

Машинное обучение

А не искусственный интеллект

Whois?

- Семен Стихин (можно на ты, Сёма или Сеня)
- Год в Контуре как Аналитик
- Сейчас почти год в Мегафоне как Аналитик данных
- Закончил ШАД Яндекса по треку Data Scientist
- 3-ий год веду семинары по терверу и статистике на MatMechе
- Аспирант УрФУ, занимаюсь МЛ + управление

Контакты: [samstikhin](#) ([telegram](#), [твиттер](#) ...)

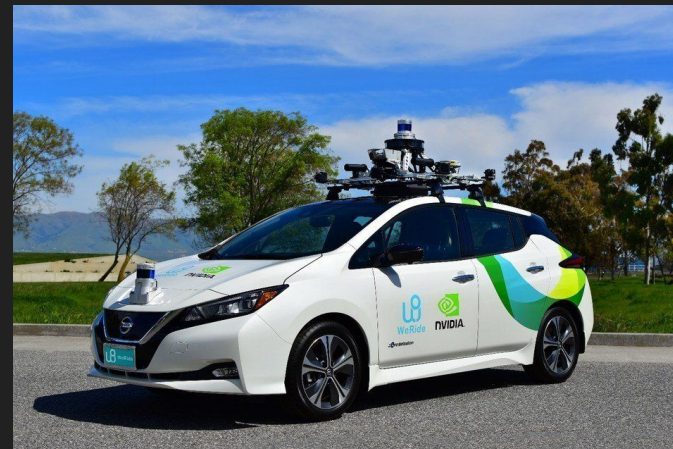
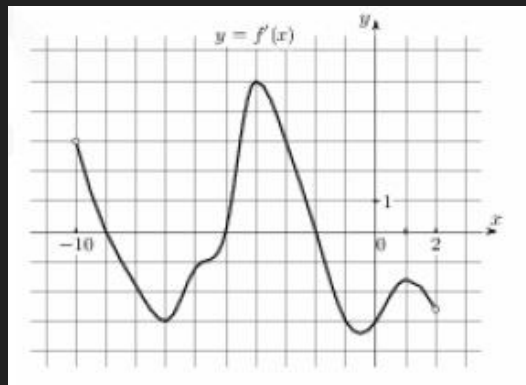
Почта: samstikhin@gmail.com

Регламент

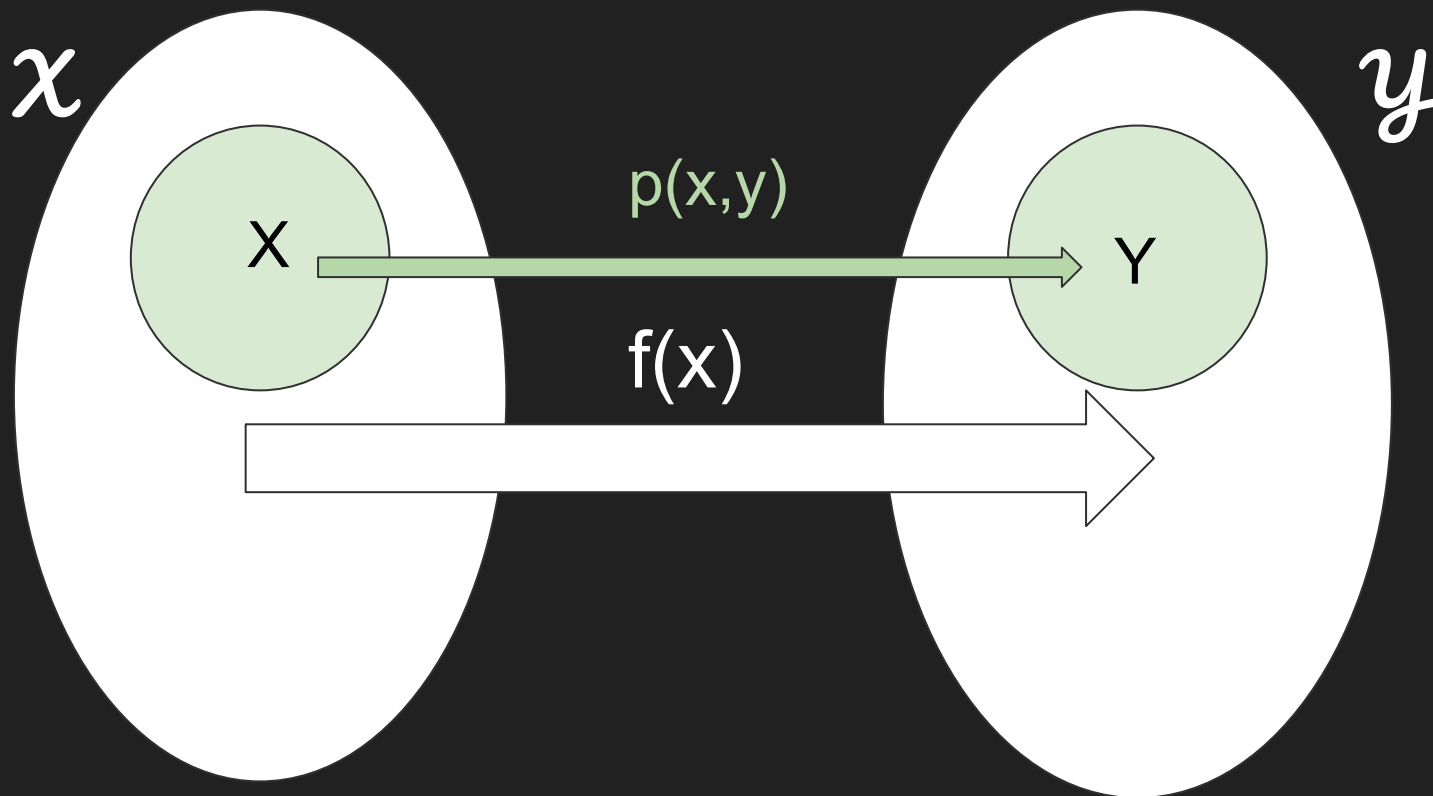
- **Задавать глупые вопросы!**
- Всем записаться в группы
- Срок ДЗ - 1 неделя, после только 50%
- Видеозаписи будут
- Длительность: 3 часа + 2 перерыва на 10 мин
- Курс стоит 6 зет
 - 3-4 курс ФИИТ: “Машинное обучение и питон” + “Теория игр”
 - Остальные сами по себе

Вопросы?

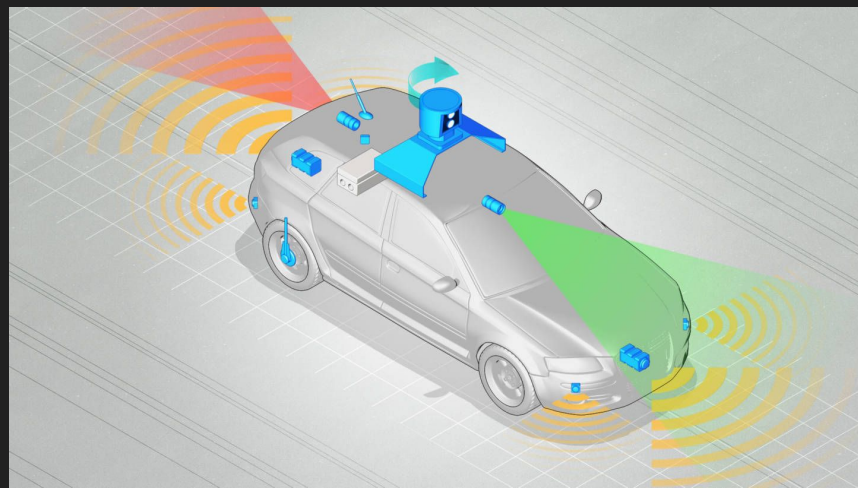
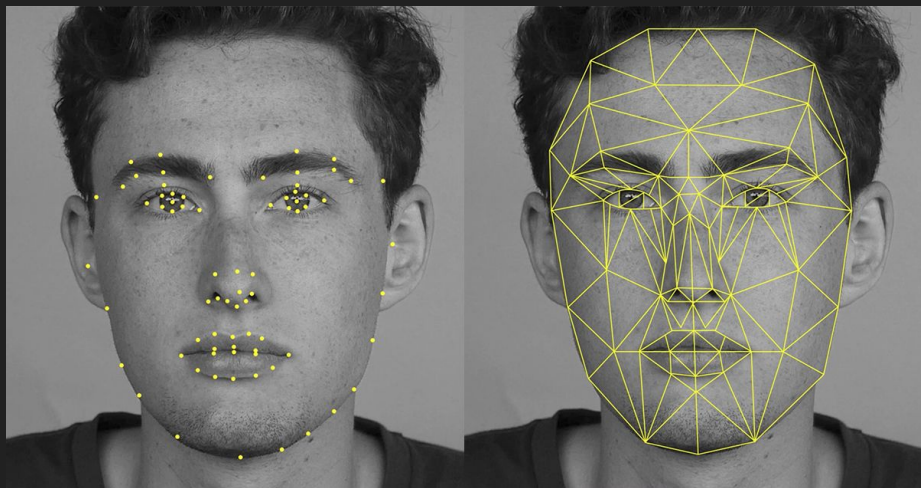
Примеры?



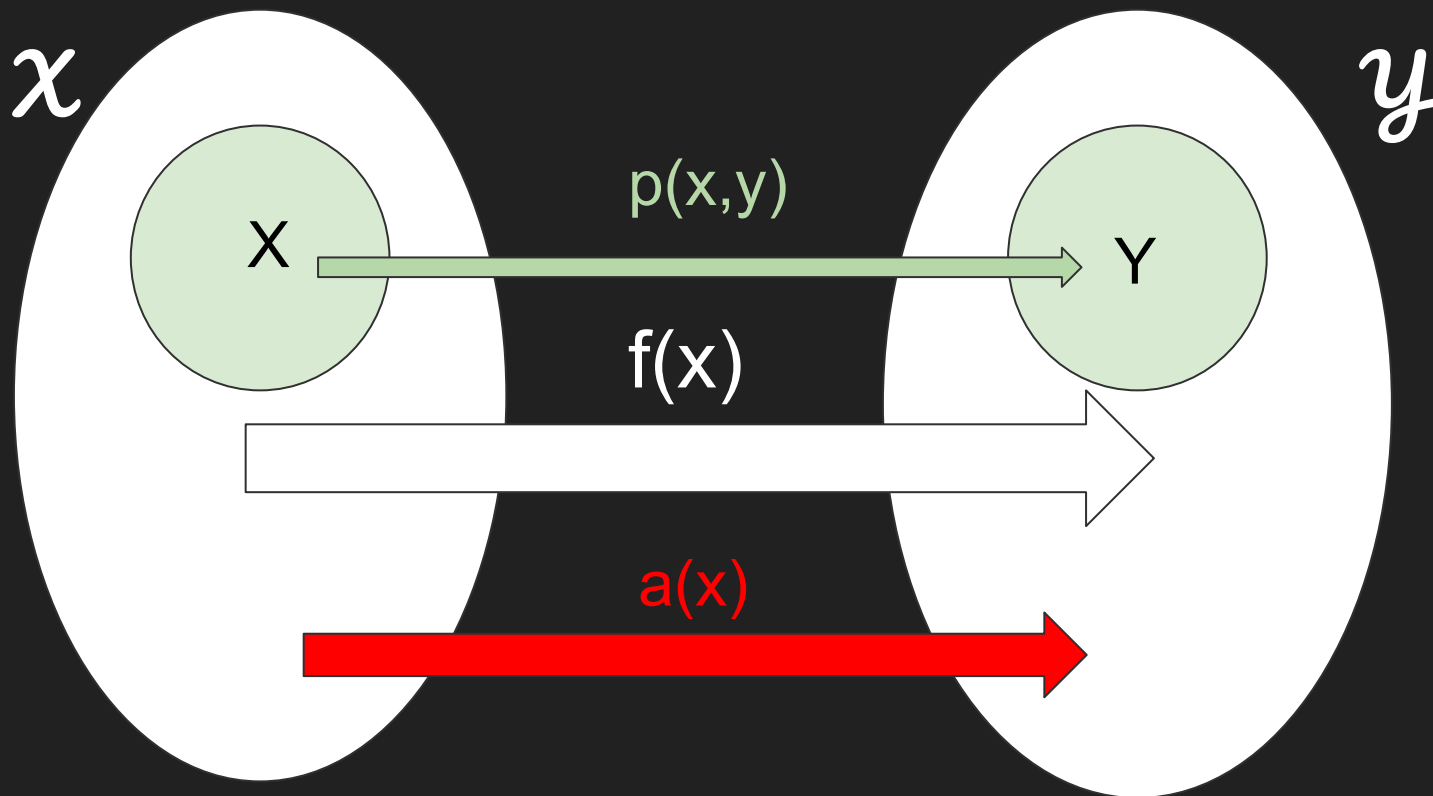
Общая схема



Разбор кейсов



Алгоритм обучения



Пояснения

$p(x,y)$ или (X,Y) - выборка (sample)

$f(x)$ - некоторая “идеальная” зависимость

$a(x; (X,Y))$ - модель (алгоритм), обученная на (X,Y)

Что такое $a(x)$?

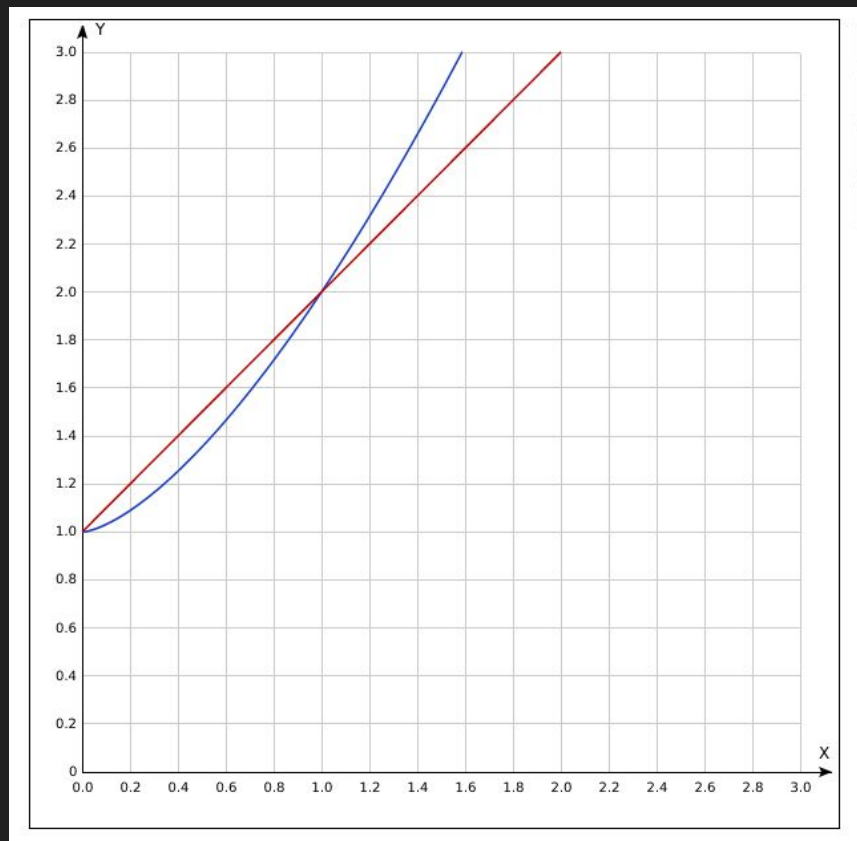
Нейронка, Регрессия, Деревья ...

Функция с параметрами (X, Y)

$$p = \{(0,1), (1,2)\}$$

$$f = x^{1.2} + 1$$

$a = ?$



Процесс обучения

Подбор наилучшей функции $a(x)$ в некотором классе по параметрам (X, Y)

$$\sum_{(x,y) \in (X,Y)} dist(a(x), y) \rightarrow min$$

$$\sum_{(x,y) \in (X,Y)} |a(x) - y| \rightarrow min$$

Профессии

- Data Engineer (DE) - (Дата) Инженер
- Data Analytics (DA) - Аналитик (больших) данных
- Data Scientist (DS)
- ML Engineer - Разработчик машинного обучения

Профессии

- Data Engineer - собирает/хранит/обрабатывает данные
- Data Analytics - анализирует данные, обучает модели
- Data Scientist - придумывает новые классы алгоритмов
- ML Engineer - пишет алгоритмы для применения на практике

Очень классные ресурсы

- [slack ods.ai](#) (ru)
- [twitter omarsar](#) (en)
- [youtube Deep Learning School](#) (ru)
- [youtube Jeremy Howard](#) (en)
- [Книга по нейронкам](#) (en, есть и ru)
- [Книга по классическому мл](#) (en)