

Міністерство освіти і науки України  
Запорізький національний технічний університет

**МЕТОДИЧНІ ВКАЗІВКИ**  
до виконання лабораторних і самостійних робіт  
з дисципліни  
**«ІНТЕЛЕКТУАЛЬНІ СИСТЕМИ  
ІНЖЕНЕРНОГО ЗАБЕЗПЕЧЕННЯ  
ВИРОБНИЦТВА»**  
для студентів напряму підготовки  
6.050101 "Комп'ютерні науки"  
усіх форм навчання

2013



**ЗМІСТ**

Загальні положення . . . . .	4
1. Лабораторна робота № 1. Розпізнавання образів на основі метричної класифікації. . . . .	5
2. Лабораторна робота № 2. Методи відбору ознак для побудови розпізнаючих моделей. . . . .	8
3. Лабораторна робота № 3. Самоорганізація та навчання без учителя. Кластер-аналіз . . . . .	10
4. Самостійна робота № 1. Програмні засоби для побудови інтелектуальних систем . . . . .	12
Література. . . . .	13
Додаток А. Кластер-аналіз у середовищі MATLAB . . . . .	14

## ЗАГАЛЬНІ ПОЛОЖЕННЯ

Дане видання призначене для вивчення та практичного освоєння студентами усіх форм навчання основ створення систем штучного інтелекту.

Відповідно до графіка студенти перед виконанням лабораторної або самостійної роботи повинні ознайомитися з конспектом лекцій та рекомендованою літературою.

Для одержання заліку по кожній роботі студент здає викладачу цілком оформлений звіт, а також CD (DVD)-R/W диск у форматі MS – DOS / Windows, перевірений на відсутність вірусів, з текстами розроблених програм, файлами програм, що виконуються, файлами даних і текстом звіту.

Звіт (приклад оформлення - див. додаток А) має містити:

- титульний аркуш (на ньому вказують назву міністерства, назву університету, назву кафедри, номер, вид і тему роботи, виконавця та особу, що приймає звіт, рік);

- мету, варіант і завдання роботи;

- лаконічний опис теоретичних відомостей;

- текст програми, що обов'язково містить коментарі;

- вхідні та вихідні дані програми;

- змістовний аналіз отриманих результатів та висновки.

Звіт виконують на білому папері формату А4 (210 x 297 мм). Текст розміщують тільки з однієї сторони листа. Поля сторінки з усіх боків – 20 мм. Аркуші скріплюють за допомогою канцелярських скріпок. Для набору тексту звіту використовують редактор MS Word 97: шрифт Times New Roman, 12 пунктів. Міжрядковий інтервал: полуторний – для тексту звіту, одинарний – для листингів програм, таблиць і роздруківок даних.

Під час співбесіди студент повинний виявити знання про мету роботи, по теоретичному матеріалу, про методи виконання кожного етапу роботи, по змісту основних розділів розробленого звіту з демонстрацією результатів на конкретних прикладах. Студент повинний вміти правильно аналізувати отримані результати. Для самоперевірки при підготовці до виконання і здачі роботи студент повинний відповісти на контрольні питання, приведені наприкінці опису відповідної роботи. Загальний залік студент одержує після виконання і здачі останньої роботи.

## 1 Лабораторна робота № 1

### РОЗПІЗНАВАННЯ ОБРАЗІВ НА ОСНОВІ МЕТРИЧНОЇ КЛАСИФІКАЦІЇ

**Мета роботи:** вивчити та засвоїти на практиці метричні методи розпізнавання образів у просторі ознак, навчитися створювати програмні засоби, що реалізують методи метричної класифікації.

#### Завдання до роботи

1. Ознайомитися з конспектом лекцій та рекомендованою літературою. На алгоритмічній мові програмування пакету MATLAB написати програму, що реалізує процедури для навчання та емуляції за методом еталонів.

2. Згідно з номером студента за журналом для відповідного номера варіанта  $V$  сформувати навчаючу вибірку  $\langle x, y \rangle$  обсягом  $S$  екземплярів  $x^s$ ,  $s=1, 2, \dots, S$ , що характеризуються  $N$  ознаками  $x_j^s$ ,  $j=1, 2, \dots, N$ , та зіставити кожному екземпляру значення цільової ознаки  $y^s$ :

$$x_j^s = \begin{cases} jV - 0,1s, & j = 1, 5, 9, \dots, \\ 0,01jV^{-1} + 0,3s, & j = 2, 4, 6, \dots, \\ j\text{rand}, & j = 3, 7, 11, \dots; \end{cases} \quad y^s = \begin{cases} 0, (x_1^s)^2 + (x_2^s)^2 < V^2 + 0,04S^2, \\ 1, (x_1^s)^2 + (x_2^s)^2 \geq V^2 + 0,04S^2; \end{cases}$$

$$S = \begin{cases} 10V, & V < 10, \\ 5V, & 10 \leq V < 20, \\ 3V, & V \geq 20; \end{cases} \quad N = \begin{cases} 5V, & V < 7, \\ 4V, & 7 \leq V < 10, \\ 3V, & 10 \leq V < 20, \\ 2V, & V \geq 20; \end{cases}$$

де  $\text{rand}$  - випадкове число в діапазоні  $[0, 1]$ .

3. Виконати нормування навчаючої вибірки даних.

4. На основі проноормованої вибірки побудувати розпізнаючу модель за методом еталонів, тобто визначити координати центрів класів (еталонів) у просторі ознак  $C^q_j$ , де  $q$  - номер класу,  $j$  - номер ознаки. Зафіксувати час навчання.

5. На основі побудованої моделі для екземплярів навчаючої вибірки виконати розпізнавання, тобто визначити розрахункові номери класів  $y^{s*}$ . Зафіксувати час розпізнавання.

6. Обчислити помилку розпізнавання, визначити ймовірність прийняття правильного рішення та ймовірність прийняття помилкового рішення для побудованої моделі.

7. У попередньо сформованій вибірці залишити тільки одну ознаку, номер якої у попередній вибірці дорівнює  $V$ , а також у центрів еталонів класів залишити тільки  $V$ -ту координату  $C^q_V$ .

8. Для нової вибірки та нових еталонів виконати розпізнавання, тобто визначити розрахункові номери класів  $y^{s*}$ . Зафіксувати час розпізнавання.

9. Обчислити помилку розпізнавання, визначити ймовірність прийняття правильного рішення та ймовірність прийняття помилкового рішення для нової моделі.

10. Порівняти результати проведених експериментів, зробити висновки щодо впливу параметрів навчаючої вибірки на характеристики процесів навчання та розпізнавання.

### Зміст звіту

1. Мета роботи.

2. Короткі теоретичні відомості до роботи. У звіті не слід наводити докладне викладення теоретичного матеріалу, необхідно виділити лише найголовніші ідеї, формули, що необхідні для пояснення суті методу, моделі тощо.

3. Текст розробленої програми.

4. Сформована навчаюча вибірка  $\langle x, y \rangle$ .

5. Побудована модель - координати еталонів (центрів зосередження екземплярів) класів. Час навчання.

6. Результати розпізнавання: розраховані значення номеру класу для екземплярів  $y^{s*}$ , помилка розпізнавання, ймовірності прийняття правильного та помилкового рішень, час розпізнавання.

7. Навчаюча вибірка із однією ознакою.

8. Еталони класів з однією ознакою.

9. Результати розпізнавання скороченої вибірки за скороченими еталонами: розраховані значення номеру класу для екземплярів  $y^{s*}$ , помилка розпізнавання, ймовірності прийняття правильного та помилкового рішень, час розпізнавання.

10. Висновки. У висновках треба проаналізувати результати роботи, а також лаконічно відповісти на контрольні питання.

## Контрольні питання

1. Задача розпізнавання образів.
2. Основні поняття теорії розпізнавання образів.
3. Розбиття вихідної вибірки на навчаючу та тестову.
4. Навчання з учителем.
5. Методи метричної класифікації.
6. Навчання без учителя. Кластер-аналіз.
7. Лінійна роздільність і лінійна нерозділеність класів.
8. помилка навчання / класифікації, час навчання / класифікації, цільова функція навчання.
9. Чи впливає кількість використаних ознак на швидкість навчання перцептрона? Відповідь обґрунтуйте теоретично та доведіть експериментально.
10. Що таке генеральна сукупність, вибірка, екземпляр, ознака?
11. Вимоги до навчаючих вибірок даних.
12. Що таке репрезентативна вибірка даних?
13. Чи повинна навчальна вибірка бути репрезентативною?
14. Чи повинна тестова вибірка бути репрезентативною?
15. Чи впливає обсяг навчаючої вибірки на швидкість навчання?
16. Чи впливає репрезентативність навчаючої вибірки на точність класифікації екземплярів тестової вибірки?
17. Чи впливає репрезентативність тестової вибірки на точність класифікації екземплярів тестової вибірки?
18. Чи впливає репрезентативність тестової вибірки на точність навчання перцептрона за навчаючої вибіркою?
19. Чи залежить якість навчання від якості та обсягу навчаючої вибірки?

## 2 Лабораторна робота № 2

### МЕТОДИ ВІДБОРУ ОЗНАК ДЛЯ ПОБУДОВИ РОЗПІЗНАЮЧИХ МОДЕЛЕЙ

**Мета роботи:** вивчити та засвоїти на практиці методи оцінювання інформативності та відбору ознак, для побудови розпізнаючих моделей.

#### Завдання до роботи

1. Ознайомитися з конспектом лекцій та рекомендованою літературою.

2. На алгоритмічній мові програмування пакету MATLAB написати програму, що реалізує методи оцінювання інформативності ознак: на основі модуля коефіцієнта парної кореляції, коефіцієнта кореляції знаків, коефіцієнта кореляції Фехнера, кількості інтервалів зміни номера класу, інформаційного підходу, статистичного підходу.

3. Згідно з номером студента за журналом для відповідного номера варіанта V сформувати навчаючу вибірку  $\langle x, y \rangle$  за формулами, наведеними у завданні попередньої роботи. Також для екземплярів вибірки визначити значення другої цільової ознаки  $y_2^s$ :

$$y_2^s = 2x_1^s + 0,1x_2^s, s=1,2,\dots,S.$$

4. На основі сформованої вибірки по відношенню до дискретного виходу  $y^s$  оцінити інформативність ознак екземплярів вибірки за допомогою коефіцієнта кореляції знаків, коефіцієнта кореляції Фехнера, кількості інтервалів зміни номера класу, інформаційного підходу, статистичного підходу.

5. На основі сформованої вибірки по відношенню до дійсного виходу  $y_2^s$  оцінити інформативність ознак екземплярів вибірки за допомогою на основі модуля коефіцієнта парної кореляції, кількості інтервалів зміни номера класу (для цього попередньо дискретизувати  $y_2^s$ ).

6. Побудувати таблицю з оцінками інформативності ознак відносно дискретного  $y^s$ , стовпці якої повинні мати назви: номер ознаки, коефіцієнт кореляції знаків, коефіцієнт кореляції Фехнера, оцінка за



кількістю інтервалів зміни номера класу, оцінка за інформаційним підходом, оцінка на основі статистичного підходу.

7. Побудувати таблицю з оцінками інформативності ознак відносно дійсного  $y^s_2$ , стовпці якої повинні мати назви: номер ознаки, модуль коефіцієнта парної кореляції, оцінка за кількістю інтервалів зміни номера класу

8. Проаналізувати за побудованими таблицями оцінки інформативності ознак. Зробити висновки щодо важливості ознак окремо для  $y^s$  та  $y^s_2$ .

### Зміст звіту

1. Мета роботи.
2. Короткі теоретичні відомості до роботи.
3. Текст розробленої програми.
4. Сформована навчаюча вибірка  $\langle x, y, y_2 \rangle$ .
5. Таблиці з оцінками інформативності ознак по відношенню до  $y^s$  та  $y^s_2$ .
6. Рішення щодо важливості ознак окремо для  $y^s$  та  $y^s_2$ .
7. Висновки. У висновках треба проаналізувати результати роботи, а також лаконічно відповісти на контрольні питання.

### Контрольні питання

1. Задача відбору ознак.
2. Критерії оцінювання інформативності ознак на основі евристичного, інформаційного, статистичного та імовірнісного підходів.
3. Оцінювання ознак за допомогою коефіцієнта кореляції знаків, коефіцієнта кореляції Фехнера, кількості інтервалів зміни номера класу.
4. Комбінована оцінка інформативності ознак.
5. Що таке генеральна сукупність, вибірка, екземпляр, ознака?
6. Вимоги до навчаючих вибірок даних.
7. Що таке репрезентативна вибірка даних?

### 3 Лабораторна робота № 3

## САМООРГАНІЗАЦІЯ ТА НАВЧАННЯ БЕЗ УЧИТЕЛЯ. КЛАСТЕР-АНАЛІЗ

**Мета роботи:** вивчити та засвоїти на практиці методи кластер-аналізу та його використання для навчання з учителем та без учителя.

### Завдання до роботи

1. Ознайомитися з конспектом лекцій та рекомендованою літературою. На алгоритмічній мові програмування пакету MATLAB написати програму, що використовує функції кластер-аналізу (див. додаток Б).

2. Згідно з номером студента за журналом для відповідного номера варіанта V сформулювати навчаючу вибірку за допомогою формул лабораторної роботи № 1.

3. На основі вибірки виділити центри кластерів шляхом чіткої та нечіткої кластеризації.

4. Виконати нормування навчаючої вибірки даних.

5. На основі пронормованої вибірки виділити центри кластерів шляхом чіткої та нечіткої кластеризації.

6. Порівняти результати для чіткої та нечіткої кластеризації для нормованої та ненормованої вибірок.

### Зміст звіту

1. Мета роботи.
2. Короткі теоретичні відомості до роботи.
3. Текст розробленої програми.
4. Сформована навчаюча вибірка.
5. Пронормована вибірка.
6. Координати центрів кластерів для ненормованої та нормованої вибірок після чіткої та нечіткої кластеризації.
7. Висновки. У висновках треба проаналізувати результати роботи, а також лаконічно відповісти на контрольні питання.

### Контрольні питання

1. Задача кластер-аналізу.
2. Чіткий кластер-аналіз.
3. Нечіткий кластер-аналіз.
4. Використання кластер-аналізу при побудові систем розпізнавання образів.
5. Навчання з учителем.
6. Навчання без учителя.
7. Подібність кластер-аналізу і метричної класифікації.
8. Лінійна роздільність і лінійна нерозділеність класів.
9. Чи впливає кількість використаних ознак на швидкість кластер-аналізу?
10. Що таке генеральна сукупність, вибірка, екземпляр, ознака?
11. Чи впливає обсяг навчаючої вибірки на швидкість кластер-аналізу?
12. Функції кластер-аналізу у пакеті MATLAB.
13. Функція `fcm`.
14. Функція `subclust`.
15. Застосування кластер-аналізу для категоризації даних.

## 4 Самостійна робота № 1

### ПРОГРАМНІ ЗАСОБИ ДЛЯ ПОБУДОВИ ІНТЕЛЕКТУАЛЬНИХ СИСТЕМ

**Мета роботи:** освоїти засоби середовища MATLAB для створення на його основі прикладних інтелектуальних систем.

#### Завдання до роботи

1. Використовуючи рекомендовану літературу вивчити основні принципи побудови програм для пакету MATLAB.
2. Розробити на мові для пакету MATLAB функції, що демонструють роботу з матрицями (із використанням матричних та масивних операцій), будують графіки, зберігають змінні на диску. Розроблені функції інтегрувати до графічної форми, яку створити за допомогою GUI.
3. Порівняти можливості пакету MATLAB із відомими системами програмування.

#### Зміст звіту

1. Мета роботи.
2. Короткі теоретичні відомості.
3. Тексти програм та інтерфейсні форми, розроблені студентом.
4. Вхідні дані та результати роботи програми.
5. Висновки, що містять відповіді на контрольні питання, а також відображують результати виконання роботи та їх критичний аналіз.

#### Контрольні питання

1. Архітектура та характеристики пакету MATLAB.
2. Інтелектуальний аналіз даних у пакеті MATLAB.
3. Поняття та властивості інтелектуальної системи.
4. Агентний підхід до інтелектуальних систем.
5. Архітектура інтелектуальної системи.

## ЛІТЕРАТУРА

1. Джонс М.Т. Программирование искусственного интеллекта в приложениях / М.Т. Джонс. Пер. с англ. Осипов А.И.–М.: ДМК Пресс, 2004.– 312 с.
2. Дубровин В.И. Интеллектуальные средства диагностики и прогнозирования надежности авиадвигателей / В.И. Дубровин, С.А. Субботин, А.В. Богуслаев, В.К. Яценко. Монография.– Запорожье: ОАО "Мотор–Сич", 2003.– 279 с.
3. Зайченко Ю.П. Основы проектирования интеллектуальных систем. / Ю.П. Зайченко Навчальний посібник.– К.: Слово, 2004.– 352 с.
4. Кричевский М.Л. Интеллектуальные методы в менеджменте. / М.Л. Кричевский – СПб.: Питер, 2005.– 304 с.
5. Леоненков А.В. Нечеткое моделирование в среде MATLAB. / А.В. Леоненков. – СПб.: БХВ–Петербург, 2003. – 736 с.
6. Люгер Дж.Ф. Искусственный интеллект: стратегии и методы решения сложных проблем / Дж.Ф. Люгер Пер. с англ.– М.: Вильямс, 2005.– 864 с.
7. Рассел С., Норвиг П. Искусственный интеллект: современный подход, 2–е изд.: / С.Рассел Пер с англ.–М.: Вильямс, 2006.– 1408 с.
8. Рідкокаша А.А., Основы систем штучного інтелекту. / А.А. Рідкокаша, К.К. Голдер. Навчальний посібник.– Черкаси, "Відлуння–Плюс", 2002.–240 с.
9. Руденко О.Г. Основы теории искусственных нейронных сетей / О.Г. Руденко, Е.В. Бодянский. – Харьков Телетех, 2002. – 317с.
10. Субботін С. О. Подання й обробка знань у системах штучного інтелекту та підтримки прийняття рішень : Навчальний посібник / С. О. Субботін. – Запоріжжя: ЗНТУ, 2008. –341 с.
11. Субботін С. О. Неітеративні, еволюційні та мультиагентні методи синтезу нечіткологічних і нейромережних моделей : Монографія / С. О. Субботін, А. О. Олійник, О. О. Олійник ; під заг. ред. С. О. Субботіна. – Запоріжжя: ЗНТУ, 2009. – 375 с.
12. Прогрессивные технологии моделирования, оптимизации и интеллектуальной автоматизации этапов жизненного цикла авиационных двигателей: Монография / А. В. Богуслаев, Ал. А. Олейник, Ан. А. Олейник, Д. В. Павленко, С. А. Субботин; Под ред. Д. В. Павленко, С. А. Субботина. – Запорожье: ОАО "Мотор Сич", 2009. – 468 с.
13. Хайкин С. Нейронные сети: полный курс, 2–е изд. / С. Хайкин: Пер. с англ.– М.: Вильямс, 2006.– 1104 с.

## Додаток А

### Кластер-аналіз у середовищі MATLAB

Модуль Fuzzy Logic Toolbox пакету MATLAB містить функції для виділення кластерів.

Функція *subclust* визначає координати центрів кластерів шляхом чіткої кластеризації зі зменшенням кількості кластерів.

Функція *subclust* знаходить оптимальну точку даних для визначення центра кластера ґрунтуючись на щільності оточення точок даних. Усі точки даних у межах відстані *RADII* до цієї точки видаляються, щоб визначити наступний кластер даних та його центр. Цей процес повторюється поки усі дані не знаходяться у межах відстані *RADII* до центра кластера.

$[C] = \text{SUBCLUST}(X, \text{RADII})$  кластеризує точки даних  $S \times N$  матриці  $X$ , де  $S$  - кількість точок даних,  $N$  - кількість координат точок даних, *RADII* - значення між 0 та 1, що визначає розмір кластера в кожному з вимірювань даних, приймаючи, що дані знаходяться у межах діапазону  $[0, 1]$  (Встановлення меншого радіуса кластера буде звичайно створювати більше менших за розміром кластерів. Коли *RADII* є скаляром, він застосовується до усіх вимірів даних. Коли *RADII* є вектором, він має окреме значення для кожного виміру даних), та повертає центри кластерів як рядки матриці  $C$ , що має розмір  $V \times N$ , де  $V$  - кількість кластерів.

$[C] = \text{SUBCLUST}(\dots, \text{XBOUNDS})$  також визначає матрицю *XBOUNDS*, розміром  $2 \times N$ , що використовується для нормалізації даних  $X$  у діапазон  $[0, 1]$ . Кожний стовпець *XBOUNDS* містить мінімальні та максимальні значення для відповідної розмірності даних. Якщо *XBOUNDS* - порожня матриця або не використовується, тоді за замовчуванням використовуються мінімальні та максимальні значення даних  $X$ .

$[C] = \text{SUBCLUST}(\dots, \text{OPTIONS})$  визначає вектор для зміни значень за замовчуванням параметрів алгоритму: *OPTIONS*(1) - коефіцієнт, що використовується для множення на значення *RADII* для визначення осередку центру кластера, у межах якого існування інших центрів кластерів заборонено; *OPTIONS*(2) - коефіцієнт прийняття, що встановлює потенціал як частку потенціалу центра першого кластера, вище якої інша точка даних буде прийнята як центр кластера; *OPTIONS*(3) - коефіцієнт відхилення, що встановлює потенціал як частку потенціалу центра першого кластера, нижче якої інша точка даних буде відхилена як центр кластера; *OPTIONS*(4) - ознака відображення поточної інформації, якщо не встановлена як 0.

Значеннями вектора *OPTIONS* за замовчуванням є  $[1.25 \ 0.5 \ 0.15 \ 0]$ .

Приклад використання *subclust*.

```

X1 = 10*rand(50,1);
X2 = 5*rand(50,1);
X3 = 20*rand(50,1)-10;
X = [X1 X2 X3]; % генеруємо вибірку даних
[C] = subclust(X,0.5); % знаходимо центри кластерів

```

Функція *fcm* здійснює нечітку кластеризацію на основі методу нечітких *c*-середніх та має формат виклику:

```
[CENTER, U, OBJ_FCN] = fcm(DATA, N_CLUSTER, OPTIONS)
```

де *N\_CLUSTER* - кількість кластерів в наборі даних масиву *DATA*, який має розміри  $S \times N$ , *S* - кількість точок даних, *N* - кількість координат точок, *CENTER* - матриця з координатами центрів кластерів (кластери містяться у рядках, ознаки - у стовпцях), *U* - матриця функції приналежності, що містить рівні приналежності кожної точки масиву *DATA* до кожного кластера, *OBJ\_FCN* - значення цільової функції для центрів кластерів, *OPTIONS* - необов'язковий параметр, що задає вектор опцій для процесу кластеризації: *OPTIONS*(1) - експонента для матриці *U* (за замовчуванням: 2.0), *OPTIONS*(2) - максимальна кількість ітерацій (за замовчуванням: 100), *OPTIONS*(3) - мінімально прийнятне покращення цільової функції (за замовчуванням:  $10^{-5}$ ), *OPTIONS*(4): ознака відображення проміжних результатів (за замовчуванням: 1).

На кожній ітерації цільова функція мінімізується для знаходження кращого розташування кластерів. Процес кластеризації зупиняється, коли максимально прийнятна кількість ітерацій є досягнутою, або коли покращення цільової функції між двома послідовними ітераціями змінюється меншою ніж мінімально прийнятний приріст.

#### Приклад використання *fcm*.

```

data = rand(100,2); % генеруємо вибірку
[center,U,obj_fcn] = fcm(data,2); % виконуємо кластер-аналіз
plot(data(:,1), data(:,2), 'o'); % зображуємо дані на графіку
hold on;
maxU = max(U);
% Знаходимо точки з найвищим рівнем приналежності до першого кластера
index1 = find(U(1,:) == maxU);
% Знаходимо точки з найвищим рівнем приналежності до другого кластера
index2 = find(U(2,:) == maxU);
line(data(index1,1),data(index1,2), 'marker','*','color','g');
line(data(index2,1),data(index2,2), 'marker','*','color','r');
% Зображуємо центри кластерів на графіку
plot([center([1 2],1)], [center([1 2],2)], '*', 'color','k');

```