

Системы искусственного интеллекта

ИТМО, 3 курс, 5 семестр

Запорожцев Иван Федорович
zaporozhtsev.if.work@gmail.com

Запорожцев Иван Федорович

Кандидат технических наук
(«Математическое моделирование, численные
методы и комплексы программ»),

(м)учитель студентов (всегда),

Научный сотрудник, разработка WebGIS
в НИИ морской биологии (до 2020),

ML Engineer, контроль и прогноз добычи
в «Газпром нефть» (до 2021)

ML Team Lead,
оптимизация в металлургии,
«Норильский никель» (с 2021)



Системы искусственного интеллекта

Кол-во: 8 лекций, 16 практик

Что: формальные задачи, матметоды, код, примеры из жизни

Зачем: мотивация к изучению DS, ML, DL, написанию кода => участвовать в хакатонах, пилить пэт-проекты, готовить резюме

Где: платформы, мессенджеры

Полезные ресурсы

Open Data Science

<https://ods.ai/>

<https://www.youtube.com/@ODSAIRu>

[Хакатоны](#) || [Hackathon list](#) || [it-соревнования \(vk.com\)](#)

Блог Александра Дьяконова

<https://alexanderdyakonov.wordpress.com/>

Лекции Сергея Николенко (СПбГУ)

<https://www.youtube.com/@snikolenko/videos>

Нестареющая классика

<https://habr.com/ru/companies/vk/articles/344044/>

[Сбер, Москва - Технологии, меняющие мир / Статьи / Хабр \(habr.com\)](#)

А знаете ли вы ...?

Train and Test Datasets, Inference

Fit, Overfit, Underfit

Loss Function and Metrics

Supervised, Unsupervised, Semi-supervised Learning

Supervised: Regression, Classification

Unsupervised: Clustering, Dimensionality Reduction

Reinforcement Learning

3 “V” of Big Data: Volume, Velocity, Variety

Computer Vision: Detection, Segmentation, Tracking

Natural Language Processing, Large Language Models,

Recommendation Systems

ИИ, ML, DL, разработка



Когнитивные способности

Восприятие: зрение, слух

Распознавание образов

Классификация и категоризация

Анализ смысла и контекста

Анализ ситуации и сопоставление с целью

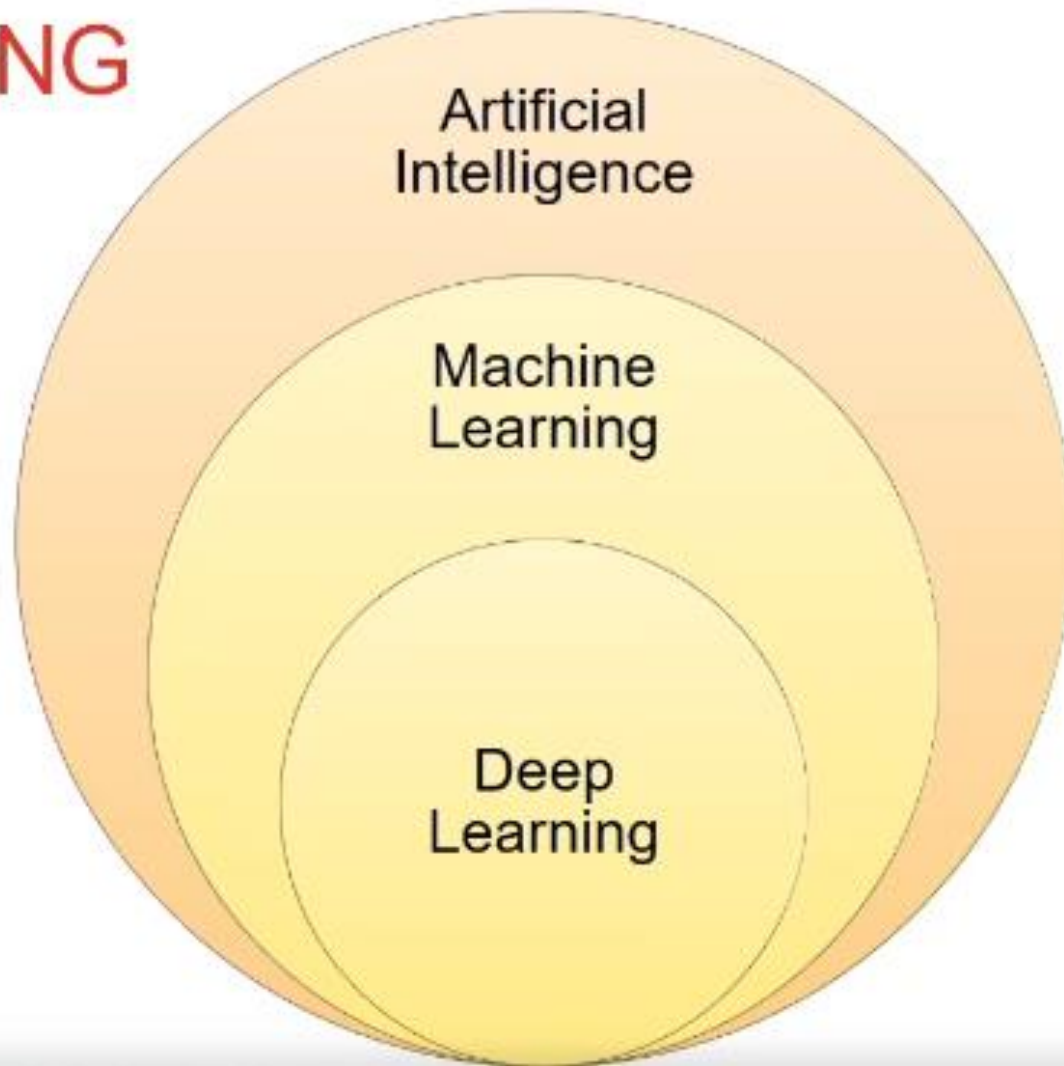
Моделирование вариантов

Выбор наилучшего и построение стратегии

Превращение стратегии в отдельные действия

Сознание и постановка целей

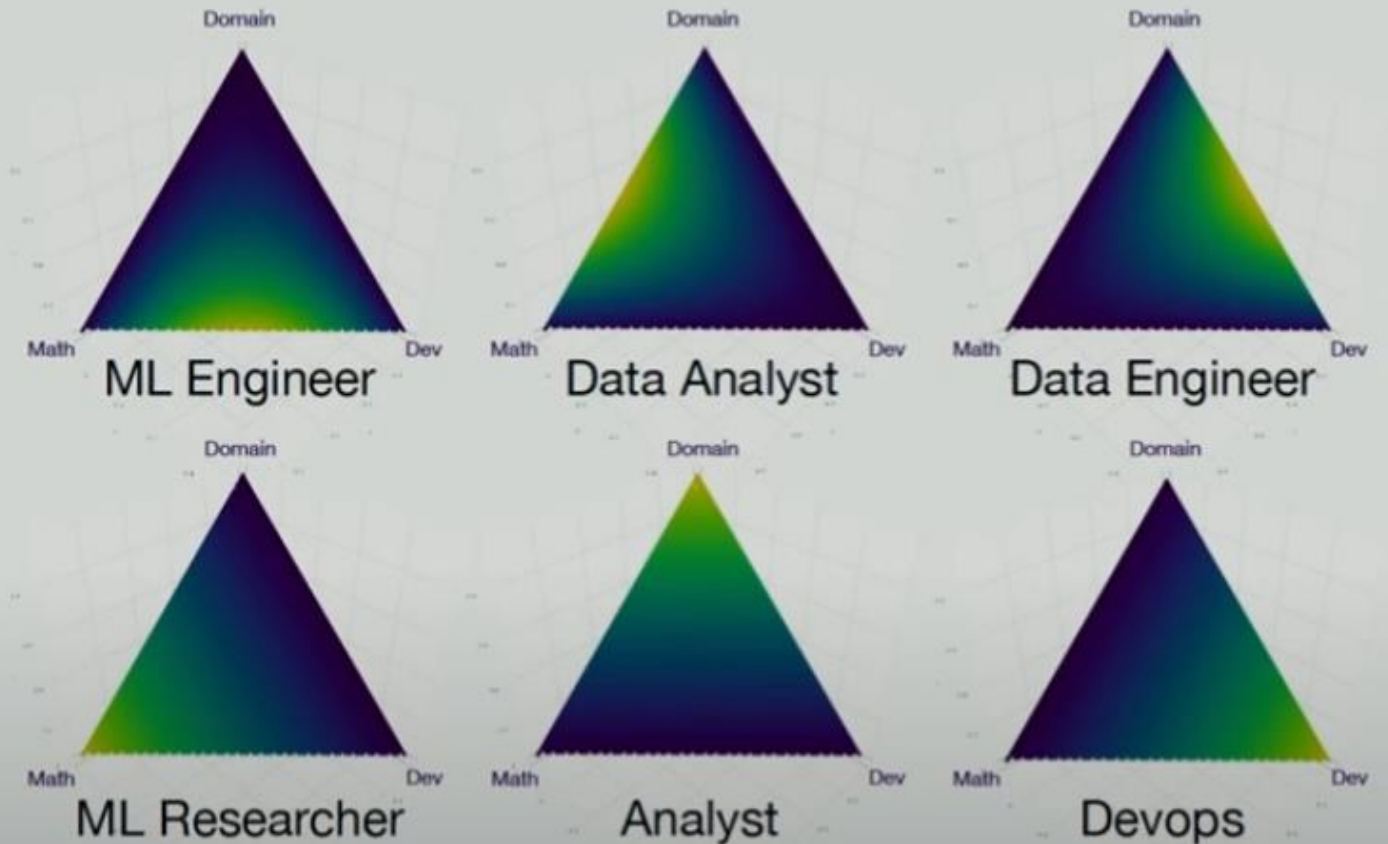
AI & MACHINE LEARNING



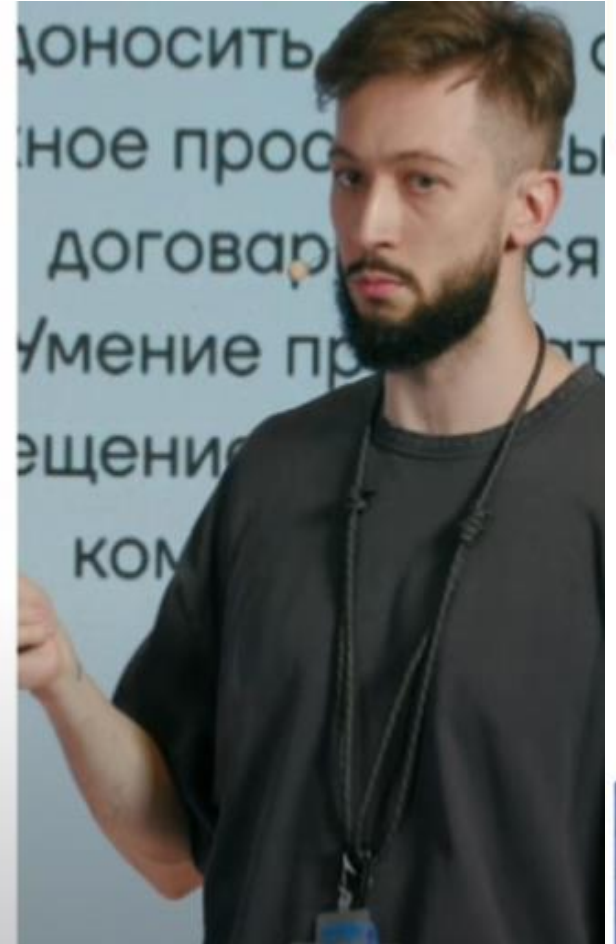
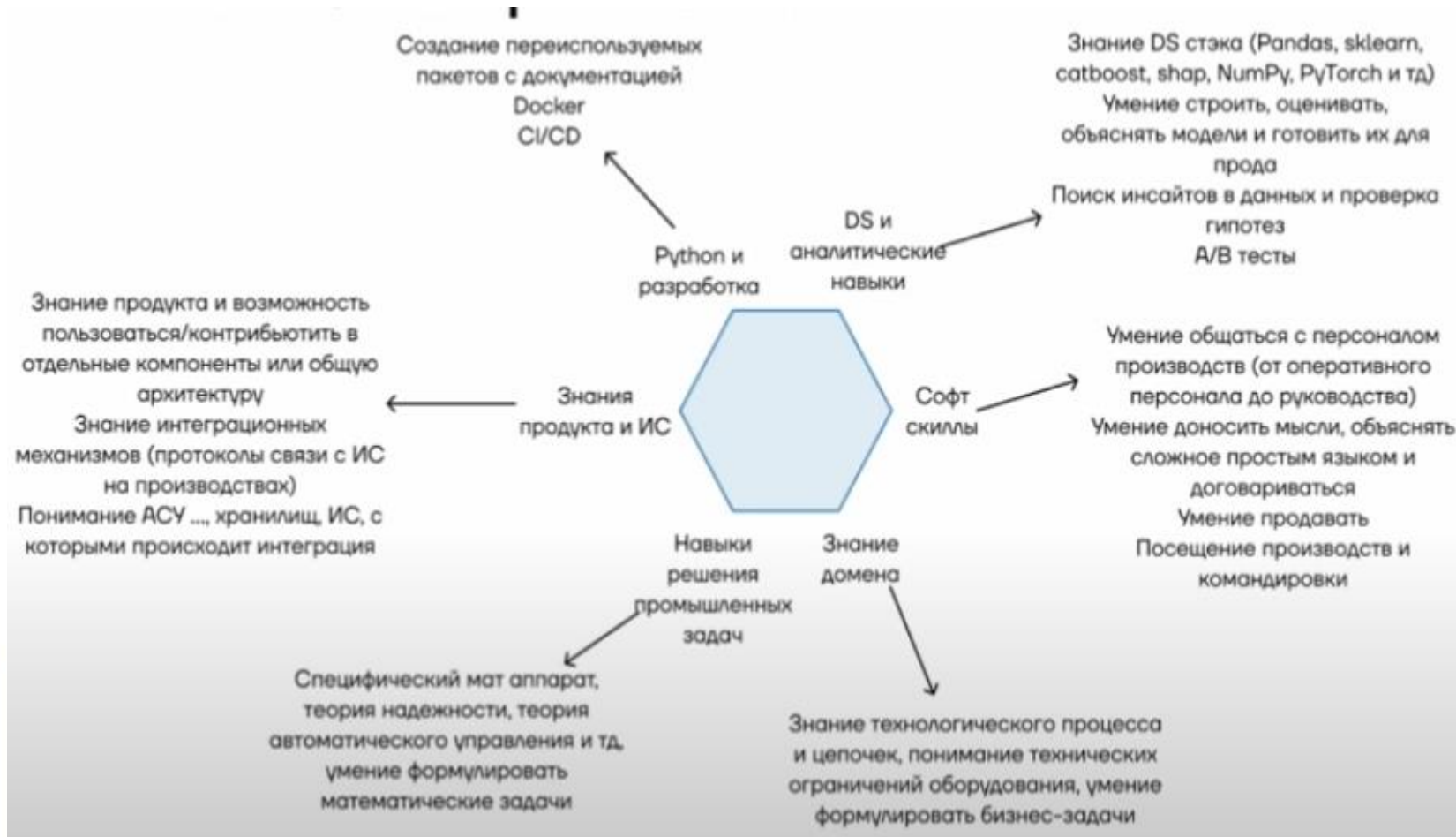
ИИ, ML, DL, разработка... и реальность



Алексей Натёкин



ИИ, ML, DL, разработка... и реальность



Юрий Кацер

Нейробиология, психология для ИИ

1. «Учить» алгоритмы, как ребенка (много попыток, случайный выбор, механизмы поощрения и наказания)
2. Человек неидеален (субъективизм, не может работать 24/7, не может держать в фокусе внимания сотни признаков одновременно)
3. Исследовать природу, поведение живых организмов (кроссинговер и мутации, движение муравьиной колонии, нейронная сеть мозга, решение задач с помощью квантового компьютера и молекул ДНК в пробирке как биологического компьютера)
4. Формализовать мышление, способ принятия решений человеком!

ИИ и общество

1. **Этический и юридический уровень**
2. Накапливать данные о конкретном человеке, о популяции для решения задач конкретного человека (советчик по диете, двигательной активности)
3. Автоматизация части рутины, коммуникации, ручной работы (в том числе «Умный дом», IoT, экзоскелеты, AR/VR, нейроинтерфейсы)
4. Готовить специалиста из любого домена к data-driven-коммуникации/управлению
5. Перенимать опыт решения задач одного домена в другом (U-net для клеток, LLM как пролог к мультимодальной обработке и управлению!)

Машинное обучение

Совокупность математических методов, позволяющих компьютеру (программно-аппаратному средству!) обучаться на предоставленных данных (train dataset) и применять знания, полученные при обучении, на новых данных

Результат научный – методика, метод, алгоритм, модель

Результат прикладной – обученная модель для использования в конкретном бизнес-кейсе, приносящая value (benefit, profit)
[value = полученная прибыль – расходы на модель]

Inference:

«модель только вычисляет выход по входу, работа на проде»

Машинное обучение: определение через пример

$$z = 2x + 3y = a_1x + a_2y$$

x, y, z

| Номер итерации обучения | a1 | a2 | Ошибка (error, loss) |
|-------------------------------|---------|----------|-------------------------|
| 0 | 0.1 | 0.4 | 1000 |
| ... | ... | ... | ... |
| 19 | 1.99 | 2.99 | 0.001 |
| 20 | 1.9999 | 3.00001 | 0.000001 |
| 21 | 1.99999 | 3.000001 | 0.0000009 |

Машинное обучение: введение

Матрица «Объект – признак»

| | Признак 1 | Признак 2 | | Признак n |
|-------------|--------------|--------------|-------|--------------|
| Объект 1 | x11 | x12 | | x1m |
| Объект 2 | x21 | x22 | | x2m |
| | | | | |
| Объект m | xm1 | xm2 | | xmm |

MyShared

$X - CB, X \in U$

$X = x_1, X = x_2, \dots$

$\{x_1, x_2, x_3, \dots, x_n\} \subset U$

$X \sim$ закон распределения

$X \sim N(\mu, \sigma)$

$P(X = x) = f(x)$

- i -тый признак,
- i -тое значение (скаляр) как реализация СВ,
- i -тый компонент вектора,
- i -тый вектор,
- скаляр/вектор на i -той итерации

x_i

x_j^i – j -тый компонент i -го вектора

Машинное обучение: основные понятия



Машинное обучение: основные понятия

$$y = f(x), \quad y = a(x, w)$$

- *Прямая задача моделирования:*

$$x^* \Rightarrow y^* : y^* = f(x^*)$$

$$x^*, w^* \Rightarrow y^* = a(x^*, w^*)$$

- *Обратная задача моделирования:*

$$y^* \Rightarrow x^* : y^* = f(x^*)$$

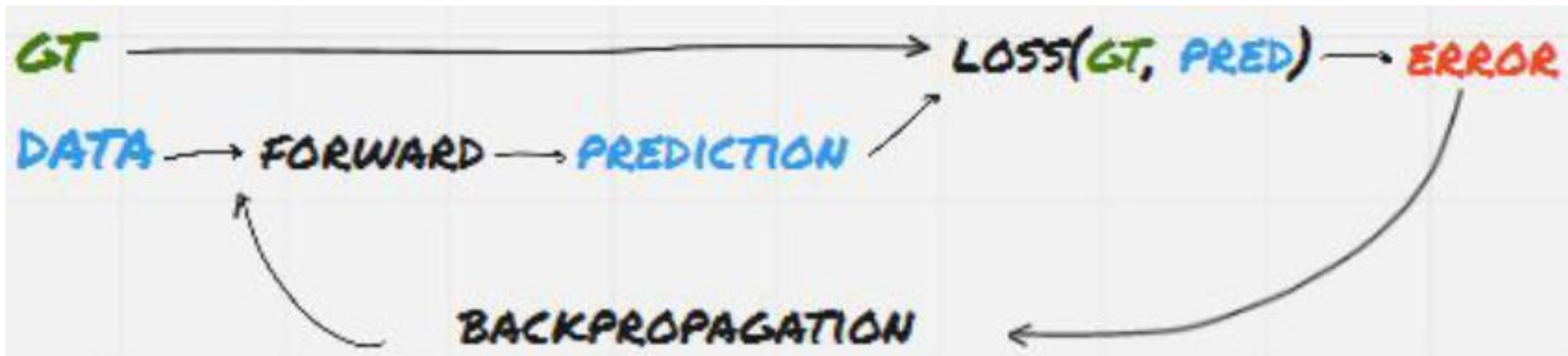
$$y^*, x^* \Rightarrow w^* : y^* = a(x^*, w^*)$$

$$y^*, x^* \Rightarrow w^* = \arg \min_w |y^* - a(x^*, w)|$$

$$y^*, x^*, g^* \Rightarrow w^* = \arg \min_w |y^* - a(x^*, g^*, w)|$$

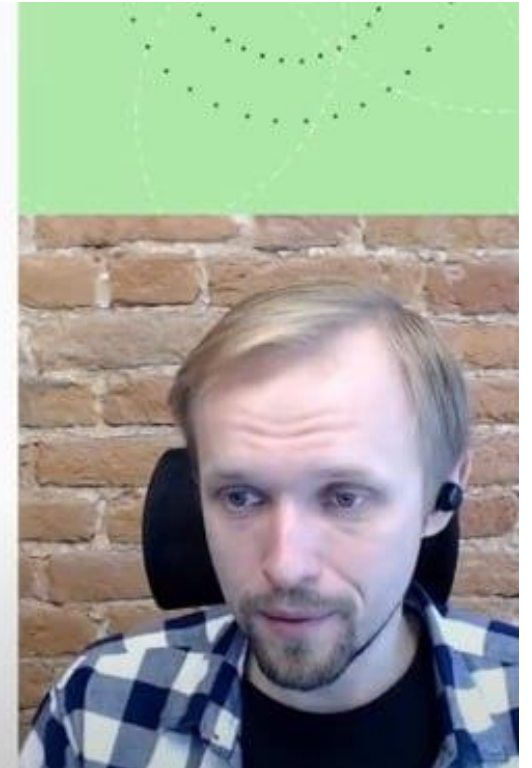
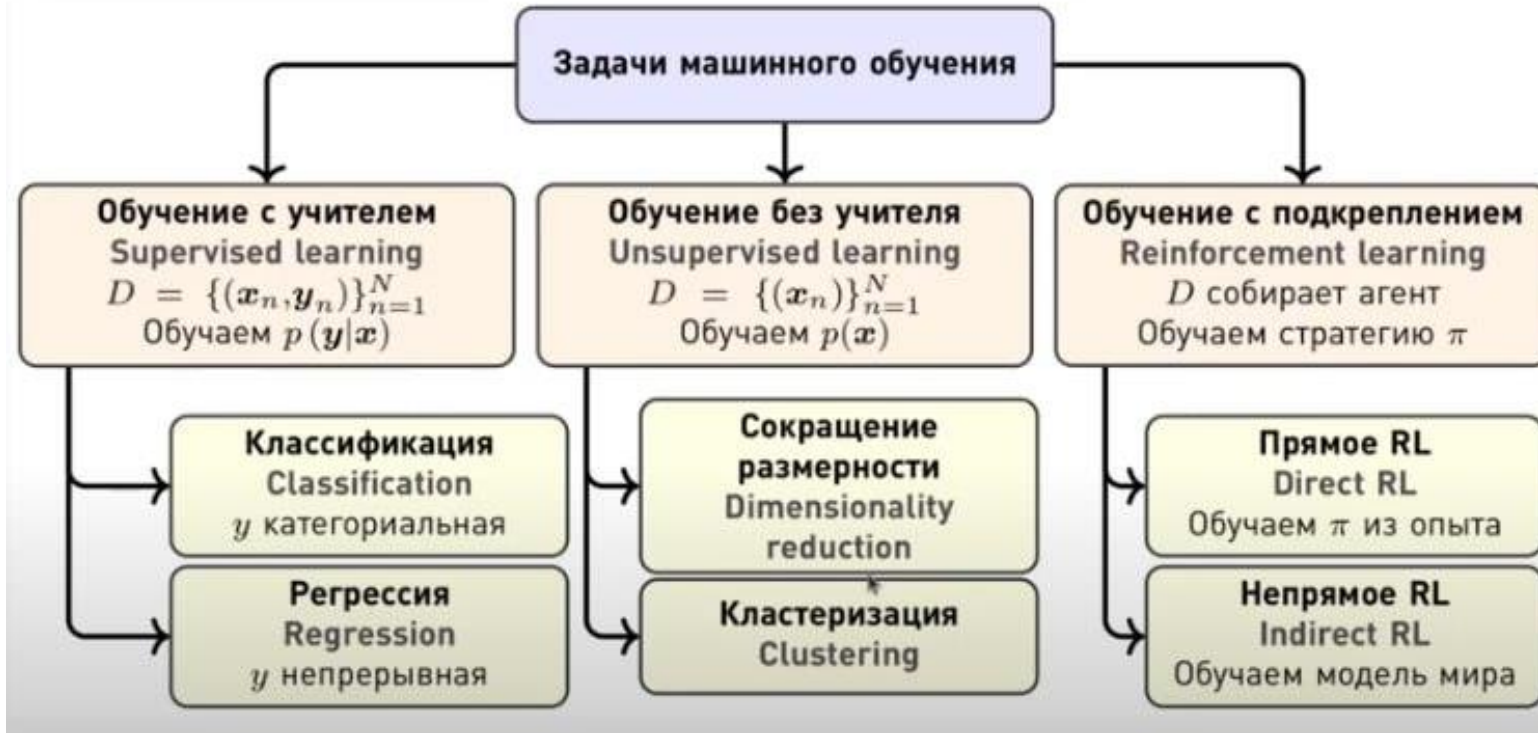
Векторы: x – входы (признаки одного объекта), w – веса, g – гиперпараметры
 y – выход (скаляр),

Машинное обучение: Supervised Learning



Матрица, вектор, скаляр?

Машинное обучение

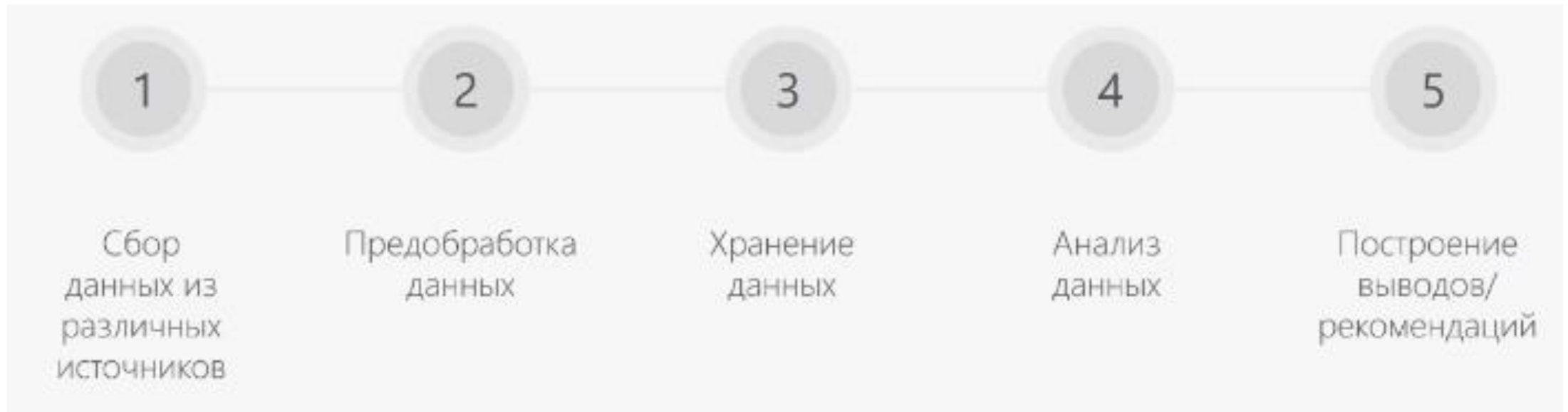


Сергей
Николенко

Машинное обучение: Loss vs Metrics

| Loss | Metrics |
|--|--|
| Функция, которую оптимизируем | Функция, через которую оцениваем |
| Наилучшее качество в <i>минимуме</i> | Нет требования |
| Выбирается одна функция (может быть комплексной) | <i>Неограниченное</i> количество, но есть приоритеты |
| ... | ... |

Машинное обучение: Pipeline



Трудозатраты?

Инновации в «Норильском никеле»

Какие задачи моделирования и управления решаются?

Какие разделы/формальные задачи ИИ задействованы?

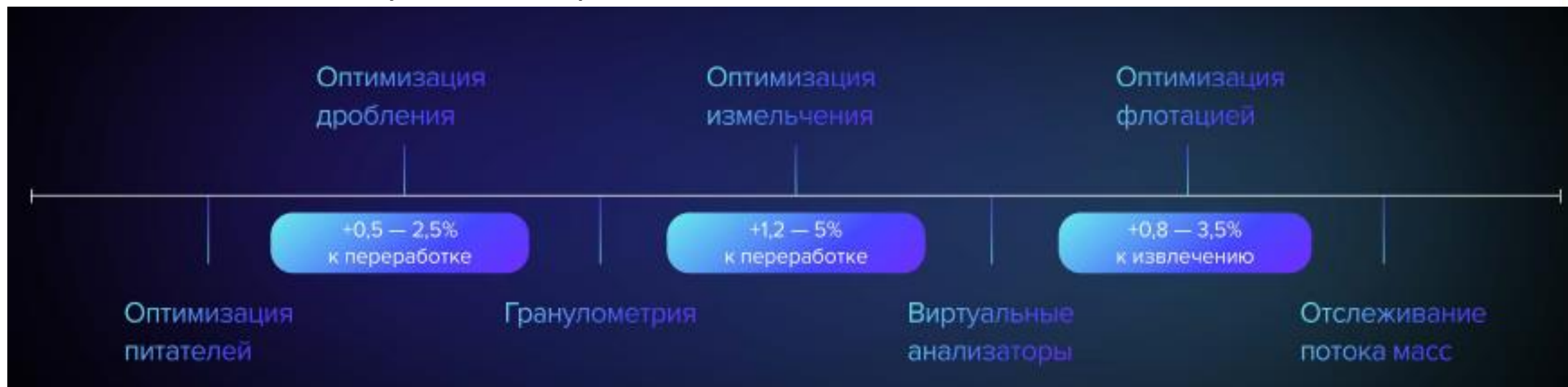
Какие физико-химические признаки для них используются?

Предположите для нескольких признаков: числовые или категориальные, временная и пространственная дискретность регистрации, сырой признак или агрегат по нескольким, сильно зашумленный, много ли пропусков, особенности закона распределения, потенциал для data leakage

Смотрим видео и пробуем ответить!

ИИ в «Норильском никеле»

Зачем нужно внедрять системы на базе ИИ?



100\$ млн

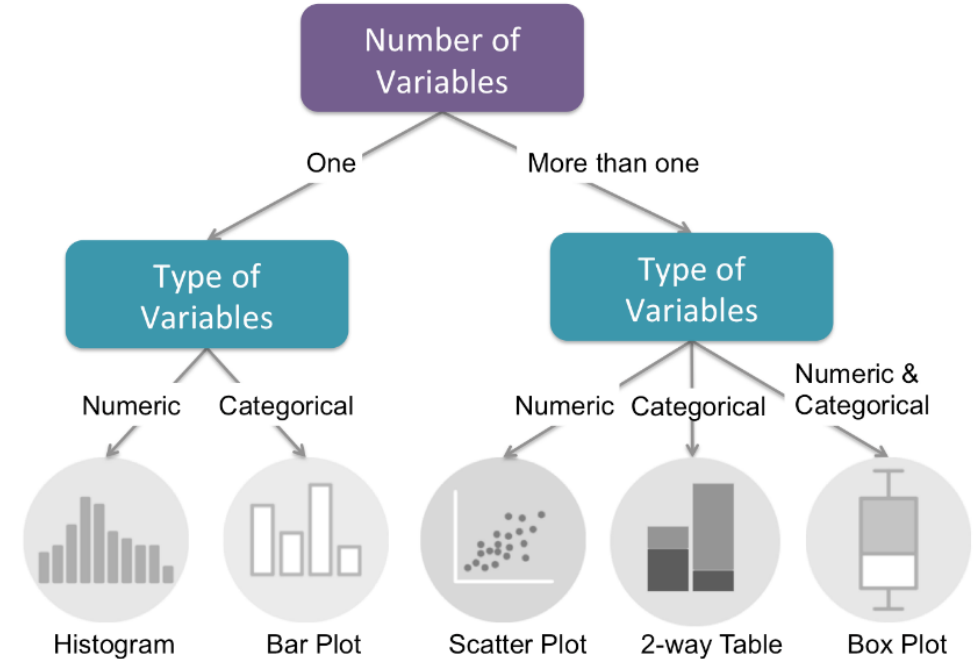
- Больше тонн металла или меньше расход сырья
- Дополнительный контроль, повод для рефлексии
- Упорядочение управления, помощь в оценке
- Помощь стажёрам
- Питательный бульон для будущих результатов

Exploratory Data Analysis

Descriptive statistics

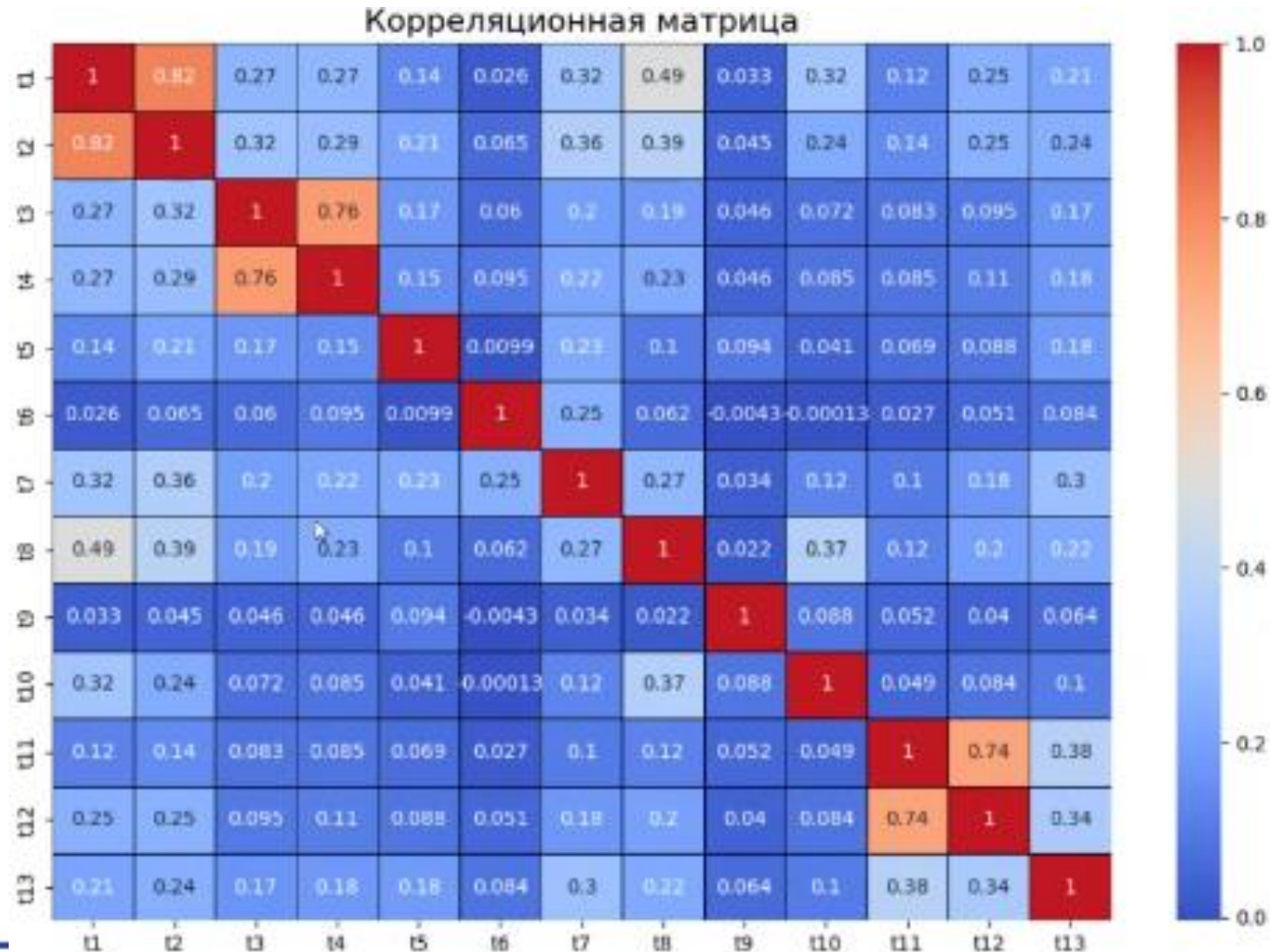
EDA Methods

Visualizations



Designed by Andy Hong (andyhong.org)

Exploratory Data Analysis



Спасибо за внимание!



Запорожцев Иван Федорович
zaporozhtsev.if.work@gmail.com