

Обучение с учителем. Классификация. Дискриминантный анализ.

Е. Ларин, Ф. Ежов, И. Кононыхин

Санкт-Петербургский государственный университет
Прикладная математика и информатика
Вычислительная стохастика и статистические модели

Обучение с учителем

Выборка из генеральной случайной величины

- Для задачи регрессии: $\mathbf{X} \in \mathbb{R}^{n \times p}$, $\mathbf{y} \in \mathbb{R}^n$
- Для задачи классификации: $\mathbf{X} \in \mathbb{R}^{n \times p}$, $\mathbf{y} \in \mathbb{A}^n$

Обучение с учителем: формальная постановка

- *Вход:* \mathbf{X} — выборка ξ , \mathbf{y} — выборка η . Предполагаем, что существует неизвестное отображение $y^* : \xi \rightarrow \eta$ (гипотеза непрерывности или компактности)
- *Задача:* По \mathbf{X} и \mathbf{y} найти такое отображение $\hat{y}^* : \xi \rightarrow \eta$, которое приблизит отображение y^* .
- *Оценка:* Функция потерь $\mathcal{L}(y^*(x), \hat{y}^*(x))$. Здесь x — реализация ξ

Классификация

$$\mathbf{X} \in \mathbb{R}^{n \times p}, \mathbf{y} \in \mathbb{A}^n \quad (1)$$

Гипотеза компактности

«Близкие» объекты, как правило принадлежат одному классу