### 1.3.2. ЛЕКЦИЯ-8. Автоматизированные системы распознавания образов

##### *Учебные вопросы*

*1. Основные понятия и определения, связанные с системами распознавания образов.*

*2. Проблема распознавания образов.*

*3. Классификация методов распознавания образов.*

*4. Применение распознавания образов для идентификации и прогнозирования. Сходство и различие в содержании понятий "идентификация" и "прогнозирование".*

*5. Роль и место распознавания образов в автоматизации управления сложными системами.*

*6. Методы кластерного анализа.*

#### *1.3.2.1. Основные понятия*

***Системой распознавания образов*** будем называть класс систем искусственного интеллекта, обеспечивающих:

– формирование конкретных образов объектов и обобщенных образов классов;

– обучение, т.е. формирование обобщенных образов классов на основе ряда примеров объектов, классифицированных (т.е. отнесенных к тем или иным категориям – классам) учителем и составляющих обучающую выборку;

– самообучение, т.е. формирование кластеров объектов на основе анализа неклассифицированной обучающей выборки;

– распознавание, т.е. идентификацию (и прогнозирование) состояний объектов, описанных признаками, друг с другом и с обобщенными образами классов;

– измерение степени адекватности модели;

– решение обратной задачи идентификации и прогнозирования (обеспечивается не всеми моделями).

##### *1.3.2.1.1. Признаки и образы конкретных объектов, метафора фазового пространства*

Признаками объектов будем называть конкретные результаты измерения значений их свойств.

Свойства объектов отличаются своим качеством и измеряются с помощью различных органов восприятия или измерительных приборов в различных единицах измерения.

Результатом измерения является снижение неопределенности в наших знаниях о значении свойств объекта. Значения свойств конкретизируются путем их сопоставления определенным градациям соответствующих измерительных шкал: номинальных, порядковых или отношений.

В номинальных шкалах отсутствуют отношения порядка, начало отсчета и единица измерения.

На порядковых шкалах определены отношения "больше – меньше", но отсутствуют начало отсчета и единица измерения.

На шкалах отношений определены отношения порядка, все арифметические операции, есть начало отсчета и единица измерения.

Можно представить себе, что шкалы образуют оси координат некоторого абстрактного многомерного пространства, которое будем называть "фазовым пространством".

В этом фазовом пространстве каждый конкретный объект представляется определенной точкой, имеющей координаты, соответствующие значениям его свойств по осям координат, т.е. градациям описательных шкал.

Оси координат фазового пространства в общем случае не являются взаимно-перпендикулярными шкалами отношений, т.е. в общем случае это пространство неортонормированное, более того – неметрическое. Следовательно, в нем в общем случае не применима Евклидова мера расстояний, т.е. не действует Евклидова метрика. Применение этой меры расстояний корректно, если одновременно выполняются два условия:

1. Все оси координат фазового пространства являются шкалами отношений.

2. Все оси координат взаимно-перпендикулярны или очень близки к этому.

##### *1.3.2.1.2. Признаки и обобщенные образы классов*

Обобщенный образ класса формируется из нескольких образов конкретных объектов, относящихся к данному классу, т.е. одной градации некоторой классификационной шкалы.

Обобщенные образы классов формализуются (кодируются) путем использования классификационных шкал и градаций, которые могут быть тех же типов, что и описательные, т.е. номинальные, порядковые и отношений.

Сама принадлежность конкретных объектов к данному классу определятся либо человеком-учителем, после чего фиксируется в обучающей выборке, либо самой системой автоматически на основе кластерного анализа конкретных объектов.

##### *1.3.2.1.3. Обучающая выборка и ее репрезентативность по отношению к генеральной совокупности. Ремонт (взвешивание) данных*

Обучающая выборка является некоторым подмножеством исследуемой совокупности, которая называется "генеральная совокупность". На основе изучения обучающей выборки мы хотели бы сделать выводы о генеральной совокупности, причем важно знать степень достоверности этих выводов.

Рассмотрим, как зависит степень достоверности выводов о генеральной совокупности от объема обучающей выборки.

Если обучающая выборка включает все объекты генеральной совокупности, т.е. они совпадают, то достоверность выводов будет наиболее высокой (при всех прочих равных условиях).

Если же обучающая выборка очень мала, то вряд ли на ее основе могут быть сделаны достоверные выводы о генеральной совокупности, т.к. в этом случае в обучающую выборку могут даже не входить примеры объектов всех или подавляющего большинства классов.

Под репрезентативностью обучающей выборки будем понимать ее способность адекватно представлять генеральную совокупность, так что изучение самой генеральной совокупности можно корректно заменить исследованием обучающей выборки.

Но репрезентативность зависит не только от объема, но и от структуры обучающей выборки, т.е. от того, насколько полно представлены все категории объектов генеральной совокупности (классы) и от того, насколько полно они описаны признаками.

***Взвешивание данных или ремонт обучающей выборки*** – *это операция, в результате которой частное распределение объектов по классам в обучающей выборке максимально, на сколько это возможно, приближается либо к частотному распределению генеральной совокупности (если оно известно из независимых источников), либо к равномерному.*

В системе "Эйдос" режим взвешивания данных реализован.

##### *1.3.2.1.4. Основные операции: обобщение и распознавание*

***Обобщение****– это операция формирования обобщенных образов классов на основе описаний конкретных объектов, входящих в обучающую выборку.*

Сразу необходимо отметить, что операция обобщения реализуется далеко не во всех моделях систем распознавания (например, в методе k-ближайших соседей), а в тех, в которых оно реализуется, – это делается по-разному.

Обычно, пока не реализовано обобщение нет возможности определить ценность признаков для решения задачи идентификации.

Например, если у нас есть 10 конкретных мячей разного размера и цвета, состоящих из разных материалов и предназначенных для разных игр, и мы рассматриваем их как совершенно независимые друг от друга объекты, наряду с другими, то у нас нет возможности определить, какие признаки являются наиболее характерными для мячей и наиболее сильно отличают их от этих других объектов. Но как только мы сформируем обобщенные образы "мяч", "стул", и т.д., сразу выясниться, что цвет мяча и материал, из которого он сделан, не является жестко связанными с обобщенным образом класса "мяч", а наиболее существенно то, что он круглый и его можно бросать или бить во время игры.

***Распознавание****– это операция сравнения и определения степени сходства образа данного конкретного объекта с образами других конкретных объектов или с обобщенными образами классов, в результате которой формируется рейтинг объектов или классов по убыванию сходства с распознаваемым объектом.*

Ключевым моментом при реализации операции распознавания в математической модели является выбор вида ***интегрального критерия или меры сходства***, который бы на основе знания о признаках конкретного объекта позволил бы количественно определить степень его сходства с другими объектами или обобщенными образами классов.

В ортонормированном пространстве, осями которого являются шкалы отношений, вполне естественным является использовать в качестве такой меры сходства Евклидово расстояние. Однако, такие пространства на практике встречаются скорее как исключение из правила, а операция ортонормирования является довольно трудоемкой в вычислительном отношении и приводит к обеднению модели, а значит ее не всегда удобно и целесообразно осуществлять.

Поэтому актуальной является задача выбора или конструирования интегрального критерия сходства, применение которого было бы корректно и в неортонормированных пространствах. Кроме того, этот интегральный критерий должен быть устойчив к наличию ***шума, т.е. к неполноте и искажению*** как в исходных данных, так и самой численной модели.

Требование устойчивости к наличию шума математически означает, что результат применения интегрального критерия к сигналу, состоящему только из белого шума, должен быть равным нулю. Это значит, что в качестве интегрального критерия может быть применена функция, используемая при определении самого понятия "белый шум", т.е. свертка, скалярное произведение, корреляция.

Такой интегральный критерий предложен в математической модели системно-когнитивного анализа и реализован в системе "Эйдос".

##### *1.3.2.1.5. Обучение с учителем (экспертом) и самообучение (кластерный анализ)*

***Обучение с учителем****– это процесс формирования обобщенных образов классов, на основе обучающей выборки, содержащей характеристики конкретных объектов как в описательных, так и в классификационных шкалах и градациях.*

Причем, если описательные характеристики могут формироваться с помощью информационно-измерительной системы автоматически, то классификационные – представляют собой результат вообще говоря неформализуемого процесса оценки степени принадлежности данных объектов к различным классам, который осуществляется человеком-экспертом или, как традиционно говорят специалисты по распознаванию образов, "учителем". В этом случае не возникает вопроса о том, для формирования обобщенного образа каких классов использовать описание данного конкретного объекта.

***Обучение без учителя или самообучение****– это процесс формирования обобщенных образов классов, на основе обучающей выборки, содержащей характеристики конкретных объектов, причем только в описательных шкалах и градациях.*

Поэтому этот процесс реализуется в три этапа:

1. Кластерный анализ объектов обучающей выборки, в результате которого определяются группы наиболее сходных их них по их признакам (кластеры).

2. Присвоение кластерам статуса обобщенных классов, для формирования обобщенных образов которых используются конкретные объекты, входящие именно в эти кластеры.

3. Формирование обобщенных образов классов, аналогично тому, как это делалось при обучении с учителем.

##### *1.3.2.1.6. Верификация, адаптация и синтез модели*

Как только произнесено или написано слово "модель", сразу неизбежно возникает вопрос о степени ее адекватности.

***Верификация модели****– это операция установления степени ее адекватности (валидности) путем сравнения результатов идентификации конкретных объектов с их фактической принадлежностью к обобщенным образам классов.*

Различают внутреннюю и внешнюю, интегральную и дифференциальную валидность.

***Внутренняя валидность*** – это способность модели верно идентифицировать объекты обучающей выборки.

Если модель имеет низкую внутреннюю валидность, то модель нельзя считать удачно сформированной.

***Внешняя валидность*** – это способность модели верно идентифицировать объекты, не входящие в обучающую выборку.

***Интегральная валидность*** – это средневзвешенная достоверность идентификации по всем классам и распознаваемым объектам.

***Дифференциальная валидность*** – это способность модели верно идентифицировать объекты в разрезе по классам.

***Адаптация модели*** – это учет в модели объектов, не входящих в обучающую выборку, но входящих в генеральную совокупность, по отношению к которой данная обучающая выборка репрезентативна.

Если моделью верно идентифицируются объекты, не входящие в обучающую выборку, то это означает, что эти объекты входят в генеральную совокупность, по отношению к которой данная обучающая выборка репрезентативна. Следовательно, на основе обучающей выборки удалось выявить закономерности взаимосвязей между признаками и принадлежностью объектов к классам, которые действуют не только в обучающей выборке, но имеют силу и для генеральной совокупности.

***Адаптация****модели* не требует изменения классификационных и описательных шкал и градаций, а лишь объема обучающей выборки, и *приводит к****количественному****изменению модели*.

***Синтез (или повторный синтез – пересинтез) модели*** – это учет в модели объектов, не входящих ни в обучающую выборку, ни в генеральную совокупность, по отношению к которой данная обучающая выборка репрезентативна.

Это объекты с новыми, ранее неизвестными закономерностями взаимосвязей признаков с принадлежностью этих объектов к тем или иным классам. Причем и признаки, и классы, могут быть как те, которые уже были отражены в модели ранее, так и новые. ***Пересинтез****модели приводит к ее****качественному****изменению.*

#### *1.3.2.2. Проблема распознавания образов*

Проблема распознавания образов сводится к двум задачам: обучения и распознавания. Поэтому, прежде чем сформулировать задачу обучения распознаванию образов уточним, в чем смысл их распознавания.

Простейшим вариантом распознавания является ***строгий запрос*** на поиск объекта в базе данных по его признакам, который реализуется в информационно-поисковых системах. При этом каждому полю соответствует признак (описательная шкала), а значению поля – значение признака (градация описательной шкалы). Если в базе данных есть записи, ***все****значения заданных полей которых****точно****совпадают* со значениями, заданными в запросе на поиск, то эти записи извлекаются в отчет, иначе запись не извлекается.

Более сложными вариантами распознавания является ***нечеткий запрос с неполнотой информации***, когда не все признаки искомых объектов задаются в запросе на поиск, т.к. не все они известны, и ***нечеткий запрос с шумом***, когда не все признаки объекта известны, а некоторые считаются известными ошибочно. В этих случаях из базы данных извлекаются все объекты, у которых совпадает хотя бы один признак и в отчете объекты сортируются (ранжируются) в порядке убывания количества совпавших признаков. При этом при определении ранга объекта в отсортированном списке все признаки считаются имеющими одинаковый "вес" и учитывается только их количество.

Однако:

– во-первых, на самом деле признаки имеют разный вес, т.е. *один и тот же признак в****разной****степени характерен для различных объектов*;

– во-вторых, нас могут интересовать не столько сами объекты, извлекаемые из базы данных прецедентов по запросам, сколько *классификация самого запроса*, т.е. отнесение его к определенной категории, т.е. к тому или иному ***обобщенному*** образу класса.

Если реализация строгих и даже нечетких запросов не вызывает особых сложностей, то распознавание как идентификация с *обобщенными* образами классов, причем с учетом *различия весов признаков* представляет собой определенную проблему.

Обучение осуществляется путем предъявления системе отдельных объектов, описанных на языке признаков, с указанием их принадлежности тому или другому классу. При этом сама принадлежность к классам сообщается системе человеком – Учителем (экспертом).

В результате обучения распознающая система должна приобрести способность:

1. Относить объекты к классам, к которым они принадлежат (идентифицировать объекты верно).

2. Не относить объекты к классам, к которым они не принадлежат (неидентифицировать объекты ошибочно).

Эта и есть проблема обучения распознаванию образов, и состоит она в следующем:

1. В разработке математической модели, обеспечивающей: обобщение образов конкретных объектов и формирование обобщенных образов классов; расчет весов признаков; определение степени сходства конкретных объектов с классами и ранжирование классов по степени сходства с конкретным объектом, включая и положительное, и отрицательное сходство.

2. В наполнении этой модели конкретной информацией, характеризующей определенную предметную область.

#### *1.3.2.3.**Классификация методов распознавания образов*

Распознаванием образов называются задачи установления отношений эквивалентности между конкретными и обобщенными образами-моделями объектов реального или идеального мира.

Отношения эквивалентности выражают принадлежность оцениваемых объектов к каким–либо классам, рассматриваемым как самостоятельные семантические единицы.

При построении алгоритмов распознавания классы эквивалентности могут задаваться исследователем, который пользуется собственными содержательными представлениями или использует внешнюю дополнительную информацию о сходстве и различии объектов в контексте решаемой задачи. Тогда говорят о "распознавании с учителем". В противном случае, т.е. когда автоматизированная система решает задачу классификации без привлечения внешней обучающей информации, говорят об автоматической классификации или "распознавании без учителя".

Большинство алгоритмов распознавания образов требует привлечения весьма значительных вычислительных мощностей, которые могут быть обеспечены только высокопроизводительной компьютерной техникой.

Различные авторы (Ю.Л. Барабаш, В.И. Васильев, А.Л. Горелик, В.А. Скрипкин, Р. Дуда, П. Харт, Л.Т.Кузин, Ф.И. Перегудов, Ф.П. Тарасенко, Темников Ф.Е., Афонин В.А., Дмитриев В.И., Дж. Ту, Р. Гонсалес, П. Уинстон, К. Фу, Я.З. Цыпкин и др.) дают различную типологию методов распознавания образов. Одни авторы различают параметрические, непараметрическиеи эвристические методы, другие – выделяют группы методов, исходя из исторически сложившихся школ и направлений в данной области.

Например, в работах В.А. Дюка [32, 33], в которых дан академический обзор методов распознавания, используется следующая типология методов распознавания образов:

– методы, основанные на принципе разделения;

– статистические методы;

– методы, построенные на основе "потенциальных функций";

– методы вычисления оценок (голосования);

– методы, основанные на исчислении высказываний, в частности на аппарате алгебры логики.

В основе данной классификации лежит различие в формальных методах распознавания образов и поэтому опущено рассмотрение эвристического подхода к распознаванию, получившего полное и адекватное развитие в экспертных системах.

Эвристический подход основан на трудно формализуемых знаниях и интуиции исследователя. При этом исследователь сам определяет, какую информацию и каким образом система должна использовать для достижения требуемого эффекта распознавания.

Подобная типология методов распознавания с той или иной степенью детализации встречается во многих работах по распознаванию. В то же время известные типологии не учитывают одну очень существенную характеристику, которая отражает специфику способа представления знаний о предметной области с помощью какого–либо формального алгоритма распознавания образов.

В.А. Дюка [32, 33] выделяет два основных способа представления знаний:

– интенсиональное, в виде схемы связей между атрибутами (признаками).

– экстенсиональное, с помощью конкретных фактов (объекты, примеры).

Интенсиональное представление фиксируют закономерности и связи, которыми объясняется структура данных. Применительно к диагностическим задачам такая фиксация заключается в определении операций над атрибутами (признаками) объектов, приводящих к требуемому диагностическому результату. Интенсиональные представления реализуются посредством операций над значениями атрибутов и не предполагают произведения операций над конкретными информационными фактами (объектами).

В свою очередь, экстенсиональные представления знаний связаны с описанием и фиксацией конкретных объектов из предметной области и реализуются в операциях, элементами которых служат объекты как целостные системы.

*На наш взгляд, можно провести глубокую и далеко идущую аналогию между интенсиональными и экстенсиональными представлениями знаний и механизмами, лежащими в основе деятельности левого и правого полушарий головного мозга человека. Если для правого полушария характерна целостная прототипная репрезентация окружающего мира, то левое полушарие оперирует закономерностями, отражающими связи атрибутов этого мира.*

Описанные выше два фундаментальных способа представления знаний позволяют предложить следующую классификацию методов распознавания образов:

– интенсиональные методы, основанные на операциях с признаками.

– экстенсиональные методы, основанные на операциях с объектами.

*Необходимо особо подчеркнуть, что существование именно этих двух (и только двух) групп методов распознавания: оперирующих с признаками, и оперирующих с объектами, на наш взгляд, глубоко закономерно. С этой точки зрения ни один из этих методов, взятый отдельно от другого, не позволяет сформировать адекватное отражение предметной области. Между этими методами существует отношение дополнительности в смысле Н.Бора, поэтому перспективные системы распознавания должны обеспечивать реализацию обоих этих методов, а не только какого–либо одного из них.*

Таким образом, в основу классификации методов распознавания, предложенной В.А. Дюка [32, 33], положены фундаментальные закономерности, лежащие в основе человеческого способа познания вообще, что ставит ее в совершенно особое (привилегированное) положение по сравнению с другими классификациями, которые на этом фоне выглядят более легковесными и искусственными.

#### *1.3.2.4. Применение распознавания образов для идентификации и прогнозирования. Сходство и различие в содержании понятий "идентификация" и "прогнозирование"*

Термины "Распознавание образов" и "Идентификация" являются синонимами.

Идентификация и прогнозирование часто практически ничем друг от друга не отличаются по математическим моделям и алгоритмам. Основное различие между ними состоит в том, что при идентификации признаки и состояния объекта относятся к одному времени, тогда как при прогнозировании признаки (факторы) относятся к прошлому, а состояния объекта – к будущему.

Это означает, что системы распознавания образов с успехом могут применяться не только для решения задач идентификации, но и прогнозирования.

#### *1.3.2.5. Роль и место распознавания образов в автоматизации управления сложными системами*

##### *1.3.2.5.1. Обобщенная структура системы управления*

Автоматизированная система управления состоит из двух основных частей: объекта управления и управляющей системы (рисунок 71).

Управляющая система осуществляет следующие функции:

– идентификация состояния объекта управления;

– выработка управляющего воздействия исходя из целей управления с учетом состояния объекта управления и окружающей среды;

– оказание управляющего воздействия на объект управления.

|  |
| --- |
| http://lc.kubagro.ru/aidos/aidos04/1.3.2.files/image002.gif |
| **Рисунок 71. Обобщенная схема рефлексивной системы управления активными объектами** |

##### *1.3.2.5.2. Место системы идентификации в системе управления*

Распознавание образов есть не что иное, как идентификация состояния некоторого объекта. Автоматизированная система управления АСУ), построенная на традиционных принципах, может работать только на основе параметров, закономерности связей которых уже известны, изучены и отражены в математической модели. В итоге АСУ, основанные на традиционном подходе, практически не эффективны с активными многопараметрическими слабодетерминированными объектами управления, такими, например, как макро– и микро– социально-экономические системы в условиях динамичной экономики "переходного периода", иерархические элитные и этнические группы, социум и электорат, физиология и психика человека, природные и искусственные экосистемы и многие другие.

*Поэтому, в состав перспективных АСУ, обеспечивающих устойчивое управление активными объектами в качестве существенных функциональных звеньев должны войти подсистемы идентификации и прогнозирования состояний среды и объекта управления, основанные на методах искусственного интеллекта (прежде всего распознавания образов), методах поддержки принятия решений и теории информации.*

##### *1.3.2.5.3. Управление как задача, обратная идентификации и прогнозированию*

Кратко рассмотрим вопрос о применении систем распознавания образов для принятия решений об управляющем воздействии. Очевидно, что применение систем распознавания для прогнозирования результатов управления при различных сочетаниях управляющих факторов позволяет рассмотреть и сравнить различные варианты управления и выбрать наилучшие из них по определенным критериям. Однако, этот подход на практике малоэффективен, особенно если факторов много, т.к. в этом случае количество сочетаний их значений может быть чрезвычайно большим.

Если в качестве классов распознавания взять целевые и иные будущие состояния объекта управления, а в качестве признаков – факторы, влияющие на него, то в модели распознавания образов может быть сформирована количественная мера причинно-следственной связи факторов и состояний.

Это позволяет по заданному целевому состоянию объекта управления получить информацию о силе и направлении влияния факторов, способствующих или препятствующих переходу объекта в это состояние, и, на этой основе, выработать решение об управляющем воздействии.

Задача выбора факторов по состоянию является обратной задачей прогнозирования, т.к. при прогнозировании, наоборот, определяется состояние по факторам.

Факторы могут быть разделены на следующие группы:

– характеризующие предысторию объекта управления и его актуальное состояние управления;

– технологические (управляющие) факторы;

– факторы окружающей среды;

Таким образом, системы распознавания образов могут быть применены в составе АСУ в подсистемах:

– идентификации состояния объекта управления;

– выработки управляющих воздействий.

Это целесообразно в случае, когда объект управления представляет собой сложную или активную систему.

#### *1.3.2.6. Методы кластерного анализа*

Термин "Кластерный анализ"впервые ввел Tryon в 1939.

***Кластеризация****– это операция автоматической классификации, в ходе которой  объекты объединяются в группы (кластеры) таким образом, что внутри групп различия между объектами минимальны, а между группами – максимальны. При этом в ходе кластеризации не только определяется состав кластеров, но и сам их набор и границы.*

Поэтому вполне обоснованно считается, что методы кластерного анализа используются в большинстве случаев тогда, когда нет каких-либо априорных гипотез относительно классов, т.е. исследование находится на первой эмпирической стадии: описательной.

Существует большое количество различных алгоритмов кластеризации, которые обычно связаны с полным перебором объектов и весьма трудоемки в вычислительном отношении, здесь же мы упомянем лишь о трех из них:

– объединение (древовидная кластеризация);

– двухвходовое объединение;

– метод K средних.

Рассмотрим кратко эти алгоритмы (описание взято с сайта http://StatSoft.ru).

##### *1.3.2.6.1. Древовидная кластеризация*

Древовидная диаграмма (*диаграмму*(рисунок 72) начинается с конкретных объектов (в левой части диаграммы). Теперь представим себе, что постепенно (очень малыми шагами) вы "ослабляете" ваш критерий о том, какие объекты являются уникальными, а какие нет. Другими словами, вы понижаете порог, относящийся к решению об объединении двух или более объектов в один кластер.

В результате, вы *связываете*вместе всё большее и большее число объектов и агрегируете (*объединяете*) все больше и больше кластеров, состоящих из все сильнее различающихся элементов. Окончательно, на последнем шаге все объекты объединяются вместе.

|  |
| --- |
| http://lc.kubagro.ru/aidos/aidos04/1.3.2.files/image004.jpg |
| **Рисунок 72. Древовидная диаграмма последовательной кластеризации** |

##### *1.3.2.6.2. Двухвходовое объединение*

Исследователь может кластеризовать конкретные образы наблюдаемых объектов для определения кластеров объектов со сходными признаками.

Он может также кластеризовать признаки для определения кластеров признаков, которые связаны со сходными конкретными объектами.

В двувходовом алгоритме эти процессы осуществляются одновременно.

##### *1.3.2.6.3. Метод K средних*

В этом методе принадлежность объектов к кластерам определяется таким образом, чтобы:

– минимизировать изменчивость (различия) объектов внутри кластеров;

– максимизировать изменчивость объектов между кластерами.

##### *Контрольные вопросы*

*1. Основные понятия и определения, связанные с системами распознавания образов.*

*2. Признаки и образы конкретных объектов, метафора фазового пространства.*

*3. Признаки и обобщенные образы классов.*

*4. Обучающая выборка и ее репрезентативность по отношению к генеральной совокупности. Ремонт (взвешивание) данных.*

*5. Основные операции: обобщение и распознавание.*

*6. Обучение с учителем (экспертом) и самообучение (кластерный анализ).*

*7. Верификация, адаптация и синтез модели.*

*8. Проблема распознавания образов.*

*9. Классификация методов распознавания образов.*

*10. Применение распознавания образов для идентификации и прогнозирования. Сходство и различие в содержании понятий "идентификация" и "прогнозирование".*

*11. Роль и место распознавания образов в автоматизации управления сложными системами.*

*12. Обобщенная структура системы управления.*

*13. Место системы идентификации в системе управления.*

*14. Управление как задача, обратная идентификации и прогнозированию.*

*15. Методы кластерного анализа.*

*16. Метод кластеризации: "Древовидная кластеризация".*

*17. Метод кластеризации: "Двувходовое объединение".*

*18. Метод кластеризации: "Метод K средних".*

##### *Рекомендуемая литература*

1. Луценко Е.В. Теоретические основы и технология адаптивного семантического анализа в поддержке принятия решений (на примере универсальной автоматизированной системы распознавания образов "ЭЙДОС-5.1"). - Краснодар: КЮИ МВД РФ, 1996. - 280с.

2. Луценко Е. В.       Автоматизированный системно-когнитивный анализ в управлении активными объектами (системная теория информации и ее применение в исследовании экономических, социально-психологических, технологических и организационно-технических систем): Монография (научное издание). – Краснодар:  КубГАУ. 2002. – 605 с.