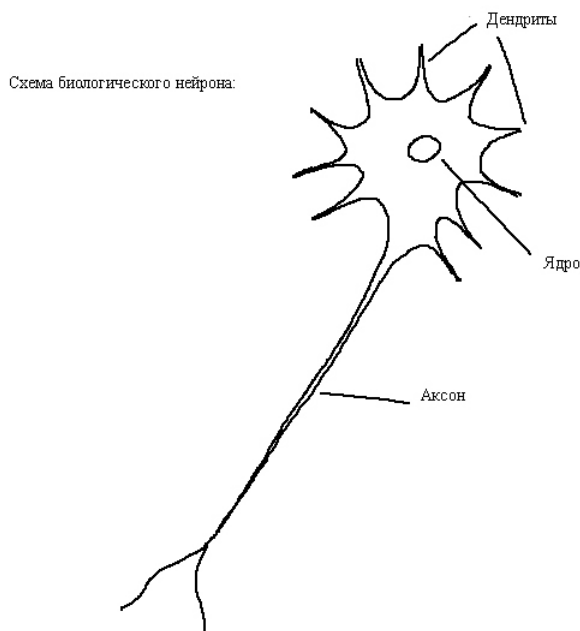


ЛЕКЦИЯ 7

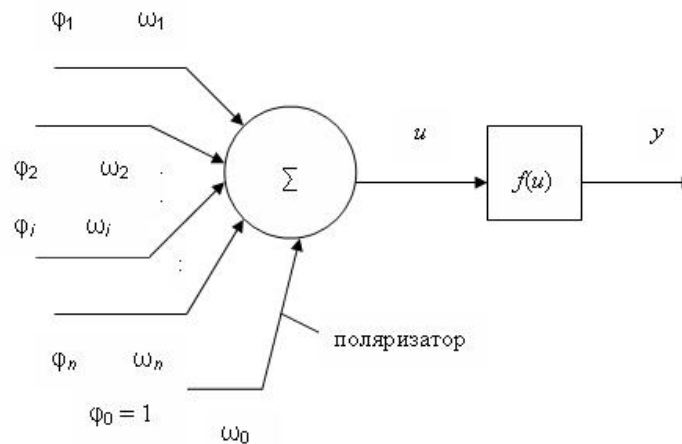
НЕЙРОННЫЕ СЕТИ

Нервная система и мозг человека состоят из нейронов, соединенных между собой нервными волокнами, способными передавать электрические импульсы между нейронами. Нейрон (нервная клетка) является особой биологической клеткой, которая обрабатывает информацию. Он состоит из тела и отростков нервных волокон двух типов - дендритов, по которым принимаются импульсы, и единственного аксона, по которым нейрон может передавать импульс. Тело нейрона включает ядро, которое содержит информацию о наследственных свойствах, и плазму, обладающую молекулярными средствами для производства необходимых нейрону материалов. Нейрон получает сигналы (импульсы) от аксонов других нейронов через дендриты (приемники) и передает сигналы, генерированные телом клетки, вдоль своего аксона (передатчик), который в конце разветвляется на волокна. На окончаниях этих волокон находятся специальные образования - синапсы, которые влияют на силу импульса.

Синапс является элементарной структурой и функциональным узлом между двумя нейронами (волокно аксона одного нейрона и дендрит другого). Когда импульс достигает синаптического окончания, высвобождаются химические вещества, называемые нейротрансмиттерами. Нейротрансмиттеры диффундируют через синаптическую щель, возбуждая или затормаживая, в зависимости от синапса, способность нейрона-приемника генерировать электрические импульсы. Результативность синапса может настраиваться проходящими через него сигналами, так что синапсы могут обучаться в зависимости от активности процессов, в которых они участвуют. Веса синапсов могут изменяться со временем, что изменяет и поведение соответствующего нейрона.



Математическая модель искусственного нейрона была предложена У. Маккалоком и У. Питтсом вместе с моделью сети, состоящей из этих нейронов. Они показали, что сеть на таких элементах может выполнять числовые и логические операции.



Текущее состояние нейрона определяется как взвешенная сумма его входов:

$$u = \sum_{i=1}^n \omega_i \varphi_i.$$

ω_i - вес синапса, число, характеризующее силу связи.

$f(u)$ - функция активации нейронов (это есть нелинейная функция одного аргумента - выхода сумматора).

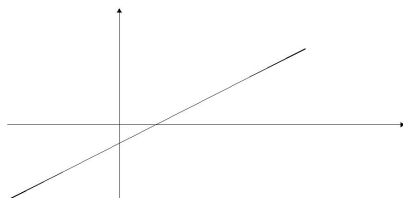
$y = f(u)$, y - выход.

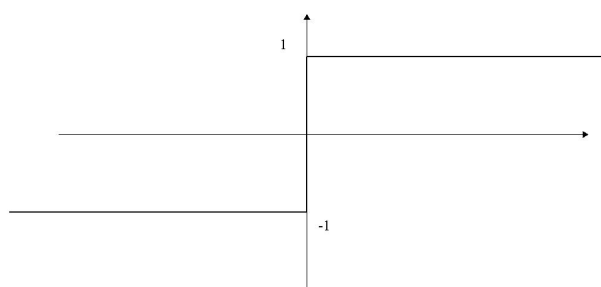
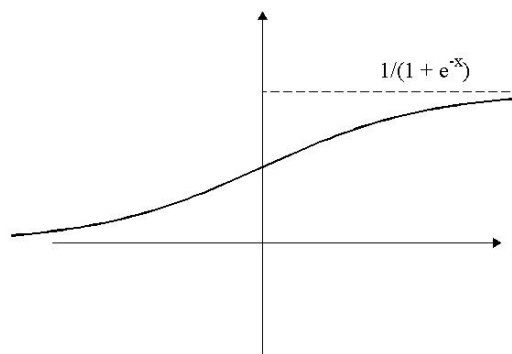
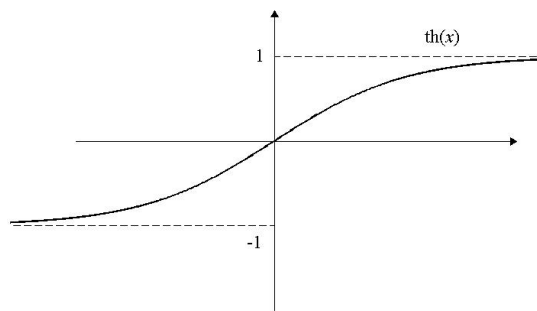
С учетом поляризатора:

$$u = \sum_{i=0}^n \omega_i \varphi_i.$$

Функция активации может иметь различный вид. Одной из наиболее распространенных является нелинейная функция с насыщением, так называемая логистическая функция или сигмоид (т.е. функция S-образного вида): $f(x) = \frac{1}{1+e^{-ax}}$.

Примеры активационных функций:





Существует два подхода к обучению искусственных нейронных сетей - обучение с учителем и обучение без учителя.

Обучение состоит в поиске (синтезе) значений, близких к эталонным в смысле некоторой функции ошибки E .

$$E(\omega) = \frac{1}{2} \sum_{k=1}^p (y^k - d^k)^2.$$

y^k - выход для сети k -го примера.

d^k - эталонное значение.

При обучении с учителем каждый обучающий пример представляет собой следующую структуру:

$$\langle (\varphi_1^k, \varphi_2^k, \dots, \varphi_n^k), (d_1^k, d_2^k, \dots, d_m^k) \rangle.$$

d_m^k - эталонные выходы, m - количество выходов, k - номер примера.

При обучении без учителя пропадает эталонный выход, остается только вход.

Рассмотрим обучение с учителем.

- Правило Хебба:

$$\omega_{ij}[t+1] = \omega_{ij}[t] + \eta d_i \varphi_j,$$

где i - номер нейрона, j - вход, t - номер текущей итерации обучения. Вес входного сигнала нейрона изменяется в сторону уменьшения ошибки пропорционально величине суммарной ошибки нейрона. Часто вводят коэффициент пропорциональности η , на который умножается величина ошибки. Этот коэффициент называют скоростью обучения.

Чем больше корреляция, тем больше надо показывать связь.

- Правило Уидроу-Хоффа:

$$\omega_{ij}[t+1] = \omega_{ij}[t] + \varphi_j (d_i - y_i) \eta.$$

Персептрон - одиночный нейрон с функцией активации.

Возьмем функцию активации $f(u) = \frac{1}{1+e^{-u}}$.

$$f'(u) = -\frac{1}{(1+e^{-u})^2} - e^{-u} = \frac{e^{-u}}{(1+e^{-u})^2}$$

Производная по соответствующему весу:

$$\frac{\partial E(\omega)}{\partial \omega_{ij}} = \sum_{k=1}^p (y^k - d^k) \frac{\partial y^k}{\partial \omega_{ij}} \varphi_j =$$

$$[u_i = \sum_{j=1}^n \omega_{ij} \varphi_j, y_i = f(u_i), \sum_{k=1}^p (y^k - d^k) \frac{\partial y^k}{\partial \omega_{ij}} = \sum_{k=1}^p (y^k - d^k) \frac{\partial f(u_i)}{\partial u_i} \cdot \frac{\partial u_i}{\partial \omega_{ij}}, \frac{\partial u_i}{\partial \omega_{ij}} = \varphi_j]$$

$$= \sum_{k=1}^p (y^k - d^k) \frac{\partial f(u_i)}{\partial u_i} \varphi_j.$$

Это есть градиент ∇E (или $gradE$).

$$\omega_{ij}[t+1] = \omega_{ij}[t] - \eta(\nabla E).$$

Минус в данной формуле означает, что движение в сторону противоположную градиенту.

Иногда обозначают

$$\delta = \sum_{k=1}^p (y^k - d^k) \frac{\partial f(u_i)}{\partial u_i}.$$

Есть способ вместо y использовать взвешенный вход (относится к эмпирическим правилам).

Рассмотрим некоторые типа нейронов.

Адаптивный линейный элемент (Адаптивный линейный нейрон или ADALINE) - частный случай линейного классификатора или искусственной нейронной сети с одним слоем. Был предложен Уидроу и Хоффом в 1960 году, развивая математическую модель нейрона МакКаллока-Питтса. Нейрон типа ADALINE обучается следующим образом:

$$E(\omega) = \frac{1}{2} \sum_{k=1}^p (u^k - d^k)^2,$$

где u^k - взвешенный вход.

$$\nabla E = \sum_{k=1}^p (u^k - d^k) \varphi^k.$$

Здесь φ^k - вектор.

MADALINE - множество нейронов ADALINE.

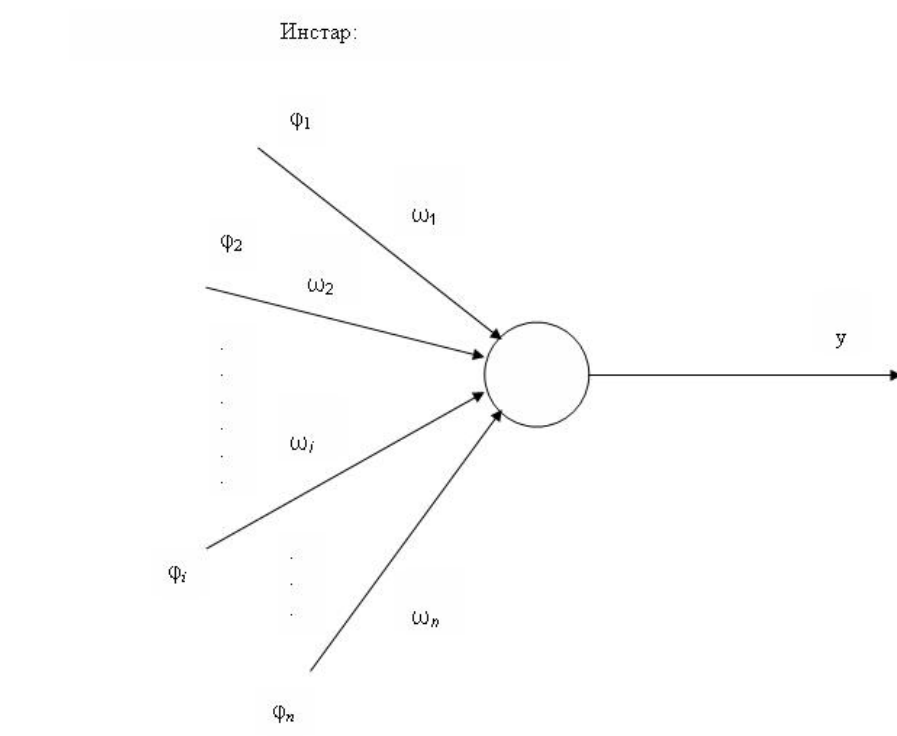
Нейрон Гроссберга: инстар, оутстар.

Правило определения весов для нейронов типа инстар:

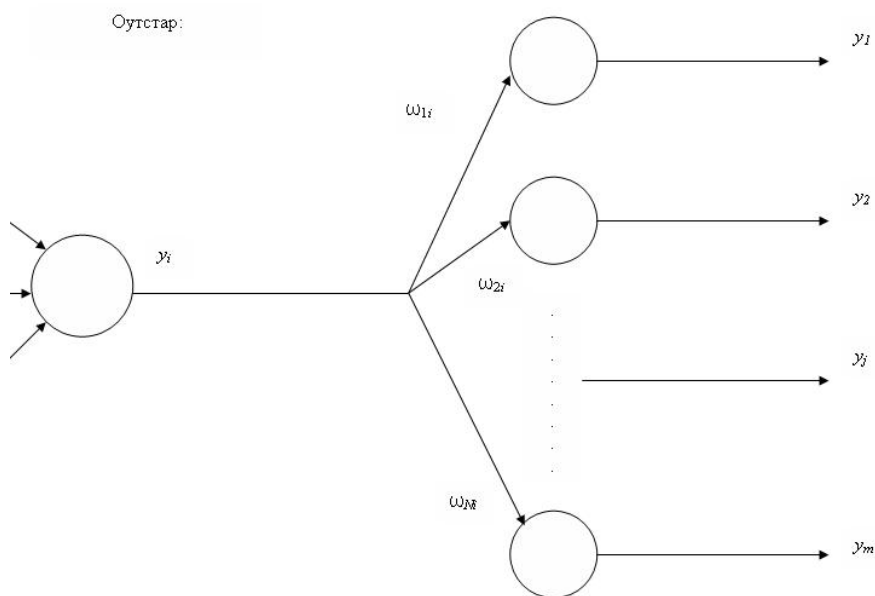
$$\omega_{ij}[t+1] = \omega_{ij}[t] + \eta y_i (\varphi_j - \omega_{ij}[t])$$

Когда получаем $\varphi_j - \omega_{ij}[t] = 0$, алгоритм заканчивается.

Решает первую часть задачи кластеризации.



Оутстар - сеть из нескольких нейронов.

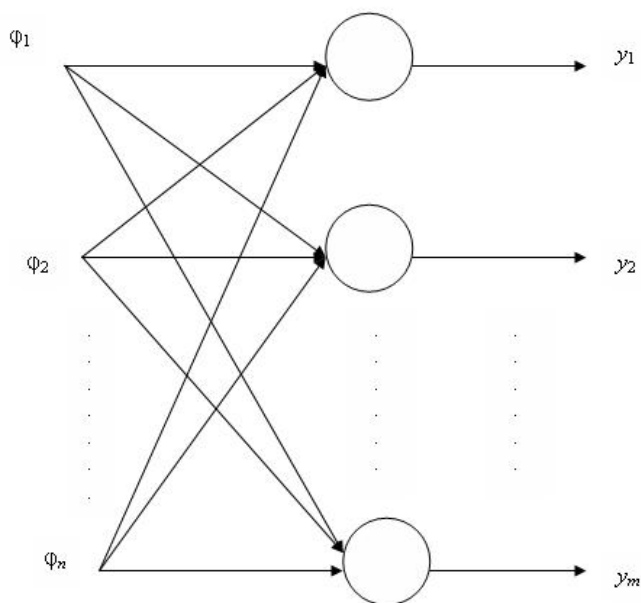


Алгоритм обучения такой сети:

$$\omega_{ji}[t+1] = \omega_{ji}[t] + \eta y_i (y_j - \omega_{ji}[t]).$$

Решает вторую часть задачи кластеризации.

Нейроны типа WTA (Winner Takes All) и WTM (Winner Takes Most) используются группами, в которых конкурируют между собой. Процесс обучения происходит без учителя.



Вариант: усиление связи между входом и выходом.

Алгоритм WTM:

$$\omega_{ij}[t+1] = \omega_{ij}[t] + \eta y_i \phi_{ij}.$$

В WTA обучается один нейрон, у которого выход максимальный, а алгоритм тот же, что WTM.