1. **Методология решения прикладных задач.**

Распознавание образов – это научная дисциплина, целью которой является классификация объектов по нескольким категориям или классам. Объекты называются образами.

Классификация основывается на прецедентах.

Прецедент – это образ, правильная классификация которого известна.

Прецедент – ранее классифицированный объект, принимаемый как образец при решении задач классификации. Идея принятия решений на основе прецедентности – основополагающая в естественно-научном мировоззрении.

Будем считать, что все объекты или явления разбиты на конечное число классов. Для каждого класса известно и изучено конечное число объектов – прецедентов. Задача распознавания образов состоит в том, чтобы отнести новый распознаваемый объект к ка- кому-либо классу.

Задача распознавания образов является основной в большинстве интеллектуальных систем. Рассмотрим примеры интеллектуальных компьютерных систем.

1. 1)  Машинное зрение. Это системы, назначение которых состоит в получении изображения

через камеру и составление его описания в символьном виде (какие объекты

присутствуют, в каком взаимном отношении находятся и т.д.).

1. 2)  Символьноераспознавание–этораспознаваниебуквилицифр.
   1. Optical Character Recognition (OCR);
   2. Ввод и хранение документов;
   3. Pen Computer;
   4. Обработка чеков в банках;
   5. Обработка почты.
2. 3)  Диагностикавмедицине.
   1. Маммография, рентгенография;
   2. Постановка диагноза по истории болезни;
   3. Электрокардиограмма.
3. 4)  Геология.
4. 5)  Распознаваниеречи.
5. 6)  Распознавание в дактилоскопии (отпечатки пальцев), распознавание лица, подписи,

жестов.

**1.2 Проблема обработки информации**

**1.** В середине прошлого столетия наметился переход современного общества от пост-индустриального к информационному этапу своего развития.

И к 80-ым годам возник информационный взрыв.

Темпы этой угрозы с каждым годом увеличиваться.

Проблемы, вызванные информационным взрывом, хорошо иллюстрируют положение дел в издательской, почтовой, банковской и других сферах человеческой деятельности.

Прогресс “технологической” цивилизации и развития науки породили информационную проблему. (У первобытного общества в силу незначительного объема информации таких проблем не было).

2.

Уровень развития современного общества можно оценивать количеством генерируемой им информации и знаний.

Без информации общество в настоящее время существовать не может.

Потому что решение многих современных социальных проблем напрямую зависит от эффективного разрешения информационной проблемы.

Характерная особенность современного общества (технологически ориентированного - все большее количество людей и организаций занимаются обработкой информации и все меньшее – обработкой материальных объектов.

Возможности человеческого мозга воспринимать информацию, по мнению ученых, могут оказаться исчерпанными. Так по утверждению одного из профессоров психологии Гарвардского университета способности мозга, даже у самых способных из людей, в настоящее время близки к пределу.

3.

Информация – ключевой элемент процесса принятия решений, а количество разнохарактерной и разной степени сложности информации постоянно растет.

Крайне ощутимой стала проблема в создании информационных системах, способных собирать и быстро обрабатывать большие объемы информации.

В середине прошлого века стали активно развиваться вычислительные машины (одно из важнейших технических достижений 20 века). Они и стали средством решения быстро расширяющейся информационной бреши.

 Информационные системы стали важнейшим средством для эффективного сбора, хранения и использования информации и стали играть ключевую роль в процессах принятия решений, как на общегосударственном, так и на бытовом уровне.

4.

Окончательная оценка любых решений остается за человеком.  Однако появляющиеся концепции и методы, и использование вычислительных машин обеспечивают выдвижение и оценку такого количества вариантов, которые не один руководитель или простой человек не мог бы их осмыслить.

Возможность использовать методы анализа риска позволяют также оценивать потенциальную опасность, связанную с каждой из множества возможных стратегий.

Информационные системы обеспечивают возможность более ясно представлять результаты принимаемых решений.

5.

Таким образом, современный мир вступает в эпоху, когда человек приобретает нового партнера - созданные им информационные системы.

Новый вид сотрудничества позволяет человеку решать более сложные задачи, делает человеческое общество более совершенным, а жизнь людей значительно лучшей.

Однако одной из сложнейших проблем, стоящих на пути создания автоматизированных информационных систем, является автоматизация процесса распознавания образов.

6.

 Это область, изучением которой занимаются многие исследовательские группы (инженеры, специалисты в области вычислительной математики и информатики, физики, статистики, психологии, биологии, лингвисты и др.)

Каждая из таких групп посвящает свою работу какому-то определенному аспекту проблемы распознавания.

Цель данного курса – рассмотреть фундаментальные (теоретические) принципы, лежащие в основе разработки автоматических систем распознавания образов.

* В середине XX века общество начало переходить к информационному этапу развития.
* В 1980-х годах возник "информационный взрыв", и проблема информации стала нарастать.
* Проблемы из-за избытка информации заметны в разных сферах, таких как издательства и банки.
* Развитие науки и технологий привело к возникновению информационной проблемы, которой не было у древних обществ.

**Абзац 2**

* Развитие общества измеряется количеством создаваемой информации и знаний.
* Современное общество не может существовать без информации.
* Решение многих социальных проблем зависит от управления информацией.
* Технологическое общество: все больше людей работают с информацией, а не с материальными объектами.
* Мозг человека может не справляться с растущим объемом информации.

**Абзац 3**

* Информация важна для принятия решений, и ее объемы постоянно растут.
* Возникла проблема создания систем для быстрого сбора и обработки большого количества информации.
* В середине XX века начали развиваться компьютеры, чтобы решить эту проблему.
* Информационные системы стали важными для принятия решений на всех уровнях.

**Абзац 4**

* Люди принимают окончательные решения, но компьютеры помогают анализировать множество вариантов.
* Методы анализа риска помогают оценить потенциальные опасности различных стратегий.
* Информационные системы делают результаты решений более понятными.

**Абзац 5**

* Человек и информационные системы начинают работать вместе, что улучшает жизнь и решает сложные задачи.
* Одна из главных проблем – автоматизация процесса распознавания образов.

**Абзац 6**

* Многие специалисты работают над распознаванием образов (инженеры, математики, физики и др.).
* Каждая группа фокусируется на своем аспекте проблемы.
* Цель курса – изучить теоретические основы систем распознавания образов.

***постиндустриальн общ -> информационное общ***

*кибернетика               информация*

*первый этап в обработке информации - это классификация информации*

***ии*** *- раздел информатики, связанный автоматизация человеческого мозга*

***образ*** *- описание обьекта или процесса позволяющего выделять его из окружающей среды, отнести к классам*

***классы*** *- категории обьектов(больной/здоровый)*

​​

**1.2 Предмет распознавания образов, основные понятия и определения**

***1.***

*Распознавание образов* - повседневная неотъемлемая составляющая *деятельности человеческого мозга*.

***Распознавание образов*** – научная дисциплина, которая занимается разработкой методов для решения *задач классификации объектов* произвольной природы (естественных и технических явлений, процессов, сигналов и т. д.) на основе их *формализованного представления (*описания*)*.

В обыденном понимании образ (*понятие образа*) включает целую совокупность наших индивидуальных ощущений, представлений и умозаключений. Поэтому в спектре компьютерных дисциплин задачи распознавания относятся к проблематике *искусственного интеллекта*.

2.

В самом общем случае в качестве *образа* может рассматриваться любая *информационная модель* объекта или процесса абстрактного или реального мира.

Отличительная особенность такой модели в задаче распознавания   –использование *подмножества характеристик* объектов исследования, которое обеспечивает *выделение* одной или нескольких групп объектов совершенно *определенного типа*.

Целью процедуры (задачи) распознавания является ответ на вопрос: относится ли объект, описанный заданными характеристиками, к интересующим нас категориям и если относится, то к какой именно.

3.

Таким образом, ***образ*** - это описание объекта или процесса, позволяющее выделять его из окружающей среды и группировать с другими объектами или процессами для принятия необходимых решений.

Категории объектов, которые необходимо выделить или на которые хотим разделить все множество образов в процессе распознавания, называют ***классами***.

*Понятие класса* в распознавании появилось существенно раньше, чем в объектных информационных моделях. И способ задания классов здесь определяется не столько сущностью предмета исследования, сколько особенностями имеющейся информации об объектах и способами ее представления.

3а.

Для системы обработки информации ***образ* -** это совокупность данных об объекте или явлении, включающая параметры и связи.

*Параметры* представляют собой количественные характеристики, полученные с помощью измерительных систем или математических моделей.

*Связи* могут описывать как внутреннюю структуру образа, так и особенности его поведения, когда образ является динамическим объектом или процессом.

4.

Любой алгоритм распознавания можно представить как:

 абстрактную функциональную систему ***R***, из трех компонент ***R={A,S,P}:***

-      *A={Ak }, k=1,...,h* – алфавит классов – множество категорий, по которым необходимо все  распределить образы,

-   *S={Sj}, j=1,...,n* - словарь признаков - множество характеристик, из которых составляется описание образа,

-       *P={Pl}, l=1,...,m* - множество правил принятия решения.

5.

*Процесс функционирование* *системы* ***R*** сводится к следующим действиям:

-на вход подается образ – некоторая конфигурация элементов множества *S*,

-к конфигурации применяется определенная последовательность правил из *Р*,

-в результате конфигурации присваивается индекс, соответствующий одному из элементов множества *A.*

Качество функционирования системы определяется тем, насколько часто присвоенный образу индекс совпадает с ожидаемым нами результатом.

6.

Компоненты системы разделяются на две составляющие:

-  *A,S*  представляют собой информационную часть,

-   *P* – методологическую.

*Понятие* ***класса*** (смысл) зависит от*способов описания образов* (для различных способов описания будет различным).

*Способ описания* ***образа*** зависит от *физической природы объектов* распознавания и возможностей формализации, соответствующих им понятий.

*Методы принятия решений* взаимосвязаны со способом представления объектов распознавания.

7.

Любая система распознавания включает два основных процесса (2 задачи):

-  ***синтеза образов*** - формирование описаний объектов и классов;

-  ***анализа образов*** - процесс принятия решений.

В зависимости от особенностей информационных компонент системы *R,* выделяют *три подхода* (основанные на трех принципах) к решению задачи распознавания образов:

- принцип *сравнения с эталоном*;

- принцип *кластеризации*;

- принцип *общности свойств*.

**1.3. Принципы построения систем распознавания.**

**Принцип сравнения с эталоном.**

*Для каждого класса задан формальный эталон  
Все основано на сопоставлении(мере)  
ввести меру подобия для этого, например, расстояние (Евклидово, Хейдово?); корреляцию (то есть угол чем он менше тем обьекты ближе)*

*Найти адекватную меру близости/меру различия  
min P()*

*Примеры: распознавания печатных шрифтов, букв*

Применяется, когда каждому классу Ak можно сопоставить конечный набор эталонных образов.

Процесс распознавания заключается в простом сопоставлении образов, поступающих на вход распознающего устройства или алгоритма, с эталонами классов Ak , на основе выбранной меры сходства.

Принцип сравнения с эталоном - один из первых подходов возникших, когда возможности   вычислительных устройств были весьма ограничены.

Применяется в аналоговых и аналогово-цифровых системах распознавания.

Пример применения - распознавание печатных шрифтов, распознавание текстур, выделение объектов определенной формы.

**Принцип кластеризации.**

Применяется, когда признаки представляют собой наборы измерений (параметров) без заданных взаимосвязей и образ представлен *n-*мерным вектором в признаковом пространстве ***Х***:

каждому классу ***Ak*** в пространстве ***Х*** сопоставляется множество векторов.

В результате *признаковое пространство* разбивается на *области*, соответствующие классам *Аk,* которые называют *кластерами или таксонами*.

*Применение.* Принцип кластеризации (таксономии) широко применяется при обработке количественных данных, в частности, в системах *компьютерного анализа* многозональных и спектрозональных аэрокосмических изображений (классификация по спектральным признакам).

как решать задачу?  
построить формальную границу собственной области каждого класса

граница - область решений

Примеры: обработка кол-ых данных, распознавнаие аэрокосмических изображений

**Принцип общности свойств**

1. **Применение при больших множествах образов**:
   * Когда множество образов (примеров) каждого класса слишком велико для прямого анализа, применяется этот принцип. Вместо анализа каждого отдельного образа, рассматриваются только характерные свойства, которые достаточно различают классы.
2. **Выявление отличительных особенностей**:
   * Из больших множеств образов выделяются конечные выборки, на основе которых выявляются характерные особенности или свойства, отличающие один класс от другого.
3. **Кодирование и хранение свойств**:
   * Выявленные свойства классов кодируются с использованием подходящей модели (например, математической или статистической) и хранятся в памяти в виде структур, функций или отношений. Это делает последующий анализ более эффективным.
4. **Процесс распознавания**:
   * В процессе распознавания происходит анализ нового образа. Выявляются его свойства, которые затем сопоставляются со свойствами, хранящимися в памяти для различных классов (обозначенных как Ak).
5. **Обобщающее свойство – алгоритм порождения образов**:
   * Обобщающим свойством может быть сам алгоритм, который порождает образы. В этом случае классы образов задаются алгоритмами, генерирующими структуры определенного вида. То есть, классы образов могут быть описаны алгоритмами, которые определяют, каким образом создаются структуры этих образов.

**Пример пояснения**

Предположим, мы хотим распознавать типы транспортных средств на изображениях (автомобили, мотоциклы, автобусы). Принцип общности свойств можно применить следующим образом:

1. **Выборка и анализ**:
   * Вместо анализа всех возможных изображений транспортных средств, мы берем выборки изображений каждого типа и выявляем характерные особенности (например, форма, количество колес, размеры).
2. **Кодирование свойств**:
   * Выявленные особенности (например, "автомобили обычно имеют четыре колеса и определенную форму кузова") кодируются и хранятся в системе.
3. **Распознавание**:
   * Когда система получает новое изображение, она анализирует его и определяет свойства (например, "это транспортное средство имеет четыре колеса и форму, похожую на автомобиль").
   * Система сравнивает выявленные свойства с закодированными свойствами классов (автомобилей, мотоциклов, автобусов) и выбирает наиболее подходящий класс.
4. **Алгоритм порождения**:
   * Если обобщающее свойство – алгоритм порождения образов, то классы могут задаваться тем, как эти образы создаются. Например, "автомобильные образы порождаются алгоритмами, которые формируют кузов с четырьмя колесами", "мотоциклы – с двумя колесами" и т.д.

свойства классов записыватются на основе функц зависимости  
как решать задачу?  
для обьекта проверяется выполняется ли это свойство или нет (например, стул это или нет)

**1.4. Методология построения систем распознавания**

Перечисленные подходы относятся к процессу ***синтеза образов***, к способу представления информации компонент *A и S* системы *R*.

*Методы* ***анализа образов*** - правила принятия решений, в большинстве случаев взаимосвязаны со способом синтеза образов.

В методологии принятия решений при распознавании образов выделяют три основных направления:

- эвристические методы;

- математические методы;

- лингвистические (синтаксические) методы.

**12.**

**Эвристические методы**

Основываются на опыте и интуиции разработчика системы распознавания.

Методы ориентированы на решение конкретного типа задач распознавания и непосредственно привязаны к способу синтеза образов.

Часто применяются при использовании принципа сравнения с эталоном и принципа общности свойств.

**13.**

**Математические методы**

Опираются на использование классического математического аппарата: методов линейного программирования, корреляционного анализа, теории статистических решений и т.п.

Математические методы применяются в случаях, когда признаки представлены измерениями (параметрами), а их связи могут быть описаны в виде аналитических зависимостей.

Подразделяют на детерминистские и статистические.

Статистические методы применяют в задачах анализа данных ДЗ. Это обусловлено высокой степенью изменчивости условий наблюдения и самих объектов исследования, а также потерей информации на разных этапах проведения ДЗ.

**14.**

**Лингвистические (синтаксические, структурные)** методы

Применяются, когда образ представляет собой некоторую структуру, состоящую из так называемых непроизводных (первичных) элементов и признаков, описывающих связи между ними.

В методах широко используется аппарат алгебры логики и теории формальных языков.

Применяются при обработке данных ДЗ и в геоинформационных технологиях. Использоваться при анализе пространственных структур, некоторые из них применяются также в экспертных системах.

15.

Характерной особенностью задач распознавания образов является ***слабая*** их ***формализация*** или отсутствие строгого обоснования используемых методов.

В распознавании образов можно выделить два основных направления:

- *теоретическое* - разработка принципов и методов распознавания;

- *прикладное* - решение конкретных практических задач, разработка программных, технических средств и технологий.

16.

Исходным в распознавании образов является понятие **объекта**, который описывается конечным набором числовых или нечисловых характеристик (**признаков**).

На заданном множестве объектов предполагается существование некоторого свойства, в соответствии с которым объекты объединяются в подмножества (**классы**) и объекты различных классов отличаются друг от друга.

Свойство, как правило, неизвестно, и может задаваться набором ”типичных” объектов из каждого класса. Такой набор ”типичных” объектов называют **обучающей выборкой**.

16.

***Методы распознавания образов*** широко применяются в *слабо формализованных* с точки зрения математики областях науки и техники:

*медицине, геологии, социологии, технической диагностике  и т. д.*

***Задачу*** ***распознавания образов*** можно трактовать как *“качественную” задачу принятия решений*, в которой действие определяется на основе оценки ситуации.

С использованием *методов распознавания образов* решаются многие проблемы *информатики и искусственного интеллекта*:

- ввод *звуковых, текстовых* и др. данных;

- построение *информационных  автоматизированных и робототехнических* систем;

- *новых информационных технологий и т.д.*

*-эвристические методы(логические приемы возникающие из опыта)  
принцип сравнение с эталоном, принцип общности свойств  
-математические методы(классический мат аппарат)*

*-лингвистические методы(логика и теория формальных языков грамматики)*

*Задачи распознания образов возникают там где мат аппарат слаб(слабая формализация)*

**1.5. Геометрическая интерпретация задачи распознавания.**

**​​​​**

Первым в последовательности приема, хранения и обработки информации во всякой технической системе или живом организме является этап распознавания.

На этом этапе определяется, с объектами какого класса осуществляется общение.

В соответствии с геометрическим подходом любой ***исследуемый объект*** можно представить ***точкой (вектором)*** в пространстве измерений Компоненты вектора определяют ***количественную характеристику*** из *п* ***параметров (признаков).***



2.

Признаки могут быть получены либо в результате измерения, либо предварительного преобразования данных наблюдений.

Если измерения проведены, то не имеет значения, какова физическая сущность объекта:

Набор полученных признаков задает его ***образ - описание объекта в виде вектора признаков.***

Каждый из объектов принадлежит к одному из конечного числа ***классов*** объединяющих объекты с некоторой общностью свойств.



3.

Концы векторов любого класса располагаются в пространстве внутри ***собственной области*** (рис. 1).



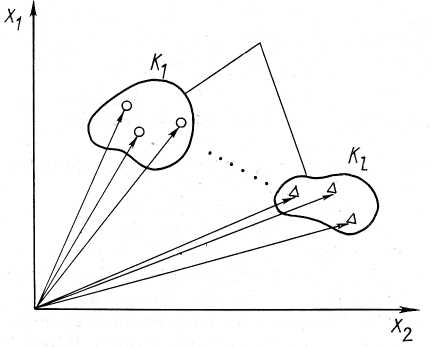


Рис. 1. Собственные области классов

4.

Эти области, как правило, пересекаются друг с другом,

(*причины:* неполнота описания объектов, искажения значений отдельных признаков при измерении и т. д).

***Задачей распознавания*** является:

отнесение предъявленных объектов с неизвестной классификацией к одному из рассматриваемых классов.

Распознавание производится по признаку попадания конца вектора в соответствующую ***область решения,*** которая некоторым образом аппроксимирует собственную область данного класса.

5.

Области решений для каждого из классов строят как области пространства в соответствии с ***определенными критериями*** и на основе ***ранее полученных данных***.



Из условий однозначности классификации объектов, ***области решений*** формируются как ***непересекающиеся.***

Вводится нуль-класс как множество точек , не вошедших ни в одну из областей решения классов (неопознанные объекты).



***Функции,*** описывающие поверхности, разделяющие области решений, ***называются решающими***.

**6. Пример.** Рассмотрим простейшую модель системы распознавания, типичную для дистанционного зондирования.

На борту летательного аппарата (самолета или спутника) размещен многоспектральный сканер, который служит датчиком для съема информации.

Выходом такого датчика является набор из *п* измерений, каждое из которых соответствует одному из каналов сканера.

**7.**

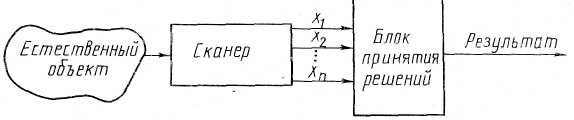


Рис.2 Модель системы распознавания образов

8.

В результате объект наблюдения представляется *n*-мерным вектором *.*



Блок принятия решений (классификатор) относит каждый такой вектор к одному из предварительно заданных классов в соответствии с некоторым классифицированным правилом.

В качестве классов, например, могут задаваться различные типы покрова земной поверхности: растительность, вода и т. д.

9.

Измерения нескольких объектов, условно относящихся к одному типу покрова, но полученных для различных участков Земли при различных условиях, представляются не одной точкой, а локализуются в пространстве измерений в виде кластера (рис.3)

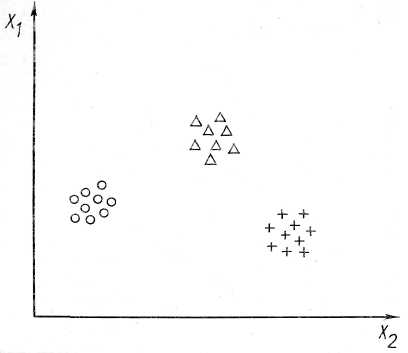


Рис.3 Кластеры объектов данных

10.

Это является результатом

- присущих природе случайностей,

- шума аппаратуры дистанционного зондирования

и т. д.

Вместе с тем кластеры с элементами определенного типа покрова обычно более или менее различимы.

В результате можно установить связь между локализованными ***областями пространства измерений*** (признаков) и конкретными ***типами земного покрова***.

11.

Для построения ***классификатора*** системы распознавания, необходимо:

разбить пространство измерений на области решений так, чтобы каждая область соответствовала конкретно различимому классу.

Под классификатором понимается устройство или алгоритм, отображающие ***образ объекта*** в числа которыми пронумерованы выделенные классы *.*



Идеология распознавания образов позволяет характеризовать, в частности, множество данных дистанционного зондирования путем ***разделения пространства измерений на непересекающиеся области решений***.

12.

Таким образом, определяются спектральные классы, которые разделяются благодаря спектральным свойствам соответствующих типов земного покрова.

Как показывает практика, ***спектральные классы совпадают с информационными***, с классами покрова земной поверхности, имеющими определенный смысл (вид сельхозкультур, тип по и т.д.).

Спектральные классы должны быть преобразованы в информационные путем идентификации типа покрова, соответствующего каждому спектральному классу.

***13.***

***Задача построения классификатора системы распознавания***

Построение оптимальных решающих процедур, необходимых для проведения идентификации или классификации:

- необходимо соответствующим образом разбить пространство признаков на области решений;

- создать классификатор, который отождествляет вектор измерений с классом, соответствующим той области решения, в которую он попал.

***14.***

***Общая методика решения задачи распознавания***

Пусть

- задано *l* классов  *;*



- определены ответствующие им области решений;

- выбрано *l* функций *.*



Функции выбираются таким образом, что всякая принимает большее значение по отношению к остальным функциям всегда, когда X принадлежит области решения класса.



15.

Если предположить, что функции непрерывны на границах областей,



то поверхности, разделяющие смежные классы и можно определить уравнением вида



.



Такие функции, как мы уже отмечали, ***называются решающими***.

Выбирая вид функции, можно строить различные по сложности разделяющие поверхности: *линейные, кусочно-линейные, полиномиальные и т. д.*

16.

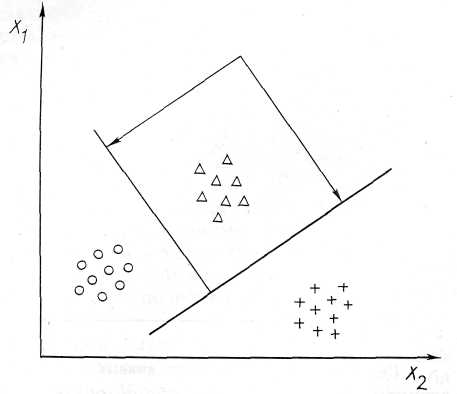


Рис. 4. Области и границы решений

17.

На рис.4 представлен простейший случай разделения линейными гиперплоскостями.

Решающие функции имеют вид



После определения величины , их можно использовать для построения классификатора.



Если, например, необходимо классифицировать произвольный объект Х

- вычисляют значения функций ;



- объект заносят в класс, решающая функция которого имеет наибольшее значение.

18.

Правило классификации можно записать следующим образом:

Пусть зафиксирован некоторый класс *.*



Полагают, что тогда и только тогда, когда



Чтобы исключить случай равенства двух и более значений решающих функций, вводят нуль-класс



При отнесении объекта к такому классу он считается неопознанным (система распознавания дает отказ).



19.

Блок принятия решений, с учетом введенных понятий, имеет вид рис. 5.

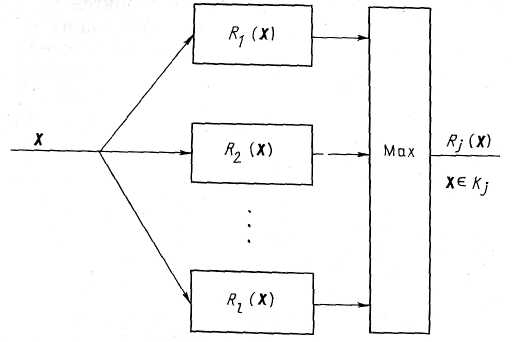


Рис. 5. Блок принятия решений в системе распознавания

**Основная идея**

1. **Этап распознавания**: Это первый этап в процессе обработки информации. Здесь определяется, к какому классу относится объект. С геометрической точки зрения объект представляется как точка (вектор) в многомерном пространстве признаков.
2. **Признаки объектов**: Признаки объектов могут быть получены либо путем измерений, либо через предварительное преобразование данных. Эти признаки формируют вектор, описывающий объект. Каждый объект принадлежит к одному из нескольких классов, объединенных общими свойствами.
3. **Расположение векторов**: Векторы объектов каждого класса располагаются в пространстве измерений в виде собственных областей классов. Эти области могут пересекаться из-за неполноты описания и искажений при измерениях.
4. **Задача распознавания**: Цель – отнести объект с неизвестной классификацией к одному из классов, используя попадание вектора объекта в соответствующую область решения. Эти области строятся на основе определенных критериев и данных.
5. **Области решений**: Области решений для каждого класса строятся как непересекающиеся области. Нуль-класс вводится для точек, не попавших ни в одну область решений. Функции, разделяющие области решений, называются решающими.
6. **Пример системы распознавания**: В системах дистанционного зондирования используются многоспектральные сканеры, которые создают многомерные векторы измерений для классификации объектов (например, типов земной поверхности).
7. **Модель системы распознавания**: Объект представляется n-мерным вектором. Классификатор относит каждый вектор к определенному классу на основе правил классификации. Измерения нескольких объектов, относящихся к одному классу, формируют кластеры в пространстве измерений.
8. **Кластеры объектов**: Измерения объектов формируют кластеры из-за случайностей и шума. Эти кластеры помогают установить связь между признаками и типами объектов (например, типами земного покрова).
9. **Построение классификатора**: Необходимо разбить пространство измерений на области решений, соответствующие различимым классам. Классификатор отображает образ объекта в числа, которые пронумерованы по классам.
10. **Методика решения задачи распознавания**: Задача распознавания включает в себя выбор функций, которые максимизируются для соответствующих классов, когда объект принадлежит их области решений. Решающие функции описывают поверхности, разделяющие области решений.
11. **Пример классификации**: На рис. 4 показано простейшее разделение линейными гиперплоскостями. Решающие функции имеют вид линейных уравнений. Для классификации объекта вычисляют значения решающих функций и определяют класс по максимальному значению.
12. **Правило классификации**: Правило записывается как выбор класса с максимальной значением решающей функции. Вводят нуль-класс для объектов, не попавших ни в одну область решений, что означает неопознанные объекты.
13. **Блок принятия решений**: Представлен блок принятия решений в системе распознавания на рис. 5, который визуализирует процесс классификации.

***Синтез и анализ образов***

* ***Синтез образов: Процесс создания векторов признаков, которые описывают объекты. Эти векторы формируются из измерений или преобразованных данных и представляют объекты в многомерном пространстве.***
* ***Анализ образов: Процесс интерпретации и классификации векторов признаков. Включает определение, к какому классу принадлежит объект, на основе его вектора признаков и областей решений.***

*Связь между синтезом и анализом: Синтез создает векторы признаков для объектов, а анализ интерпретирует эти векторы, определяя класс объекта на основе решающих функций и областей решений.*

**1.6 Классификация задач.**

**находим функцию которая представляет наши данные и применяем ее к остальному множеству - основная в ТРО**

**Обучающая выборка- совокупность объектов классов, для которых известна принадлежность к классам.(известна их функ зависимость)**

1.

Для построения ***областей решений*** (явно или неявно) могут использоваться следующие ***виды данных***:

- ***критерий оптимальности***, которому удовлетворяют области решения; процент правильно распознанных обьектов по отношению к неправильным

- ***совокупность параметров*** (признаков) исследуемого объекта:

- количество и условные ***обозначения*** ***классов***;

- указатель ***принадлежности*** объектов к классам.

2.

В задачах распознавания свойство классов, как правило, неизвестно, и может задаваться некоторым набором ”*типичных*” объектов из каждого класса.

Такой набор ”типичных” объектов называют ***обучающей выборкой.***

В зависимости от того, какие из перечисленных выше данных используются, решаются *задачи распознавания*, основанные:

- на предварительном обучении (обучение с учителем);

- на самообучении (распознавание без учителя).

3.

***Задача распознавания с обучением***

При наличии всех перечисленных выше условий получаем задачу распознавания с обучением.

Свойство классов в этом случае задаваться набором ”типичных” объектов из каждого класса, то есть  *обучающей выборкой***.**

Задача распознавания в этом случае формулируется следующим образом:

Для *произвольного объекта* необходимо определить, используя обучающую выборку, его принадлежность к одному из заданных классов.

Такая задача называется *задачей с обучением (с учителем)* и считается основной в теории распознавания образов.

4.

В данном случае *решающие функции*, с помощью которых определяются *разделяющие поверхности*, строятся на основе информации, выделенной из набора обучающих образов (векторов измерений с известной классификацией - типичных представителей классов).

*Критерий оптимальности*, которому должны удовлетворять формируемые по обучающей выборке области решений, определяется характером задачи распознавания.

Типичным показателем качества является вероятность правильного распознавания.

На практике эта величина обычно заменяется ее оценкой ‒ долей правильно распознанных объектов обучающей выборки.

5.

***Решение задачи распознавания*** ***с обучением*** производится в два этапа.

Этап ***обучения.***

По данным обучающей выборки в соответствии с заданным критерием строят *области решения*.

Этап ***распознавания.***

Предъявляются объекты с неизвестной классификацией, в отношении которых принимается решение о принадлежности их к классу, в область которого они попадают.

6.

При решении задачи распознавания с обучением принимаются следующие допущения (предположения):

- Построенный аглоритм обладает хорошей обобщающей способностью

что качество распознавания объектов обучающей выборки с помощью построенных областей решения *в достаточной степени сохраняется*и для объектов с неизвестной классификацией.

- Обучающая выборка является представительной

что при правильно выбранных:

   - *формы областей решения*;

   - *критерия оптимальности;*

*- объема и состава обучающей выборки,*

построенные области окажутся удовлетворительными при классификации объектов на этапе распознавания (допущение о *представительности обучающей выборки*).

7.

Выбор формы областей решения (*типа решающей функции*) во многом определяет успех решения задачи в целом.

Целесообразно (на практике) стремиться к *простым областям решения*, что упрощает реализацию задачи на ЭВМ и повышает степень уверенности в принятых допущениях.

Выбор сложных конфигураций (в стремлении повысить качество распознавания) допущения (из-за детального учета индивидуальных особенностей обучающей выборки) могут оказаться невыполненными.

***Задачу распознавания с обучением*** называют контролируемой классификацией.

Исследователь, определяя обучающие выборки, в некотором смысле контролирует получение границ решения и как бы обучает классификатор распознаванию информационных классов.

***8.***

***Задача распознавания, основанная на самообучении***

*Здесь задано только множество обьектов. Нужно найти естественную группировку обьектов по признакам, зная только сам набор обьектов.*

*Как решать такие задачи?*

*Выдвигается гипотеза, что обьекты одного класса имеют совокупные признаки. То есть вводится мера близости. И выдвигается гипотезза о том, что расстоянние между обьектами одного класса имеют расстояние между собой более близкое, чем по отношщению к другому классу.*

*Задача выбора системы признаков:*

*Инфор модель - набор признаков объекта*

*Решается экспертным путем.*

*Есть предметная область и есть эксперт, который изучает ее.*

*Пример эксперт врач в эпидемию. У врача два класса k1 - болен, k2- здоров. Обьекты - пациенты, разделяются на два класса. Следующий шаг - определить набор признаков - симптомов. Эксперт 1) выделяет классы и 2)определяет альфавит признаков/сво-в, потом 3) устанавливает обучающую выборку. Решает задачу распознавания и по качеству распознавания определеляет является ли этот набор признаков информативным. Основная проблема найти меру важности признаков и затем упорядочить их по убыванию. И затем эти признаки выбираются и строится модель инф.*

*Замечание. Ни один алгоритм не компенсирует небрежный выбор признаков.*

*Допущения:*

*1 На качество влияет информативность признаков.*

*2 Обучающая выборка является представительной.  
3 Построенный аглоритм обладает хорошей обобщающей способностью.*

Данный класс задач приходится решать, когда имеются исходные данные первых двух (или трех) перечисленных выше видов данных.

В этом случае *задача распознавания* формулируется следующим образом:

Задан набор объектов, который необходимо разбить (используя только описание самих объектов) на конечное число классов.

Такие задачи называется *задачами с самобучением (без учителя, кластеризации или таксономии).*

9.

Данная задача может использоваться как вспомогательная для, ***задачи распознавания с обучением,*** когда необходимо построить, или уточнить обучающую выборку.

В этом случае имеем соответственно задачи распознавания объектов:

- когда количество классов известно;

- когда количество классов неизвестно.

Результатом классификации в первом случае является распределение объектов по классам, а во втором ‒ еще и их количество.

10.

При решении данного типа задач объекты в качестве критерия выбираются функции, которые связывают различное местоположение объектов в пространстве с их принадлежностью к тому, или иному классу.

Чем точнее учитывается эта связь, тем выше получаемое качество распознавания.

Задачи распознавания с самообучением называют ***неконтролируемым анализом -*** исследователь мало влияет на установление границ решений.

Задач могут использоваться:

- для распределения объектов по различным классам, когда отсутствуют обучающие выборки;

- для анализа структуры собственных областей классов (проверить их много или одно связность).

11.

***Задача выбора системы признаков для описания объектов распознавания***

*Выбор совокупности признаков* объектов - серьезная научная проблема, предваряющая разработку методов распознавания образов.

Отбор признаков связан с качеством классификации: необходимо выбирать наиболее «информативные» признаки.

Решение задачи основано на понятии *меры информативности признака* (или набора признаков) с точки зрения различения классов.

***Замечание.*** Использование наилучшего классификатора объектов не может компенсировать небрежный выбор *признаков*.

!2.

Разработано большое количество алгоритмов и методов решения задач распознавания.

Однако не существует универсальных рекомендаций об использовании того или иного алгоритма или метода.

 Все зависит от априорной информации о предметной области, от интуиции и опыта исследователя и т. д.

Жизнеспособность алгоритмов, как правило, определяется экспериментально.

Методы решения задач определения признаков (для описания объектов) и снижения размерности векторов образов являются специфическими и существенно зависят от предметной области.

**9. Линейный классификатор.**

***Линейный классификатор* — это один из базовых алгоритмов в машинном обучении, используемый для классификации объектов на различные категории или классы.**

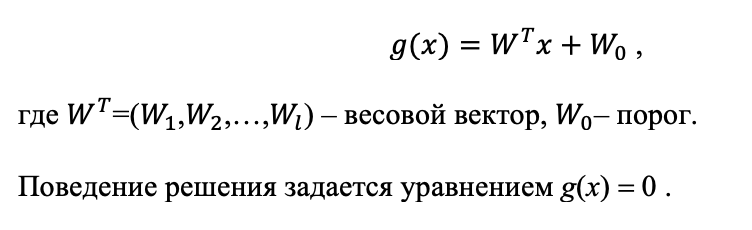
**G = w1x1+w2x2+…+wnxn - параметрическое семейство гиперплоскостей**

**Где w - угол поворота гиперплоскости**

**А свободные члены - расположение гиперплоскости относительно 0**

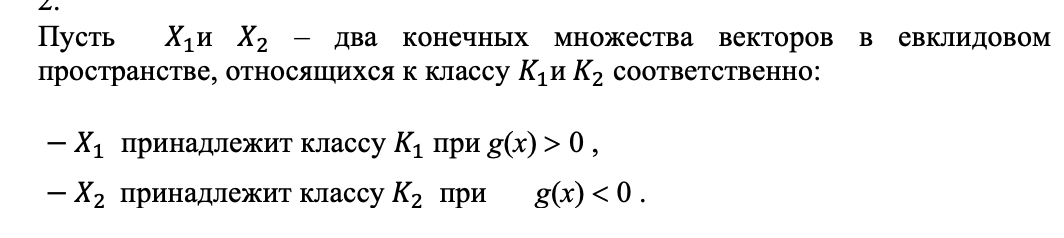
***Линейный классификатор***  строится на основе линейной разделяющей гиперповерхности.

Главным достоинством линейного классификатора является его простота и вычислительная эффективность.

Рассмотрим задачу построения линейной разделяющей гиперповерхности в  виде линейной дискриминантной функции:  


Поведение решения задается уравнением *g*(*x*) = 0 .

2.



***Задача***: Необходимо установить разделимость множестви ,

то есть найти гиперплоскость, разделяющую эти множества.

Рассмотрим решение на примере двумерной задачи, когда образы представляются точками на плоскости.

***3.***

***Определение.*** Множество, содержащее отрезок, соединяющий две произвольные внутренние точки, называются выпуклыми

***Определение.*** Выпуклая оболочка – это минимальное выпуклое множество, содержащее данное множество.

***Утверждение 1.*** Два множества на плоскости линейно разделимы тогда и только тогда, когда их выпуклые оболочки не пересекаются.

Из утверждения следует очевидная процедура проверки разделимости множеств на плоскости.

***4.***

***Процедура 1*** (проверка  условия разделимости)

Шаг 1. Построить выпуклые оболочки.

Шаг 2. Проверить пересечение выпуклых оболочек.

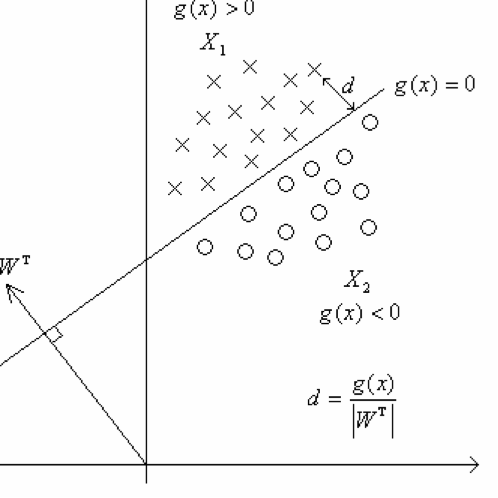
Если они не пересекаются - множества разделимы.

***Процедура 2*** (построение разделяющей прямой)

Шаг 1.  Найти ближайшую пару точек в выпуклых оболочках обоих множеств.

Шаг 2. Построить срединный перпендикуляр к отрезку, соединяющему эти точки.

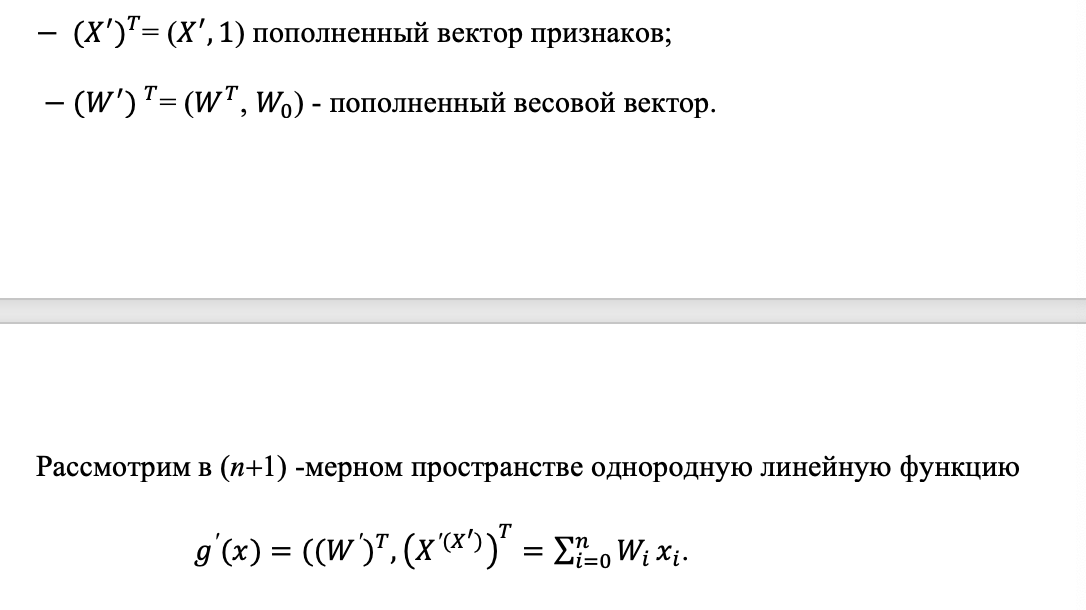
Этот перпендикуляр и будет разделяющей прямой.



6.

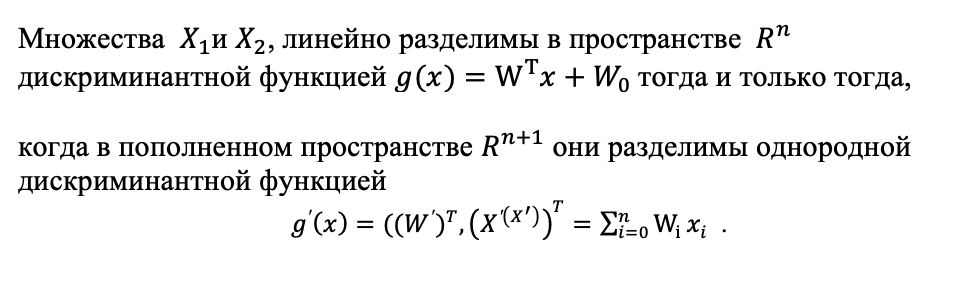
Пусть размерность векторов признаков *X* и коэффициентов *W* равна *n*.

Рассмотрим «пополненные» вектора *X* ¢, *W*¢  следующего вида:



Справедливо следующее утверждение.

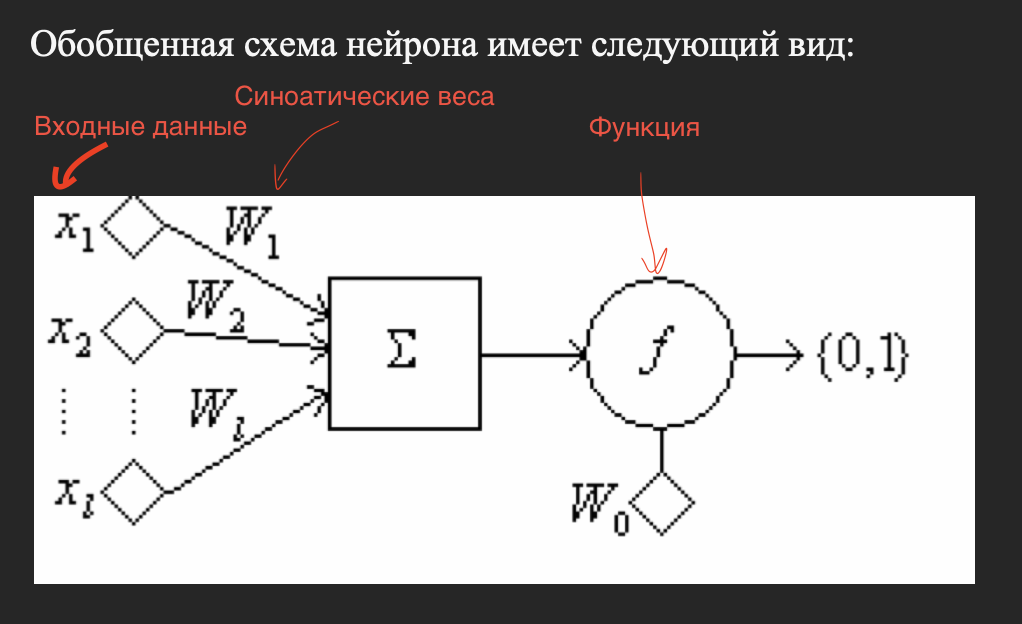
***Утверждение 2***.

Множества                         .

**10. Математическая модель нейрона.**

**Математическая модель нейрона - это формальное математическое представление, описывающее поведение биологического нейрона.**

​​Обобщенная схема нейрона имеет следующий вид:



9.

- x\_1,x\_2,…,x\_n– компоненты вектора〖 x=(x〗\_1,x\_2,…,x\_n);  – сумматор;

- W\_1,W\_2,…W\_n  –синоптические веса; W\_0– порог;  f – функция активации.

Выходом сумматора является величина  ∑\_(i=0)^n▒W\_i  x\_i ,

∑\_(i=0)^n▒W\_i  x\_i является входом (аргументом) функции активации.

Значение функции активации вычисляется на основе определения знака суммы    ∑\_(i=0)^n▒W\_i  x\_i+W\_0:

                         f(v)=0    если    v<0  и  f(v)=1 если    v>0.

10.

Таким образом, нейрон - линейный классификатор с дискриминантной функцией

g(x)=∑\_(i=0)^n▒W\_i  x\_i+W\_0

Задача построения линейного классификатора сводится к задаче обучения нейрона, т.е. подбору соответствующих весов и порога.

Обучение состоит в коррекции весов W\_1,W\_2,…W\_n  и порога  W\_0.

**11. Алгоритм персептрона.**

Алгоритм персептрона - последовательная итерационная процедура.

Общий шаг состоит в предъявлении нейрону очередного вектора-прецедента и коррекции весов по результатам классификации.



Прецеденты предъявляются циклически - после предъявления последнего снова предъявляется первый.

Процесс обучения заканчивается, когда нейрон правильно классифицирует все прецеденты.

Пусть весовой вектор после *k-й* итерации, – прецедент, предъявляемый на *k-й* итерации.



***Общий k-й шаг алгоритма***:

Если ;



Если +;



Если ;



Если +





***g (x) – дискриминантная функция*** после *k*-го шага алгоритма;

***– весовой вектор*** после *k*-го шага алгоритма.



***Основной вопрос*** связан со сходимостью алгоритмом персептрона:

конечен ли итерационный процесс обучения?

***Теорема Новикова.*** Пусть - {} – бесконечная последовательность векторов из двух непересекающихся множеств ;

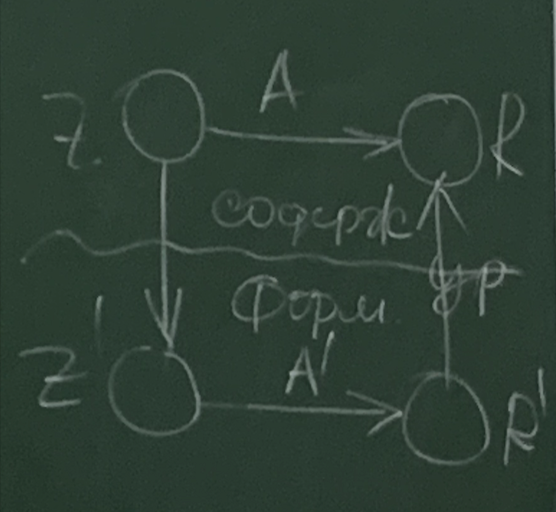


- существует гиперплоскость, проходящая через начало координат и разделяющая (не имеющая с ними общих точек).



Тогда число шагов алгоритма персептрона (число коррекций весового вектора) конечно.

**12. Формальная постановка задачи с обучением.**

****

Пусть: - -система характеристик (признаков), описывающих объект , - - множество допустимых значений -ого признака . Набор , где ,,

называется ***стандартным описанием допустимого объекта .***

-- признак выполнен или не выполнен (бинарные признаки);

-, элемент - означает отсутствие информации о данном признаке;

- - признак имеет несколько градаций, ;

- - признак принимает конечное число значений;

- - значениями признака являются функции;

-- значениями признака являются функции распределения не­которой случайной величины (возможны и другие множества ).

Пусть - множество допустимых объектов , для элементов которого определено описание.

Предположим, что множество  покрыто конечным числом подмножеств, называемых классами, и .

Пусть также задана совокупность допустимых объектов

,,

среди которых есть представителя всех классов.

Каждому  поставлен в соответствие вектор ,

где - значение элементарного предиката, которое означает вхождение или не вхождение объекта  в класс , .

 называется ***информационным вектором***  по разбиению .

Совокупность множеств ;, называется ***стандартной обучающей информацией*** и обозначается .

**Формальная постановка задачи**.

Пусть задана обучающая информация о классах  и описание допустимого объекта .

Требуется установить, используя только информацию  и описание , к каким клас­сам из  принадлежит объект .

Алгоритм  называется распознающим, если, он по паре  (где *S*- допустимый объект) вычисляет вектор

, где .

Требуется построить алгоритм , для которого

.

 называется *классификационным* вектором объекта  в алгоритме .

1. **Статистические алгоритмы.**

Рассмотрим подход к построению алгоритмов распознавания, основанный на использовании статистического анализа.

Статистический анализ играет в распознавании важную роль в случаях, когда случайные факторы оказывают влияние на порождение объектов.

***Пример.***

Случайный характер природных процессов, шум аппаратуры дистанционного зондирования, некоторая неопределенность относительно правильной идентификации объектов обучающей выборке и т. д. ‒ факторы, позволяющие рассматривать набор признаков *,*  как п-мерную случайную величину, а принадлежность объекта к какому-либо классу с некоторой долей вероятности.



2. *Статистический анализ* предоставляет возможность:

построения алгоритма оптимального в том смысле, что его применение обеспечивает в среднем наименьшую вероятность совершения ошибок распознавания.

В методах *статистического распознавания* обычно используют:

функции распределения вероятностей, связанные с классами объектов. Эти функции, как правило, не известны и должны оцениваться с помощью обучающей выборки.

Покажем, как можно использовать теорию статистических решений для построения *алгоритмов распознавания*.

***Пусть:***

- имеется классов ;



- объект ‒ случайная *n*-мерная величина.



***Сделаем следующие предположения:***

) - известна или может быть оценена;



*-* известна или может быть оценена по  
обучающей выборке;



- не известна.



‒ априорная вероятность класса ,вероятность наблюдения объекта из независимо от любой другой информации;



‒ функция плотности вероятностей, зависящая от объекта при условии, что он принадлежит ,



‒ вероятность принадлежности к классу (апостериорная вероятность).



Если алгоритм распознавания (классификатор) принимает решение, что объект , а он принадлежит ,то классификатор несет потери .



Так как может принадлежать любому из классов, то матожидание потерь, связанных с решением можно определить выражением:



Эта величина называется *условным средним риском*.

Плотность распределения (или плотность вероятности) — это функция, которая описывает вероятность того, что случайная величина примет определённое значение в заданном интервале. В контексте непрерывных случайных величин плотность распределения позволяет оценить вероятность попадания случайной величины в любой конкретный интервал значений.

Математическое ожидание — это основная характеристика случайной величины, которая описывает её среднее значение в долгосрочной перспективе или ожидаемый результат при повторении эксперимента множество раз. Формально математическое ожидание случайной величины XXX, обозначаемое как E[X]\mathbb{E}[X]E[X] или μX\mu\_XμX​, определяется следующим образом:

1. **Нормальное распределение (Гауссово распределение)**: Это одно из основных распределений, используемых в статистическом анализе для моделирования случайных величин. В контексте распознавания образов, когда объекты признаются на основе их признаков, предполагается, что признаки объектов могут быть распределены нормально в пространстве признаков. Это позволяет использовать параметры таких распределений для оценки вероятностей принадлежности объектов к различным классам.
2. **Метод опорных векторов (SVM)**: Этот метод также используется в задачах классификации объектов. SVM строит гиперплоскость в многомерном пространстве признаков, которая максимально разделяет объекты разных классов. Он является не только методом линейной классификации, но и может использоваться с различными функциями ядра для работы в нелинейных пространствах признаков. В контексте статистического распознавания SVM также может быть использован для минимизации вероятности ошибок классификации, что соответствует идее оптимального алгоритма, упомянутой в тексте.

Отношение нормального распределения и метода опорных векторов к тексту заключается в том, что оба метода могут быть применены для статистического анализа и построения алгоритмов распознавания на основе вероятностных соображений. Нормальное распределение позволяет моделировать признаки объектов, а SVM — строить эффективные границы между классами для минимизации ошибок распознавания.

* + **Случайная величина** — это числовая характеристика, которая принимает различные значения в результате случайного эксперимента или процесса. Она формализует случайные явления в математической модели.
  + Примеры случайных величин:
    - Бросок монеты (может принять значения "орел" или "решка").
    - Выигрыш в лотерее (может быть различным денежным значением).
    - Рост человека в определенной популяции (принимает значения в пределах определенного диапазона).
  + Случайные величины могут быть дискретными (принимают счетное количество значений) или непрерывными (принимают значения на непрерывном интервале).

1. **Дисперсия**:
   * **Дисперсия** — это мера разброса случайной величины относительно ее математического ожидания (среднего значения).
   * Математически дисперсия случайной величины XXX, обозначается как Var(X)\text{Var}(X)Var(X), и вычисляется как среднее значение квадрата отклонения случайной величины от ее математического ожидания.
   * Формула дисперсии для дискретной случайной величины XXX: Var(X)=∑i(xi−μ)2⋅P(X=xi),\text{Var}(X) = \sum\_{i} (x\_i - \mu)^2 \cdot P(X = x\_i),Var(X)=i∑​(xi​−μ)2⋅P(X=xi​), где xix\_ixi​ — значения случайной величины, μ\muμ — математическое ожидание, P(X=xi)P(X = x\_i)P(X=xi​) — вероятность того, что случайная величина примет значение xix\_ixi​.
   * Дисперсия показывает, насколько сильно значения случайной величины разбросаны относительно ее среднего значения. Чем больше дисперсия, тем больше разброс значений относительно среднего, и наоборот.

Эти концепции используются для анализа случайных процессов, моделирования вероятностных явлений и принятия статистических решений в различных областях, включая экономику, физику, биологию, исследования операций и машинное обучение.

**14. Байесовский классификатор.**

Работа алгоритма распознавания заключается в получении решения, которое бы минимизировало условный риск.

Для этого необходимо выполнить следующие действия:

**Шаг 1.** Для каждого объекта вычисляют значения условного риска



.



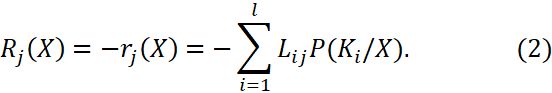
**Шаг 2.** Принимается решение, что ,



если для любых ).



Таким образом, при построении описанного алгоритма используются решающие функции вида



Классификатор минимизирует условные потери.

Математическое ожидание полных потерь на множестве всех решений также будет минимизировано.

Такой классификатор соответствует оптимальному качеству распознавания и называется ***Байессовским классификатором.***

Для реализации алгоритмов с решающими функциями (2) необходимо определить:

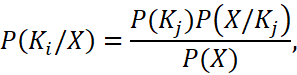
- способ задания функции потерь  *;*



- эффективные выражения для вычисления неизвестной величины .



Согласно известной формуле Байесса



где ‒ вероятность появления объекта *X*.



Тогда



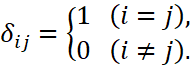
Множитель является общим для и им можно пренебречь.



Расчеты упрощаются, если функцию потерь выбрать в виде:



‒ символ Кронекера:



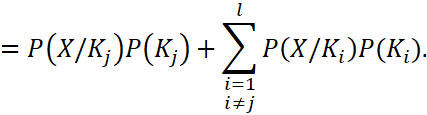
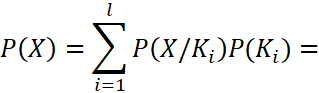
Величина называется нуль-единичной функцией потерь:



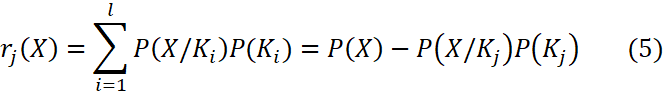
- стоимость потерь равна нулю при правильном распознавании;

- равна единице при ошибке.

По закону вероятности:



Подставляя (4) в (3), получаем



минимизируется по *j* максимизацией члена



Таким образом, при построении алгоритмов распознавания в качестве решающих можно использовать функции вида



Решение, полученное с помощью таких функций, называется решением по максимуму правдоподобия.

На практике удобно использовать другое выражение:



которое приводит к тому же самому решению, так как логарифм является монотонно возрастающей функцией.

*Функции распределения* не известны - могут быть оценены с помощью объектов *обучающей выборки*.

Вид функции распределения известен - *по обучающей* *выборке оцениваются параметры функций* (математическое ожидание, дисперсия и др.).

Хорошей моделью многих статистических процессов является многомерное нормальное распределение:

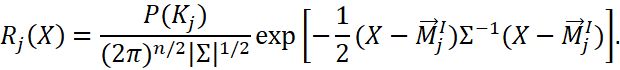
- просто реализуется на ЭВМ;

- удобно при аналитической обработке.

Пусть функция распределения вероятностей, связанных с классами , представляет многомерные нормальные плотности,



тогда решающая функция (6) примет вид:



‒ *n*-мерный вектор среднего значения;



/ ‒ математическое ожидание данных *i*-гo канала сканера для класса



‒ ковариационная матрица размерности



‒ ковариация *i* и *k* каналов по классу



‒ транспонированный вектор;



‒ матрица, обратная , а ‒ ее детерминант.

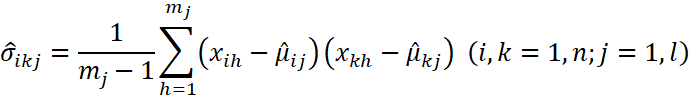
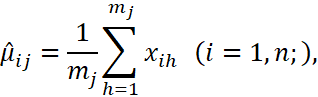


На практике математическое ожидание и ковариационные матрицы классов не известны, но могут быть оценены по обучающей выборке.

15. Пусть и ‒ несмещенные оценки соответственно для , и ,



тогда эти оценки имеют вид:



, — число объектов обучающей выборки из класса , т. е., используя полученные оценки, можно оценивать и .



*Замечание.*

Для оценки матожидания и ковариационных матриц классов, необходимо иметь обучающую выборку соответствующей мощности.

Например, для данных с *n* диапазонов длин волн минимальное число обучающих объектов для класса должно быть , иначе ковариационная матрица будет вырожденной.



*На практике* для хорошей оценки параметров, необходимо иметь число объектов в диапазоне .



17. Если тип распределения неизвестен, тогда необходимо оценить .



Одним из способов решения является *функциональная аппроксимация.*

Покажем как аппроксимировать с помощью множества функций:



(7)



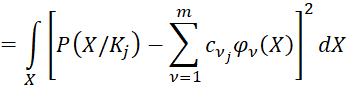
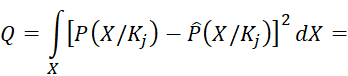
‒ неизвестные коэффициенты;



‒ множество заданных базисных функций (например, полиномы Эрмита).



18. Необходимо определить коэффициенты , для которых значение квадратичной оценки:



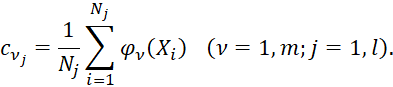
для всех из класса минимально.



Необходимое условие минимума для ‒



19. Доказано, что если базисные функции ортонормированы, то



После определения коэффициентов, с помощью формулы (7) формируется оценка плотности распределения *.*



Если объекты обучающей выборки поступают последовательно, то коэффициенты можно определить итеративно по формуле:



и - коэффициенты, полученные с помощью ; и объектов обучающей выборки соответственно.



20.***Замечание.***

Качество аппроксимации с помощью выбранной системы базисных функций зависит от *т*членов разложения.

Если для некоторой оценки качество алгоритма распознавания оказывается неудовлетворительным, то следует увеличить число базисных функций.



При отсутствии априорной информации о характере плотности распределения базисные функции надо выбирать, исходя из простоты их реализации.

1. Детерминистские алгоритмы.

Алгоритм градиентного спуска

Случайного поиска

1. Подход базируется на математическом аппарате, не использующем (в явном виде) статистические свойства классов.

Требуется определить решающие функции, которые в пространстве признаков порождают границы, отделяющие объекты классов.

Выдвигается предположение и выбирается параметрическое семейство решающих функций.

На основе обучающей выборки определяют значения параметров конкретной функции.

Целесообразно искать решающие функции, обладающие простыми свойствами (например, линейные).

2. Рассмотрим алгоритмы распознавания с линейными и нелинейными решающими функциями.

Пусть задано *l* классов .



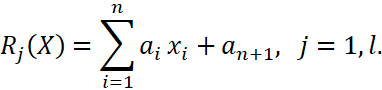
Необходимо построить *l* функций *)*, таких



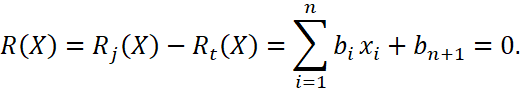
что если



Линейная форма решающей функции имеет следующий вид



3. Тогда уравнение границы между классами и имеет вид:



***Замечание.*** В дальнейшем для упрощения изложения полагаем

что не снижает общности приведенных результатов.



4. Пусть и – обучающие выборки и .



Если выборки линейно разделимы, то для построения алгоритма необходимо определить вектор параметров ,для которого



Умножив вектор на -1, получим эквивалентную систему



(8)



5. Для удобства записи системы неравенств используется расширенный вектор описания объекта *.*



Тогда решением системы (8) является вектор *,* который строится на основе объектов обучающей выборки.



Для нахождения решения используются итерационные методы обучения с коррекцией.

6. ***Общая схема метода.***

*Предварительный шаг.*

Выбор начального значения вектора .



*Общий к-ый шаг.* Коррекция вектора параметров.

Объекты обучающей выборки предъявляются последовательно по одному. Если необходимо (зависит от результата распознавания объекта), то в соответствии некоторым *правилом* корректируется.   
*Замечание.* Объекты могут предъявляться в любом порядке, но обязательно каждый из них используется неоднократно.

*Принцип останова*.

*Общий шаг* повторяется до тех пор, пока на некотором *к-ом* шаге ***значение вектора для всех объектов обучающей выборки не изменяется***.



7. Пример, градиентной схемы минимизации функционала



Этот функционал имеет один минимум, равный нулю при , т. е. минимум эквивалентен решению системы (8).



Напомним, что градиент произвольной функции определяется по формуле



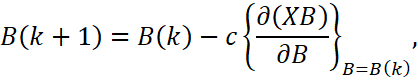
8. *Свойство вектора градиентов*:

направление вектора градиентов указывает направление наискорейшего роста функции при увеличении ее аргумента.

В данном случае увеличение значения вектора в направлении отрицательного градиента позволяет найти минимум функционала (9).



9. Алгоритм градиентного спуска определяется рекуррентным соотношением:



‒ значение вектора на *k-*мшаге; *c* ‒ коэффициент коррекции.



Так как



То



где ‒ объект обучающей выборки, предъявленный на *k-*мшаге.



10. Если - объект распознан правильно, выражение в {...} равно нулю и корректировка не производится, в противном случае



типичное рекуррентное соотношение итерационного метода с коррекцией.

*Примечание.* Если объекты обучающей выборки линейно разделимы, то за конечное число шагов можно построить решение, удовлетворяющее системе неравенств (8).

16. Метод потенциальных функций.

Рассмотрим способ определения решающих функций, основанный на использовании понятия ***потенциальной функции.***

Полагаем, что объекты обучающей выборки создают вокруг себя некоторое поле.



Например, можно считать, что в каждой точке пространства признаков, соответствующей объекту , находится единичный заряд.



Тогда, по аналогии с физикой, поле, создаваемое зарядами, описывается потенциалом, создаваемым системой зарядов \пространстве .



12. Введем функцию

,  *-* параметр.



Функция называется *потенциальной*. Она должна удовлетворять условиям:

*-* достигает максимума при и - убывает по мере удаления от *.*



Примеры потенциальных функций:



‒ положительная константа; ‒ метрика Евклида.



13. Пусть обучающей выборке



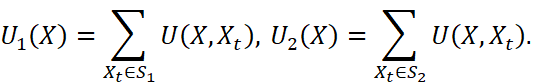
соответствует последовательность потенциальных функций



Рассмотрим потенциал, создаваемый в объектами классов и , полагая его аддитивной функцией.



Получим



14. Объект относится к классу ,если



Тогда уравнение границы между классами имеет вид



Таким образом решающая функция может быть построена на основе объектов обучающей выборки.

15. На этапе обучения последовательно рассматриваются

объекты и вычисляются значения *.*



Функция  *-* на *k-*мшаге определяется совокупностью значений отдельных потенциальных функций.



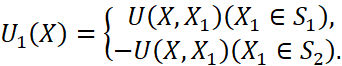
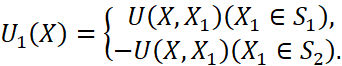
При неправильном распознавании очередного объекта обучающей выборки корректируется текущее значение функции .



16. Вначале полагают .



*Правило коррекции* (при неправильном распознавании :



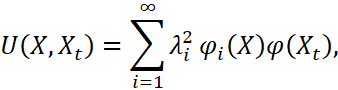
Для -го шага (обучающая процедура принимает вид):



коэффициент , если объект распознается правильно, и или ‒ в противном случае, в зависимости от принадлежности к или .



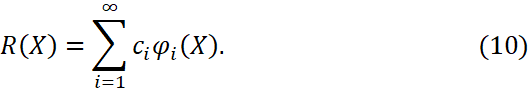
16. Потенциальную функцию для любого можно представить выражением:



полагаются ортонормированными; ‒ действительные числа такие, что функция ограничена для .



решающая функция строится по последовательности имеет вид;



- строятся в процессе обучения.



17. ***Метод потенциальных функций*** предполагает существование в пространстве системы функций , что для каждой пары разделяемых множеств найдется число *т,* при котором решающую функцию можно представить в виде



Т.е. необходимо, чтобы разлагалась в ряд с конечным числом членов.

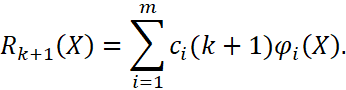


При этом решающая функция должна быть относительно гладкой с не большим числом перегибов в малой окрестности.

18. Таким образом, если в существует полная система функций, то можно считать ее элементом и любая функция из этой системы, в том числе и , может быть представлена в виде (10).



Тогда можно записать следующее рекуррентное соотношение:



Коэффициенты зависят от числа итераций и вычисляются как:



19. *Замечание.*

Система , как правило, задается априорно.



Выбор системы полностью зависит от класса решаемых задач и интуиции исследователя.17. Эвристические алгоритмы.

Методы распознавания применяются в слабо формализованных областях науки и техники, где невозможно построить точную математическую модель предметной области.

***Пример.*** Задача классификации природных покровов:

*Необходимо установить функциональную зависимость между спектрограммами и подстилающими покровами земной поверхности.*

*Зависимости заранее не известны и в большинстве случаев на практике их явный вид не удается установить.*

Подход к решению:

***Выдвигается предположение*** (о типе распределения объектов в классе, о компактности их расположения и т.п.).

Получают формализм, основанный на идеи распознавания образов.

В практических приложениях все алгоритмы распознавания, по сути, являются эвристическими:

*невыполнение выдвинутой гипотезы приводит к ошибкам распознавания и отказам.*

3. На первом этапе развития теории и практики распознавания возникло большое количество ***эвристических алгоритмов***, для решения *конкретных практических задач*.

По мере их накопления появилась возможность описывать не только сами алгоритмы, но и принципы их формирования.

Эти принципы (уже над множеством алгоритмов), высказанные в слабо формализованном виде, получили точное математическое описание.

Эвристическим является выбор принципа, а алгоритмы, построенные на их основе, уже могут быть получены стандартным образом. Формализация принципов приводит к появлению моделей эвристических алгоритмов распознавания.

4. Таким образом, определяется набор допустимых в некотором смысле процедур, с помощью которого можно описать модель (семейство) алгоритмов.

Построенная модель характеризуется системой параметров.

Для построения алгоритма, необходимо зафиксировать набор процедур и конкретные значения параметров.

- Выбор вида модели проводит исследователь.

- Значения параметров получают в результате решения оптимизационной задачи на основе обучающей выборки

5 **Модель** *позволяет:*

- единообразно описывать класс алгоритмов распознавания;

- устанавливать взаимно однозначное соответствие между алгоритмом и набором числовых параметров.

*делает возможным*:

- ставить задачу нахождения оптимального алгоритма;

- применять для ее решения математические методы.

18. Алгоритмы вычисления оценок.

6. ***Модель алгоритмов вычисления оценок***

Данный класс алгоритмов реализует *принцип прецедентности* (принятия решения по аналогии) или *частичной прецедентности*.

***Предположение:*** разделительные свойства объектов содержатся не в отдельных признаках, а в различных их сочетаниях.

*Принцип действия алгоритмов*

- вычисляют оценки похожести (близость) распознаваемого объекта с объектами-эталонами по системе ансамблей признаков;

- по частичной близости вычисляют обобщенную близость.

7. ***Описание модели (семейства)*** ***алгоритмов***.

Пусть Эталонные объекты: - обучающая выборка задана в виде таблицы (классы непересекающиеся).



Из набора признаков выделяют всевозможные подмножества множества *N* - *систему опорных множеств*.



В алгоритмах вычисления оценок в качестве системы опорных множеств часто используют все подмножества мощности *k.*

8. Каждому опорному множеству взаимно однозначно соответствует булев вектор с единичными координатами .



Из таблицы удаляют все столбцы, кроме ,



Получают (так называемую) -часть таблицы.



Объекты-эталоны этой таблицы обозначают:

(-части объектов).



9. Функцию или - степень похожести -частей, называют *функцией близости объектов*.



Например, для



в качестве можно выбрать число выполненных неравенств



,



где  ‒ положительное число, т. е.



‒ число выполненных неравенств, .



Функция близости выделяет части строк, у которых большинство признаков достаточно близки.

10. Вычисление оценок осуществляют по значениям функции близости на объектах и по опорному множеству *I*.



Оценка может зависеть и от внешних параметров:

например, - ‒ степень важности эталона ‒ веса признаков.



Оценка (мера близости) имеет общий вид

, т. е.



или



11. Таким образом, вычисляют оценки для классов по заданному опорному множеству *I*.



Пусть для класса ; с эталонными объектами вычислены значения .



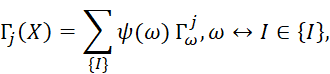
Тогда оценку для класса вычисляют по общей формуле:



Например: .



12. Далее вычисляют оценки для класса по системе опорных множеств {*I*}.



‒ параметр (степень важности опорного множества).



В результате для объекта *X* получают *l* функций (*j*=1, *l*),



Их можно рассматривать в качестве ***решающих функции***.

Для построения алгоритма достаточно определить классификационное правило *r* как функцию от оценок .

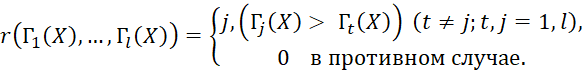
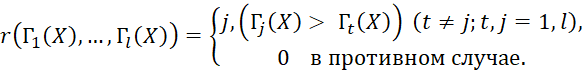


13. Пусть - *X* ‒ объект, предъявленный на распознавание,

- ‒ оценки, вычисленные по системе опорных множеств {*I*},



Тогда



Если значение функции *r* равно нулю, то происходит отказ от распознавания (класс ).



14. Таким образом, выбирая:

*- фиксируя систему опорных множеств;*

*- определяя функцию близости;*

*- задавая правила вычисления всех типов указанных оценок;*

*- назначая классификационное правило*,

Таким образом, ***алгоритмы вычисления оценок*** определяются шестью типами процедур. Варьируя ими, можно строить различные по сложности модели (семейства алгоритмов).

15. Эффективность данных алгоритмов зависит от того, насколько просто решается задача вычисления оценок .



Например, для модели с параметрами:

- *k* ‒ длина опорного множества;

- , ‒ пороги функции близости;



- и ‒ веса важности эталонов и признаков.



‒ число слагаемых в выражении равно ,



где - число объектов таблицы из класса .



17. В частности, можно получить эффективные формулы для .



Так, для описанной модели имеет место следующее выражение:



где ‒ число опорных множеств из {*I*}, содержащих признак *h* ( в соответствующем векторе отличен от нуля), таких, что .



Вычисления формулы не вызывают никаких затруднений.

18. Преимуществами алгоритмов вычисления оценок являются:

- более слабые требования к исходной информации;

- не обязательность наличия сведений о статистических характеристиках;

- небольшие объемы априорных данных (по сравнению с традиционными алгоритмами с обучением).

19. Задача корректировки.

Основной **задачей** (при реализации принципа корректировки) является:

построение оператора (корректора), который минимизирует заданный функционал качества (мера близости) корректировки .

***Для решения такой задачи необходимо располагать:***

- информацией  об элементах множеств  и ;

- информацией  об используемых алгоритмах.

20. Двухуровневые схемы распознавания.

**Основная идея**

Двухуровневые схемы распознавания используются для улучшения точности работы некорректных алгоритмов (т.е. алгоритмов, которые иногда могут ошибаться). Идея заключается в том, чтобы объединить результаты нескольких таких алгоритмов и затем скорректировать их, чтобы получить более надежное решение. Этот подход был первоначально разработан в теории автоматических систем.

**Основные компоненты задачи**

1. **Некорректные алгоритмы**:
   * Это алгоритмы, которые решают задачу, но их результаты могут быть ненадежными.
   * Например, несколько моделей машинного обучения, которые классифицируют данные, но не всегда делают это правильно.
2. **Функция (оператор) коррекции**:
   * Это механизм, который принимает на вход результаты некорректных алгоритмов и преобразует их в более надежное решение.
   * Оператор работает на основе информации о том, насколько близки полученные результаты к правильным.
3. **Качество корректировки**:
   * Оценивается с помощью меры близости или расстояния между корректированными решениями и истинными значениями.
   * Например, с помощью функционала качества, который показывает, насколько хорошо скорректированные решения соответствуют правильным.

**Построение оптимального корректора**

1. **Набор алгоритмов**:
   * Имеется конечный набор некорректных алгоритмов.
   * Каждый алгоритм выдает свои решения на основе контрольной выборки объектов.
2. **Оптимальный корректор**:
   * Задача состоит в построении такого корректора, который минимизирует ошибки на контрольной выборке.
   * Корректор выбирает лучшие результаты из набора решений, чтобы свести к минимуму ошибки классификации.
3. **Вычисление функционала качества**:
   * Определяется таблица штрафов, которая показывает, какие ошибки классификации (и насколько) должны быть наказаны.
   * Штрафы могут быть разными в зависимости от типа ошибки (например, пропущенная классификация или неправильная классификация).

**Алгоритм построения оптимального корректора**

1. **Просмотр наборов решений**:
   * Группируются пары результатов, полученных некорректными алгоритмами.
   * Объединяются одинаковые наборы решений для упрощения анализа.
2. **Вычисление функционала для каждого набора**:
   * Для каждой группы решений вычисляется значение функционала качества.
   * Определяются лучшие значения корректора, которые минимизируют штрафы.
3. **Присвоение значений корректору**:
   * Корректору присваиваются значения, которые минимизируют функционал для каждой группы.
   * Если минимальное значение достигается для нескольких вариантов, выбирается любой из них произвольно.
4. **Принцип останова**:
   * Алгоритм заканчивает работу, когда все группы решений просмотрены и обработаны.
   * Доказано, что такой корректор будет оптимальным для заданной контрольной выборки.

**Заключение**

Двухуровневые схемы распознавания позволяют улучшить точность работы некорректных алгоритмов путем объединения их результатов и последующей коррекции. Основная цель — построить оптимальный корректор, который минимизирует ошибки на основе контрольной выборки. Это достигается через итеративный процесс, включающий просмотр, группировку и анализ результатов, с последующим выбором наилучших корректировок.

Для повышения точности результатов работы некорректных алгоритмов применяется подход, основанный на двухуровневых схемах распознавания.

Некорректные алгоритмы рассматриваются как *ненадежно* работающие блоки в системе переработки информации.

Каждый из них решает одну и ту же задачу, а затем определенным образом произво­дится корректировка полученных результатов.

Подобный принцип был впервые разработан в теории автоматических систем при синтезе надежных схем из ненадежных элементов.

1. ***Двухуровневые схемы распознавания***

Пусть заданы:

-  - некоторая совокупность задач и- их решения;

-- множество (некорректно работающих) алгоритмов;

 - решений алгоритмами задач из множества.

В силу некорректности алгоритмов,  и  не совпадают.

Рассмотрим функцию (оператор) 

- с областью определения  и

- с областью значений .

2. Оператор  переводит совокуп­ность решений , полученных алгоритмами , в элементы множества ).

Элементы называются **скорректированными решениями задачи** .

Качество корректировки можно контролировать с помощью задания *меры близости (расстояния)* между множествами  и .

**21. Построение оптимального корректора.**

***Построение оптимального корректора***

Пусть задано:

-  - конечный набор алгоритмов из множества ;

- - контрольная выборка объектов .

Из классификационных векторов

11.



Каждому набору таблицы 

соответствует пара .

12.

Построим оптимальный корректор , определенный на наборах таблицы , принимающий значения из множества .

Введем функцию  такую, что 

где  и  - координата информационного вектора , соответствующая паре .

Если все наборы таблицы *Т* - различны, то полагаем .

13.

Однако встре­чаются ситуации:

когда наборы и  равны при , а .

Возможны следующие ошибки классификации объекта:

а) объект , но корректор не относит его к классу ;\_

б) объект , но корректор относит его к классу .

Кроме того, возможны ситуации, когда корректор отказывается от классификации объекта.

14.

Введем функционал качества , определяемый таблицей штрафов, как функцию от значений  и .



- при правильном распознавании объекта , штраф равен нулю;

- при неправильном - величина штрафа больше нуля.

15.

Возможны, например, следующие виды штрафов:

а) ;

б) .

Над значениями  и  введем операцию  по правилу:



16.

Тогда функционал качества можно задать в виде следующего выражения:

.

Учитывая значения  и , функционал можно записать в виде:

,

где  - значение корректора  на наборе .

17.

Пусть - множество допустимых корректоров, определенных на наборах таблицы *Т*.

Корректор является оптимальным в , если:

.

Тогда для построения оптимального (по заданной контрольной выборке ) корректора  возможен следующий алгоритм.

***18.***

***Алгоритм***.

*Шаг 1.* Просматриваем наборы таблицы  и группируем пары  с равными наборами, получаем  - совокупности пар  с равными наборами.

*Шаг 2.* Рассматриваем совокупность для каждого  и вычисляем значение функционала



*Шаг 3.* Присваиваем корректору  на наборе  значение, для которого достигается .

19.

Если минимум достигается для нескольких значений , то корректору  присваивается одно из этих значений произвольно.

*Шаг 4.* (принцип останова) Алгоритм заканчивает работу, когда все  просмотрены.

Нетрудно доказать следующее:

***Утверждение.*** Построенный данным алгоритмом корректор  является оптимальным для заданной контрольной выборки .

**22. Постановка задачи кластеризации.**

Основным условием решения задач распознавания с обучением является наличие априорной информации о предметной области, в виде *обучающей выборки.*

*Кластеризация* ‒ это классификация объектов без обучения, когда процесс генерации классов происходит без априорной информации о прототипах классификации.

2.

***Формальная постановка задачи кластеризации***.

Пусть в пространстве признаков задано *m* объектов .



Кластеризацию формально можно записать следующим образом.

Определить области такие, что любой входит только в одну из областей , т.е.



для любых .



3.

Алгоритмы кластеризации включают объекты в классы (кластеры) в соответствии с *мерой сходства*, которая отражает *естественные связи* между объектами.

Предполагается, что степень этих связей намного выше среди объектов внутри группы и ниже среди объектов из разных групп.

Поэтому для классификации на основе кластерного анализа необходимо ввести меру сходства или различия.

4.

Например, если определяет меру сходства между объектами, то очевидно, что



Мера сходства (различия) задается в числовой форме и указывает степень естественных связей:

- между объектами в группе;

- между объектами и группой объектов;

- между группами объектов.

**5.**

**23. Меры сходства.**

***Примеры функций (***используемых в качестве мер сходства***).***

***Евклидово расстояние*** (простейшая мера).

Для многомерного случая, когда все координаты равнозначны, она имеет следующий вид:



Или рассматривается взвешенное расстояние:



где ‒ весовой коэффициент для *k*-й координаты.



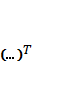
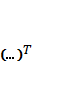
**6.**

***Расстояние Махаланобиса.***

Квадратичное расстояние между объектами записывается в виде



где ‒трансформированный вектор; ‒ обратная ковариационная матрица.



**7.**

***Коэффициент Танимото.***

Танимото предложил отношение подобия:



где ‒ соответственно общее число совпадающих признаков у объектов и .



8.

Проблема разбиения исходных данных на кластеры остается открытой и после выбора меры сходства.

*Качество разбиения* определяется *критерием*, который:

- либо воспроизводить некоторые эвристические соображения (опыт и интуицию исследователя),

- либо основываться на оптимизации показателя качества.

8а.

*Евклидово расстояние* хорошо для эвристического подхода, что связано с естественностью его интерпретации как меры сходства (близости).

Однако поскольку близость двух объектов является относительной мерой их подобия, то вводят порог (параметр).

Порог позволяет установить приемлемые степени сходства при построении кластеров.

9.

Показателем качества является сумма квадратов ошибки:



где *l* ‒ число кластеров; ‒ множество объектов кластера *j*;



‒ вектор его выборочных средних значений.



Критерий определяет совокупное расстояние между объектом и математическим ожиданием кластера, к которому отнесен объект *X*,

Объекты включают в кластер так, чтобы кластеры имели наибольшую возможную «плотность».

**24. Общая схема кластеризации.**

***Классификация алгоритмов кластеризации.***

Алгоритмы кластеризации можно разделить на 2 класса:

- *прямые* (конструктивные, «собирательные»);

- *обратные* (оптимизационные, «разделительные»).

В первом случае процесс кластеризации начинается с изолированного объекта и в соответствии с заданным порогом объединяет ближайшие объекты в группы (направление кластеризации снизу вверх).

Во втором ‒ в зависимости от значения (максимума или минимума) критерия объекты распределяются в оптимальные кластеры (направление сверху вниз).

Рассмотрим примеры типичных схем.

**25. Алгоритмы с произвольным числом разбиений.**

Пусть ‒ исходное множество объектов.



***Алгоритм 1 (число классов неизвестно).***

**Шаг 1.** Первый объект выбирается как представитель первого кластера: , где ‒ центр кластера.



**Шаг 2.** Рассматривают очередной объект *X* и вычисляют его расстояние до всех существующих кластеров:

- *X* включают в , если ;



- *X* не включают в , если ;



- решение не принимается, если *X* попадает в зону неопределенности кластера (рис. 6).

13.

**Шаг 3.**

а) когда новый объект *X* включен в кластер , вычисляют его параметры:



*k* ‒ число объектов в классе *X* ‒ -й объект.



и - оценки матожидания и дисперсии.



б) формируют новый кластер , если для любого .



**14.**

**Шаг 4.** Шаги 2, 3 повторяют пока все объекты не распределены по кластерам.

При повторном рассмотрении данных в исходном порядке возможны перераспределения отдельных объектов.

**Шаг 5.** После завершения обучения (когда объекты не изменяют своей классовой принадлежности), полученную систему распознавания можно использовать для классификации большего числа объектов.

К данному моменту зон неопределенности уже не должно быть.

*Замечание.* Все объекты, не попавшие ни в одну из зон, могут, например, включаться в ближайший по минимуму расстояния кластер.

15.

*Недостатки* описанной схемы:

- зависимость результатов от выбора первого центра;

- от порядка рассмотрения объектов;

- от значения порогов.

*Преимущества*:

- позволяет быстро получать приблизительные оценки основных характеристик заданного набора данных;

- является привлекательной с вычислительной точки зрения.

Для понимания геометрии распределения объектов, необходимо проводить эксперименты с различными значениями порога х и исходными точками кластеризации.

16.

**25. Алгоритмы с заданным числом классов.**

Пусть, число классов, на которое необходимо разбить множество объектов , заранее определено.



Предлагается схема, основанная на минимизации суммы квадратов расстояний от всех объектов кластера до его центра:



‒ область -го кластера с центром , полученная на -й итерации.



***Алгоритм 2.***

**Шаг 1.** Произвольным образом выбирают *l* исходных центров кластеров :



**Шаг 2.** На *k*-й итерации распределяют объекты по кластерам в соответствии с правилом:



для всех .



В случае равенства объекты распределяются произвольно.

18.

**Шаг 3.**

Корректируют центры кластеров по формуле:



где ‒ число объектов в кластере .



*Примечание.* Уточненный центр кластера будет минимизировать сумму квадрата расстояний между всеми объектами из .



19.

Центры новых кластеров выбираются таким образом, чтобы минимизировать показатель качества



**Шаг 4.**

Когда , алгоритм заканчивается (условие сходимости).



В противном случае переходим к шагу 2.

20.

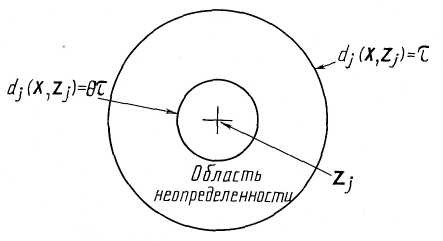


Рис. Зоны принятия решений

*Замечание.* Практическое применение алгоритма также требует проведения численных экспериментов.

**27. Нейронные сети: Общие сведения**

**Нейронные сети** — это вычислительные модели, вдохновленные биологическими нейронными сетями, которые составляют мозг. Они состоят из множества взаимосвязанных узлов или "нейронов", работающих совместно для решения задач, таких как классификация, прогнозирование и распознавание образов.

**Основные понятия:**

* **Нейрон**: Основной элемент нейронной сети, который получает входные данные, обрабатывает их и передает результат дальше.
* **Слои**: Нейронные сети состоят из слоев нейронов: входной слой, скрытые слои и выходной слой.
* **Весы и смещения**: Входные данные умножаются на веса, к результату добавляется смещение и передаются через активационную функцию.
* **Активационная функция**: Преобразует выход нейрона, добавляя нелинейность (например, сигмоидная функция, ReLU).
* **Обучение**: Процесс настройки весов и смещений с целью минимизации ошибки сети с использованием алгоритмов, таких как градиентный спуск.

**28. Основные архитектуры нейронных сетей**

**Различные архитектуры нейронных сетей предназначены для различных типов задач**:

* **Полносвязные сети (Fully Connected Networks)**: Каждый нейрон в одном слое связан со всеми нейронами следующего слоя. Применяются в задачах общего назначения.
* **Сверточные нейронные сети (Convolutional Neural Networks, CNN)**: Используются для обработки данных, имеющих топологическую структуру, таких как изображения и видео. Включают слои свертки и подвыборки (пулинга).
* **Рекуррентные нейронные сети (Recurrent Neural Networks, RNN)**: Обрабатывают последовательные данные (тексты, временные ряды) путем сохранения состояния (памяти) в виде скрытых состояний.
* **Долгосрочная краткосрочная память (Long Short-Term Memory, LSTM)**: Улучшенная версия RNN, способная сохранять долгосрочные зависимости.
* **Автокодировщики (Autoencoders)**: Используются для задач, связанных с уменьшением размерности и извлечением признаков.

**29. Глубокие нейронные сети**

**Глубокие нейронные сети (Deep Neural Networks, DNN)** — это нейронные сети с множеством скрытых слоев. Они могут обучаться более сложным представлениям данных благодаря большому количеству уровней абстракции.

**Основные аспекты:**

* **Глубокое обучение**: Обучение глубоких нейронных сетей, часто с использованием больших объемов данных и мощных вычислительных ресурсов.
* **Обратное распространение ошибки (Backpropagation)**: Алгоритм, используемый для обучения нейронных сетей путем вычисления градиентов ошибки и обновления весов.
* **Регуляризация**: Методы предотвращения переобучения, такие как Dropout и L2-регуляризация.

**30. Градиентные методы обучения**

**Градиентные методы обучения** — это подходы к оптимизации параметров модели путем минимизации функции потерь.

**Основные методы:**

* **Градиентный спуск (Gradient Descent)**: Итеритивный алгоритм, который обновляет параметры в направлении отрицательного градиента функции потерь.
* **Случайный градиентный спуск (Stochastic Gradient Descent, SGD)**: Обновляет параметры по одному обучающему примеру за раз, что может ускорить обучение и обеспечить лучшее обобщение.
* **Адаптивные методы (Adam, RMSprop)**: Улучшают SGD, адаптируя скорость обучения для каждого параметра на основе статистики градиентов.

**31. Генетические алгоритмы**

**Генетические алгоритмы (GA)** — это эвристические методы поиска, вдохновленные процессами естественного отбора и генетики.

**Основные компоненты:**

* **Популяция**: Набор потенциальных решений.
* **Хромосомы**: Представление решений.
* **Оценка приспособленности**: Функция, которая оценивает, насколько хорошо каждое решение решает задачу.
* **Операторы генетики**: Включают селекцию, кроссовер (рекомбинацию) и мутацию для создания новых решений.

**32. Метод отжига**

**Метод отжига (Simulated Annealing)** — это вероятностный метод оптимизации, который ищет глобальный минимум функции, имитируя процесс отжига в металлах.

**Основные идеи:**

* **Температура**: Управляет вероятностью принятия худших решений на ранних этапах поиска для предотвращения локальных минимумов.
* **Охлаждение**: Постепенное уменьшение температуры для сужения поиска к оптимальному решению.

**33. Метод опорных векторов**

**Метод опорных векторов (Support Vector Machine, SVM)** — это алгоритм классификации, который ищет гиперплоскость, максимально разделяющую классы в пространстве признаков.

**Основные понятия:**

* **Гиперплоскость**: Поверхность, разделяющая классы.
* **Опорные векторы**: Точки данных, наиболее близкие к гиперплоскости, которые определяют ее положение.
* **Ядра (Kernels)**: Функции для преобразования данных в более высокую размерность, позволяющие SVM решать нелинейные задачи.

**34. Метод решающих деревьев**

**Решающие деревья (Decision Trees)** — это модели, использующие деревообразную структуру для принятия решений на основе значений признаков.

**Основные аспекты:**

* **Узлы и листья**: Внутренние узлы представляют собой проверки признаков, а листья — конечные решения или классы.
* **Обучение дерева**: Процесс разделения данных на основе наилучшего критерия (например, прироста информации, индекса Джини).

**35. Алгоритмы случайного леса**

**Случайный лес (Random Forest)** — это ансамблевый метод, использующий множество решающих деревьев для улучшения точности и устойчивости модели.

**Основные особенности:**

* **Бэггинг (Bagging)**: Процесс создания множества деревьев путем обучения на случайных подвыборках данных.
* **Агрегация**: Объединение предсказаний всех деревьев путем голосования для классификации или усреднения для регрессии.

Эти темы охватывают широкий спектр современных методов и моделей машинного обучения, которые находят применение в различных областях, от анализа данных до разработки интеллектуальных систем.