

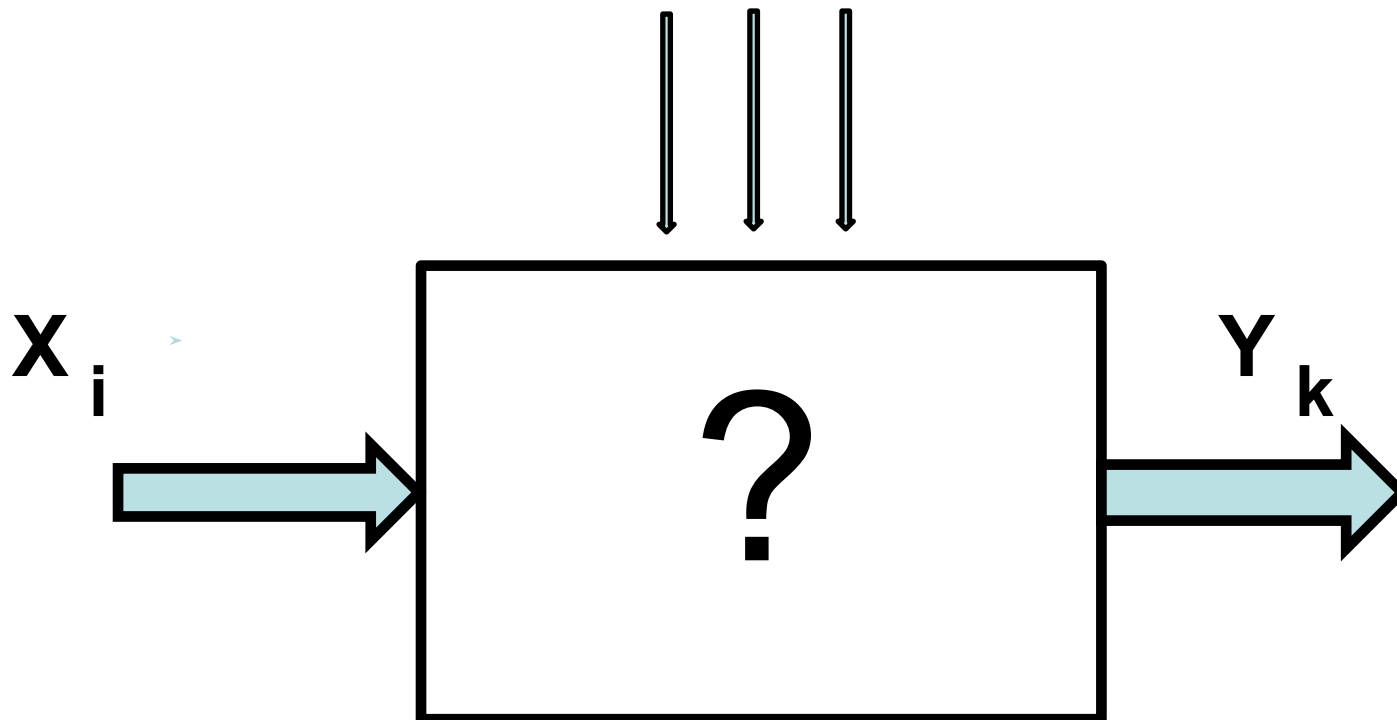
Распознавание Образов

Pattern Recognition

Общая модель кибернетики: Черный ящик

• 1

Дестабилизирующие воздействия



Общая модель кибернетики

- Общая ситуация – черный ящик:

Варианты задач:

- Построение модели – регрессия.

$$Y = F (\mathbf{x})$$

- Построение модели – нейрокомпьютинг.

Несколько десятков парадигм!

- Задачи распознавания образов – классификация.
- Задача кластеризации.

Определения

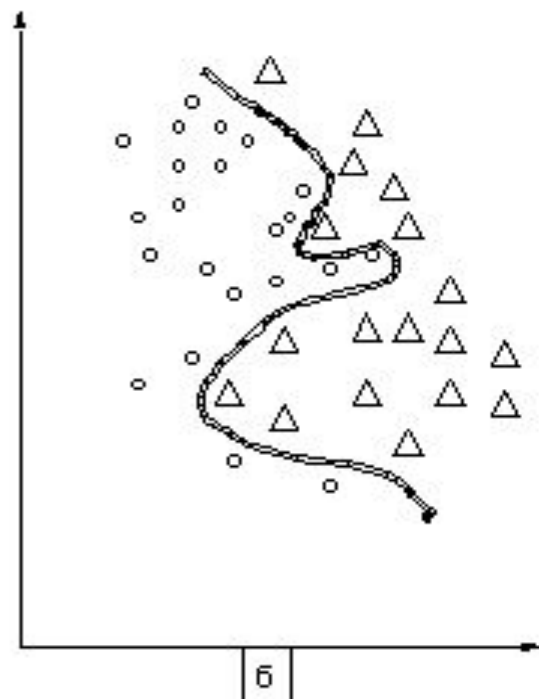
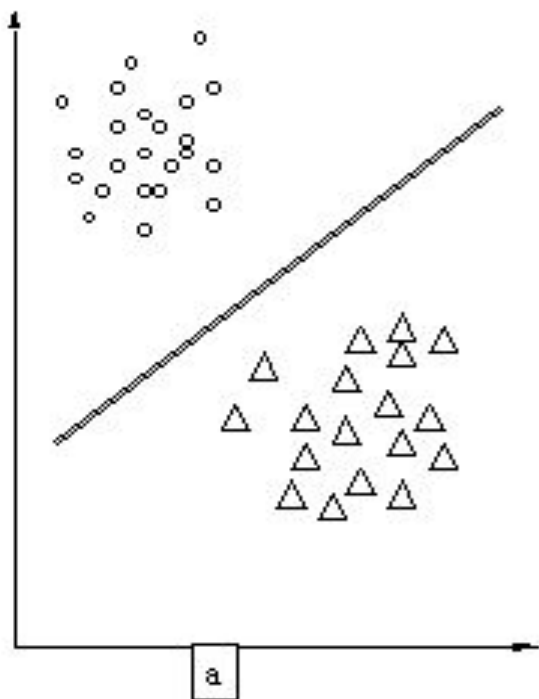
- Предмет распознавания образов (классификация) объединяет ряд научных дисциплин. Их связывает поиск решения общей задачи - выделить элементы, принадлежащие конкретному классу, среди множества элементов, относящихся к нескольким классам. Под классом образов понимается некоторая категория, с рядом свойств, общих для всех ее элементов.
- Образ (класс) - это описание любого элемента как представителя соответствующего класса образов. Элементы класса имеют общую метку, метку принадлежности данному классу.

Простой пример

Распознавание больных гриппом:

X- температура; Y-число лейкоцитов

Пространство признаков – решающее правило



Непараметрические методы

- Метод ближайшего соседа.
- Метод k -ближайших соседей.
- Метод потенциальных функций.
- Метод «парзеновских» окон.

Идеи эвристик распознавания_1

- **Метод ближайшего соседа** является, пожалуй, самым простым алгоритмом классификации. Классифицируемый объект относится к тому классу, которому принадлежит ближайший объект обучающей выборки.
- **Метод k-ближайших соседей.** Для повышения надёжности классификации объект относится к тому классу, которому принадлежит большинство из его *соседей* — ближайших к нему объектов обучающей выборки. В задачах с двумя классами число соседей берут нечётным, чтобы не возникало ситуаций неоднозначности, когда одинаковое число соседей принадлежат разным классам.

Идеи эвристик распознавания_2

- **Метод взвешенных ближайших соседей.** В задачах с числом классов 3 и более нечётность уже не помогает, и ситуации неоднозначности всё равно могут возникать. Тогда i -му соседу приписывается вес, как правило, убывающий с ростом ранга соседа. Объект относится к тому классу, который набирает больший суммарный вес среди ближайших соседей.

Примеры задач классификации

**Содержательный
характер задачи
классификации**

Вид исходных данных

**Вид ответа системы
распознавания**

**Распознавание
символов**

**Оптические : сигналы
или элементы
развертки**

Название символа

Распознавание речи

**Акустические сигналы.
Голос**

**Слова
говорящего человека**

Прогноз погоды

Синоптические карты

Прогноз погоды

**Медицинская или
техническая
диагностика**

Симптомы заболевания

**Вид заболевания или
неисправности**

**Прогноз состояния
фондовой биржи**

**Финансовые новости и
сводки**

**Прогноз повышения
или понижения цен на
рынке**

Способы определения классов объектов

Перечисление.

Каждый класс задаётся путём прямого указания его членов. Такой подход используется в том случае, если доступна полная априорная информация о всех возможных объектах распознавания. Предъявляемые системе образы сравниваются с заданными описаниями представителей классов и относятся к тому классу, которому принадлежат наиболее сходные с ними образцы. Такой подход называют методом ***сравнения с эталоном***. Его недостатком является слабая устойчивость к шумам и искажениям в распознаваемых образах.

Способы определения классов объектов

Пример. Распознавание машинопечатного шрифта. Все символы имеют чётко заданное шрифтом начертание. Следовательно, необходимо обучить систему путём прямого указания изображений всех распознаваемых символов (т.е. путём задания эталонов):

А Б В а б в ... 1 2 3

Необходимо отметить, что если предполагается распознавание курсивного, полужирного или иного начертания символов шрифта, то при таком подходе будет необходимо представить каждый вариант начертания каждого символа. Кроме того, способность распознавания линейных трансформаций данных эталонов требует определённых усилий на этапе предобработки.

Способы определения классов объектов

Задание общих свойств.

Класс задаётся указанием некоторых признаков, присущих всем его членам. Распознаваемый объект в таком случае не сравнивается напрямую с группой эталонных объектов. В его первичном описании выделяются значения определённого набора признаков, которые затем сравниваются с заданными признаками классов. При этом для каждого признака может задаваться требование либо к его наличию/отсутствию, либо к нахождению его числового значения в установленных пределах.

Такой подход называется **сопоставлением по признакам**. Он экономичнее метода сравнения с эталоном в вопросе количества памяти, необходимой для хранения описаний классов. Кроме того, он допускает некоторую вариативность распознаваемых образов. Однако, **главной сложностью является определение полного набора признаков, точно отличающих членов одного класса от членов всех остальных.**

Способы определения классов объектов

Пример. Распознавание цифр почтовых индексов.

Рассматривается набор из 10-ти цифр почтового индекса.

Каждый символ - представляет класс распознаваемых объектов — одну из цифр. Все эти изображения построены по одному принципу — с помощью комбинирования вертикальных, горизонтальных и диагональных сегментов в определённых позициях знакомест. Для описания классов предлагаются следующие признаки:

x1 — количество вертикальных линий минимального размера;

x2 — количество горизонтальных линий;

x3 — количество наклонных линий;

x4 — количество горизонтальных линий снизу объекта.

С помощью этих признаков можно следующим образом задать классы цифр:

Способы определения классов объектов

	<i>x1</i>	<i>x2</i>	<i>x3</i>	<i>x4</i>
0	4	2	0	1
1	2	0	1	0
2	1	2	1	1
3	0	2	2	0
4	3	1	0	0
5	2	3	0	1
6	2	2	1	1
7	1	1	1	0
8	4	3	0	1
9	2	2	1	0

Заметим, что набор выбранных признаков не является единственным. Качество распознавания во многом зависит от того, насколько удачно выбран набор признаков.

Общие принципы распознавания

- Методика отнесения элемента к какому-либо образу называется **решающим правилом**.
- Еще одно важное понятие - **метрика**, способ определения расстояния между элементами универсального множества. Чем меньше это расстояние, тем более похожими являются образы, символы, звуки - то, что мы распознаем.

Гипотеза компактности

Классификация, распознавание, кластеризация - неявно опираются на одно важное предположение, называемое гипотезой компактности: если система признаков и мера сходства объектов введена достаточно удачно, то схожие объекты гораздо чаще лежат в одном классе, чем в разных. В этом случае граница между классами имеет достаточно простую форму, а классы образуют компактно локализованные области в пространстве признаков.

Проблема выбора метрики_1

- В практических задачах классификации редко встречаются такие «идеальные случаи», когда заранее известна хорошая функция расстояния. Если объекты описываются числовыми векторами, часто берут евклидову метрику. Этот выбор, как правило, ничем не обоснован — просто это первое, что приходит в голову. При этом необходимо помнить, что все признаки должны быть измерены «в одном масштабе», т.е. **отнормированы**. В противном случае признак с наибольшими числовыми значениями будет доминировать в метрике, остальные признаки, фактически, учитываться не будут.

Проблема выбора метрики_2

- Однако и нормировка является весьма сомнительной эвристикой, так как остаётся вопрос: «неужели все признаки одинаково значимы и должны учитываться примерно с одинаковым весом?»

Об информативности признаков

- Если признаков слишком много, а расстояние вычисляется как сумма отклонений по отдельным признакам, то возникает проблема проклятия размерности. Суммы большого числа отклонений с большой вероятностью имеют очень близкие значения (согласно закону больших чисел). Получается, что в пространстве высокой размерности все объекты примерно одинаково далеки друг от друга; выбор ближайших соседей становится практически произвольным.
- Проблема определения информативности признаков - отбор относительно небольшого числа информативных признаков (features selection).

Задача кластеризации

Задачи кластеризации (clustering) отличаются от классификации (classification) тем, что в них не задаются ответы $y(i) = y^*x(i)$. Известны только сами объекты $x(i)$, и требуется разбить выборку на подмножества (кластеры) так, чтобы каждый кластер состоял из схожих объектов, а объекты разных кластеров существенно отличались.

Методологии распознавания

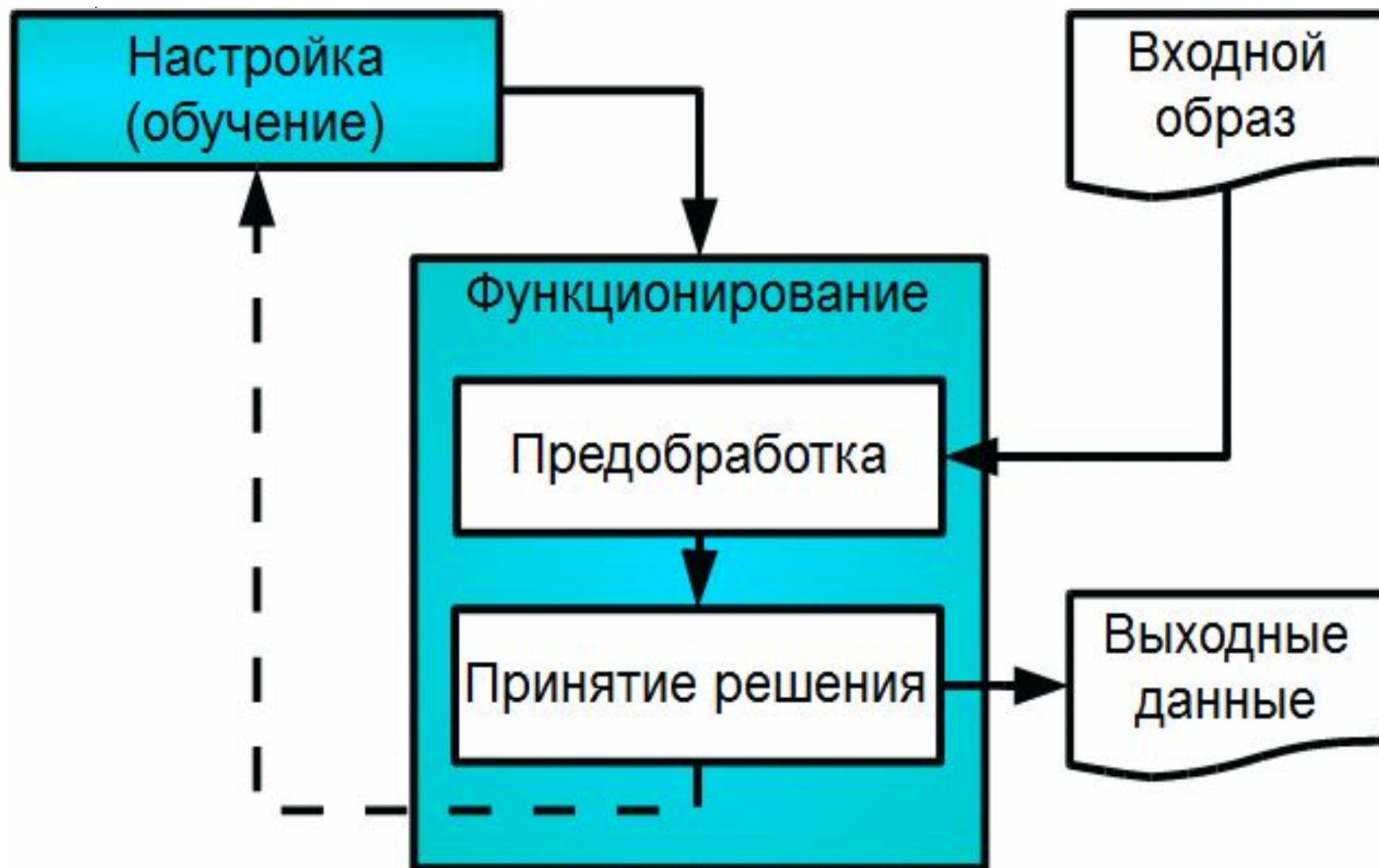
- Статистический подход:
непараметрические методы.
- Распознавание по образцу
- Статистический подход: основан на теории принятия решений.
- Структурно-лингвистический подход
- Нейросетевой подход.

Алгоритм «ближайшего соседа» или распознавание «по образцу»

Простейший алгоритм на основе метода множества эталонов.

1. На входе имеется обучающая выборка - набор примеров A'_{ij} для каждого образа A_i , метрика d и сам распознаваемый объект x .
2. С помощью метрики вычисляем расстояние от x до каждого элемента обучающей выборки $d(x, a_{ij})$
3. Элемент x относится к образу, который окажется ближе всех.

Распознавание по образцу



Этап обучения – этап распознавания

Этап обучения – этап распознавания !!!

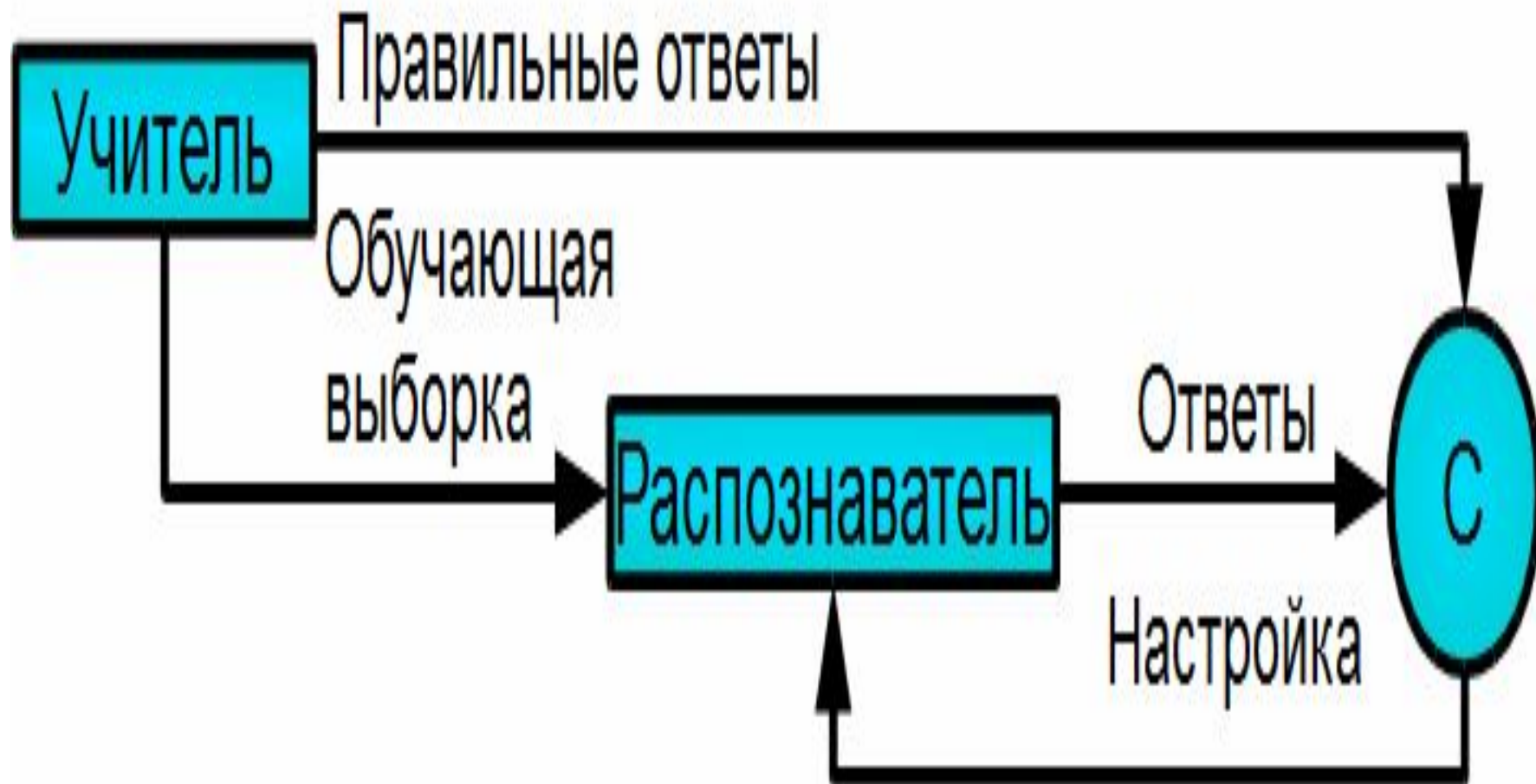
Обучение с учителем – обучение без учителя.

Множество выборок Y_k разбивается на c классов: $Y_1, Y_2, \dots, Y_c$.

Для выборок каждого класса считаются соответствующие векторы признаков z .

Получаем так называемое «помеченное обучающее множество».

Обучение с учителем



Статистический подход

- Основу статистического подхода к задаче распознавания образов, составляет Байесовская теория принятия решений.
- Подход основан на предположении, что задача выбора решения сформулирована в терминах теории вероятностей и известны все необходимые вероятностные величины.
- Предполагается, что каждый объект с некоторой заранее неизвестной вероятностью может принадлежать к любому из имеющихся классов. Решение при этом выдается в виде вероятности принадлежности распознаваемого объекта к различным классам.

Априорная вероятность

- В байесовском статистическом выводе **априорное распределение** вероятностей (prior probability distribution) неопределённой величины p — распределение вероятностей, которое выражает предположения о p до учёта экспериментальных данных. Например, если p — доля избирателей, готовых голосовать за определённого кандидата, то априорным распределением будет предположение о p до учёта результатов опросов или выборов.

Апостериорная вероятность

- Априорное распределение часто задается субъективно опытным экспертом или по известной выборке.
- **Апостериорная вероятность** — условная вероятность случайного события при условии того, что известны апостериорные данные, т.е. полученные после опыта.

Обозначения

- Ω - множество состояний природы.
- A – множество из a возможных действий.
- $\lambda_{ij}(\alpha_i | \Omega_j)$ – потери, связанные с принятием α_i , когда Ω_j - (вероятность правильного распознавания или другое).
- z – вектор признаков, случайная величина.
- $p(z | \Omega_j)$ – условная плотность распределения.
- $P(\Omega_j)$ – априорная вероятность, что состояние природы Ω_j .

Правило Байеса

$$\mathcal{P}(\Omega_j | z) = p(z | \Omega_j) \mathcal{P}(\Omega_j) / p(z),$$

$$p(z) = \sum p(z | \Omega_j) \mathcal{P}(\Omega_j)$$

Позволяет, после измерения z , из априорной вероятности получить апостериорную.

Классификация для 2-х классов

- Апостериорная вероятность
- Ожидаемые потери – риск – R .
- Найти $\min R$.
- Классификация в случае 2-х классов:

принять решение Ω_1

$$p(z | \Omega_1) / p(z | \Omega_2) > (\lambda_{12} - \lambda_{22}) / (\lambda_{21} - \lambda_{11}) \times \\ \times P(\Omega_2) / P(\Omega_1), \text{ при } \lambda_{21} > \lambda_{11}$$

Что есть в теории ?

- Вероятности ошибок.
- Нормальный закон распределения.
- Разделяющие функции – разделяющие плоскости.
- Случай зависимости $\mathbf{P}(\Omega_j)$ от контекста.
- Параметризация при недостаточной статистике (нормальный закон) - критерий максимума правдоподобия.

Рассуждения на основе прецедентов

Парадигма CBR (Case Based-Reasoning), возникла как одно из направлений ИИ. В ЭС важно не только классифицировать объекты, но и выдавать пользователю объяснение предлагаемой классификации. В методе ближайшего соседа такие объяснения выглядят весьма разумно: «Объект отнесён к классу потому, что к этому же классу относился близкий объект обучающей выборки». Такая «прецедентная» логика хорошо понятна экспертам во многих предметных областях.

Структурно-лингвистический подход

- Основан на теории формальных грамматик. (К.Фу)
- **Буква**(символ) — простой неделимый знак.
- **Алфавит** — множество букв (символов)
 $A = \{a, b, c\}$.
- **Конкатенация** — операция слияния символов в языке.

Структурно-лингвистический подход_2

- **Слово** (строка) — упорядоченная совокупность букв из алфавита.
 - Множество всех строк (включая пустую), которые могут быть построены из символов алфавита A , называется **замыканием** A , и обозначается A^* .
Положительное замыкание A (обозначается A^+) — множество всех строк, которые могут быть построены из символов алфавита A , за исключением пустой строки.
- Язык** — в общем случае, совокупность слов или предложений, сформированных набором правил и ограничений.

Классификация грамматик

- **Иерархия Хомского** — классификация формальных языков и формальных грамматик, согласно которой они делятся на 4 типа по их условной сложности. Предложена профессором Массачусетского технологического института, лингвистом Ноамом Хомским.
- Наиболее известная работа Хомского «Синтаксические структуры» (1957).

Обозначения

V_T — множество (словарь) терминальных символов — неизменяемый элемент (буква) алфавита.

V_N — множество (словарь) дополнительных нетерминальных символов.

Имена символов могут содержать буквы, цифры (но не начинаться с цифры), знаки подчёркивания и точки.

P — множество продукций (правил подстановки) грамматики.

S — начальный символ.

Грамматика – Тип 0

Тип 0 — неограниченные

- К типу 0 по классификации Хомского относятся неограниченные грамматики (также известные как грамматики с фразовой структурой). Это — все без исключения формальные грамматики. Для грамматики $G(V_T, V_N, P, S)$, $V = V_T \cup V_N$ все правила имеют вид:

$\alpha \rightarrow \beta$, где α — любая цепочка, содержащая хотя бы один нетерминальный символ, а β — любая цепочка символов из алфавита V .

Грамматика – Тип 1

- **Тип 1 — контекстно-зависимые**

- К этому типу относятся контекстно-зависимые (КЗ) грамматики и неукорачивающие грамматики. Для грамматики $G(V_T, V_N, P, S)$, $V = V_T \cup V_N$ все правила имеют вид:
- $\alpha A \beta \rightarrow \alpha \gamma \beta$, где $\alpha, \beta \in V^*$, $\gamma \in V^+$, $A \in V_N$. Такие грамматики относят к контекстно-зависимым.
- $\alpha \rightarrow \beta$, где $\alpha, \beta \in V^+$, $|\alpha| \leq |\beta|$. Такие грамматики относят к неукорачивающим.
- Эти классы грамматик эквивалентны. Могут использоваться при анализе текстов на естественных языках, однако при построении компиляторов практически не используются в силу своей сложности. Для контекстно-зависимых грамматик доказано утверждение: по некоторому алгоритму за конечное число шагов можно установить, принадлежит цепочка терминальных символов данному языку или нет.

Грамматика – Тип 2

- **Тип 2 — контекстно-свободные**
- К этому типу относятся контекстно-свободные (КС) грамматики. Для грамматики $G(V_T, V_N, P, S)$, $V = V_T \cup V_N$ все правила имеют вид:
- $A \rightarrow \beta$, где $\beta \in V^+$ (для неукорачивающих КС-грамматик, $\beta \in V^*$ - для укорачивающих), $A \in V_N$.
То есть грамматика допускает появление в левой части правила только нетерминального символа.
- КС-грамматики широко применяются для описания синтаксиса компьютерных языков.

Грамматика – Тип 3

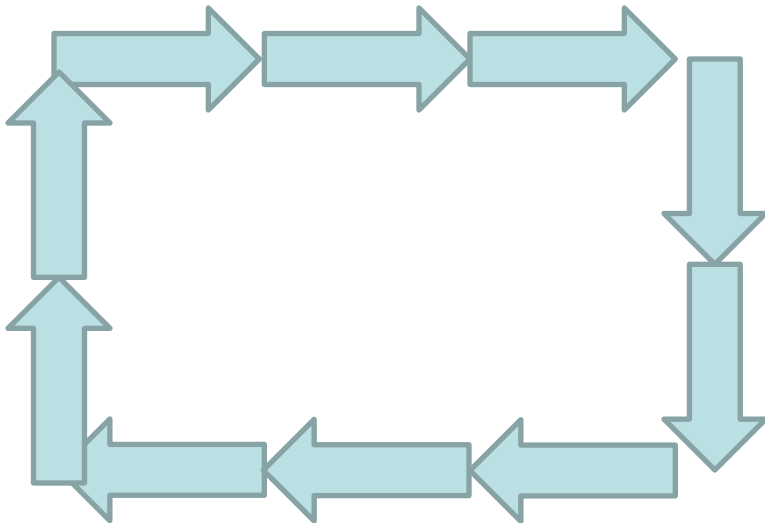
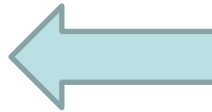
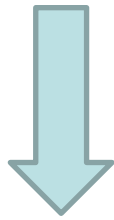
- **Тип 3 — регулярные**
- К третьему типу относятся регулярные грамматики (автоматные) — самые простые из формальных грамматик. Они являются контекстно-свободными, но с ограниченными возможностями.
- Все регулярные грамматики могут быть разделены на два эквивалентных класса, которые для грамматики вида III будут иметь правила следующего вида:
- $A \rightarrow B\gamma$ или $A \rightarrow \gamma$, где $\gamma \in V_T^*$, $A, B \in V_N$ (для левوليнейных грамматик).
- $A \rightarrow \gamma B$; или $A \rightarrow \gamma$, где $\gamma \in V_T^*$, $A, B \in V_N$ (для правوليнейных грамматик).
- Регулярные грамматики применяются для описания простейших конструкций: идентификаторов, строк, констант, а также языков ассемблера, командных процессоров и др.

Структурно-лингвистический подход_3

- Выделяется набор исходных понятий - типичных фрагментов, встречающихся на изображениях, и характеристик взаимного расположения фрагментов - "слева", "снизу", "внутри" и т. д.
- Эти исходные понятия образуют словарь, позволяющий строить различные логические высказывания, иногда называемые предположениями.
- Задача состоит в том, чтобы из большого количества высказываний, которые могли бы быть построены с использованием этих понятий, отобрать наиболее существенные для данного конкретного случая.

«Простенький примерчик»

• a b c d



Цепочки (слова):

aaabbccdd

aaaabbbbccccddd

aabbbbccccddd

Грамматика для «примерчика»

- $V_N = \{ S, A, B, C, D \}$ – словарь нетерминалов
- $V_T = \{ a, b, c, d \}$ – словарь терминалов

Правила подстановки:

- $S \rightarrow aA$ $S \rightarrow AB$ $S \rightarrow bB$
- $A \rightarrow aB$ $B \rightarrow bC$ $C \rightarrow cD$
- $A \rightarrow a$ $B \rightarrow b$ $C \rightarrow c$ $D \rightarrow d$

Продолжение примерчика

abcd – слово (образ) – «четыреугольник».

a(3)b(5)c(3)d(5) – с атрибутами

Что за слово ? Ccdaabbbcc

c(2)d(1)a(2)b(3)c(2)

Глобальная проблема – синтез автомата !!!!

Благодарности - литература

- Р. Дуда, П. Харт. Распознавание образов и анализ сцен. Пер. с англ., М., Мир, 1976.
- Д. Форсайт, Ж. Понс. Компьютерное зрение. Современный подход. 2004г.
- У.Прэтт - Цифровая обработка изображений. В двух книгах.
- Кривопуск Т.И. ДонНТУ
- К. В. Воронцов - <http://www.ccas.ru/voron>
- <http://www.it-claim.ru/> - сайт по Интеллектуальным технологиям