

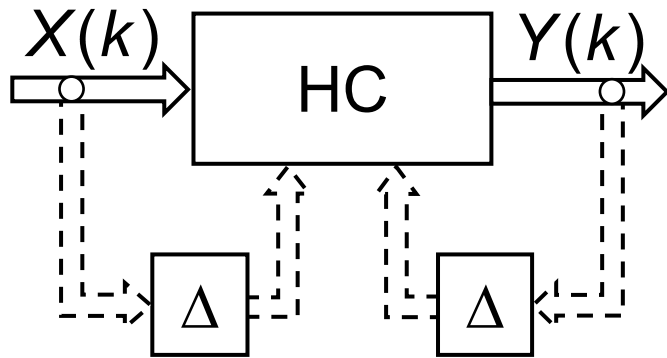
Лекция 7

РЕКУРРЕНТНЫЕ НС. НС ГЛУБОКОГО ОБУЧЕНИЯ (DEEP LEARNING).



РЕКУРРЕНТНАЯ НЕЙРОННАЯ СЕТЬ

(Recurrent Neural Network, RNN) – это НС с элементами памяти (временной задержки) на входе и в цепи обратной связи.



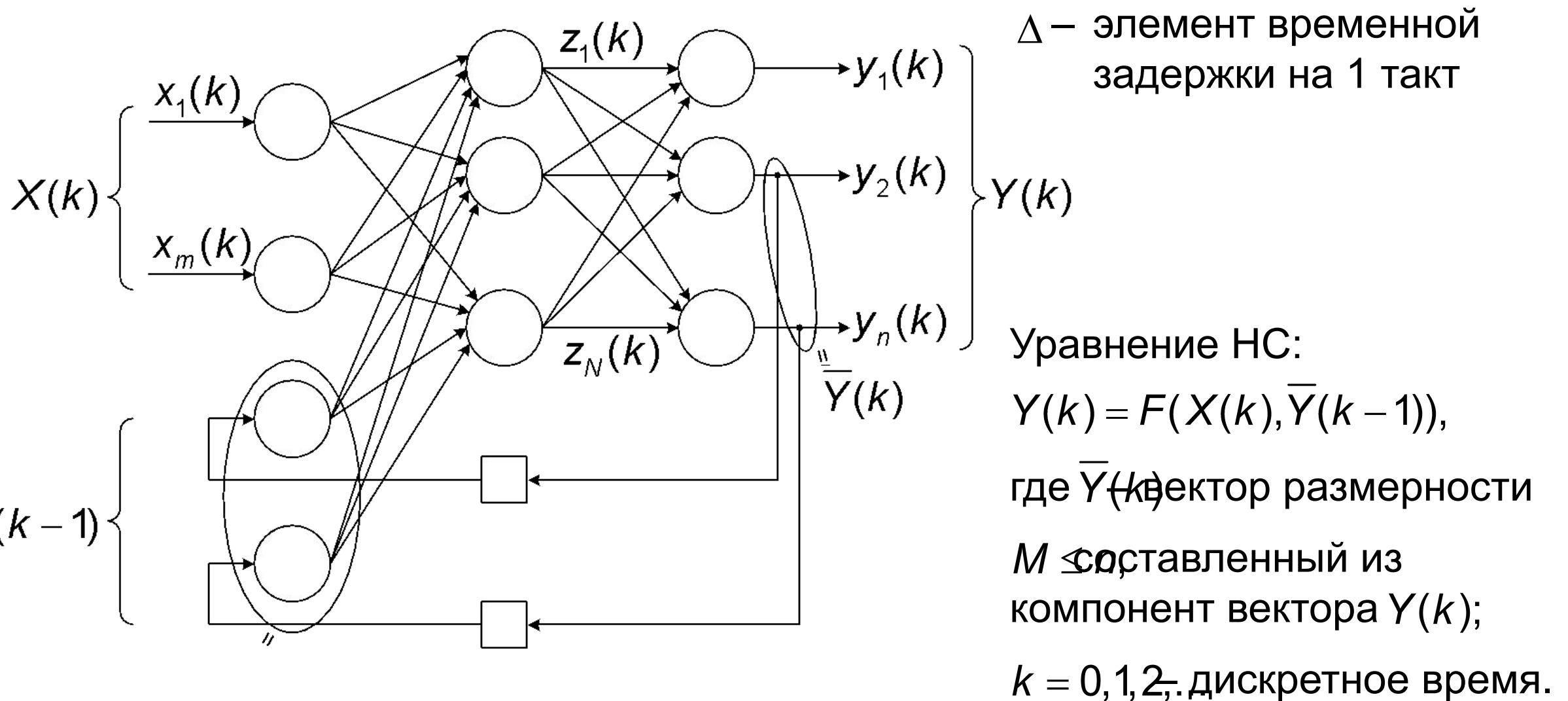
$RNN = \underline{TDNN}$
(Time Delay NN)

Уравнение НС ($k = 0, 1, 2, \dots$):

$$\underbrace{Y(k)}_{\text{выход НС}} = F(\underbrace{Y(k-1), \dots, Y(k-M)}_{\text{предыстория выхода НС}}, \underbrace{X(k)}_{\text{вход НС}}, \underbrace{X(k-1), \dots, X(k-L)}_{\text{предыстория входа НС}})$$

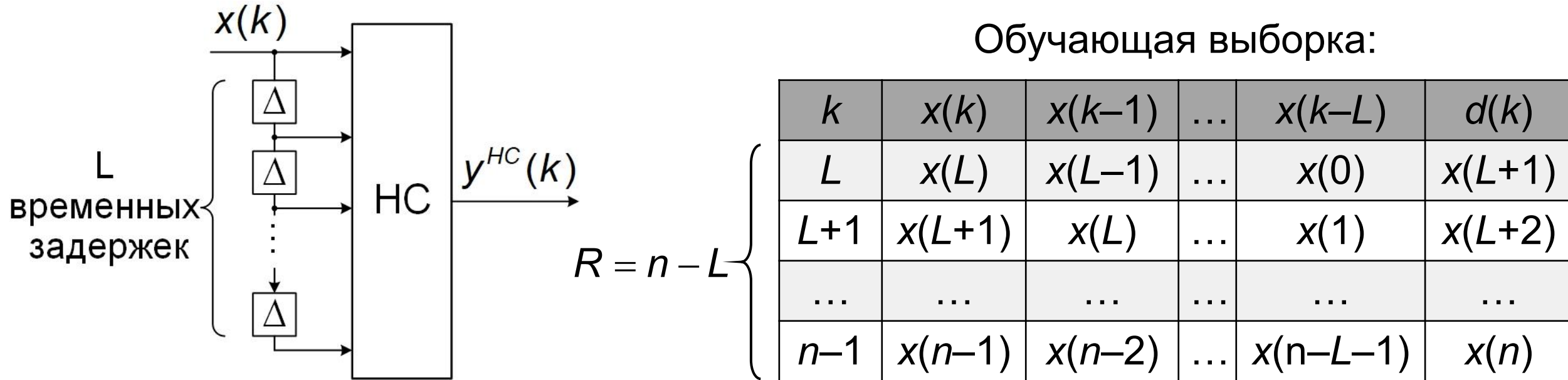
НС ДЖОРДАНА (Jordan M., 1986)

– НС с обратными связями с выходного слоя нейронов.



ПРОГНОЗИРОВАНИЕ ВРЕМЕННЫХ РЯДОВ

Решение: НС-предиктор на базе персептрона



СКО:
$$E = \sum_k \left[y^{HC}(k) - d(k) \right]^2$$

Точность прогноза:

$$\delta_{\text{прогн.}} = \frac{|y^{HC}(n) - x(n+1)|}{x(n+1)} \cdot 100\%$$

Количество нейронов:

$$m = L + 1; \quad n = 1; \quad N = \left\lceil \frac{R}{L + 2} \right\rceil$$

ГЛУБОКОЕ ОБУЧЕНИЕ (Deep Learning)

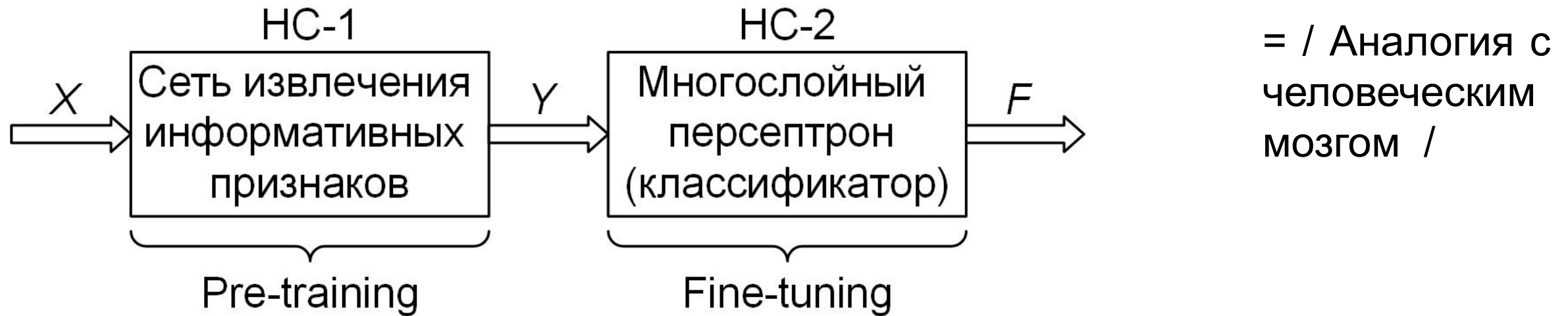
– это набор алгоритмов машинного обучения, которые моделируют множество уровней представления (абстракций) данных с использованием глубоких (многомерных) НС.

2005-2006 гг. – «Глубинная революция» в ИИ:

- Йошуа Бенджио (Университет Монреаля, Канада) – Автоенкодеры;
- Ян ЛеКун (Facebook) – Сверточные НС;
- Джеффри Хинтон (Университет Торонто, Канада) – Ограниченная машина Больцмана.

/ Глубокие НС = НС 3-го поколения /

ОСНОВНАЯ ИДЕЯ ГЛУБОКОГО ОБУЧЕНИЯ

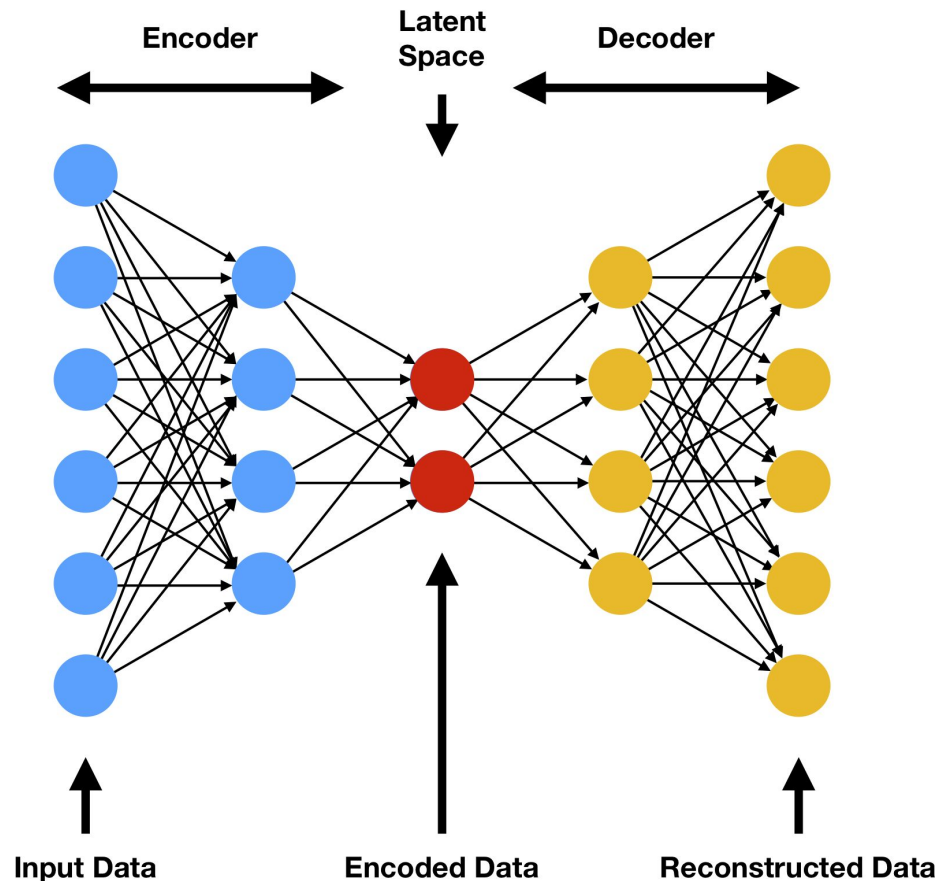


2 этапа обучения:

- 1) Извлекается «глубинная» информация о внутренней структуре входных данных (feature extraction);
- 2) Принимается решение о принадлежности входного вектора X определенному классу (classification).

ГЛУБОКИЕ НС НА ОСНОВЕ АВТОЕНКОДЕРА

Автоенкодер (автокодировщик, autoencoder) – это НС, которая пытается максимально приблизить значение выходного сигнала к входному, т.е. наилучшим образом аппроксимировать тождественное преобразование.



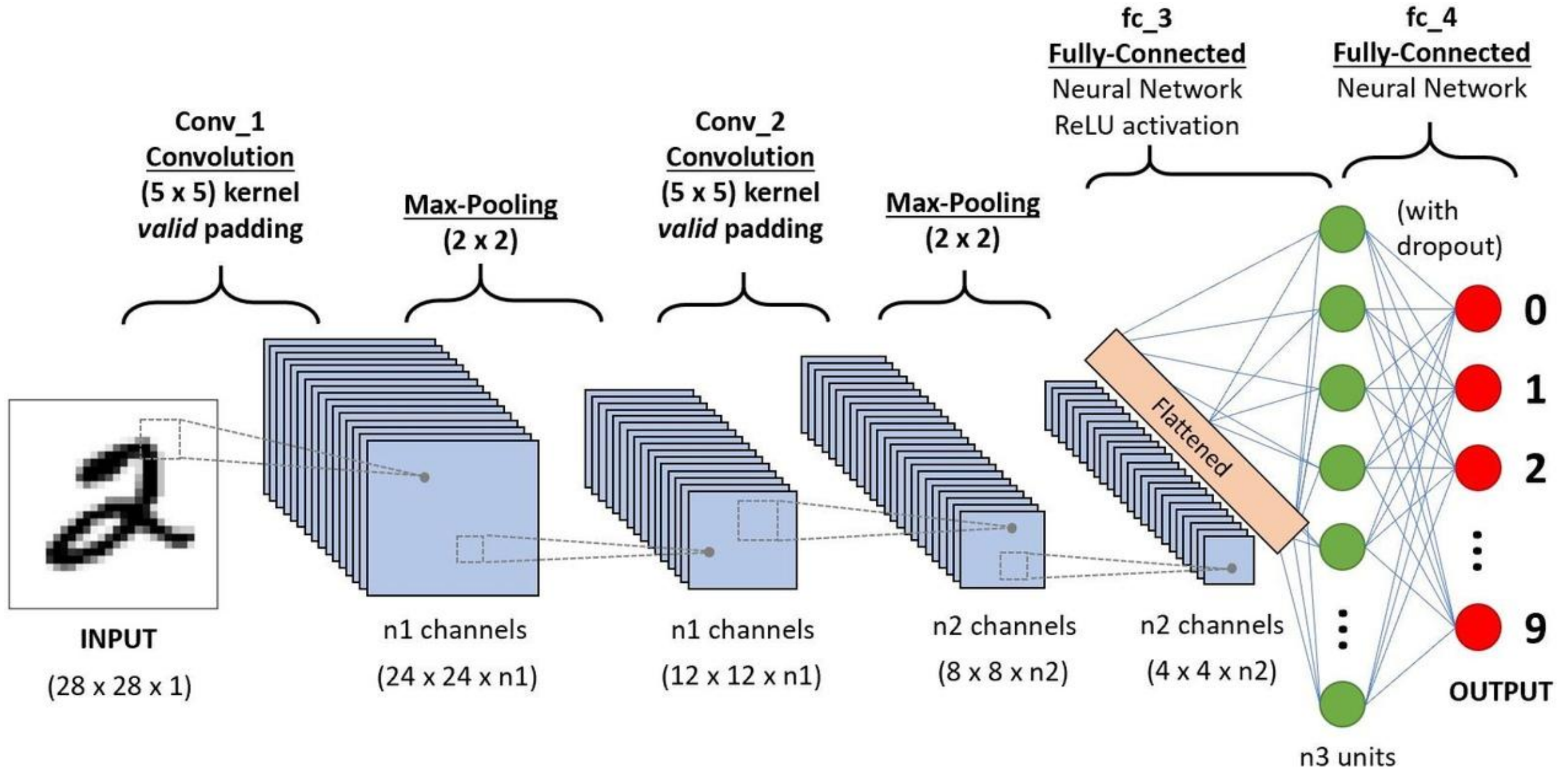
Цель обучения:

$$|Y - X| \leq \epsilon,$$

где X – входной вектор; Y – выходной вектор; ϵ – малая величина.

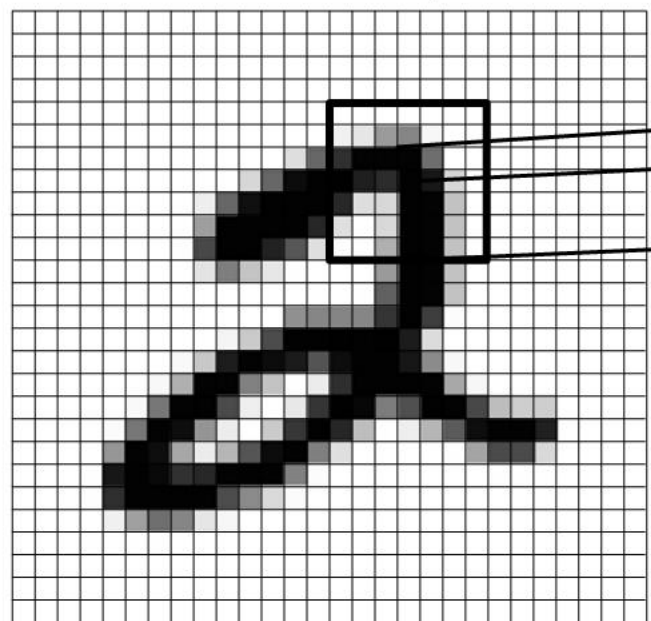
Центральный скрытый слой («bottleneck») – выделяет наиболее существенные признаки входного объекта (образа).

АРХИТЕКТУРА СВЕРТОЧНОЙ НС

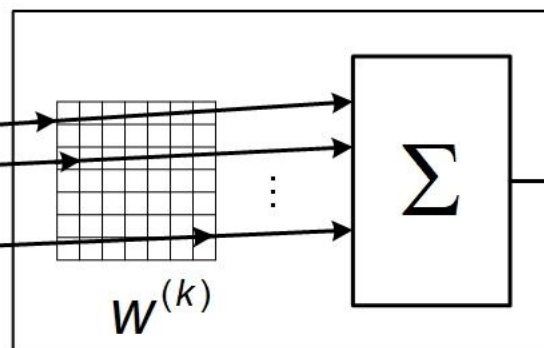


ОПЕРАЦИЯ СВЕРТКИ (Convolutional)

Исходное изображение



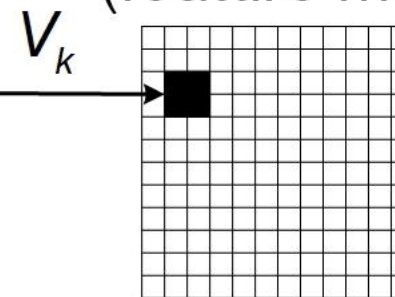
28x28



Фильтр
(шаг = 2)

$$V_k = \sum_{i,j=1}^5 w_{ij}^{(k)} x_{ij}$$

Карта признаков
(feature map)



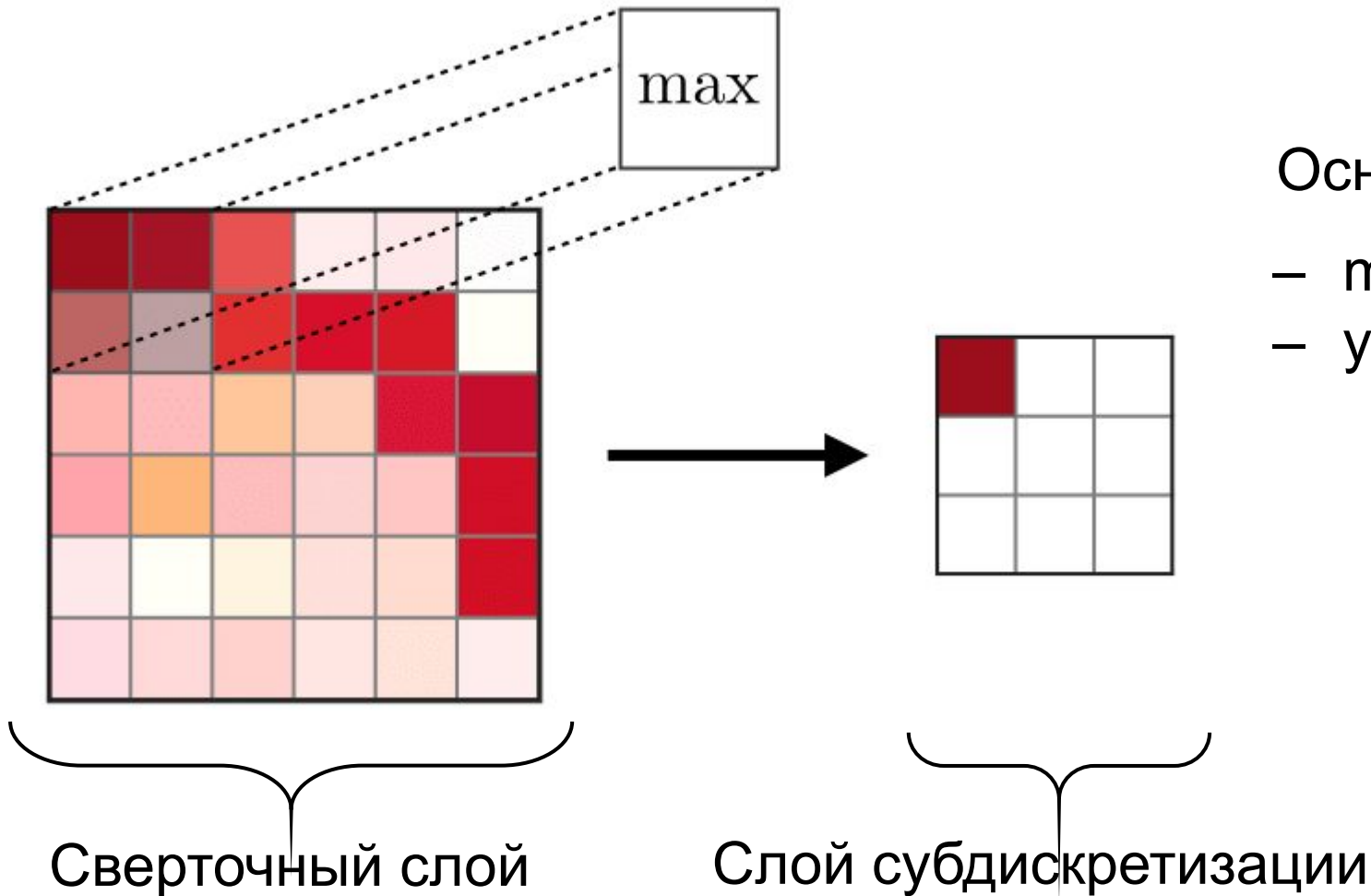
13x13

$w^{(k)} = \|w_{ij}^{(k)}\|_{5 \times 5}$ — ядро свертки

6 фильтров $\rightarrow$ 6 карт признаков ($k = 1, 2, \dots, 6$).

Каждая карта признаков имеет свое ядро свертки (свой фильтр).

СЛОЙ СУБДИСКРЕТИЗАЦИИ (Subsampling / pooling layer)



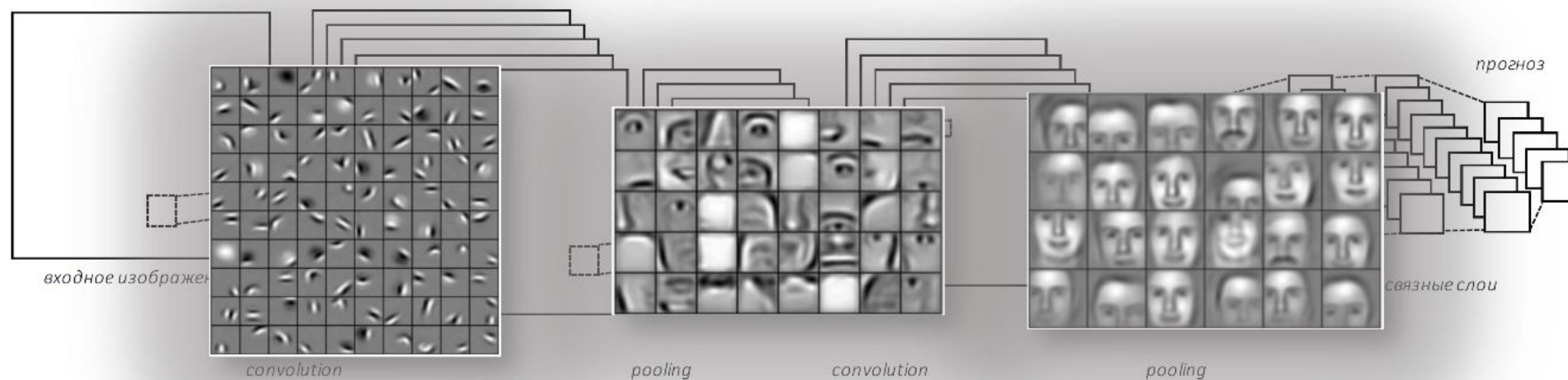
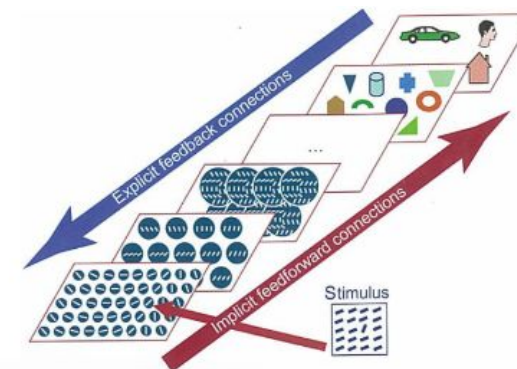
Основные операции:

- max;
- усреднение.

Цель: снижение размерности карт признаков.

ВИЗУАЛИЗАЦИЯ СЛОЕВ СВЕРТОЧНОЙ (иерархия признаков)

Распознавание



ЛИТЕРАТУРА

(по глубокому обучению)

1. Николенко С., Кадурин А., Архангельская Е. Глубокое обучение. Погружение в мир нейронных сетей. – СПб.: Питер, 2019. – 480 с.
2. Траск Э. Грокаем глубокое обучение / Пер. с англ. – СПб.: Питер, 2019. – 352 с.
3. Головки В.А. От многослойных персептронов к нейронным сетям глубокого доверия: парадигмы обучения и применения // Лекции по нейроинформатике. – М., 2017. – С. 47-84.
4. Визильтер Ю.В. Глубокое обучение и глубокая оптимизация: современное состояние и перспективы развития технологий ИИ // Семинар НИУ ВШЭ по высокопроизводительным вычислениям. – М.: ГосНИИАС, 2020. (Презентация / 80 слайдов).