Моделирование

Лекция 2. Распознавание образов.

2018

Построение модели

$$E = \Phi(y, x, a, \xi)$$

Объекты моделирования в системах

- статические сущности (свойства, элементы, образы),
- динамические сущности (процессы, тенденции, поведение, сигналы),
- сущности, выражающие зависимости, в виде семантически различных структур (подсистема).

Какой у вас объект моделирования??

Два основных этапа моделирования

1. Построение модели объекта
2. Применение и анализ модели объекта

Две цели моделирования объектов

1. Для анализа (напр. в задаче распознавания образов и анализа сигналов)
2. Для синтеза (напр. в задаче оптимизации)

Какая у вас цель моделирования??

Цели моделирования

1. Модели анализа. Получить новое знание о свойствах моделируемой системы:
    1. О природе ее входов и выходов
    2. О зависимостях и свойствах, отражающих поведение системы
    3. О зависимостях и свойствах в структуре элементов и процессов
2. Модели синтеза. Получить новое знание, улучшающее моделируемую систему по выбранному критерию (минимизирующие или максимизирующие значения критерия)
    1. Оптимальные значения параметров системы (входов, процессов, элементов)
    2. Оптимальную структуру системы (количество элементов, количество связей)

Какие у вас цели моделирования??

Этапы построения модели анализа

1. Уяснение природы исследуемой системы, входов и выходов. Содержательное описание системы и ее архитектуры. Можно исследовать всю систему, а можно отдельный объект системы (входы, выходы, внутреннюю структуру, процесс, поведение). **Цель моделирования формулируется в отношении объекта исследования**

2. Формализация

- формальное описание входов, выходов объекта исследования (количественное или качественное моделирование входов и выходов), в зависимости от целей моделирования и исследовательских вопросов,
- **формальное описание зависимости выходов от входов (модель Y=H(X)),**
- **формальное описание зависимости переменных от времени (модель Y=F(T)),**
- **формальное описание критериев эффективности и целевой функции** в зависимости от целей моделирования;

Способ описания объектов

- евклидово пространство;
- объекты представляются точками в евклидовом пространстве их вычисленных параметров, представление в виде набора измерений;
- списки признаков;
- выявление качественных характеристик объекта и построение характеризующего вектора;
- структурное описание;
- выявление структурных элементов объекта и определение их взаимосвязи.

Количественные модели входов и выходов

- Детерминированные:
  - В виде одного значения
  - в виде векторов значений одного типа
  - В виде матрицы значений одного типа
  - В виде структуры гетерогенных значений
  - в виде функции
- Вероятностные, допускающие стохастическую неопределенность
  - В виде функции распределения вероятности для совокупности наблюдений (выборки из генеральной совокупности наблюдений) и ее характеристик:
    - среднее, медиана,
    - размах
    - вариация (среднеквадратическое отклонение или дисперсия)
  - В виде стохастических временных рядов, цифровых сигналов

Количественные модели входов и выходов

- Нечеткие, допускающие лингвистическую неопределенность
  - В виде множества функций (лингвистической переменной) для совокупности наблюдений X и ее характеристик:
    - количество термов,
    - виды функции распределения нечеткости термов и
    - ее носители в виде интервалов.
  - В виде нечетких временных рядов
- С неизвестной природой неопределенности в значениях
  - В виде множества однотипных значений, сформированных в процессе эволюции и адаптации к внешней среде

Распознавание образов – задача анализа

- **Теория распознавания образов** — раздел информатики и смежных дисциплин, развивающий основы и методы классификации и идентификации предметов, явлений, процессов, сигналов, ситуаций - объектов, - которые характеризуются конечным набором некоторых свойств и признаков.
- Такие задачи решаются довольно часто, например, при медицинской и технической диагностике, в криминалистике, в информационной безопасности, в задаче информационного поиска похожих объектов, **для оценки и ранжирования**, в системах доступа, для выявления аномальных объектов.

Понятие образа

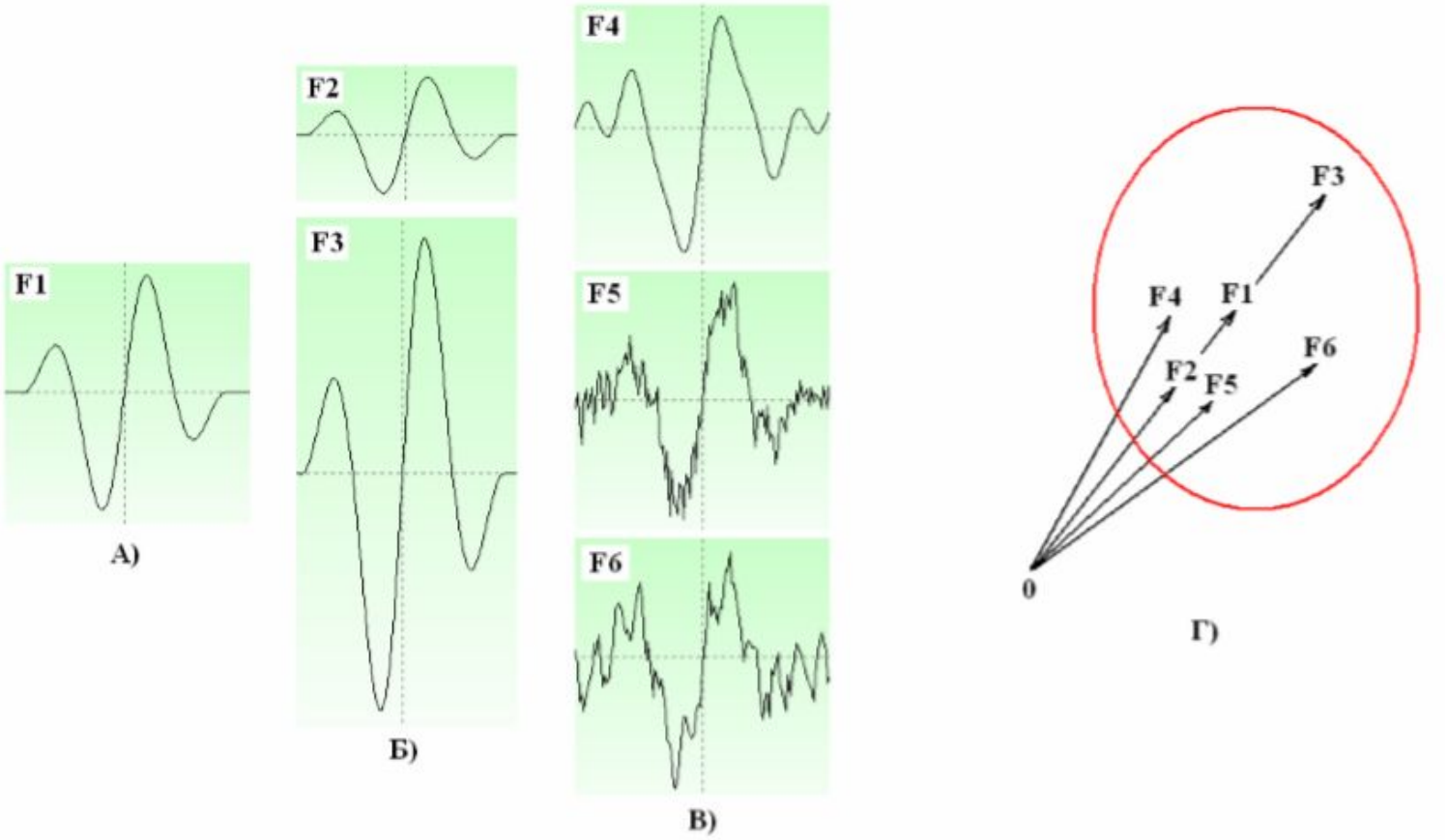


Рис.1 Иллюстрации к определению понятия «образ».

А) Ожидаемый сигнал – образ.

Б) Сигналы с различной амплитудой.

В) Примеры сигнала с искажениями.

Г) Множество сигналов в информационном пространстве, ассоциированных с образом.

Распознавание

Это способность живых организмов обнаруживать в потоке информации, поступающей от органов чувств, определённые объекты, закономерности, явления.

Оно может осуществляться на основе зрительной, слуховой, тактильной информации. Так, человек без труда может узнать другого знакомого ему человека, взглянув на него или услышав его голос.

Некоторые животные активно используют обоняние для узнавания других особей и поиска пищи.

Возможность распознавания опирается на схожесть однотипных объектов. Несмотря на то, что все предметы и ситуации уникальны в строгом смысле, между некоторыми из них всегда можно найти сходства по тому или иному признаку.

Отсюда возникает понятие классификации — разбиения всего множества объектов на непересекающиеся подмножества - классы, элементы которых имеют некоторые схожие свойства, отличающие их от элементов других классов.

И, таким образом, задачей распознавания является отнесение рассматриваемых объектов или явлений по их описанию к нужным классам.

Распознавание образов – задача анализа

- Проблема распознавания образов приобрела значение в условиях информационных перегрузок, когда человек не справляется с линейно-последовательным пониманием поступающих к нему сообщений.
- Проблема распознавания образов это сфера междисциплинарных исследований - в том числе в связи с работой по созданию искусственного интеллекта, а создание систем *распознавания образов* привлекает к себе всё большее внимание.

Формальная постановка задачи

Распознавание образов — это задача отнесения исходного экземпляра объекта *s** к некоторому классу (аномальности, поведения, типу).

1. При этом возможные классы *C* могут быть заданы или не заданы:

   $C = \varnothing$ *или* $C \neq \varnothing$

1. При этом доступно (или не доступно) множество других экземпляров *S* объектов: $S = \varnothing$ ***или*** $S \neq \varnothing$.
2. При этом известны (или не известны) существенные признаки *F*, характеризующих эти объекты: $F = \varnothing$ *или* $F \neq \varnothing$.

Сколько вариантов задач можно определить?

Формальная постановка задачи

Классическая постановка задачи распознавания образов определяется в виде задачи классификации:

- Дано множество объектов S в виде набора атрибутов. Множество объектов может быть представлено подмножествами похожих объектов, которые называются классами C.
- Имеется или извлекается: информация о классах C, информация об распознаваемом объекте s^*, принадлежность которого к определенному классу неизвестна.
- Требуется по описанию объекта s^* **построить модель** m для определения принадлежности s^* к некоторому классу c^*.

9 вариантов моделей распознавания образов

- Формальная постановка задачи как задачи поиска похожего образа.

$$\exists\, S = \{s_i\}, i = 1,2, \ldots Sk;\ s^* \notin S$$

1 вариант. *S ≠ ∅, C = ∅. Решение есть всегда.*

Требуется построить модель $m =< S, s^*, s', dist >$

для определения $s' \in S$, такого, что $dist(s', s^*) \rightarrow$ min.

Или (функциональное описание). Требуется определить

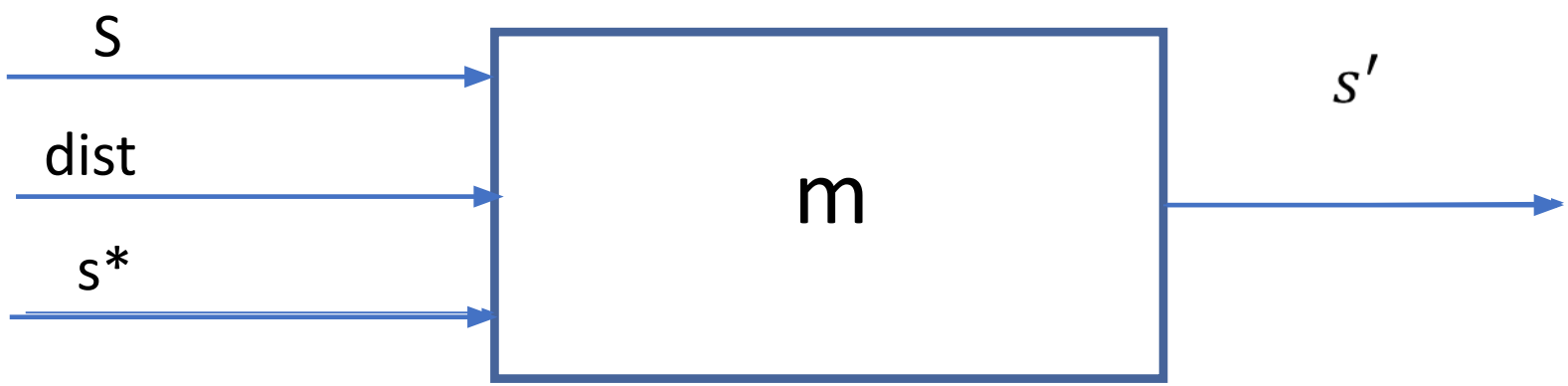
$$s' = m(s^*, S, dist),$$ такое, что

$$s' = s_z, \text{где } z = arg\min_i(g_i),\ g_i = dist(s^*, s_i),\ i = 1,2, \ldots Sk.$$

Тогда $s^* \backsimeq s'$.

Постройте черный ящик системы моделирования решения задачи распознавания как задачи поиска

Вариант 1.Модель распознавания образов как «черный ящик»



Изобразите архитектуру этой системы

Добавьте критерий эффективности распознавания

критерий эффективности распознавания

- $if\ dist(s^*, s') < Krit,\ then\ s^* \backsimeq s'$

$Else\ null$

Добавьте этот критерий в архитектуру системы моделирования

Варианты моделей распознавания образов

Формальная постановка задачи классификации

$$\exists\, S = \{s_i\}, i = 1,2, \dots Sk;\ s^* \notin S$$

2 вариант. $S \neq \emptyset$, $C \neq \emptyset$.

Предположим $\exists$ классов $C = \{c_j\}, j = 1,2, \dots Ck$; и определено отношение $R \subset S \times C$;

2.1.Требуется построить модель $m =< s^*, S, R, C, c^*, dist >$:

$$s^* \to c^*, c^* \in C$$

Или (функциональное описание). Требуется определить

$c^* = m(s^*, S, R, C, dist)$, такое, что

$c^* = c_z$, где $z = arg\min_i(g_i);\ g_i = dist(s^*, s_i),\ i = 1,2, \dots$Sk.

Постройте черный ящик

Постройте архитектуру

Варианты моделей распознавания образов

- 2.2.Требуется построить модель

$$m =< s^*, S, R1, C, F, c^*, dist >$$

$$m1: S \to F, m1: s^* \to f^*, \text{где } F = \{fn\}, n = 1,2, \dots, Nk; f^* \in F;$$

$$m2: f^* \to c^*, c^* \in C, R1 \subset F \times C$$

Функциональный вид. Требуется определить

$$c^* = m(s^*, S, R1, C, F, dist), \text{ такое, что}$$

$$c^* = c_z, \text{где } z = arg\min_i(g_i); g_i = dist(f^*, f_i), f_i = m1(s_i); f^* = m1(s^*), i = 1,2, \dots \text{Sk.}$$

Этапы построения модели распознавания образов. Эти этапы, что напоминают?

1) Выбор и анализ экземпляров образов(объектов) S, определение множества классов C и способов их описания (моделей).
2) Формирование словаря признаков F. Определении моделей признаков F (вероятностные, нечеткие, детерминированные, логические). Построение модели объектов: $m1: S \to F$
3) Описание классов на языке признаков или построение отношения: $R1 \subset F \times C$
4) Построение и анализ модели: $m2: F \to C$
5) Применение модели

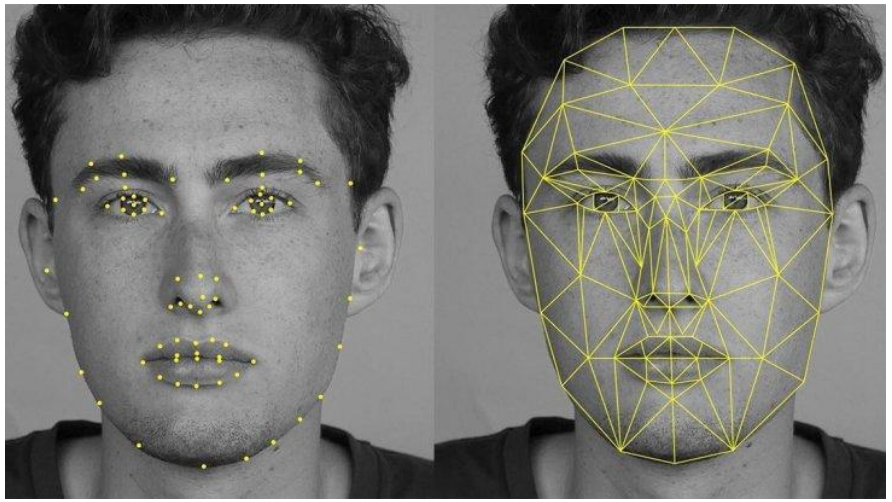
Эти этапы, что напоминают?

Стандартный процесс Data Mining: CRISP-DM

1) Business understanding. Выбор и анализ экземпляров образов(объектов) S, определение множества классов C и способов их описания (моделей).
2) Data understanding. Формирование словаря признаков F. Определении моделей признаков F (вероятностные, нечеткие, детерминированные, логические). Построение модели объектов: $m1: S \to F$
3) Data Preparation. Описание классов на языке признаков или построение отношения: $R1 \subset F \times C$
4) Modeling&Evaluation.Построение и анализ модели: $m2: F \to C$
5) Deployment. Применение модели

Пример распознавания образов

- Распознавание лица — последний тренд в авторизации пользователя. Apple использует Face ID, OnePlus — технологию Face Unlock. Baidu использует распознавание лица вместо ID-карт для обеспечения доступа в офис, а при повторном пересечении границы в ОАЭ вам нужно только посмотреть в камеру.



Модель объекта: статическое изображение

- Наиболее часто в задачах распознавания образов рассматривается задача поиска изображения. Образ – это изображение. Оно представимо как функция на плоскости. Если рассмотреть точечное множество на плоскости T, где функция f(x,y) выражает в каждой точке изображения его характеристику — яркость, прозрачность, оптическую плотность, то такая функция есть формальная запись изображения. Каждая f(x,y) может быть представлена ВР.

- Множество же всех возможных функций {f(x,y)} на плоскости T есть модель статического изображения. Вводя понятие сходства между образами можно поставить задачу поиска как задачу распознавания. Конкретный вид такой постановки сильно зависит от последующих этапов при распознавании в соответствии с тем или иным методом.

FaceNet

- FaceNet — нейронная сеть, которая учится преобразовывать изображения лица в компактное евклидово пространство, где эвклидово расстояние соответствует мере схожести лиц. Проще говоря, чем более похожи лица, тем они ближе.
- Здесь Ссылка на Гитхаб, кому нужен код
- FaceNet использует особую функцию потерь называемую TripletLoss. Она минимизирует расстояние между искомым изображением и изображениями, которые содержат похожую внешность, и максимизирует расстояние между разными.

FaceNet

- FaceNet (на python) — сиамская сеть. Сиамская сеть — тип архитектуры нейросети, который обучается на входных данных. То есть, позволяет научиться понимать какие изображения похожи, а какие нет.
- Сиамские сети состоят из двух идентичных нейронных сетей, каждая из которых имеет одинаковые точные веса. Сначала, каждая сеть принимает одно из двух входных изображений в качестве входных данных. Затем выходы последних слоев каждой сети отправляются в функцию, которая определяет сходство.
- В FaceNet это делается путем вычисления расстояния между двумя выходами.

Алгоритм распознавания лиц

- Подготовка базы данных личностей $S = \{s_i\}, i = 1,2, \ldots$, из которых сеть будет распознавать новое лицо s^*. *FaceNet достаточно мощна, чтобы распознать человека по одной фотографии.*
- Определение словаря признаков $F = \{f_i\}, i = 1,2, \ldots.$; в виде 128 float чисел. Преобразование каждого изображения в набор признаков, $f_i = m1(s_i)$.
- Установка параметров. Инициализация сети размерности (3, 96, 96). Это означает, что картинка передается в виде трех каналов RGB и размерности 96x96 пикселей.
- Загрузка и преобразование нового изображения $f^* = m1(s^*)$. Функция обрабатывает изображение, возвращает закодированное изображение тоже в виде 128 float чисел.
- Вычисление расстояний $g_i = dist(f^*, f_i),\ i = 1,2, \ldots,$
- Если $\min_i(g_i)$<krit, то определение похожего изображения

$$s^* = s_z, \text{где } z = arg\min_i(g_i)$$

иначе отказ от распознавания.

Siamese network

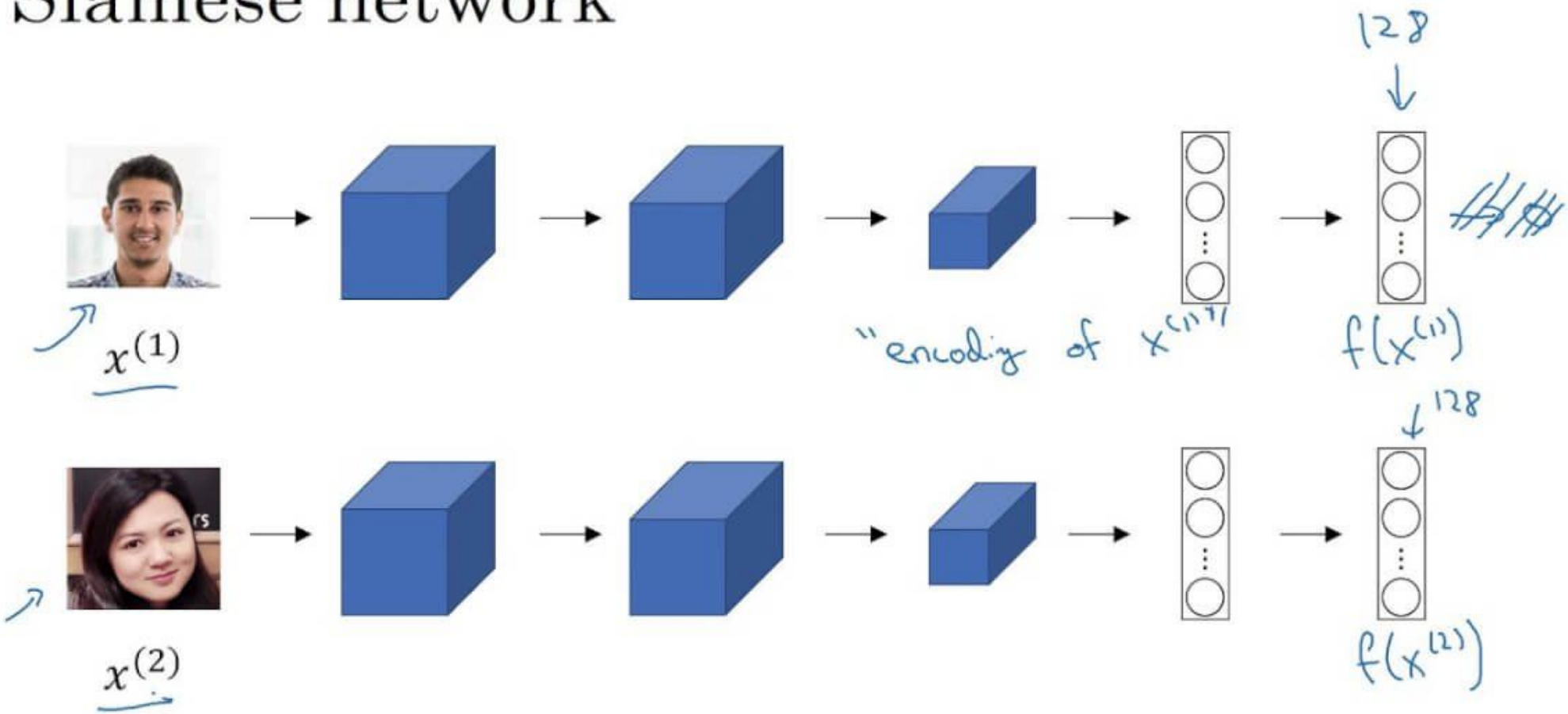
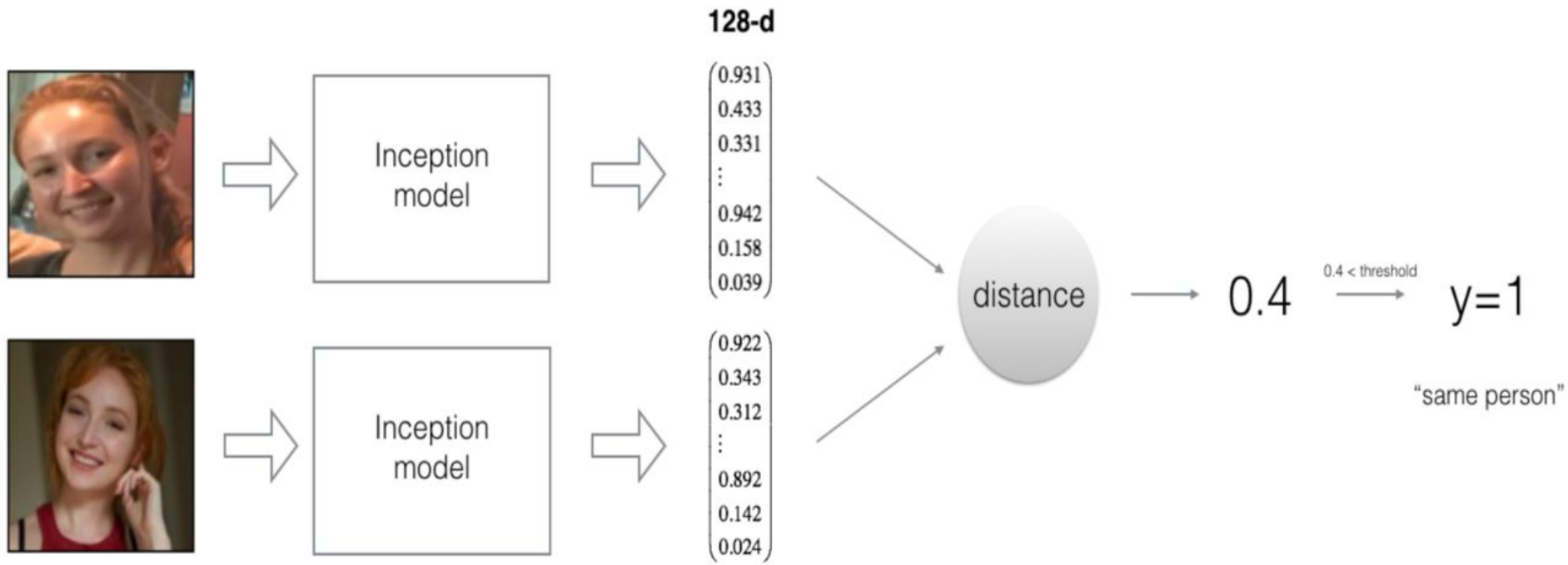


Figure 8. Face normalization for people in the LFW dataset [22]. Top to bottom: input photographs, result of our method using FaceNet


128-d
Inception model
Inception model
0.931
0.433
0.331
0.942
0.158
0.039
0.922
0.343
0.312
0.892
0.142
0.024
distance
0.4
0.4 < threshold
y=1
"same person"

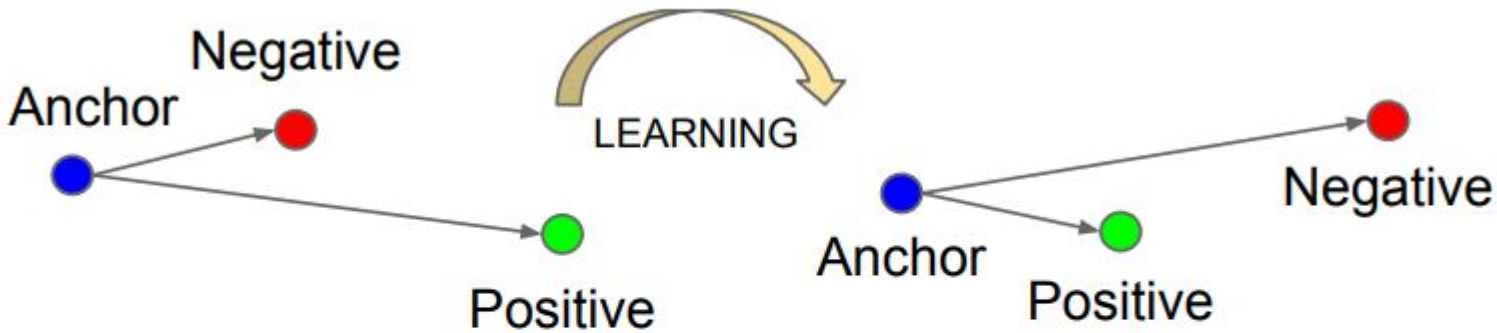


Figure 3. The **Triplet Loss** minimizes the distance between an *anchor* and a *positive*, both of which have the same identity, and maximizes the distance between the *anchor* and a *negative* of a different identity.

3 D распознавание

- **Обнаружение**: получение снимка при помощи цифрового сканирования существующей фотографии (2D) или видео для получения живой картинки субъекта (3D).
- **Центровка**: определив лицо, система отмечает положение головы, размер и позу.
- **Измерение**: система измеряет кривые на лице с точностью до миллиметра и создает шаблон.
- **Репрезентация**: система переводит шаблон в уникальный код. Этот код задает каждому шаблону набор чисел, представляющих особенности и черты лица.
- **Сопоставление**: если снимок в 3D и база данных содержит трехмерные изображения, сопоставление пройдет без изменений снимка. Но если же база данных состоит из двумерных снимков, трехмерное изображение раскладывается на разные составляющие (словно сделанные под разными углами двумерные снимки одних и тех же черт лица), и они конвертируются в 2D-изображения. И затем находится соответствие в базе данных.
- **Верификация или идентификация**: в процессе верификации снимок сравнивается только с одним снимков в базе данных (1:1). Если целью же стоит идентификация, снимок сравнивается со всеми снимками в базе данных, что приводит к ряду возможных совпадений (1:N). Применяется тот или иной другой метод по необходимости.

- Российский банк «Открытие» представил собственное уникальное решение, разработанное под технологическим брендом Open Garage: перевод денег по фотографии в мобильном приложении «Открытие.Переводы». Вместо того чтобы вбивать номер карты или телефона, достаточно просто сфотографировать человека, которому нужно сделать перевод. Система распознавания лиц сравнит фото с эталонным (делается, когда банк выдает карту) и подскажет имя и фамилию. Останется только выбрать карту и ввести сумму. Что особенно важно, клиенты сторонних банков также могут использовать эту функцию для переводов клиентам «Открытия» — отправитель переводов может пользоваться картой любого российского банка.

- В задаче распознавания объектов существует несколько серьезных проблем:
- сильная зависимость от начальных параметров (необходимо знать достаточно много о тех объектах, которые собираемся искать в видеопотоке)
- идентификация объекта (различные похожие объекты, могут быть похожи, например мотоцикл и велосипед)
- внутриклассовая изменчивость

Направления в распознавании образов

- Развитие методов обучения для распознавания в условиях неполноты информации;
- Изучение способностей к распознаванию, которыми обладают живые существа(человек, птицы, млекопитающие, насекомые), объяснение и моделирование их;
- Развитие теории и методов построения устройств, предназначенных для решения отдельных задач в прикладных целях.

Обработка цифровых сигналов

- Анализ сигналов[править | править код]
- **Анализ сигналов** — извлечение информации из сигнала, например, выявление и обособление интересующих особенностей в экспериментально полученной функции. Существуют корреляционный анализ сигналов и спектральный анализ сигналов.
- **Спектральный анализ сигналов.** Вероятно, наиболее распространённым видом анализа сигналов является преобразование Фурье временного сигнала в частотную область для получения спектра частот сигнала. Для анализа сигналов, в частности для получения временно-частотного представления также могут быть использованы другие преобразования, такие как оконное преобразование Фурье и непрерывное вейвлет-преобразование. Другие разновидности анализа сигналов включают подбор параметров, например поиск наилучшего приближения методом наименьших квадратов.

- Для аналоговых сигналов обработка может включать усиление и фильтрацию, модуляцию и демодуляцию. Для цифровых сигналов также осуществляется сжатие, обнаружение и исправление ошибок и пр.
- Аналоговая обработка сигналов — для неоцифрованных сигналов, таких как радио-, телефонные или телевизионные сигналы.
- Цифровая обработка сигналов — для оцифрованных сигналов. Обработка осуществляется с помощью цифровых схем, в том числе с помощью программных решений.
- Статистическая обработка сигналов — включает анализ и получение информации из сигналов, основываясь на их статистических свойствах.
- Обработка звука — для электрических сигналов, представляющих звук, например, музыку.
- Распознавание речи — для обработки и интерпретации речи.
- Обработка изображений — в цифровых камерах, компьютерах и подобных системах.
- Обработка видео — для обработки движущихся изображений.

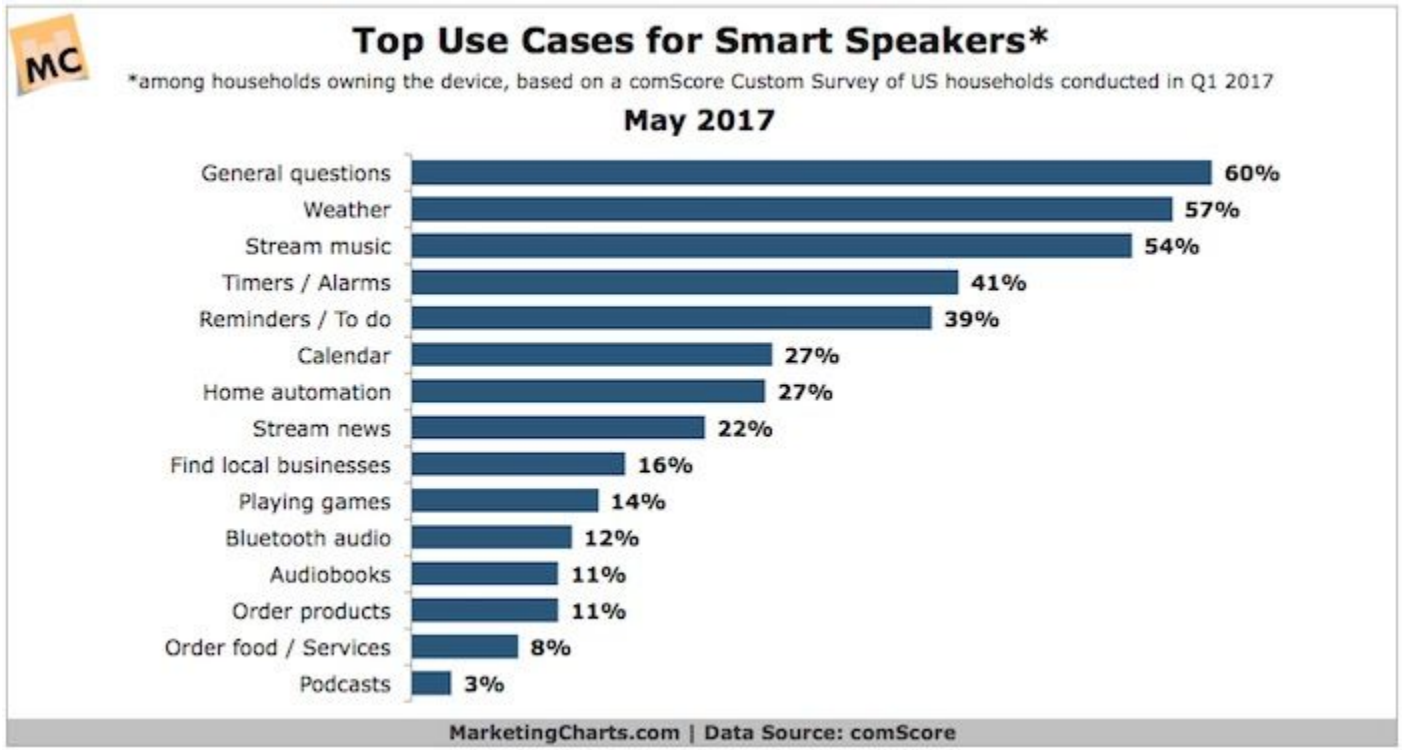
- Обнаружение сигнала — задача обнаружения сигнала на фоне шумов и помех.
- Различение сигнала — задача распознавания сигнала на фоне других сигналов, с подобными характеристиками.

распознавание речи, изображений, распознавание образов, подавление шумов, адаптивные антенные решётки.

Распознавание речи — процесс преобразования речевого сигнала в цифровую информацию (например, текстовые данные). Обратной задачей является синтез речи.

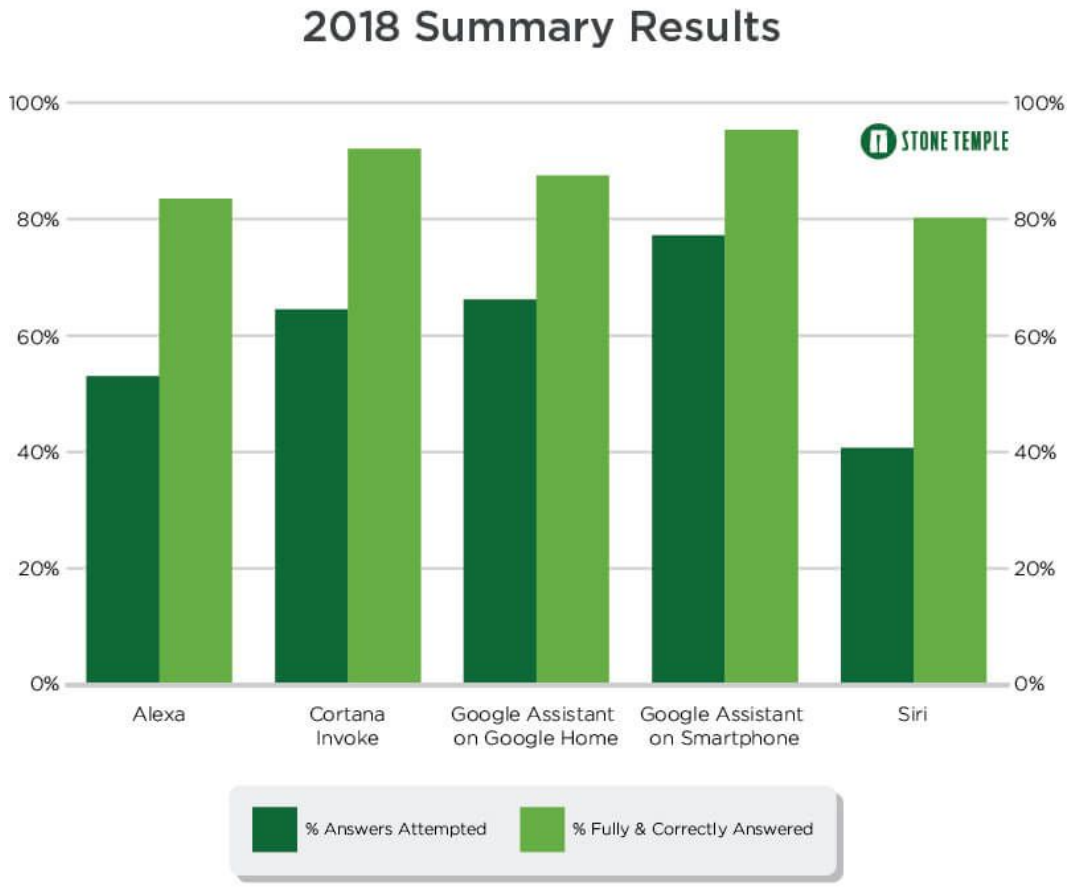
- Распознавание речи — одна из самых интересных и сложных задач искусственного интеллекта. Здесь задействованы достижения весьма различных областей: от компьютерной лингвистики до цифровой обработки сигналов.

цифровые голосовые помощники



https://www.stonetemple.com/digital-personal-assistants-study/

1. Alexa
2. Cortana Invoke
3. Google Assistant on Google Home
4. Google Assistant on a Smartphone
5. Siri



- Knowledge Graph – когда люди ищут информацию, они ищут не только сайты на которых может быть эта информация, но и ответы нужные вопросы. Knowledge Graph решает эти задачи.
- **Knowledge Graph** (рус. *Сеть знаний*; дословно *Граф знаний*) — семантическая технология и база знаний, используемая Google для повышения качества своей поисковой системы с семантическо-розыскной информацией, собранной из различных источников. Граф знаний был добавлен в поисковую систему Google в 2012 году, сначала в США, о чём было объявлено 16 мая2012 года[1]. Граф знаний предоставляет структурированную и подробную информацию о теме в дополнение к списку ссылок на другие сайты. Цель состоит в том, что пользователи смогут использовать эту информацию для решения своих запросов без необходимости перехода на другие сайты и сбора информации самостоятельно[2][3].
- По словам представителей Google, база знаний Knowledge Graph сформирована из многих источников, включая CIA World Factbook, Freebase и Wikipedia. По функциям Knowledge Graph похож на поисковики, дающие ответ, такие как Ask Jeeves, Wolfram Alpha, Linked data и DBpedia. С 2012 в семантической сети содержится более 570 миллионов объектов и более 18 млрд фактов и отношения между этими различными объектами, которые используются, чтобы понять смысл запроса.

Структура теста

- Мы собрали набор из 4 952 вопросов, чтобы спросить каждого личного помощника. Мы задали каждому из пяти участников одинаковый набор вопросов и отметили множество различных возможных категорий ответов, в том числе:

https://roem.ru/31-07-2018/272581/google-assistant-rus/

- 2018 г. Голосовой помощник Google Assistant, появившийся в 2016 году, стал доступен на русском языке, рассказали представители компании. Чтобы начать работу с Ассистентом на Android, нужно сказать «Окей, Гугл» или долго удерживать кнопку главного экрана, а чтобы начать работу в iOS — скачать приложение в App Store.
- Также по словам представителей Google, помощник отвечает за секунду в формате обычного диалога. С помощью Ассистента можно звонить, отправлять сообщения, например, через приложения, управлять устройством (устанавливать таймер, будильник, включать плейлист, делать фотографии), искать места,узнавать погоду, переводить речь на другой язык. Кроме того, у помощника есть виртуальная личность, он может отвечать на философские вопросы и рассказывать о себе.
- **Google Assistant** (Google Ассистент) — облачный сервис персонального ассистента, разработанный компанией Google может использоваться в смартфонах, также он включен в Google Allo — приложение для мгновенного обмена сообщениями, Google Home — умный голосовой Wi-Fi динамик для управления вашим домом, Android Wear — умные часы от Google.

- В октябре 2017 года «Яндекс» запустил голосового помощника «Алису» в мобильном приложении, который в разговоре также не ограничивается шаблонными ответами. Она может найти информацию в интернете, подсказать погоду, открыть приложение или сайт и просто общаться.
- Алиса — первый голосовой помощник, который в разговоре не ограничивается шаблонными ответами. Она может найти информацию в интернете, подсказать погоду, открыть приложение или сайт и просто поболтать. Мобильное приложение «Яндекса» обновилось автоматически, поэтому, чтобы воспользоваться голосовой помощницей, отдельно ничего скачивать не нужно. В перспективе владельцы сторонних приложений или сервисов смогут добавить в них возможности Алисы, пообещал «Яндекс».

- «Алиса» разговаривает голосом российской актрисы дубляжа Татьяны Шитовой, дублировавшей на русский язык большинство ролей актрисы Скарлетт Йоханссон, в том числе роль виртуальной помощницы Саманты из фильма «Она»[11]. Синтезатор речи использует специально подготовленные записи Шитовой[13].
- Многие особенности личности «Алисы» заданы набором фраз, сочинённых редакторами «Яндекса». Одним из авторов персонажа «Алисы» стал журналист и писатель Владимир Гуриев. Однако создатели подчёркивают[1], что «Алиса» не ограничивается набором заранее заданных редакторских ответов: нейронная сеть помощницы обучена на большом массиве русскоязычных текстов, в том числе сетевых диалогов. Это сказалось на характере программы: некоторые пользователи сталкиваются[12] с тем, что она отказывается отвечать на вопросы или дерзит. Разработчики «Алисы» постоянно наблюдают за её поведением и корректируют его[14].
- «Алиса» умеет отвечать эмоционально: например, в зависимости от контекста она может проявлять жизнерадостность или грустить

- Мы, наверное, первые в мире пытаемся сделать вот что: мы тоже используем редакторские ответы на вопросы, но добавляем специальную нейронную сеть, обученную на свободную беседу. Она может подобрать ответ или втянуть пользователя в болтовню ни о чем. В этом, наверное, кардинальное отличие,потому что людям, помимо поиска каких-то фактов, иногда хочется с кем-то поболтать. Алиса уже сейчас способна поболтать и будет в этом только совершенствоваться, — рассказал руководитель разработки голосовых технологий «Яндекса» и один из главных идеологов «Алисы» Денис Филиппов в интервью «Ленте.ру».
- В мае стало известно, что «Яндекс» работает над голосовой помощницей «Алиса», аналог Siri в iOS. Сегодня «Ведомости» сообщили, что компания разрабатывает аналог Amazon Echo и Google Home — «умную» голосовую колонку. Об этом изданию рассказали представитель крупной медиакомпании и представители крупных звукозаписывающих лейблов.
- В «Яндексе» рассказали, что прототип колонки уже существует и называется Artificial Intelligence speaker. «Мы уверены, что за голосовым интерфейсом будущее, у нас наработаны необходимые технологии: распознавание и синтез речи, продвинутые диалоговые системы, искусственный интеллект», — прокомментировал представитель компании.

- **Корта́на** — виртуальная голосовая помощница с элементами искусственного интеллекта от Microsoft для Windows Phone 8.1[1], Microsoft Band[2][3], Windows 10, Android, Xbox One и iOS.
- Впервые была продемонстрирована во время конференции Build в Сан-Франциско 2 апреля 2014 года[4][5]. Кортана была названа в честь героини серии компьютерных игр Halo — голос помощницы в версии для американского рынка принадлежит Джен Тейлор, которая также озвучивала Кортанув оригинальной игре

- Персональная помощница Кортана призвана предугадывать потребности пользователя. При желании ей можно дать доступ к вашим личным данным, таким как электронная почта, адресная книга, история поисков в сети и т. п. — все эти данные она будет использовать для упреждения ваших нужд. Кортана заменит стандартную поисковую систему и будет вызываться нажатием кнопки «Поиск». Нужный запрос можно как напечатать вручную, так и задать голосом. Необходимую информацию она будет находить, опираясь на результаты поиска в системе Bing[7], Foursquare и среди личных файлов пользователя. Также виртуальный ассистент не лишена чувства юмора: она может поддерживать с вами беседу, петь песенки и рассказывать анекдоты. Она заранее напомнит вам о запланированной встрече, дне рождения друга и других важных событиях. Кортана сообщит, если ваш авиарейс отменили или на дорогах много пробок.[8] Её интерфейс имеет очень гибкие настройки конфиденциальности, позволяющие пользователю самому определять, какого рода информацию предоставлять виртуальному ассистенту. По словам разработчиков, таким уровнем контроля не может похвастаться ни Siri, ни Google Now

- Голосовая помощница Кортана интегрирована в Windows 10, но не доступна на русском. Она не выступает в роли отдельного приложения, а интегрирована в поиск Windows 10.

цифровые голосовые помощники

- Алекса, известная просто как Алекса, является виртуальным помощником, разработанным Amazon, впервые используемым в Amazon Echo и Amazon Echo Dot, разработанном Amazon Lab126.
- 30 ноября 2016 года Amazon объявила о том, что они сделают технологию распознавания речи и технологию обработки естественного языка за Alexa доступной для разработчиков под названием **Amazon Lex** . Эта новая услуга позволит разработчикам создавать свои собственные чаты,
- Amazon позволяет разработчикам создавать и публиковать навыки для Alexa, используя Alexa Skills Kit. [47] Эти сторонние разработанные навыки, которые были опубликованы, доступны на устройствах, поддерживающих Alexa. Пользователи могут использовать эти навыки, используя приложение Alexa.
- Доступен «Smart Home Skill API» [48] , предназначенный для использования производителями оборудования, чтобы позволить пользователям управлять интеллектуальными домашними устройствами . [49]
- Большинство навыков запускают код почти полностью в облаке , используя сервис Amazon AWS Lambda . [50]
- В апреле 2018 года Amazon запустила Blueprints, инструмент для людей, чтобы создавать навыки для личного использования.

- Алекса, известная просто как Алекса, является виртуальным помощником, разработанным Amazon, впервые используемым в Amazon Echo и Amazon Echo Dot, разработанном Amazon Lab126.
- 30 ноября 2016 года Amazon объявила о том, что они сделают технологию распознавания речи и технологию обработки естественного языка за Alexa доступной для разработчиков под названием **Amazon Lex** . Эта новая услуга позволит разработчикам создавать свои собственные чаты,
- Amazon позволяет производителям устройств интегрировать голосовые возможности Alexa в свои собственные продукты с помощью службы Alexa Voice Service (AVS), облачной службы, которая предоставляет API для взаимодействия с Alexa. Продукты, созданные с использованием AVS, имеют доступ к растущему списку Alexa, включая все навыки Alexa. AVS обеспечивает автоматическое распознавание речи на основе облачных вычислений (ASR) и понимание естественного языка (NLU). Для компаний, которые хотят интегрировать Alexa в свои продукты, нет платы за использование AVS

Amazon Echo Dot

•Благодаря семи микрофонам, технологии шумоподавлению, Echo слышит вас в любом направлении - даже во время воспроизведения музыки

- Просто попросите Alexa проверить свой календарь, погоду, трафик и спортивные результаты, управлять списками дел и покупок, контролировать умный дом, термостаты, гаражные ворота, контролировать ваш телевизор, заказать Убер, пиццу и многое другое.

- Переводы с помощью голосовых команд пока доступны только для владельцев iOS-устройств. «Бинбанк» реализовал опцию на основе технологий Yandex SpeechKit, а «Открытие» —сервиса Siri.
- Чтобы воспользоваться таким способом перевода средств и оплаты услуг пользователю нужно встроить поддержку платежных интентов SiriKit или Yandex SpeechKit в мобильное приложение банка и создать шаблон с нужными реквизитами для переводов. После этого при совершении операций в мобильном банке достаточно будет дать голосовую команду — например, о переводе средств.
- О планах внедрения такой опции заявили и в Промсвязьбанке. Представитель организации пообещал, что банк запустит подобный сервис на основе Siri до конца года.
- Представитель «Открытия» отметил, что новая опция позволяет осуществить более быстрый перевод средств между клиентами. В Уральском банке реконструкции и развития предупредили, что нововведение может нести в себе и риски. Злоумышленники могут воспользоваться уязвимостью сервиса Siri и получить доступ к средствам пользователей.

- «Умные» голосовые помощники базируются на архитектуре нейронных сетей и технологии машинного обучения. При этом надо понимать, что в мозгу человека около 86 млрд нейронов, а в современном ИИ их всего несколько сот тысяч. Если посчитать количество нейронов в нервной системе различных животных, то выяснится, что, как отметил основатель и глава компании ABBYY Дэвид Ян, сейчас искусственный интеллект глупее пчелы.
- Для каждого языка требуется обучение программ распознавания, когнитивной обработки распознанного текста и синтеза речи по сформированному тексту. И поскольку русский язык в мире востребован существенно меньше, чем английский, далеко не все компании готовы тратить время и средства на разработки в этом направлении. Не хватает и технологической базы, например, мал корпус русского языка, нужный для алгоритмов машинного обучения», — считает аналитик агентства MForum Analytic Алексей Бойко.

- Все вышесказанное не значит, что современные голосовые помощники на базе ИИ бесперспективны и бесполезны. Уже сейчас на базе ИИ можно создавать более или менее удобные сценарии работы голосовых помощников, совмещая их с визуальными. Именно в этой логике были разработаны представленные в январе 2018 года умные колонки, оснащенные дисплеем, — например, вы запрашиваете у голосового помощника рецепт блюда, а он выводит его на экран.
- Таких сценариев уже сейчас можно придумать немало, особенно тех, которые связаны с распознаванием образов — это именно та у область, где ИИ достиг наибольшего прогресса.
- Периодически различные исследователи проводят тесты ИИ, выясняя кто же из них умнее, но тут важны даже не абсолютные цифры, а то что
- IQ искусственного интеллекта удваивается примерно каждые два года.

- Речевые технологии распознают, анализируют и синтезируют голос человека. Имитация речи, восприятие смысла фраз, конвертация речи в текст, работа с голосом как с биометрической характеристикой – все это разные типы речевых технологий. Этот раздел компьютерной науки считается одним из сложнейших, поскольку находится на стыке нескольких комплексных дисциплин: лингвистики, математики и программирования.
- **Где нужны технологии распознавания речи**
- Прежде всего, технологии распознавания речи используются для голосового набора команд, в ситуациях, при которых говорить намного проще, чем печатать. Распознавание речи применяется в системах интерактивного речевого самообслуживания, когда, например, на телефонные звонки в компании отвечает робот, который может разобраться со стандартными вопросами из области поддержки. Еще одно применение технологий распознавания голоса — диктовка текстов, своеобразный автоматический секретарь. Наконец, все чаще появляются системы с голосовым управлением любой техникой, например «умный дом», или автомобилем. Область применения будет в ближайшее время непрерывно расширяться в связи как с несомненным удобством для пользователя голосовых команд, так и с прогрессом в точности распознавания речи.

- Речевые технологии демонстрируют впечатляющие результаты в разных сферах. Так, в области трансформации речи в текст тон продолжает задавать Dragon Natural Speaking.
- Новейшая 13-я версия этого ПО, помимо стандартной функции диктовки, понимает голосовые команды для управления компьютером, например, открывает программы или переключает окна в браузере. Это ПО может конвертировать в текст подкасты и аудиоклипы или с помощью одной команды вставлять в письмо электронную подпись.
- Распознавание по голосу – другое обширное направление развития речевых технологий, связанное с идентификацией и верификацией личности. Они подразделяются на зависимые от текста, когда человеку необходимо назвать определенное слово или повторить фразу, и не зависимые от текста, когда идентификация производится просто на основе речи.
- Голос считается менее надежным биометрическим параметром, чем, например, отпечатки пальцев.

Используемые технологии и методы

- **Распознавание речи –** это процесс преобразования речевого сигнала в цифровую информацию. Именно этот процесс позволяет организовать речевое управление компьютером или программой и осуществить ввод текста с микрофона. Эта технология позволяет создавать голосовое командное управление ПК, системы диктовки текста или средства идентификации по образцу речи.
- **Понимание речи –** процесс, при котором компьютер или программа воспринимает смысл сказанного. Такая возможность стала реальной благодаря технологии искусственного интеллекта (ИИ). Благодаря ИИ речевой интерфейс может не только дублировать голосовые команды.

- **Голосовой поиск** (или голосовая команда) – функция поиска информации без использования клавиатуры. Пользователь произносит фразу, а приложение распознает текст, выполняет поиск и предоставляет результаты на странице поисковой выдачи. Голосовой поиск, в отличие от классического, взаимодействует с пользователем с помощью диалогов, а не посредством ключевых слов и фраз.
- ***Голосовой интерфейс*** – это программный продукт, который при помощи голосовой или речевой платформы позволяет взаимодействовать пользователю и компьютеру, запуская автоматизированные процессы. Задача таких интерфейсов – распознать и генерировать голос человека.
- Голосовые интерфейсы удобны, когда вводить текст сложно или неудобно. Например, во время вождения автомобиля пользователь может проговорить свой запрос, продиктовать нужный адрес, проверить пробки в приложении навигатора. Или же если пользователь выполняет слишком много задач и не может сконцентрироваться на одной.
- UX-исследователь и экс-специалист по речевым интерфейсам в Google ***Константин Самойлов*** в своем докладе, подготовленном для UX-марафона ***«Взаимодействие будущего»***, назвал три важных признака, которыми должны обладать голосовые интерфейсы:
- естественный язык,
- диалог,
- неограниченный словарный запас и грамматика.

- **Интеллектуальные голосовые помощники** (или голосовые ассистенты) – это веб-сервисы, которые объединяют технологию распознавания речи и текста и поиска информации по ключевым словам. Голосовые помощники умеют распознавать речь, определять значение сказанного и синтезировать голос для ответа. Основные приложения: ***Alexa*** Amazon, ***Siri*** Apple, ***OK Google***, ***Кортана*** Microsoft, ***«Алиса»*** Яндекса.
- Голосовые ассистенты используются не только в мобильных приложениях и персональных компьютерах, но и в устройствах умного дома. Они могут быть внедрены в холодильники, бытовую технику, машины. Или же представляют собой беспроводные динамики, снабженные голосовым управлением.

- Звучащая речь для нас — это, прежде всего, цифровой сигнал. И если мы посмотрим на запись этого сигнала, то не увидим там ни слов, ни четко выраженных фонем — разные «речевые события» плавно перетекают друг в друга, не образуя четких границ. Одна и та же фраза, произнесенная разными людьми или в различной обстановке, на уровне сигнала будет выглядеть по-разному. Вместе с тем, люди как-то распознают речь друг друга: следовательно, существуют инварианты, согласно которым по сигналу можно восстановить, что же, собственно, было сказано. Поиск таких инвариантов — задача акустического моделирования.

  Предположим, что речь человека состоит из фонем (это грубое упрощение, но в первом приближении оно верно). Определим фонему как минимальную смыслоразличительную единицу языка, то есть звук, замена которого может привести к изменению смысла слова или фразы. Возьмем небольшой участок сигнала, скажем, 25 миллисекунд. Назовем этот участок «фреймом». Какая фонема была произнесена на этом фрейме? На этот вопрос сложно ответить однозначно — многие фонемы чрезвычайно похожи друг на друга. Но если нельзя дать однозначный ответ, то можно рассуждать в терминах «вероятностей»: для данного сигнала одни фонемы более вероятны, другие менее, третьи вообще можно исключить из рассмотрения. Собственно, акустическая модель — это функция, принимающая на вход небольшой участок акустического сигнала (фрейм) и выдающая распределение вероятностей различных фонем на этом фрейме. Таким образом, акустическая модель дает нам возможность по звуку восстановить, что было произнесено — с той или иной степенью уверенности.

- Еще один важный аспект акустики — вероятность перехода между различными фонемами. Из опыта мы знаем, что одни сочетания фонем произносятся легко и встречаются часто, другие сложнее для произношения и на практике используются реже. Мы можем обобщить эту информацию и учитывать ее при оценке «правдоподобности» той или иной последовательности фонем.

  Теперь у нас есть все инструменты, чтобы сконструировать одну из главных «рабочих лошадок» автоматического распознавания речи — скрытую марковскую модель (HMM, Hidden Markov Model). Для этого на время представим, что мы решаем не задачу распознавания речи, а прямо противоположную — преобразование текста в речь. Допустим, мы хотим получить произношение слова «Яндекс». Пусть слово «Яндекс» состоит из набора фонем, скажем, [й][а][н][д][э][к][с]. Построим конечный автомат для слова «Яндекс», в котором каждая фонема представлена отдельным состоянием. В каждый момент времени находимся в одном из этих состояний и «произносим» характерный для этой фонемы звук (как произносится каждая из фонем, мы знаем благодаря акустической модели). Но одни фонемы длятся долго (как [а] в слове «Яндекс»), другие практически проглатываются. Здесь нам и пригодится информация о вероятности перехода между фонемами. Сгенерировав звук, соответствующий текущему состоянию, мы принимаем вероятностное решение: оставаться нам в этом же состоянии или же переходить к следующему (и, соответственно, следующей фонеме).

- Более формально HMM можно представить следующим образом. Во-первых, введем понятие эмиссии. Как мы помним из предыдущего примера, каждое из состояний HMM «порождает» звук, характерный именно для этого состояния (т. е. фонемы). На каждом фрейме звук «разыгрывается» из распределения вероятностей, соответствующего данной фонеме. Во-вторых, между состояниями возможны переходы, также подчиняющиеся заранее заданным вероятностным закономерностям. К примеру, вероятность того, что фонема [а] будет «тянуться», высока, чего нельзя сказать о фонеме [д]. Матрица эмиссий и матрица переходов однозначно задают скрытую марковскую модель.

  Хорошо, мы рассмотрели, как скрытая марковская модель может использоваться для порождения речи, но как применить ее к обратной задаче — распознаванию речи? На помощь приходит алгоритм Витерби. У нас есть набор наблюдаемых величин (собственно, звук) и вероятностная модель, соотносящая скрытые состояния (фонемы) и наблюдаемые величины. Алгоритм Витерби позволяет восстановить наиболее вероятную последовательность скрытых состояний.

- Пусть в нашем словаре распознавания всего два слова: «Да» ([д][а]) и «Нет» ([н'][е][т]). Таким образом, у нас есть две скрытые марковские модели. Далее, пусть у нас есть запись голоса пользователя, который говорит «да» или «нет». Алгоритм Витерби позволит нам получить ответ на вопрос, какая из гипотез распознавания более вероятна.

  Теперь наша задача сводится к тому, чтобы восстановить наиболее вероятную последовательность состояний скрытой марковской модели, которая «породила» (точнее, могла бы породить) предъявленную нам аудиозапись. Если пользователь говорит «да», то соответствующая последовательность состояний на 10 фреймах может быть, например, [д][д][д][д][а][а][а][а][а][а] или [д][а][а][а][а][а][а][а][а][а]. Аналогично, возможны различные варианты произношения для «нет» — например, [н'][н'][н'][е][е][е][е][т][т][т] и [н'][н'][е][е][е][е][е][е][т][т]. Теперь найдем «лучший», то есть наиболее вероятный, способ произнесения каждого слова. На каждом фрейме мы будем спрашивать нашу акустическую модель, насколько вероятно, что здесь звучит конкретная фонема (например, [д] и [а]); кроме того, мы будем учитывать вероятности переходов ([д]->[д], [д]->[а], [а]->[а]). Так мы получим наиболее вероятный способ произнесения каждого из слов-гипотез; более того, для каждого из них мы получим меру, насколько вообще вероятно, что произносилось именно это слово (можно рассматривать эту меру как длину кратчайшего пути через соответствующий граф). «Выигравшая» (то есть более вероятная) гипотеза будет возвращена как результат распознавания.

- Однако акустическая модель — это всего лишь одна из составляющих системы. Что делать, если словарь распознавания состоит не из двух слов, как в рассмотренном выше примере, а из сотен тысяч или даже миллионов? Многие из них будут очень похожи по произношению или даже совпадать. Вместе с тем, при наличии контекста роль акустики падает: невнятно произнесенные, зашумленные или неоднозначные слова можно восстановить «по смыслу». Для учета контекста опять-таки используются вероятностные модели. К примеру, носителю русского языка понятно, что естественность (в нашем случае — вероятность) предложения «мама мыла раму» выше, чем «мама мыла циклотрон» или «мама мыла рама». То есть наличие фиксированного контекста «мама мыла ...» задает распределение вероятностей для следующего слова, которое отражает как семантику, так и морфологию. Такой тип языковых моделей называется n-gram language models (триграммы в рассмотренном выше примере); разумеется, существуют куда более сложные и мощные способы моделирования языка.

- В последние годы основной тренд исследований в области распознавания речи смещается в сторону отказа от использования скрытых марковских моделей. Согласно марковскому свойству, следующее состояние — в данном случае звуковая единица типа фонемы — в цепи зависит только от предыдущего состояния и не зависит от всех остальных состояний в прошлом. Конечно, такая модель является очень упрощенной, поэтому для построения акустических моделей в настоящее время стали использоваться рекуррентные нейронные сети, которые позволяют сохранить долговременные зависимости. Первые результаты в 2014 году показали, что такой подход позволяет решать задачи распознавания речи даже лучше, чем описанные ранее скрытые марковские модели с глубокими нейронными сетями прямого распространения.Развитие современных речевых технологий идет в сторону реализации полного цикла обучения систем распознавания спонтанной речи без выделения отдельных акустических и лингвистических моделей. Вместо предварительного отбора акустических признаков, таких как кепстральные коэффициенты, все участки речевого сигнала представляются своими спектрограммами, которые подаются на вход одной большой нейронной сети. Примером здесь может являться система *Deep Speech 2*, в которой спектрограммы обрабатываются как изображения с помощью последовательности сверточных слоев, соединяющихся с последовательностью рекуррентных блоков. На выходе нейронной сети появляется результат распознавания — последовательность символов. Такой подход пока не реализован в существующих программных библиотеках, но это будущее, которое нас наверняка ждет.
- Наконец, следующий шаг — это разработка высокоточных технологий понимания речи, когда нужно не только распознать речь, перевести ее в текст, но и понять содержание разговора, чтобы отвечать на вопросы и поддерживать диалог. Подобные системы уже появились, например «Алиса», и будут в ближайшее время становиться все более развитыми.

о алгоритма диаризации — разделения
ветствии с принадлежностью слов тому
ая технология более эффективна, чем
tala_novyj_algoritm_diarizacii.html).

(RNN). Такая архитектура позволяет
довательностей произвольной длины и
удиопотоком. В разработке Google для
NN, вычленяющий высказывания.

тма диаризации с помощью теста NIST SRE
6 %. Использовавшиеся ранее методы
показывали погрешность 8,8 % и 9,9 %
лгоритм обладает производительностью,
достаточной для обработки потока в реальном времени

- Google активно развивает технологии распознавания речи и привлекает к этому процессу сторонних разработчиков. В апреле 2017 года компания открыла доступ к Cloud Speech API — технологии распознавания речи, лежащей в основе Google Ассистента.

- Разработчики из Университета Цинхуа разработали голосовой помощник для смартфонов, который распознаёт команды по движениям губ пользователя. Эта технология может применяться в общественных местах без риска помешать другим.
- Юаньчунь Ши (Yuanchun Shi) с коллегами представили на конференции UIST 2018 статью, в которой описали технологию распознавания движений губ и перевода их в текст. Такой голосовой помощник использует фронтальную камеру и свёрточную нейросеть. Алгоритм отслеживает 20 контрольных точек, которые достаточно точно описывают форму губ, а также определяет насколько открыт рот пользователя. Это позволяет распознать начало и конец команды. Второй алгоритм расшифровывает данные. При этом пока все вычисления происходят отдельно на мощном ПК.
- Для распознавания используется ограниченный набор команд — всего 44, которые относятся как к отдельным приложениям, так и к конкретным функциям, вроде включения и выключения Wi-Fi. Также поддерживаются и общесистемные задачи, вроде ответа на сообщение или выделения текста.
- **Насколько точно голосовой помощник распознаёт команды?**
- Разработчики утверждают, что средняя точность распознавания составила 95,5 % по результатам обучения на речи 21 человека. Тесты проводились в метро Пекина. В результате оказалось, что такой метод считается пользователями более комфортным.

Стохастические модели процессов

- Стохастические процессы, задающие зависимость между четким временем и случайной величиной с известным законом распределения вероятностей.
- Стохастические процессы указанного вида описывают поведение процесса в условиях риска и неопределенности, то есть случайные и недетерминированные ее изменения.

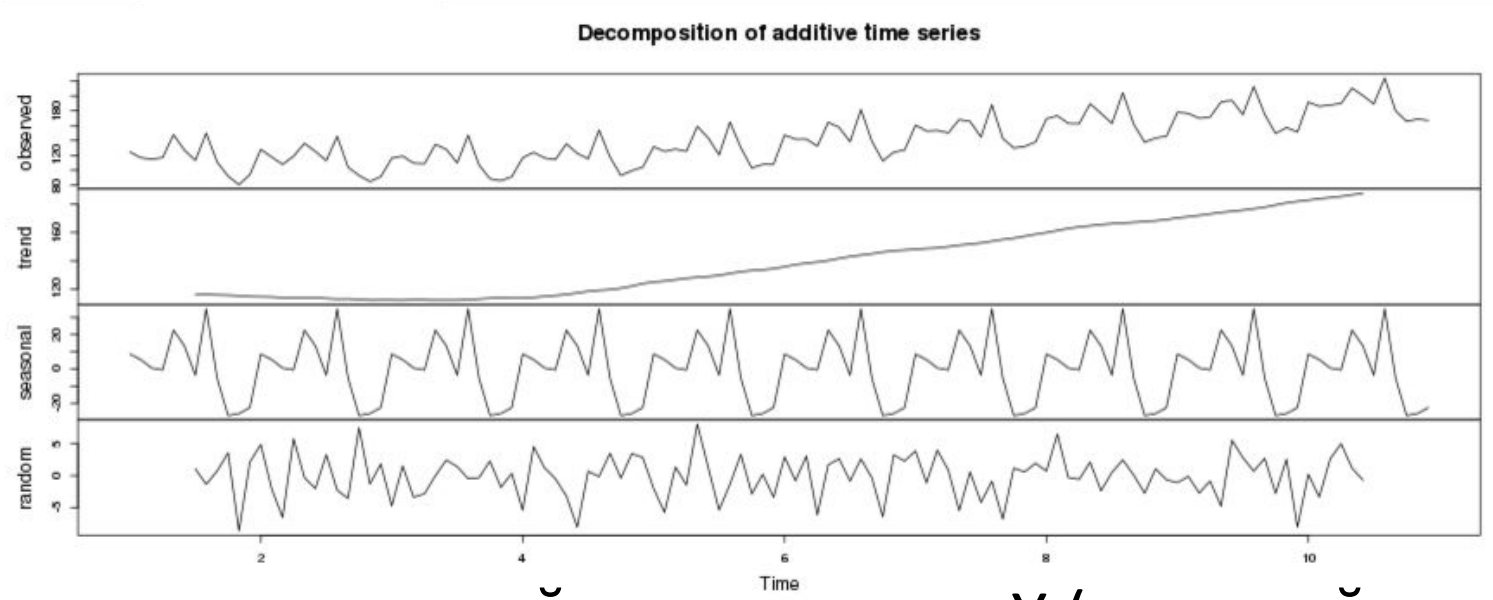
Модель временного ряда

$$y_t = f_t + \psi_t + \xi_t$$

- y_t – заданный временной ряд, f_t –тренд-цикл временного ряда, ψ_t – сезонность, ξ_t – случайная компонента.
- Модель тренда :

1. $y_t = f(t) + u_t, \quad y_t = at + u_t$ [*Андерсон*,1976]

Анализ поведения путем декомпозиции и ВР



Динамические свойства выхода Y (верхний график):

- Наличие восходящего тренда
- Наличие сезонности

$$\begin{cases} x(t) = x_{tr}(t) + s(t), \\ s(t) = x_{sea}(t) + a(t) \\ a(t) = x_{ARMA}(t) + g(t), \\ g(t) = x_{GARCH}(t) + e(t); \end{cases} \quad t = t_1, \dots, t_n$$

Нечеткие модели входов и выходов

Л. Заде предложил по аналогии с теорией вероятности использовать функцию в качестве математической модели лингвистической неопределенности

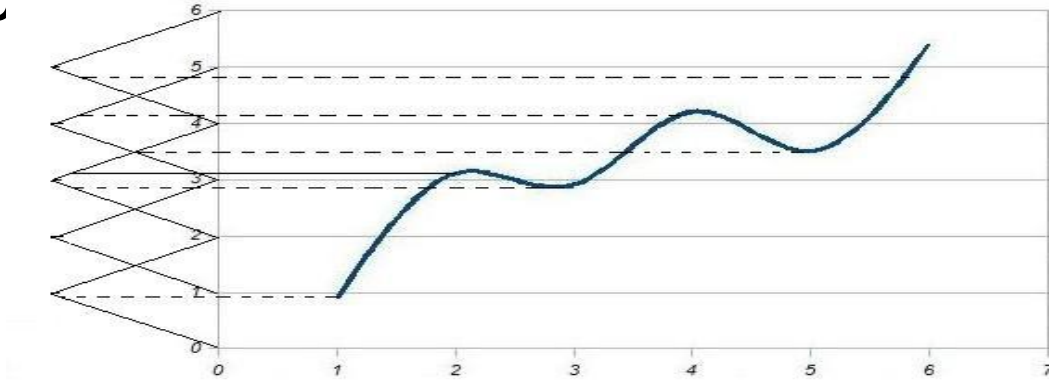
$$Y = \mu(x, B),$$

где Y – результат вычисления функции, выражающий меру неопределенности (нечеткости) для конкретного объекта X.

μ – непрерывная функция

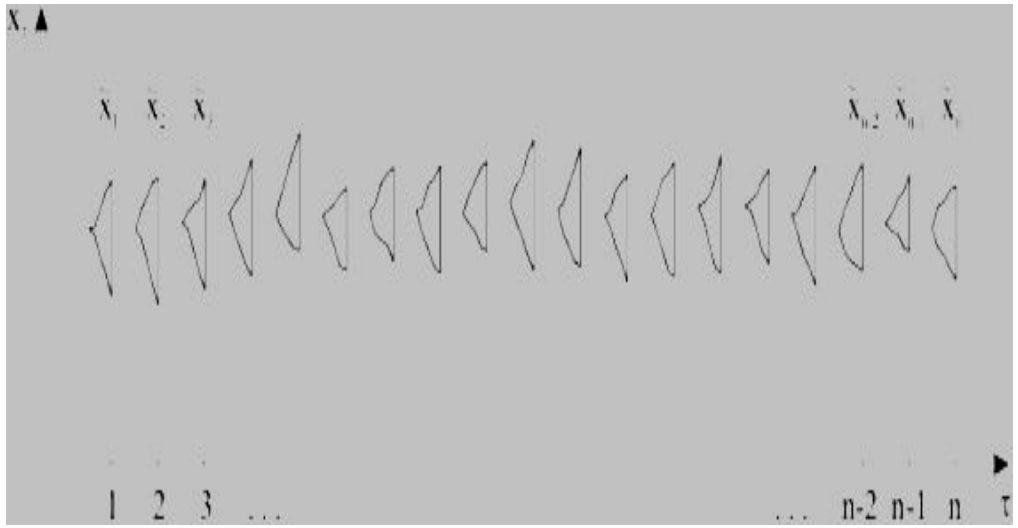
$$\mu : X \to [0,1].$$

- В том случае, если значения функции принадлежности нечеткого множества представлены точными числовыми значениями, такие нечеткие множества относят к *нечетким множествам типа 1*.
- Если значения функции принадлежности нечеткого множества моделируются другими нечеткими множествами, то такое нечеткое множество относят к *нечетким множествам типа 2*.



Нечеткий временной ряд как модель процесса

$$\tilde{x}_i = Fuzzy(\mu_{\tilde{x}_i}(w), x_i)$$



- Предположим, что задан процесс, состояния которого описываются *n* значениями одной переменной.
- В результате наблюдения получен временной ряд этой переменной, он преобразован в нечеткий

Нечеткие процессы

1. Нечеткие процессы. Нечеткие процессы представляют группу зависимостей между категориями x_i, t_i (время – значение), каждая их которых может быть задана нечеткими значениями. Рассмотрим их как разные классы отношений «время» – «значение».

а. Отношение «Четкое время t» – «Нечеткая переменная $\tilde{x}$ ».

В этом отношении каждое нечеткое значение $\tilde{x}_i$ представляется нечетким термом, определенным на некотором базовом множестве и функцией принадлежности , выражающей количественную зависимость между элементами базового множества и нечетким значением $\tilde{x}_i$. Это отношение соответствует представлению нечеткого ВР в виде $\tilde{x}_i = Fuzzy(x_i, t_i)$.

Нечеткие процессы

a. Отношение «Нечеткое время $\tilde{t}$» – «четкая переменная x». В этом случае можно предположить, что четкое значение времени было предварительно преобразовано в нечеткие значения $\tilde{t}_i = Fuzzy\_t(x_i, t_i)$. Идентификация нечетких тенденций для указанного отношения может проводиться двояким образом: по отношению к нечеткому времени $\tau_i = Tend(\tilde{x}_i, \tilde{t}_i)$ и по отношению к приближенному моменту времени, полученному в результате его дефаззификации $\tau_i = Tend(\tilde{x}_i, t_i')$.

b. Отношение «Нечеткое время $\tilde{t}$» – «нечеткая переменная $\tilde{x}$». Данное отношение может использовать без предварительных преобразований алгоритм $\tau_i = Tend(\tilde{x}_i, \tilde{t}_i)$.

Методологические основы нечеткого моделирования ВР

- Состояния динамического процесса, порождающего ВР, представляются нечеткими множествами, введенными Заде.
- Идентификация модели процесса в виде $f: A \times B$ на основе нечеткой импликации $R: A \to B$ и нечетких конечно-разностных уравнений (Song, Chissom, 1993)
- Выражение модели в форме базы нечетких правил «Если – То»
- Нечеткий логический вывод по базе нечетких правил, как численный алгоритм приближения нелинейных зависимостей

Теоретические основы нечеткого моделирования

- Теорема нечеткой аппроксимации (Kosko, 1992)
- Теорема аппроксимации любой непрерывной функции с произвольной точностью нечеткой моделью и нечетким логическим выводом (Vang,1992), (Kastro, 1995)

$$\begin{cases} R_1: \text{ЕСЛИ } x_1 \text{ есть } A_{11} \dots \text{И} \dots x_n \text{ есть } A_{1n}, \text{ТО } y \text{ есть} \\ R_j: \text{ЕСЛИ } x_1 \text{ есть } A_{j1} \dots \text{И} \dots x_n \text{ есть } A_{jn}, \text{ТО } y \text{ есть} \\ \dots \\ R_m: \text{ЕСЛИ } x_1 \text{ есть } A_{m1} \dots \text{И} \dots x_n \text{ есть } A_{mn}, \text{ТО } y \text{ ест} \end{cases}$$

Пример. Определите цель моделирования

- **Объектом исследования** является сложный технический объект – сервер предприятия

- Исходными данными для вычислительных экспериментов послужил массив данных $\{x_{\text{сеть}}, x_{\text{цп}}, x_{\text{оп}}\}$, представляющий собой значения параметров сервера: загрузка сети, процент загрузки процессора и процент загрузки оперативной памяти, измеренных в течение 90 дней.

Обобщенная регрессионно-нечеткая модель сервера

$$\tilde{S} = F(\tilde{Z}, A, \tilde{E})$$

$\tilde{S}$ - лингвистическая оценка состояния сервера; $\tilde{E}$ - лингвистическая экспертная оценка состояния предприятия, являющаяся для сервера внешним фактором; $\tilde{Z}$ - лингвистическая оценка состояния отдельных характеристик сервера; A - набор правил для определения состояния сервера.

$$\tilde{Z} = F(\tilde{V}, R)$$

R - набор правил для определения состояния отдельных характеристик сервера, $\tilde{V}$ - лингвистическая оценка локальной тенденции характеристик сервера.

$$\tilde{V} = GTend(TTend(\tilde{X}))$$

$\tilde{X} = \{\tilde{x}_i\}, i = \overline{1,n}$ - множество нечетких значений характеристик сервера, получаемых по формуле:

$$\tilde{x}_i(t) = Fuzzy(x_i(t))$$

x_i - прогнозные значения характеристик сервера, получаемые по модели:

$$x_i(t) = f(t) + \varphi(t) + e(t)$$

- Нечеткое шкалирование – построение ACL-шкалы для каждой из характеристик.
  - $C_x = \{\tilde{X}, \tilde{V}, Fuzzy, TTend, GTend\}$
  - $\tilde{X} = \{\text{Хороший}, \text{Нормальный}, \text{Плохой}\}$
  - $\tilde{V} = \{\text{Падение}, \text{Стабильность}, \text{Рост}\}$

- Фаззификация прогнозов характеристик сервера.
  - $\tilde{x}_i(t) = Fuzzy\left(\hat{x}_i(t)\right)$

Введены три лингвистические переменные

Нечеткие множества		Нечеткие тенденции		Нечеткие интенсивности	
1	Очень низкий	1	Рост	1	Нулевой
2	Низкий	2	Падение	2	Очень малый
3	Ниже среднего	3	Стабильность	3	Малый
4	Средний			4	Средний
5	Выше среднего			5	Больше среднего
6	Высокий			6	Большой
7	Очень высокий			7	Очень большой

Результаты фаззификации значений выходных характеристик сервера

№	Значение	Степень принадлежности	Нечеткое значение	Нечеткая тенденция	Интенсивность
82	77,3	0,76041	Выше среднего	Рост	Очень малый
83	54,71	0,79087	Низкий	Падение	Средний
84	76,438	0,89196	Выше среднего	Рост	Средний
85	68,159	0,84419	Средний	Стабильность	Очень малый
86	64,029	0,78638	Ниже среднего	Падение	Очень малый
87	73,36	0,6382	Выше среднего	Рост	Малый
88	67,396	0,72763	Средний	Падение	Очень малый
89	65,366	0,58223	Ниже среднего	Стабильность	Очень малый
90	71,681	0,61818	Средний	Рост	Очень малый

Правила нечеткой модели

$$\begin{cases} R_1: \text{Если } \tilde{x}_i(t) \text{ есть } \tilde{X}_{i,1} \text{ и } \tilde{v}_{Gt,i} \text{ есть } \tilde{V}_{i,1}, \text{то } z_i \text{ есть } \tilde{Z}_1 \\ R_2: \text{Если } \tilde{x}_i(t) \text{ есть } \tilde{X}_{i,2} \text{ и } \tilde{v}_{Gt,i} \text{ есть } \tilde{V}_{i,2}, \text{то } z_i \text{ есть } \tilde{Z}_2 \\ \dots \\ R_k: \text{Если } \tilde{x}_i(t) \text{ есть } \tilde{X}_{i,k} \text{ и } \tilde{v}_{Gt,i} \text{ есть } \tilde{V}_{i,k}, \text{то } z_i \text{ есть } \tilde{Z}_k \end{cases}$$

Нечеткое моделирование

Для всех технических параметров построены правила нечеткого вывода:

Нечеткое значение	Общая тенденция	Решение по параметру
Низкий	Падение	**Недостаток загрузки**
Низкий	Рост	**Норма**
Средний	Стабильность	**Норма**
…	…	**…**
Высокий	Рост	**Превышение загрузки**

Сформированы правила нечеткого вывода для прогноза состояния сервера

Загрузка ЦП	Загрузка ОП	Загрузка по трафику	Экспертная оценка нагрузки	Состояние сервера
Норма	Норма	Норма	Стабильность	**Норма**
…	…	…	…	**…**
Норма	Недостаток загрузки	Превышение нагрузки	Стабильность	**Возможно превышение**
Норма	Недостаток загрузки	Превышение нагрузки	Снижение нагрузки	**Норма**
…	…	…	…	**…**
Недостаток загрузки	Норма	Норма	Снижение нагрузки	**Возможен простой ресурсов**

№	Загрузка ЦП	Загрузка ОП	Загрузка сети
82	Рост	Стабильность	Рост
83	Падение	Стабильность	Падение
84	Рост	Стабильность	Рост
85	Стабильность	Падение	Стабильность
86	Падение	Стабильность	Падение
87	Рост	Стабильность	Рост
88	Падение	Рост	Падение
89	Падение	Падение	Стабильность
90	Рост	Стабильность	Рост
ОТ	**Рост**	**Стабильность**	**Стабильность**

- Нечеткое моделирование всего сервера на основе предложенной методики

$$A_m: \text{Если } \tilde{z}_1 \text{ есть } \tilde{Z} \text{ и } \tilde{z}_2 \text{ есть } \tilde{Z} \text{ и } \dots \ \tilde{z}_n \text{ есть } \tilde{Z} \text{ и } E_{\text{пр}} \text{ есть } \tilde{E}, \text{то } S \text{ есть } \tilde{S}$$

$$\tilde{E} = \{\text{Уменьшение нагрузки, Стабильность, Увеличение нагрузки}\}$$

$$\tilde{S} = \{\text{Простой, Норма, Превышение}\}$$

						
Загрузка ЦП	Выше среднего	0,971	Рост	0,925	**Превышение загрузки**	**0,925**
Загрузка ОП	Выше среднего	0,729	Рост	0,957	**Превышение загрузки**	**0,729**
Загрузка сети	Выше среднего	0,618	Стабильность	0,985	**Норма**	**0,618**

	Загрузка ЦП	**Загрузка ОП**	**Загрузка сети**	**Экспертная оценка нагрузки**	**Состояние сервера**
Лингвистическая оценка	Превышение загрузки	Превышение загрузки	Норма	Рост	**Превышение загрузки**
Степень принадлежности	0,925	0,729	0,618	1	0,618