

МІНІСТЕРСТВО ОСВІТИ І НАУКИ
Запорізький національний технічний університет

МЕТОДИЧНІ ВКАЗІВКИ

до лабораторних робіт з дисципліни
«Теорія прийняття рішень»
для студентів спеціальності
121 “Інженерія програмного забезпечення”
122 “Комп’ютерні науки”
всіх форм навчання

2018

Методичні вказівки до лабораторних робіт з дисципліни «Теорія прийняття рішень» для студентів спеціальності 121 “Інженерія програмного забезпечення”, 122 “Комп’ютерні науки” всіх форм навчання / Укладачі: В.І. Дубровін, Л.Ю. Дейнега, А.В. Белова. – Запоріжжя: ЗНТУ, 2018. – 32 с.

Укладачі: Дубровін В.І., канд.техн.наук, професор
Дейнега Л.Ю., ст. викладач
Белова А.В., асистент

Рецензент: Притула А. В., канд. техн. наук, доцент

Відповідальний за випуск: В.І. Дубровін

Затверджено
на засіданні кафедри
програмних засобів

Протокол № 6 від 15.02.2018 р.

ЗМІСТ

Вступ	2
Лабораторна робота № 1. Евристичний алгоритм класифікації з використанням даних навчального експерименту	5
Лабораторна робота № 2. Оптимальна оцінка значення прогнозованого параметра.....	9
Лабораторна робота № 3. Оптимальна класифікація	13
Лабораторна робота № 4. Метод дискримінантних функцій	20
Лабораторна робота № 5. Метод потенціальних функцій.....	24
Література	29

ВСТУП

Для завоювання українським виробником твердих позицій на європейському ринку необхідно пройти певні етапи, а саме, позбутися такого серйозного недоліку, як високий рівень дефектності вітчизняної продукції.

Ефективним засобом проти даного недоліку є діагностика процесів і виробів [1-13]. Ефективним елементом технічної діагностики є методи й алгоритми розпізнавання.

Великий внесок у розвиток теорії й практики розпізнавання образів внесли вітчизняні й закордонні вчені: В.М. Глушков, Н.П. Бусленко, О.Г. Івахненко, М.А. Айзерман, Е.М. Браверман, В.Н. Вапник, М.Г. Загоруйко, В.С. Михалевич, В.С. Пугачов, Ю.І. Журавльов, Я.З. Ципкін, М.М. Бонгард, В.Л. Матросов, Л.А. Растрігін, О.Л. Горелик, І.Б. Гурвич, В.О. Скрипкін, О.Я. Червоненкіс, С.А. Айвазян, В.М. Васильєв, О.О. Харкевич, Ф. Розенбат, Р. Гонсалес, Х. Гренандер, Р. Дуда, Г. Себестіан, Дж. Ту, К. Фу, П. Харт, К. Фукунага та ін. [14-59].

У даному лабораторному практикумі представлено ряд методів і алгоритмів теорії розпізнавання образів.

При відборі цих методів і алгоритмів для включення в даний цикл лабораторних робіт враховувався той факт, що зусилля потрібно спрямовувати, у першу чергу, на створення простих не елітарних методик, адресно розрахованих на певні категорії працівників (топ-менеджерів, інженерів, технологів, робітників). Спрощена й навіть менш ефективна методика за рахунок масовості її застосування виграє у більш ефективної, але складної.

Лабораторна робота № 1.
**ЕВРИСТИЧНИЙ АЛГОРИТМ КЛАСИФІКАЦІЇ
З ВИКОРИСТАННЯМ ДАНИХ НАВЧАЛЬНОГО ЕКСПЕРИМЕНТУ**

Мета роботи – вивчити один із евристичних алгоритмів індивідуального прогнозування – класифікацію за однією ознакою з використанням даних навчального експерименту; навчитися обирати оптимальне порогове значення ознаки.

1. КОРОТКІ ТЕОРЕТИЧНІ ВІДОМОСТІ

Дана робота відкриває цикл робіт, присвячених вирішенню задачі індивідуального прогнозування. У наш час чималі можливості прогнозування привертають увагу широкого кола фахівців як в області дослідження, розробки й теоретичного обґрунтування методів прогнозування, так і в області їхнього практичного застосування у виробництві й експлуатації.

Використання індивідуального прогнозування у виробництві дозволяє усунути потенційно ненадійні вироби з готової продукції, що саме по собі є дуже важливим, тому що сприяє підвищенню ефективності експлуатації. Крім того, аналіз причин випуску потенційно ненадійних виробів дозволяє оперативно впливати на виробництво, тому що з'являється можливість науково обґрунтованого керування якістю продукції, що випускається, за рахунок введення зворотного зв'язку від прогнозування до виробництва.

Мета індивідуального прогнозування в експлуатації полягає в запобіганні відмовам і збільшенні строків між профілактичними роботами шляхом виявлення й виключення з експлуатації потенційно ненадійних екземплярів з погіршеними значеннями параметрів і інтенсивним старінням.

При оптимальній класифікації за однією ознакою для вирішення задачі індивідуального прогнозування необхідно знати двовимірну густину розподілу ознаки й параметру, що прогнозується. В методі оптимальної класифікації аналітично знаходиться таке порогове значення ознаки, при якому величина середніх втрат мінімальна. Однак аналітичні перетворення, необхідні при цьому, навіть при одній ознаці й нормальному спільному законі є досить складними. Якщо ж спільна густина невідома, то необхідно провести статистичний експеримент для визначення оцінки густини розподілу.

У даній лабораторній роботі розв'язання задачі класифікації за однією ознакою здійснюється за даними навчального експерименту. При цьому немає необхідності прибгати до складних аналітичних перетворень

і статистичних експериментів з визначення оцінки спільної густини, якщо вона невідома. В даній роботі безпосередньо за даними навчального експерименту в ході іспиту підбирається таке порогове значення ознаки, при якому необхідна ймовірність помилкових рішень мінімальна відповідно до обраного критерію (зазвичай в якості такого критерію для підбора порогового значення ознаки обирається мінімум ризику споживача).

2. ДОМАШНЄ ЗАВДАННЯ

Використовуючи конспект лекцій та рекомендовану літературу, вивчити метод класифікації за однією ознакою з використанням даних навчального експерименту.

Ознайомитися зі змістом і порядком виконання роботи.

3. ПОРЯДОК ВИКОНАННЯ РОБОТИ

3.1. Провести класифікацію резисторів, прийнявши в якості ознаки, що характеризує їхню стабільність, напругу фліккерного шуму $U_{ш}$, а в якості прогнозованого параметру – коефіцієнт старіння α .

Масив вихідних даних, отриманих у результаті навчального експерименту, приведено в таблиці 1.

Таблиця 1 – Дані навчального експерименту

№ тр-ра	$U_{ш}$	$\alpha \cdot 10^3$	№ тр-ра	$U_{ш}$	$\alpha \cdot 10^3$	№ тр-ра	$U_{ш}$	$\alpha \cdot 10^3$	№ тр-ра	$U_{ш}$	$\alpha \cdot 10^3$
1	2.5	0.07	16	8.4	0.18	31	6.5	0.18	46	6.3	0.16
2	3.8	0.07	17	3.7	0.05	32	5.7	0.12	47	8.6	0.31
3	3.8	0.09	18	8.5	0.28	33	4.7	0.2	48	13	0.34
4	3.8	0.15	19	5.5	0.1	34	8.5	0.22	49	10.2	0.26
5	6.3	0.15	20	4.3	0.08	35	5.5	0.11	50	12	0.35
6	6.3	0.18	21	11.1	0.3	36	4.6	0.11	51	5.7	0.17
7	4.6	0.1	22	5.3	0.08	37	8.8	0.3	52	10.3	0.30
8	8.6	0.28	23	5.5	0.15	38	8.5	0.27	53	5.5	0.1
9	4.6	0.15	24	10.1	0.24	39	5.8	0.16	54	3.4	0.03
10	7.7	0.23	25	5.7	0.15	40	7.6	0.22	55	7.9	0.24
11	6.5	0.22	26	4.9	0.17	41	3.7	0.04	56	4.9	0.18
12	10.2	0.27	27	10.3	0.27	42	6.5	0.17	57	10.3	0.28
13	4.6	0.18	28	7.9	0.24	43	6.7	0.18	58	4.6	0.09
14	8.4	0.26	29	2.7	0.06	44	11.3	0.31	59	8.4	0.27
15	8.5	0.27	30	2.3	0.05	45	10.3	0.29	60	8.5	0.29

Примітка: напругу шуму $U_{ш}$ приведено в умовних одиницях.

При виконання завдання значення $U_{ш}$ помножити на номер варіанту. Порогове значення прогнозованого параметру $U_{ш,кл}$ дорівнює номеру варіанту.

3.2. Написати й налагодити програму в пакеті Matlab, яка реалізує:

- побудову поля кореляції;
- визначення числа вірних і помилкових рішень про віднесення екземплярів до того або іншого класу;
- визначення ймовірностей помилкових і вірних рішень;
- протоколювання результатів розрахунків у вигляді таблиці .

3.3. За даними навчального експерименту програмно реалізувати побудову поля кореляції ознаки X і прогнозованого параметра Y .

3.4. Задавшись одним з можливих порогових значень ознаки $X_{кл}$, для заданого порогового значення прогнозованого параметра $Y_{пр}$ виділити області рішень за класами K_1 і K_2 , а також області (ріш K_1/K_2), (ріш K_1/K_1), (ріш K_2/K_2) , (ріш K_2/K_1).

3.5. Для заданого в п.3.4 $X_{кл}$ знайти значення імовірності прийняття помилкових рішень $P_{пом}$.

3.6. Для значень $X_{кл}$ у діапазоні від X_{min} до X_{max} визначити число вірних і помилкових рішень, що полягають у віднесенні екземплярів до класів K_1 або K_2 , тобто n (ріш K_1/K_1), n (ріш K_2/K_2), n (ріш K_1/K_2), n (ріш K_2/K_1), а також загальне число рішень, прийнятих про віднесення екземплярів, відповідно, до класу K_1 і K_2 – n (ріш K_1), n (ріш K_2).

3.7. Для значень $X_{кл}$ у діапазоні від X_{min} до X_{max} знайти імовірності помилкових і вірних рішень $P(K_1/\text{ріш } K_2)$, $P(K_2/\text{ріш } K_1)$, $P(\text{ріш } K_1/K_2)$, $P(\text{ріш } K_2/K_1)$, $P_{пом}$, $P_{вірн}$, а також імовірності $P(\text{ріш } K_1)$, $P(\text{ріш } K_2)$.

3.8. Результати, отримані в п.п. 3.6 та 3.7, програмно представити у вигляді таблиці.

3.9. Представити результати, отримані у п.3.8, у графічному вигляді.

3.10. Проаналізувати отримані результати й вибрати оптимальне значення $X_{кл}^{(opt)}$, при якому імовірність прийняття вірних рішень буде максимальною.

4. ЗМІСТ ЗВІТУ

- 4.1. Сформульована мета роботи.
- 4.2. Текст програми, що реалізує етапи роботи, зазначені в п.п. 3.2, 3.6, 3.7, 3.8.
- 4.3. Роздруківки результатів роботи програми, таблиця, графіки.
- 4.4. Аналіз отриманих результатів і висновки.

КОНТРОЛЬНІ ЗАПИТАННЯ

1. Для вирішення задачі оптимальної класифікації за однією ознакою повинна бути відома двовимірна густина розподілу ознаки й прогнозованого параметру. Що є вихідною інформацією при вирішенні задачі класифікації за однією ознакою з використанням евристичного алгоритму?

2. Що є критерієм для вибору порогового значення ознаки при класифікації за однією ознакою з використанням евристичного алгоритму?

3. У чому полягає навчальний експеримент при прогнозуванні стабільності резисторів? Що при цьому є ознакою й прогнозованим параметром?

4. Охарактеризуйте імовірності прийняття помилкових і вірних рішень при класифікації за однією ознакою з використанням евристичного алгоритму.

5. Назвіть способи зменшення ризику споживача й імовірності прийняття помилкових рішень

6. Яким повинен бути об'єм навчальної вибірки для вирішення задачі класифікації? Від яких факторів він залежить?

Лабораторна робота № 2.
**ОПТИМАЛЬНА ОЦІНКА ЗНАЧЕННЯ
ПРОГНОЗОВАНОГО ПАРАМЕТРА**

Мета роботи – вивчити індивідуальне прогнозування за ознаками з оцінкою значення прогнозованого параметра із використанням теорії статистичних оцінок; для оптимального прогнозування у випадку, коли ознака й прогнозований параметр мають спільний нормальний розподіл, детально дослідити властивості оптимальної оцінки прогнозованого параметра, а також вплив на неї конкретного значення ознаки.

1. КОРОТКІ ТЕОРЕТИЧНІ ВІДОМОСТІ

У даній лабораторній роботі досліджується класичний статистичний метод, що базується на теорії статистичних оцінок і дає оптимальне розв'язання задачі індивідуального прогнозування. Для дослідження цього методу фахівець повинен мати у своєму розпорядженні багатовимірні умовні густини розподілу прогнозованого параметра та ознак.

2. ДОМАШНЄ ЗАВДАННЯ

Використовуючи конспект лекцій та рекомендовану літературу, вивчити теорію статистичного оцінювання і її використання при вирішенні задачі індивідуального прогнозування за ознаками з оцінкою значення прогнозованого параметра. Ознайомитися зі змістом і порядком виконання роботи.

3. ПОРЯДОК ВИКОНАННЯ РОБОТИ

3.1. За номером у журналі обрати варіант завдання для оптимального оцінювання прогнозованого параметра Y , коли початковий стан виробу оцінюється однією ознакою X , і спільна густина розподілу цієї ознаки й прогнозованого параметра підкоряється двовимірному нормальному закону.

Вихідними даними є наступні параметри:

- математичне сподівання ознаки M_x ;
- математичне сподівання прогнозованого параметру M_y ;
- дисперсія ознаки $D[x]$;
- дисперсія прогнозованого параметра $D[y]$;

- коефіцієнт кореляції між ознакою й прогнозованим параметром r ;
- значення ознаки $X^{(j)}$.

3.2. Написати й налагодити програму в пакеті Matlab, яка знаходить:

- спільну густину розподілу ознаки й прогнозованого параметру $W(x,y)$;
- одновимірні густини розподілу ознаки й параметру $W(x)$, $W(y)$;
- умовну густину розподілу прогнозованого параметра за умови, що ознака прийняла деяке значення $W(y/x)$;
- дисперсію похибки прогнозування $D[\tilde{y}]$;
- умовне математичне сподівання прогнозованого параметра $M[\tilde{y} / \tilde{x}]$;
- умовну дисперсію прогнозованого параметру $D[\tilde{y} / \tilde{x}]$.

3.3. Визначити залежність дисперсії похибки прогнозування від величини коефіцієнта кореляції r між ознакою й прогнозованим параметром і від дисперсії прогнозованого параметру $D[y]$. Значення дисперсії похибки прогнозування визначити для $|r| = 0; 0.1; 0.2; 0.3; 0.4; 0.5; 0.6; 0.7; 0.8; 0.9; 1.0$ і для $D[y]_1 = D[y]$, $D[y]_2 = 2 \cdot D[y]$, $D[y]_3 = 4 \cdot D[y]$.

Розраховані залежності надати у вигляді таблиці й графіка.

Проаналізувати одержані результати.

3.4. Розрахувати значення кривої безумовної густини розподілу прогнозованого параметра $W(y)$ для $y = M_y$; $M_y \pm 0.2\sigma_y$; $M_y \pm 0.4\sigma_y$; $M_y \pm 0.6\sigma_y$; $M_y \pm 0.8\sigma_y$; $M_y \pm \sigma_y$; $M_y \pm 2\sigma_y$; $M_y \pm 3\sigma_y$.

3.5. Розрахувати значення кривих густин умовного розподілу $W(y/x^{(j)})$ для двох значень коефіцієнта кореляції між ознакою й прогнозованим параметром r_1 і r_2 і чотирьох значень ознаки $X^{(1)}$, $X^{(2)}$, $X^{(3)}$, $X^{(4)}$ для $y = M[\tilde{y} / x^{(j)}]$; $M[\tilde{y} / x^{(j)}] \pm 0.2\sigma_y$; $M[\tilde{y} / x^{(j)}] \pm 0.4\sigma_y$; $M[\tilde{y} / x^{(j)}] \pm 0.6\sigma_y$; $M[\tilde{y} / x^{(j)}] \pm 0.8\sigma_y$; $M[\tilde{y} / x^{(j)}] \pm \sigma_y$; $M[\tilde{y} / x^{(j)}] \pm 2\sigma_y$; $M[\tilde{y} / x^{(j)}] \pm 3\sigma_y$.

Результати розрахунків надати у вигляді таблиці.

3.6. Показати графічно розраховані в п.п.3.4. і 3.5. густини розподілу. Проаналізувати, як змінюється розташування умовної густини відносно одновимірної безумовної густини розподілу прогнозованого параметру в залежності від коефіцієнта кореляції r і конкретного значення ознаки j -го екземпляру $X^{(j)}$. Проаналізувати, як впливає ступінь залежності ознаки й прогнозованого параметра на дисперсію умовного розподілу параметра.

3.7. Використовуючи спільну густину розподілу ознаки й прогнозованого параметру $W(x,y)$ для значення ознаки $X^{(2)}$ визначити оптимальну оцінку значення прогнозованого параметра $Y^{*(j)}_{opt}$. Нагадаємо, що в якості $Y^{*(j)}_{opt}$ прогнозованого параметра $Y^{(j)}$ береться

найбільш імовірне значення випадкової величини $\tilde{y}$ – її мода, тобто таке, при якому густина розподілу максимальна. Така оцінка має найменшу дисперсію похибки в порівнянні з усіма іншими можливими оцінками (наприклад, математичне сподівання, медіана розподілу).

Для знаходження оптимальної оцінки прогнозованого параметру використати наступні методи одномірного пошуку оптимуму:

- метод розподілу інтервалу навпіл;
- метод Пауела;
- метод Ньютона-Рафсона.

3.8. Беручи до уваги той факт, що для нормального розподілу мода умовної густини розподілу $W(y/x^{(j)})$ (найбільш імовірне значення випадкової величини $\tilde{y}$) збігається з умовним математичним сподіванням прогнозованого параметру $M[\tilde{y}/x^{(j)}]$, знайти оптимальну оцінку прогнозованого параметра, використовуючи вираз для умовного математичного сподівання (для значення ознаки $X^{(2)}$).

3.9.3 огляду на те, що знайдене в пункті 3.8. значення є оптимальною оцінкою прогнозованого параметру, оцінити точність використаних (див. п. 3.7) методів пошуку оптимуму (знайти абсолютну й відносну похибки).

4. ЗМІСТ ЗВІТУ

- 4.1. Сформульована мета роботи.
- 4.2. Текст програми, що реалізує етапи роботи зазначені в п.3.2.
- 4.3. Роздруківки результатів роботи програми, таблиці, графіки.
- 4.4. Аналіз отриманих результатів і висновки.

КОНТРОЛЬНІ ЗАПИТАННЯ

1. Яку інформацію повинен мати дослідник для вирішення задачі індивідуального прогнозування за ознаками з оцінкою значення прогнозованого параметра методами теорії статистичних оцінок?

2. У якому випадку доцільно ставити задачі прогнозування значення параметра за ознаками?

3. Який вигляд має багатовимірна густина розподілу при незалежності прогнозованого параметра від кожної з ознак?

4. Що береться в якості кількісної міри точності прогнозування?

5. Який вигляд має аналітична залежність між багатовимірною спільною густиною розподілу значень ознак і прогнозованого параметра,

умовною густиною розподілу прогнозованого параметра та густиною спільного розподілу ознак?

6. Чому дисперсія, обчислена за умовною густиною розподілу прогнозованого параметра, буде менше дисперсії, обчисленої за одновимірною густиною розподілу прогнозованого параметра? Від чого залежить розходження між цими дисперсіями?

7. Який параметр обирається в якості кількісної міри оптимальної оцінки прогнозованого параметра?

8. Проілюструйте графічно ефективність прогнозування за допомогою умовної густини в порівнянні з оцінкою параметра за безумовною густиною розподілу прогнозованого параметра?

9. Чим визначається розташування умовної густини відносно безумовної густини розподілу прогнозованого параметра?

10. У якому випадку моди умовної й безумовної густин збігаються?

11. З якими проблемами доводиться мати справу дослідникові при практичному використанні оптимальної оцінки прогнозованого параметра?

Лабораторна робота № 3. ОПТИМАЛЬНА КЛАСИФІКАЦІЯ

Мета роботи – вивчити теорію статистичної класифікації; навчитися застосовувати теорію статистичної класифікації для вирішення задач індивідуального прогнозування за ознаками.

1. КОРОТКІ ТЕОРЕТИЧНІ ВІДОМОСТІ

При індивідуальному прогнозуванні за ознаками із класифікацією задача полягає в розподілі досліджуваної сукупності виробів на класи й немає необхідності в оцінці конкретного значення прогнозованого параметра. У більшості практичних застосувань цього методу число класів дорівнює двом. Так буває, наприклад, коли досліджувану сукупність необхідно за заданим правилом розділити на клас придатних і клас дефектних виробів.

У цій лабораторній роботі зазначена задача вирішується методами теорії статистичної класифікації, для чого необхідно мати у своєму розпорядженні умовні багатовимірні густини розподілу ознак для кожного класу. Задача полягає у знаходженні способу прийняття оптимального рішення про приналежність екземпляра, що перевіряється, до того або іншого класу в умовах невизначеності, тобто в умовах дії випадкових факторів, що маскують зв'язок між ознаками й класом екземпляра. Умовимося, що екземпляр, який перевіряється, належить до класу K_1 , якщо значення прогнозованого параметра на момент часу прогнозування буде більше деякого порогового значення; будемо вважати такі вироби придатними. У протилежному випадку екземпляр належить до класу K_2 (дефектних).

2. ДОМАШНЄ ЗАВДАННЯ

Використовуючи конспект лекцій та рекомендовану літературу, вивчити теорію статистичної класифікації і її застосування для вирішення задачі індивідуального прогнозування. Ознайомитися зі змістом і порядком виконання роботи.

3. ПОРЯДОК ВИКОНАННЯ РОБОТИ

3.1. За номером у журналі обрати варіант завдання для вирішення задачі прогнозування методом оптимальної класифікації, коли число

класів дорівнює двом (клас придатних і клас дефектних виробів). Початковий стан виробу характеризується однією ознакою, спільна густина розподілу значень ознаки та прогнозованого параметра задається двовимірним нормальним законом розподілу.

Вихідними даними є наступні параметри:

- порогове значення прогнозованого параметра $Y_{пор}$;
- математичне сподівання ознаки M_x ;
- математичне сподівання прогнозованого параметра M_y ;
- дисперсія ознаки $D[x]$;
- дисперсія прогнозованого параметра $D[y]$;
- коефіцієнт кореляції між ознакою й прогнозованим параметром r ;
- значення ознаки $X^{(j)}$;
- втрати, пов'язані з помилковими рішеннями: ціна перейменування

класу екземпляра з K_1 у K_2 ($\Pi_{1 \rightarrow 2}$) і ціна перейменування класу екземпляра з K_2 у K_1 ($\Pi_{2 \rightarrow 1}$).

3.2. Написати й налагодити програму в пакеті Matlab, яка реалізує процедуру оптимальної класифікації. Основні кроки алгоритму розв'язання задачі індивідуального прогнозування за ознаками методом оптимальної класифікації наведено нижче.

3.3. Визначити апіорні імовірності приналежності екземпляра виробу до класу K_1 або K_2 : $P(K_1)$ і $P(K_2)$.

3.4. Визначити умовне математичне сподівання ознаки за умови, що екземпляр належить до класу K_1 – $M[\tilde{x} / K_1]$ і умовне математичне сподівання ознаки за умови, що екземпляр належить до класу K_2 – $M[\tilde{x} / K_2]$.

3.5. Розрахувати значення безумовної густини розподілу ознаки $W(x)$ для значень ознаки $x=M_x$; $M_x \pm 0.2\sigma_x$; $M_x \pm 0.4\sigma_x$; $M_x \pm 0.6\sigma_x$; $M_x \pm 0.8\sigma_x$; $M_x \pm \sigma_x$; $M_x \pm 2\sigma_x$; $M_x \pm 3\sigma_x$. Результати розрахунків представити у вигляді таблиці.

3.6. Розрахувати значення умовних густин розподілу ознаки за умови, що екземпляр належить до класу K_1 або K_2 : $W(x / K_1)$ і $W(x / K_2)$. Значення умовних густин визначити для значень ознаки $x=M[\tilde{x} / K_1]$; $M[\tilde{x} / K_1] \pm 0.2\sigma_x$; $M[\tilde{x} / K_1] \pm 0.4\sigma_x$; $M[\tilde{x} / K_1] \pm 0.6\sigma_x$; $M[\tilde{x} / K_1] \pm 0.8\sigma_x$; $M[\tilde{x} / K_1] \pm \sigma_x$; $M[\tilde{x} / K_1] \pm 2\sigma_x$; $M[\tilde{x} / K_1] \pm 3\sigma_x$. Результати розрахунків представити у вигляді таблиці.

3.7. Знайти порогове значення ознаки $X_{кл}$.

3.8. Визначити апіорні імовірності ухвалення рішення про віднесення екземпляра до класу K_1 або K_2 : $P(\text{ріш } K_1)$ і $P(\text{ріш } K_2)$.

3.9. Знайти умовне математичне сподівання прогнозованого параметра при умовах, що ухвалено рішення про віднесення екземпляра відповідно до класу K_1 – $M[\tilde{y} / \text{ріш } K_1]$ і класу K_2 – $M[\tilde{y} / \text{ріш } K_2]$.

3.10. Визначити значення безумовної густини прогнозованого параметра $W(y)$ для значень $y=M_y$; $M_y \pm 0.2\sigma_y$; $M_y \pm 0.4\sigma_y$; $M_y \pm 0.6\sigma_y$; $M_y \pm 0.8\sigma_y$; $M_y \pm \sigma_y$; $M_y \pm 2\sigma_y$; $M_y \pm 3\sigma_y$. Результати розрахунків представити у вигляді таблиці.

3.11. Розрахувати умовні розподіли прогнозованого параметра за умови, що приймається рішення про віднесення екземпляра до класу K_1 або K_2 : $W(y/\text{ріш } K_1)$ і $W(y/\text{ріш } K_2)$. Значення умовних густин визначити для наступних значень прогнозованого параметра $y=M[\tilde{y}/\text{ріш } K_i]$;
 $M[\tilde{y}/\text{ріш } K_i] \pm 0.2\sigma_y$; $M[\tilde{y}/\text{ріш } K_i] \pm 0.4\sigma_y$; $M[\tilde{y}/\text{ріш } K_i] \pm 0.6\sigma_y$;
 $M[\tilde{y}/\text{ріш } K_i] \pm 0.8\sigma_y$; $M[\tilde{y}/\text{ріш } K_i] \pm \sigma_y$; $M[\tilde{y}/\text{ріш } K_i] \pm 2\sigma_y$;
 $M[\tilde{y}/\text{ріш } K_i] \pm 3\sigma_y$. Результати розрахунків представити у вигляді таблиці.

3.12. Знайти відношення правдоподібності $\lambda(x)$ для чотирьох значень ознаки $X^{(1)}$, $X^{(2)}$, $X^{(3)}$, $X^{(4)}$.

3.13. Розрахувати порогове значення відношення правдоподібності й прийняти рішення про віднесення кожного із чотирьох екземплярів виробу (див. п. 3.11) до того або іншого класу .

3.14. Використовуючи порогове значення $X_{кр}$, зробити класифікацію виробів, значення ознак яких відповідають $X^{(1)}$, $X^{(2)}$, $X^{(3)}$, $X^{(4)}$. Порівняти результати класифікації в п.3.12 і п. 3.13.

3.15. Розрахувати значення модуля різниці умовних математичних сподівань, знайдених у п.3.9. При збільшенні зазначеного модуля різниці ризику виробника $P(K_1/\text{ріш } K_2)$ та споживача $P(K_2/\text{ріш } K_1)$ зменшуються. Значення модуля різниці умовних математичних сподівань визначити для наступних значень тісноти зв'язку між ознакою й прогнозованим параметром: r ; $0.8r$; $0.6r$; $0.4r$; $0.2r$; $0.1r$. Результати розрахунків представити у вигляді таблиці й графіка.

3.16. Використовуючи результати розрахунків, отримані в п.п. 3.4, 3.5, 3.6, 3.9, 3.10, 3.11, в системі координат ознаки й прогнозованого параметра надати взаємне розташування відповідних безумовних і умовних густин розподілу. В разі необхідності для більшої наочності знайти значення даних густин розподілу для додаткових значень ознаки й прогнозованого параметра, не зазначених вище у відповідних розділах завдання.

3.17. Відзначити в системі координат (див. п. 3.16) порогове значення прогнозованого параметра й порогове значення ознаки. Дати геометричне тлумачення наступним ймовірностям: $P(K_1)$, $P(K_2)$, $P(\text{ріш } K_1)$, $P(\text{ріш } K_2)$, $P(K_1/\text{ріш } K_2)$, $P(K_2/\text{ріш } K_1)$, $P(\text{ріш } K_1/K_2)$, $P(\text{ріш } K_2/\text{ріш } K_1)$.

3.18. Показати в системі координат (див. п.п. 3.16, 3.17) області, які відносяться до екземплярів, що фактично належать до класів K_1 і K_2 , а

також області, які належать екземплярам, щодо яких ухвалене рішення про віднесення до указаних класів.

3.19. Задане порогове значення прогнозованого параметра й знайдене порогове значення ознаки ділять площину ознаки та параметра на чотири області. Імовірність того, що екземпляр буде характеризуватися визначеними значеннями ознаки й прогнозованого параметра (у цьому випадку екземпляр буде належати одній із чотирьох областей) відповідає ймовірності спільного настану двох подій – екземпляр фактично належить до визначеного класу (K_1 або K_2) і ухвалене рішення про його віднесення до якого-небудь класу (K_1 або K_2). Співвіднесіть зазначені ймовірності з кожною із чотирьох областей.

4. ЗМІСТ ЗВІТУ

- 4.1. Сформульована мета роботи.
- 4.2. Алгоритм і програма процедури вирішення задачі індивідуального прогнозування методом оптимальної класифікації.
- 4.3. Роздруківки результатів роботи програми.
- 4.4. Рисунок, що представляє взаємне розташування безумовних і умовних густин розподілу ознаки й прогнозованого параметра (див. п.п. 3.16, 3.17, 3.18, 3.19)
- 4.5. Аналіз отриманих результатів і висновки.

КОНТРОЛЬНІ ЗАПИТАННЯ

1. Яку інформацію повинен мати дослідник для вирішення задачі індивідуального прогнозування методами теорії статистичної класифікації?
2. Як можна визначити апріорні ймовірності приналежності даного екземпляра до визначеного класу?
3. У чому полягає задача індивідуального прогнозування методами теорії статистичної класифікації? У якому випадку доцільно ставити цю задачу?
4. Якими співвідношеннями пов'язані між собою умовні спільні густини розподілу ознак за умови, що екземпляр належить до даного класу, і спільна густина розподілу ознак і прогнозованого параметра?
5. Дайте характеристику ризиків виробника й споживача та інших ймовірностей прийняття помилкових рішень при оптимальній класифікації.
6. У чому полягає критерій Байєса?

7. Яким чином можна мінімізувати середній ризик (середні втрати) при розпізнаванні?

8. Що називається відношенням правдоподібності?

9. Сформулюйте алгоритм оптимальної класифікації.

10. Як параметри спільного розподілу впливають на імовірності прийняття помилкових рішень при оптимальній класифікації?

11. Зобразіть графічно взаємне розташування безумовних і умовних густин розподілу для індивідуального прогнозування методом оптимальної класифікації за однією ознакою. За допомогою цього малюнка дайте графічну інтерпретацію ймовірностей прийняття правильних і помилкових рішень.

Варіанти завдань для лабораторних робіт №2 та №3

Варіант №1	Варіант №2	Варіант №3	Варіант №4	Варіант №5
$M_x=5.0$	$M_x=4.0$	$M_x=4.0$	$M_x=10.0$	$M_x=10.0$
$M_y=15.0$	$M_y=12.0$	$M_y=20.0$	$M_y=12.0$	$M_y=15.0$
$\sigma[x]=0.5$	$\sigma[x]=0.4$	$\sigma[x]=0.6$	$\sigma[x]=0.8$	$\sigma[x]=1.0$
$\sigma[y]=1.0$	$\sigma[y]=1.6$	$\sigma[y]=1.2$	$\sigma[y]=1.6$	$\sigma[y]=2.0$
$R_1=-0.5$	$R_1=-0.4$	$R_1=-0.5$	$R_1=-0.5$	$R_1=0.7$
$R_2=-0.8$	$R_2=-0.9$	$R_2=-0.8$	$R_2=-0.7$	$R_2=0.9$
$X(1)=2.0$	$X(1)=6.0$	$X(1)=5.0$	$X(1)=2.0$	$X(1)=10.5$
$X(2)=4.0$	$X(2)=5.0$	$X(2)=4.0$	$X(2)=3.0$	$X(2)=11.0$
$X(3)=6.0$	$X(3)=4.0$	$X(3)=3.0$	$X(3)=4.0$	$X(3)=10.0$
$X(4)=8.0$	$X(4)=3.0$	$X(4)=2.0$	$X(4)=5.0$	$X(4)=9.0$

Варіант №6	Варіант №7	Варіант №8	Варіант №9	Варіант №10
$M_x=6.0$	$M_x=10.0$	$M_x=4.0$	$M_x=6.0$	$M_x=10.0$
$M_y=13.0$	$M_y=10.0$	$M_y=15.0$	$M_y=13.0$	$M_y=15.0$
$\sigma[x]=0.5$	$\sigma[x]=0.3$	$\sigma[x]=0.5$	$\sigma[x]=0.5$	$\sigma[x]=0.5$
$\sigma[y]=1.2$	$\sigma[y]=0.8$	$\sigma[y]=1.5$	$\sigma[y]=1.2$	$\sigma[y]=1.5$
$R_1=-0.7$	$R_1=-0.7$	$R_1=-0.7$	$R_1=-0.7$	$R_1=-0.7$
$R_2=-0.9$	$R_2=-0.9$	$R_2=-0.9$	$R_2=-0.9$	$R_2=-0.9$
$X(1)=6.5$	$X(1)=10.9$	$X(1)=4.5$	$X(1)=6.5$	$X(1)=11.5$
$X(2)=7.0$	$X(2)=10.6$	$X(2)=5.0$	$X(2)=7.0$	$X(2)=11.0$
$X(3)=6.0$	$X(3)=10.0$	$X(3)=4.0$	$X(3)=6.0$	$X(3)=10.0$
$X(4)=5.0$	$X(4)=9.4$	$X(4)=3.0$	$X(4)=5.0$	$X(4)=9.0$

Варіант №11	Варіант №12	Варіант №13	Варіант №14	Варіант №15
$M_x=55.0$	$M_x=5.0$	$M_x=21.0$	$M_x=15.0$	$M_x=10.0$
$M_y=2.0$	$M_y=16.0$	$M_y=17.0$	$M_y=25.0$	$M_y=15.0$
$\sigma[x]=0.3$	$\sigma[x]=1.2$	$\sigma[x]=0.5$	$\sigma[x]=1.0$	$\sigma[x]=0.5$
$\sigma[y]=1.8$	$\sigma[y]=0.8$	$\sigma[y]=1.1$	$\sigma[y]=0.5$	$\sigma[y]=1.0$
$R_1=0.02$	$R_1=-0.5$	$R_1=-0.4$	$R_1=0.5$	$R_1=0.5$
$R_2=-0.9$	$R_2=-0.4$	$R_2=-0.8$	$R_2=-0.9$	$R_2=-0.2$
$X(1)=55.9$	$X(1)=8.6$	$X(1)=22.5$	$X(1)=16.0$	$X(1)=10.5$
$X(2)=55.6$	$X(2)=7.4$	$X(2)=20.0$	$X(2)=15.0$	$X(2)=11.0$
$X(3)=55.0$	$X(3)=5.0$	$X(3)=21.0$	$X(3)=14.0$	$X(3)=12.0$
$X(4)=54.4$	$X(4)=2.6$	$X(4)=20.0$	$X(4)=13.0$	$X(4)=13.0$

Варіант №16	Варіант №17	Варіант №18	Варіант №19	Варіант №20
-------------	-------------	-------------	-------------	-------------

Mx=15.0	Mx=10.0	Mx=9.0	Mx=10.0	Mx=12.0
My=10.0	My=20.0	My=13.0	My=20.0	My=25.0
$\sigma[x]=0.4$	$\sigma[x]=0.6$	$\sigma[x]=0.5$	$\sigma[x]=0.7$	$\sigma[x]=0.6$
$\sigma[y]=0.9$	$\sigma[y]=0.8$	$\sigma[y]=0.8$	$\sigma[y]=0.9$	$\sigma[y]=0.8$
R1=-0.8	R1=-0.6	R1=-0.4	R1=-0.7	R1=-0.7
R2=-0.5	R2=-0.7	R2=-0.5	R2=-0.9	R2=-0.8
X(1)=16.2	X(1)=11.8	X(1)=10.5	X(1)=12.1	X(1)=13.8
X(2)=15.8	X(2)=11.2	X(2)=10.0	X(2)=11.4	X(2)=13.2
X(3)=15.0	X(3)=10.0	X(3)=9.0	X(3)=10.0	X(3)=12.0
X(4)=14.2	X(4)=8.8	X(4)=8.0	X(4)=8.6	X(4)=10.8

Варіант №21	Варіант №22	Варіант №23	Варіант №24	Варіант №25
-------------	-------------	-------------	-------------	-------------

Mx=15.0	Mx=15.0	Mx=10.0	Mx=5.0	Mx=5.0
My=10.0	My=10.0	My=20.0	My=20.0	My=20.0
$\sigma[x]=0.4$	$\sigma[x]=0.4$	$\sigma[x]=0.7$	$\sigma[x]=0.3$	$\sigma[x]=0.5$
$\sigma[y]=0.9$	$\sigma[y]=0.9$	$\sigma[y]=0.9$	$\sigma[y]=0.8$	$\sigma[y]=0.9$
R1=-0.8	R1=-0.8	R1=-0.7	R1=-0.7	R1=-0.7
R2=-0.5	R2=-0.5	R2=-0.9	R2=-0.9	R2=-0.9
X(1)=16.2	X(1)=16.2	X(1)=12.1	X(1)=5.9	X(1)=6.5
X(2)=15.8	X(2)=15.8	X(2)=11.4	X(2)=5.6	X(2)=6.0
X(3)=15.0	X(3)=15.0	X(3)=10.0	X(3)=5.0	X(3)=5.0
X(4)=14.2	X(4)=14.2	X(4)=8.6	X(4)=4.7	X(4)=4.0

Завдання до лабораторної роботи №3

- 1) $\Pi_{1 \rightarrow 2} = \Pi_{2 \rightarrow 1}$
- 2) $r = -0.9$
- 3) $Y_{\text{гр}} = My - \sigma[x]$

Лабораторна робота № 4.
МЕТОД ДИСКРИМІНАНТНИХ ФУНКЦІЙ

Мета роботи – вивчити метод індивідуального прогнозування за допомогою дискримінантних функцій для вирішення задач класифікації.

1. КОРОТКІ ТЕОРЕТИЧНІ ВІДОМОСТІ

Якщо представити кожний екземпляр, що характеризується значеннями K ознак, деякою точкою в K -вимірному просторі ознак, то задача індивідуального прогнозування із класифікацією на основі теорії розпізнавання образів буде полягати в розподілі цього K -вимірного простору ознак за допомогою якоїсь $(K-1)$ -вимірної поверхні на дві області, що відповідають класам K_1 і K_2 . Ця поділяюча поверхня в загальному випадку задається рівнянням $g(x_1, x_2, \dots, x_k) = \text{const}$. Функція $g(x_1, x_2, \dots, x_k)$ називається дискримінантною. Для розпізнавання класу якого-небудь екземпляра досить за виміряними значеннями його ознак визначити, у якій області K -вимірного простору перебуває точка, координати якої задаються цими значеннями.

У даній лабораторній роботі задача класифікації вирішується для випадку, коли поверхнею, що розділяє простір на дві області, є гіперплощина.

2. ДОМАШНЄ ЗАВДАННЯ

Використовуючи конспект лекцій та рекомендовану літературу, вивчити метод дискримінантних функцій.

Ознайомитися зі змістом і порядком виконання роботи.

3. ПОРЯДОК ВИКОНАННЯ РОБОТИ

3.1. {Одержати варіант завдання у викладача.}

Вихідними даними є:

- масив даних навчального експерименту (для кожного екземпляра навчальної вибірки відомі значення ознак і фактичний клас);
- значення ознак екземплярів, що не входять у навчальну вибірку.

3.2. Написати й налагодити програму в пакеті Matlab, що реалізує розв'язання задачі індивідуального прогнозування із класифікацією методом дискримінантних функцій. Опис алгоритму рішення даної задачі наведено нижче.

3.3. За даними навчального експерименту знайти оцінки умовних математичних сподівань і дисперсії кожної ознаки за умови, що екземпляр належить до класу K_1 : $M^*[\tilde{x}_i/K_1]$ і $D^*[\tilde{x}_i/K_1]$.

3.4. Розрахувати оцінки умовних математичних сподівань і дисперсії кожної ознаки за умови, що екземпляр належить до класу K_2 : $M^*[\tilde{x}_i/K_2]$ і $D^*[\tilde{x}_i/K_2]$.

3.5. Визначити значення коефіцієнтів парної кореляції між ознаками за умови, що екземпляр належить, відповідно, до класу K_1 або K_2 : $r^*[x_i, x_l/K_1]$ і $r^*[x_i, x_l/K_2]$ ($i, l=1, K$) ($i \neq l$).

3.6. Зробити статистичну оцінку значимості коефіцієнтів парної кореляції, одержаних у п.3.5. Ознаки, для яких коефіцієнти кореляції є статистично незначущими, вважати некоррельованими між собою.

3.7. Знайти оцінки умовних математичних сподівань дискримінантної функції за умови, що екземпляр належить до класу K_1 або K_2 : $M^*[G/K_1]$ і $M^*[G/K_2]$. При розрахунку умовних математичних сподівань використати теореми про математичне сподівання.

3.8. Розрахувати оцінки умовних дисперсій дискримінантної функції за умови, що екземпляр належить відповідно до класу K_1 або K_2 : $D^*[G/K_1]$ і $D^*[G/K_2]$. При розрахунку умовних дисперсій використати теореми про дисперсію.

3.9. Підставити у вираз для критерію оптимізації, що визначає найкращий “нахил” поділяючої гіперплощини в просторі ознак, оцінки умовних математичних сподівань і дисперсій дискримінантної функції (див. п.п. 3.7, 3.8). Знайти частинні похідні критерію оптимізації за коефіцієнтами дискримінантної функції та дорівняти їх до нуля. Вирішити одержану систему K алгебраїчних рівнянь із K невідомими коефіцієнтами й визначити оптимальні оцінки коефіцієнтів дискримінантної функції.

3.10. Знайдені на попередньому етапі оцінки коефіцієнтів дискримінантної функції визначають оптимальний нахил гіперплощини в просторі ознак. Далі необхідно знайти порогове значення Pg для дискримінантної функції, що задає найкраще розташування поділяючої гіперплощини. Очевидно, повинне виконуватися одне із двох умов: $M^*[G/K_1] > Pg > M^*[G/K_2]$ або $M^*[G/K_1] < Pg < M^*[G/K_2]$. При змінюванні порога будуть змінюватися імовірності помилкових рішень. Оптимальна величина порога може бути знайдена шляхом послідовних прорахунків імовірності помилкових рішень за даними навчального експерименту для різних Pg і вибором такого з них, при якому імовірність помилкових рішень буде мінімальною. Опис процедури знаходження оптимальної величини порога наведено нижче в п.п. 3.11, 3.12, 3.13, 3.14.

3.11. Для кожного j -го екземпляра навчальної вибірки обчислити дискримінантну функцію $G^{(j)} = g(x_1^{(j)}, x_2^{(j)}, \dots, x_k^{(j)})$.

3.12. Оскільки задача визначення оптимальної величини порога є типовою задачею оптимізації (функцією, що підлягає мінімізації, є

ймовірність помилкових рішень), використати для її розв'язання один із заданих викладачем методів одномірного пошуку оптимуму.

3.13. Для обраних відповідно до заданої стратегії пошуку (див. п. 3.12) значень порога Pg і розрахованих значень дискримінантної функції (див. п. 3.11) визначити прогнозований клас екземплярів навчальної вибірки. Якщо має місце нерівність $M*[G/K_1] > M*[G/K_2]$ і при цьому $G^{(j)} > Pg$, то приймається рішення про віднесення j -го екземпляра до класу K_1 , якщо $G^{(j)} < Pg$, приймається рішення про віднесення його до класу K_2 . Якщо $M*[G/K_1] < M*[G/K_2]$, умови, при яких приймаються рішення про віднесення екземплярів до того або іншого класу, змінюються на протилежні.

3.14. Порівнюючи для екземплярів навчальної вибірки прогнозовані та фактичні класи, оцінити імовірності помилкових рішень для заданих значень порога. Визначити оптимальну величину порога.

3.15. Зменшення ймовірності помилкових рішень при класифікації може бути досягнуте шляхом відбору більш інформативних ознак. Використати для цієї мети заданий викладачем алгоритм вибору інформативних ознак.

3.16. Для отриманого оптимального набору інформативних ознак перерахувати оцінки умовних математичних сподівань і дисперсій дискримінантної функції (п.п.3.7, 3.8), оптимальні оцінки коефіцієнтів дискримінантної функції (п.3.9) і оптимальну величину порога (п.п.3.11, 3.12, 3.13, 3.14).

3.17. Оцінити імовірність помилкових рішень для знайденого оператора прогнозування.

3.18. Оцінити клас екземплярів, що не входять у навчальну вибірку.

4. ЗМІСТ ЗВІТУ

- 4.1. Сформульована мета роботи.
- 4.2. Алгоритм і програма вирішення задачі індивідуального прогнозування із класифікацією методом дискримінантних функцій.
- 4.3. Роздруківки результатів роботи програми.
- 4.4. Аналіз отриманих результатів і висновки.

КОНТРОЛЬНІ ЗАПИТАННЯ

1. У чому полягає задача індивідуального прогнозування із класифікацією методом дискримінантних функцій?
2. Якою є розмірність коефіцієнтів дискримінантної функції?

3. Напишіть вираз для умовних математичних сподівань і умовних дисперсій ознак.

4. Що є критерієм оптимізації при знаходженні оцінок коефіцієнтів дискримінантної функції?

5. З яких міркувань вибирається порогове значення для дискримінантної функції?

6. Як прогнозується клас об'єктів, що не брали участь у навчальному експерименті, методом дискримінантних функцій?

7. Які дії варто вживати в тому випадку, якщо зміною порогового значення для дискримінантної функції не вдається досягти прийнятного значення ймовірності помилкових рішень?

8. Назвіть переваги й недоліки методу дискримінантних функцій.

9. При яких умовах метод дискримінантних функцій є окремим випадком оптимальної класифікації?

10. При яких характеристиках розподілюваних класів найбільш ефективно застосування методу дискримінантних функцій?

11. Значення яких параметрів мають бути оцінені за даними навчального експерименту при класифікації за двома ознаками методом дискримінантних функцій?

12. Проілюструйте малюнком, як зміна порогового значення для дискримінантної функції при класифікації за двома ознаками впливає на ймовірності помилкових рішень.

Лабораторна робота № 5.
МЕТОД ПОТЕНЦІАЛЬНИХ ФУНКЦІЙ

Мета роботи – вивчити методику вирішення задач класифікації з використанням потенціальних функцій.

1. КОРОТКІ ТЕОРЕТИЧНІ ВІДОМОСТІ

Метод потенціальних функцій є одним з найпоширеніших серед методів індивідуального прогнозування за ознаками. Зазвичай його застосовують, коли класи мають взаємопроникнення, тобто є важко розділними, і в цьому випадку він ефективніше методу дискримінантних функцій. Можливості нелінійного перетворення простору ознак, які закладені в методі потенційних функцій і які порівняно легко реалізуються, роблять його більш гнучким у порівнянні з іншими методами й дозволяє у більшості випадків досягти задовільних результатів. Однак обчислювальні процедури тут складніше, ніж у методі дискримінантних функцій.

2. ДОМАШНЄ ЗАВДАННЯ

Використовуючи конспект лекцій та рекомендовану літературу, вивчити метод потенціальних функцій. Ознайомитися зі змістом і порядком виконання роботи.

3. ПОРЯДОК ВИКОНАННЯ РОБОТИ

3.1. {Одержати варіант завдання у викладача.}

Вихідними даними є:

- масив даних навчального експерименту (для кожного екземпляра навчальної вибірки відомі значення ознак і фактичний клас);
- значення ознак екземплярів, що не входять у навчальну вибірку;
- порогове значення сумарного потенціалу;
- коефіцієнти, що визначають міру нелінійного перетворення простору ознак;
- вагові множники, що характеризують порівняльний вплив інформативних ознак.

3.2. Написати й налагодити програму в пакеті Matlab, що реалізує процедуру класифікації з використанням методу потенціальних функцій. Схема алгоритму методу потенціальних функцій представлена нижче.

3.3. За даними навчальної вибірки обчислити оцінки математичного сподівання й дисперсії кожної ознаки.

3.4. Визначити нормовані значення ознак.

3.5. Розрахувати узагальнену відстань для j -го й l -го екземплярів ($j, l=1, n$), ($j \neq l$), де n -число екземплярів навчальної вибірки.

3.6. Знайти потенціал кожного j -го екземпляра, який "наводиться" на нього кожним l -им екземпляром.

3.7. Обчислити сумарний потенціал для кожного екземпляра, що належить до класу K_1 і до класу K_2 .

3.8. Для заданих значень порога, аналізуючи значення сумарних потенціалів, прийняти рішення про віднесення кожного з екземплярів навчальної вибірки до того або іншого класу.

3.9. Результати, отримані в п.п.3.4, 3.7, 3.8, надати у вигляді таблиці. У цій же таблиці навести виміряні (ненормовані) значення ознак екземплярів навчальної вибірки й фактичний клас екземплярів.

3.10. Для кількісної оцінки результатів навчання для заданих значень порога, обчислити кількість рішень n (ріш K_1), n (ріш K_2), $n(K_2/\text{ріш } K_1)$, $n(K_1/\text{ріш } K_2)$ і імовірності $P(\text{ріш } K_1)$, $P(\text{ріш } K_2)$, $P(K_2/\text{ріш } K_1)$, $P(K_1/\text{ріш } K_2)$, $P_{\text{пом}}$. Результати розрахунків надати у вигляді таблиці та графіка.

3.11. За заданим критерієм, що характеризує імовірність помилкових рішень (ризиком споживача, ризиком виробника або $P_{\text{пом}}$), зробити вибір порога.

3.12. Оскільки обране в п.3.11. значення порога є тільки кращим із заданих, питання про знаходження його оптимального значення залишається відкритим.

Крім варіювання величини порога, оптимізація оператора прогнозування може бути досягнута підбором коефіцієнтів у виразі для потенціалів, визначених у п.3.6., тобто шляхом зміни ступеня нелінійності перетворення простору ознак. Крім того, якість прогнозування можна підвищити введенням вагових множників μ_i ($i=1, K$); ($\sum \mu_i=1$), (де K - кількість ознак) у виразі для узагальненої відстані (див.п.3.5.). Значення вагових множників підбираються таким чином, щоб підкреслити вплив більш інформативних ознак.

Із сказаного вище випливає, що задача визначення оптимального значення порога, оптимальних значень коефіцієнтів у виразі для потенціалів і оптимальних вагових множників є задачею багатомірної оптимізації. Використати для її вирішення один із заданих викладачем методів оптимізації.

3.13. Оцінити імовірність помилкових рішень для знайденого в п. 3.12. оператора прогнозування.

3.14. Оцінити клас екземплярів, що не входять у навчальну вибірку.

4. ЗМІСТ ЗВІТУ

- 4.1. Сформульована мета роботи.
- 4.2. Алгоритм і програма вирішення задачі індивідуального прогнозування із класифікацією методом потенційних функцій.
- 4.3. Роздруківки результатів роботи програми, таблиці, графіки.
- 4.4. Аналіз отриманих результатів і висновки.

КОНТРОЛЬНІ ЗАПИТАННЯ

1. У чому полягає суть індивідуального прогнозування методом потенціальних функцій?
2. З якою метою застосовується нормування ознак?
3. Яким чином визначається одинична відстань між значеннями однієї ознаки для двох екземплярів?
4. Яким чином визначається узагальнена відстань по всіх ознаках для двох екземплярів?
5. Напишіть вираження для потенціалу екземпляра виробу, що наводиться на нього іншим екземпляром.
6. Напишіть вираз для сумарного потенціалу.
7. Яким чином перевіряється придатність знайденого в процесі навчання оператора прогнозування для класифікації екземплярів, що не використовувалися в навчальному експерименті?
8. Яким чином виробляється оцінка класу методом потенціальних функцій?
9. Назвіть способи оптимізації оператора прогнозування при класифікації методом потенціальних функцій.
10. Як впливає величина порога на якість прогнозування?
11. Назвіть переваги та недоліки методу потенціальних функцій.

Завдання для лабораторних робіт №4 та №5

Таблиця 2 – Навчальна вибірка

Номер прибору	Ознаки									Надійність прибору
	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	
1	1.64	1.20	76.8	71.2	48.0	4.05	5.35	20.5	1.85	0.5
2	1.72	1.20	79.4	73.7	51.0	3.55	5.20	32.0	2.30	1
3	1.36	1.20	85.5	79.1	37.0	3.90	4.95	20.5	1.85	1
4	1.48	1.36	82.6	76.5	37.0	3.55	4.75	20.5	1.85	0.5
5	2.28	1.92	91.0	84.3	45.0	3.25	4.45	20.5	1.85	1
6	2.08	1.84	88.1	81.6	43.0	3.05	4.35	20.5	1.85	1
7	1.64	1.20	78.8	73.0	42.0	3.70	4.85	35.0	1.85	0.5
8	1.52	1.16	82.6	76.5	50.0	3.20	4.80	18.5	1.85	1
9	2.72	1.90	52.1	76.0	57.0	2.75	3.85	18.0	1.85	1
10	1.65	1.05	76.2	65.0	43.0	3.55	5.40	18.5	1.85	0.5
11	1.88	1.52	97.6	59.3	46.0	2.90	4.00	18.0	1.85	1
12	2.88	2.40	70.6	83.0	43.0	3.25	4.35	22.5	1.85	1
13	1.72	1.24	82.7	76.5	43.0	4.10	5.15	22.5	1.85	0
14	2.52	2.16	93.8	86.3	45.0	3.45	4.60	20.5	1.85	1
15	2.20	1.68	91.2	84.3	36.0	3.55	4.55	20.5	1.85	0.5
16	2.24	1.76	88.3	81.7	45.0	3.25	4.55	20.5	1.85	1
17	2.56	2.08	78.8	73.0	40.0	3.10	4.85	11.5	1.85	0
18	1.38	1.16	89.5	79.1	49.0	3.55	4.65	21.0	1.85	1
19	2.52	1.88	90.9	84.2	45.0	3.25	4.65	19.5	2.60	1
20	2.12	1.52	89.6	83.0	45.0	3.25	4.65	23.0	2.30	0
21	2.28	1.60	105	79.6	51.0	3.35	4.70	22.5	1.55	1
22	2.56	2.24	87.5	81.0	49.0	3.25	4.25	22.0	1.85	1
23	1.68	1.26	91.0	84.2	40.0	3.35	4.60	32.0	2.00	0
24	2.70	1.95	82.0	76.0	51.0	3.15	3.90	11.5	1.85	1

Надійність прибору:

0, 0.5 – прибор негідний

1 – прибор гідний

Таблиця 3 –Тестова вибірка (дефектні прибори № 3, 12, 14, 17, 19)

Номер прибору	Ознаки								
	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉
1	2.76	2.44	81.6	79.0	49.0	2.97	4.25	4.00	2.15
2	2.64	2.28	83.0	76.4	43.0	3.00	3.77	39.0	2.0
3	2.64	2.04	81.6	73.8	40.0	3.50	4.20	39.0	2.67
4	2.56	2.44	81.6	79.0	45.0	2.90	3.86	39.0	2.67
5	2.80	2.52	84.3	79.0	50.0	2.95	4.05	39.0	2.67
6	2.18	2.06	79.0	79.0	43.0	3.10	4.15	39.0	2.67
7	2.04	2.08	81.6	79.0	48.0	3.10	3.95	21.5	2.30
8	2.56	2.40	84.3	79.0	44.0	3.53	3.67	21.0	1.80
9	2.58	2.24	81.6	79.0	42.0	3.02	4.25	49.5	2.34
10	2.32	2.30	81.6	80.4	44.0	3.20	4.15	21.0	1.80
11	2.44	2.12	81.6	73.8	43.5	2.75	4.15	21.0	1.80
12	2.20	2.12	81.4	74.4	36.0	3.33	4.10	21.5	2.30
13	2.22	1.88	79.0	68.4	43.0	3.36	4.40	26.0	2.70
14	2.34	1.96	81.6	73.8	37.0	3.23	4.60	21.0	2.00
15	2.36	1.84	77.8	83.8	47.0	3.35	4.60	30.0	2.45
16	2.30	1.96	81.6	71.9	36.0	3.25	4.60	26.0	2.70
17	2.24	1.94	82.9	73.8	38.0	3.50	4.43	21.5	2.30
18	2.68	2.52	84.0	79.0	50.0	2.75	3.95	39.0	2.67
19	2.08	1.96	83.0	76.4	46.0	3.15	4.35	39.0	2.00

Завдання до лабораторної роботи №5

ЛІТЕРАТУРА

1. Основы технической диагностики / П.П. Пархоменко, В.В. Карибский, Е.С. Согомонян, В.Ф. Халчев. – М.: Энергия, 1976. – 462 с.
2. Пархоменко П.П., Согомонян Е.С. Основы технической диагностики. – М.: Энергия, 1981. – 320 с.
3. Мозгалеvский А.В., Гаскаров Д.В. Техническая диагностика. – М.: Высшая школа, 1975. – 207 с.
4. Мозгалеvский А.В., Койда А.Н. Вопросы проектирования систем диагностирования. – Л.: Энергоатомиздат, 1985. – 112 с.
5. Мозгалеvский А.В., Волинский В.И., Гаскаров Д.В. Техническая диагностика судовой автоматики. – Л.: Судостроение, 1972.
6. Сотсков Б.С. Основы теории и расчета надежности элементов и устройств автоматики и вычислительной техники. – М.: Высшая школа, 1970. – 270 с.
7. Тоценко В.Г. Алгоритмы технического диагностирования дискретных устройств. – М.: Радио и связь, 1985. – 240 с.
8. Павлов Б.В. Акустическая диагностика машин. – М.: Машиностроение, 1971. – 222 с.
9. Павлов Б.В. Кибернетические методы технического диагноза. – М.: Машиностроение, 1966. – 147 с.
10. Сердаков А.С. Автоматический контроль и техническая диагностика. – К.: Техника, 1971. – 241 с.
11. Силин В.Б., Заковряшин А.М. Автоматическое прогнозирование состояния аппаратуры управления и наблюдения. – М.: Энергия, 1973. – 334 с.
12. Введение в техническую диагностику / Г.Ф. Верзаков, Н.В. Кишт, В.И. Рабинович, Л.С. Тимонен. – М.: Энергия, 1968. – 224 с.
13. Дубровин В., Степаненко О., Луценко Д. Діагностика процесів та керування якістю. // Досвід розробки і застосування САПР в мікроелектроніці: Матеріали п'ятої міжнародної науково-технічної конференції CADSM'99. – Львів, Державний університет “Львівська політехніка”, 1999, С.22-24.
14. Айвазян С.А., Бежаева З.И., Староверов О.В. Классификация многомерных наблюдений. – М.: Статистика, 1974. – 240 с.
15. Айзерман М.А., Браверман Э.М., Розоноэр Л.И. Метод потенциальных функций в теории обучения машин. – М.: Наука, 1970. – 383 с.
16. Алгоритмы и программы восстановления зависимостей / Под ред. В.Н. Вапника. – М.: Наука, 1984. – 816 с.
17. Анисимов Б.В., Курганов В.Д., Злобин В.К. Распознавание и цифровая обработка изображений: Учеб. пособие для студентов вузов. – М.: Высшая школа, 1983. – 295 с.

18. Аркадьев А.Г., Браверман Э.М. Обучение машины классификации объектов. – М.: Наука, 1971. – 172 с.
19. Афифи А., Эйзен С. Статистический анализ: Подход с использованием ЭВМ.: Пер. с англ. – М.: Мир, 1982. – 488 с.
20. Бонгард М.М. Проблема узнавания. – М.: Наука, 1967. – 320с.
21. Вапник В.Н. Задача обучения распознаванию образов. – М.: Знание, 1971. – 60 с.
22. Вапник В.Н., Червоненкис А.Ф. Теория распознавания образов. – М.: Наука, 1974.
23. Вапник В. Н. Восстановление зависимостей по эмпирическим данным. – М.: Наука, 1979. – 448 с.
24. Васильев В.И. Проблема обучения распознаванию образов. – К.: Выща школа, 1989. – 64 с.
25. Васильев В.И. Распознающие системы (Справочник). – К.:Наукова думка, 1983. – 422 с.
26. Галушкин А.И. Синтез многослойных систем распознавания образов. – М.: Энергия, 1974. – 368 с.
27. Гренандер У. Лекции по теории образов. Т. 1. Синтез образов: Пер. с англ. – М.: Мир, 1979. – 384 с.
28. Гренандер У. Лекции по теории образов. Т. 2. Анализ образов: Пер. с англ. – М.: Мир, 1981. – 448 с.
29. Гроп Д. Методы идентификации систем. – М.: Мир, 1979. – 302 с.
30. Головкин Б.А. Машинное распознавание и линейное программирование. – М.: Сов. радио, 1973. – 100 с.
31. Горелик А.Л., Скрипкин В.А. Методы распознавания: Учеб. пособие для вузов.-3-е изд. перераб. и доп. – М.: Высшая школа, 1989. – 232с.
32. Горелик А.Л., Гуревич И.Б., Скрипкин В.А. Современное состояние проблемы распознавания: Некоторые аспекты. – М.: Радио и связь, 1985. – 160 с.
33. Горелик А. Л. Общая постановка задачи распознавания объектов и явлений. – Кибернетика, 1980, №6, С. 72-75.
34. Горелик А. Л., Скрипкин В. А. Об одном методе решения задач классификации объектов или явлений. – Техническая кибернетика, 1965, № 1, С. 56-64.
35. Горелик А. Л., Эпштейн С. С. Об аддитивности информации в задачах распознавания объектов и явлений. – Кибернетика, 1983, С. 85-88.
36. Горелик А. Л. Об одном подходе к выбору пространства признаков, используемого при построении системы распознавания объектов и явлений. – Кибернетика, 1972, №4, С. 142-146.
37. Горелик А. Л. Игровой подход к построению пространства признаков систем распознавания объектов и явлений. – Кибернетика, 1973, № 5. С. 114-116.

38. Дуда Р., Харт П. Распознавание образов и анализ сцен. – М.: Мир, 1976. – 512 с.
39. Журавлев Ю. И., Никифоров В. В. Алгоритмы распознавания, основанные на вычислении оценок. – Кибернетика, 1971, №3, С. 1–11.
40. Журавлев Ю. И. Непараметрические задачи распознавания образов. – Кибернетика, 1976, № 6, С. 93 – 103.
41. Загоруйко Н.Г. Методы распознавания и их применение. – М.: Советское радио, 1972. – 208 с.
42. Загоруйко Н.Г., Елкина В.Н., Лбов Г.С. Алгоритмы обнаружения эмпирических закономерностей. – Новосибирск: Наука, 1985. – 110 с.
43. Ивахненко А.Г. Самообучающиеся системы распознавания и автоматического регулирования. – Киев: Техника, 1969. – 392 с.
44. Круг Г.К., Кабанов В.А., Фомин Г.А., Фомина Е.С. Планирование эксперимента в задачах нелинейного оценивания и распознавания образов. – М.: Наука, 1981. – 172 с.
45. Кузин Л.Т. Основы кибернетики: В 2-х тт. Учеб. пособие для вузов. – М: Энергия, 1976.
46. Патрик Э. Основы теории распознавания образов / Пер. с англ. – М.: Сов. радио, 1980. – 408с.
47. Прикладная статистика: Классификация и снижение размерности: Справ. изд./ С.А. Айвазян, В.М. Бухштабер, И.С. Енюков, Л.Д. Мешалкин. Под ред. С.А. Айвазяна. – М.: Финансы и статистика, 1989. – 607с.
48. Перцептрон – система распознавания образов/А. Г. Ивахненко, В. И. Васильев, К. Г. Агабабян и др. / Под ред. А. Г. Ивахненко. – Киев: Наукова думка, 1975. – 432 с.
49. Распознавание образов: состояние и перспективы: Пер. с англ./ К.Верхаген, Р. Дейн, Ф. Грун и др. – М.: Радио и связь, 1985. – 104 с.
50. Растринин Л. А., Эренштейн Р. Х. Метод коллективного распознавания. – М.: Энергоиздат, 1981. – 79 с.
51. Розенблат Ф. Принципы нейродинамики. Перцептрон и теория механизмов мозга. – М.: Мир, 1965. – 480 с.
52. Себестиан Г. Процессы принятия решений при распознавании образов. – К.: Техніка, 1965. – 150с.
53. Ту Дж., Гонсалес Р. Принципы распознавания образов: Пер. с англ./Под ред. Ю. И. Журавлева. – М.: Мир, 1978. – 416 с.
54. Фор А. Восприятие и распознавание образов. – М.: Машиностроение, 1989. – 272с.
55. Фукунага К. Введение в статистическую теорию распознавания образов. – М.: Наука, 1979. – 368 с.
56. Фу К. Структурные методы в распознавании образов: Пер. с англ./Под ред. А. М. Айзермана. – М.: Мир, 1977. – 320 с.

57. Цыпкин Я.З. Основы теории обучающихся систем. М.: Наука, 1970. 252 с.
58. Цыпкин Я. З. Адаптация и обучение в автоматических системах. – М.: Наука, 1968. – 400 с.
59. V. Vapnic. Statistical larning theory. A Wiley – Interscience publication. John Wiley & Sons, New York, 1998, 732p.