

Методы машинного обучения

Лекция 10

Решающие деревья

Эльвира Зиннурова

elvirazinnurova@gmail.com

НИУ ВШЭ, 2019

Параметры и гиперпараметры

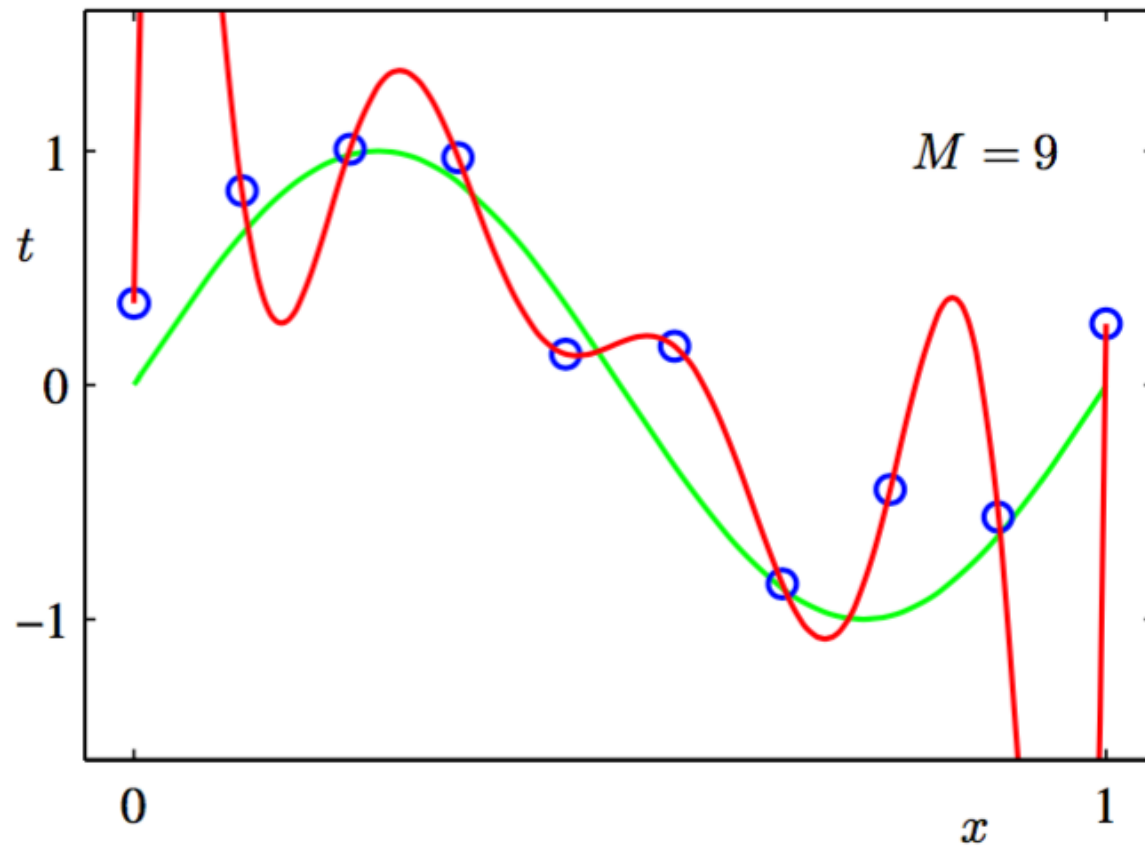
Простой пример

- Максимизируем удовлетворённость студентов
- Обучающая выборка — время до сессии
- Контрольная выборка — сессия
- Параметр — продолжительность лекции
- Гиперпараметр — минимальная продолжительность лекции

Простой пример

- Максимизируем удовлетворённость студентов
 - Обучающая выборка — время до сессии
 - Контрольная выборка — сессия
 - Параметр — продолжительность лекции
 - Гиперпараметр — минимальная продолжительность лекции
-
- Максимальная удовлетворённость на обучении — если не ограничивать продолжительность
 - Но оценки во время сессии будут ужасными

Переобучение



Регуляризация

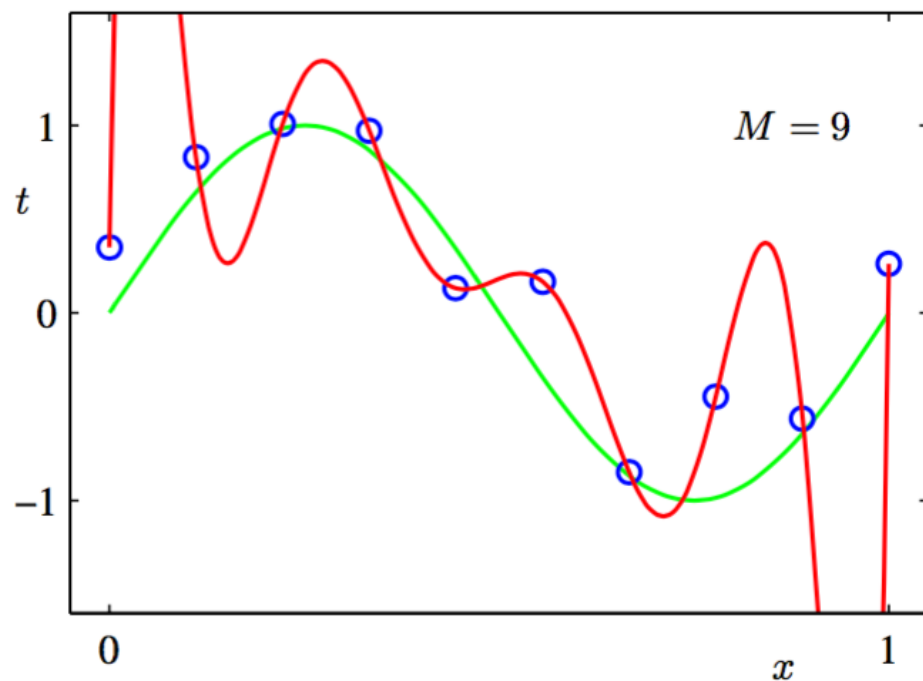
$$\frac{1}{\ell} \sum_{i=1}^{\ell} (\langle w, x_i \rangle - y_i)^2 + \lambda \|w\|^2 \rightarrow \min_w$$

Гиперпараметры

- Параметры модели — веса w
 - Позволяют подогнать модель под обучающую выборку
 - Настраиваются по обучающей выборке
- Гиперпараметр модели — коэффициент регуляризации λ
 - Определяют сложность модели
 - Лучшее качество на обучении достигается при $\lambda = 0$
 - Необходимо настраивать по другим данным — отложенная выборка, кросс-валидация и т.п.

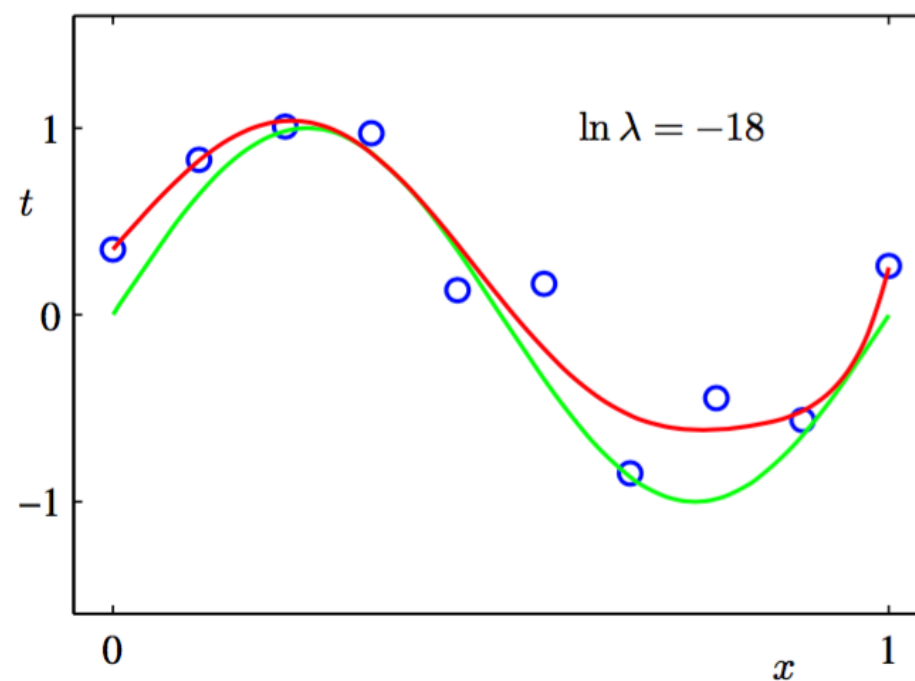
Гиперпараметры

Без регуляризации



Высокое качество на обучении

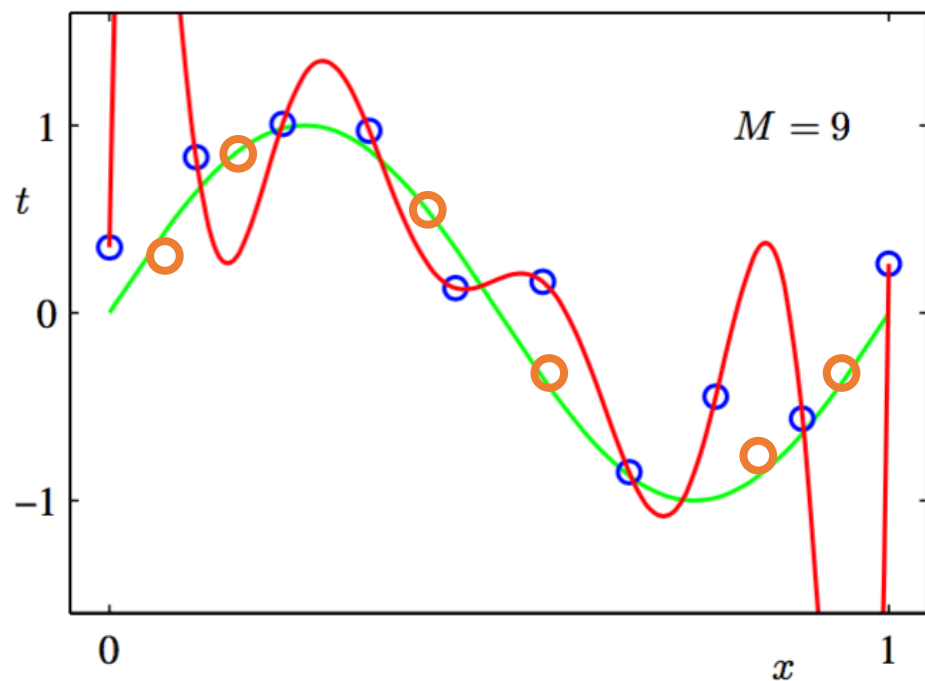
С регуляризацией



Качество на обучении ниже

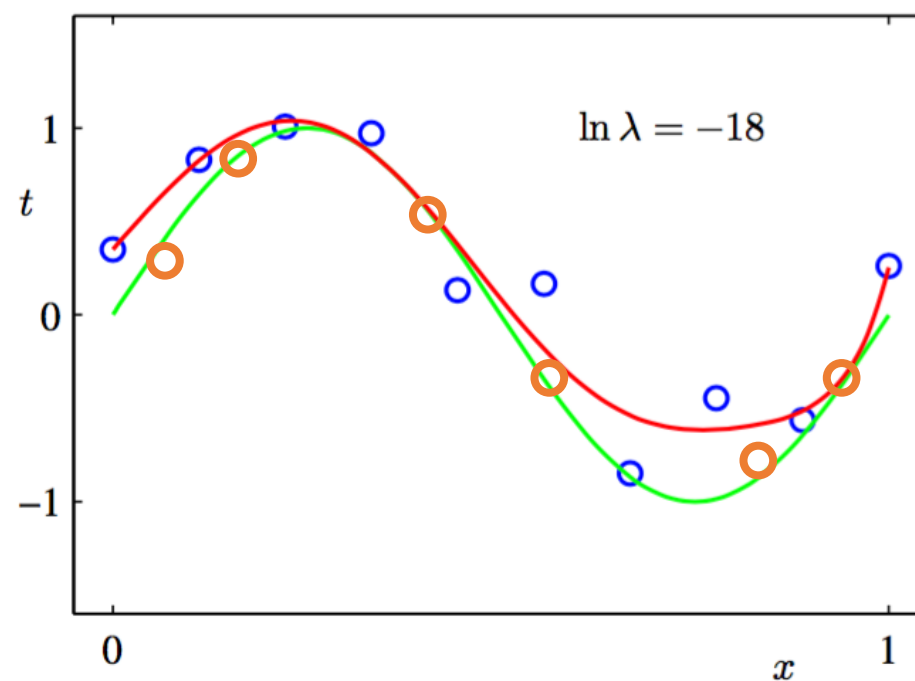
Гиперпараметры

Без регуляризации



Низкое качество на
тесте

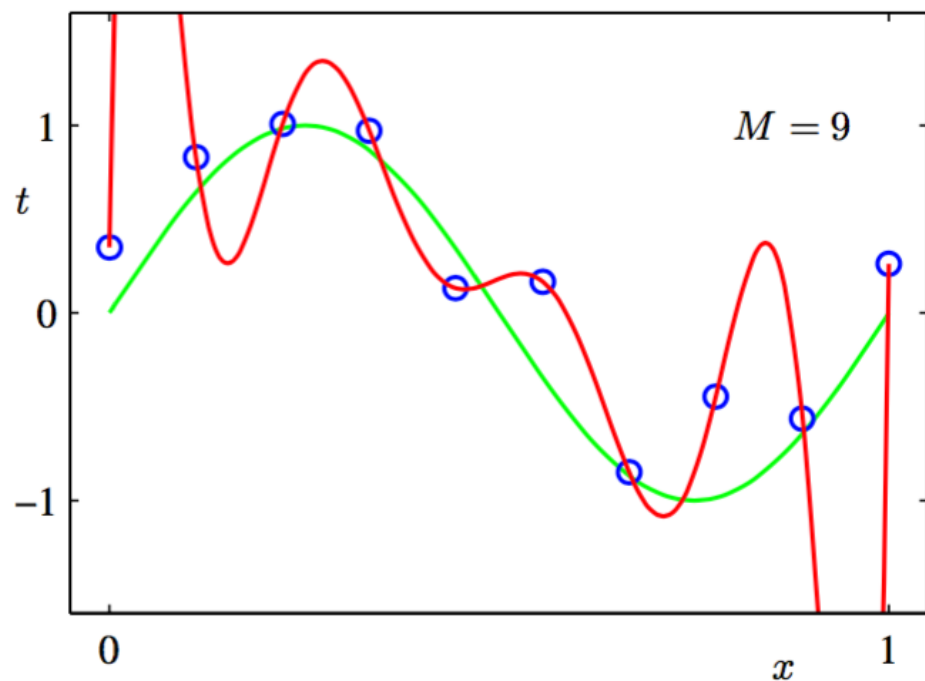
С регуляризацией



Высокое качество на
тесте

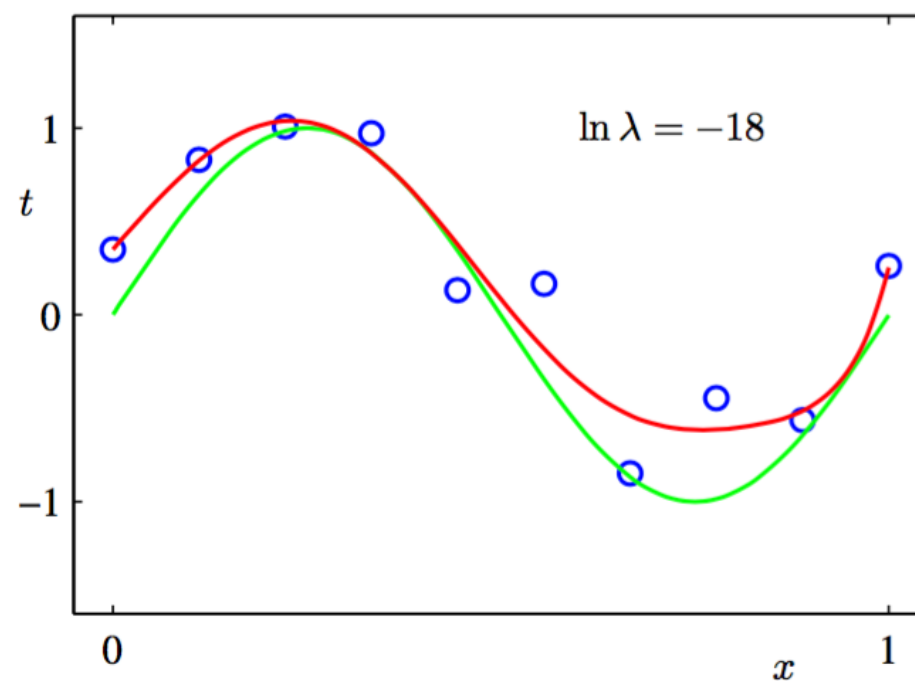
Гиперпараметры

Без регуляризации



Низкая обобщающая
способность

С регуляризацией

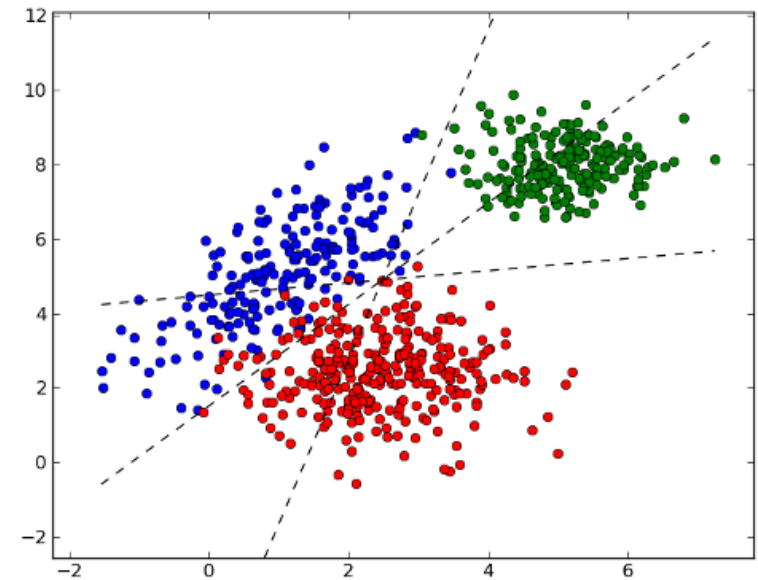


Высокая обобщающая
способность

Многоклассовые задачи

Многоклассовая классификация

- $\mathbb{Y} = \{1, 2, \dots, K\}$



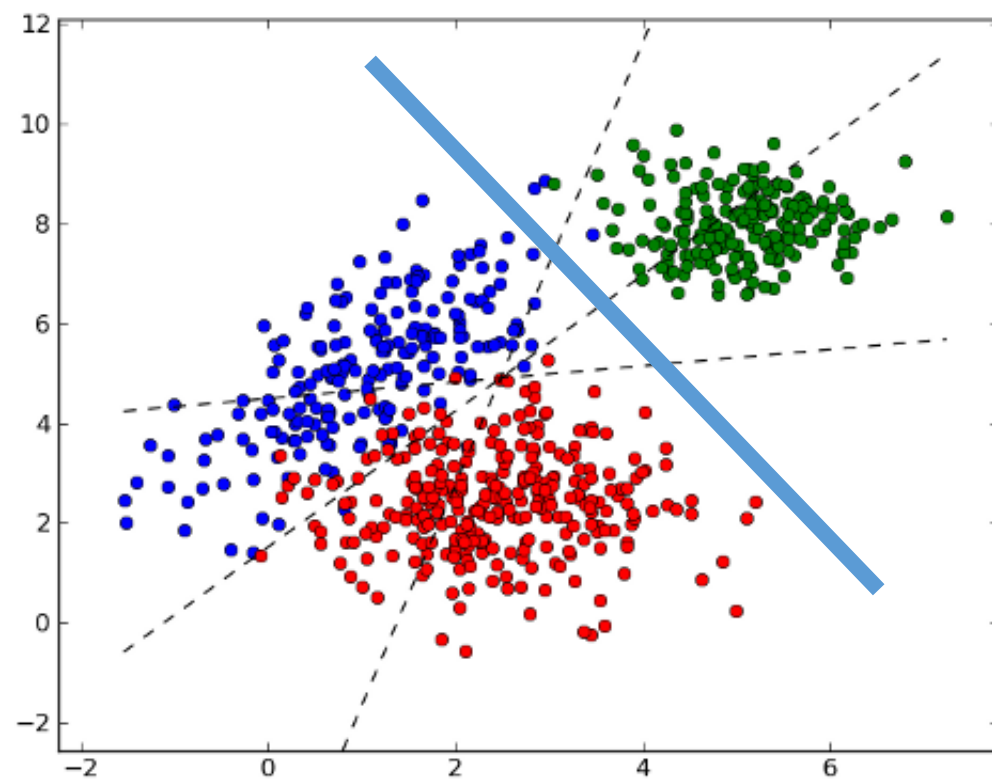
Бинарная классификация

$$a(x) = \text{sign } \langle w, x \rangle$$

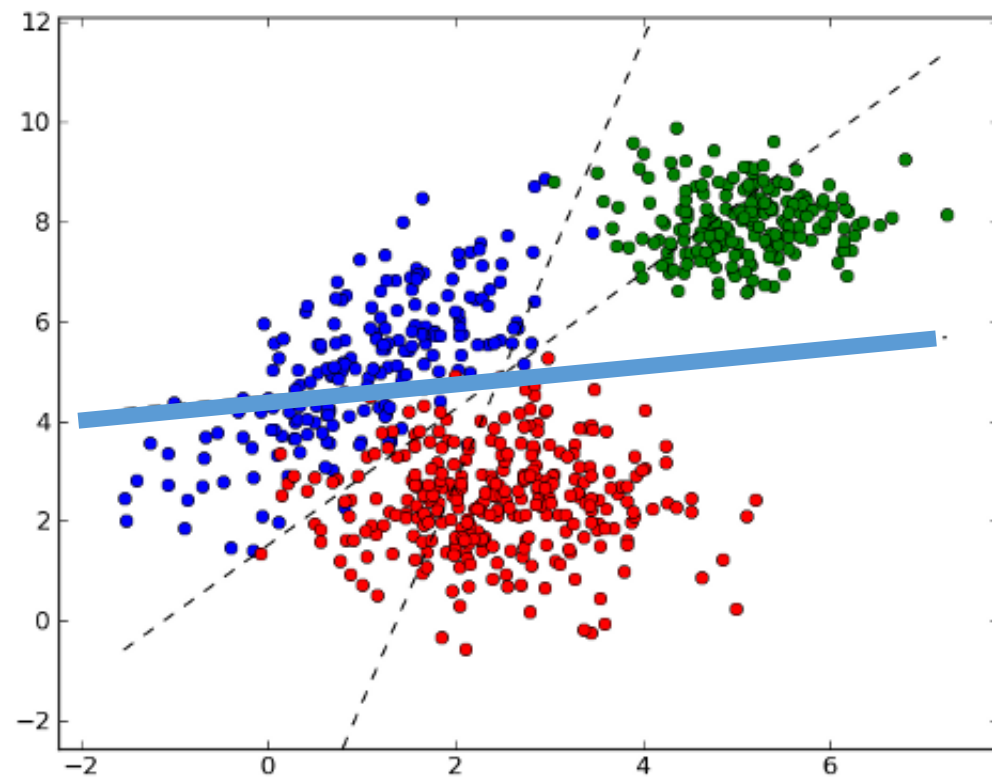
One-vs-all

- Способ сведения многоклассовой задачи к набору бинарных классификаций
- Обучаем свой классификатор для каждого класса
- Задача: отделение класса от всех остальных

One-vs-all



One-vs-all



One-vs-all

- K задач бинарной классификации
- k -я задача:
 - $X = (x_i, [y_i = k])_{i=1}^{\ell}$
 - Классификатор $a_k(x) = \text{sign } \langle w_k, x \rangle$
- Алгоритм:

$$a(x) = \arg \max_{k \in \{1, \dots, K\}} \langle w_k, x \rangle$$

Матрица ошибок

	$y = 1$	$y = 2$...	$y = K$
$a(x) = 1$	q_{11}	q_{12}	...	q_{1K}
$a(x) = 2$	q_{21}	q_{22}	...	q_{2K}
...
$a(x) = K$	q_{K1}	q_{K2}	...	q_{KK}

Доля правильных ответов

$$\text{accuracy}(a, X) = \frac{1}{\ell} \sum_{i=1}^{\ell} [a(x_i) = y_i]$$

Точность и полнота

- Относительно каждого класса
- Можно усреднить точность и полноту по всем классам
- Можно усреднить F-меру