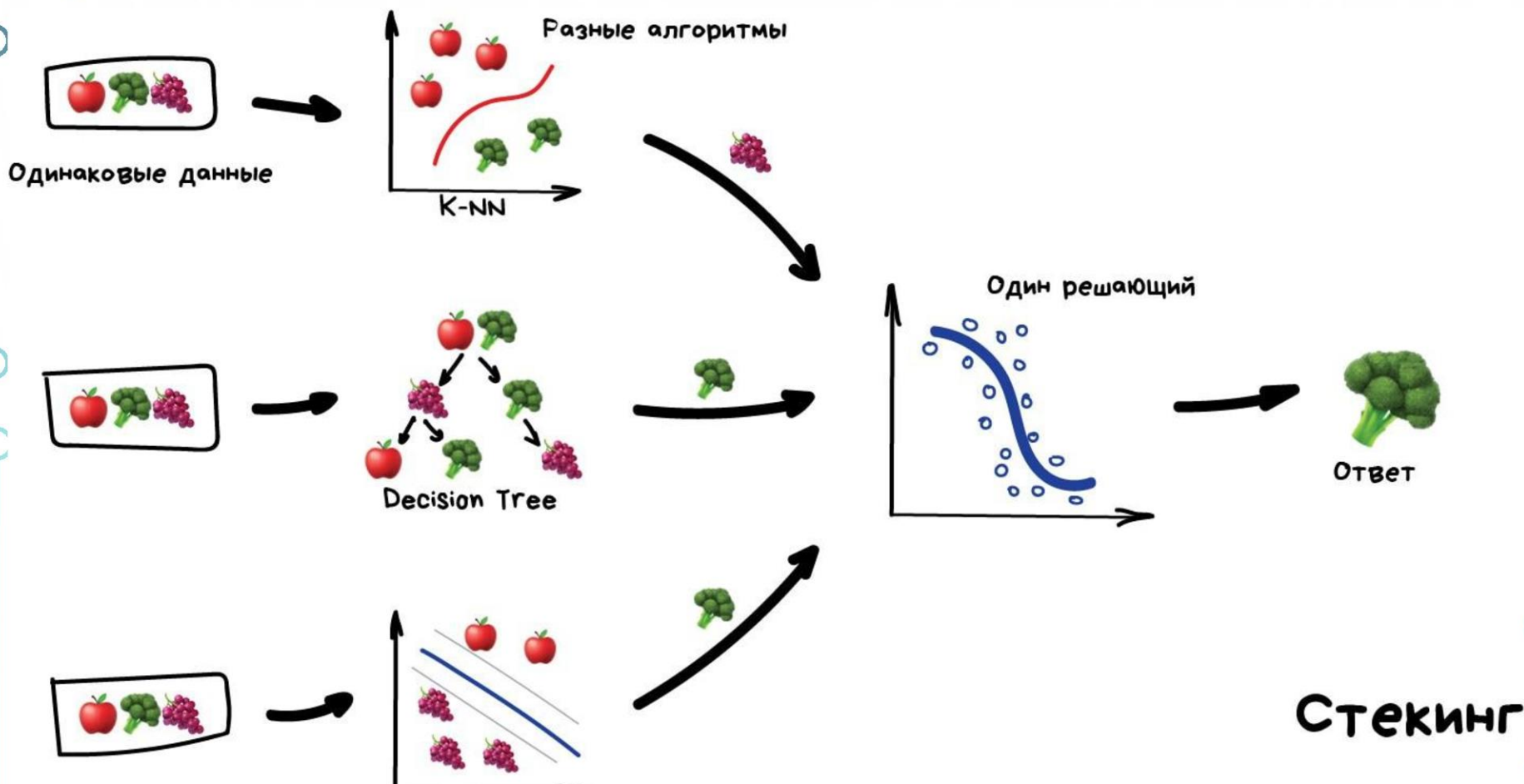
A decorative graphic on the left side of the slide consists of a network of thin, dark blue lines. These lines form a complex, branching pattern that resembles a circuit board or a neural network. The lines are interconnected by small, empty circles, creating a series of nodes and paths. The overall shape is vertical, extending from the top to the bottom of the slide, and is positioned on the left side, leaving the right side for the text.

# Blending, Stacking

Кантонистова Е.О.

# СТЕКИНГ (STACKING)

Идея: обучаем несколько разных алгоритмов и передаём их результаты на вход последнему, который принимает итоговое решение.



# СТЕКИНГ (STACKING)

- Пусть мы обучили  $N$  базовых алгоритмов  $b_1(x), b_2(x), \dots, b_N(x)$  на выборке  $X$ .
- Обучим теперь мета-алгоритм  $a(x)$  на прогнозах этих алгоритмов (т.е. прогнозы алгоритмов – это по сути новые признаки):

$$\sum_{i=1}^l L(y_i, \mathbf{a}(b_1(x_i), b_2(x_i), \dots, b_N(x_i))) \rightarrow \min_a$$

- алгоритм  $a(x)$  будет больше опираться на предсказание тех алгоритмов, которые сильнее подошлись под обучающую выборку  $\Rightarrow$  будет переобучен.

# СТЕКИНГ (STACKING)

Решение: будем обучать базовые алгоритмы и мета-алгоритм на разных выборках.

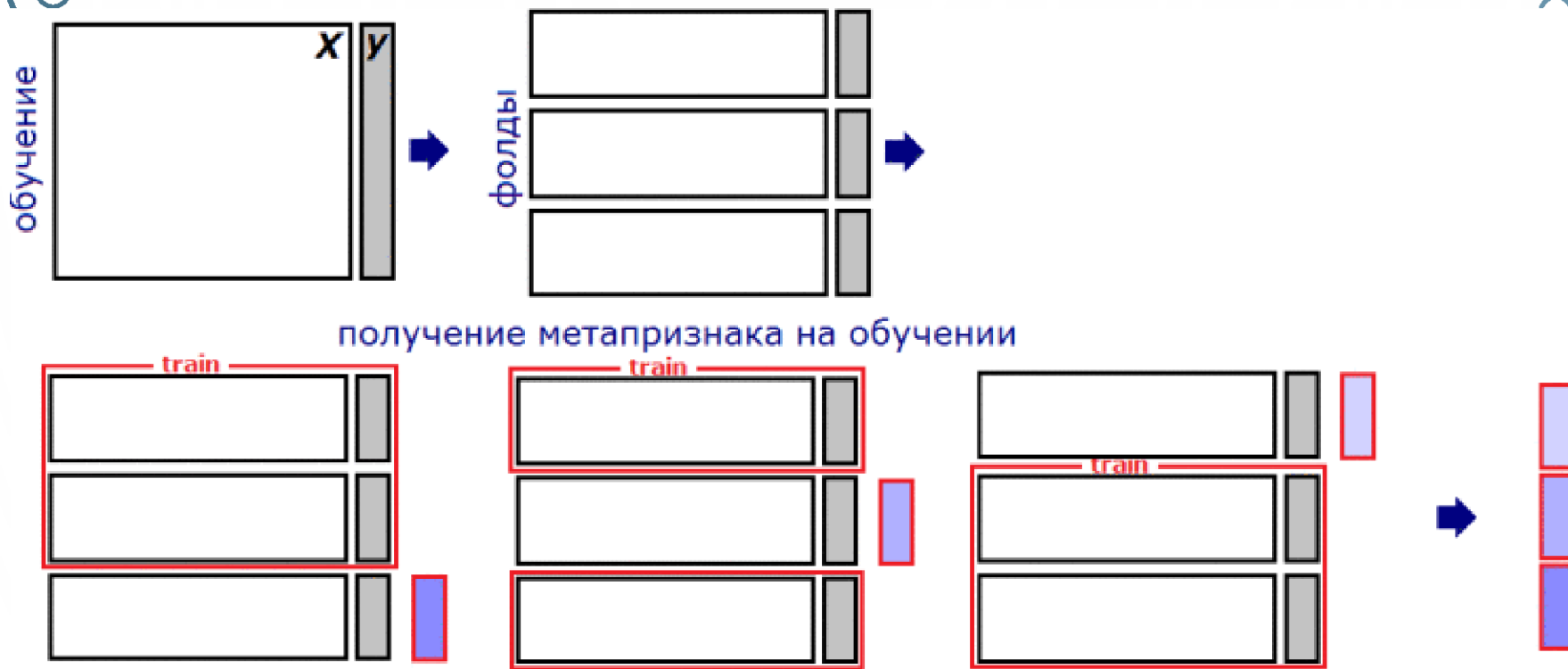
- Разобьем выборку на  $K$  частей:  $X_1, X_2, \dots, X_K$ .
- Пусть  $b_j^{-k}(x)$  -  $j$ -й алгоритм, обученный на всех блоках, кроме  $k$ -го.

Для обучения мета-алгоритма будем минимизировать функционал:

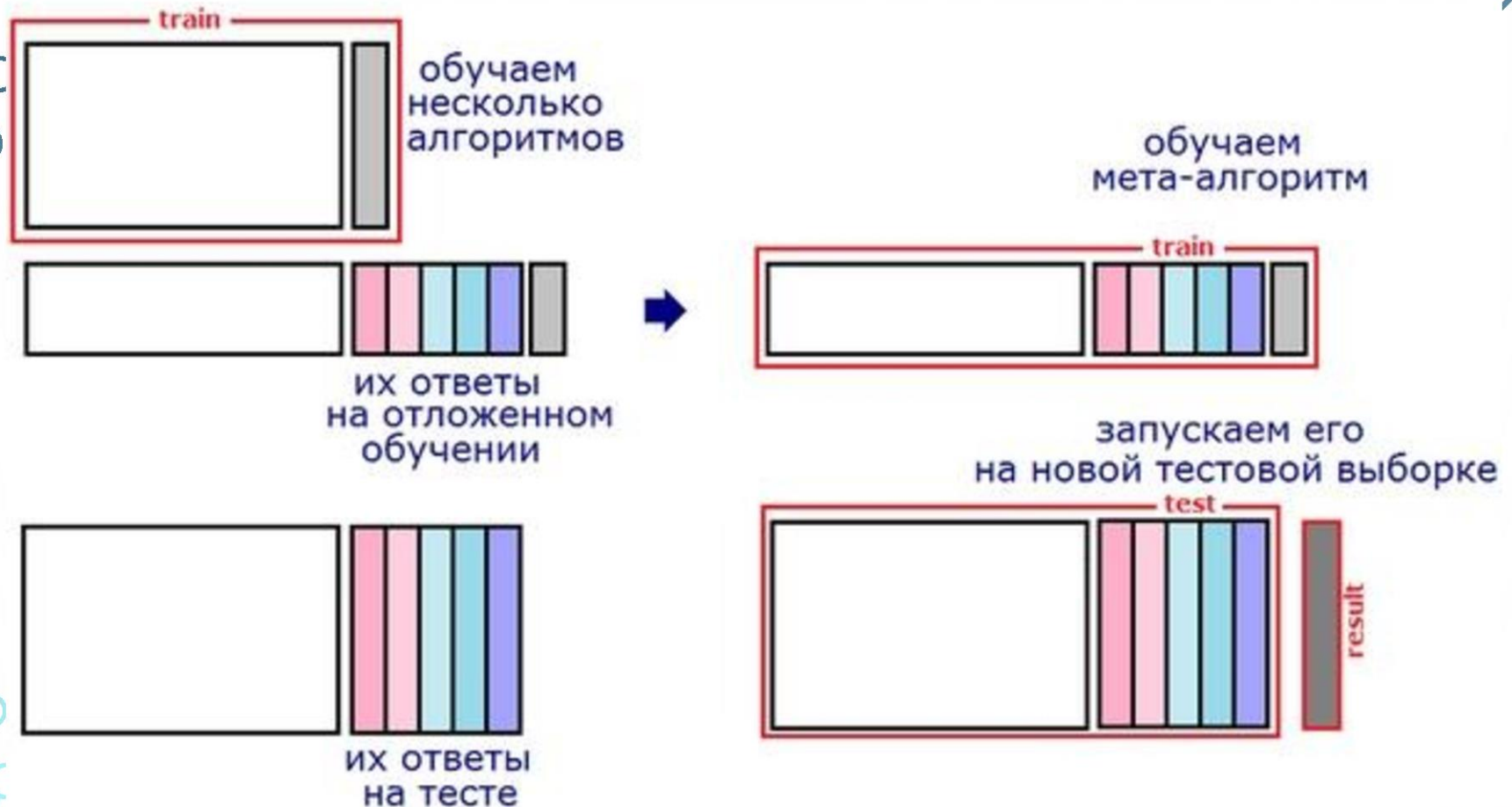
$$\sum_{k=1}^K \sum_{(x_i, y_i) \in X_k} L\left(y_i, a\left(b_1^{-k}(x_i), b_2^{-k}(x_i), \dots, b_N^{-k}(x_i)\right)\right) \rightarrow \min_a$$

- теперь алгоритм  $a$  обучается на объектах, на которых не обучались базовые алгоритмы  $\Rightarrow$  нет переобучения.

# СТЕКИНГ (STACKING)



# ИСПОЛЬЗОВАНИЕ МЕТАПРИЗНАКОВ ВМЕСТЕ С ПРИЗНАКАМИ

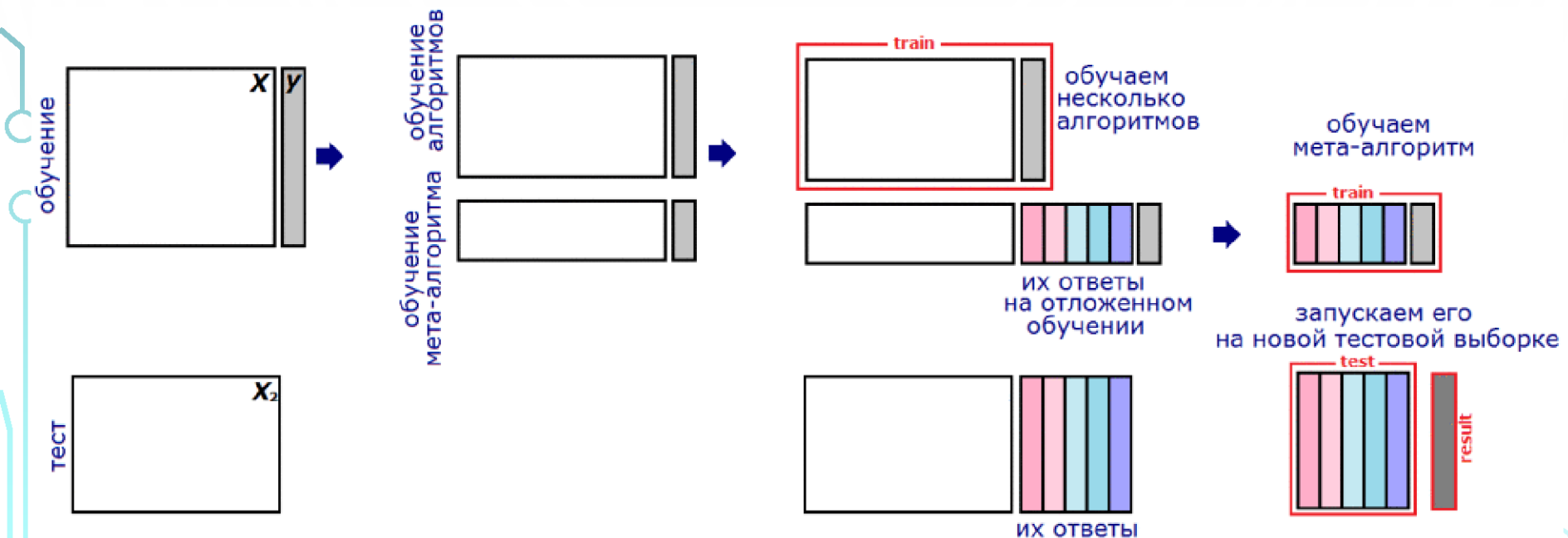


<https://dyakonov.org/2017/03/10/стекинг-stacking-и-блендинг-blending/>

# БЛЕНДИНГ (BLENDING)

Блендинг – это частный случай стекинга, в котором мета-алгоритм линеен:

$$a(x) = \sum_{n=1}^N w_n b_n(x)$$



# BLENDING И STACKING

Характеристика	Blending	Stacking
Разделение данных	Использует holdout-набор	Использует k-fold cross-validation
Использование данных	Потеря части данных (на holdout-набор)	Эффективно использует всю тренировочную выборку
Сложность реализации	Простая	Более сложная
Выбор метамоделей	Простая модель (например, линейная регрессия)	Любая модель
Риск переобучения	Ниже	Выше
Точность	Обычно ниже	Обычно выше