

# Основы машинного обучения

Поляк Марк Дмитриевич

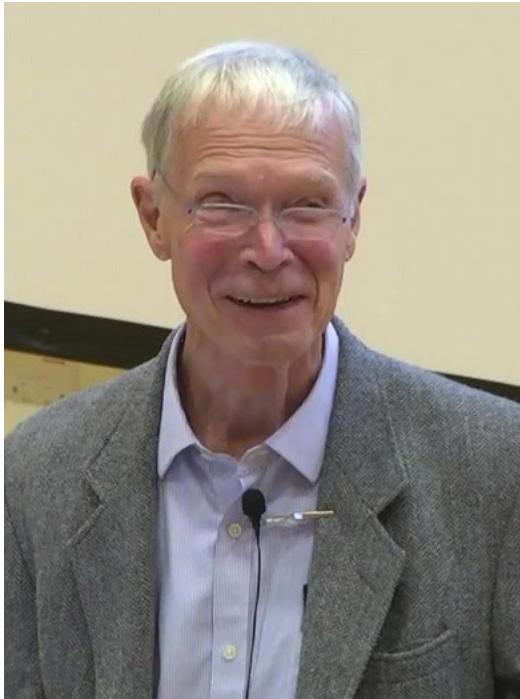
2025

# За что дают Нобелевскую премию

Дополнительная лекция

# Лауреаты Нобелевской премии по физике 2024 года

**Джон Хопфилд (р. 1933)**

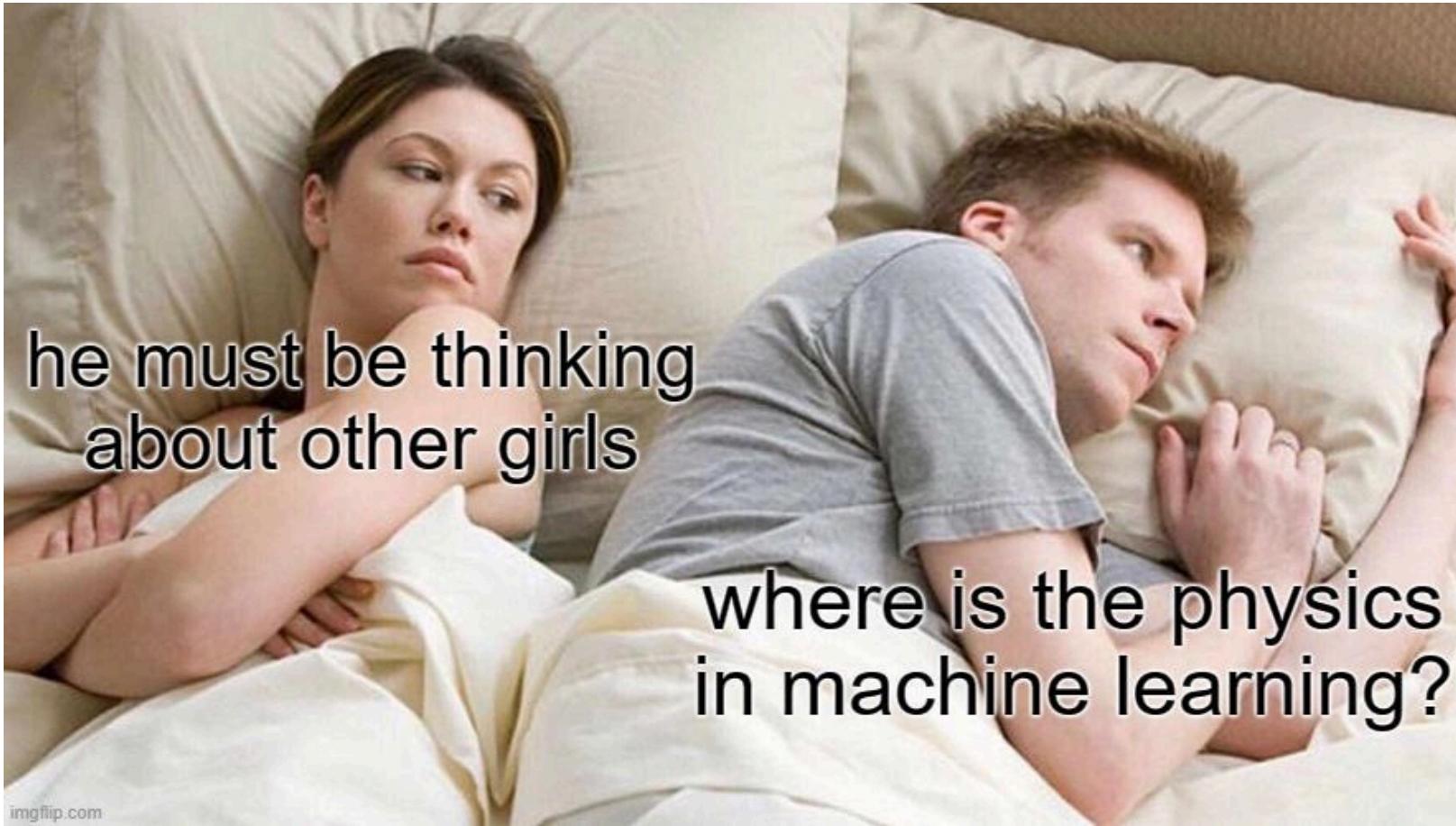


**Джеффри Хинтон (р. 1947)**



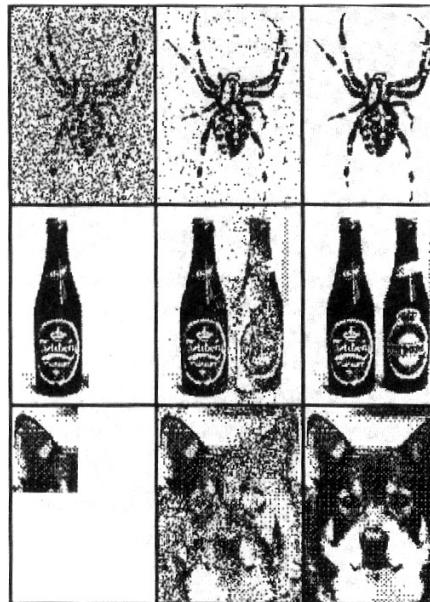
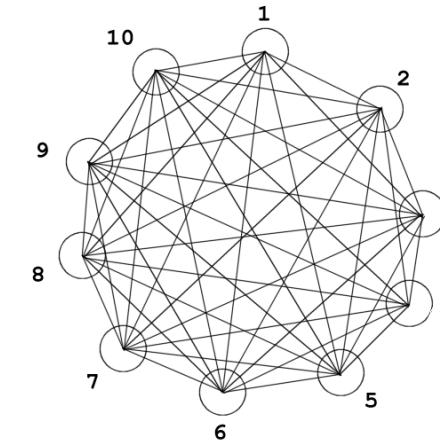
... за фундаментальные открытия и изобретения, которые позволяют применять машинное обучение с использованием искусственных нейронных сетей

# Основной вопрос

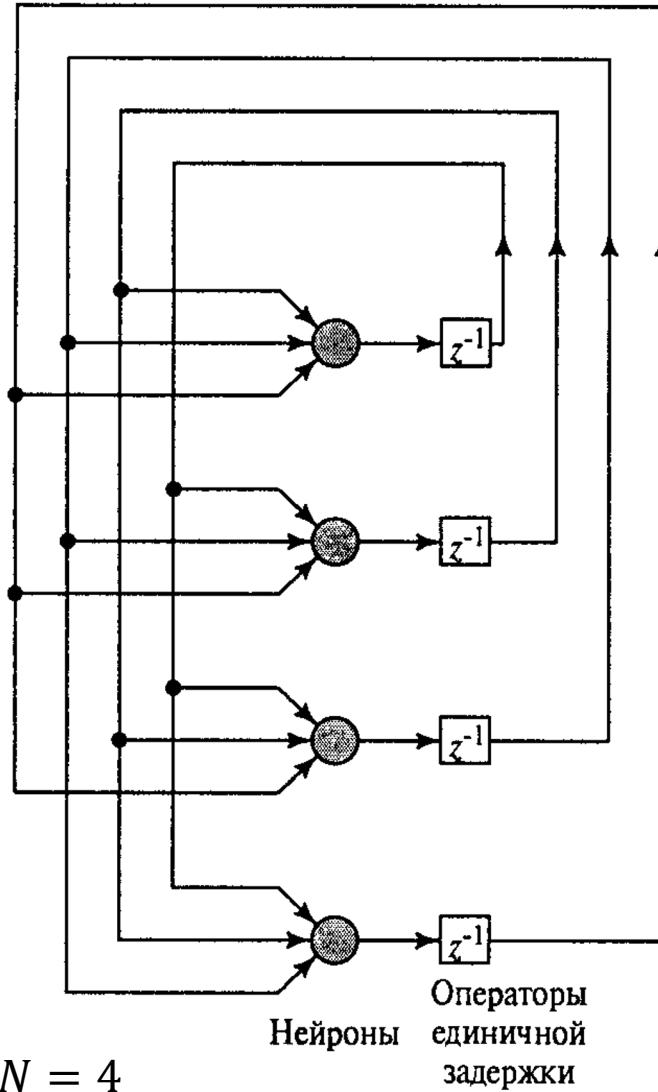


# Сеть Хопфилда

- Один из видов рекуррентных сетей
    - Полностью рекуррентный
    - Веса симметричны
    - Узлы могут быть только в состояниях *вкл* или *выкл*
    - Случайный порядок перехода от одного узла к другому
  - Обучение: **Правило Хебба** (cells that fire together wire together)
  - Может вспомнить и восстановить образ по его искаженной или неполной версии
- ассоциативная память



# Структура сети Хопфилда



- Состояние сети (модели) из  $N$  нейронов:

$$\mathbf{y} = (y_1, y_2, \dots, y_N)^T$$

- Дискретная сеть Хопфилда:

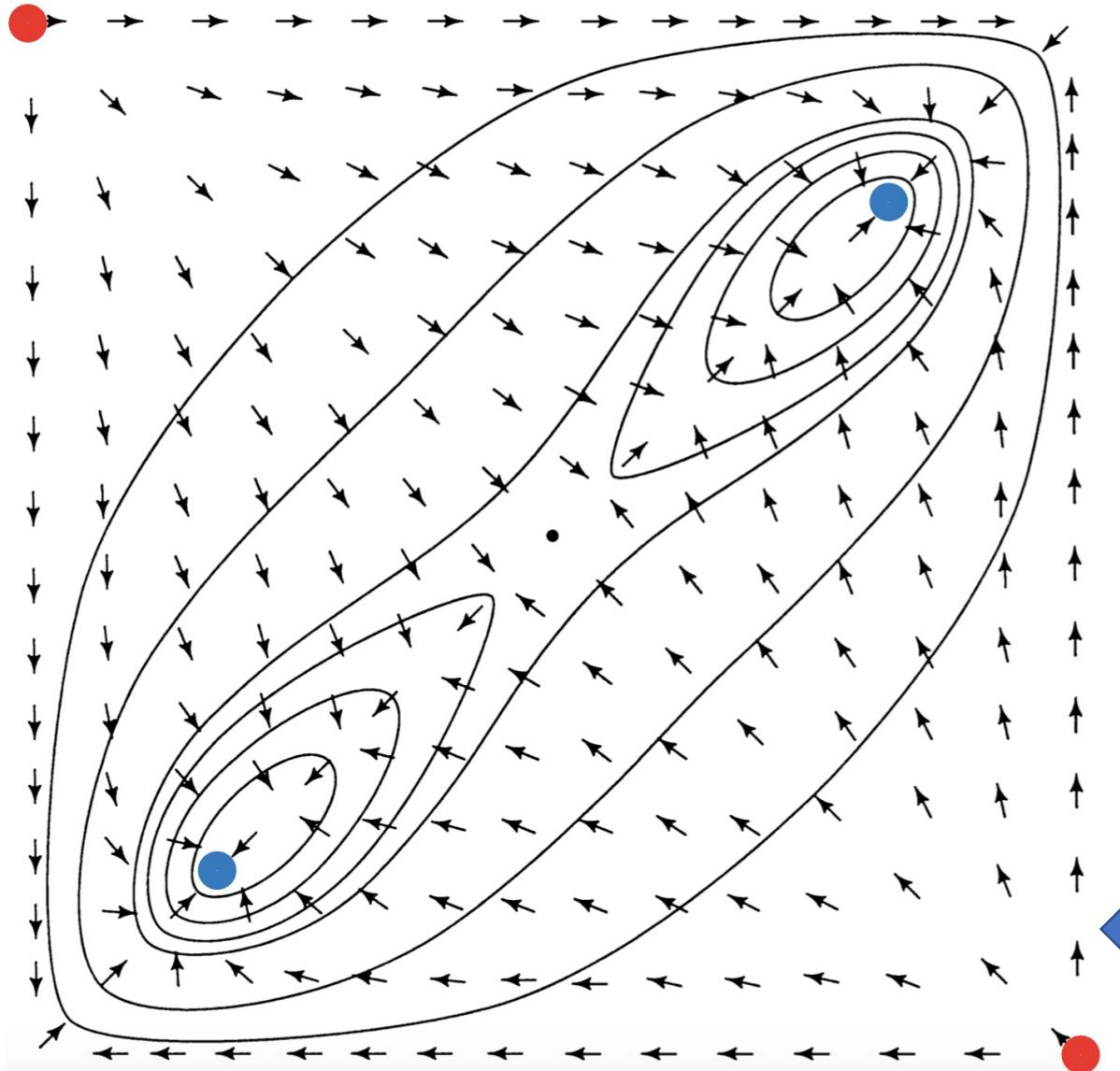
$$y_j = \text{sign } u_j = \begin{cases} +1, & u_j > 0 \\ -1, & u_j < 0 \end{cases}$$

$$u_j = \sum_{i=1}^N w_{ji} y_i + b_j$$

где  $w_{ji} = \theta_{ji}$  – параметры модели Хопфилда,  $w_{ii} = 0$ .

- Непрерывная сеть Хопфилда: система ДУ 1-го порядка, функция активации – сигмоида

# Энергия сети Хопфилда



- Функция энергии (Ляпунова) дискретной сети:

$$E = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1, j \neq i}^N w_{ji} y_i y_j$$

- Динамика сети Хопфилда описывается механизмом, обеспечивающим нахождение минимумов  $E$ .

Контурная карта энергии непрерывной модели Хопфилда с двумя нейронами и двумя устойчивыми состояниями. Ординаты и абсциссы являются выходами двух нейронов. Устойчивые состояния соответствуют минимумам энергии и показаны синим, неустойчивые экстремумы – красным.

# Обучение сети Хопфилда

- Правило Хебба (ассоциативное обучение):  
одновременная активация клеток приводит к выраженному  
увеличению синаптической силы между этими клетками
- Инициализация весов в дискретной модели Хопфилда:

$$\theta_{ji} = w_{ji} = \frac{1}{N} \sum_{k=1}^{\ell} x_{k,j} x_{k,i}$$

где  $x_{k,j}$  –  $j$ -ый элемент бинарного вектора  $\mathbf{x}_k$  из обучающей выборки

- Пример:  $N = 4; \ell = 2; \mathbf{x}_1 = (-1, +1, -1, -1); \mathbf{x}_2 = (+1, -1, +1, +1)$   
 $w_{12} = w_{21} = \frac{1}{4}(-1 \times 1 + 1 \times (-1)) = -0.5$        $w_{14} = w_{41} = 0.5$   
 $w_{13} = w_{31} = 0.25 \times (-1 \times (-1) + 1 \times 1) = 0.5$        $w_{24} = w_{42} = -0.5$   
 $w_{23} = w_{32} = 0.25 \times (1 \times (-1) + (-1) \times 1) = -0.5$        $w_{34} = w_{43} = 0.5$

# Векторизация

- Инициализация весов (фаза сохранения информации в ассоциативной памяти):

$$\mathbf{W} = \frac{1}{N} \left( \sum_{k=1}^{\ell} \mathbf{x}_k \mathbf{x}_k^T - \ell \mathbf{E} \right)$$

где  $\mathbf{E}$  – единичная матрица,  $\mathbf{x}_k$  – вектор для сохранения в памяти сети.

- Пример:

$$\mathbf{W} = \begin{pmatrix} 0 & w_{12} & w_{13} & w_{14} \\ w_{21} & 0 & w_{23} & w_{24} \\ w_{31} & w_{32} & 0 & w_{34} \\ w_{41} & w_{42} & w_{43} & 0 \end{pmatrix} = \begin{pmatrix} 0 & -0.5 & 0.5 & 0.5 \\ -0.5 & 0 & -0.5 & -0.5 \\ 0.5 & -0.5 & 0 & 0.5 \\ 0.5 & -0.5 & 0.5 & 0 \end{pmatrix}$$

- Вывод с помощью сети (фаза извлечения информации):

- Задать начальное состояние сети (образец для поиска в памяти)

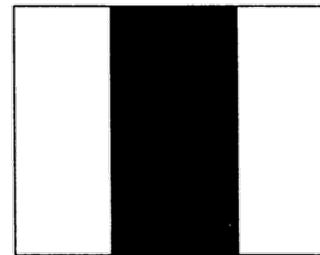
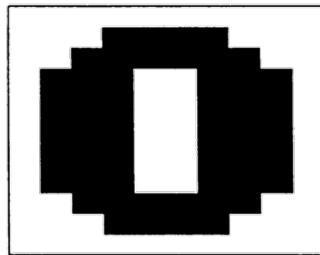
$$\mathbf{y}_0 = (y_{0,1}, y_{0,2}, \dots, y_{0,N})^T$$

- Обновлять состояние сети до сходимости:

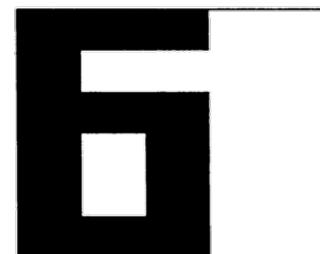
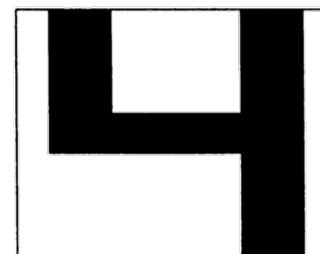
$$\mathbf{y} = \text{sign}(\mathbf{W}\mathbf{y} + \mathbf{b})$$

# Пример моделирования работы сети Хопфилда

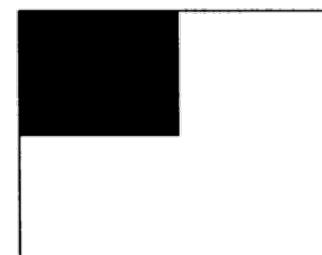
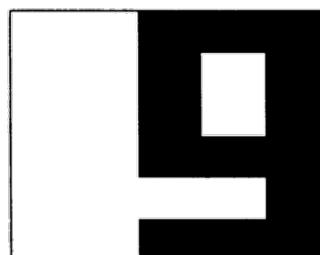
Исходные образы, использовавшиеся при обучении сети\*:



$\ell = 8$  изображений



$N = 120 = 10 \times 12$  пикселей

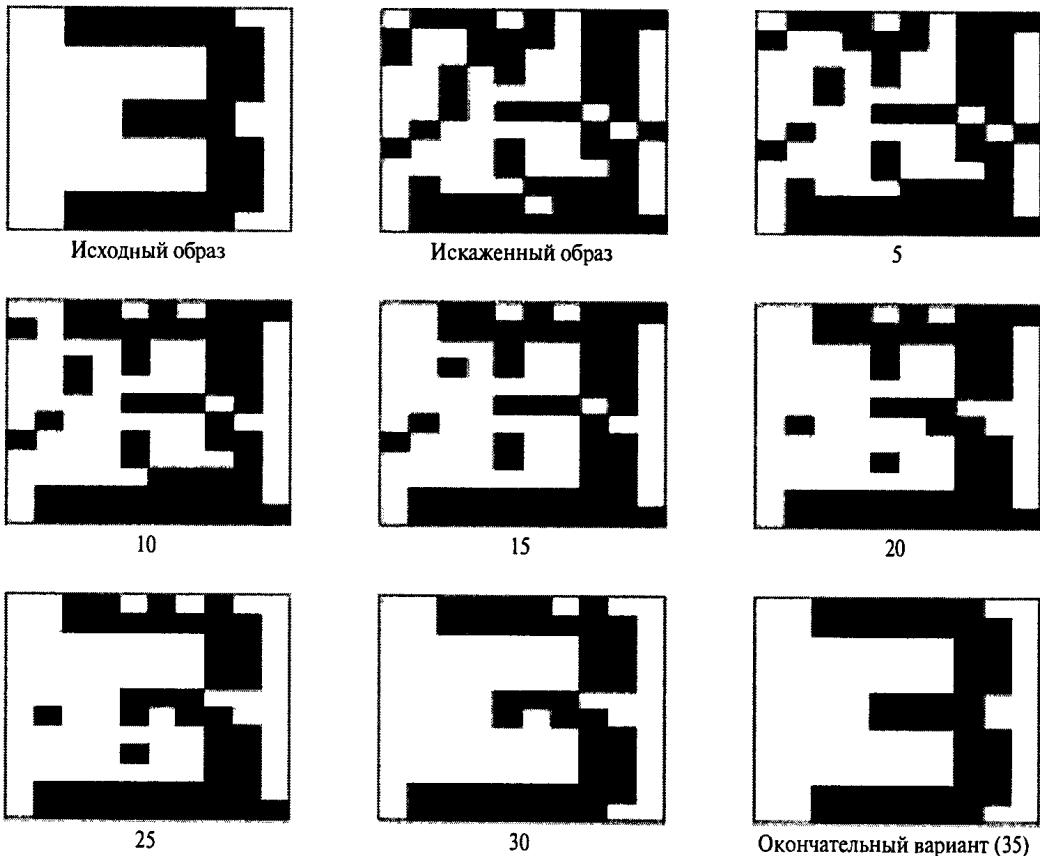


$N^2 - N = 12280$

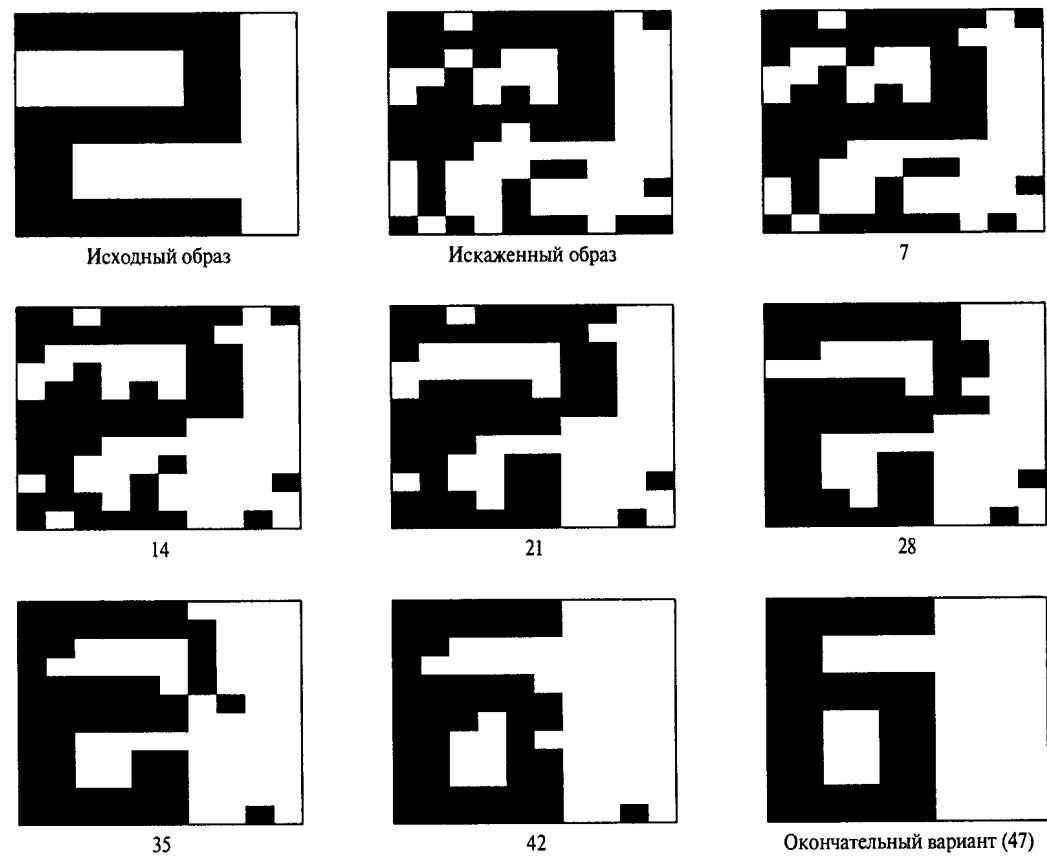
синаптических весов

# Пример моделирования работы сети Хопфилда

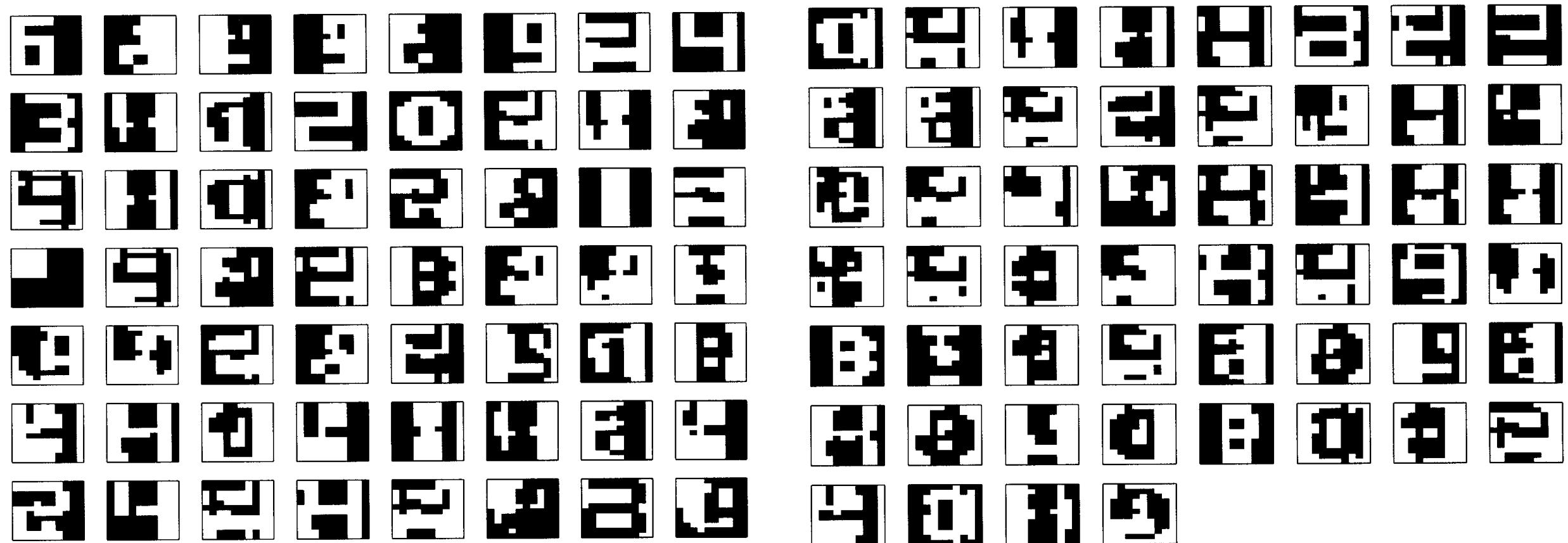
**Корректировка искаженного образа  
цифры 3**



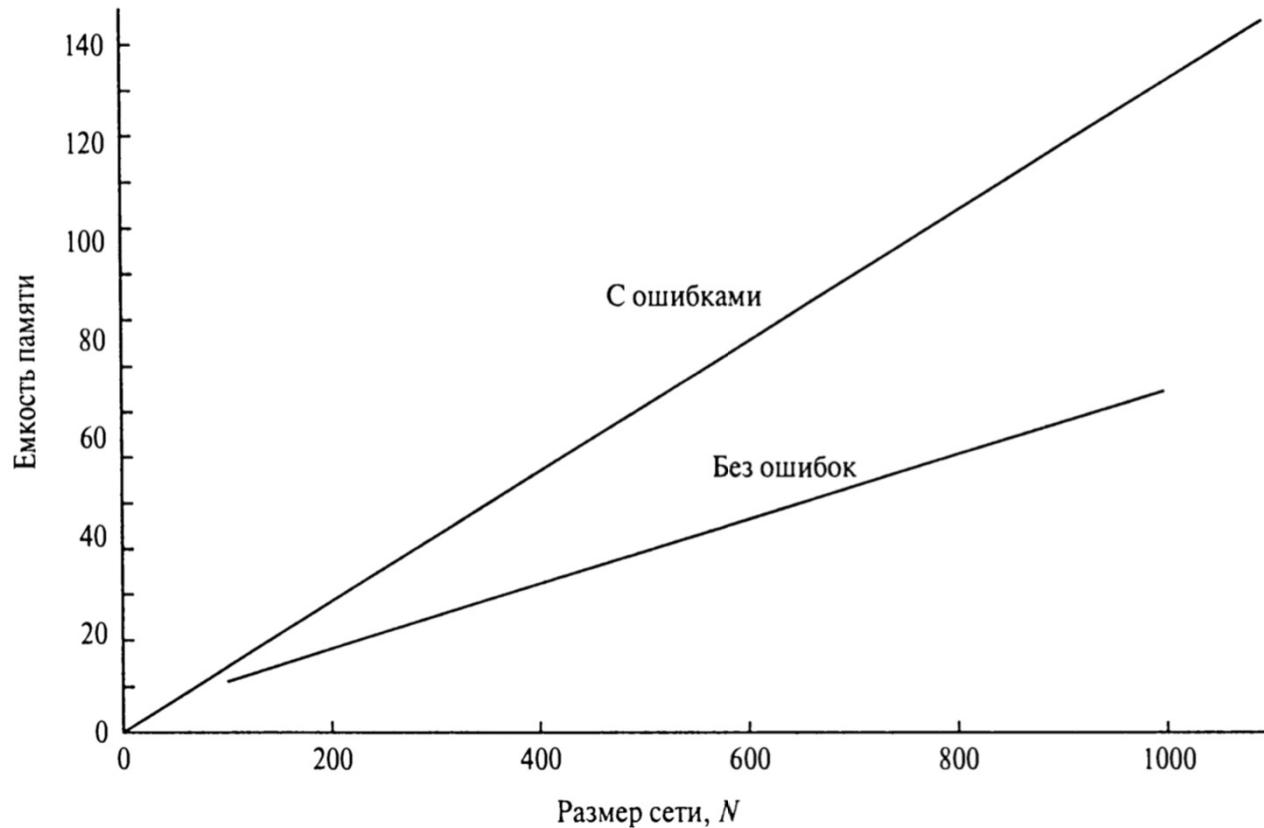
**Некорректное восстановление  
цифры 2**



# Ложные состояния сети Хопфилда



# Ёмкость памяти сети Хопфилда



- Восстановление образов из памяти невозможно, если:

$$\frac{\ell}{N} > 0.14$$

- Восстановить  $\ell$  образов с ошибками можно, если:

$$\ell = 0.14N$$

- Восстановить  $\ell$  образов почти без ошибок:

$$\ell = \frac{N}{2 \ln N}$$

# Распределение Больцмана

- Распределение Больцмана (реже – распределение Гиббса):

$$p_i \propto e^{-\frac{\varepsilon_i}{kT}}$$

где  $p_i$  – вероятность пребывания системы в определенном  $i$ -ом состоянии, символ  $\propto$  означает пропорциональность,  $\varepsilon_i$  – энергия системы,  $k = 1,380649 \cdot 10^{-23}$  – постоянная Больцмана,  $T$  – температура системы.

- Интерпретация р. Б. в статистической механике и термодинамике:

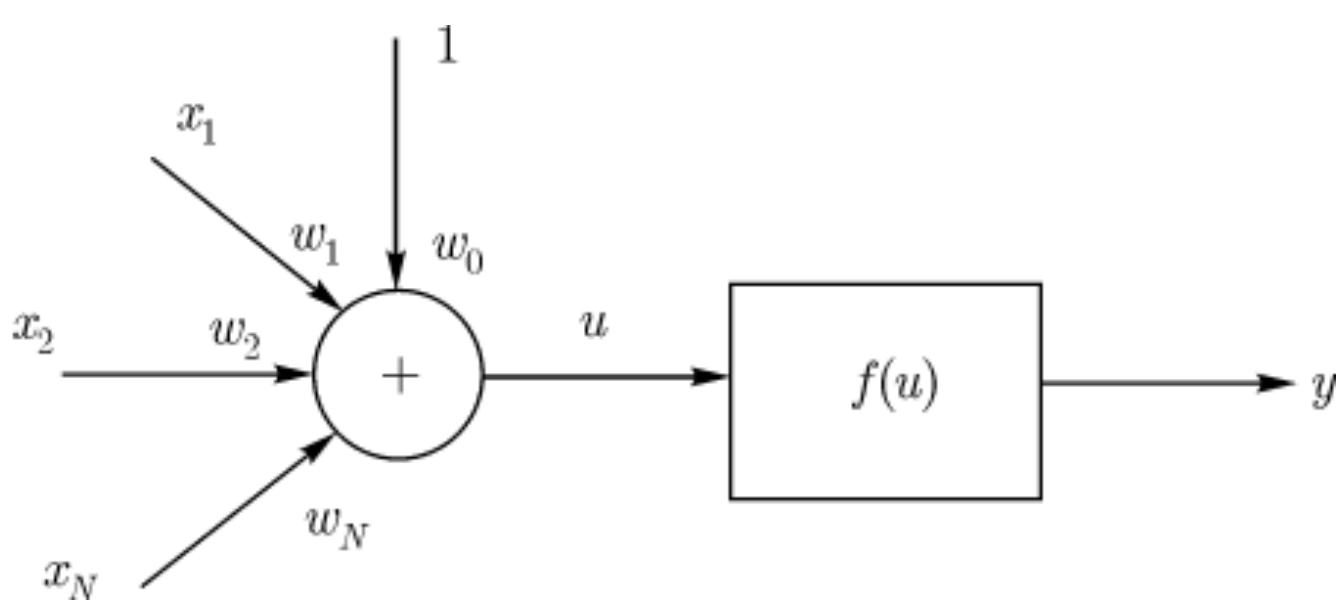
Концентрация молекул газа тем выше, чем меньше их потенциальная энергия.

Кроме этого, с понижением температуры увеличивается отличие концентраций в точках с различными значениями потенциальной энергии молекул. А при стремлении температуры к абсолютному нулю, молекулы начинают скапливаться в месте, где их потенциальная энергия принимает наименьшее значение.

# Классическая модель нейрона

Модель нейрона МакКаллока-Питтса:

$$y = f(u)$$



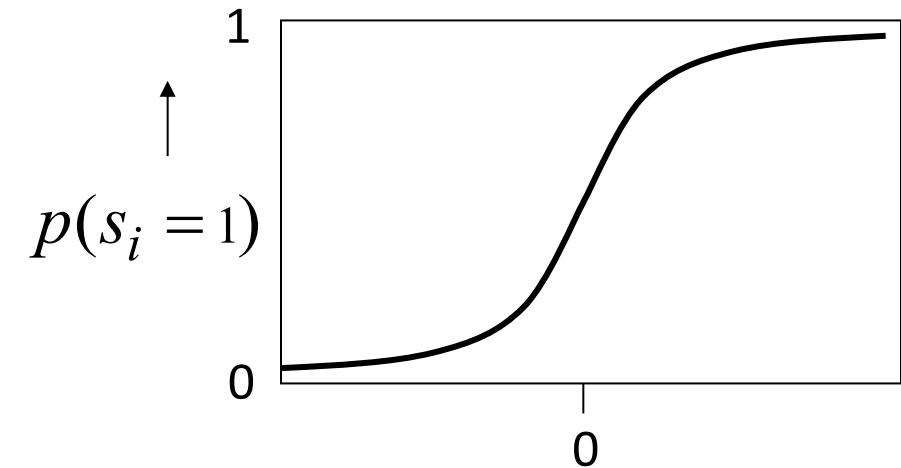
$$\begin{aligned} u &= \sum_{j=1}^N w_j x_j + b = \\ &= \sum_{j=0}^N w_j x_j, w_0 = b, x_0 = 1 \end{aligned}$$

где  $y$  – выход,  $x_j$  – входы,  
 $w_j$  – веса,  $b$  – смещение,  
 $f(u)$  – функция активации

# Стохастический двоичный нейрон

- Всегда находится в одном из состояний:  $s = 1$  или  $s = 0$  ( $s$  – выход нейрона)
- Вероятность активации  $i$ -го нейрона ( $s_i = 1$ ) определяется взвешенными входными сигналами от других нейронов плюс смещение:

$$p(s_i = 1) = \frac{1}{1 + \exp\left(-\frac{E_i}{T}\right)} =$$
$$= \frac{1}{1 + \exp\left(-T^{-1}(b_i + \sum_j w_{ji} s_j)\right)}$$



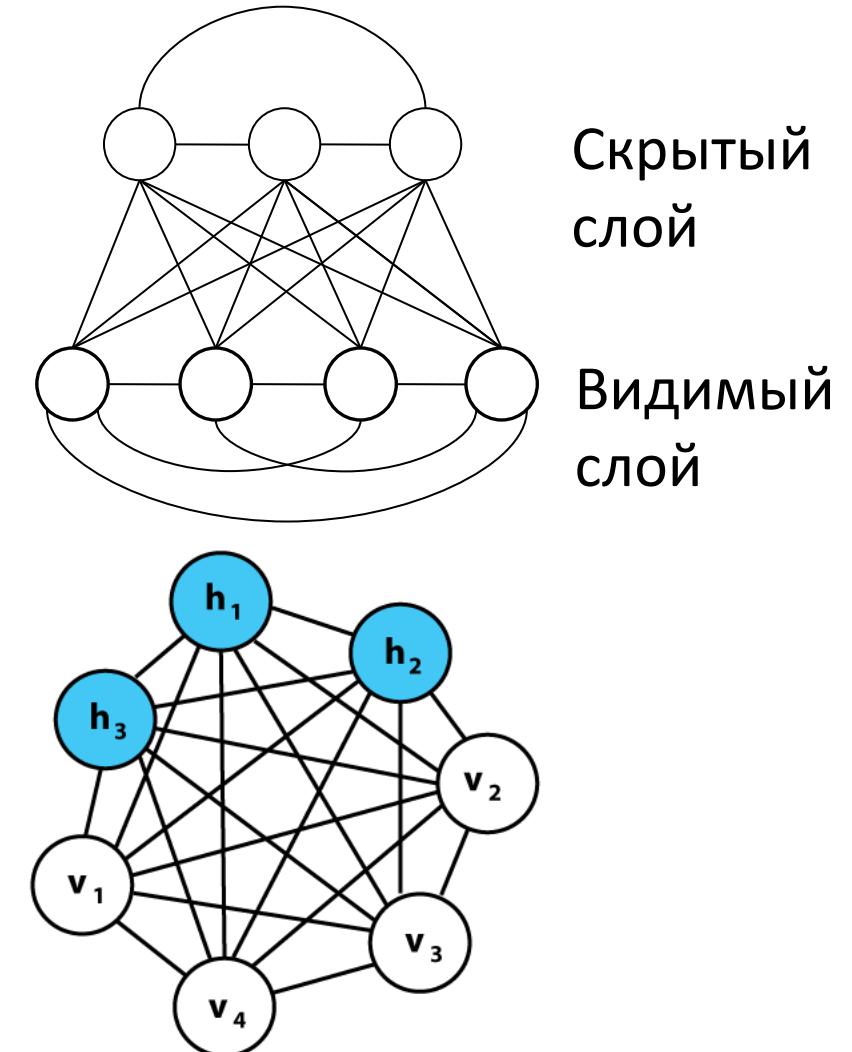
$$b_i + \sum_j s_j w_{ji} \rightarrow$$

$E_i$  – «энергия»  $i$ -го нейрона

$$E_i = \sum_{j=1}^N w_{ji} s_j + b_i; \quad E = - \sum_{i \neq j}^N E_i$$

# Машина Больцмана

- Полносвязный граф с двумя слоями: видимые  $v$  и скрытые  $h$  нейроны
- Все связи двунаправленные
- Узлы графа – стохастические двоичные нейроны
- Видимые нейроны отражают состояние МБ, доступное наблюдателю
- Скрытые нейроны образуют вектор признаков

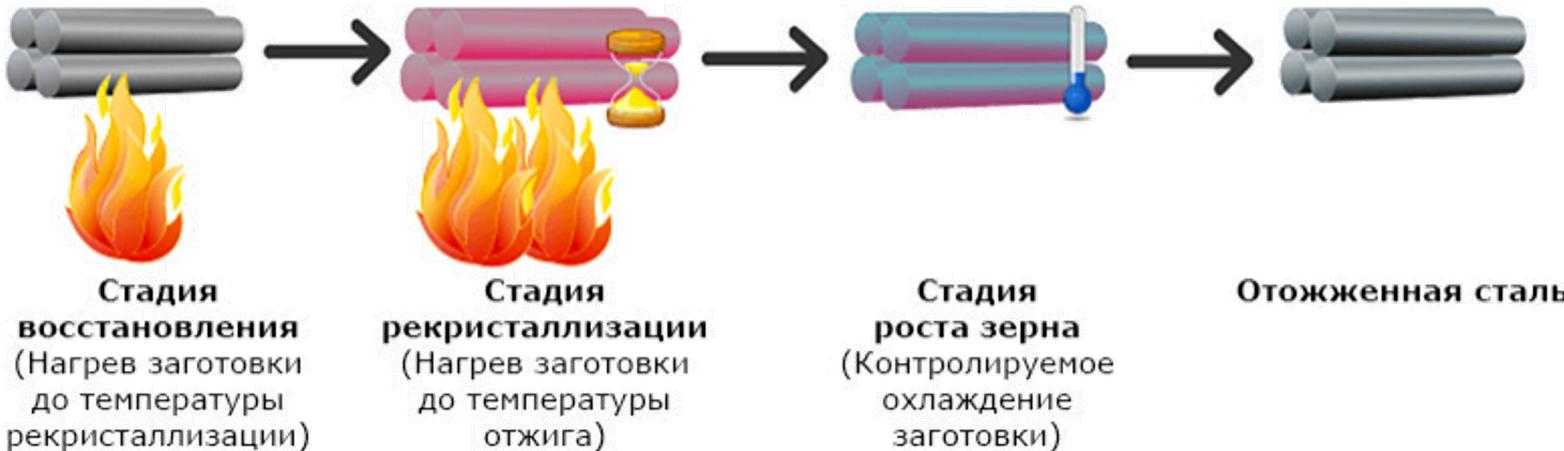


# Машина Больцмана. Продолжение

- Выборочная функция имеет распределение Больцмана, отсюда название
- Термодинамическое равновесие сети – среднее значение вектора состояния сети не меняется
- Обучение разбито на две фазы:
  - Положительная фаза (фаза «бодрствования»), используется правило Хебба
  - Отрицательная фаза (фаза «сна»), используется обратное правило Хебба
- Поиск термодинамического равновесия осуществляется алгоритмом имитации отжига

# Алгоритм имитации отжига

- Источник вдохновения  
– металлургия:

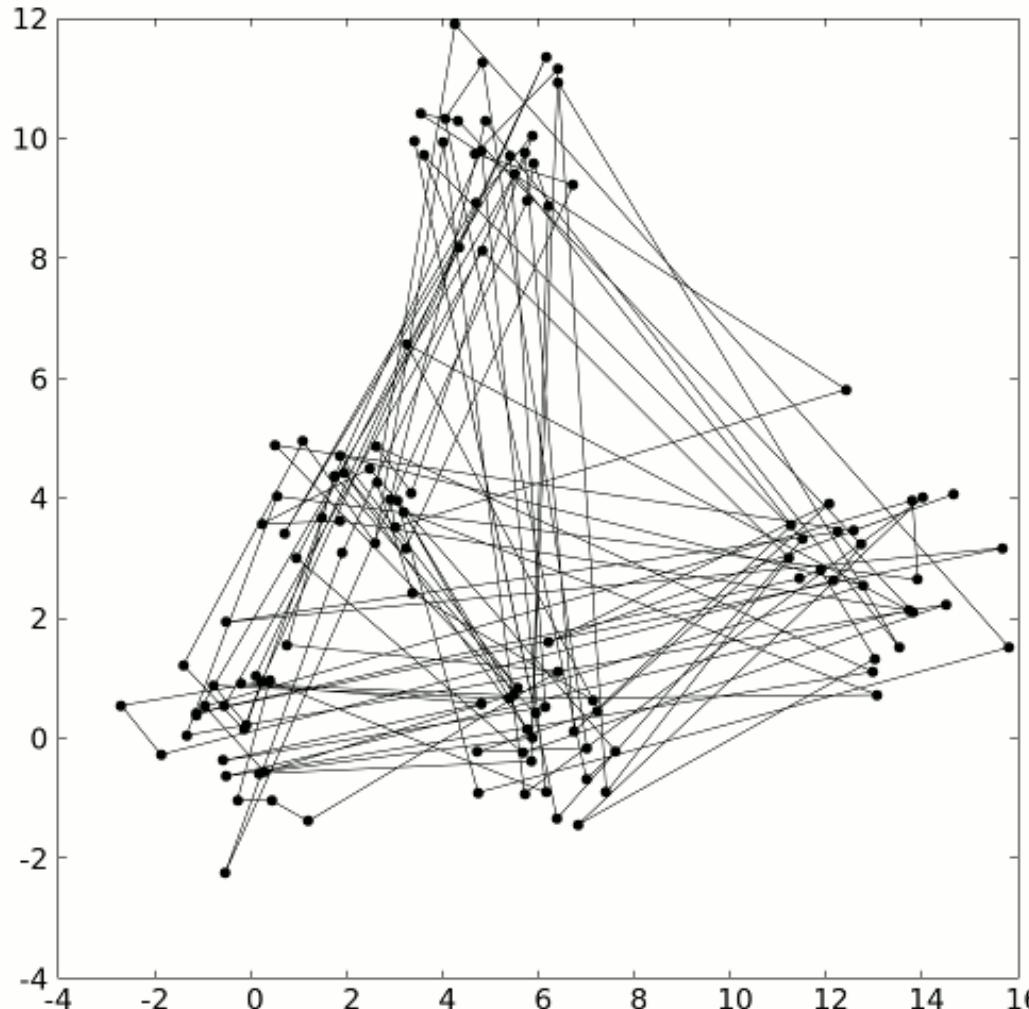


- Псевдокод:

1. `let  $s = s_0$  # задаём начальное состояние`
2. `for  $k = 0 .. k_{\max}$  (не включительно):`
  1.  `$T = \text{temperature}(1 - (k+1)/k_{\max})$  # уменьшаем температуру`
  2.  `$s_{\text{new}} = \text{neighbour}(s)$  # выбираем новое "соседнее" состояние`
  3. `If  $P(E(s), E(s_{\text{new}}), T) \geq \text{random}(0, 1)$ : # вычисление вероятности перехода в новое состояние и делаем бросок кубика, чтобы решить: переходим или нет?`
    - `$s = s_{\text{new}}$`
3. Возвращаем итоговое состояние  $s$

# Пример работы алгоритма имитации отжига: задача коммивояжера

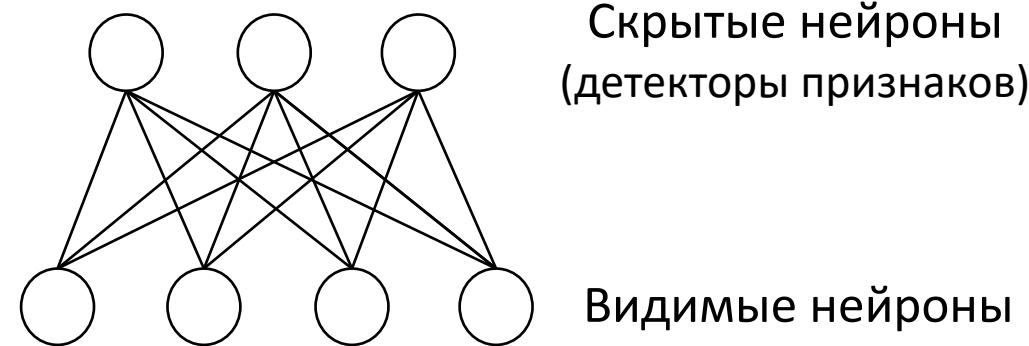
$E = 852 T = 125$



[https://commons.wikimedia.org/wiki/File:Travelling\\_salesman\\_problem\\_solved\\_with\\_simulated\\_annealing.gif#/media/File:Travelling\\_salesman\\_problem\\_solved\\_with\\_simulated\\_annealing.gif](https://commons.wikimedia.org/wiki/File:Travelling_salesman_problem_solved_with_simulated_annealing.gif#/media/File:Travelling_salesman_problem_solved_with_simulated_annealing.gif)

# Ограниченнная Машина Больцмана

- Restricted Boltzmann Machine (RBM)
  - Наложенные ограничения на связи между нейронами делают обучение сети проще
  - Можно легко получить выборку из апостериорного распределения скрытых признаков при заданном векторе входных значений



# Product of Experts

- Произведение экспертов – Product of Experts, PoE

Вероятность активации нейронов:

$$P(h_j = +1 | \mathbf{v}) = \frac{1}{1 + \exp\left(\sum_i w_{ij} v_i + a_j\right)}$$
 
$$P(v_i = +1 | \mathbf{h}) = \frac{1}{1 + \exp\left(\sum_j w_{ij} h_j + b_i\right)}$$

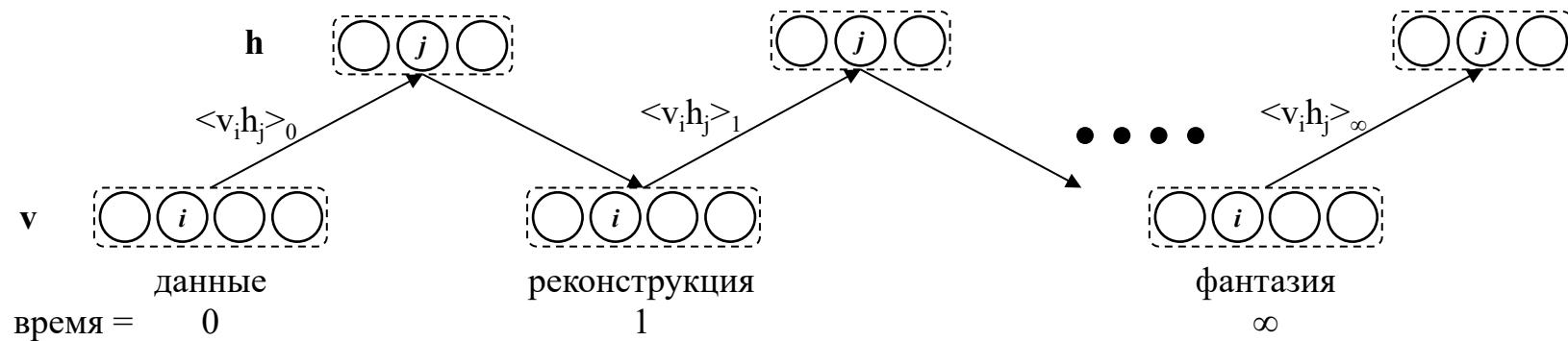
Условные распределения для видимых и скрытых нейронов:

$$p(\mathbf{v} | \mathbf{h}; W) = \prod_i P(v_i | \mathbf{h})$$

$$p(\mathbf{h} | \mathbf{v}; W) = \prod_j P(h_j | \mathbf{v})$$

# MCMC & Gibbs sampling

- RBM реконструирует данные с помощью метода Монте Карло для марковской цепи (Markov Chain Monte Carlo, MCMC) и квантизатора Гиббса (Gibbs sampling)



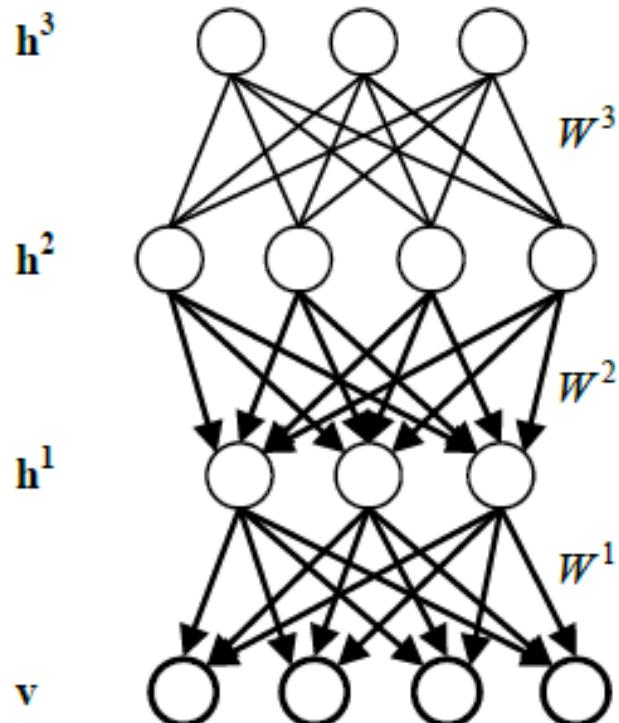
# Contrastive Divergence

- Обучение путем минимизации сравнительного отклонения  
(contrastive Divergence, CD)

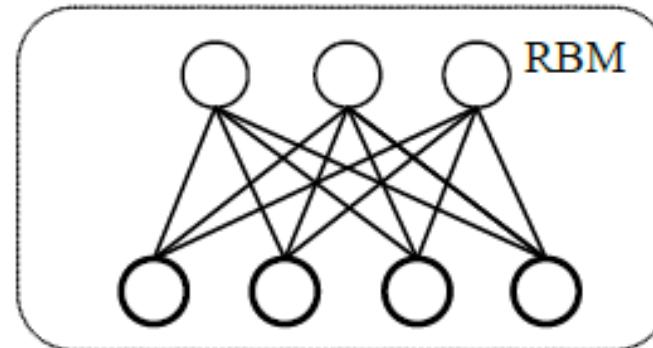
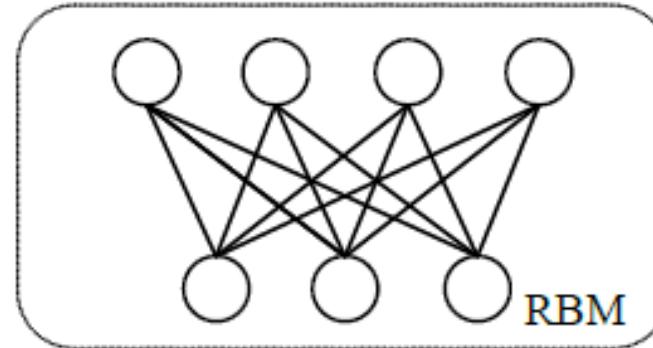
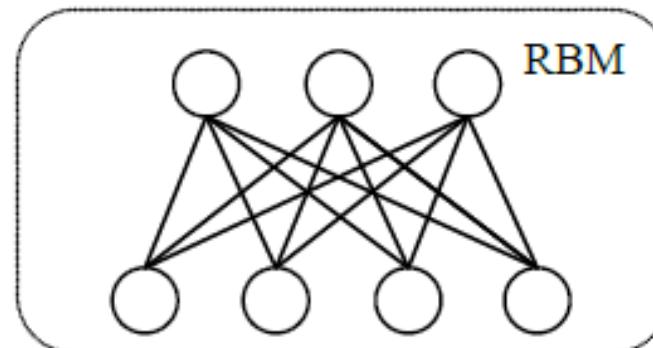
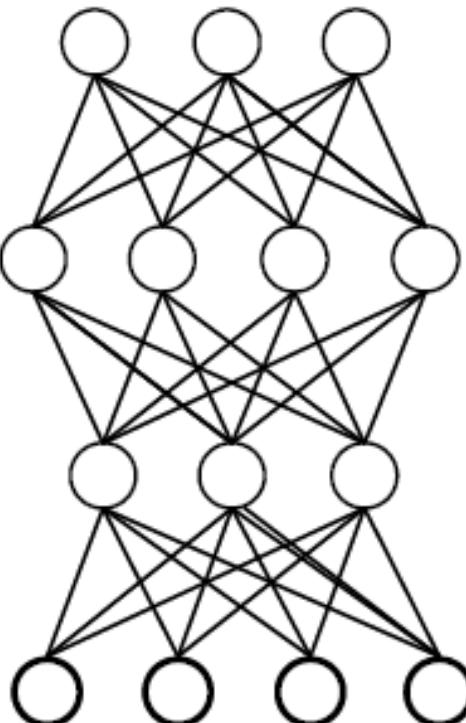
$$\Delta w_{ij} \approx \langle v_i h_j \rangle_0 - \langle v_i h_j \rangle_1$$

# Глубокие сети

**Deep Belief Network**



**Deep Boltzmann Machine**

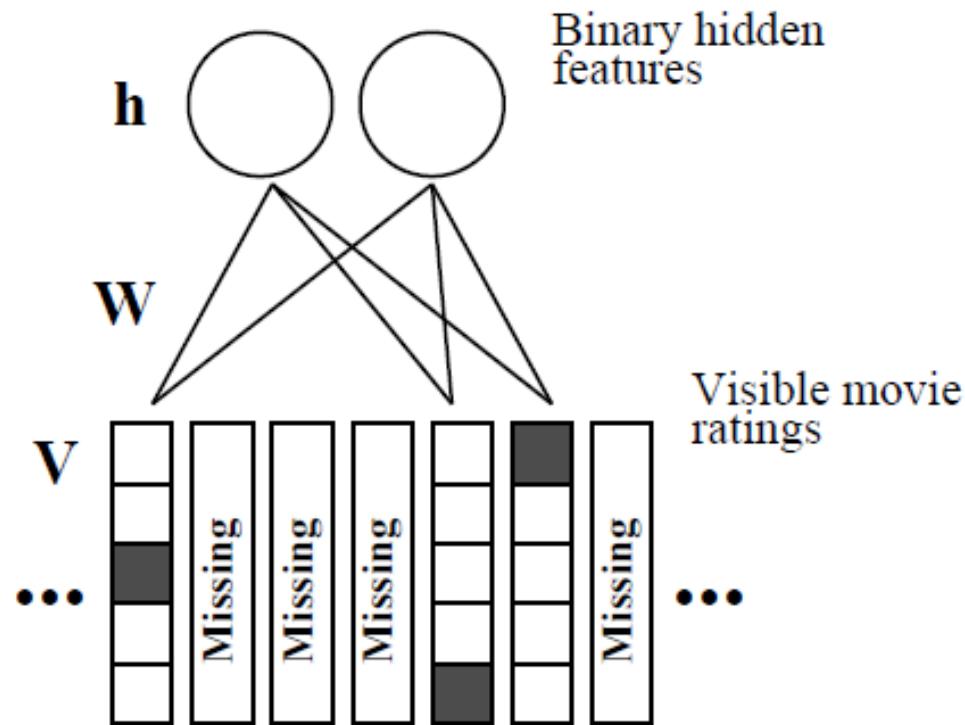


# Анализ данных с помощью сетей на основе машины Больцмана

- Коллаборативная фильтрация
- Классификация и моделирование рукописных цифр
- Моделирование походки и движений человека (motion capture)
- Генерация видеопотока на уровне отдельных пикселей
- Моделирование ЭКГ
- Распознавание речи (по расширенному словарю)
- Распознавание 3-D образов по стереопаре
- Классификация цветных изображений

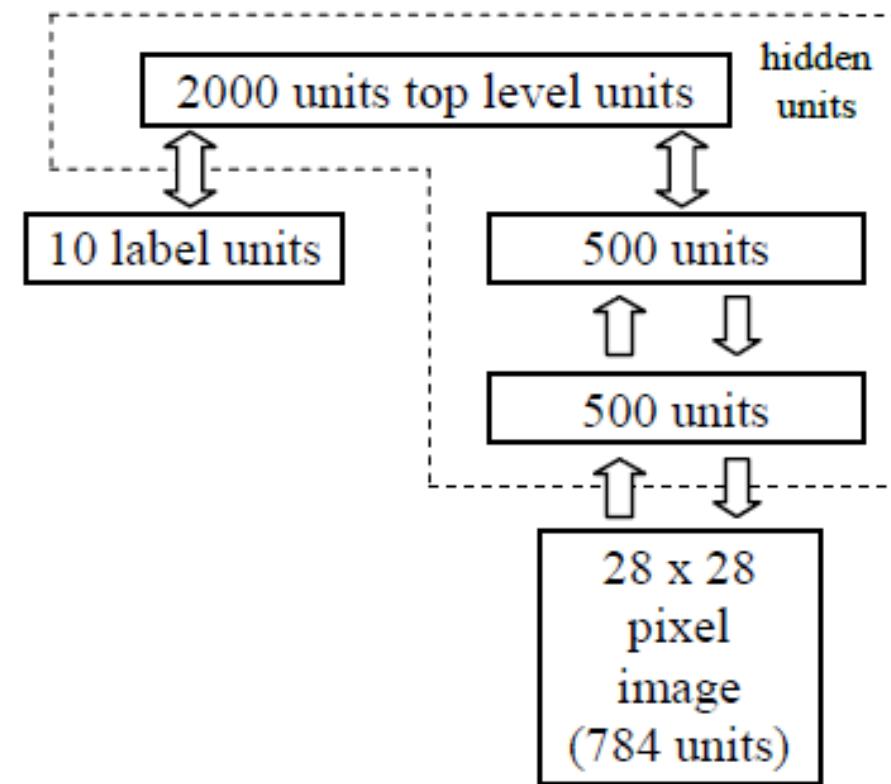
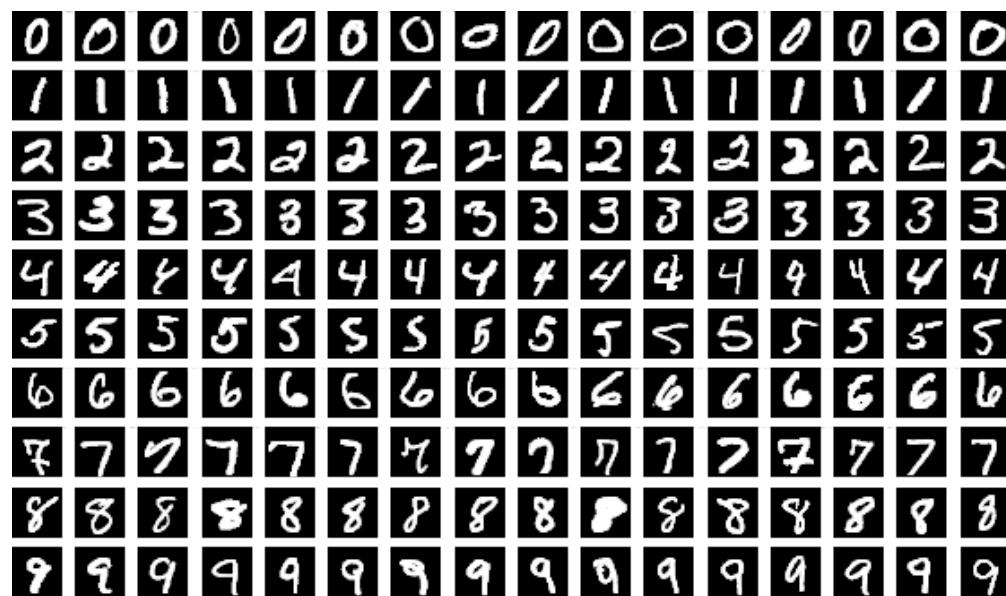
# Коллаборативная фильтрация

- Используется в рекомендательных системах для фильтрации информации и составления автоматических прогнозов



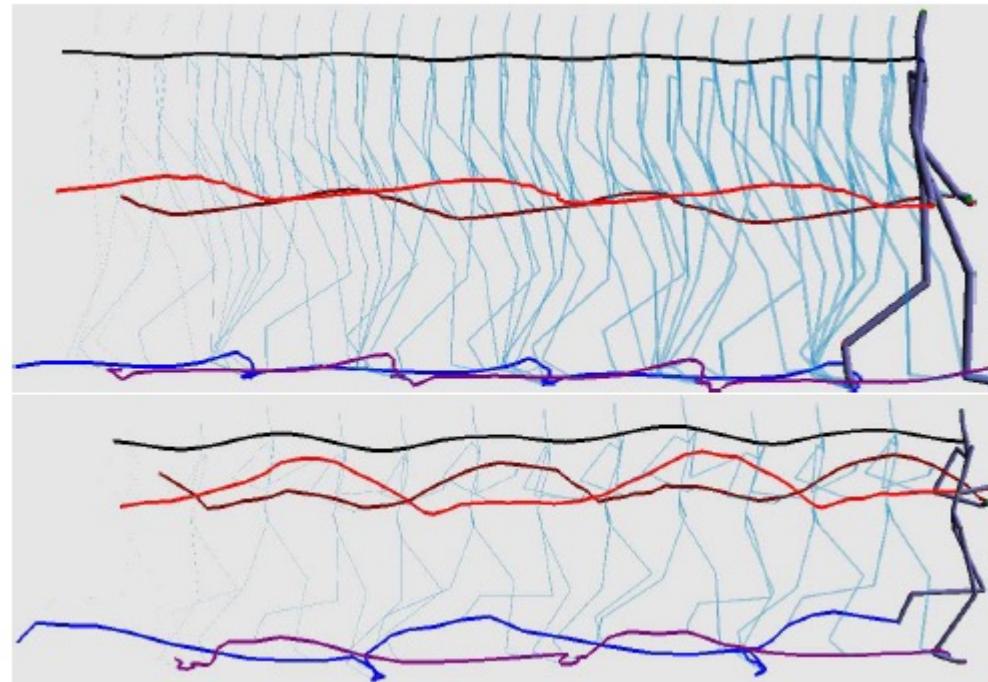
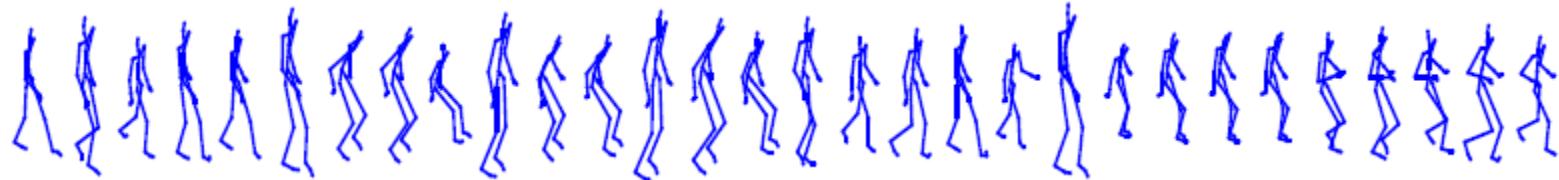
# MNIST

- MNIST – 60000 изображений рукописных чисел



# Моделирование походки

- Сеть обучалась на данных motion capture



# Continuous RBM

- Модель нейрона с вещественным выходом

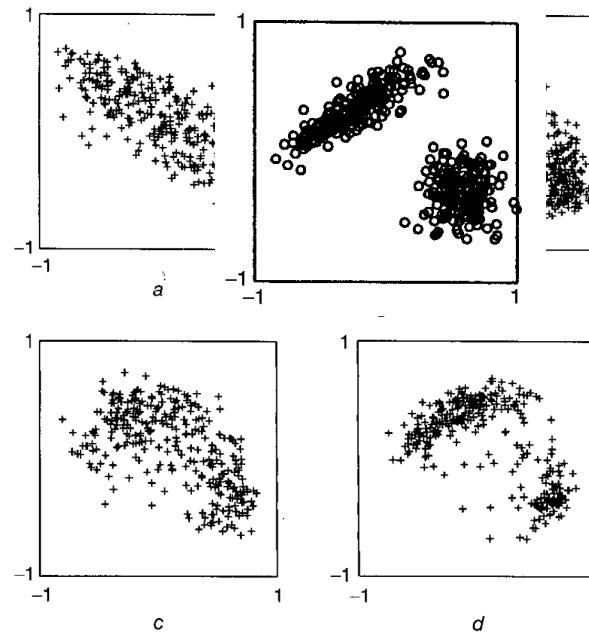
$$s_j = \varphi_j \left( \sum_i w_{ij} s_i + \sigma \cdot N_j(0, 1) \right)$$

$$\varphi_j(x_j) = \theta_L + (\theta_H - \theta_L) \cdot \frac{1}{1 + \exp(-a_j x_j)}$$

$N(0,1)$  – случайный шум с нормальным распределением

# Моделирование с помощью CRBM

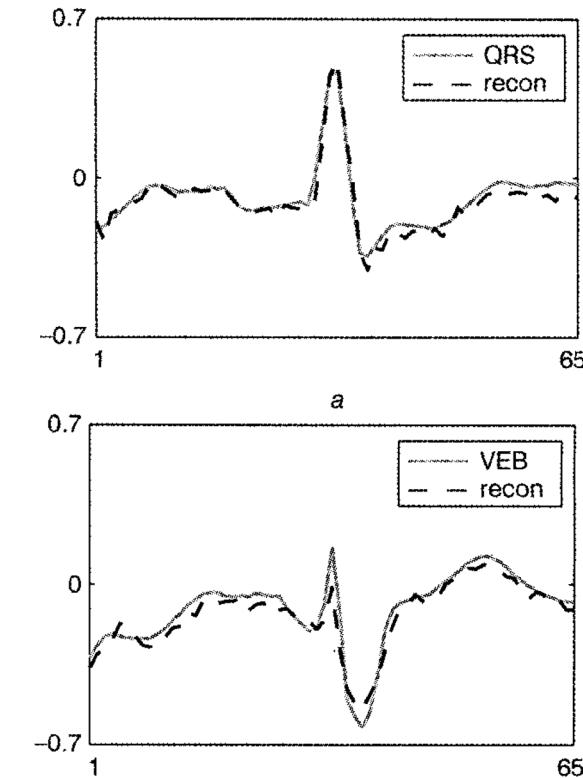
## Искусственные выборки



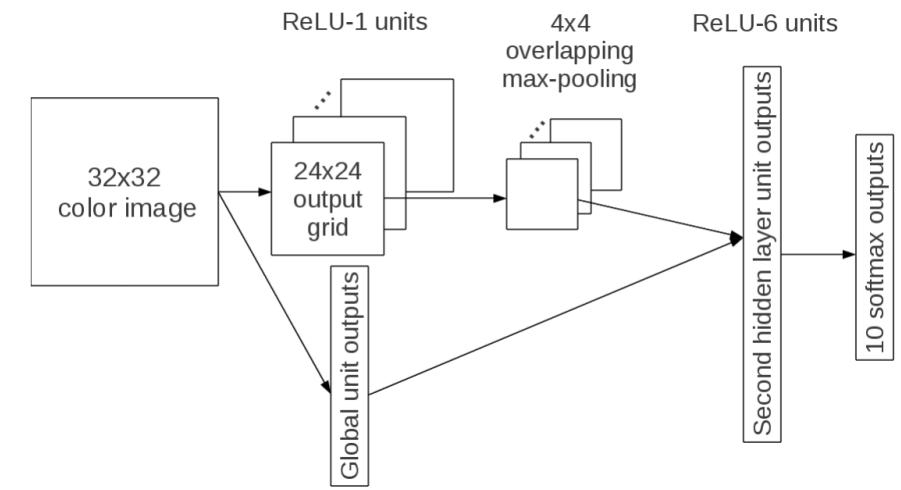
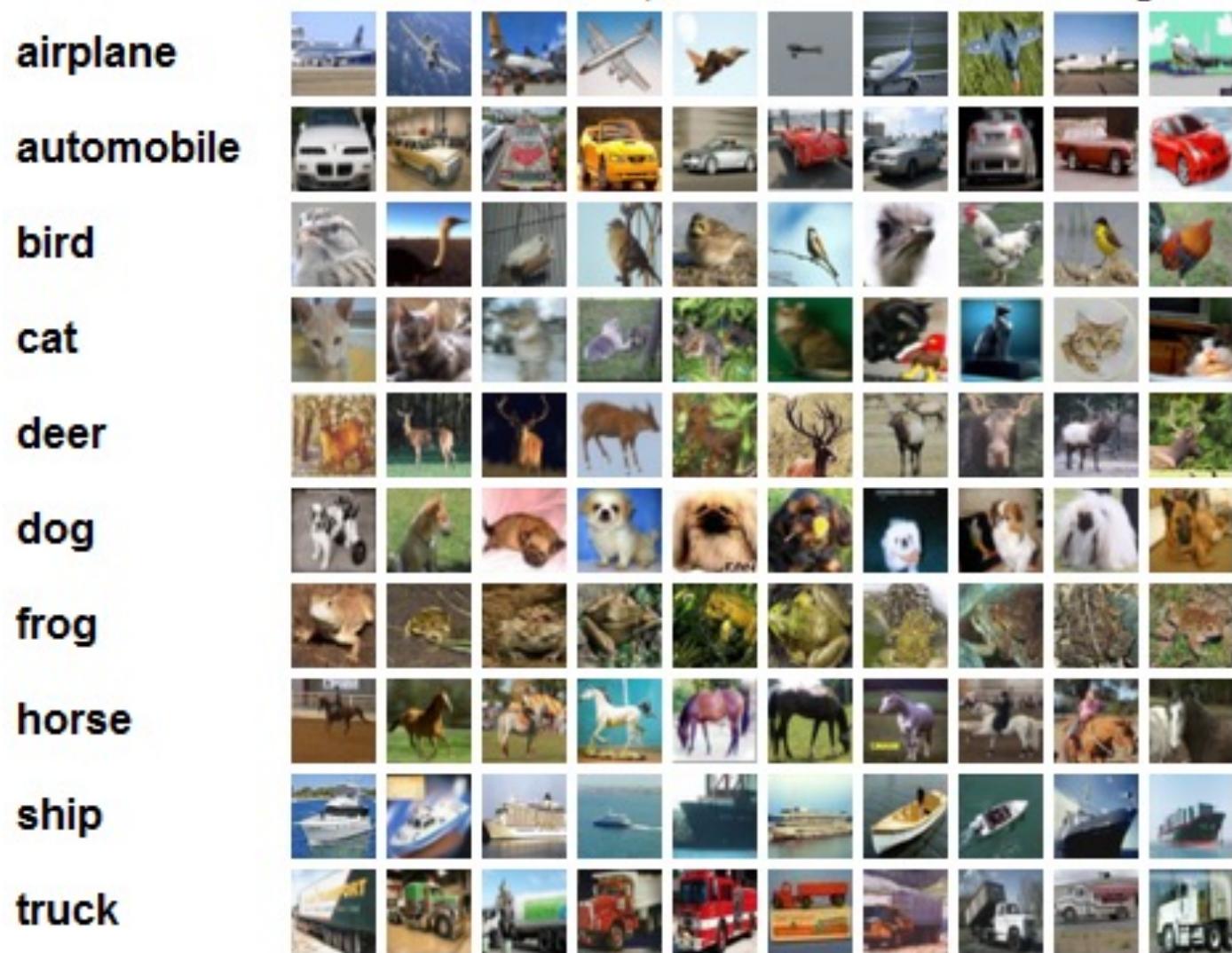
**Fig. 5** 20-step Gibbs sampled reconstructions by the CRBM

- a After 50 training epochs
- b After 395 training epochs
- c After 1800 training epochs
- d After 2700 training epochs

## Моделирование ЭКГ

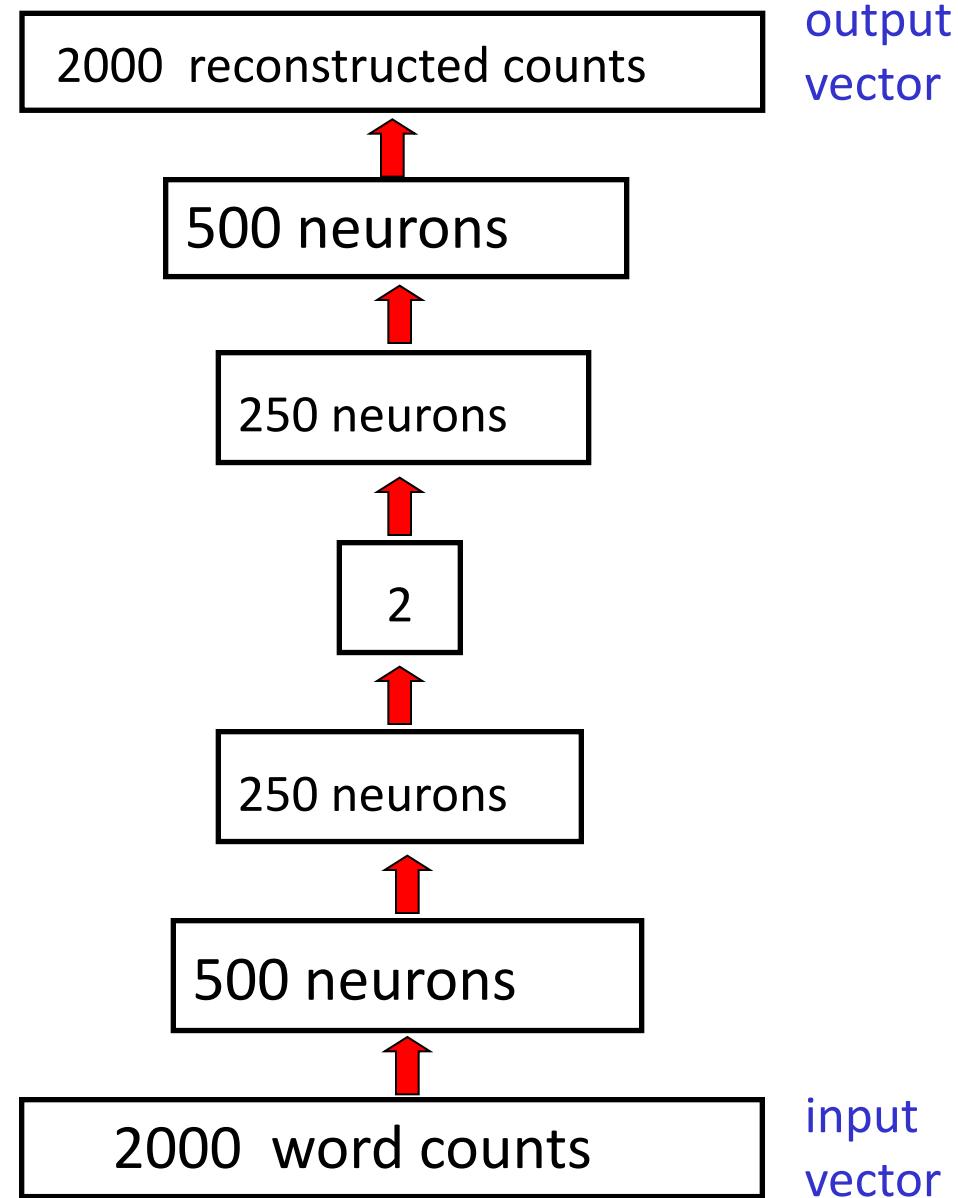


# Классификация изображений: CIFAR 10



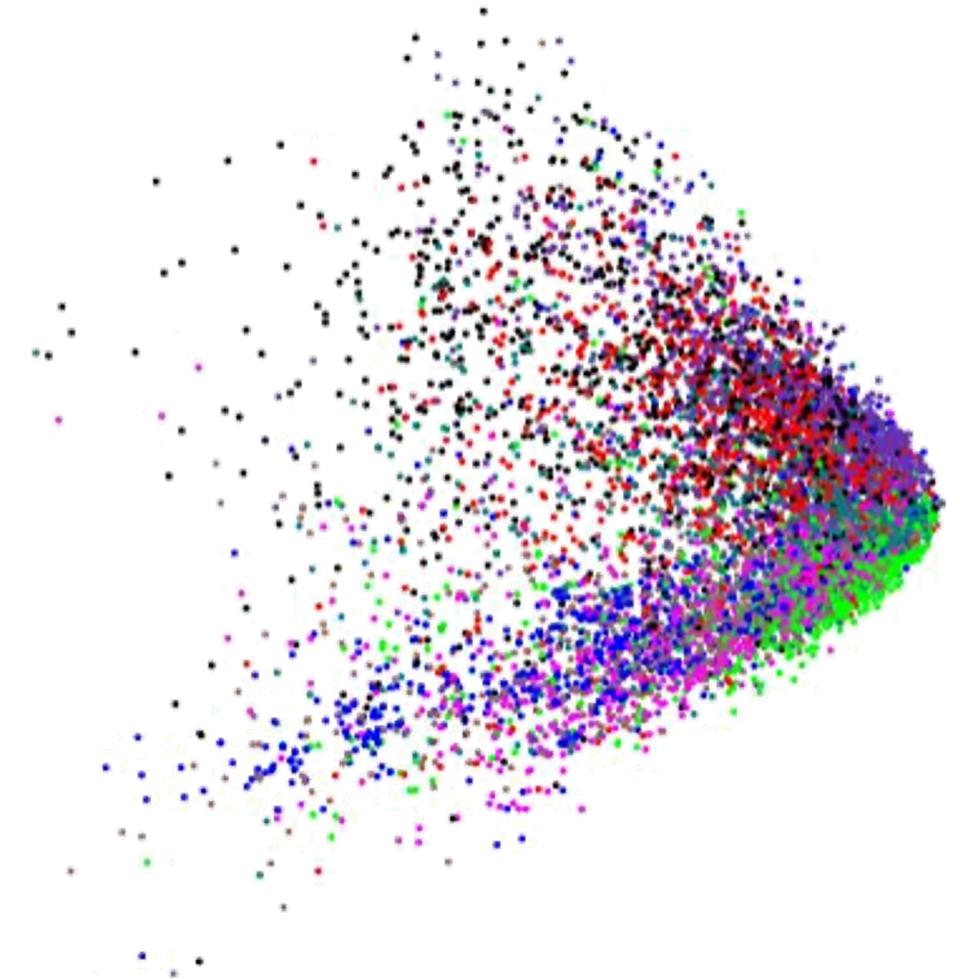
A. Krizhevsky. Convolutional Deep Belief Networks on  
CIFAR-10. UoT, 2010  
<https://www.cs.toronto.edu/~kriz/conv-cifar10-aug2010.pdf>

# Уменьшение размерности данных



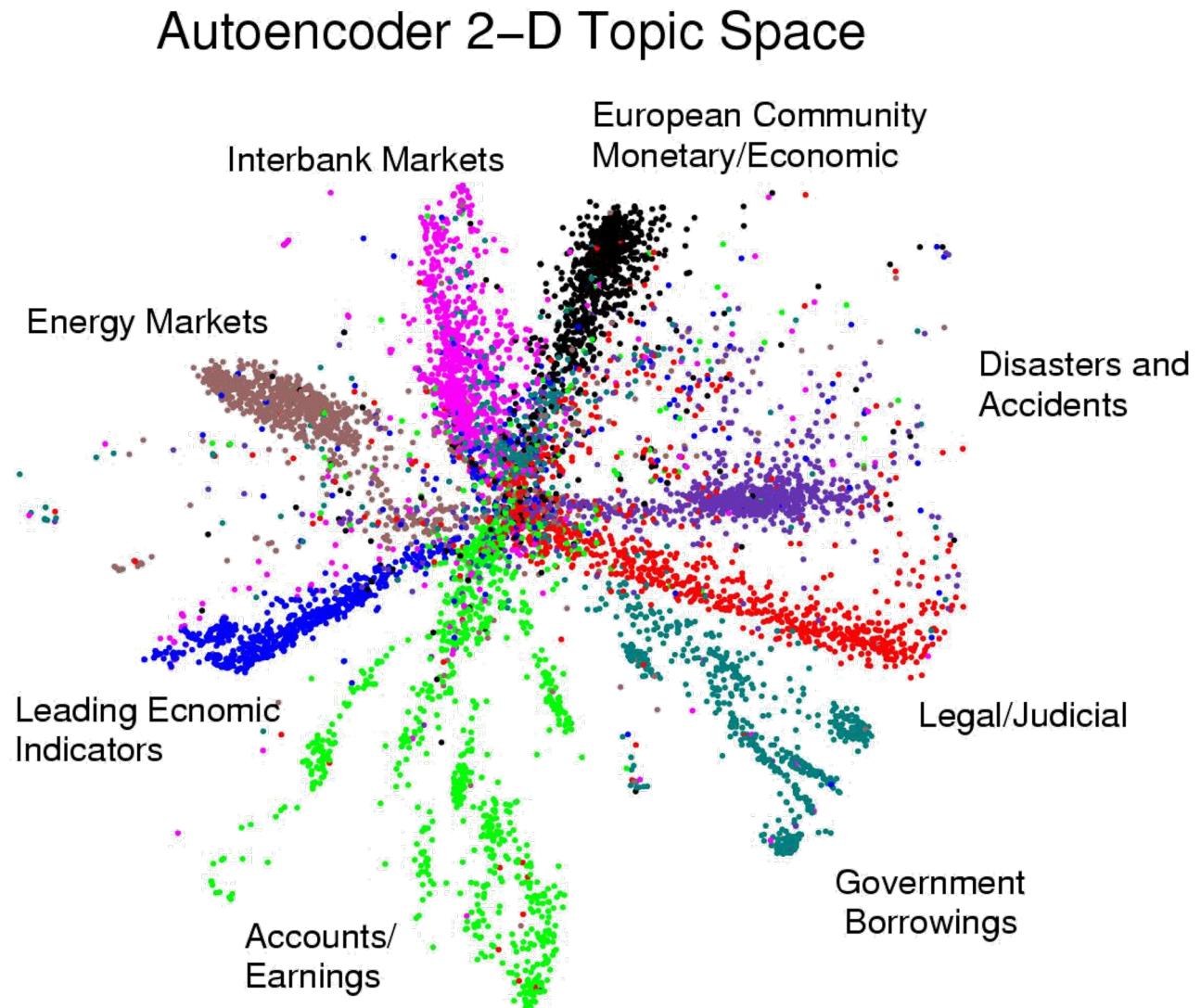
# Уменьшение размерности

- Тексты на разные темы сжимаются до 2 вещественных чисел с помощью разновидности **метода главных компонент**
- С помощью цвета отображены различные категории текстов
- Вывод: классификация в двумерном пространстве невозможна

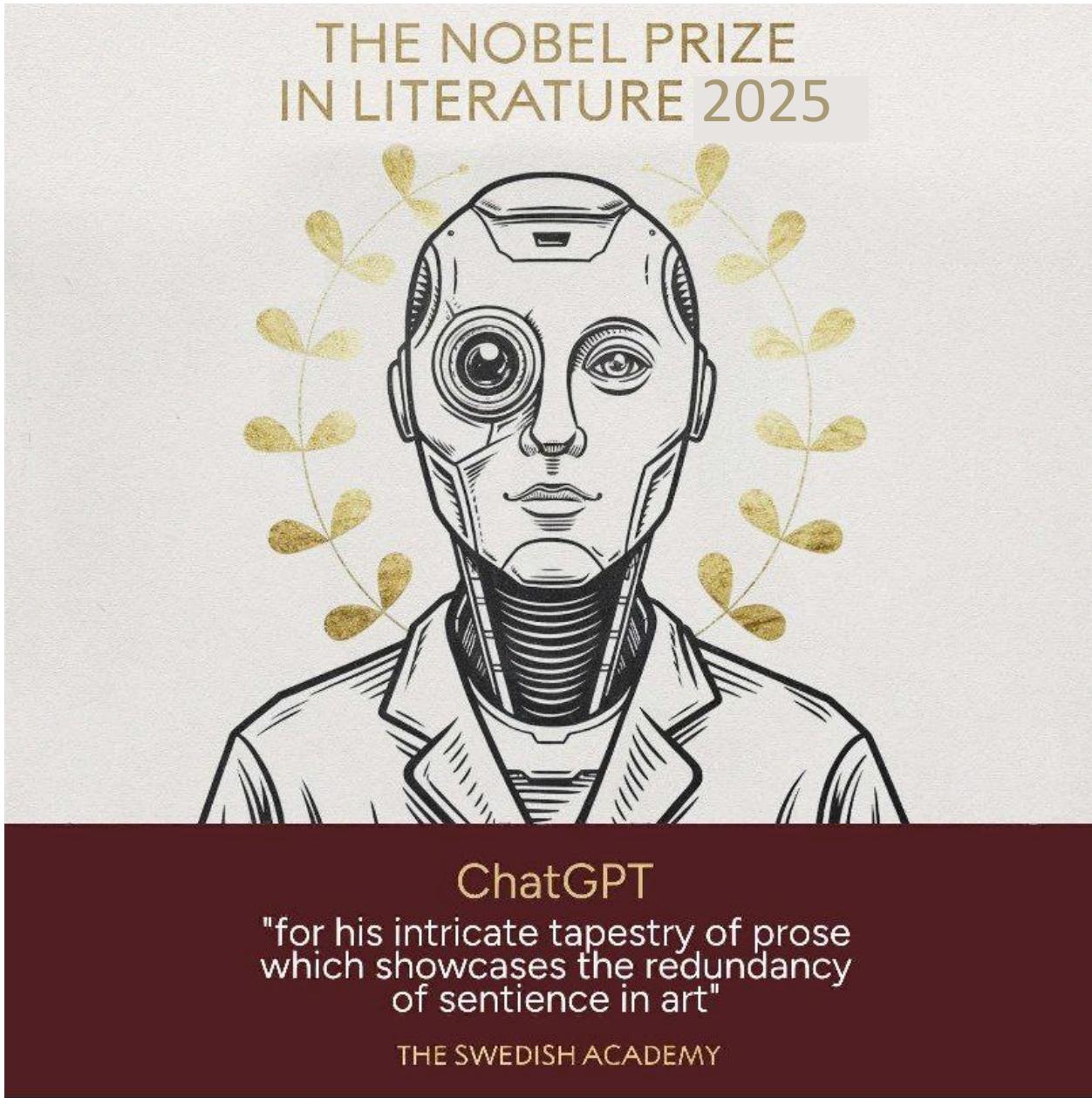


# Уменьшение размерности

- Тексты на разные темы сжимаются до 2 вещественных чисел с помощью **автоэнкодера**
- С помощью цвета отображены различные категории текстов
- Вывод: классификация в двумерном пространстве элементарна



# Что дальше?



"Нобелевскую премию по литературе 2025 года получает ...

**ChatGPT**

... за его сложное полотно прозы, демонстрирующее избыточность сознания в искусстве"

/s