модели ассоциативной памяти, сеть Хопфилда, машина Больцмана

Евгений Борисов

Нейросети

Ассоциативная память - адресация по содержанию,

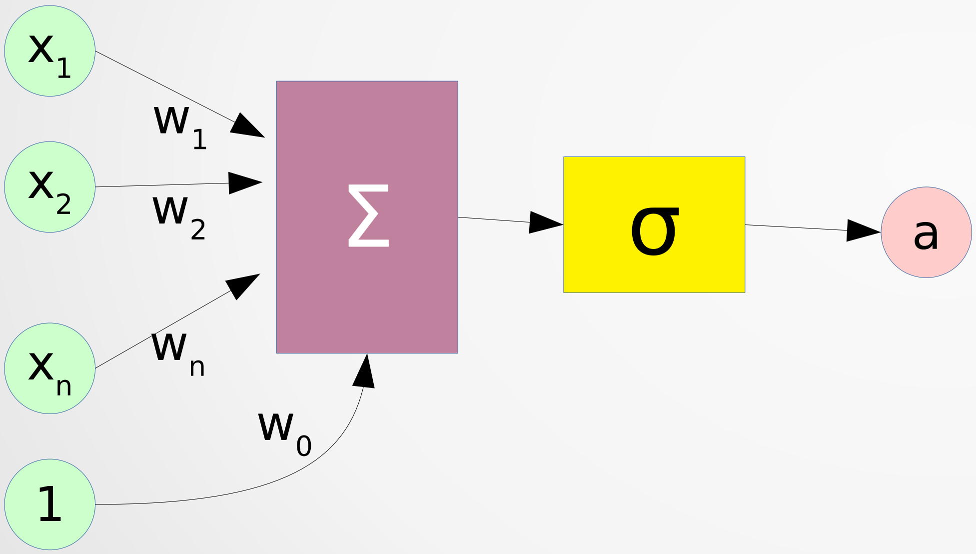
чтение и запись в ячейки такой памяти
выполняется в зависимости от их содержимого

[входной образ] → [выходной образ (ассоциация)]

Нейросети

модель нейрона

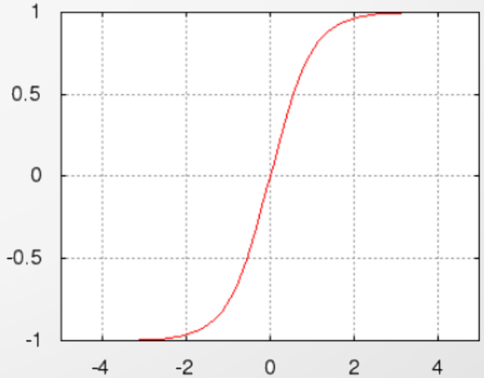
$$a(x,w)=\sigma\left(\sum_{i=1}^{n} x_i\cdot w_i - w_0\right)=\sigma(\langle x,w\rangle)$$



x_i - вход
w_i - вес связи
σ - функция активации
s - состояние нейрона

$$s(x,w)=\sum_{i=1}^{n} x_i\cdot w_i - w_0$$

σ - функция активации

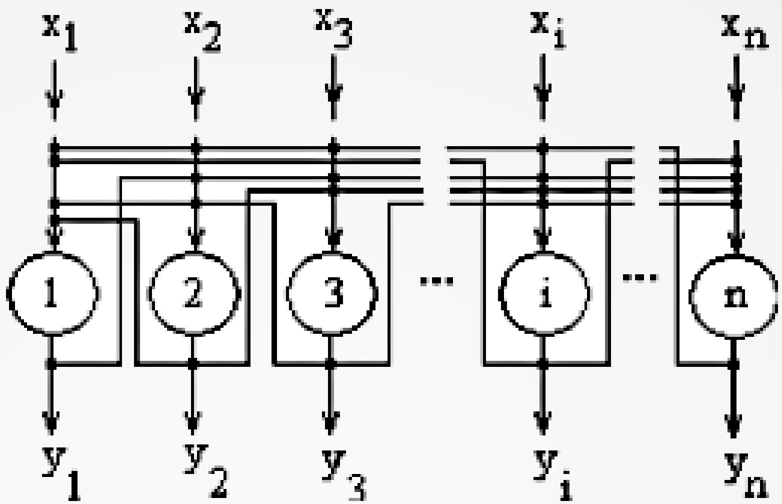


биполярный сигмоид

Нейросети

сеть Хопфилда

(J.J.Hopfield 1982)



один слой

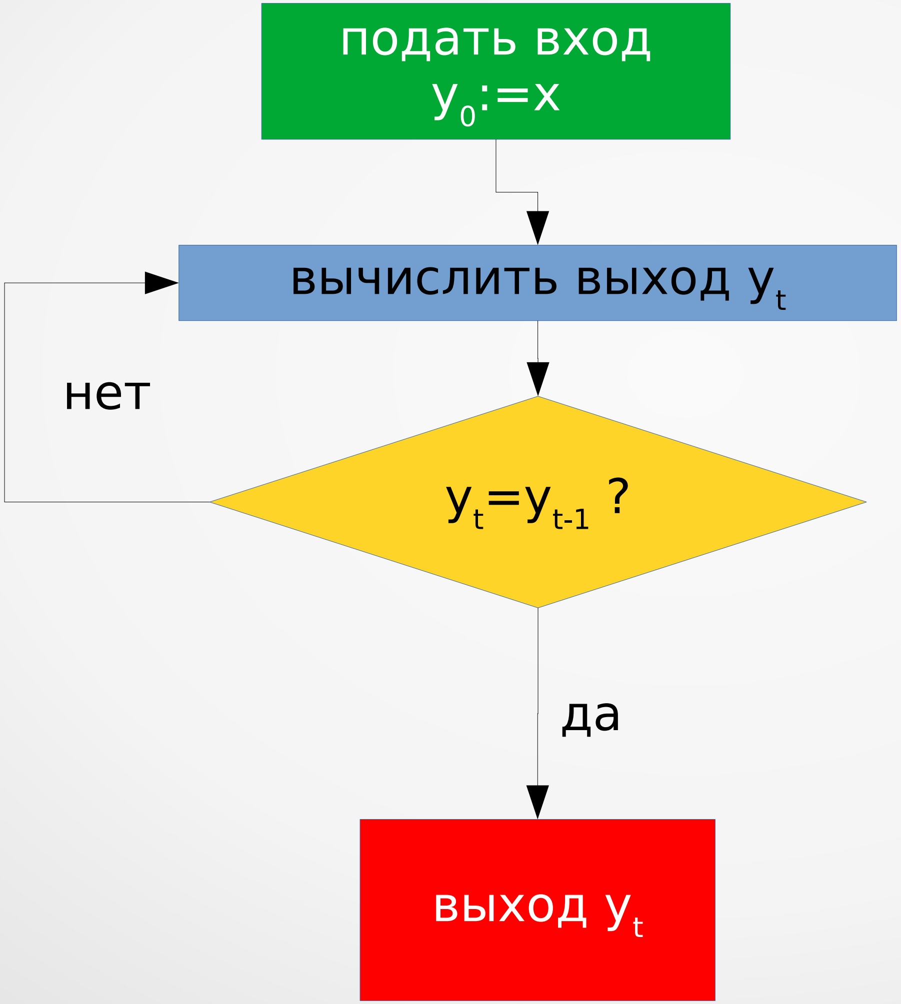
обратные связи

обратная связь нейрона на себя отсутствует

матрица весов симметрична $\boldsymbol{w_{ij}=w_{ji}}$
и имеет нулевую главную диагональ $\boldsymbol{w_{ii}=0}$

Нейросети

схема работы сети Хопфилда



Нейросети

сеть Хопфилда

энергия сети $$E = -\sum_i s_i b_i - \sum_{i>j} s_i s_j w_{ij}$$

s_i - состояние нейрона i
b_i - сдвиг нейрона i
w_{ij} - вес связи i j

Нейросети

сеть Хопфилда

энергия сети $$E=-\sum_i s_i b_i-\sum_{i>j} s_i s_j w_{ij}$$

s_i - состояние нейрона i
b_i - сдвиг нейрона i
w_{ij} - вес связи i j

в процессе переходов состояний сети энергия уменьшается и достигает локального минимума - **аттрактора**

Нейросети

сеть Хопфилда

энергия сети $$E=-\sum_i s_i b_i-\sum_{i>j} s_i s_j w_{ij}$$

s_i - состояние нейрона i
b_i - сдвиг нейрона i
w_{ij} - вес связи i j

в процессе переходов состояний сети энергия уменьшается и достигает локального минимума - **аттрактора**

в случае ассоциативной памяти:
  аттракторами являются хранимые сетью образцы

Нейросети

сеть Хопфилда

энергия сети $$E = -\sum_i s_i b_i - \sum_{i>j} s_i s_j w_{ij}$$

s_i - состояние нейрона i
b_i - сдвиг нейрона i
w_{ij} - вес связи i j

в процессе переходов состояний сети энергия уменьшается и достигает локального минимума - **аттрактора**

в случае ассоциативной памяти:
аттракторами являются хранимые сетью образцы

ёмкость памяти ~0.15 * количество нейронов сети

ложные аттракторы

Нейросети

сеть Хопфилда

обучение - метод Хэбба

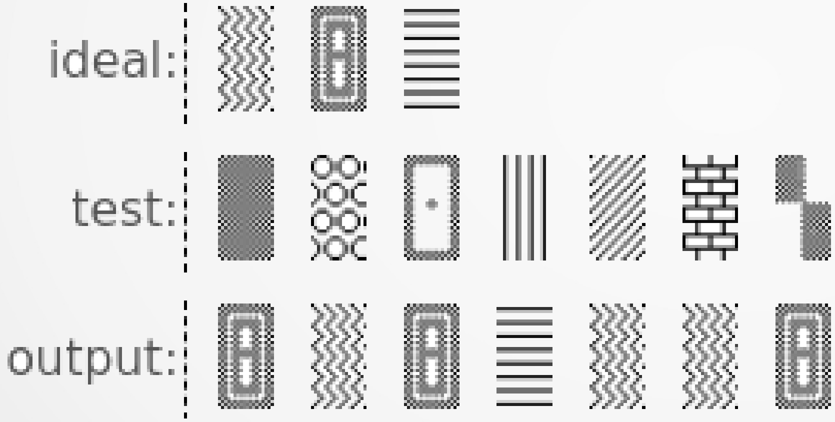
$$W = X^T \cdot X$$

обнуляем главную диагональ матрицы весов
т. е. убираем связь нейрона на себя

Нейросети

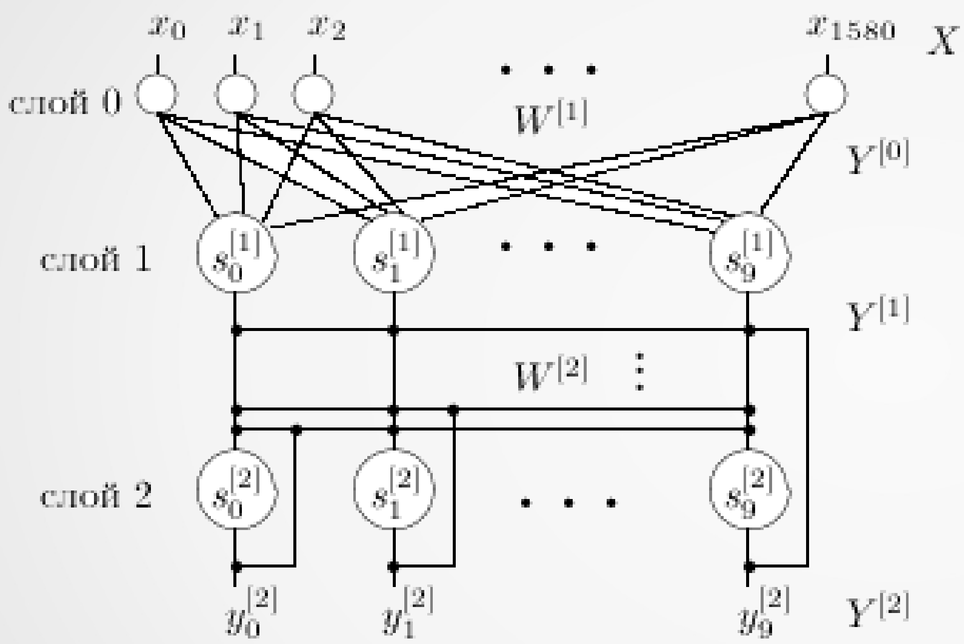
сеть Хопфилда

пример результатов работы



Нейросети

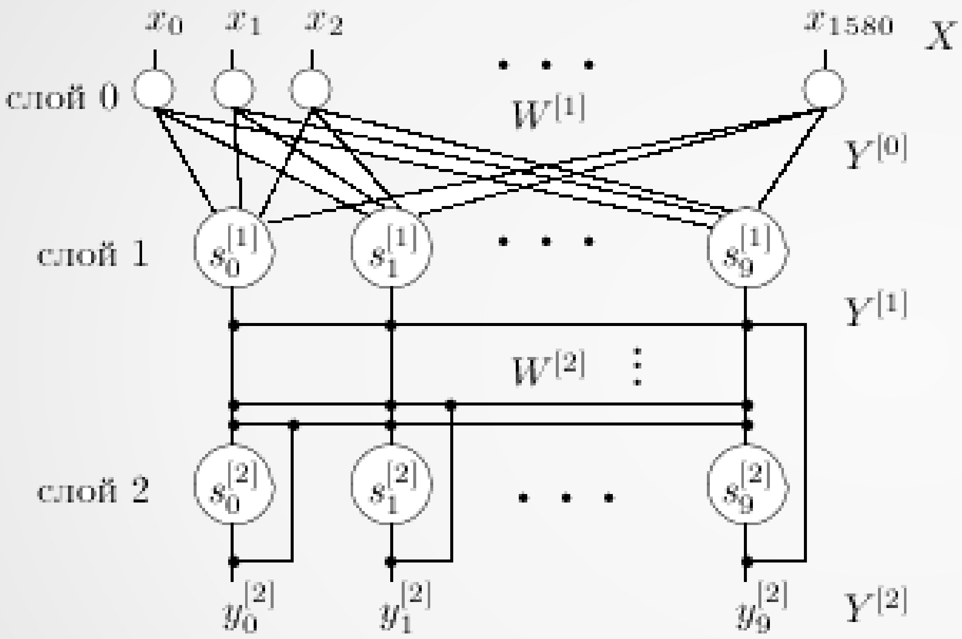
сеть Хэминга **(Lipman R. 1987)**



Вход сети — вектор из {-1,+1}

Нейросети

сеть Хэминга   (Lipman R. 1987)



Вход сети — вектор из {-1,+1}

Первый слой - имнеет линейную активацию,

вычисляет расстояние Хемминга между всеми эталонными образцами, хранящимися в сети, и текущим входом сети.

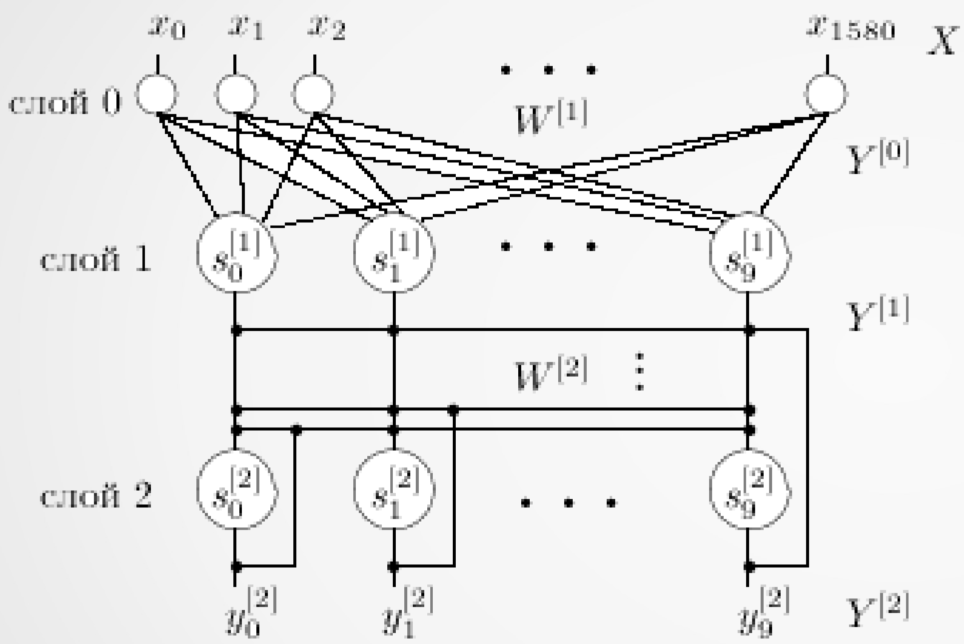
Расстояние Хемминга - количество отличающихся компонент во входном и эталонном векторах.

Веса первого слоя: $W^{[1]} = \frac{1}{2} X^T$

где X — матрица из векторов-образов, которые записываем в память

Нейросети

сеть Хэминга (Lipman R. 1987)



Вход сети — вектор из {-1,+1}

Первый слой - имнеет линейную активацию,

вычисляет расстояние Хемминга между всеми эталонными образцами, хранящимися в сети, и текущим входом сети.

Расстояние Хемминга - количество отличающихся компонент во входном и эталонном векторах.

Веса первого слоя: $W^{[1]} = \frac{1}{2} X^T$

где X — матрица из векторов-образов, которые записываем в память

Второй слой - сеть Хопфилда, используется для разрешения конфликтов, когда входной вектор оказывается похож по Хеммингу на более чем один идеал.

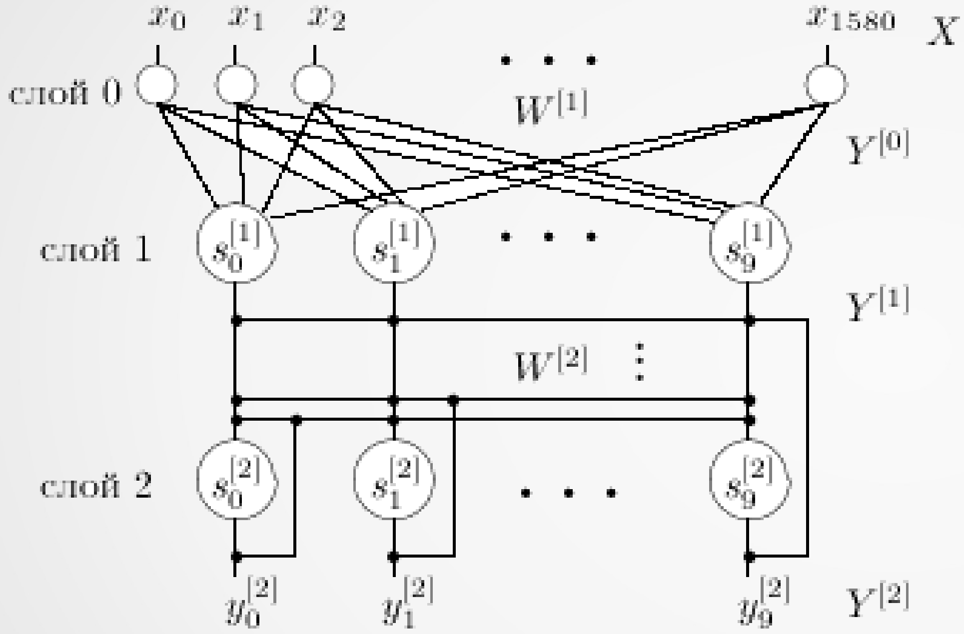
Веса второго слоя фиксированны

$$w_{ij}^{[2]} = \begin{cases} 1 & , \quad i = j \\ -c & , \quad i \neq j \end{cases}$$

где c = 1/(2m) - коэффициент торможения,
m — количество образов

Нейросети

сеть Хэминга   (Lipman R. 1987)



Вход сети — вектор из {-1,+1}

Первый слой - имнеет линейную активацию,

вычисляет расстояние Хемминга между всеми эталонными образцами, хранящимися в сети, и текущим входом сети.

Расстояние Хемминга - количество отличающихся компонент во входном и эталонном векторах.

Веса первого слоя: $W^{[1]} = \frac{1}{2} X^T$

где X — матрица из векторов-образов, которые записываем в память

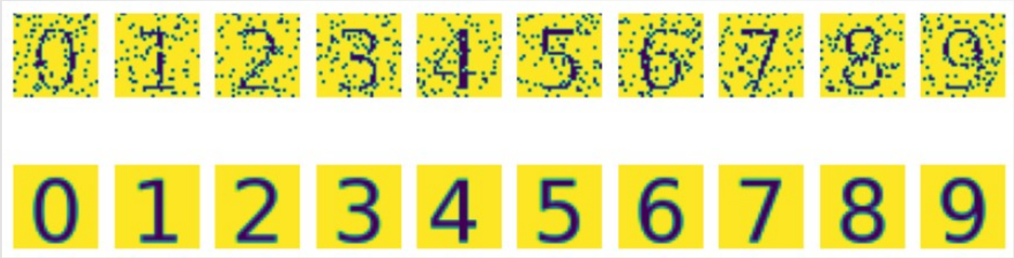
Второй слой - сеть Хопфилда, используется для разрешения конфликтов, когда входной вектор оказывается похож по Хеммингу на более чем один идеал.

Веса второго слоя фиксированны

$$w_{ij}^{[2]} = \begin{cases} 1 & , \; i = j \\ -c & , \; i \neq j \end{cases}$$

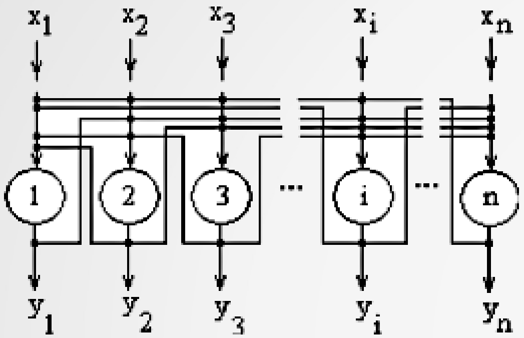
где c = 1/(2m) - коэффициент торможения,
m — количество образов

Пример работы сети Хэминга как фильтра шумов

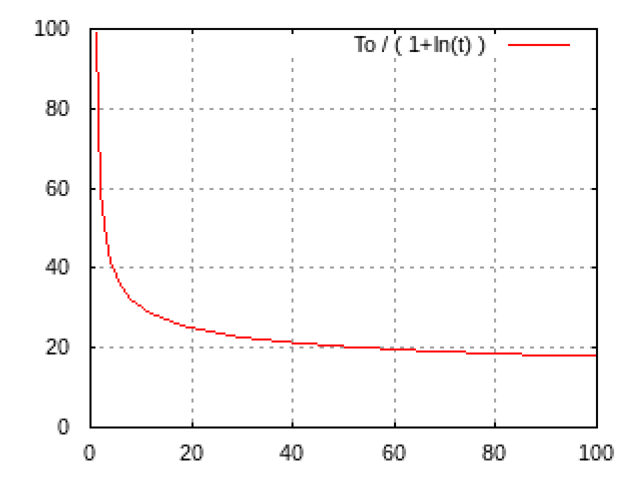


Нейросети

машина Больцмана (G.Hinton, 1985)



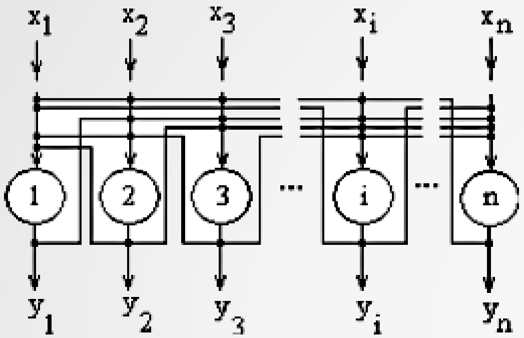
моделирует процесс отжига металла - со снижением температуры вероятность измененения состояния уменьшается



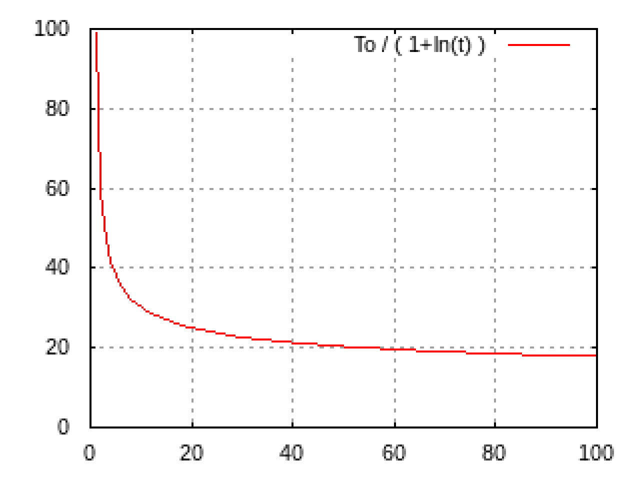
температура сети $T = \dfrac{T_0}{1 + \ln(t)}$

Нейросети

машина Больцмана (G.Hinton, 1985)



моделирует процесс отжига металла - со снижением температуры вероятность изменениия состояния уменьшается



Вероятность перехода нейрона i в новое состояние

$$P_i = 1/\left(1 + \exp\left(-\frac{\Delta E_i}{T}\right)\right)$$

изменение энергии

$$\Delta E_i(t) = E_i(t) - E_i(t-1)$$

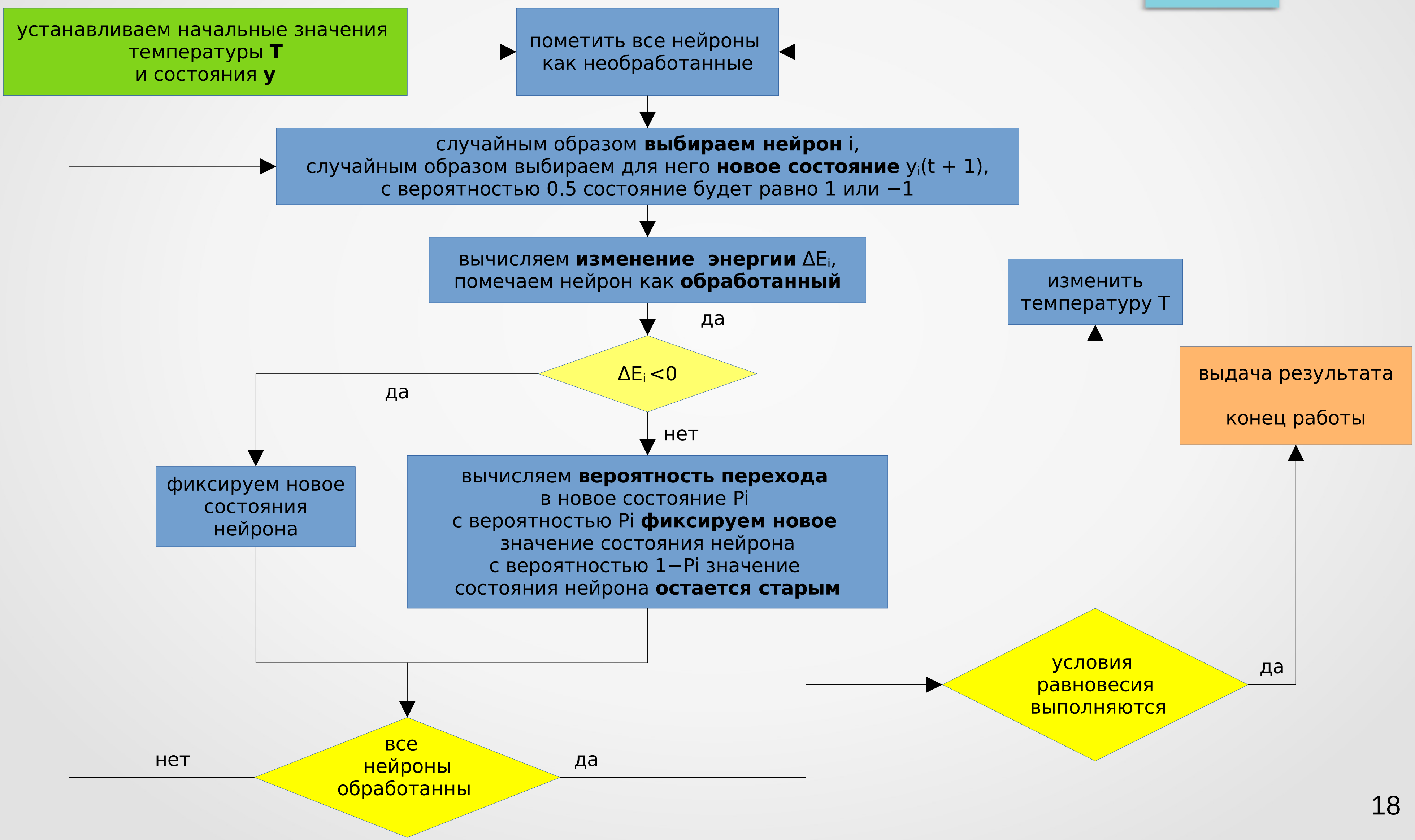
температура сети $T = \dfrac{T_0}{1 + \ln(t)}$

энергия сети на итерации t

$$E_i(t) = -\frac{1}{2} \cdot \left(\sum_j w_{ij} \cdot y_j(t) \cdot y_i(t)\right) - y_i(t) \cdot y_i(t-1)$$

Нейросети

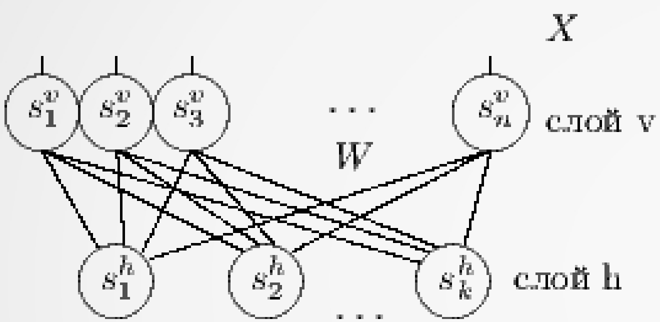
машина Больцмана (G.Hinton, 1985)



Нейросети

ограниченная машина Больцмана (RBM)
(P.Smolensky, 1986, G.Hinton, 2006)

RBM - модификация машины Больцмана
нейроны были разделены на две группы,
убраны некоторые связи,
таким образом был образован второй (скрытый) слой.
http://mechanoid.su/neural-net-boltzman-restr.html



$$L(W, b_v, b_h | v) = p(v | W, b_v, b_h)$$

$$\ln L(W, b_v, b_h | v) = \ln \sum_{h} \exp(-E(v,h)) - \ln \sum_{v,h} \exp(-E(v,h))$$

$$E(v,h) = -(b_v \cdot v + b_h \cdot h + v \cdot h \cdot W)$$

$$p(h = 1 | v, W, b_h) = \sigma(W \cdot v + b_h)$$

$$\sigma(x) = \frac{1}{1 + \exp(-x)}$$

$$\begin{cases} \frac{\partial \ln L(W,b_v,b_h|v)}{\partial W} = \nabla W & = & (v \cdot h)_{data} - (v \cdot h)_{model} \\ \frac{\partial \ln L(W,b_v,b_h|v)}{\partial b_v} = \nabla b_v & = & (v)_{data} - (v)_{model} \\ \frac{\partial \ln L(W,b_v,b_h|v)}{\partial b_h} = \nabla b_h & = & (h)_{data} - (h)_{model} \end{cases}$$

$(\cdot)_{data}$ - значение слоёв в начальном состоянии сети,
$(\cdot)_{model}$ - мат.ожидание состояния слоёв

Нейросети: литература

Борисов Е.С. Методы машинного обучения. 2024
https://github.com/mechanoid5/ml_lectorium_2024_I

- Е.С.Борисов Ассоциативная память на основе нейронной сети Хопфилда. http://mechanoid.su/neural-net-hopfield-associative-memory.html
- Е.С.Борисов Классификатор на основе нейронной сети Хемминга. http://mechanoid.su/neural-net-hamming-classifier.html
- Е.С.Борисов Нечеткий поиск на основе нейронной сети Хемминга. http://mechanoid.su/neural-net-hamming-fuzzy-search.html
- Ассоциативная память на основе машины Больцмана. http://mechanoid.su/neural-net-boltzman.html
- Е.С.Борисов Ассоциативная память на основе ограниченной машины Больцмана (RBM). http://mechanoid.su/neural-net-boltzman-restr.html
- Саймон Хайкин   Нейронные сети. Полный курс : пер. с англ. -- Москва:Вильямс, 2006.   (глава 14 нейродинамика)
- Осовский С. Нейронные сети для обработки информации. — М.: Финансы и статистика, 2002. (глава 7 рекурентные сети как ассоциативные запоминающие устройства)