

Анализ данных **и** Машинное обучение в гидрологии

...

Неделя 6

План

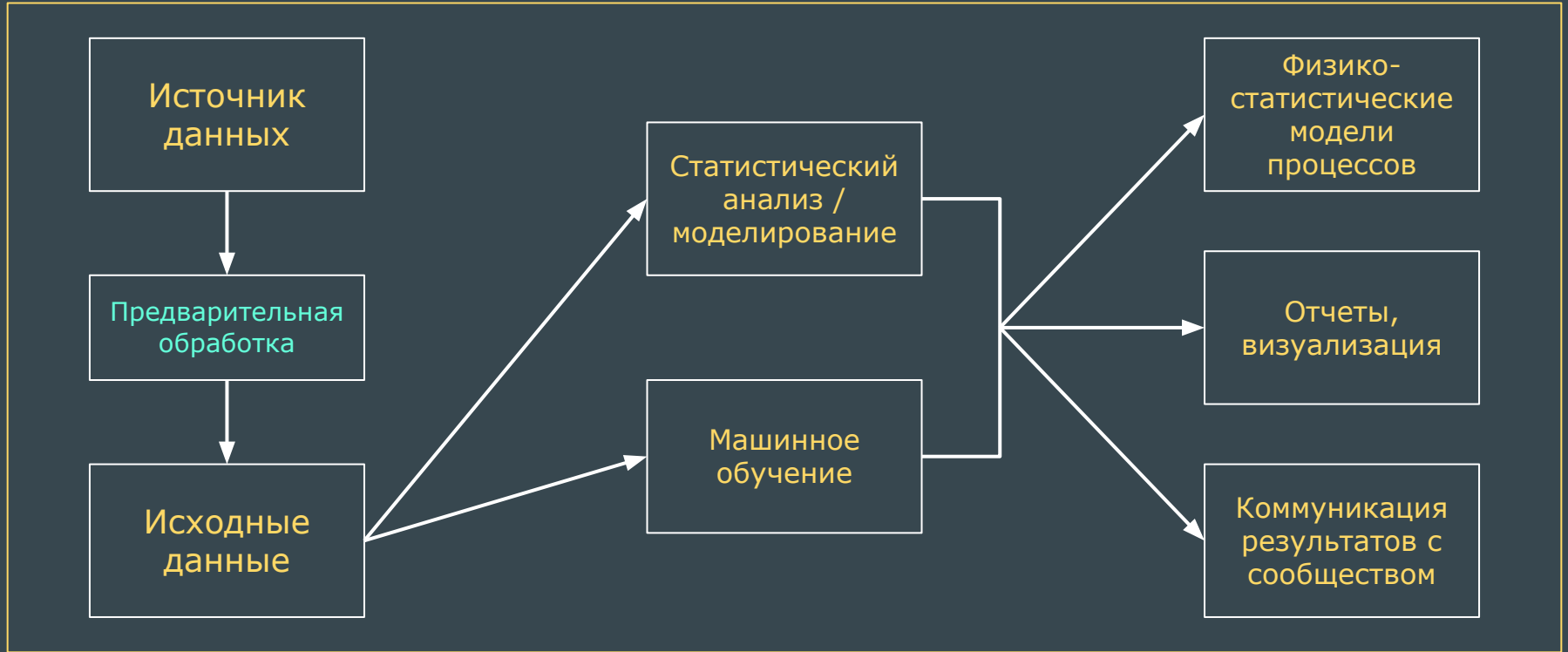
Лекция

- искусственный нейрон
- нейронные сети
- многослойный персептрон
- обучение нейронных сетей
- создание нейронных сетей ---
искусство машинного обучения

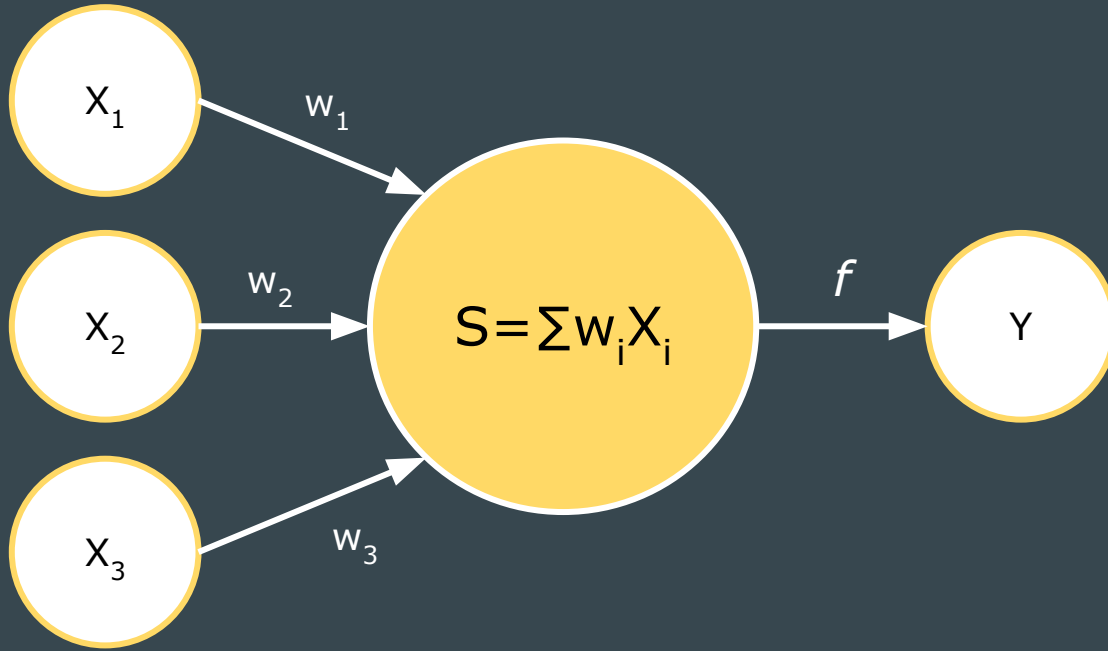
Практикум

- подготовка данных
- конструирование нейронной сети
- обучение
- настройка гиперпараметров
- проверка устойчивости

Research workflow



Искусственный нейрон



X_i -- ВХОДЫ

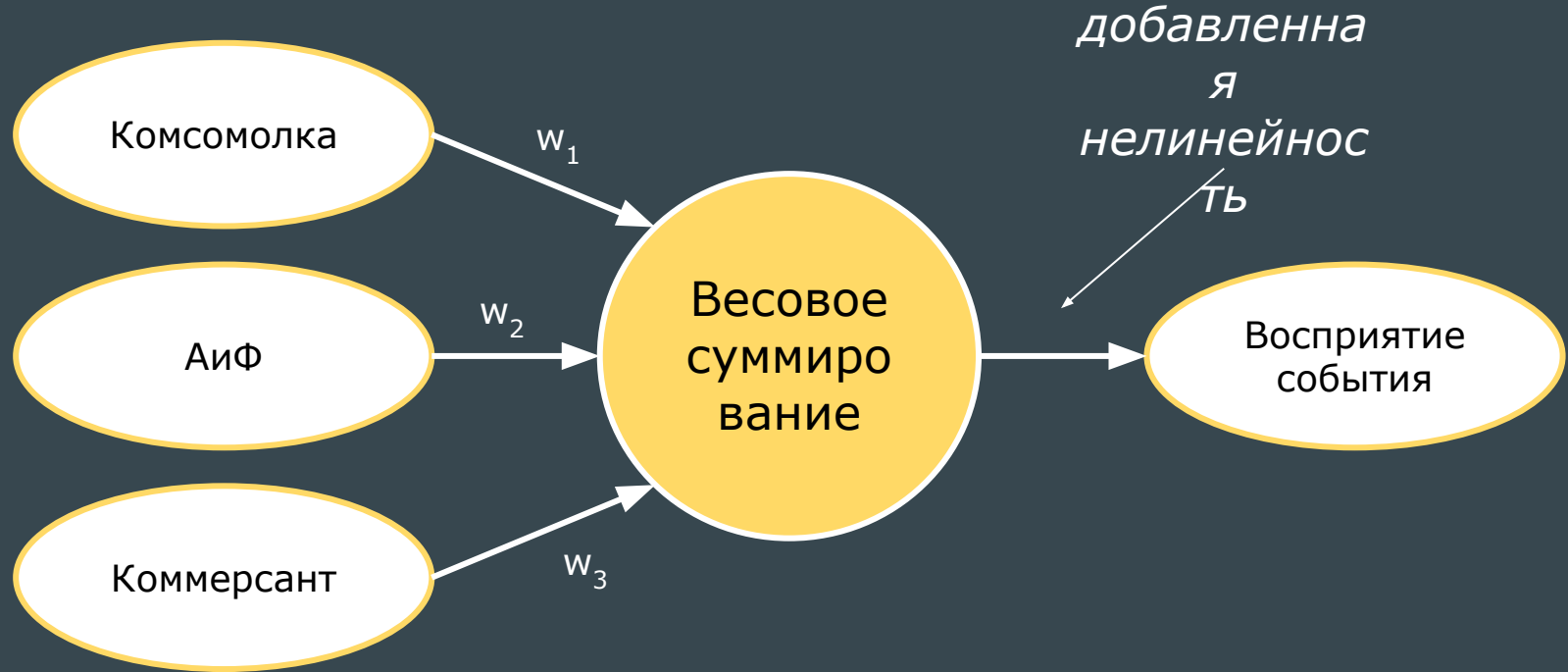
w_i -- веса

S -- сумматор

f -- функция
активации

Y -- ВЫХОД

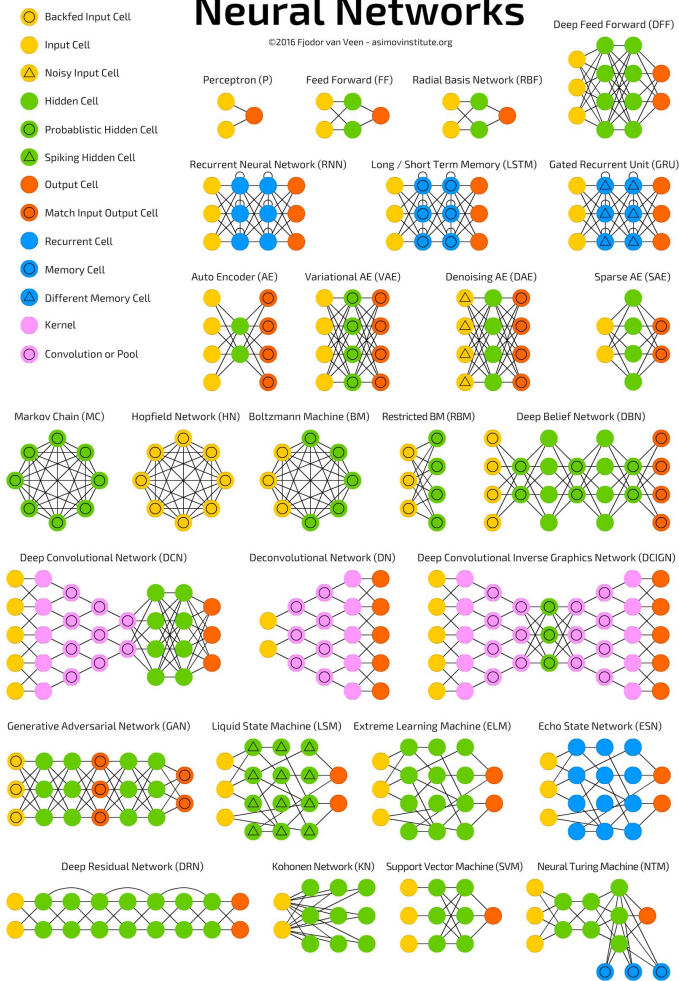
Искусственный нейрон. Смысл



A mostly complete chart of

Neural Networks

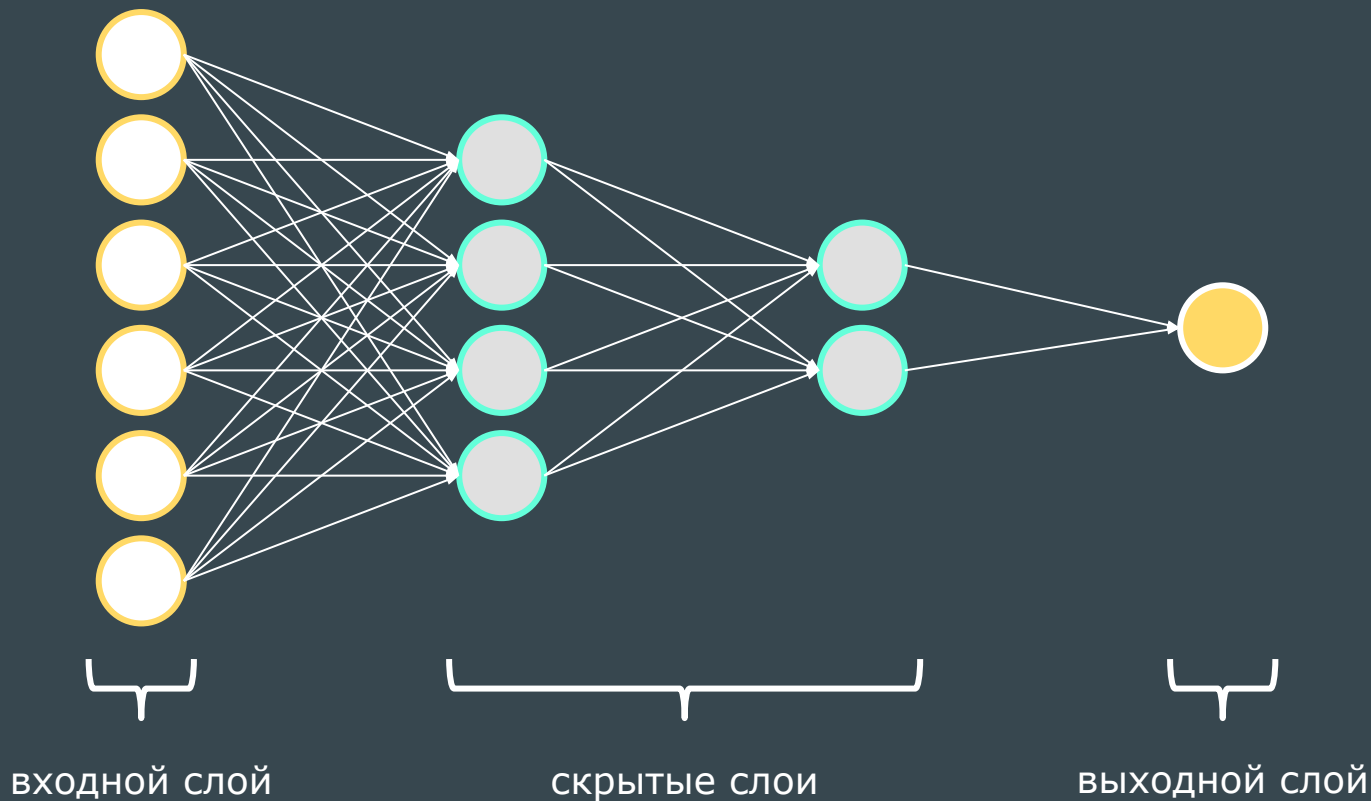
©2016 Fjodor van Veen - asimovinstitute.org



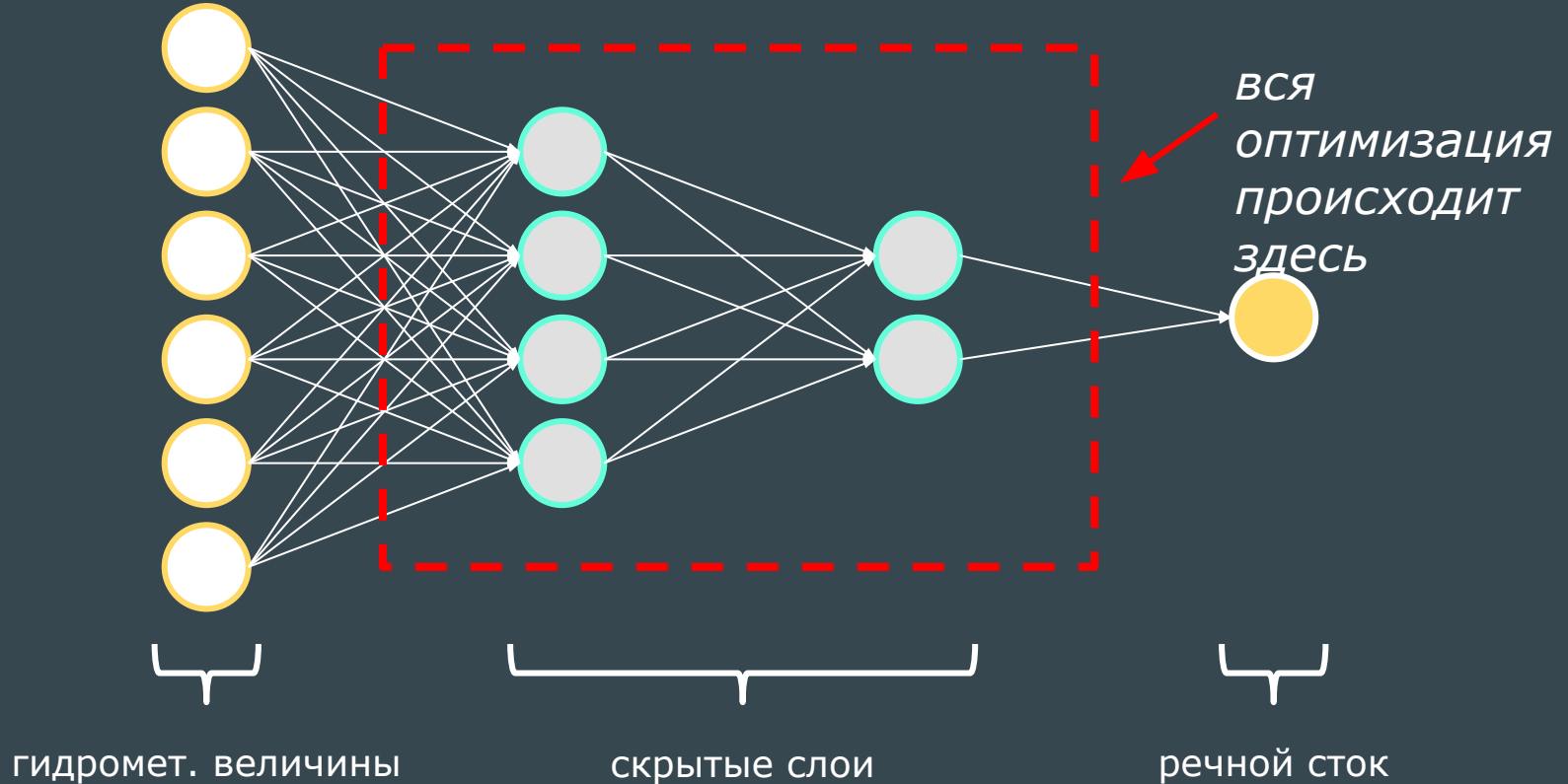
Зоопарк нейронных сетей

asimovinstitute.org/neural-network-zoo/

Многослойный персептрон (MLP, FFNN)



Многослойный персептрон в действии



Теорема сходимости Розенблатта

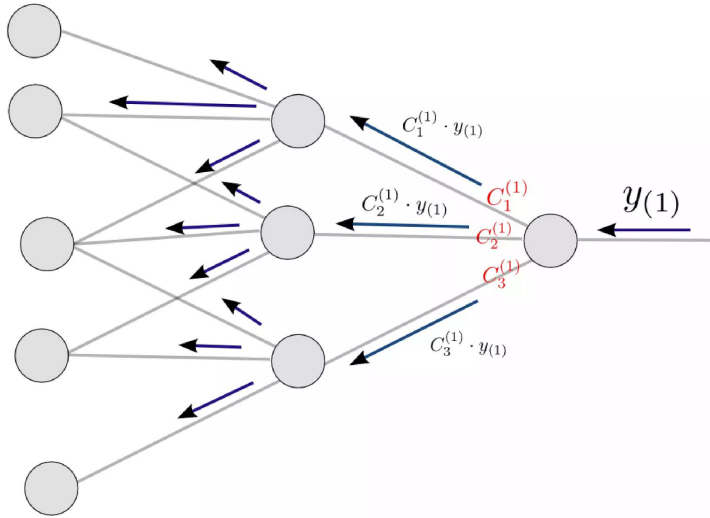
Под любую обучающую выборку можно подобрать архитектуру и веса нейронной сети с одним внутренним слоем, такую чтобы обучающая выборка классифицировалась с 100% точностью.

Теорема Цыбенко

Искусственная нейронная сеть прямой связи с одним скрытым слоем может аппроксимировать любую непрерывную функцию многих переменных с любой точностью.

Обучение нейронных сетей

Самый простой алгоритм -- обратного распространения ошибки



+ медленно, но верно

- паралич сети
- локальные минимумы
- размер шага обучения

Обучение нейронных сетей. Формулы

$$\begin{aligned}\delta_i^{(n_l-1)} &= \frac{\partial}{\partial z_i^{n_l-1}} J(W, b; x, y) = \frac{\partial}{\partial z_i^{n_l-1}} \frac{1}{2} \|y - h_{W,b}(x)\|^2 = \frac{\partial}{\partial z_i^{n_l-1}} \frac{1}{2} \sum_{j=1}^{S_{n_l}} (y_j - a_j^{(n_l)})^2 \\&= \frac{1}{2} \sum_{j=1}^{S_{n_l}} \frac{\partial}{\partial z_i^{n_l-1}} (y_j - a_j^{(n_l)})^2 = \frac{1}{2} \sum_{j=1}^{S_{n_l}} \frac{\partial}{\partial z_i^{n_l-1}} (y_j - f(z_j^{(n_l)}))^2 \\&= \sum_{j=1}^{S_{n_l}} -(y_j - f(z_j^{(n_l)})) \cdot \frac{\partial}{\partial z_i^{(n_l-1)}} f(z_j^{(n_l)}) = \sum_{j=1}^{S_{n_l}} -(y_j - f(z_j^{(n_l)})) \cdot f'(z_j^{(n_l)}) \cdot \frac{\partial z_j^{(n_l)}}{\partial z_i^{(n_l-1)}} \\&= \sum_{j=1}^{S_{n_l}} \delta_j^{(n_l)} \cdot \frac{\partial z_j^{(n_l)}}{\partial z_i^{n_l-1}} = \sum_{j=1}^{S_{n_l}} \left(\delta_j^{(n_l)} \cdot \frac{\partial}{\partial z_i^{n_l-1}} \sum_{k=1}^{S_{n_l-1}} f(z_k^{n_l-1}) \cdot W_{jk}^{n_l-1} \right) \\&= \sum_{j=1}^{S_{n_l}} \delta_j^{(n_l)} \cdot W_{ji}^{n_l-1} \cdot f'(z_i^{n_l-1}) = \left(\sum_{j=1}^{S_{n_l}} W_{ji}^{n_l-1} \delta_j^{(n_l)} \right) f'(z_i^{n_l-1})\end{aligned}$$

Обучение нейронных сетей. Картинки

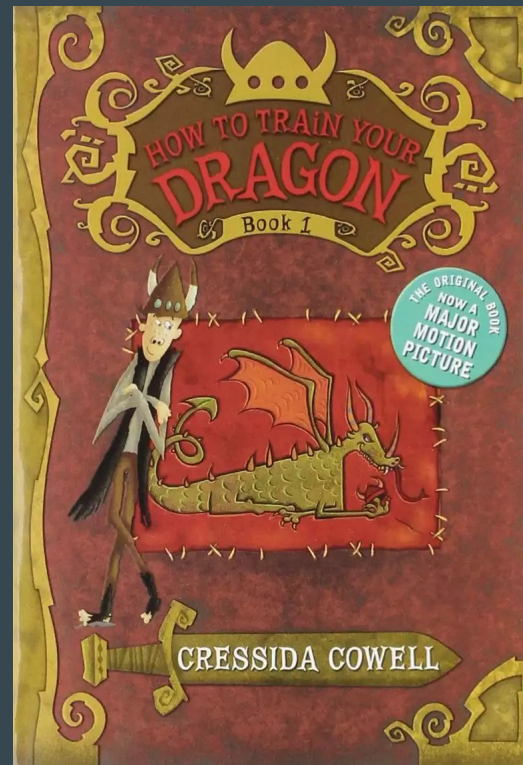


Обучение нейронных сетей. Яндекс

How to train your ~~dragon~~ ANN?

Кричите на неё!

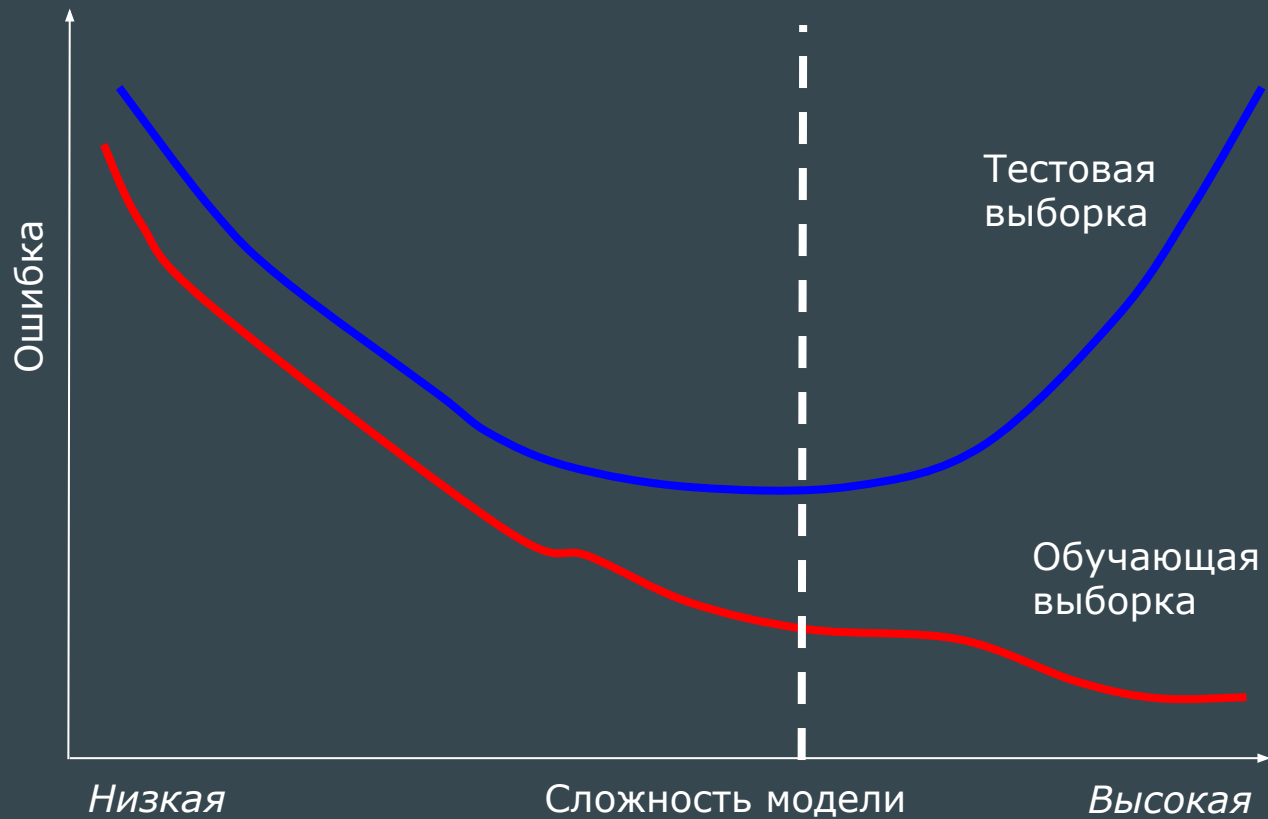
@Максим Мусин



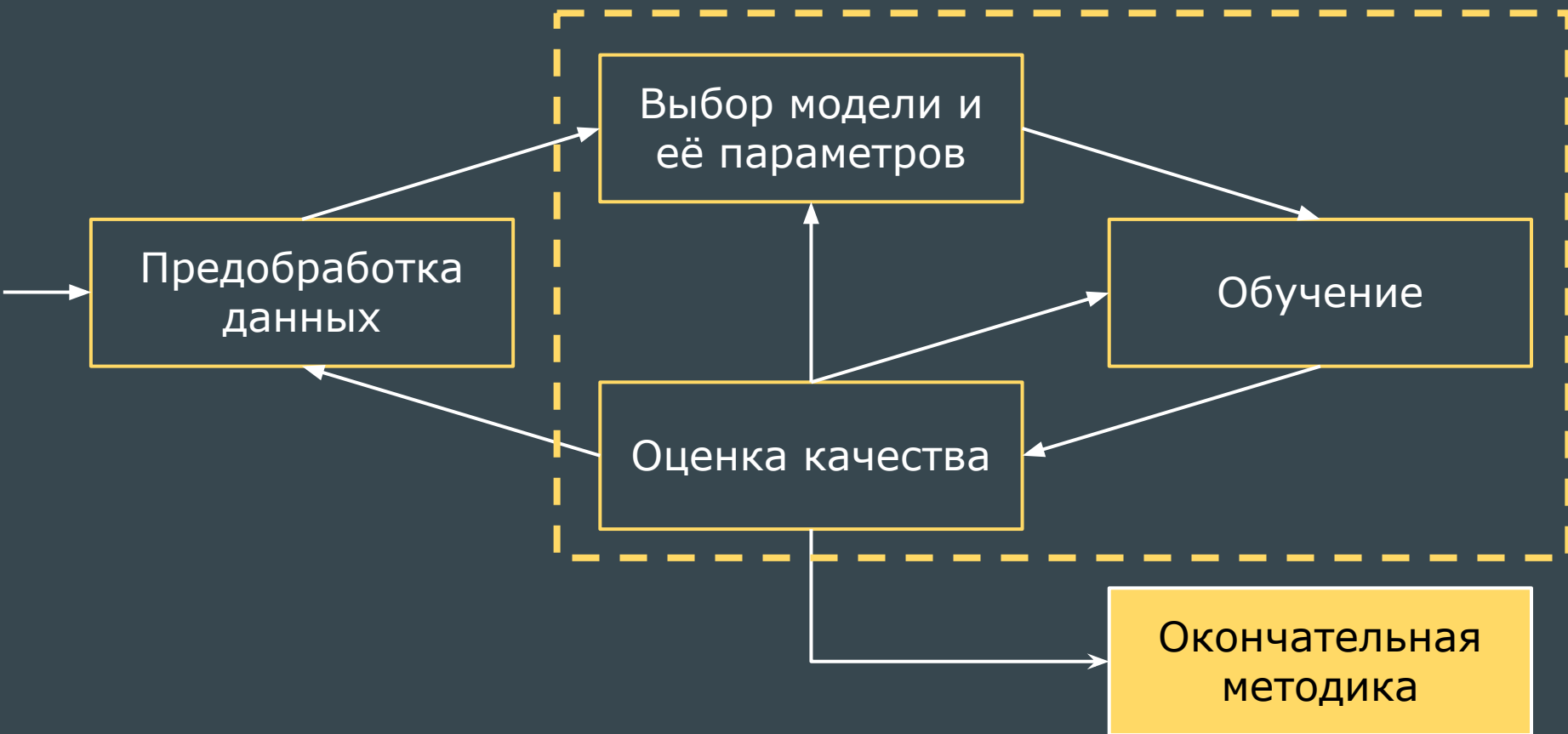
Обучение нейронных сетей. Реальность



Поиск баланса



Workflow



Подведем итоги

- + искусственный нейрон -- сумматор с добавленным свойством нелинейности (определяется функцией активации)
- + нейронная сеть -- полносвязная совокупность нейронов
- + тонкая настройка параметров нейронной сети -- искусство
- + время обучения сети сложной топологии огромно (дни)

Чем занять себя две недели?

- + Константин Лахман, малый ШАД. "Нейронные сети";
- + playground.tensorflow.org -- интерактивный подбор параметров;
- + визуализация обучения нейронных сетей на Keras;
- + методы оптимизации нейронных сетей (habrahabr);
- + что действительно скрывают нейронные сети (habrahabr).

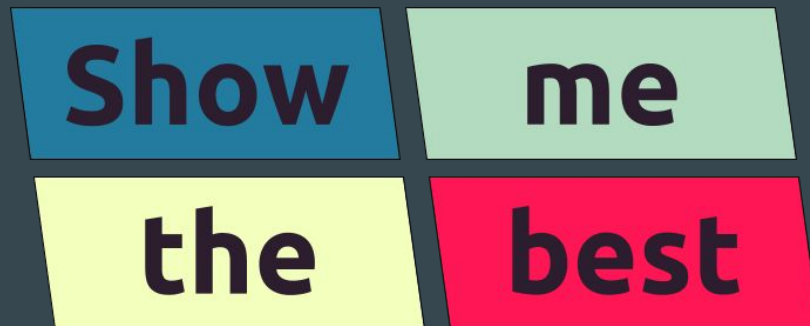
На правах рекламы

Plovcast



новости науки глазами молодых
ученых ИВП и ИО РАН

Show me the best



все яркие события из мира машинного
обучения, python и науки

Важно

Вы можете помочь существенно улучшить этот курс!

- ayzelgv@gmail.com, hydrogo@yandex.ru
- vk.com/ayzelgv, facebook.com/ayzelgv
- ИВП РАН, кабинет 617