



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

# **ВВЕДЕНИЕ В АНАЛИЗ ГЕОДАННЫХ**

## **Машинное обучение для анализа геоданных**

**Николай Никифоров  
Елизавета Хазиева  
Николай Билев**

***Москва 2018***

# ВВЕДЕНИЕ: Что такое искусственный интеллект (AI)?

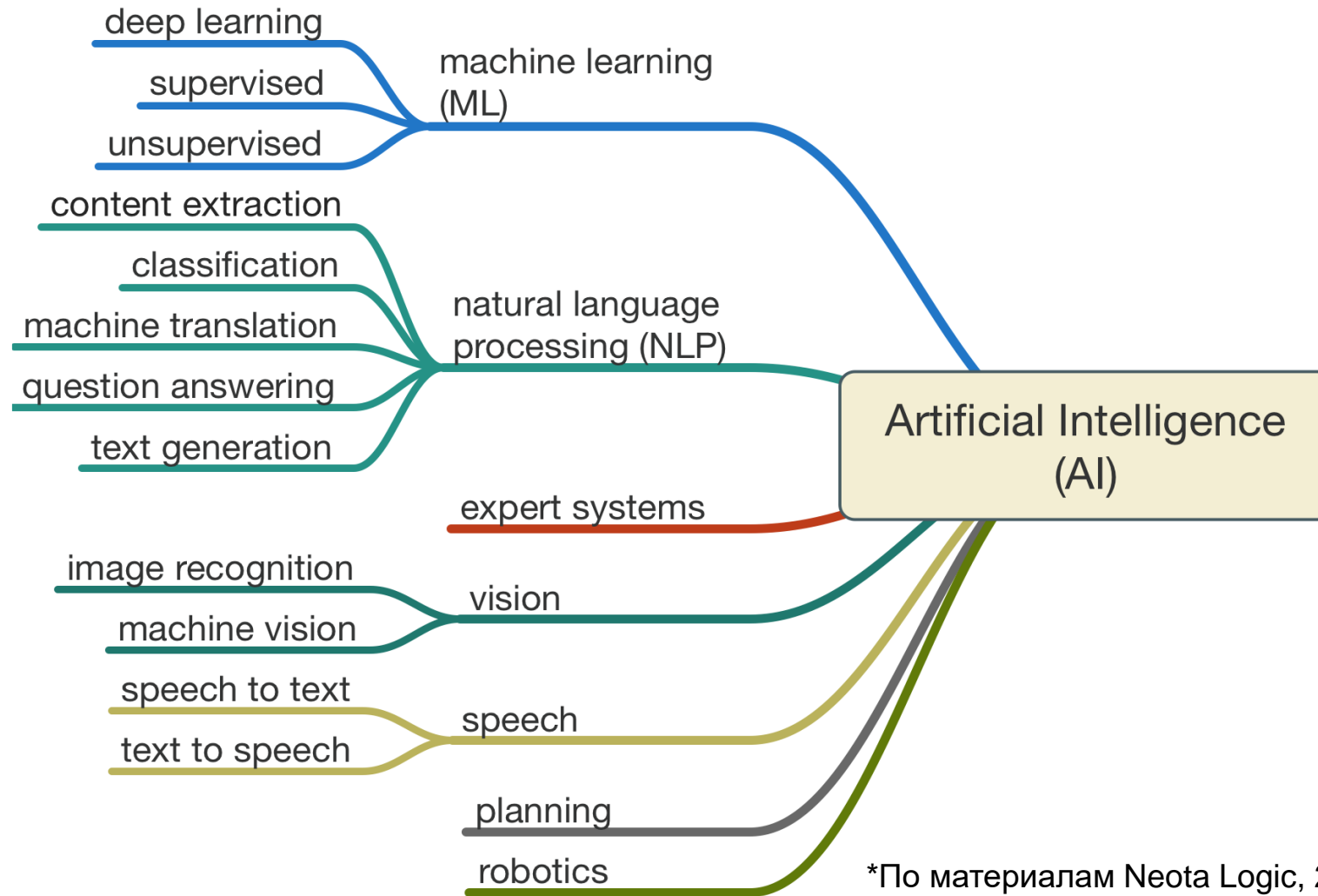
## ARTIFICIAL INTELLIGENCE

1: a branch of computer science dealing with the simulation of intelligent behavior in computers

2: the capability of a machine to imitate intelligent human behavior <sup>1</sup>

1. "Artificial Intelligence." *Merriam-Webster.com*. Merriam-Webster, n.d. Web. 13 Mar. 2018.

# ВВЕДЕНИЕ: Разделы искусственного интеллекта



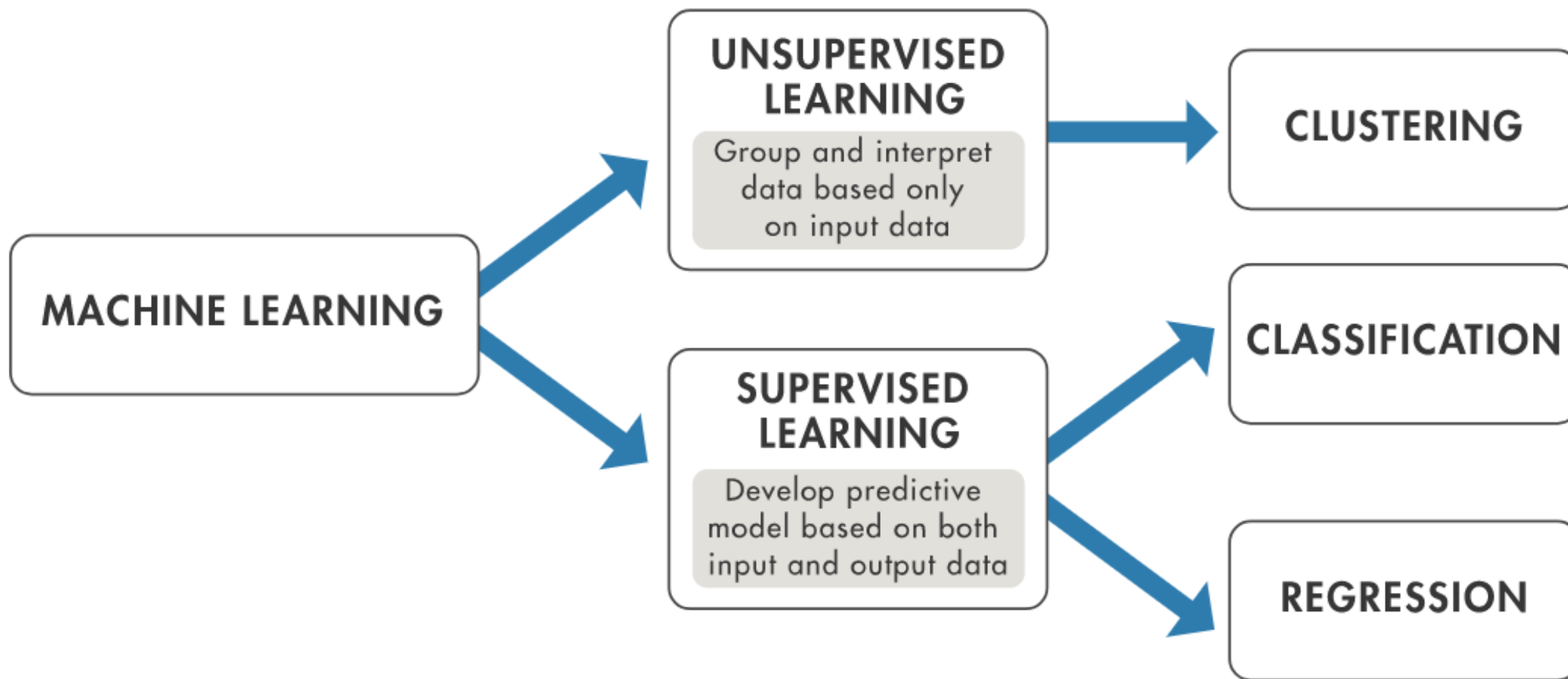
\*По материалам Neota Logic, 2015

# ВВЕДЕНИЕ: Что такое машинное обучение?

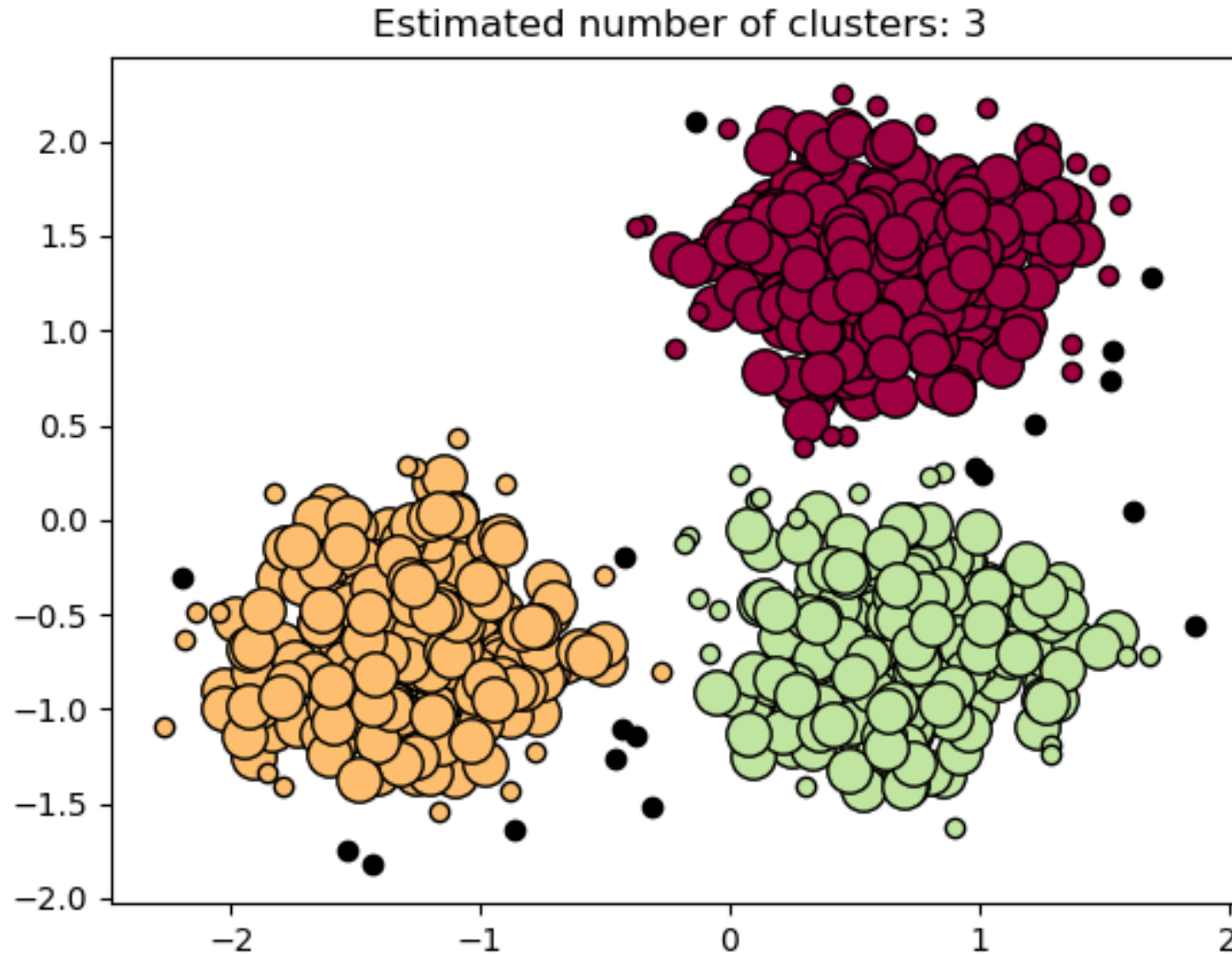
**Machine Learning:** Field of study that gives computers the ability to learn without being explicitly programmed (Arthur Samuel, 1959)

**Машинное обучение (англ. machine learning, ML)** — класс методов искусственного интеллекта, характерной чертой которых является не прямое решение задачи, а обучение в процессе применения решений множества сходных задач. Для построения таких методов используются средства математической статистики, численных методов, методов оптимизации, теории вероятностей, теории графов, различные техники работы с данными в цифровой форме (Wikipedia)

# ВВЕДЕНИЕ: Что такое машинное обучение?



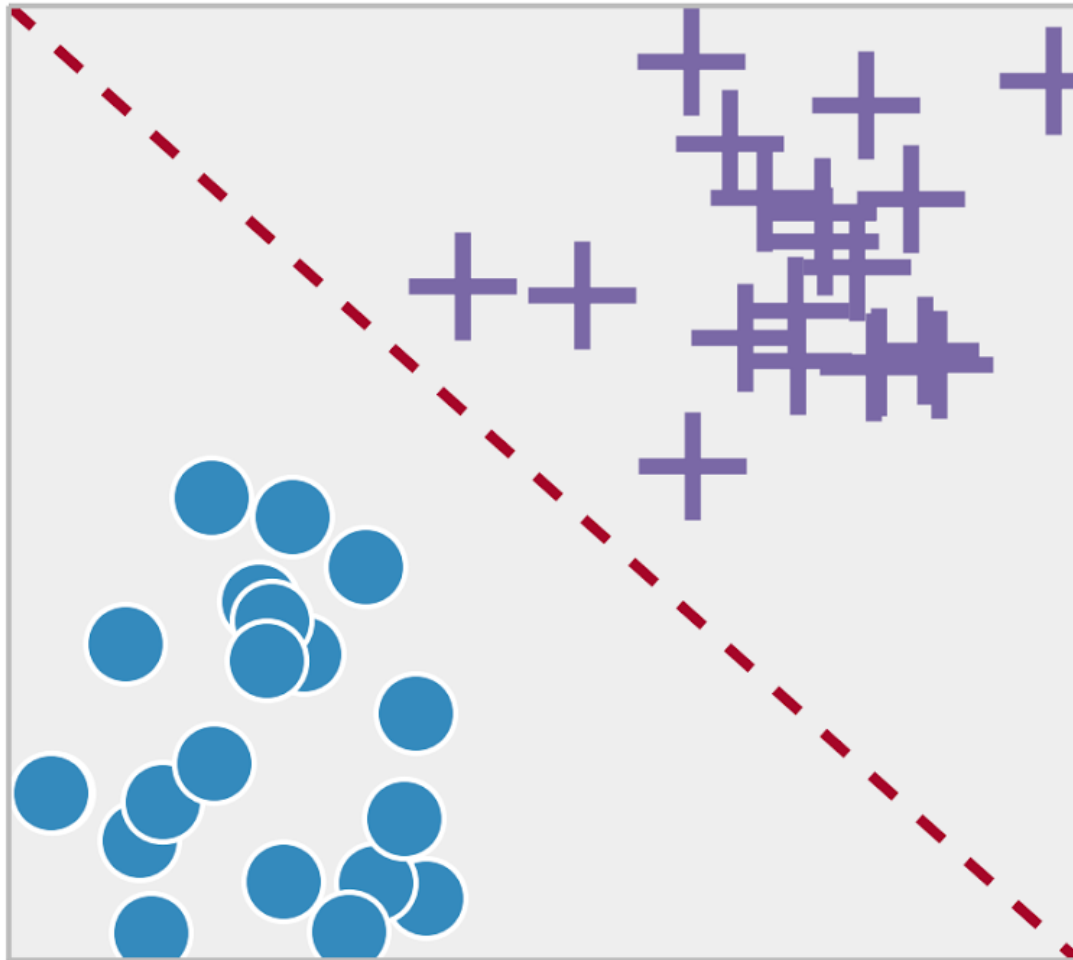
# МАШИННОЕ ОБУЧЕНИЕ БЕЗ УЧИТЕЛЯ (UNSUPERVISED LEARNING)



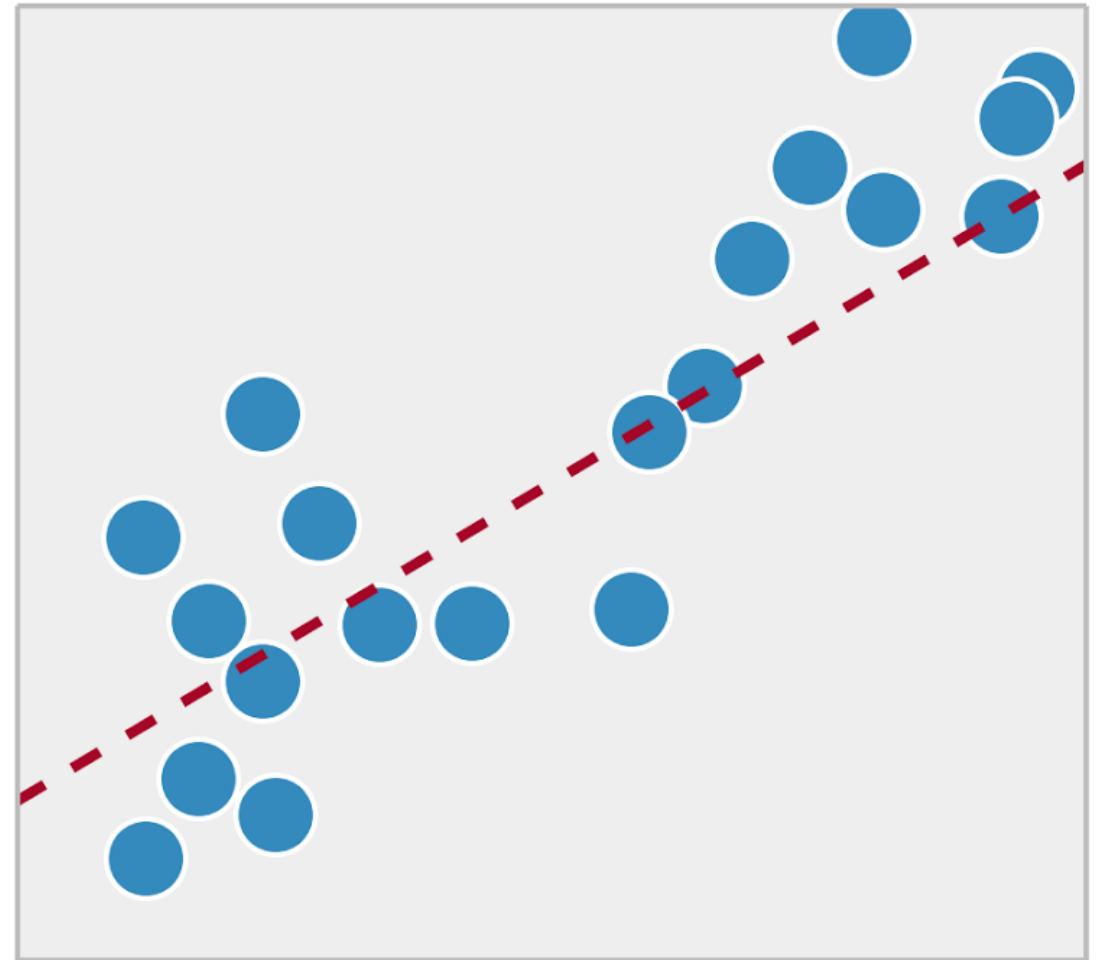


# МАШИННОЕ ОБУЧЕНИЕ С УЧИТЕЛЕМ (SUPERVISED LEARNING)

## Classification



## Regression



# ПОСТАНОВКА ЗАДАЧИ МАШИННОГО ОБУЧЕНИЯ

$X$  – Набор объектов

$Y$  – Набор ответов

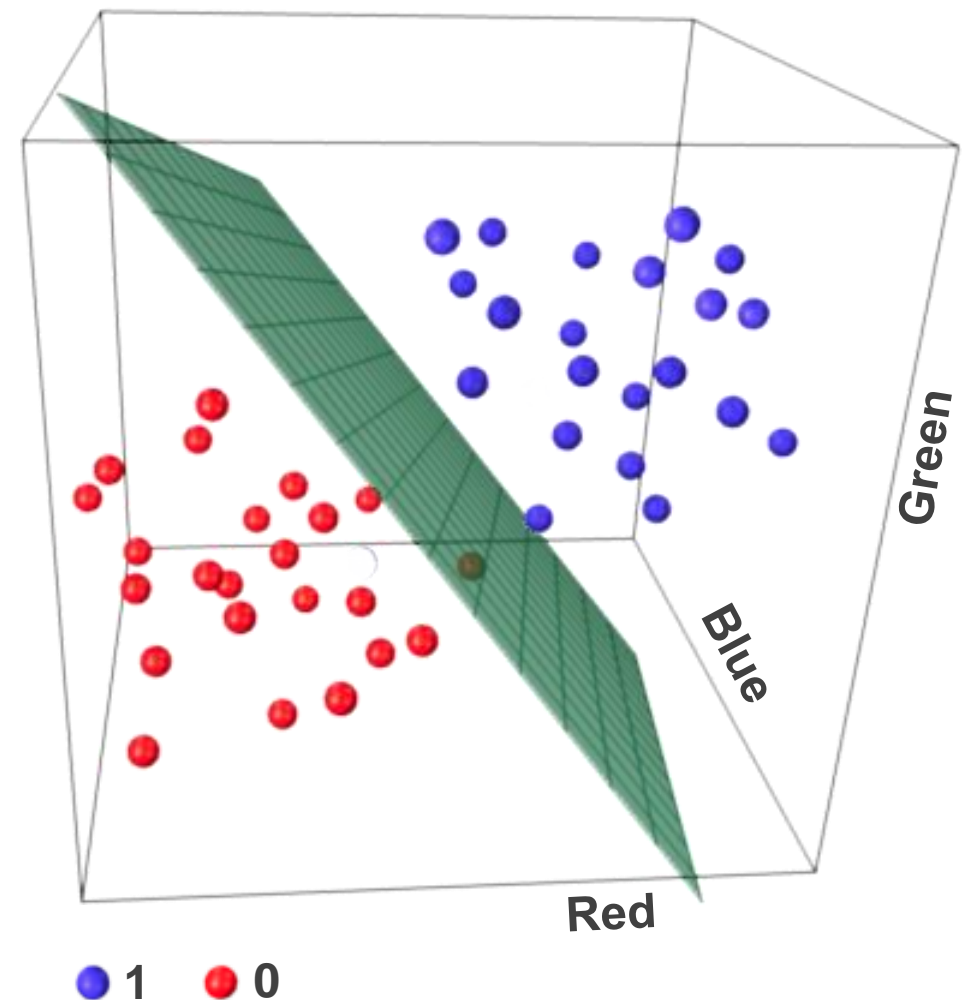
$y: X \rightarrow Y$  – Искомая функция

$\{x_1, \dots, x_l\} \subset X$  – тренировочная выборка

$y_i = y(x_i), i = 1, \dots, l$  – известные ответы

$a: X \rightarrow Y$  – алгоритм – функция решения,  
апрксимирующая  $y$  через набор значений  $X$

Разделяющая гиперплоскость

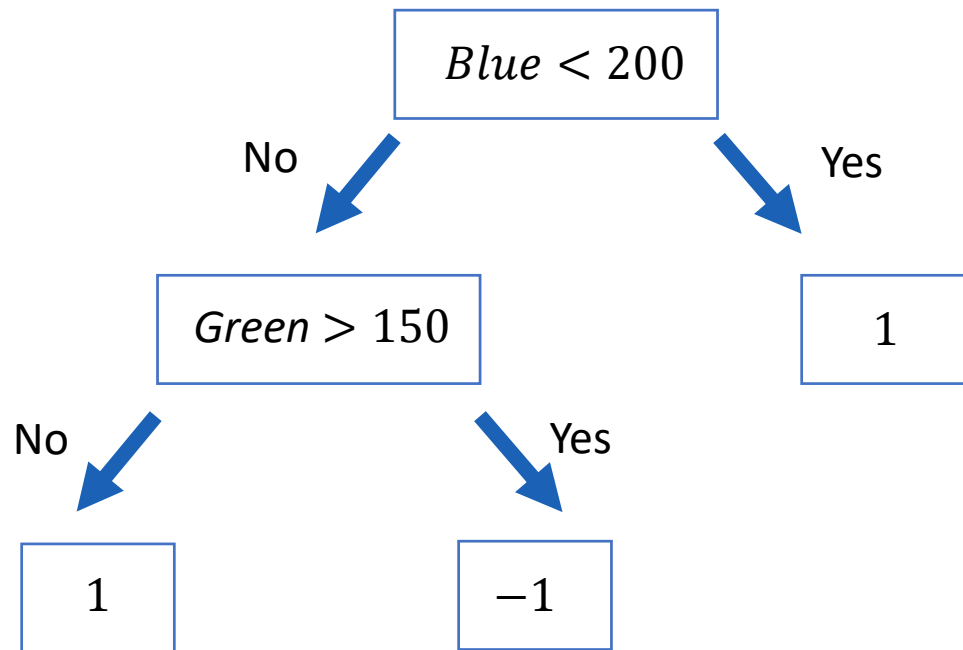




# МАШИННОЕ ОБУЧЕНИЕ: RANDOM FOREST

Decision Tree  
Using random sample of features  
and random sample of objects  $x \in X$

$b_1(x)$ :



$$I_G(p) = \sum_{i=1}^J p_i(1 - p_i) \quad \text{Gini Impurity}$$

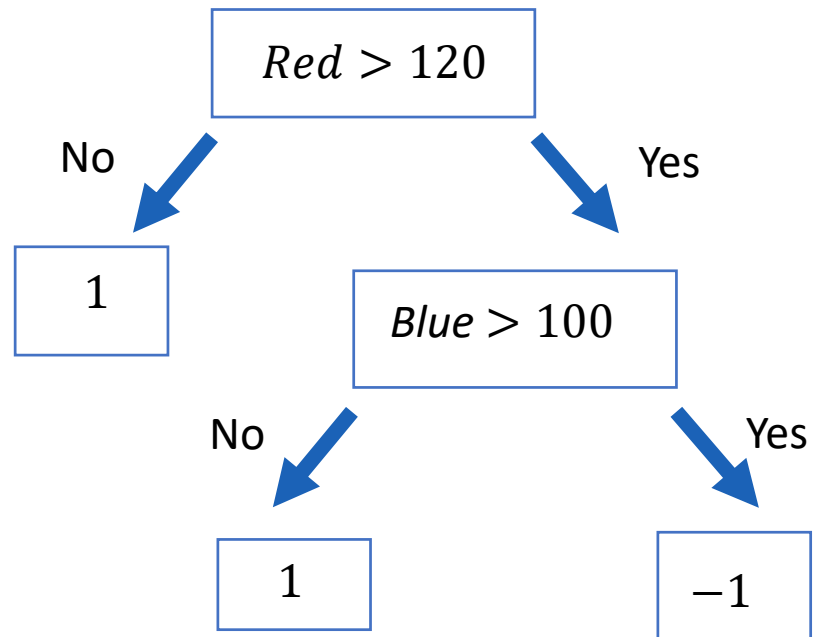
$$a(t) = \text{sign} \frac{1}{T} \sum_{t=1}^1 b_t(x) \quad \text{Ensemble of decision trees}$$

| Pixel Id | Features |       |      | True | $b_1$ |
|----------|----------|-------|------|------|-------|
|          | Red      | Green | Blue |      |       |
| 1        | 121      | 181   | 201  | -1   | 1     |
| 2        | 118      | 209   | 66   | 1    | 1     |
| 3        | 23       | 57    | 119  | -1   | -1    |
| 4        | 150      | 120   | 167  | 1    | -1    |

# МАШИННОЕ ОБУЧЕНИЕ: RANDOM FOREST

Decision Tree  
Using random sample of features  
and random sample of objects  $x \in X$

$b_2(x)$ :



$$I_G(p) = \sum_{i=1}^J p_i(1 - p_i) \quad \text{Gini Impurity}$$

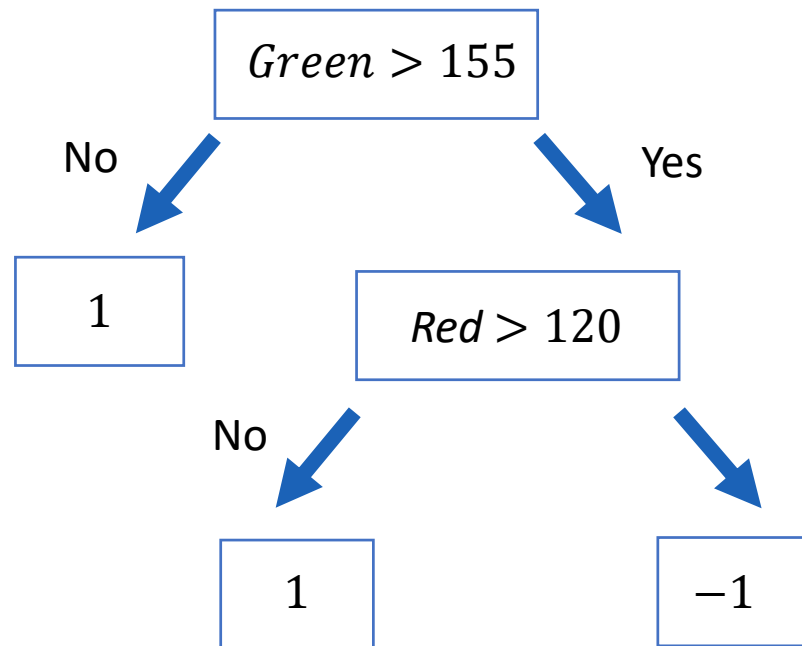
$$a(t) = \text{sign} \frac{1}{T} \sum_{t=1}^1 b_t(x) \quad \text{Ensemble of decision trees}$$

| Pixel Id | Features |       |      | True | $b_1$ | $b_2$ |
|----------|----------|-------|------|------|-------|-------|
|          | Red      | Green | Blue |      |       |       |
| 1        | 121      | 181   | 201  | -1   | 1     | -1    |
| 2        | 118      | 209   | 66   | 1    | 1     | 1     |
| 3        | 23       | 57    | 119  | -1   | -1    | 1     |
| 4        | 150      | 120   | 167  | 1    | -1    | 1     |

# МАШИННОЕ ОБУЧЕНИЕ: RANDOM FOREST

Decision Tree  
Using random sample of features  
and random sample of objects  $x \in X$

$b_3(x)$ :



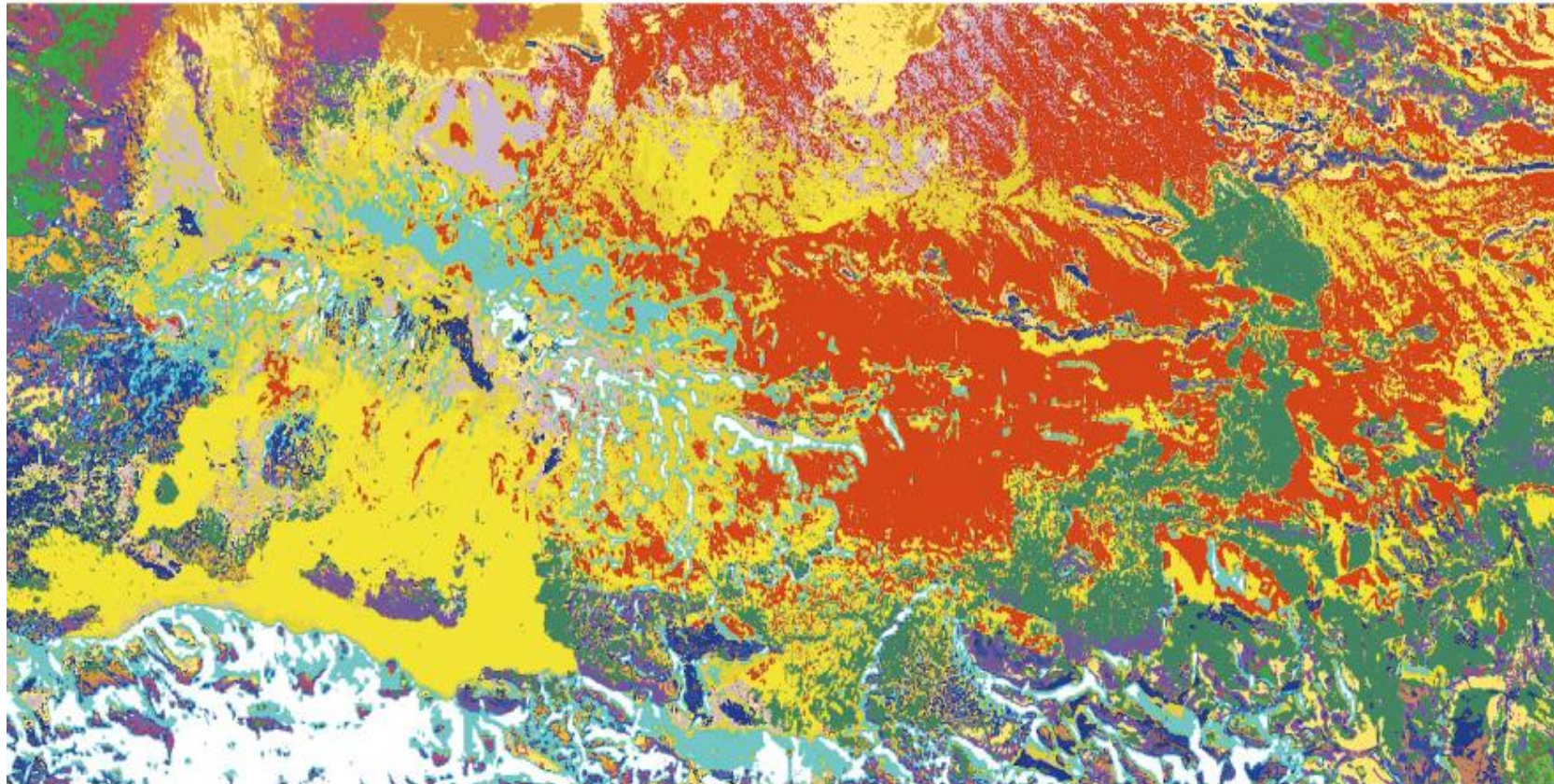
$$I_G(p) = \sum_{i=1}^J p_i(1 - p_i) \quad \text{Gini Impurity}$$

$$a(t) = \text{sign} \frac{1}{T} \sum_{t=1}^1 b_t(x) \quad \text{Ensemble of decision trees}$$

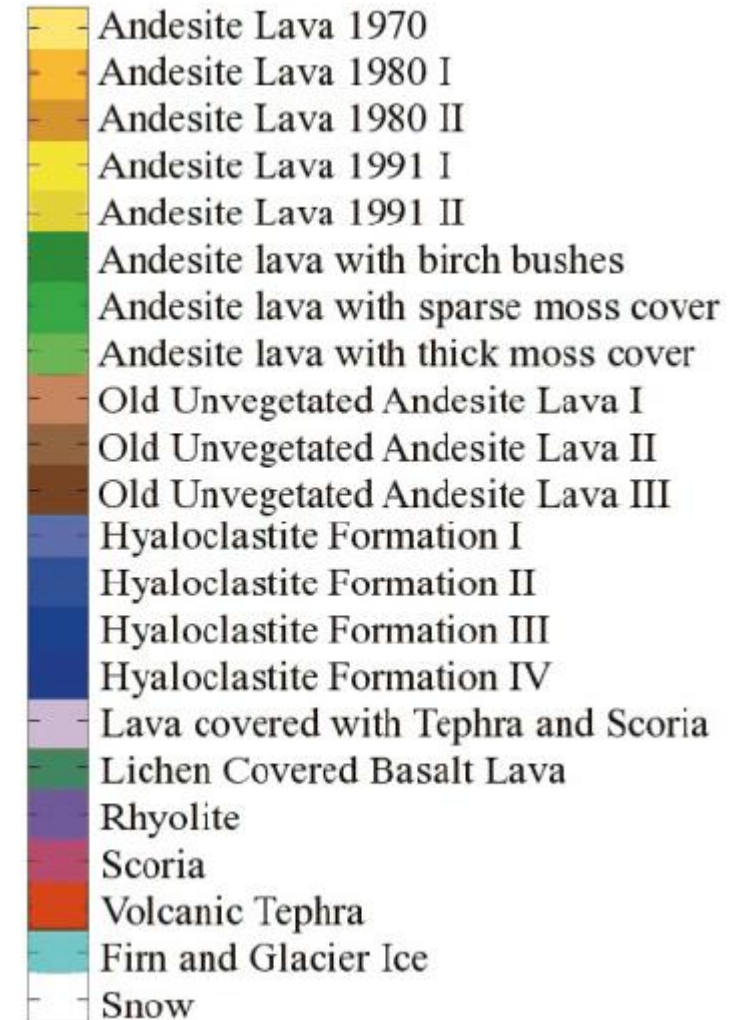
| Pixel Id | Features |       |      | True | $b_1$ | $b_2$ | $b_3$ | Avg  | Sign |
|----------|----------|-------|------|------|-------|-------|-------|------|------|
|          | Red      | Green | Blue |      |       |       |       |      |      |
| 1        | 121      | 181   | 201  | -1   | 1     | -1    | -1    | -0.3 | -1   |
| 2        | 118      | 209   | 66   | 1    | 1     | 1     | 1     | 1    | 1    |
| 3        | 23       | 57    | 119  | -1   | -1    | 1     | 1     | 0.3  | 1    |
| 4        | 150      | 120   | 167  | 1    | -1    | 1     | 1     | 0.3  | 1    |

# ПРИМЕРЫ МАШИННОГО ОБУЧЕНИЯ С УЧИТЕЛЕМ (ЗАДАЧА КЛАССИФИКАЦИИ)

# ПРИМЕРЫ РЕШЕНИЯ ПРИКЛАДНЫХ ЗАДАЧ



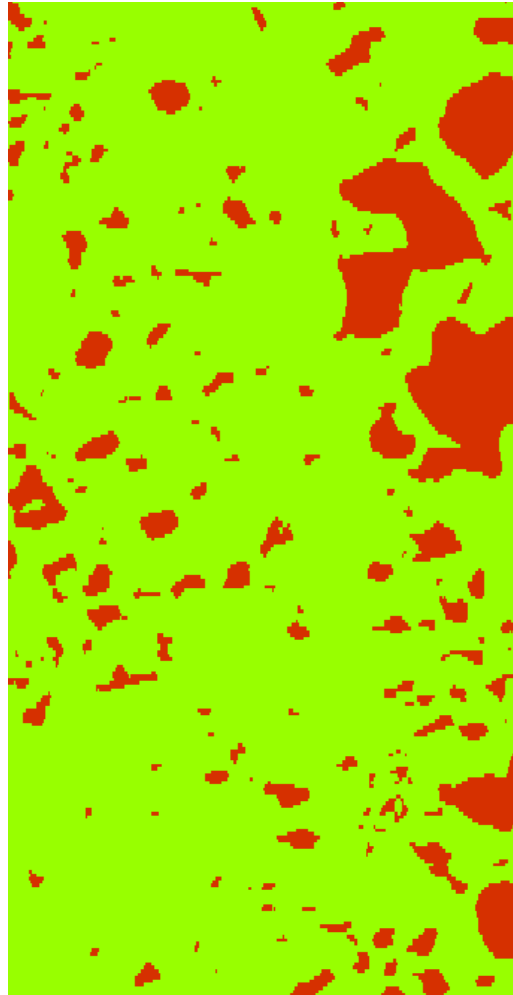
Результаты классификации геологических формаций, полученных с помощью машинного обучения. Район влк. Гекла, Исландия.  
Бьёрн Вак, 2015



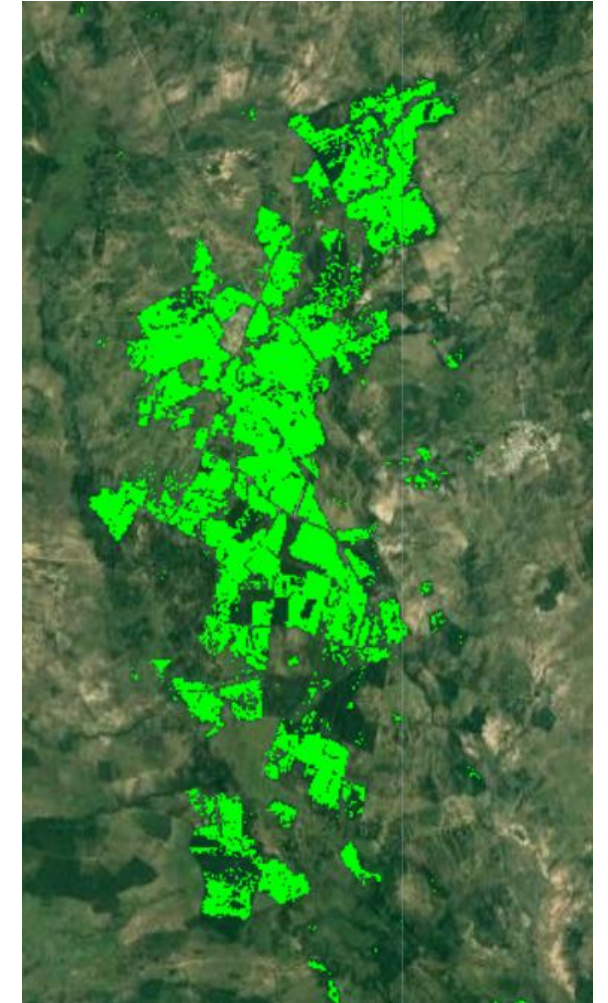
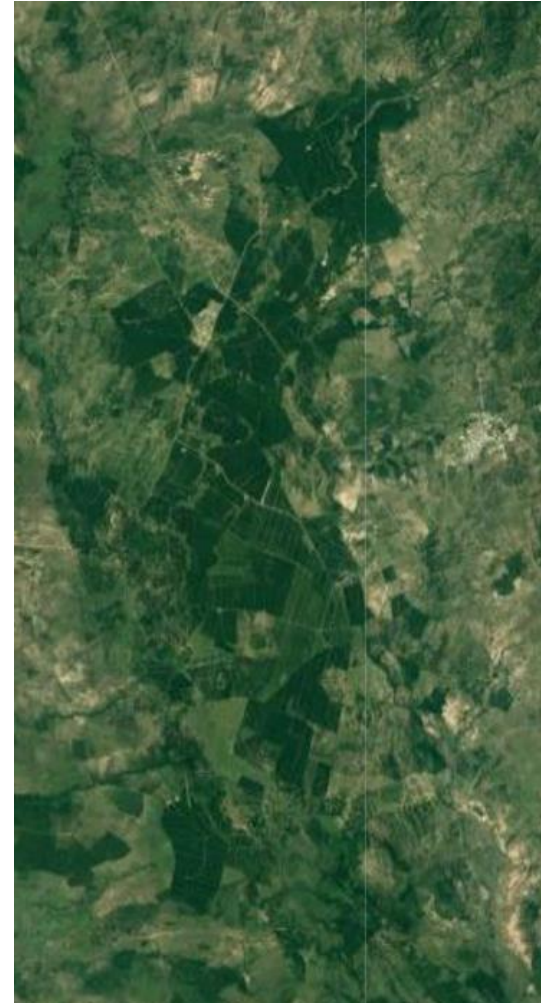


# ПРИМЕРЫ РЕШЕНИЯ ПРИКЛАДНЫХ ЗАДАЧ

## Интеллектуальная классификация и выделение ландшафтов



Анализ пораженности местности термокарстовыми озерами с помощью алгоритмов Random Forest



Выделение техногенно-преобразованных зон с помощью алгоритмов Random Forest



# ПРИМЕРЫ РЕШЕНИЯ ПРИКЛАДНЫХ ЗАДАЧ

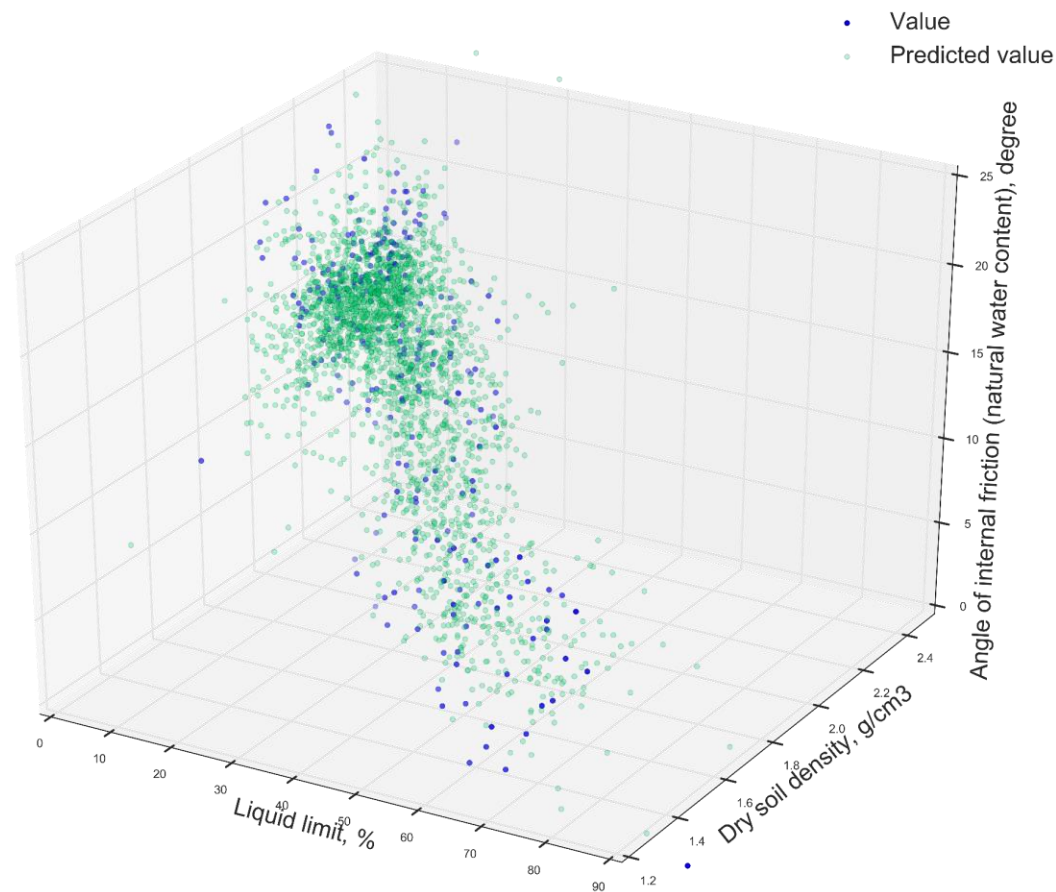


Определение участков с повышенной облачностью с помощью алгоритмов компьютерного зрения

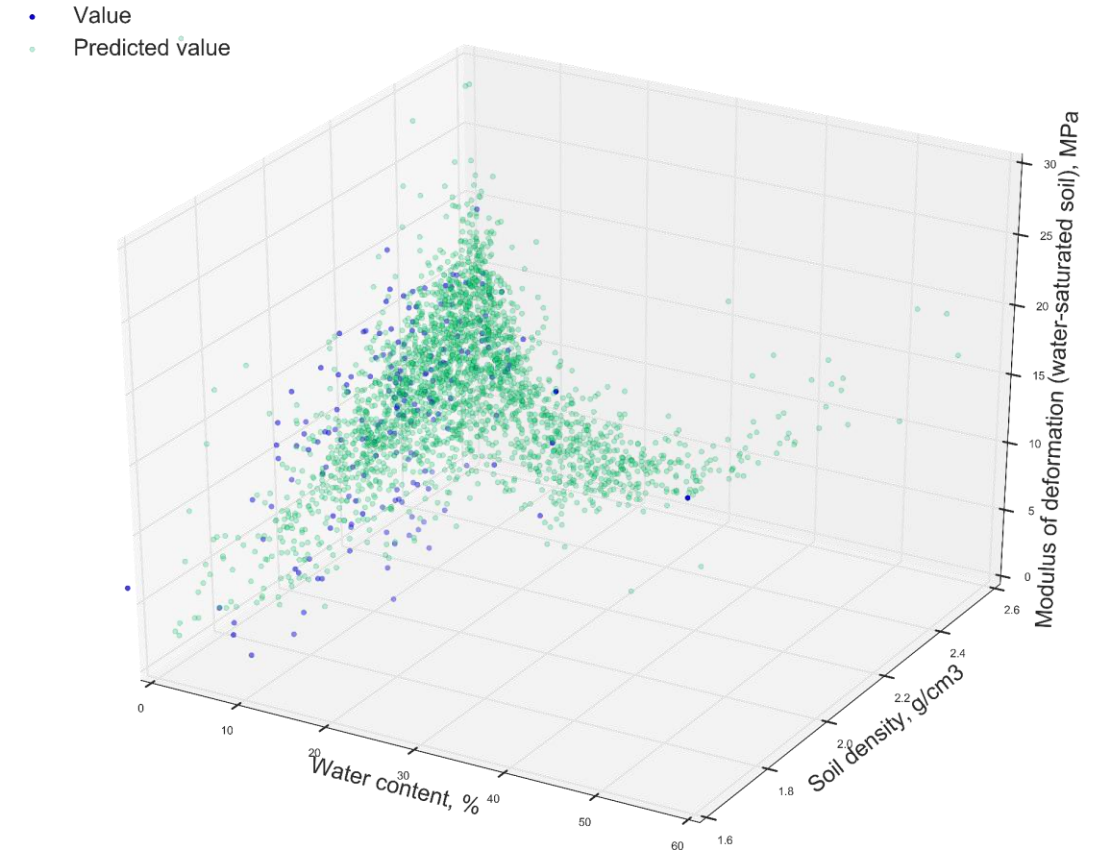
# ПРИМЕР МАШИННОГО ОБУЧЕНИЯ С УЧИТЕЛЕМ (ЗАДАЧА РЕГРЕССИИ)

# ПРИМЕРЫ РЕШЕНИЯ ПРИКЛАДНЫХ ЗАДАЧ

## Решение задач регрессии



Моделирование угла внутреннего трения с помощью алгоритмов машинного обучения

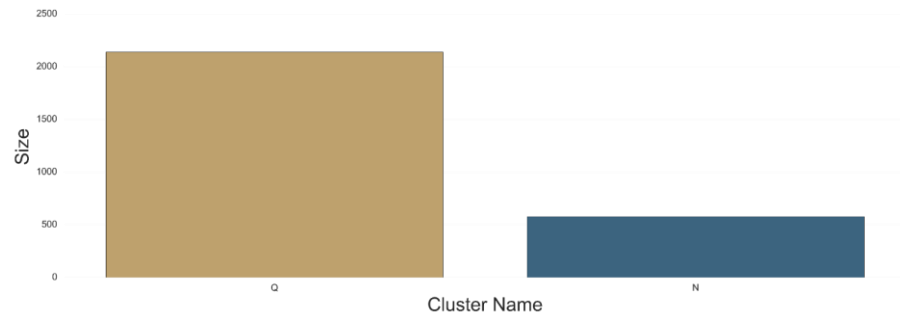


Моделирование модуля деформации с помощью алгоритмов машинного обучения

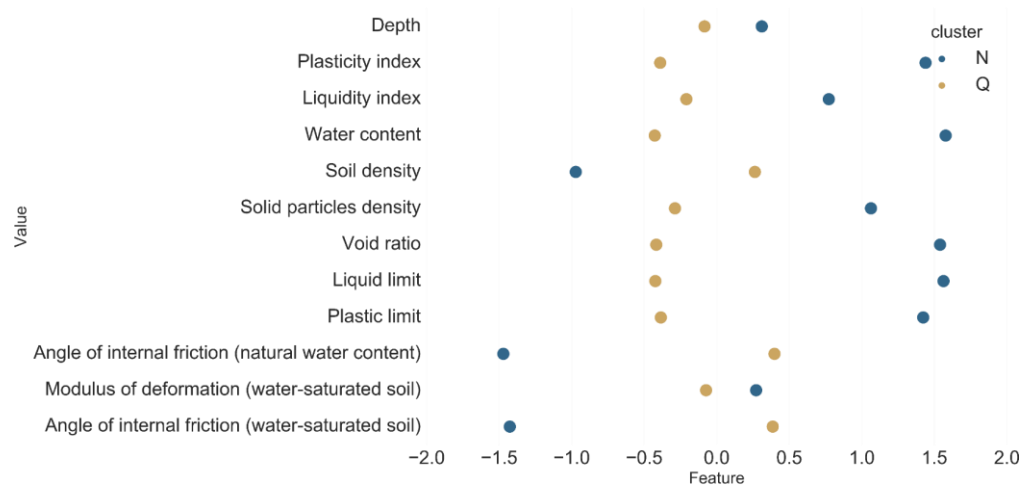
# ПРИМЕР МАШИННОГО ОБУЧЕНИЯ БЕЗ УЧИТЕЛЯ (ЗАДАЧА КЛАСТЕРИЗАЦИИ)

# ПРИМЕРЫ РЕШЕНИЯ ПРИКЛАДНЫХ ЗАДАЧ

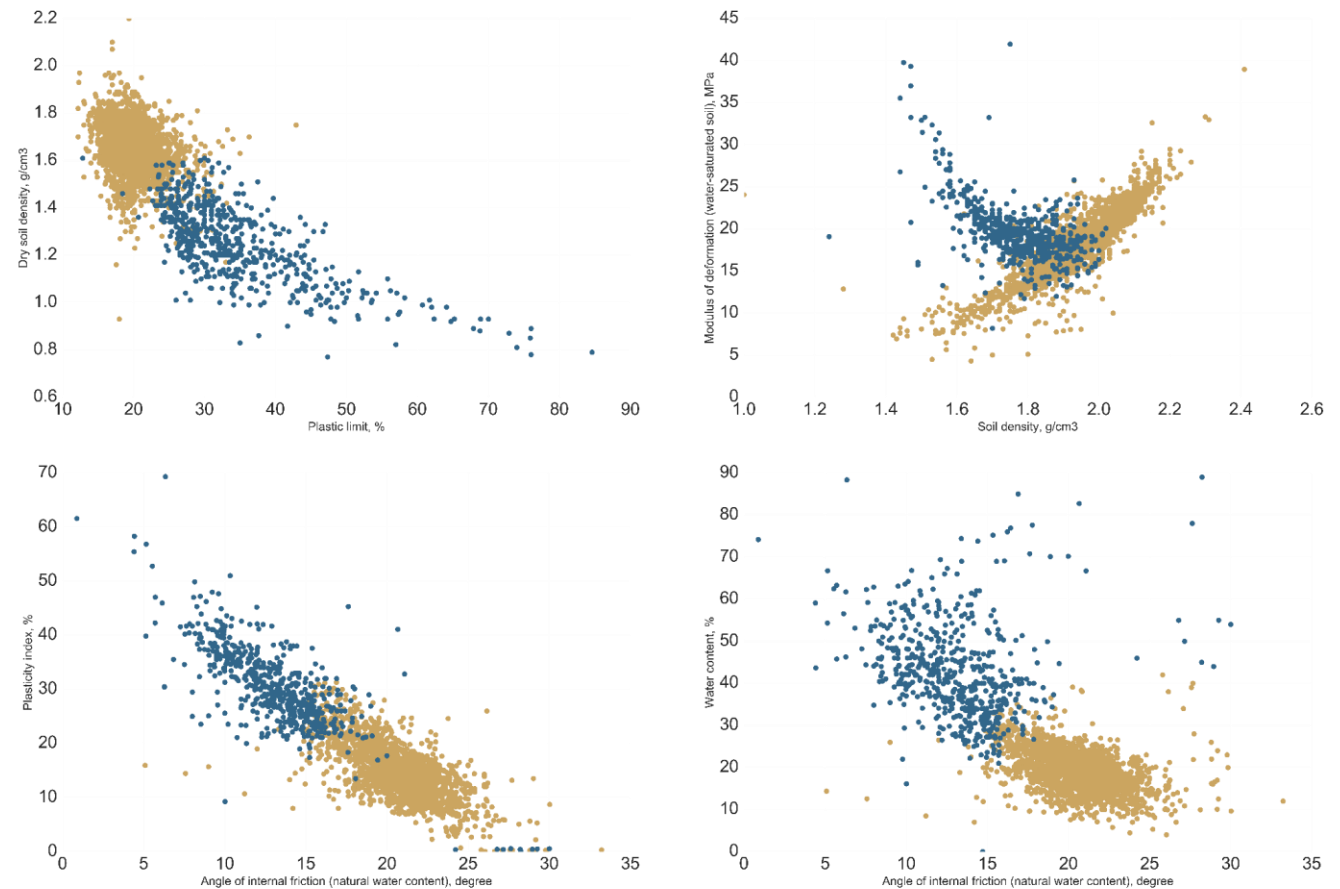
## Решение задач кластеризации данных



### Выделение размеров кластеров



### Проверка обоснованности выделения кластеров

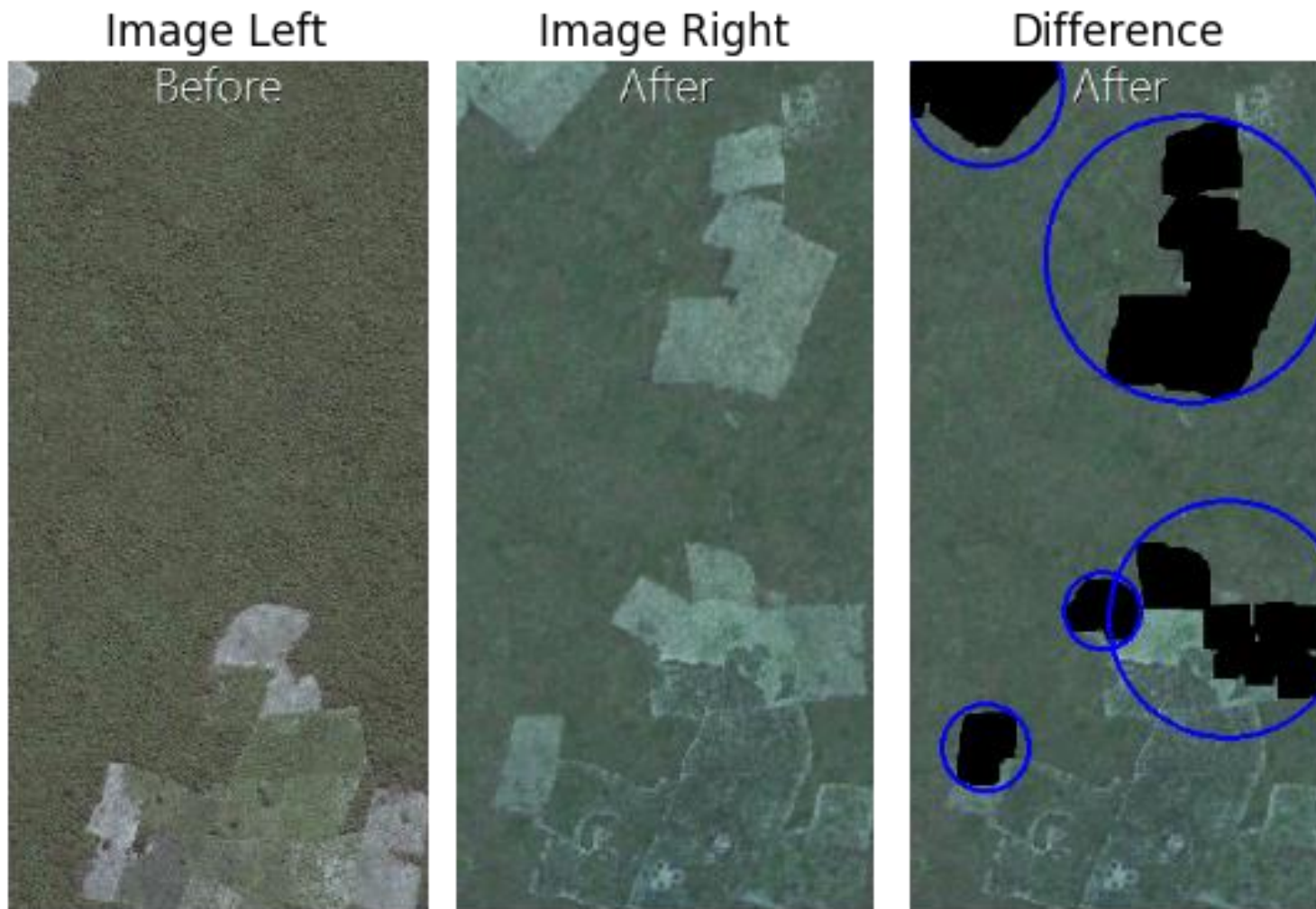


### Графическая визуализация и интерпретация данных

# ПРИМЕР ЗАДАЧИ МАШИННОГО ЗРЕНИЯ



# ПРИМЕРЫ РЕШЕНИЯ ПРИКЛАДНЫХ ЗАДАЧ



Пример автоматического определения преобразования местности с помощью алгоритмов компьютерного зрения

Бинаризация изображений с помощью алгоритмов компьютерного зрения



Расчет геомасок



Расчет пиксельной разности между масками



Автоматическое определение изменений



Московский Государственный Университет  
им. М.В. Ломоносова

СПАСИБО!