

Минимизация эмпирического риска

Метод упорядоченной минимизации риска

Определения

Метод минимизации эмпирического риска - это общий подход к решению широкого класса задач обучения по прецедентам

Задача обучения по прецедентам.

X - множество описаний объектов.

Y - множество допустимых ответов.

Существует неизвестная целевая зависимость $y^* : X \rightarrow Y$, значения которой известны только на объектах конечной обучающей выборки $X^m = \{(x_1, y_1), \dots, (x_m, y_m)\}$.

Задача - построение алгоритма $a: X \rightarrow Y$, который бы приближал неизвестную целевую зависимость.

Функция потерь и эмпирический риск

Функция потерь - $\mathcal{L}(y, y')$, величина отклонения ответа $y = a(x)$ от правильного ответа $y' = y^*(x)$ на произвольном объекте $x \in X$.

Вводится **модель алгоритмов** $A = \{a: X \rightarrow Y\}$ в рамках которой будет вестись поиск отображения, приближающего неизвестную целевую зависимость.

Эмпирический риск — это функционал качества, характеризующий среднюю ошибку алгоритма a на выборке X^m :

$$Q(a, X^m) = \frac{1}{m} \sum_{i=1}^m \mathcal{L}(a(x_i), y^*(x_i)).$$

Минимизация эмпирического риска

Метод минимизации эмпирического риска заключается в том, чтобы в заданной модели алгоритмов \mathbf{A} найти алгоритм, доставляющий минимальное значение функционалу эмпирического риска:

$$\alpha = \operatorname{argmin}_{\alpha \in A} Q(\alpha, X^m).$$

Разновидности функции потерь:

1. Классификация $\mathcal{L}(y, y') = [y' \neq y]$.
2. Регрессия $\mathcal{L}(y, y') = (y' - y)^2$.

Достоинства:

Конструктивный и универсальный подход, позволяющий сводить задачу обучения к задачам численной оптимизации.

Недостатки:

Переобучение, которое возникает практически всегда при использовании метода минимизации эмпирического риска.

