



**КУРСЫ ПОВЫШЕНИЯ КВАЛИФИКАЦИИ  
(ФАКУЛЬТАТИВНЫЙ КУРС)  
“МАШИНОЕ ОБУЧЕНИЕ.  
ИСКУССТВЕННЫЕ НЕЙРОННЫЕ СЕТИ  
И ГЕНЕТИЧЕСКИЕ АЛГОРИТМЫ”**

---

**НИИЯФ МГУ**

**Лаборатория адаптивных методов обработки данных**

**С.А.Доленко, А.О.Ефиторов, В.Р.Широкий  
И.В.Исаев, И.М.Гаджиев, Р.Д.Владимиров**

**Ю.В.Орлов, И.Г.Персианцев, О.А.Агапкин, А.Г.Гужва**

**<http://kpk-nnga.sinp.msu.ru/>**

**E-mail: [kpk\\_nnga@sinp.msu.ru](mailto:kpk_nnga@sinp.msu.ru), [nnga@yandex.ru](mailto:nnga@yandex.ru)**

**тел. (495) 939-46-19**

**Занятия ведут:**

**зав. лаб. к.ф.-м.н. Сергей Анатольевич Доленко  
инженер Владимир Романович Широкий  
н.с. Игорь Викторович Исаев  
программист Исмаил Маратович Гаджиев  
инженер Роман Дмитриевич Владимиров**

*В курсе  
использованы  
материалы,  
разработанные  
авторами  
при поддержке*



# ПРЕДМЕТ КУРСА: МАШИННОЕ ОБУЧЕНИЕ

- **Машинное обучение = обучение ЭВМ на примерах**
- **Адаптивные методы анализа данных**  
**адаптивный = самоприспосабливающийся**
- **Другими словами, это**  
**методы, управляемые данными**  
**(data-driven methods)**
- **В курсе будут рассматриваться**
  - (искусственные) Нейронные сети (НС или ИНС)
  - Методы подготовки и предобработки данных
  - Генетические алгоритмы
  - Нечёткая логика
  - Современный инструментарий для практической работы
  - И многое другое...

# ПРЕДМЕТ КУРСА

---

**Машинное обучение (МО):**

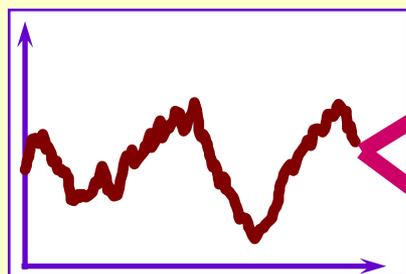
**обучение машин, а не обучение при помощи машин!!!**



**Будет рассмотрена  
общая методология МО:  
что, зачем и  
в каком порядке  
надо делать**

# ПЛОХО ФОРМАЛИЗУЕМЫЕ ЗАДАЧИ

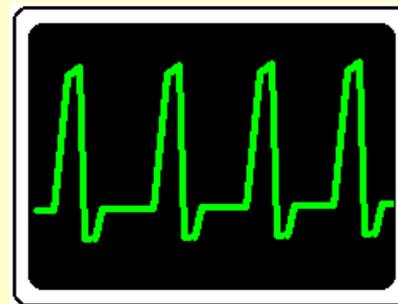
- финансы



рост

падение

- медицинская диагностика

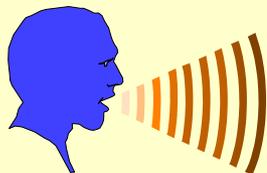


здоров

болен

- распознавание речи и текста

Da



“Да”

- распознавание изображений



# **ПРИМЕРЫ ЗАДАЧ, РЕШАЕМЫХ С ПОМОЩЬЮ ИНС**

---

## **Интернет, телекоммуникации и связь**

- распознавание речи (устной и письменной), изображений
- ассоциативный поиск в базах данных, рубрикация новостей
- компрессия данных
- маршрутизация пакетов

## **Физический эксперимент**

- обработка сигналов в реальном времени
- обнаружение и классификация событий
- решение обратных задач в реальном времени

## **Безопасность**

- идентификация личности
- детектирование взрывчатых веществ

## **Производство**

- адаптивное управление в робототехнике
- прогноз потребления ресурсов
- контроль качества продукции

## **Финансы**

- прогнозирование поведения рынка
- выработка сигналов торговых стратегий
- предсказание банкротств, оценка надежности клиентов

**ПЛОХО ФОРМАЛИЗУЕМЫЕ ЗАДАЧИ**

# АППРОКСИМАЦИОННЫЕ МЕТОДЫ И ИХ МЕСТО

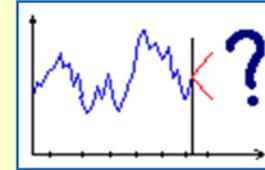
---

- Большинство рассматриваемых задач представляют собой задачи построения **функционального отображения**
- Наилучший вариант – когда функцию можно задать **аналитически** на основе физической модели объекта
- Если аналитическое решение невозможно – пытаемся решить задачу **численно** на основе математической модели
- Если содержательных моделей нет – остается построение **аппроксимационной** модели на основе имеющихся примеров
- Аппроксимационная модель = **разложение** искомого отображения
- Таким образом, «**обучение на примерах**» = построение аппроксимационной модели, т.е. «**черного ящика**»
- Если есть возможность, ящик лучше **покрасить**, т.е. использовать имеющуюся априорную информацию

# ТИПОЛОГИЯ ЗАДАЧ, РЕШАЕМЫХ МЕТОДАМИ МО

## ❑ Регрессия (оценка значения величины)

- Количественная оценка  $Y = 378.14 \pm 0.21$
- Прогнозирование временных рядов
- Аппроксимация зависимостей

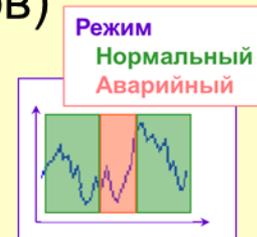


$$Y = F(X_1, X_2, \dots, X_n)$$

## ❑ Классификация (распознавание образов)

- Бинарная
- Многоклассовая
- Многометочная

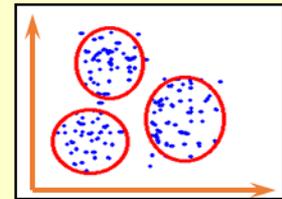
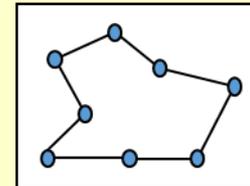
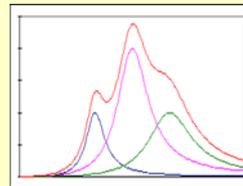
А Б В Г Д



1 2 3 4 5 6

## ❑ Оптимизация (поиск оптимальных значений / комбинаций)

- Поиск оптимальных значений
- Комбинаторная оптимизация



## ❑ Кластеризация (разбиение данных на группы)

Границы между типами задач могут быть размыты, одну и ту же задачу можно ставить как задачи разных типов

# ВИДЫ ЗАДАЧ КЛАССИФИКАЦИИ

---

## ❑ Бинарная

- **Обнаружение** аномалии (*установка неисправна*)
- Прогнозирование **факта** наступления события (*распад произойдёт*)
- Диагностика принадлежности объекта множеству (*вы беременны!*)

## ❑ Многоклассовая

- **Распознавание** символов (*это Ъ*)
- Распознавание **вида** объекта (*это вода*)
- Выбор **одного из** множеств, которым может принадлежать объект (*агрегатное состояние – жидкость*)

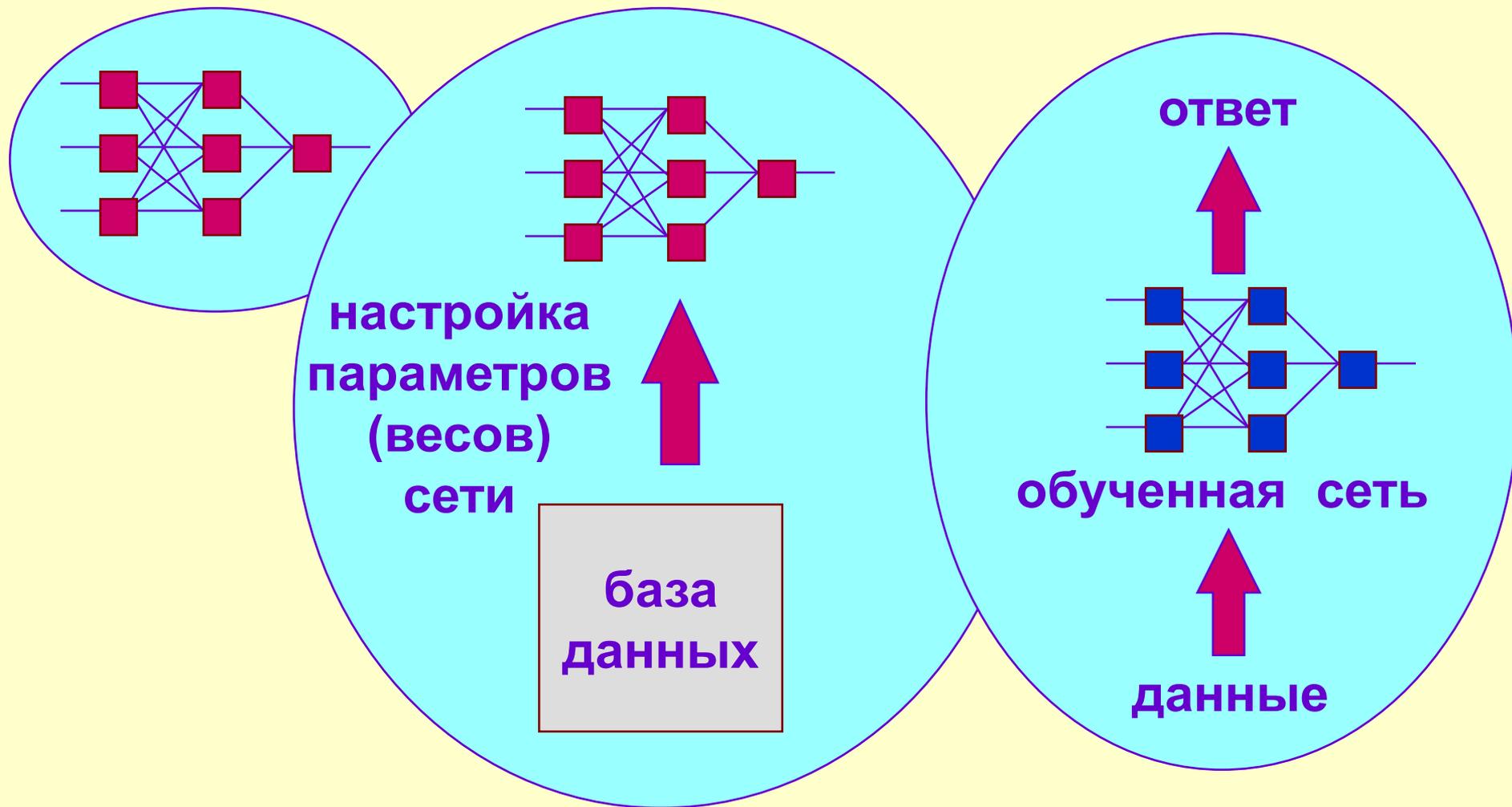
## ❑ Многометочная

- Рубрикация (*статья по физике и методам МО*)
- Диагностика принадлежности объекта **одновременно нескольким** множествам (*в растворе присутствуют ионы меди, натрия и хлора*  
(*животное умеет плавать, летать и имеет 2 ноги*)

# ИСПОЛЬЗОВАНИЕ НЕЙРОННОЙ СЕТИ (МЕТОДА МО)

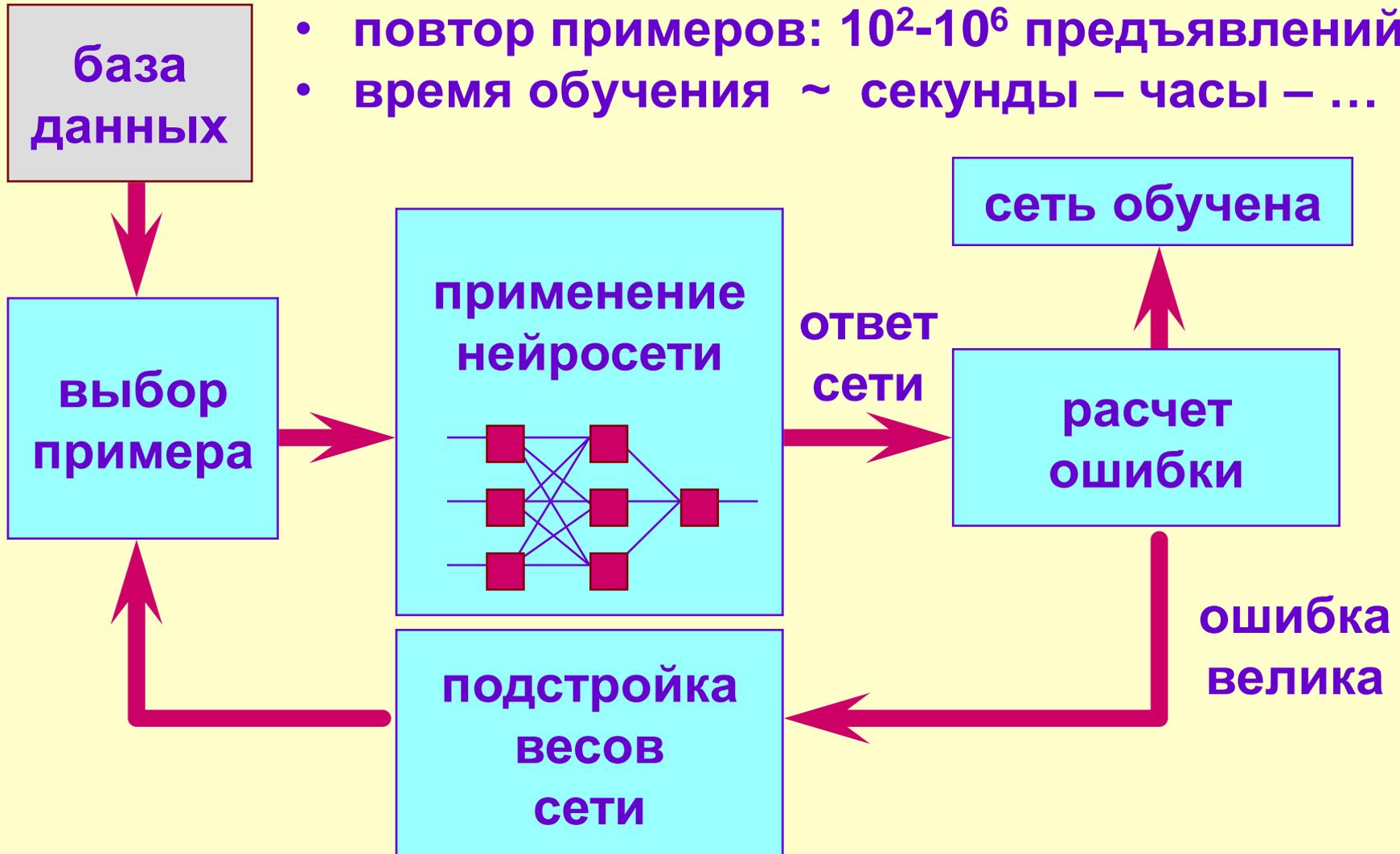
Подготовка данных →

Выбор типа сети → Обучение сети → Применение



# ОБУЧЕНИЕ НЕЙРОННОЙ СЕТИ

- постепенная подстройка весов ИНС
- повтор примеров:  $10^2$ - $10^6$  предъявлений
- время обучения ~ секунды – часы – ...



# ПРЕИМУЩЕСТВА НЕЙРОСЕТЕЙ

---

- + универсальность технологии
- + нет априорных предположений о модели
- + **обучение на примерах**
  - не требует "ТОЧНЫХ" знаний и СФОРМУЛИРОВАННЫХ экспертом правил принятия решения
  - доучивание в изменяющемся окружении без модификации алгоритма
- + **устойчивость к шуму**
  - работают с неполными, противоречивыми, зашумленными данными
- + **обобщение полученной информации**
  - разумный ответ на незнакомых данных ("здравый смысл")
- + **учет многих разнородных факторов**
- + **параллельность алгоритмов**
  - скорость вычисления, устойчивость к повреждениям

---

**= Возможность решения задач,  
не решаемых традиционными методами**

# НЕДОСТАТКИ НЕЙРОСЕТЕЙ

---

- выбор структуры сети - метод проб и ошибок
- субоптимальность решения
- не слишком высокая точность ответа, предпочтение "качественным" задачам
- трудно объяснить причину принятия того или иного решения
- требуется **представительный набор примеров**
- оценка качества работы **по независимым данным**
- относительно высокая вычислительная стоимость обучения

# ЗАКОН МУРА

Гордон Эрл Мур (Gordon Earle Moore), р.1929  
"Количество элементов на кристаллах  
электронных микросхем будет и далее  
удваиваться каждый год« (1965)

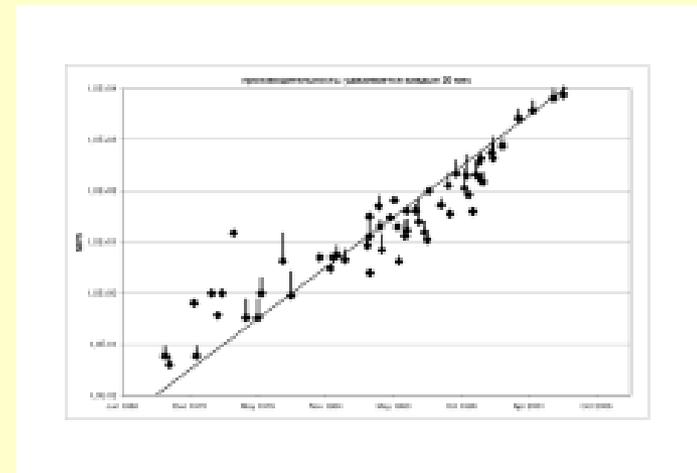


Это эмпирическая закономерность

"Доступная вычислительная мощность  
удваивается каждые 18 (20) месяцев"

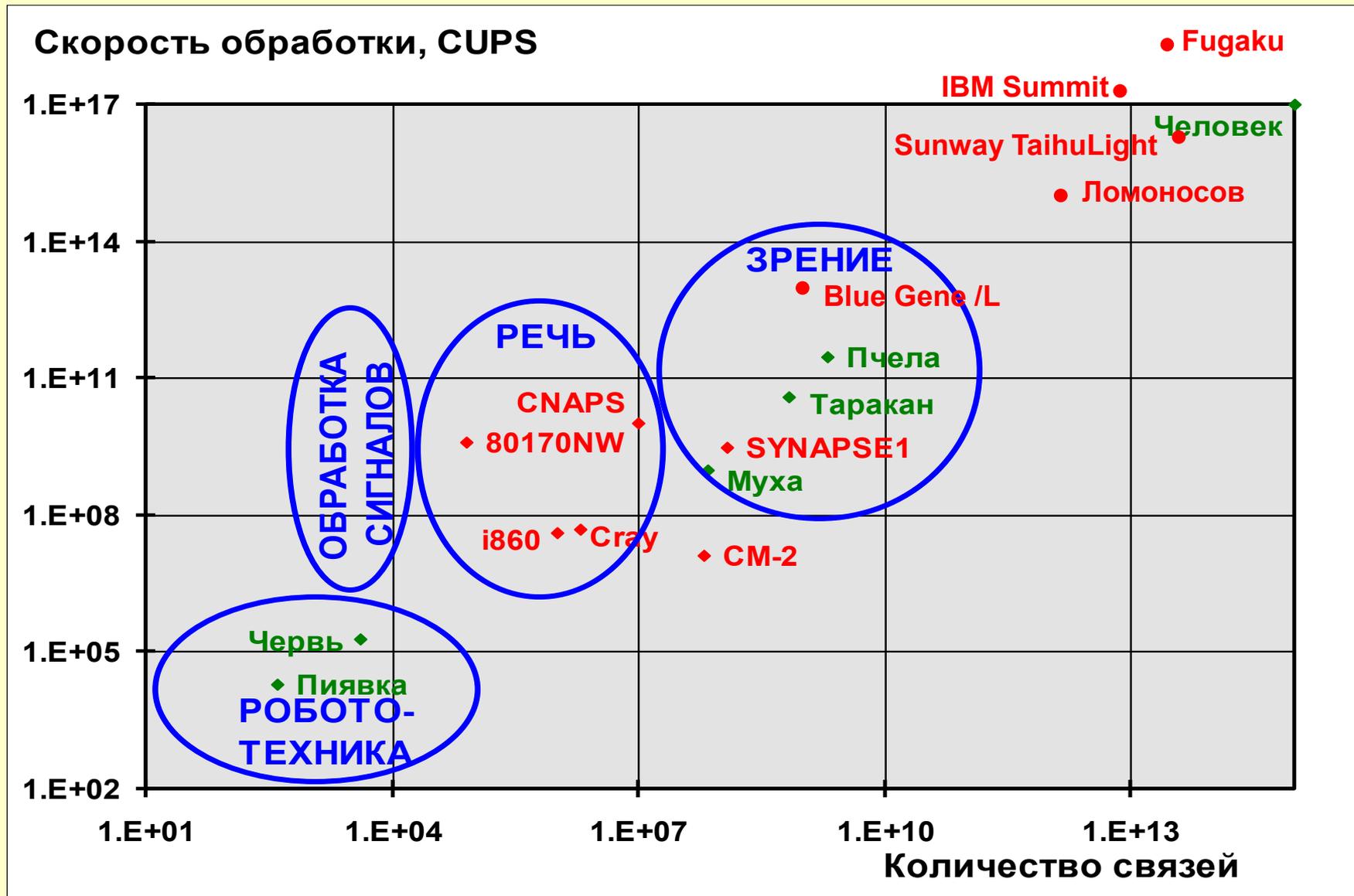
Увеличение производительности  
на порядок ~ каждые 5 лет

С некоторыми оговорками,  
закон Мура выполняется  
уже почти 60 лет!



За это время произошло несколько технологических революций

# ТЕХНИЧЕСКИЕ ТРЕБОВАНИЯ К ИНС



# УРОВНИ ВЗАИМОДЕЙСТВИЯ ИНТЕЛЛЕКТА С ОКРУЖАЮЩИМ МИРОМ

---

1. Перцептивный (лат. *perceptio* – восприятие)
  - Восприятие окружающего мира
  - Осуществление простейших ответных действий – реакция на внешние раздражители
2. Когнитивный (лат. *cognitio* – познание)
  - Познание окружающего мира
  - Установление закономерностей
  - Обучение действиям в ситуациях, схожих с известными
3. Креативный (лат. *creatio* – творчество)
  - Творческое осмысление и освоение окружающего мира
  - Изобретение нового
  - Умение адекватно действовать в незнакомых ситуациях

На повестке дня – переход с когнитивного на креативный уровень.

Узенькие мостики через пропасть уже перекинуты, и они всё шире...

Мы уже всерьёз говорим о создании искусственного интеллекта (ИИ)

# ДОБЫЧА ДАННЫХ, МАШИННОЕ ОБУЧЕНИЕ И ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

---

## ❑ Добыча данных (Data Mining, DM)

- Сбор данных, их очистка и первичная обработка
- Простые виды анализа данных без применения обучаемых алгоритмов

## ❑ Машинное обучение (МО) (Machine Learning, ML)

- Применение адаптивных методов, управляемых данными
- Установление закономерностей и построение моделей
- Применение полученных моделей в ситуациях, **схожих с известными**

## ❑ Искусственный интеллект (ИИ) (Artificial Intelligence, AI)

- Комплексное применение различных методов, основанных на разных данных, разных подходах, различных областях знания
- Получение универсальных комплексных моделей, пригодных для применения в широком круге разных ситуаций, в том числе новых

**Слабый ИИ** (Narrow AI) – эффективно решает **конкретный** круг задач

**Сильный ИИ** (General AI, AGI) – решает **разные** задачи

# ДОБЫЧА ДАННЫХ, МАШИННОЕ ОБУЧЕНИЕ И ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

---

## ❑ Добыча данных (Data Mining, DM)

- **Перцептивный уровень** автоматизированная обработка
- Простые виды анализа данных без применения обучаемых алгоритмов

## ❑ Машинное обучение (МО) (Machine Learning, ML)

- **Когнитивный уровень** управляемых данными
- Установление закономерностей и построение моделей
- Применение полученных моделей в ситуациях, **схожих с известными**

## ❑ Искусственный интеллект (ИИ) (Artificial Intelligence, AI)

- **Креативный уровень** различных методов, основанных на различных областях знания
- Получение универсальных комплексных моделей, пригодных для применения в широком круге работ: **Соответствие – условное!!!**

**Слабый ИИ** (Narrow AI) – эффективно решает **конкретный** круг задач

**Сильный ИИ** (General AI, AGI) – решает **разные** задачи

# ПРОГРАММА КУРСА

## План лекций

- Часть 1. Основы машинного обучения (2 лекции)
- Часть 2. Искусственные нейронные сети (ИНС) (4 лекции)
- Часть 3. Основы предобработки и анализа данных (3 лекции)
- Часть 4. Генетические алгоритмы (ГА) (3 лекции)
- Часть 5. Нечеткая логика (1 лекция)
- Часть 6. Другие алгоритмы анализа данных.  
Решение практических задач (2 лекции)

## План практических занятий

- Основные модели ИНС – наглядная демонстрация
- Современный инструментарий для практического решения задач анализа данных: языки R и Python
- Решение практических задач анализа данных (работа с ИНС, обработка изображений, обработка текстовой информации)
- Самостоятельная работа под руководством преподавателя

# ОБЩАЯ ИНФОРМАЦИЯ О КУРСЕ

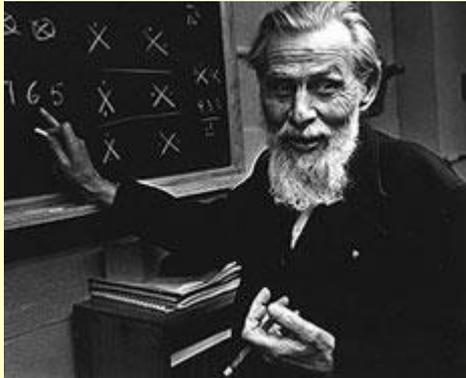
---

- ❑ Составные части курса:
  - 16 лекций
  - 12 практических занятий с домашними заданиями
  - Зачёт (с возможностью внесения в диплом для студентов)
- ❑ Два занятия в неделю по вторникам и пятницам в 19-00 (вторник – лекция, пятница – практическое занятие)
- ❑ Очные занятия с трансляцией в Интернете
- ❑ Выкладывание видеозаписей занятий на платформе Teach-In
- ❑ Для допуска к зачёту необходимо посетить не менее половины лекций
- ❑ На зачёте можно пользоваться любыми источниками

# МАТЕМАТИЧЕСКАЯ МОДЕЛЬ НЕЙРОНА

---

Предложена в 1943 г. У. Мак-Каллоком и У. Питтсом



**Warren Sturgis McCulloch**  
1898-1969

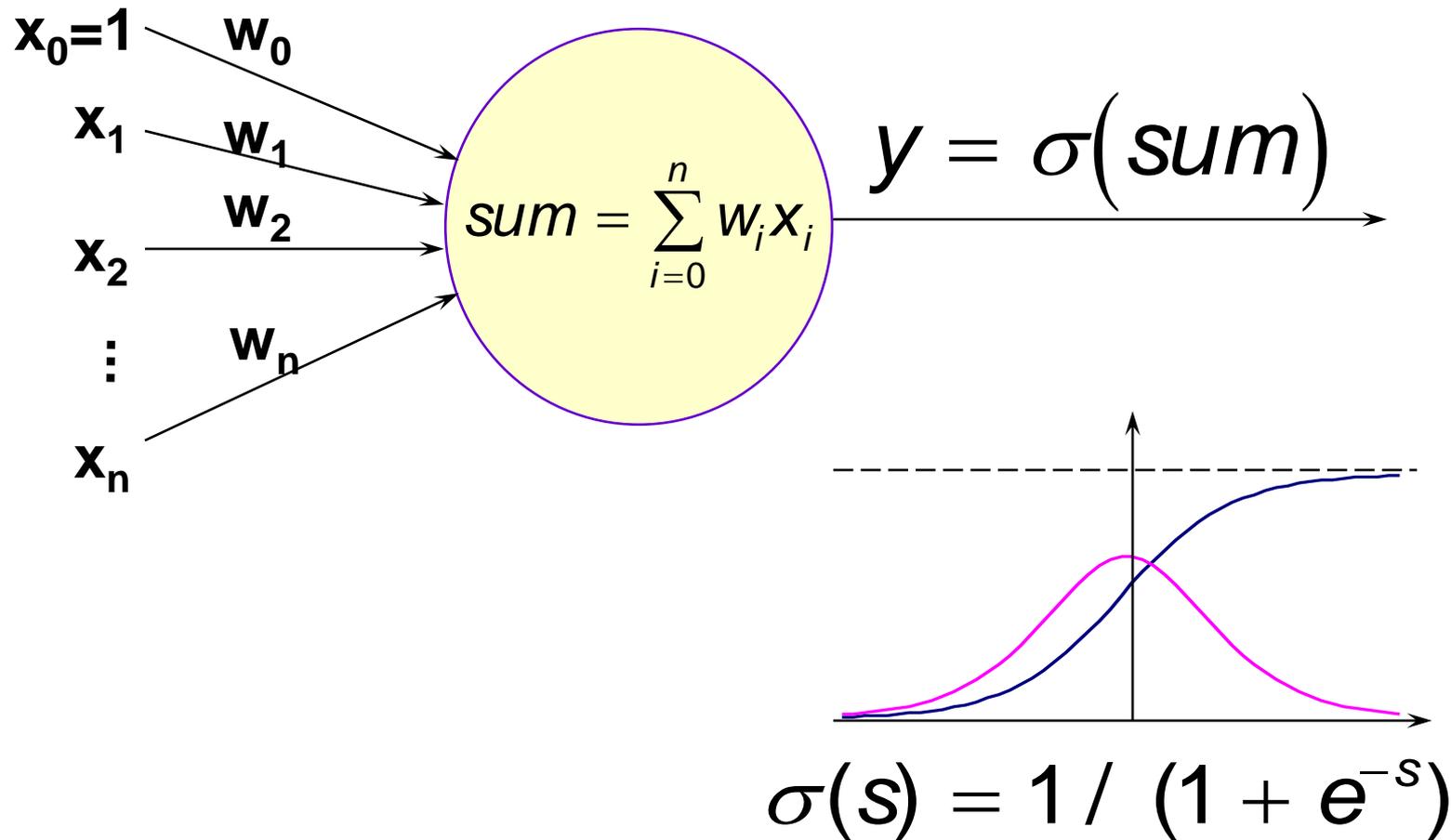


**Walter Pitts**  
1923-1969

1943, "A Logical Calculus of the Ideas Immanent in Nervous Activity".  
In: *Bulletin of Mathematical Biophysics*, Vol. 5, pp. 115–133.

**Нейрон как пороговое устройство с несколькими входами и одним бинарным выходом**

# МАТЕМАТИЧЕСКАЯ МОДЕЛЬ НЕЙРОНА

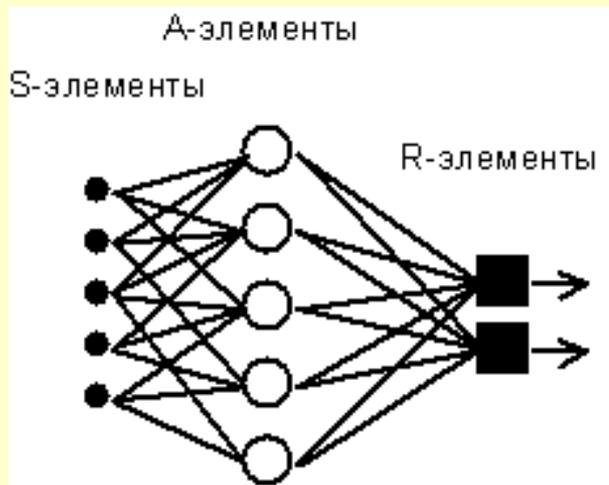


# ПЕРСЕПТРОН РОЗЕНБЛАТТА



Frank Rosenblatt (1928 – 1971)

1958, “The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain”, Cornell Aeronautical Laboratory, Psychological Review, v.65, No. 6, pp. 386-408.



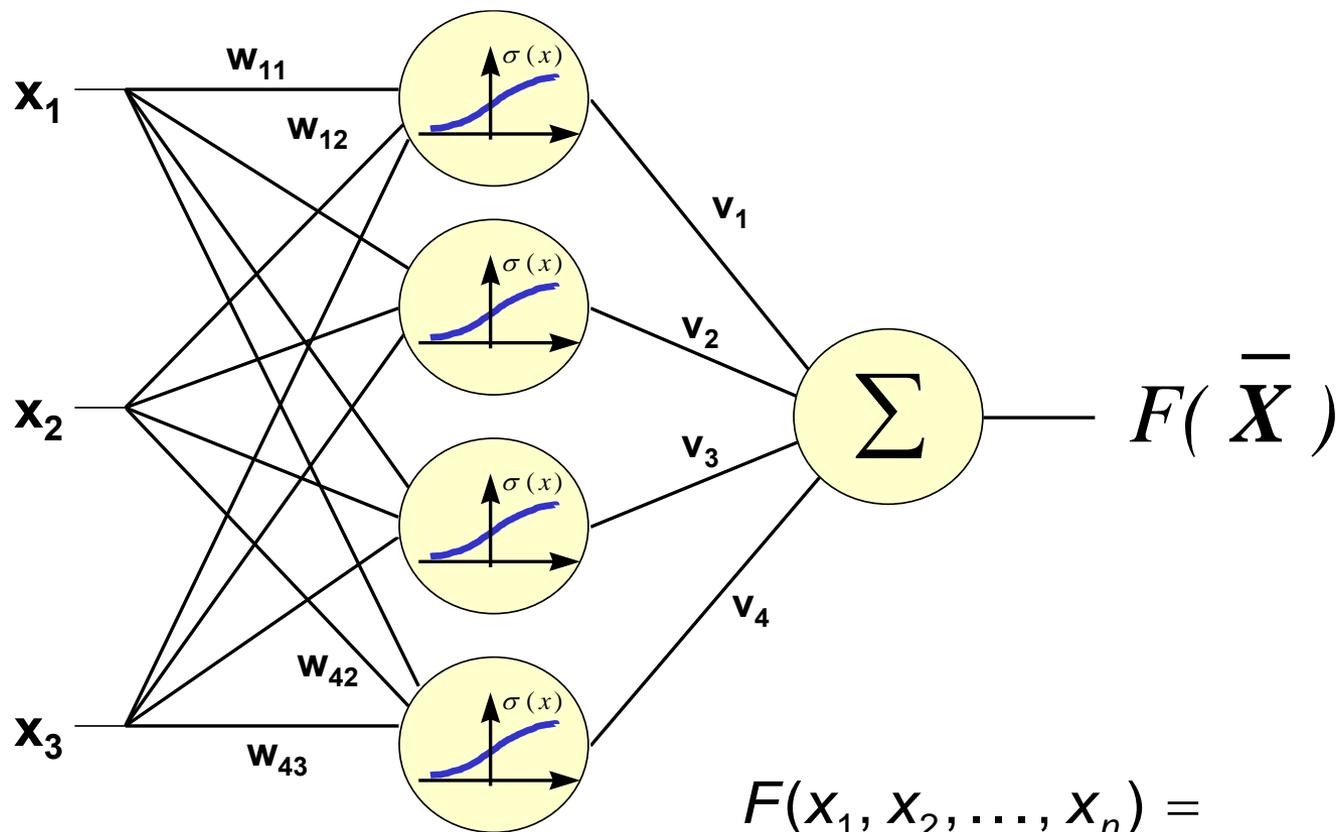
**S-элементы** – сенсорные клетки

**A-элементы** – ассоциативные, имеют переменные веса

**R-элементы** – формируют реакцию, имеют фиксированные веса

Персептрон MARK-1 распознавал буквы алфавита с матрицы 20x20

# МНОГОСЛОЙНЫЙ ПЕРСЕПТРОН



$$F(x_1, x_2, \dots, x_n) =$$

$$\sum_{i=1}^H v_i \cdot \sigma(w_{i1}x_1 + w_{i2}x_2 + \dots + w_{in}x_n + u_i)$$

Может ли персептрон реализовать достаточно сложную функцию?

# 13-Я ПРОБЛЕМА ГИЛЬБЕРТА



David Hilbert  
(1862 – 1943)



Владимир Игоревич  
Арнольд  
(1937 – 2010)

На Втором международном математическом конгрессе в 1900 г. сформулировал список из 23 нерешённых проблем математики

XIII проблема:

Верно ли, что существует непрерывная функция от трех переменных, которая не может быть представлена в виде композиции непрерывных функций от двух переменных?

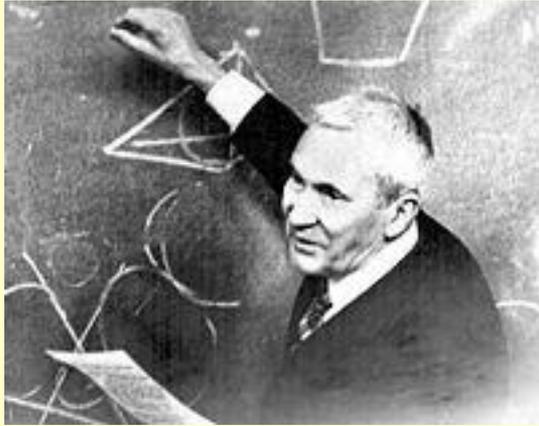
$$F(x,y,z) = xz + yz = S(M(x,z), M(y,z));$$

$$M(x,z) = xz, S(a,b) = a + b$$

В.И.Арнольд в 1957 г. показал, что:

Любая непрерывная функция трех переменных представляется в виде композиции непрерывных функций двух переменных

# ТЕОРЕМА КОЛМОГорова



Андрей Николаевич  
Колмогоров  
(1903-1987)

Любая непрерывная функция от  $n$  переменных  $F(x_1, x_2, \dots, x_n)$  может быть представлена в виде:

$$F(x_1, x_2, \dots, x_n) = \sum_{j=1}^{2n+1} g_j \left( \sum_{i=1}^n h_{ij}(x_i) \right),$$

где  $g_j$  и  $h_{ij}$  – непрерывные функции, причём  $h_{ij}$  не зависят от функции  $F$ . (1957)

Это означает, что для реализации функции многих переменных достаточно операций суммирования и композиции функций одной переменной, т.е. любую непрерывную функцию от нескольких переменных можно точно реализовать с помощью простой нейросети на основе персептрона с одним скрытым слоем

К сожалению, теорема ничего не говорит о том, как это сделать.

# МОДИФИЦИРОВАННАЯ ТЕОРЕМА КОЛМОГОРОВА

Пусть  $F(x_1, x_2, \dots, x_n)$  – любая непрерывная функция, определённая на ограниченном множестве, а  $\varepsilon$  – сколь угодно малое число.

Теорема. Существуют такое число  $H$ , набор чисел  $w_{ij}$ ,  $u_i$ , и набор чисел  $v_i$ , что функция

$$f(x_1, x_2, \dots, x_n) = \sum_{i=1}^H v_i \cdot \sigma(w_{i1}x_1 + w_{i2}x_2 + \dots + w_{in}x_n + u_i)$$

приближает данную функцию  $F(x_1, x_2, \dots, x_n)$  с погрешностью не более  $\varepsilon$  на всей области определения.

Это означает, что любую непрерывную функцию от нескольких переменных можно с любой наперёд заданной точностью реализовать с помощью обычного персептрона с достаточным количеством нейронов в единственном скрытом слое

1. Hornik, Stinchcombe, White. Multilayer Feedforward Networks are Universal Approximators. Neural Networks, 1989, v. 2, № 5.
2. Cybenko. Approximation by Superpositions of a Sigmoidal Function. Mathematical Control Signals Systems, 1989, v. 2, № 4.
3. Funahashi. On the Approximate Realization of Continuous Mappings by Neural Networks. Neural Networks, 1989, v. 2, № 3.

# ТРИ ВОЛНЫ ИНТЕРЕСА К НЕЙРОННЫМ СЕТЯМ

---

## Краткая историческая справка

### ❑ Первая волна (1957 – 1969)

- Теорема Колмогорова о представлении функции  $n$  переменных (1957)
- Персептрон Розенблатта (1958)
- Минский, Пейперт, «Персептроны» (1969)

### ❑ Вторая волна (1986 – середина 90-х)

- Алгоритм обратного распространения ошибки (1986)
- Теорема об универсальной аппроксимации персептроном (1989)

### ❑ Третья волна (середина 2000-х – наши дни)

- Глубокие НС
- Свёрточные НС (с 1990-х!)
- Рекуррентные сети с долгой краткосрочной памятью (LSTM)
- Генеративные сети
- Обучение с подкреплением (с 1980-х!)

# НЕКОТОРЫЕ УСПЕХИ ТРЕТЬЕЙ ВОЛНЫ

---

- ❑ Решена задача **дикторонезависимого** распознавания **речи**
- ❑ Решена задача **идентификации** (классификации) объектов на **изображениях** и в видеопотоке
- ❑ Решена задача **самообучения сложным играм** на основе знания лишь правил игры (шахматы, сёги, го)
- ❑ Достигнуты существенные успехи в машинном **анализе текстов**
- ❑ Достигнуты существенные успехи в машинном **переводе текстов**
- ❑ Достигнуты существенные успехи в **генерации новых данных**
- ❑ Достигнуты заметные успехи в общении **на естественном языке**

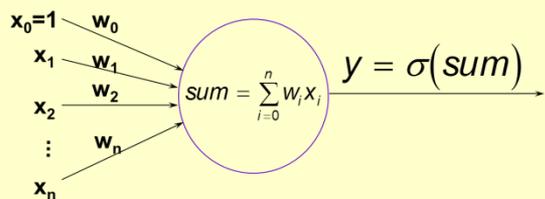
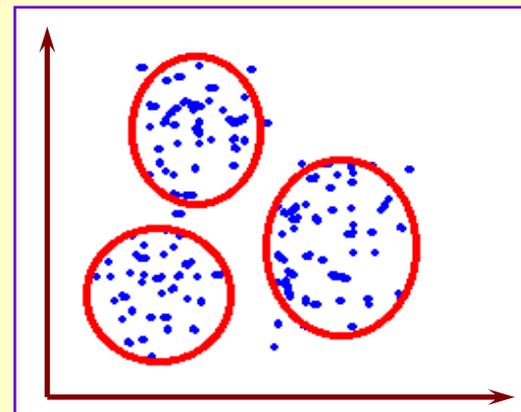
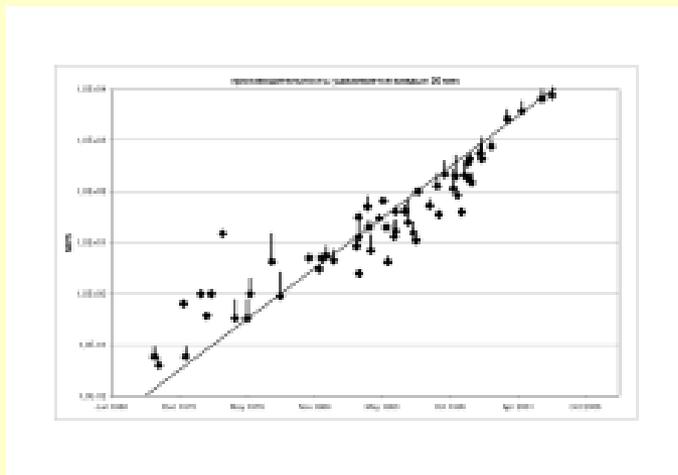
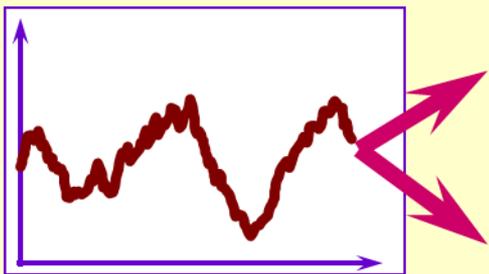
Полученные результаты основаны на доступности

- Существенно возросших **вычислительных мощностей** («закон Мура»)
- Практически неограниченного **объёма данных** некоторых типов
- Серьёзных **инвестиций**, связанных с уже достигнутыми успехами

# НЕЙРОННЫЕ СЕТИ: ИТОГИ ВВЕДЕНИЯ

---

- Нейронные сети – не аналог искусственного интеллекта...  
...но его важнейшая составная часть!
- Нейронные сети – мощная технология обработки данных
- К настоящему моменту позволили достичь уровня человека в задачах распознавания речи и изображений или превзойти его
- Для обучения нейронных сетей требуются объёмные массивы данных и большие вычислительные мощности
- С помощью нейросетевых методов могут решаться задачи, не решаемые традиционными методами
- Имеют обширные перспективы применения для решения разнообразных задач обработки и анализа данных



**Спасибо за внимание !**

