

Введение в нейронные сети. Основы обучения нейронных сетей.



GeekBrains

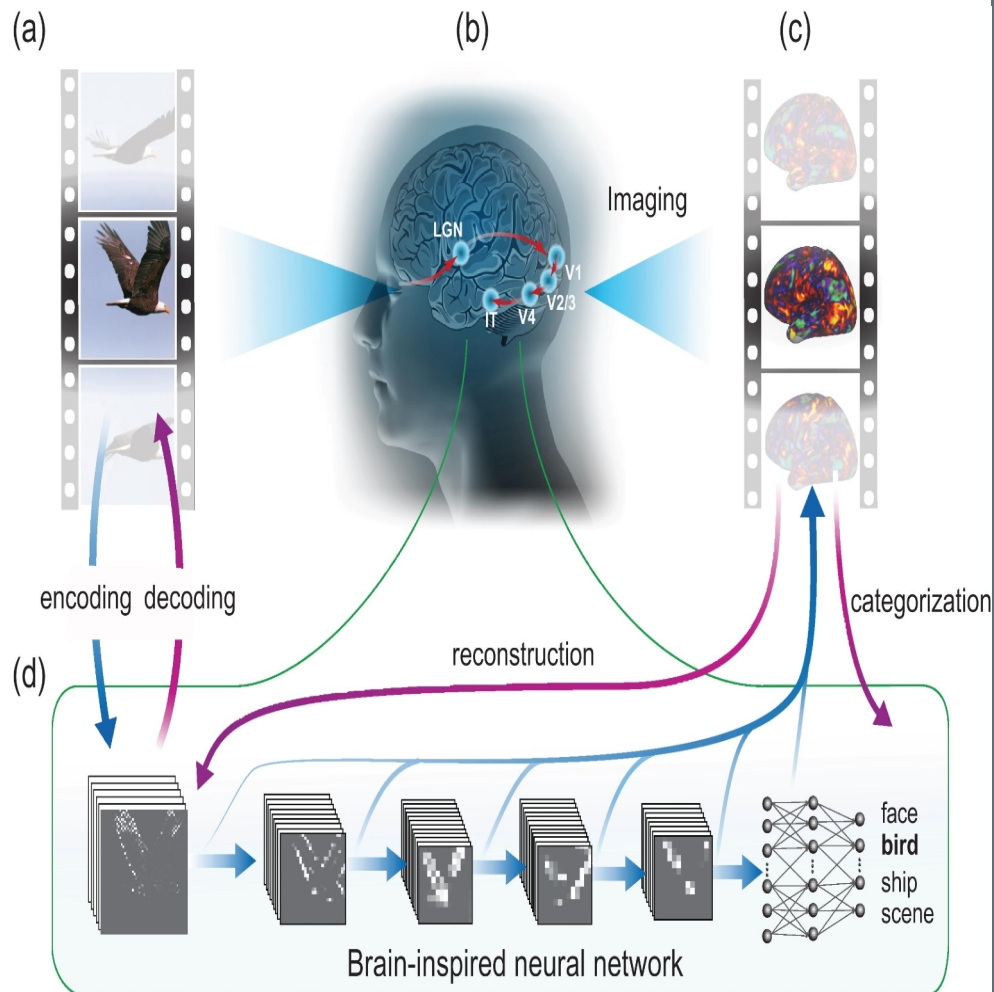
План вебинара



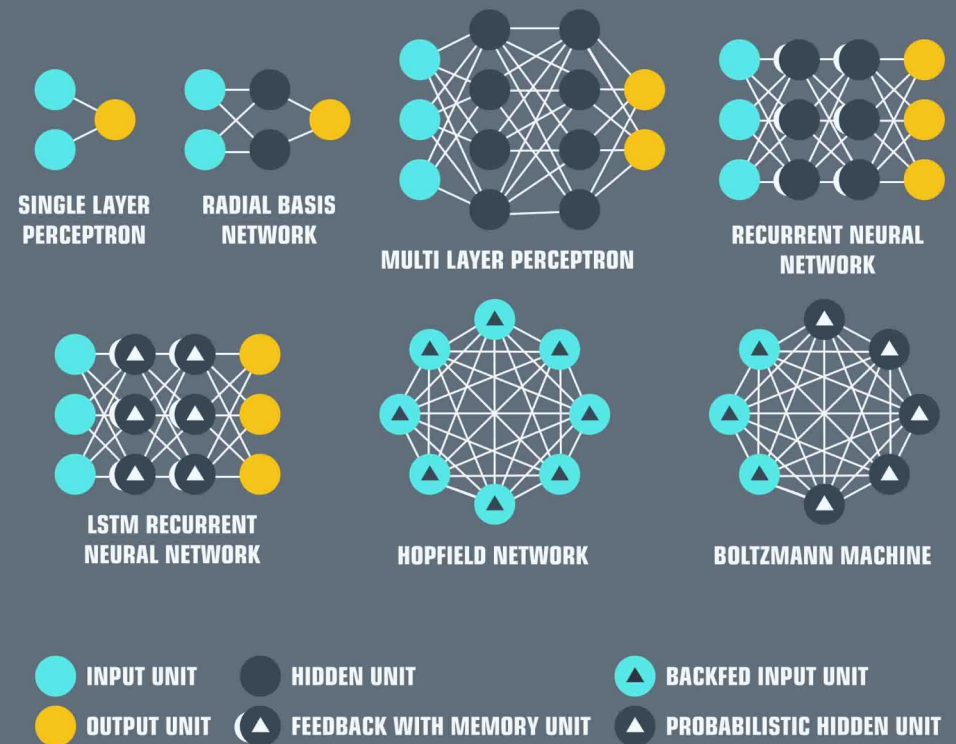
1. Общие сведения о нейронных сетях.
2. История
3. Основы обучения нейронных сетей
4. Практика



Общие сведения



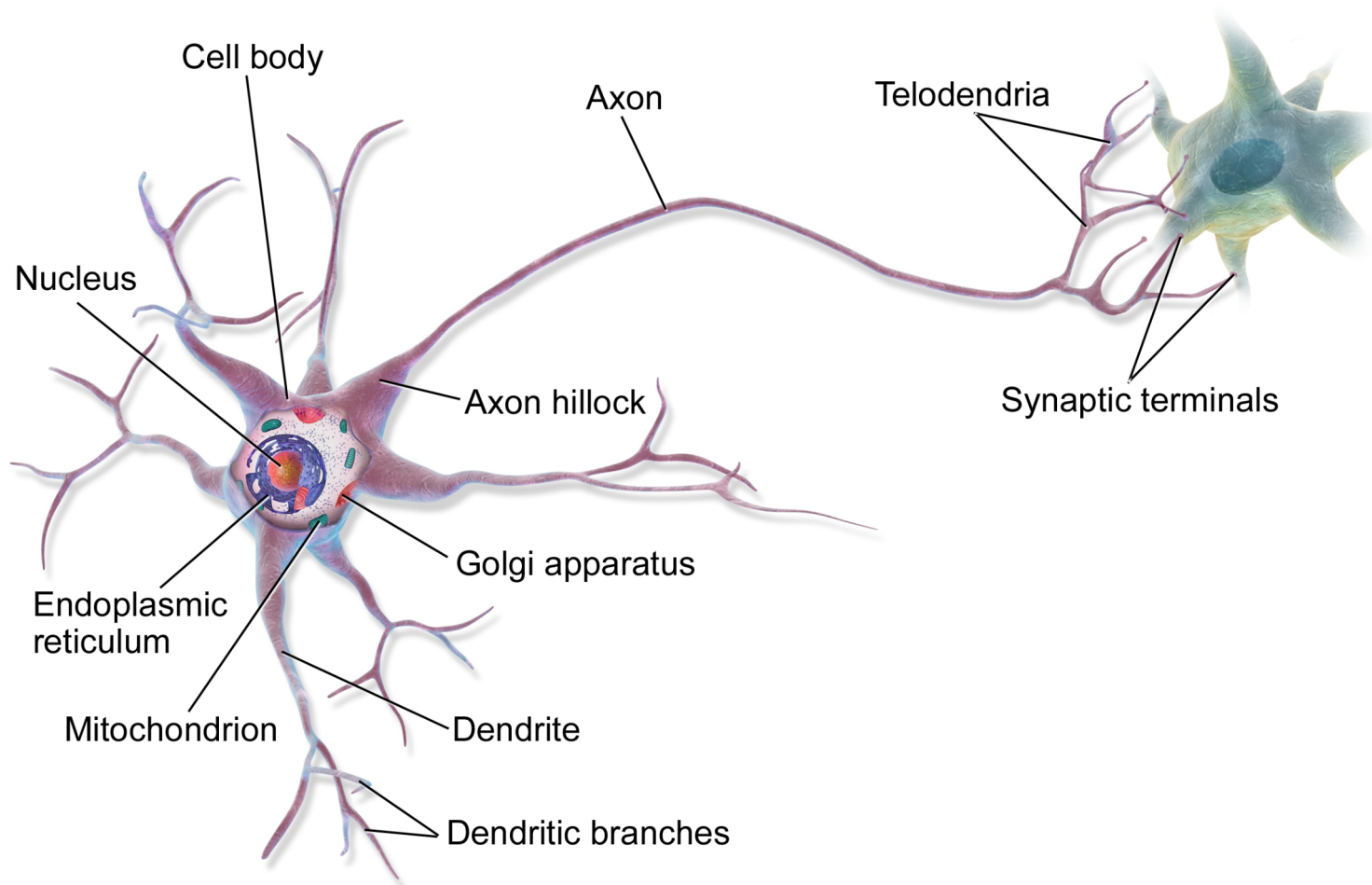
NEURAL NETWORK ARCHITECTURE TYPES



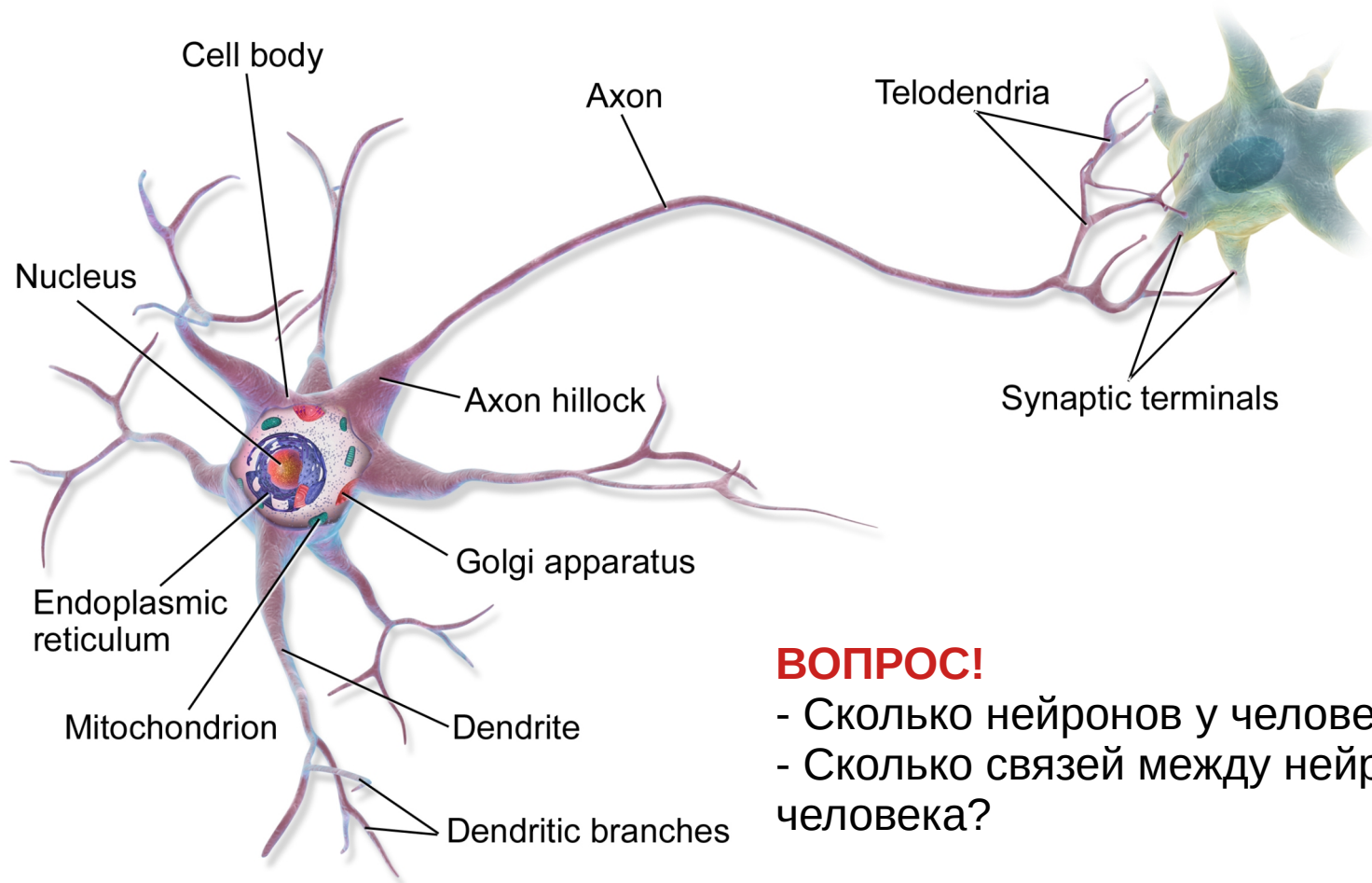
Место искусственных нейронных сетей в современном мире



Биологическая нейронная сеть



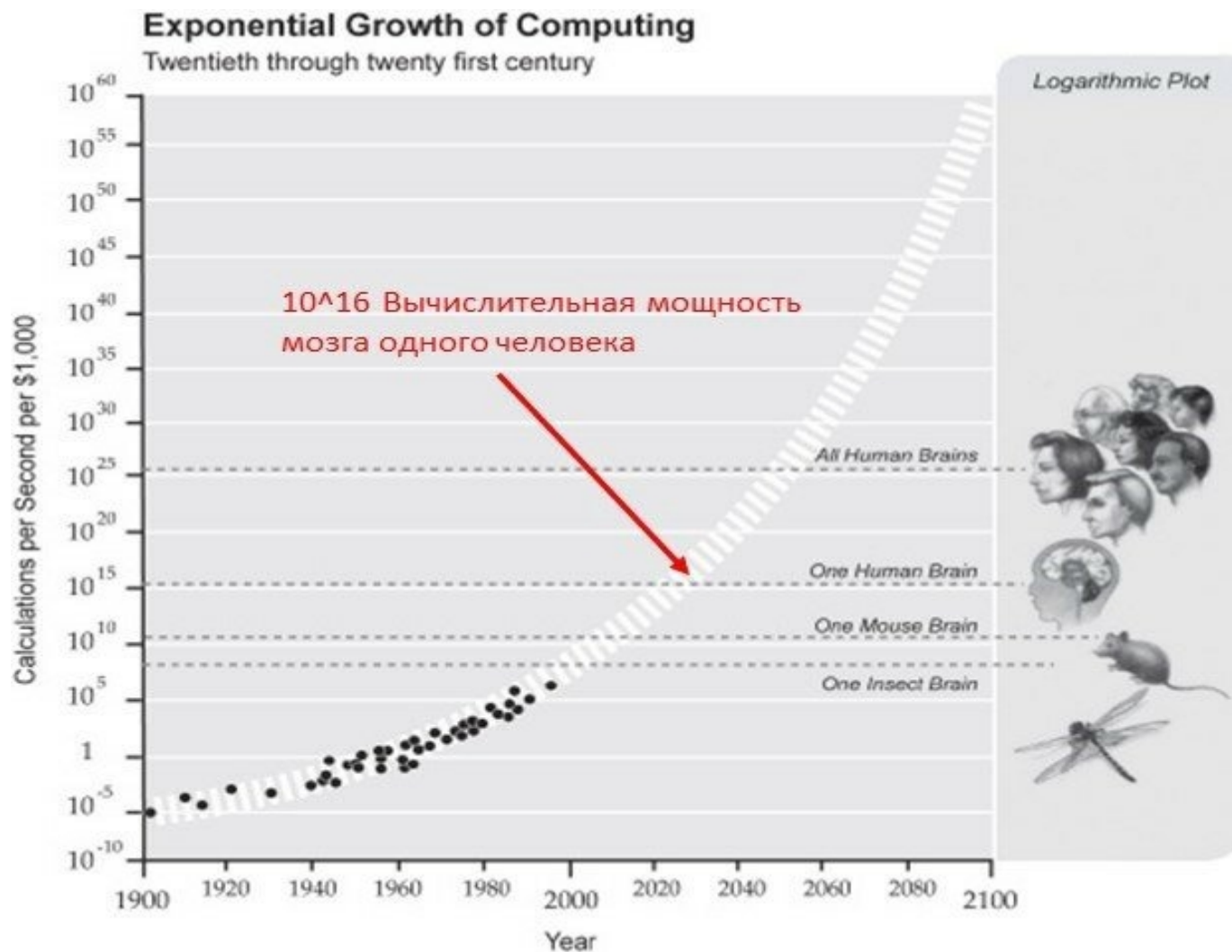
Биологическая нейронная сеть



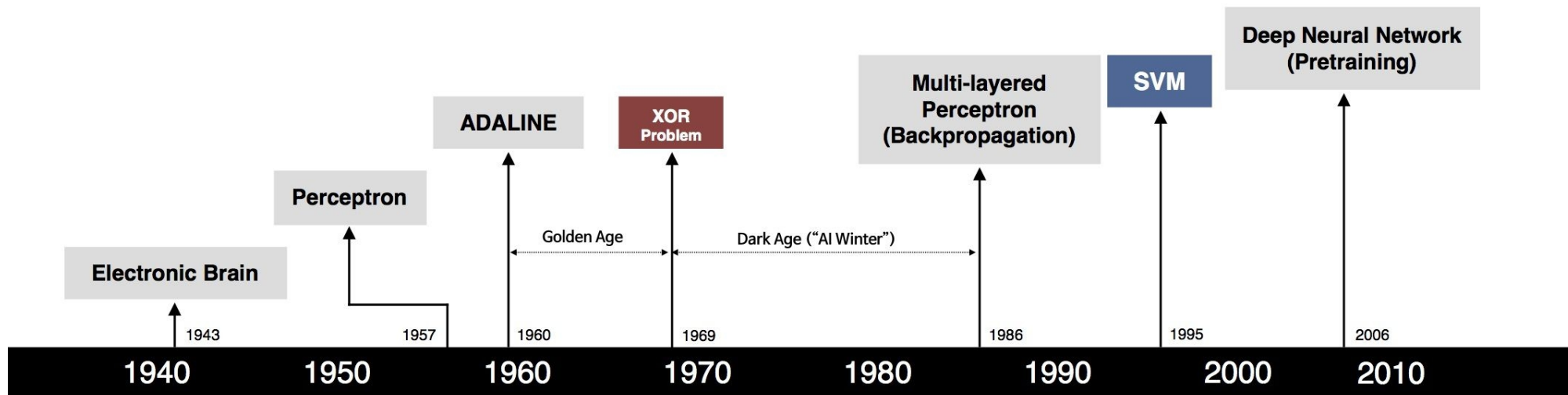
ВОПРОС!

- Сколько нейронов у человека?
- Сколько связей между нейронами у человека?

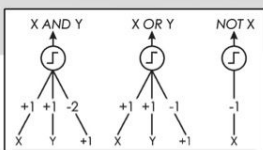
Биологическая нейронная сеть



История нейронных сетей



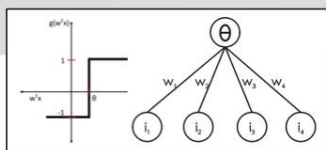
S. McCulloch – W. Pitts



- Adjustable Weights
- Weights are not Learned



F. Rosenblatt



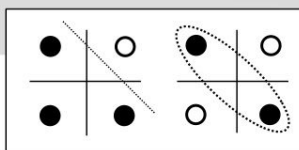
- Learnable Weights and Threshold



B. Widrow – M. Hoff



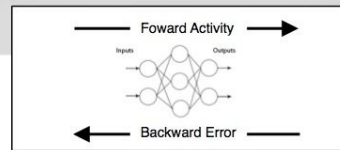
M. Minsky – S. Papert



- XOR Problem



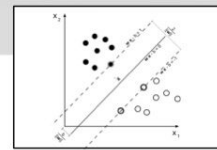
D. Rumelhart – G. Hinton – R. Williams



- Solution to nonlinearly separable problems
- Big computation, local optima and overfitting



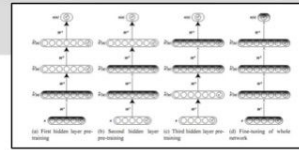
V. Vapnik – C. Cortes



- Limitations of learning prior knowledge
- Kernel function: Human Intervention



G. Hinton – S. Ruslan

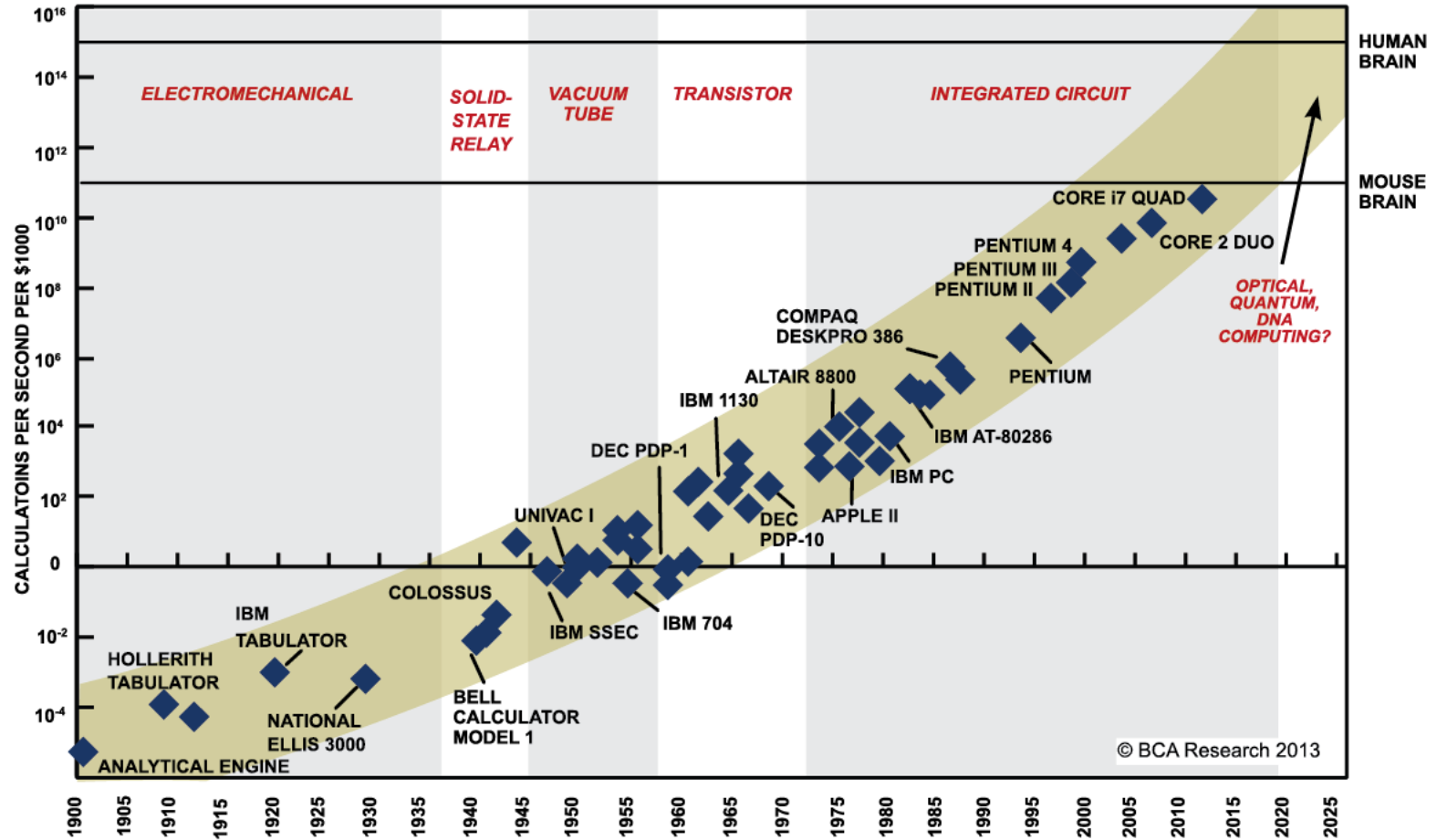


- Hierarchical feature Learning

Закон Мура



GeekBrains

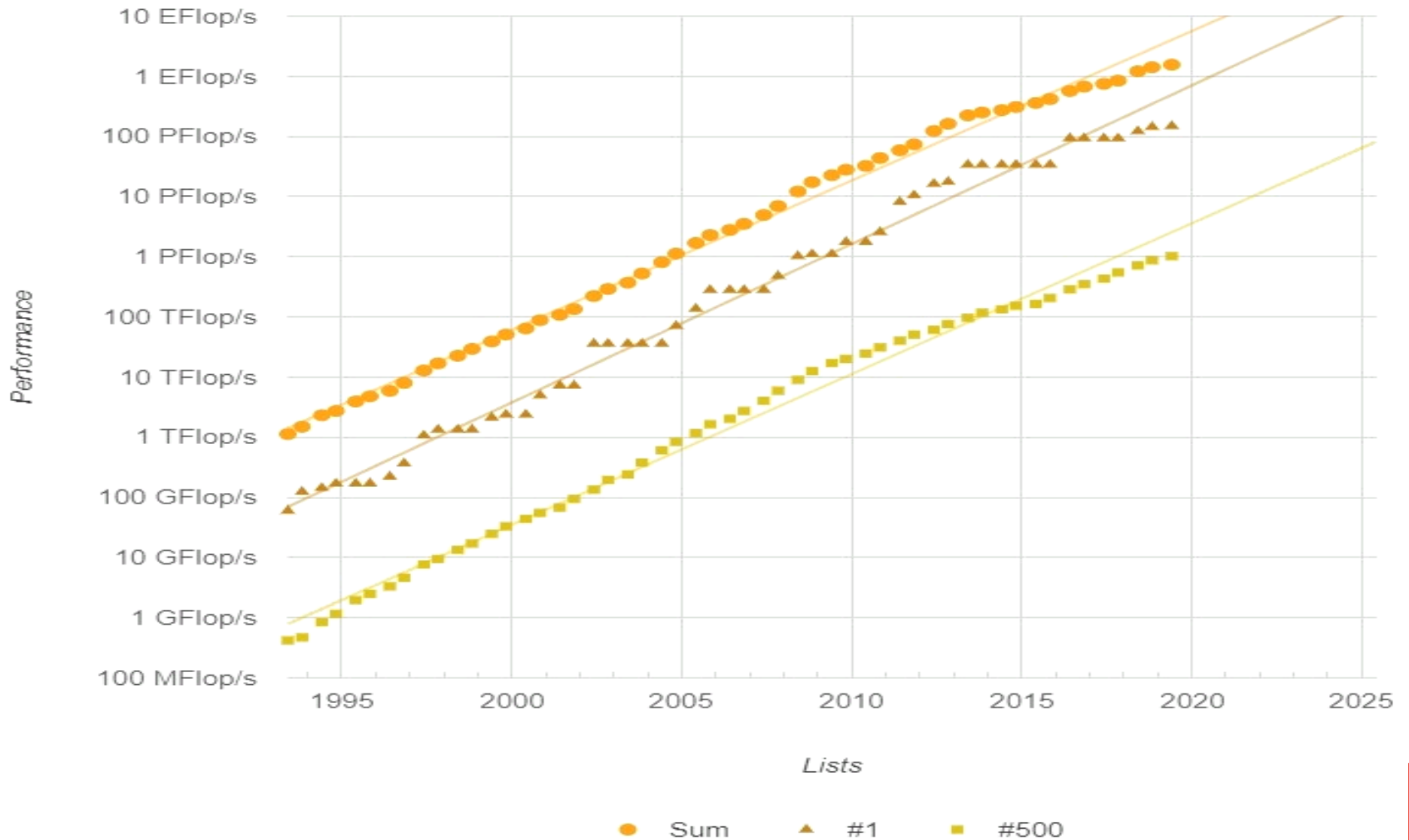


SOURCE: RAY KURZWEIL, "THE SINGULARITY IS NEAR: WHEN HUMANS TRANSCEND BIOLOGY", P.67, THE VIKING PRESS, 2006. DATAPOINTS BETWEEN 2000 AND 2012 REPRESENT BCA ESTIMATES.

Производительность для Top500

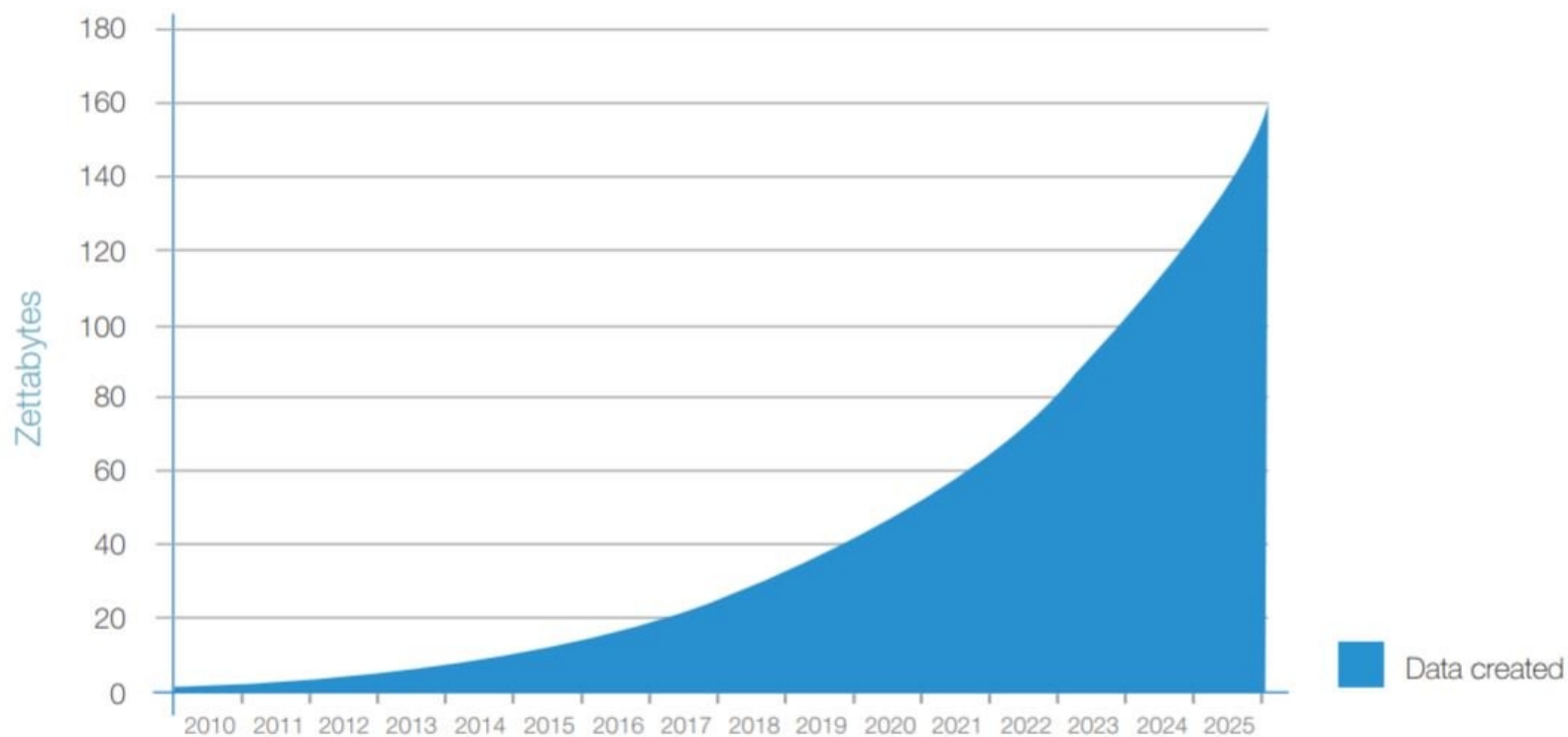


GeekBrains



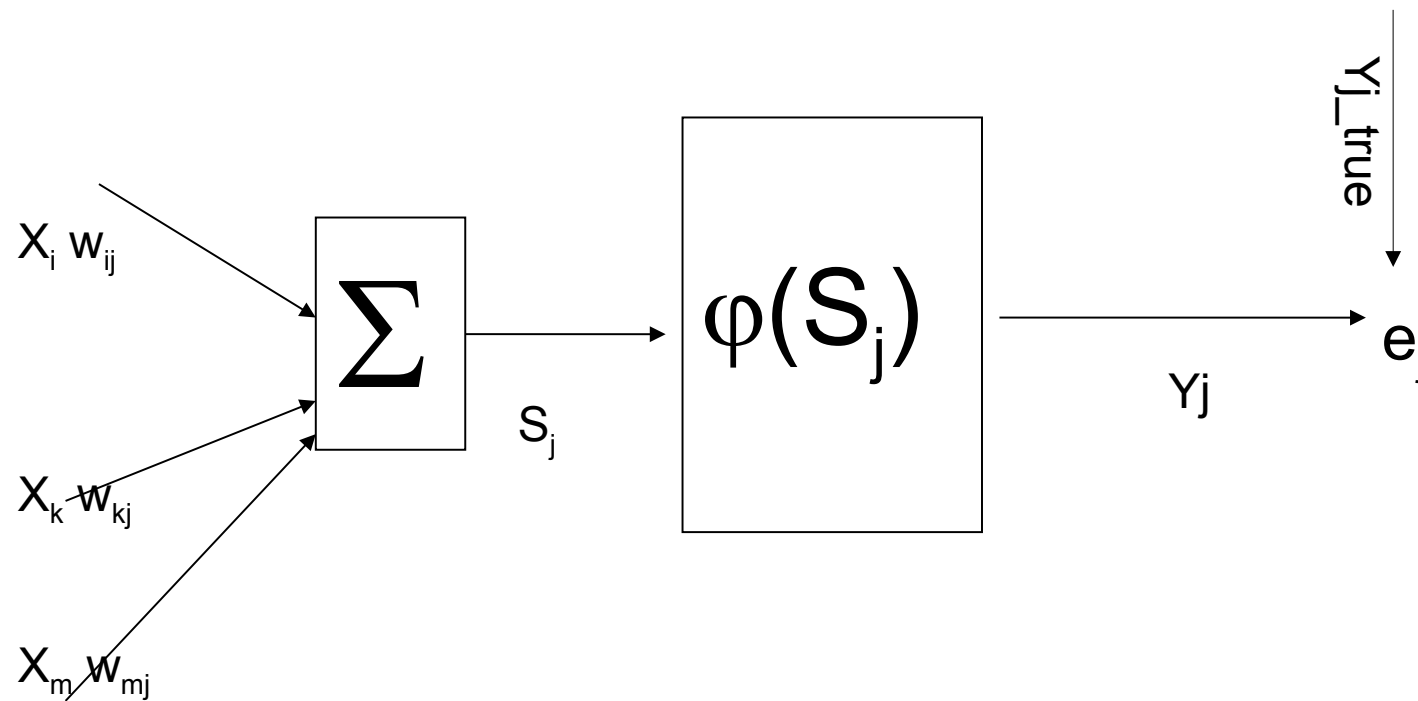
Рост числа данных

Figure 2. Annual Size of the Global Datasphere



Source: IDC's Data Age 2025 study, sponsored by Seagate, April 2017

Нейрон



$$S_j = \text{sum}(X_i w_{ij})$$

$$Y_j = \varphi(S_j)$$

Обучение нейронных сетей

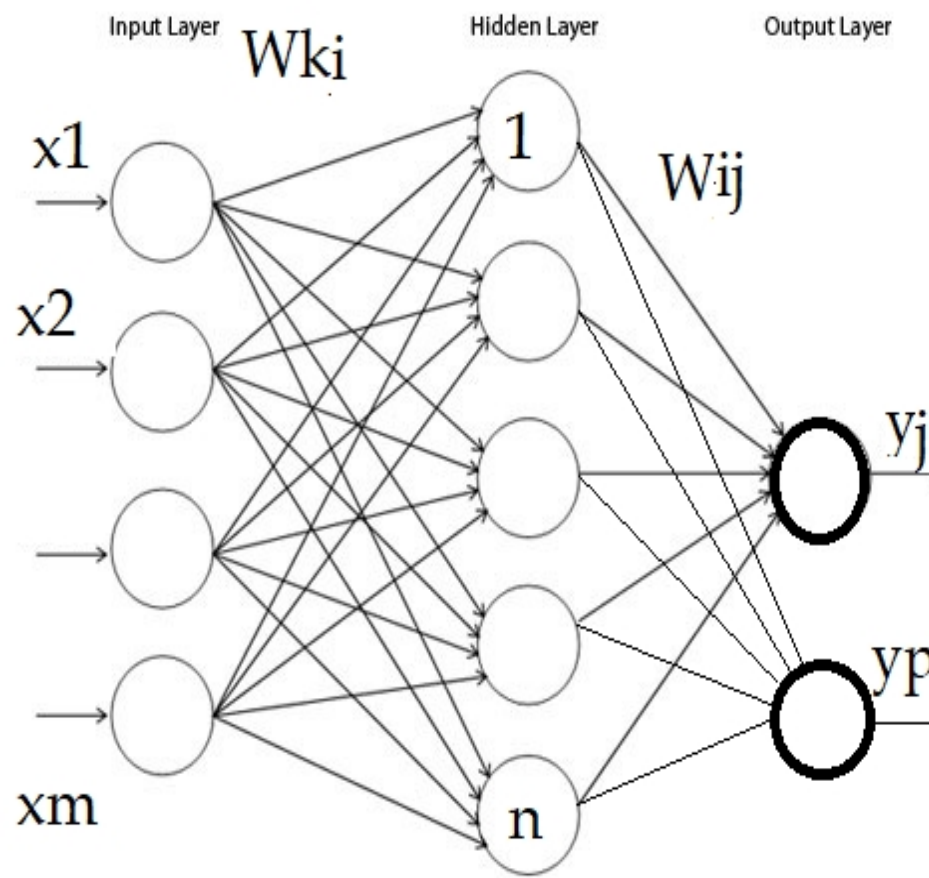


GeekBrains

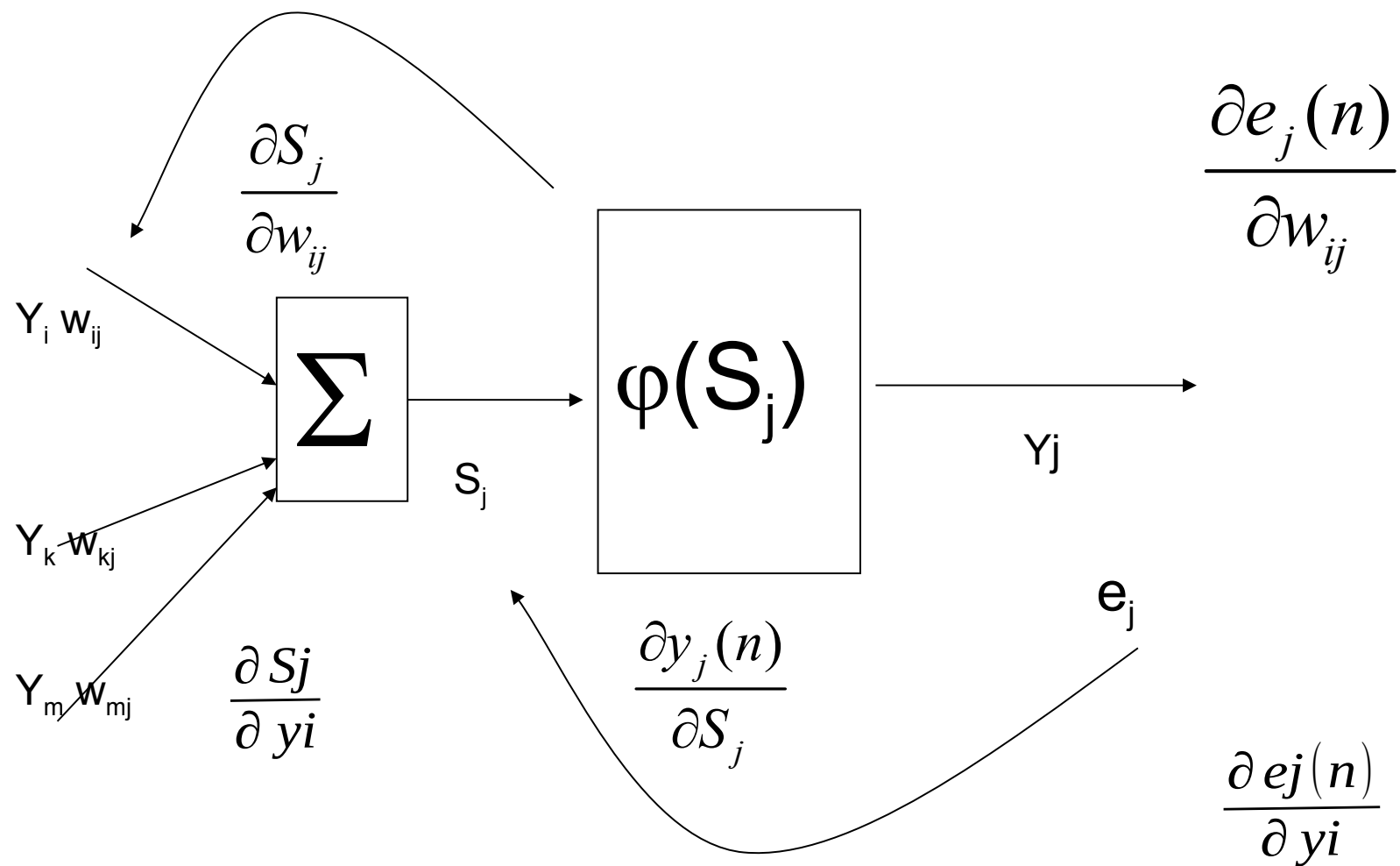
$$Q(W, X) = \sum_{i=1, N} L(X_i, W)$$

$$L(X_i, W) = (y_i(X, W) - \hat{y})^2$$

$$y_j = f_j \left(\sum_{i=0, n} w_{ij} f_i \left(\sum_{k=0, m} w_{ki} x_k \right) \right)$$



Нейрон j



Математика BackProp

шаг коррекции

$$\frac{\partial E(n)}{\partial w_{ij}(n)} = \frac{\partial E(n)}{\partial e_j(n)} \frac{\partial e_j(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial S_j(n)} \frac{\partial S_j(n)}{\partial w_{ij}(n)}$$

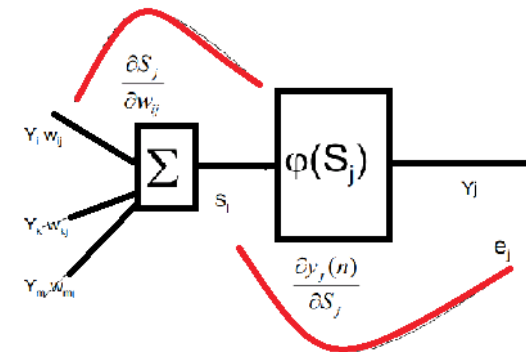
$$E(n) = \frac{1}{2} \sum_{j=1}^k e_j(n)^2$$

$$\frac{\partial E(n)}{\partial e_j(n)} = e_j(n) = d_j - y_j(n), \quad \frac{\partial e_j(n)}{\partial y_j(n)} = -1$$

$$y_j(n) = \varphi(S_j) \Rightarrow \frac{\partial y_j(n)}{\partial S_j(n)} = \varphi'(S_j(n))$$

$$S_j(n) = \sum w_{ij} y_i(n) \Rightarrow \frac{\partial S_j(n)}{\partial w_{ij}(n)} = y_i(n)$$

$$\frac{\partial E(n)}{\partial w_{ij}(n)} = -e_j(n) \varphi'(S_j(n)) y_i(n) \Rightarrow \Delta w_{ij}(n) = -\eta \frac{\partial E(n)}{\partial w_{ij}(n)}$$



Математика BackProp

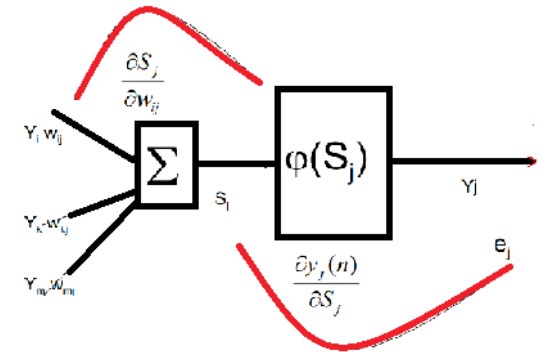
обратный шаг

$$E(n) = \frac{1}{2} \sum_k e_k^2(n)$$

$$\frac{\partial E(n)}{\partial y_j(n)} = \sum_k e_k \frac{\partial e_k(n)}{\partial y_j(n)}$$

$$\delta_{ij}(n) = \frac{\partial e_j}{\partial y_i} = \frac{\partial e_j(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial S_j(n)} \frac{\partial S_j(n)}{\partial y_i} = -\varphi'_j(S_j(n)) w_{ij}$$

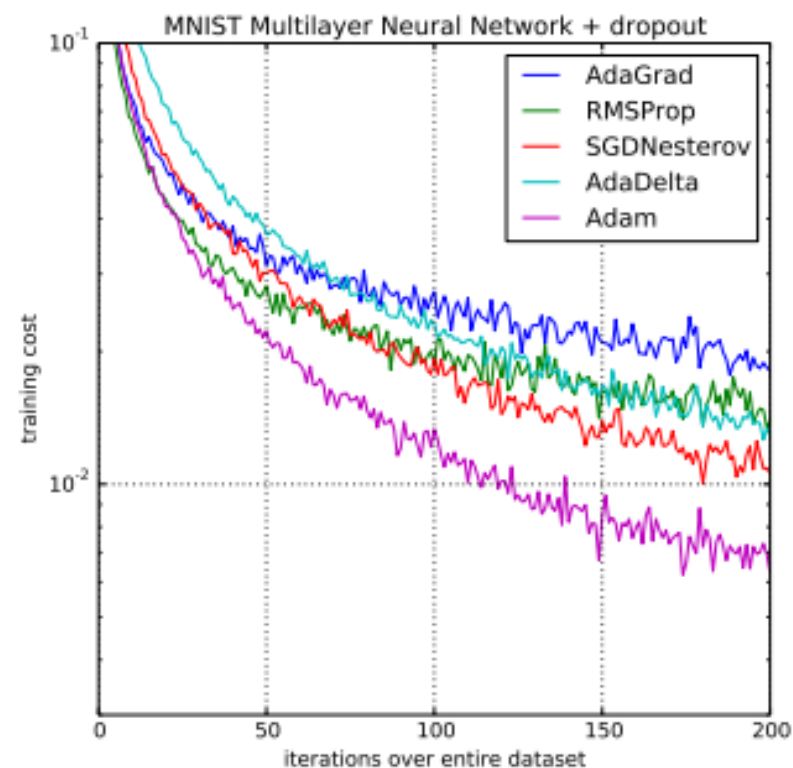
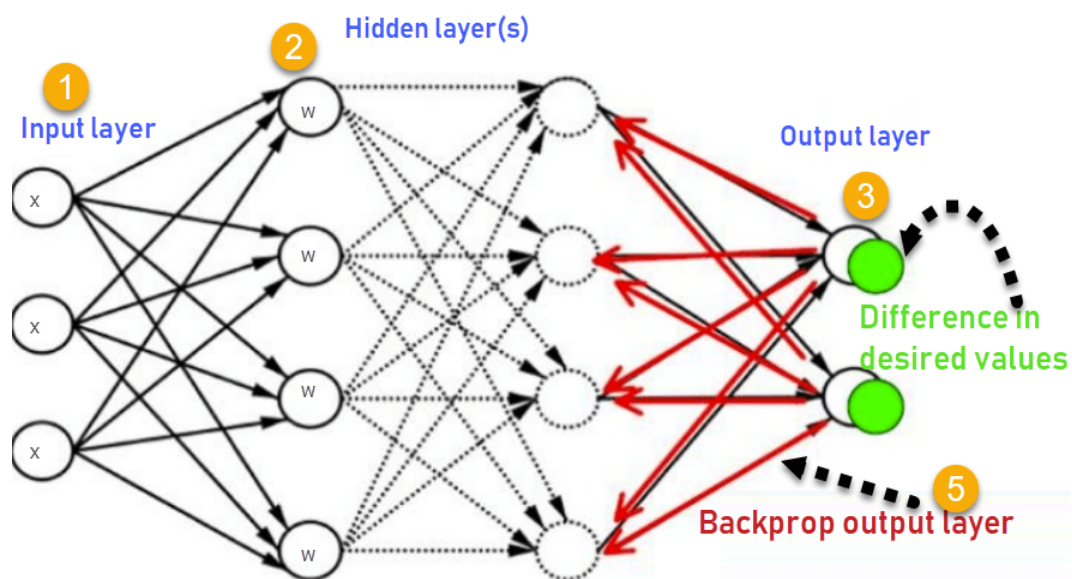
$$e_i = \sum_k e_k \delta_{ik}(n)$$



Обучение нейронных сетей



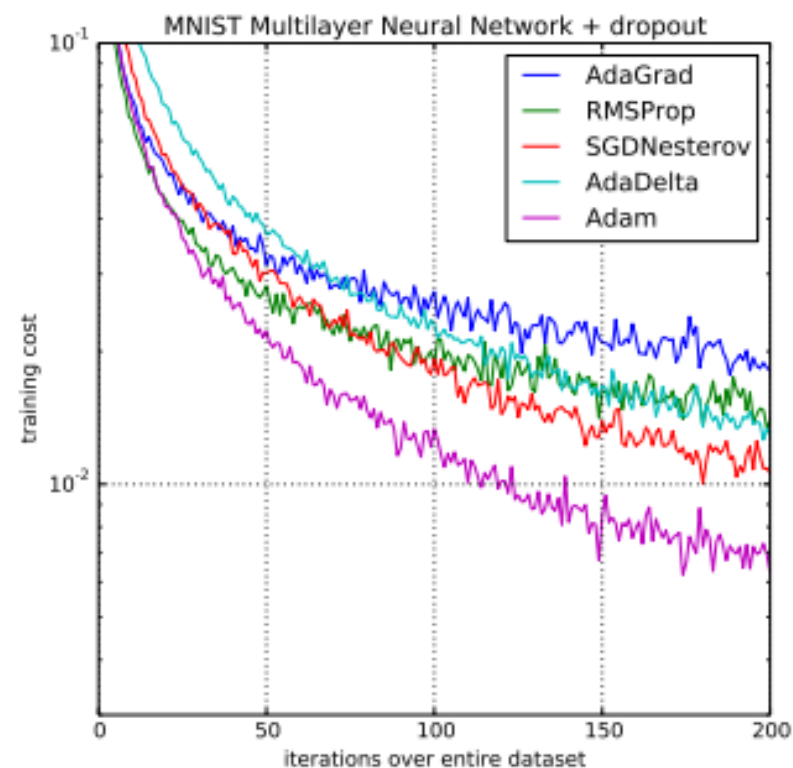
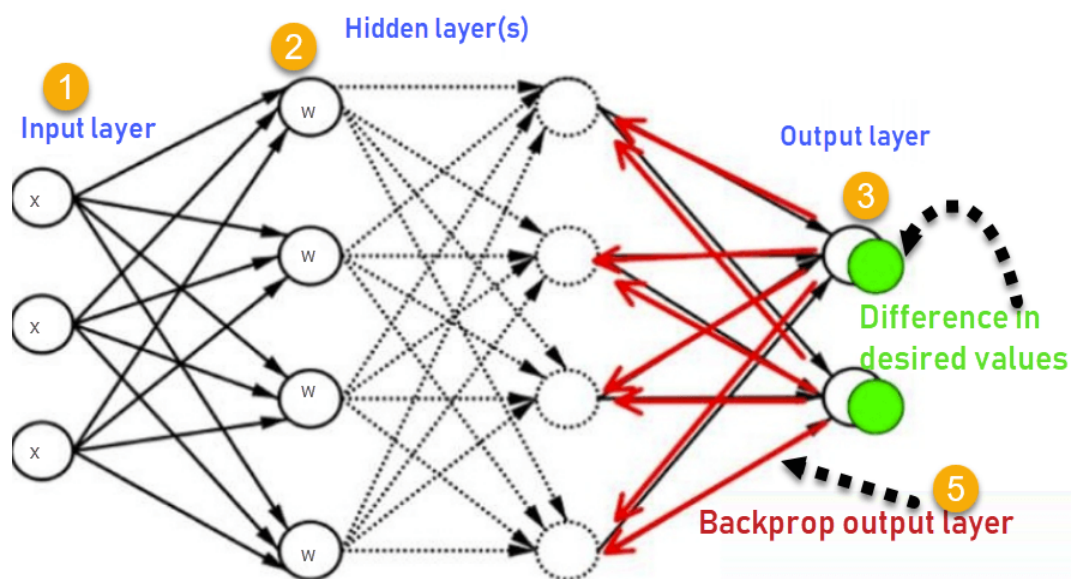
GeekBrains



Обучение нейронных сетей



GeekBrains



ВОПРОС!

Сколько в этой сети
настраиваемых параметров?

Вопросы



GeekBrains



Практическое задание



1. Попробуйте видоизменить параметры разобранный на уроке нейронной сети таким образом, чтобы улучшить ее точность.

Проведите анализ:

Что приводит к ухудшению точности нейронной сети? Что приводит к увеличению ее точности?

