МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Юго-Западный государственный университет»
(ЮЗГУ)

Кафедра биомедицинской инженерии

УТВЕРЖДАЮ
Проректор по учебной работе
О.Г. Локтионова
«5» мая 2017 г.

ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ КЛАССИФИКАЦИИ И РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЙ

Методические рекомендации по выполнению лабораторных работ для студентов направления 12.04.04 - Биотехнические системы и технологии

Курск 2017

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Локтионова Оксана Геннадьевна
Должность: проректор по учебной работе
Дата подписания: 31.12.2020 13:36:44
Уникальный программный ключ:
0b817ca911e6668abb13a5d426d39e5f1c11eabbf73e943df4a4851fda56d089

УДК 504.064.38
Составитель Бурмака А.А., Говорухина Т.Н.

Рецензент
Цыплаков Ю.В., начальник отдела НИИЦ(г.Курск)ФГУП«18 ЦНИИ» МО РФ

Интеллектуальные системы классификации и распознавания изображений: методические рекомендации по выполнению лабораторных работ / Юго-Зап. гос. ун-т; сост.: А.А. Бурмака, Т.Н., Говорухина, Курск, 2017. 42 с. с ил.

Содержат методические рекомендации к проведению лабораторных работ по дисциплине «Интеллектуальные системы классификации и распознавания изображений».

Методические указания по структуре, содержанию и стилю изложения материала соответствуют методическим и научным требования, предъявляемым к учебным и методическим пособиям.

Предназначены для студентов направления подготовки 12.04.04 - Биотехнические системы и технологии.

Текст печатается в авторской редакции

Подписано в печать 5.05.17 Формат 60x84 1/16
Усл.печ.л.2,44. Уч.-изд.л.2,21.Тираж 30 экз. Заказ:855.Бесплатно.
Юго-Западный государственный университет.
305040, г. Курск, ул. 50 лет Октября, 94.

Оглавление

Лабораторная работа №1 4
Пример оформления лабораторной работы 11
Лабораторная работа №2 28
Лабораторная работа №3 31
Приложение 34

Лабораторная работа №1

СОСТАВЛЕНИЕ ОБУЧАЮЩИХ ТАБЛИЦ ЭКСПЕРЕМЕНТАЛЬНЫХ ДАННЫХ И ИХ ПРЕДВАРИТЕЛЬНАЯ ОБРАБОТКА

1. Краткие теоретические сведения.

Независимо от вида и свойств анализируемого объекта, процесса или явления, приемом, предшествующим процессу распознавания объекта (процесса, явления, далее – объекта), является получение текущей информации о свойствах объекта. Эта информация извлекается в виде аналоговых, дискретных или цифровых сигналов, формируемых различными датчиками, прямо или косвенно соприкасающимися с объектом и сообщающими в виде параметров сигналов различным измерительным устройствам, которые, как правило, преобразуют входную информацию в цифровые данные, которые в дальнейшем системе распознавания представляются в виде графиков, таблиц или гистограмм.

Сформированные таким образом данные о свойствах (параметрах, характеристиках) объекта являются исходным моментом получения и формирования данных, служащих в дальнейшем материалом для решения задач распознавания объектов, в том числе – элементов изображений (визуализированной информации). При этом стандартными приемами являются:

- Получение текущей измерительной информации посредством датчиков и измерительных устройств, наполнение массивов (выборок) цифровых данных (текущих статистик);
- Упорядочивание информации, содержащейся в выборке и определяющей одно из свойств (параметров, характеристик) объекта (процесса, явления), подвергающегося анализу с целью последующего его распознавания и (или) идентификации;
- Устранение артефактов аналитическим способом (построение доверительных интервалов, вариационных статистических рядов, гистограмм и т.д.);
- Устранение пропусков значений параметров для идентичных адресных выборочных номеров;
- Восстановление пропущенных значений параметров по методу наибольшего подобия;
- Восстановление выборочных массивов после устранения артефактов.

Восстановленные таким образом выборки могут быть в дальнейшем исходным материалом для формирования признакового пространства, используемого в алгоритмах распознавания.

Типичная форма при сборе экспериментальных данных – это таблица «объект – признак», в которую заносятся значения признаков (свойств), характеризующие каждый исследуемый объект. Примерами признаков могут быть наличие или отсутствие симптома, «вес», «температура», «давление», «частота сердечных сокращений" и. т.д. Под объектами может

рассматриваться любые проявления реального мира - люди, нозологические формы, животные, изделия и пр. Таблицу такого вида принято называть таблицей экспериментальных данных (ТЭД). Это название следует трактовать более широко, говоря не об экспериментальных данных, а о данных научного исследования. В качестве примера рассмотрим ТЭД, в которой приводятся сведения о пациентах:

№ пациента	Температура (°С) (x1)	Давление НД (мм.рт.ст) (x2)	Давление ВС (мм.рт.ст) (x3)	ЧСС (уд/мин) (x4)	№ группы
1.	36,6	60	115	65	1
2.	36,8	65	118	68	1
3.	36,5	70	117	72	1
4.	38,1	80	130	95	2
5.	39,4	70	140	105	2
6.	39,6	90	160	98	2
7-	36,8	95	144	72	3
8.	36,2	115	176	68	3
9	36,5	95	159	69	3

В данной таблице в качестве примера приводятся данные по 9 обследованным больным, относящимся к 3-м различным группам. Например группа 1 - здоровые люди, 2- больные гриппом, 3 - больные с артериальной гипертензией. В данном случае диагностические показатели x1, x2, x3, x4 являются признаками, а исследуемые пациенты - объектами.

При проведении реальных вычислительных экспериментов нельзя исключить возможность получения «артефактных» данных в ТЭД, например в случае возникновения ошибок при регистрации соответствующих показателей. Для поиска в ТЭД искаженных данных можно использовать построение и визуальный анализ гистограмм распределений признаков. При этом ошибочным может считаться тот показатель, который не вписывается в общую картину закона распределения соответствующих признаков. Например, если признаки подчинены нормальному (Гауссовскому) распределению, то за «истинные» данные можно принять интервал $\pm 2\sigma$ (где σ – стандартное среднеквадратическое отклонение). В этом случае все результаты за пределом данного интервала могут быть исключены из ТЭД. В тоже время при исключении ошибочных данных, а также в ряде других случаев (данные невозможно было собрать, или они утеряны) в ТЭД возникают пропуски, которые необходимо заполнить. Если при возникновении в строке таблицы пропусков потери информации не превышают 30% (например, если в таблице 10 столбцов, и в какой либо строке возникает 3 пропуска), то для заполнения пропусков можно использовать метод «максимального подобия». Суть метода состоит в том, что для восстановления утраченных данных в строках таблицы производится последовательный просмотр всех строк ТЭД с поиском такой строки, которая не имеет пропусков и максимально похожа на восстанавливаемую строку с

пропусками. При нахождении строки ТЭД удовлетворяющей указанным требованиям в строку с пропущенными данными копируются ячейки из найденной строки. Для нахождения строки с максимальным подобием можно использовать метод наименьших квадратов, в этом случае каждой строке ТЭД ставится в соответствие сумма квадратов отклонений (СКО) элементов заполняемой строки с соответствующими элементами всех строк ТЭД,

$$k_{.j} = \sum_{i=1}^{n} (x_{m_i} - x_{ij})^2,$$

где i –номер столбца, j - номер строки, m - номер строки с пропущенными ячейками. При этом в качестве наиболее «похожей» строки выбирается строка ТЭД с наименьшем показателем СКО.

Состав данных - это прежде всего состав признаков, которые характеризуют объекты. В целом, разрабатываемая система исходных признаков должна удовлетворять следующим требованиям:

1. Полнота описания. Система исходных признаков должна охватывать все выделенные аспекты измеряемого понятия.

2. Экономность описания. При разработке системы признаков следует избегать излишнего объема исходной информации, который может затруднить дальнейший эмпирико-статистический анализ параметров диагностической модели.

3. Четкая структурированность системы признаков. Признаки должны группироваться, относительно равномерно описывая все стороны измеряемого явления.

4. Количественная определенность отбираемых признаков. Эта определенность требуется для проведения эмпирико-статистического анализа. Признаки должны быть выражены в номинальной, качественной или количественной шкале.

Каждый реальный объект имеет бесконечное число различных свойств, отражающих его различные стороны. Естественно, что в каждом конкретном исследовании существенными являются не все свойства, а их ограниченный набор, определяющий наиболее важные признаки. Таким образом, после формирования исходной ТЭД необходимо решить задачу анализа информативности признаков таблицы. Другими словами: необходимо из имеющихся признаков отобрать некоторое их количество, которое бы описывало структуру данных так же, как и вся совокупность. Если исходить из того, что оставшегося набора признаков должно быть достаточно для восстановления исходной ТЭД, то, очевидно, тот признак, любое значение которого можно найти, зная значение другого признака, является лишним. Так бывает, например при функциональной, а гораздо чаще при статистической (корреляционной) зависимости. Кроме того, определенные признаки (и соответствующие им столбцы данных в ТЭД) могут не содержать в себе информации о различии объектов или их свойств в разных классах, т.е. быть неинформативными. Таким образом, вторая задача (после удаления артефактов и заполнения пропусков), решаемая на этапе предварительного анализа данных сводится к оценке степени корреляции

между столбцами ТЭД с целью исключения наименее коррелированных между собой, а также селекции наиболее информативных признаков. Для решения данной задачи используются различные методы:

1. Селекция признаков по относительному дисперсионному разбросу (внутриклассовой вариации). При использовании данного метода каждый признак характеризуется величиной $V_i = 100 D_i / E_i$.

Здесь D_i, E_i - оценки математического ожидания и дисперсии соответственно, i=1...n, где n - общее количество признаков. После определения величи-ны Vi строится гистограмма, по которой, начиная с нижней границы, производится исключение факторов, имеющих значение ниже порогового $V_{пор}$.

2. По парной корреляции, путем построения векторов парной корреляции признаков $R/k,i/$, k=1,n; i=1,n, где n-количество признаков, и $R/k,i/$- ранжированное с точностью до 0,1 определяется абсолютное значение коэффициента парной корреляционной связи между признаками k и i. Селекция производится с выбором факторов, у которых значение парной корреляции составляет значение меньше порогового $R_{пор}$.;

Использование первых двух методов обычно производится в качестве предварительного анализа. Так, использование селекции признаков по относительному дисперсионному разбросу позволяет выявить представительность отдельных признаков как внутри класса, так и вне его, построение векторов парной корреляции - оценить величины линейных зависимостей между исходными признаками. Однако, первый метод не позволяет оценить величины линейной зависимости между исходными классификационными переменными, а также оставляет в генерируемой совокупности признаки, положительно не влияющие на вероятность правильной классификации в конечном итоге. Использование второго метода позволяет выявить существующие линейные зависимости между исходными классификационными переменными, однако данный метод не обладает возможностью оценки информативности отдельного признака. Раздельное применение данных методик в качестве критерия отбора иногда приводит к получению противоречивых результатов, в то время, как и совместно взятые приемы в рамках одного метода, не лишены определенных недостатков. Так, метод селекции признаков по относительному дисперсионному разбросу не столько отражает свойства признака в смысле деления структуры формируемого пространства на классы, сколько говорит о равномерности или неравномерности распределения значений признака по классам. Более значимые результаты получаются при использовании следующих методов:

3. По информационному показателю силы влияния (ИПВ). Данный критерий отсева основывается на свойствах информационной энтропии и заключается в следующем: для каждого признака x_i (i=1..n, где n - общее количество признаков) строится гистограмма его значений, по которой оценивается общая энтропия $Э_i$. Поступая аналогичным образом в каждом классе, получаем внутриклассовые энтропии $Э_k$ (k=1...m, где v - общее

количество классов). Тогда показатель негэнтропии будет найден исходя из следующего выражения:

$$НЭ_i = Э_i \frac{(\sum_{i=1}^{k} n_i \cdot Э_{ki})}{N}, \quad i = \overline{1,k}$$

где N - общее количество объектов во всех классах, k - количестве классов, n_i - количество объектов в i - м классе (i=1...k), $Э_{ki}$ - внутриклассовая энтропия. Отсюда имеем показатель информативной силы влияния (ИПВi):

$$ИПВ_i = \sqrt{\frac{НЭ_i}{Э_i}}$$

, на основании которого выбирается оптимальный набор признаков, с максимальными значениями показателя ИПВi.

4. Метод отбора переменных на основании величины отношения внутренней обобщенной дисперсии к общей обобщенной дисперсии для отбираемых переменных. При использовании данного метода внутренняя обобщенная дисперсия вычисляется как определитель внутригрупповой матрицы перекрестных произведений $W(x)$ для переменных $x=(x_1,x_2,...,x_p)$, и аналогично, общая обобщенная дисперсия есть определитель общей матрицы перекрестных, произведений $T(x)$ для этих переменных. Отношение

$$\lambda(x) = \det W(x) / \det T(x)$$

называется λ - статистикой Уилкса. Оно принимает значения между нулем и единицей. Большие значения указывают на слабое разделение между группами, в то время как малые значения на хорошее разделение между группами (по крайней мере ,между некоторыми из них). Мультипликативное приращение

$$\lambda(ux) = \lambda(x,u) / \lambda(x)$$

в $\lambda(x)$, получаемое при добавлении переменной и к множеству $x=(x_1,x_2,..,x_p)$, называется частной λ статистикой. Соответствующая F - статистика есть

$$F = \frac{n \quad q \quad p \quad 1}{q \quad 1} \cdot \frac{\lambda(ux)}{\lambda(ux)}$$

где n - общее числе наблюдений, q - число групп, p - число переменных в текущем анализе. Используется для проверки значимости изменений в λ(x) в результате добавления u. Рассматриваемое в данной работе значение F статистики было использовано нами как одно из средств для отбора переменных. При этом F - статистика может рассматриваться как F - статистика включения (при включении переменной u в множество $x=(x_1,x_2,...x_p)$), так к как F - статистика исключения (при исключении переменной и из множестве $x=(x_1,x_2,...,x_p,u)$). При этом вводится ряд правил на запрещение включения и исключения переменных:

1. Переменная не исключается, если значение ее F - статистики исключения больше или равно установленному порогу.

2. Переменная не включается, если значение ее F - статистики включения ниже установленного порога.

3. Переменная не включается, если значение ее толерантности ниже установленного порога (в данном случае толерантность для включения классификационной переменной равна единице за вычетом квадрата ее внутригрупповой корреляции с текущими включенными переменными).

Для реализации вышеизложенных правил был использован следующий пошаговый алгоритм:

1. Исключается переменная с наименьшим значением F - статистики исключения, если это значение больше или равно пороговому значению F - статистики исключения.

2. Если исключение невозможно, то находится переменная с наибольшим значением F - статистики включения среди всех переменных, толерантность которых больше или равна порогу толерантности. Эта переменная включается, если значение ее F - статистики включения ниже порога для F - статистики включения.

3. Если включение или исключение переменной невозможно, то пошаговая процедура заканчивается, в противном случае, производится переход на п 1.

2. Цель работы.

Целью работы является приобретение навыков в формировании и предварительной обработке таблиц экспериментальных данных.

3. Порядок выполнения работы.

1. С использованием программы Excel сформировать исходный набор экспериментальных данных для 2-х классов с указанными преподавателем параметрами: количество признаков, вид функции плотности, координаты центров классов, дисперсия для каждого признака.

2. Удалить данные из ТЭД в указанных преподавателем позициях.

3. Провести предварительную обработку таблицы, а именно удалить артефакты и заполнить пропуски.

4. С использованием программы Excel провести анализ данных ТЭД, а также селекцию наиболее коррелированных и наименее информативных признаков.

4. Содержание отчета

1. Данные исходной ТЭД, с указанием количества классов, функции плотностей, математического ожидания и дисперсии для каждого класса ТЭД.

2. Гистограммы распределения объектов обучающей выборки для каждого признака ТЭД с указанием местоположения артефактов, а также порядок заполнения пропусков в ТЭД.

3. Описание анализа информативности признаков и этапов их селекции.

4. Выводы по работе.

5. Контрольные вопросы.

1. Дайте определение классов, объектов и признаков.

2. Что такое информативность признаков и в каких оценочных шкалах можно ее представить?

3. Каким образом производится заполнение пропусков в ТЭД и нахождение артефактов?

4. Какие методы оценки информативности признаков вы знаете?

Пример оформления лабораторной работы

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение высшего образования
«Юго-Западный государственный университет»

Кафедра биомедицинской инженерии

Лабораторная работа №1
По дисциплине «Интеллектуальные системы классификации и теории распознавания образов»

СОСТАВЛЕНИЕ ОБУЧАЮЩИХ ТАБЛИЦ ЭКСПЕРЕМЕНТАЛЬНЫХ ДАННЫХ И ИХ ПРЕДВАРИТЕЛЬНАЯ ОБРАБОТКА

Вариант ___

Выполнил: студент группы _____

Проверил: профессор кафедры БМИ
….._____

Курск-2016

1. Цель работы.

Приобретение навыков в формировании и предварительной обработки таблиц экспериментальных данных.

2. Выполнение работы.

1. Сформировать исходный набор экспериментальных данных для 2-х классов с указанными преподавателем параметрами: количество признаков, вид функции плотности, координаты центров классов, дисперсия для каждого признака.

21	100	-15	518	204
19	99	-11	482	208
19	96	-5	477	209
20	95	-18	517	204
15	89	-3	475	203
21	94	-7	479	193
25	88	-7	549	203
21	89	-11	494	194
24	110	-18	487	195
23	88	-17	486	190
21	66	-31	544	213
19	98	-16	465	190
19	100	-22	482	113
23	82	-14	485	211
21	99	-14	492	204
22	79	-8	516	195
20	109	-9	493	197
20	96	-10	481	196
18	85	-21	479	205
19	105	-14	497	189
29	111	-29	417	159
29	110	-31	427	160
28	128	-27	434	171
26	126	-38	427	163
35	106	-37	390	173
30	123	-34	387	145
33	121	-30	421	165
30	114	-39	460	169
26	118	-20	452	159
30	122	-31	440	157
33	117	-32	356	164

30	114	-40	427	162
27	107	-39		161
28	104	-34	398	159
29	123	-32	392	156
37	103	-35	438	162
28			417	165
31	116	-34	461	166
30	124	-30	417	154
29	105	-37	422	161

2. Провести предварительную обработку таблицы, а именно удалить артефакты и заполнить пропуски.

2.1. Заполнение пропусков возможно, если потеря информативности не превышает 30%. Это значит, что если в строке пропущено 2 и более значений, то она теряет информативность.

Необходимо найти наиболее похожую строку данных и скопировать оттуда интересующее значение.

Для этого необходимо найти среднеквадратическое отклонение строк от строки с пропуском по следующей формуле:

$$\sigma_x = \sqrt{\frac{\left(x_1^i - x_1^j\right)^2 + \left(x_2^i - x_2^j\right)^2 + \left(x_3^i - x_3^j\right)^2 + \left(x_4^i - x_4^j\right)^2 + \left(x_5^i - x_5^j\right)^2}{n}}$$

где x_1^i - значение в строке с пропуском; x_1^j - значение текущей строки; n -количество строк в таблице.

Находим из получившегося столбца σ_x наименьшее значение и помещаем его в строку с пропуском.

Допустимая потеря информативности, %	Реальная потеря информативности,%	Подсчет информативных строк	Проверка на расчет СКО	Метод максимального подобия, СКО
30	0	1		-
30	0	2		-
30	0	3		-
30	0	4		-
30	0	5		-
30	0	6		-
30	0	7		-
30	0	8		-
30	0	9		-
30	0	10	СКО не	-
30	0	11	рассчитывается	-
30	0	12		-
30	0	13		-

30	0	14		-
30	0	15		-
30	0	16		-
30	0	17		-
30	0	18		-
30	0	19		-
30	0	20		-
30	0	1	СКО рассчитывается	3,01749286
30	0	2		2,645751311
30	0	3		6,692493831
30	0	4		5,893796918
30	0	5		2,695024656
30	0	6		6,236564506
30	0	7		4,43668914
30	0	8		3,516876605
30	0	9		5,525062515
30	0	10		4,872047014
30	0	11		3,441236008
30	0	12		3,203616378
30	20	13		2,645751311
30	0	14		2,200478417
30	0	15		5,175854875
30	0	16		0
30	40	16		Не заполняется
30	0	17		3,418217449
30	0	18		5,520297475
30	0	19		1,960128889

Наименьшее СКО	Наименьшее значение столбца
1,960128889	422

Полученная матрица после заполнения пропусков:

Класс 1 после поиска пропусков				
x1	x2	x3	x4	x5
21	100	-15	518	204
19	99	-11	482	208
19	96	-5	477	209
20	95	-18	517	204
15	89	-3	475	203
21	94	-7	479	193
25	88	-7	549	203
21	89	-11	494	194
24	110	-18	487	195

23	88	-17	486	190
21	66	-31	544	213
19	98	-16	465	190
19	100	-22	482	113
23	82	-14	485	211
21	99	-14	492	204
22	79	-8	516	195
20	109	-9	493	197
20	96	-10	481	196
18	85	-21	479	205
19	105	-14	497	189
Класс 2 после поиска пропусков				
x1	x2	x3	x4	x5
29	111	-29	417	159
29	110	-31	427	160
28	128	-27	434	171
26	126	-38	427	163
35	106	-37	390	173
30	123	-34	387	145
33	121	-30	421	165
30	114	-39	460	169
26	118	-20	452	159
30	122	-31	440	157
33	117	-32	356	164
30	114	-10	427	162
27	107	-39	422	161
28	104	-34	398	159
29	123	-32	392	156
37	103	-35	438	162
28			417	165
31	116	-34	461	166
30	124	-30	417	154
29	105	-37	422	161

2.1. Поиск артефактов, т.е. тех значений, где данные неверны.

В таблице данных находим среднеквадратическое отклонение по формуле:

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^{n}\left(\overline{x_1} - x_{1i}\right)^2}{n-1}}\ ;\ \overline{x_1} = \frac{\sum_{i=1}^{n} x_{1i}}{n}$$

где $\overline{x_1}$ - математическое ожидание столбца; x_{1i} - текущее значение столбца; n - количество элементов выборки.

Если отклонение значения в столбце от соответствующего ему математического ожидания по абсолютной величине превышает удвоенное среднеквадратическое отклонение, то это значение является артефактом и его необходимо удалить.

Поиск артефактов класса 1

м-х1	х1	м-х2	х2	м-х3	х3	м-х4	х4	м-х5	х5
0,5	не артефакт	6,65	не артефакт	1,45	не артефакт	23,1	не артефакт	8,2	не артефакт
1,5	не артефакт	5,65	не артефакт	2,55	не артефакт	12,9	не артефакт	12,2	не артефакт
1,5	не артефакт	2,65	не артефакт	8,55	не артефакт	17,9	не артефакт	13,2	не артефакт
0,5	не артефакт	1,65	не артефакт	4,45	не артефакт	22,1	не артефакт	8,2	не артефакт
5,5	артефакт	4,35	не артефакт	10,55	не артефакт	19,9	не артефакт	7,2	не артефакт
0,5	не артефакт	0,65	не артефакт	6,55	не артефакт	15,9	не артефакт	2,8	не артефакт
4,5	не артефакт	5,35	не артефакт	6,55	не артефакт	54,1	артефакт	7,2	не артефакт
0,5	не артефакт	4,35	не артефакт	2,55	не артефакт	0,9	не артефакт	1,8	не артефакт
3,5	не артефакт	16,65	не артефакт	4,45	не артефакт	7,9	не артефакт	0,8	не артефакт
2,5	не артефакт	5,35	не артефакт	3,45	не артефакт	8,9	не артефакт	5,8	не артефакт
0,5	не артефакт	27,35	артефакт	17,45	артефакт	49,1	артефакт	17,2	не артефакт
1,5	не артефакт	4,65	не артефакт	2,45	не артефакт	29,9	не артефакт	5,8	не артефакт
1,5	не артефакт	6,65	не артефакт	8,45	не артефакт	12,9	не артефакт	82,8	не артефакт
2,5	не артефакт	11,35	не артефакт	0,45	не артефакт	9,9	не артефакт	15,2	не артефакт
0,5	не артефакт	5,65	не артефакт	0,45	не артефакт	2,9	не артефакт	8,2	не артефакт
1,5	не артефакт	14,35	не артефакт	5,55	не артефакт	21,1	не артефакт	0,8	не артефакт
0,5	не артефакт	15,65	не артефакт	4,55	не артефакт	1,9	не артефакт	1,2	не артефакт
0,5	не артефакт	2,65	не артефакт	3,55	не артефакт	13,9	не артефакт	0,2	не артефакт
2,5	не артефакт	8,35	не артефакт	7,45	не артефакт	15,9	не артефакт	9,2	не артефакт
1,5	не артефакт	11,65	не артефакт	0,45	не артефакт	2,1	не артефакт	6,8	не артефакт

Поиск артефактов класса 2

м-х1	х1	м-х2	х2	м-х3	х3	м-х4	х4	м-х5
0,9	не артефакт	4,368421053	не артефакт	4,105263158	не артефакт	3,25	не артефакт	2,55
0,9	не артефакт	5,368421053	не артефакт	2,105263158	не артефакт	6,75	не артефакт	1,55
1,9	не артефакт	12,63157895	не артефакт	6,105263158	не артефакт	13,75	не артефакт	9,45
3,9	не артефакт	10,63157895	не артефакт	4,894736842	не артефакт	6,75	не артефакт	1,45
5,1	не артефакт	9,368421053	не артефакт	3,894736842	не артефакт	30,25	не артефакт	11,45

0,1	не артефакт	7,631578947	не артефакт	0,894736842	не артефакт	33,25	не артефакт	16,55
3,1	не артефакт	5,631578947	не артефакт	3,105263158	не артефакт	0,75	не артефакт	3,45
0,1	не артефакт	1,368421053	не артефакт	5,894736842	не артефакт	39,75	не артефакт	7,45
3,9	не артефакт	2,631578947	не артефакт	13,10526316	артефакт	31,75	не артефакт	2,55
0,1	не артефакт	6,631578947	не артефакт	2,105263158	не артефакт	19,75	не артефакт	4,55
3,1	не артефакт	1,631578947	не артефакт	1,105263158	не артефакт	64,25	артефакт	2,45
0,1	не артефакт	1,368421053	не артефакт	6,894736842	не артефакт	6,75	не артефакт	0,45
2,9	не артефакт	8,368421053	не артефакт	5,894736842	не артефакт	1,75	не артефакт	0,55
1,9	не артефакт	11,36842105	не артефакт	0,894736842	не артефакт	22,25	не артефакт	2,55
0,9	не артефакт	7,631578947	не артефакт	1,105263158	не артефакт	28,25	не артефакт	5,55
7,1	артефакт	12,36842105	не артефакт	1,894736842	не артефакт	17,75	не артефакт	0,45
1,9	не артефакт	115,3684211	артефакт	33,10526316	артефакт	3,25	не артефакт	3,45
1,1	не артефакт	0,631578947	не артефакт	0,894736842	не артефакт	40,75	не артефакт	4,45
0,1	не артефакт	8,631578947	не артефакт	3,105263158	не артефакт	3,25	не артефакт	7.55
0,9	не артефакт	10,36842105	не артефакт	3,894736842	не артефакт	1,75	не артефакт	0,55

Полученная матрица данных после нахождения артефактов будет иметь вид:

Класс 1 после нахождения артефактов				
x1	x2	x3	x4	x5
21	100	-15	518	204
19	99	-11	482	208
19	96	-5	477	209
20	95	-18	517	204

	89	-3	475	203
21	94	-7	479	193
25	88	-7		203
21	89	-11	494	194
24	110	-18	487	195
23	88	-17	486	190
21				213
19	98	-16	465	190
19	100	-22	482	
23	82	-14	485	211
21	99	-14	492	204
22	79	-8	516	195
20	109	-9	493	197
20	96	-10	481	196
18	85	-21	479	205
19	105	-14	497	189

Класс 2 после нахождения артефактов				
x1	x2	x3	x4	x5
29	111	-29	417	159
29	110	-31	427	160
28	128	-27	434	171
26	126	-38	427	163
35	106	-37	390	173
30	123	-34	387	145
33	121	-30	421	165
30	114	-39	460	169
26	118		452	159
30	122	-31	440	157
33	117	-32	356	164
30	114	-40	427	162
27	107	-39	422	161
28	104	-34	398	159
29	123	-32	392	156
	103	-35	438	162
28			417	165
31	116	-34	461	166
30	124	-30	417	154
29	105	-37	422	161

После нахождения артефактов вновь образовались пропуски - заполним по описанной выше схеме.

Для класса 1:

Допустимая потеря информативности, %	Реальная потеря информативности, %	Подсчет информативных строк	Проверка на расчет СКО	Метод максимального подобия, СК01	Метод максимального подобия, СКО2	Метод максимального подобия, СКО3
30	0	1	СКО рассчитывается	5,501196042	7,535879092	8,762960327
30	0	2		9,553175278	3,402785237	2,78151795
30	0	3		10,29818584	5,078333751	3,332456025
30	0	4		4,801315609	7,749363302	8,657214819
30	20	5		9,920738512	8,54092316	5,301439725
30	0	6		9,17376121	4,866642634	1,100239208
30	20	7		118,4109703	111,8681084	110,3687599
30	0	8		5,591346698	5,37048464	3,034884893
30	0	9		10,00526177	0	2,481510576
30	0	10		7,563764031	5,062763963	2,502630195
30	60	10		119,8430992	114,6165231	110,4419352
30	0	11		12,67197488	5,880386668	4,161224516
30	20	12		45,77864361	2,955814965	45,05084846
30	0	13		8,149911204	6,509102938	3,741657387
30	0	14		7,584610457	3	3,260610211
30	0	15		0	10,01577703	8,058993015
30	0	16		8,687559163	2,655679118	2,772040859
30	0	17		8,941329231	4,052289801	0
30	0	18		9,386440388	6,215430713	3,324549831
30	0	19		7,639922843	2,955814965	4,116718161

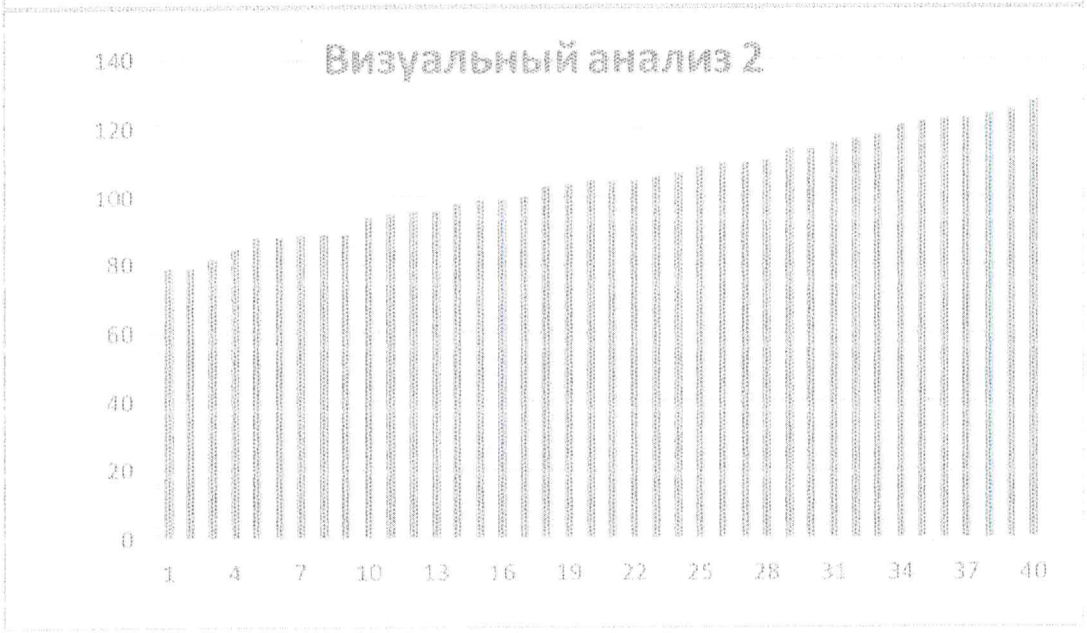
Наименьшее СКО1	Наименьшее значение столбца	Наименьшее СКО2	Наименьшее значение столбца	Наименьшее СКО3	Наименьшее значение столбца
4,801315609	20	2,655679118	109	1,100239208	193

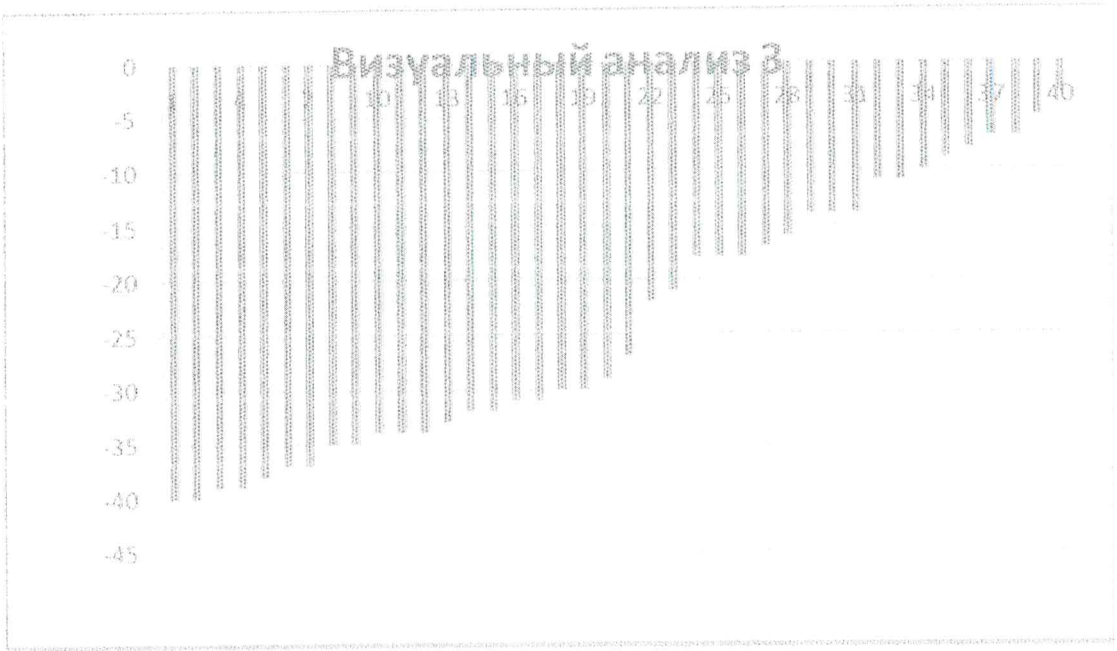
Для класса 2:

Допустимая потеря информативности, %	Реальная потеря информативности, %	Подсчет информативных строк	Проверка на расчет СКО	Метод максимального подобия, СКО1	Метод максимального подобия, СКО2	Метод максимального подобия, СКО3
30	0	1	СКО рассчитывается -	2,762531257	5,32125826	2,973567767
30	0	2		2,282657731	3,069373329	3,228165521
30	0	3		6,605420422	3,641861872	3,831998242
30	0	4		4,53640247	3,769685175	2,449489743
30	0	5		7,857279561	12,17633598	8,124038405
30	0	6		9,629996994	12,48578139	9,087295933
30	0	7		3,933660409	4,784844495	0
30	0	8		9,052478579	5,656854249	9,162279652
30	20	9		11,5758369	7,694837641	7,451598204
30	0	10		5,753717334	0	4,784844495
30	0	11		15,41359619	19,35132254	14,9419931
30	0	12		2	3,804429551	2,328315406
30	0	13		0	4,662278639	3,620119219
30	0	14		5,684697835	9,681778871	6,801702573
30	0	15		8,045920836	11,01912214	7,04123195
30	20	16		3,900067476	7,05243519	9,489606503
30	40	16		26,16848245	9,055385138	27,79814458
30	0	17		9,324557277	5,291502622	9,262260722
30	0	18		4,834089475	5,326201372	2,856202853
30	0	19		0,648885685	4,454447807	3,900067476

Наименьшее СКО1	Наименьшее значение столбца	Наименьшее СКО2	Наименьшее значение столбца	Наименьшее СКО3	Наименьшее значение столбца
0,648885685	29	3,069373329	110	2,328315406	-40

Ниже приведен визуальный анализ поиска артефактов для 5 признаков (x1-x5) по 2 классам одновременно:




Визуальный анализ 3
0
-5
-10
-15
-20
-25
-30
-35
-40
-45

Визуальный анализ 4
600
500
400
300
200
100
0
1 4 7 10 13 16 19 22 25 28 31 34 37 40

Визуальный анализ 5
250
200
150
100
50
0
1 4 7 10 13 16 19 22 25 28 31 34 37 40

Класс 1 после нахождения пропусков				
x1	x2	x3	x4	x5
21	100	-15	518	204
19	99	-11	482	208
19	96	-5	477	209
20	95	-18	517	204
20	89	-3	475	203
21	94	-7	479	193
25	88	-7	482	203
21	89	-11	494	194
24	110	-18	487	195
23	88	-17	486	190
21	89	-18	475	213
19	98	-16	465	190
19	100	-22	482	191
23	82	-14	485	211
21	99	-14	492	204
22	79	-8	516	195
20	109	-9	493	197
20	96	-10	481	196
18	85	-21	479	205
19	105	-14	497	189
Класс 2 после нахождения пропусков				
x1	x2	x3	x4	x5
29	111	-29	417	159
29	110	-31	427	160
28	128	-27	434	171
26	126	-38	427	163
35	106	-37	390	173
30	123	-34	387	145
33	121	-30	421	165
30	114	-39	460	169
26	118	-33	452	159
30	122	-31	440	157
33	117	-32	356	164
30	114	-40	427	162
27	107	-39	422	161
28	104	-34	398	159
29	123	-32	392	156
27	103	-35	438	162
28	105	-35	417	165
31	116	-34	461	166

30	124	-30	417	154
29	105	-37	422	161

3. Провести статистический анализ данных (найти основные статистические параметры рядов).

Для каждого столбца обоих класса найдем основные статистические показатели, используя следующие аналитические выражения.

3.1. Математическое ожидание

$$\bar{x} = \frac{\sum_{i=1}^{n} x_{1i}}{n}$$

, где x_i – текущее значение признака, n – число значений столбца для $\{x_i\}$.

3.2. Дисперсия

$$D = \frac{\sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2}{n}$$

3.3. Среднеквадратическое отклонение

$$\sigma_x = \sqrt{D}$$

Стандартное отклонение — это мера того, насколько широко разбросаны точки данных относительно их среднего.

3.4. Коэффициент вариации

$$V = \frac{\sigma_x}{\bar{x}} \cdot 100\%$$

3.5. Коэффициент асимметрии

$$As = \frac{\sum_{i=1}^{n} \left(x_i - \bar{x}\right)^3 \cdot \frac{1}{n}}{\sigma_x^3}$$

Асимметрия характеризует степень несимметричности распределения относительно его среднего. Положительная асимметрия указывает на отклонение распределения в сторону положительных значений. Отрицательная асимметрия указывает на отклонение распределения в сторону отрицательных значений.

3.6. Эксцесс

$$Ex = \frac{\sum_{i=1}^{n} \left(x_i - \bar{x}\right) \cdot \frac{1}{n}}{\sigma_x^4}$$

Эксцесс характеризует относительную остроконечность или сглаженность распределения по сравнению с нормальным распределением. Положительный эксцесс обозначает относительно остроконечное

распределение. Отрицательный эксцесс обозначает относительно сглаженное распределение.

Эксцесс характеризует форму распределения данных. Если он равен нулю, то данные подчинены нормальному закону распределения.

Для класса 1:

м.о.	20,75	94,5	-12,9	488,1	199,7
дисперсия	3,460526316	70,57894737	28,72631579	210,7263158	57,16842105
с.к.о.	1,860248993	8,401127744	5,359693628	14,51641539	7,560980165
V,%	8,965055388	8,890082269	-41,54801262	2,974065845	3,786169336
асимметрия	0,786926015	0,089946851	0,111759823	1,011242833	0,156683461
эксцесс	0,061465822	-0,443281046	-0,876239735	0,512172988	-1,269114888

Для класса 2:

м.о.	29,4	114,85	-33,85	420,25	161,55
дисперсия	5,410526316	65,81842105	13,39736842	671,5657895	38,57631579
с.к.о.	2,326053808	8,112855296	3,66024158	25,91458642	6,21098348
V,%	7,911747645	7,063870523	-10,81312136	6,166469107	3,84461992
асимметрия	0,740498086	-0,00571516	-0,040087388	-0,618356207	-0,57008406
эксцесс	0,586371704	-1,396964289	-0,848775562	0,712892364	1,739492631

4. Методом внутриклассовой вариации провести анализ данных ТЭД и селекцию наиболее информативных признаков.

Селекция признаков по относительному дисперсионному разбросу (внутриклассовой вариации).

При использовании данного метода каждый признак характеризуется величиной $V_i = 100 * СКО_i/МО_i$ После определения величины V_i строится гистограмма, по которой, начиная с нижней границы, производится исключение факторов, имеющих значение ниже порогового $V_{пор}$.

x1	x2	x3	x4	x5
21	•100	-15	518	204
19	99	-11	482	208
19	%	-5	477	209
20	95	-18	517	204
20	89	-3	475	203
21	94	-7	479	193
25	88	-7	482	203
21	89	-11	494	194
24	ПО	-18	487	195
23	88	-17	486	190

	21	89	-18	475	213
	19	98	-16	465	190
	19	100	-22	482	191
	23	82	-14	485	211
	21	99	-14	492	204
	22	79	-8	516	195
	20	109	-9	493	197
	20	96	-10	481	196
	18	85	-21	479	205
	19	105	-14	497	189
	29	111	-29	417	159
	29	110	-31	427	160
	28	128	-27	434	171
	26	126	-38	427	163
	35	106	-37	390	173
	30	123	-34	387	145
	33	121	-30	421	165
	30	114	-39	460	169
	26	118	-33	452	159
	30	122	-31	440	157
	33	117	-32	356	164
	30	114	-40	427	162
	27	107	-39	422	161
	28	104	-34	398	159
	29	123	-32	392	156
	27	103	-35	438	162
	28	105	-35	417	165
	31	116	-34	461	166
	30	124	-30	417	154
	29	105	-37	422	161
м.о.	25,075	104,675	-23,375	454,175	180,625
с.к.о.	4,848407087	13,13907365	11,53520253	40,12792844	20,48975667
V,%	19,33562148	12,5522557	49,34846003	8,835345063	11,34380992

Наиболее информативные признаки: x1, x3, x5

Селекция признаков по парной корреляции.

Случайным образом, используя ДСЧ, строятся векторы парной корреляции признаков R[k,i], k=1,n; i=1,n, где n-количество признаков, и R[k,i] - ранжированное с точностью до 0.1 абсолютное значение коэффициента парной корреляционной связи между признаками k и i. Селекция производится с выбором факторов, у которых значение парной корреляции составляет значение меньше порогового $R_{пор}$ =0,14 (взято из методических указаний к лабораторной работе №1).

	X1	X2	X3	X4	X5	∑(X1-X5)
X1	1	0,685456408	-0,813268787	-0,824692735	-0,831846084	-0,784351197
X2	0,685456408	1	-0,700885784	-0,663864772	-0,792697005	-0,471991153
X3	-0,813268787	-0,700885784	1	0,769735495	0,867385596	1,12296652
X4	-0,824692735	-0,663864772	0,769735495	1	0,827213241	1,10839123
X5	-0,831846084	-0,792697005	0,867385596	0,827213241	1	1,070055748

3. Выводы.

В ходе выполнения лабораторной работы приобретены навыки в формировании и предварительной обработке таблиц экспериментальных данных, а именно: навыки по восстановлению пропущенных данных в ТЭД при помощи метода максимального подобия; навыки по удалению артефактов, которые не удовлетворяю правилу 2-х сигм; навыки построения гистограмм распределения объектов по каждому признаку

Лабораторная работа №2

ПОСТРОЕНИЕ ЛИНЕЙНЫХ РАЗДЕЛЯЮЩИХ ПОВЕРХНОСТЕЙ

1. Краткие теоретические сведения.

Основной задачей в теории распознавания образов (ТРО) является построение решающего правила (РП) - математического выражения, позволяющего отличить один класс от другого. Приведем для примера наиболее простое решающее правило - отметку 37° на шкале градусника. Температура выше данной отметки позволяет судить принадлежности исследуемого объекта (человека) к классу больных людей, в противном случае объект принадлежит к классу здоровых людей. Одним из способов построение РП является построение линейных разделяющих поверхностей вида $y= a_1x_1+a_2x_2+a_3x_3+\ldots a_nx_n$, (1), где $a_1\ldots a_n$ - настраиваемые параметры, $x_1\ldots x_n$ - классификационные признаки. Уравнение (1) представляет собой ничто иное как каноническое уравнение плоскости в пространстве вида $a_1x_1+a_2x_2+\ldots a_nx_n+C=0$, в котором свободный член уравнения C вынесен в правую часть и обозначен через y. Таким образом, при изменении параметра y уравнение (1) задаст семейство параллельных плоскостей, при этом параметры $a_1\ldots a_n$, определяют углы наклона плоскости к осям координат, y - есть параметр, пропорциональный Евклидовому расстоянию от начала координат до плоскости (или гиперплоскости, в зависимости от размерности пространства). В другой постановке задачи, при подстановке в уравнение (1) классификационных признаков $x_1\ldots x_n$, будет получено расстояние y, пропорциональное расстоянию от начала координат до плоскости, проходящей через точку $x_1\ldots x_n$. Данный параметр y, в данном случае, будет являться решающим правилом, на основании которого принимается решение о принадлежности объекта с координатами $x_1\ldots x_n$ к классу ω_k. Рассмотрим геометрическую интерпретацию разделения классов поверхностью (рис. 1): пусть классификационные признаки - оси n мерного пространства. Каждый объект (наблюдение), является точкой этого пространства с координатами, представляющими собой наблюдаемые значения этой переменной. Если классы отличаются друг от друга по наблюдаемым переменным, их можно представить как скопление точек. Так как классы частично пересекаются, соответствующие им «территории» не совпадают. Для определения положения класса можно использовать их центры (центроиды).

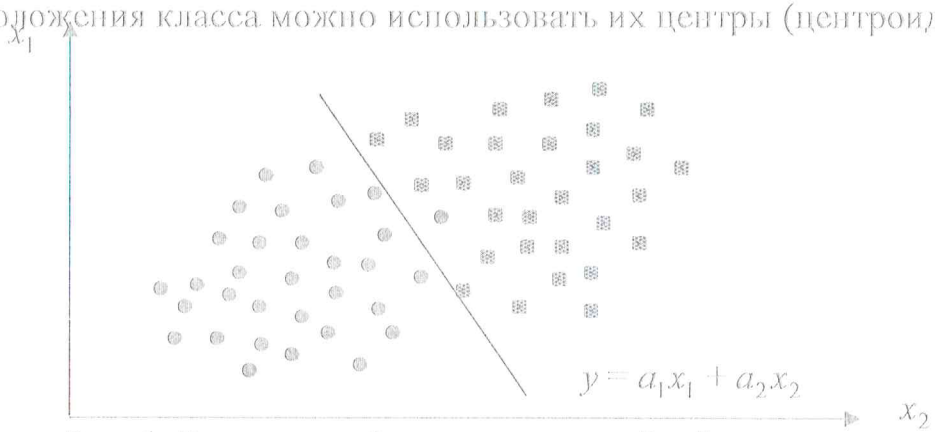


Рис. 1. Разделение 2-х классов линейной поверхностью

Роль числа классов становится понятной, если обратиться к геометрическим аналогам. Для любых пространств, где применимы аксиомы Евклидовой геометрии, 2 точки определяют положение прямой линии, 3-плоскость, 4- 3-х мерную поверхность и.т.д. Принцип сводится к тому, что точки определяют пространство (Линию, плоскость и. т.д.), имеющие размерность на единицу меньшую, чем число точек.

Процесс нахождения настраиваемых параметров $a_1 \ldots a_n$ называется процессом обучения. При этом подбираются такие значения параметров, при котором будет достигнут верхний критерий качества разделения классов. Данная задача может решаться как аналитически (например, с использованием методов аналитической геометрии в пространстве), так и итерационно. В последнем случае производится пошаговое настраивание параметров модели до достижения верхнего экстремума качества. При использовании аналитических методов можно использовать следующий подход. Плоскость, проходящая через точку $M_0(x_0, y_0, z_0)$ и перпендикулярная к вектору N(A,B,C) представляется уравнением вида $A(x-x_0)+B(y-y_0)+C(z-z_0)=0$ (2), или $Ax+By+Cz+D=0$, (3), где через D обозначена величина $-(Ax_0+By_0+Cz_0)$. Вектор N(A, B, C) называется нормальным вектором разделяющей плоскости. Полученное уравнение вида (3) легко преобразуется в уравнение (1) путем изменения обозначений переменной и переноса свободного члена уравнения (3) в правую часть. Рассмотренный случай с использованием 3-х мерного пространства может быть перенесен на пространство размерностью n.

2. Цель работы.

Целью работы является приобретение навыков в построении решающего правила в виде линейных разделяющих поверхностей.

3. Порядок выполнения работы.

1. С использованием программы Excel сформировать исходный набор экспериментальных данных для 2-х классов с указанными преподавателем параметрами: количество признаков, вид функции плотности, координаты центров классов, дисперсия для каждого признака.

2. На основании полученных данных найти координаты центров классов.

3. С использованием описанной выше методики найти уравнение линейной разделяющей поверхности аналитическим методом.

4. С использованием программы Excel провести поиск уравнения линейной разделяющей поверхности итерационным методом, а также определить время поиска. Сравнить полученное уравнение с уравнением, полученным в п.п.3.

5. Провести отображение данных в 2-мерное пространство. Для этого используя первые 2 координаты признакового пространства отобразить объекты классов ТЭД в декартовом пространстве x_1, x_2. Также отобразить полученные уравнения.

6. Используя данные ТЭД в качестве контрольной выборки провести оценку качества полученных моделей.

4. Содержание отчета

1. Данные исходной ТЭД, с указанием количества классов, функции плотностей, математического ожидания и дисперсии для каждого класса ТЭД.

2. Результаты моделирования: промежуточные этапы формирования уравнения поверхности аналитическим методом и уравнение, полученное итерационным методом.

3. Результаты отображения данных и результат оценки качества модели

4. Выводы по работе.

5. Контрольные вопросы.

1. Общий вид уравнения разделяющей поверхности.

2. Что называется решающим правилом?

3. Для каких структур данных рассмотренный метод наиболее подходит?

4. Каким образом производится поиск решающего правила аналитически итерационно? Написать алгоритм итерационного поиска решающего правила.

5. Какие методы анализа информативности признаков вы знаете?

Лабораторная работа №3

ИЗУЧЕНИЕ ЭТАЛОННЫХ КЛАССИФИКАТОРОВ

1. Краткие теоретические сведения.

Среди задач принятия решений связанных с обработкой ТЭД для получения решающих правил, одним из методов классификации является построение решающего правила в виде эталонных классификаторов. Эталоны, как правило, состоят из набора элементарных замкнутых объемов многомерного пространства, образованного выбранными, на этапе предварительного анализа данных, классификационными переменными. Чаще всего эталоны представляют в виде гиперкубов, гиперпараллелипипедов, гиперсфер, гиперэлипсоидов и.т.д.

В практических приложениях используют гиперпараллелипипеды, описываемые системой неравенств типа (1) и гиперсферы типа (2).

$$\bigcap_{i=1}^{n} a_{ik} \le x_i \le b_{ik} \qquad (1),$$

где i - номер классификационной переменной, a_{ik} и b_{ik} - нижняя и верхняя граница ограничения w_k по признаку x_i.

$$\sqrt{\sum_{i=1}^{n}(x_i - m_i)^2} \le R_k, \qquad (2)$$

где m_i - коэффициент, определяющий центр гиперсферы, R_k - её радиус.

На рисунке 1 и 2 представлена геометрия тип темой выражений типа 1 и 2 соответственно.

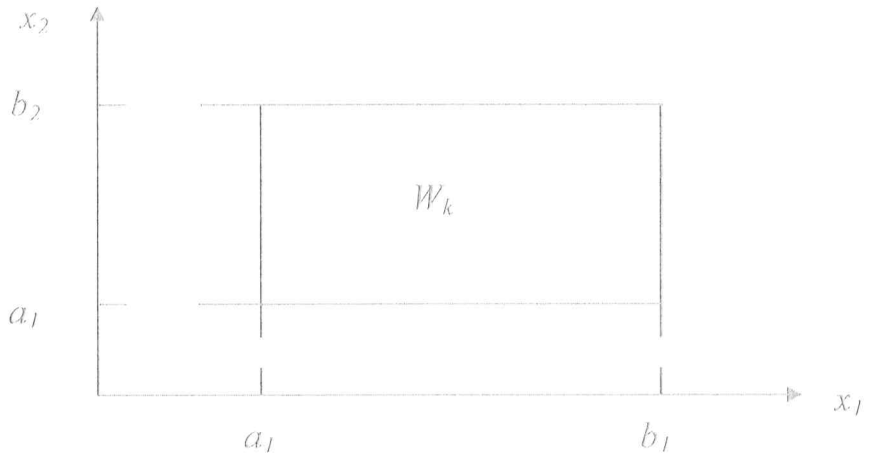


Рис. 2.9. Геометрия гиперобъема, описываемого системой выражений типа (2.6)

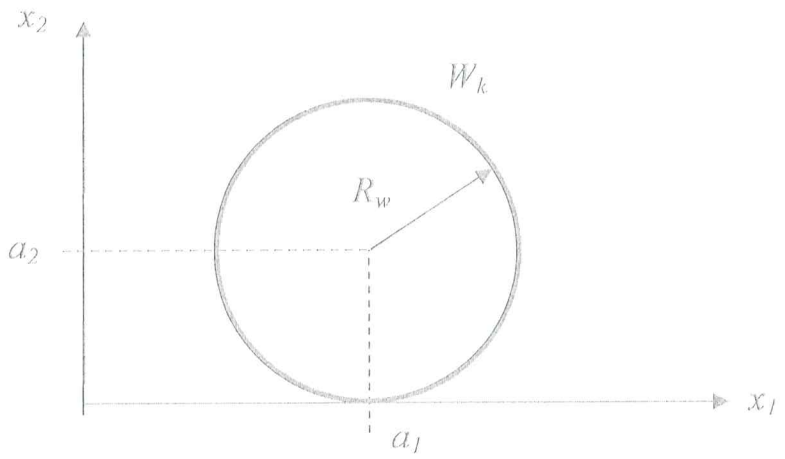


Рис. 2.10 Геометрия гиперобъема, описываемого выражением (2.5)

Классы со сложной геометрической структурой могут быть представлены в виде целого набора элементарных гиперобъемов, например, так, как это представлено на рисунке 3.

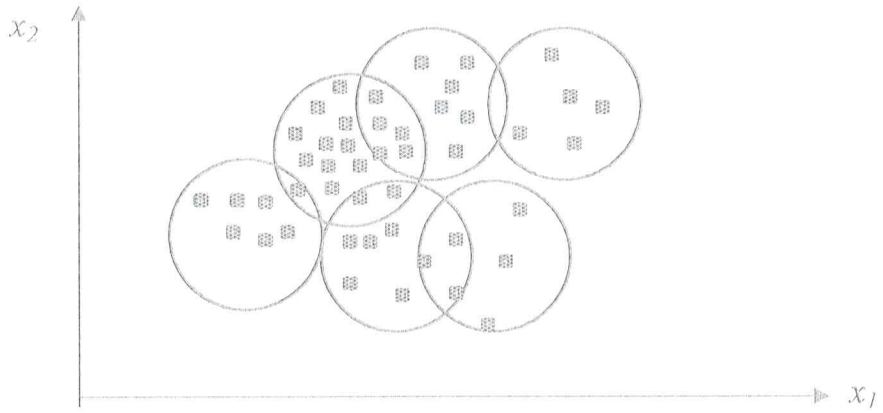


Рис 3. Представление классов со сложной геометрической борами элементарных гиперобъемов.

Построение решающих правил в виде эталонных классификаторов сводится к нахождению настраиваемых параметров модели типа 1 и 2. Для гиперкуба типа 1 - настраиваются параметры a_{ik} и b_{jk}, которые представляют собой максимальное и минимальное значения элементов ТЭД для k-го класса i-го признака, для гиперсферы типа 2 определяются параметры центра гиперсферы, например как центр масс данного класса, а также радиуса R_k. Нахождение данного параметра сводится к поиску наиболее удаленной от центра точки в ТЭД для класса номер k. При этом найденное расстояние и будет соответствовать заданному параметру.

2. Цель работы.

Целью работы является приобретение навыков в построении решающего правила в виде эталонных классификаторов.

3. Порядок выполнения работы.

1. С использованием программы Excel сформировать исходный набор экспериментальных данных для 2-х классов с указанными преподавателем параметрами: количество признаков, вид функции плотности, координаты центров классов, дисперсия для каждого признака.

2. На основании полученных данных найти координаты центров классов.

3. С использованием описанной выше методики найти решающие правила, представленные гиперсферами и гиперпараллелипипедами.

4. Провести отображение данных в 2-мерное пространство. Для этого используя первые 2 координаты признакового пространства отобразить объекты классов ТЭД в декартовом пространстве x_1, x_2. Также отобразить проекции полученных решающих правил.

6. Используя данные ТЭД в качестве контрольной выборки провести оценку качества полученных моделей.

4. Содержание отчета.

1. Данные исходной ТЭД, с указанием количества классов, функции плотностей, математического ожидания и дисперсии для каждого класса ТЭД.

2. Результаты моделирования: промежуточные этапы формирования решающих правил.

3. Результаты отображения данных и результат оценки качества модели

4. Выводы по работе.

5. Контрольные вопросы.

1. Общий вид рассмотренных в работе эталонных классификаторов.

2. Что называется решающим правилом?

3. Для каких структур данных рассмотренные методы распознавания образов наиболее подходят?

4. Каким образом производится поиск решающего правила аналитически

5. Напишите алгоритм поиска решающего правила в виде эталонов представленных гиперсферами.

Приложение

Вариант 1

21	100	-15	518	204
19	99	-11	482	208
19	96	-5	477	202
20	95	-18	517	204
15	89	-3	475	207
21	94	-7	479	192
25	88	-7	549	203
21	89	-11	494	194
24	110	-18	487	195
23	88	-17	486	196
21	66	-31	544	213
19	98	-16	465	196
19	100	-22	482	112
23	82	-14	485	211
21	99	-14	492	204
22	79	-8	516	195
20	109	-9	493	197
20	96	-10	481	196
18	85	-21	479	205
19	105	-14	497	189
29	111	-29	417	152
29	110	-31	427	160
28	128	-27	434	171
26	126	-38	427	163
35	106	-37	390	173
30	123	-34	387	145
	121	-30	421	165
30	114	-39	460	162
26	118	-20	452	152
30	122	-31	440	157
33	117	-32	356	164
30	114	-40	427	162
27	107	-39		161
28	104	-34	398	152
29	123	-32	392	156
37	103	-35	438	162
28			417	165
31	116	-34	461	166
30	124	-30	417	154
29	105	-37	422	161

Вариант 2

-18	1078	-6	2089	210
-21	1060	2	300	220
-21	991	14	32	224
-20	948	-14	2065	210
-26	780	20	-74	208
-19	921	10	159	182
-15	742	10	3635	208
-19	773	3	904	184
-16	1364	-13	540	188
-17	755	-10	491	175
-18	857	1	3395	232
-21	1047	-9	-544	176
-22	1083	-21	307	231
-16	579	-5	434	229
-19	1063	-4	791	211
-17	500	8	2012	188
-20	1354	7	861	192
-20	983	5	265	191
-23	663	-19	134	212
-21	1242	-4	1031	172
-31	2	55	5631	158
-31	-29	50	6146	160
-32	482	58	6506	187
-34	424	34	6139	166
-24	-128	38	4307	292
-31	345	45	4127	124
-26	282	53	5839	172
-30	85	33	7815	184
-35	203	74	7406	158
-30	325	49		153
-26	158	48	4095	171
-30	82	30	6128	166
-34	-106	33	5468	164
-33	-205	43	4699	158
-31	341	48	4406	151
-22	-231	42	6714	166
-32	225	43	5662	173
-29	136	44	7854	176
-30	01	52	5645	145
-31	-167	38	5892	163

Вариант 3

21	1060	2	300	220
27	1120	2	301	240
21	1000	2	302	241
27	1000	3	314	247
15	1120	4	312	270
20	1070	7	300	272
14	1260	2	270	276
15	1070	2	301	245
24	1060	5	312	271
26	1072	5	307	240
29	1100	3	310	290
24	1060	8	310	300
24	1090	2	275	312
26	1060	7	301	301
22	1120	5	315	490
21	1063	4	306	270
19	1000	5	247	290
19	1072	4	301	276
21	1060	7	300	247
23	1090	9	306	254
23	1105	8	315	290
17	900	2	309	297
18	999	2	301	301
17	1092	5	900	304
4	1100	4	302	301
20	1067	3	274	249
21	1092	3	304	234
24	1010	5	333	291
18	1111	3	307	301
18	1090	4	305	255
19	997	3	303	281
24	1070	5	334	241
22	990	2	308	279
14	1061	2	301	249
27	1114	5	309	277
27	1069	3	315	250
21	1030	2	292	290
17	1001	3	311	284
23	1060	3	300	249
18	1100	2	315	292
17	1067	5	309	249

Вариант 4

518	2009	81	68	117
490	2008	92	62	119
483	2003	109	61	120
496	2000	72	68	117
438	1988	116	61	117
515	1998	104	62	111
556	1985	103	73	117
513	1987	93	64	112
548	2030	73	63	153
532	1986	77	63	110
518	1993	92	72	121
488	2007	78	59	110
483	2009	62	62	121
539	1973	83	62	121
511	2008	85	64	117
529	1968	100	68	113
497	2029	99	64	113
499	2002	97	62	113
471	1979	64	61	117
489	2021	85	64	109
498	1980	25	74	122
497	1974	21	76	122
495	2066	27	78	127
491	2055	11	76	1
513	1957	13	69	18
499	2041	18	68	115
509	2030	23	75	124
499	1995	10	83	127
490	2016	37	81	122
501	2037	21	79	121
508	2008	20	68	124
501	1994	8	76	123
491	1961	10	74	123
494	1943	17	71	122
498	2040	20	69	120
517	1938	16	79	123
	2020	17	74	125
503	2004	17	83	125
501	2046	23	74	119
498	1950	13	75	123

Вариант 5

486	87	59	101	11
430	86	67	89	12
417	82	79	87	12
443	80	51	101	11
326	70	85	87	11
479	78	75		11
563	68	75	111	11
477	70	68	93	11
546	104	52	91	11
514	69	55	90	10
485	75	66		12
425	86	56	83	10
416	88	44	89	12
528	59	60	90	12
473	86	61	92	11
507	54	73	100	11
443	103	72	93	11
448	82	70	89	11
	64	46		11
427	97	61	94	10
439	47	45	74	14
434	46	41	76	14
	60	47	77	14
406	59	31	76	14
516	43	33	70	14
444	57	38	69	13
493	55	43	75	14
446	49	30	82	146
398	53	57	80	142
453	56	41	78	14
491	51	40	69	14
454	49	28	76	14
407	44	30	74	143
420	41	37		142
439	56	40	70	141
534	40	36	78	143
429	53	37	75	144
465	51	37	82	155
455	57	43	74	140
438	42	33	75	142

Вариант 6

993	4987	810	98	499
997	5000	801	110	508
994	4989	814	99	512
996	4999	795	113	490
	4996	806	108	489
997	4996	790	105	493
1016	4991	781	90	510
998	5003	788	101	503
1002	5003	786	99	508
1001	5009	806	101	496
1001	4986	802	108	498
998	5000	802	102	499
1003		805	95	498
1004	5003	794	146	499
1009	5001	802	98	499
1002	5008	814	101	501
996	5005	804	95	485
1002	5005	797	88	507
994	4993	794	101	499
989	4995	801	90	502
1016	5017	829	126	487
1026	5021	788	74	472
1039	5004	804	96	469
1003	4989	848	82	436
1028	4962	786	113	472
1045	5038	802	94	688
1020	4956	812	87	524
1040	5054	783	61	533
1015	5022	813	198	519
1019	5012	839	132	541
938	5041	785	105	548
982	4979	869	96	502
1012	4995	739	117	463
1049	5021	828	121	528
1019	5055	778	44	526
991	4997	748	73	472
1016	5059	828	44	481
1042	4974	778	96	454
1015	5012	752	96	529
990	4991	840	95	462

Вариант 7

88	-60	922	477	200
118	-51	896	506	189
99	-49	891	527	191
91	-49	870	483	201
92	-40	895	473	220
115	-46	895	472	215
133	-66	867	511	197
81	-43	892	493	207
112	-49	906	487	170
124	-42	913	526	197
120	-60	915	467	206
104	-52	893	513	195
108	-58	871	480	199
119	-55	906	489	215
59	-54	889	476	208
422	-45	880	506	208
118	-32	906	510	195
91	-85	902	515	217
108	-54	896	502	195
94	-41	881	539	203
178	-24	887	528	209
30	-4	922	534	200
133	-82	864	598	238
168	-51	884	553	250
97	-84	860	595	169
148	-40	951	457	180
96	-86	994	519	126
80	-95	867	429	222
45	-71	863	486	112
105	35	871	537	240
104	-86	874	460	263
164	-14	919	454	110
16	-78	859	537	190
97	-107	849	533	165
128	-68	958	593	249
136	-50	834		245
173	-58	865	484	118
124	-109	851	445	237
73	-80	984	488	224
125	3	849	566	153

Вариант 8

492	-47	804		
500	-53	796	198	105
501	-55	806	195	94
500	-54	798	201	93
496	-53	790	199	100
501	-44	801	199	89
504	-56	807	201	99
504	-53	799	210	101
502	-50	793	201	108
504	-43	799	200	104
499	-48	795	193	99
494	-35	803	205	97
504	-52	800	203	104
495	-60	797	197	93
500	-52	797	208	103
500	-48	801	206	102
502	-53	797	197	94
507	-50	797	195	108
499	-60	796	204	106
496	-55	802	194	105
494	-73	820	182	124
551	-62	819	200	143
497	-76	776	215	156
485	-98	814	191	123
498	-111	999	165	52
	-61	791	189	88
532	-90	845	191	61
533	-10	797	180	128
503	-81	770	212	78
521	-50	824	220	115
485	-62	761	187	93
517	-79	767	192	113
514	-93	784	184	119
509	-35	812	201	66
531	-64	800	210	87
458	-62	772	193	83
479	-54	769	180	94
475	-71	799	182	112
459	-28	784	192	128
527	-11	816	194	115

Вариант 10

6400	4300	50	90	8700
6300	4100	51	94	8720
6100	4200	54	87	8300
9400	7100	150	904	8380
6402	4010	52	83	8400
6100	4000	47	81	8200
6150	4050	57	93	8800
6500	4400	60	82	8900
6000	4201	43	80	8450
6280	4120	53	96	8770
6660	4670	43	97	8940
6010	4120	47	120	8990
6120	4440	55	101	8380
6222	4180	50	84	8880
6501	4611	60	104	8999
7770	4120	900	106	8380
6590	4600	47	99	8701
6540	4390	43	85	8444
6300	4010	56	97	8120
6720	4290	57	103	8170
6390	4017	43	89	8950
6190	4301	44	88	8895
6401	4500	45	91	8702
6700	4600	59	99	8777
6800	4009	60	74	8000
6610	4155	43	78	8100
6090	4099	44	91	8120
6044	4709	51	93	8220
6601	4770	47	99	8280
6809	4308	43	105	8991
6605	4800	49	104	8855
6000	5001	60	100	9000
6900	4833	53	90	8702
6904	4099	51	94	8720
6305	4100	44	81	8120
6301	4180	43	82	8128
6280	4000	51	101	8902
1105	4000	20	11	10000