

2.3 Персептрон

2.3.1 Структура и основные особенности персептрона

Простейшей искусственной нейронной сетью является персептрон (Single layer perceptron - SLP). Название «персептрон» (англ. – **perceptron**, нем. – **perzeptron**) происходит от латинского слова **perceptio**, которое означает восприятие.

Искусственную нейронную сеть в виде персептрона предложил в 1957 году американский нейрофизиолог **Фрэнк Розенблатт**. И это была первая попытка создания компьютерной модели мозга. В 1960 году Розенблатт реализовал персептрон в виде первого нейрокомпьютера, который получил название Марк-1. Эксперименты показали, что этот нейрокомпьютер обладал элементами искусственного интеллекта, поскольку мог распознавать некоторые буквы английского алфавита.

О терминологии. В отечественной литературе используются два термина: «персептрон» и «перцептрон». Вариант «перцептрон» появился раньше, в переводе книги Розенблатта (1965). Он также присутствует в справочнике «Толковый словарь по искусственному интеллекту», выпущенном в 1992 году. Но термин «персептрон» в настоящее время встречается чаще.

Персептрон Розенблатта состоит из трёх типов элементов. Это элементы **S**-типа, **A**-типа и **R**-типа.

Элементы S-типа, стоящие на входе, представляют собой **сенсоры**. Входов, а следовательно, и используемых сенсоров, может быть много.

Поступающие от сенсоров сигналы передаются **ассоциативным элементам** (элементам A-типа). Они называются ассоциативными, поскольку ассоциированы с определённым набором

сенсоров, отвечающих за восприятие определённых признаков объектов внешней среды. Связи элементов персептрона обладают коэффициентами передачи (весовыми коэффициентами). Коэффициенты передачи в каналах связи выходов элементов S со входами элементов A могут иметь значения -1 , $+1$ или 0 . Если сумма сигналов от сенсоров превышает заданный порог ($+1$), то A-элемент активизируется и создаёт на своём выходе сигнал, равный $+1$. В противном случае этот сигнал равен нулю.

Сигналы, сформированные элементами A-типа, передаются на *реагирующие элементы* (элементы R-типа). Значения коэффициентов передачи сигналов с выходов элементов A к входам элемента R могут быть любыми.

В общем случае персептрон Розенблатта имеет несколько выходов. Выходные сигналы R-элементов могут иметь значения $+1$ или -1 . Значения коэффициентов связи между элементами персептрона *определяются в процессе обучения*.

Благодаря своим структурным особенностям персептроны позволяют сформировать и запомнить ассоциации между входными сигналами (сигналами-стимулами) и выходными сигналами (сигналами-реакциями) искусственной нейронной сети. Можно заметить, что сигналы передаются только от входов к выходам сети. *Обратные связи в такой сети отсутствуют*. Поэтому персептрон относится к классу искусственных нейронных сетей с *прямым распространением сигналов*.

Более простой структурой обладает так называемый *элементарный персептрон* Розенблатта. Единственное его отличие от классического персептрона Розенблатта состоит в том, что он имеет всего один R-выход. Пример логической схемы элементарного персептрона, имеющего шесть элементов S-типа и три элемента A-типа, показан на рис. 2.9.

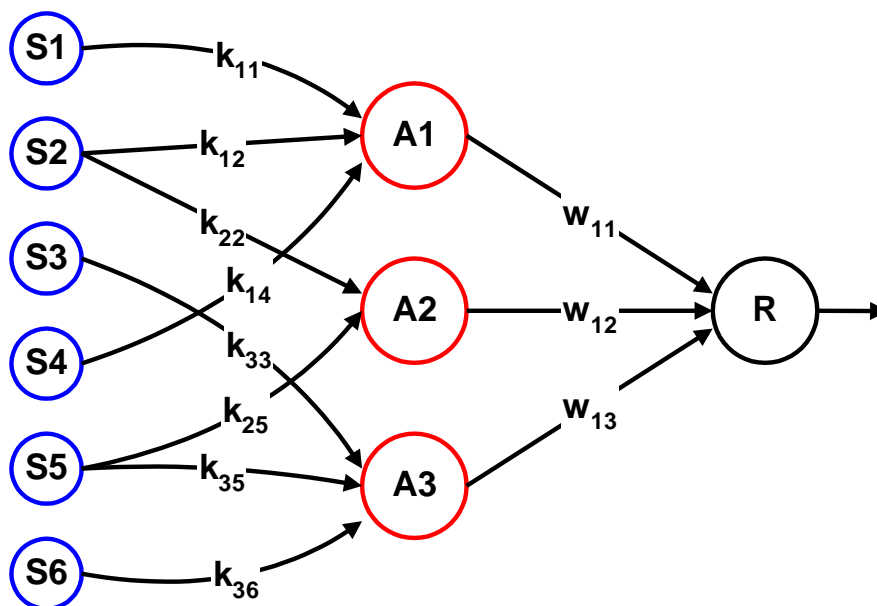


Рисунок 2.9 Логическая схема элементарного персептрона

Коэффициенты k_{ij} , связывающие выходы элементов S-типа со входами элементов A-типа, могут иметь значения 0, -1 или +1. Здесь i – номер элемента A-типа от 1 до 3, а j – номер элемента S-типа от 1 до 6. Связи элементов A-типа с выходным элементом R-типа характеризуются коэффициентами w_{1i} , которые могут иметь любые значения.

Наиболее простой формой ИНС и частным случаем классического персептрона Розенблатта является *простейший персептрон*. Он является однослойным, состоит из единственного искусственного нейрона, реализуемого в виде одного процессора, и в нём отсутствуют элементы A-типа. Сенсоры непосредственно связаны со входами классического нейрона Мак-Каллока и Питтса (рис. 2.10).

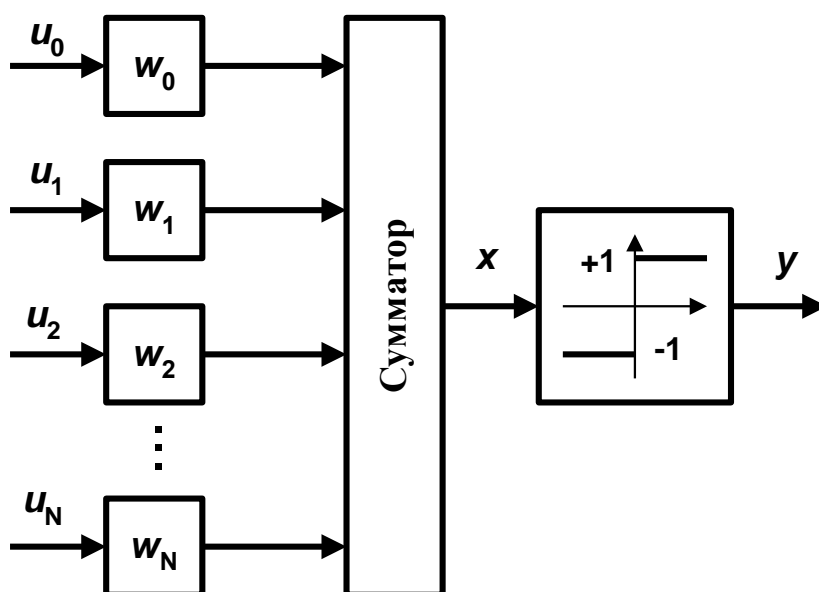


Рисунок 2.10 Схема простейшего персептрона

В общем случае простейший персептрон имеет N входов, на которые поступают сигналы u_1, u_2, \dots, u_N . Они умножаются на весовые коэффициенты w_1, w_2, \dots, w_N , суммируются и определяют состояние персептрона

$$x = \sum_{i=0}^N w_i u_i.$$

Если используется биполярная функция активации $f(x) = \text{sign}(x)$, то выходной сигнал персептрона y будет иметь только два значения: 1 или -1.

Персептрон решает задачу классификации входных векторов. Он может отнести входной вектор $U = [u_1, u_2, \dots, u_N]^T$ к одному из двух классов, обозначенных L_1 и L_2 . Если выходной сигнал персептрона принимает значение 1, то персептрон относит вектор U к классу L_1 , а если выходной сигнал равен -1, то к

классу L_2 . В результате этого персептрон разделяет N -мерное пространство входных векторов U на два пространства. Границей между ними является $(N-1)$ -мерная гиперплоскость, задаваемая уравнением

$$\sum_{i=0}^N w_i u_i = 0.$$

Такая гиперплоскость называется *решающей границей*.

Рассмотрим наиболее простой частный случай. Пусть персептрон имеет только два входа. Тогда решающая граница представляет собой прямую линию, которая определяется уравнением

$$w_1 u_1 + w_2 u_2 + w_0 u_0 = 0.$$

или

$$u_2 = \frac{x_0}{w_2} - \frac{w_1}{w_2} u_1.$$

Любой совокупности входных сигналов, представленной в виде вектора $U = [u_1, u_2]^T$, можно поставить в соответствие точку на плоскости (рис. 2.11).

Координатами этой точки будут значения переменных u_1 и u_2 . Если эта точка лежит выше прямой, являющейся решающей границей, то совокупность сигналов принадлежит классу L_1 . А если точка лежит ниже этой прямой, то совокупность сигналов принадлежит классу L_2 .

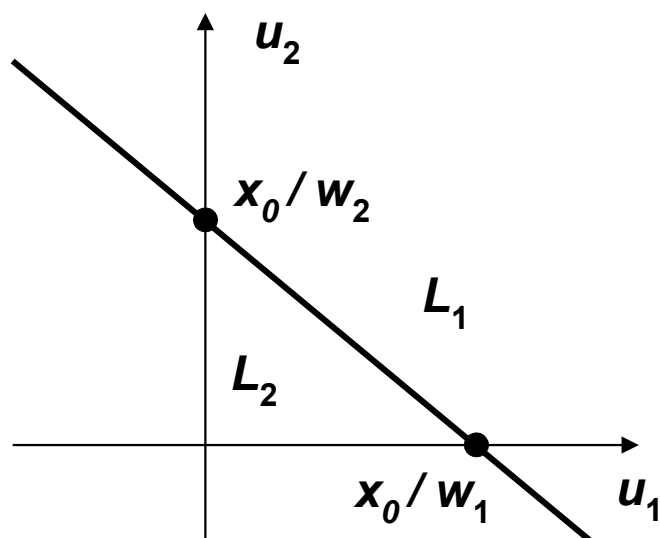


Рисунок 2.11 Иллюстрация решения задачи классификации

Точки, лежащие на решающей границе, нельзя строго отнести к тому или иному классу.

2.4 Обучение персептрона

Персептрон может правильно решать задачу классификации входных векторов только в том случае, если он имеет правильно подобранные значения коэффициентов w_1, w_2, \dots, w_N . Но первоначально эти коэффициенты имеют случайные и потому неоптимальные значения. Для правильной работы персептрона его надо обучить, подобрав значения компонентов вектора коэффициентов

$$W = [w_1, w_2, \dots, w_N]^T.$$

Существуют различные методы обучения ИНС. В частности, используется так называемый способ *обучения с учителем*.

На вход персептрона последовательно подаются обучающие наборы сигналов (шаблоны), для которых известен правильный результат классификации, и производится автоматическая подборка значений весовых коэффициентов. Обучающие последовательности можно разделить на две последовательности. Одна из них соответствует классу L_1 , а другая – классу L_2 .

На некотором n -ом этапе обучения вектор коэффициентов имеет текущие значения

$$W(n) = [w_1(n), w_2(n), \dots, w_N(n)]^T,$$

на вход персептрона подаётся обучающий вектор

$$U(n) = [u_1(n), u_2(n), \dots, u_N(n)]^T,$$

а текущее состояние персептрона описывается уравнением

$$x(n) = W^T(n)U(n).$$

Введём в рассмотрение так называемый эталонный сигнал

$$d(n) = \begin{cases} +1 & \text{при } U(n) \in L_1, \\ -1 & \text{при } U(n) \in L_2. \end{cases}$$

Тогда рассогласование между эталонным сигналом и выходом персептрона определяется на основании уравнения

$$\varepsilon(n) = d(n) - y(n),$$

где

$$y(n) = \text{sign}\{x(n)\}.$$

Это рассогласование можно использовать для корректировки значений весовых коэффициентов.

Обучение персептрона осуществляется в результате *рекуррентной процедуры* коррекции значений вектора $W = [w_1, w_2, \dots, w_N]^T$ в соответствии с алгоритмом

$$W(n+1) := W(n) + \eta \varepsilon(n) U(n),$$

где η – коэффициент, значение которого лежит в диапазоне от 0 до 1 и влияет на устойчивость и быстроту сходимости процесса настройки коэффициентов. Схема персептрона с блоком настройки параметров показана на рис. 2.14.

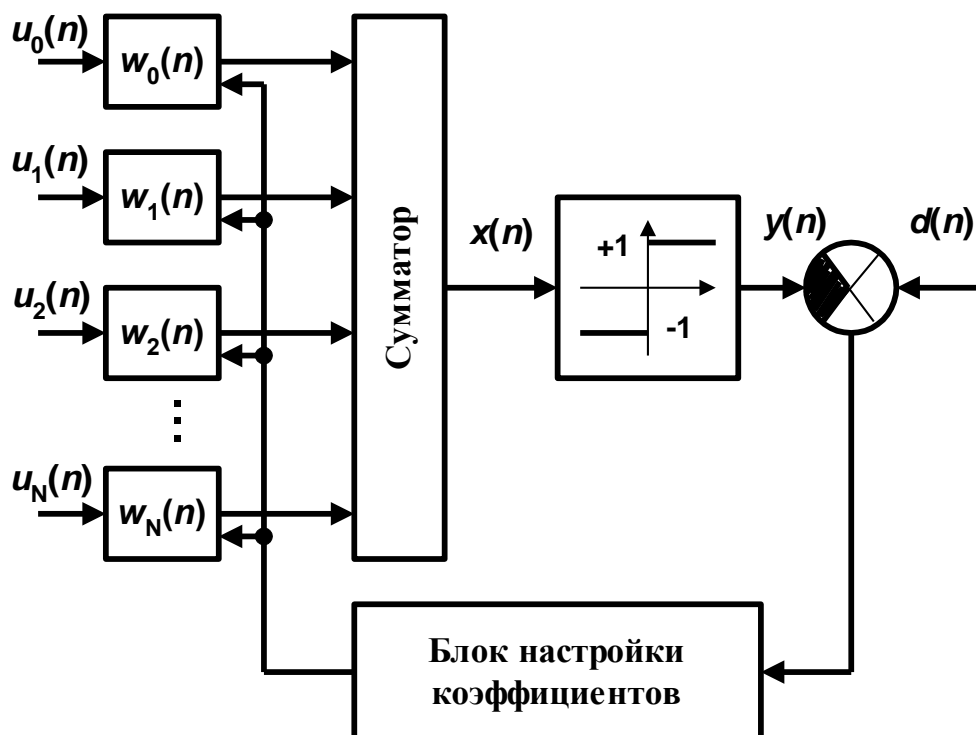


Рисунок 2.14 Схема обучения простейшего персептрона

Процесс обучения считается завершённым, когда изменения значений коэффициентов на последующих шагах становятся незначительными. С этого момента персептрон обучен и способен правильно классифицировать входные векторы, причём не только те, которые использовались при обучении, но и многие из тех, которые отсутствовали в обучающей последовательности. Таким образом, ИНС приобретает свойство обобщения.

Пример процесса обучения простейшего персептрона с двумя входами показан в таблице 2.3. Настраиваются два коэффициента w_1 и w_2 . Решающая граница представляет собой прямую линию, которая описывается уравнением

$$w_1 u_1 + w_2 u_2 - x_0 = 0.$$

Обучающая последовательность входных сигналов u_1 и u_2 , а также соответствующая им последовательность эталонных сигналов d формируются в соответствии с истинными значениями параметров $w_1 = 0,5$; $w_2 = 0,5$; $x_0 = 2,5$.

В таблице 2.3 приведены значения весовых коэффициентов w_1 и w_2 , состояние персептрона x и ошибка, которая обозначена «delta». Начальные значения коэффициентов равны 0,2.

Пример показывает, что по завершении 26 шагов обучения удалось найти значения параметров w_1 и w_2 , несущественно отличающиеся от их истинных значений. Относительные погрешности определения коэффициентов w_1 и w_2 составляют 2,07% и -5,98% соответственно.

Таблица 2.3 Процесс обучения простейшего персептрона

	1	2	3	4	5	6	7	
1	k =	0,01		x0 =	2,5			
2	w1	w2	u1	u2	d	x	delta	
3	0,2	0,2	1	1	-1	-2,1	1,1	
4	0,211	0,211	3	3	1	-1,234	2,234	
5	0,27802	0,27802	2	2	-1	-1,38792	0,38792	
6	0,285778	0,285778	4	4	1	-0,21377	1,213773	
7	0,334329	0,334329	0,5	0,5	-1	-2,16567	1,165671	
8	0,340158	0,340158	5	5	1	0,901577	0,098423	
9	0,345079	0,345079	1	3	-1	-1,11968	0,119685	
10	0,346276	0,348669	1	5	1	-0,41038	1,410377	
11	0,360379	0,419188	3	1	-1	-0,99967	-0,00033	
12	0,36037	0,419185	2	4	1	-0,10252	1,102521	
13	0,38242	0,463286	0,25	3	-1	-1,01454	0,014538	
14	0,382456	0,463722	1	6	1	0,664788	0,335212	
15	0,385809	0,483835	0	1,5	-1	-1,77425	0,774248	
16	0,385809	0,495448	0	5,5	1	0,224966	0,775034	
17	0,385809	0,538075	3	0,5	-1	-1,07354	0,073537	
18	0,388015	0,538443	4	1,5	1	-0,14028	1,140277	
19	0,433626	0,555547	2,5	1	-1	-0,86039	-0,13961	
20	0,430135	0,554151	2,5	3,5	1	0,514867	0,485133	
21	0,442264	0,571131	1	3,8	-1	0,11256	-1,11256	
22	0,431138	0,528853	2	3,3	1	0,107492	0,892508	
23	0,448988	0,558306	3	1,7	-1	-0,20391	-0,79609	
24	0,425106	0,544773	4	1,1	1	-0,20033	1,200327	
25	0,473119	0,557976	0,5	4,3	-1	0,135857	-1,13586	
26	0,46744	0,509134	0,75	4,5	1	0,141684	0,858316	
27	0,473877	0,547759	0,9	4	-1	0,117524	-1,11752	
28	0,463819	0,503058	2,5	2,6	1	-0,0325	1,032502	
29	0,489632	0,529903						

Положения истинной решающей границы и решающей границы, полученной в результате обучения персептрона, показаны на рис. 2.15. Из него следует, что зафиксированные погрешности настройки параметров в процессе обучения невелики, и обучен-

ная сеть может использоваться для классификации входных воздействий.

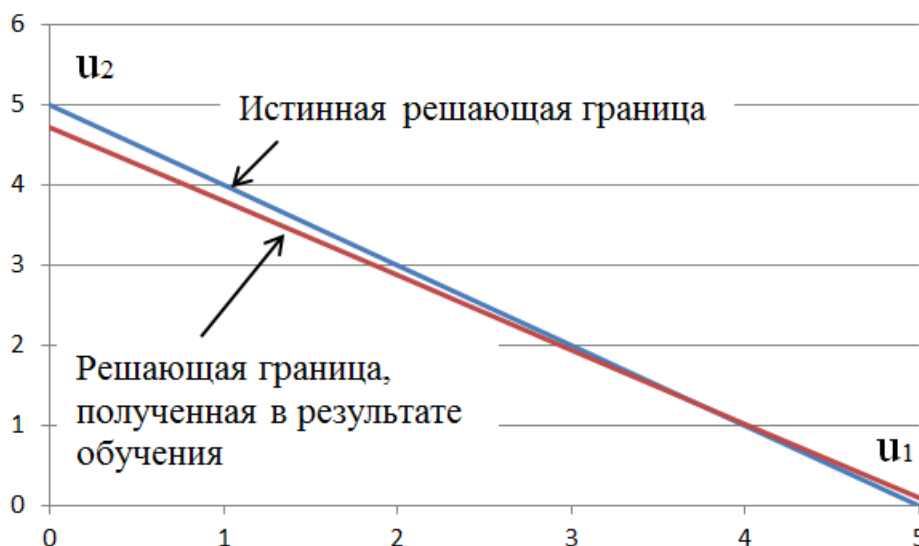


Рисунок 2.15 Истинная и найденная в результате обучения решающие границы

Поступающие на вход обученного персептрона входные сигналы в основном будут приводить к правильным результатам классификации. Но входные сигналы, значения которых принадлежат областям между этими двумя решающими границами, приведут к ошибочным решениям.

Ошибка настройки параметров ИНС, а следовательно, и ошибка обобщения есть всегда. Они обусловлены, прежде всего, ограниченным по времени процессом обучения, недостаточно корректно выбранным значением шага подстройки η и ограниченным объёмом информации в обучающей выборке, предъявленной сети в процессе её обучения.