Лекция 5

План лекции:

1. Модель системы персептронного типа для классификации объектов.
2. Метод коррекции весовых коэффициентов в процессе обучения системы.
3. Алгоритм классификации объектов на N классов.

5.1 Модель системы автоматической классификации объектов

Рассмотрим задачу, состоящую в том, чтобы выработать процедуру, по возможности автоматическую, позволяющую отнести предъявляемый незнакомый объект к множеству известных образов.

Каждый класс будем обозначать C_i, число классов k, а множество классов C состоит из их совокупности: $C=\{C_1, C_2, \ldots C_k\}$. Каждый объект представляется точкой в пространстве признаков X. Число существенных параметров равно числу составляющих каждого вектора $\vec{x}$. Пусть это число равно n. Тогда каждый объект представляется точкой $\vec{x} = \langle x_1, x_2, \ldots, x_n \rangle$.

Классификация считается выполненной, если найдены такие решающие правила, которые могут быть представлены в общем виде:

$g_i(\vec{x}) = \sum_{i=1}^{n} \omega_{il} \Phi_l(\vec{x}) \quad i = 1,2 \ldots k.$ Где ω_{il} – весовые коэффициенты составляющих вектора, а Φ_l – функции от $\vec{x}$. Их выбор определяется характером решаемой задачи. Каждому классу C_j соответствует свой вектор весовых коэффициентов $\omega_l : \vec{\omega} = \langle \omega_{1j}, \omega_{2j} \ldots \omega_{nj} \rangle$. Векторы ω_l при фиксированных Φ_l определяют так, чтобы получить классификацию, наилучшую из всех возможных. Весовые коэффициенты находятся в процессе адаптации, или «обучения», которое может проходить в одном из двух вариантов.

- *С учителем.* В этом случае параметры оцениваются при помощи алгоритмов с использованием измерений, выполненных на объектах, принадлежность которых к определенному классу априори известна «учителю». Параметры уточняются последовательно до получения правильных результатов классификации.
- *Без учителя.* Этот вариант соответствует ситуации, когда классы априори не известны.

В первом случае процедура в значительной степени зависит от выбора начальных образов, используемых для обучения. Это – так называемая обучающая выборка. В частности, входящие в нее элементы должны быть легко и четко разделимы.

В середине XX века ученый Розенблат предложил модель обучаемой машины, известной под названием «Персептрон». Машины такого типа не получили широкого практического применения, однако персептроны – это первые технические устройства, позволяющие поставить и решить задачу автоматической классификации. На рис. 1 приведена схема такой машины. Где 1 – стимулы, 2 – датчика, 3 – суммирующие ячейки, 4 – весовые множители, 5 – блок решения, 6 – управление изменением весов, 7 – вычисление новых весовых множителей.

Датчики, составляющие слой чувствительных элементов, соединены случайным образом с ассоциативным слоем, имеющим фиксированную структуру. Выходными сигналами этого слоя являются признаки $x_1, x_2, ..., x_n$, которые взвешиваются с весами $\omega_1, \omega_2 ... \omega_n$, а затем объединяются, в результате чего получается отклик R, поступающий на блок решения 5. В частном случае, когда классификация выполняется на два класса, это пороговое устройство, дающее на выходе номера классов 1 или 2.

Поскольку система является обучаемой, это решение используется для корректировки весов ($\omega_1, \omega_2 ... \omega$) после сравнения его с заведомо известным, При этом решающему устройству дается «поощрение», если класс определен верно, и «наказание» – если неверно. Это выражается в том, что в первом случае решение подкрепляется, во втором – ослабляется.

Структура модели персептронного типа может быть положена в основу системы распознавания. Ее можно улучшить за счет изменения связей между датчиками и суммирующими ячейками, путем возбуждения одних связей и затормаживания других. Система, «обученная» распознаванию некоторых образов, распознает только их. При смене класса предъявляемых образов обучение должно начинаться заново. Тем не мене модель персептрона достаточно удобна, так как ее можно приспособить, например, для моделирования нелинейных систем. Она допускает также определенное обобщение, в результате которого ее можно применять для распознавания абстрактных понятий.

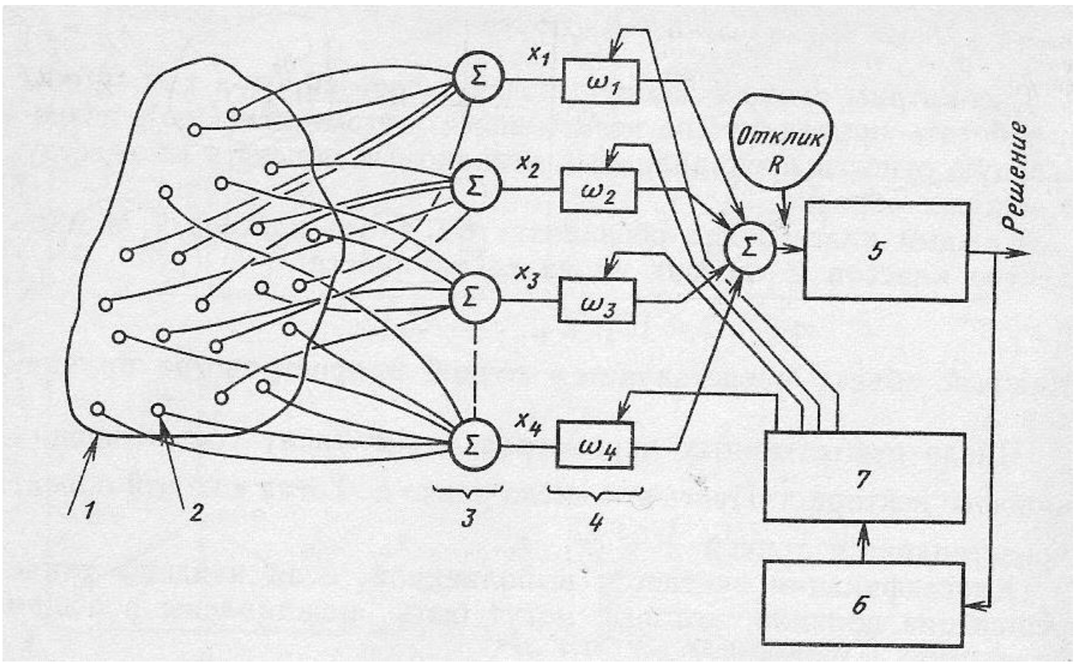

Рис. 1 – Схема персептрона

Рассмотрим это на примере классификации треугольников и квадратов. Вначале системе предъявляется квадрат, и она автоматически вырабатывает набор весов ω_i. Затем предъявляют треугольник и вырабатывают новые веса ω_i. Обучение продолжают с различными экземплярами треугольников и квадратов. Когда автоматическая классификация предъявляемых объектов становится безошибочной, машина считается способной различать понятия «треугольник» и «квадрат», поскольку конкретные реализации имеют различные размеры, ориентацию и другие индивидуальные черты. Задача, возникающая на данном этапе, состоит из двух частей: во-первых, необходимо найти закон изменения параметров систем и, во-вторых, обеспечить сходимость разрабатываемого алгоритма.

5.2 Метод коррекции весовых коэффициентов

Будем рассматривать фундаментальную задачу классификации на примере разделения на два класса, к которой сводятся многие другие задачи. Расширение на большее число классов можно выполнить, например, путем параллельного соединения таких систем.

Пусть имеются два класса – C_1 и C_2, а также векторы $\vec{v}$ и $\vec{y}$ соответственно. Критерий классификации можно представить в виде:

если $\vec{v}\vec{y} > 0,$ то класс C_1;

если $\vec{v}\vec{y} \le 0,$ то класс C_2.

Принцип обучения состоит в следующем: если результат классификации неверный, то применяют «наказание», состоящее в том, что изменяют веса, увеличивая или уменьшая их путем умножения на соответствующий коэффициент; если же результат классификации верный, то применяют «поощрения», или «отсутствия наказания», оставляя веса неизменными. В табл. 1 показана последовательность корректирующих действий.

Таблица 1 Коррекция весовых коэффициентов в процессе обучения

Объект известного класса	Восприятие		Классификация	Изменение весов (c — константа)
	Отклик	Решение		
$\vec{x}(i) \in C_1$	$\leqslant 0$	C_2	Неверно	$\vec{w}(i+1) = \vec{w}(i) + c\vec{x}(i)$
$\vec{x}(i) \in C_1$	>0	C_1	Верно	$\vec{w}(i+1) = \vec{w}(i)$
$\vec{x}(i) \in C_2$	>0	C_1	Неверно	$\vec{w}(i+1) = \vec{w}(i) - c\vec{x}(i)$
$\vec{x}(i) \in C_2$	$\leqslant 0$	C_2	Верно	$\vec{w}(i+1) = \vec{w}(i)$

Исходное состояние системы обозначено индексом i. Задача состоит в отыскании для $(i+1)$-го этапа своего вектора весовых коэффициентов $\omega(i+1)$, выразив его через веса $\omega(i)$. Задача считается решенной, когда все предъявлемые объекты классифицированы правильно. Здесь, по-прежнему, рассматриваются два класса, но следует иметь в виду, что один и тот же образ, принадлежащий одному и тому же классу, может иметь различную ориентацию, поэтому требуется выполнять ряд последовательных испытаний. Эта процедура сводится к конечному числу этапов, если классы линейно сепарабельны.

5.3 Алгоритм классификации объектов на N классов

В разделе 2.3 лекции №2 были рассмотрены три случая разделения объектов на несколько классов. В данной лекции рассмотрим алгоритм, который можно применить для определения решающих функций в случае 3, когда допускается существование M решающих функций, характеризующихся тем свойством, что при $x \in \omega_i$, где x – объект, ω_i – класс $d_i(x) > d_j(x)$ для всех $i \neq j$.

Рассмотрим M классов $\omega_1, \omega_2, \ldots \omega_M$. Пусть на $к$-м шаге процедуры обучения системе предъявляется образ $x(к)$, принадлежащий классу ω_i. Вычисляются значения M решающих функций

$d_j[x(k)] = w_j(k)x(k),\ j = 1,2...,M.$ Затем если выполняются условия $d_i[x(k)] > d_j[x(k)],\ j = 1,2,...,M;\ j \neq i,$ то векторы весов не изменяются, т.е. $w_j(k+1) = w_j(k),\ j = 1,2,...M.$

С другой стороны, допустим, что для некоторого l $d_i[x(k)] \leq d_l[x(k)].$ В этом случае выполняются следующие коррекции весов:

$$
\begin{aligned}
&w_i(k+1) = w_i(k) + cx(k),\\
&w_l(k+1) = w_l(k) - cx(k), \qquad\qquad (5.1)\\
&w_j(k+1) = w_j(k),\ j = 1,2,\ ...M;\ j \neq i,\ j \neq l,
\end{aligned}
$$

где с – положительная константа. Если при рассмотрении случая 3 классы разделимы, то доказано, что этот алгоритм сходится за конечное число итераций при произвольных начальных векторах. Рассмотрим это на примере.

Даны классы, причем каждый из них содержит один образ: ω_1: {(0, 0)}, ω_2: {(1,1}, ω_3: {(-1, 1)}. Дополним заданные образы: (0, 0, 1), (1, 1, 1), (-1, 1, 1). Выберем в качестве начальных векторов весов $w_1(1) = w_2(1) = w_3(1) = (0,0,0),$ положим c=1 и, предъявляя образы в указанном порядке, получим следующее:

$$
\begin{aligned}
&d_1[x(1)] = w_1(1)x(1) = 0,\\
&d_2[x(1)] = w_2(1)x(1) = 0,\\
&d_3[x(1)] = w_3(1)x(1) = 0.
\end{aligned}
$$

Поскольку $x(1) \in \omega_1$ и $d_2[x(1)] = d_3[x(1)] = d_1[x(1)],$ первый весовой вектор увеличивается, а два других уменьшаются в соответствии с соотношениями (5.1), т.е.

$$
w_1(2) = w_1(1) + x(1) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},
$$

$$
w_2(2) = w_2(1) - x(1) = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix},
$$

$$
w_3(2) = w_3(1) - x(1) = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}.
$$

Следующий предъявляемый образ x(2)=(1, 1, 1) принадлежит классу ω_2. Для него получаем

$w_1(2)x(2)=1,\ w_2(2)x(2)=-1, w_3(2)x(2)=-1.$ Поскольку все произведения больше либо равны $w_2(2)x(2)$, вводятся коррекции

$$w_1(3)=w_1(2)-x(2)=\begin{pmatrix}-1\\-1\\0\end{pmatrix},$$

$$w_2(3)=w_2(2)+x(2)=\begin{pmatrix}1\\1\\0\end{pmatrix},$$

$$w_3(3)=w_3(2)-x(2)=\begin{pmatrix}-1\\-1\\-2\end{pmatrix}.$$

Следующий предъявленный образ x(3)=(-1, 1, 1) принадлежит классу ω_3. Для него получаем $w_1(3)x(3)=0,\ w_2(3)x(3)=0, w_3(3)x(3)=-2.$ Все эти произведения опять требуют корректировки

$$w_1(4)=w_1(3)-x(3)=\begin{pmatrix}0\\-2\\-1\end{pmatrix},$$

$$w_2(4)=w_2(3)-x(3)=\begin{pmatrix}2\\0\\-1\end{pmatrix},$$

$$w_3(4)=w_3(3)+x(3)=\begin{pmatrix}-2\\0\\-1\end{pmatrix}.$$

Поскольку в данном цикле итерации присутствовали ошибки, следует провести новый цикл. Положив $x(4)=x(1)$, $x(5)=x(2)$, $x(6)=x(3)$, получим $w_1(4)x(4)=-1,\ w_2(4)x(4)=-1, w_3(4)x(4)=-1.$ Так как образ $x(4)$ принадлежит классу ω_1, то все произведения «неверны». Поэтому

$$w_1(5)=w_1(4)+x(4)=\begin{pmatrix}0\\-2\\0\end{pmatrix},$$

$$w_2(5)=w_2(4)-x(4)=\begin{pmatrix}2\\0\\-2\end{pmatrix},$$

$$w_3(5) = w_3(4) - x(4) = \begin{pmatrix} -2 \\ 0 \\ -2 \end{pmatrix}.$$

Следующий предъявленный образ $x(5)$=(1, 1, 1) принадлежит классу ω_2. Соответствующие скалярные произведения равны $w_1(5)x(5) = -2,\ w_2(5)x(5) = 0, w_3(5)x(5) = -4.$ Образ $x(5)$ классифицирован правильно. Поэтому

$$w_1(6) = w_1(5) = \begin{pmatrix} 0 \\ -2 \\ 0 \end{pmatrix},$$

$$w_2(6) = w_2(5) = \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix},$$

$$w_3(6) = w_3(5) = \begin{pmatrix} -2 \\ 0 \\ -2 \end{pmatrix}.$$

Следующий образ $x(6)$=(-1, 1, 1) принадлежит классу ω_3, для него получаем $w_1(6)x(6) = -2,\ w_2(6)x(6) = -4, w_3(6)x(6) = -0.$ Этот образ также классифицирован правильно, так что коррекции не нужны, т.е.

$$w_1(7) = w_1(6) = \begin{pmatrix} 0 \\ -2 \\ 0 \end{pmatrix},$$

$$w_2(7) = w_2(6) = \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix},$$

$$w_3(7) = w_3(6) = \begin{pmatrix} -2 \\ 0 \\ -2 \end{pmatrix}.$$

Если продолжить процедуру обучения, рассматривая образы $x(7)$, $x(8)$, $x(9)$, можно убедиться, что в следующем полном цикле никакие коррекции не производятся. Поэтому искомые решающие функции имеют следующий вид:

$$d_1(x) = 0 \cdot x_1 - 2x_2 + 0 = -2x_2,$$

$$d_2(x) = 2x_1 - 0 \cdot x_2 - 2 = 2x_1 - 2,$$

$$d_3(x) = -2x_1 + 0 \cdot x_2 - 2 = -2x_1 - 2.$$