

Лекция по дисциплине «Нейронные сети»

Лекция 1.11 Модели нейронных сетей- 1 час

Цель: - Рассмотреть основные понятия моделей нейронных сетей;

План: - Многослойные персептроны.

- Радиальные нейроны

Многослойные персептроны являются многослойными нейронными сетями с прямыми связями (с прямым распространением сигнала) и состоят из нескольких слоев нейронов: *входного слоя, нескольких скрытых слоев и выходного слоя.*

На входной слой поступает вектор $X = (x_1, x_2, \dots, x_n)$, который не выполняет никаких вычислений, а просто распределяет вектор X на следующий слой и служит приемником информации. Каждый скрытый элемент можно рассматривать как персептрон, который представляет мягкую пороговую функцию в пространстве входов. Каждый S -элемент связан с совокупностью ассоциативных элементов (A -элементов) первого промежуточного слоя, а A -элементы последнего слоя соединены с реагирующими элементами (R -элементами).

Структура многослойного персептрона с прямыми связями представлена на рисунке П.4.4.

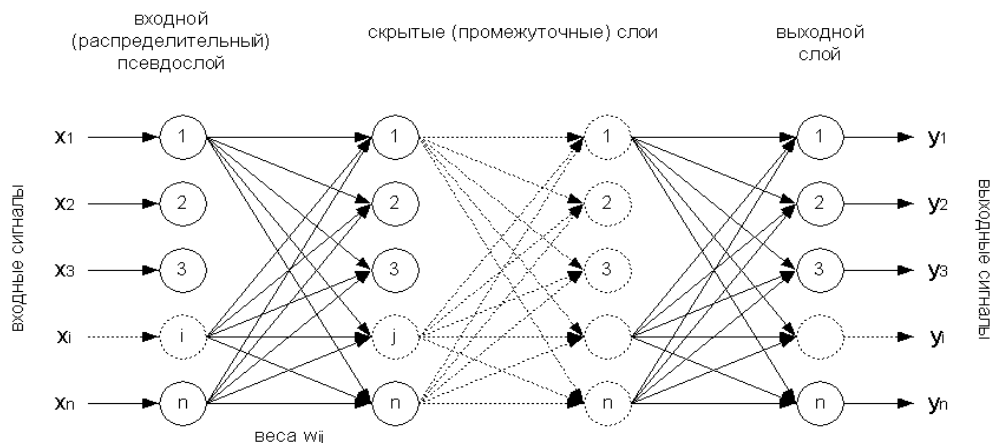


Рисунок П.4.4. Структура многослойного персептрона.

Если скрытых слоев больше, чем два, то во всех скрытых слоях одинаковое количество нейронов. При наличии большего числа скрытых элементов появляется возможность создавать еще больше выступов различных размеров, расположенных в разных местах.

Входные сигналы принимаются нейронами входного слоя, передаются нейронам первого скрытого слоя, и так далее вплоть до выходного, который выдает выходные сигналы всей нейронной сети. Каждый выходной сигнал k -го слоя подается на вход всех нейронов $(k+1)$ -го слоя, образуя последовательные связи. Однако возможен скачок, когда k -ый слой соединяется с произвольным слоем $(k+q)$, где $k > 1$ & $q \in N$.

Введем обозначения: W^k – матрица весов синаптических связей, соединяющих нейроны $k-1$ -го слоя с нейронами k -го слоя; w_{ij}^k – вектор-столбец связи, соединяющий i -й нейрон $k-1$ -го слоя с j -м нейроном k -го слоя $(y_j^{k-1} \in y^{k-1}, y_j^k \in y^k)$.

Аналогично x_j^1 – входной вектор-столбец слоя 1. Тогда прямое функционирование многослойного персептрона описывается соотношениями:

$$s_j^k = f\left(\sum_{i=1}^{n^{(k-1)}} w_{ij}^k y_i^{k-1}\right), \quad y_j^k = f(s_j^k + w_{0j}^k), \quad x_{ij}^{k+1} = y_j^k$$

где i – номер входа, j – номер нейрона в слое, k – номер слоя;

x_{ij}^k – i -й входной сигнал j -го нейрона в слое k ;

w_{ij}^k – вес синапса i -го входа j -го нейрона в слое k ;

s_j^k – взвешенная сумма j -го нейрона в слое k ;

w_{0j}^k – порог j -го нейрона в слое k .

Каждый слой рассчитывает нелинейное преобразование от линейной комбинации сигналов предыдущего слоя. Многослойная нейронная сеть может формировать на выходе произвольную многомерную функцию при соответствующем выборе количества слоев, диапазона изменения сигналов и параметров нейронов. Многослойные нейронные сети являются универсальным аппроксиматором функций. Общая формула, описывающая многослойный персептрон, определяется как:

$$f(s) = f \left(\sum_{i_K} w_{i_K j_K}^K \cdots f \left(\sum_{i_2} w_{i_2 j_2}^2 f \left(\sum_{i_1} w_{i_1 j_1} + w_{j_1 2} \right) + w_{j_2 2} \right) \cdots + w_{j_K K} \right)$$

слой 1
слой 2
слой K

За счет поочередного расчета линейных комбинаций и нелинейных преобразований достигается аппроксимация произвольной многомерной функции при соответствующем выборе параметров сети. Персептроны дают возможность строить более сложные разделяющие поверхности, вследствие чего, и более распространены.

Многослойные персептроны имеют три отличительных признака. Каждый нейрон имеет нелинейную функцию активации. Данная нелинейная функция является гладкой (всюду дифференцируемой), в отличие от жесткой пороговой функции, используемой в персептроне Розенблатта. Самой популярной формой функции, удовлетворяющей этому требованию, является сигмоидальная, определяемая логистической функцией:

$$f(s) = \frac{1}{1 + e^{-s}}$$

Наличие нелинейности играет очень важную роль, так как в противном случае отображение «вход-выход» сети можно свести к обычному однослойному персептрону. Более того, использование логистической функции мотивировано биологически, так как в ней учитывается восстановительная фаза реального нейрона.

Несмотря на то, что возможности многослойных сетей были известны давно, в течение многих лет не было теоретически обоснованного алгоритма для настройки их весов. Далее мы детально изучим многослойные обучающие алгоритмы, на данном этапе достаточно понимать проблему и знать, что исследования привели к определенным результатам.

Многослойные нейронные сети с прямыми связями обучаются с помощью дельта-правила, которое используется при обучении с учителем и реализуется следующим образом:

$$w_{ij}(t+1) = w_{ij}(t) + \eta x_i (d_j - y_j),$$

η – параметр (шаг обучения); d – эталонное (требуемое) значение выхода элемента. выхода элемента.

Таким образом, изменение силы связей происходит в соответствии с ошибкой выходного сигнала ($d - y$) и уровнем активности входного элемента x . Обобщение дельта-правила, называемое обратным распространением ошибки (Back – propagation), применимо к сетям с любым числом слоев. Обучение сети в этом случае состоит из следующих шагов:

1. Выбрать очередную обучающую пару (x, d). Подать входной вектор на вход сети.
2. Вычислить выход сети y .
3. Вычислить разность между выходом сети и требуемым выходом (ошибку).
4. Подкорректировать веса сети так, чтобы минимизировать ошибку.
5. Повторять шаги с 1-го по 4-й для каждой обучающей пары, пока ошибка не достигнет приемлемого уровня.
6. Ошибка функционирования сети обычно определяется как

$$E = \frac{1}{2} \cdot \sum_{j=1}^{n(k)} (d_j - y_j)^2,$$

где $y_j = y_j^k$ – выход сети. Для уменьшения этой ошибки следует изменить веса сети по правилу

$$w^k(t+1) = w^k(t) - \eta \cdot \frac{dE}{dw^k},$$

Для выходного слоя определение ошибки δ_j нейрона j тривиально:

$$\delta_j = (d_j - y_j) \cdot f'(s_j) = (d_j - y_j) \cdot C \cdot f(s_j) \cdot (1 - f(s_j))$$

и напоминает систему поощрений-наказаний, используемую при обучении однослойных сетей. Для каждого предыдущего слоя ошибка определяется рекурсивно через ошибку следующего слоя:

$$\delta_j = \left(\sum_{i=1}^m \delta_i w_{ij} \right) \cdot f'(s_j),$$

т.е. для каждого j -го нейрона ошибки следующего слоя распространяются к нему обратно сквозь соответствующие веса. — Этот механизм обратного распространения дополнен традиционными для многих градиентных методов оптимизации процедурами оценки вектора кратчайшего спуска, изменения величины шага пропорционально крутизне склона и др.

Если при прямом функционировании входной сигнал распространяется по сети от входного слоя к выходному, то при подстройке весов ошибка сети распространяется от выходного слоя к входному.