

ЛЕКЦИЯ 16

План лекции:

1. ИНС в виде многослойного персептрона.
2. Решение прикладных задач с помощью многослойного персептрона.

16.1 Многослойный персептрон

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

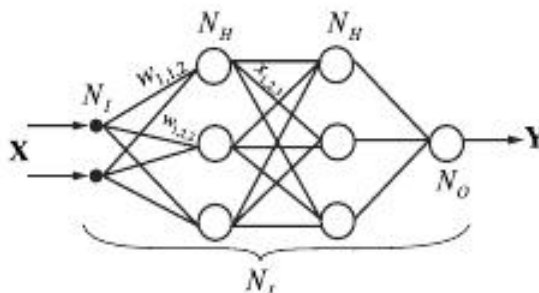


Рис. 1 – ИНС в виде многослойного персептрона

ФН 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, т.к. 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 рисунке 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:

$$NET_{jl} = \sum_i w_{ijl} x_{ijl} \quad (1)$$

$$OUT_{jl} = F(NE_{Tjl} - \theta_{jl}) \quad (2)$$

$$x_{ij(l+1)} = OUT_{il} \quad (3),$$

где индексом i будем обозначать 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 слоя l . В столбцах матрицы расположены вектора 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. Видно отличие работы нейронной сети от разложения функции в ряд:

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.

16.2 Решение прикладных задач с помощью многослойного персептрона

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, т.к. 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.