

Лабораторная работа №9

РАСПОЗНАВАНИЕ ОБРАЗОВ НА ОСНОВЕ ИСКУССТВЕННОЙ НЕЙРОННОЙ СЕТИ

Цель работы: научиться классифицировать объекты с помощью искусственной нейронной сети (ИНС).

Порядок выполнения работы

1. Ознакомление с теоретической частью работы.
2. Реализация алгоритма автоматической классификации объектов с помощью ИНС в виде многослойного персептрона.
3. Оформление отчета по лабораторной работе.

Исходные данные: обучающая выборка, количество классов.

Выходные данные: классы и отнесенные к ним объекты.

Одной из наиболее популярных моделей ИНС с контролируемым обучением считается модель в виде многослойного персептрона (рис. 1).

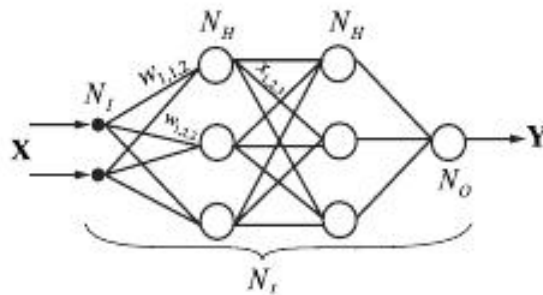


Рис. 1. ИНС в виде многослойного персептрона

Формальные нейроны могут объединяться в сети различным образом. Сеть состоит из произвольного количества слоев нейронов, в простейшем случае – однослойная сеть. Первый слой называется сенсорным или входным, внутренние слои называются скрытыми или ассоциативными, последний – выходным или результативным. Количество нейронов в слоях может быть произвольным. Обычно во всех скрытых слоях одинаковое количество нейронов. В каждом слое выполняется нелинейное преобразование линейной комбинации сигналов предыдущего слоя. Следовательно, в тех сетях, где требуется последовательное соединение слоев нейронов друг за другом, необходима нелинейная функция активации. В противном случае многослойность оказывается ненужной, так как ее можно заменить эквивалентной однослойной сетью с соответствующими весовыми коэффициентами. Многослойная сеть может формировать на выходе произвольную многомерную функцию при соответствующем выборе количества слоев, диапазона изменения сигналов и параметров нейронов.

Сеть состоит из произвольного количества слоев нейронов. Нейроны каждого слоя соединяются с нейронами предыдущего и последующего слоев по принципу «каждый с каждым». Первый слой (слева) называется сенсорным или входным, внутренние слои называются скрытыми или ассоциативными, последний (самый правый, на рис. 1 состоит из одного нейрона) – выходным или результативным. Количество нейронов в слоях может быть произвольным. Обычно во всех скрытых слоях одинаковое количество нейронов.

Введем обозначение количества слоев и нейронов в слое. Входной слой: N_I нейронов; N_H нейронов в каждом скрытом слое; N_O выходных нейронов; x – вектор входных сигналов сети; y – вектор выходных сигналов.

Входной слой не выполняет никаких вычислений, а лишь распределяет входные сигналы, поэтому иногда его не учитывают, считая количество слоев в сети. Обозначим через N_L полное количество слоев в сети, считая и входной. Работа многослойного персептрона (МСП) описывается формулами:

$$\begin{aligned} NET_{jl} &= \sum_i w_{ijl} x_{ijl}, \\ OUT_{jl} &= F(NE_{jl} - \theta_{jl}), \\ x_{ij(l+1)} &= OUT_{il}, \end{aligned}$$

где индексом i обозначен номер входа; j – номер нейрона в слое; l – номер слоя.

Кроме того, x_{ijk} – i -й входной сигнал j -го нейрона в слое l ; w_{ijk} – весовой коэффициент i -го входа j -го нейрона в слое l ; NET_{jl} – сигнал NET j -го нейрона в слое l ; OUT_{jl} – выходной сигнал нейрона; θ_{jl} – пороговый уровень j -го нейрона в слое l .

Введем еще некоторые обозначения: w_{jl} – вектор-столбец весов для всех входов нейрона j в слое l ; W_l – матрица весов всех нейронов в слое l , в столбцах матрицы расположены векторы w_{jl} ; x_{jl} – входной вектор-столбец слоя l .

В каждом слое рассчитывается нелинейное преобразование от линейной комбинации сигналов предыдущего слоя. Отсюда видно, что линейная функция активации может применяться только для тех моделей сетей, где не требуется последовательное соединение слоев нейронов друг за другом. Для многослойных сетей функция активации должна быть нелинейной, иначе можно построить эквивалентную однослойную сеть и многослойность оказывается ненужной. Если применена линейная функция активации, то каждый слой будет давать на выходе линейную комбинацию входов. Следующий слой даст линейную комбинацию выходов предыдущего, а это эквивалентно одной линейной комбинации с другими коэффициентами и может быть реализовано в виде одного слоя нейронов. Многослойная сеть может формировать на выходе произвольную многомерную функцию при соответствующем выборе количества слоев, диапазона изменения сигналов и

параметров нейронов. Как и ряды, многослойные сети оказываются универсальным инструментом аппроксимации функций.

В нейронной сети за счет поочередного расчета линейных комбинаций и нелинейных преобразований достигается аппроксимация произвольной многомерной функции при соответствующем выборе параметров сети. В многослойном персептроне нет обратных связей. Такие модели называются сетями прямого распространения.

Для построения персептрона необходимо выбрать его параметры. Чаще всего выбор пороговых значений и весовых коэффициентов происходит в процессе обучения, т. е. веса и коэффициенты пошагово изменяются в ходе обучения сети. Рассмотрим по шагам общий алгоритм решения задачи с помощью персептрона.

1. Определить смысл, вкладываемый в компоненты входного вектора x . Вектор x должен содержать формализованное условие задачи, т. е. всю информацию, необходимую для получения требуемого результата.

2. Выбрать выходной вектор y таким образом, чтобы его компоненты содержали ответ поставленной задачи.

3. Выбрать вид нелинейности в нейронах (функцию активации). При этом желательно учесть специфику задачи, так как удачный выбор сократит время обучения.

4. Выбрать число слоев и нейронов в слое.

5. Задать диапазон изменения входов, выходов, весов и пороговых уровней, учитывая множество значений выбранной функции активации.

6. Присвоить начальные значения весовым коэффициентам и пороговым уровням и дополнительным параметрам (например, крутизне функции активации, если она будет настраиваться при обучении). Начальные значения не должны быть большими, чтобы нейроны не оказались в насыщении (на горизонтальном участке функции активации), иначе обучение окажется фиктивным: пороговые значения и веса будут достигнуты, а сеть не обучится. Начальные значения не должны быть и слишком малыми, чтобы выходы большей части нейронов не были равны нулю, иначе обучение замедлится.

7. Провести обучение, т. е. подобрать параметры сети так, чтобы задача решалась наилучшим образом. По окончании обучения сеть готова решать задачи того типа, которым она обучена.

8. Подать на вход сети условия задачи в виде вектора x . Рассчитать выходной вектор y , который и предоставит искомое решение задачи.

Существенное значение на решение поставленной задачи оказывает выбор количества слоев в сети и нейронов в слоях. Нет строгого правила для выбора этих параметров сети. Однако, чем больше количество нейронов и слоев, тем шире возможности сети, тем медленнее она обучается и работает и тем более сложной может быть зависимость вход-выход. Количество нейронов и слоев связано:

- со сложностью задачи;
- с количеством данных для обучения;
- с требуемым количеством входов и выходов сети;

– с имеющимися ресурсами: памятью и быстродействием машины, на которой моделируется сеть.

Поскольку Персептрон обучается с учителем, должно быть задано множество пар векторов $\{x, d\}$, где x – условие задачи, d – известное решение для этого условия. Количество элементов в обучающем множестве должно быть достаточным для обучения сети, чтобы под управлением алгоритма сформировать набор параметров сети, дающий нужное отображение $X \rightarrow Y$. При этом количество элементов в обучающей выборке не регламентируется. Выберем один из векторов x^S (S указывает на то, что x принадлежит обучающей выборке) и подадим его на вход сети. На выходе получится некоторый вектор y^S . Тогда ошибкой сети можно считать $E = |d - y|$ для каждой пары (x, d) . Для оценки качества обучения выбирают суммарную квадратическую ошибку или среднюю относительную ошибку.

Приведем общую схему обучения персептрона:

1. Инициализировать веса и параметры функции активации в малые ненулевые значения;
2. Подать на вход один образ и рассчитать выход;
3. Посчитать ошибку E^S , сравнив d^S и y^S .
4. Изменить веса и параметры функции активации так, чтобы ошибка E^S уменьшилась.
5. Повторить шаги 2 – 4 до тех пор, пока ошибка не перестанет убывать или не станет достаточно малой.

Рассмотрим пример классификации объектов с помощью ИНС персептронного типа (рис. 2).

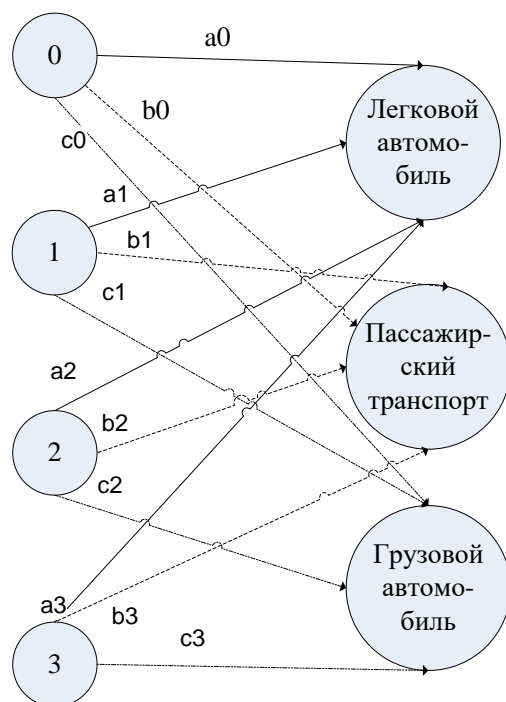


Рис. 2. ИНС персептронного типа

Сеть предназначена для разделения объектов на три класса: *легковой автомобиль, пассажирский транспорт, грузовой автомобиль*. Для обучения используется выборка, каждый объект которой имеет четыре признака.

В качестве функции активации выбрана кусочно-заданная функция «линейный порог», изображенная на рис. 3. Функция активации определяет уровень возбуждения нейрона в зависимости от суммарного уровня входного сигнала S .

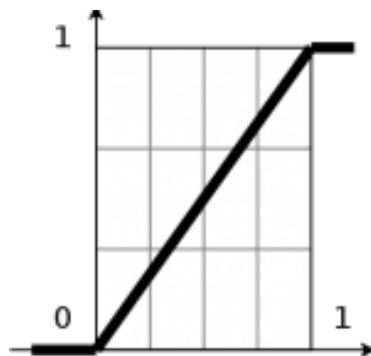


Рис. 3. Функция «линейный порог»

В зависимости от S функция принимает следующие значения

$$Y = \begin{cases} 0, & S < 0; \\ S, & 0 \leq S \leq 1; \\ 1, & S > 1. \end{cases}$$

Начальные значения коэффициентов сети приведены в табл. 1.

Таблица 1

Класс	Признак			
	0	1	2	3
Легковой автомобиль (класс 1, коэффициенты a_i)	– 0,3	0,1	0,4	– 0,5
Пассажирский транспорт (класс 2, коэффициенты b_i)	0,1	0,3	– 0,3	0,2
Грузовой автомобиль (класс 3, коэффициенты c_i)	0,4	– 0,2	0,1	0,2

Обучающая выборка представлена тремя объектами (табл. 2), каждый из которых принадлежит одному из трех классов.

Таблица 2

Объект	Признак			
	Масса	Двигатель	Пассажировместимость	Грузоподъемность
Ока (класс 1)	0,25	0,25	0,25	0
Газель пассажирская (класс 2)	0,75	0,75	0,75	0,25
Камаз (класс 3)	1	1	0,25	1

Рассмотрим процесс обучения ИНС с помощью обучающей выборки. Поскольку в основе ИНС лежит модель «Персептрон», для ИНС применяются те же условия и правила корректировки коэффициентов, что и для персептрона.

Результаты обучения для объекта «КаМаз».

$$\begin{aligned}\text{класс 1: } & -0,3 \cdot 1 + 0,1 \cdot 1 + 0,4 \cdot 0,25 + (-0,5) \cdot 1 = -0,6; \\ \text{класс 2: } & 0,1 \cdot 1 + 0,3 \cdot 1 + (-0,3) \cdot 0,25 + 0,2 \cdot 1 = 0,525; \\ \text{класс 3: } & 0,4 \cdot 1 + (-0,2) \cdot 1 + 0,1 \cdot 0,25 + 0,2 \cdot 1 = 0,65.\end{aligned}$$

Обучающий объект принадлежит к 3-му классу, на котором достигается максимальное значение, поэтому не требуется коррекция коэффициентов сети. Все коэффициенты остаются теми же, что и на предыдущем шаге.

Результаты обучения для объекта «Ока».

$$\begin{aligned}\text{класс 1: } & -0,3 \cdot 0,25 + 0,1 \cdot 0,25 + 0,4 \cdot 0,25 + 0 \cdot (-0,5) = 0,05; \\ \text{класс 2: } & 0,1 \cdot 0,25 + 0,3 \cdot 0,25 + (-0,3) \cdot 0,25 + 0,2 \cdot 0 = 0,025; \\ \text{класс 3: } & 0,4 \cdot 0,25 + (-0,2) \cdot 0,25 + 0,1 \cdot 0,25 + 0,2 \cdot 0 = 0,075;\end{aligned}$$

Поскольку обучающий объект принадлежит классу 1, а максимум достигается для класса 3, эти два класса требуют коррекции коэффициентов. Она выполняется методом обратного распространения ошибки. Коэффициенты класса 2 не требуют исправления. Новые коэффициенты ИНС приведены в табл. 4.

$$\begin{aligned}\text{класс 1: } & 1 - 0,05 = 0,95; \\ \text{класс 2: } & 0 - 0,025 = -0,025; \\ \text{класс 3: } & 0 - 0,075 = -0,075.\end{aligned}$$

Таблица 4

Класс	Признак			
	0	1	2	3
Легковой автомобиль (1 класс)	$-0,3 + 0,95 =$ 0,65	$0,1 + 0,95 =$ 1,05	$0,4 + 0,95 =$ 1,35	$-0,5 + 0,95 =$ 0,45
Пассажирский транспорт (2 класс)	0,1	0,3	-0,3	0,2
Грузовой автомобиль (3 класс)	$0,4 - 0,075 =$ 0,325	$-0,2 - 0,075 =$ -0,275	$0,1 - 0,075 =$ 0,025	$0,2 - 0,075 =$ 0,125

Результат обучения для объекта «Газель пассажирская»:

класс 1: $0,65 \cdot 0,75 + 1,05 \cdot 0,75 + 1,35 \cdot 0,75 + 0,45 \cdot 0,25 = 2,475$;
класс 2: $0,1 \cdot 0,75 + 0,3 \cdot 0,75 - 0,3 \cdot 0,75 + 0,2 \cdot 0,25 = 0,125$;
класс 3: $0,325 \cdot 0,75 - 0,275 \cdot 0,75 + 0,025 \cdot 0,75 + 0,125 \cdot 0,25 = 0,0875$.

Поскольку обучающий объект принадлежит классу 2, а максимум достигается для класса 1, эти два класса требуют коррекции коэффициентов. Она выполняется методом обратного распространения ошибки. Коэффициенты класса 3 не требуют исправления. Новые коэффициенты ИНС приведены в табл. 5.

класс 1: $0 - 2,475 = - 2,475$;
класс 2: $1 - 0,125 = 0,875$;
класс 3: $0 - 0,0875 = - 0,0875$.

Таблица 5

Класс	Признак			
	0	1	2	3
Легковой автомобиль (1 класс)	$0,65 - 2,475 =$ 1,825	$1,05 - 2,475 =$ - 1,425	$1,35 - 2,475 =$ - 1,125	$0,45 - 2,475 =$ - 2,025
Пассажирский транспорт (2 класс)	$0,1 + 0,875 =$ 0,975	$0,3 + 0,875 =$ 1,175	$-0,3 + 0,875 =$ 0,575	$0,2 + 0,875 =$ 1,075
Грузовой автомобиль (3 класс)	0,325	- 0,275	0,025	0,125

На следующей итерации опять выполняется классификация объектов обучающей выборки. Если не потребовалось ни одной коррекции коэффициентов сети, то считается, что сеть обучена и может использоваться для классификации тестовых объектов. В противном случае вся обучающая выборка повторно подается на вход ИНС для выполнения следующей итерации обучения.