



**Московский Государственный Университет
им. М.В. Ломоносова**

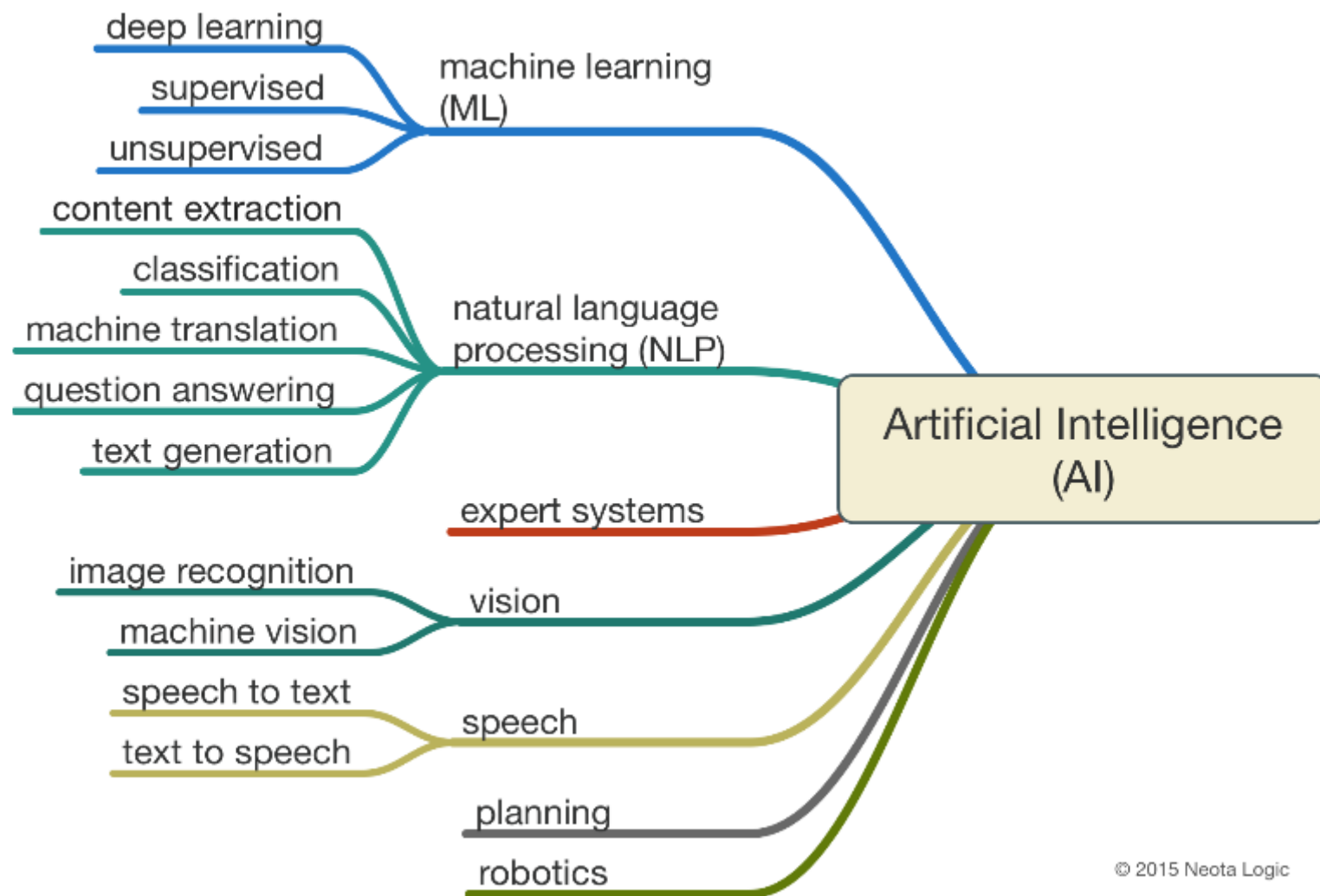
ВВЕДЕНИЕ В АНАЛИЗ ПРОСТРАНСТВЕННЫХ ДАННЫХ

Лекция №9. Основы интеллектуального анализа данных

**Николай Никифоров
Елизавета Хазиева
Николай Билев**

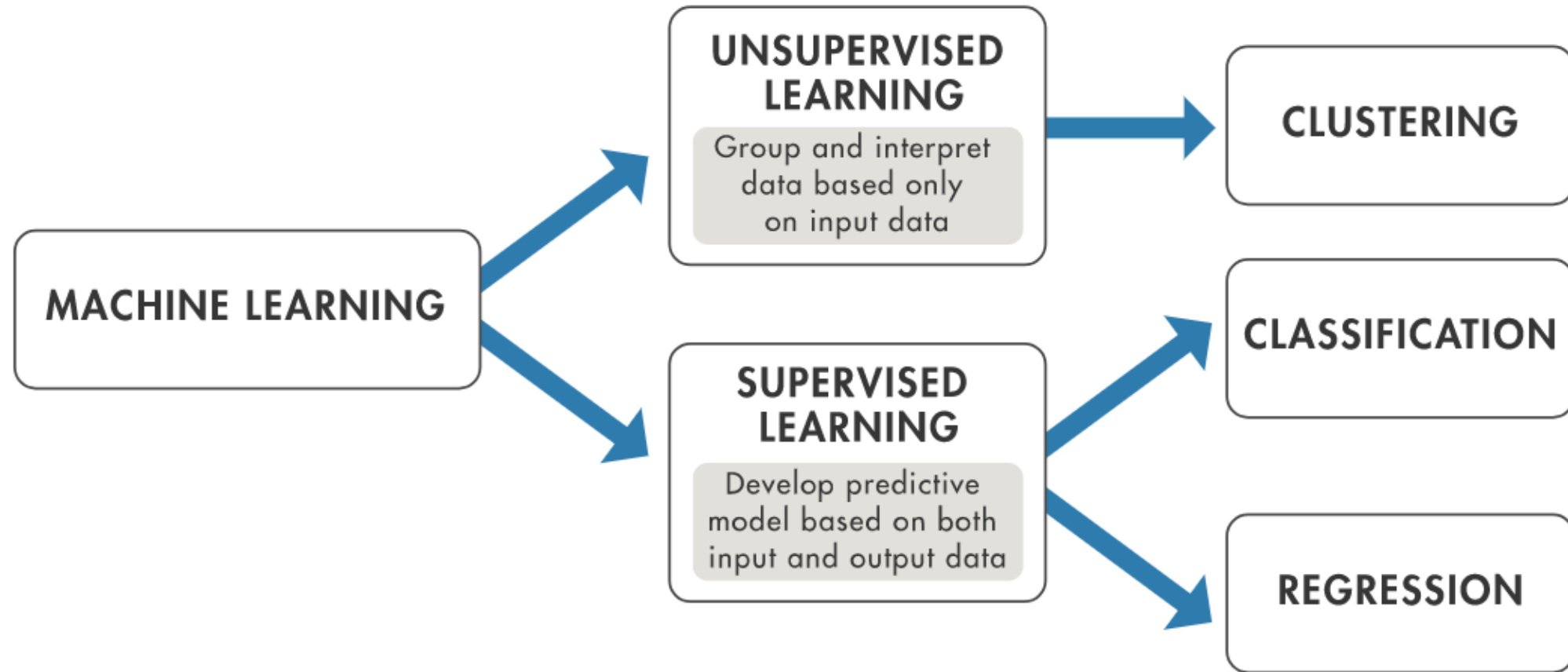
Москва 2019

Разделы искусственного интеллекта (AI)?



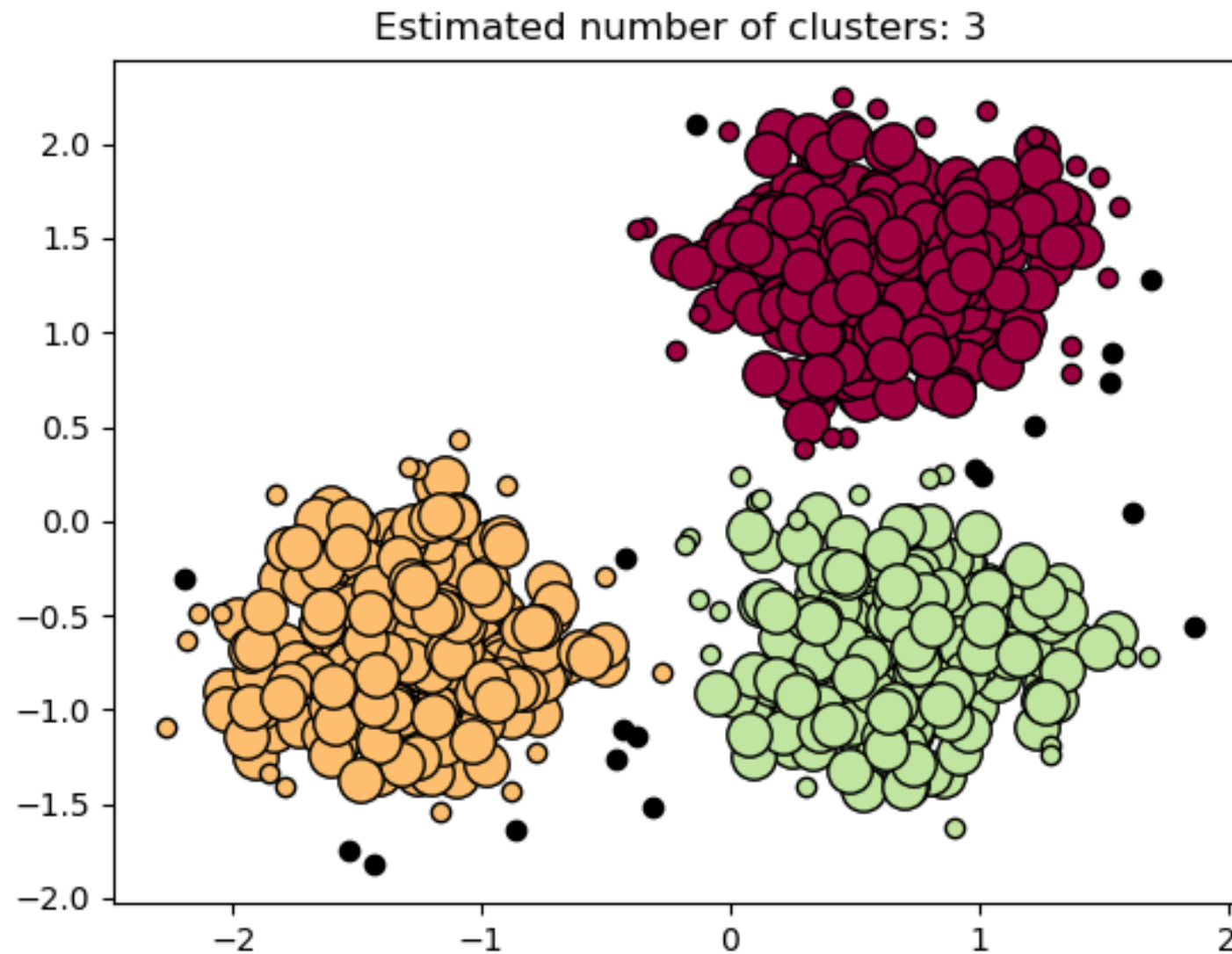


Машинное обучение



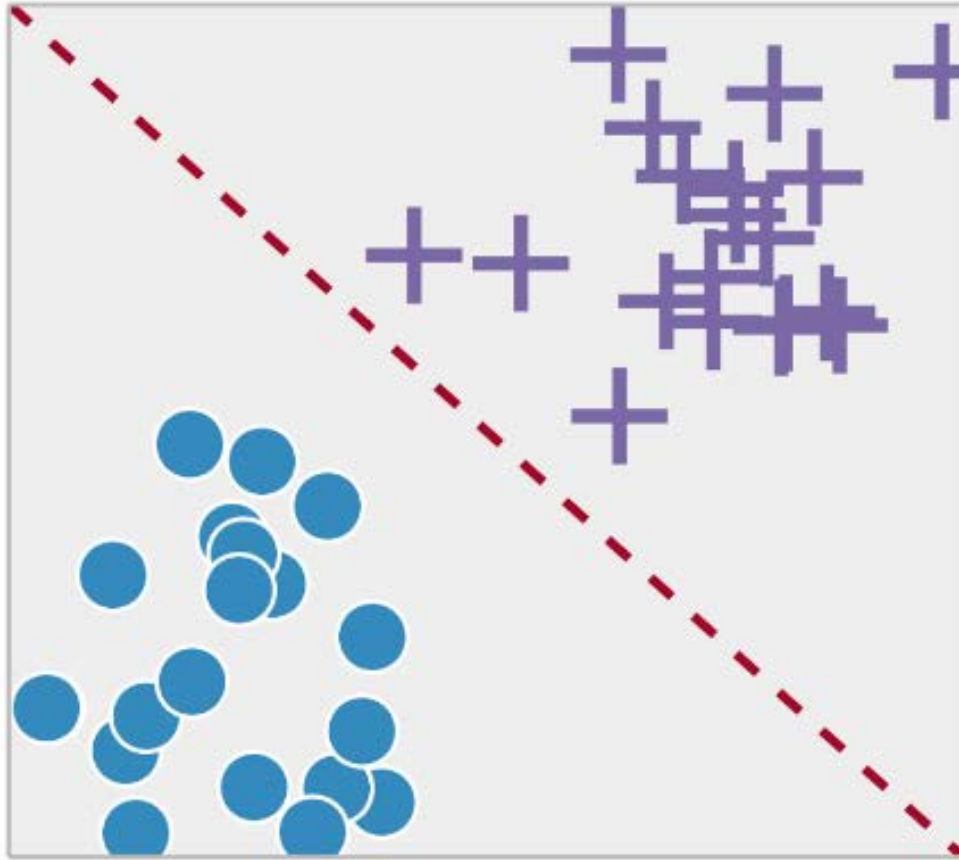


Машинное обучение (Unsupervised)

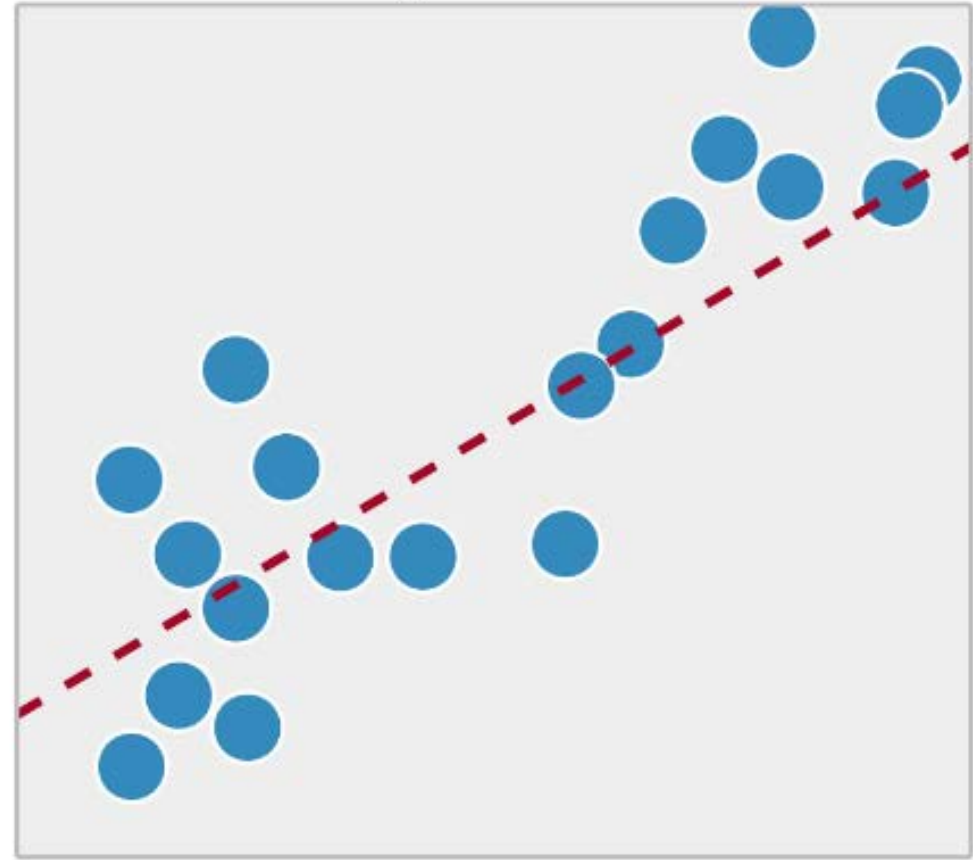


Машинное обучение (Supervised)

Classification



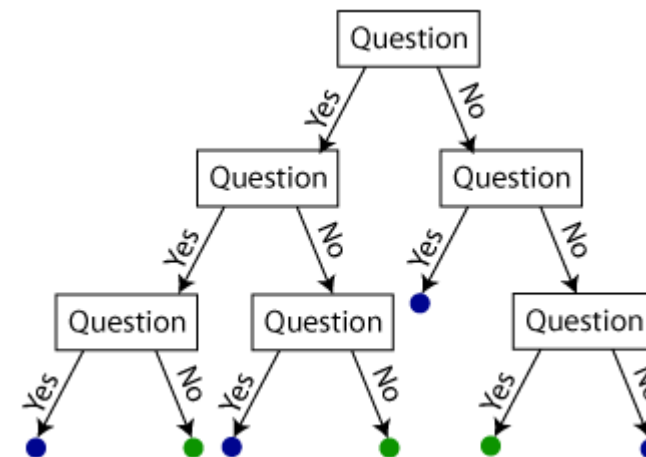
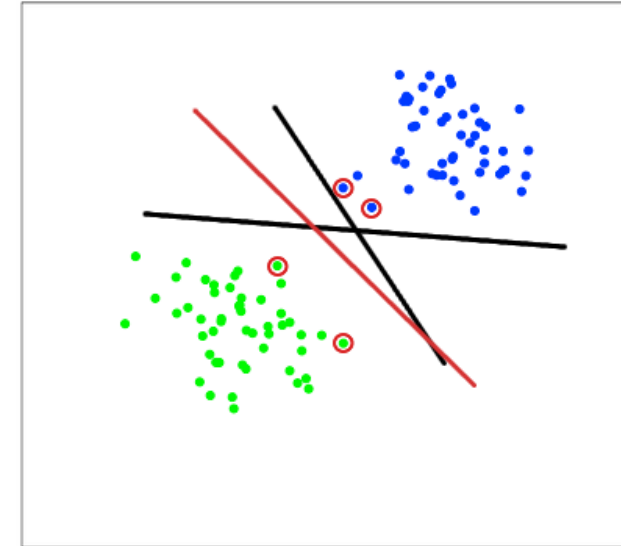
Regression





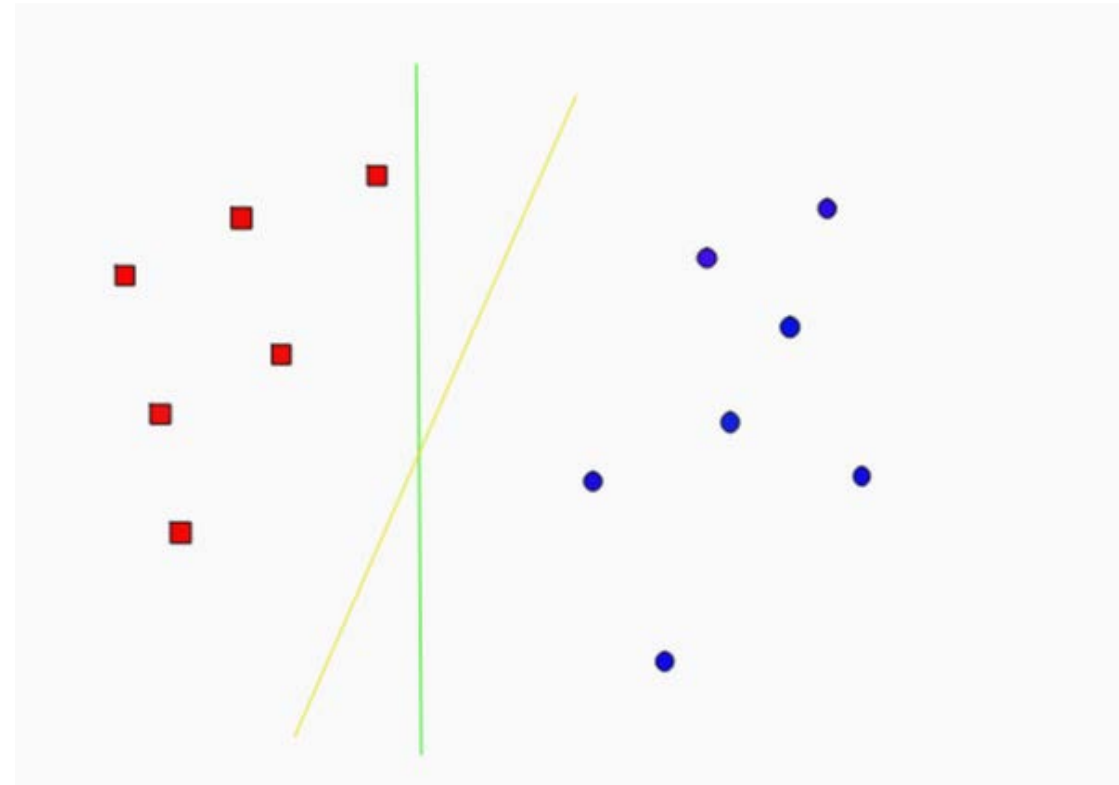
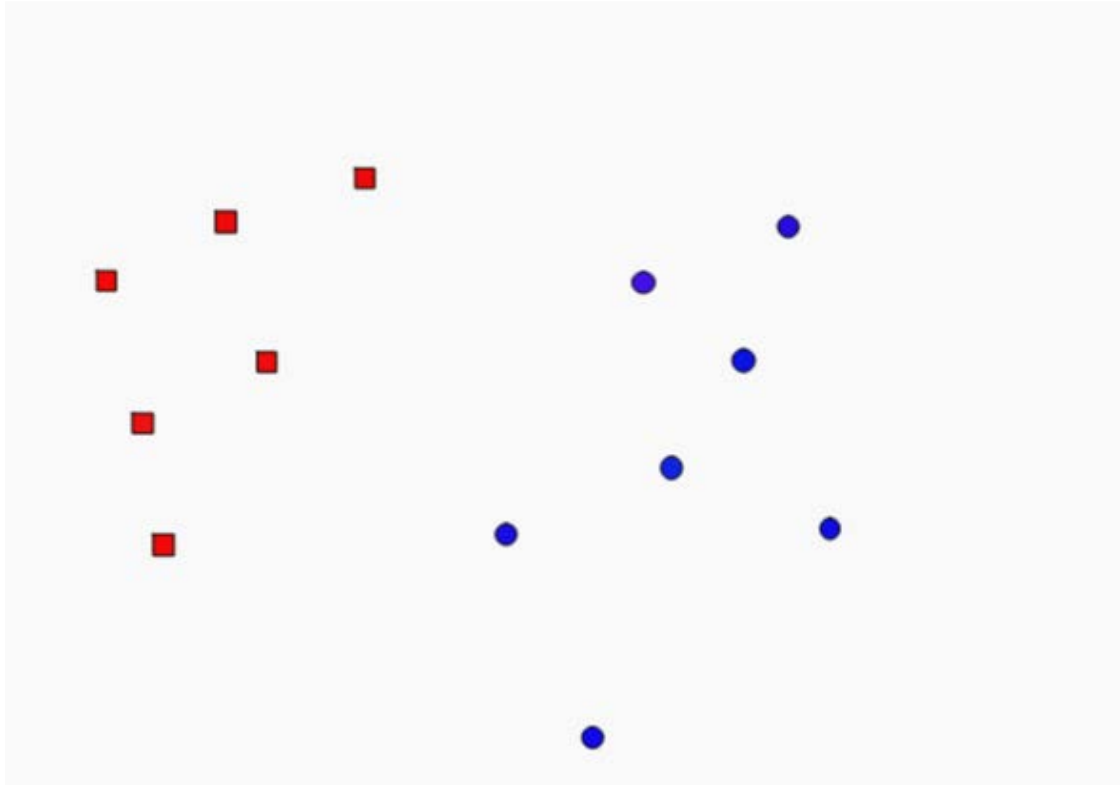
Алгоритмы машинного обучения (примеры)

- Метод опорных векторов (SVM)
- Наивные байесовские классификаторы (Naive Bayes classifier)
- Деревья принятия решений и случайные леса (Decision Tree and Random Forest)





Метод опорных векторов (SVM)





Наивные байесовские классификаторы (Naive Bayes classifier)



$$P(\text{Class } A | \text{Feature } 1, \text{Feature } 2) = \frac{P(\text{Feature } 1 | \text{Class } A) \cdot P(\text{Feature } 2 | \text{Class } A) \cdot P(\text{Class } A)}{P(\text{Feature } 1) \cdot P(\text{Feature } 2)}$$



Наивные байесовские классификаторы (Naive Bayes classifier)

Class	Long	Sweet	Yellow	Total
Banana	400	350	450	500
Orange	0	150	300	300
Other	100	150	50	200
Total	500	650	800	1000

$$P(\text{Class } A | \text{Feature } 1, \text{Feature } 2) = \frac{P(\text{Feature } 1 | \text{Class } A) \cdot P(\text{Feature } 2 | \text{Class } A) \cdot P(\text{Class } A)}{P(\text{Feature } 1) \cdot P(\text{Feature } 2)}$$

Шаг 1

$P(\text{Banana} | \text{Long, Sweet, Yellow})$

Шаг 2: Начнем с числителя и подставим все значения в уравнение:

$P(\text{Long} | \text{Banana}) = 400/500 = 0.8$

$P(\text{Sweet} | \text{Banana}) = 350/500 = 0.7$

$P(\text{Yellow} | \text{Banana}) = 450/500 = 0.9$

$P(\text{Banana}) = 500/1000 = 0.5$

Перемножив значения (согласно уравнению), мы получим:

$0.8 \times 0.7 \times 0.9 \times 0.5 = 0.252$

Шаг 3: Проигнорируем знаменатель, поскольку он будет одинаковым для всех последующих вычислений.

Шаг 4: Проделаем те же вычисления для других классов:

$P(\text{Orange} | \text{Long, Sweet, Yellow}) = 0$

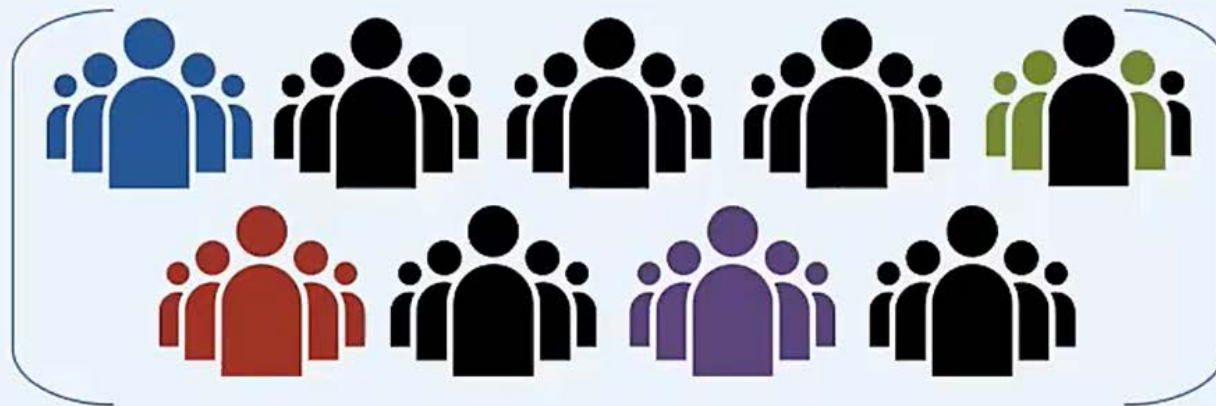
$P(\text{Other} | \text{Long, Sweet, Yellow}) = 0.01875$

Алгоритм Random Forest



Алгоритм Random Forest

- Training data



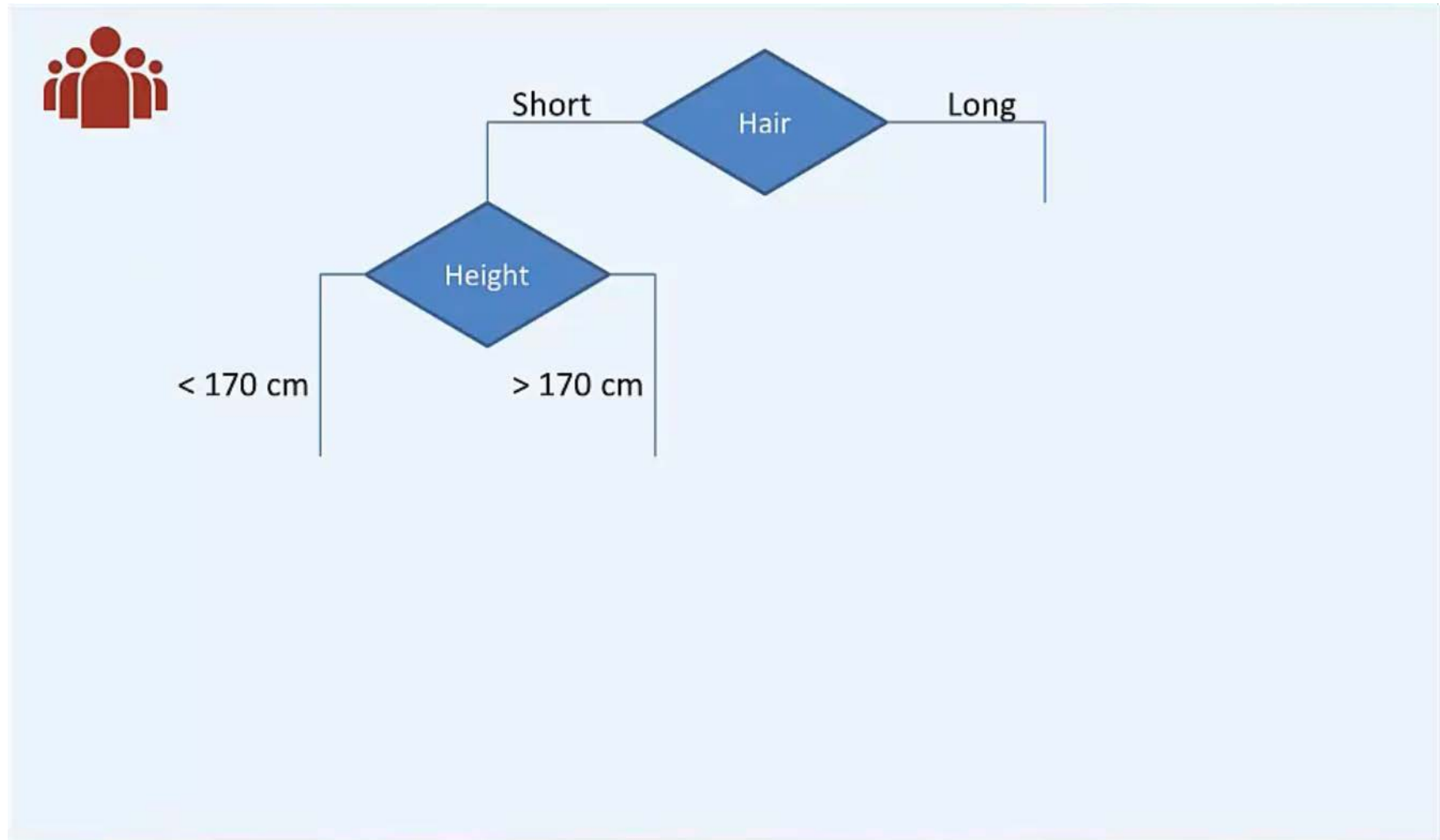
- Subset



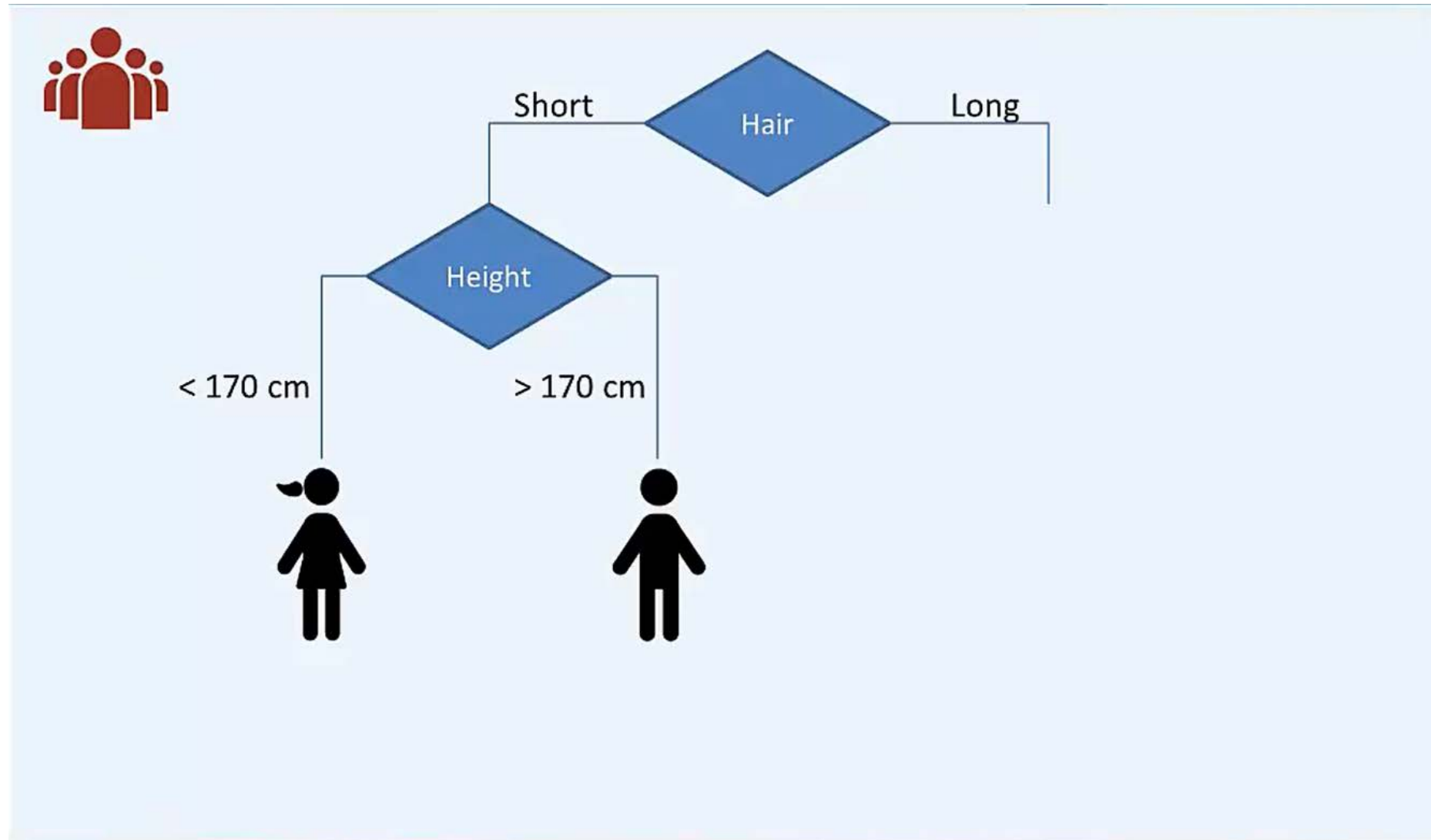
Алгоритм Random Forest



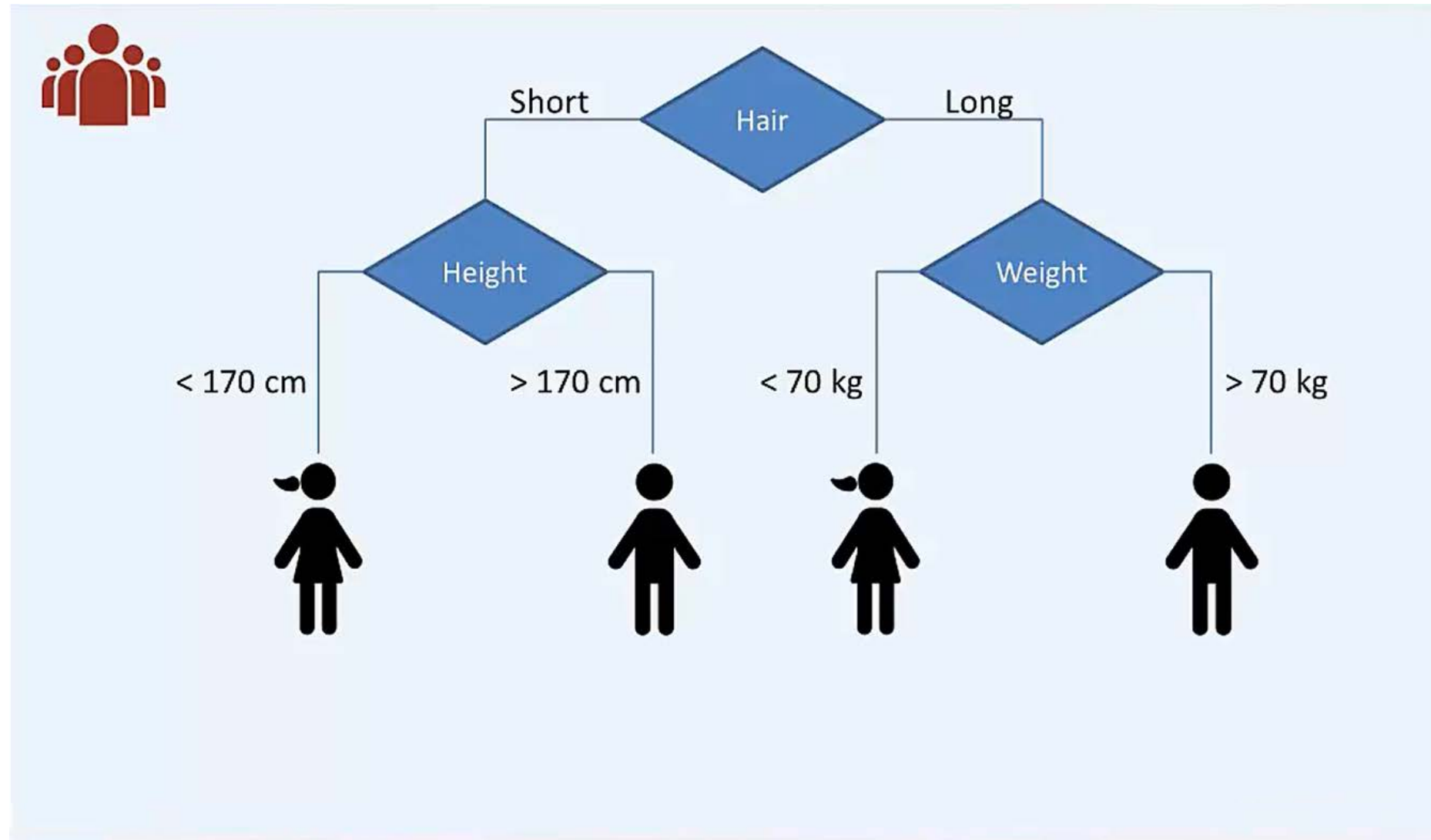
Алгоритм Random Forest



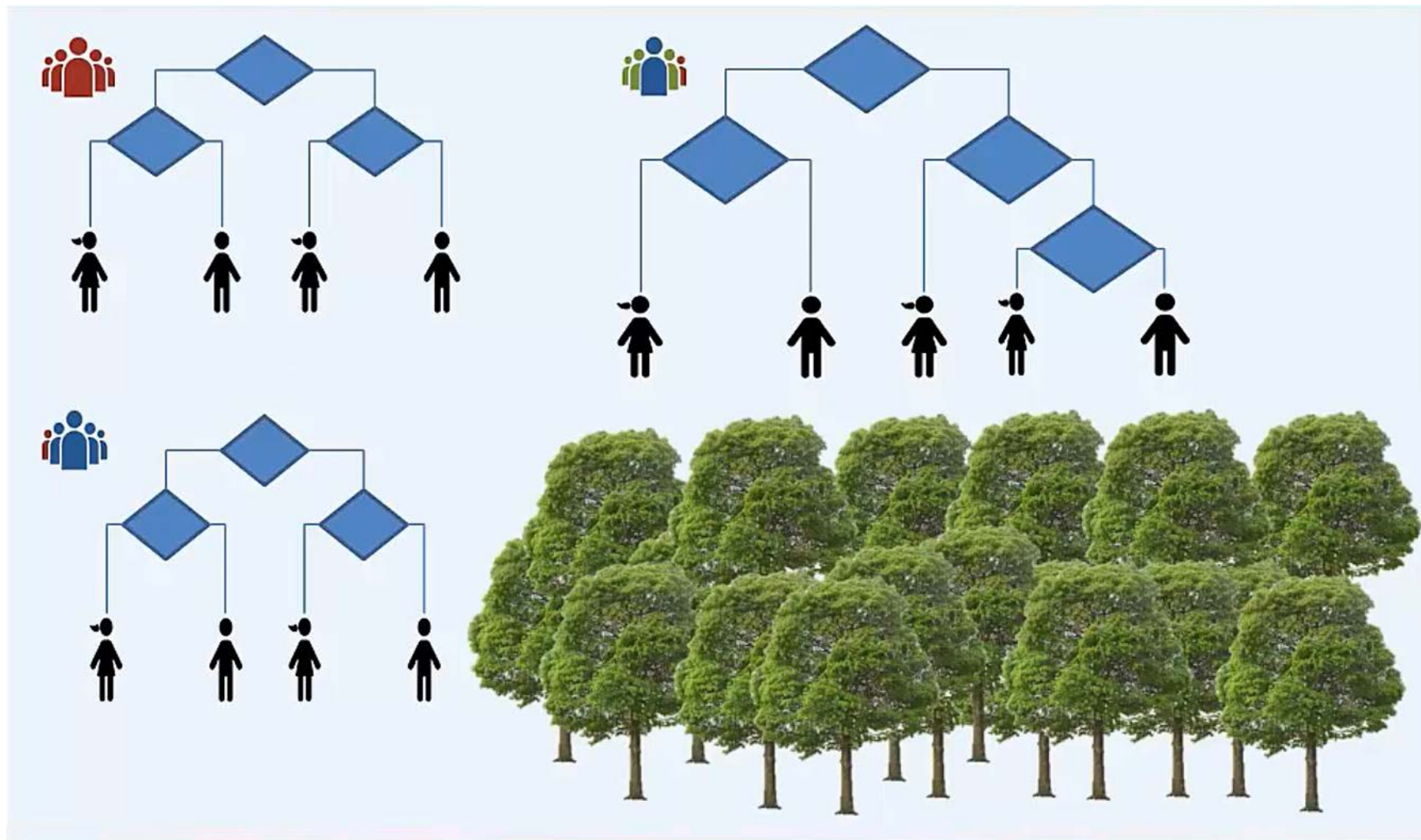
Алгоритм Random Forest



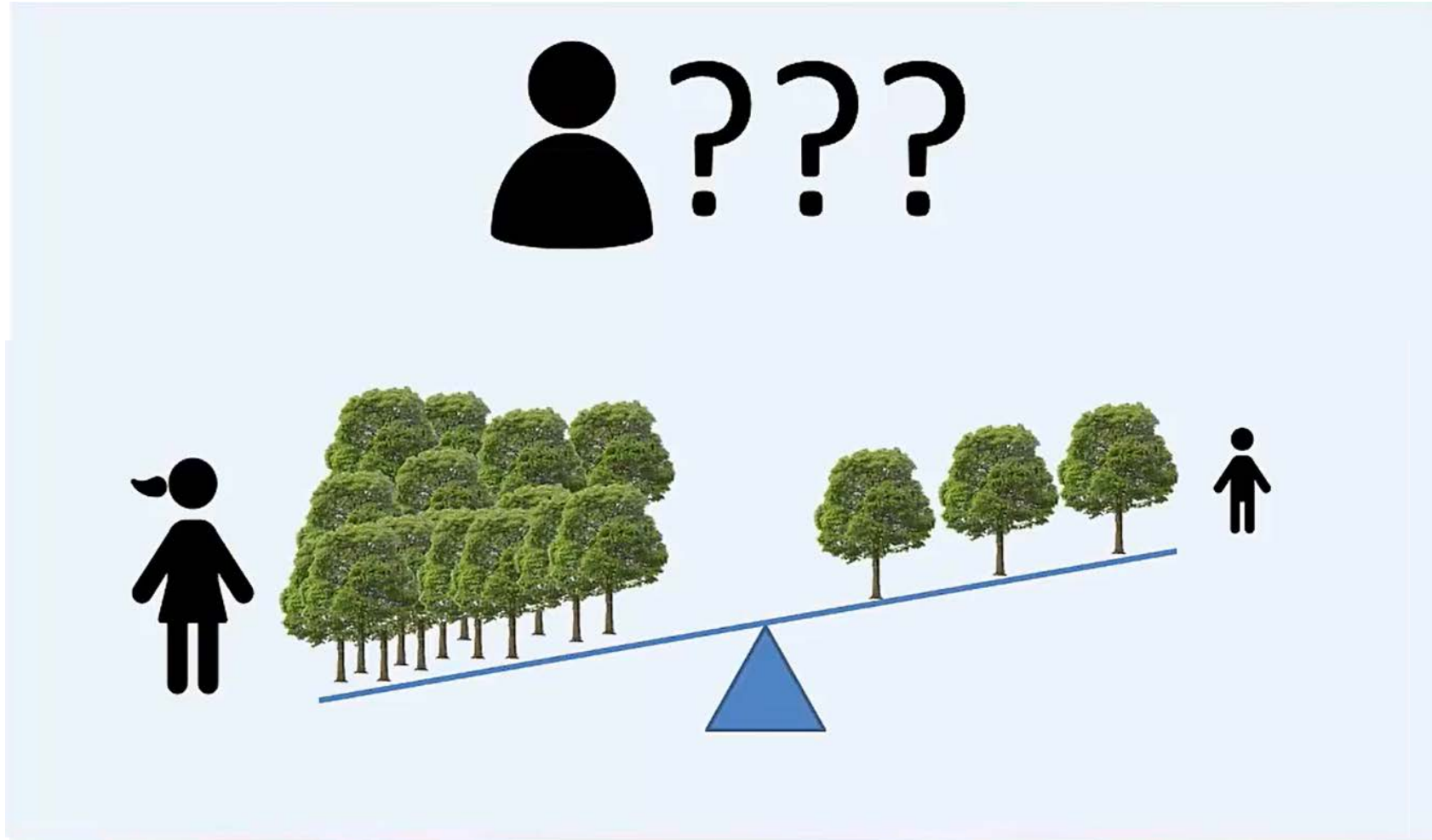
Алгоритм Random Forest



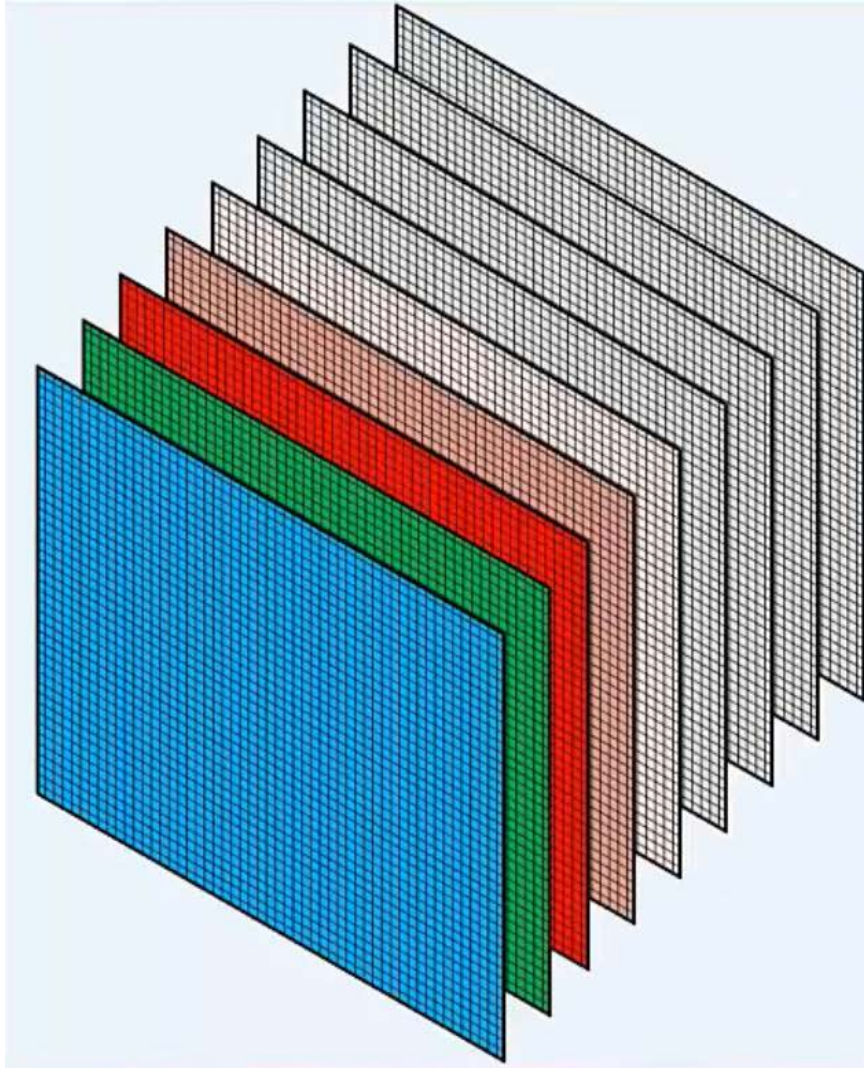
Алгоритм Random Forest



Алгоритм Random Forest

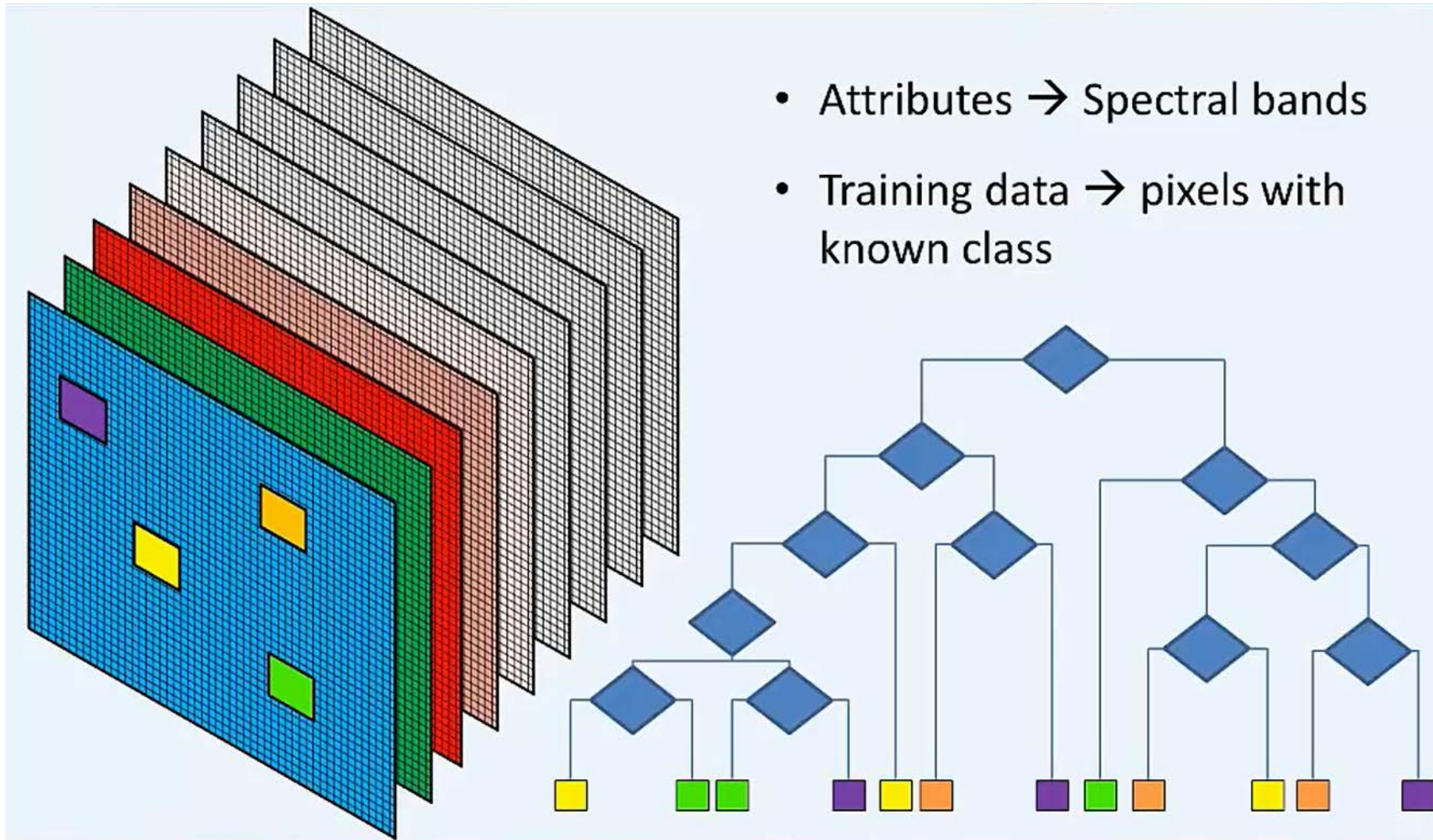


Алгоритм Random Forest

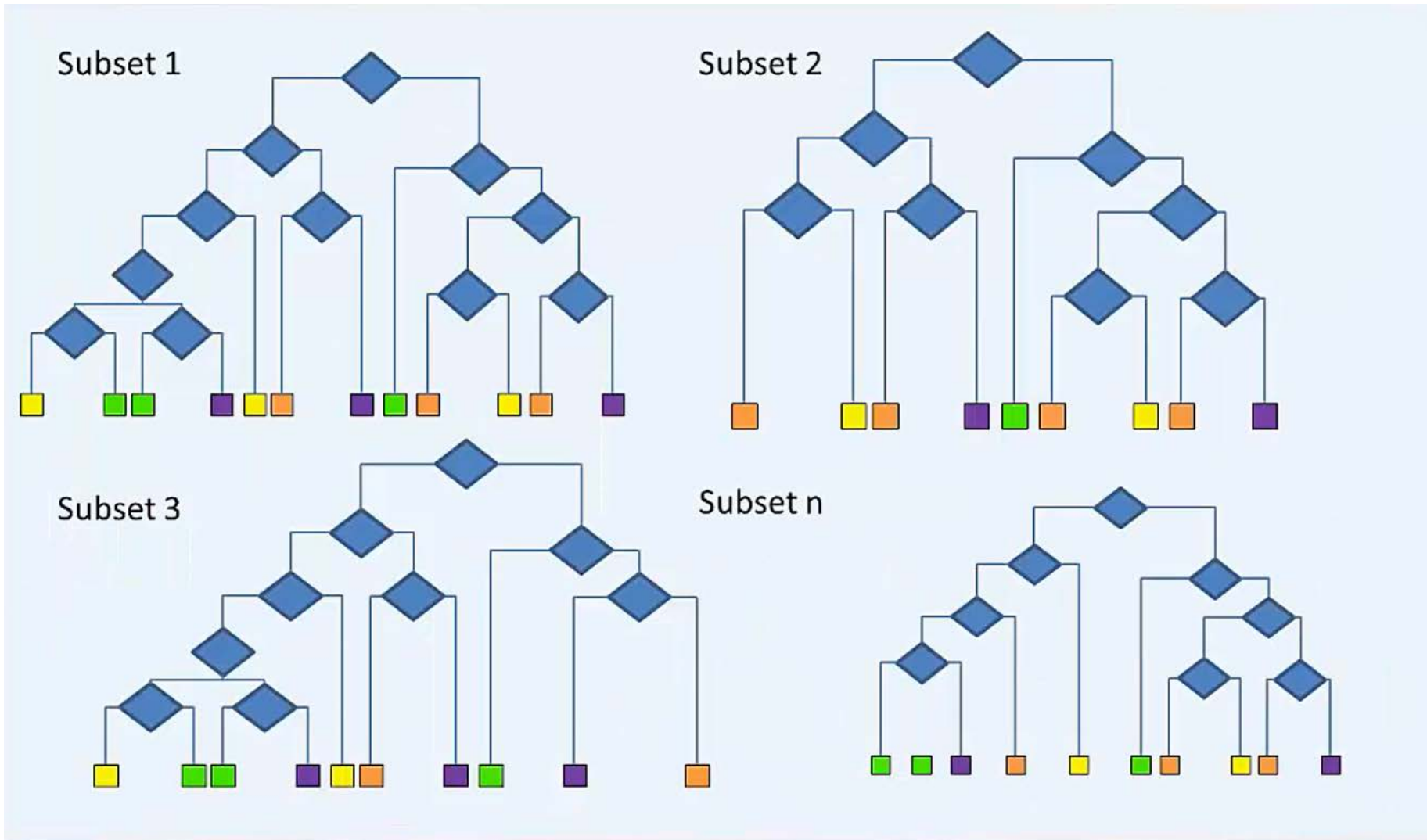


- Attributes → Spectral bands

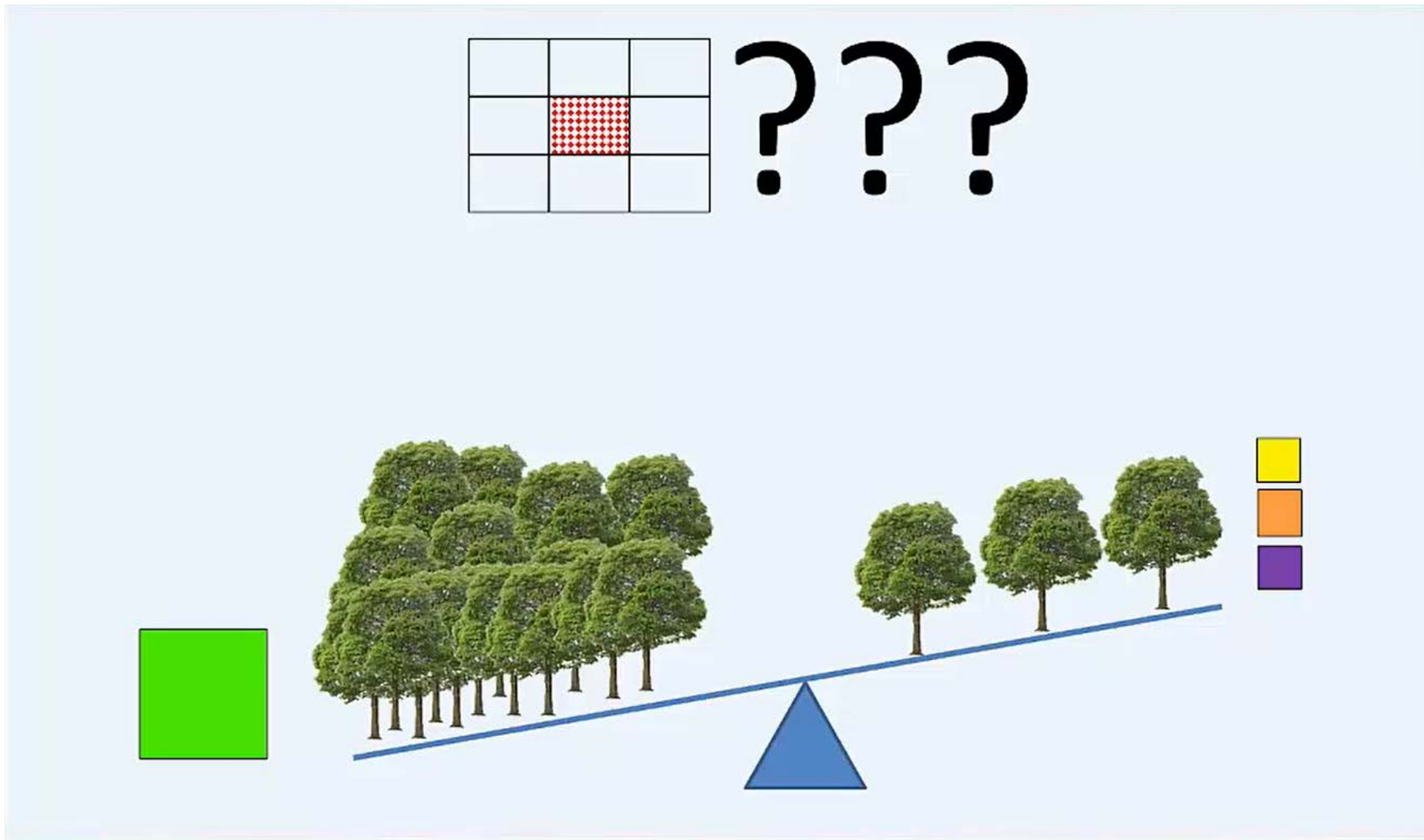
Алгоритм Random Forest



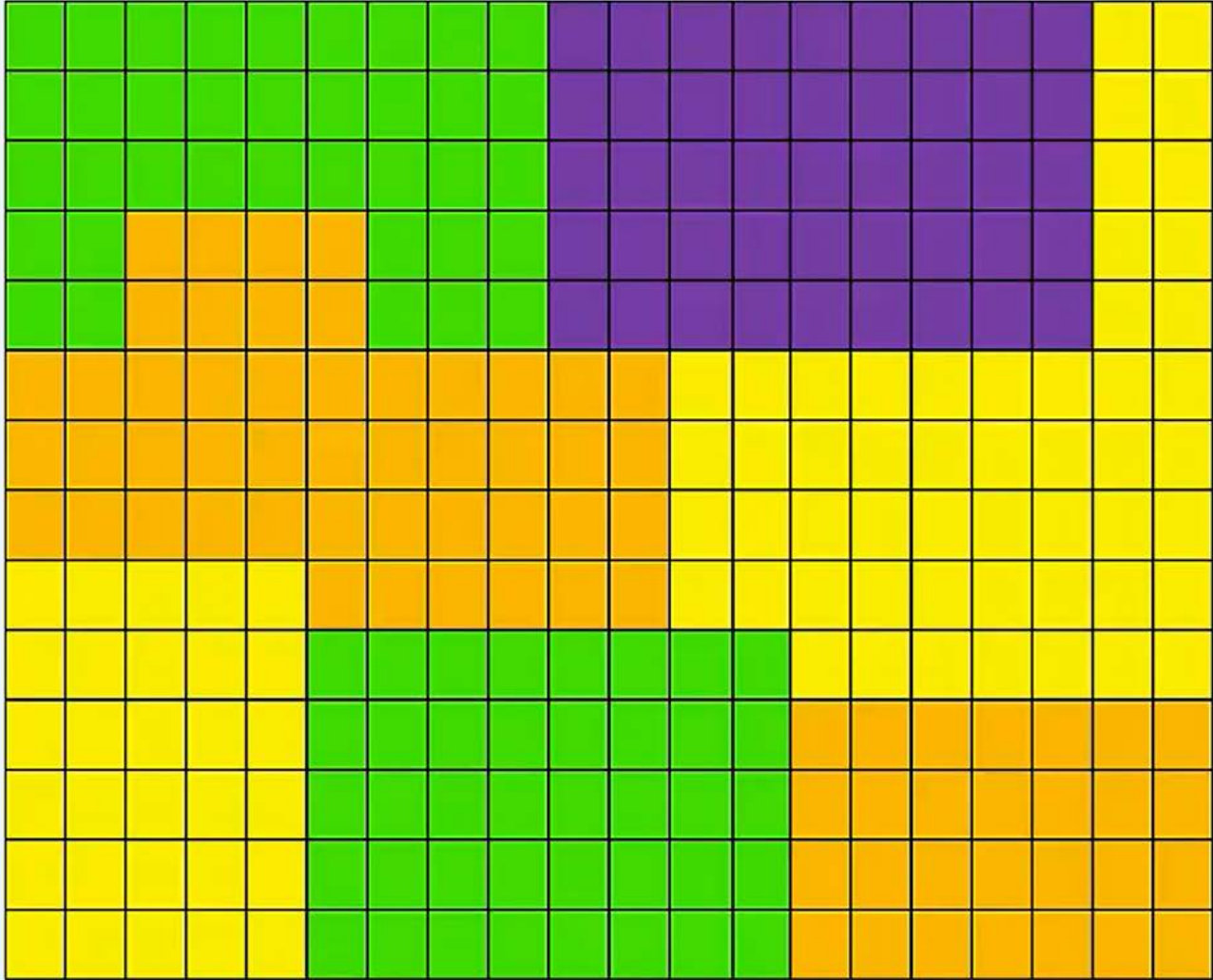
Алгоритм Random Forest



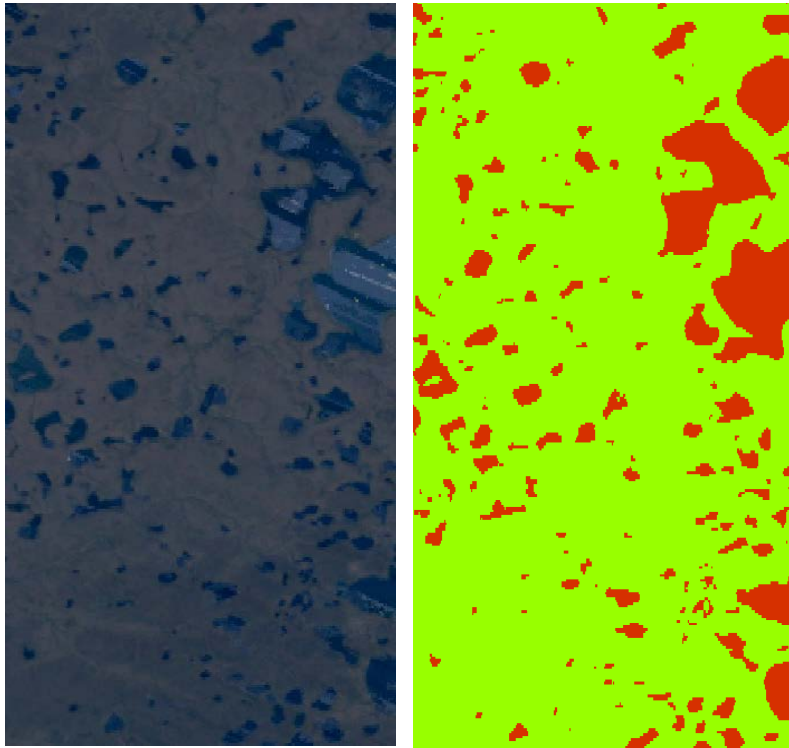
Алгоритм Random Forest



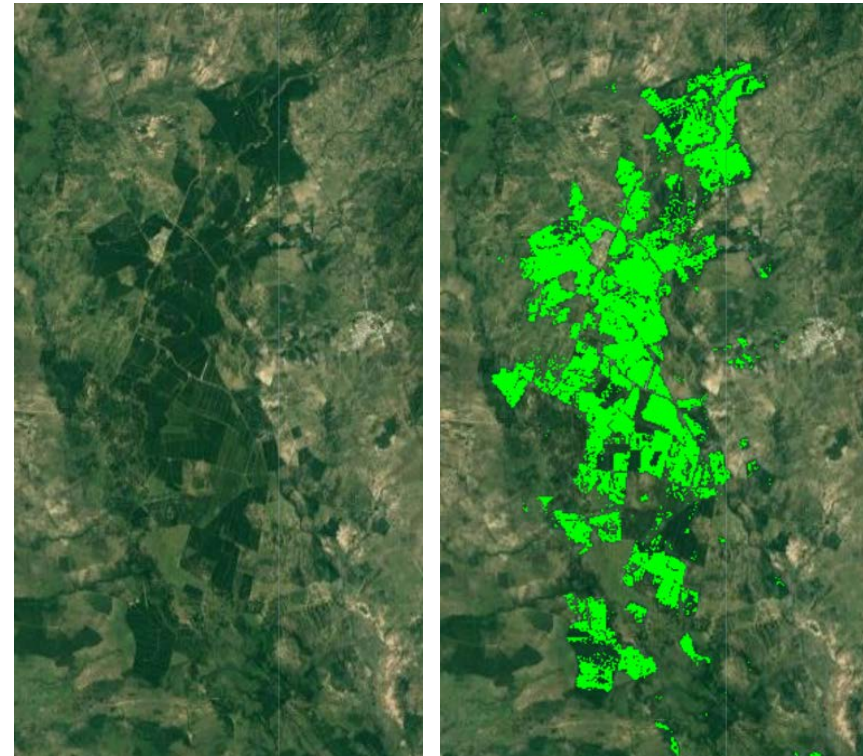
Алгоритм Random Forest



Примеры решения прикладных задач

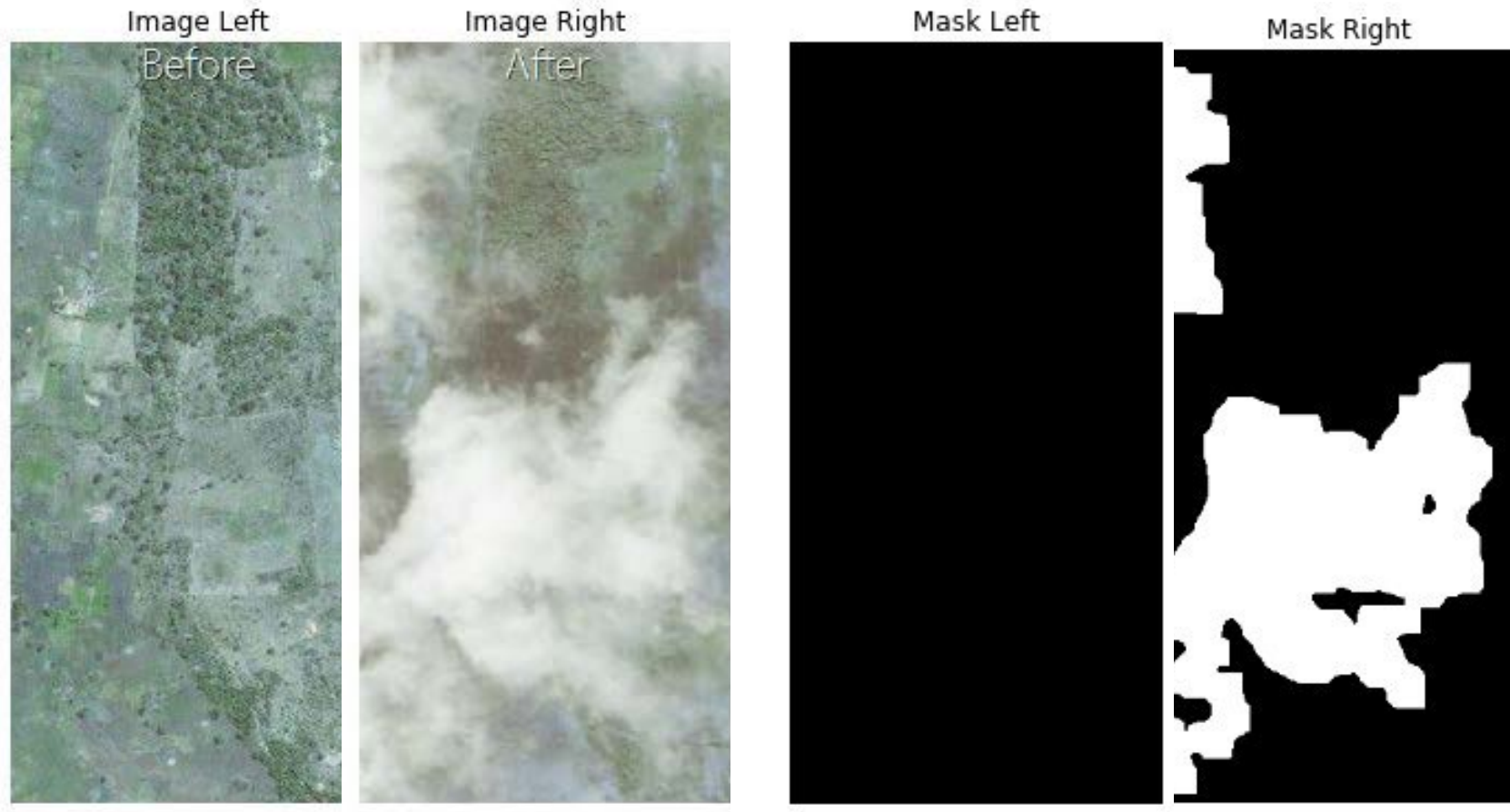


Анализ пораженности местности
термокарстовыми озерами с помощью
алгоритмов Random Forest



Выделение техногенно-
преобразованных зон с помощью
алгоритмов Random Forest

Примеры решения прикладных задач



Определение участков с повышенной облачностью с помощью алгоритмов компьютерного зрения



Московский Государственный Университет
им. М.В. Ломоносова

СПАСИБО!