ОСНОВЫ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ

ЛЕКЦИЯ 1

История и перспективы развития нейронных сетей. Области применения ИНС. Биологический нейрон и его математическая модель.

к.т.н., доцент Савченко Е.Ю.

В истории развития нейронных сетей выделяют несколько этапов:

1943-1958г.г.
ПРОРЫВ

- формулировка понятия «нейронная сеть» представление модели нейронной сети У. Маккалоком, У. Питтсом (1943 г.);
- публикация работ о кибернетике Н. Винером (1948 г.);
- создание первого алгоритма обучения Д. Хеббом (1949 г.);
- изобретение однослойного персептрона Ф. Розенблаттом (1958 г.).

1969Г.
ПЕССИМИЗМ

- невозможность решения сетями задач, которые внешне весьма сходны с успешно решаемыми;
- невозможность решения однослойными сетями простых задач, в том числе реализации функции «исключающее ИЛИ»;
- публикация М. Минским (1969 г.) формального доказательства ограниченности персептрона, его неспособности решать широкий круг стоящих задач.

1974г.-настоящее время
ОПТИМИЗМ

- разработан П. Вербосом в 1974 г. алгоритма обратного распространения ошибки для обучения многослойных персептронов (переоткрытый в 1982 г. в исследованиях Д. Паркера, в 1986 г. Д. Румельхартом, Дж. Хинтоном, Р. Вильямсом и независимо одновременно С.И. Барцевым, В.А. Охониным)
- Дальше исследования показали не универсальность предложенного метода (долгий процесс обучения, возможное не обучение сети в результате паралича сети, попадания в локальный минимум).
- 1975 г. Фукусимой представлен Когнитрон – самоорганизующаяся сеть, инвариантное распознавание образов;

1974г.-настоящее время
ОПТИМИЗМ

- 1982 г. разработка Дж. Хопфилдом нейронной сети с обратными связями (основа нейронных рекуррентных сетей – представление о нейронных сетях как об ассоциативной памяти);
- 1982 г. представление Кохоненом моделей сети обучающейся на основе самоорганизации (без учителя);
- разработка Р. Хехт-Нильсоном сетей встречного распространения;

1974г.-настоящее время
ОПТИМИЗМ

- решение проблемы стабильности-пластичности в 1987 г. Гроссбергом при создании адаптивной резонансной теории, построенных на ее основе моделей сетей.
- В 2000-е годы решена проблема попадания в локальный минимум применением стохастических методов обучения (обучение Коши, Больцмановское), создание алгоритмов глубокого обучения нейронных сетей.

Основные проблемы, решаемые искусственными нейронными сетями

- ***Классификация образов***
- ***Кластеризация / категоризация***
- ***Аппроксимация функций***
- ***Предсказание/прогноз***
- ***Оптимизация***

Классификация образов

Задача классификáции — задача, в которой имеется множество объектов (ситуаций), разделённых некоторым образом на классы. Задано конечное множество объектов, для которых известно, к каким классам они относятся.

Ирисы Фишера - самый популярный в статистической литературе набор данных, часто используемый для иллюстрации работы различных алгоритмов классификации. При всем желании мы не смогли без него обойтись, поскольку в современных реальных приложениях редко встречаются такие компактные наборы данных, позволяющие построить хороший классификатор при минимуме исходных признаков.

Выборка состоит из 150 экземпляров ирисов трех видов, для которых измерялись четыре характеристики: длина и ширина чашелистика, длина и ширина лепестка.

Ирис щетинистый (*Iris setosa*) | Ирис разноцветный (*Iris versicolor*) | Ирис виргинский (*Iris virginica*)

Кластеризация / категоризация

Кластеризация (или кластерный анализ) — это задача разбиения множества объектов на группы, называемые кластерами. Внутри каждой группы должны оказаться «похожие» объекты, а объекты разных группы должны быть как можно более отличны. Главное отличие кластеризации от классификации состоит в том, что перечень групп четко не задан и определяется в процессе работы алгоритма.

Применение

Маркетинг

- Кластеризация широко используется при изучении рынка для обработки данных, полученных из различных опросов.
- Может применяться для выделения типичных групп покупателей, разделения рынка для создания персонализированных предложений, разработки новых линий продукции.

Компьютерные науки

- Подбор рекомендаций для пользователя на основе предпочтений других пользователей в данном кластере.

Медицина

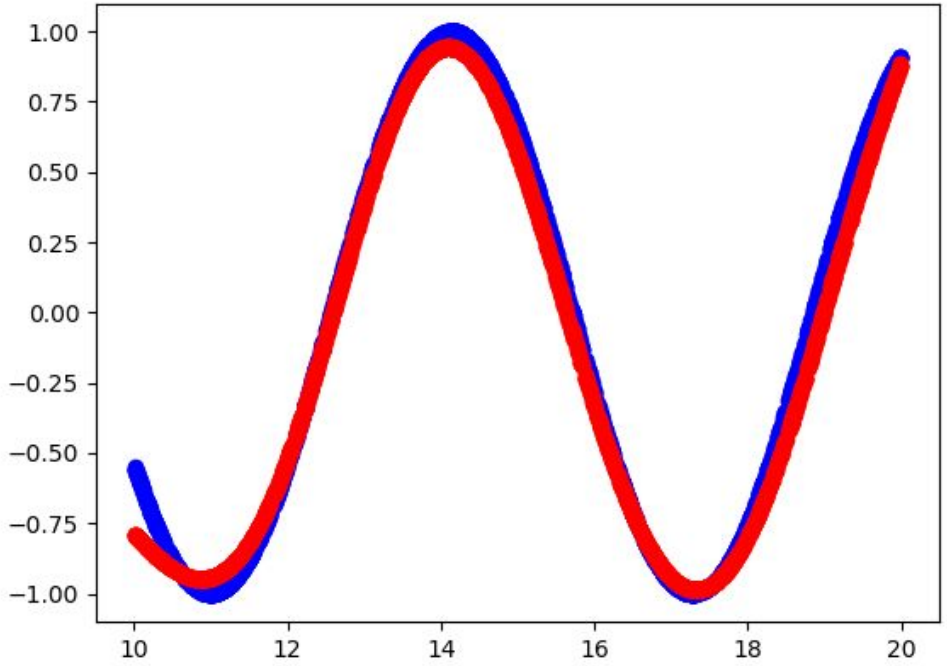
- Применяется для выявления шаблонов устойчивости к антибиотикам;
- для кластеризации антибиотиков по типу антибактериальной активности.

Назад

Аппроксимация функций

Основная ***задача аппроксимации*** — построение приближенной (аппроксимирующей) **функции**, в целом наиболее близко проходящей около данных точек или около данной непрерывной **функции**.

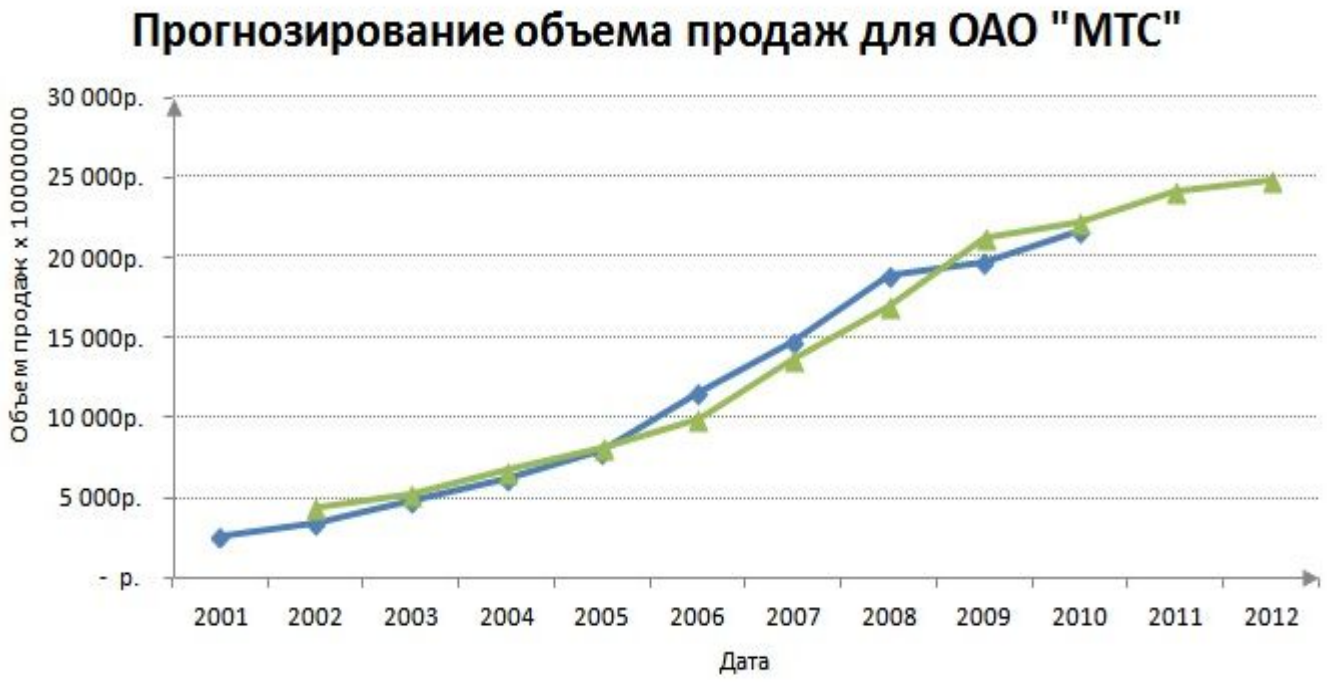
Предположим, что имеется обучающая выборка, заданная парами вход-выход: $((x_1, y_1), (x_2, y_2), \ldots (x_n, y_n))$ полученными от системы, описываемой неизвестной функцией *f*. Задача аппроксимации состоит в нахождении такой ИНС, поведение которой соответствует данной функции.



- Синий цвет — исходная функция
- Красный цвет — аппроксимация функции

Предсказание / прогноз

Пусть заданы *n* дискретных отсчетов$\{y(t_1), y(t_2) \ \dots y(t_n)\}$ в последовательные моменты времени t_1, $t_2 \dots t_n$. Задача состоит в предсказании значения $y(t_{n+1})$. Прогнозирование имеет большое значение при принятии решений в разных областях человеческой деятельности



Оптимизация

Многие проблемы в науке, технике, медицине и экономике могут рассматриваться как задачи оптимизации. Под задачей оптимизации понимается нахождение такого решения, которое удовлетворяет системе ограничений и обеспечивает экстремум заданной целевой функции

Под оптимизацией понимают процесс выбора наилучшего варианта из всех возможных. Большее количество задач сводятся к задачам оптимизации:

- безусловная оптимизация нелинейных функций;
- метод наименьших квадратов;
- решение нелинейных уравнений;
- линейное программирование;
- квадратичное программирование;
- условная минимизация нелинейных функций.

Области применения искусственных нейронных сетей

- Контроль операций с кредитными карточками
- Медицинская диагностика
- Распознавание образов. Классификация образов.
- Анализ потребительского рынка
- Спортивное прогнозирование
- Оценка стоимости продуктов
- Прогнозирование потребления услуг
- Кластеризация /категоризация

Научные исследования

- Прогнозирование магнитных бурь
- Прогнозирование ливневых дождей
- Оптимизация
- Управление

Производство

- Определение качества пива
- Производство микросхем
- Применение в строительстве

Биологический нейрон и его математическая модель.

***Нейрон* (нервная клетка)** является особой биологической клеткой, которая обрабатывает информацию (рис. 1.).

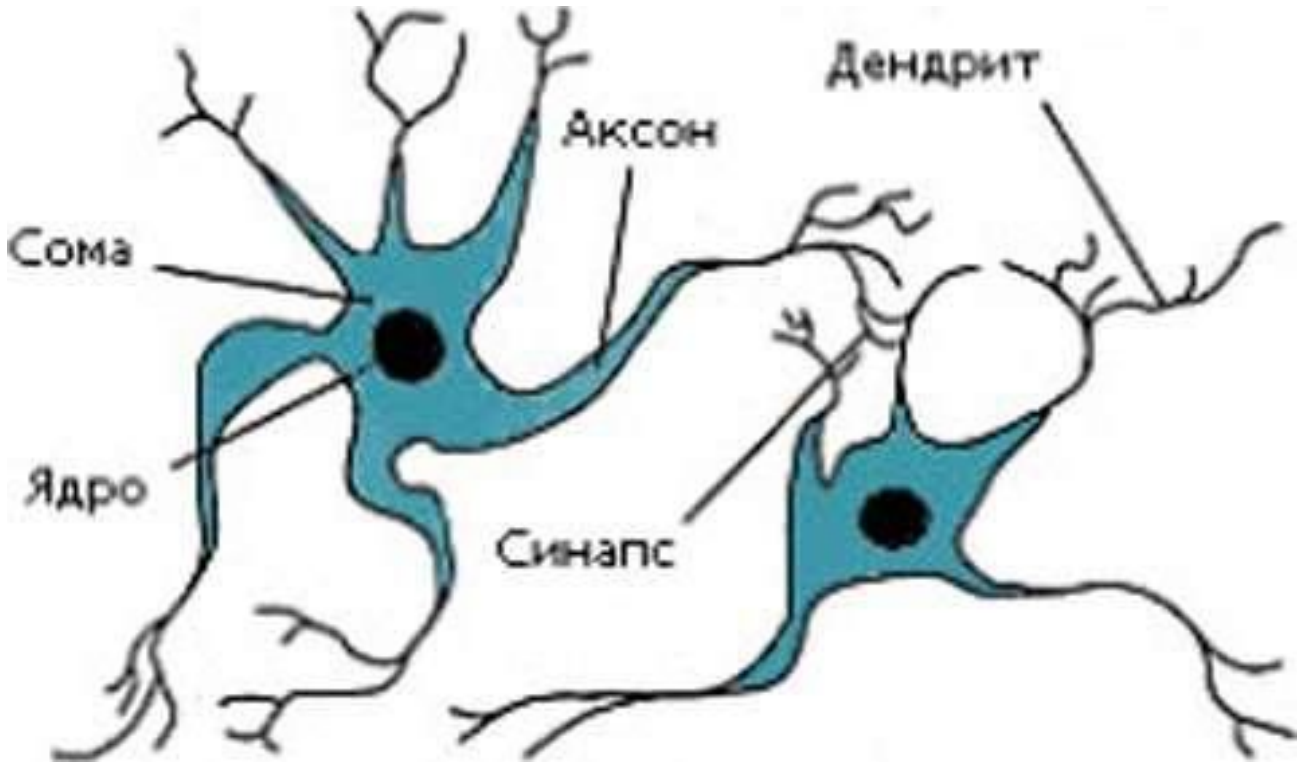


Рис. 1. Взаимосвязь биологических нейронов

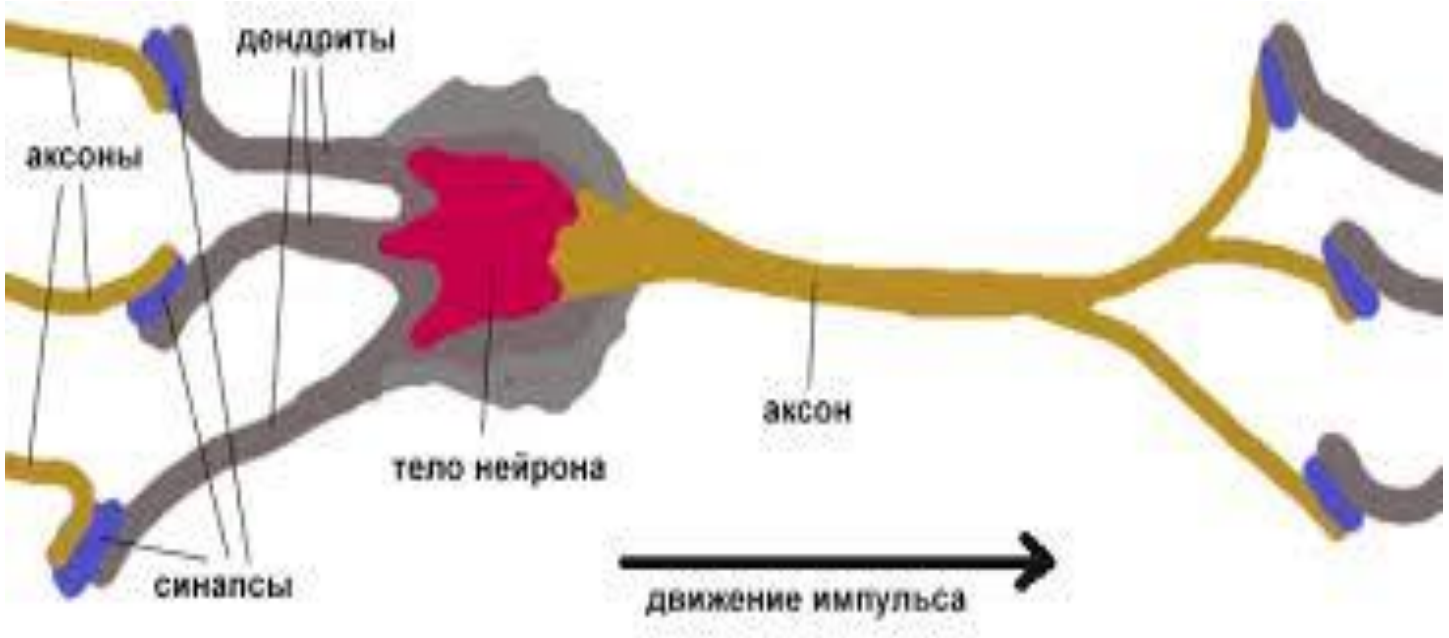


Рис. 2. Биологический нейрон

Искусственный нейрон

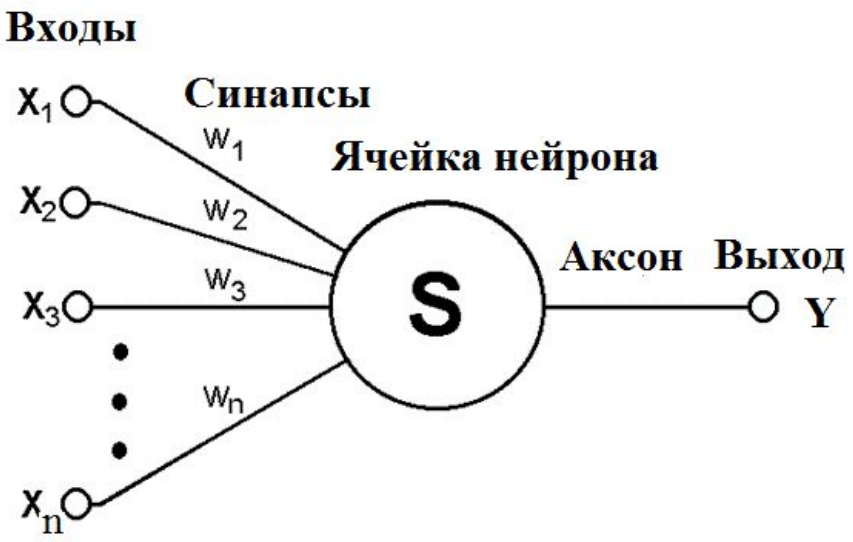


Рис.3 Искусственный нейрон

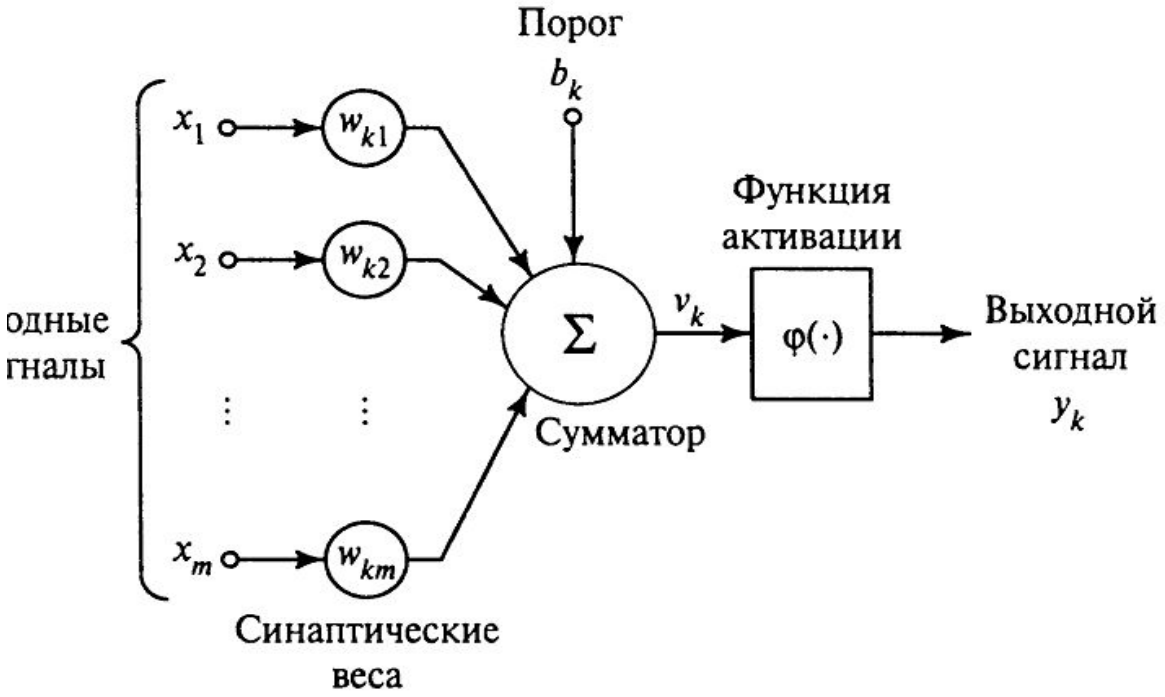


Рис. 4. Нелинейная модель нейрона

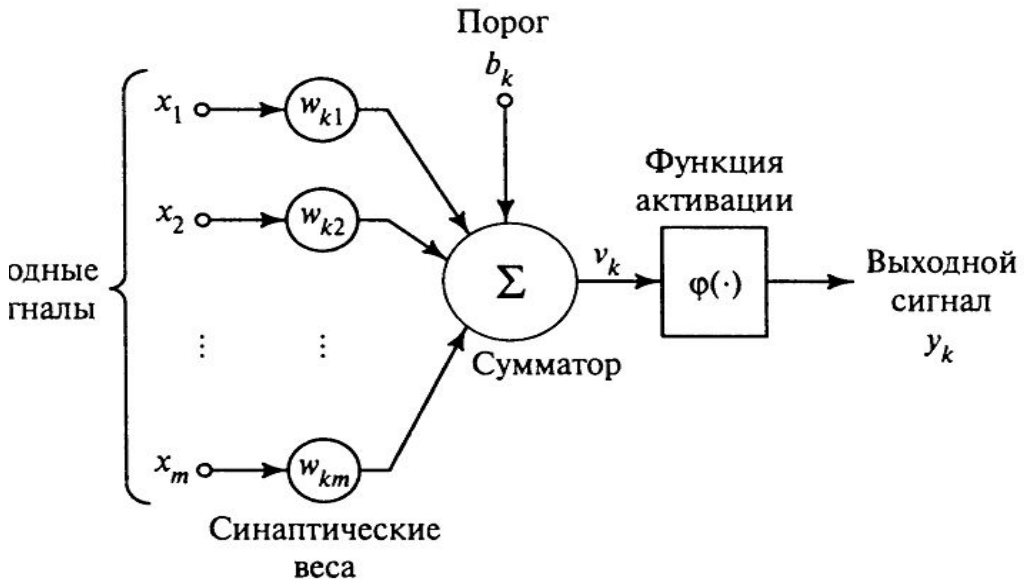


Рис. 5. Нелинейная модель нейрона

$$u_k = \sum_{j}^{m} w_{kj} x_j \qquad (1)$$

$$v_k = u_k + b_k \qquad (2)$$

$$y_k = \varphi(u_k + b_k) \qquad (3)$$

$x_1, x_2, \ldots, x_m$-входные сигналы;
$w_{k1}, w_{k2}, \ldots, w_{km}$- синаптические веса нейрона k;
u_k- линейная комбинация входных воздействий;
b_k- порог;
$\varphi(\cdot)$ - функция активации;
y_k- выходной сигнал нейрона

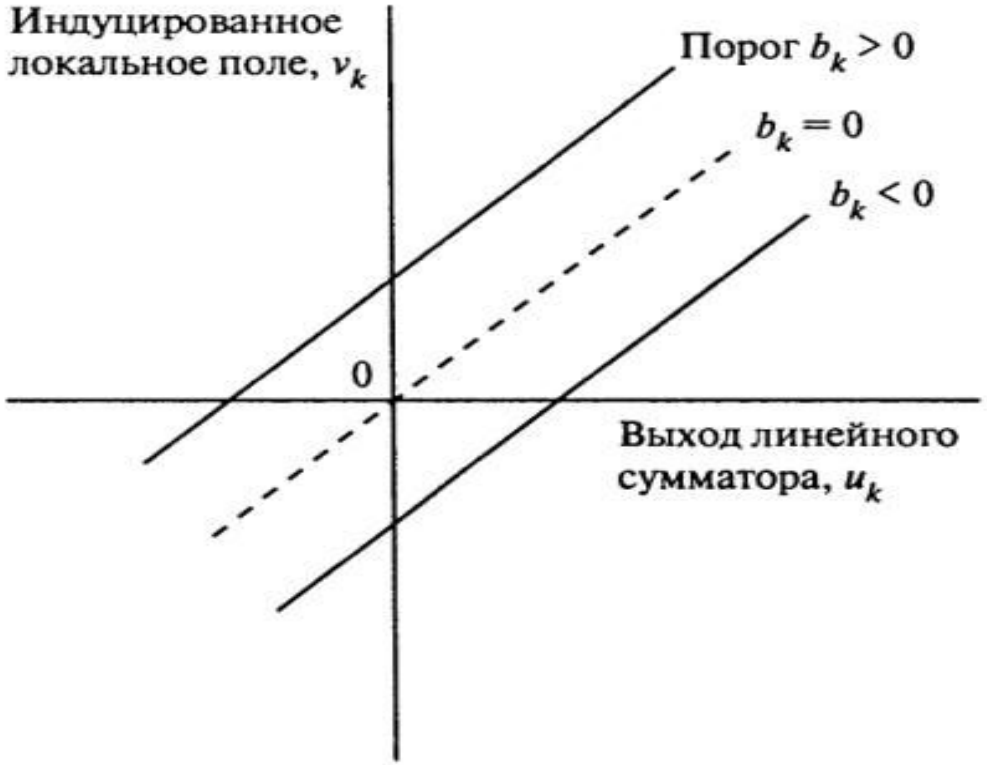


Рис.6 Аффинное преобразование, вызванное наличием порога.

где $u_k = 0, \quad v_k = b_k$ (4)

Тогда формулу (1) можно преобразовать к следующему виду:

$$v_k = \sum_{j=0}^{m} w_{kj} x_j, \quad (4)$$

$$y_k = \varphi(v_k) \quad (5)$$

В выражении 4 добавился новый синапс. Его входной сигнал равен:

$$x_0 = +1, \ (6)$$

а его вес : $w_{k0} = b_k$ (7)

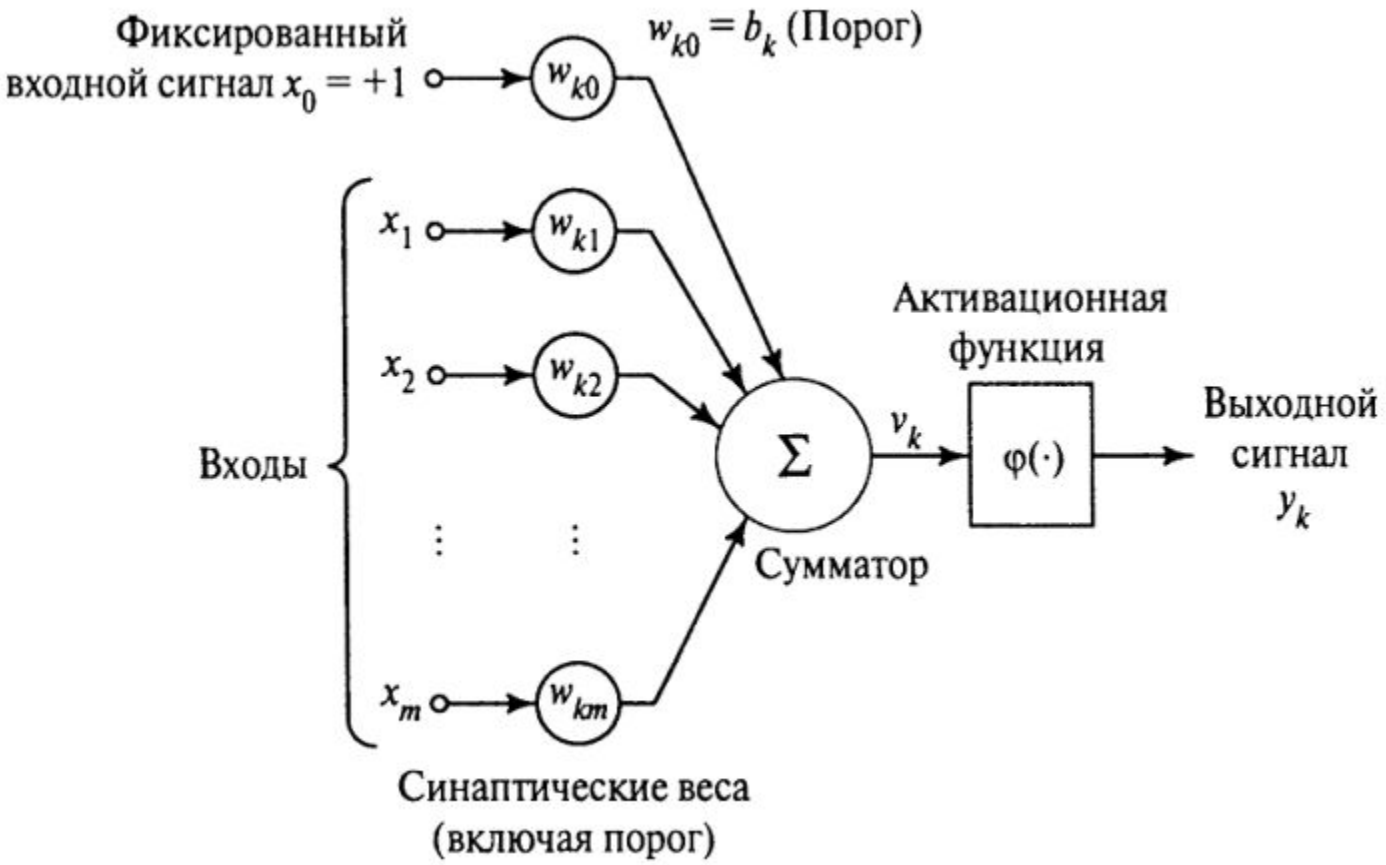


Рис.7 Нелинейная модель нейрона

Типы функций активации

Функции активации, представленные в формулах как $\varphi(.)$, определяют выходной сигнал нейрона в зависимости от индуцировано локального поля v

№	Название	Название функции в Matlab	Формула	Область значений
1	Пороговая	hardlim(v)		0, 1
2	Знаковая	hardlims(v)		-1, 1
3	Сигмоидальная	logsig(v)		(0, 1)
4	Полулинейная	poslin(v)		(0, ∞)
5	Линейная	purelin(x)		(-∞, ∞)

Типы функций активации

Функции активации, представленные в формулах как $\varphi(.)$, определяют выходной сигнал нейрона в зависимости от индуцировано локального поля v

№	Название	Название функции в Matlab	Формула	Область значений
6	Радиальная базисная	radbas(v)		(0, 1)
7	Полулинейная с насыщением	satlin(v)		(0, 1)
8	Линейная с насыщением	satlins(x)		(-1, 1)
9	Гиперболический тангенс	tansig(v)		(-1, 1)

Задача 1.

Нейрон получает входной сигнал x_m, где $m = 4$, блочная диаграмма данной модели представлена на рисунке

$x_1 = 10, x_2 = -20,\ x_3 = 4, x_4 = -2.$

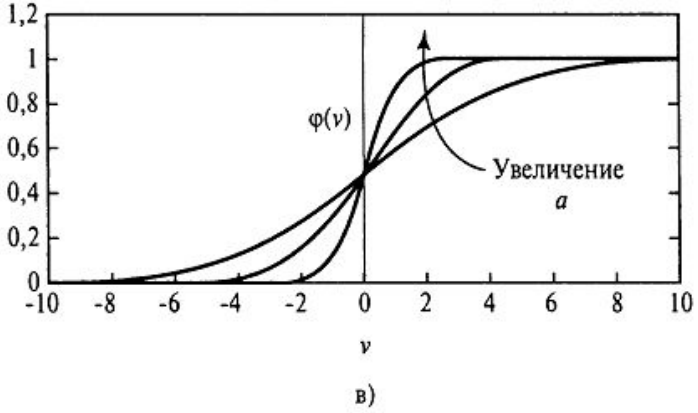
Соответствующие весовые коэффициенты нейрона равны

$w_{11} = 0.8, w_{12} = 0.2,\quad w_{13} = 0.4,$

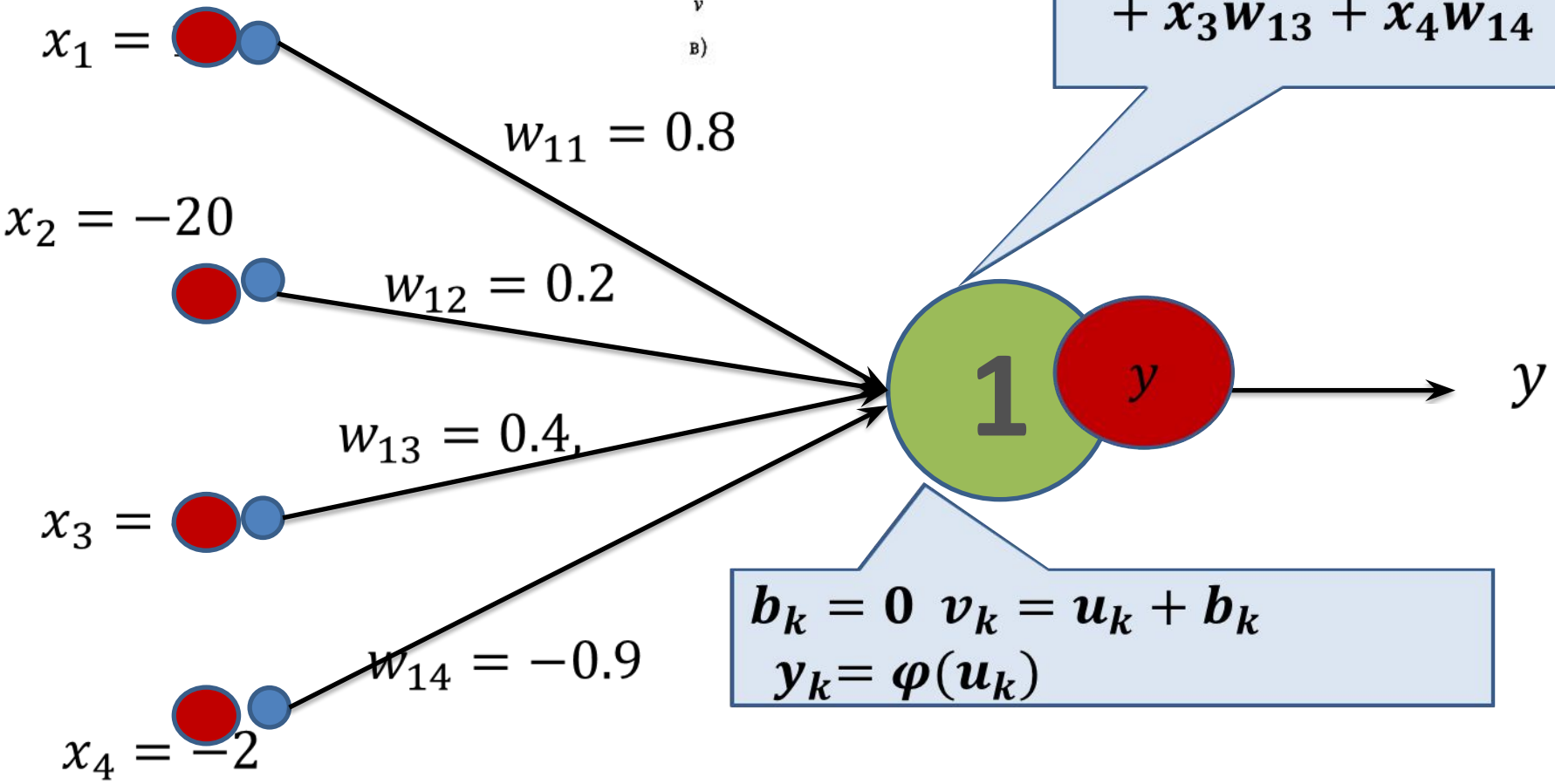
$$w_{14} = -0.9$$

Вычислить выходное значение нейрона, модель которого описывается сигмоидальной функцией активации. Предполагается, что порог отсутствует.

$$\varphi(v) = \frac{1}{1+e^{-v}}.$$



$$\boldsymbol{u_1 = x_1 w_{11} + x_2 w_{12} + x_3 w_{13} + x_4 w_{14}}$$



$$\boldsymbol{b_k = 0 \quad v_k = u_k + b_k}$$
$$\boldsymbol{y_k = \varphi(u_k)}$$

Решение.

Сигмоидальная функция активации имеет вид: $\varphi(v) = \frac{1}{1+e^{-v}}$.

Сделаем вычисления согласно следующим формулам

$u_k = \sum_{j=0}^{m} w_{kj} x_j$,

$v_k = u_k + b_k$

$y_k = \varphi(u_k)$

Тогда,

$$u_1 = x_1 w_{11} + x_2 w_{12} + x_3 w_{13} + x_4 w_{14}$$

$$u_1 = 10 * 0.8 + (-20 * 0.2) + 4 * 0.4 + (-2 * -0.9) = 7.4$$

$$v_1 = u_1 + b_1$$

$b_1 = 0$, т.к. по условию порог отсутствует

$$v_1 = 7.4$$

$y_1 = \frac{1}{1+e^{-7.4}}$=0,99

Таким образом, выход нейронной сети на данное входное множества равен $y_1 = 0{,}99$