

Фундаментальные основы искусственного интеллекта

Воронцов Константин Вячеславович

д.ф.-м.н., профессор РАН (МГУ, МФТИ, ФИЦ ИУ РАН)

k.v.vorontsov@phystech.edu

Докладчик: *Воронцов Константин Вячеславович*

http://www.MachineLearning.ru/wiki?title=User:Vokov

Участник:Vokov

Участник:Vokov

Воронцов Константин Вячеславович

профессор РАН, д.ф.-м.н.,
руководитель [лаборатории машинного интеллекта МФТИ](#),
проф. каф. «Интеллектуальные системы» [ФУПМ МФТИ](#),
с.н.с. отдела «Интеллектуальные системы» [Вычислительного центра ФИЦ ИУ РАН](#),
доц. каф. «Математические методы прогнозирования» [ВМК МГУ](#),
преподаватель [Школы анализа данных Яндекса](#),
зам. директора по науке [ЗАО «Форексис»](#), [www.forecsys.ru](#),
один из идеологов и [Администраторов](#) ресурса [MachineLearning.RU](#),
прочие подробности — на подстранице [Curriculum vitae](#).

- [Профиль ORCID = 0000-0002-4244-4270](#)
- [Профиль SCOPUS ID = 6507982932](#)
- [Профиль WoS ResearcherID = G-7857-2014](#)
- [Профиль Google Scholar](#)
- [Профиль DBLP](#)
- [Профиль РИНЦ ID = 15081](#)
- [Профиль в системе ИСТИНА](#)
- [Профиль MathNet.ru](#)

[Мне можно написать письмо.](#)

Содержание [убрать]

- 1 Учебные материалы
 - 1.1 Курсы лекций
 - 1.2 Рекомендации для студентов и аспирантов
- 2 Интервью
 - 2.1 Российский радиоуниверситет, Радио России
 - 2.2 Газеты, журналы, электронные СМИ
 - 2.3 Видеоинтервью
- 3 Доклады на конференциях и семинарах
- 4 Научные интересы
 - 4.1 Анализ текстов и информационный поиск
 - 4.2 Диагностика заболеваний по ЭКГ
 - 4.3 Теория обобщающей способности
 - 4.4 Комбинаторная (перестановочная) статистика
 - 4.5 Прогнозирование объёмов продаж
 - 4.6 Другие проекты и семинары
- 5 Публикации
- 6 Софт
- 7 Аспиранты и студенты
 - 7.1 Бакалаврские диссертации
 - 7.2 Магистерские диссертации

Фундаментальные основы искусственного интеллекта

1. Задачи машинного обучения

- Бум искусственного интеллекта и нейронных сетей
- Терминология машинного обучения
- Модели машинного обучения

2. Методология машинного обучения

- Нейронные сети и глубокое обучение
- Оптимизационные задачи машинного обучения
- Задачи машинного обучения с векторизацией объектов

3. Вопросы практического применения

- Особенности реальных данных
- Необходимые условия применения ИИ
- Распределение ролей в проектах ИИ

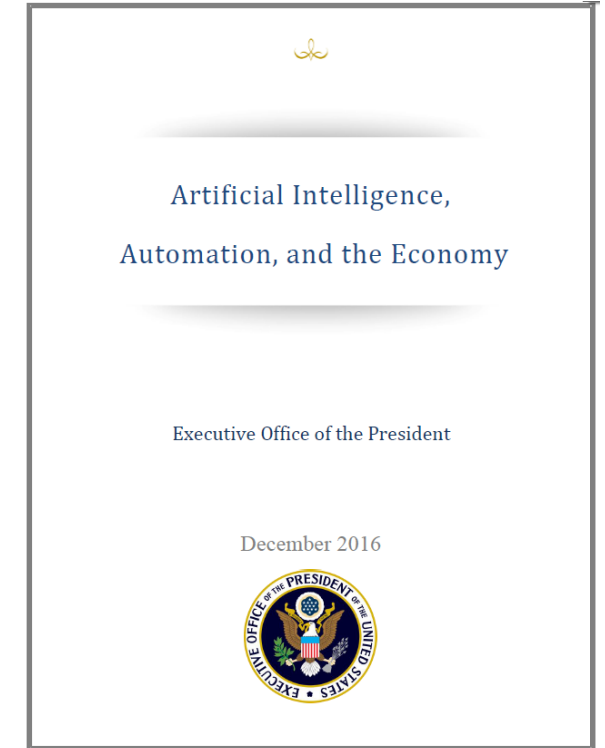
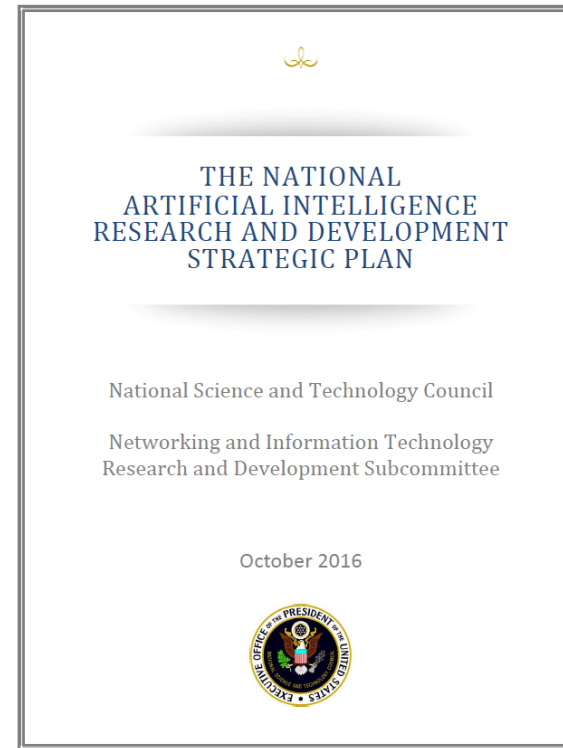
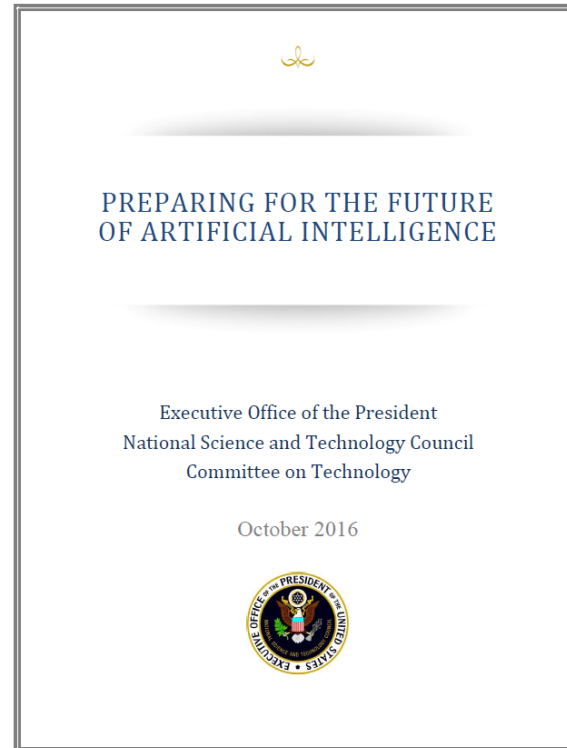
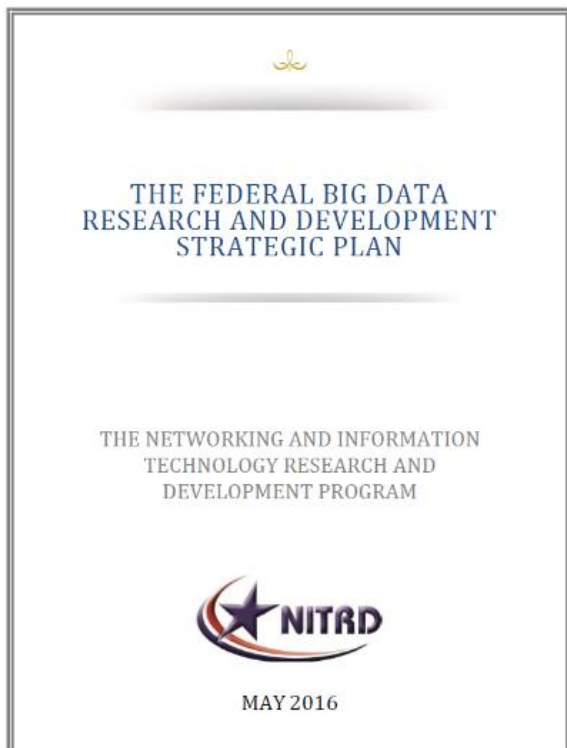
«Четвёртая технологическая революция строится на вездесущем и мобильном Интернете, *искусственном интеллекте* и *машинном обучении*» (2016)

Клаус Мартин Шваб,
президент Всемирного
экономического форума



Отчёты Белого дома США, май-октябрь 2016

«Nations with the strongest presence in AI R&D will establish leading positions in the automation of the future»



Некоторые из 23 рекомендаций

- #1. Организации должны активно развивать партнёрство с научными коллективами для эффективного использования данных.
- #2. В приоритетном порядке развивать стандарты *открытых данных* для привлечения научного сообщества к решению задач.
- #8. Инвестировать в разработку систем автоматического управления воздушным трафиком.
- #11. Вести постоянный мониторинг развития ИИ в других странах.
- #13. Приоритетно поддерживать фундаментальные и долгосрочные исследования в области искусственного интеллекта.
- #14. Развивать образовательные программы по ИИ и курсы повышения квалификации для прикладных специалистов.
- #20. Развивать международную кооперацию по ИИ.
- #22. Учитывать взаимовлияние ИИ и кибербезопасности.

Национальная стратегия развития ИИ в РФ

10 октября 2019 г.



УКАЗ

ПРЕЗИДЕНТА РОССИЙСКОЙ ФЕДЕРАЦИИ

О развитии искусственного интеллекта в Российской Федерации

В целях обеспечения ускоренного развития искусственного интеллекта в Российской Федерации, проведения научных исследований в области искусственного интеллекта, повышения доступности информации и вычислительных ресурсов для пользователей, совершенствования системы подготовки кадров в этой области **п о с т а н о в л ю**:

1. Утвердить прилагаемую Национальную стратегию развития искусственного интеллекта на период до 2030 года.

2. Правительству Российской Федерации:

а) до 15 декабря 2019 г. обеспечить внесение изменений в национальную программу "Цифровая экономика Российской Федерации", в том числе разработать и утвердить федеральный проект "Искусственный интеллект";

б) представлять Президенту Российской Федерации ежегодно доклад о ходе реализации Национальной стратегии развития искусственного интеллекта на период до 2030 года;

в) предусматривать при формировании в 2020 - 2030 годах проектов федеральных бюджетов на очередной финансовый год и на плановый период бюджетные ассигнования на реализацию настоящего Указа.

УТВЕРЖДЕНА
Указом Президента
Российской Федерации
от 10 октября 2019 г. № 490

НАЦИОНАЛЬНАЯ СТРАТЕГИЯ развития искусственного интеллекта на период до 2030 года

1. Общие положения

1. Настоящей Стратегией определяются цели и основные задачи развития искусственного интеллекта в Российской Федерации, а также меры, направленные на его использование в целях обеспечения национальных интересов и реализации стратегических национальных приоритетов, в том числе в области научно-технологического развития.

2. Правовую основу настоящей Стратегии составляют Конституция Российской Федерации, Федеральный закон от 28 июня 2014 г. № 172-ФЗ "О стратегическом планировании в Российской Федерации", указы Президента Российской Федерации от 7 мая 2018 г. № 204 "О национальных целях и стратегических задачах развития Российской Федерации на период до 2024 года", от 9 мая 2017 г. № 203 "О Стратегии развития информационного общества в Российской Федерации на 2017 - 2030 годы", от 1 декабря 2016 г. № 642 "О Стратегии научно-технологического развития Российской Федерации" и иные нормативные правовые акты Российской Федерации, определяющие направления применения информационных технологий в Российской Федерации.

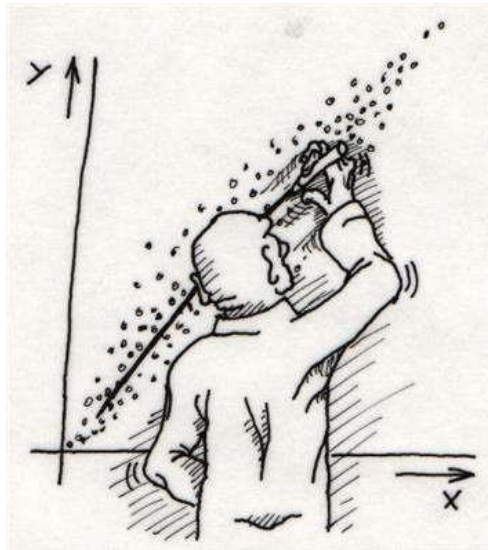
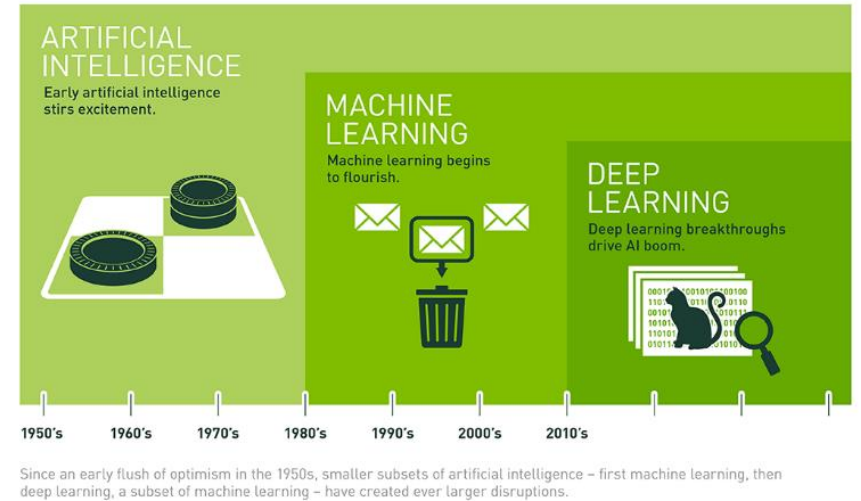
3. Настоящая Стратегия является основой для разработки (корректировки) государственных программ Российской Федерации, государственных программ субъектов Российской Федерации, федеральных и региональных проектов, плановых и программно-целевых документов государственных корпораций, государственных компаний, акционерных обществ с государственным участием,



2 100011 98860 6

Машинное обучение (Machine Learning, ML)

- одна из ключевых информационных технологий будущего
- наиболее успешное направление ИИ, вытеснившее экспертные системы и инженерию знаний



- проведение функции через заданные точки в сложно устроенных пространствах
- математическое моделирование в условиях, когда знаний мало, данных много
- тысячи различных методов и алгоритмов
- около 100 000 научных публикаций в год

Задача машинного обучения с учителем

Этап №1 – обучение (train)

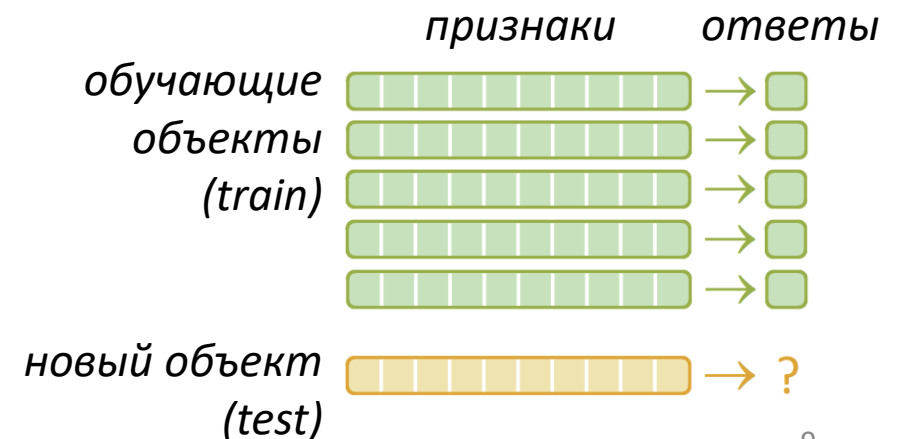
- **На входе:**
данные – выборка пар «*объект* → *ответ*»,
каждый объект описывается набором *признаков*
- **На выходе:**
модель, предсказывающая ответ по объекту

Задача поставлена,
если у неё есть «**ДНК**»:

- **Дано**
- **Найти**
- **Критерий**

Этап №2 – применение (test)

- **На входе:**
данные – **новый объект**
- **На выходе:**
предсказание ответа на новом объекте



Шаги практического решения задач ML

Формализация постановки задачи

- Описание данных: **дано** (объекты, признаки), **найти** (ответы)
- **Критерии** для измерения качества ответов

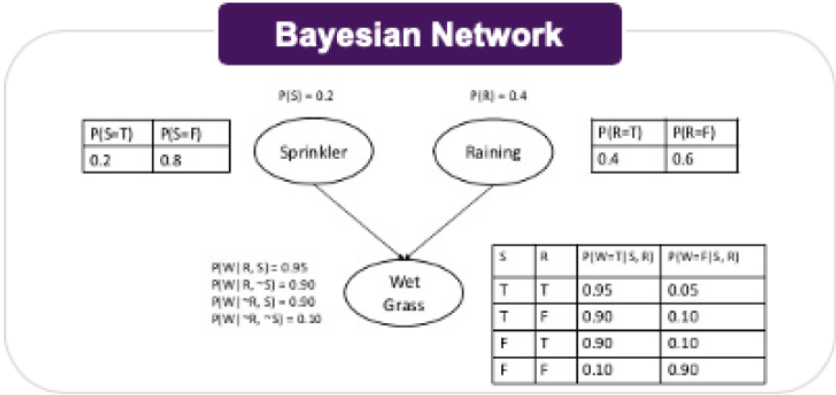
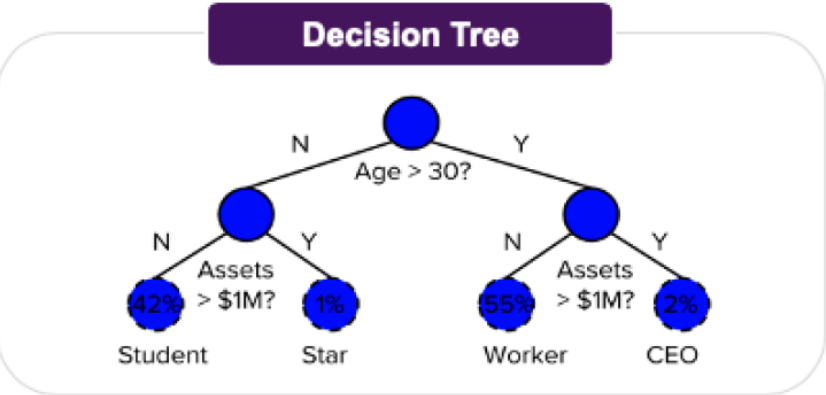
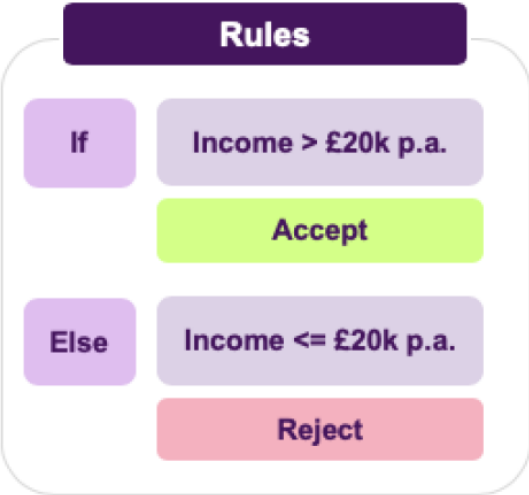
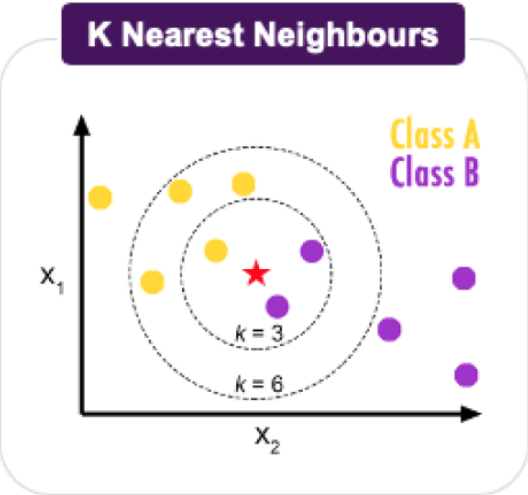
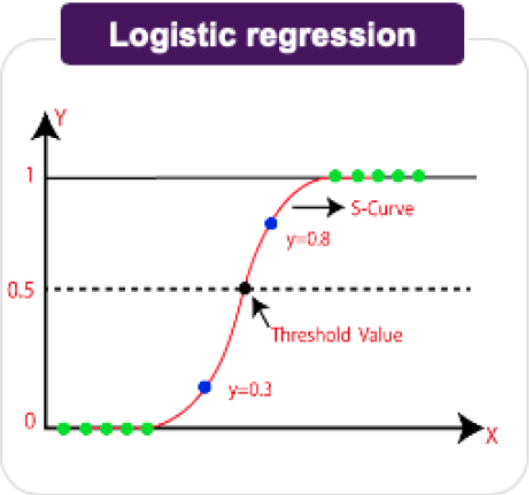
Моделирование

- Методы предобработки или векторизации данных
- Предсказательные модели или их комбинации
- Критерий оптимизации
- Численные методы и инструментальные средства

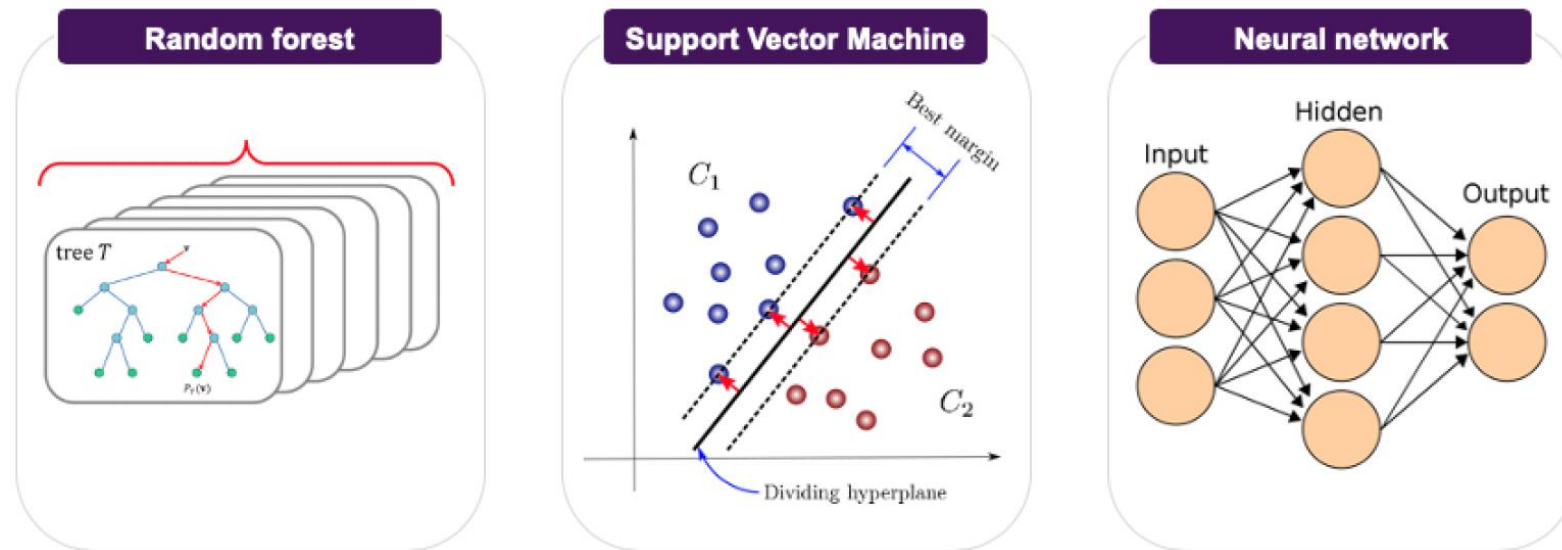
Внедрение

- Скорость обучения и применения
- Интеграция с информационными системами и бизнес-процессами

Интерпретируемые модели ML



Неинтерпретируемые модели ML



- + качество, как правило, лучше, чем у интерпретируемых моделей
- невозможно понять, правильно ли работает этот «чёрный ящик»
- невозможно объяснить предсказания на конкретных объектах

Фундаментальные основы искусственного интеллекта

1. Задачи машинного обучения

- Бум искусственного интеллекта и нейронных сетей
- Терминология машинного обучения
- Модели машинного обучения

2. Методология машинного обучения

- Нейронные сети и глубокое обучение
- Оптимизационные задачи машинного обучения
- Задачи машинного обучения с векторизацией объектов

3. Вопросы практического применения

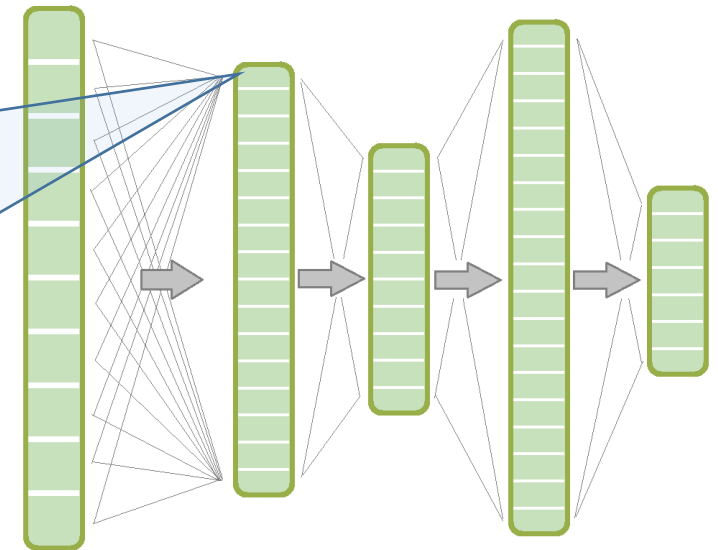
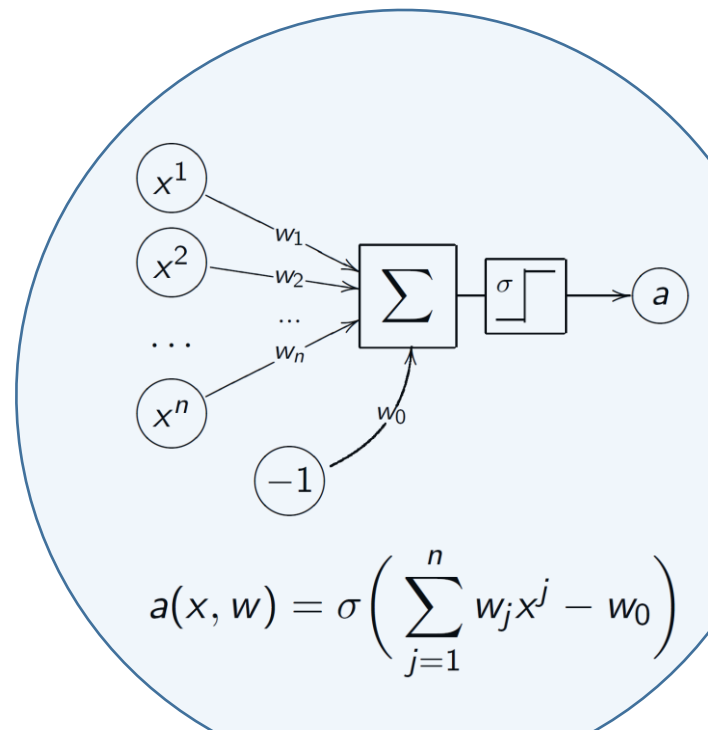
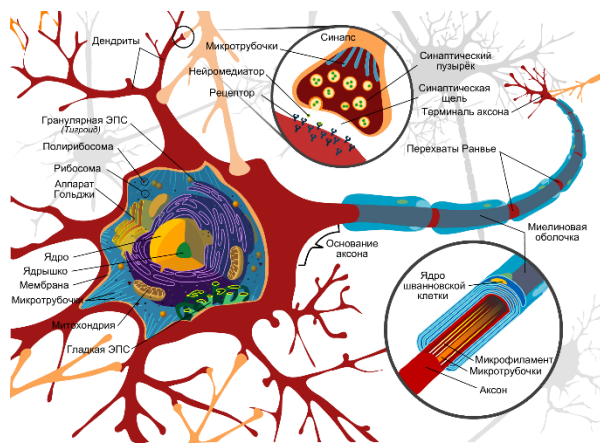
- Особенности реальных данных
- Необходимые условия применения ИИ
- Распределение ролей в проектах ИИ

Модели искусственных нейронных сетей

На каждом слое сети вектор объекта преобразуется в новый вектор

Каждое преобразование (нейрон) – взвешенная сумма признаков

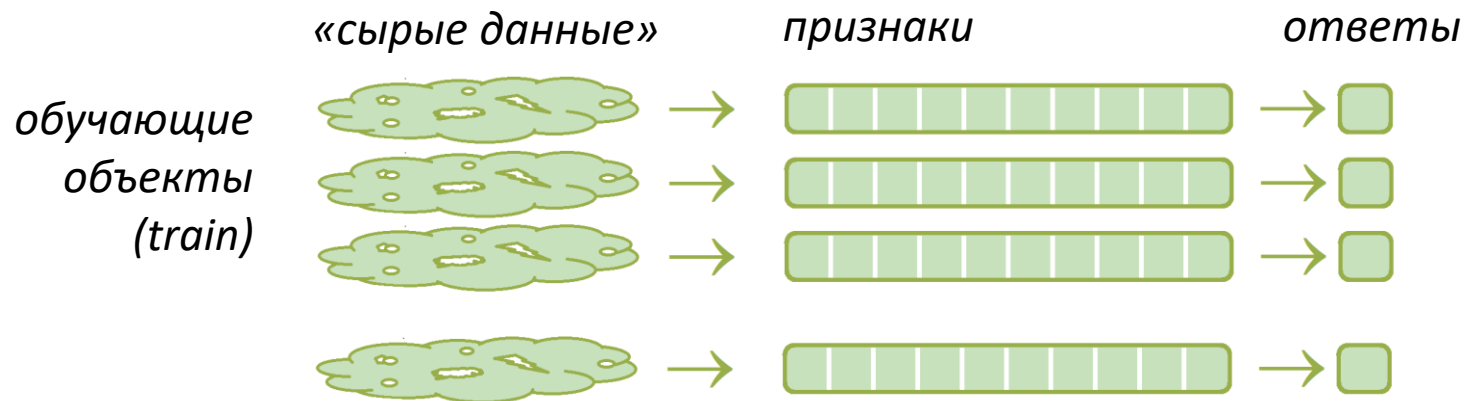
Веса w являются обучаемыми параметрами модели



Глубокие нейронные сети (Deep Learning, DL)

Вход: сложно структурированные «сырые» данные объектов

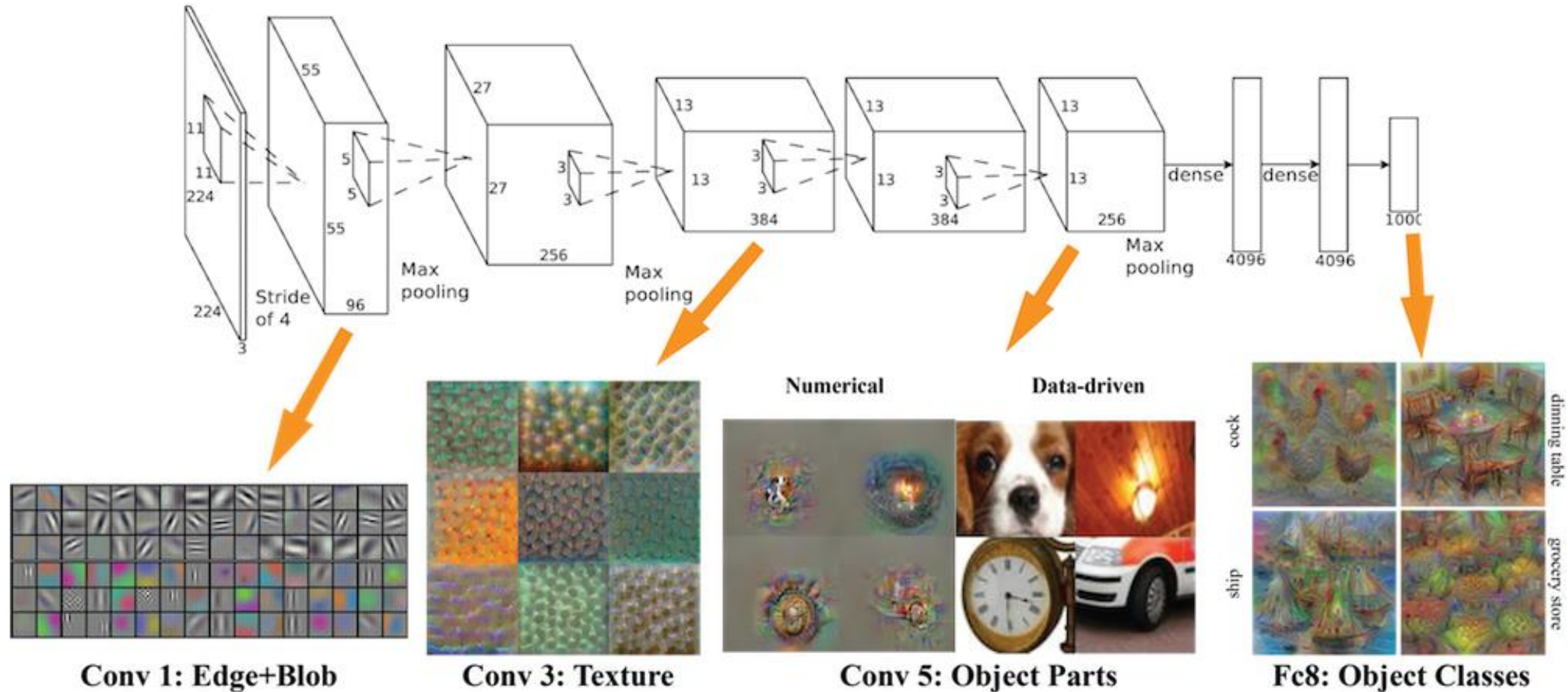
Выход: векторные представления объектов и ответы



*Deep Learning – это
всего лишь обучаемая
векторизация
сложных объектов*

Примеры сложно структурированных объектов:
тексты, изображения, видео, временные ряды, транзакции, графы, ...

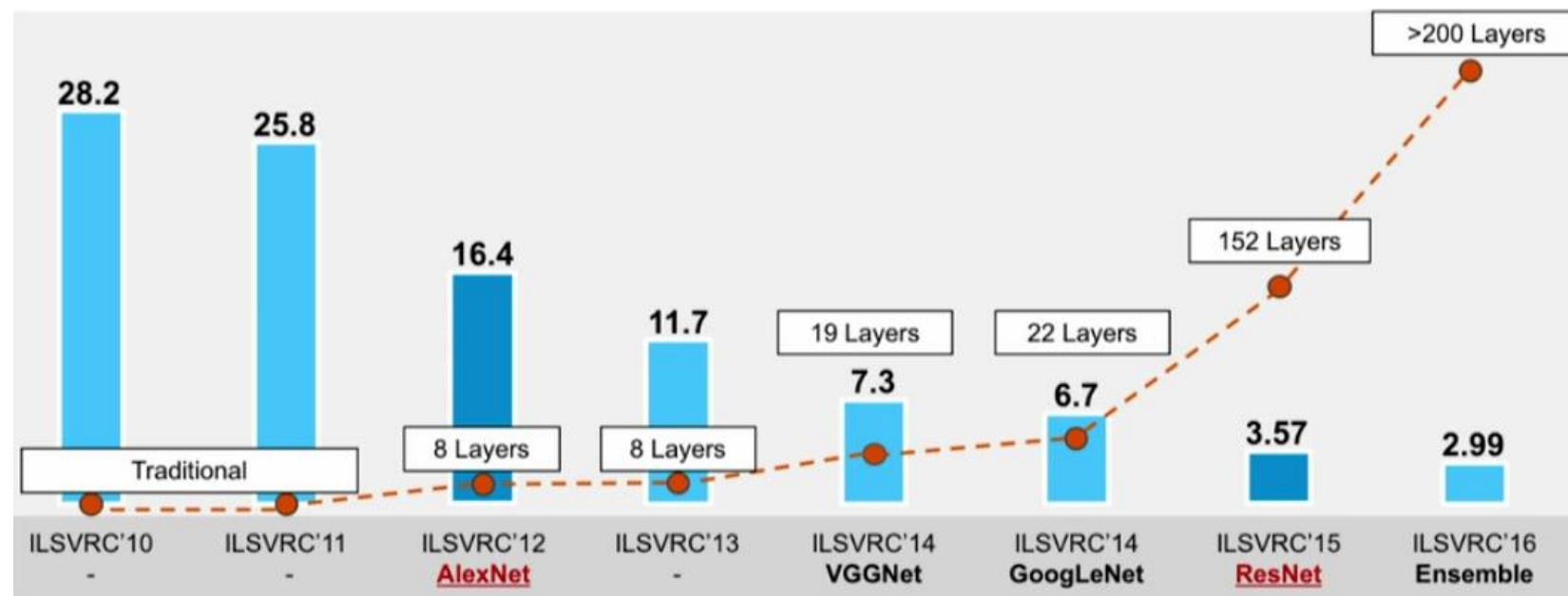
Глубокие свёрточные нейронные сети для классификации изображений



Роль больших данных

ImageNet: открытая выборка 14М изображений, 20К категорий

IMAGENET

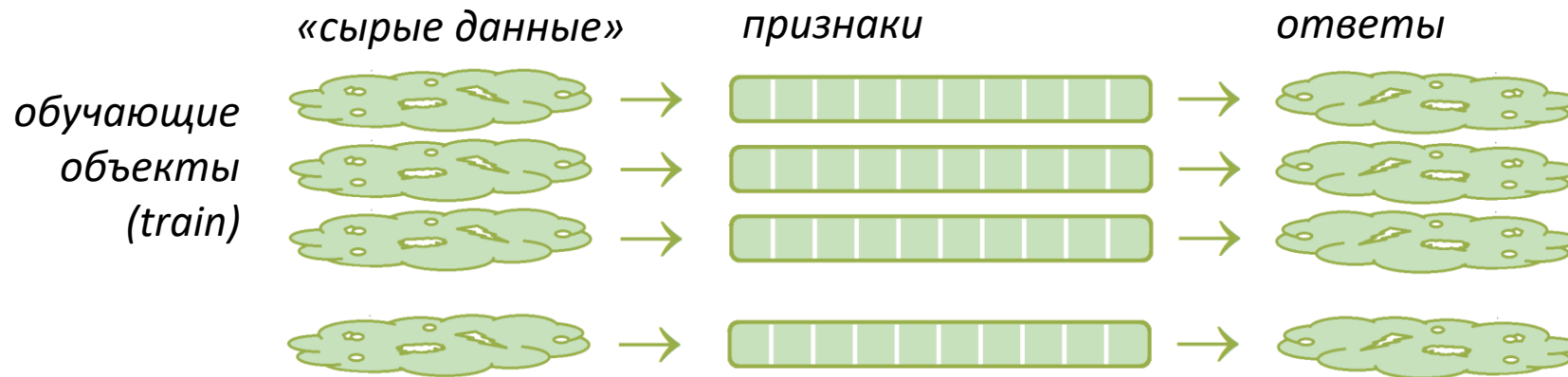


Старт в 2009 г. Человеческий уровень ошибок 5% пройден в 2015 г.

Нейронные сети для синтеза объектов

Вход: сложно структурированные объекты

Выход: сложно структурированные ответы



Примеры: синтез изображений, перенос стиля, машинный перевод, суммаризация текстов, распознавание и синтез речи

Модели: seq2seq, CNN, RNN, LSTM, GAN, BERT, GPT-3 и др.

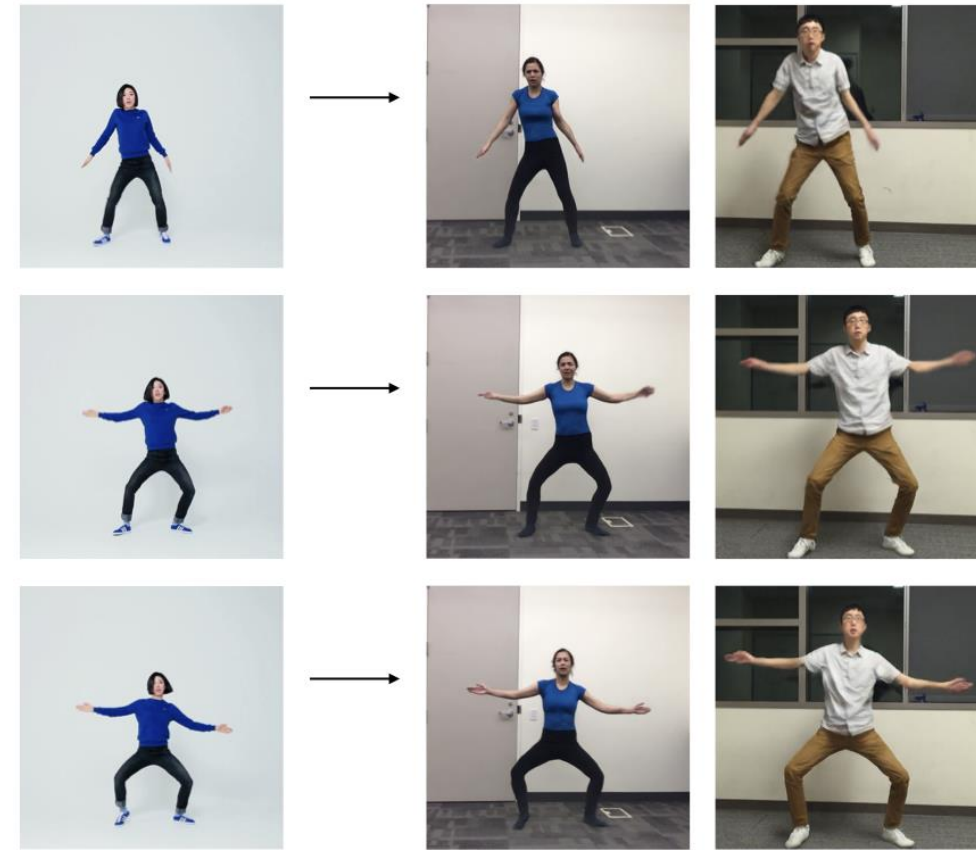
Синтез изображений и видео



(d) input image

(e) output 3d face

(f) textured 3d face



Source Subject

Target Subject 1

Target Subject 2

Машинное обучение – это оптимизация

x – вектор объекта обучающей выборки

w – параметры модели

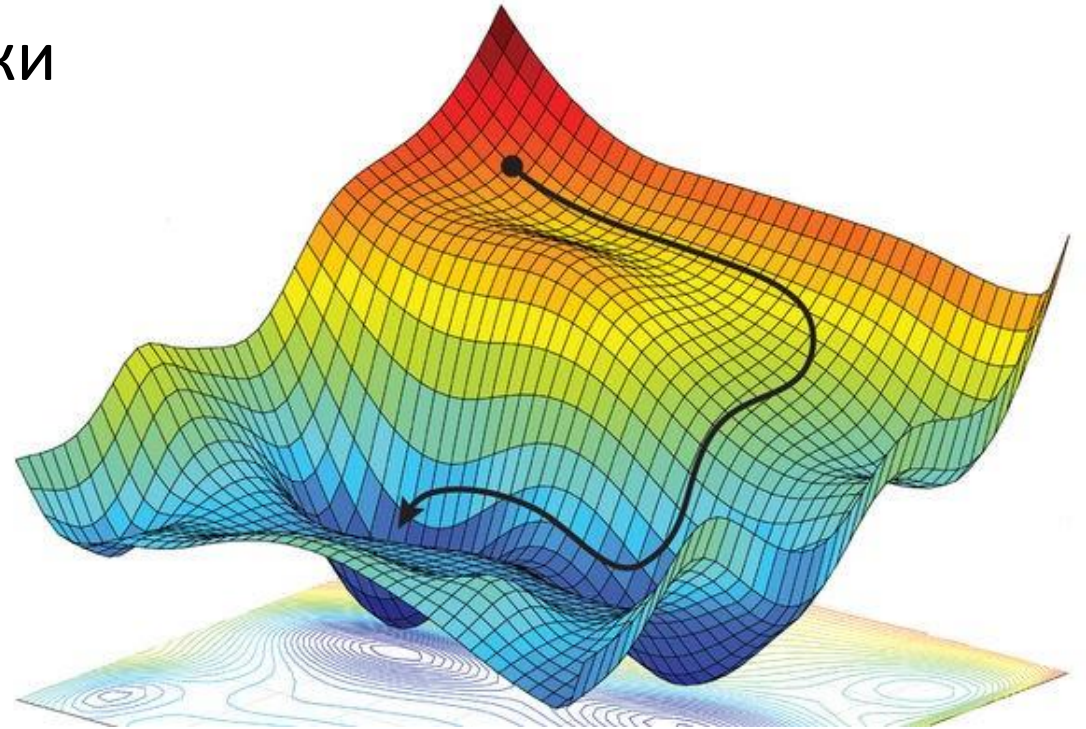
$\text{Loss}(x, w)$ – функция потерь

$Q(w)$ – критерий качества модели

Задача на этапе обучения модели:

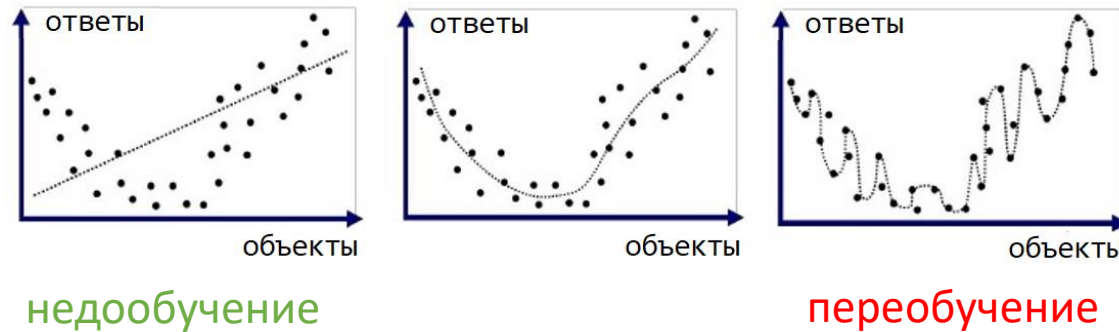
$$Q(w) = \sum_x \text{Loss}(x, w) \rightarrow \min$$

Способ решения – численные методы оптимизации

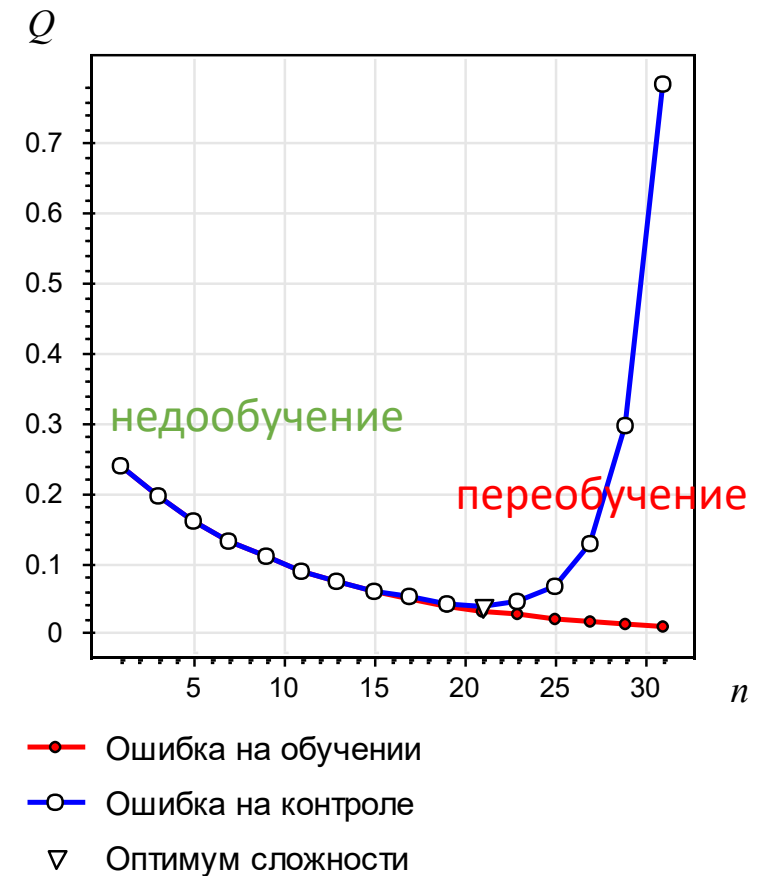


Переобучение – основная трудность ML

Причина переобучения – избыточная сложность модели



- **Внутренние критерии:**
для оптимизации параметров модели
- **Внешние критерии:**
для оценивания обобщающей способности модели и контроля *переобучения*



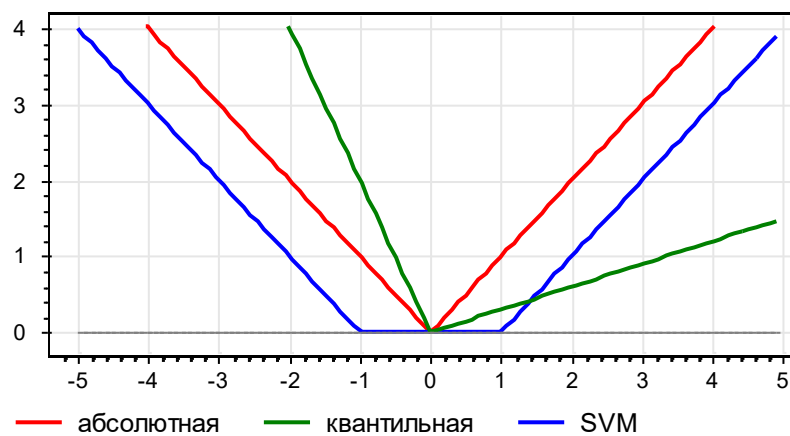
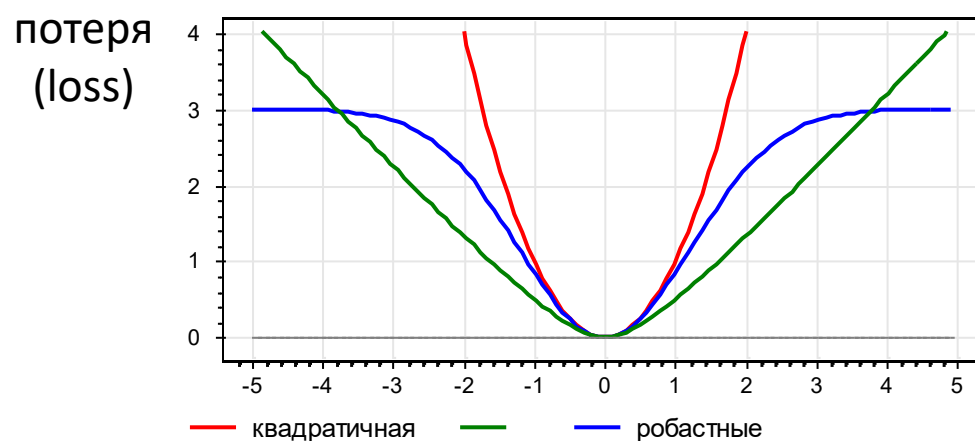
Восстановление регрессии (regression)

x – вектор объекта обучающей выборки, y – числовой ответ

$a(x, w)$ – модель регрессии с параметрами w

Например, $a(x, w) = \sum_j w_j x_j$ – линейная модель регрессии

$\text{Loss}(x, w) = (a(x, w) - y)^2$ – квадратичная функция потерь



НЕВЯЗКА
(error)

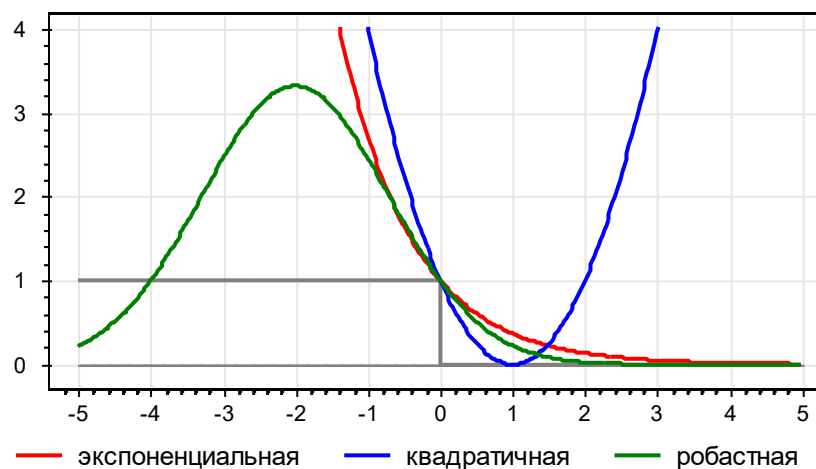
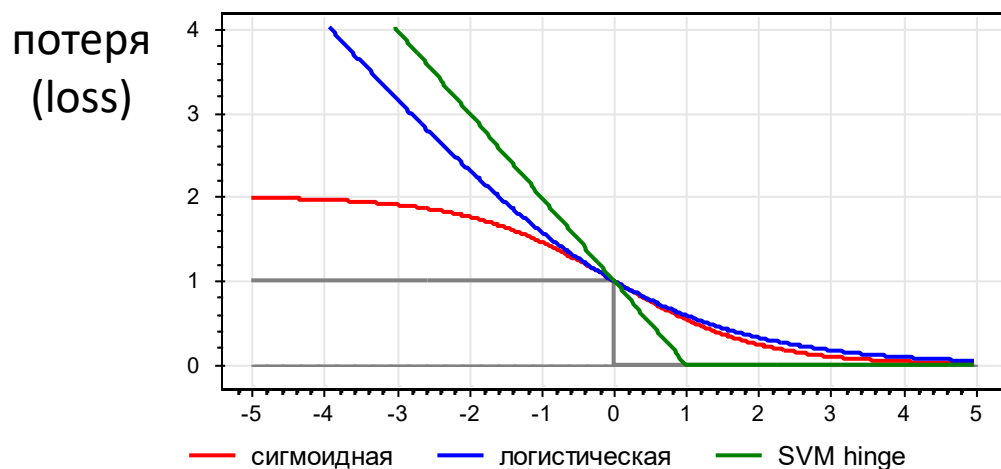
Классификация (classification)

x – вектор объекта обучающей выборки, y – ответ (+1 или -1)

$a(x, w)$ – модель классификации с параметрами w

Например, $a(x, w) = \text{sign}(\sum_j w_j x_j)$ – линейная модель

$\text{Loss}(x, w) = \max(0, 1 - y \sum_j w_j x_j)$ – функция потерь hinge



отступ
(margin)

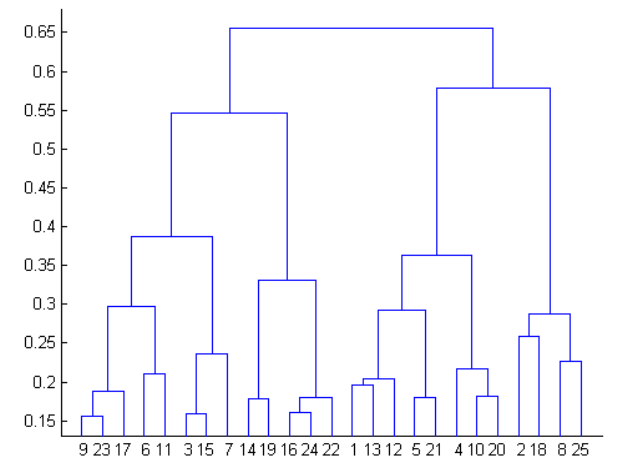
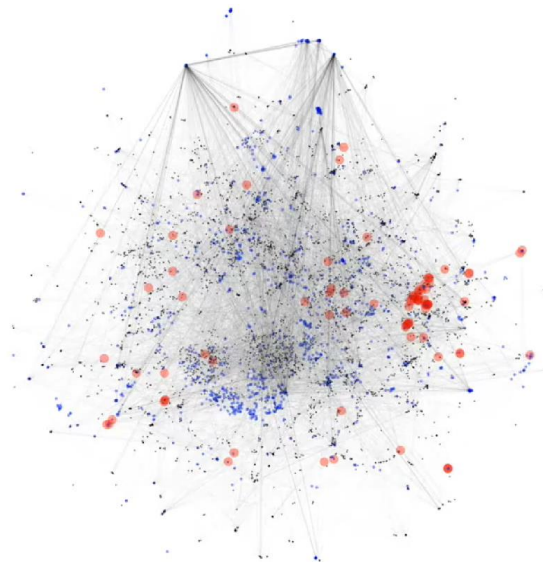
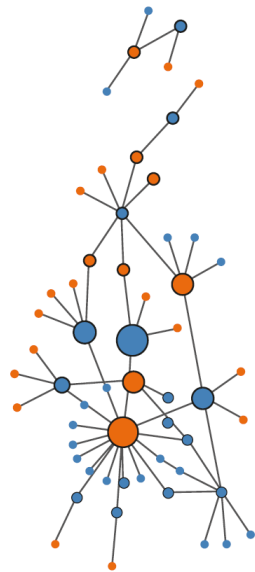
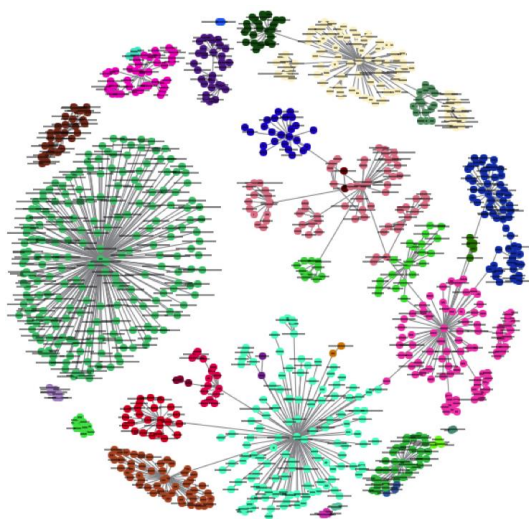
Кластеризация (clustering)

x – вектор объекта обучающей выборки, ответов не дано

$a(x, w)$ – ближайший к x центр кластера

$w = \{c_1, \dots, c_K\}$ – векторы центров всех кластеров

$\text{Loss}(x, w) = \min_k \|x - c_k\|$ – расстояние до ближайшего кластера



Ранжирование (learning to rank)

x – вектор пары «запрос-документ», y – оценка релевантности

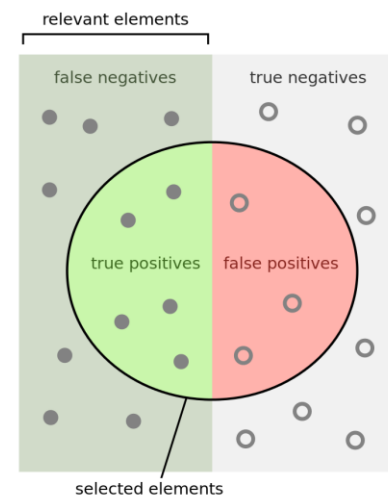
$a(x, w)$ – модель ранжирования документов по запросу, параметр w

Например, $a(x, w) = \sum_j w_j x_j$ – линейная модель

$\text{Loss}(x, x', w) = \max\left(0, 1 - [y > y'](a(x, w) - a(x', w))\right)$ – ф.потерь

историческая информатика

- Информатика историческая** litres. Без подписок
litres.ru > Историческая-информа...
Более 1 000 000 книг в форматах FB2, EPUB, TXT, PDF, Аудиокниги. Выберите и читайте! · Без подписок. Книга ваша навсегда. Все аудиокниги. Без скрытых платежей ·
- Историческая информатика** — Википедия
ru.wikipedia.org > Историческая информатика
Историческая информатика — междисциплинарная область исторических исследований, целью которой является расширение информационного...
- Журнал "Историческая информатика"**
kleio.asu.ru
Историческая информатика. Информационные технологии и математические методы в исторических исследованиях и образовании. Читать ещё >
- Методологические проблемы исторической информатики
nbpublish.com > e_jstinf/
Ключевые слова: виртуальные исторические реконструкции, историческая информатика, источниковедение, методология, исторические источники, классификация, научно-техническая документация, электронные... Читать ещё >
- Историческая информатика.**
ost-talent.org > 40526-istoricheskaya-informatika-...



$$\text{Precision} = \frac{\text{true positives}}{\text{true positives} + \text{false positives}}$$

$$\text{Recall} = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}}$$

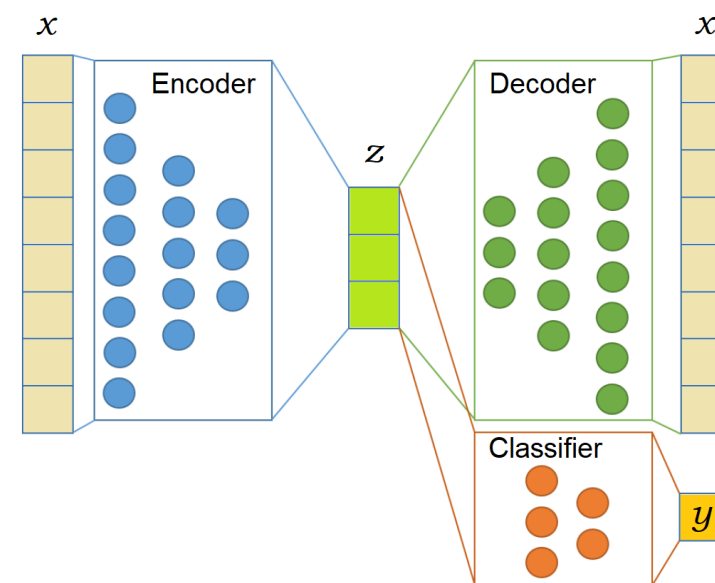
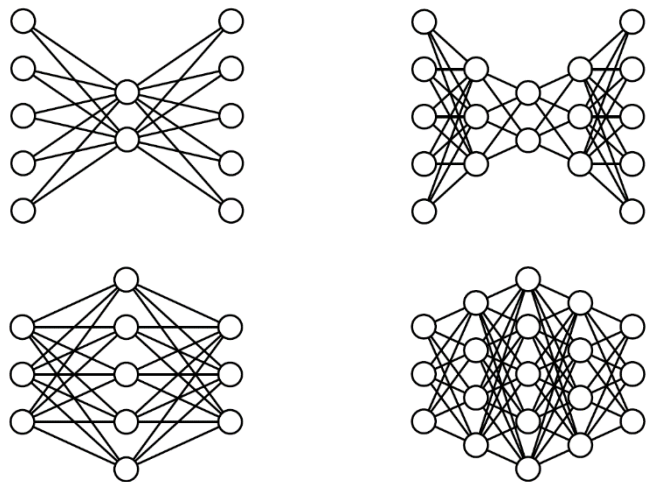
Векторизация объектов (autoencoders)

x – вектор объекта обучающей выборки, ответов не дано

$z = f(x, w)$ – модель кодирования x в векторное представление z

$x' = g(z, w)$ – модель декодирования z в реконструкцию x'

$\text{Loss}(x, w) = \|x'(w) - x\|$ – точность реконструкции объекта



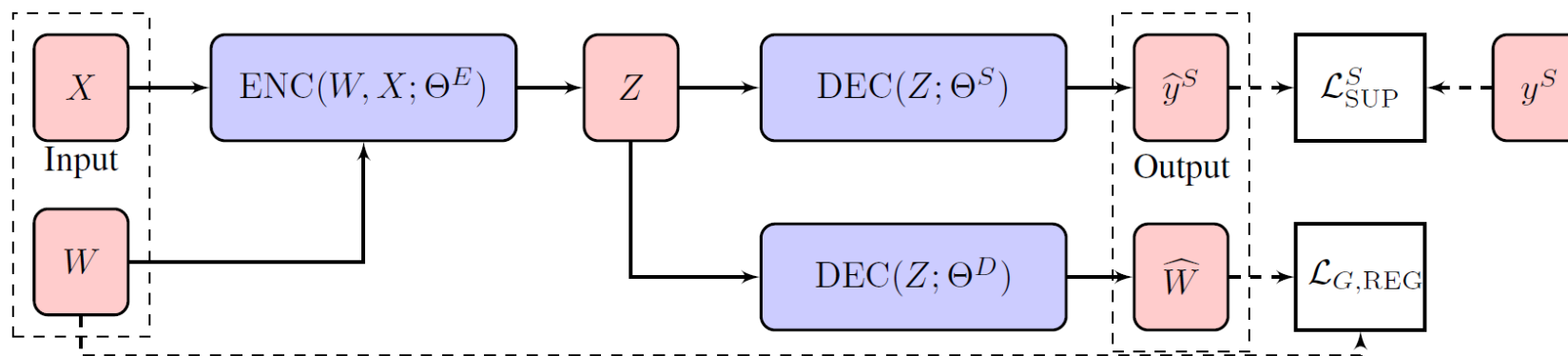
Векторизация графов (graph embeddings)

$x; (x, x')$ – данные об объектах и взаимодействиях между объектами

$z = f(x, w)$ – модель кодирования x в векторное представление z

$x' = g(z, w)$ – модель декодирования z в реконструкцию x'

$\text{Loss}(x, w) = \|x'(w) - x\| + \tau L_{\text{SUP}}(x, w_S)$ – сумма двух критериев



T.Mikolov et al. Efficient estimation of word representations in vector space, 2013.

I.Chami et al. Machine learning on graphs: a model and comprehensive taxonomy. 2020.

Перенос обучения (transfer learning)

$f(x, w)$ – часть модели, универсальная для широкого класса задач

$g(x, w')$ – часть модели, специфичная для своей задачи

$\min_{w, w'} \sum_x \text{Loss}_1(f(x, w), g_1(x, w'))$ – обучение по большим данным

$\min_{w'} \sum_{x'} \text{Loss}_2(f(x', w), g_2(x', w'))$ – обучение по своим данным

Sinno Jialin Pan, Qiang Yang.
A Survey on Transfer
Learning. 2009



Многозадачное обучение (multi-task learning)

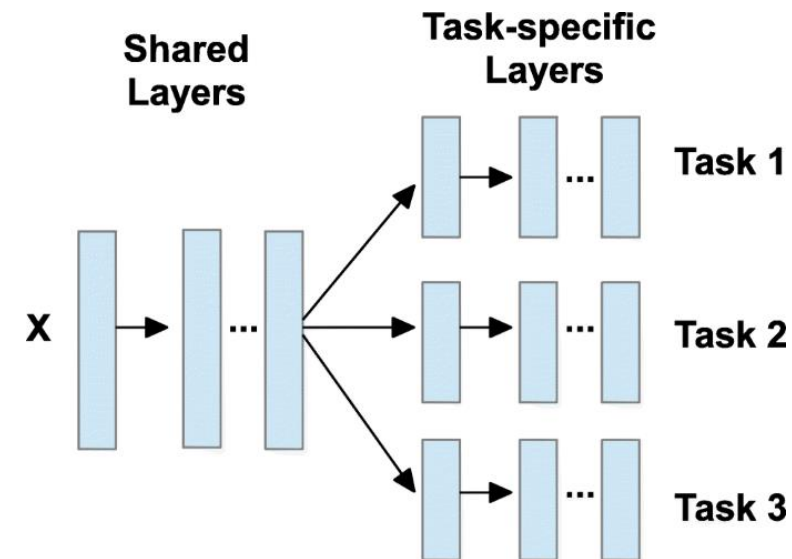
$f(x, w)$ – часть модели, универсальная для всех задач

$g_t(x, w'_t)$ – часть модели, специфичная для t -й задачи

$\min_{w, w'_t} \sum_t \sum_x \text{Loss}_t(f(x, w), g_t(x, w'_t))$ – обучение по всем задачам

M.Crawshaw. Multi-task learning with deep neural networks: a survey. 2020

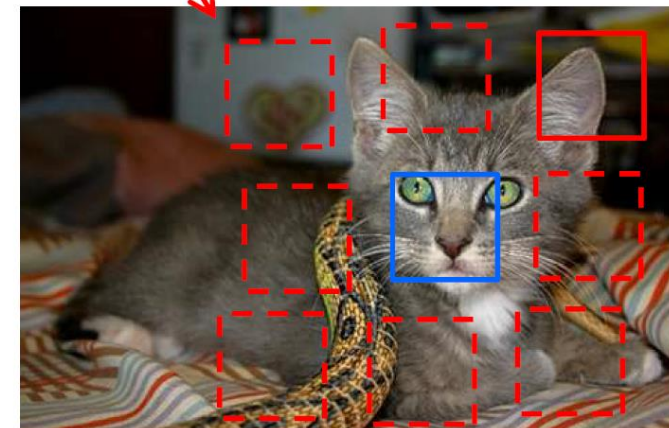
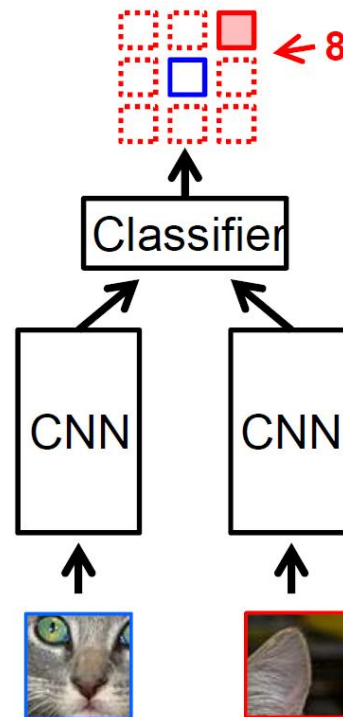
Y.Wang et al. Generalizing from a few examples: a survey on few-shot learning. 2020



Самостоятельное обучение (self-supervised)

Модель векторизации $z = f(x, w)$ обучается предсказывать взаимное расположение пар фрагментов одного изображения

Преимущество: сеть выучивает векторные представления объектов без размеченной обучающей выборки



Randomly Sample Patch
Sample Second Patch

Unsupervised visual representation learning by context prediction,
Carl Doersch, Abhinav Gupta, Alexei A. Efros, ICCV 2015

Обучение с привилегированной информацией

$g(x, x^*, w')$ – модель-учитель, имеет доступ к x^*

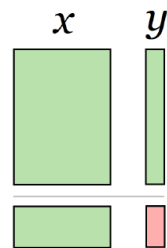
$f(x, w)$ – модель-ученик, учится повторять ошибки учителя

$$\min_{w, w'} \sum_x \text{Loss}(f(x, w), y) + \text{Loss}(g(x, x^*, w'), y) + \text{Loss}(f(x, w), g(x, x^*, w'))$$

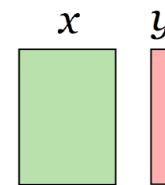
V.Vapnik, A.Vashist. A new learning paradigm: Learning Using Privileged Information. 2009

D.Lopez-Paz, L.Bottou, B.Scholkopf, V.Vapnik. Unifying distillation and privileged information. 2016

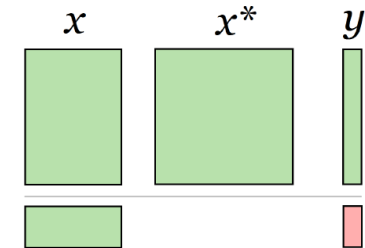
с учителем



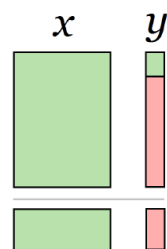
без учителя



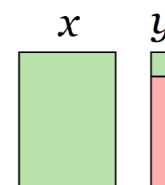
привилегированное (LUPI)



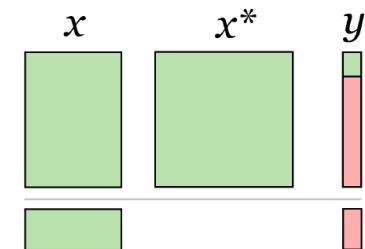
частичное



трандуктивное



частичное LUPI



Генеративная состязательная сеть (GAN)

$x = g(z, w)$ – модель генерации реалистичного объекта x из шума z

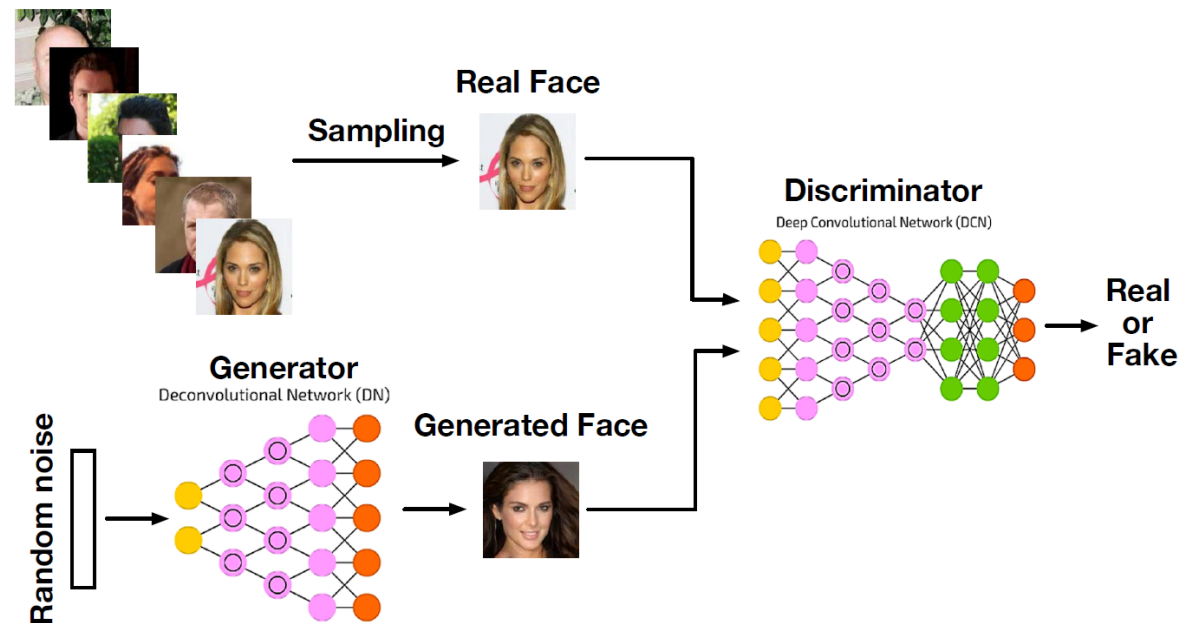
$f(x, w')$ – модель классификации x «реальный/сгенерированный»

$\min_w \max_{w'} \sum_x \ln f(x, w') + \ln (1 - f(g(z, w), w'))$ – совместное обучение

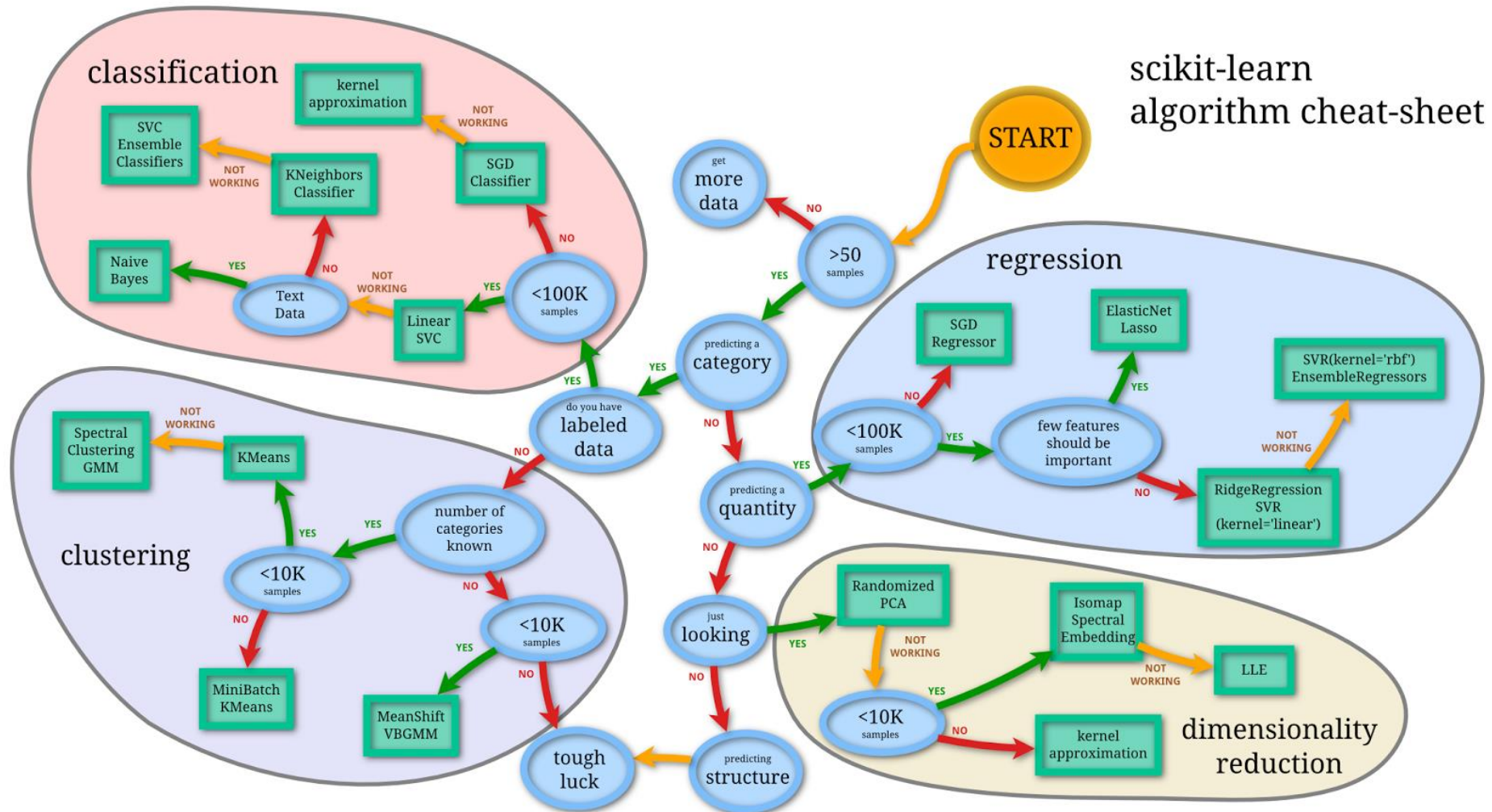
Antonia Creswell et al. Generative Adversarial Networks: an overview. 2017.

Zhengwei Wang et al. Generative Adversarial Networks: a survey and taxonomy. 2019.

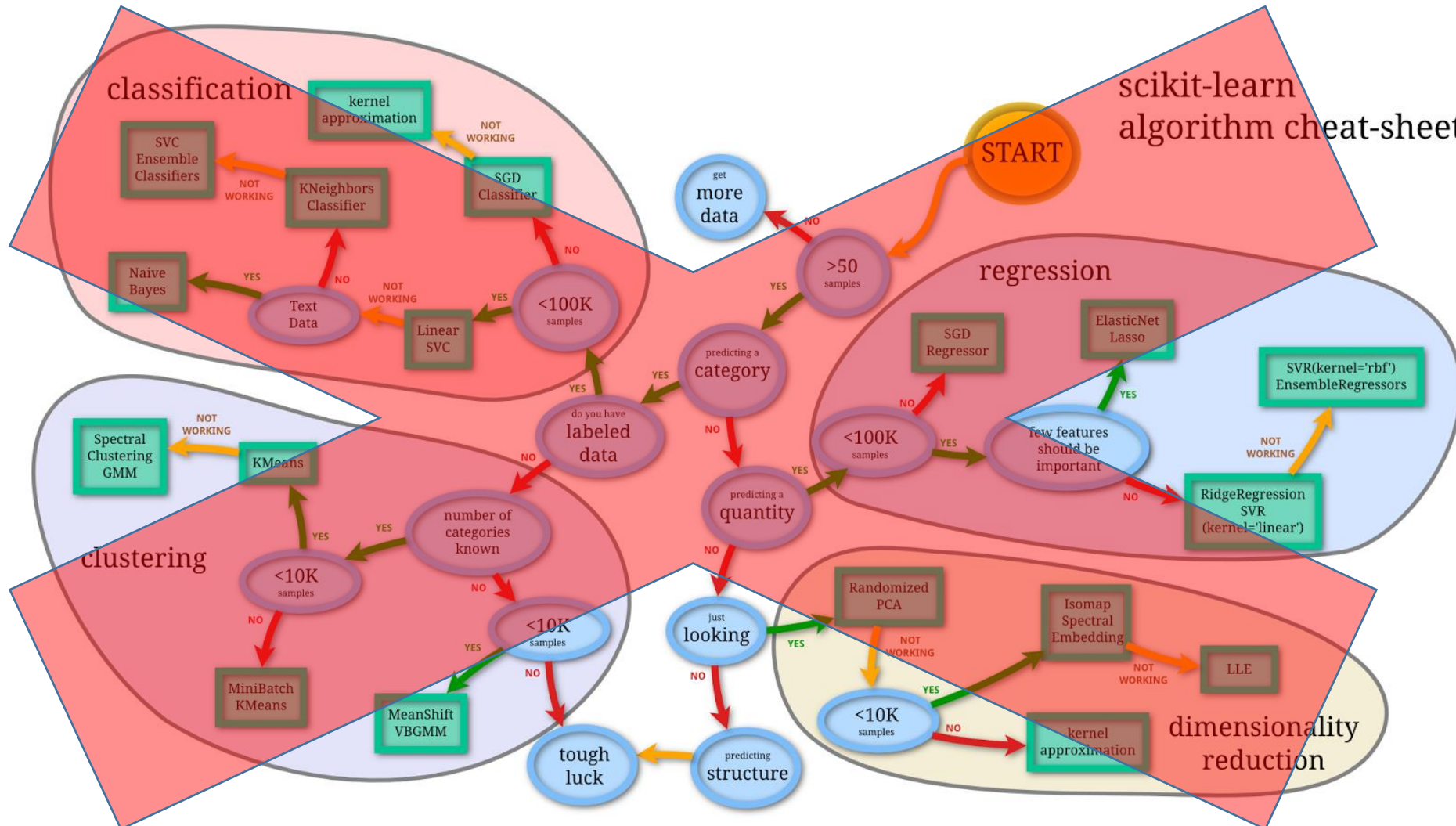
Chris Nicholson. A Beginner's Guide to Generative Adversarial Networks. 2019.



Задачи и методы машинного обучения



Избегайте пользоваться чужими шпаргалками!



Фундаментальные основы искусственного интеллекта

1. Задачи машинного обучения

- Бум искусственного интеллекта и нейронных сетей
- Терминология машинного обучения
- Модели машинного обучения

2. Методология машинного обучения

- Нейронные сети и глубокое обучение
- Оптимизационные задачи машинного обучения
- Задачи машинного обучения с векторизацией объектов

3. Вопросы практического применения

- Особенности реальных данных
- Необходимые условия применения ИИ
- Распределение ролей в проектах ИИ

Особенности реальных данных

В реальных приложениях данные бывают ...

- разнородные (признаки измерены в разных шкалах)
- неполные (признаки измерены не все, имеются пропуски)
- неточные (признаки измерены с погрешностями)
- противоречивые (объекты одинаковые, ответы разные)
- избыточные (сверхбольшие, не помещаются в память)
- недостаточные (объектов меньше, чем признаков)
- неструктурированные (нет признаков описаний)
- «грязные» (ошибочные, грубо не соответствующие истине)

*со всем этим
можно
работать*



*но только не
с грязными
данными!*

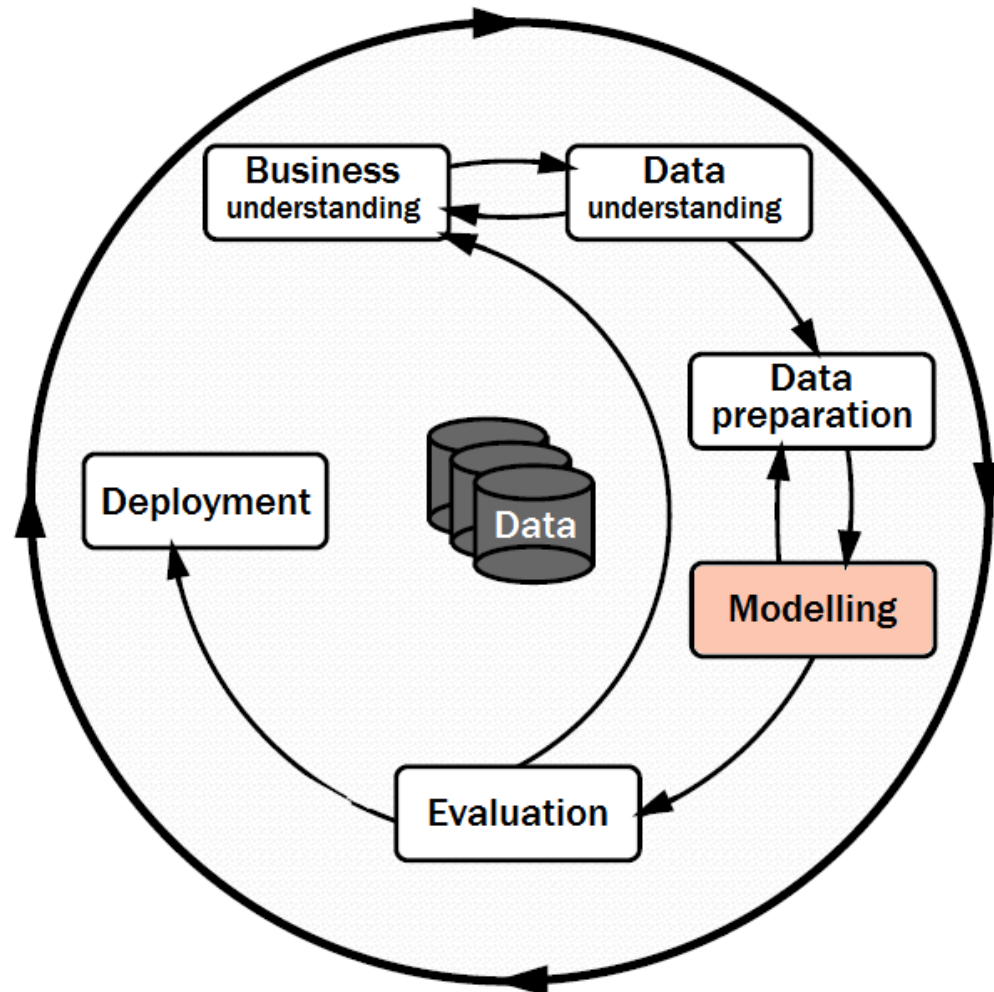


Необходимые условия применения ИИ

- **Полнота, чистота, достоверность данных**
 - Автоматизация и цифровизация процессов, порождающих данные
 - Контроль качества данных (цифровой двойник или «цифровое чучело»?)
 - Трудовая и технологическая дисциплина при работе с данными
- **Культура постановки задач**
 - Предметная экспертиза вместо «абстрактной веры во всемогущий ИИ»
 - Понимание целей анализа и их формализация в критериях качества
 - Готовность пилотировать новые технологии (кто на деле «data-driven»?)
- **Культура анализа данных**
 - Владение средствами визуализации и понимания данных
 - Грамотный и тщательный анализ ошибок при выборе моделей
 - Умение находить «простые но гениальные» решения

Этапы решения задач ML/DS/AI

CRISP-DM: Cross Industry Standard Process for Data Mining (1999)



- понимание прикладной задачи
- понимание данных
- предобработка данных
- инженерия признаков
- построение моделей
- оптимизация параметров
- контроль переобучения
- (кросс-)валидация решения
- внедрение и эксплуатация

Рынок труда в области анализа данных

Инженер по данным (Data Engineer)

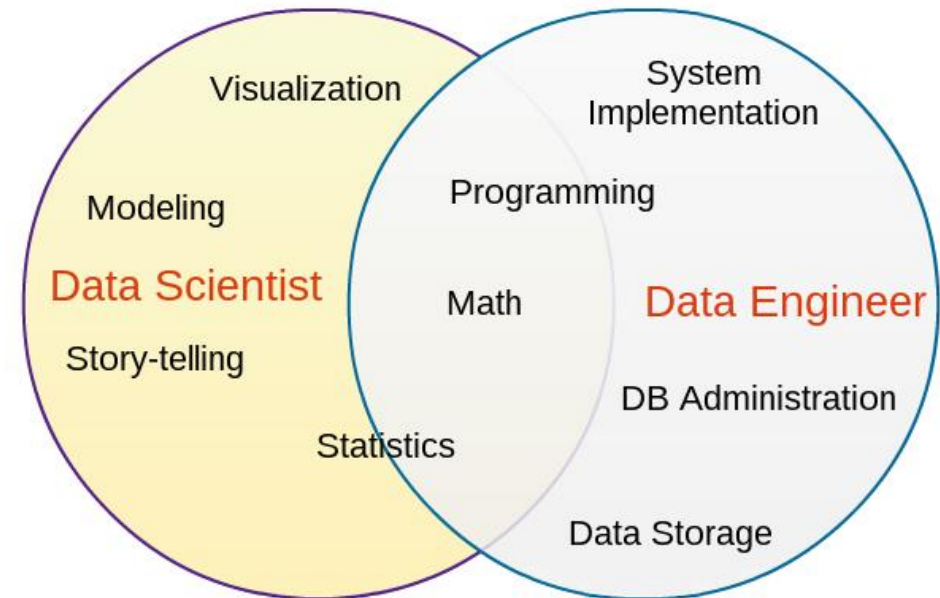
- Понимает бизнес-процессы, порождающие данные
- Работает с данными в различных форматах
- Визуализирует, понимает, очищает, готовит данные

Исследователь данных (Data Scientist)

- Моделирует, строит признаки (feature engineering)
- Выбирает модели и методы, оценивает решения
- Ходит по кругу CRISP-DM

Менеджер проектов по анализу данных

- Организует бизнес-процессы сбора и очистки данных
- Видит бизнес задачи и формализует их в терминах «Дано-Найти-Критерий»
- Организует открытые конкурсы и пилотные проекты
- Реалистично оценивает сложность задач и трудозатраты



Сухой остаток

- ИИ = Имитация Интеллекта
- ИИ = Машинное Обучение + устаревшие или нишевые методы
- Машинное Обучение = разновидность математ. моделирования
- Методы ИИ = методы аппроксимации и численной оптимизации
- Прорыв в ИИ = оптимизация в пространствах большой размерности
- Глубокие нейронные сети = автоматическая векторизация объектов
- Формализация задачи и чистота данных важнее методов
- Знания прикладной области ценнее навыков обучения нейросетей

Рекомендуемая литература

- *Андрей Бурков.* Машинное обучение без лишних слов. Питер, 2020.
- *Николенко С. И., Кадурин А. А., Архангельская Е. О.* Глубокое обучение. Питер, 2018.
- *Бенджио И., Гудфеллоу Я., Курвилль А.* Глубокое обучение. ДМК-Пресс, 2018.
- *Воронцов К. В.* Лекции по машинному обучению. www.MachineLearning.ru, 2004-2021.
- *Коэльо Л. П., Ричарт В.* Построение систем машинного обучения на языке Python. 2016.
- *Hastie T., Tibshirani R., Friedman J.* The Elements of Statistical Learning. Springer, 2014.
- *Мерков А. Б.* Распознавание образов. Введение в методы статистического обучения. 2011.
- *Мерков А. Б.* Распознавание образов. Построение и обучение вероятностных моделей. 2014.
- *Домингос П.* Верховный алгоритм. 2016.