



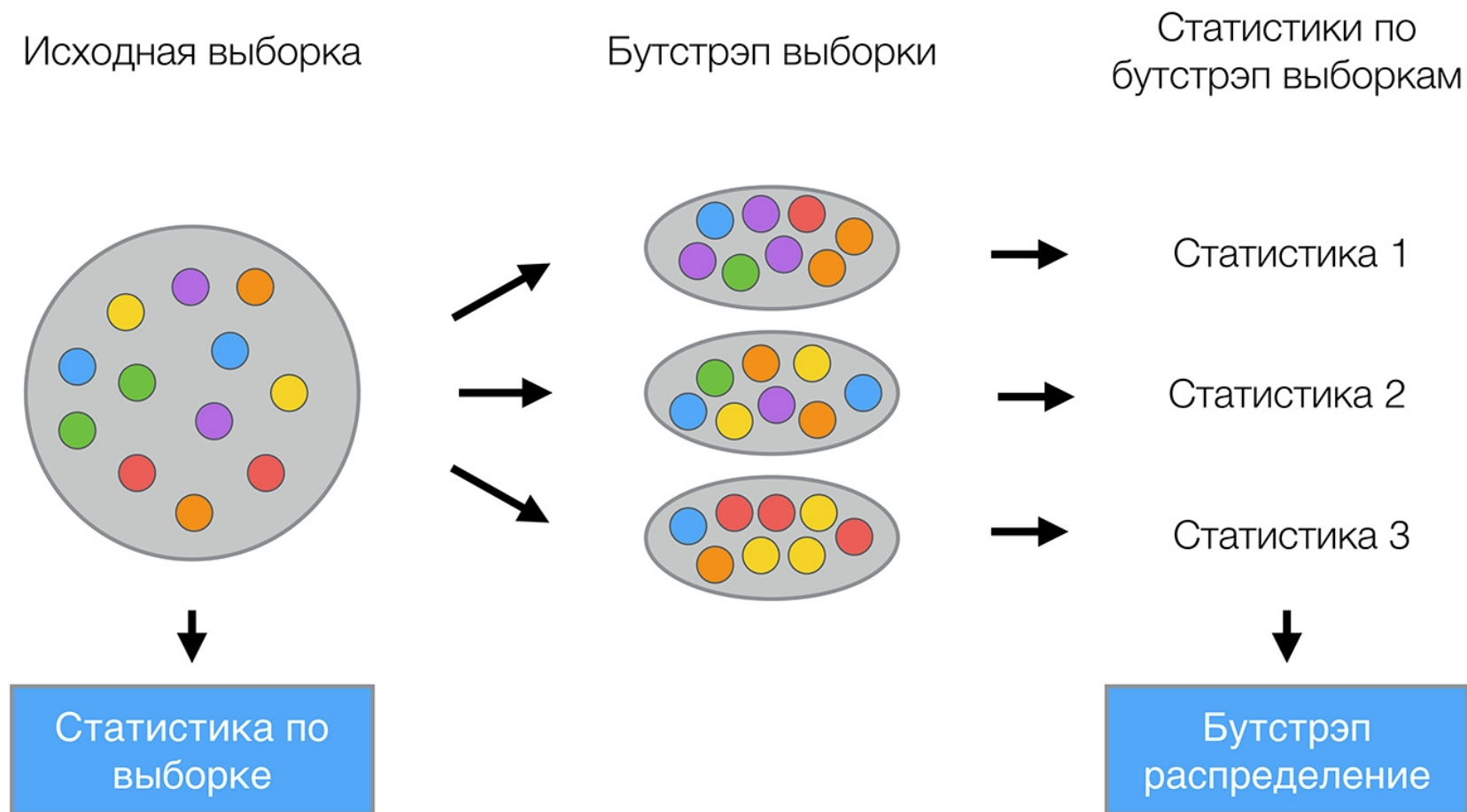
Ансамбли Бустинг

Корлякова Мария.
2021

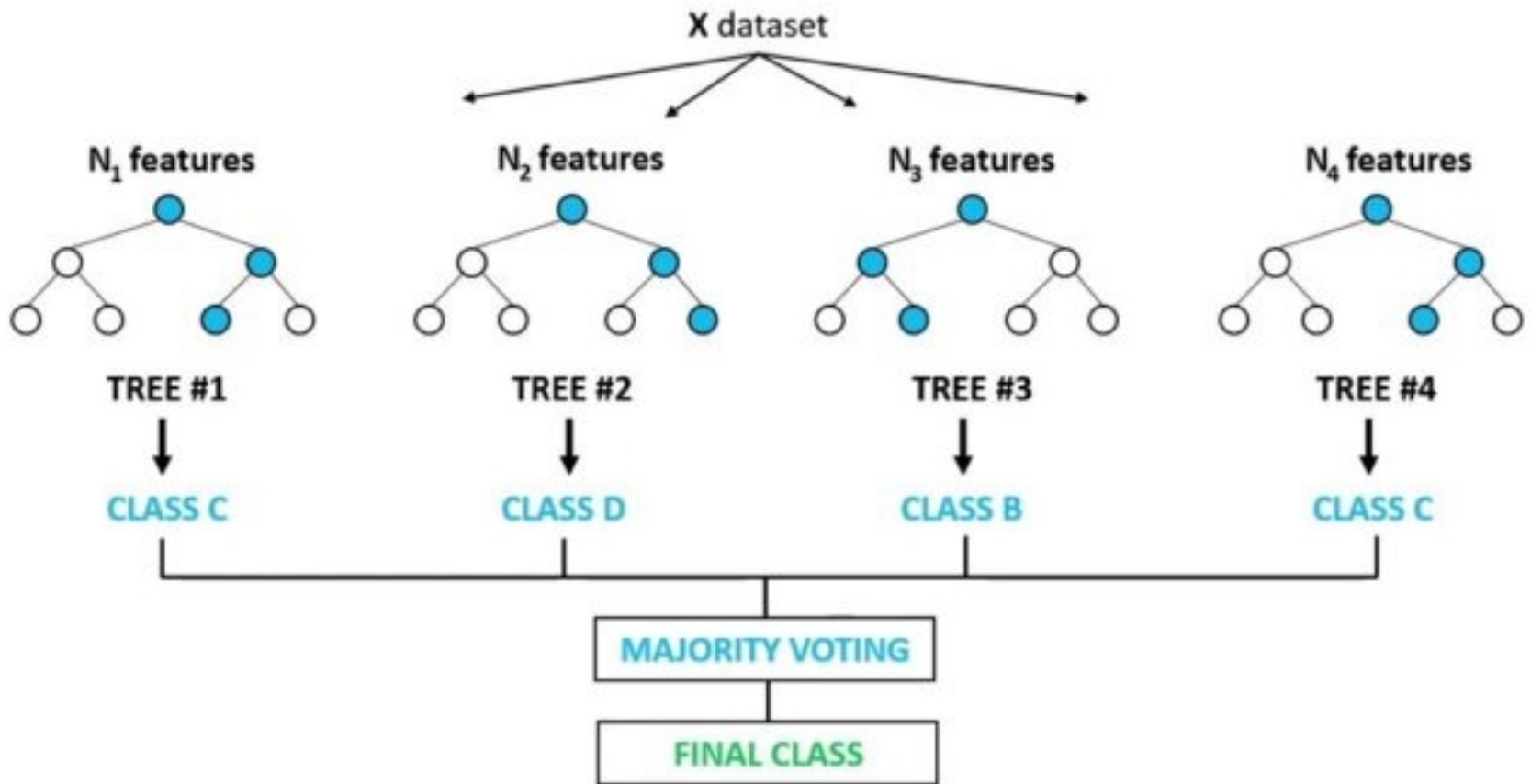
Как решить дилемму дисперсии-смещения:

- Композиции алгоритмов
 - Алгебраический подход к построению корректных алгоритмов
 - Области компетентности
 - Багинг - bagging
 - Бустинг - boosting

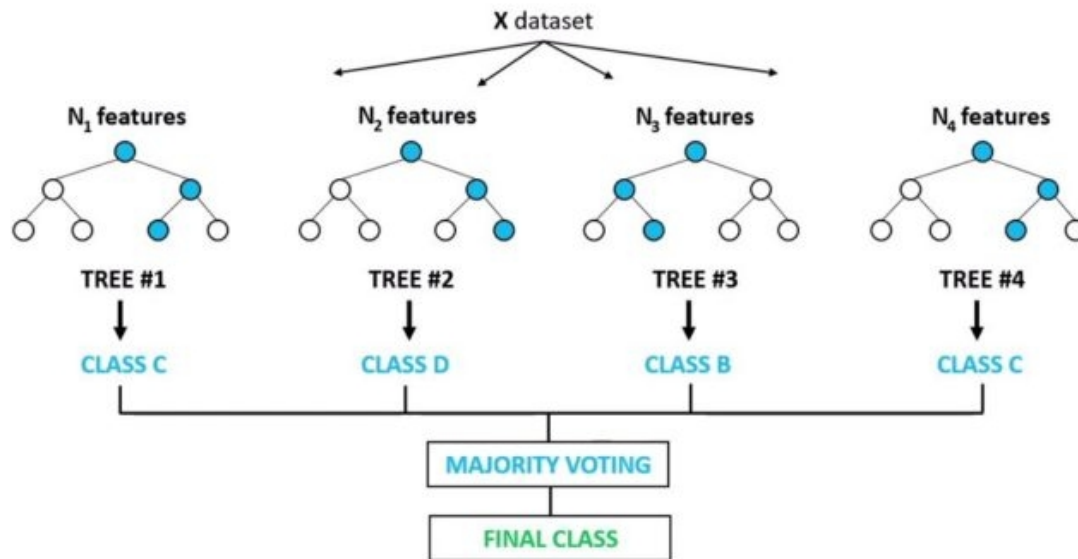
Бэгинг



Ансамбль Random Forest



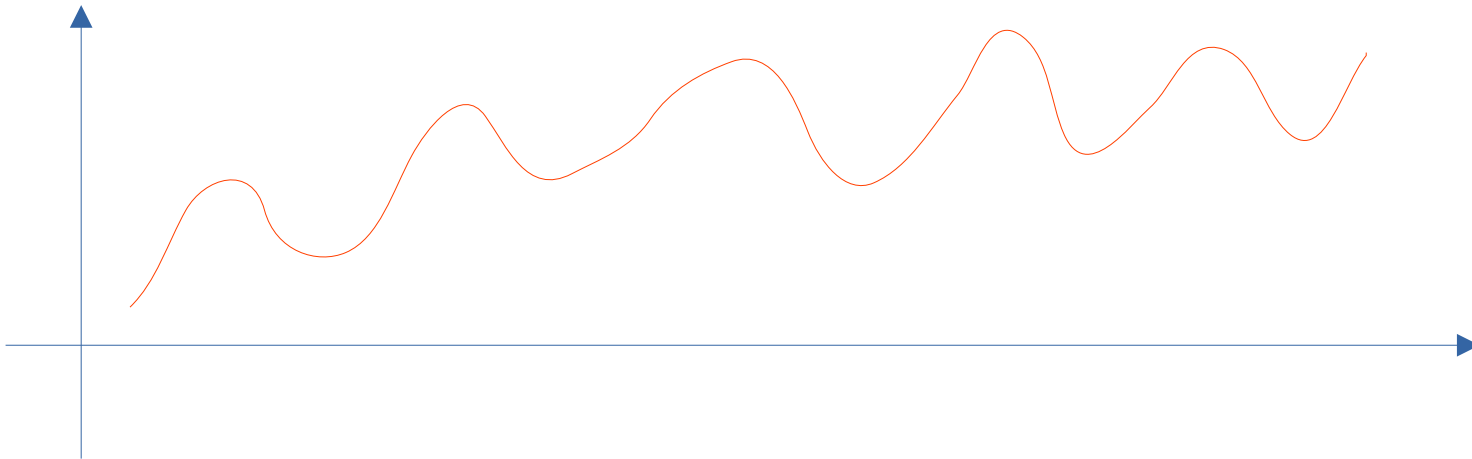
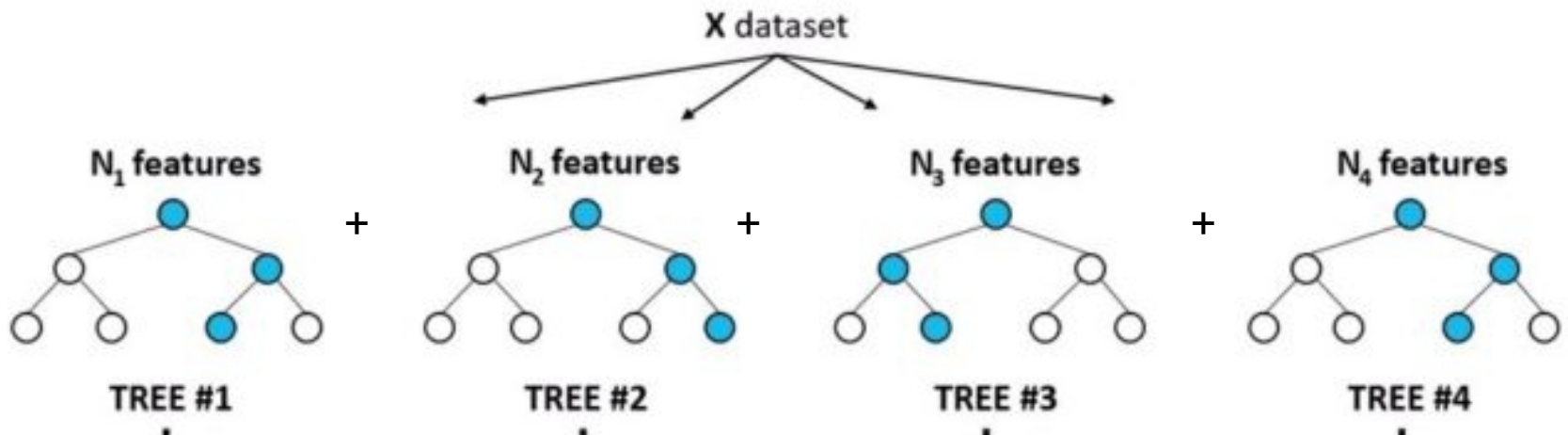
Ансамбль Random Forest



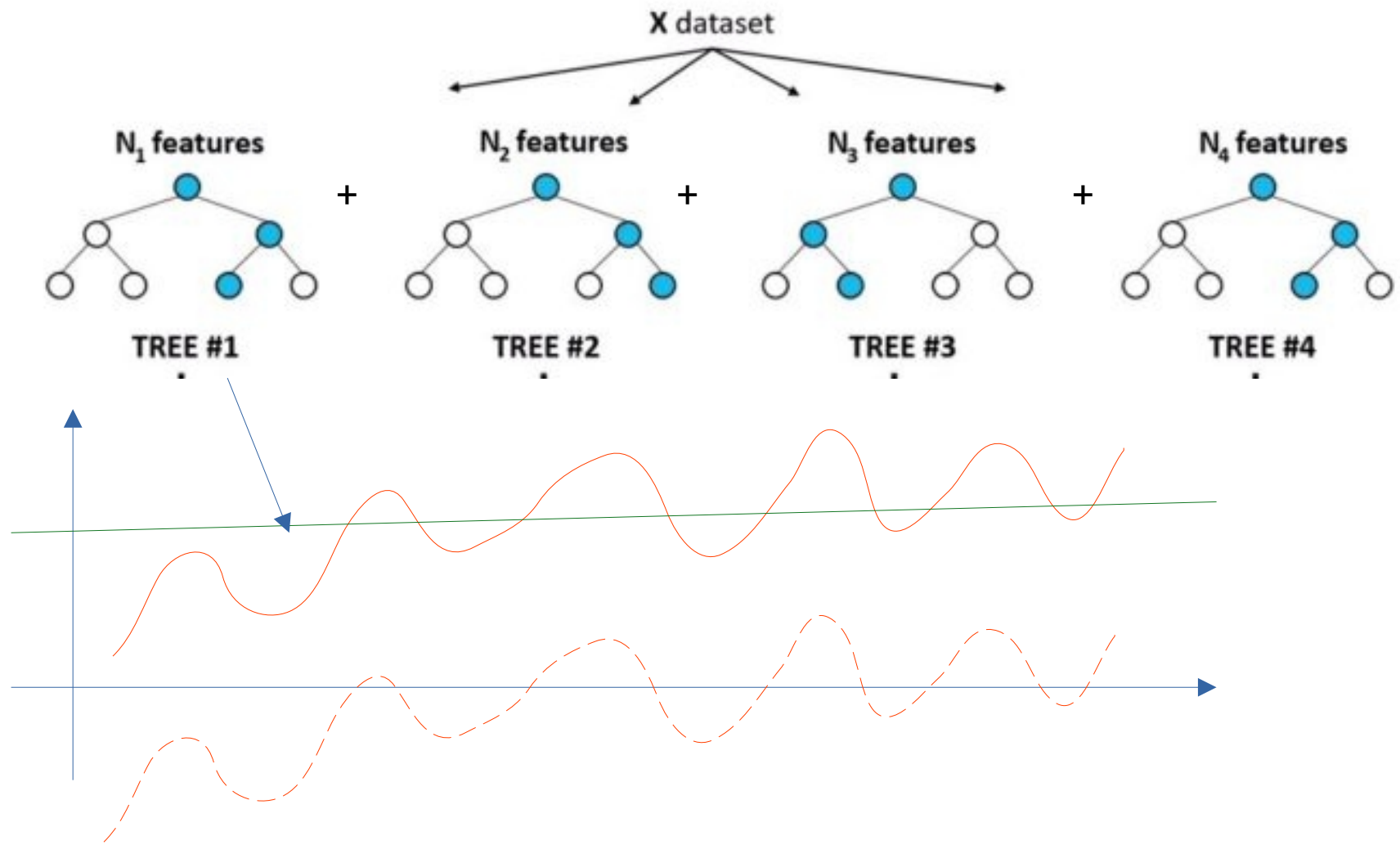
$$a_N(x) = \frac{1}{N} \sum_{n=1}^N b_n(x)$$

решающее дерево $b_n(x)$ по выборке \tilde{X}_n :

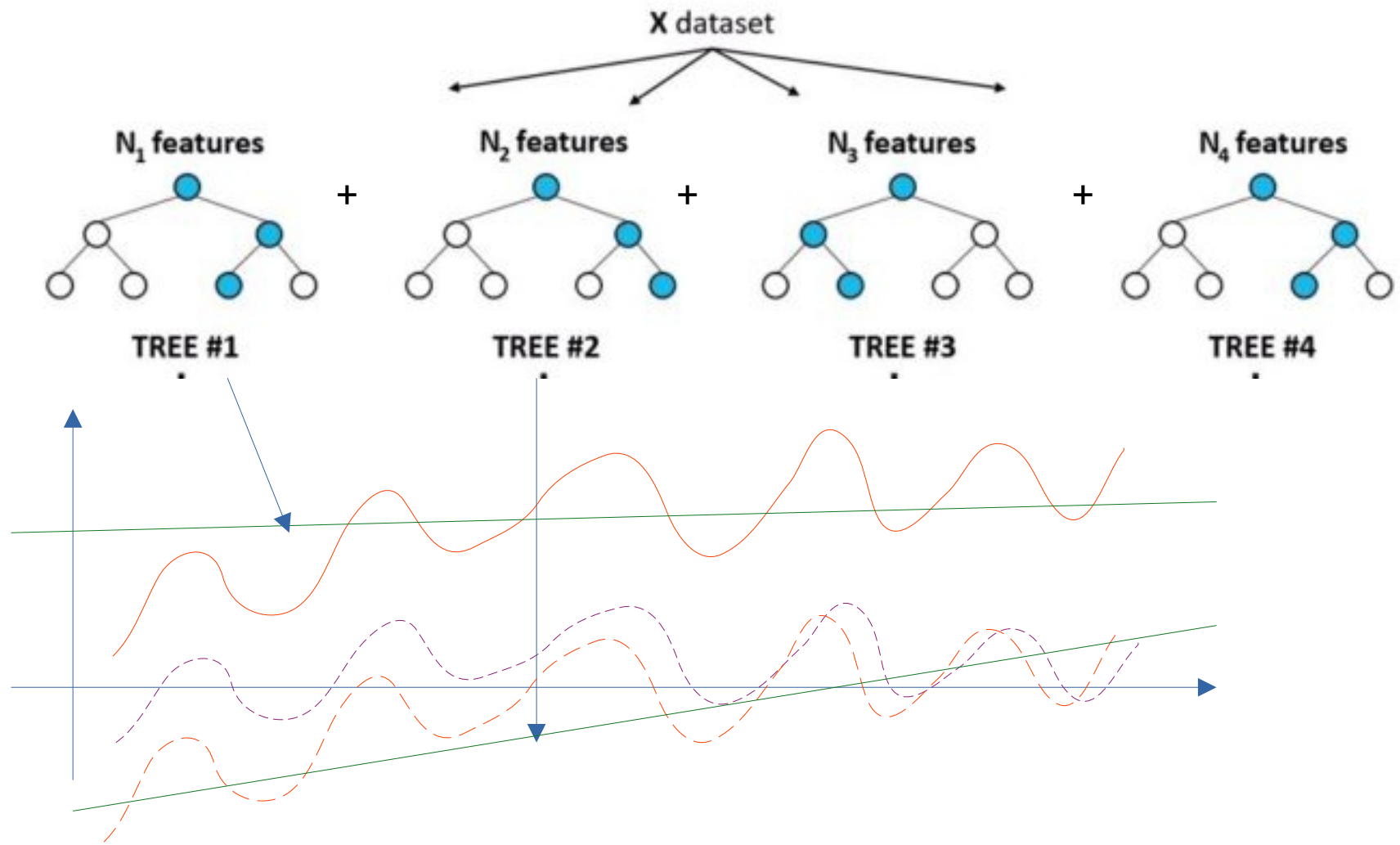
Ансамбль усиления



Ансамбль усиления



Ансамбль усиления



Ансамбль усиления Boosting

Потери:

$$\frac{1}{2} \sum_{i=1}^{\ell} (a(x_i) - y_i)^2 \rightarrow \min_a$$

Искомый алгоритм:

$$a_N(x) = \sum_{n=1}^N b_n(x),$$

b_n принадлежат некоторому семейству \mathcal{A} .

Ансамбль усиления Boosting

Первый шаг

$$b_1(x) := \arg \min_{b \in \mathcal{A}} \frac{1}{2} \sum_{i=1}^{\ell} (b(x_i) - y_i)^2$$

Остатки:

$$s_i^{(1)} = y_i - b_1(x_i)$$

Ансамбль усиления Boosting

Новый шаг :

$$b_2(x) := \arg \min_{b \in \mathcal{A}} \frac{1}{2} \sum_{i=1}^{\ell} (b(x_i) - s_i^{(1)})^2$$

Ансамбль усиления Boosting

Все следующие:

$$s_i^{(N)} = y_i - \sum_{n=1}^{N-1} b_n(x_i) = y_i - a_{N-1}(x_i), \quad i = 1, \dots, \ell;$$

$$b_N(x) := \arg \min_{b \in \mathcal{A}} \frac{1}{2} \sum_{i=1}^{\ell} (b(x_i) - s_i^{(N)})^2$$

Остатки

$$s_i^{(N)} = y_i - a_{N-1}(x_i) = - \left. \frac{\partial}{\partial z} \frac{1}{2} (z - y_i)^2 \right|_{z=a_{N-1}(x_i)}$$

Ансамбль усиления AdaBoost

Взвешивает примеры выборки X весами D :

1. $D = 1/l$

2. $b_n = \operatorname{argmin} \epsilon_j$

$$\epsilon_j = \sum_{i=1}^l D_n(i) [y_i \neq b_j(x)]$$

пока не наступит $\epsilon_j \geq 0.5$.

Алгоритм $\alpha_n = \frac{1}{2} \ln \frac{1 - \epsilon_n}{\epsilon_n}$

, Веса

$$D_{n+1}(i) = \frac{D_n(i) e^{-\alpha_n y_i b_n(x_i)}}{Z_n}$$

$$a(x) = \operatorname{sign} \left(\sum_{n=1}^N \alpha_n b_n(x) \right)$$

Ансамбль усиления Градиентный бустинг

Модель:

$$a_N(x) = \sum_{n=1}^N \gamma_n b_n(x).$$

Инициализация b_1 :

- константа
- среднее (самое частое)

Ансамбль усиления

Градиентный бустинг

Цель:

$$\frac{1}{l} \sum_{i=1}^l (a(x_i) - y_i)^2 \rightarrow \min.$$

b1:

$$b_1(x) = \operatorname{argmin}_b \frac{1}{l} \sum_{i=1}^l (b(x_i) - y_i)^2.$$

Остатки s1:

$$s_i^{(1)} = y_i - b_1(x_i).$$

Ансамбль усиления Градиентный бустинг

b2:

$$b_2(x) = \operatorname{argmin}_b \frac{1}{l} \sum_{i=1}^l (b(x_i) - s_i^{(1)})^2 = \operatorname{argmin}_b \frac{1}{l} \sum_{i=1}^l (b(x_i) - (y_i - b_1(x_i)))^2.$$

bN:

$$b_N(x) = \operatorname{argmin}_b \frac{1}{l} \sum_{i=1}^l (b(x_i) - s_i^{(N)})^2,$$

$$s_i^{(N)} = y_i - \sum_{n=1}^{N-1} b_n(x_i) = y_i - a_{N-1}(x_i).$$

Ансамбль усиления Градиентный бустинг

Вектор сдвига s_N :

$$\sum_{i=1}^I L(y_i, a_{N-1}(x_i) + s_i) \rightarrow \min_s,$$

$$z = a_{N-1}(x_i)$$

$$s_i = - \left. \frac{\partial L}{\partial z} \right|_{z=a_{N-1}(x_i)}.$$

Ансамбль усиления

Градиентный бустинг

Алгоритм bN:

$$b_N(x) = \underset{s}{\operatorname{argmin}} \frac{1}{l} \sum_{i=1}^l (b(x_i) - s_i)^2.$$

$$\gamma_N = \underset{\gamma}{\operatorname{argmin}} \sum_{i=1}^l L(y_i, a_{N-1}(x_i) + \gamma b_N(x_i)).$$

$$a_N(x) = a_{N-1}(x) + \eta \gamma_N b_N(x).$$