

Машинное обучение.  
Нейронные сети  
(часть 2)

# Задача машинного обучения (на примере классификации)

$U$  – множество объектов

$X(U)$  - признаки объектов

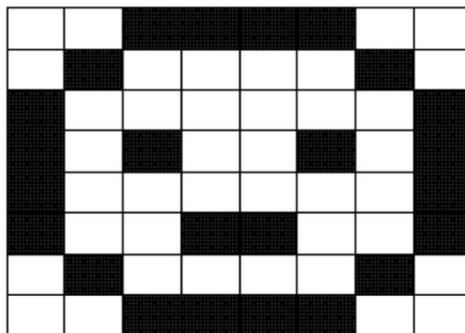
$Y$  – классы объектов

# Задача машинного обучения (на примере классификации)

$U$

$X(U)$

$Y$



Человек

Птица

Дом

⋮

# Задача машинного обучения (на примере классификации)

Решающая функция

$$F: X \rightarrow Y$$

Обучающая выборка

|          |       |   |       |       |
|----------|-------|---|-------|-------|
| $D$      | {     | <table border="1"><tr><td><math>x_1</math></td><td><math>y_1</math></td></tr></table> | $x_1$ | $y_1$ |
|          |       | $x_1$   | $y_1$ |       |
|          |       | <table border="1"><tr><td><math>x_2</math></td><td><math>y_2</math></td></tr></table> | $x_2$ | $y_2$ |
| $x_2$    | $y_2$ |   |       |       |
| $\vdots$ |       |   |       |       |
|          |       | <table border="1"><tr><td><math>x_N</math></td><td><math>y_N</math></td></tr></table> | $x_N$ | $y_N$ |
| $x_N$    | $y_N$ |   |       |       |

# Задача машинного обучения (на примере классификации)

Функция ошибки

$$L(D, F) = \sum_i [F(x_i) \neq y_i]$$

Оптимальная решающая функция

$$W^* = \operatorname{argmin}_W L(D, F(W))$$

# Проблема исключающего ИЛИ

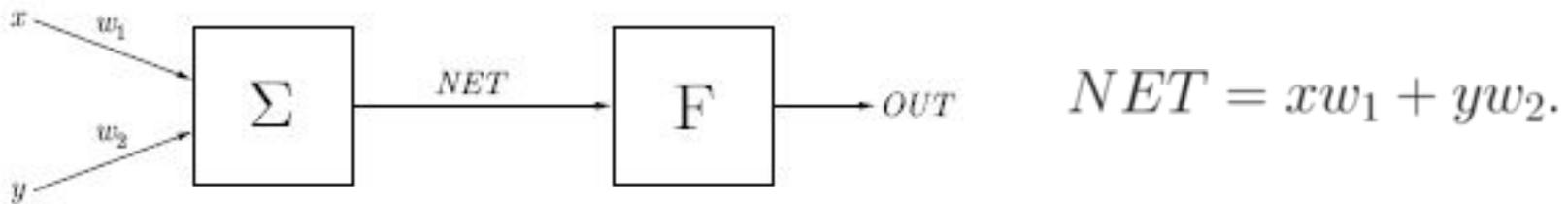
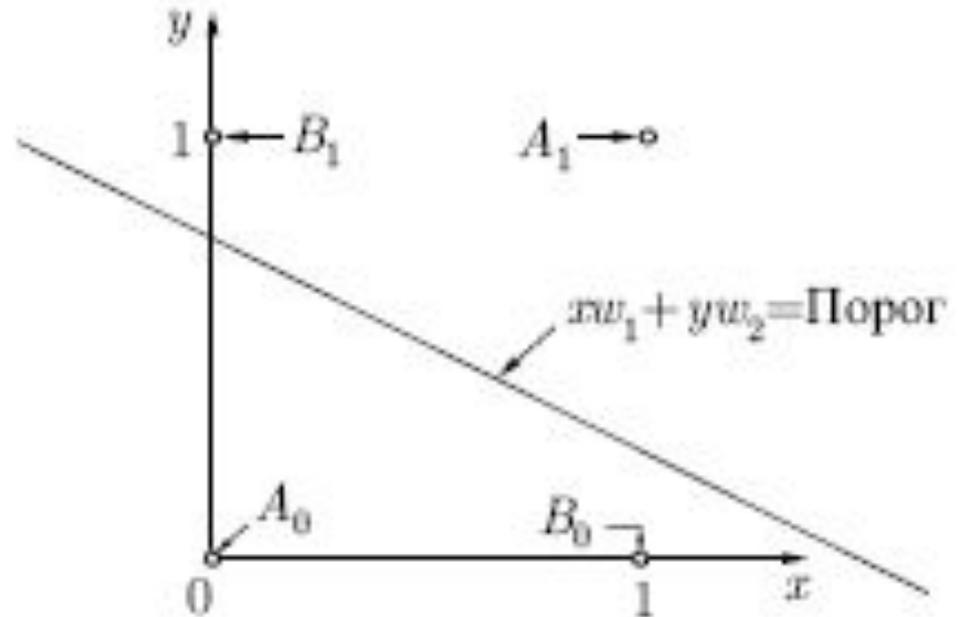
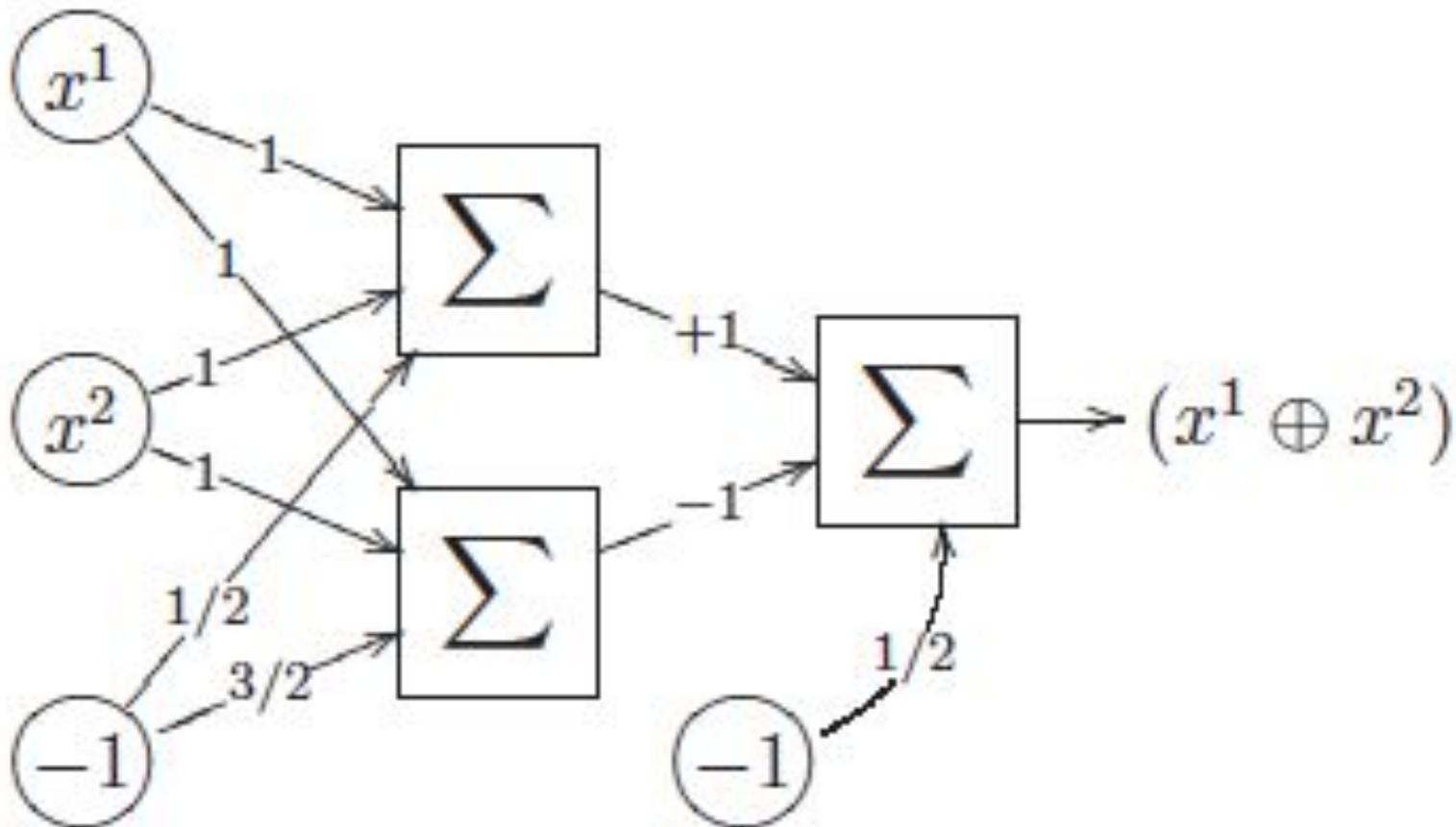


Таблица истинности

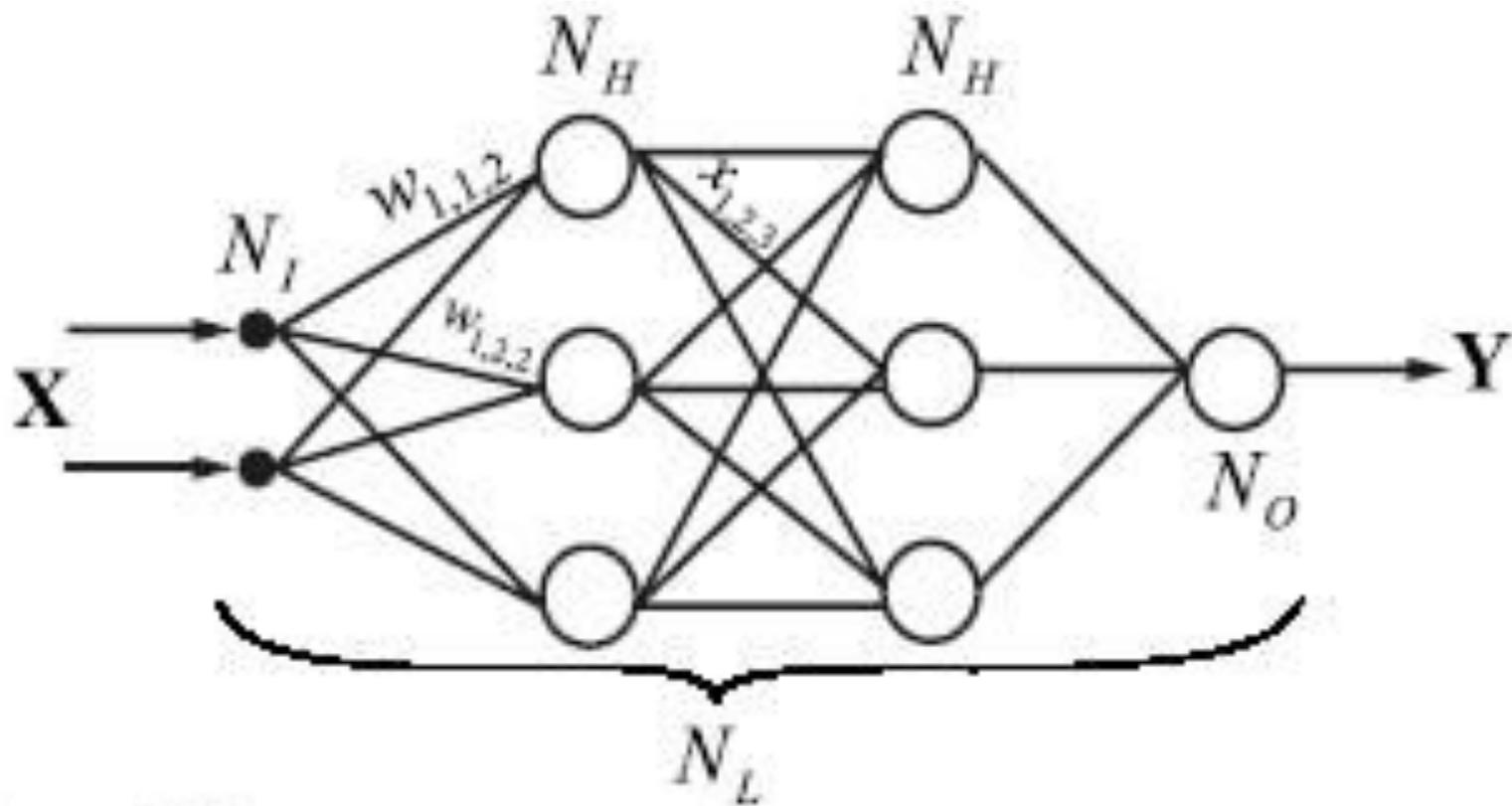
|   |   |   |
|---|---|---|
| 0 | 0 | 0 |
| 1 | 0 | 1 |
| 0 | 1 | 1 |
| 1 | 1 | 0 |



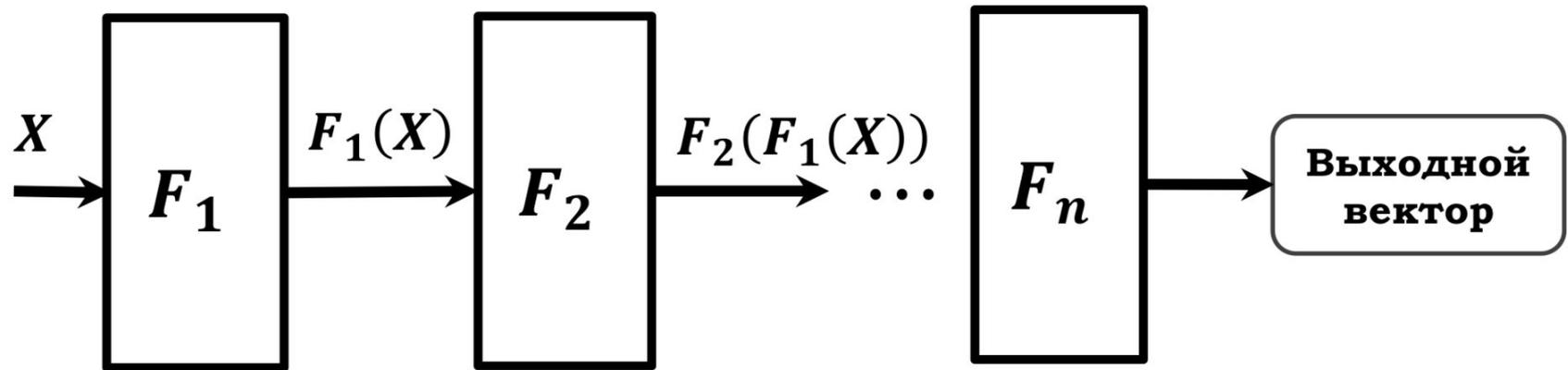
# Персептрон. Искл. ИЛИ



# Многослойный персептрон



# Персептрон общего вида



$$\Delta = \text{Выходной вектор} - \text{Эталонный выход}$$

$$\frac{\partial \Delta}{\partial W_n} = \frac{\partial \Delta(F_n)}{\partial F_n} * \frac{\partial F_n(F_{n-1}, W_n)}{\partial W_n}$$

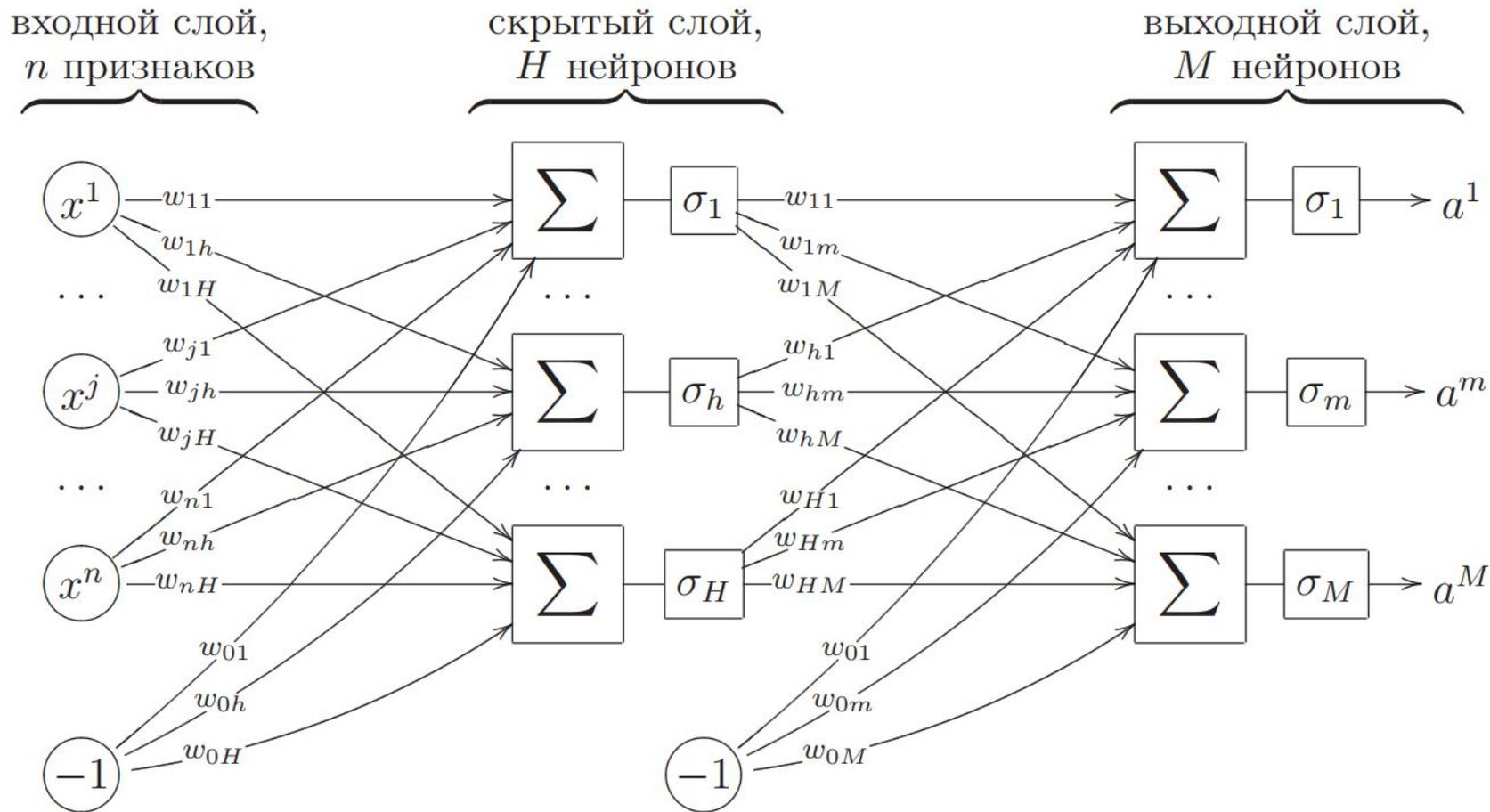
# Многослойный персептрон

**Теорема 1.2 (Колмогоров, 1957).** *Любая непрерывная функция  $n$  аргументов на единичном кубе  $[0, 1]^n$  представима в виде суперпозиции непрерывных функций одного аргумента и операции сложения:*

$$f(x^1, x^2, \dots, x^n) = \sum_{k=1}^{2n+1} h_k \left( \sum_{i=1}^n \varphi_{ik}(x^i) \right),$$

где  $h_k, \varphi_{ik}$  — непрерывные функции, причём  $\varphi_{ik}$  не зависят от выбора  $f$ .

# Многослойная сеть с одним скрытым слоем



**Алгоритм 1.3.** Обучение двухслойной сети методом back-propagation — обратного распространения ошибки

---

**Вход:**

$X^\ell = (x_i, y_i)_{i=1}^\ell$  — обучающая выборка,  $x_i \in \mathbb{R}^n$ ,  $y_i \in \mathbb{R}^M$ ;

$H$  — число нейронов в скрытом слое;

$\eta$  — темп обучения;

**Выход:**

синаптические веса  $w_{jh}$ ,  $w_{hm}$ ;

---

1: инициализировать веса небольшими случайными значениями:

$$w_{jh} := \text{random} \left( -\frac{1}{2n}, \frac{1}{2n} \right);$$

$$w_{hm} := \text{random} \left( -\frac{1}{2H}, \frac{1}{2H} \right);$$

2: **повторять**

3: выбрать объект  $x_i$  случайным образом;

4: прямой ход:

$$u_i^h := \sigma_h \left( \sum_{j=0}^J w_{jh} v^j(x_i) \right), \text{ для всех } h = 1, \dots, H;$$

$$a_i^m := \sigma_m \left( \sum_{h=0}^H w_{hm} u^h(x_i) \right), \text{ для всех } m = 1, \dots, M;$$

$$\varepsilon_i^m := a_i^m - y_i^m, \text{ для всех } m = 1, \dots, M;$$

$$Q_i := \sum_{m=1}^M (\varepsilon_i^m)^2;$$

5: обратный ход:

$$\varepsilon_i^h := \sum_{m=1}^M \varepsilon_i^m \sigma'_m w_{hm}, \text{ для всех } h = 1, \dots, H;$$

6: градиентный шаг:

$$w_{hm} := w_{hm} - \eta \varepsilon_i^m \sigma'_m u^h, \text{ для всех } h = 0, \dots, H, m = 1, \dots, M;$$

$$w_{jh} := w_{jh} - \eta \varepsilon_i^h \sigma'_h x^j, \text{ для всех } j = 0, \dots, n, h = 1, \dots, H;$$

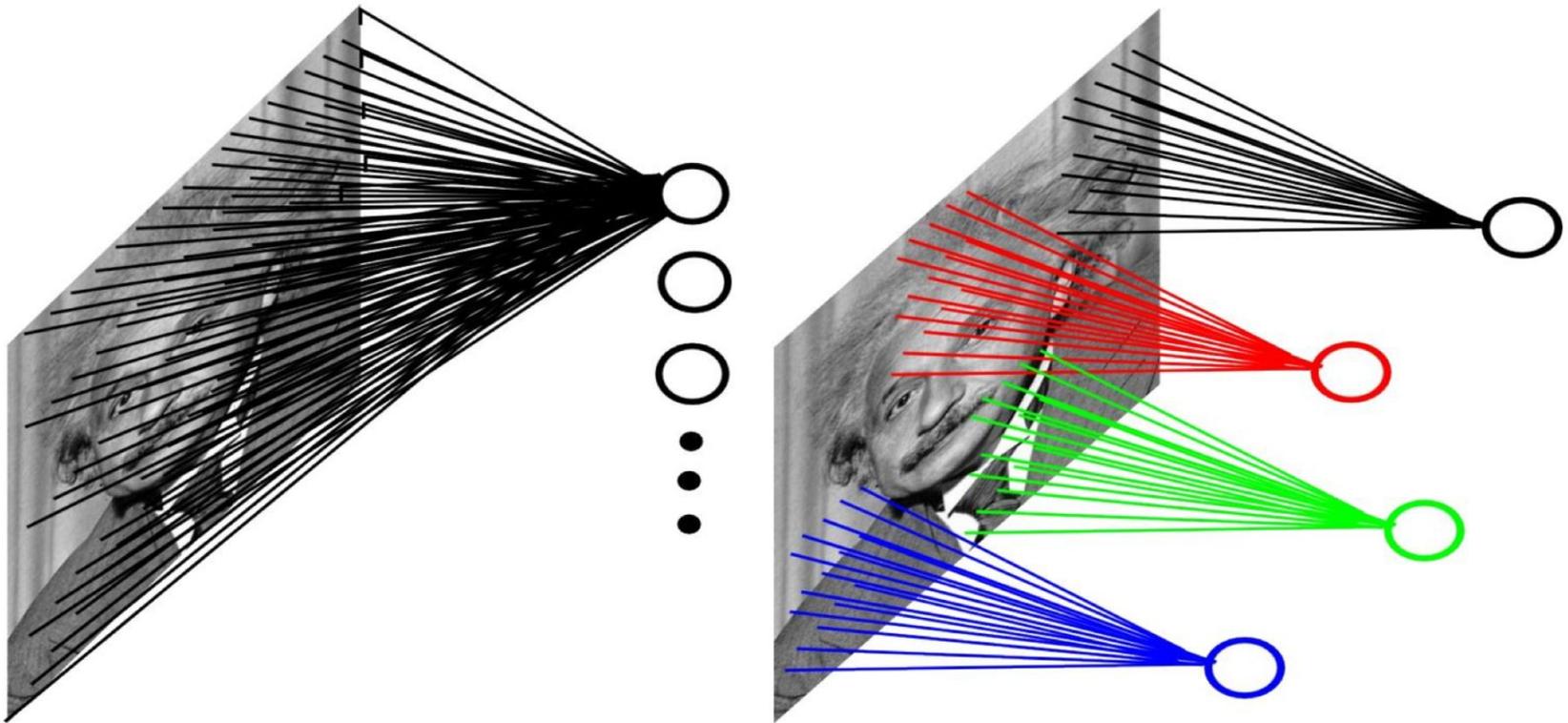
7:  $Q := \frac{\ell-1}{\ell} Q + \frac{1}{\ell} Q_i$ ;

8: **пока**  $Q$  не стабилизируется;

---

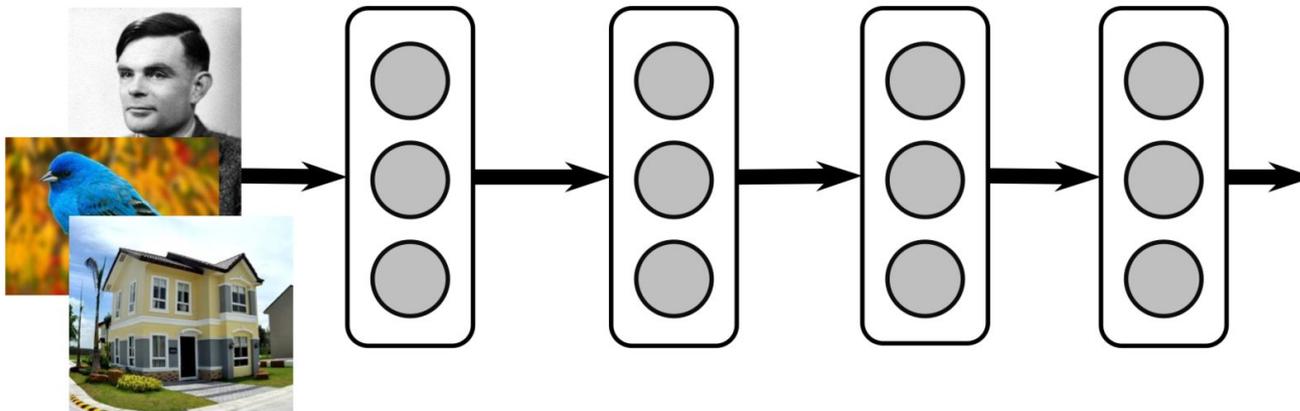
- К.В. Воронцов «Лекции по искусственным нейронным сетям», §1.2.1
- Самостоятельно

# Сверточные сети

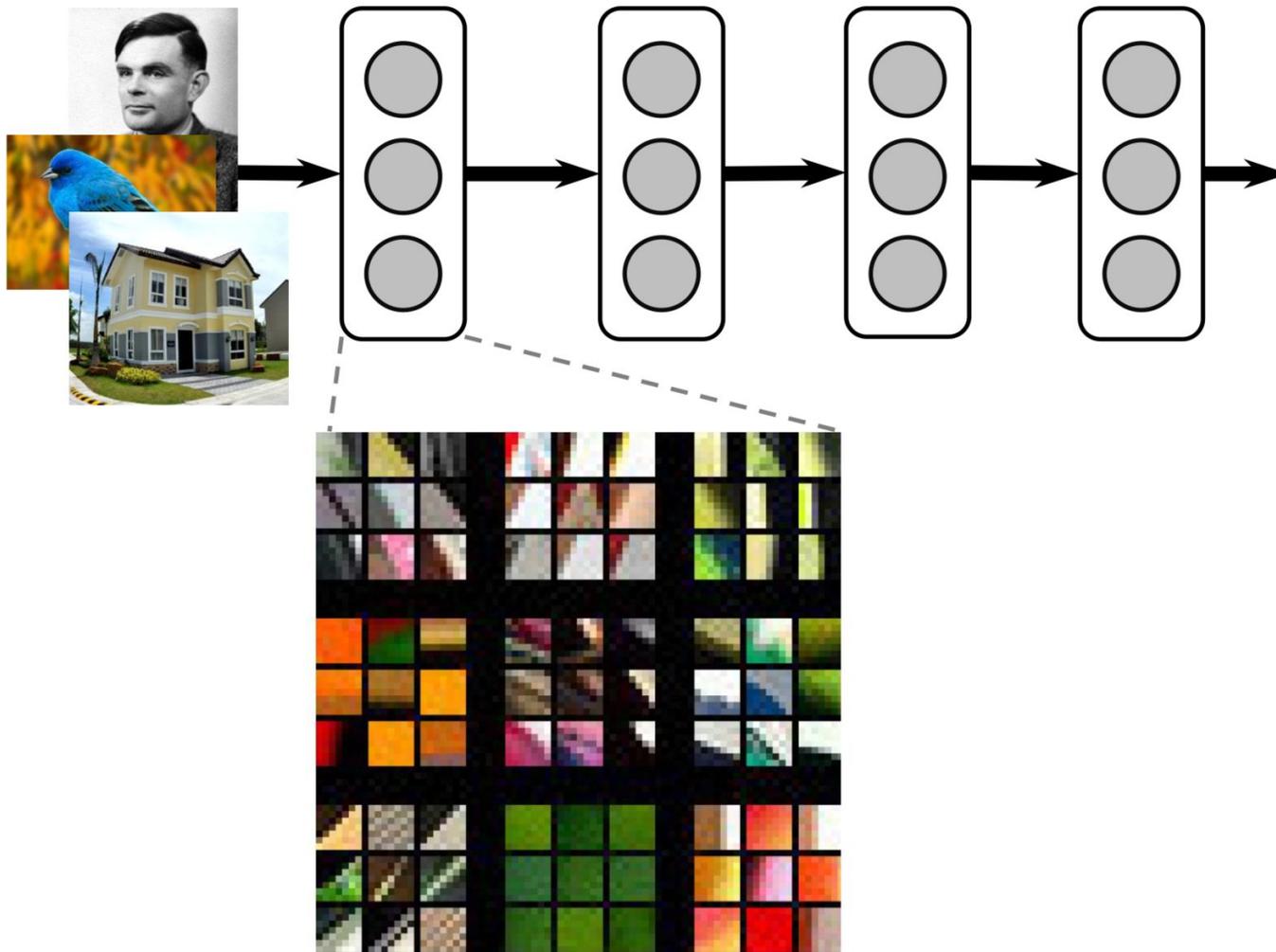




# Обучение иерархии признаков

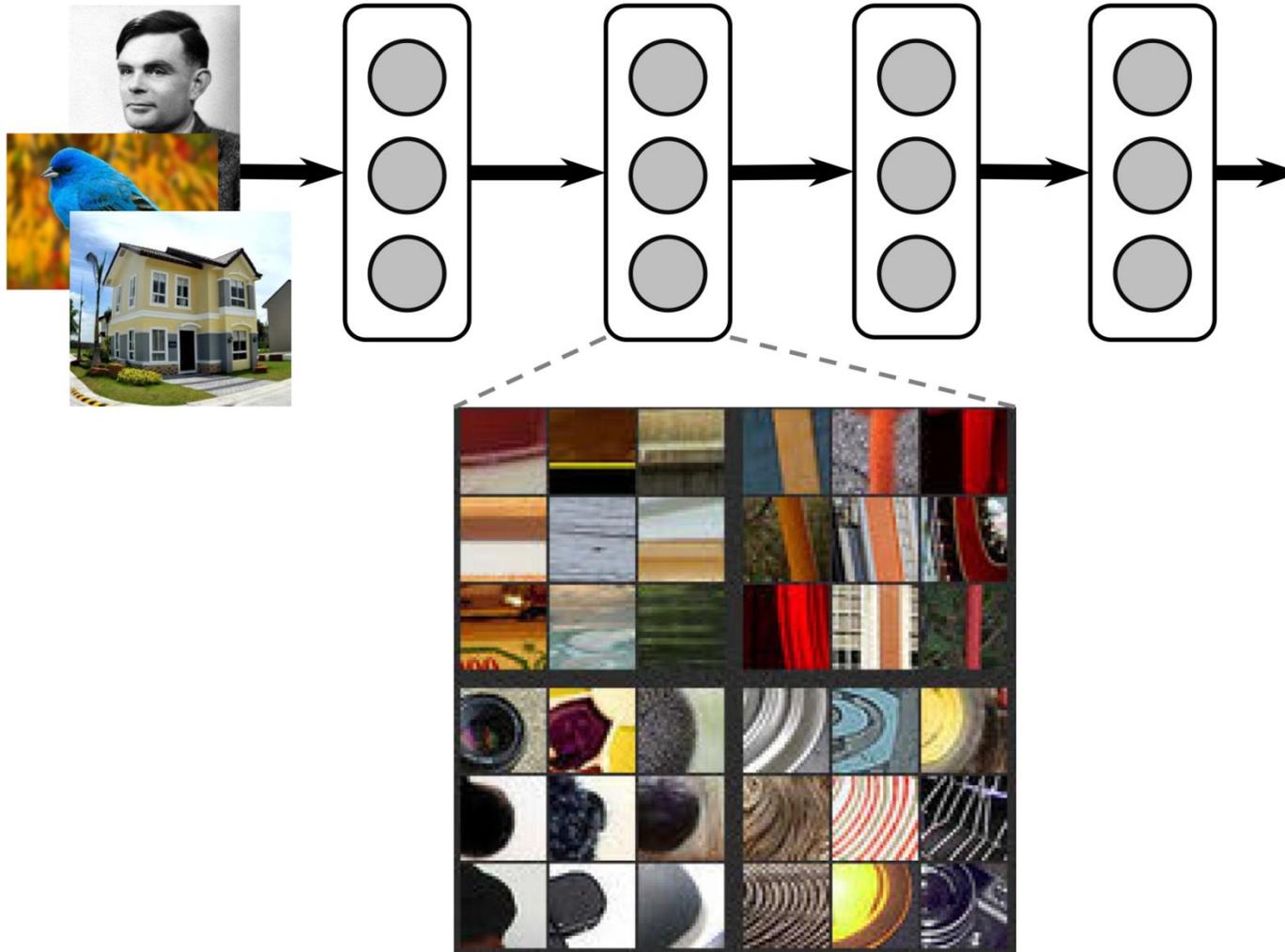


# Обучение иерархии признаков



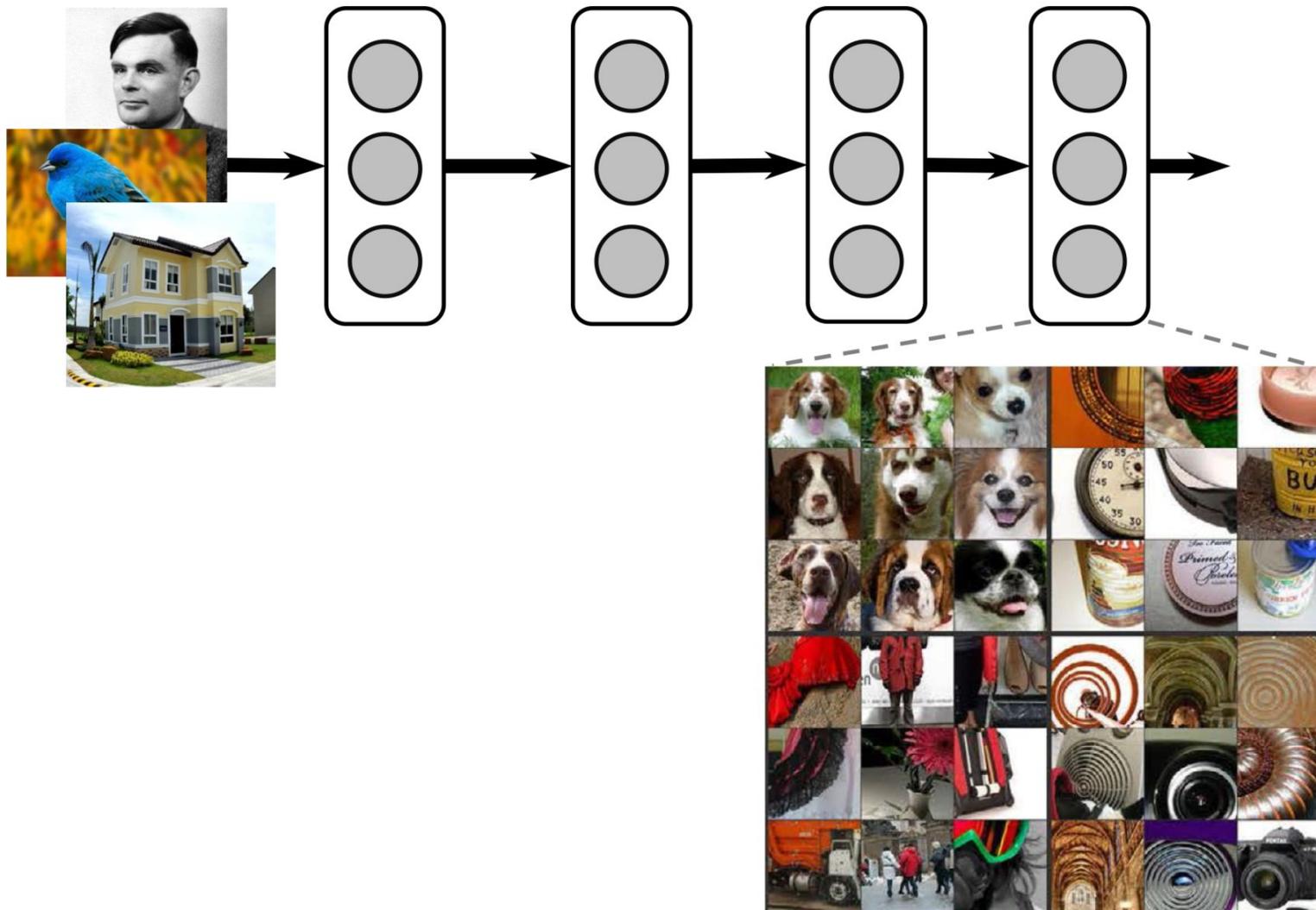
Zeiler M., Fergus R. <http://arxiv.org/pdf/1311.2901v3.pdf>, 2013

# Обучение иерархии признаков



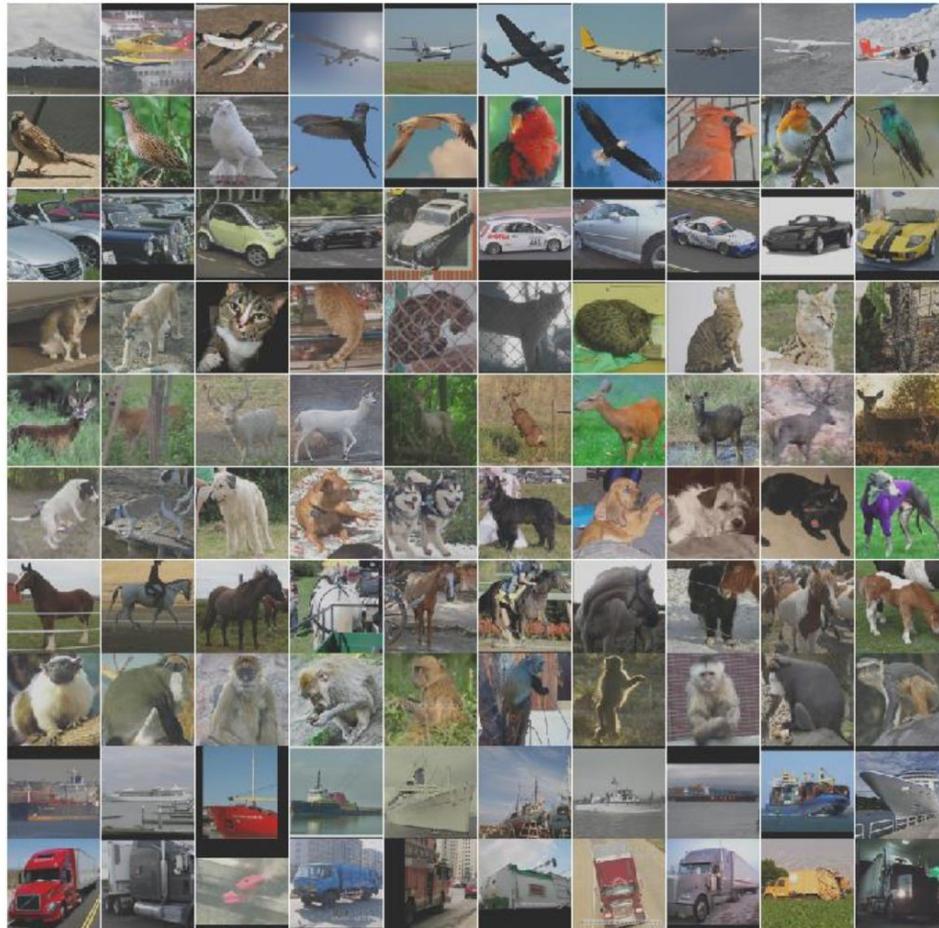
Zeiler M., Fergus R. <http://arxiv.org/pdf/1311.2901v3.pdf>, 2013

# Обучение иерархии признаков



Zeiler M., Fergus R. <http://arxiv.org/pdf/1311.2901v3.pdf>, 2013

# ImageNet



**1.2 миллиона изображений, 1000 классов**

# ImageNet

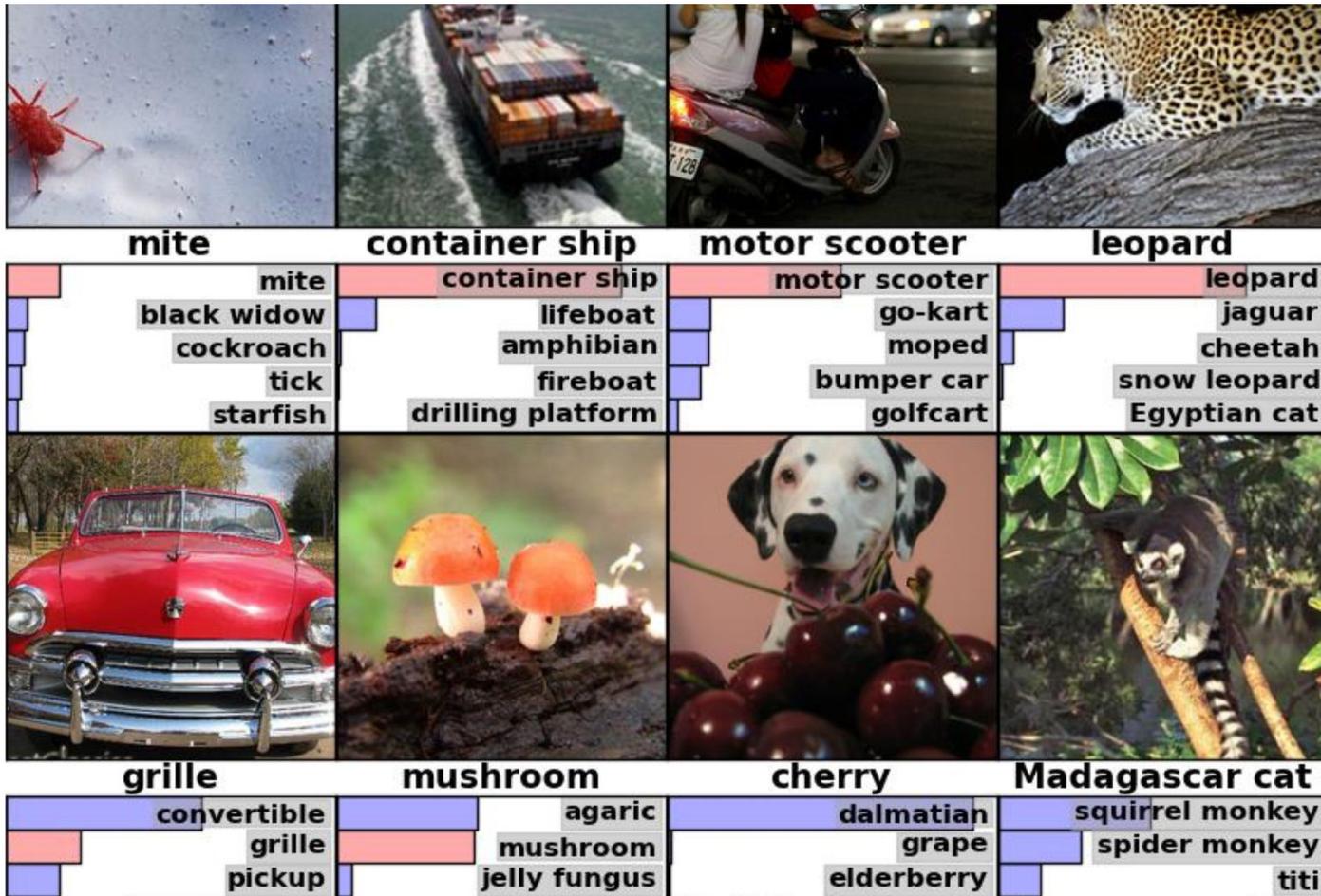
**Хаски**



**Сибирский хаски**

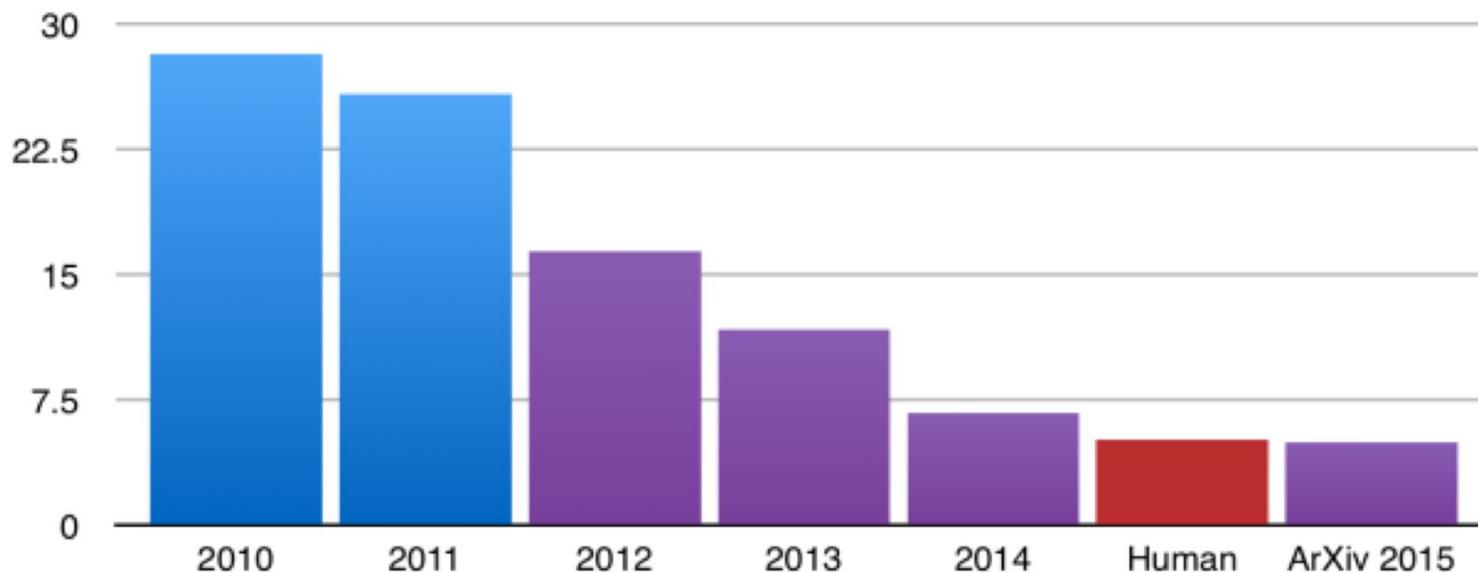


# Вероятностное распознавание



# Прорыв в распознавании изображений

ILSVRC top-5 error on ImageNet





# Похожие изображения

Яндекс

Картинки



Загруженная картинка ×

Найти



klakhman

[← Вернуться назад](#)



# Стилизация изображений

Пабло Пикассо

Уильям Тёрнер

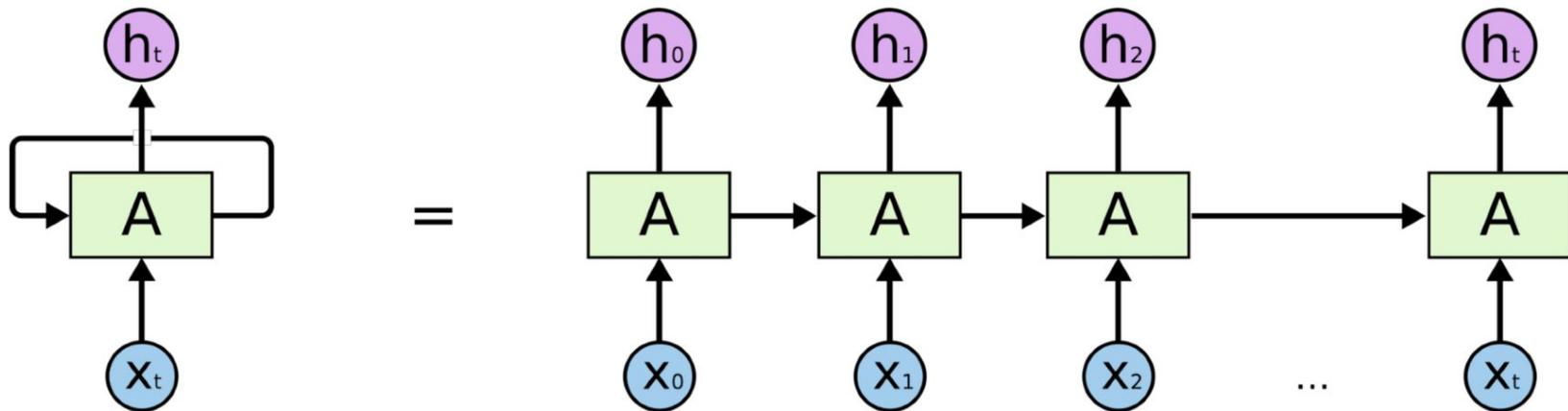


Винсент Ван Гог

Василий Кандинский

Эдвард Мунк

# Рекуррентные сети



# Диалог с рекуррентной сетью

**Machine:** hi

**Human:** hello

**Machine:** could you please let me know what are the operating systems you are using ?

**Human:** i am using linux

**Machine:** ok it was in the ssl vpn on which you stated it was broken

**Machine:** please check if you are logged in to corp network

**Machine:** what is the error that you are running please

**Human:** i am seeing an error related to vpn

**Machine:** what is the error message that you are getting when connecting to vpn using network connect ?

**Human:** connection refused or something like that

**Machine:** may i know the version of network connect you connect ?

**Human:** i am not sure i know that

**Machine:** is the network connect prompting for any user certificate

**Machine:** ?

**Human:** yes

**Machine:** are you at home or at work at the moment ?

**Machine:** office ?

**Human:** i am at home

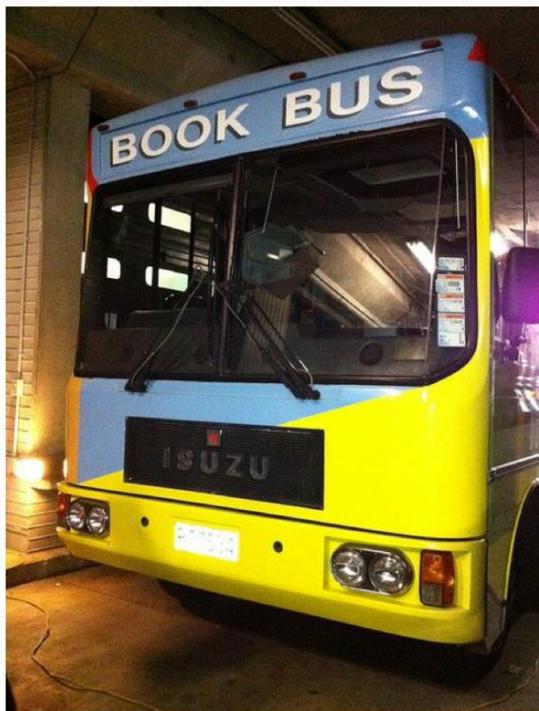
# Генерация описания изображения



Later on the eighth day , Billy was a friend of a man who lived on his own . He did n't know how far away they were , and if he was to survive the fall . His mind raced , trying not to show any signs of weakness .

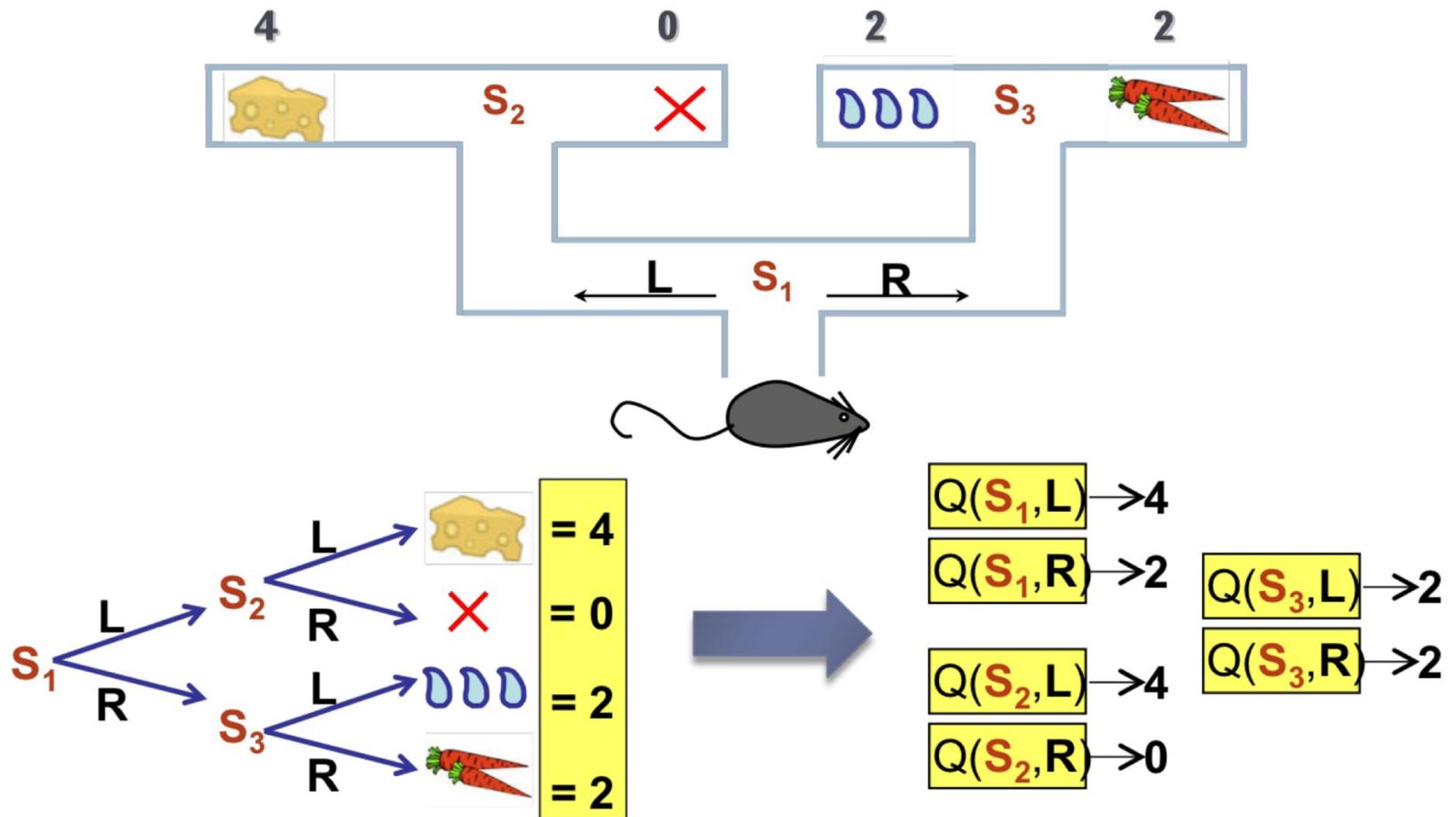
The wind ruffled the snow and ice in the snow . He had no idea how many times he was going to climb into the mountains

# Генерация описания изображения



In the second day , I panicked . I haven t seen him enter the city , so I quickly corrected the bus ride to the city . I was hoping to see that they were on such a bus , and as soon as we got within the city limits , I had no idea where I was . The bus driver was on the verge of tears running down his nose and mouth . I picked up the rental car and headed back towards London . The ferry never faltered in such a state , but it s not .

# Обучение с подкреплением



# Обучение с подкреплением

$$Q^*(s, a) = R + \gamma \max_{a'} Q^*(s', a')$$

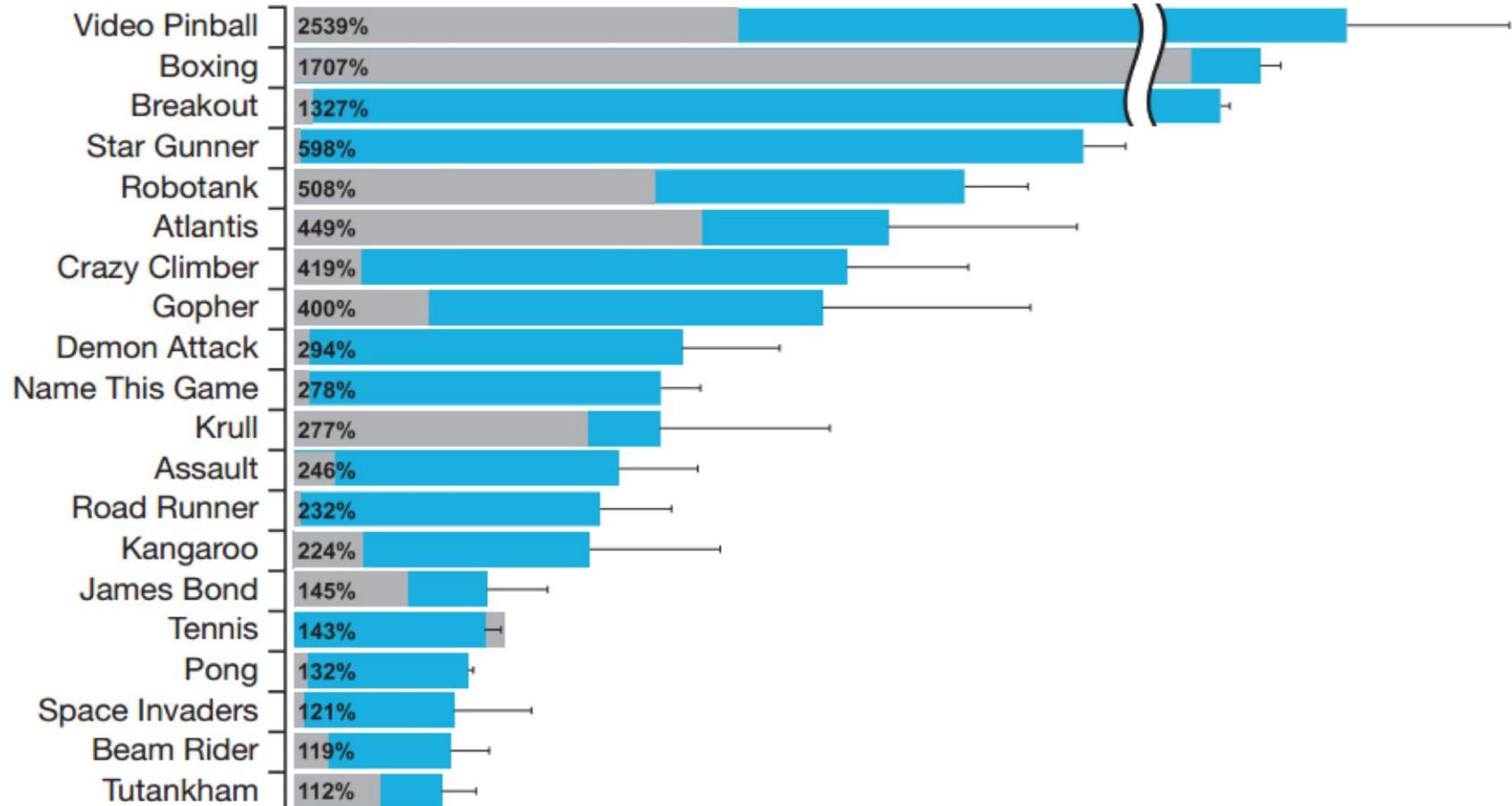
$$\underbrace{Q_{t+1}(s_t, a_t)}_{\text{новое значение}} = \underbrace{Q_t(s_t, a_t)}_{\text{старое значение}} + \alpha \left( R_{t+1} + \gamma \max_a Q_t(s_{t+1}, a_t) - Q_t(s_t, a_t) \right)$$

# Обучение с подкреплением

$$Q^*(s, a) = R + \gamma \max_{a'} Q^*(s', a')$$

$$Q_{t+1}(s_t, a_t) = Q_t(s_t, a_t) + \underbrace{\alpha \left( \underbrace{R_{t+1}}_{\text{награда}} + \gamma \max_a \underbrace{Q_t(s_{t+1}, a_t)}_{\text{оценка будущих наград}} - Q_t(s_t, a_t) \right)}$$

# Atari 2600 games



Mnih V. et al. Human-level control through deep reinforcement learning. *Nature*, 2015