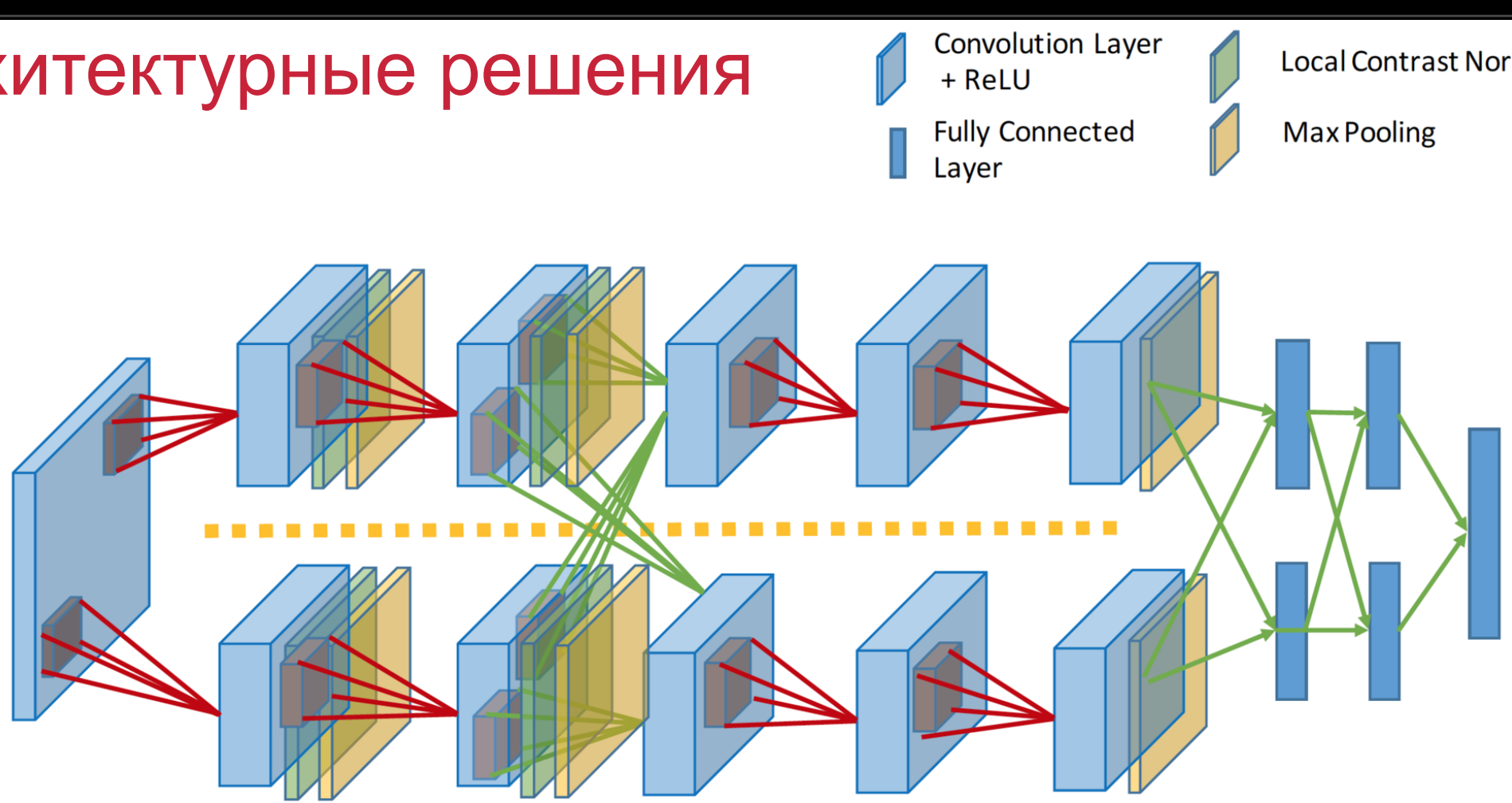
DL: Повышение качества обучения

План

- Способы повышения качества обучения
- Байесовская оптимизация

Архитектурные решения

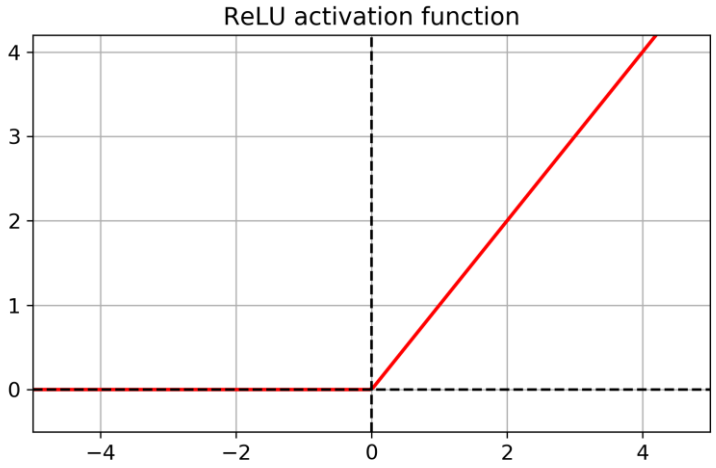


Архитектурные решения

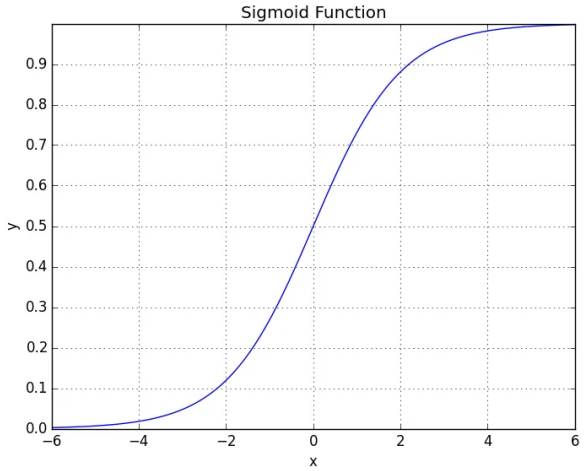
- Input
- Convolution
- Pooling
- Nonlinear
- Normalization
- Loss

Нелинейность

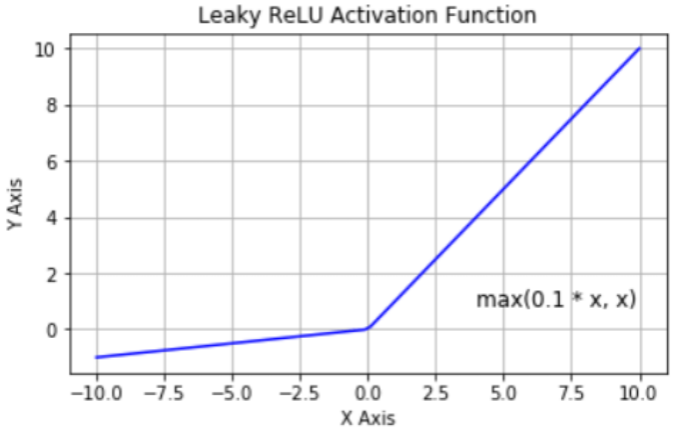
$$relu = max(0, x)$$



$$\sigma = \frac{1}{1 + e^{-x}}$$



$$f = max(0.1x, x)$$



Loss

$$L_2 = \sum (y - t)^2$$

$$H = \sum t \, log y$$

$$H(t, y) = H(t) + D_{KL}(t, y)$$

$$D_{KL}(t, y) = \sum t_i \, log \frac{y_i}{t_i}$$

$$H(t) = \sum t_i \, log t_i$$

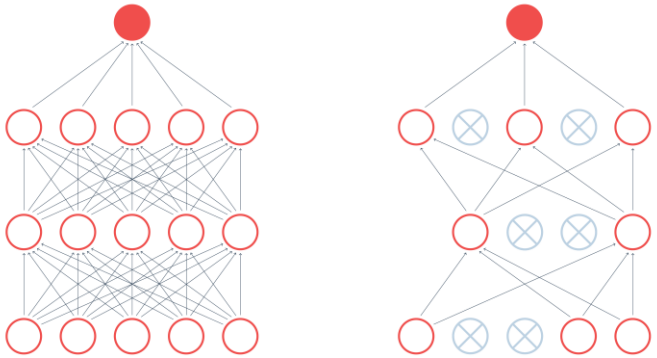
Аугментация

Регуляризация

$$L_1 = \sum (y - t)^2 + \lambda \sum |w_i|$$

$$L_2 = \sum (y - t)^2 + \lambda \sum w_i^2$$

Dropout

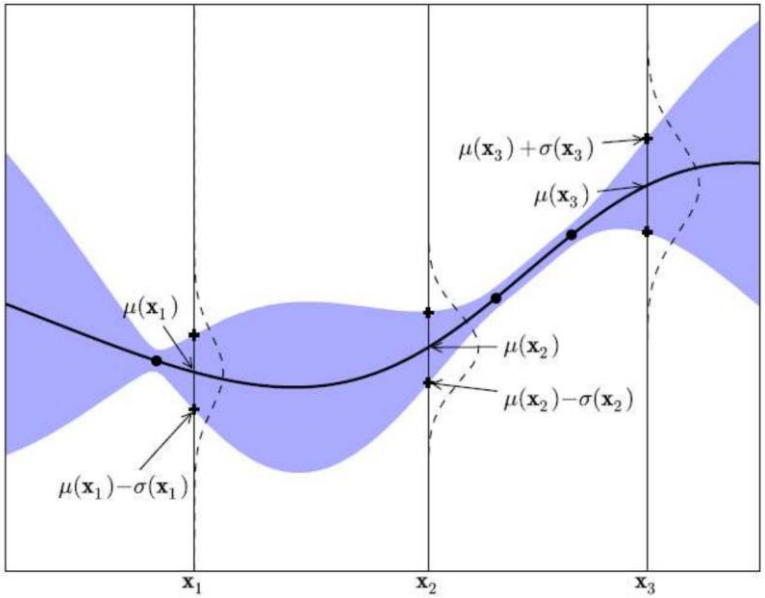


Оптимизация гиперпараметров

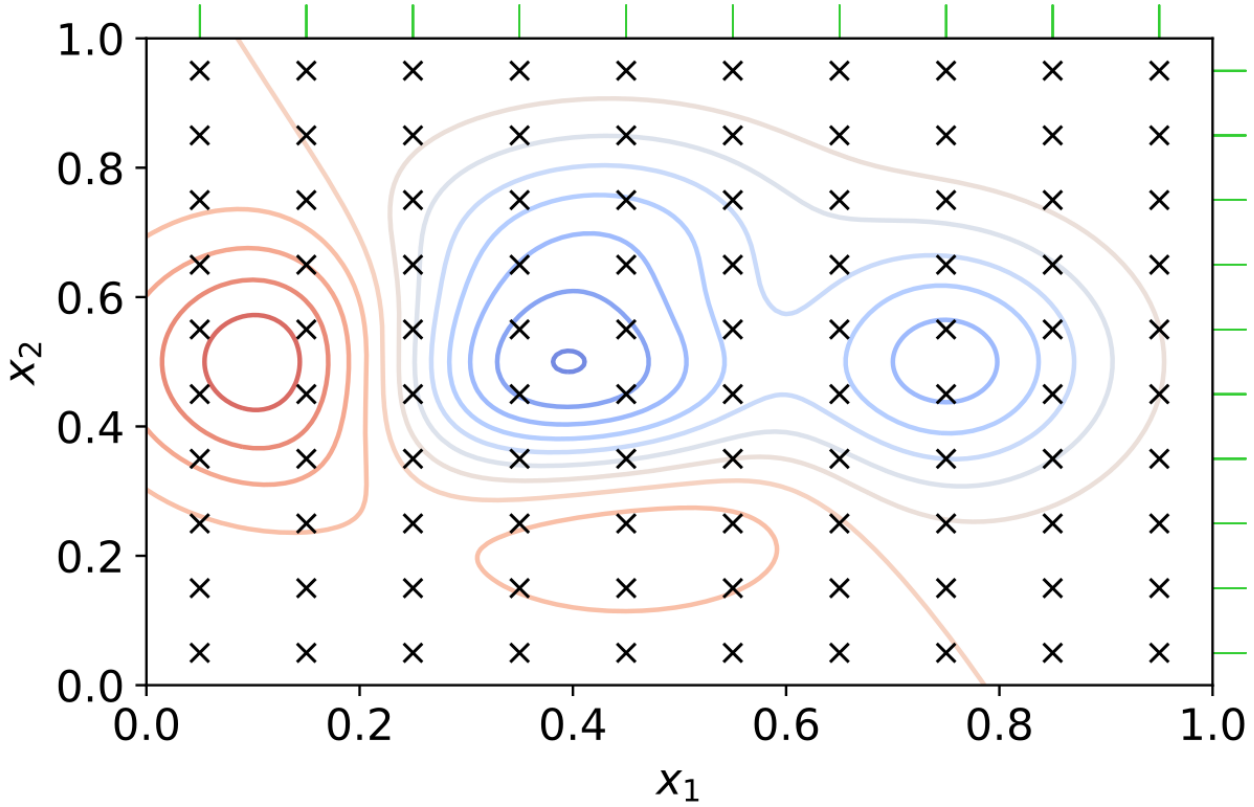
Поиск по сетке

Случайный поиск

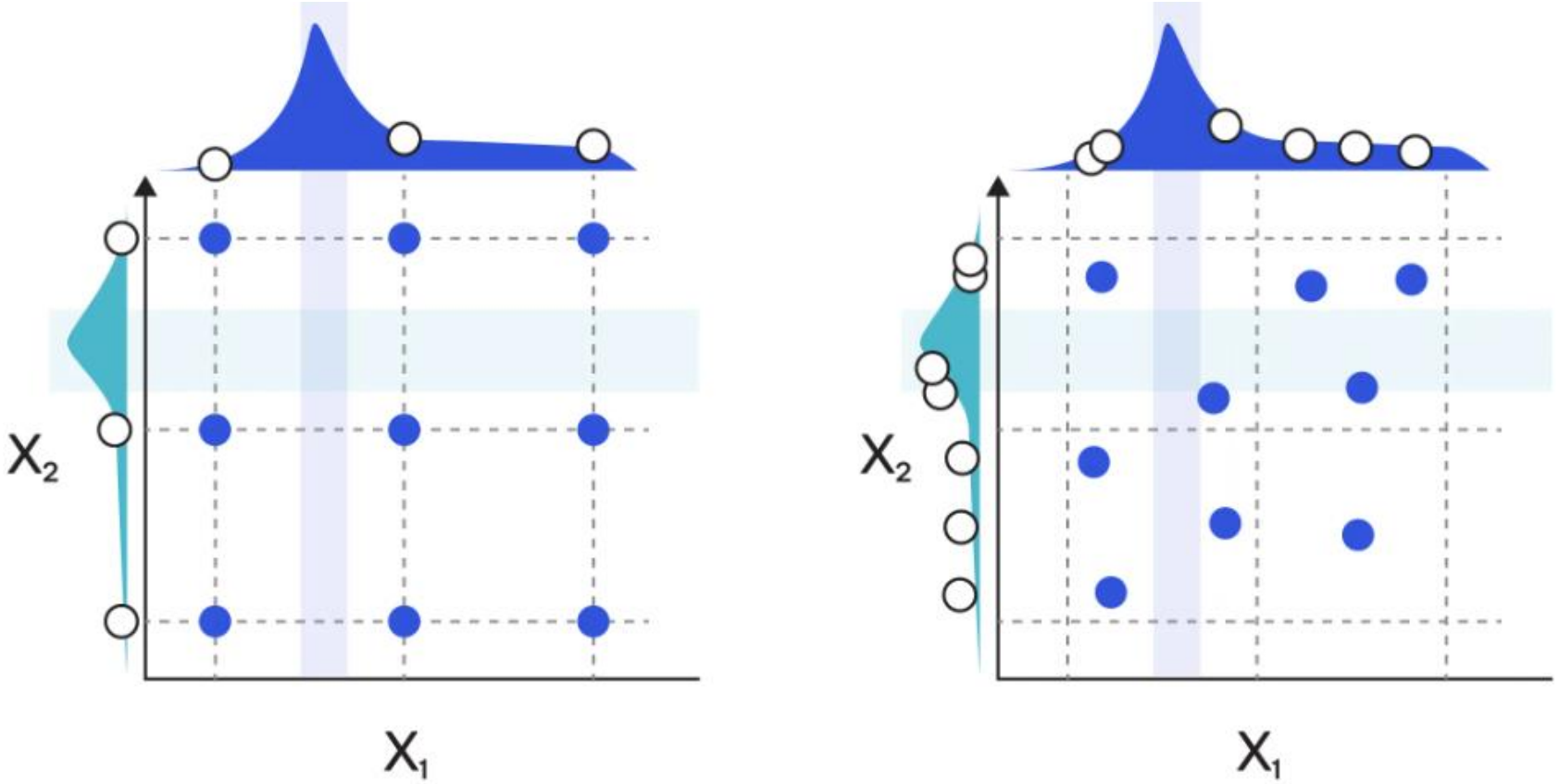
Байесовская оптимизация



Поиск по сетке

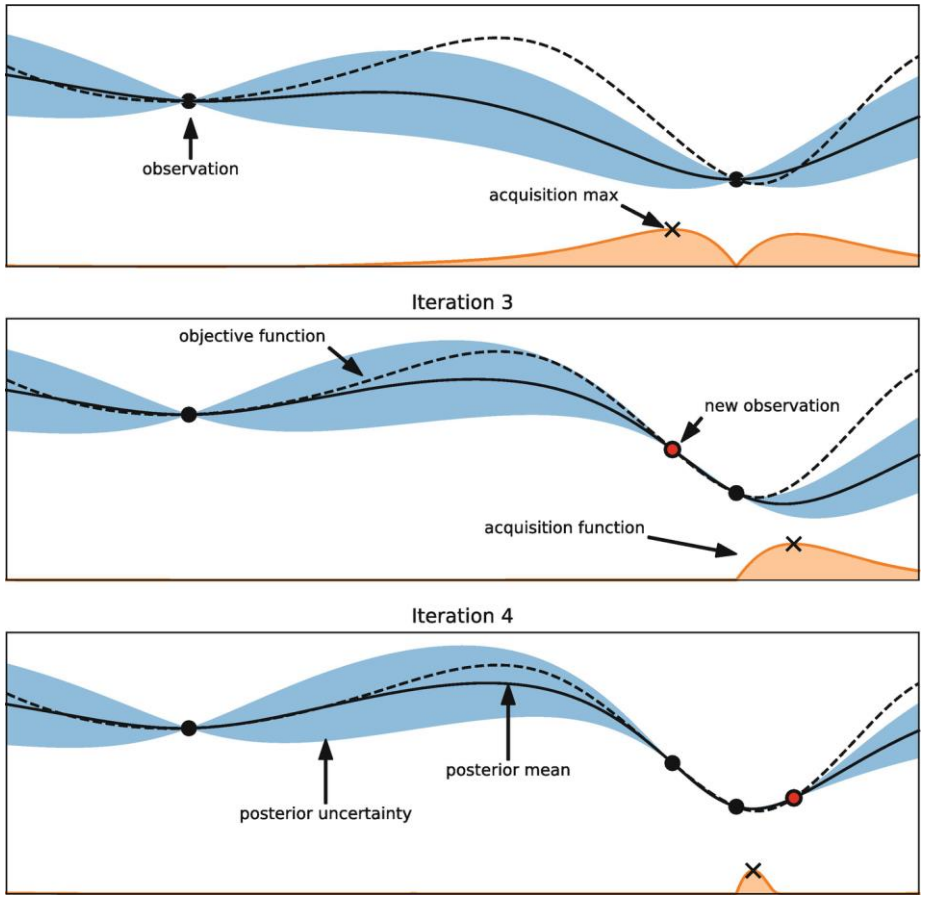


Случайный поиск



Байесовская оптимизация

acquisition function
surrogate function



Acquisition function (функция выбора)

Expected Improvement
Probability Improvement
UCB

- x_{best} as the location of the lowest posterior mean.
- $\mu_Q(x_{best})$ as the lowest value of the posterior mean.

Then the expected improvement

$$EI(x, Q) = E_Q\left[max\left(0, \mu_Q(x_{best}) - f(x)\right)\right]$$

$f(x)$ - суррогатная модель (GP модель)