

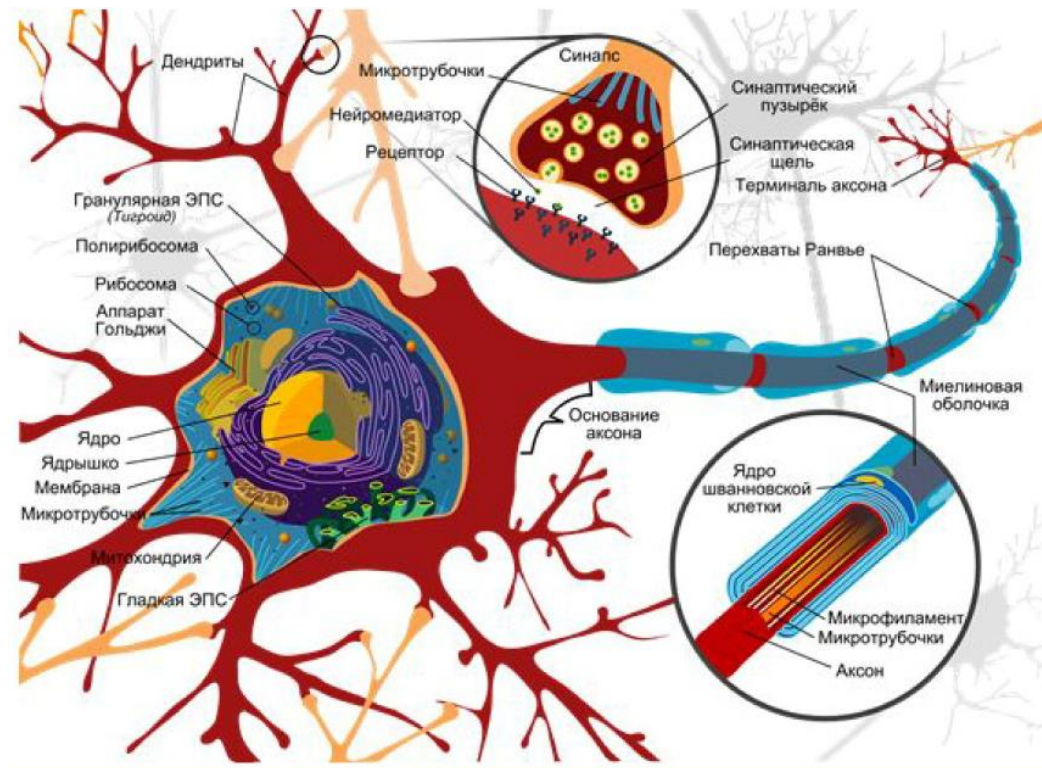
10 misconceptions about Neural Networks

[Stuart Reid](#) | May 8, 2014

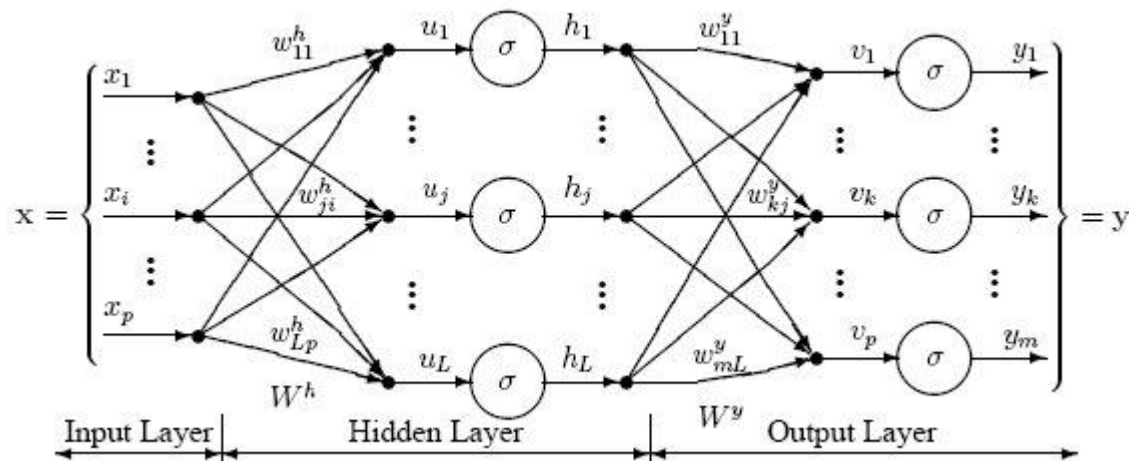
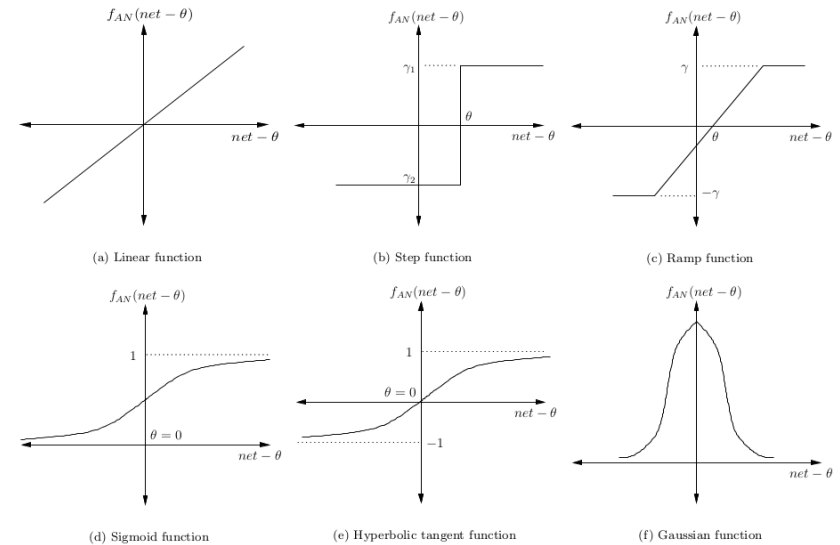
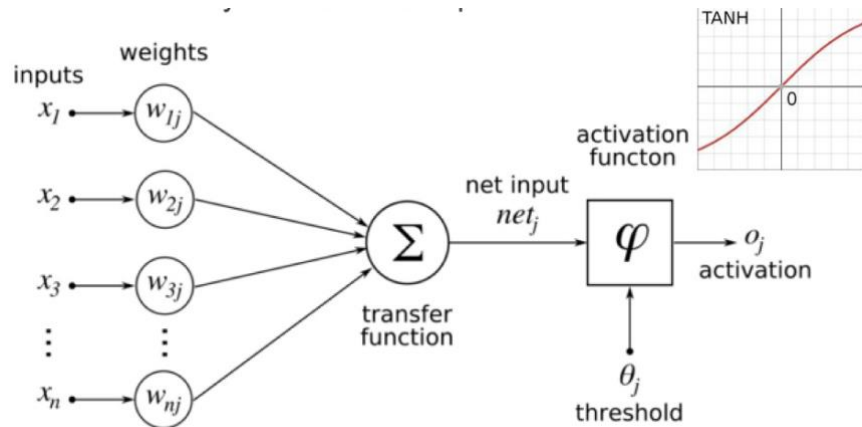
Синиця Анастасія | ядерна енергетика

Neural networks are not models of the human brain

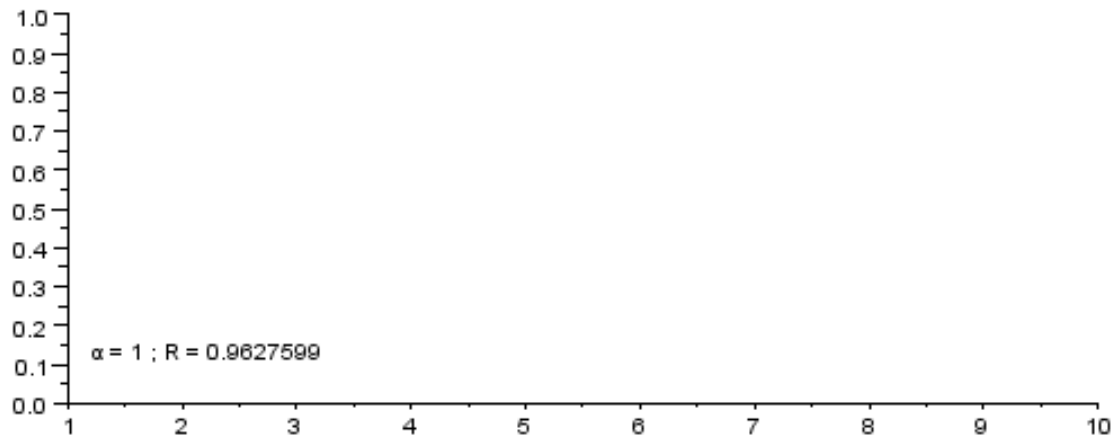
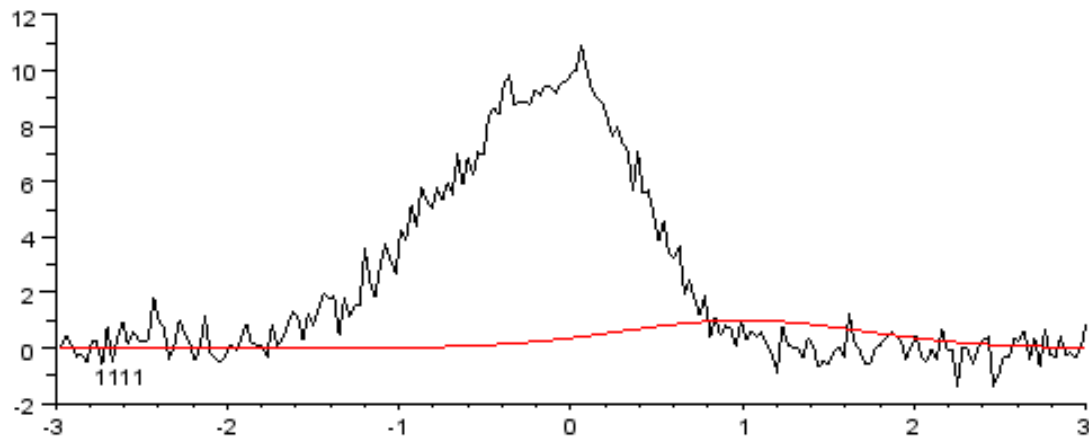
- distributed representation theory
- grandmother cell theory



Neural networks aren't a "weak form" of statistics



Neural networks are not models of the human brain



Learning Rules

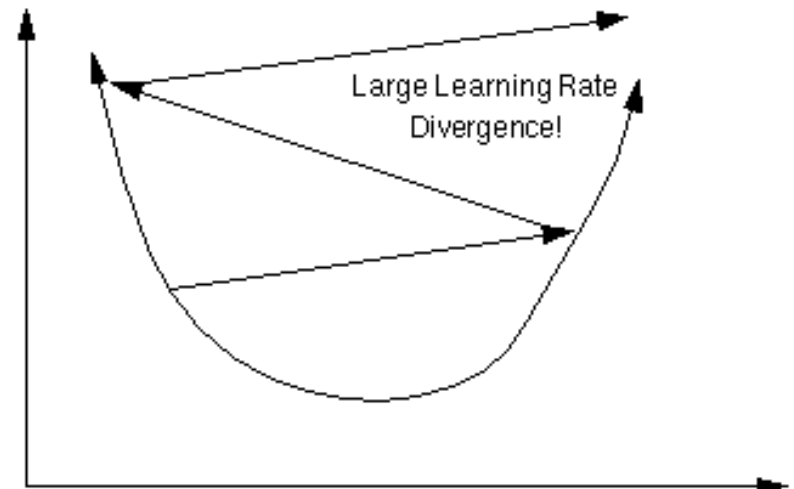
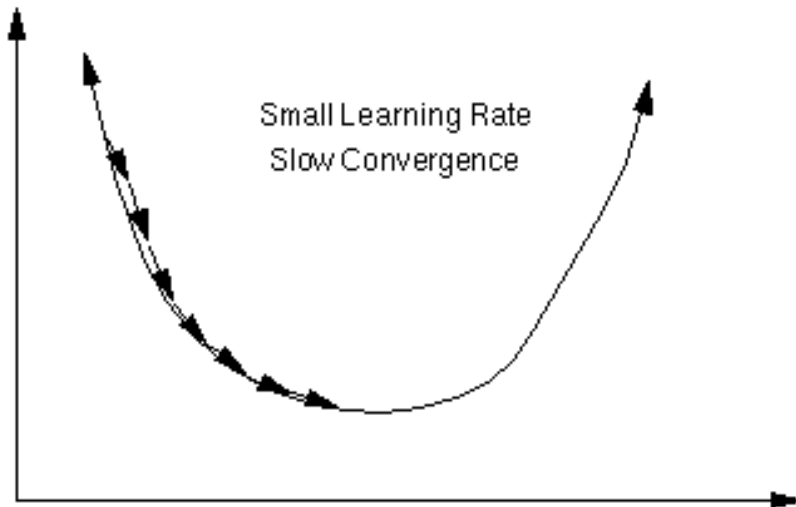
- ▶ Мета: мінімізувати похибку
- ▶ Метод градієнтного спуску

$$\sum_{p=1}^{P_T} (t_p - o_p)^2$$

$$v_i(t) = v_i(t - 1) + \delta v_i(t)$$

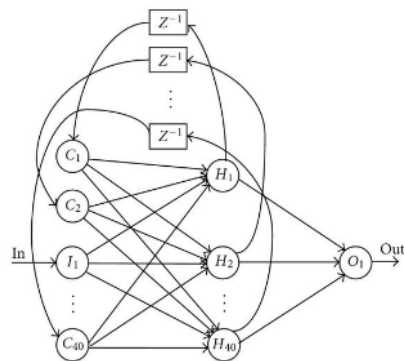
$$\delta v_i(t) = \eta \left(-\frac{\partial \epsilon}{\partial v_i} \right)$$

$$\frac{\partial \epsilon}{\partial v_i} = -2(t_p - o_p) \frac{\partial f}{\partial \text{net}_p} z_{i,p}$$

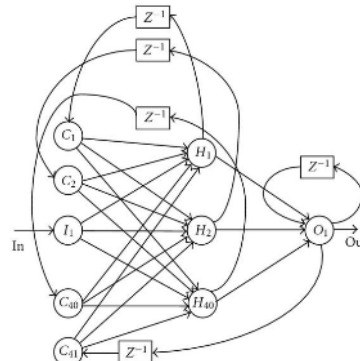


Neural networks come in many architectures

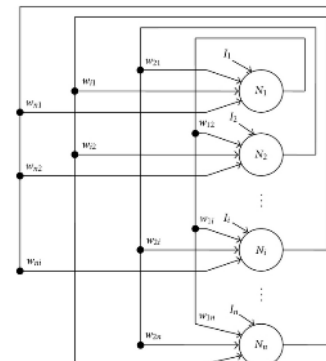
► Recurrent Neural Networks



Elman Neural Network

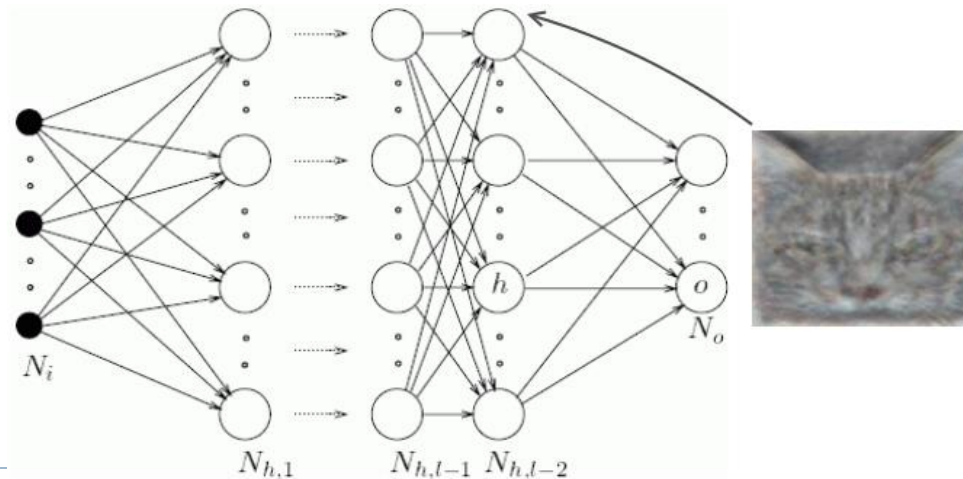


Jordan Neural Network

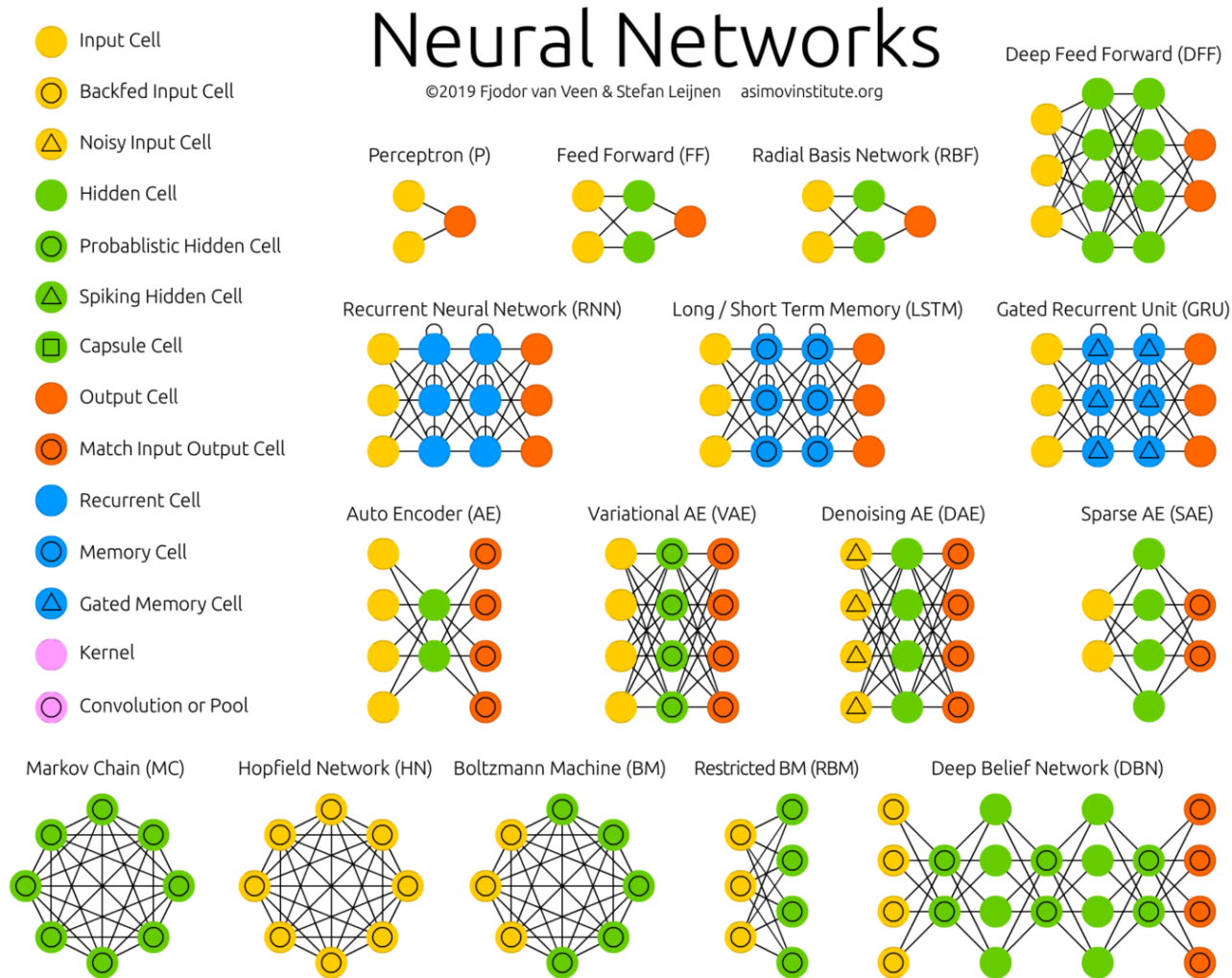


Hopfield Neural Network

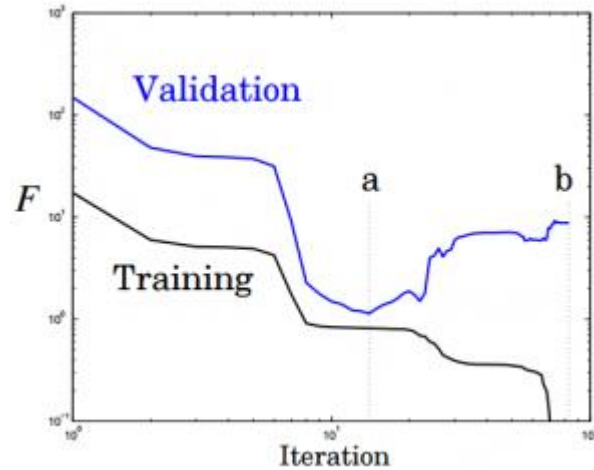
► Deep neural networks



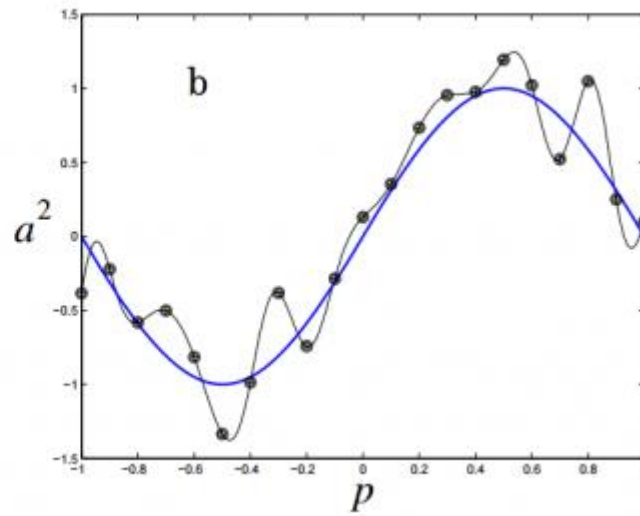
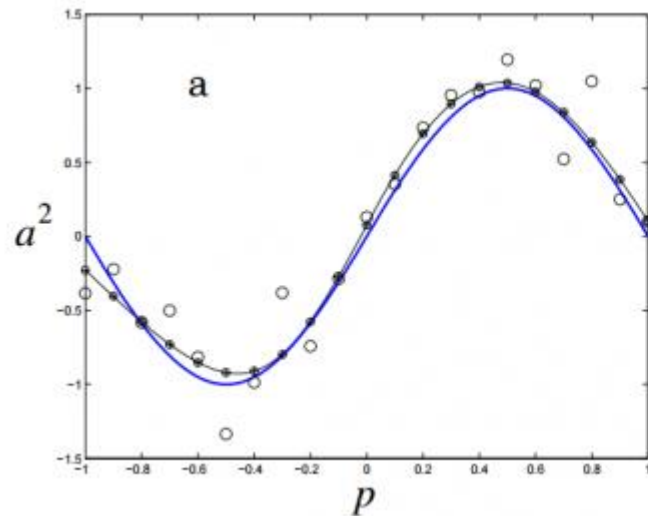
Neural networks come in many architectures



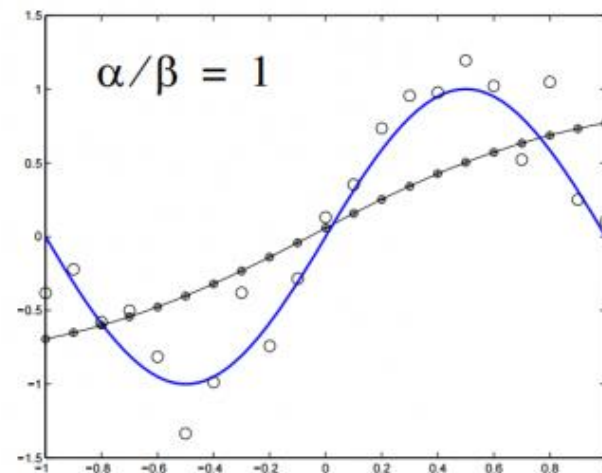
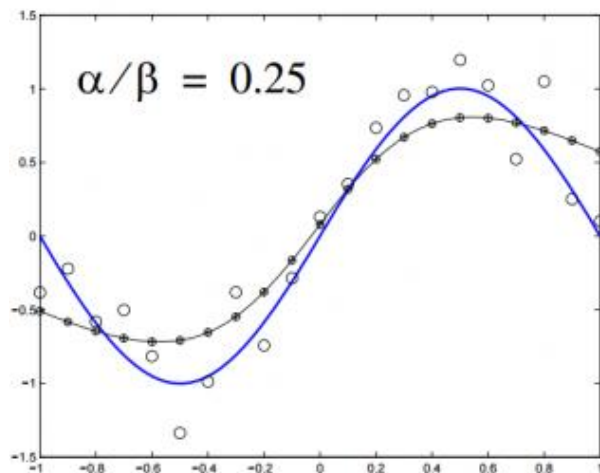
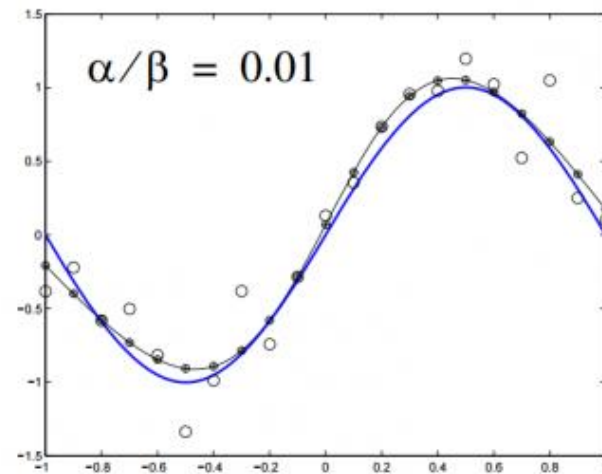
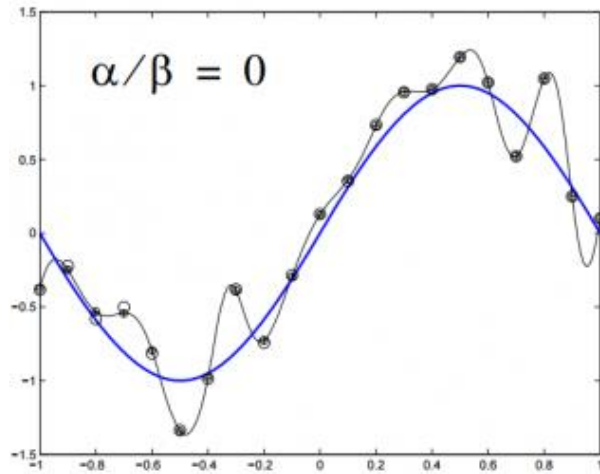
Size matters, but bigger isn't always better



$$\epsilon = \beta \sum_{p=1}^{P_T} (t_p - o_p)^2 + \alpha \sum_{j=1}^n v_j^2$$



Size matters, but bigger isn't always better

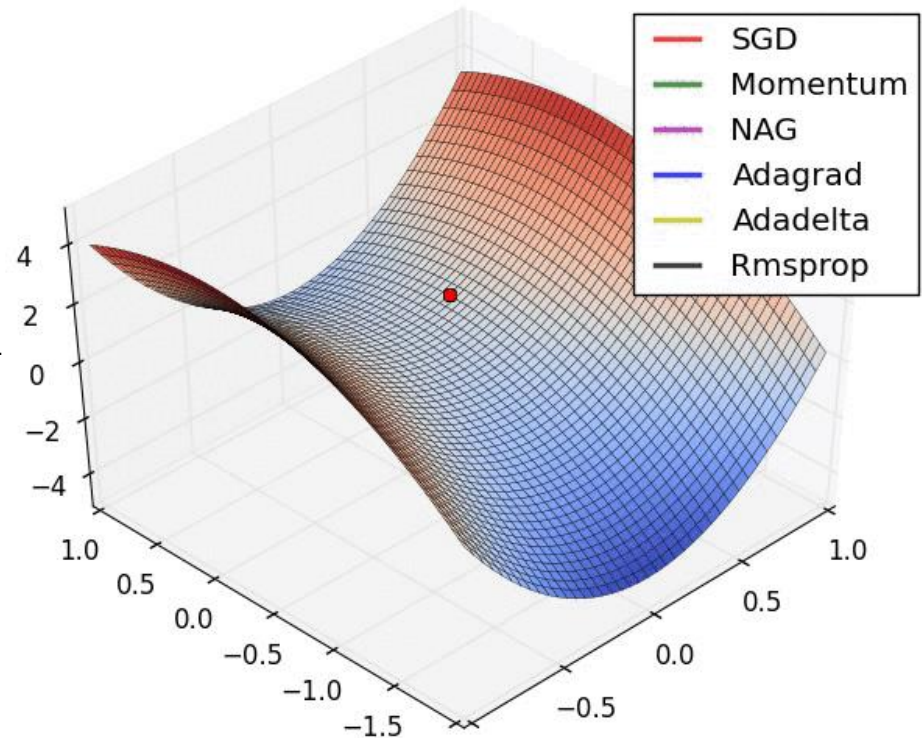


Many training algorithms exist for neural networks

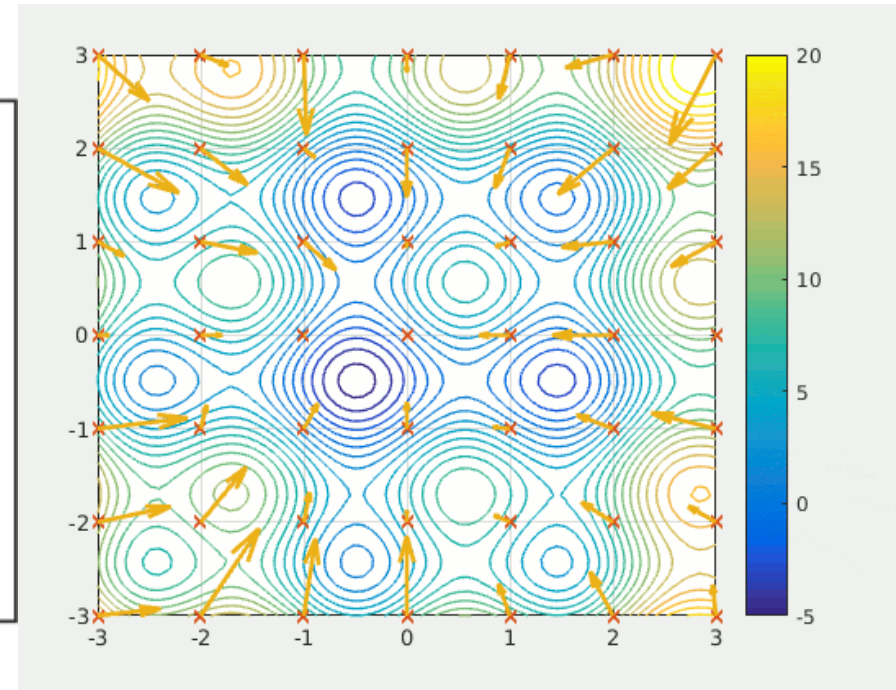
Метод зворотного поширення помилки:

Пряме проходження: дані проходять через мережу, записується вихідний сигнал і підраховуються похибки.

Зворотне поширення: сигнал помилки проходить назад через мережу, вага мережі оптимізується з використанням градієнтного спуску.



Метод рою частинок



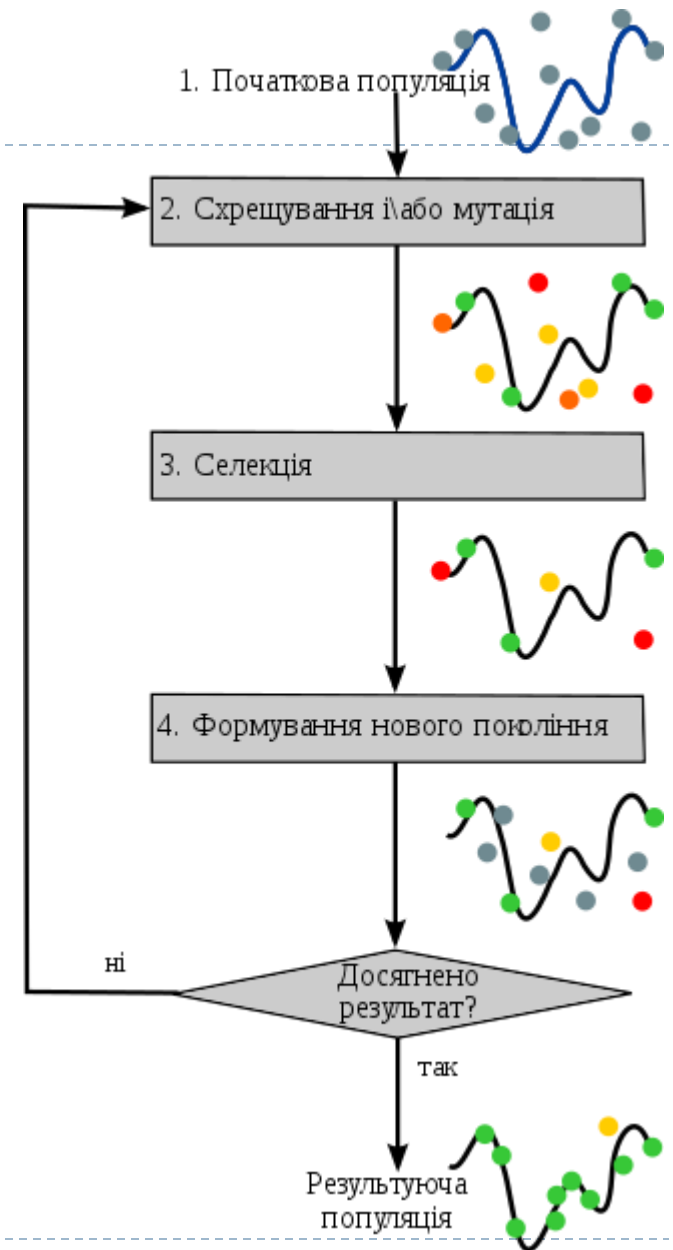
$$\mathbf{v}_i \leftarrow \omega \mathbf{v}_i + \phi_p \mathbf{r}_p \times (\mathbf{p}_i - \mathbf{x}_i) + \phi_g \mathbf{r}_g \times (\mathbf{g} - \mathbf{x}_i)$$

Генетичний алгоритм

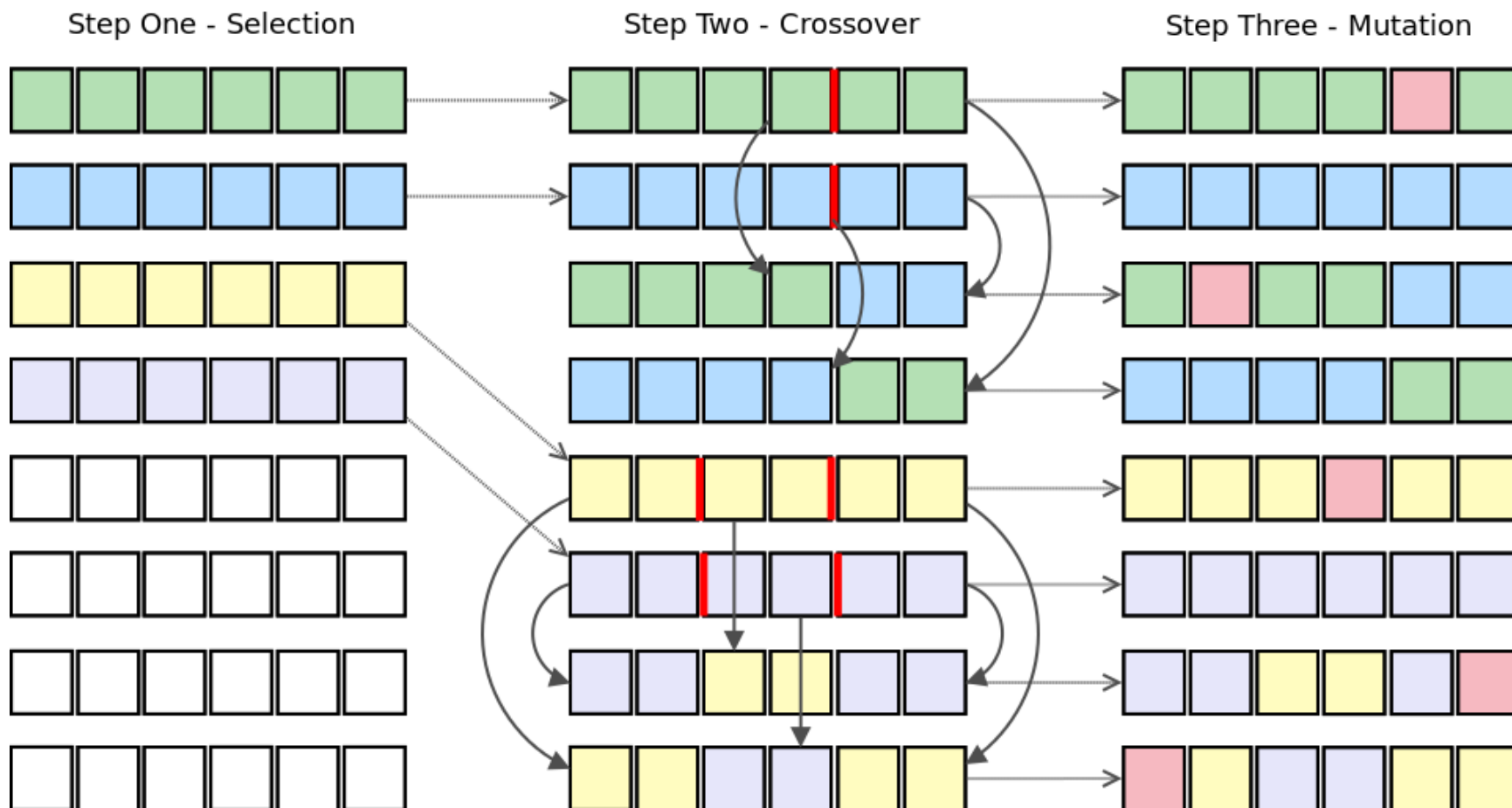
Селекція: після кожного прямого проходження підраховується сума квадратів помилок, популяція нейронної мережі ранжируется. Верхній відсоток популяції вибирається для виживання і використовується для кросовера.

Кросовер: верхній $x\%$ генів популяції змагається між собою, отримуємо якесь нове потомство, кожне потомство представляє, по суті, нову нейронну мережу.

Мутація: цей оператор вимагає підтримки генетичної різноманітності в популяції, невеликий відсоток її відбирається для проходження мутації, тобто деякі ваги мережі будуть регулюватися випадково.

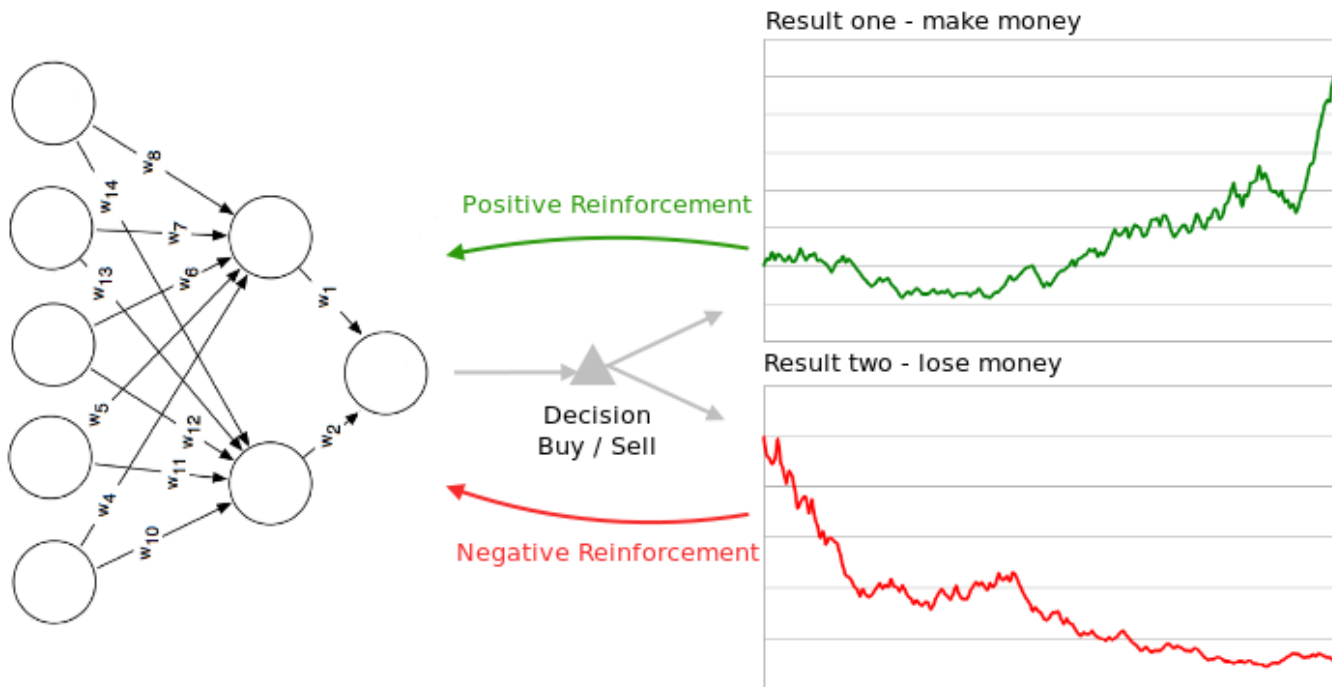


Генетичний алгоритм



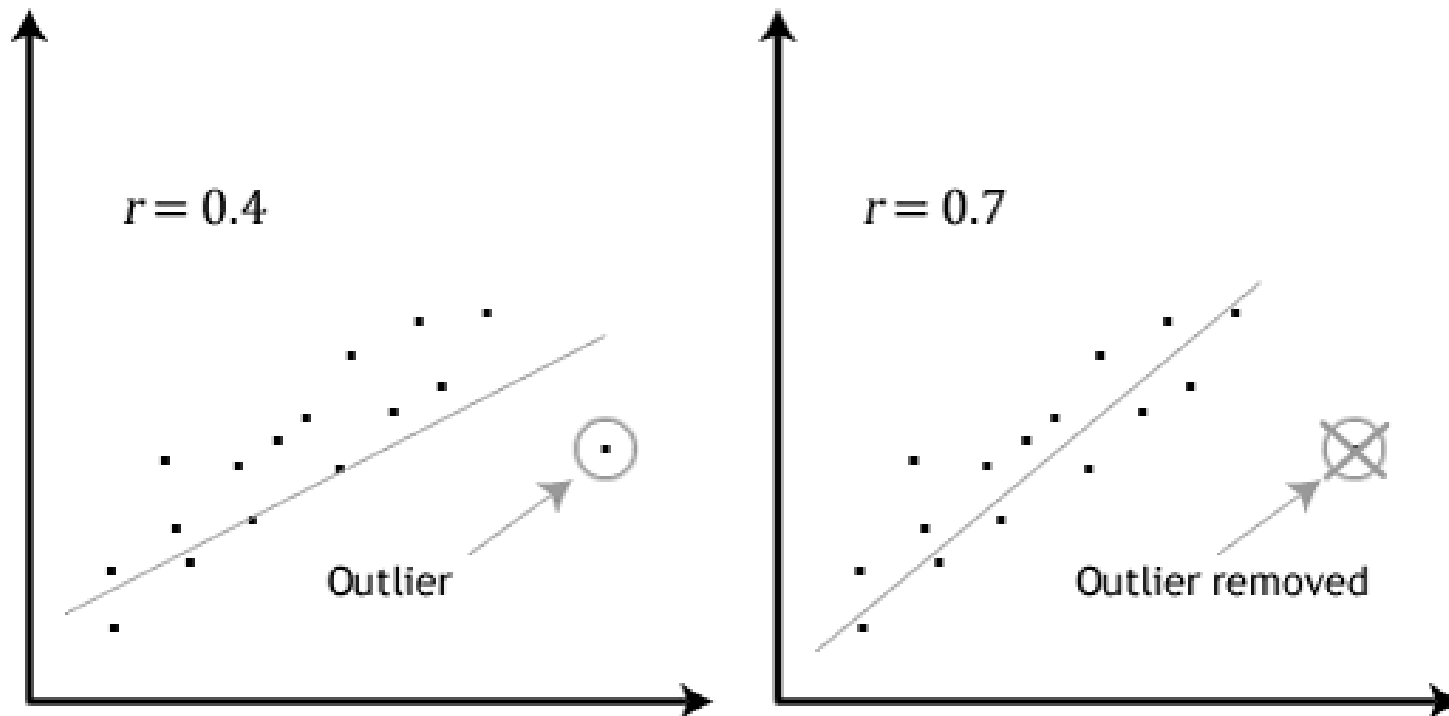
Neural networks do not always require a lot of data

- ▶ Контрольоване навчання
- ▶ Неконтрольоване навчання
- ▶ Навчання з підкріпленням



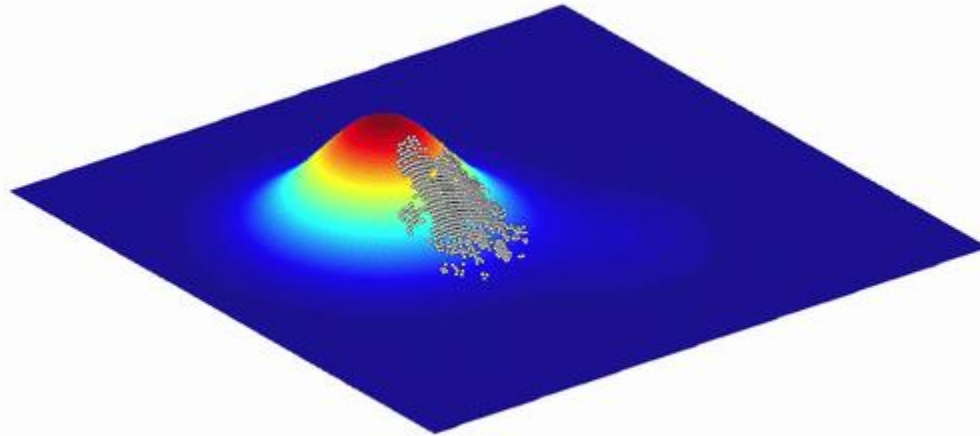
Neural networks cannot be trained on any data

- ▶ Нормалізація даних
- ▶ Видалення даних, що сильно вирізняються
- ▶ Видалення даних, що корелюють



Neural networks may need to be retrained

Dynamic fitness landscape

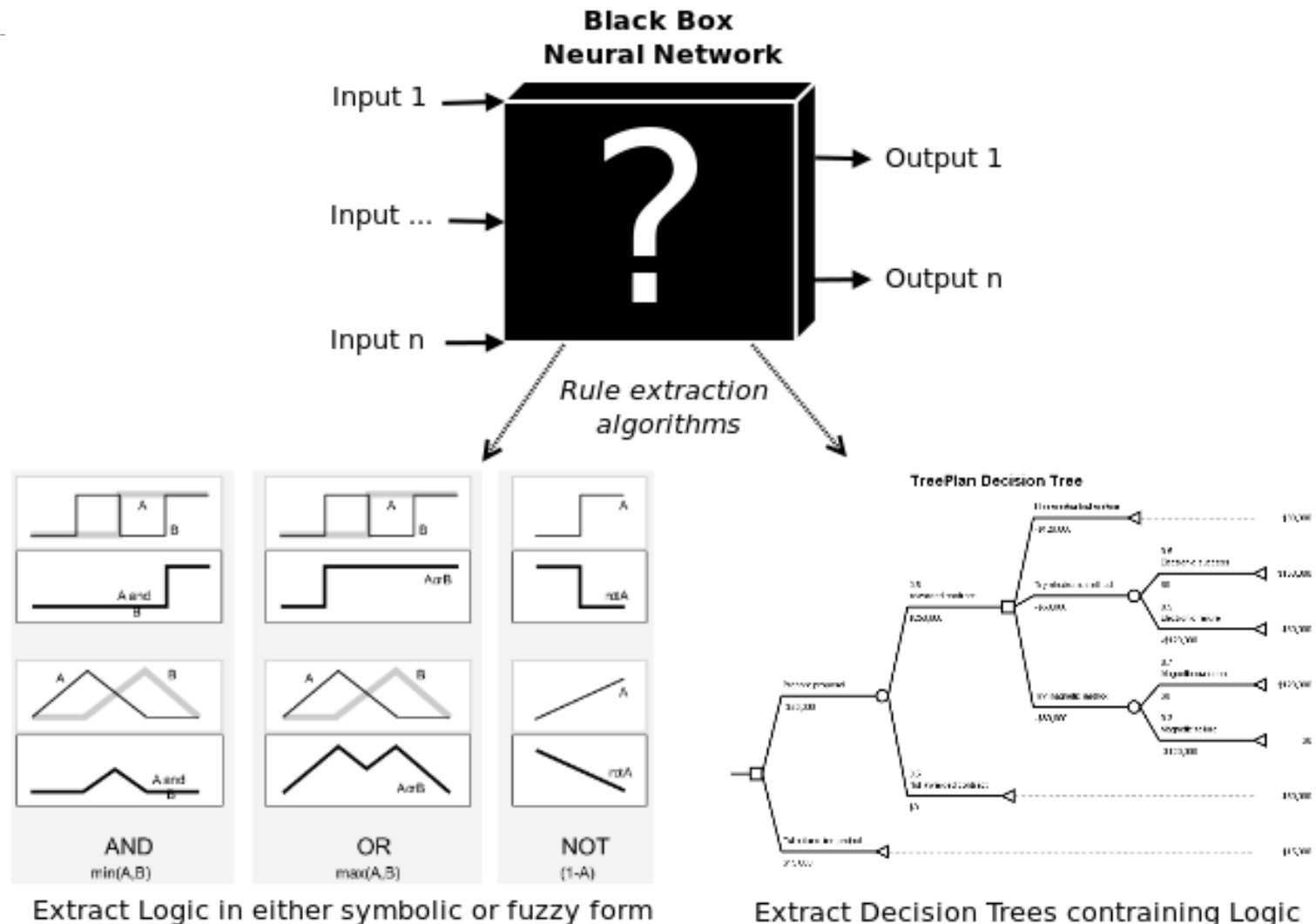


Population size, $N = 2,304$
Mutation rate, $\mu = 0.5$ per trait

© Randy Olson and Bjørn Østman



Neural networks are not black boxes



Дякую за увагу!

