

# Оптимизация и Интерпретация моделей ML

2020

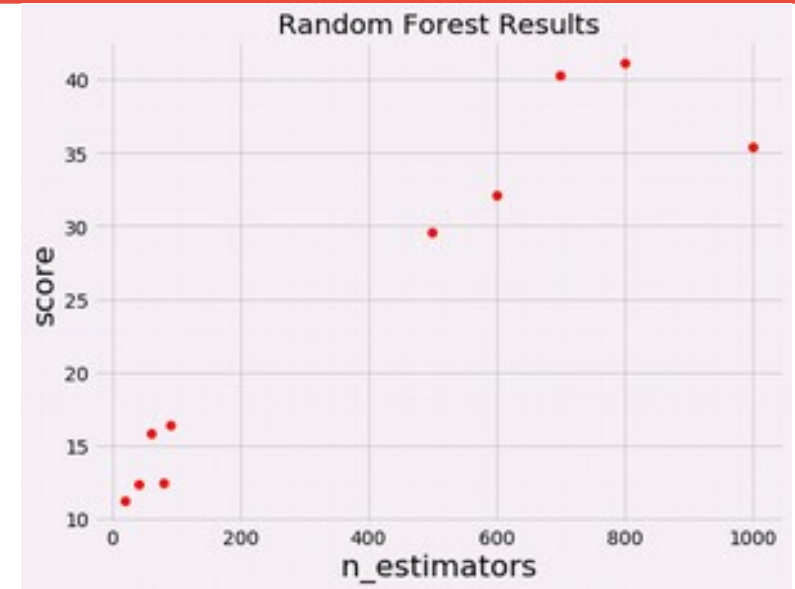
# Поиск гиперпараметров

$$\text{CVLoss}(\mathbf{h}) := \text{Loss}(\mathbf{X}_v, \mathbf{Y}_v; \mathbf{a}^*(\mathbf{h}), \mathbf{h}),$$
$$\mathbf{a}^*(h) := \operatorname{argmin}_a \{ \text{Loss}(\mathbf{X}_t, \mathbf{Y}_t, \mathbf{a}, \mathbf{h}) \}$$

- Поиск на сетке (Grid Search, Randomized Search)
- Покоординатный спуск (Coordinate Descent)
- Генетические алгоритмы (Genetic Algorithms)
- Байесовская оптимизация

# Поиск гиперпараметров

$$\text{CVLoss}(\mathbf{h}) := \text{Loss}(\mathbf{X}_v, \mathbf{Y}_v; \mathbf{a}^*(\mathbf{h}), \mathbf{h}),$$
$$\mathbf{a}^*(\mathbf{h}) := \operatorname{argmin}_{\mathbf{a}} \{ \text{Loss}(\mathbf{X}_t, \mathbf{Y}_t, \mathbf{a}, \mathbf{h}) \}$$



Байесовская оптимизация:

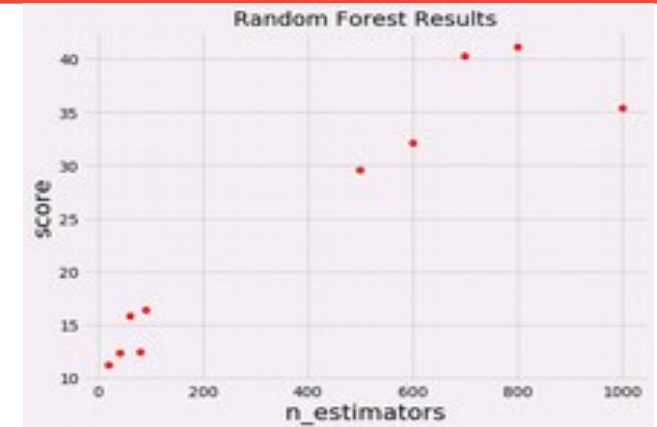
- Построить суррогатную вероятностную модель целевой функции
- Найдите гиперпараметры, которые лучше всего работают на суррогате
- Примените эти гиперпараметры к истинной целевой функции
- Обновление суррогатной модели, включающей новые результаты
- Повторите шаги 2–4, пока не будет достигнуто максимальное количество итераций или время

# Поиск гиперпараметров

$$\text{CVLoss}(\mathbf{h}) := \text{Loss}(\mathbf{X}_v, \mathbf{Y}_v; \mathbf{a}^*(\mathbf{h}), \mathbf{h}),$$
$$\mathbf{a}^*(\mathbf{h}) := \operatorname{argmin}_{\mathbf{a}} \{ \text{Loss}(\mathbf{X}_t, \mathbf{Y}_t, \mathbf{a}, \mathbf{h}) \}$$

## Методы последовательной оптимизации на основе моделей (SMBO) :

- Определить область для поиска гиперпараметров
- Определить Целевую функцию, которая принимает гиперпараметры и выводит оценку, которую мы хотим минимизировать (или максимизировать)
- Собрать суррогатная модель целевой функции
- Определить Критерий (функцию выбора), для оценки того, какие гиперпараметры выбрать следующим из суррогатной модели.
- История, состоящая из пар (оценка, гиперпараметр), используемых алгоритмом для обновления суррогатной модели



# Поиск гиперпараметров

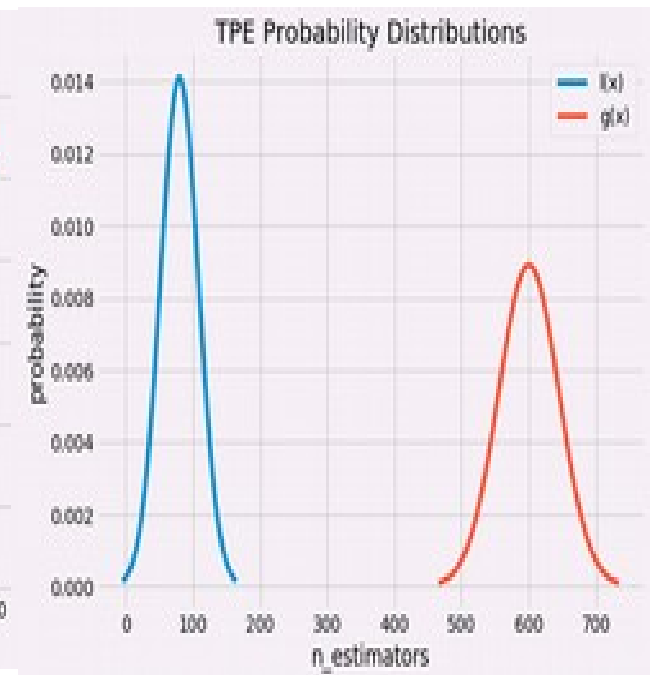
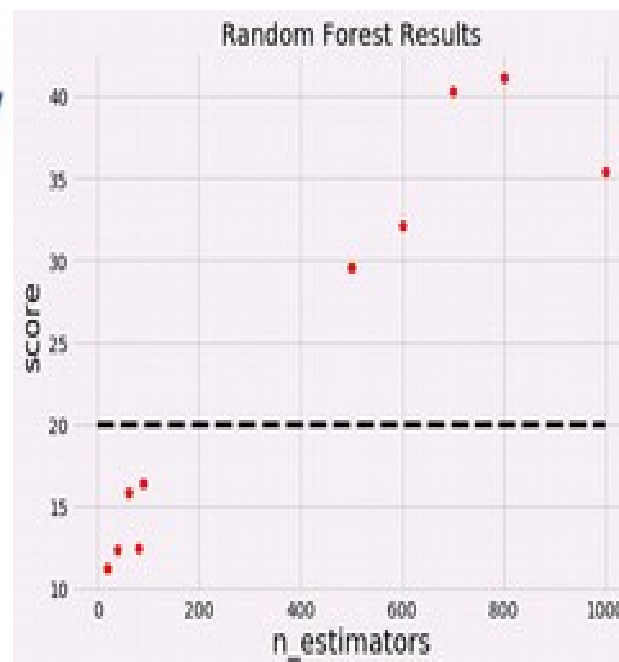
## Методы последовательной оптимизации на основе моделей (SMBO) :

- Древовидная структура Парзена (TPE)

$$EI_{y^*}(x) = \int_{-\infty}^{y^*} (y^* - y)p(y|x)dy$$

$$p(y|x) = \frac{p(x|y) * p(y)}{p(x)}$$

$$p(x|y) = \begin{cases} \ell(x) & \text{if } y < y^* \\ g(x) & \text{if } y \geq y^* \end{cases}$$



# Интерпретация моделей ML

Yellowbrick

ELI5

MLxtend

LIME

SHAP,

# Вектор ШЕПЛИ

SHAP - SHapley Additive explanation

$$\Phi(v)_i = \sum_{K \ni i} \frac{(k-1)!(n-k)!}{n!} (v(K) - v(K \setminus i)),$$

где  $n$  - количество игроков,  $k$  - количество участников коалиции  $K$ .

# Вектор ШЕПЛИ - интерпретация ML-моделей

- Результат обучения с учителем (на основе заданного примера) – это игра;
- Выигрыш – это разница между матожиданием результата на всех имеющихся примерах и результатом, полученном на заданном примере;
- Вклады игроков в игру – влияние каждого значения признака на выигрыш, т.е. результат.

$$\Phi(v)_i = \sum_{K \ni i} \frac{(k-1)!(n-k)!}{n!} (v(K) - v(K \setminus i)),$$

где  $n$  – количество признаков,

$k$  – количество признаков коалиции  $K$ .