
**МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ
ФЕДЕРАЦИИ**

Федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования

**«ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ СИСТЕМ
УПРАВЛЕНИЯ И РАДИОЭЛЕКТРОНИКИ» (ТУСУР)**

УТВЕРЖДАЮ

Заведующий кафедрой ЭМИС

_____ И. Г. Боровской

«__» _____ 2012 г.

С.И. КОЛЕСНИКОВА

Научный семинар «Распознавание образов»

Методические указания к практическим работам

2012

А Н Н О Т А Ц И Я

Цели настоящих методических указаний: 1) освоение основных понятий и определений теории распознавания образов; 2) приобретение практических навыков в построении алгоритмов распознавания и анализ их качества. В четырех частях указаний приведены примеры задач и методов их решения (анализа возможного решения) на следующие темы:

1. Математические основы теории распознавания образов.
2. Методы распознавания образов.
3. Алгебраический подход к задаче распознавания.
4. Распознавание образов и распознавание изображений.

Теоретический материал приведен *только тот и в том объеме*, который необходим для решения предлагаемых задач. Задачи контрольных заданий являются весьма простыми, они предназначены для усвоения основных начальных понятий и основ теории массового обслуживания. Предполагается, что студенты знают математику в объеме, требуемом в техническом ВУЗе.

Методические указания предназначены для студентов экономического факультета.

СОДЕРЖАНИЕ ПРАКТИЧЕСКИХ ЗАНЯТИЙ
по дисциплине «Научный семинар «Распознавание образов»»
и руководство по выполнению
для студентов направления 230100.68 – Информатика и вычислительная
техника. Профиль - Информационное и программное обеспечение
автоматизированных систем

Краткое содержание тем и планируемых результатов их освоения	4
ПОРЯДОК ПРОВЕДЕНИЯ Практических работ	4
ПОРЯДОК ПРОВЕДЕНИЯ Интерактивных занятий-СЕМИНАРОВ.....	5
Раздел 1. Математические основы теории распознавания образов	6
ПРИМЕРЫ ТИПОВЫХ АУДИТОРНЫХ ЗАДАНИЙ	6
Варианты домашних Заданий к разделу 1	8
Варианты контрольных Заданий к разделу 1	8
Контрольные вопросы к разделу 1	9
Раздел 2. Методы распознавания образов	9
Интерактивное занятие-семинар №3: Детерминистские методы распознавания образов	9
ПРИМЕРЫ ТИПОВЫХ АУДИТОРНЫХ ЗАДАНИЙ	9
Интерактивное занятие-семинар №4 по теме: Статистические методы распознавания образов	15
ПРИМЕРЫ ТИПОВЫХ АУДИТОРНЫХ ЗАДАНИЙ	15
Варианты домашних Заданий к разделу 2	20
Варианты контрольных Заданий к разделу 2.....	21
Контрольные вопросы к разделу 2	21
Раздел 3. Алгебраический подход к задаче распознавания	21
ПРИМЕРЫ ТИПОВЫХ АУДИТОРНЫХ ЗАДАНИЙ	21
Интерактивные занятия-семинары №7, 8 по теме: Алгебраические методы в задачах распознавания и классификации. Эффективность систем распознавания с коллективным распознаванием	23
ПРИМЕРЫ ТИПОВЫХ АУДИТОРНЫХ ЗАДАНИЙ	23
Варианты домашних Заданий к разделу 3	25
Варианты контрольных Заданий к разделу 3.....	25
Контрольные вопросы к разделу 3	26
Раздел 4. Распознавание образов и распознавание изображений	26
Интерактивные занятия-семинары №9, 10 по теме: Распознавание образов и распознавание изображений. Системы РО на основе нейросети	26
ПРИМЕРЫ ТИПОВЫХ АУДИТОРНЫХ ЗАДАНИЙ	26
Варианты домашних Заданий к разделу 4	29
Варианты контрольных Заданий к разделу 4.....	29
Контрольные вопросы к разделу 4	29
Использованная литература.....	30

Обозначения: ИДЗ - индивидуальные домашние задания
 ИГЗ - индивидуальные групповые задания
 СРС - самостоятельная работа студентов
 ИнЗ - интерактивное занятие
 ТРО - теория распознавания образов

 3-Эл – знания элементарные (определения, понятия, умение приводить иллюстрирующие примеры);

 3-Пр – знания продуктивные (умение применить знания элементарные для решения учебных задач);

У-Эл – «умения» элементарные (уметь пользоваться готовыми частными алгоритмами для решения типовых задач), умение решать задачи по шаблону (копировать);

У-Пр – «умения» продуктивные (применять положения и известные частные алгоритмы дисциплины для решения практических задач);

В-Эл – элементарное владение методами дисциплины и уверенное осуществление (построение) основных операций для решения типовых задач;

В-Пр – продуктивно распознавать проблемы, алгоритмизировать их анализ и применять методы дисциплины для решения практических задач;

С.в. - случайная величина.

Краткое содержание тем и планируемых результатов их освоения

Тема практических занятий	Деятельность студента. Решая задачи, студент:	Отрабатываемые компетенции/ ожидаемый уровень освоения
1. Математические основы теории распознавания образов	<ul style="list-style-type: none"> • <i>знакомится</i> с принципами теории распознавания образов; • <i>использует</i> определения и понятия теории распознавания образов; • <i>использует</i> знания, полученные ранее в курсе теории вероятностей и математической статистики, в курсе дискретной математики; • <i>строит</i> модель тестового распознавания для конкретной предметной области; 	ОК-1, ОК-2, / 3-Эл, У-Эл, В-Эл ПК-5/ 3-Пр, У-Пр, В-Пр
2. Методы распознавания образов	<ul style="list-style-type: none"> • <i>выбирает</i> тип метода РО для текстовой задачи; • <i>определяет</i> основные признаки объектов для конкретного метода РО; • <i>учится</i> применять критерии эффективности для оптимизации методов ТРО; • <i>применяет</i> статистические методы распознавания; • <i>применяет</i> детерминистские методы распознавания; • <i>применяет</i> логические методы распознавания; 	ОК-1, ОК-2, / 3-Эл, У-Эл, В-Эл ПК-5/ 3-Пр, У-Пр, В-Пр
3. Алгебраический подход к задаче распознавания	<ul style="list-style-type: none"> • <i>учится</i> применять критерии эффективности для оптимизации алгоритмов вычисления оценок; • <i>изучает</i> алгебраический подход к задаче распознавания; • <i>учится</i> применять и строить алгоритмические композиции для решения практических задач. 	ОК-1, ОК-2, / 3-Эл, У-Эл, В-Эл ПК-5/ 3-Пр, У-Пр, В-Пр
4. Распознавание образов и распознавание изображений	<ul style="list-style-type: none"> • <i>разрабатывает</i> совместно с преподавателем алгоритм решения задач распознавания изображений на базе известных подходов. • <i>знакомится</i> с принципом перцептрона; • <i>строит</i> однослойную сеть для решения простых задач. 	ОК-1, ОК-2, / 3-Эл, У-Эл, В-Эл ПК-5/ 3-Пр, У-Пр, В-Пр

ПОРЯДОК ПРОВЕДЕНИЯ ПРАКТИЧЕСКИХ РАБОТ

1. Ознакомиться с нижеуказанной темой в основной и дополнительной литературе.
2. Ознакомиться со справочными интернет-сведениями (СРС).
3. Ознакомиться с принципом решения задач аудиторных.
4. Рекомендуется решить задачи домашние (в рамках СРС).
5. Ознакомиться с планом проведения интерактивных занятий в случае их проведения, прилагающегося к каждому разделу, и принципом подготовки к нему.
6. Составить и предоставить преподавателю отчет о работе, если он входит в форму отчетности по данному разделу знаний.

ПОРЯДОК ПРОВЕДЕНИЯ ИНТЕРАКТИВНЫХ ЗАНЯТИЙ-СЕМИНАРОВ

1. Ознакомиться со справочными интернет-сведениями (подготовка к ИнЗ в указаниях по СРС).
2. Ознакомиться с указанной темой в основной и дополнительной литературе.

Основная литература

1. Горелик А. Л., Скрипкин В. А. Методы распознавания: Учебное пособие для вузов. - 4-е изд., испр. - М.: Высшая школа, 2004. – 260 с.
2. Лапко А.В. Непараметрические системы обработки информации : Учебное пособие для вузов / А. В. Лапко, С. В. Ченцов; Российская Академия наук. Сибирское отделение, Институт вычислительного моделирования. - М. : Наука, 2000. - 349 с.
3. Воронцов К.В. Лекции по методам оценивания и выбора моделей. 2007. Режим доступа: www.ccas.ru/voron/download/Modeling.pdf.

Дополнительная литература

1. Р. Гонсалес. Цифровая обработка изображений в среде MATLAB: Пер. с англ. / Р. Гонсалес, Р. Вудс, С. Эддинс ; пер. : В. В. Чепыжов. - М. : Техносфера, 2006. – 615 с.
2. Ту Д., Гонсалес Р. Принципы распознавания образов. – М.: Мир, 1978, 2008.
3. Вапник В.Н. и др. Алгоритмы и программы восстановления зависимостей: Практическое руководство. - М. : Наука. Физматлит, 1984. - 816 с.
4. Вапник В.Н., Червоненкис А.Я. Теория распознавания образов. М.: Наука, 1974, 2002.- 415 с.
5. Загоруйко Н.Г. Прикладные методы анализа данных и знаний. // Новосибирск. Изд-во института математики. 1999, 2008.
6. Дадашев Т.М. Теория распознавания образов (логические методы): Учебное пособие. - М.: МФТИ, 1982, 2006. - 84 с.
7. Айзерман А.А., Браверман Э.М., Розоноэр Э.И. Метод потенциальных функций в теории обучения машин. – М.: Наука, 1970.
8. Патрик Э. Основы теории распознавания образов. – М.: Сов. радио, 1980.
9. Фу К.С. Структурные методы в распознавании образов. – М.: Мир, 1977.
10. Дуда Р., Харт П. Распознавание образов и анализ сцен. - М.: Мир, 1976.- 511 с.
11. Дюкова Е.В., Песков Н.В. Построение распознающих процедур на базе элементарных классификаторов // [www.ccas.ru /frc/papers /djukova05 construction.pdf](http://www.ccas.ru/frc/papers/djukova05construction.pdf).
12. Воронцов К.В. Обзор современных исследований по проблеме качества обучения алгоритмов. Таврический вестник информатики и математики. – 2004. – № 1. – С. 5 – 24. <http://www.ccas.ru/frc/papers/voron04twim.pdf>.

Электронный учебно-методический комплекс курса

Программное обеспечение: электронный учебно-методический комплекс курса, размещенный на сервере ЭФ по адресу: student\Колесникова\ТРО

Базы данных, информационно-справочные и поисковые системы

www.ccas.ru/voron/

[www.ccas.ru /frc/papers /djukova05 construction.pdf](http://www.ccas.ru/frc/papers/djukova05_construction.pdf).

<http://www.all-library.com/obrazovanie/nauka/42843-osnovy-teorii-raspoznaniya-obrazov.html>

<http://window.edu.ru/resource/738/20738>

<http://www.bsu.by/Cache/pdf/229903.pdf>

<http://sumschool.sumdu.edu.ua>

3. Ознакомиться с принципом решения задач аудиторных.
4. Рекомендуется решить задачи домашние (в рамках СРС).
5. Ознакомиться с планом проведения интерактивных занятий и принципом подготовки к нему. Обсудить с преподавателем частные вопросы, прилагающиеся к каждому ИнЗ.
6. Ознакомиться с формой текущего контроля освоения компетенций ОК-1, ОК-2 уровня З-Эл, У-Эл, В-Эл; ПК-5 уровня З-Пр, У-Пр, В-Пр (см. табл.1): *отчет* по решению следующих практических текстовых задач:
7. Составить и предоставить преподавателю отчет о работе по установленной форме.

Раздел 1. Математические основы теории распознавания образов

ПРИМЕРЫ ТИПОВЫХ АУДИТОРНЫХ ЗАДАНИЙ

ПРАКТИЧЕСКОЕ ЗАНЯТИЕ 1. Постановка задачи распознавания. Основы теории распознавания образов (данные, знания, гипотеза, закономерность, признак).

Цель работы

Знакомство с основными понятиями математической дисциплины «Теория распознавания образов», изучающей закономерности массовых случайных явлений (процессов).

Задача 1.1. Перечислите меры расстояния и укажите условия применимости.

Евклидово расстояние. Это вероятно наиболее часто используемый тип расстояния. Оно является простым геометрическим расстоянием в многомерном

пространстве и вычисляется как: $d_2(x, y) = \left(\sum_i (x_i - y_i)^2 \right)^{1/2}$.

Используется и *квадрат евклидова расстояния*, если мы хотим придать прогрессивно возрастающий вес объектам, которые являются более удаленными.

Это расстояние вычисляется как: $d'_2(x, y) = \sum_i (x_i - y_i)^2$.

Метрика Хемминга (покоординатное расстояние, городских кварталов, манхэттенское расстояние). Это расстояние в некотором смысле усредняет разницу между различными компонентами векторов. В большинстве случаев, эта мера расстояния дает результаты, подобные простому евклидову расстоянию. Однако, отметим, что при данной мере, эффект привносимый отдельными большими компонентами демпфируется (так как они не возводятся в квадрат). Покоординатное расстояние вычисляется так: $d_1(x, y) = \sum_i |x_i - y_i|$.

Расстояние Чебышева. Эта мера расстояния может подойти в случае, когда нам потребуется определить два объекта как различные, если они различны хотя бы по одному измерению: $d_{\infty}(x,y) = \max_i |x_i - y_i|$.

Степенное расстояние. Иногда может потребоваться увеличить или уменьшить вес увеличения расстояний по измерениям. Это может быть достигнуто путем использования степенного расстояния. Расстояние это вычисляется как:

$$d_{pr}(x,y) = \left(\sum_i (x_i - y_i)^p \right)^{1/r} \text{ где } r \text{ и } p - \text{определяемые пользователем параметры.}$$

Поведение данной меры выглядит следующим образом: Параметр p контролирует вес разностей по отдельным компонентам, параметр r контролирует вес придаваемый расстоянию между объектами в целом. Если r и p равны 2, то это расстояние равно Евклидову расстоянию, при $r=p$ – мера Минковского.

Мера «доли рассогласования». Мера целесообразна для номинальных признаков.

Это расстояние вычисляется как: $d(x,y) = \frac{\text{quantity } x \neq y}{i}$.

Обобщённая мера расстояний Минковского $d_p(x,y) = \left(\sum_i (x_i - y_i)^p \right)^{1/p}$ (при $p=1$ –

метрика Хемминга, при $p=2$ – метрика Евклида, при $p=\infty$ – метрика Чебышева).

Существуют и другие виды расстояний (Махаланобиса) и функции близости-различия объектов (FRiS-функция (Загоруйко Н.Г.)), не являющиеся расстоянием в общепринятом смысле (симметричность $d(x,y) = d(y,x)$, $d(x,y) = 0$ при $x=y$, неотрицательность $d(x,y) \geq 0$). Ознакомьтесь с другими подходами по оценке расстояний самостоятельно.

Задача 1.2. [8] Дана обучающая выборка двух образов (I и II) в пространстве двух бинарных признаков Z_1 и Z_2 . Сформулировать решающее правило для разделения двух классов.

	Z_1	Z_2
I	0	0
	1	1
II	0	1
	1	0

Решение. Проекция реализаций на каждую ось показывают, что оба признака Z_1 и Z_2 по отдельности неинформативны. Использование этих признаков в системе позволяет найти правило для распознавания двух образов: признаки Z_1 и Z_2 у реализаций I-го образа имеют одинаковые значения, а у II-го образа – разные.

Задача 1.3. Дано множество прямоугольников со сторонами, параллельными осям координат как множество точек в двумерном признаковом пространстве. Сформулировать решающее правило для разделения возможных классов. Указать число классов.

Решение. В случае двух образов – вертикально (I-й образ) и горизонтально (II-й образ) вытянутые прямоугольники, решающее правило может быть выбрано в виде биссектрисы угла в начале координат. Все точки (объекты), лежащие выше биссектрисы, относятся к образу I, ниже – к образу II.

ВАРИАНТЫ ДОМАШНИХ ЗАДАНИЙ К РАЗДЕЛУ 1

Задача Д1.1. Составить алгоритм распознавания 6-ти английских букв (блок-схему), определив предварительно признаковое пространство: Буквы: V O J I 6 8.

Задача Д1.2. Дан многоугольник, заданный координатами точек-вершин (x_i, y_i) . Дана точка, заданная координатами (x_0, y_0) . Сформулировать решающее правило для определения, принадлежит ли заданная точка внутренности многоугольника. Рассмотреть отдельно случаи: а) выпуклый многоугольник; б) невыпуклый многоугольник. Для ответа достаточно изложить алгоритм в виде блок-схемы. Программа оценивается дополнительно.

Задача Д1.3. Предложить алгоритм отнесения какого-либо конкретного плода к определенной группе (классу). Решить, сколько будет классов. Выбор признаков и ответ обосновать. Объекты, подлежащие классификации: арбуз, дыня, апельсин, лимон, грейпфрут, яблоко, огурец, груша, кабачок, баклажан, клюква, брусника, облепиха.

ВАРИАНТЫ КОНТРОЛЬНЫХ ЗАДАНИЙ К РАЗДЕЛУ 1

Следует отобрать признаки, при помощи которых можно отличить левые шесть картинок от правых шести. Составить алгоритм распознавания.

I вариант

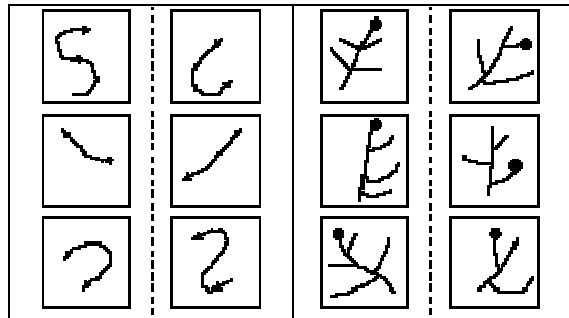


Рис. 2.1.

II вариант

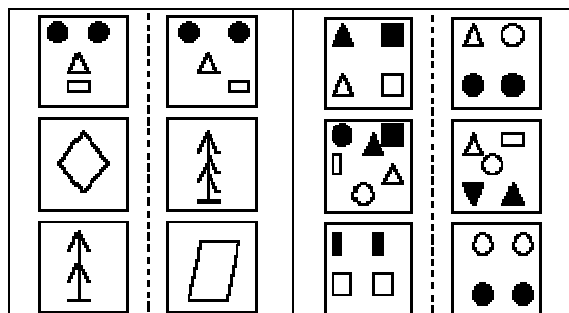


Рис. 2.2.

III вариант

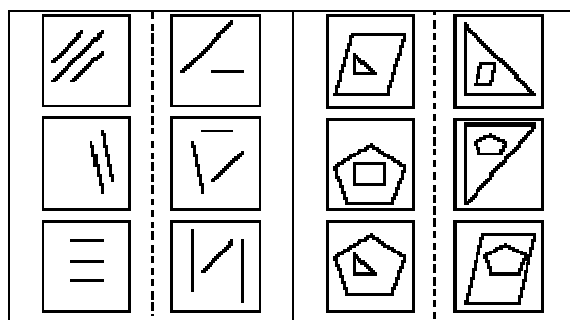


Рис. 2.3.

Контрольные вопросы к разделу 1

1. Дайте определение терминам: алгоритм распознавания, алфавит; признаковое пространство, выборочное пространство, обучение без учителя, самообучение, кластерный анализ, таксономия.
2. Назовите критерии информативности признаков.
3. Решающее правило, риск потерь при распознавании.

Раздел 2. Методы распознавания образов

Интерактивное занятие-семинар №3: Детерминистские методы распознавания образов

Цель занятия: активное воспроизведение полученных знаний на лекциях по разделу 1 в «незнакомых» условиях: применение основных понятий ТРО для решения практических задач; построение детерминированных моделей для текстовых задач и расчет числовых характеристик эффективности метода распознавания с применением вычислительных средств (Excel, MatLab).

Планируемые к приглашению на семинар специалисты-эксперты: Матросова А.Ю., д.т.н., профессор ТГУ, зав. каф. программирования, специалист по методам математической логики; Цой Ю.Р., к.т.н., доцент ТПУ, специалист по методам интеллектуального анализа данных (ИАД);

ПРИМЕРЫ ТИПОВЫХ АУДИТОРНЫХ ЗАДАНИЙ

Задача 2.1. Задана следующая таблица обучения и подлежащая распознаванию строка ω' . Выбрать подходящий алгоритм, выбор обосновать, определить принадлежность данной строки какому-либо образу на основе обучения и выбранного алгоритма.

Классы	Объекты	Значения признаков					
		X ₁	X ₂	X ₃	X ₄	X ₅	X ₆
Ω_1	ω_{11}	5	11	9	3	3	1
	ω_{12}	4	10	2	7	12	1
	ω_{13}	9	5	4	6	11	1
	ω_{14}	7	13	3	4	6	2
	ω_{15}	2	14	8	5	9	1
Ω_2	ω_{12}	5	9	2	8	14	1
	ω_{22}	4	6	7	3	13	1
	ω_{23}	6	11	9	11	5	1
	ω_{24}	7	10	4	2	12	1

	ω_{25}	3	10	5	9	7	1
	ω	3	13	7	4	8	2

Решение. Для решения задачи выберем метод эталонов. Для объектов каждого класса-образа строился объект-эталон, после чего находилось расстояние от каждого из эталонов до распознаваемого объекта. Листинг программы приводится ниже:

```
#include <vector>
#include <math.h>
class obj
{
public:
    std::vector<float> v;
    obj() {};
    obj(float i1, float i2, float i3, float i4, float i5, float i6)
    {
        v.push_back(i1);
        v.push_back(i2);
        v.push_back(i3);
        v.push_back(i4);
        v.push_back(i5);
        v.push_back(i6);
    }
};
int _tmain(int argc, _TCHAR* argv[])
{
    obj o1[5] = {
        obj(5, 11, 9, 3, 3, 1),
        obj(4, 10, 2, 7, 12, 1),
        obj(9, 5, 4, 6, 11, 1),
        obj(7, 13, 3, 4, 6, 2),
        obj(2, 14, 8, 5, 9, 1)
    };
    obj o2[5] = {
        obj(5, 9, 2, 8, 14, 1),
        obj(4, 6, 7, 3, 13, 1),
        obj(6, 11, 9, 11, 5, 1),
        obj(7, 10, 4, 2, 12, 1),
        obj(3, 10, 5, 9, 7, 1)
    };
    obj e1, e2;
    for(int i = 0; i < 6; ++i)
    {
        float x_mean1 = 0, x_mean2 = 0;
        for(int j = 0; j < 5; ++j)
        {
            x_mean1 += o1[j].v[i];
            x_mean2 += o2[j].v[i];
        }
        x_mean1 /= 5;
        x_mean2 /= 5;
        e1.v.push_back(x_mean1);
    }
}
```

```

        e2.v.push_back(x_mean2);
    }
    float r1, r2;
    obj r(3, 13, 7, 4, 8, 2);
    r1 = fabs((e1.v[0] - r.v[0])*(e1.v[1] - r.v[1])*(e1.v[2] - r.v[2])*(e1.v[3] -
r.v[3])*(e1.v[4] - r.v[4])*(e1.v[5] - r.v[5]));
    r2 = fabs((e2.v[0] - r.v[0])*(e2.v[1] - r.v[1])*(e2.v[2] - r.v[2])*(e2.v[3] -
r.v[3])*(e2.v[4] - r.v[4])*(e2.v[5] - r.v[5]));
    r1 = powf(r1, 1.f/6);
    r2 = powf(r2, 1.f/6);
    return 0;
}

```

Результаты работы программы:

Расстояние от распознаваемого объекта до первого эталона - 1.0880172

Расстояние от распознаваемого объекта до второго эталона - 2.0279393

Таким образом, по методу эталонов распознаваемый объект относится к классу Ω_1 .






Задача 2.2. Составить алгоритм распознавания подсчета углов в многоугольниках (блок-схему и/или программу), определив предварительно признаковое пространство.

По конвейеру движутся детали и заготовки, имеющие форму многоугольников (точнее сказать, их проекции на плоскость конвейера представляют собой многоугольники). Над конвейером расположен робот, который с помощью камеры осматривает каждую деталь и определяет количество углов в ней. После этого он сортирует детали: треугольники к треугольникам, четырехугольники к четырехугольникам и т.д.

Имеется *вариант* решения задачи. Есть ли в нем ошибки, приводящие к неверному ответу? Является ли предложенный алгоритм оптимальным (в смысле минимального признакового описания)?

Вариант решения задачи. Предположим, что все распознаваемые многоугольники – правильные. Известно, что угол правильного n -угольника вычисляется по формуле $\frac{180^\circ(n-2)}{n}$, где n – число его сторон. Для рассмотрения примера алгоритма решения задачи ограничимся первыми 4-мя правильными многоугольниками – треугольник, квадрат, пятиугольник и шестиугольник.

Основной словарь признаков:

- a -  (горизонтальная черта, угол 0°)
- b -  (наклонная черта под углом 60°)
- c -  (вертикальная черта, угол 90°)
- d -  (наклонная черта под углом 108°)
- e -  (наклонная черта под углом 120°)

Алгоритм может быть следующим:

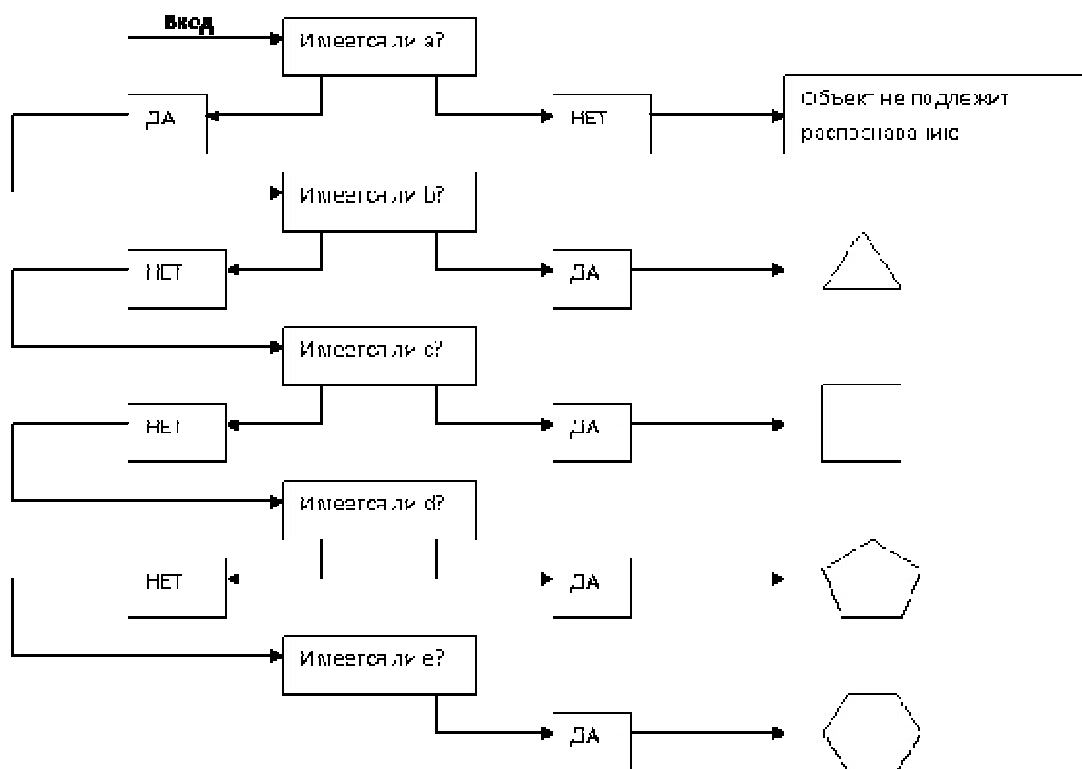


Рис. 2.4.

Задача 2.3. [1] Предположим, что у противника имеются три типа мин: K_0 – осколочного действия, $K_{оф}$ – осколочно-фугасного действия, $K_{ф}$ – фугасного действия. В инструкции по применению этих мин сказано, что:

- осколочные мины применяются на равнинной местности с каменистым грунтом или же на песчаных холмах;
- осколочно-фугасные мины применяются либо на равнине, либо на местности с каменистым грунтом;
- фугасные мины не применяются, если грунт каменистый, а местность – холмистая.

Требуется определить:

- а) какие мины будут применены противником в зависимости от вида ландшафта и типа грунта?
- б) что можно сказать о свойствах местности, если известно, что противник применяет только осколочно-фугасные мины?

Решение. Предлагается вариант решения, который следует проанализировать и, если обнаружится противоречие, найти способ его исправления.

Для решения задачи введем обозначения в терминах алгебры логики: A – равнинная местность; \bar{A} – холмистая местность; B – каменистый грунт; \bar{B} – песчаный грунт.

Тогда условия можно записать следующим образом:

- $A \cdot B + \bar{A} \cdot \bar{B} \rightarrow K_0$
- $A + B \rightarrow K_{оф}$
- $A \cdot \bar{B} \rightarrow K_{ф}$

Всего можно составить 4 пары сочетаний местность-грунт:

$$\begin{cases} A \cdot B \\ A \cdot \bar{B} \\ \bar{A} \cdot B \\ \bar{A} \cdot \bar{B} \end{cases}$$

Исходя из определений конъюнкции и дизъюнкции (в частности, из определений истинности конъюнктивно и дизъюнктивно связанных высказываний), можно сделать следующие выводы:

$$\begin{cases} A \cdot B \rightarrow K_o, K_{of} \\ A \cdot \bar{B} \rightarrow K_{of}, K_{\phi} \\ \bar{A} \cdot B \rightarrow K_{of} \\ \bar{A} \cdot \bar{B} \rightarrow K_o \end{cases}$$

Таким образом, если местность равнинная, а грунт каменистый, противник будет применять осколочные или осколочно-фугасные мины. Если местность равнинная, а грунт песчаный – осколочно-фугасные или фугасные мины. Если местность холмистая, а грунт каменистый – осколочно-фугасные мины. Если местность холмистая, а грунт песчаный – осколочные мины.

Ход занятий №ИЗ, 4.

Вступление. Сообщение темы и обоснование ее актуальности через вышеуказанные задачи. Ведущий студент, ответственный за выбор и подачу необходимой информации, согласует алгоритм занятия.

Основная часть:

- I. Сообщение в виде доклада-презентации ответственными двумя студентами за проведение занятия, в котором излагается суть обсуждаемых положений:
 - 1) Детерминистские методы распознавания образов (обзор).
 - 2) Логические методы распознавания образов.

Задача И2.1. Задана следующая таблица обучения и подлежащая распознаванию строка ω' . Определить принадлежность данной строки какому-либо образу.

Классы	Объекты	Значения признаков			
		X ₁	X ₂	X ₃	X ₄
Ω ₁	ω ₁	1	1	1	1
	ω ₂	1	0	1	1
	ω ₃	0	0	1	0
Ω ₂	ω ₄	1	0	1	0
	ω ₅	0	0	0	0
	ω ₆	0	0	1	1
	ω'	1	1	0	0

Задача И2.2. [1] *Логический метод распознавания.* Предположим, что на острове находятся 2 самолетные опознавательные башни. В течение нескольких дней в небе летают одни и те же вражеские самолеты. Опознать тип наблюдаемых самолетов трудно, и это привело к некоторой полемике между двумя наблюдательными пунктами. Тем не менее, было сделано предположение (далекое от определенности), что это наблюдаются 4 типа вражеских самолетов:

А, В, Х, Y (причем определенно известно, что типы А и В существуют). На протяжении 3-х дней от каждого поста поступают следующие сообщения:

Пост 1	Пост 1
1-й день. Самолеты типов Х и Y. 2-й день. Самолеты типа А или типа В, или же как типа А, так и типа В одновременно. 3-й день. Самолеты типа Х и одновременно типа А, или типа В; или типа А и В; или же самолеты типа А и типа Y..	1-й день. Самолеты типов А и не В. 2-й день. Самолеты типа Y и не А или же типа Х. 3-й день. Самолеты типа А

Требуется определить, можно ли на основе только этих сообщений заключить, что самолеты типов Х и Y в действительности являются самолетами типов А и В.

Задача И2.3. Предположим, что на основе данных, полученных из разных источников, были составлены следующие высказывания:

- 1) Самолет с реактивным двигателем и малым радиусом действия – бомбардировщик.
- 2) Поршневые двигатели бомбардировщиков покрыты тяжелой броней.
- 3) Поршневые двигатели истребителей рассчитаны на малый радиус действия.
- 4) Поршневые самолетные двигатели, рассчитанные на большой радиус действия, имеют легкую броню.
- 5) Реактивные самолеты имеют тяжелую броню.
- 6) Истребители представляют собой самолеты, покрытые тяжелой броней и с малым радиусом действия.
- 7) Легкую броню имеют или самолеты с большим радиусом действия или истребители.
- 8) Тяжелую броню имеют или самолеты с поршневым двигателем или самолеты с малым радиусом действия.

На основании анализа этих высказываний необходимо дать ответы на следующие вопросы:

- 1) Все ли утверждения совместны (непротиворечивы)?
- 2) Если высказывания несовместны, то будем предполагать, что только одно из них неправильно. Может ли быть одно утверждение отброшено с тем, чтобы оставшиеся высказывания были совместны, и если да, то какое это высказывание?
- 3) Зависимы ли какие-либо высказывания?
- 4) Не являются какие-либо высказывания избыточными?
- 5) Какие заключения можно сделать при различных предположениях об ошибочности отдельных высказываний?

II. Выяснение позиций участников с зафиксированными точками зрения на решение вышеизложенных задач.

Итог II-го этапа: формирование целевых групп по общности позиций каждой из групп.

III. Организация коммуникации между группами: 1) выяснение позиции-варианта решения выявленных групп и защита занятой позиции; 2) формирование нового набора вариантов решений на основании общего обсуждения; 3) выбор одного решения голосованием;

IV. Повторная защита позиций-вариантов групп после проведения расчетов с целью оценки отклонения от «истинного» решения (парное оценивание).

Выводы: реализован самостоятельный поиск учащимися путей и вариантов решения поставленной учебной задачи (выбор одного из предложенных вариантов или нахождение собственного варианта и обоснование решения на базе коллективной интерактивной работы).

Итог занятий №ИЗ: Оценивание компетенций (табл.1) по результатам работы на занятиях (активность, инициативность, грамотность, обоснованность защищаемой позиции) и своевременности сдачи отчета по решению практических задач И2.1-И2.3.

Таблица 1

№	№ задач и	Вид (совмещение нескольких видов) интерактивной работы	Трудоемкость (час)	Отрабатываемые компетенции/ожидаемый уровень освоения	Оценка личностных качеств	Контроль выполнения работы (участие в полемике, индивидуальные групповые задания (ИГЗ) и т.д)
1	И2.1	Работа в команде. Решение ситуационных задач.	1	ОК-1, ОК-2/ З-Эл, У-Эл, В-Эл ПК-5/ З-Пр, У-Пр, В-Пр	Качество работы; своевременность сдачи отчета по решению ИГЗ	ИГЗ. Критерии оценивания поведения на занятии: активность, инициативность, грамотность, обоснованность защищаемой позиции.
2	И2.2-И2.3	Работа в команде. Решение ситуационных задач. Исследовательский метод	3	ОК-1, ОК-2/ З-Эл, У-Эл, В-Эл ПК-5/ З-Пр, У-Пр, В-Пр	Качество работы; своевременность сдачи отчета по решению ИГЗ	ИГЗ. Критерии оценивания поведения на занятии: активность, инициативность, грамотность, обоснованность защищаемой позиции.
Всего			4			

Интерактивное занятие-семинар №4 по теме: Статистические методы распознавания образов

Цель занятия: активное воспроизведение полученных знаний на лекциях по разделу 1 в «незнакомых» условиях: применение основных понятий ТРО для решения практических задач; построение статистических моделей для текстовых задач и расчет числовых характеристик эффективности метода распознавания с применением вычислительных средств (Excel, MatLab).

Планируемые к приглашению на семинар специалисты-эксперты: Конев В.В., д.ф.-м.н., профессор ТГУ, зав. каф. ВМиММ, специалист по методам математической статистики; Цой Ю.Р., к.т.н., доцент ТПУ, специалист по методам интеллектуального анализа данных (ИАД).

ПРИМЕРЫ ТИПОВЫХ АУДИТОРНЫХ ЗАДАНИЙ

Задача 2.4. Составить алгоритм метода «ФОРЕЛЬ» [8].

Решение. В методе используется критерий, оценивающий плотность распределений образов в объеме, ограниченном сферой фиксированного радиуса.

При переходе от шага к шагу центр этой сферы движется в сторону увеличения плотности точек - образов. Сфера стабилизируется в определенном положении, когда плотность точек внутри ее становится максимальной, и любое перемещение сферы ведет к ухудшению ситуации. Точки-образы, попадающие внутрь стабилизировавшейся сферы, образуют класс K_1 . Процесс повторяется с любой точки из оставшихся и завершается новой стабилизацией сферы, внутри которой находится класс K_2 и т.д. Формируемое число классов зависит от величины радиуса сферы.

Алгоритм

Инициализация. Задается радиус сферы R .

1. Строится сфера радиуса R с центром в любой точке так, чтобы внутри среды попала хотя бы одна точка - образ Z_1 . Определяется центр сферы $C^{(1)}$.

2. Определяются точки - образы $Z_1^{(1)}$, для которых $|Z_{1(1)} - C_1^{(1)}| < R$ т.е. эти точки попадают внутрь сферы.

3. Вычисляется центр $C^{(2)}$ тяжести $Z_1^{(1)}$ с координатами $X_C = 1/k \sum_{I=1}^k x_I$;

$$Y_C = 1/k \sum_{I=1}^k y_I^{(1)}.$$

4. Строится сфера радиуса R с центром в точке $C^{(2)}$. Определяются точки $Z_1^{(2)}$, для которых $|Z_1^{(2)} - C_1^{(2)}| < R$.

5. Вычисляется центр $C^{(3)}$ тяжести $Z_1^{(2)}$.

Процесс переходов от $C^{(k)}$ к $C^{(k+1)}$ останавливается тогда, когда расстояние $|C^{(k)} - C^{(k+1)}| < \Delta$.

Сфера с центром в $C^{(k+1)}$ представляет собой таксон K_1 .

6. Точки - образы, принадлежащие K_1 из рассмотрения исключаются. В качестве начальной берётся любая точка - образ из оставшихся, и процесс продолжается с шага 1.

7. Последовательность таксонов K_1, K_2, \dots , является последовательностью классов, соответствующих радиусу сферы $= R$.

Замечание 2.3. «Оптимальное» разбиение набора образов на классы диктуется минимумом суммарного объема сфер - таксонов.

Задача 2.5. Найти дискриминантную функцию для классификации двух образов:

$$X_1 = \begin{Bmatrix} 1 \\ 4 \end{Bmatrix}; X_2 = \begin{Bmatrix} 4 \\ 2 \end{Bmatrix}.$$

Решение. Дискриминантная функция — это функция $d(\bar{X})$, которая определяет решающую поверхность (для $n=3$ — линия, для $n=3$ - плоскость, для $n=4$ или больше - гиперплоскость). В двухмерном признаковом пространстве $d_k(X)$ и $d_j(X)$ -дискриминантные функции для образа X в классах k и j соответственно (рис.2.5).

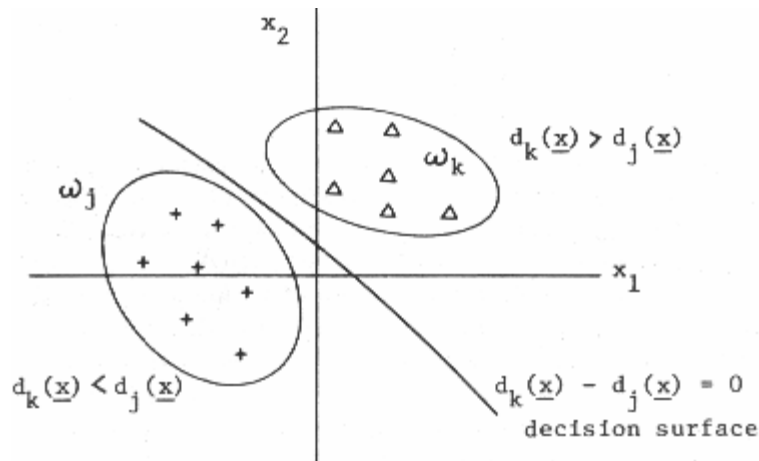


Рис. 2.5.

Здесь $d(X) = d_k(X) - d_j(X) = 0$ будет уравнением, которое определяет поверхность, разделяющую классы k и j . Уравнение линии: $w_1 \cdot x_1 + w_2 \cdot x_2 + w_{n+1} = 0$, или в матричной форме:

$$W^T \cdot X = 0, \text{ где } W = \begin{Bmatrix} w_1 \\ \dots \\ w_n \\ w_{n+1} \end{Bmatrix}, X = \begin{bmatrix} x_1 \\ \dots \\ x_n \\ 1 \end{bmatrix},$$

где W и X называются *расширенным весовым вектором* и *расширенным вектором образа*.

Здесь для удобства скаляр w_{n+1} добавлен в весовой вектор и, соответственно, вектор \bar{X} становится $(n+1)$ -мерным путем добавления $x_{n+1} = 1$. Этот прием позволяет представить все линейные дискриминантные функции проходящими через начало координат.

Выберем $d(X) = X_1 - 0.5X_2 - 2$ и вычислим величины $d(X_1)$, $d(X_2)$:

$$d(X_1) = w' \cdot \bar{X}_1 = (1 - 0.5 - 2) \begin{pmatrix} 1 \\ 4 \\ 1 \end{pmatrix} = -3 < 0; \quad d(X_2) = w' \cdot \bar{X}_2 = (1 - 0.5 - 2) \begin{pmatrix} 4 \\ 2 \\ 1 \end{pmatrix} = +1 > 0$$

Таким образом, $d(X) = X_1 - 0.5X_2 - 2$ можно использовать как дискриминантную функцию.

Задача 2.6. Дано два набора точек:

$$(1, 2), (2, 2), (3, 1), (3, 2), (2, 3) \in K_1, (7, 9), (8, 9), (9, 8), (9, 9), (8, 10) \in K_2.$$

Найти разделяющую границу - решающую поверхность.

Решение. Находим средние:

$$m_1 = \frac{1}{N_1} \sum_{j=1}^{N_1} x_{1j} = \frac{1}{5} \begin{bmatrix} 11 \\ 10 \end{bmatrix}$$

$$m_2 = \frac{1}{N_2} \sum_{j=1}^{N_2} x_{2j} = \frac{1}{5} \begin{bmatrix} 41 \\ 45 \end{bmatrix}$$

Находим матрицы ковариаций:

$$C_1 = \frac{1}{N_1} \sum_{j=1}^{N_1} x_{1j} x_{1j}^T - m_1 m_1^T = \frac{1}{5} \left[\begin{pmatrix} 1 \\ 2 \end{pmatrix} (1 \ 2) + \begin{pmatrix} 2 \\ 2 \end{pmatrix} (2 \ 2) + \begin{pmatrix} 3 \\ 1 \end{pmatrix} (3 \ 1) + \begin{pmatrix} 3 \\ 2 \end{pmatrix} (3 \ 2) + \begin{pmatrix} 2 \\ 3 \end{pmatrix} (2 \ 3) \right] - \frac{1}{25} \begin{pmatrix} 11 \\ 10 \end{pmatrix} (11 \ 10) = \frac{1}{25} \begin{pmatrix} 14 & -5 \\ -5 & 10 \end{pmatrix}; C_2 = \frac{1}{25} \begin{pmatrix} 14 & -5 \\ -5 & 10 \end{pmatrix}.$$

Получаем, $C_1 = C_2 = C$.

Для случая равных матриц ковариаций для различных классов дискриминантная функция имеет вид:

$$d_k(x) = -\frac{1}{2} x^T C^{-1} x - \frac{1}{2} m_k^T C^{-1} m_k + x^T C^{-1} m_k + \log p(\omega_k) - \frac{1}{2} \log |C|,$$

для двухклассовой задачи:

$$d(x) = x^T C^{-1} (m_1 - m_2) + \log \frac{p(\omega_1)}{p(\omega_2)} - \frac{1}{2} [(m_1 - m_2)^T C^{-1} (m_1 + m_2)].$$

Детерминант и C' (приведенная матрица) имеют вид:

$$|C| = \frac{1}{25} \begin{vmatrix} 14 & -5 \\ -5 & 10 \end{vmatrix} = \frac{23}{5}; C' = \begin{pmatrix} \frac{10}{25} & \frac{5}{25} \\ \frac{5}{25} & \frac{14}{25} \end{pmatrix}; C^{-1} = \frac{1}{|C|} C' = \frac{1}{115} \begin{pmatrix} 10 & 5 \\ 5 & 14 \end{pmatrix};$$

$$C^{-1} m_1 = \frac{1}{115} \begin{pmatrix} 32 \\ 39 \end{pmatrix}; m_1^T C^{-1} m_1 = \frac{742}{5 \cdot 115};$$

Дискриминантная функция для класса 1:

$$d_1(x) = x^T C^{-1} m_1 - \frac{1}{2} m_1^T C^{-1} m_1 = \frac{32}{115} x_1 + \frac{39}{115} x_2 - 0.65.$$

$$C^{-1} m_2 = \frac{1}{115} \begin{pmatrix} 127 \\ 167 \end{pmatrix}; m_2^T C^{-1} m_2 = 22.$$

Дискриминантная функция для класса 2:

$$d_2(x) = x^T C^{-1} m_2 - \frac{1}{2} m_2^T C^{-1} m_2 = \frac{127}{115} x_1 + \frac{167}{115} x_2 - 11.$$

Решающая поверхность: $d(x) = -0.826x_1 - 1.1|x_2| - 10.35 = 0$.

Задача 2.7. [14] Даны матрицы Q - описание 6-ти объектов; R - соответствие номеров объектов и классов.

$$Q = \begin{matrix} & z_1 & z_2 & z_3 & z_4 \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{matrix} & \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 2 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 2 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 2 \end{bmatrix} \end{matrix}; R = \begin{matrix} & \text{Тупиковые тесты:} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{matrix} & \begin{bmatrix} 1 \\ 1 \\ 1 \\ 2 \\ 2 \\ 2 \end{bmatrix} \end{matrix}.$$

Тупиковые тесты:

$$\Theta_1 = \{z_1, z_2, z_3\},$$

$$\Theta_2 = \{z_1, z_2, z_4\}$$

$$\Theta_3 = \{z_2, z_3, z_4\}.$$

Найдены по одному из методов тупиковые тесты: $\Theta_1, \Theta_2, \Theta_3$. Сформируйте решающее правило, по которому можно отнести объект $S=(0, 1, 2, 2)$ (не входящий в обучающую выборку) к одному из выделенных классов.

Решение. Набор тестов порождает описание обучающих объектов:

$$Q_1(\Theta_1) = \begin{matrix} & z_1 & z_2 & z_3 & z_4 \\ 1 & \begin{bmatrix} 0 & 1 & 1 & - \end{bmatrix} \\ 2 & \begin{bmatrix} 1 & 2 & 0 & - \end{bmatrix} \\ 3 & \begin{bmatrix} 0 & 1 & 0 & - \end{bmatrix} \\ 4 & \begin{bmatrix} 1 & 2 & 1 & - \end{bmatrix} \\ 5 & \begin{bmatrix} 1 & 1 & 0 & - \end{bmatrix} \\ 6 & \begin{bmatrix} 1 & 1 & 1 & - \end{bmatrix} \end{matrix}; \quad Q_2(\Theta_2) = \begin{matrix} & z_1 & z_2 & z_3 & z_4 \\ 1 & \begin{bmatrix} 0 & 1 & - & 0 \end{bmatrix} \\ 2 & \begin{bmatrix} 1 & 2 & - & 1 \end{bmatrix} \\ 3 & \begin{bmatrix} 0 & 1 & - & 1 \end{bmatrix} \\ 4 & \begin{bmatrix} 1 & 2 & - & 0 \end{bmatrix} \\ 5 & \begin{bmatrix} 1 & 1 & - & 1 \end{bmatrix} \\ 6 & \begin{bmatrix} 1 & 1 & - & 2 \end{bmatrix} \end{matrix}; \quad Q_3(\Theta_3) = \begin{matrix} & z_1 & z_2 & z_3 & z_4 \\ 1 & \begin{bmatrix} - & 1 & 1 & 0 \end{bmatrix} \\ 2 & \begin{bmatrix} - & 2 & 0 & 1 \end{bmatrix} \\ 3 & \begin{bmatrix} - & 1 & 0 & 1 \end{bmatrix} \\ 4 & \begin{bmatrix} - & 2 & 1 & 0 \end{bmatrix} \\ 5 & \begin{bmatrix} - & 1 & 0 & 1 \end{bmatrix} \\ 6 & \begin{bmatrix} - & 1 & 1 & 2 \end{bmatrix} \end{matrix}$$

Реакция описания $S=(0, 1, 2, 2)$ на тесты $\Theta_1, \Theta_2, \Theta_3$:

$$\Theta_1 = \{0, 1, 2, -\},$$

$$\Theta_2 = \{0, 1, -, 2\},$$

$$\Theta_3 = \{-, 1, 2, 2\}.$$

Новый объект $S=(0, 1, 2, 1)$ тестовый алгоритм не отнесет S ни к одному из классов; однако по набору $\Theta=\{z_1, z_2\}$ (*части теста*) по фрагменту $(0, 1)$ (содержится в S и объектах из 1-го класса и не содержится в объектах из 2-го класса) возможно отнесение объекта S к первому классу.

Ход занятия №4.

Вступление. Сообщение темы и обоснование ее актуальности через вышеуказанные задачи. Ведущий студент, ответственный за выбор и подачу необходимой информации, согласует алгоритм занятия.

Основная часть:

I. Сообщение в виде доклада-презентации ответственными двумя студентами за проведение занятия, в котором излагается суть обсуждаемых положений:

- 1) Статистические методы распознавания образов.
- 2) «Пригодность» методов распознавания образов для решения задачи прогнозирования.
- 3) Методики оценки эффективности статистических методов распознавания образов.

Задача И2.4. Взять случайным образом 100 точек из 1-го квадрата плоскости $ХОУ$ и классифицировать их методом «ФОРЕЛЬ».

Задача И2.5. Провести разделение классов K_1 и K_2 с использованием статистических характеристик:

K_1 : A1(3,1); A2(4,2); A3(5,2); A4(2,1); A5(4,1); A6(3,2);

K_2 : B1(3,4); B2(4,4); B3(5,4); B4(3,5); B5(5,3); B6(5,5).

Задача И2.6. Дана база данных, содержащая сведения о состоянии почв Томской области (химический, минералогический) во времени. Выявить классы почв и построить решающее правило (рекомендовать метод его построения), позволяющее по текущему состоянию почвы прогнозировать характер изменения экологического развития анализируемого района.

II. Выяснение позиций участников с зафиксированными точками зрения на решение вышеизложенных задач.

Итог II-го этапа: формирование целевых групп по общности позиций каждой из групп.

- III. Организация коммуникации между группами: 1) выяснение позиции-варианта решения выявленных групп и защита занятой позиции; 2) формирование нового набора вариантов решений на основании общего обсуждения; 3) выбор одного решения голосованием;
- IV. Повторная защита позиций-вариантов групп после проведения расчетов с целью оценки отклонения от «истинного» решения (парное оценивание).

Выводы: реализован самостоятельный поиск учащимися путей и вариантов решения поставленной учебной задачи (выбор одного из предложенных вариантов или нахождение собственного варианта и обоснование решения на базе коллективной интерактивной работы).

Итог занятия №4: Оценивание компетенций (табл.2) по результатам работы на занятиях (активность, инициативность, грамотность, обоснованность защищаемой позиции) и своевременности сдачи отчета по решению практических задач И2.4-И2.6.

Таблица 2

№	№ задачи	Вид (совмещение нескольких видов) интерактивной работы	Трудоемкость (час)	Отрабатываемые компетенции/ожидаемый уровень освоения	Оценка личностных качеств	Контроль выполнения работы (участие в полемике, индивидуальные групповые задания (ИГЗ) и т.д)
1	И2.4-И2.5	Работа в команде. Решение ситуационных задач.	2	ОК-1, ОК-2/ З-Эл, У-Эл, В-Эл ПК-5/ З-Пр, У-Пр, В-Пр	Качество работы; своевременность сдачи отчета по решению ИГЗ	ИГЗ. Критерии оценивания поведения на занятии: активность, инициативность, грамотность, обоснованность защищаемой позиции.
2	И2.6	Работа в команде. Решение ситуационных задач. Исследовательский метод	2			
Всего			4			

ВАРИАНТЫ ДОМАШНИХ ЗАДАНИЙ К РАЗДЕЛУ 2

Задача Д2.1. Даны матрицы Q - описание 6-ти объектов; R - соответствие номеров объектов и классов.

$$Q = \begin{matrix} & z_1 & z_2 & z_3 & z_4 \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{matrix} & \begin{bmatrix} 2 & 1 & 1 & 2 \\ 1 & 3 & 2 & 1 \\ 2 & 1 & 2 & 1 \\ 1 & 3 & 1 & 2 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 3 \end{bmatrix} \end{matrix}; R = \begin{matrix} \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{matrix} & \begin{bmatrix} 1 \\ 1 \\ 1 \\ 2 \\ 2 \\ 2 \end{bmatrix} \end{matrix}.$$

Сформировать решающее правило, по которому можно отнести объект $S=(2, 1, 2, 3)$ (не входящий в обучающую выборку) к одному из выделенных классов.

Задача Д2.2. Дано два набора точек:

$$(1,3), (3,3), (4,1), (4,3), (3,4) \in K_1, (8,6), (9,6), (6,9), (6,6), (9,10) \in K_2..$$

Найти разделяющую границу - решающую поверхность.

Задача Д2.3. Составить алгоритм метода « k -средних».

ВАРИАНТЫ КОНТРОЛЬНЫХ ЗАДАНИЙ К РАЗДЕЛУ 2

Найти дискриминантные функции для классификации двух образов по представительным обучающим объектам:

I вариант

$$X_1 = \begin{Bmatrix} 2 \\ 5 \end{Bmatrix}; X_2 = \begin{Bmatrix} 6 \\ 2 \end{Bmatrix}$$

II вариант

$$X_1 = \begin{Bmatrix} 3 \\ 4 \end{Bmatrix}; X_2 = \begin{Bmatrix} 7 \\ 3 \end{Bmatrix}$$

III вариант

$$X_1 = \begin{Bmatrix} 4 \\ 7 \end{Bmatrix}; X_2 = \begin{Bmatrix} 8 \\ 3 \end{Bmatrix}$$

Контрольные вопросы к разделу 2

1. Охарактеризуйте методы построения эталонов.
2. В чем заключается метод дробящихся эталонов?
3. Метод ближайших соседей и влияние метрики на качество распознавания.
4. Метод потенциальных функций и его модификации.
5. Метод дискриминантных функций: в чем его «слабое» место?

Раздел 3. Алгебраический подход к задаче распознавания

ПРИМЕРЫ ТИПОВЫХ АУДИТОРНЫХ ЗАДАНИЙ

ПРАКТИЧЕСКОЕ ЗАНЯТИЕ 7. Алгебраические методы в задачах распознавания и классификации.

Цель занятия: воспроизведение полученных знаний на лекции по разделу 3 «Алгебраический подход к задаче распознавания», применение основных понятий темы раздела 3 для решения задач распознавания.

Задача 3.1. Исследовать ряд методов вычисления информативности признаков и показать, что они дают существенно отличающиеся решения.

Решение. Рассмотрим ряд известных в научной литературе подходов (мер сравнения) к поиску информативных признаков.

А) В случае известного закона распределения значений и вида признаков применимы показатели информативности: по коэффициенту корреляции: Пирсона (вычисление силы линейной связи между количественными признаками),

- Фехнера, Кендалла, Спирмена (определение монотонных зависимостей) и др. (Гайдышев И.).
- Б) Информативность по Шеннону как мера трудности распознавания (Вагин В.Н.).
- В) Метрические методы. Так, мера Хемминга применяется для подсчета числа попарно одинаковых битов векторов.
- Г) Метод на основе мультимножественного представления данных с последующим использованием метрик на мультимножествах и метода парных сравнений (Колесникова С.И., Янковская А.Е.).
- Д) Построение решающего правила в виде дерева дихотомических делений выборки по отдельным признакам (Загоруйко Н.Г.).
- Е) Другие методы....

Пример. [25] Результаты численного эксперимента по формированию списка 10 наиболее важных признаков с точки зрения разных мер информативности приведены в

Таблица 4.1. На одном и том же наборе данных был получен существенно отличающийся порядок признаков с точки зрения информативности

Таблица 4.1.

№ признака Мера информ.	Сортировка признаков в порядке важности									
	1	2	3	4	5	6	7	8	9	10
Пирсон	82	58	74	10	50	6	66	34	65	14
Фехнер	2	6	10	14	18	26	30	34	38	41
Спирмен	2	6	10	14	18	22	26	30	34	38
Кендалл	98	34	37	81	17	90	78	89	41	82
Шеннон	1	2	18	66	10	5	22	35	3	15
Хемминг	1	2	3	4	5	6	7	8	9	10
Дихотом.	1	2	3	4	5	6	7	8	9	10

Задача 3.2. *Пример композиции алгоритмов.* Дайте определение композиции алгоритмов по Ю.И. Журавлёву и покажите, что возможно построить коллективное правило, по которому классификация (распознавание образов) дает существенно меньшую ошибку, чем при использовании отдельных алгоритмов.

Решение. Пусть X – пространство объектов, Y – пространство ответов алгоритмов. По Ю.И. Журавлёву композицией $C(A(x))$ базовых алгоритмов $A_k, k = \overline{1, g}$, где $A_k: X \rightarrow R$, называется алгоритм (*алгоритмический оператор*) $A: X \rightarrow Y$ вида: $A(x) = C(F(A_1(x), \dots, A_g(x)))$, $x \in X$, где $F: R^g \rightarrow R$ - корректирующая операция и $C: R \rightarrow Y$ - решающее правило; им же впервые были введены линейные корректирующие операции:

$$\left\{ F(A_1 \dots A_g) = \sum_{x=1}^g \omega_x A_x \mid \omega_x \in R, x = \overline{1, g} \right\}.$$

Пример композиции алгоритмов для решения задачи прогнозирования по временным рядам можно посмотреть в [26]. В СРС в разделе с аналогичным названием будет изложен пример композиций на базе метода анализа иерархий.

Замечание. Многие алгоритмы классификации имеют именно такую структуру: сначала вычисляются оценки принадлежности объекта классам, затем решающее правило переводит эти оценки в номер класса. Значение оценки, как правило, характеризует степень уверенности классификации.

Интерактивные занятия-семинары №7, 8 по теме: Алгебраические методы в задачах распознавания и классификации. Эффективность систем распознавания с коллективным распознаванием.

Цель занятия: продолжение занятия 7: применение основных понятий темы раздела 1 для решения задач: построение алгебраических композиций с целью повышения надежности принятия решений в практических задачах.

Планируемые к приглашению на семинар специалисты-эксперты: Буймов А.Г., д.т.н., профессор ТУСУР; Цой Ю.Р., к.т.н., доцент ТПУ, специалист по методам интеллектуального анализа данных (ИАД);

ПРИМЕРЫ ТИПОВЫХ АУДИТОРНЫХ ЗАДАНИЙ

Задача 3.1. Приведите примеры переобучения. В чем заключается «негатив» этого явления?

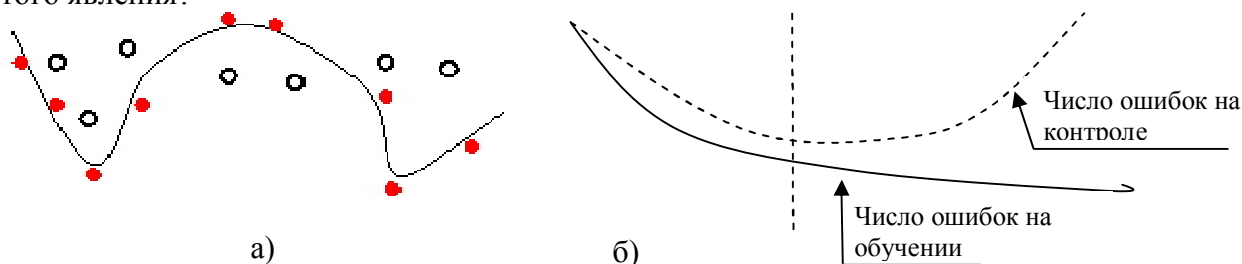


Рис. 3.1. а) Обучающее (закрашенные точки) множество точек и множество точек из контрольного множества (незакрашенные точки); б) «нестыковка» в моментах прекращения обучения на обучающем множестве и контрольном

Ход занятия №И7, 8.

Вступление. Сообщение темы и обоснование ее актуальности через вышеуказанные задачи. Ведущий студент, ответственный за выбор и подачу необходимой информации, согласует алгоритм занятия.

Основная часть:

- I. Сообщение в виде доклада-презентации ответственными двумя студентами за проведение занятия, в котором излагается суть обсуждаемых положений:
 - 1) Алгебраические методы в задачах распознавания и классификации.
 - 2) История коллективных решений.
 - 3) Комитеты решающих правил.

Задача ИЗ.1. В чем заключается технология бустинга (Р. Шапир). Используйте данную технологию (Boost1) для решения задачи обработки изображений.

Решение. Бустинг (boosting – усиление, улучшение) – процедура последовательного построения композиции алгоритмов машинного обучения, в основе которой лежит построение цепочки (каскада) классификаторов-алгоритмов, каждый из которых (за исключением первого) обучается на ошибках предыдущего. Boost1 (исторически 1-й алгоритм бустинга) использует каскад из 3-х моделей, первая из которых обучается на всем наборе данных X , вторая – на выборке объектов (обучающих примеров), в половине из которых первая дала правильные ответы, а третья – на объектах, где «ответы» первых двух разошлись. В такой последовательной обработке объектов каскадом классификаторов задача для каждого последующего классификатора-алгоритма

становится труднее. Результат определяется путем голосования: объект относится к тому классу, который выдан большинством моделей каскада.

Рассмотрим задачу классификации на два класса, $Y = \{-1, +1\}$. Пусть базовые алгоритмы возвращают только два ответа -1 и $+1$, и решающее правило фиксировано: $C(b) = \text{sign}(b)$. Искомая алгоритмическая композиция имеет вид:

$$a(x) = C(F(b_1(x), \dots, b_g(x))) = \text{sign}(\sum_{t=1}^T \omega_t b_t(x)), x \in X.$$

Функционал качества композиции Q_T определяется как число ошибок, допускаемых композицией на обучающей выборке:

$$Q_T(\mathbf{b}, W^m) = \sum_{t=1}^T [y_t \sum_{i=1}^T \omega_i b_i(x)],$$

где $W^m = (w_1, \dots, w_m)$ – вектор весов объектов. На практике используют экспоненциальную аппроксимацию пороговой функции потерь $[z < 0] e^{-z}$

Основные этапы алгоритма бустинга для распознавания изображений [28]:

Дано изображение с потенциально распознаваемыми объектами, представленное двумерной матрицей пикселей размером $w \times h$, в которой каждый пиксель имеет значение: $0 \div 255$ для черно-белое изображений; $0 \div 255 \times 3$, для цветных (компоненты R, G, B).

1. Формируется «прямоугольный признак» $\text{rectangle}(i) = \{x, y, w, h, a\}$, где x, y – координаты центра i -го прямоугольника, w – ширина, h – высота, a – угол наклона прямоугольника к вертикальной оси изображения.

2. Сканируется изображение окном поиска с одновременным формированием интегрального представления изображения – матрицы L , в каждом элементе $L(x, y)$ которой хранится суммарная яркость каждого прямоугольника на данном изображении (сумма интенсивностей всех пикселей в прямоугольнике от $(0, 0)$ до (x, y) , то есть находящихся левее и выше данного элемента согласно методу Виолы-Джонса):

$$L(x, y) = I(x, y) - L(x-1, y-1) + L(x, y-1) + L(x-1, y),$$

где $I(x, y)$ — яркость пикселя исходного изображения.

3. Применяется классификатор к каждому положению окна сканирования.

При этом на обучении осуществляется итеративный процесс:

- 1). Определяются слабые классификаторы по прямоугольным признакам на каждом примере с выбором «подходящего порога» для каждого признака;
- 2). Отбираются лучшие признаки и лучший подходящий порог;
- 3). Пересчитываются веса объектов выборки.

Замечание 3.1. В стандартном методе Виолы – Джонса используются прямоугольные признаки - примитивы Хаара.

Замечание 3.2. Использовать для работы базу данных ORL Database of Faces, содержащую набор лиц, сфотографированных в исследовательской лаборатории Кембриджа (1992-1994гг).

Задача ИЗ.2. Постройте коллективное правило для прогнозирования по временным рядам изменения минералогических и химических свойств торфов по предоставленной базе данных.

II. Выяснение позиций участников с зафиксированными точками зрения на решение вышеизложенных задач.

Итог II-го этапа: формирование целевых групп по общности позиций каждой из групп.

III. Организация коммуникации между группами: 1) выяснение позиции-варианта решения выявленных групп и защита занятой позиции; 2) формирование нового набора вариантов решений на основании общего обсуждения; 3) выбор одного решения голосованием;

IV. Повторная защита позиций-вариантов групп после проведения расчетов с целью оценки отклонения от «истинного» решения (попарное оценивание).

Выводы: реализован самостоятельный поиск учащимися путей и вариантов решения поставленной учебной задачи (выбор одного из предложенных вариантов или нахождение собственного варианта и обоснование решения на базе коллективной интерактивной работы).

Итог занятия №И7,8: Оценивание компетенций (табл.3) по результатам работы на занятиях (активность, инициативность, грамотность, обоснованность защищаемой позиции) и своевременности сдачи отчета по решению практической задачи ИЗ.1.

Таблица 2

№	№ задач и	Вид (совмещение нескольких видов) интерактивной работы	Трудоемкость (час)	Отрабатываемые компетенции/ожидаемый уровень освоения	Оценка личностных качеств	Контроль выполнения работы (участие в полемике, индивидуальные групповые задания (ИГЗ) и т.д)
1	ИЗ.1, ИЗ.2	Работа в команде. Решение ситуационных задач.	3	ОК-1, ОК-2/ З-Эл, У-Эл, В-Эл ПК-5/ З-Пр, У-Пр, В-Пр	Качество работы; своевременность сдачи отчета по решению ИГЗ	ИГЗ. Критерии оценивания поведения на занятии: активность, инициативность, грамотность, обоснованность защищаемой позиции.
Всего			3			

ВАРИАНТЫ ДОМАШНИХ ЗАДАНИЙ К РАЗДЕЛУ 3

Задача ДЗ.1. Подготовить сообщение по одной из тем (на выбор) по книгам:

1) Воронцов К.В. Обзор современных исследований по проблеме качества обучения алгоритмов. Таврический вестник информатики и математики. – 2004. – № 1. – С. 5 – 24. <http://www.ccas.ru/frc/papers/voron04twim.pdf>.

2) Воронцов К.В. Лекции по алгоритмическим композициям <http://www.machinelearning.ru/wiki/images/0/0d/Voron-ML-Compositions.pdf>.

Задача ДЗ.2. Подготовить ответ на вопрос: в чем принципиальная разница в подходах: гибридное правило распознавания, коллективное распознавание, комитет решающих правил, метод группового учета аргументов. Выстроить историческую преемственность. Привести примеры.

Задача ДЗ.3. Реализовать программно алгоритм построения алгоритмической композиции по временным рядам – реализациям процесса работы асинхронного двигателя согласно работе [26]: Воронцов К.В., Егорова Е.В. Динамически адаптируемые композиции алгоритмов прогнозирования // Искусственный Интеллект. – № 10. - 2006. – С. 277–280.

ВАРИАНТЫ КОНТРОЛЬНЫХ ЗАДАНИЙ К РАЗДЕЛУ 3

I вариант

Метод бустинга: условия применения, ограничения. Привести пример программы, его реализующий на выборке из 50-ти объектов (примеры выборок - student\Колесникова\ TPO).

II вариант

Линейные алгоритмические композиции. Привести пример программы, его реализующий на выборке из 50-ти объектов (примеры выборок - student\Колесникова\ TPO).

III вариант

Нелинейные алгоритмические композиции. Привести пример программы, его реализующий на выборке из 50-ти объектов (примеры выборок - student\Колесникова\ TPO).

Контрольные вопросы к разделу 3

1. Алгебраический подход к задаче распознавания.
2. Коллективы алгоритмов: методики построения.
3. Комитетные решающие правила.
4. Суть гибридного распознавания.

Раздел 4. Распознавание образов и распознавание изображений

Интерактивные занятия-семинары №9, 10 по теме: Распознавание образов и распознавание изображений. Системы РО на основе нейросети.

Цель занятия: Знакомство с методами распознавания изображений. Знакомство с существующим доступным программным обеспечением для распознавания изображений и обсуждение границ их применимости.

Планируемые к приглашению на семинар специалисты-эксперты: Спицын В.Г., д.т.н., профессор ТПУ, специалист по методам математической логики; Цой Ю.Р., к.т.н., доцент ТПУ, специалист по методам интеллектуального анализа данных (ИАД);

Дополнительная информация.

1. <http://www.delphikingdom.com/asp/viewitem.asp?catalogid=1299>
2. <http://www.delphikingdom.com/asp/viewitem.asp?catalogid=1203>
3. http://www.forekc.ru/Ns/index_6.htm

ПРИМЕРЫ ТИПОВЫХ АУДИТОРНЫХ ЗАДАНИЙ

Задача №4.1. Рассмотреть базовый алгоритм SOM (самоорганизующиеся карты Кохонена), являющийся составной частью нейронных систем обработки информации.

Решение. Самоорганизующаяся карта (SOM) - это процесс обучения без учителя и классификации ряда образов без какой-либо информации о классах. Элемент проецируется из входного множества R^n на позицию в карте – информация кодируется как позиция активированного узла, обеспечивая топологический «заказ» классов - набор узлов в пространстве, имеющем меньше измерений.

Основные положения алгоритма SOM.

1. Элемент проецируется из входного множества R^n на позицию в карте.
2. Каждый из узлов описывается двумя векторами, первый — вектор веса $\mathbf{m}_i \in R^n$, имеющий такую же размерность, что и входные данные (вектор-ссылка привязан к каждому узлу в SOM). Второй — координаты узла на карте.
3. Во время обучения каждый входной вектор \mathbf{x} сравнивается со всеми \mathbf{m}_i в поиске размещения наиболее сходного \mathbf{m}_c : $|\mathbf{x} - \mathbf{m}_c| = \min_i \{|\mathbf{x} - \mathbf{m}_i|\}$.
4. Векторы весов узлов в SOM пересчитываются согласно формуле:

$$\mathbf{m}_i(t+1) = \mathbf{m}_i(t) + h_{ci}(t)[\mathbf{x}(t) - \mathbf{m}_i(t)],$$

где t - время, в течении которого уже происходит обучение, $h_{ci}(t)$ – сглаживающая функция, максимум которой достигается на \mathbf{m}_c . Обычно полагают $h_{ci}(t) = h(|\mathbf{r}_c - \mathbf{r}_i|, t)$, где \mathbf{r}_c и \mathbf{r}_i - положения узлов в выходном пространстве SOM, с наиболее близким весовым вектором к входному шаблону и текущего при «пробегании» по всем узлам, соответственно, например:

$$h_{ci} = \alpha(t) \exp\left(-\frac{|\mathbf{r}_c - \mathbf{r}_i|^2}{2\sigma^2(t)}\right)$$

где $\alpha(t)$ - это коэффициент обучаемости, а $\sigma(t)$ определяет длину ряда. Они обычно оба монотонно убывают с течением времени. Использование аппроксимирующей функции означает, что узлы, которые располагаются в SOM структуре в соответствии с положением «победившего» узла. Это создаёт сглаживающий эффект, который приводит к глобальной организации карты. SOM может быть представлена, как нелинейная проекция плотности вероятностей.

Задача №4.2. Охарактеризуйте критерии оценки потерь качества изображения. Назовите источники потерь качества изображения.

Решение. Источники потерь качества изображения: процессы оцифровки, перевода в ограниченную палитру цветов, перевода в другую систему цветопродставления для печати, архивации с потерями.

Критерии оценки потерь качества изображения [27]:

- 1) среднеквадратичное отклонение значений пикселей (L_2 мера):

$$d(x, y) = \left(\sum_{i,j=1}^n \frac{(x_i - y_i)^2}{n^2} \right)^{1/2}$$

Согласно этому критерию изображение будет сильно испорчено при понижении яркости всего на 5%.

- 2) мера *Чебышева* $d(x, y) = \max_{i,j} |x_{ij} - y_{ij}|$ крайне чувствительна к биению

отдельных пикселей (во всем изображении может существенно измениться только значение одного пикселя, но согласно этой мере изображение будет сильно испорчено).

- 3) мера отношения сигнала к шуму, похожая на среднеквадратичное отклонение, но удобнее за счет логарифмического масштаба шкалы:

$$d(x, y) = 10 \log_{10} \frac{(255n)^2}{\sum_{i,j=1}^n (x_i - y_i)^2}$$

Задача №4.3. Составьте программу – простую однослойную нейросеть для распознавания двух объектов, заданных векторами-признаками.

Ход занятия №И9, 10.

Вступление. Сообщение темы и обоснование ее актуальности через вышеуказанные задачи. Ведущий студент, ответственный за выбор и подачу необходимой информации, согласует алгоритм занятия.

Основная часть:

I. Сообщение в виде доклада-презентации ответственными двумя студентами за проведение занятия, в котором излагается суть обсуждаемых положений:

- 1) Перцептрон Розенблатта и принцип работы.
- 2) Теорема Новикова А.Б. о сходимости перцептрона.
- 3) Основные принципы нейросетевого распознавания образов.
- 4) Пример алгоритма по распознаванию образов (изображения) нейросетью.

Задача № И4.1. Составить (выбрать) алгоритм для практической реализации распознавания устной речи. Реализовать соответствующую программу. Сравнить существующие программы по распознаванию речи.

Задача № И4.2. Составить (выбрать) алгоритм для практической реализации распознавания изображения. Реализовать соответствующую программу. Сравнить существующие программы по распознаванию изображений.

II. Выяснение позиций участников с зафиксированными точками зрения на решение вышеизложенных задач.

Итог II-го этапа: формирование целевых групп по общности позиций каждой из групп.

III. Организация коммуникации между группами: 1) выяснение позиции-варианта решения выявленных групп и защита занятой позиции; 2) формирование нового набора вариантов решений на основании общего обсуждения; 3) выбор одного решения голосованием;

IV. Повторная защита позиций-вариантов групп после проведения расчетов с целью оценки отклонения от «истинного» решения (попарное оценивание).

Выводы: реализован самостоятельный поиск учащимися путей и вариантов решения поставленной учебной задачи (выбор одного из предложенных вариантов или нахождение собственного варианта и обоснование решения на базе коллективной интерактивной работы).

Итог занятия №И9,10: Оценивание компетенций (табл.4) по результатам работы на занятиях (активность, инициативность, грамотность, обоснованность защищаемой позиции) и своевременности сдачи отчета по решению практических задач типа И4.1.

Таблица 4

№	№ задач и	Вид (совмещение нескольких видов) интерактивной работы	Трудоемкость (час)	Отрабатываемые компетенции/ожидаемый уровень освоения	Оценка личностных качеств	Контроль выполнения работы (участие в полемике, индивидуальные групповые задания (ИГЗ) и т.д)
1	И4.1, И4.2	Работа в команде. Решение ситуационных	4	ОК-1, ОК-2/ З-Эл, У-Эл, В-Эл	Качество работы; своевременность сдачи отчета	ИГЗ. Критерии оценивания поведения на занятии: активность, инициативность,

	задач.		ПК-5/ 3-Пр, У- Пр, В-Пр	по решению ИГЗ	грамотность, обоснованность защищаемой позиции.
Всего		4			

ВАРИАНТЫ ДОМАШНИХ ЗАДАНИЙ К РАЗДЕЛУ 4

Задача Д4.1. Сопоставьте архитектуры биологической нейронной системы и машины фон Неймана.

Задача Д4.2. Преобразование Карунена-Лоэва для реализации самоорганизующейся карты и условия (ограничения) его применимости.

Задача Д4.3. Предложите алгоритм классификации сигнала электрокардиограммы на базе нейросети.

ВАРИАНТЫ КОНТРОЛЬНЫХ ЗАДАНИЙ К РАЗДЕЛУ 4

I вариант. Составить программу, предназначенную для распознавания рукописного ввода двух китайских иероглифов.

II вариант

Составить программу, предназначенную для распознавания рукописного ввода двух цифр.

III вариант

Составить программу, предназначенную для распознавания рукописного ввода двух букв.

Контрольные вопросы к разделу 4

1. Дайте основные положения структурных (лингвистических) методов.
2. В чем заключается особенность метода потенциальных функций, позволяющая его использование в распознавании на базе нейросетей.
3. Основные возможности языка описания изображений PDL.

При составлении методических указаний использовался материал нижеуказанной литературы, а также материал интернет-ресурсов.

ИСПОЛЬЗОВАННАЯ ЛИТЕРАТУРА

1. Горелик А. Л., Скрипкин В. А. Методы распознавания: Учебное пособие для вузов. - 4-е изд., испр. - М.: Высшая школа, 2004. – 260 с.
2. Лапко А.В. Непараметрические системы обработки информации: Учебное пособие для вузов / А. В. Лапко, С. В. Ченцов; Российская Академия наук. Сибирское отделение, Институт вычислительного моделирования. - М. : Наука, 2000. - 349 с.
3. Воронцов К.В. Лекции по методам оценивания и выбора моделей. 2007. Режим доступа: www.ccas.ru/voron/download/Modeling.pdf.
4. Р. Гонсалес. Цифровая обработка изображений в среде MATLAB: Пер. с англ. / Р. Гонсалес, Р. Вудс, С. Эддинс ; пер. : В. В. Чепыжов. - М. : Техносфера, 2006. – 615 с.
5. Ту Д., Гонсалес Р. Принципы распознавания образов. – М.: Мир, 1978, 2008.
6. Вапник В.Н. и др. Алгоритмы и программы восстановления зависимостей: Практическое руководство. - М. : Наука. Физматлит, 1984. - 816 с.
7. Вапник В.Н., Червоненкис А.Я. Теория распознавания образов. М.: Наука, 1974, 2002.- 415 с.
8. Загоруйко Н.Г. Прикладные методы анализа данных и знаний. // Новосибирск. Изд-во института математики. 1999, 2008.
9. Дадашев Т.М. Теория распознавания образов (логические методы): Учебное пособие. - М.: МФТИ, 1982, 2006. - 84 с.
10. Айзерман А.А., Браверман Э.М., Розоноэр Э.И. Метод потенциальных функций в теории обучения машин. – М.: Наука, 1970.
11. Патрик Э. Основы теории распознавания образов. – М.: Сов. радио, 1980.
12. Фу К.С. Структурные методы в распознавании образов. – М.: Мир, 1977.
13. Дуда Р., Харт П. Распознавание образов и анализ сцен. - М.: Мир, 1976.- 511 с.
14. Дюкова Е.В., Песков Н.В. Построение распознающих процедур на базе элементарных классификаторов // [www.ccas.ru /frc/papers /djukova05 construction.pdf](http://www.ccas.ru/frc/papers/djukova05construction.pdf).
15. Воронцов К.В. Обзор современных исследований по проблеме качества обучения алгоритмов. Таврический вестник информатики и математики. – 2004. – № 1. – С. 5 – 24. <http://www.ccas.ru/frc/papers/voron04twim.pdf>.
16. Геппенер В.В. Лекционный курс «Распознавание изображений и речевых сигналов». <http://www.studfiles.ru/dir/cat32/subj1011/file4179/view34452.html>
17. [www.ccas.ru/voron/www.ccas.ru /frc/papers /djukova05 construction.pdf](http://www.ccas.ru/voron/www.ccas.ru/frc/papers/djukova05construction.pdf).
18. <http://www.all-library.com/obrazovanie/nauka/42843-osnovy-teorii-raspoznavaniya-obrazov.html>
19. <http://window.edu.ru/resource/738/20738>
20. <http://www.bsu.by/Cache/pdf/229903.pdf>
21. <http://www.delphikingdom.com/asp/viewitem.asp?catalogid=1299>
22. <http://www.delphikingdom.com/asp/viewitem.asp?catalogid=1203>
23. http://www.forekc.ru/Ns/index_6.htm
24. <http://www.uran.donetsk.ua/~masters/2010/fknt/kostetskaya/library/art03/index.htm>
1
25. Корлякова М.О., Твердохлеб Н.С. Анализ подходов к определению информативности признаков. // Научная сессия МИФИ-2006. Сборник

- научных трудов. В 16 томах. Т.3. Интеллектуальные системы и технологии. М.: МИФИ, 2006. 256 с. С. 146-147
26. Воронцов К.В., Егорова Е.В. Динамически адаптируемые композиции алгоритмов прогнозирования // Искусственный Интеллект. – № 10. - 2006. – С. 277–280.
 27. <http://algotlist.manual.ru/compress/image/fractal/index.php>
 28. <http://habrahabr.ru/post/133826/>
 29. Волошин Г.Я. Методы распознавания образов (конспект лекций) <http://abc.vvsu.ru>
 30. Айзерман М.А., Браверман Э.М., Розоноэр Л.И. Метод потенциальных функций в теории обучения машин – изд. «Наука», Москва, 1970г.
 31. <http://sumschool.sumdu.edu.ua>