\left(\sum_{j=1}^N w_{ij} x_j\right). \quad (4.1)$$

Принимая во внимание, что при обратной связи типа «один ко многим» роль возбуждающих импульсов для нейрона играют выходные сигналы других нейронов, изменение его состояния может быть описано системой дифференциальных нелинейных уравнений

$$\tau_i \frac{du_i}{dt} = \sum_{j=1, j \neq i}^N w_{ij} f(u_j) - u_i - b_i \quad (4.2)$$

где b_i - пороговое значение, заданное внешним источником (в однонаправленных сетях это показатель поляризации). Коэффициент τ_i является числовой константой, а его интерпретация аналогична постоянной времени в уравнениях, описывающих динамическое состояние. Состояние нейрона рассчитывается в результате решения дифференциального уравнения как $y_i = f(u)$.

Рекуррентной сети можно поставить в соответствие энергетическую функцию Ляпунова

$$E = -(1/2) \sum_j \sum_{i \neq j} w_{ij} y_i y_j + \sum_{i=1, N} (1/R_i) \int_0^{y_i} f_i^{-1}(y_i) dy_i + \sum_{i=1, N} b_i y_i. \quad (4.3)$$

Изменение состояния какого-либо нейрона инициализирует изменение энергетического состояния сети в направлении минимума E энергии вплоть до его достижения. В пространстве состояний локальные энергетические минимумы E

представлены точками стабильности, называемыми аттракторами из-за тяготения к ним ближайшего окружения. Аттрактор (англ. attract — привлекать, притягивать) — множество точек в фазовом пространстве динамической системы, к которым стремятся траектории системы. Если траектория прошла достаточно близко к аттрактору, то со временем она уже не покинет окрестность аттрактора и даже будет подходить к нему всё ближе и ближе, то есть будет наблюдаться эффект притяжения к аттрактору. Благодаря наличию аттракторов, рекуррентные сети могут быть использованы как устройства ассоциативной памяти.

Главная задача ассоциативной памяти сводится к запоминанию входных (обучающих) выборок таким образом, чтобы при представлении новой выборки система смогла сгенерировать ответ — какая из запомненных ранее выборок наиболее близка к вновь поступившему образу. Наиболее часто в качестве меры близости отдельных множеств применяется мера Хемминга.

При использовании двоичных значений (0,1) расстояние Хемминга между двумя векторами $\mathbf{y} = [y_1, y_2, \dots, y_n]^T$ и $\mathbf{d} = [d_1, d_2, \dots, d_n]^T$ определяется в виде

$$d_H(\mathbf{y}, \mathbf{d}) = \sum_{i=1}^n [d_i(1 - y_i) + (1 - d_i)y_i] \quad (4.4)$$

При биполярных (± 1) значениях элементов обоих векторов расстояние Хемминга рассчитывается по формуле

$$d_H(\mathbf{y}, \mathbf{d}) = \frac{1}{2} \left[n - \sum_{i=1}^n y_i d_i \right] \quad (4.5)$$

Мера Хемминга равна нулю только тогда, когда $\mathbf{y} = \mathbf{d}$. В противном случае она равна количеству битов, на которое различаются оба вектора.

Автоассоциативная сеть Хопфилда

Структура сети Хопфилда представляется в виде системы с непосредственной обратной связью выхода со входом. Выходные сигналы нейронов являются одновременно входными сигналами сети: $x_i(k) = y_i(k-1)$. В классической сети Хопфилда отсутствует автосвязь (связь выхода нейрона с его собственным входом), что соответствует $w_{ii} = 0$, а матрица весов является

симметричной: $\mathbf{W} = \mathbf{W}^T$. Отсутствие автосвязи и симметричность матрицы весов являются достаточными (но не необходимыми!!!!) условиями сходимости итерационных (переходных) процессов в сети Хопфилда.

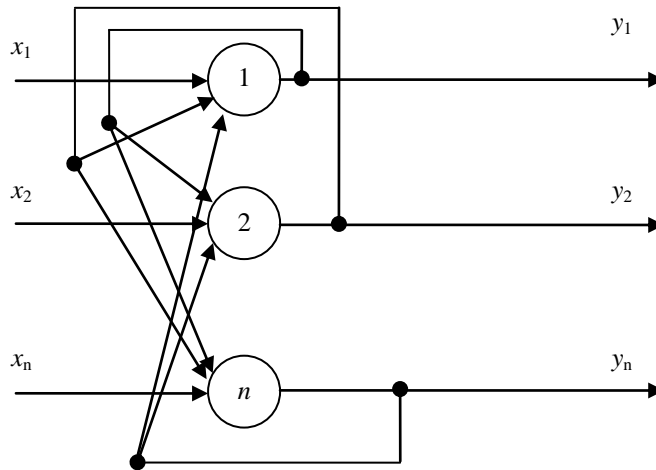


Рис. 4.1 Сеть Хопфилда

Кратко рассмотрим принцип действия сети. Если каждый нейрон имеет биполярную ступенчатую функцию активации со значениями ± 1 , то выходной сигнал i -го нейрона определяется функцией

$$y_i = \text{sgn} \left(\sum_{j=0, N} w_{ij} x_j + b_i \right). \quad (4.6)$$

Далее допустим, что порог срабатывания является компонентой вектора x . Тогда основную зависимость, определяющую сеть Хопфилда, можно представить в виде:

$$y_i(k) = \text{sgn} \left(\sum_{j=0, N} w_{ij} y_j(k-1) \right). \quad (4.7)$$

с начальным условием $y_j(0) = x_j$.

В процессе функционирования сети Хопфилда можно выделить два режима: обучения и классификации. В режиме обучения на основе известных векторов подбираются весовые коэффициенты сети. В режиме классификации при фиксированных значениях весов и вводе конкретного начального состояния нейронов возникает переходный процесс вида (4.7), завершающийся в одном из локальных минимумов, для которого $y(k) = y(k-1)$.

Существует несколько методов обучения сети Хопфилда (по Хеббу, метод проекций, Δ -проекций), однако нами будет рассмотрен только один.

Метод проекций

Отправной точкой этого метода считается предположение, что при правильно подобранных весах каждая поданная на вход выборка \mathbf{x} генерирует на выходе саму себя, мгновенно приводя к искомому состоянию. В матричной форме это можно представить в виде:

$$\mathbf{W}\mathbf{X} = \mathbf{X}, \quad (4.8)$$

где \mathbf{W} - матрица весов сети размерностью $N \times N$, а \mathbf{X} - прямоугольная матрица размерностью $N \times p$, составленная из p последовательных обучающих векторов $\mathbf{x}^{(i)}$, т.е. $\mathbf{X} = [\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(p)}]$. Решение такой линейной системы уравнений имеет вид:

$$\mathbf{W} = \mathbf{X}\mathbf{X}^+, \quad (4.9)$$

где знак $+$ означает псевдоинверсию.

После завершения подбора весов сети их значения «замораживаются», и сеть может использоваться в режима распознавания. В этой фазе на вход подается тестовый вектор \mathbf{x} и рассчитывается ее отклик в виде:

$$y(i) = \text{sgn}(\mathbf{W}y(i-1)) \quad (4.10)$$

(в начальный момент $y(0) = x$), причем итерационный процесс повторяется для последовательных значений $y(i)$ вплоть до стабилизации отклика. Итерационный процесс стабилизации отклика системы состоит из определенного количества циклов и в значительной степени зависит от размеров сети и от распределения локальных минимумов [3].

Сети Хемминга

Предложенная Р. Липпманном сеть Хемминга – это трехслойная рекуррентная структура, которую считают развитием сети Хопфилда. Она позиционируется как специализированное гетероассоциативное запоминающее устройство. Основная идея функционирования данной сети состоит в минимизации

расстояния Хемминга между тестовым вектором, подаваемым на вход сети, и векторами обучающих выборок, закодированными в структуре сети.

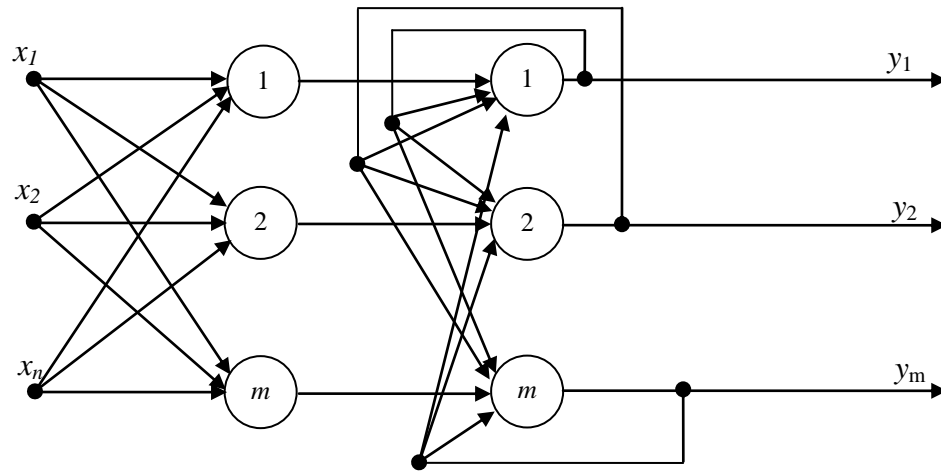


Рис. 4.2. Сеть Хемминга

Первый слой имеет однонаправленное распространение сигналов от входа к выходу и фиксированные значения весов. Второй слой состоит из нейронов, связанных обратными связями по принципу «каждый с каждым», при этом существует ненулевая связь входа нейрона со своим собственным выходом. Весовые коэффициенты этого слоя тоже постоянны. Разные нейроны связаны отрицательной обратной связью с весом $-\varepsilon$, при этом обычно величина ε обратно пропорциональна количеству образцов. С собственным выходом нейрон связан положительной обратной связью с весом $+1$. Веса поляризации нейронов принимают значения, соответствующие нулю. Нейроны этого слоя функционируют в режиме «победитель получает все», при этом в каждой фиксированной ситуации активизируется только один нейрон, а остальные пребывают в состоянии покоя. Выходной однонаправленный слой формирует выходной вектор, соответствующий входному вектору. Веса нейронов этого слоя подбираются в зависимости от входных обучающих выборок.

Сеть Хемминга считается гетероассоциативным запоминающим устройством с парой связанных между собой векторов (y, x) , где x и y - это входной и выходной биполярные вектора со значениями элементов ± 1 . Входные узлы сети $1, 2, \dots, N$

принимают значения, задаваемые аналогичными компонентами вектора x . Нейроны первого слоя рассчитывают расстояние Хемминга между фактически предъявленным входным вектором x и каждым из p закодированных векторов-образцов $x^{(i)}$, образующих веса нейронов первого слоя. Нейроны во втором слое выбирают вектор с наименьшим расстоянием Хемминга, определяя таким образом класс, к которому принадлежит предъявленный входной вектор x . Веса нейронов выходного слоя формируют вектор, соответствующий предъявленному входному вектору [3].

Процесс обучения:

Весовые коэффициенты первого слоя выбираются по следующему правилу:

$$w_{ik}^{(1)} = \frac{x_i^k}{2}, \text{ для } i=1,2,\dots,n, k=1,2,\dots,m \quad (4.11)$$

m - число образцов, n - размер образа.

$$T_k = \frac{n}{2}, k=1,2,\dots,m.$$

Весовые коэффициенты второго слоя выбираются следующим образом:

$$w_{ij}^{(2)} = \begin{bmatrix} 1 & i=j \\ -\varepsilon & i \neq j \end{bmatrix} 0 < \varepsilon < \frac{1}{m}. \quad (4.12)$$

Функционирование сети:

На вход сети подается неизвестный вектор x . Рассчитывается состояние нейронов первого слоя:

$$y_j^{(1)} = \sum_{i=1}^n w_{ij}^{(1)} x_i + T_j, j=1,2,\dots,m. \quad (4.13)$$

Этими значениями инициализируются значения выходов второго слоя.

Вычисляются новые состояния нейронов второго слоя.

$$s_j^{(2)}(t+1) = y_j^{(2)} - \varepsilon \sum_{i \neq j}^m y_i^{(2)}(t), j=1,2,\dots,m. \quad (4.15)$$

$$y_j^{(2)}(t+1) = f(s_j^{(2)}(t+1)). \quad (4.16)$$

Если произошло изменение y , то п.2 повторяется, если нет, то конец [3].

ПРИМЕРЫ РАСЧЕТА

Задача 1.

Нейрон j получает на вход сигнал, уровни которого равны \mathbf{X} . Соответствующие веса связей равны \mathbf{W} . Вычислить выходное значение нейрона j для заданных моделей нейрона.

Дано:

$$\mathbf{X} = [-1,3; 2,3; -3]$$

$$\mathbf{W} = [0,2; 0,6; -0,02], \quad w_0 = -1,2$$

Модель нейрона: полулинейный нейрон $\alpha = 3$

Решение:

В математическом представлении функционирование нейрона можно описать следующей парой уравнений

$$u^j = \sum_{i=1}^n w_i^j x_i,$$

$$y^j = \varphi(u + b^j)$$

где $\mathbf{X} = [x_1, x_2, \dots, x_n]$ - входные сигналы; $\mathbf{W}^j = [w_1^j, w_2^j, \dots, w_n^j]$ - синаптические веса j -го нейрона; u - линейная комбинация входных воздействий; b^j - порог; $\varphi(\bullet)$ - функция активации; y^j - выходной сигнал j -го нейрона.

Так как в данном примере рассматривается только один нейрон, то верхний индекс можно опустить.

Тогда имеем:

$$x_1 = -1,3 \quad w_1 = 0,2$$

$$x_2 = 2,3, \quad w_2 = 0,6, \quad w_0 = b = -1,2, \quad \text{функция активации } \varphi(u) = \begin{cases} \alpha \cdot u, & \text{если } u > 0; \\ 0, & \text{если } u \leq 0 \end{cases}.$$

$$x_3 = -3 \quad w_3 = -0,02$$

$$u = \sum_{i=1}^n w_i x_i = x_1 w_1 + x_2 w_2 + x_3 w_3 = -1,3 \cdot 0,2 + 2,3 \cdot 0,6 + (-3) \cdot (-0,02) = -0,26 + 1,38 + 0,06 = 1,18$$

$$y = \varphi(u + b) = \varphi(1,18 - 1,2) = \varphi(-0,02) = 0$$

Ответ: $y = 0$

Задача 2.

Задан набор точек S_1 , принадлежащий первому классу (выходной сигнал – 1) и набор точек S_2 , принадлежащий второму классу (выходной сигнал – 0). Составить уравнение гиперплоскости, линейно разделяющее пространство признаков на две части (найти вектор W и w_0).

Дано:

$S_1 : (2, 4); (0, 3); (0, 0)$

$S_2 : (-2, 0); (-1, -5); (3, -6); (2, -5)$

Найти: уравнение гиперплоскости.

Решение:

1. Изобразим заданные множества на плоскости

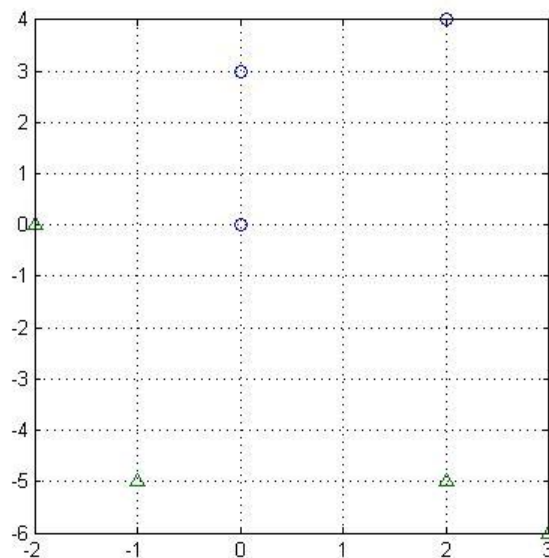


Рис. 5.1 Исходные точки

2. Построим линейное разделяющее правило при помощи разделения центров масс. Определим начальное значение вектора весов.

$$W = \left(\sum_{i=1}^n S_1^i - \sum_{j=1}^m S_2^j \right) / (n + m),$$

где S_1^i , $i = 1, \dots, n$ вектора из множества первого класса, а S_2^j , $j = 1, \dots, m$ - ко второму классу.

$$\begin{aligned} W &= \left(\sum_{i=1}^3 S_1^i - \sum_{j=1}^4 S_2^j \right) / (3 + 4) = \frac{(2, 4) + (0, 3) + (0, 0) - (-2, 0) - (-1, -5) - (3, -6) - (2, -5)}{7} \\ &= \frac{(2, 0) - (2, -16)}{7} = \frac{(0, 16)}{7} = (0, 2,285) \end{aligned}$$

Начальное значение вектора весов $w_0 = 0; w_1 = 0; w_2 = 2,285$.

3. Произведем обучение персептрона по отдельным примерам, для этого возьмем обучающую выборку и поочередно рассчитаем значение выходного сигнала.

$$\mathbf{S}_1^1 = (2, 4)$$

$$u = 1 \cdot w_0 + s_1 \cdot w_1 + s_2 \cdot w_2 = 0 + 0 + 4 \cdot 2,285 = 9,14$$

$$\varphi(u) = \frac{1}{1 + e^{-au}} = \frac{1}{1 + e^{-9,14}} = 0,9999 \approx 1$$

$$\mathbf{S}_1^2 = (0, 3)$$

$$u = 1 \cdot w_0 + s_1 \cdot w_1 + s_2 \cdot w_2 = 0 + 0 + 3 \cdot 2,285 = 6,857$$

$$\varphi(u) = \frac{1}{1 + e^{-au}} = \frac{1}{1 + e^{-6,857}} = 0,9989 \approx 1$$

$$\mathbf{S}_1^3 = (0, 0)$$

$$u = 1 \cdot w_0 + s_1 \cdot w_1 + s_2 \cdot w_2 = 0 + 0 + 0 \cdot 2,285 = 0$$

$$\varphi(u) = \frac{1}{1 + e^{-au}} = \frac{1}{1 + e^{-0}} = 0,5$$

Производим корректировку весов по формуле (2.2). $\gamma = 0,8$

$$w_0(t+1) = w_0(t) + \gamma s_0 = 0 + 0,8 \cdot 1 = 0,8$$

$$w_1(t+1) = w_1(t) + \gamma s_1 = 0 + 0,8 \cdot 0 = 0$$

$$w_2(t+1) = w_2(t) + \gamma s_2 = 2,285 + 0,8 \cdot 0 = 2,285$$

Теперь $\mathbf{W} = [0,8 \ 0 \ 2,285]$

$$\mathbf{S}_1^1 = (2, 4)$$

$$u = 1 \cdot w_0 + s_1 \cdot w_1 + s_2 \cdot w_2 = 1 \cdot 0,8 + 2 \cdot 0 + 4 \cdot 2,285 = 9,9428$$

$$\varphi(u) = \frac{1}{1 + e^{-au}} = \frac{1}{1 + e^{-9,9428}} = 1$$

$$\mathbf{S}_1^2 = (0, 3)$$

$$u = 1 \cdot w_0 + s_1 \cdot w_1 + s_2 \cdot w_2 = 1 \cdot 0,8 + 0 \cdot 0 + 3 \cdot 2,285 = 7,657$$

$$\varphi(u) = \frac{1}{1 + e^{-au}} = \frac{1}{1 + e^{-7,657}} = 0,9995 \approx 1$$

$$\mathbf{S}_1^3 = (0, 0)$$

$$u = 1 \cdot w_0 + s_1 \cdot w_1 + s_2 \cdot w_2 = 1 \cdot 0,8 + 0 + 0 \cdot 2,285 = 0,8$$

$$\varphi(u) = \frac{1}{1 + e^{-au}} = \frac{1}{1 + e^{-0,8}} = 0,69$$

$$\mathbf{S}_2^1 = (-2, 0)$$

$$u = 1 \cdot w_0 + s_1 \cdot w_1 + s_2 \cdot w_2 = 1 \cdot 0,8 + (-2) \cdot 0 + 0 \cdot 2,285 = 0,8$$

$$\varphi(u) = \frac{1}{1+e^{-au}} = \frac{1}{1+e^{-0,8}} = 0,69$$

Для второй группы данных ответ должен соответствовать ≈ 0 .

Коррекция весов осуществляется по формуле 2.3.

$$w_i(t+1) = w_i(t) - \gamma s_i, \quad \gamma = 0.8$$

$$w_0(t+1) = w_0(t) - \gamma s_0 = 0,8 - 0,8 \cdot 1 = 0$$

$$w_1(t+1) = w_1(t) - \gamma s_1 = 0 - 0,8 \cdot (-2) = 1,6$$

$$w_2(t+1) = w_2(t) - \gamma s_2 = 2,285 - 0,8 \cdot 0 = 2,285$$

Теперь $\mathbf{W} = [0 \ 1,6 \ 2,285]$

$$\mathbf{S}_1^1 = (2, \ 4)$$

$$u = 1 \cdot w_0 + s_1 \cdot w_1 + s_2 \cdot w_2 = 1 \cdot 0 + 2 \cdot 1,6 + 4 \cdot 2,285 = 12,34$$

$$\varphi(u) = \frac{1}{1+e^{-au}} = \frac{1}{1+e^{-12,34}} = 1$$

$$\mathbf{S}_1^2 = (0, \ 3)$$

$$u = 1 \cdot w_0 + s_1 \cdot w_1 + s_2 \cdot w_2 = 1 \cdot 0 + 0 \cdot 1,6 + 3 \cdot 2,285 = 7,157$$

$$\varphi(u) = \frac{1}{1+e^{-au}} = \frac{1}{1+e^{-7,157}} \approx 1$$

$$\mathbf{S}_1^3 = (0, \ 0)$$

$$u = 1 \cdot w_0 + s_1 \cdot w_1 + s_2 \cdot w_2 = 1 \cdot 0 + 0 \cdot 0 + 0 \cdot 2,285 = 0$$

$$\varphi(u) = \frac{1}{1+e^{-au}} = \frac{1}{1+e^{-0}} = 0,5$$

Производим корректировку весов по формуле (2.2). $\gamma = 0.8$

$$w_0(t+1) = w_0(t) + \gamma s_0 = 0 + 0,8 \cdot 1 = 0,8$$

$$w_1(t+1) = w_1(t) + \gamma s_1 = 1,6 + 0,8 \cdot 0 = 1,6$$

$$w_2(t+1) = w_2(t) + \gamma s_2 = 2,285 + 0,8 \cdot 0 = 2,285$$

Теперь $\mathbf{W} = [0,8 \ 1,6 \ 2,285]$

$$\mathbf{S}_1^1 = (2, \ 4)$$

$$u = 1 \cdot w_0 + s_1 \cdot w_1 + s_2 \cdot w_2 = 1 \cdot 0,8 + 2 \cdot 1,6 + 4 \cdot 2,285 = 13,12$$

$$\varphi(u) = \frac{1}{1+e^{-au}} = \frac{1}{1+e^{-13,12}} = 1$$

$$\mathbf{S}_1^2 = (0, \ 3)$$

$$u = 1 \cdot w_0 + s_1 \cdot w_1 + s_2 \cdot w_2 = 1 \cdot 0,8 + 0 \cdot 1,6 + 3 \cdot 2,285 = 7,657$$

$$\varphi(u) = \frac{1}{1+e^{-au}} = \frac{1}{1+e^{-7,657}} = 0,9995 \approx 1$$

$$\mathbf{S}_1^3 = (0, 0)$$

$$u = 1 \cdot w_0 + s_1 \cdot w_1 + s_2 \cdot w_2 = 1 \cdot 0,8 + 0 \cdot 1,6 + 0 \cdot 2,285 = 0,8$$

$$\varphi(u) = \frac{1}{1 + e^{-au}} = \frac{1}{1 + e^{-0,8}} = 0,69$$

$$\mathbf{S}_2^1 = (-2, 0)$$

$$u = 1 \cdot w_0 + s_1 \cdot w_1 + s_2 \cdot w_2 = 1 \cdot 0,8 + (-2) \cdot 1,6 + 0 \cdot 2,285 = -2,4$$

$$\varphi(u) = \frac{1}{1 + e^{-au}} = \frac{1}{1 + e^{2,4}} = 0,08 \approx 0$$

$$\mathbf{S}_2^2 = (-1, -5)$$

$$u = 1 \cdot w_0 + s_1 \cdot w_1 + s_2 \cdot w_2 = 1 \cdot 0,8 + (-1) \cdot 1,6 + (-5) \cdot 2,285 = -12,228$$

$$\varphi(u) = \frac{1}{1 + e^{-au}} = \frac{1}{1 + e^{12,228}} \approx 0$$

$$\mathbf{S}_2^3 = (3, -6)$$

$$u = 1 \cdot w_0 + s_1 \cdot w_1 + s_2 \cdot w_2 = 1 \cdot 0,8 + 3 \cdot 1,6 + (-6) \cdot 2,285 = -8,114$$

$$\varphi(u) = \frac{1}{1 + e^{-au}} = \frac{1}{1 + e^{8,114}} \approx 0$$

$$\mathbf{S}_2^4 = (2, -5)$$

$$u = 1 \cdot w_0 + s_1 \cdot w_1 + s_2 \cdot w_2 = 1 \cdot 0,8 + 2 \cdot 1,6 + (-5) \cdot 2,285 = -7,428$$

$$\varphi(u) = \frac{1}{1 + e^{-au}} = \frac{1}{1 + e^{7,428}} \approx 0$$

Сеть обучена.

Ответ: $\mathbf{W} = [0,8 \ 1,6 \ 2,285]$ $0,8 + 1,6 \cdot s_1 + 2,285 \cdot s_2 = 0$ уравнение разделяющей
гиперплоскости.

Задача 3.

Задан набор точек S_1 , принадлежащий первому классу (выходной сигнал – 1) и набор точек S_2 , принадлежащий второму классу (выходной сигнал – 0). Перенести множество точек в пространство с большей размерностью.

Дано:

$$S_1 = [(-2, 0); (-1, 4); (1, 1)]$$

$$S_2 = [(-3, -5); (-1, 7); (7, 0)]$$

Найти: перенос вектора в пространство с большей размерности

Решение:

1. Построим точки в координатной плоскости

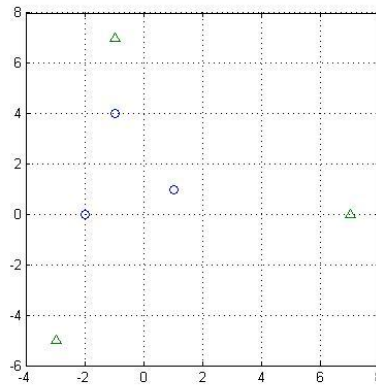


Рис. 5.2 Исходные данные в двумерном пространстве

В данном случае невозможно линейное разделение классов. Следовательно, возможно два варианта: построение многослойного персептрона или радиальной сети.

2. Применим радиальные функции для перехода из 2-мерного пространства в 3-мерное с координатами $(\varphi_1, \varphi_2, \varphi_3)$.

$$\varphi_1(\|s - c_1\|) = \exp\left(-\frac{\|s - c_1\|^2}{2\sigma_1^2}\right) \quad \varphi_2(\|s - c_2\|) = \exp\left(-\frac{\|s - c_2\|^2}{2\sigma_2^2}\right) \quad \varphi_3(\|s - c_3\|) = \exp\left(-\frac{\|s - c_3\|^2}{2\sigma_3^2}\right)$$

Пусть $\sigma_1 = \sigma_2 = \sigma_3 = \frac{1}{\sqrt{2}}$, а центры c_1, c_2, c_3 соответствуют координатам точек, принадлежащих первому классу.

Входной вектор	φ_1	φ_2	φ_3
$(-2, 0)$	1	0	0
$(-1, 4)$	0	1	0
$(1, 1)$	0	0	1
$(-3, -5)$	0	0	0
$(-1, 7)$	0	0	0
$(7, 0)$	0	0	0

Изобразим полученные точки в пространстве.

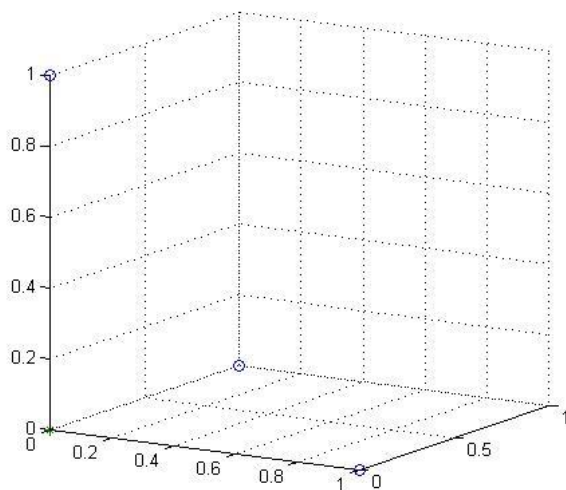
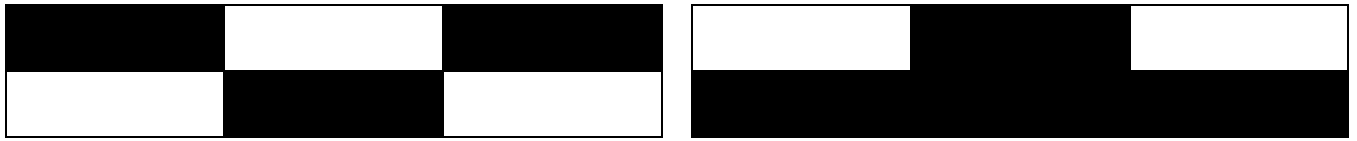


Рис. 5.3. Точки в трехмерном пространстве

Задача 4.

Дано два растровых черно-белых изображений. Обучить сеть на распознавание этих изображений. Выполнить проверку.



Сеть Хемминга

Обучающую выборку можно представить в следующем виде (черный пиксель +1, белый пиксель 0):

$$\mathbf{S}_1 = [+1 \ 0 \ +1 \ 0 \ +1 \ 0]$$

$$\mathbf{S}_2 = [0 \ +1 \ 0 \ +1 \ +1 \ +1]$$

Весовые коэффициенты первого слоя выбираются по следующему правилу:

$$w_{ik}^{(1)} = \frac{x_i^k}{2}, \text{ для } i=1,2,\dots,n, \ k=1,2,\dots,m$$

$$m=2, \ n=6.$$

$$T_k = \frac{n}{2} = \frac{6}{2} = 3, \ k=1,2,\dots,m.$$

$$\mathbf{W}_1^{(1)} = [w_1, w_2, \dots, w_n] = \left[\frac{\mathbf{S}_1}{2} \right] = \left[\frac{s_{11}}{2}, \frac{s_{21}}{2}, \dots, \frac{s_{n1}}{2} \right] = [0,5, \ 0, \ 0,5, \ 0, \ 0,5, \ 0]$$

$$\mathbf{W}_2^{(1)} = [w_1, w_2, \dots, w_n] = \left[\frac{\mathbf{S}_2}{2} \right] = \left[\frac{s_{12}}{2}, \frac{s_{22}}{2}, \dots, \frac{s_{n2}}{2} \right] = [0, \ 0,5, \ 0, \ 0,5, \ 0,5, \ 0,5]$$

Весовые коэффициенты второго слоя выбираются следующим образом:

$$w_{ij}^{(2)} = \begin{bmatrix} 1 & i=j \\ -\varepsilon & i \neq j \end{bmatrix} 0 < \varepsilon < \frac{1}{m}.$$

$$\mathbf{W}^{(2)} = \begin{bmatrix} 1 & -\varepsilon \\ -\varepsilon & 1 \end{bmatrix}, \ \varepsilon = 0.1$$

$$\mathbf{W}_1^{(2)} = [1 \ -0,1]$$

$$\mathbf{W}_2^{(2)} = [-0,1 \ 1]$$

Выполним проверку:

Подадим на вход вектор $\mathbf{S} = [0 \ 1 \ 0 \ 1 \ 1 \ 1]$.

На выходах нейронов первого слоя получим отклик $\mathbf{Y}^{(1)} = [y_1^{(1)}, y_2^{(1)}]$

$$y_j^{(1)} = \sum_{i=1}^n w_{ij}^{(1)} x_i + T_j$$

$$y_1^{(1)} = \sum_{i=1}^n w_{i1}^{(1)} x_i + T_1 = 0,5 \cdot 0 + 0 \cdot 1 + 0,5 \cdot 0 + 0 \cdot 1 + 0,5 \cdot 1 + 0 \cdot 0,5 + 3 = 3,5$$

$$y_2^{(1)} = \sum_{i=1}^n w_{i2}^{(1)} x_i + T_2 = 0 \cdot 0 + 0,5 \cdot 1 + 0 \cdot 0 + 0,5 \cdot 1 + 0,5 \cdot 1 + 0,5 \cdot 1 + 3 = 5$$

Зададим порог насыщения $L = 10000$. Вычислим состояния нейронов второго слоя.

$$s_j^{(2)}(t+1) = y_j^{(2)} - \varepsilon \sum_{i \neq j}^m y_i^{(2)}(t),$$

$$y_j^{(2)}(t+1) = f(s_j(t+1)) = \begin{cases} 4 \cdot s_j(t+1), & 0 < s_j(t+1) < L \\ 10000, & s_j(t+1) \geq L \end{cases}.$$

$t = 0$

$$s_1^{(2)}(0) = y_1^{(1)}$$

$$y_1^{(2)}(0) = f(s_1^{(2)}(0)) = 4 \cdot 3,5 = 14.$$

$$s_2^{(2)}(0) = y_2^{(1)}$$

$$y_2^{(2)}(0) = f(s_2^{(2)}(0)) = 4 \cdot 5 = 20.$$

$t = 1$

$$s_1^{(2)}(0+1) = y_1^{(2)} - \varepsilon \sum_{i \neq j}^m y_i^{(2)}(0) = 14 - 0,1 \cdot 14 = 14 - 1,4 = 12,6,$$

$$y_1^{(2)}(0+1) = f(s_1(0+1)) = 4 \cdot 8,4 = 33,6.$$

$$s_2^{(2)}(0+1) = y_2^{(2)} - \varepsilon \sum_{i \neq j}^m y_i^{(2)}(0) = 20 - 0,4 \cdot 20 = 20 - 8 = 12,$$

$$y_2^{(2)}(0+1) = f(s_2(0+1)) = 4 \cdot 12 = 48.$$

$t = 2$

$$s_1^{(2)}(1+1) = y_1^{(2)} - \varepsilon \sum_{i \neq j}^m y_i^{(2)}(1) = 33,6 - 0,4 \cdot 33,6 = 33,6 - 13,44 = 20,16,$$

$$y_1^{(2)}(1+1) = f(s_1(1+1)) = 4 \cdot 20,16 = 80,64.$$

$$s_2^{(2)}(1+1) = y_2^{(2)} - \varepsilon \sum_{i \neq j}^m y_i^{(2)}(1) = 48 - 0,4 \cdot 48 = 48 - 19,2 = 28,8,$$

$$y_2^{(2)}(1+1) = f(s_2(1+1)) = 4 \cdot 28,8 = 115,2.$$

$t = 3$

$$s_1^{(2)}(2+1) = y_1^{(2)} - \varepsilon \sum_{i \neq j}^m y_i^{(2)}(2) = 80,64 - 0,4 \cdot 80,64 = 80,64 - 32,256 = 48,384,$$

$$y_1^{(2)}(2+1) = f(s_1(2+1)) = 4 \cdot 48,384 = 193,536.$$

$$s_2^{(2)}(2+1) = y_2^{(2)} - \varepsilon \sum_{i \neq j}^m y_2^{(2)}(2) = 115,2 - 0,4 \cdot 115,2 = 115,2 - 46,08 = 69,12 ,$$

$$y_2^{(2)}(2+1) = f(s_2(2+1)) = 4 \cdot 69,12 = 276,48 .$$

$$t = 4$$

$$s_1^{(2)}(3+1) = y_1^{(2)} - \varepsilon \sum_{i \neq j}^m y_1^{(2)}(3) = 193,536 - 0,4 \cdot 193,536 = 193,536 - 77,4144 = 116,1216 ,$$

$$y_1^{(2)}(3+1) = f(s_1(3+1)) = 4 \cdot 116,1216 = 464,4864 .$$

$$s_2^{(2)}(3+1) = y_2^{(2)} - \varepsilon \sum_{i \neq j}^m y_2^{(2)}(3) = 276,48 - 0,4 \cdot 276,48 = 276,48 - 110,592 = 165,888 ,$$

$$y_2^{(2)}(3+1) = f(s_2(3+1)) = 4 \cdot 165,888 = 663,552 .$$

$$t = 5$$

$$s_1^{(2)}(4+1) = y_1^{(2)} - \varepsilon \sum_{i \neq j}^m y_1^{(2)}(4) = 464,48 - 0,4 \cdot 464,48 = 464,48 - 185,792 = 278,688 ,$$

$$y_1^{(2)}(4+1) = f(s_1(4+1)) = 4 \cdot 278,688 = 1114,752 .$$

$$s_2^{(2)}(4+1) = y_2^{(2)} - \varepsilon \sum_{i \neq j}^m y_2^{(2)}(4) = 663,552 - 0,4 \cdot 663,552 = 663,552 - 265,4208 = 398,1312 ,$$

$$y_2^{(2)}(4+1) = f(s_2(4+1)) = 4 \cdot 398,1312 = 1592,525 .$$

$$t = 6$$

$$s_1^{(2)}(5+1) = y_1^{(2)} - \varepsilon \sum_{i \neq j}^m y_1^{(2)}(5) = 1114,752 - 0,4 \cdot 1114,752 = 1114,752 - 445,9 = 668,85 ,$$

$$y_1^{(2)}(5+1) = f(s_1(5+1)) = 4 \cdot 668,85 = 2675,405 .$$

$$s_2^{(2)}(5+1) = y_2^{(2)} - \varepsilon \sum_{i \neq j}^m y_2^{(2)}(5) = 1592,525 - 0,4 \cdot 1592,525 = 1592,525 - 637 = 955,52 ,$$

$$y_2^{(2)}(5+1) = f(s_2(5+1)) = 4 \cdot 955,52 = 3822,06 .$$

$$t = 7$$

$$s_1^{(2)}(6+1) = y_1^{(2)} - \varepsilon \sum_{i \neq j}^m y_1^{(2)}(6) = 2675,405 - 0,4 \cdot 2675,405 = 2675,405 - 1070,162 = 1605,243 ,$$

$$y_1^{(2)}(6+1) = f(s_1(6+1)) = 4 \cdot 1605,243 = 6420,97 .$$

$$s_2^{(2)}(6+1) = y_2^{(2)} - \varepsilon \sum_{i \neq j}^m y_2^{(2)}(6) = 3822,06 - 0,4 \cdot 3822,06 = 3822,06 - 1528,82 = 2293,23 ,$$

$$y_2^{(2)}(6+1) = f(s_2(6+1)) = 4 \cdot 2293,23 = 9172,94 .$$

$$t = 8$$

$$s_1^{(2)}(7+1) = y_1^{(2)} - \varepsilon \sum_{i \neq j}^m y_1^{(2)}(7) = 6420,97 - 0,4 \cdot 6420,97 = 6420,97 - 2568,388 = 3852,582 ,$$

$$y_1^{(2)}(7+1) = f(s_1(7+1)) = 4 \cdot 3852,582 = 15410,328 .$$

$$s_2^{(2)}(7+1) = y_2^{(2)} - \varepsilon \sum_{i \neq j}^m y_2^{(2)}(7) = 9172,94 - 0,4 \cdot 9172,94 = 9172,94 - 3669,177 = 5503,76 ,$$

$$y_2^{(2)}(7+1) = f(s_2(7+1)) = 4 \cdot 5503,76 = 22015,05 .$$

$$t = 9$$

$$s_1^{(2)}(8+1) = y_1^{(2)} - \varepsilon \sum_{i \neq j}^m y_1^{(2)}(8) = 15410 - 0,4 \cdot 15410 = 15410 - 6164 = 9246 ,$$

$$y_1^{(2)}(8+1) = f(s_1(8+1)) = 4 \cdot 9246 = 36984 .$$

$$s_2^{(2)}(8+1) = y_2^{(2)} - \varepsilon \sum_{i \neq j}^m y_2^{(2)}(8) = 22015,05 - 0,4 \cdot 22015,05 = 22015,05 - 8806,02 = 13208,98 ,$$

$$y_2^{(2)}(8+1) = f(s_2(8+1)) = 10000 .$$

$$t = 10$$

$$s_1^{(2)}(9+1) = y_1^{(2)} - \varepsilon \sum_{i \neq j}^m y_1^{(2)}(9) = 36984 - 0,4 \cdot 36984 = 36984 - 14793,6 = 22190,4 ,$$

$$y_1^{(2)}(9+1) = f(s_1(9+1)) = 10000 .$$

$$s_2^{(2)}(9+1) = y_2^{(2)} - \varepsilon \sum_{i \neq j}^m y_2^{(2)}(9) = 10000 - 0,4 \cdot 10000 = 10000 - 4000 = 6000 ,$$

$$y_2^{(2)}(8+1) = f(s_2(8+1)) = 4 * 6000 = 24000 .$$

Ответ: подавался образ №2.

ЗАДАНИЯ НА КОНТРОЛЬНУЮ РАБОТУ

Задача 1.

Нейрон j получает на вход сигнал, уровни которого равны \mathbf{X} . Соответствующие веса связей равны \mathbf{W} . Вычислить выходное значение нейрона j для заданных моделей нейрона. Номер варианта зависит от двух последних цифр зачетной книжки. Исходные данные приведены в таблице 1.

Задача 2.

Задан набор точек \mathbf{S}_1 , принадлежащий первому классу (выходной сигнал – 1) и набор точек \mathbf{S}_2 , принадлежащий второму классу (выходной сигнал – 0). Составить уравнение гиперплоскости, линейно разделяющее пространство признаков на две части (найти вектор \mathbf{W} и w_0). Данные даны в таблице 2.

Задача 3.

Задан набор точек \mathbf{S}_1 , принадлежащий первому классу (выходной сигнал – 1) и набор точек \mathbf{S}_2 , принадлежащий второму классу (выходной сигнал – 0). Перенести множество точек в пространство с большей размерностью. Исходные данные даны в таблице 3.

Задача 4.

Дано два растровых черно-белых изображений. Обучить сеть на распознавание этих изображений. Выполнить проверку. Изображения и вид сети даны в таблице 4.

Таблица 1.

№ п/п	X	W	Модель	№ п/п	X	W	Модель
01/51	[3 -1 4,3 -0,9]	[0,8 0,2 -0,2 1] $w_0=-1,8$	Адалайн	26/76	[6 -3,1 0,4 0,9 -1,3]	[0,85 -1,3 2,1 -0,4 0,98]	Адалайн
02/52	[4,5 -0,5 4,1 -2,3 0]	[-0,7 -0,4 1,6 2,1 -1]	Сигмоидальная биполярная $\alpha=1$	27/77	[-1,27 -4 0,4 0,8 -3,4]	[1,2 1,18 -0,05 0,3 0,17] $w_0=1,8$	Сигмоидальная биполярная $\alpha=0,8$
03/53	[4,1 -0,3 1,24 -2 2 1]	[0,1 -1 -1,2 0,3 0,4 -0,5]	Мак-Каллока-Питца	28/78	[-0,75 2,9 -2,35 8,7 -0,5 0,45]	[1,6 0,05 -0,1 -0,8 0,29 -1,33]	Линейный $\alpha=2$
04/54	[-7,1 3 0,5 1,6]	[0,02 -0,7 2,3 1,7]	Сигмоидальная унополярная $\alpha=0,5$	29/79	[4,3 9,74 -4,15 5,1 7,52]	[-1,5 -1,44 0,57 -0,4 0,69]	Мак-Каллока-Питца
05/55	[-1,29 -4,99 0,37 0,86 - 3,43]	[0,5 1,69 0,5 -0,64 0,38]	Адалайн	30/80	[3,2 2,8 5,16 2,6 4,76 - 4,8]	[-0,01 -0,15 -1,6 0,25 -1,1 1,4]	Сигмоидальная унополярная $\alpha=0,9$
06/56	[4,3 9,7 -4,15 5,15 7,5]	[-1,59 -1,4 0,5 -0,4 0,69] $w_0=-0,87$	Сигмоидальная биполярная $\alpha=0,5$	31/81	[-1,6 1,05 0,43 -1,8]	[-2,17 -0,05 -1,01 0,61]	Адалайн
07/57	[-0,02 -0,05 0 -0,31 1,09 -1,87]	[-0,2 -0,15 -1,6 0,25 -1 1,41]	Мак-Каллока-Питца	32/82	[0,67 0,56 -0,2 -0,37 - 0,29]	[-1,47 -0,23 0,12 0,3 1,44]	Линейный $\alpha=4$
08/58	[0,23 -1 -0,74 1,08 -0,13 0,38]	[-0,35 0,62 0,79 0,94 -0,9 0,21]	Сигмоидальная унополярная $\alpha=1$	33/83	[-0,5 -0,04 0 -0,3 1,09 - 1,87]	[0,5 1,69 0,59 -0,64 0,38 - 1,01]	Сигмоидальная унополярная $\alpha=0,5$
09/59	[-0,08 -2,25 -2,7 -0,52 1,9]	[0,23 0,02 -1 -0,95 -0,37]	Адалайн	34/84	[0,25 1,3 -2,33 -0,9 - 0,52 -2,42]	[-1,3 0,93 0,01 -0,64 0,8 0,2]	Мак-Каллока-Питца
10/60	[-3,28 -1 -4,75 -8,8 3,94 - 2,1 1,3]	[0,08 -0,6 -0,5 0,44 -0,9 0,78 0,56]	Мак-Каллока-Питца	35/85	[-0,98 1,34 0,28 1,4 1,13]	[0,66 -0,85 -1,2 -0,12 -0,06]	Сигмоидальная биполярная $\alpha=0,7$
11/61	[0,48 -0,6 -0,14 -0,43]	[0,46 -0,9 0,03 -0,62]	Сигмоидальная унополярная $\alpha=0,4$	36/86	[-2,73 -5,16 -0,3 -1,3 - 3,37]	[0,49 1,48 -0,5 -0,84 -0,2]	Адалайн
12/62	[1,86 -0,3 -1,14 -0,2 1,2]	[-1,35 -1,6 1,03 -0,3 -1,3]	Сигмоидальная биполярная $\alpha=0,2$	37/87	[-1,39 1,11 2,1 6,33]	[0,13 1,59 1,01 -1,4]	Линейный $\alpha=1$
13/63	[-1,2 -4,5 -2,4 2,1 -0,22 0,75]	[0,17 -4 -0,98 0,92 -0,64 2,4]	Адалайн	38/88	[-3,36 1,9 -1,8 1,65 - 3,29]	[0,31 1,55 0,7 1,95 0,5]	Сигмоидальная унополярная $\alpha=0,4$
14/64	[0,5 0,55 -0,27 -2,05]	[-0,07 -0,68 -1,02 -1,23]	Линейный $\alpha=2$	39/89	[1,88 -4,24 -1,28 -1,4 - 2,03 -0,36]	[1,52 -0,03 1,22 -0,69 0,01 - 0,78]	Мак-Каллока-Питца
15/65	[0,26 1,21 -0,27 -0,13 - 1,27]	[-1,66 -0,7 0,28 -0,54 -1,33]	Мак-Каллока-Питца	40/90	[-3,76; -1,04; -2,82; - 3,52; -3,06]	[-0,4; 0,17; -0,11; 1,06; -0,24]	Адалайн
16/66	[9,23 2,09 -0,04 3,65 0,28 -4,42]	[0,58 -0,25 0,48 0,66 -0,07 0,82]	Сигмоидальная унополярная $\alpha=0,5$	41/91	[1,18 -0,79 -4,9 -3,08 0,72]	[0,26 1,21 -0,2 -0,13 -1,2]	Сигмоидальная биполярная $\alpha=0,5$
17/67	[1 2,55 -1,09 0,52]	[-1,5 0,009 0,07 0,31]	Адалайн	42/92	[-1,17 0,86 0,04 -0,75 - 0,1 -2]	[-0,2 -0,02 0,27 1,05 0,62 - 1,75]	Линейный $\alpha=4$
18/68	[1,18 -0,79 -4,9 -3 0,72]	[-0,4 0,17 -0,11 1,06 -0,2]	Сигмоидальная биполярная $\alpha=0,1$	43/93	[-0,01 0,53 -0,7 -0,66 0,34 0,1 1,84]	[0,62 1,26 -0,8 0,13 -0,13 - 1,16 1,18]	Адалайн
19/69	[1,07 -0,71 -0,01 0 -0,24]	[0,26 1,21 -0,2 -0,13 -1,2]	Адалайн	44/94	[0,59 -0,25 0,48 0,66 - 0,07 0,88]	[0,45 -0,005 -0,27 1,27 1,86 - 0,52]	Мак-Каллока-Питца
20/70	[0,4 -0,05 -0,2 1,27 1,86 -	[0,14 -0,8 0,68 -2,36 0,99]	Мак-Каллока-Питца	45/95	[-2,8 3,54 0,65 -2,23]	[-1,5 0,07 0,17 0,31]	Сигмоидальная

	0,52]	0,21]					унополярная $\alpha=0,3$
21/71	[-0,27 2,26 1,5 -1,94 -1,68 -0,57 -0,18]	[-0,01 0,53 -0,71 -0,65 0,31 0,1 1,84]	Линейный $\alpha=3$	46/96	[1,07 -0,71 -0,01 -0,8 -0,24]	[-1,6 -0,7 0,2 -0,54 -1,3]	Адалайн
22/72	[0,89 0,83 -0,7 -0,72]	[-1,51 0,01 0,07 0,31]	Мак-Каллока-Питца	47/97	[0,9 2,56 -1,09 0,52]	[-2,81 3,54 0,65 -2,23]	Сигмоидальная биполярная $\alpha=0,1$
23/73	[-0,17 -0,95 1,29 0,49 1,28 -0,47]	[-1,08 -0,04 0,37 -0,33 -0,4 -0,03]	Сигмоидальная унополярная $\alpha=0,5$	48/98	[1 -0,8 0,68 -2,36 0,99 0,21]	[0,58 -0,25 0,48 0,66 -0,07 0,88]	Линейный $\alpha=1$
24/74	[9,23 2,09 -0,04 3,65 0,22 -4,42]	[0,48 -0,005 -0,27 1,27 1,86 -0,52]	Сигмоидальная биполярная $\alpha=1$	49/99	[1,08 -0,98 -0,68 1,33 -0,9]	[-0,41 -0,5 1,67 0,08 -1,08]	Мак-Каллока-Питца
25/75	[-3,37 5,2 5,81 4,9 -3,76]	[1,08 -0,98 -0,68 1,33 -0,9]	Адалайн	50/00	[-0,21 -0,19 0,35 -0,57 -0,97 -0,44 1,08]	[2,37 0,22 -0,26 0,7 -0,48 1,86 1,1]	Сигмоидальная унополярная $\alpha=0,8$

Таблица 2.

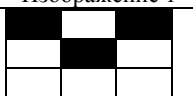
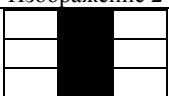

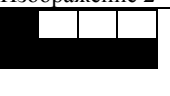
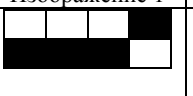
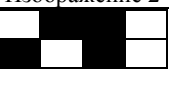

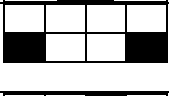
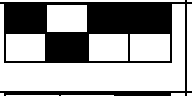
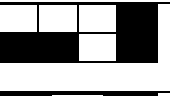
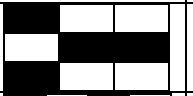
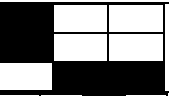
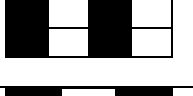
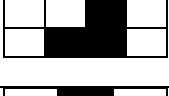
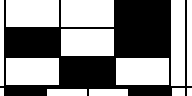
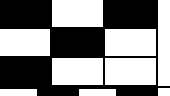


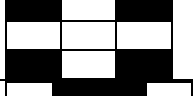
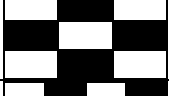

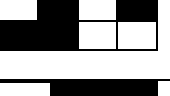
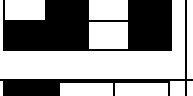


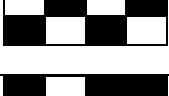
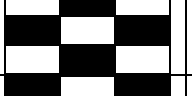
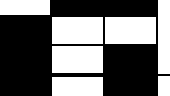
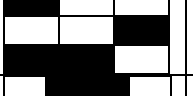
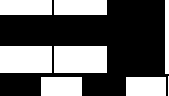


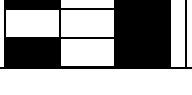
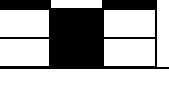
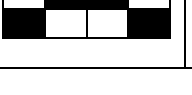

№ п/п	S_1	S_2	№ п/п	S_1	S_2
01/31/61/91	(-2 4); (0 8); (-6 -1,5); (-3 0); (1 2)	(-2 -3); (7 -2); (7 2,5); (-1 -5)	16/46/76	(0,7 0,7); (-2 -1,5); (-3,5 -2,5)	(1,5 0,5); (1,5 -3); (3,5 -0,5)
02/32/62/92	(6 3); (-1 -3); (-4 1)	(6 4); (7 2); (0 5)	17/47/77	(-13 -1); (-3 -7); (11 -7)	(-2,5 6,6); (-10 13); (3 15)
03/33/63/93	(-5 3); (6 -4); (-1 8)	(4 6); (3,5 1); (6 3)	18/48/78	(-2 20); (0,5 -6); (5 5,5)	(-1,5 -4,5); (-6 -13); (-1 15)
04/34/64/94	(-2 -4); (0 -8); (-2 -1,5); (1 -3); (-1 -5)	(0 5); (-3 5); (6 -7); (3 2)	19/49/79	(1 0,7); (-2,7 -1); (-4 -7,3)	(2,6 1,8); (0 -3,5); (3,5 -2)
05/35/65/95	(1,6 2); (2,7 2,2); (2,4 2)	(-3 1,5); (4 -2,5); (-1,5 3); (0 -5); (-3 -2)	20/50/80	(-4,5 3); (-4 -1,5); (-1 -4)	(0 0); (5 -1); (-1 5); (1 3)
06/36/66/96	(-16 2,5); (-14,5 5,7); (-7 8); (8 7)	(-0,1 -1,5); (13 6); (12 -12); (10 -5)	21/51/81	(5,5 -2,5); (1 3); (7 5,5); (4 9)	(0 -3); (-3,5 -3); (-14 -2); (10 -10)
07/37/67/97	(-2 -4); (3 -6); (-2 -1,5); (1 -2)	(0 5); (-3 5); (6 -1); (3 2)	22/52/82	(-2 7); (2,4 5,4); (3,5 0); (4 1,5)	(-0,5 0,5); (-0,5 -3); (1 -4,5)
08/38/68/98	(-10 6); (5 10); (-0,2 -0,5)	(-15 -3); (6 -6,5); (2 -9); (4 -10)	23/53/83	(3,5 1); (-6,5 -5,5); (2 2,5)	(10 -1); (4,5 -10,5); (9,5 -10); (14 -3,5)
09/39/69/99	(1 -4,5); (-0,3 -2,7); (-0,5 0,5); (-2 7)	(2,5 5,5); (3,5 0); (6 1,5); (2 2)	24/54/84	(-5,5 1); (0 -0,5); (-2,5 2); (-3,5 5)	(1 -3); (4 -4); (-4 -1)
10/40/70/00	(-1 -4); (3 -6); (2 -5); (1 -2)	(0 5); (-3 5); (-6 -1); (-3 2); (-3 -5)	25/55/85	(3,5 2,5); (8 5,5); (14 0,5); (0 -10)	(-8 -2); (-10 -10); (-15 -20)
11/41/71	(-2,5 2); (-5,5 1); (-3,5 5)	(0 -0,5); (1 -3); (2 3)	26/56/86	(-7 0); (-3,5 2); (1,5 5,5)	(0 0); (2 -2); (1,5 -3); (-1 -4)
12/42/72	(-7 0); (-3 2); (2,5 1,5)	(0 0); (1,5 -3); (2 -2)	27/57/87	(0 2,5); (1,5 3); (2,5 2); (-0,5 3,5)	(2 -1); (0 -1); (3,5 -6)
13/43/73	(10 -4); (12 -11); (15 7); (18 -4)	(6 -6); (5,5 -11); (-11 -2)	28/58/88	(-13 8); (-12 0); (-2,5 -12)	(0 -6,5); (1 6); (8 2)
14/44/74	(-1 -4); (3 -6); (2 -5); (1 -2)	(-5 5); (-3 5); (-6 -1); (-3 2); (-3 -5)	29/59/89	(0,5 1); (-1 -1); (-5 -1); (2 2,5)	(3 -3); (5 -1); (-2 0)
15/45/75	(-5 -1); (-1,5 -1); (0,5 1)	(2 2,5); (5 -1); (3 -3)	30/60/90	(-4,5 3,5); (-10 -12); (-13 -10)	(1,5 15); (0,5 3,5); (-0,5 -7)

Таблица 3.

№ п/п	S_1	S_2	Вид радиальной функции	№ п/п	S_1	S_2	Вид радиальной функции
01/31/61/91	(-1, 1); (2, -2); (1, 0)	(-5, 3); (4, -5); (3, 3)	Мульти-квадратичная	16/46/76	(-1, 1); (2, -2); (1, 0)	(-5, 3); (4, -5); (3, 3)	Обратная мультиквадратичная
02/32/62/92	(0, -2); (2, -1); (0,5, 1,5)	(-3, -2); (0, 3); (4, -3)	Мульти-квадратичная	17/47/77	(0, -2); (2, -1); (0,5, 1,5)	(-3, -2); (0, 3); (4, -3)	Обратная мультиквадратичная

03/33/63/93	(3, -1); (4, -3); (5, 0)	(0, 0); (7, -4); (5, 2)	Мульти- квадратичная	18/48/78	(3, -1); (4, -3); (5, 0)	(0, 0); (7, -4); (5, 2)	Обратная мультиквадратичная
04/34/64/94	(-3, -3); (-4, -2); (-5, -4)	(0, 0); (-7, -5); (-5, 0)	Мульти- квадратичная	19/49/79	(-3, -3); (-4, -2); (-5, -4)	(0, 0); (-7, -5); (-5, 0)	Обратная мультиквадратичная
05/35/65/95	(-1, 0); (1, 0); (0, -1)	(0, -3); (-3, 0); (2, 2)	Мульти- квадратичная	20/50/80	(-1, 0); (1, 0); (0, -1)	(0, -3); (-3, 0); (2, 2)	Обратная мультиквадратичная
06/36/66/96	(0, -1); (1, 1); (1, -1)	(0, 2); (-3, 1); (-3, 1)	Мульти- квадратичная	21/51/81	(0, -1); (1, 1); (1, -1)	(0, 2); (-3, 1); (-3, 1)	Обратная мультиквадратичная
07/37/67/97	(0, 0); (-2, 0); (2, 1)	(-3, 2); (2, -3); (2, 2)	Мульти- квадратичная	22/52/82	(0, 0); (-2, 0); (2, 1)	(-3, 2); (2, -3); (2, 2)	Обратная мультиквадратичная
08/38/68/98	(1, 5); (2, 4); (-1, 4.5)	(0, 3); (-2, 6); (3, 5)	Мульти- квадратичная	23/53/83	(1, 5); (2, 4); (-1, 4.5)	(0, 3); (-2, 6); (3, 5)	thin-plate spline
09/39/69/99	(-1, 1); (2, -2); (1, 0)	(-5, 3); (4, -5); (3, 3)	Гаусса	24/54/84	(-1, 1); (2, -2); (1, 0)	(-5, 3); (4, -5); (3, 3)	thin-plate spline
10/40/70/00	(0, -2); (2, -1); (0.5, 1.5)	(-3, -2); (0, 3); (4, -3)	Гаусса	25/55/85	(0, -2); (2, -1); (0.5, 1.5)	(-3, -2); (0, 3); (4, -3)	thin-plate spline
11/41/71	(3, -1); (4, -3); (5, 0)	(0, 0); (7, -4); (5, 2)	Гаусса	26/56/86	(3, -1); (4, -3); (5, 0)	(0, 0); (7, -4); (5, 2)	thin-plate spline
12/42/72	(-3, -3); (-4, -2); (-5, -4)	(0, 0); (-7, -5); (-5, 0)	Гаусса	27/57/87	(-3, -3); (-4, -2); (-5, -4)	(0, 0); (-7, -5); (-5, 0)	thin-plate spline
13/43/73	(-1, 0); (1, 0); (0, -1)	(0, -3); (-3, 0); (2, 2)	Гаусса	28/58/88	(-1, 0); (1, 0); (0, -1)	(0, -3); (-3, 0); (2, 2)	thin-plate spline
14/44/74	(0, -1); (1, 1); (1, -1)	(0, 2); (-3, 1); (-3, 1)	Гаусса	29/59/89	(0, -1); (1, 1); (1, -1)	(0, 2); (-3, 1); (-3, 1)	thin-plate spline
15/45/75	(0, 0); (-2, 0); (2, 1)	(-3, 2); (2, -3); (2, 2)	Гаусса	30/60/90	(0, 0); (-2, 0); (2, 1)	(-3, 2); (2, -3); (2, 2)	thin-plate spline

Таблица 4.

№ п/п	Изображение 1	Изображение 2	Вид сети	№ п/п	Изображение 1	Изображение 2	Вид сети	№ п/п	Изображение 1	Изображение 2	Вид сети
01/51			Хопфилд	18/68			Хемминг	35/85			Хопфилд
02/52			Хемминг	19/69			Хопфилд	36/86			Хемминг
03/53			Хопфилд	20/70			Хемминг	37/87			Хопфилд
04/54			Хемминг	21/71			Хопфилд	38/88			Хемминг
05/55			Хопфилд	22/72			Хемминг	39/89			Хопфилд
06/56			Хемминг	23/73			Хопфилд	40/90			Хемминг

07/57			Хопфилд	24/74			Хемминг	41/91			Хопфилд
08/58			Хемминг	25/75			Хопфилд	42/92			Хемминг
09/59			Хопфилд	26/76			Хемминг	43/93			Хопфилд
10/60			Хемминг	27/77			Хопфилд	44/94			Хемминг
11/61			Хопфилд	28/78			Хемминг	45/95			Хопфилд
12/62			Хемминг	29/79			Хопфилд	46/96			Хемминг
13/63			Хопфилд	30/80			Хемминг	47/97			Хопфилд
14/64			Хемминг	31/81			Хопфилд	48/98			Хемминг
15/65			Хопфилд	32/82			Хемминг	49/99			Хопфилд
16/66			Хемминг	33/83			Хопфилд	50/00			Хемминг
17/67			Хопфилд	34/84			Хемминг				

Вопросы по контрольной работе для самостоятельной проверки:

1. Дать определение нейрокомпьютера. В чем его отличие от других ЭВМ?
2. Воспроизведите схему абстрактного нейрокомпьютера и поясните принцип его работы.
3. Какие задачи может решать нейронная сеть?
4. Какие топологии нейронных сетей вам известны?
5. Что такое формальный нейрон? Какие модели формального нейрона вам известны?
6. Что такое функция активации? Приведите графики этих функций?
7. Что такое однослойная нейронная сеть? Как нейроны могут быть соединены между собой?
8. В чем заключается задача линейного разделения двух классов?
9. Что такое многослойная сеть? Принцип ее организации.
10. В чем заключается задача нелинейного разделения двух и более классов?
11. Что такое радиальная базисная сеть?
12. Сколько слоев в радиальной базисной сети? Поясните их назначение.
13. Перечислите радиальные базисные функции. Дайте их характеристику.
14. Сформулируйте алгоритм обучения радиальной базисной сети.
15. Назовите достоинства и недостатки радиальных базисных сетей.
16. В чем отличие многослойных персептронов и радиальных нейронных сетей?
17. Что такое рекуррентный нейрон? Принцип его работы.
18. Чем отличаются рекуррентные нейронные сети от сетей персептронного типа?
19. Чем сеть Хемминга отличается от сети Хопфилда?
20. Приведите алгоритм обучения сети Хемминга.
21. Приведите алгоритм обучения сети Хопфилда.

Литература:

Основная:

1. Комарцова Л.Г., Максимов А.В. Нейрокомпьютеры: учебное пособие для вузов. – 2-е изд., перераб. И доп. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2004. – 400 с.
2. Хайкин С. Нейронные сети: полный курс, 2-е издание, исправленное.: Пер. с англ. – М.: ООО «И.Д. Вильямс», 2006. – 1104 с.: ил. – Парал. тит. англ.
3. Осовский С. Нейронные сети для обработки информации / Пер. с польского И.Д. Рудинского. – М.: Финансы и статистика, 2002. – 344 с.: ил.
4. Тарков М.С. Нейрокомпьютерные системы: Учебное пособие / М.С. Тарков. – М.: Интернет-Университет Информационных Технологий; БИНОМ. Лаборатория знаний, 2006. – 142 с.
5. Круглов В.В., Борисов В.В. Искусственные нейронные сети. Теория и практика. – 2-е изд., стереотип. – М.: Горячая линия-Телеком, 2002.

Дополнительная:

1. Гергель В.П. Теория и практика параллельных вычислений: учебное пособие / В.П. Гергель. – М.: Интернет-Университет Информационных Технологий; БИНОМ. Лаборатория знаний, 2007. – 423 с
2. Барский А.Б. Логические нейронные сети: учебное пособие / А.Б. Барский – М.: Интернет-Университет Информационных Технологий; БИНОМ. Лаборатория знаний, 2007. – 352 с
3. Цыганков В.П. Нейрокомпьютер и его применение – М.: «Сол Системы», 1993.
4. Журнал «Нейрокомпьютеры: разработка, применение» Изд. «Радиотехника»