

# Лекция 6. Деревья решений

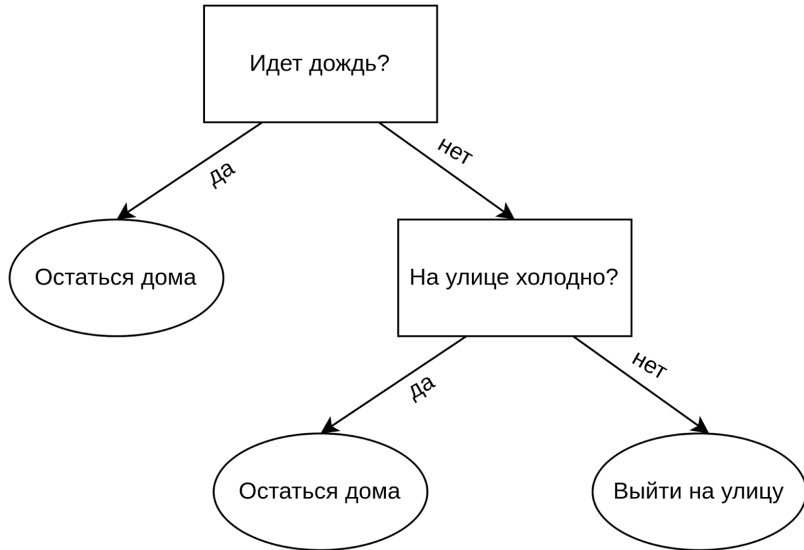
## Основы интеллектуального анализа данных

Полузёров Т. Д.

БГУ ФПМИ



# Пример дерева решений



# Компоненты бинарного дерева

Дерево  $T$  состоит из 2-х типов вершин:

- **Внутренняя** вершина  $v$  хранит в себе предикат  $b_v : \mathbb{X} \rightarrow \{0, 1\}$
- **Листовая вершина**  $v$  хранит выходное значение  $c_v \in \mathbb{Y}$

Алгоритм  $a(x)$  работает по схеме:

- 1 Стартуем из корня
- 2 Вычисляем предикат в текущей вершине
- 3 Если  $b_v = 1$  - шагаем в право,  $b_v = 0$  - в лево
- 4 Пока не дошли до листовой вершины, повторяем с шага 2
- 5 Возвращаем значение в листе  $c_v$

Предикат - любая решающая функция  $b : \mathbb{X} \rightarrow \{0, 1\}$

- Пороговая функция  $b(x) = [x_i > t]$
- Линейный  $b(x) = [\langle x, \omega \rangle > t]$
- Метрический  $b(x) = [\rho(x, x_v) > t]$ , где  $x_v$  - некоторый объект выборки

Но выбор сложных предикатов - излишен. Поэтому используются  $b(x) = [x_i > t]$

