

Машинное обучение  
Лекция №1, осень 2022

# Вводное занятие по курсу «Машинное обучение»



# О курсе

- Машинное обучение (осень) – дифференциальный зачет
  - 15 недель занятий.
    - ~7 заданий – допуск к зачету +1 доп. балл;
    - 14 тестов (перед каждым семинаром) +1 доп. балл;
    - Работа на семинарах +1 доп. балл;
  - Устный зачет.
    - 2 случайных билета по темам курса;
    - 1 тема на выбор студента вне программы курса (вопрос по выбору);
    - Доп. Вопросы по курсу.
- Глубокое обучение (весна) – экзамен

# Программа курса

1. Введение в машинное обучение.
2. Naïve Bayes, kNN.
3. Линейные модели.
4. Логистическая регрессия.
5. SVM, PCA.
6. BVD, k.
7. Деревья решений. Методы ансамблирования моделей.
8. Градиентный бустинг.
9. Введение в нейронные сети.
10. Методы кластеризации и понижения размерности.
11. Неградиентная оптимизация.
12. Задачи ранжирования и матчинга.
13. Практика решения задач.

# Формат курса

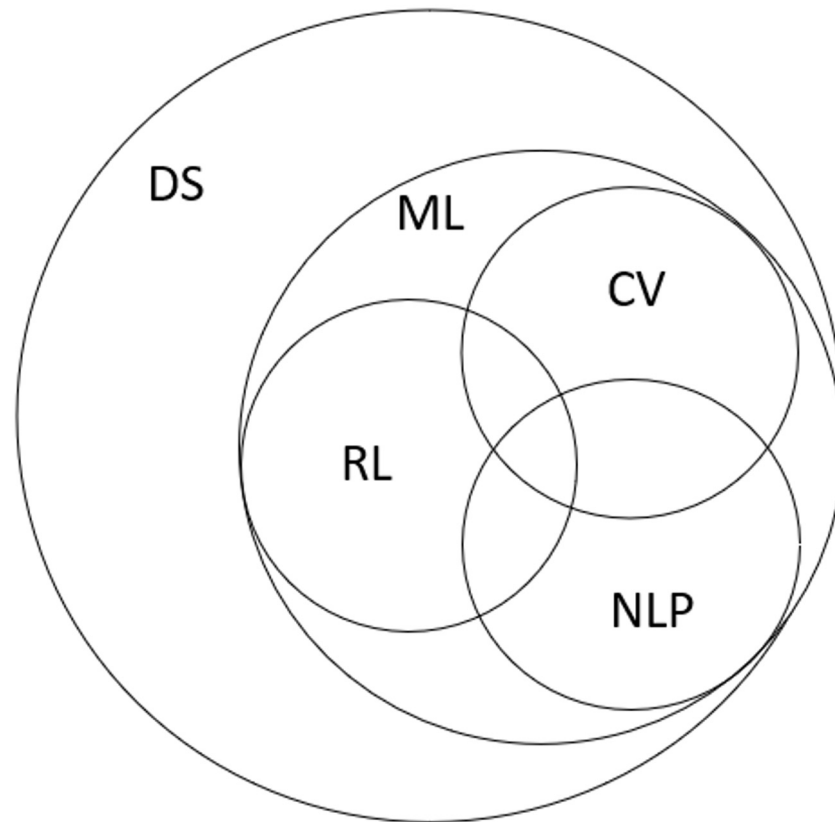
## Оценка за курс

- устный ответ на билеты по программе курса
    - 2 вопроса из программы (на зачёте)
    - 1 вопрос по выбору по теме курса
  - 3 дополнительных балла
    - работа на семинарах
    - решение дополнительных заданий
    - решение тестов в начале семинаров
- 
- Лекционные занятия: записи и очные
  - Семинарские занятия (по группам): очно
  - Домашние задания с фиксированным дедлайном
    - Проверка семинаристами по группам

# Введение

Три основных области исследований в ML (Machine learning)

1. CV (Computer Vision)
2. NLP (Natural Language Processing)
3. RL (Reinforcement Learning)



# Коротко о задачах в ML

Решим задачу

Сколько минут в 3 часах?

# Коротко о задачах в ML

Решим задачу

Сколько минут в 3 часах?

$$f(x) = 60 * x$$

$$f(3) = 60 * 3 = 180$$

# Решим другую задачу

**Мальчик на санках едет с горки. Масса мальчика вместе с санками составляет 40 кг, угол наклона горы  $30^\circ$ . Найдите ускорение, с которым съезжает мальчик, если коэффициент трения скольжения равен 0,2.**



# Решим другую задачу

**Мальчик на санках едет с горки. Масса мальчика вместе с санками составляет 40 кг, угол наклона горы  $30^\circ$ . Найдите ускорение, с которым съезжает мальчик, если коэффициент трения скольжения равен 0,2.**

Дано:

$$m = 40 \text{ кг}$$

$$\alpha = 30^\circ$$

$$\mu = 0,2$$

---

$$a = ?$$

$$m\vec{a} = \vec{N} + m\vec{g} + \vec{F}_{\text{тр}}$$

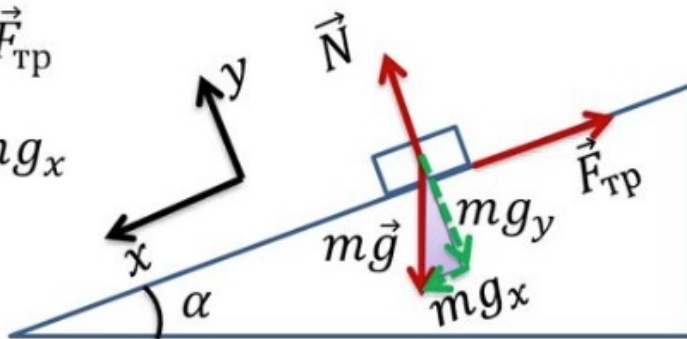
$$X: ma = -F_{\text{тр}} + mg_x$$

$$Y: 0 = N - mg_y$$

$$N = mg_y$$

$$F_{\text{тр}} = \mu N = \mu mg_y$$

$$ma = mg_x - \mu mg_y$$



$$mg_x = mg \sin \alpha$$

$$mg_y = mg \cos \alpha$$

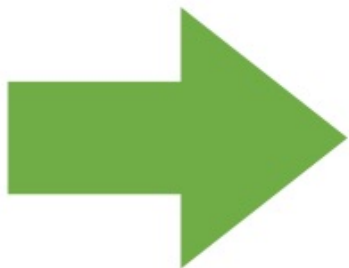
# А что если?

- система сложнее?
- процесс сложнее?
- мы не имеем представления, как он устроен?
- мы не понимаем, как параметры внутри влияют друг на друга?

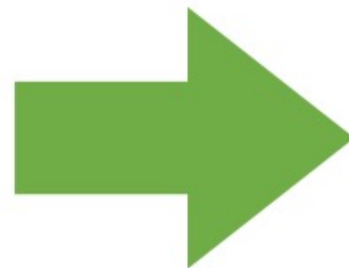
# Попробуем найти зависимость



Провести тысячи экспериментов



Вжух!



$f(x)$

Найти истинную функцию

# Попробуем найти зависимость



Провести тысячи экспериментов

Вжух!

Найти истинную функцию

**Machine  
Learning**

(Статистика)

# Определения

Машинное обучение – это процесс, в результате которого машина (компьютер) способна показывать поведение, которое в нее не было явно заложено (запрограммировано).

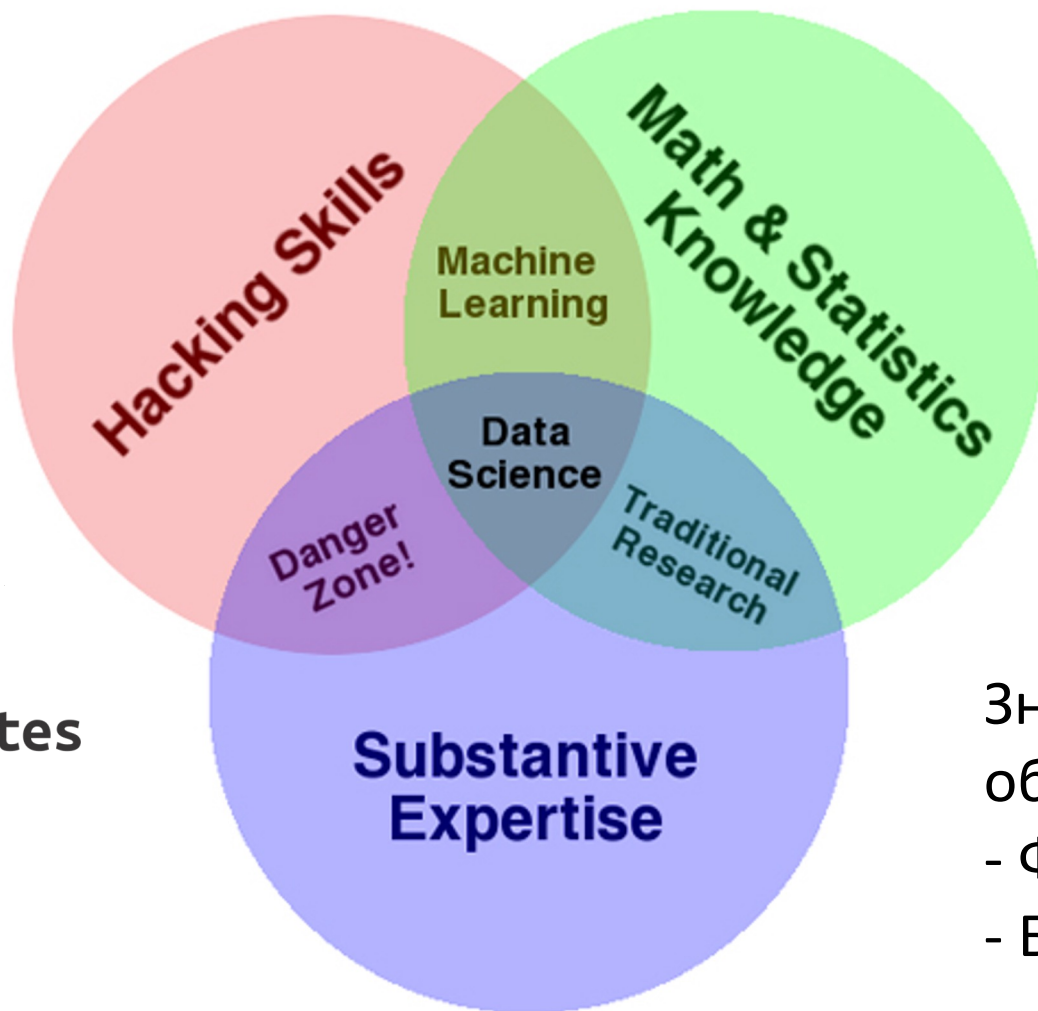
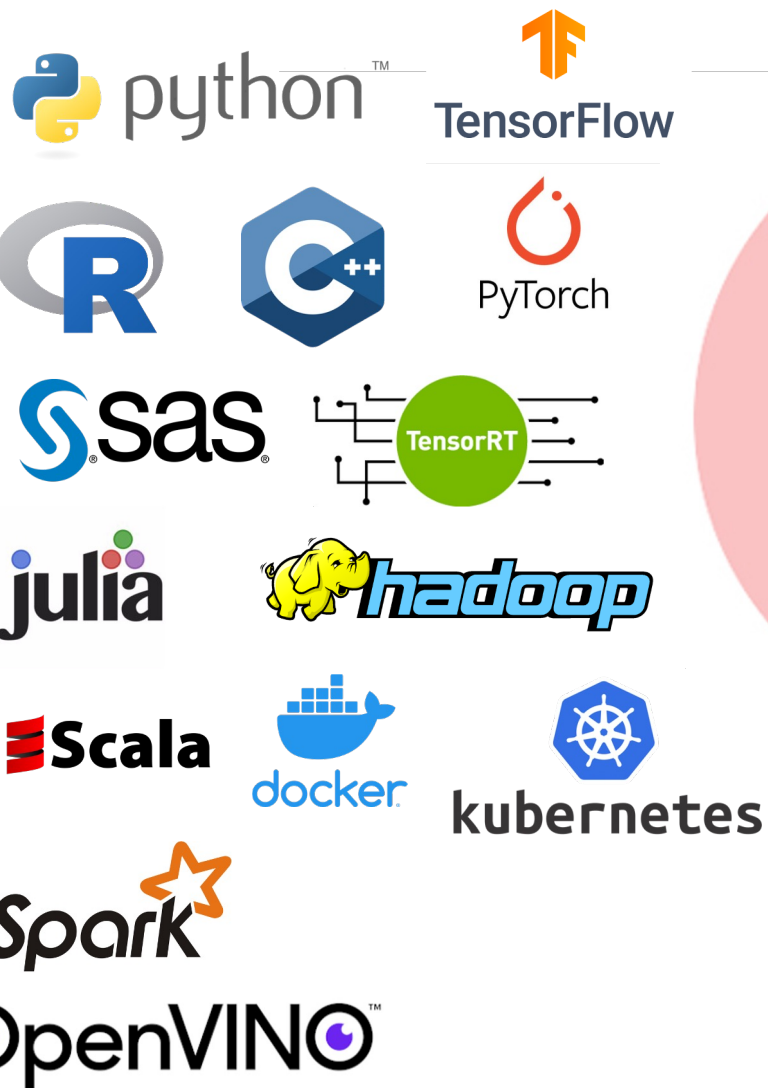
*Артур Самуэль, 1959*

Компьютерная программа обучается при решении какой-то задачи из класса  $T$ , если ее производительность, согласно метрике  $P$ , улучшается при накоплении опыта  $E$ .

*Том Митчелл, 1997*



# Что нужно знать и уметь (но это неточно)



Фундаментальное образование:

- Математический анализ
- Линейная алгебра
- Алгоритмы и структуры данных
- Вычислительная математика
- Математическая статистика

Знания в доменной области:

- Физика
- Биология
- ...

# Роли в машинном обучении

## Рабочий процесс

