



История и методология математики и компьютерных наук

«Прогресс стоит человеческому разуму меньших

усилий, чем познание самого себя»

Пьер-Симон Лаплас.

Лекция 8

Эволюция компьютерных технологий: от программирования конечных автоматов к интерпретации результатов и обучению

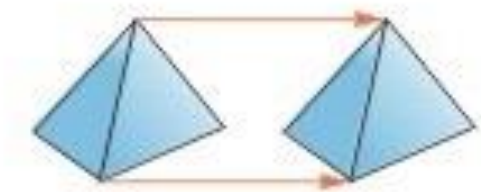
27 октября 2021 г.

- Привести примеры
 - трансляционных преобразований (функций)
 - Конечного, счетного и несчетного множества

- примеры

- трансляционных преобразований

- функция $f(x) = f(x + t)$,
- законы физики трансляционно-инвариантны при пространственном переносе
- любой объект, который по крайней мере в одном направлении объект бесконечен...



- множество всех цифр, множество целых чисел, множество вещественных чисел

Что было на прошлой лекции

- Декларативный и функциональный подход к описанию моделей данных.
- Декларативное описание – ДЕДУКЦИЯ задаёт спецификацию решения рассматриваемой задачи или описание того, что представляет собой проблема и ожидаемый результат.
- Дедукция обладает предсказательной «силой», только если описание толерантно погрешностям количественной характеристики объектов, т.е. характеризует устойчивые процессы.
- Функциональное описание – ИНДУКЦИЯ задает характер законов развития, эволюции или алгоритмов «работы» системы
- Ошибки вычислений (разрядность обрабатываемых данных) могут носить
 - случайный, т.е. устранимый,
 - методический (системный)характер

Что будем обсуждать : аксиоматика нейронауки

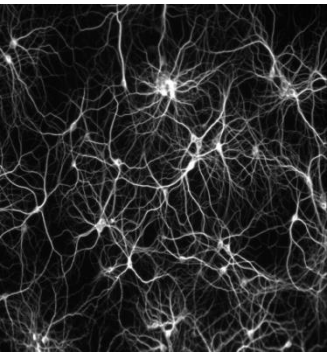
Почему программирование МТ **не vs** машинное обучение ИНС

Аксиоматика нейронауки

- любой конечный автомат, реализующий набор частично рекурсивных функций, можно заменить искусственной нейронной сетью
- искусственные нейронные сети способны хранить информацию и производить вычисления некоторого класса функций
- **любая цифровая вычислительная машина** может быть представлена в виде надлежащего соединения конечных автоматов, которые возможно заменить нейронными сетями, которые могут аппроксимировать решения, но лишь с некоторой **точностью**

Вопрос: а можно ли мозг человека отнести к категории объектов «вычислительная машина», решающая задачи предсказания ?

Ответ – определено нет. Сравнительные характеристики природных и электронных вычислителей

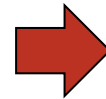


Формирование
нейронных
колонок,
цепей,
сетей

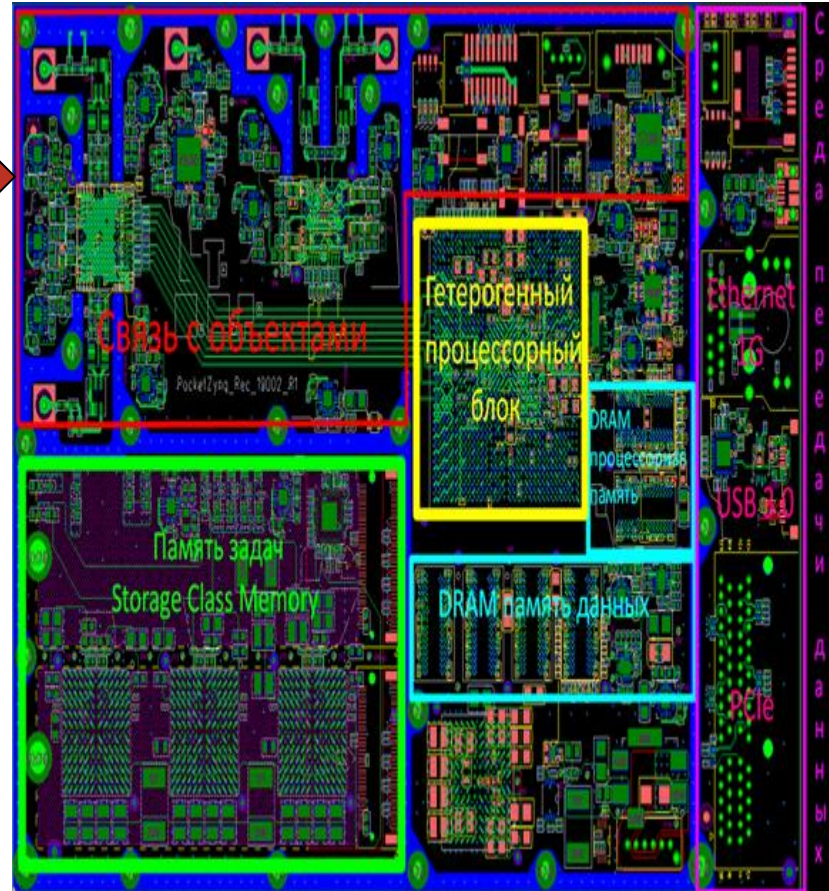


Структура мозга
МРТ снимок

Число нейронов (коммутаторов) 10^{11} шт.
Число синапсов 10^{14} их длина более 10^6 км
Объем памяти > 10^{15} байт



10^6



Структура вычислителя:

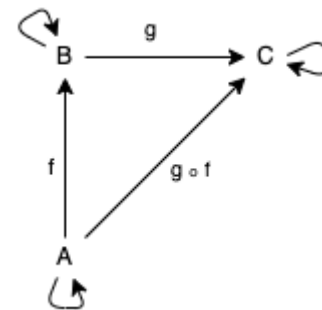
Число процессорных ядер 10^4
Длина соединительных линий 1 км
Объем памяти 10^{13} байт

Математика описания реальности : от теории множеств к теории категорий

- В современной математике, построенной на теории множеств, считается, что любой **объект** исследований должен принадлежать некоторому несущему множеству, состоящему из отдельных неделимых элементов. При выполнении операций с элементами само множество не меняется.
- В теории категорий **преобразования объектов** (объекты – аналоги множеств, преобразования – аналоги отображений) входят в аксиоматическое определение наравне с объектами.
- В итоге **объекты** оказываются предельным случаем (результатом) **процесса** некоторых преобразований. Предметом научного исследования становятся совокупности способов преобразований объектов, **т.е. процессы**.

Динамика описания: от множества vs категории

- категории являются «гибким» шаблоном или «патерном» с уже готовыми конструкциями и операциями, с помощью которых можно строить дедуктивные «теории», которые описывают различные способы преобразования(трансляции) объектов, т.е. процессы, сохраняющие некоторые инварианты, характеризующие эти объекты.
- Примеры : категория групп, множеств, топос...



ИСТИННО ЛИШЬ ТО, ЧТО МОЖНО

- Вычислить на компьютере (? теоремы Геделя)
а полученный результат
- Понять и объяснить (? неформализуемость истны)

Ключевой вопрос: что есть «инвариант» процесса цифровой трансформации

В чем суть алгоритмической вычислимости и как ее достичь, используя феномен ИИ (....) ?

1) **Вычислимость** – описание решения «**прямых**» задач с помощью программ-алгоритмов, реализующих счетный набор частично рекурсивных операций ?

2) **Интеллектуальность** - решение «**обратных**» задач путем построения АЛГОРИТМОВ выбора из множества решений одного «правильного» и объяснить (интерпретации) полученный результат

Интерпретируемость (объяснение) результатов - это степень уверенности ... в том, что вычисленный результат можно **понять**, а используемый алгоритм вычислений **уточнить** так, чтобы повысить степень уверенности и снизить «коэффициент незнания»

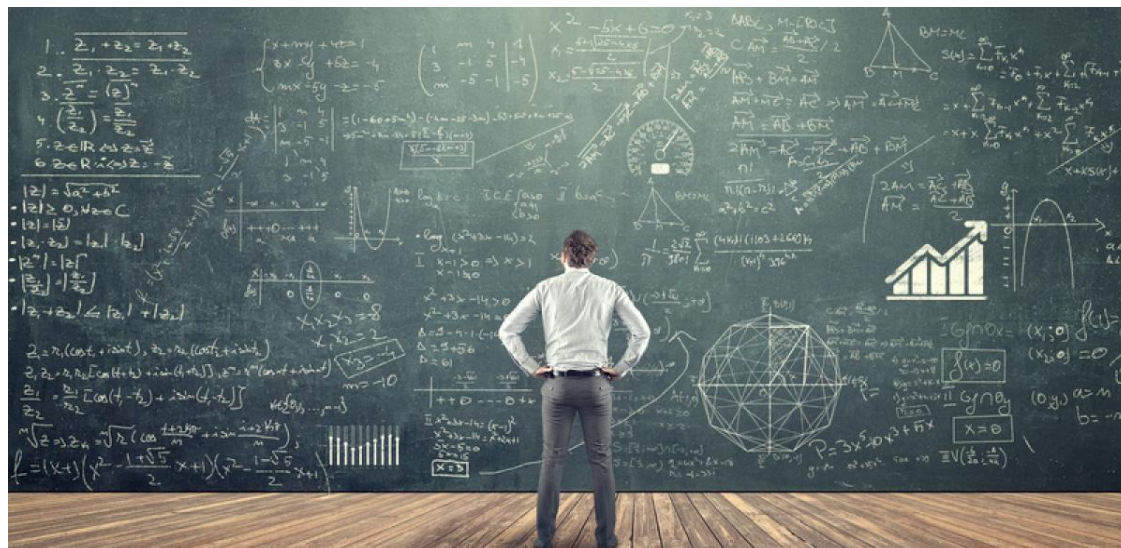
Важно понимать:
каждая вычислимая функция $f(x)$ может быть представлена **счетным множеством** различных **программ!**

и уметь отвечать на вопросы:

Какую программу из возможных следует **выбрать** (написать), чтобы **можно было понять** (объяснить) полученные результаты ?

(какую роль здесь играет математика и ее аксиомы выбора ?)

Если вы не можете **объяснить** что-либо простыми словами, вы это не **понимаете**
Р. Фейнман



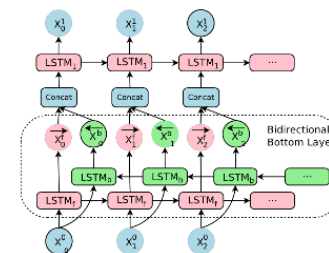
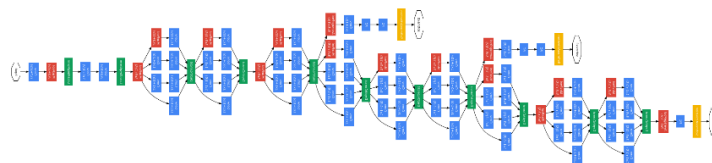
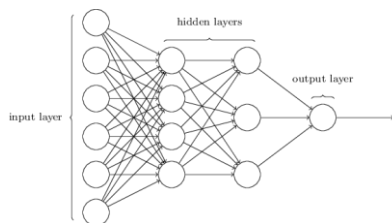
ЭВОЛЮЦИЯ СКТ В НАПРАВЛЕНИИ ПОВЫШЕНИЯ ИНТЕЛЛЕКТУАЛЬНОСТИ

За счет чего повышается интеллектуальность ?

Физика достигла таких высот, что мы можем рассчитать даже то, что невозможно себе представить

Л. Ландау





Многоуровневый перцептрон

- классификация
- функциональная аппроксимация
- автокодер

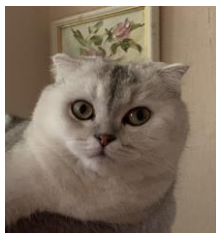
Сверточные нейронные сети

- выделение свойств
- детектирование объектов
- сегментация изображений

Рекуррентные нейронные сети

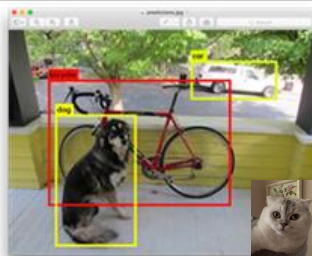
- Обработка временных потоков
- Преобразования разговора в текст
- Перевод текстов

классификация



“КОТ”

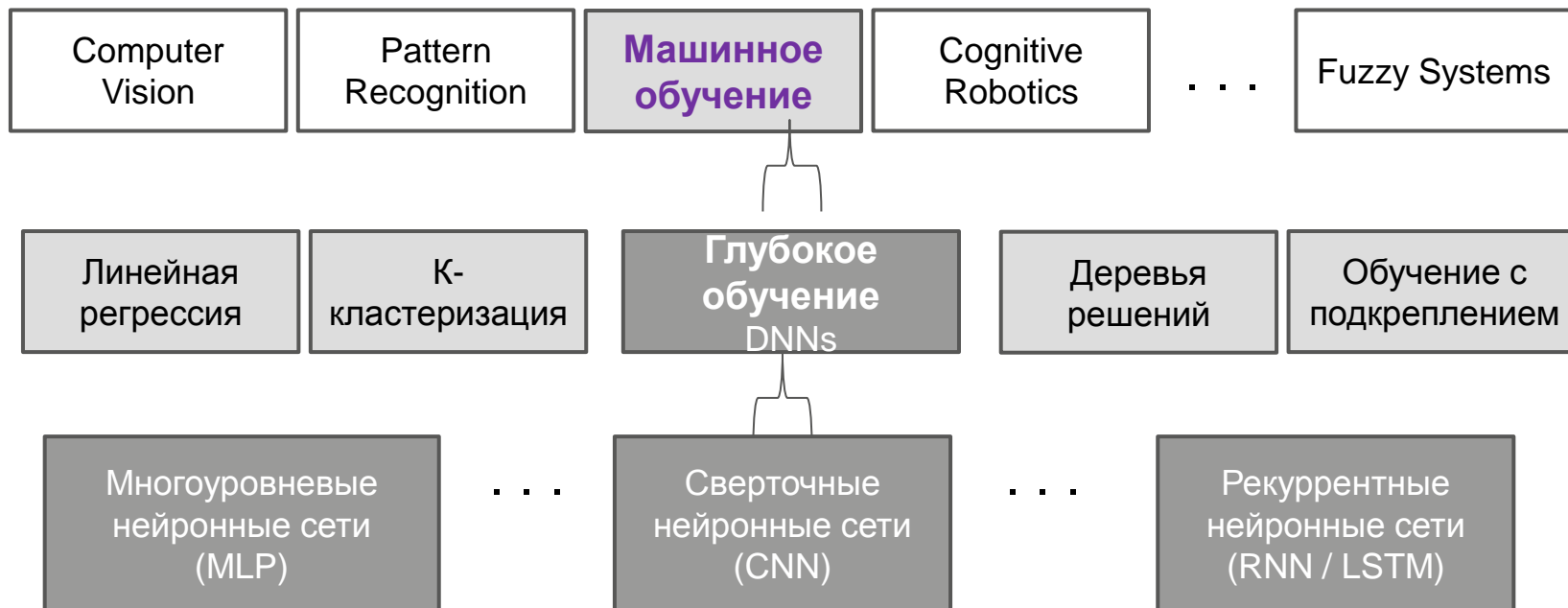
Обнаружение объектов



Сегментация данных



Используя платформы PyTorch и Tensorflow, можно распределить рабочую нагрузку по сотням (или тысячам) идентичных GPU (должны иметь одинаковую память и вычислительную производительность) **время обучения ИНС можно сократить**

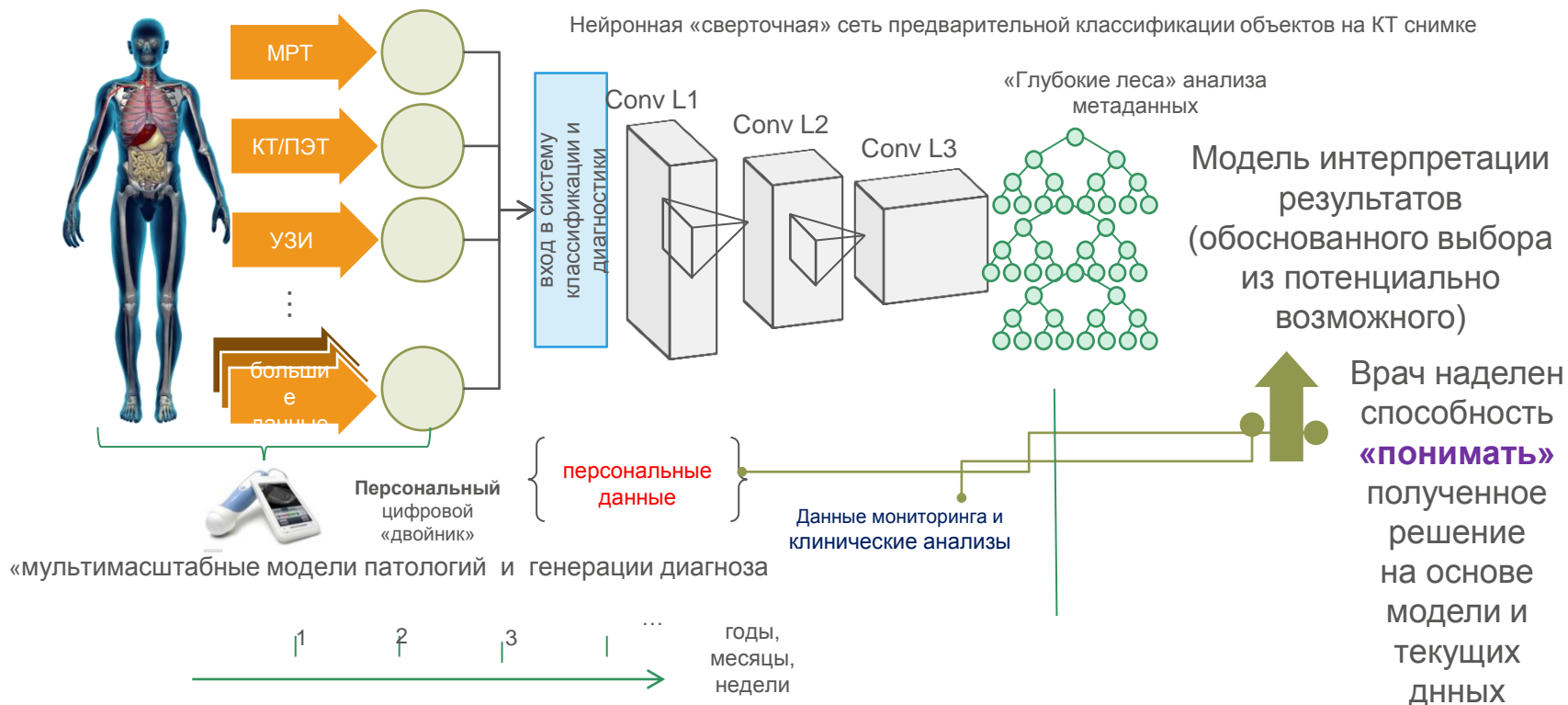


Платформы существуют, но: трудоемкость обучения высокая, а робастность результатов низкая

ПРИМЕР «ЦИФРОВОЙ ДВОЙНИК ПАЦИЕНТА»: АЛГОРИТМИЧЕСКАЯ МОДЕЛЬ + ТЕКУЩИЕ ДАННЫЕ

Потоки первичных данных, характеризующих состояние биологических тканей и процессов

Суперкомпьютерный центр хранения, анализа и предсказательной диагностики



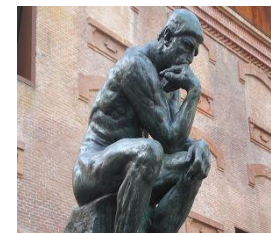
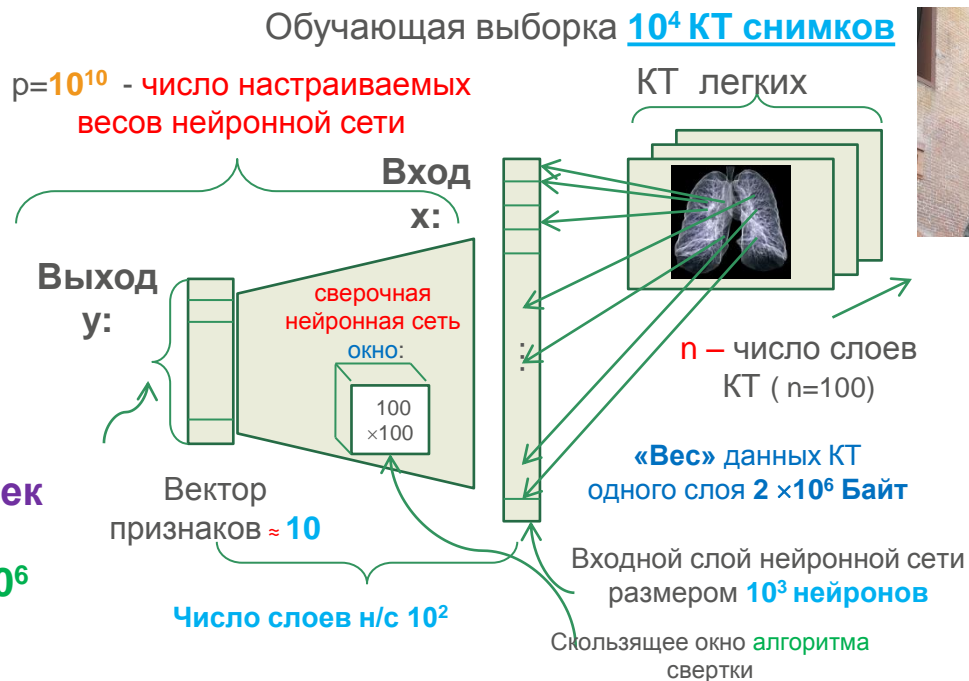
Пример: задача обучения ГЛУБОКОЙ ИНС распознавать 3D изображения компьютерной томографии (КТ)



СК: 1Пфлопс
= 1×10^{15} Флопс

Для обучения н/с с точностью классификации **90%** надо:

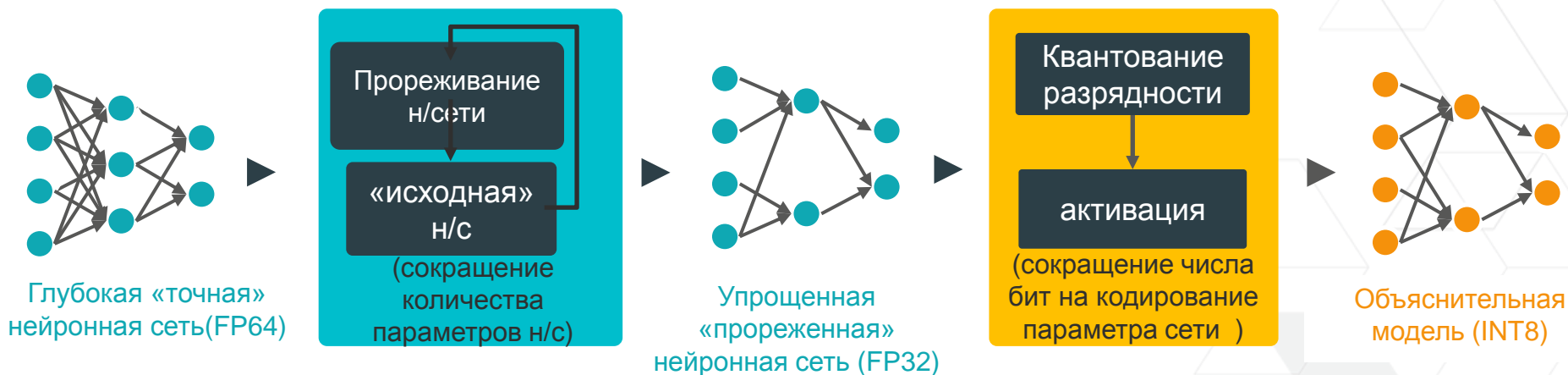
- Число операций $\approx 10^{17}$
- Время обучения ИНС на СК $\approx 10^2$ сек
- t обучения на ПК (200 ГФлопс) $\approx 2 \times 10^6$ сек



обучение ИНС даже среднего размера может занимать недели или месяцы

Градиентный алгоритм обучения глубокой сверточной нейронной сети:

- функция ошибки $F = \|y^* - y\|^2$, y^* - эталонный вектор признаков
- Алгоритм использует 10^{10} частных производных F по всем настраиваемым параметрам;
- Число операций численного дифференцирования на одну итерацию $Q = 10^{15}$



Процесс включает два инструментальных этапа

- Сокращение числа параметров н/с – p раз ?
- Сокращение разрядности данных – q раз ?

Эффекты

- Сжатие размера модели $5x \sim 100x$
- Сокращение времени вычислений $1.5x - 10x$

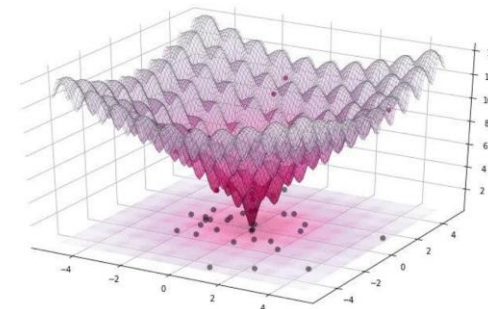
Объяснительную модель можно 1) быстро посчитать; 2) результаты легко интерпретируются и 3) используется вместе с другими метриками, которые типичны для СКТ

- Перечислимость,
 - вычислимость,
 - разрешимость ...

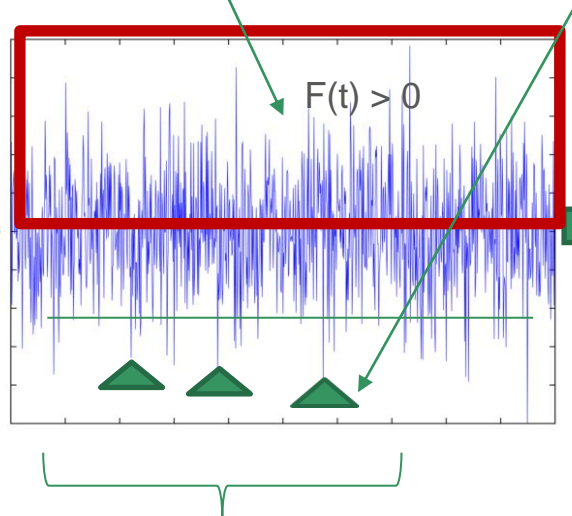
- **Объяснимость (интерпретируемость)**

В силу **сложности** применяемых алгоритмов и гетерогенности вычислительных структур имеется потребность использовать мультимодальные метрики, которые позволяют определять какие:


- параметры модели следует учесть, чтобы **объяснить** результат ?
- характеристики программы вычислений **влиять** на точность результата ?



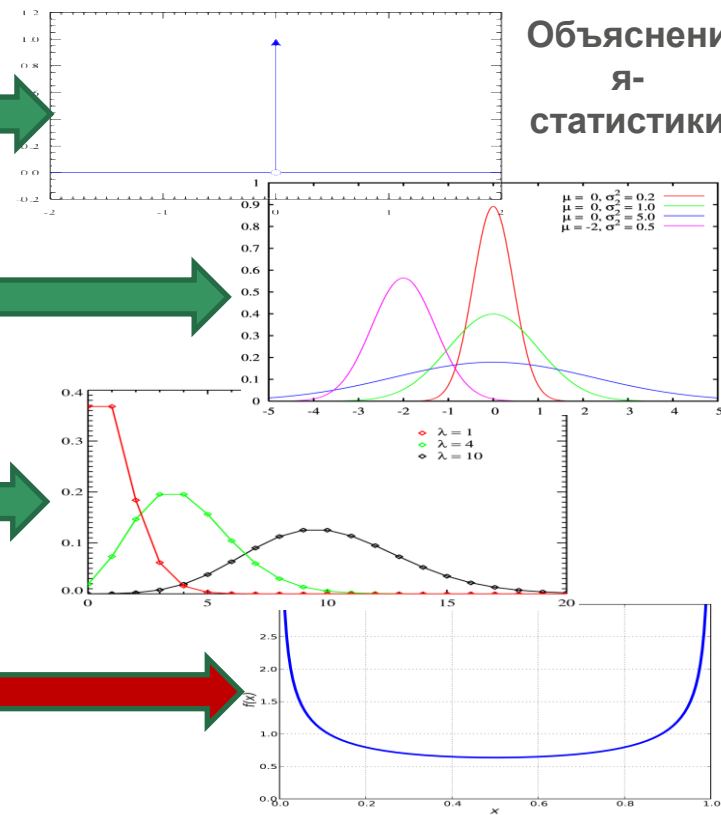
Сложный физический процесс можно моделировать в различные моменты времени



Интеллектуальный классификатор – видим то, что понимаем

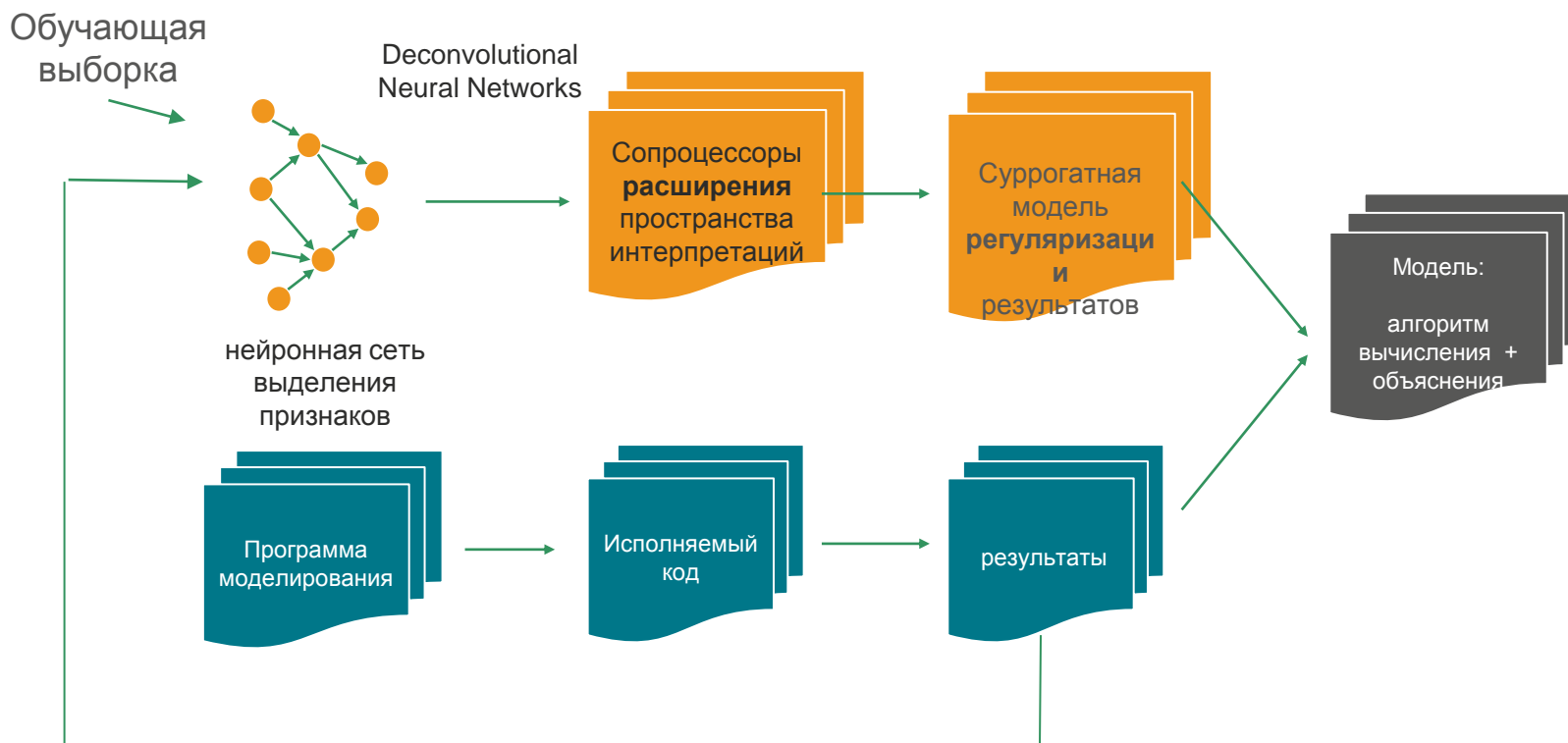


Принятие решения, какими данными можно пренебречь



На конечных интервалах наблюдения сложные процессы могут иметь различные «модели объяснения»

Вопрос: вычисления – это числовые функции или манипулирование символами (образами) ? .





ПОЛИТЕХ

СОДЕРЖАТЕЛЬНАЯ ХАРАКТЕРИЗАЦИЯ ОБСУЖДАЕМОЙ ПРОБЛЕМЫ: ХАИ - EXPLAINABLE ARTIFICIAL INTELLIGENCE

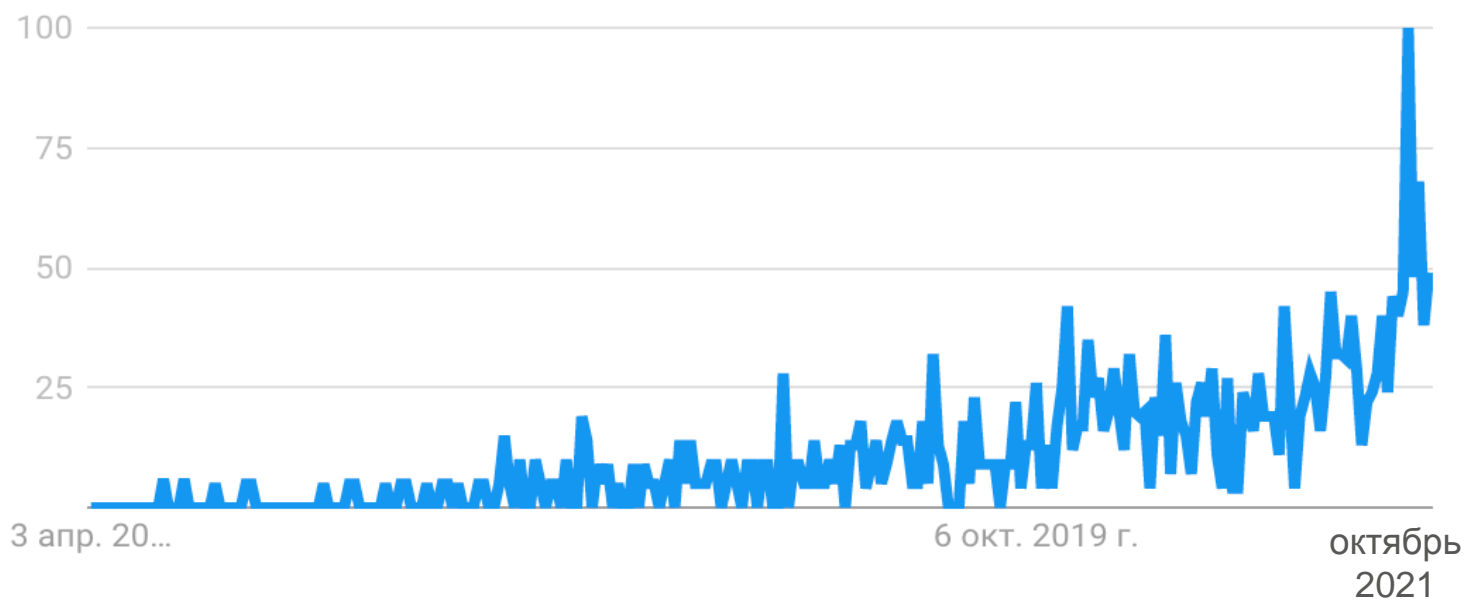
Ключевые слова: объяснение, интерпретация, понимание, доверие

- Дедуктивные **объяснения** (explanation) в терминах понятий
- Индуктивная **интерпретация** (interpretation) в терминах параметров
- **Понимание** сделанных предсказаний (understanding) например с помощью SHAP (SHapley Additive exPlanation)
- **Доверие** к результатам функционирования модели (trust)

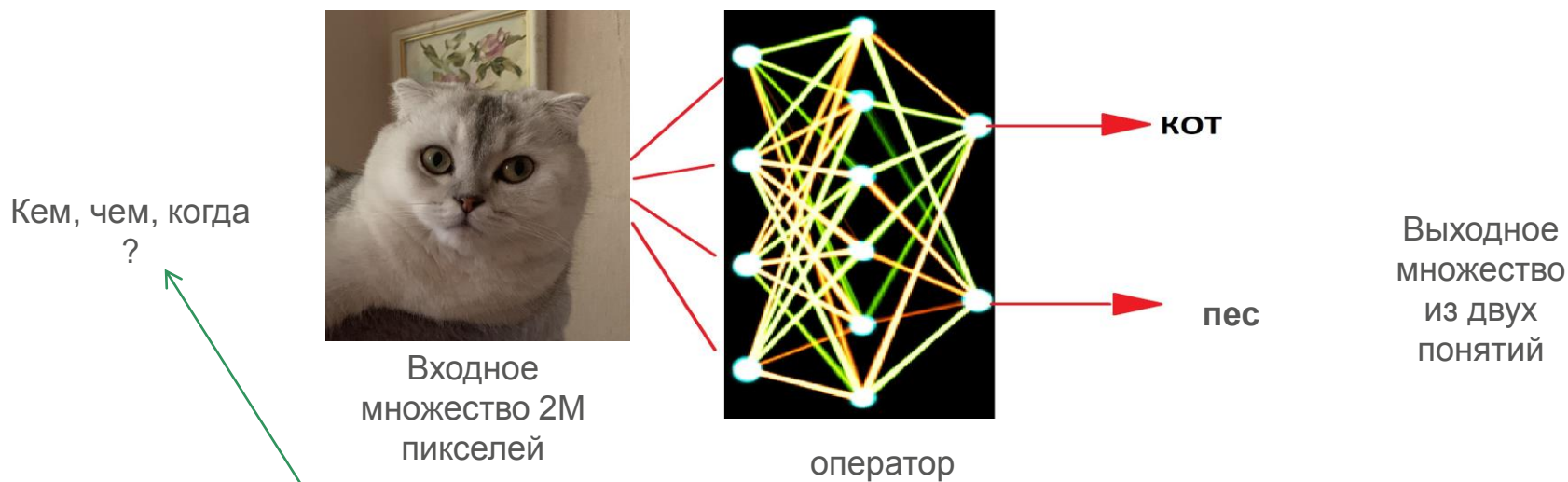
Исходим из того, что если пользователи не доверяют модели или прогнозу, то они не будут их использовать.

Публикации на тему XAI

● Explainable AI



По всему миру. Последние 5 лет. Веб-поиск.

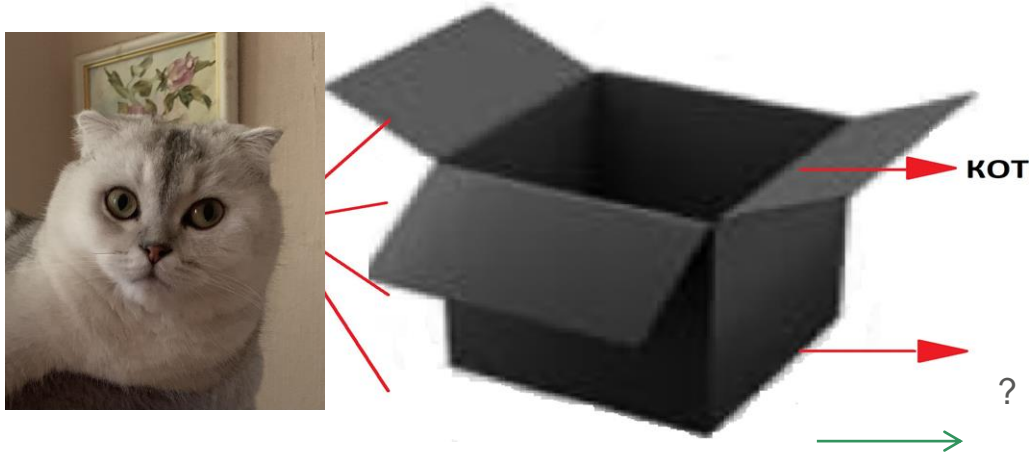


"заранее **обученная** нейроморфная модель **предсказывает**, что **2 М пикселей** входного изображения **кодируют** с помощью hidden **сигнатуры** в выходном embedding vector позицию «кот» с вероятностью **0.98**", а слово «пес» с вероятностью **2%**.

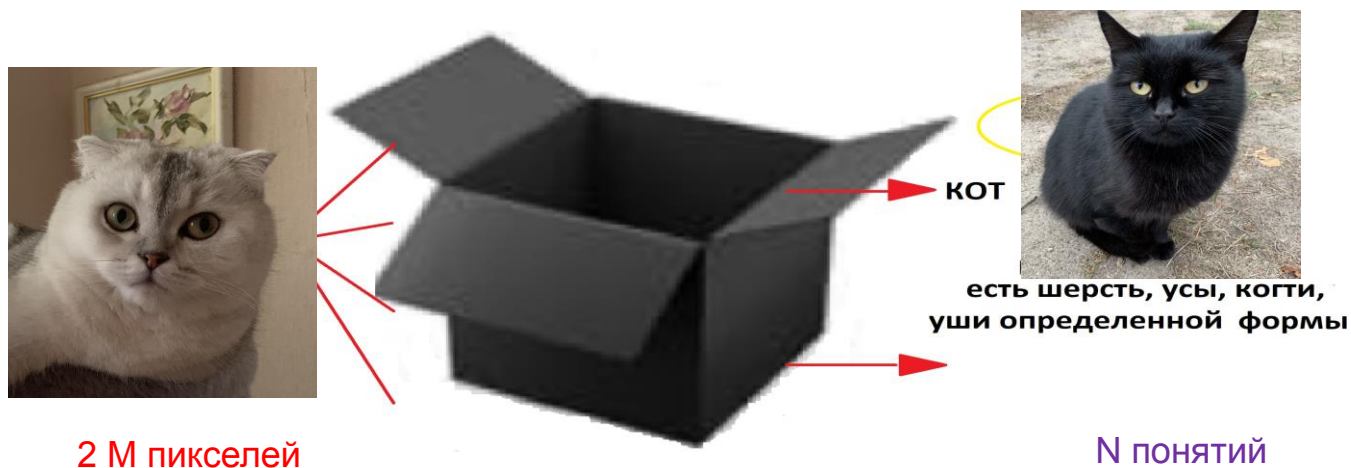


ПОЛИТЕХ

СТАНДАРТНАЯ МОДЕЛЬ МО С ТОЧКИ ЗРЕНИЯ ПОЛЬЗОВАТЕЛЯ

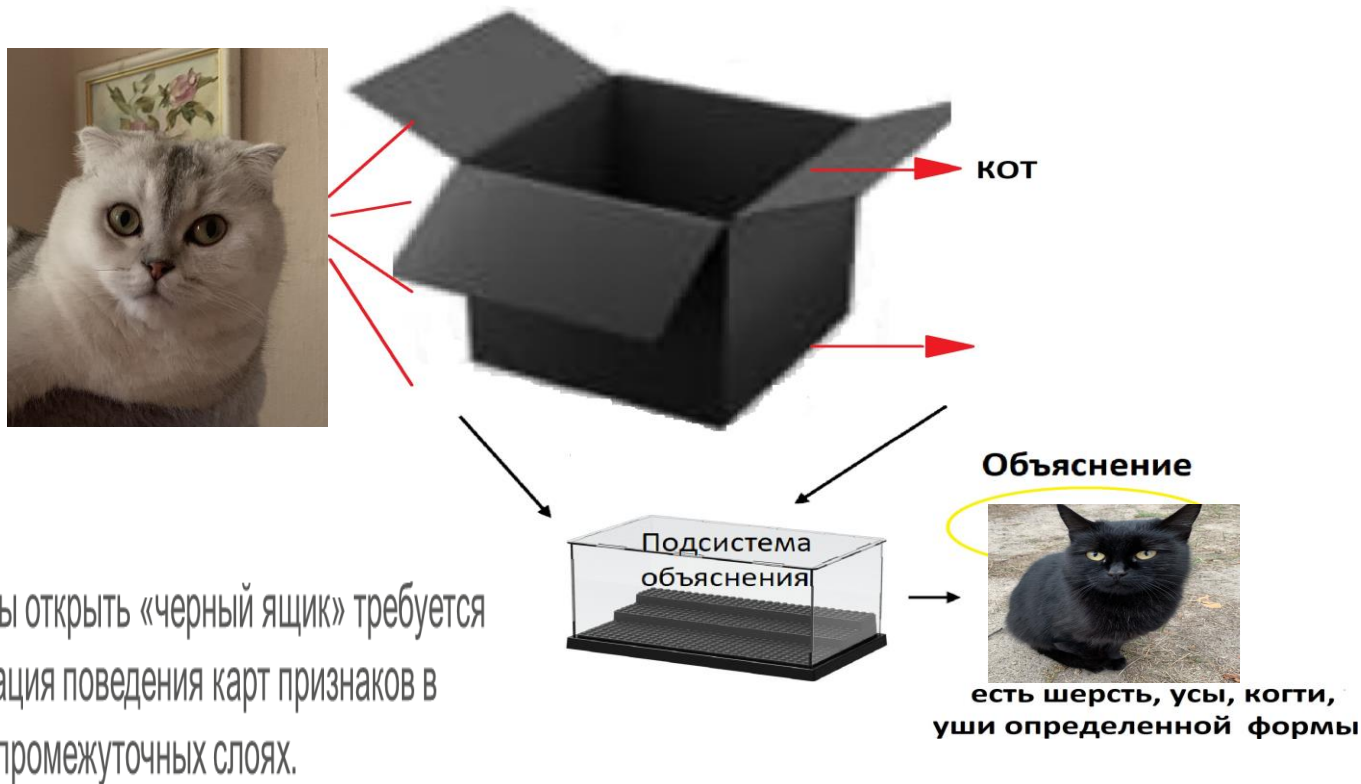


«черный ящик» **предсказывает**, что на фотографии изображен **кот** с вероятностью **0.98**



«черный ящик» предсказывает, что на фотографии - **КОТ** с вероятностью 0.98, так как у него есть шерсть, усы, уши определенной формы"

Должна быть не только система расчетов, но и подсистема дедуктивных объяснений в терминах понятий, объем которых N многократно меньше 2M



А НУЖНО ЛИ ОТКРЫВАТЬ «ЧЕРНЫЙ ЯЩИК»? ?



Да, нужно, так как **эффективные** (точные) модели «черного ящика» обычно сложны для понимания и не **самообъясняемые**, что снижает **уровень доверия к результатам**

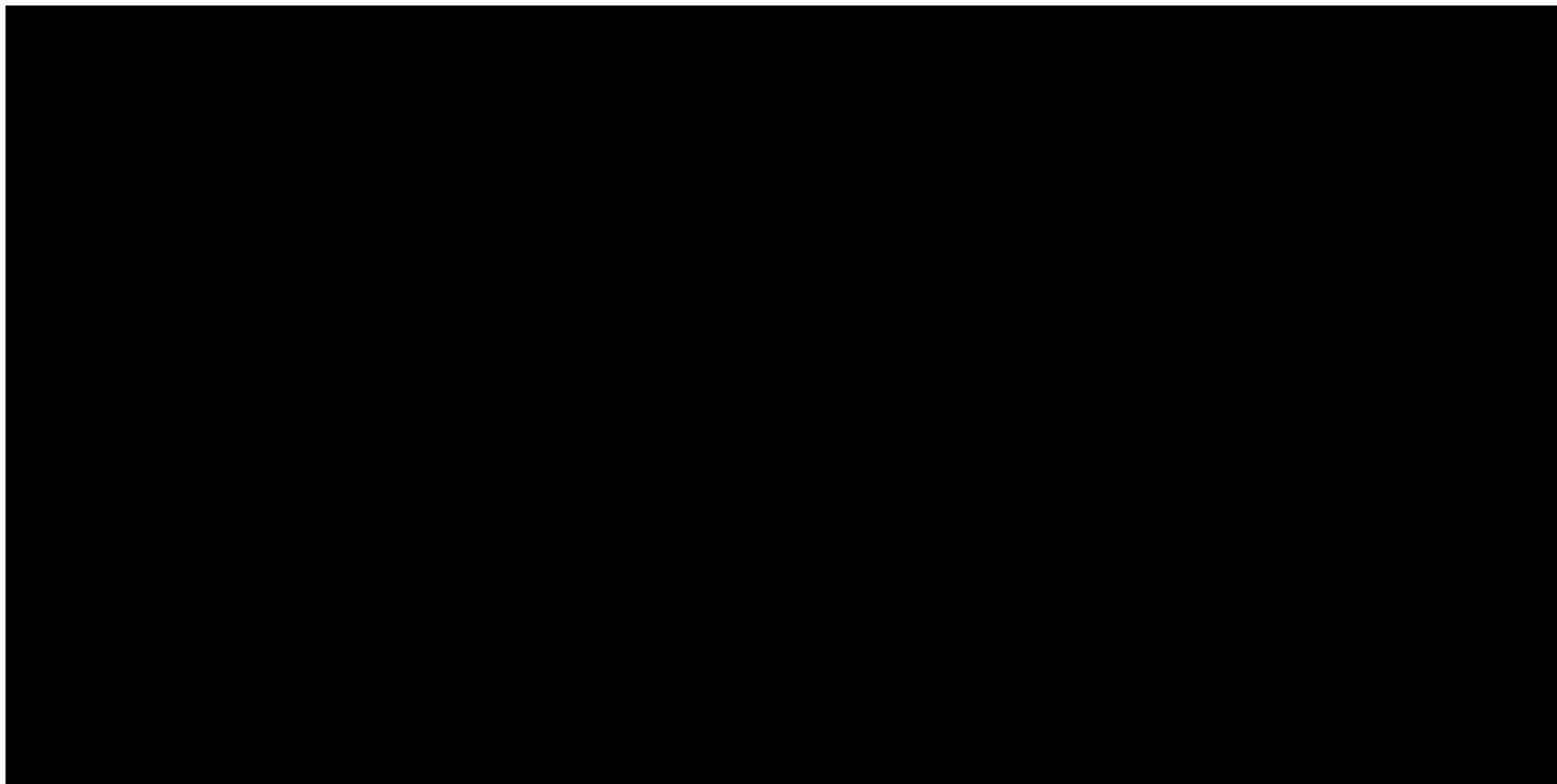
- **программирования** МТ– следование принципу хранимой программы /данных и **последовательного** выполнения **статического** множества операций булевой алгебры.
- машинное обучение искусственных нейронных сетей – «запутывание» программы /данных и **параллельное** выполнения операций отображения **динамического** множеств входных данных на множество выходных данных (аналог процессов с обострением)

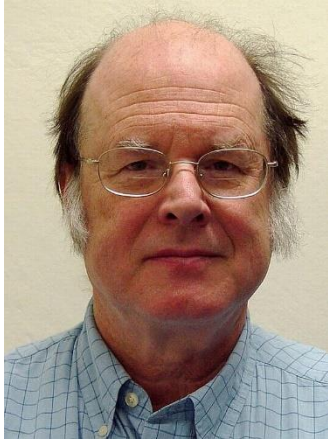
:

Актуальность этих двух подходов высокая, а повысить «остроту» компьютерных инструментов можно за счет «обучения» компьютерных систем реализовывать различные классы алгоритмов, реконфигурируя свою «вычислительную структуру».

	Машина Тьюринга (MT)	Нейронные сети (НС)
Процессор	Сложный	Простой
	Высокоскоростной	Низкоскоростной
	Один или несколько	Большое количество
Память	Отделена от процессора Локализована Адресация по расположению	Интегрирована в процессор Адресация по содержанию
Вычисления	Централизованные	Распределенные
	Последовательные	Параллельные
	Хранимые программы	Способность к обучению
Надежность	Высокая уязвимость	Высокая робастность
Специализация	Численные и символьные операции	Задачи восприятия
Среда функционирования	Строго определенная и ограниченная	Не определенная и без ограничений

Заметим, что связи между нейронами н/с играют роль "памяти» для хранения «программы» и настраиваются в процессе обучения или функционирования.





The digital computer may be thought as an engine that dissipates energy in order to perform mathematical work. Цифровой (компьютер можно представить как «двигатель», который рассеивает энергию ,выполняя математическую работу)

*Charles H . Bennett, 1981
Ч. Беннет*

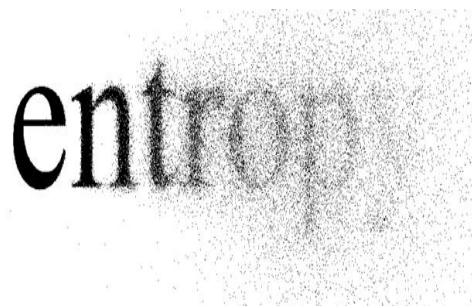
Вопрос: в чем суть «математической работы» которая производится в ИНС с точки зрения физики?

Формально любой алгоритм вычислений с точки зрения физики задает «траекторию» движения (чего) в «лабиринте» состояний конфигурационного пространства компьютера. Алгоритмом может быть «программа», но может быть сама структура «вычислителя»

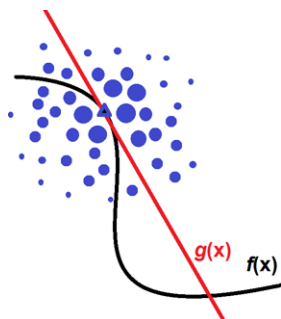
Можно ли так сформулировать *концепцию* «вычислений», так чтобы достигать «математически значимого» результата , при этом не только **потребляя энергию, но и объясняя полученные результаты** ?.

Суть интеллектуализации компьютерных технологий **«нового поколения»**:

1) новая модель память-процессор, способная накапливать информацию о произведенных вычисленных и строить на их основе «суррогатную» модель для объяснения результатов решения задач и различных вариантов их используя ...



Неопределенность
данных



Модели
объяснения

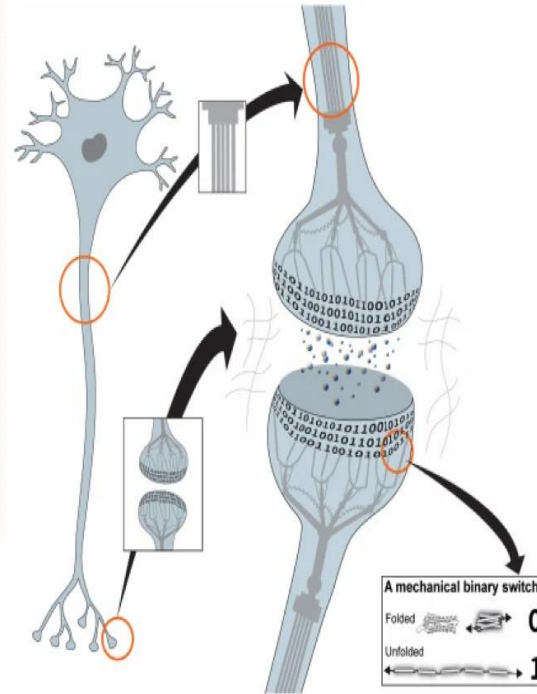
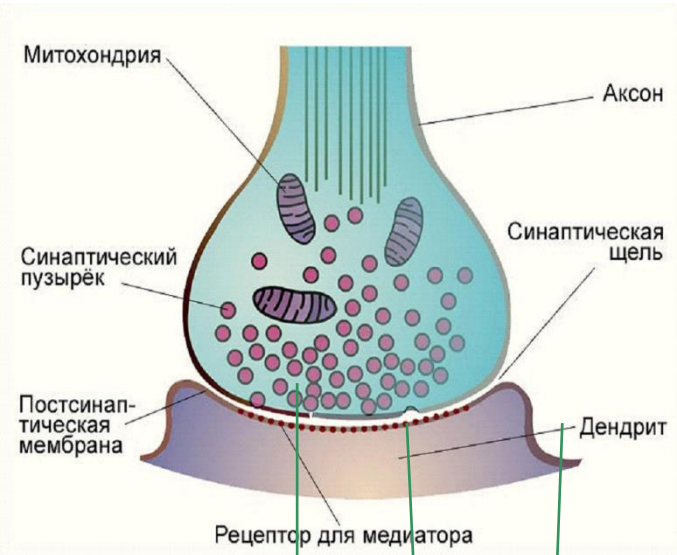
2) новый метод регуляризации с помощью использования «суррогатной» модели, используемой для объяснения того, почему полученное решение **«обратной задачи»** отвечает принятой «модели реальности»

Умение решать обратные задачи – основа интеллекта: в новых условиях «поиск» решений в пространстве возможностей организует и **объясняет сам интеллектуальный вычислитель**.

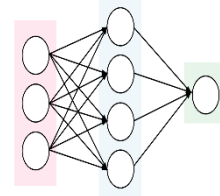


ПОЛИТЕХ

ПРИМЕР РЕШЕНИЯ : БИОИНСПИРИРОВАННАЯ СРЕДА ВИРТУАЛЬНОГО МОДЕЛИРОВАНИЯ ИНФОРМАЦИОННЫХ ПРОЦЕССОВ, ПРОТЕКАЮЩИХ В АССОЦИАТИВНОЙ ПАМЯТИ



Физические каналы передачи информации – цепочка аксонов, инициирующих нейромедиаторные **виртуальные каналы**, являющиеся частью пространства состояний для хранения «паттернов» - аналога **аддитивного ситуационного базиса**



Нейромедиаторы — целостный носитель (сборка) **данных и программ**, представляемый через множество морфизмов .

Системная конфигурация «виртуальные» информационные каналы передачи нейромедиаторов (категорная конструкция)

«Если не знаешь, к какой пристани
плыть, ни один ветер не будет
попутным».

Сенека



В нейронных сетях как пространственно-распределенных системах потенциально могут существовать **режимы с обострением** – процессы, которые можно рассматривать с позиций сгущение точек постоянной фазы, сохраняющих структуру паттерна

Суть научения - настройка **вычислителя** решать задачи исходя из требований

.... **Минимальной** диссипации энергии –

Минимального числа операций (транзакций процессор-память) доставляющих решение задачи

... **Объяснения** полученных результатов на основе конечной совокупности паттернов

Можно предположить, что именно в нейроморфной среде возможно **возникновение «процессов с обострением»** такого преобразования входных данных, при котором **размерность пространства данных сокращается, но не искажаются паттерны**, используемые в системе **«интерпретации» полученных результатов.**

- Интерпретируемость модели важна не меньше, чем числовая точность результатов расчетов.
- Совместное использование суперкомпьютерных технологий и машинного обучения позволяет выявить значение каждого их параметров программы моделирования путем интерпретации полученных результатов как с точки зрения различных числовых «мер» (площадь, длина, вес, скорость, ...), так и понятий.