






Введение

Обучающая выборка / Эмпирический риск

Обучающая выборка:

$\mathbf{z} :$

$x_1 :$		$x_2 :$		$x_3 :$		$x_4 :$		$x_5 :$	
$y_1 :$	cat	$y_2 :$	cat	$y_3 :$	dog	$y_4 :$	cat	$y_5 :$	dog

Эмпирический риск:

$$r(h, \mathbf{z}) := \frac{1}{n} \sum_{i=1}^n l(h(x_i), y_i),$$

- h – предиктор;
- l – функция потерь.

Введение

Невидимые данные / Ожидаемый риск

Невидимые (будущие) данные:



Ожидаемый риск:

$$R(h) := \mathbf{E}_{(x,y) \sim P} [l(h(x), y),].$$

- P – неизвестное, но фиксированное распределение.

Введение

Декомпозиция ошибки обучения

$$R(h_{\mathbf{z}}) = [R(h_{\mathbf{z}}) - r(h_{\mathbf{z}}, \mathbf{z})] + [r(h_{\mathbf{z}}, \mathbf{z}) - r(h^*, \mathbf{z})] + [r(h^*, \mathbf{z}) - R(h^*)] + R(h^*)$$

Bartlett P.L., Montanari A., Rakhlin A. Deep learning: a statistical viewpoint, *Acta Numerica*, 2021, vol. 30, pp. 87-201.

- \mathcal{H} – класс предикторов (модель);
- $\mathbf{z} \mapsto h_{\mathbf{z}} \in \mathcal{H}$ – алгоритм обучения;
- $h^* \in \mathcal{H}$ – минимальный предиктор.

Введение

Декомпозиция ошибки обучения

$$R(h_z) = \underbrace{[R(h_z) - r(h_z, \mathbf{z})]}_{(1)} + \underbrace{[r(h_z, \mathbf{z}) - r(h^*, \mathbf{z})]}_{(2)} + \underbrace{[r(h^*, \mathbf{z}) - R(h^*)]}_{(3)} + \underbrace{R(h^*)}_{(4)}$$

- (1), (3) – способность к обобщению;
- (2) – оптимизация;
- (4) – аппроксимация.