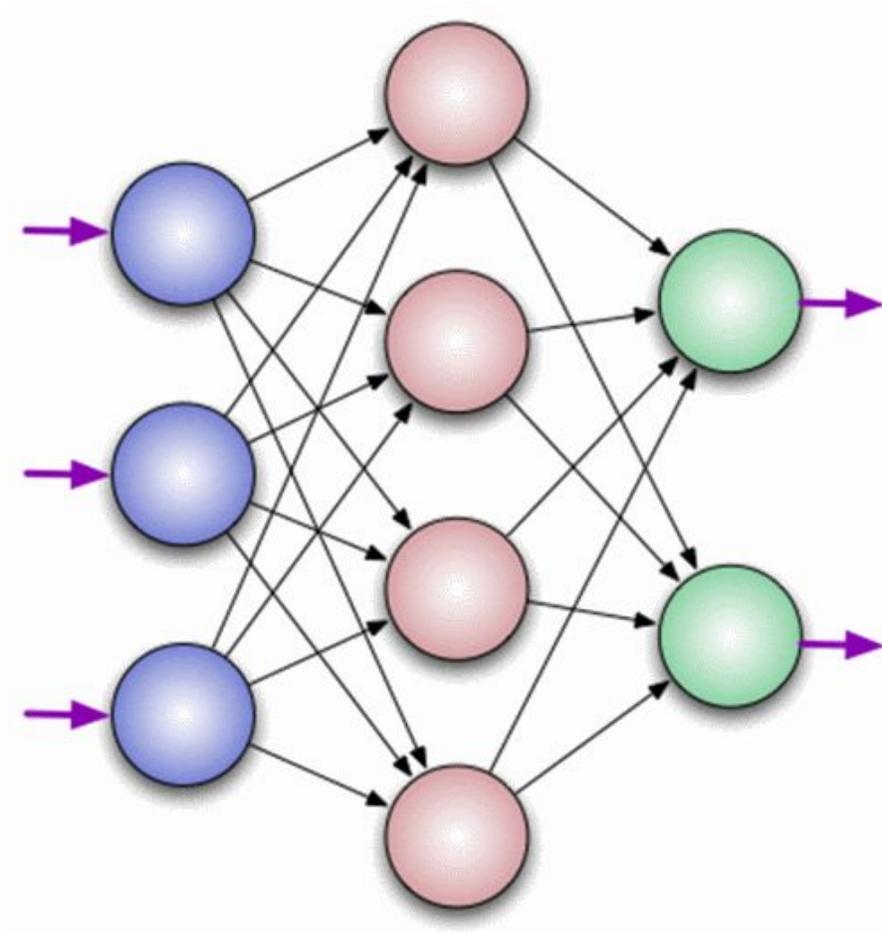


НЕЙРОННЫЕ СЕТИ

Определение

Искусственная нейронная сеть (ИНС)
(англ. *Artificial neural network (ANN)*) —
упрощенная модель биологической
нейронной сети, представляющая собой
совокупность искусственных нейронов,
взаимодействующих между собой.



Задачи, решаемые нейронными сетями

1. Классификация/распознавание образов.
2. Кластеризация/ категоризация
3. Аппроксимация функций.
4. Предсказание/ прогноз
5. Оптимизация
6. Ассоциативная память.
7. Управление.

В общем случае все вышеуказанные задачи, решаемые нейронными сетями, можно свести к двум основным:

Задача классификации

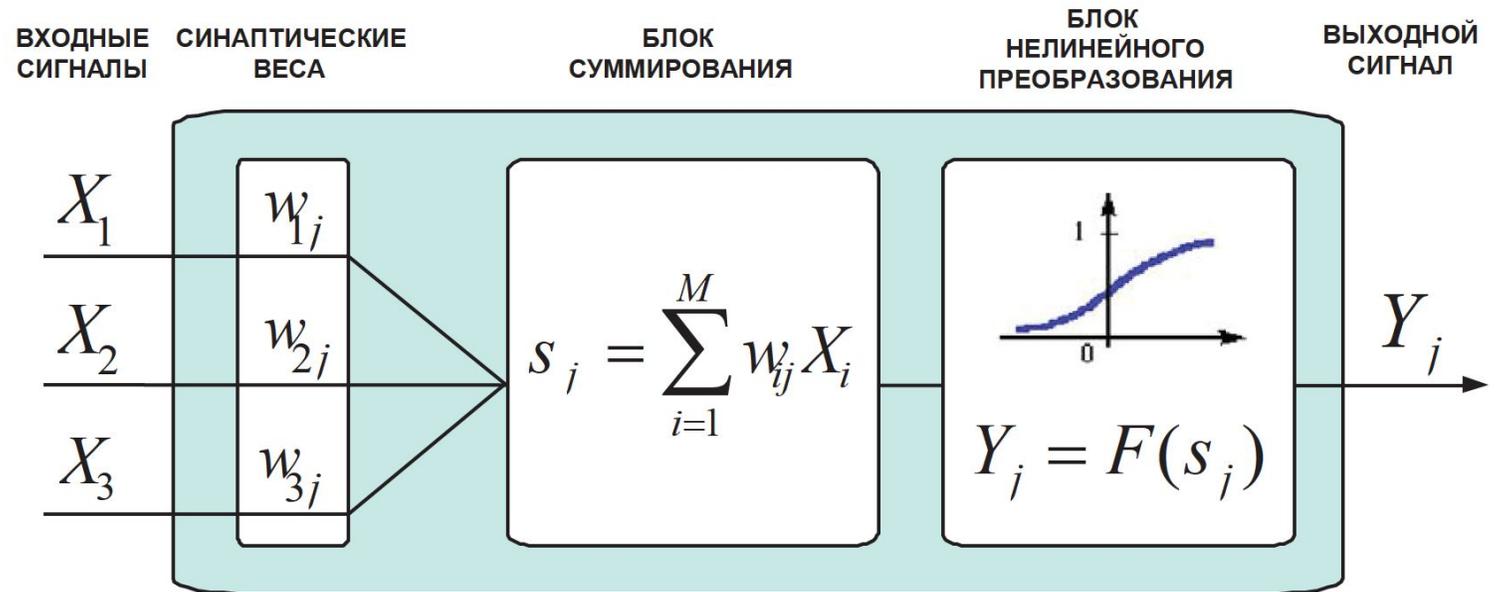
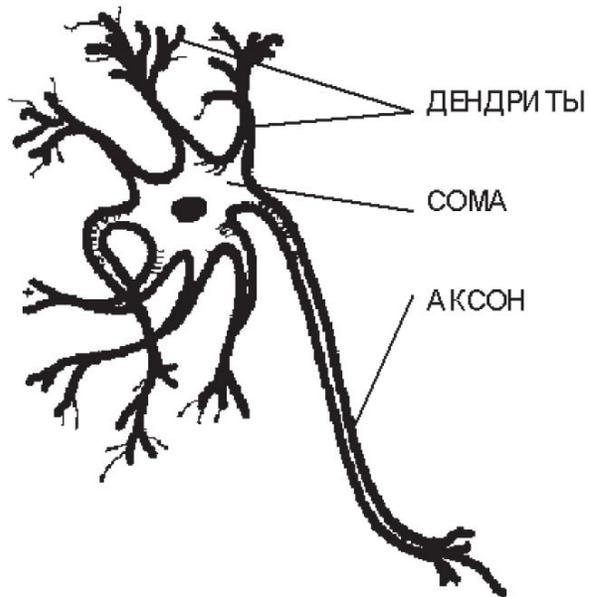
заключается в формировании нейронной сетью в процессе обучения гиперповерхности в пространстве признаков, разделяющей признаки на классы. И выходы обученной нейронной сети соответствуют распознанному классу входного вектора (набора признаков).

Задача регрессии

заключается в аппроксимации нейронной сетью произвольной нелинейной функции. В этом случае значение функции снимается с выхода нейронной сети, а входами являются аргументы.

Биологический нейрон и формальная модель нейрона Маккалоки и Питтса

Биологический нейрон имеет вид, представленный на слайде. В 1943 году Дж. Маккалоки и У. Питт предложили формальную модель биологического нейрона как устройства, имеющего несколько входов (входные синапсы – дендриты), и один выход (выходной синапс – аксон).



Дендриты получают информацию от источников информации (рецепторов) X_i , в качестве которых могут выступать и нейроны. Набор входных сигналов $\{X_i\}$ характеризует объект, его состояние или ситуацию, обрабатываемую нейроном.

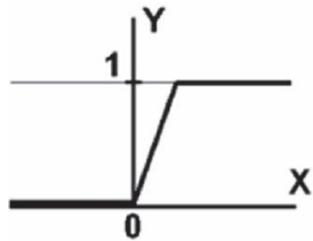
Текущее состояние нейрона определяется, как взвешенная сумма его входов:

$$s = \sum_{i=1}^n x_i \cdot w_i ,$$

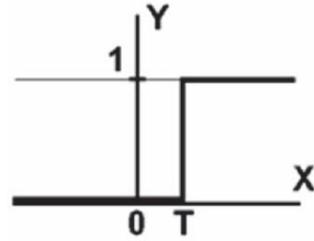
а выход нейрона есть функция его состояния: $y = f(s)$.

Активационная функция нейрона

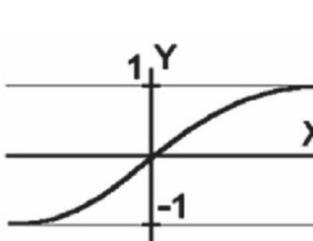
Нелинейная функция f называется активационной и может иметь различный вид:



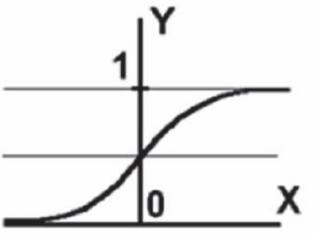
Линейный порог
(гистрезис)



Функция единичного
скачка



Гиперболический
тангенс



Сигмоид

| Название | Формула | Область значений |
|-------------------------|---|----------------------------|
| Линейная | $f(x) = kx$ | $(-\infty, \infty)$ |
| Полулинейная | $f(x) = \begin{cases} kx, & x > 0 \\ 0, & x \leq 0 \end{cases}$ | $(0, \infty)$ |
| Сигмоид | $f(x) = \frac{1}{1 + e^{-x}}$ | $(0, 1)$ |
| Гиперболический тангенс | $f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ | $(-1, 1)$ |
| Экспоненциальная | $f(x) = e^{-x}$ | $(0, \infty)$ |
| Квадратичная | $f(x) = x^2$ | $(0, \infty)$ |
| Знаковая | $f(x) = \begin{cases} 1, & x > 0 \\ -1, & x \leq 0 \end{cases}$ | $(-1, 1)$ |

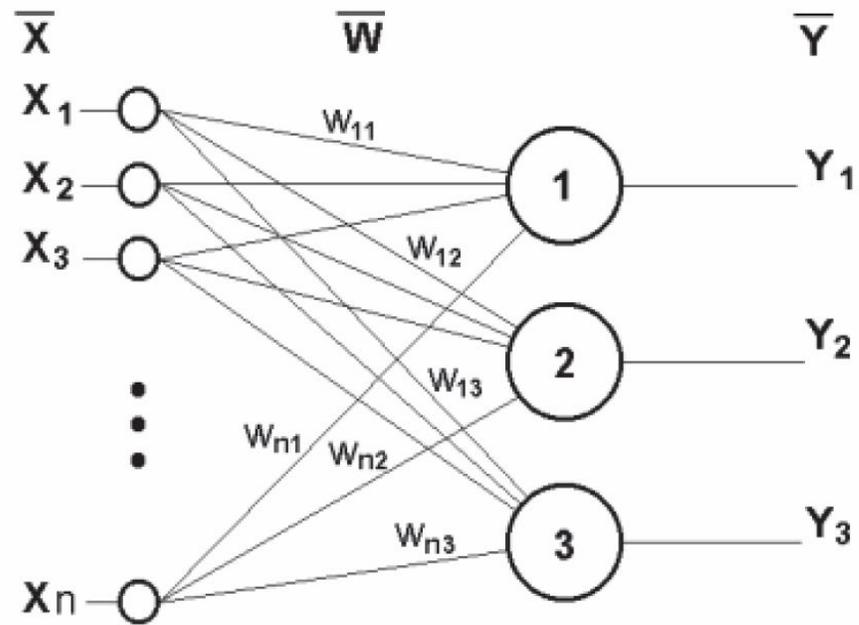
Сигмоидная функция

Одной из наиболее распространенных является нелинейная функция с насыщением, так называемая логистическая функция или сигмоид (т.е. функция S-образного вида):

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

Следует отметить, что сигмоидная функция дифференцируема на всей оси абсцисс, что используется в некоторых алгоритмах обучения.

Простейшая нейронная сеть



$$y_j = f \left[\sum_{i=1}^n x_i \cdot w_{ij} \right], j = 1 \dots 3$$

Машинное обучение нейронной сети на примерах

Обучение классической нейронной сети состоит в подстройке весовых коэффициентов каждого нейрона.

Обучающая выборка $\{x^\alpha, y^\alpha\}, \alpha = 1..p,$

Вектор $\{x^\alpha\}$ характеризует систему признаков конкретного объекта α обучающей выборки, зафиксированную S -элементами.

Вектор $\{y^\alpha\}$ характеризует картину возбуждения нейронов при предъявлении нейронной сети конкретного объекта α обучающей выборки.

$$x_i^\alpha = \begin{cases} 1, & \text{если у объекта } \alpha \text{ наблюдается } i - \text{й признак,} \\ 0, & \text{если у объекта } \alpha \text{ } i - \text{й признак не наблюдается;} \end{cases}$$

$$y_j^\alpha = \begin{cases} 1, & \text{если при предъявлении объекта } \alpha \text{ активизируется } j - \text{й нейрон,} \\ 0, & \text{если при предъявлении объекта } \alpha \text{ } j - \text{й нейрон не активизируется;} \end{cases}$$

Будем называть нейронную сеть **обученной на данной обучающей выборке**, если при подаче на вход сети вектора $\{x^\alpha\}$ на выходе всегда получается соответствующий вектор $\{y^\alpha\}$ т.е. каждому набору признаков соответствуют определенные классы.

Итерационный алгоритм обучения НС

Шаг 0: Начальные значения весов всех нейронов полагаются случайными.

Шаг 1: Сети предъявляется входной образ x , в результате формируется выходной образ.

Шаг 2: Вычисляется вектор ошибки, делаемой сетью на выходе.

Итерационный алгоритм обучения НС

Шаг 3: Вектора весовых коэффициентов корректируются таким образом, что величина корректировки пропорциональна ошибке на выходе и равна нулю если ошибка равна нулю:

- модифицируются только компоненты матрицы весов, отвечающие ненулевым значениям входов;
- знак приращения веса соответствует знаку ошибки, т.е. положительная ошибка (значение выхода меньше требуемого) проводит к усилению связи;
- обучение каждого нейрона происходит независимо от обучения остальных нейронов, что соответствует важному с биологической точки зрения, принципу локальности обучения.

Итерационный алгоритм обучения НС

Шаг 4: Шаги 1-3 повторяются для всех обучающих векторов.

Один цикл последовательного предъявления всей выборки называется эпохой. Обучение завершается по истечении нескольких эпох, если выполняется по крайней мере одно из условий:

- когда итерации сойдутся, т.е. вектор весов перестает изменяться;
- когда полная просуммированная по всем векторам абсолютная ошибка станет меньше некоторого малого значения.

Данный метод обучения был назван Ф.Розенблаттом «методом коррекции с обратной передачей сигнала ошибки». Имеется в виду передача сигнала ошибка от выхода сети на ее вход, где и определяются, и используются весовые коэффициенты. Позднее этот алгоритм назвали « α - правилом».

