

MACHINE LEARNING AND AI

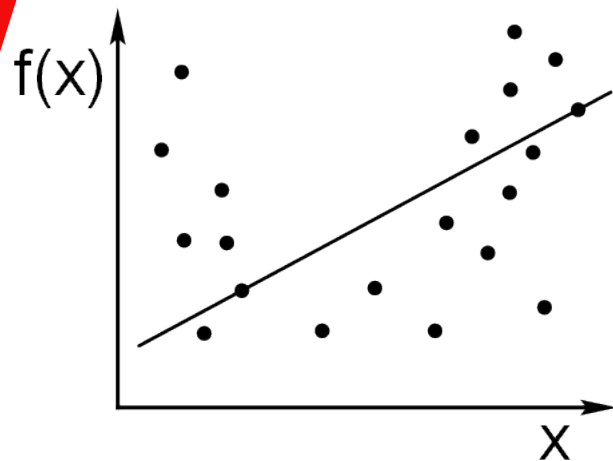
ЛЕКЦИЯ 2

ПЛАН ЛЕКЦИИ

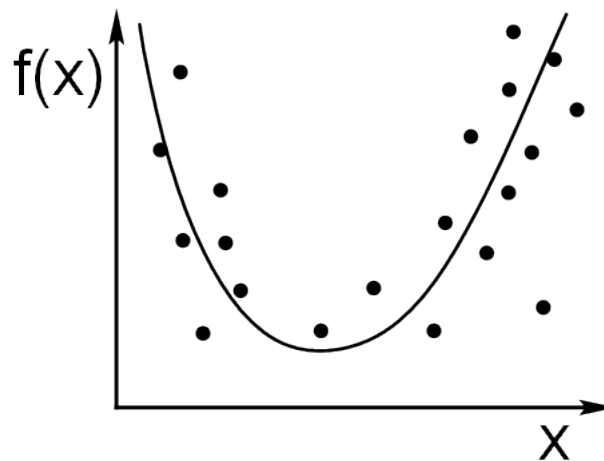
- Проблема недообучения и переобучения
- Регуляризация
- LASSO-регрессия и RIDGE-регрессия
- Кривые обучения
- Классификация: one vs all
- Метод опорных векторов

ПРОБЛЕМА НЕДООБУЧЕНИЯ И ПЕРЕОБУЧЕНИЯ

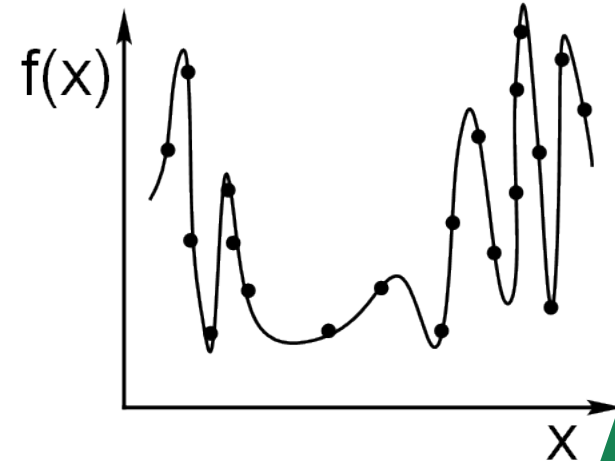
Недообучение



Оптимум



Переобучение



КАК ПОНЯТЬ, ЧТО МОДЕЛЬ НЕОПТИМАЛЬНА?

- Разбиваем множество примеров на 3 выборки:
 1. Обучающая выборка (~70 %)
 2. Валидационная выборка (~15 %)
 3. Тестовая выборка (~15 %)
- Полезно использовать стратификацию

РЕГУЛЯРИЗАЦИЯ

- Вспомним вид целевой функции

$$J(A) = \frac{1}{2} \sum_{i=1}^m (h_A(X) - Y)^2$$

$$h(X) = a_1x_1 + a_2x_2 + \dots + a_nx_n$$

- Добавляем к целевой функции добавочный член

$$J(A) = \frac{1}{2} \sum_{i=1}^m (h_A(X) - Y)^2 + \lambda \sum_{j=1}^n a_j^2$$

L1 и L2 регуляризация

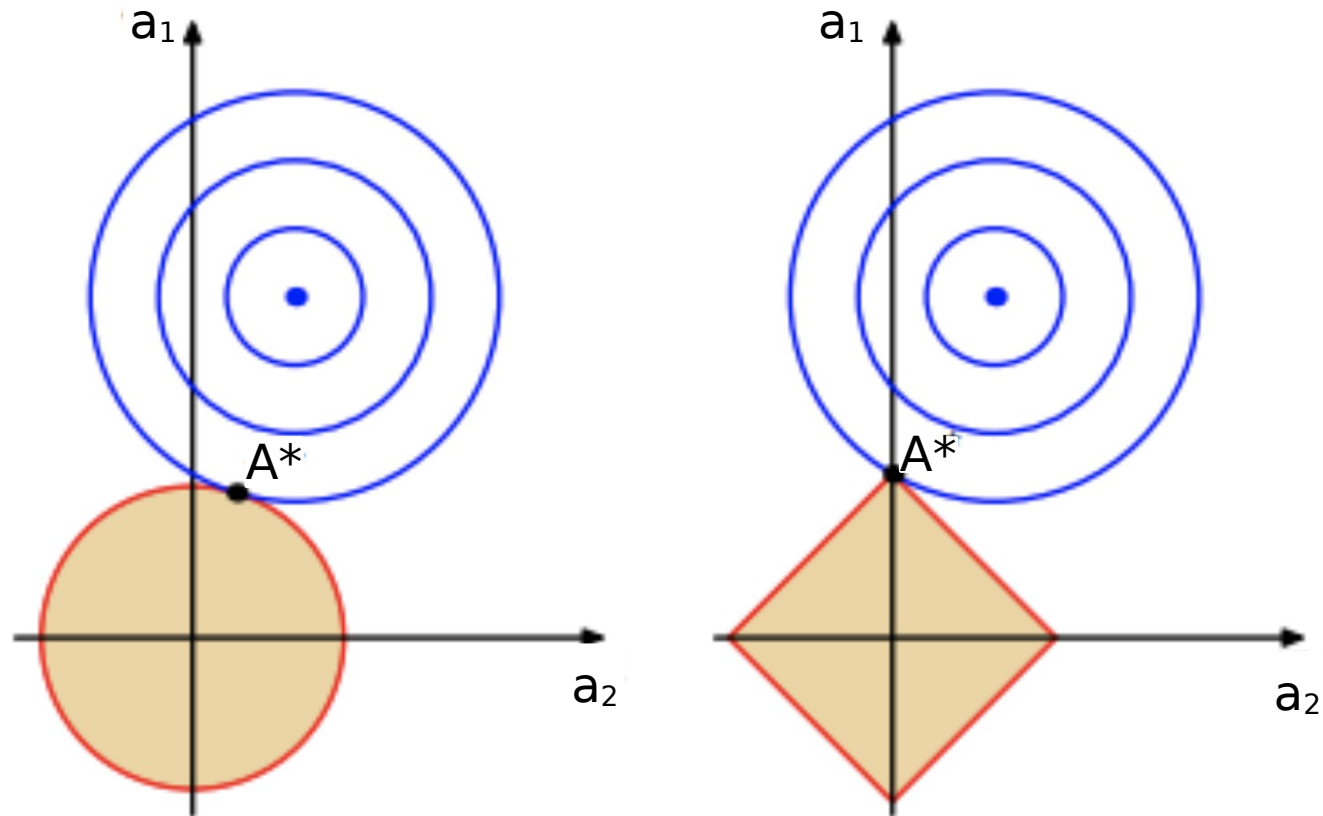
- L2 – Ridge regression

$$J(A) = \frac{1}{2} \sum_{i=1}^m (h_A(X) - Y)^2 + \lambda \sum_{j=1}^n a_j^2$$

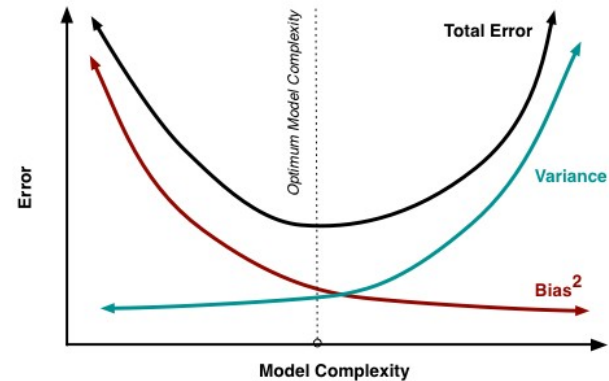
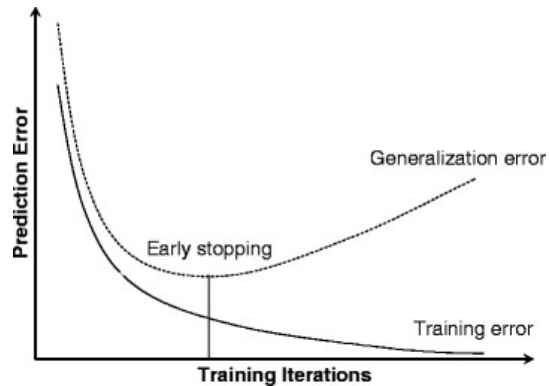
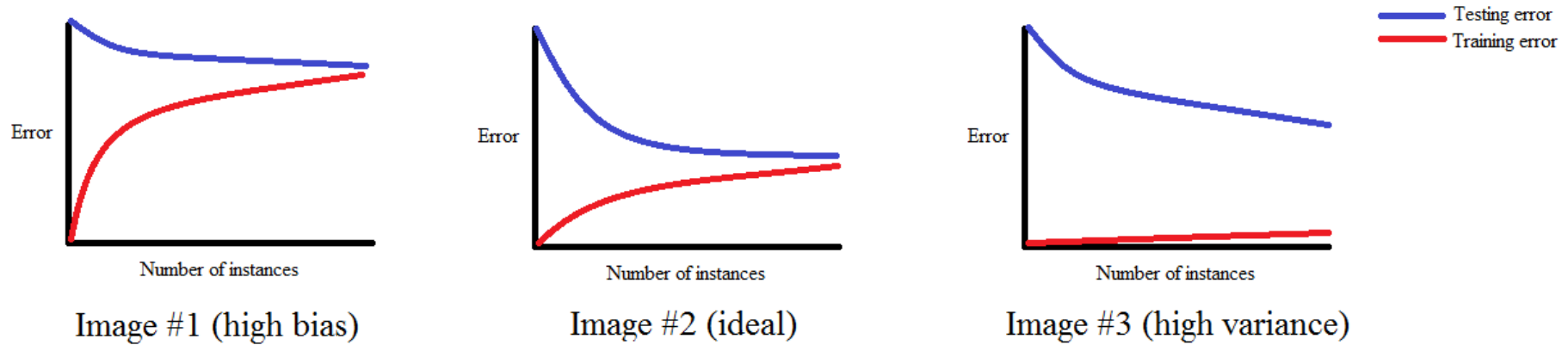
- L1 – Lasso regression

$$J(A) = \frac{1}{2} \sum_{i=1}^m (h_A(X) - Y)^2 + \lambda \sum_{j=1}^n |a_j|$$

Принципиальное отличие L1 и L2

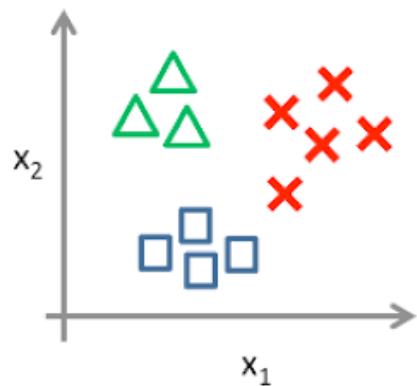


КРИВЫЕ ОБУЧЕНИЯ



Многоклассовая классификация

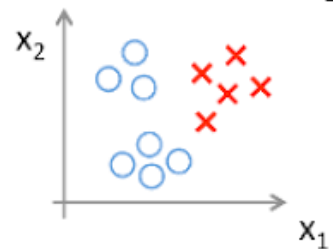
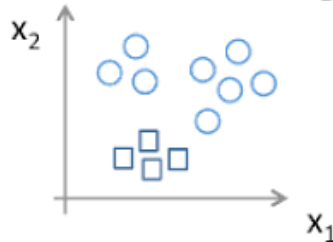
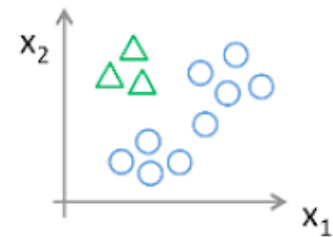
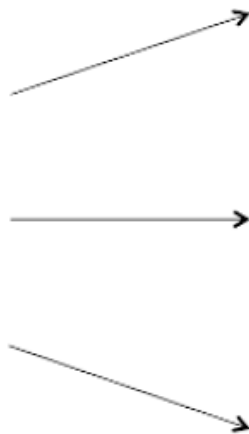
One-vs-all (one-vs-rest):



Class 1: Green

Class 2: Blue

Class 3: Red



Метод опорных векторов (SVM)

