**1.4 Классификация задач распознавания**

1.

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

- ***критерий оптимальности***, которому удовлетворяют области решения;

- ***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):

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

- 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***

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.