

ИИ_на_пальцах (
ИИ == Искусственный_Интеллект или
ИИ == Имитация_Интеллекта?
)

Воронцов Константин Вячеславович

д.ф.-м.н., профессор РАН,
руководитель лаборатории Машинного интеллекта МФТИ

k.v.vorontsov@phystech.edu

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

Клаус Мартин Шваб,
президент Всемирного
экономического форума



Бум искусственного интеллекта

1997: IBM Deep Blue обыграл чемпиона мира по шахматам

2005: Беспилотный автомобиль: DARPA Grand Challenge

2006: Google Translate – статистический машинный перевод

2011: 40 лет DARPA CALO привели к созданию Apple Siri

2011: IBM Watson победил в ТВ-игре «Jeopardy!»

2011–2018: ImageNet: 25% → 2,5% ошибок против 5% у людей

2015: Фонд OpenAI в \$1 млрд. Илона Маска и Сэма Альтмана

2016: DeepMind, OpenAI: динамическое обучение играм Atari

2016: Google DeepMind обыграл чемпиона мира по игре го

2017: OpenAI обыграл чемпиона мира по компьютерной игре Dota 2



Глубокие нейронные сети обеспечили прорыв в компьютерном зрении

ImageNet: открытая выборка 15М размеченных изображений



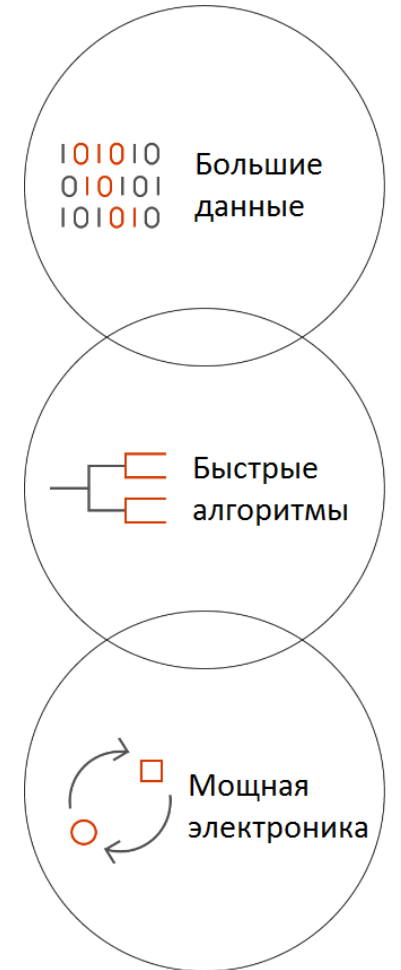
Google: Распознавание кадров с котами на видео из Youtube



Три предпосылки этого бума

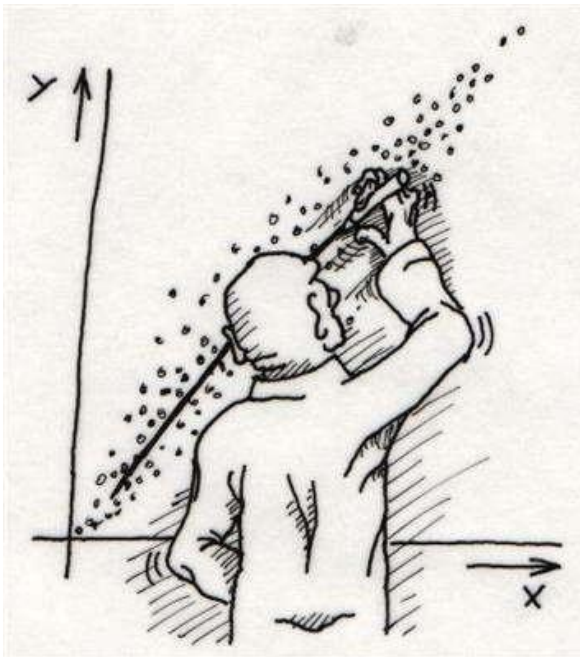
– три перехода количества в качество:

- Повсеместное применение компьютерных технологий
→ *накопление больших выборок данных
в частности, ImageNet*
- Развитие математических методов и алгоритмов
→ *накопление критической массы опыта
в частности, Deep Neural Networks*
- Достижения микроэлектроники
→ *рост вычислительных мощностей по закону Мура
в частности, GPU*



Машинное обучение (Machine Learning, ML)

- одна из ключевых информационных технологий будущего
- наиболее успешное направление *искусственного интеллекта*, 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

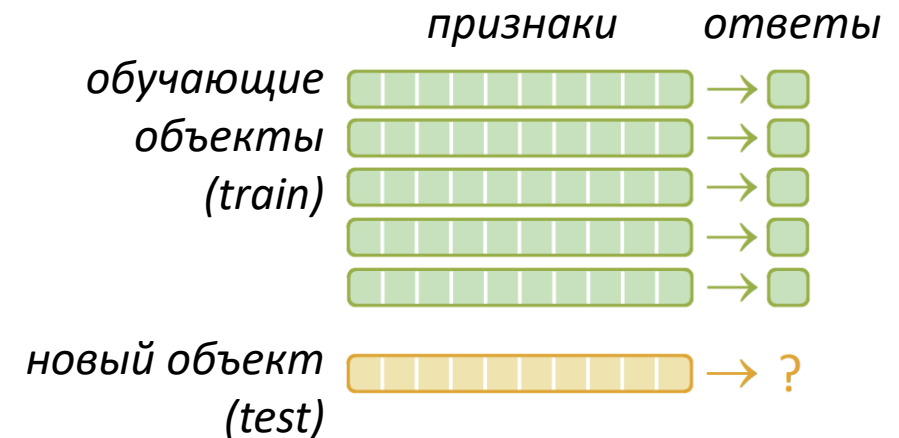
Этап №1 – обучение с учителем

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

Если нет данных,
то нет
и машинного
обучения

Этап №2 – применение

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



Примеры задач машинного обучения

- **Кредитный скоринг:**

объект – данные о заёмщике

ответ – решение по кредиту, оценка вероятности дефолта



- **Информационный поиск в Интернете:**

объект – данные о паре «запрос и документ»

ответ – оценка релевантности документа запросу



- **Рекомендательные системы в Интернете / TV:**

объект – данные о паре «пользователь, товар / фильм»

ответ – оценка вероятности покупки / просмотра

amazon.com
Prime

NETFLIX

hulu

Примеры задач ML в LegalTech

- **Поиск схожей судебной практики:**

объект – текст иска, акта или обращения заявителя

ответ – ранжированный список схожих дел



- **Рекомендательный сервис:**

объект – пара «описание дела, профиль юриста/фирмы»

ответ – ранжированный список консультантов



- **Предсказание судебного решения:**

объект – описание дела, документы по делу

ответ – вероятность выиграть дело



Примеры задач с данными сложной структуры

- **Автоматическая генерация ответов на вопросы:**

объект – текст вопроса на естественном языке

ответ – текст ответа на естественном языке

*Прогресс в этих
областях связан с
«большими данными»
(англ. «Big Data»)*

- **Перевод речи в текст:**

объект – аудиозапись речи человека

ответ – текстовая запись речи

...очень важное уточнение:

с аккуратными

большими данными

- **Компьютерное зрение:**

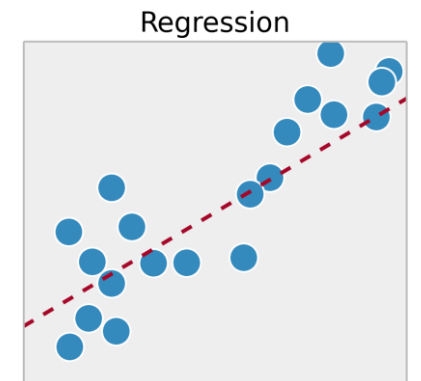
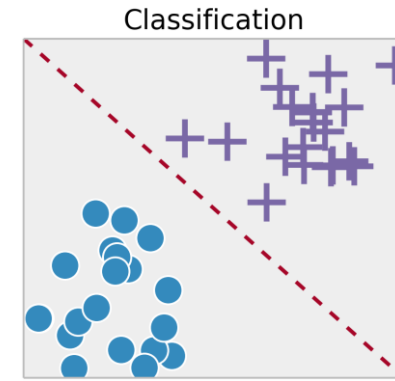
объект – фото или скан документа

ответ – текст документа, заполненные поля «ключ-значение»

Типология задач машинного обучения

Обучение с учителем (supervised learning)

- классификация (classification)
- регрессия (regression)
- ранжирование (learning to rank)
- прогнозирование (forecasting)

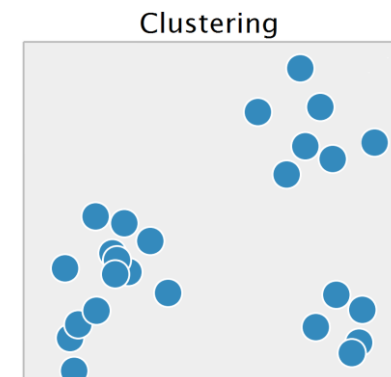


Обучение без учителя (unsupervised learning)

- кластеризация (clustering)
- поиск ассоциативных правил (association rule learning)
- восстановление плотности (density estimation)
- обнаружение аномалий (anomaly detection)

Частичное обучение (semi-supervised learning)

- обучение с положительными примерами (PU-learning)



Типология задач машинного обучения

Предварительная обработка (data preparation)

- извлечение признаков (feature extraction)
- отбор признаков (feature selection)
- восстановление пропусков (missing values)
- обнаружение выбросов (outlier detection)
- уменьшение шума (noise reduction)

Обучение представлений (representation learning)

- обучение признаков (feature learning)
- обучение многообразий (manifold learning)
- анализ главных компонент (principal component analysis)
- матричные и тензорные разложения (matrix and tensor factorization)

Типология задач машинного обучения

Динамическое обучение (online/incremental learning)

- обучение с подкреплением (reinforcement learning)
- активное обучение (active learning)

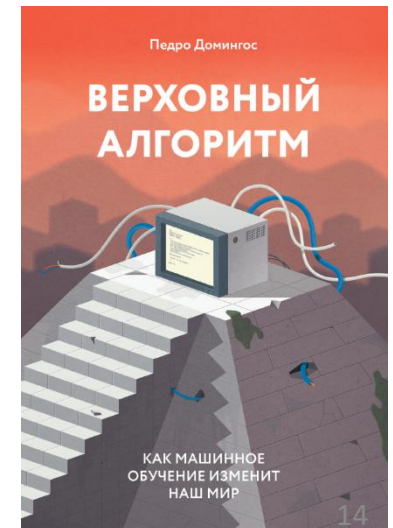
Новые и активно развивающиеся направления

- обучение глубоких сетей (deep learning)
- состязательное обучение (adversarial learning)
- обучение преобразованию последовательностей (sequence-to-sequence learning)
- привилегированное обучение (learning with privileged information)
- обучение выявлению связей (relational learning)
- обучение с переносом опыта (transfer learning)
- мета-обучение (meta-learning)

Типология методов машинного обучения

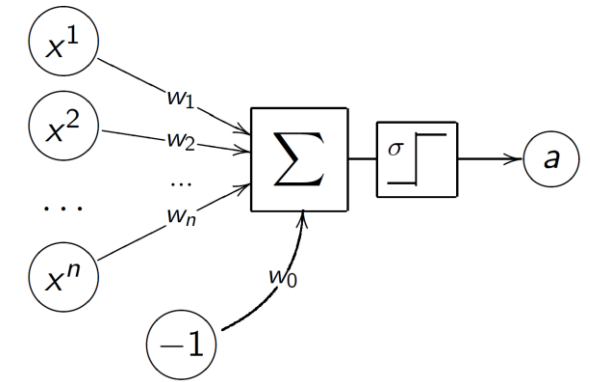
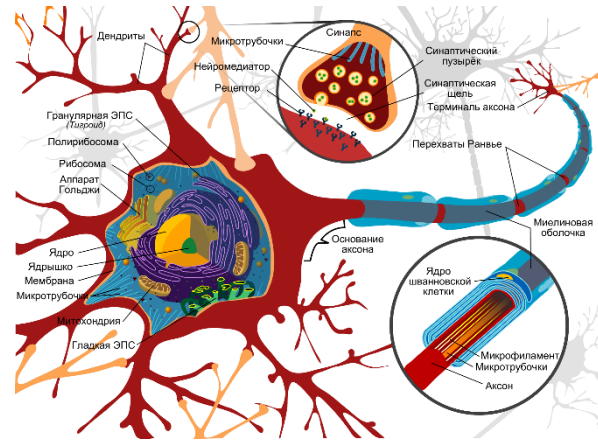
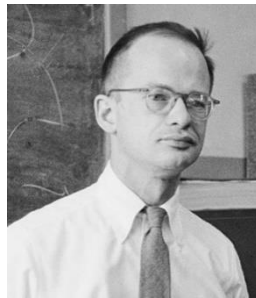
- *Символизм* – поиск логических закономерностей
- *Коннекционизм* – обучаемые нейронные сети
- *Эволюционизм* – адаптивная оптимизация сложных моделей
- *Байесионизм* – оценивание распределений над параметрами
- *Аналогизм* – «близким объектам близкие ответы»
- + *Композиционизм* – кооперация моделей

Педро Домингос. «Верховный алгоритм». 2016.



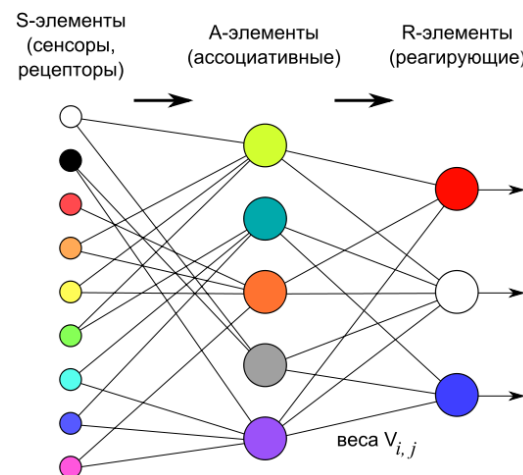
Что такое «искусственные нейронные сети»

Математическая модель нейрона
(МакКаллок и Питтс, 1943)

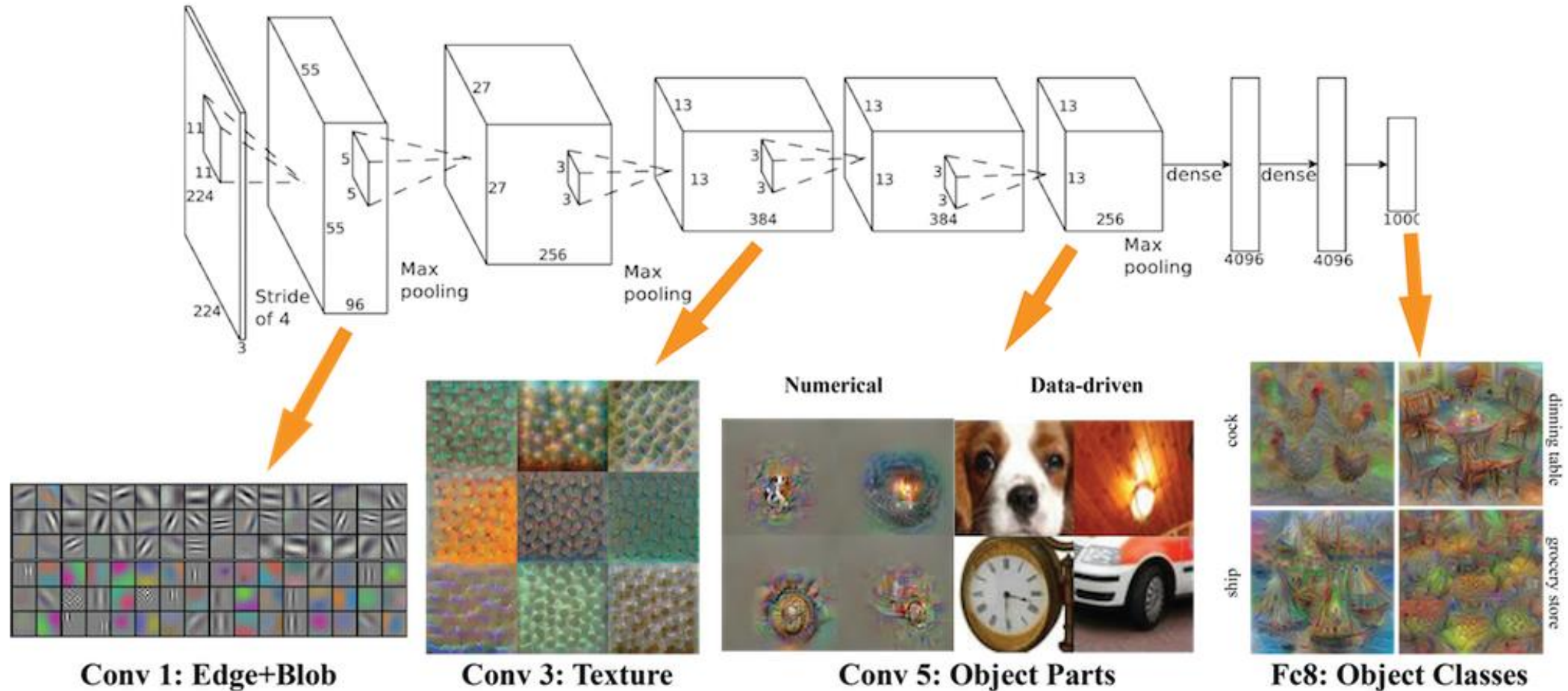


$$a(x, w) = \sigma \left(\sum_{j=1}^n w_j x^j - w_0 \right)$$

Первый нейрокомпьютер Mark-1
(Фрэнк Розенблатт, 1960)



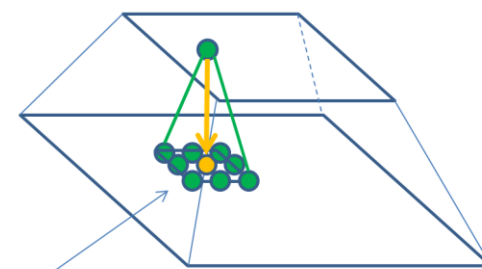
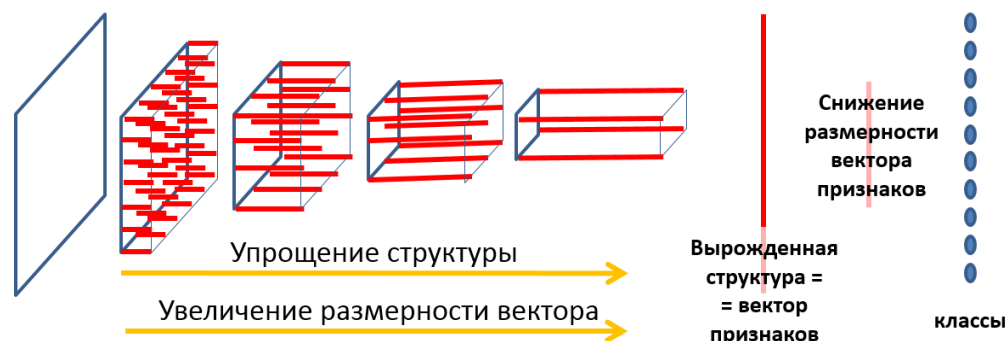
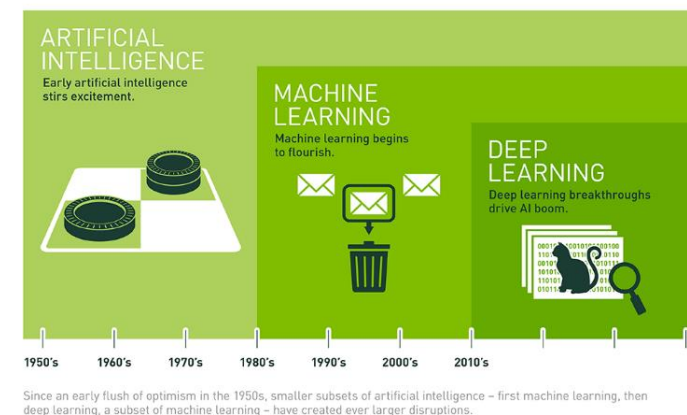
Что такое «глубокие нейронные сети»



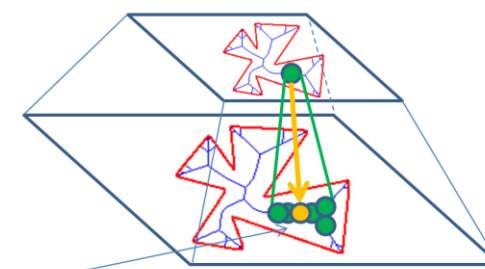
Вытеснит ли DL всё остальное ML?

Глубокие сети – это инструмент автоматизации извлечения признаков (Feature Extraction).

Ближайшее будущее: свёрточные сети обобщаются на любые данные с локальными структурами.



Прямоугольное окно заданного размера с центром в заданной точке + операция свёртки по окну



Локальная окрестность, определяемая для любой вершины графа + операция свёртки по окрестности

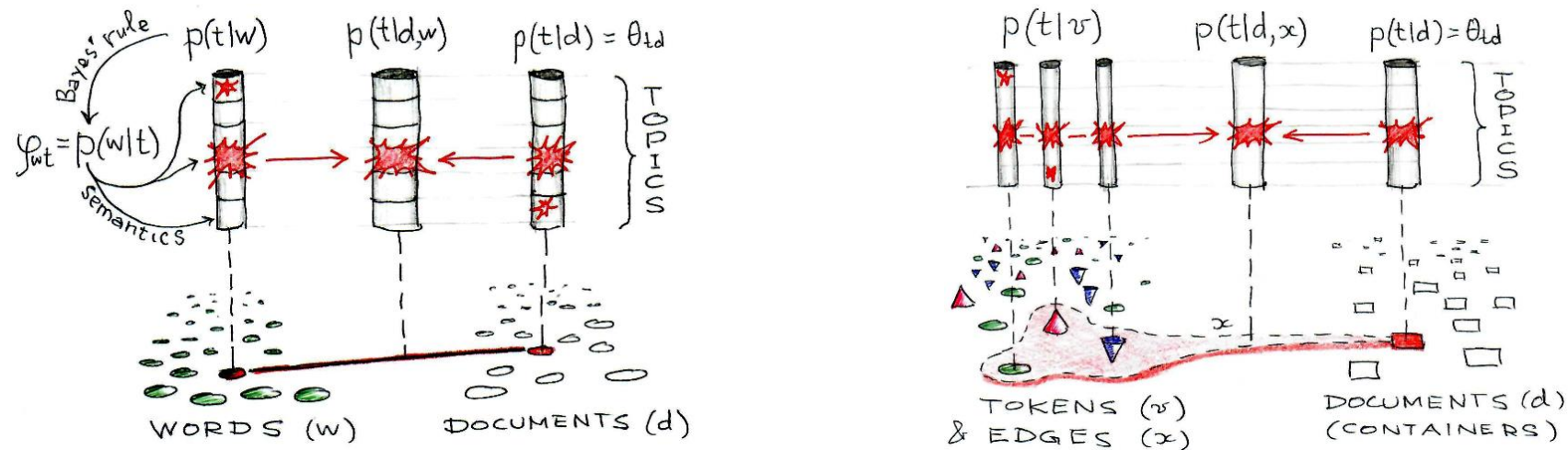
Визильтер Ю.В., Горбацевич В.С. Структурно-функциональный анализ и синтез глубоких конволюционных нейронных сетей. ММРО-2017.

Новые методы: векторизация сложных данных

Сложные данные: тексты, изображения, графы, гиперграфы, транзакции

Векторные представления объектов по наблюдаемым взаимодействиям:

- *неинтерпретируемые:* word2vec, doc2vec, node2vec, graph2vec, prod2vec, StarSpace,...
- *интерпретируемые:* тематические модели (Topic Modeling)



Воронцов К.В. Вероятностное тематическое моделирование: обзор моделей и аддитивная регуляризация. www.MachineLearning.ru. 2019.

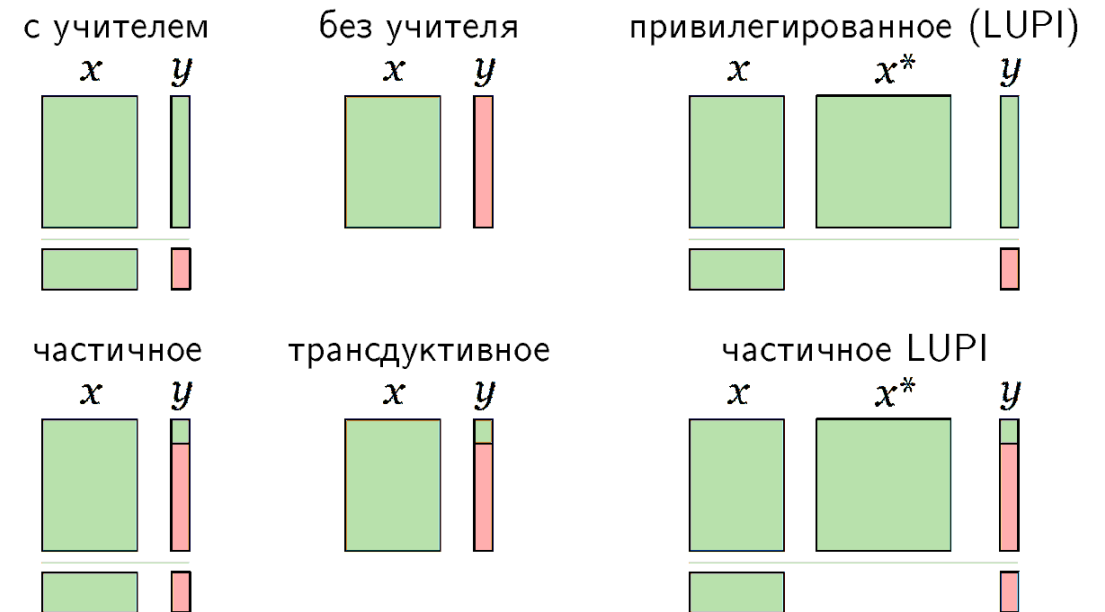
Новые методы: обучение с привилегированной информацией

Естественная модель обучения с учителем:

LUPI – Learning Using Privileged Information

учитель даёт не только правильные ответы, но и объяснения

- На стадии обучения учитель сообщает важную информацию x^* об объектах обучения
- Но на стадии тестирования этой информации не будет



Новые методы: платформы адаптивного обучения

Обычная схема решения задач DS|ML|AI:

- Забираем данные из промышленной системы (долго!)
- Строим модели, экспериментируем в удобной для нас среде
- Переносим модели обратно в пром (долго!)

Будущее – за онлайн-машинным обучением:

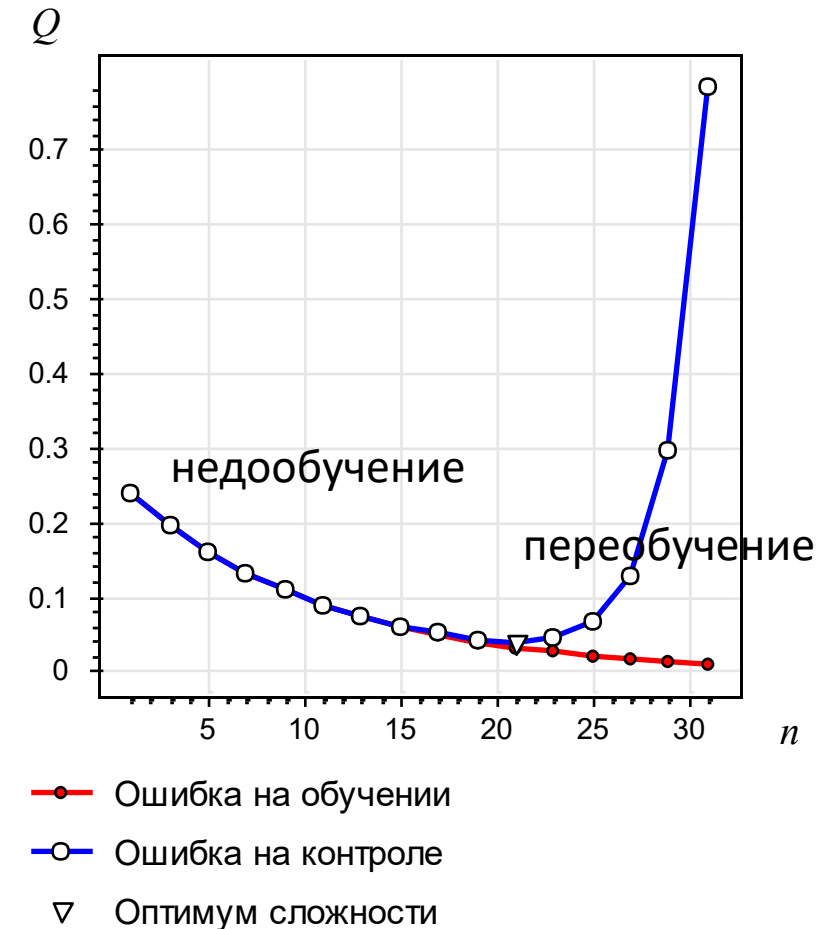
- Предобработка данных и дообучение моделей – налету
- Валидация моделей по совокупности критериев
- Адаптивная селекция и композиция моделей
- Работа аналитика – мониторинг, визуализация и доработка моделей

Методы оценивания моделей

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

Часто используемые внешние критерии:

- hold-out
- (q-fold) cross-validation, leave-one-out
- out-of-sample, out-of-time



Задачи обработки естественного языка

Вспомогательные лингвистические задачи:

Morphological Parsing
Part-of-speech Tagging
Syntax Parsing
Semantic Role Labeling
Named Entity Recognition
Automatic Term Extraction
Word Sense Disambiguation
Paraphrase Identification
Topic Modeling
Word Embedding
Intent Recognition

Конечные бизнес-задачи:

Conversational Intelligence
Machine Translation
Question Answering
Information Retrieval
Sentiment Analysis
Text Summarization
Text Segmentation
Text Classification
Text Clustering
Fact Extraction
Ontology Learning



Стек технологий NLP (см. также nlpub.ru)

- Графематический анализ (tokenization, regular expressions)
- Морфологический анализ, лемматизация (morphological parsing, lemmatization)
- Синтаксический анализ (syntax parsing)
- Автоматическое выделение терминов (automatic term extraction)
- Распознавание именованных сущностей (named entity recognition)
- Сегментация текста (text segmentation)
- Классификация текстов (text classification)
- Кластеризация текстов (text clustering)
- Тематическое моделирование (topic modeling)
- Семантические векторные представления слов (word embedding)
- Семантический анализ и построение онтологий (ontology learning)
- Аннотирование и суммаризация (text summarization)
- Обучаемое ранжирование (learning to rank)
- Ответы на вопросы, машинный перевод, чат-боты (sequence-to-sequence)

Выделение смысла? Понимание речи?

- Поиск «смысла» бессмысленен
- Что такое «понимание», не понятно
- Бизнес, технологии и математика работают только с чётко определяемыми понятиями и чётко поставленными задачами
- Задача чётко поставлена, если для неё описано **ДНК**:
«что **Д**ано – что **Н**айти – **К**ритерий качества решения»
- Измеримый критерий появляется, когда цели прагматичны:
 - автоматизация рутинных операций
 - повышение производительности труда
 - снижение издержек

Типология задач NLP (весьма условная)

- По структуре входов-выходов «чёрного ящика»:
 - классификация, предсказательное моделирование:
ВХОД – текст, **ВЫХОД** – число
 - векторное представление текста:
ВХОД – текст, **ВЫХОД** – числовой вектор
 - информационный поиск:
ВХОД – текст, **ВЫХОД** – ранжированный список документов
 - преобразование и синтез текста:
ВХОД – текст, **ВЫХОД** – текст
- По критерию качества и положению в цепочке обработки данных:
 - бизнес-задачи
 - вспомогательные задачи компьютерной лингвистики
- По уровням анализа текста (пирамида NLP)

Классификация, предсказательные модели

Вход:

- текстовый документ (e-mail, web-страница)

Выход:

- один из двух классов: спам / не-спам

Критерий:

- AUC, чувствительность и специфичность

Модель классификации строится по обучающей выборке,

Основная подзадача: преобразовать текст в векторное признаковое описание фиксированной размерности.

Задача классификации отзывов/обращений

Вход :

- текст отзыва или обращения клиента

Выход:

- класс: куда маршрутизировать запрос / о какой проблеме сообщает клиент или сотрудник

Критерий:

- AUC, чувствительность и специфичность (для многоклассовой классификации)

Модель классификации строится по обучающей выборке,

Основная подзадача: преобразовать текст в векторное признаковое описание фиксированной размерности.

Задача анализа тональности (Sentiment Analysis)

Вход:

- текст отзыва или обращения клиента

Выход:

- оценку тональности отзыва в целом, от -1 до +1

Критерий:

- точность определения тональности на размеченных данных

Модель классификации строится по обучающей выборке, могут использоваться готовые словари тональных слов.

Задача ранжирования поисковой выдачи (Learning-to-Rank, Information Retrieval)

Вход:

- пара «запрос, документ»

Выход:

- оценка релевантности документа запросу

Критерий:

- точность и полнота поиска по ассессорским данным
- качество ранжирования

Надо придумывать признаки, распознавать и исправлять опечатки, учитывать словоформы, синонимы, парафразы.

Машинный перевод (Machine Translation)

Вход:

- текст на одном языке

Выход:

- его перевод на другой язык

Критерий:

- близость к профессиональному переводу, число исправлений

Обучающие данные: большой корпус параллельных текстов, частично с выравниванием, двуязычные словари.

Суммаризация и аннотирование (Summarization)

Вход:

- документ или подборка документов

Выход:

- краткое содержание (реферат)

Критерий:

- точность соответствия (как правило, нескольким) рефератам, написанным людьми (метрики ROUGE, BLUE)

Особенности задачи: надо учитывать словоформы, синонимы, парафразы; надо выбирать самое важное, но без повторов

Ответы на вопросы (Question Answering)

Вход:

- текст вопроса

Выход:

- текст ответа на поставленный вопрос

Критерий:

- точность выделения фразы ответа на размеченной выборке пар «вопрос - текст-с-ответом»

Обучающие данные: коллекция пар «вопрос – ответ»,
большие коллекции текстов, содержащих факты (Википедия)

Разговорный интеллект (Conversational Intelligence, chatbots)

Вход:

- текст диалога бота с человеком

Выход:

- следующую реплику бота

Критерий:

- тест Тьюринга: человек-судья не может отличить собеседника-человека от собеседника-бота
- в приложениях: доля случаев, когда потребность клиента была удовлетворена / когда оператор принял подсказку бота

Обучающие данные: коллекция диалогов операторов с клиентами

Что должен уметь *инженер-исследователь* в области Data Science

- Готовить данные (*Data Engineer*):
 - Работать с сырыми данными в любых форматах
 - Визуализировать, понимать, очищать, преобразовывать данные
 - Доводить пилотное решение до внедрения (production)
- Строить модели (*Data Scientist*):
 - Выбирать инструменты и методы под задачу
 - Строить признаки/архитектуры (feature/architecture engineering)
 - Оценивать и сравнивать модели в соответствии с бизнес-целями
 - Делать анализ ошибок и корректировать модели
- Ходить по кругу CRISP-DM

Что должен уметь *менеджер* в области Data Science

- Видеть применимость машинного обучения в бизнесе
- Ставить задачи в виде «Д-Н-К» (Дано-Найти-Критерий)
- Разбираться в методах на уровне «возможности–ограничения»
- Организовывать бизнес-процессы для сбора чистых данных
- Организовывать открытые конкурсы анализа данных
- Привлекать научно-исследовательские группы
- Запускать пилотные проекты для тестирования гипотез
- Формировать проектные команды
- Адекватно оценивать сложность задач и трудозатраты

Рекомендуемая литература

- *Домингос П.* Верховный алгоритм. 2016.
- *Коэльо Л. П., Ричарт В.* Построение систем машинного обучения на языке Python. 2016.
- *Бенджио И., Гудфеллоу Я., Курвилль А.* Глубокое обучение. ДМК-Пресс, 2018.
- *Николенко С., Кадурич А., Архангельская Е.* Глубокое обучение. Питер, 2018.
- *Гольдберг Й.* Нейросетевые методы в обработке естественного языка. ДМК, 2019.
- *Воронцов К. В.* Лекции по машинному обучению. www.MachineLearning.ru, 2004-2018.
- *Hastie T., Tibshirani R., Friedman J.* The Elements of Statistical Learning. Springer, 2014.
- *Bishop C. M.* Pattern Recognition and Machine Learning. Springer, 2006.