

Міністерство освіти і науки, молоді та спорту України
Запорізький національний технічний університет

МЕТОДИЧНІ ВКАЗІВКИ
до виконання розрахунково-графічного завдання
з дисципліни
«ІНТЕЛЕКТУАЛЬНІ СИСТЕМИ»
для студентів напрямку підготовки
6.050103 “Програмна інженерія”
(усіх форм навчання)

Методичні вказівки до виконання розрахунково-графічного завдання з дисципліни «Інтелектуальні системи» для студентів напряму підготовки 6.050103 “Програмна інженерія” (усіх форм навчання) / Уклад.: С.О. Субботін. – Запоріжжя: ЗНТУ, 2012. – 15 с.

Укладач: Сергій Олександрович Субботін,
кандидат технічних наук, доцент,
доцент кафедри програмних засобів

Рецензент: А.О. Олійник, кандидат технічних наук

Відповідальний
за випуск: В.І. Дубровін, зав. каф. програмних засобів

Затверджено
на засіданні кафедри
програмних засобів

Протокол № 12
від 27.04.2012 р.

ЗМІСТ

Загальні положення	4
1. Розрахункова частина завдання.....	6
2. Графічна частина завдання	8
3. Реферативна (факультативна) частина завдання	9
Література.....	10
Додаток А. Приклад оформлення титульного аркушу звіту про виконання розрахунково-графічного завдання.....	11
Додаток Б. Кластер-аналіз у середовищі MATLAB.....	12
Додаток В. Перелік тем для виконання реферативної частини розрахунково-графічного завдання.....	15

ЗАГАЛЬНІ ПОЛОЖЕННЯ

Дане видання призначене для вивчення та практичного освоєння студентами усіх форм навчання теоретичних основ та практичних навичок і вмінь з навчальної дисципліни шляхом виконання розрахунково-графічних завдань.

Розрахунково-графічне завдання є різновидом індивідуальних завдань, що виконуються з окремих навчальних дисциплін студентами самостійно при консультуванні викладачем.

Студент виконує розрахунково-графічне завдання протягом семестру відповідно до варіанту завдання, наданого викладачем на початку семестру.

При виконанні розрахунково-графічного завдання студент має використовувати конспект лекцій та рекомендовану літературу.

Для одержання оцінки за виконане розрахунково-графічне завдання здає викладачу цілком оформлений звіт, а також оптичний носій (CD або DVD) у форматі MS-DOS / Windows, перевірений на відсутність вірусів, з текстами розроблених програм, файлами програм, що виконуються, файлами даних і текстом звіту.

Звіт має містити:

- титульний аркуш (на ньому вказують назву міністерства, назву університету, назву кафедри, вид і тему роботи, виконавця та особу, що приймає звіт, рік – приклад оформлення наведено у додатку А);
- мету, номер варіанту і формулювання вирішуваного завдання;
- лаконічний опис використовуваних 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. Дозволяється виконання звіту уручну, на друкарській машинці або шляхом комп'ютерного набору із роздруківкою на принтері (рекомендований вид виконання) – за вибором студента.

При виконанні звіту із застосуванням комп'ютерної техніки для набору тексту звіту використовують редактор MS Word (версії 97 і вище) або сумісний з ним. Параметри тексту документу: гарнітура

шрифту – Times New Roman, розмір шрифту: 14 пунктів – для тексту звіту, 10 пунктів – для роздруківок програм та великих таблиць, міжрядковий інтервал: полуторний – для тексту звіту, одинарний – для листингів програм, таблиць і роздруківок даних, абзацний відступ встановлюють 1,5 см. Таблиці і рисунки та перелік посилань оформлюють відповідно до чинних стандартів.

Поданий звіт з виконання розрахунково-графічного завдання та програмне забезпечення на відповідних носіях перевіряється викладачем та захищається студентом шляхом усної або письмової відповіді на питання викладача і демонстрації роботи розробленого програмного забезпечення на ЕОМ.

Під час співбесіди студент повинний виявити знання про мету виконуваного завдання, по теоретичному матеріалу, про методи виконання кожного етапу роботи, по змісту основних розділів розробленого звіту з демонстрацією результатів на конкретних прикладах. Студент повинен вміти правильно аналізувати отримані результати.

Оцінювання розрахунково-графічного завдання здійснюється викладачем за такою шкалою:

– "незадовільно" – студент не виконав, виконав не повністю чи неправильно виконав поставлене завдання або не здатен пояснити хід його виконання, а також у випадку виявлення плагіату (спроба здачі чужої роботи);

– "задовільно" – студент в цілому виконав і захистив поставлене завдання, але з грубими помилками (у тому числі в оформленні звіту);

– "задовільно" – студент в цілому виконав і захистив поставлене завдання, але з незначними помилками (у тому числі в оформленні звіту);

– "відмінно" – студент в цілому виконав і захистив поставлене завдання, але з незначними помилками (у тому числі в оформленні звіту).

Передача розрахунково-графічного завдання студентом здійснюється лише у випадку отримання ним оцінки "незадовільно".

1 РОЗРАХУНКОВА ЧАСТИНА ЗАВДАННЯ

Мета роботи: вивчити та засвоїти на практиці методи кластерного аналізу для розпізнавання образів у просторі ознак, навчитися створювати та використовувати програмні засоби, що реалізують методи кластер-аналізу для вирішення інтелектуальних задач.

Порядок виконання розрахункової частини завдання

1. Ознайомитися з конспектом лекцій та рекомендованою літературою.

2. Згідно з номером студента за журналом для відповідного номера варіанта V сформувати навчальну 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 :

$$x_j^s = \begin{cases} jV - 0,1s, j = 1, 5, 9, \dots, \\ 0,01jV^{-1} + 0,3s, j = 2, 4, 6, \dots, \\ j\text{rand}, j = 3, 7, 11, \dots; \end{cases}$$

$$y^s = \begin{cases} 0, (x_1^s)^2 + (x_2^s)^2 < V^2 + 0,04S^2, \\ 1, (x_1^s)^2 + (x_2^s)^2 \geq V^2 + 0,04S^2; \end{cases}$$

$$S = \begin{cases} 10V, V < 10, \\ 5V, 10 \leq V < 20, \\ 3V, V \geq 20; \end{cases}$$

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

де rand – випадкове число в діапазоні $[0, 1]$.

3. Виконати нормування навчаючої вибірки даних.

4. За допомогою функцій модуля Fuzzy Toolbox пакета MATLAB виконати кластер-аналіз пронормованої вибірки.

Зміст розділу звіту з розрахункової частини

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 навчальна вибірка $\langle x, y \rangle$.
5. Координати центрів виділених кластерів після чіткої та нечіткої кластеризації.

Контрольні питання

1. Задача розпізнавання образів.
2. Основні поняття теорії розпізнавання образів.
3. Розбиття вихідної вибірки на навчаючу та тестову.
4. Лінійна роздільність і лінійна нерозділеність класів.
5. Що таке генеральна сукупність, вибірка, екземпляр, ознака?
6. Вимоги до навчаючих вибірок даних.
7. Що таке репрезентативна вибірка даних?
8. Задача кластер-аналізу.
9. Чіткий кластер-аналіз.
10. Нечіткий кластер-аналіз.
11. Використання кластер-аналізу при побудові систем розпізнавання образів.
12. Навчання з учителем. Навчання без учителя.
13. Подібність кластер-аналізу і метричної класифікації.
14. Лінійна роздільність і лінійна нерозділеність класів.
15. Чи впливає кількість використаних ознак на швидкість кластер-аналізу?
16. Чи впливає обсяг навчаючої вибірки на швидкість кластер-аналізу?
17. Функції кластер-аналізу у пакеті MATLAB.
18. Застосування кластер-аналізу для категоризації даних.

2 ГРАФІЧНА ЧАСТИНА ЗАВДАННЯ

Порядок виконання графічної частини завдання

1. Ознайомитися з конспектом лекцій та рекомендованою літературою.
2. Побудувати за допомогою функціональних засобів пакету MATLAB двовимірні зрізи (проекції) згенерованої у розрахунковій частині вибірки даних за усіма можливими комбінаціями перших чотирьох ознак без повторів (тобто комбінації $\langle x_1, x_2 \rangle$ та $\langle x_2, x_1 \rangle$ розглядаються як тотожні), які зобразити на відповідних рисунках. Підписати програмно осі на рисунках.
3. Зобразити на усіх проекціях точки, що відповідають центрам виділених кластерів (виділити їх маркером іншої форми та іншим кольором, ніж екземпляри вибірки та різними кольорами – для різних методів кластер-аналізу).
4. Порівняти візуалізовані отримані результати кластер-аналізу.

Зміст розділу звіту з графічної частини

1. Текст розробленої програми, що візуалізує результати кластер-аналізу.
2. Рисунки (екранні форми), що зображують зрізи вибірки та центри виділених кластерів.
3. Висновки. У висновках треба проаналізувати результати роботи, а також лаконічно відповісти на контрольні питання.
4. Висновки. У висновках треба проаналізувати результати роботи, а також лаконічно відповісти на контрольні питання з розрахункової та графічної частин.

Контрольні питання

1. Засоби двовимірної графіки пакету MATLAB.
2. Засоби тривимірної графіки пакету MATLAB.
3. Порівняння графічних засобів пакету MATLAB та інтегрованих середовищ розробки програм Visual Studio.

З РЕФЕРАТИВНА (ФАКУЛЬТАТИВНА) ЧАСТИНА ЗАВДАННЯ

У випадку, якщо студент бажає більш глибоко засвоїти курс лекцій та підвищити оцінку з курсу, пропонується виконати факультативну (необов'язкову) частину розрахунково-графічного завдання шляхом підготовки короткого реферату з курсу.

Реферат з курсу готується за рекомендованими та іншими знайденими студентом джерелами і оформлюється як розділ звіту з розрахунково-графічного завдання та розміщується перед висновками.

Обсяг реферативної частини має становити 3–5 сторінок без урахування рисунків і таблиць, а також текстів програм, що займають повну сторінку звіту.

До виконання реферативної частини студент обирає тему з наведеного у додатку В переліку (або формулює власну тему) і узгоджує її з викладачем.

У переліку посилань звіту з розрахунково-графічного завдання мають бути обов'язково наведені усі використані студентом при написанні реферату джерела, упорядковані за згадуванням у тексті, а у тексті реферативної частини – наведені посилання на відповідні джерела (наводяться у квадратних дужках).

Оцінка за реферат проставляється тільки у випадку його наявності у звіті і ставиться окремо від основної оцінки за розрахунково-графічне завдання за двобальною системою:

– "не зараховано" – у випадку низького рівня виконання або плагіату реферату;

– "зараховано" – у випадку прийнятного рівня виконання.

Якщо оцінка за основну (обов'язкову) частину звіту є нижчою від "задовільно", а оцінка з реферативної частини – "зараховано", тоді загальна оцінка за виконання розрахунково-графічного завдання встановлюється як вища на один рівень від оцінки за основну частину. У випадку, якщо основну (обов'язкову) частину звіту виконано на "відмінно", то оцінка "зараховано" враховується як 10 балів за стобальною системою у загальній сумі балів з відповідного модуля курсу.

ЛІТЕРАТУРА

Основна література

1. 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.

2. Олійник А. О. Інтелектуальний аналіз даних : навчальний посібник / А. О. Олійник, С. О. Субботін, О. О. Олійник. – Запоріжжя : ЗНТУ, 2012. – 271 с.

3. Зайченко Ю.П. Основи проектування інтелектуальних систем : навчальний посібник / Ю.П. Зайченко. – К.: Слово, 2004. – 352 с.

Додаткова література

4. Рідкокаша А. А. Основи систем штучного інтелекту : навчальний посібник / А.А.Рідкокаша, К.К. Голдер. – Черкаси, "Відлуння–Плюс", 2002.–240 с.

5. Леоненков А. В. Нечеткое моделирование в среде MATLAB / А. В. Леоненков. – СПб.: БХВ–Петербург, 2003. – 736 с.

6. Дубровин В.И. Интеллектуальные средства диагностики и прогнозирования надежности авиадвигателей : монография / [В.И. Дубровин, С.А. Субботин, А.В. Богуслаев, В.К. Яценко] .– Запорожье: ОАО "Мотор–Сич", 2003.– 279 с.

7. Кричевский М.Л. Интеллектуальные методы в менеджменте. / М.Л. Кричевский – СПб.: Питер, 2005.– 304 с.

8. 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.

9. 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.

10. Интеллектуальные информационные технологии проектирования автоматизированных систем диагностирования и распознавания образов : монография / С. А. Субботин, Ан. А. Олейник, Е. А. Гофман, С. А. Зайцев, Ал. А. Олейник ; под ред. С. А. Субботина. – Харьков. : Компания СМІТ, 2012. – 318 с.

Додаток А

**ПРИКЛАД ОФОРМЛЕННЯ ТИТУЛЬНОГО АРКУШУ
ЗВІТУ ПРО ВИКОНАННЯ
РОЗРАХУНКОВО-ГРАФІЧНОГО ЗАВДАННЯ**

Міністерство освіти і науки, молоді та спорту України
Запорізький національний технічний університет

Кафедра програмних засобів

ЗВІТ

про виконання розрахунково-графічного завдання
з дисципліни "Назва курсу"
варіант № ____
на тему
"Формулювання теми реферату"

Виконав: ст. гр. ІОТ-410

І. І. Іванов

Прийняв: доцент, к.т.н.

П. П. Петренко

2012

Рисунок А.1 – Титульний аркуш звіту

Додаток Б

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 належності, що містить рівні належності 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 кла-
стер-аналіз
plot(data(:,1), data(:,2),'o'); % зображуємо дані
на графіку
hold on;
maxU = max(U);
% Знаходимо точки з найвищим рівнем належності до
першого кластера
index1 = find(U(1,:) == maxU);
% Знаходимо точки з найвищим рівнем належності до
другого кластера
index2 = find(U(2,:) == maxU);
line(data(index1,1),data(index1,2),'marker','*','c
olor','g');
line(data(index2,1),data(index2,2),'marker','*','c
olor','r');
% Зображуємо центри кластерів на графіку
plot([center([1
2],1)], [center([1
2],2)], '*','color','k');

```

Додаток В

ПЕРЕЛІК ТЕМ ДЛЯ ВИКОНАННЯ РЕФЕРАТИВНОЇ ЧАСТИНИ РОЗРАХУНКОВО-ГРАФІЧНОГО ЗАВДАННЯ

1. Архітектура інтелектуальної системи.
2. Історія розвитку штучного інтелекту.
3. Технології інтелектуального аналізу даних.
4. Класифікація методів кластер-аналізу.
5. Самоорганізація та навчання без учителя.
6. Методи оптимізації в задачах розпізнавання образів.
7. Методи витягу знань з даних.
8. Методи витягу знань з експертів.
9. Моделі подання знань.
10. Пошук у просторі станів.
11. Експертні системи.
12. Вирішення завдань технічного і біомедичного діагностування на основі методів штучного інтелекту.
13. Інтелектуальний пошук та аналіз інформації в комп'ютерних банках даних.
14. Програмні засоби з елементами штучного інтелекту.
15. Дескриптивні моделі об'єктів управління і проектування.
16. Методи обробки результатів статистичного моделювання.
17. Критерії оцінювання точності математичних моделей.
18. Стратегії та методи виведення для моделей подання знань.
19. Типологія та застосування інтелектуальних агентів.
20. Формально-логічні засади систем штучного інтелекту.
21. Типи моделей штучного інтелекту.
22. Алгоритми та розв'язання задач штучного інтелекту.
23. Відмінності між звичайними та інтелектуальними системами.
24. Принципи побудови систем штучного інтелекту.
25. Основні визначення та методи теорії розпізнавання образів.
26. Принципи створення гібридних інтелектуальних систем.
27. Критерії аналізу алгоритмів штучного інтелекту.