

**НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ БІОРЕСУРСІВ І
ПРИРОДОКОРИСТУВАННЯ УКРАЇНИ**

Кафедра автоматики та робототехнічних систем ім. акад. І.І. Мартиненка

“ЗАТВЕРДЖУЮ”

Директор ННІ енергетики,
автоматики і енергозбереження

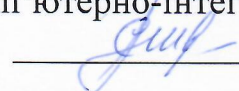

_____ (Каплун В.В.)
“17” 08 2024 р.

“СХВАЛЕНО”
на засіданні кафедри
автоматики та робототехнічних
систем ім. акад. І. І. Мартиненка
Протокол № 1 від 15.08.2024 р.

В.о. завідувача кафедри
_____ (Опришко О.О.)

”РОЗГЛЯНУТО”

Гарант ОНП підготовки докторів філософії зі
спеціальності 151 Автоматизація та
комп'ютерно-інтегровані технології


_____ (Шворов С.А.)

РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

**Обробка інформації в комп'ютерно-інтегрованих
системах автоматизації**

(назва навчальної дисципліни)

Галузь знань 15 «Автоматизація та приладобудування»

Спеціальність 151 Автоматизація та комп'ютерно-інтегровані технології

ОНП Автоматизація та комп'ютерно-інтегровані технології

ННІ Енергетики, автоматики і енергозбереження

Розробник: проф. каф., д.т.н., проф. Коваль В.В., доц. каф., доц. Кіктєв М.О.

Київ – 2024 р.

1. Опис навчальної дисципліни
ОБРОБКА ІНФОРМАЦІЇ В КОМП'ЮТЕРНО-ІНТЕГРОВАНИХ
СИСТЕМАХ АВТОМАТИЗАЦІЇ

Галузь знань, напрям підготовки, спеціальність, освітньо-кваліфікаційний рівень		
Галузь знань	15 Автоматизація та приладобудування	
Спеціальність	151 Автоматизація та комп'ютерно-інтегровані технології	
Рівень вищої освіти	Третій (освітньо-науковий) рівень	
Характеристика навчальної дисципліни		
Вид	Вибіркова	
Загальна кількість годин	90	
Кількість кредитів ECTS	3	
Кількість змістових частин	2	
Форма контролю	Екзамен	
Показники навчальної дисципліни для денної та заочної форм навчання		
	денна (вечірня) форма навчання	заочна форма навчання
Рік підготовки (курс)	1	1
Семестр	1	1
Лекційні заняття	10 год.	5 год.
Практичні, семінарські заняття	20 год.	10 год.
Лабораторні заняття	-	-
Самостійна робота	60 год.	75 год.
Кількість тижневих аудиторних годин для денної форми навчання:	3	

2. Мета та завдання навчальної дисципліни

Предметом дисципліни є методи обробки та аналізу даних в складних системах автоматизації.

Метою вивчення дисципліни “Обробка інформації в комп'ютерно-інтегрованих системах автоматизації” є теоретична і практична підготовка аспірантів щодо отримання ними знань з обробки та аналізу даних в системах автоматизації складних біотехнічних об'єктів.

Завдання дисципліни – вивчення методів обробки і аналізу даних в сучасних систем автоматизації аграрного спрямування.

Вивчення навчальної дисципліни передбачає формування у аспірантів компетентностей (та їх складових):

Інтегральна компетентність: Здатність продукувати нові ідеї, розв'язувати комплексні проблеми професійної та/або дослідницько-інноваційної діяльності у сфері автоматизації та комп'ютерно-інтегрованих технологій, застосовувати методологію наукової та педагогічної діяльності, а також проводити власне наукове дослідження, результати якого мають наукову новизну, теоретичне та практичне значення.

Загальні компетентності:

ЗК2. Здатність до пошуку, оброблення та аналізу інформації з різних джерел.

Спеціальні (фахові) компетентності:

СК1. Здатність виконувати оригінальні дослідження, досягати наукових результатів, які створюють нові знання у сфері автоматизації, комп'ютерно-інтегрованих технологій, керування складними організаційно-технічними чи кіберфізичними системами та дотичних до неї міждисциплінарних напрямках і можуть бути опубліковані у провідних наукових виданнях;

СК5. Здатність створювати новітні системи автоматизації, комп'ютерно-інтегровані технології, розробляти їх технічне, інформаційне, математичне, програмне та організаційне забезпечення із застосуванням сучасних інформаційних технологій, інструментів та компонентів.

Програмні результати навчання:

РН1. Мати передові концептуальні та методологічні знання з автоматизації та комп'ютерно-інтегрованих технологій та з дотичних міждисциплінарних напрямів, розуміти методологію наукових досліджень. Уміти застосовувати їх у власних дослідженнях, скерованих на отримання нових знань та/або здійснення інновацій, та у викладацькій практиці;

РН4. Планувати і виконувати експериментальні та/або теоретичні дослідження систем автоматизації, комп'ютерно-інтегрованих комплексів та їх складових з використанням сучасних методів дослідження, технічних та програмних засобів та з дотриманням норм академічної і професійної етики. Формулювати і перевіряти гіпотези; використовувати для обґрунтування висновків результати теоретичного аналізу, експериментальних досліджень і математичного та/або комп'ютерного моделювання, наявні літературні дані;

РН7. Застосовувати сучасні цифрові технології, мікропроцесорні засоби, мехатронні компоненти, спеціалізоване програмне забезпечення для створення новітніх систем автоматизації, комп'ютерно-інтегрованих технологій, їх технічного, інформаційного, математичного, програмного та організаційного забезпечення.

РН8. Застосовувати сучасні інструменти і технології пошуку, оброблення та аналізу інформації, зокрема, статистичні методи аналізу даних, спеціалізовані бази даних та інформаційні системи.

У результаті вивчення навчальної дисципліни аспіранти повинні **знати:**

основні положення системного аналізу об'єктів автоматизації, методи попереднього аналізу та обробки даних в складних системах автоматизації і алгоритми для реалізації вивчених методів.

Аспіранти повинні **вміти:** використовувати методи попереднього аналізу та обробки даних в складних системах автоматизації біотехнічних об'єктів.

3. Структура навчальної дисципліни

Назви змістових розділів і тем	Кількість годин								
	денна форма					Заочна форма			
	Тиж-ні	Усьо-го	у тому числі			Усьо-го	у тому числі		
			л	пр	с.р.		л	пр	с.р.
1	2	3	4	5	6	7	8	9	10
Змістовий розділ 1. «Загальні питання підготовки та попереднього аналізу даних»									
Тема 1. Методи попереднього аналізу та зниження розмірностей багатовимірних даних.	1	18	2	4	12	18	1	2	15
Тема 2. Методи обробки багатовимірних даних у OLAP-системі	4	18	2	4	12	18	1	2	15
Тема 3. Класифікація даних з використанням детермінованих моделей	5	18	2	4	12	18	1	2	15
Разом за змістовим розділом 1		54	6	12	36	54	3	6	45
Змістовий розділ 2. «Класифікація даних на основі статистичних моделей»									
Тема 4. Методи статистичного оцінювання і порівняння вибірок. Непараметричні методи перевірки однорідності вибірок	7	18	2	4	12	18	1	2	15
Тема 5. Дисперсійний аналіз.	8	18	2	4	12	18	1	2	15
Разом за змістовим розділом 2		36	4	8	24	36	4	4	30
Усього годин		90	10	20	60	90	5	10	75

4. Теми практичних занять

№ з/п	Назва теми	Кількість годин
1	Збір та підготовка даних	2 год.
2	Регресійний аналіз	4 год.
3	Бінарна класифікація	2 год.
4	Кластерний аналіз	4 год.
5	Порівняння і аналіз двох вибірок з використанням критеріїв Стюдента і хв-квадрат	2 год.
6	Робота в середовищі мови R	4 год.
7	Статистична обробка даних на мові R	2 год.
	Разом:	20 год.

5. Самостійна робота

№ з/п	Назва теми	Кількість годин
1	Вимоги до даних.	6
2	Сховища даних	6
3	Методи моделювання випадкових величин.	6
4	Робастне статистичне оцінювання	6
5	Методи багатовимірної класифікації даних.	6
6	Методи зниження розмірностей даних	6
7	Методи проведення експертних досліджень і аналіз даних оцінок експертів	6
8	Історія і основні принципи організації середовища R для аналізу і обробки даних	6
9	Робота з даними в R. Організація обчислень в R	6
10	Аналіз великих даних з використанням ресурсів середовища програмування R.	6
	Разом	60

6. Контрольні питання, комплекти тестів для визначення рівня засвоєння знань здобувачами

1. Загальна концепція методології Data Mining і технології реалізації
2. Огляд завдань, що вирішуються методами Data Mining.
3. Класифікація методів Data Mining по різним ознакам.
4. Етапи інтелектуального аналізу даних: аналіз предметної області, постановка задачі, підготовка даних.
5. Процес підготовки даних, поняття якості даних, брудних даних, етапи очищення даних.
6. Етапи процесу Data Mining, пов'язані з побудовою, перевіркою, оцінкою, вибором і корекцією моделей.
7. Процес Data Mining як послідовність етапів і як послідовність робіт, що виконуються виконавцями ролей Data Mining.
8. Аналіз взаємозв'язків (кореляцій) між змінними - коефіцієнти кореляції Пірсона, Спірмена і Кендалла.
9. Порівняння груп – критерії Стюдента, Манна - Уїтні і дисперсійний аналіз (ANOVA).
10. Інструменти очищення і редагування даних, основні функції інструментів очищення даних,
11. Класифікація помилок в даних, які виникають в результаті використання коштів очищення даних.
12. Інструменти очищення даних.
13. Завдання класифікації та методи вирішення.
14. Постановка задач, ключові поняття і визначення.
15. Методи, що застосовуються для вирішення завдань класифікації:

індукція дерев рішень.

16. Байєсовські мережі (Bayesian Networks) для задач класифікації
17. Нейронні мережі (neural networks) для задач класифікації
18. Статистичні методи для задач класифікації
19. Лінійна регресія для задач класифікації .
20. Переваги дерев рішень.
21. Інтуїтивність дерев рішень.
22. Точність дерев рішень.
23. Процес конструювання дерева рішень.
24. Метод "найближчого сусіда"
25. Регресійний аналіз. Послідовність етапів регресійного аналізу
26. Класичні методи регресійного аналізу: множинна і логістична регресії, вибір змінних для аналізу. Встановлення форми залежності.
27. Визначення функції регресії. Оцінка невідомих значень залежної змінної.
28. Рівняння регресії. Випадкове відхилення.
29. Коефіцієнт множинної кореляції R
30. Порівняння якості побудованих моделей.
31. Завдання кластеризації: постановка завдання, ключові поняття і визначення.
32. Метод k-середніх
33. ЕМ-алгоритм.
34. Завдання зниження розмірності,
35. Метод незалежних компонент (independent component analysis).
36. Міри схожості. Ієрархічний кластерний аналіз в SPSS.
37. Алгоритм k-середніх (kmeans).
38. Класичні методи кластеризації – Метод ієрархічної кластеризації (tree clustering).
39. Класичні методи зниження розмірності: метод головних компонент (principal component analysis).
40. Факторний аналіз (Factor Analysis).

7. Методи навчання:

- словесний метод (лекції, індивідуальне консультування);
- практичний метод (виконання практичних завдань, визначених по кожній темі, включаючи роботи з використанням доступних прикладних програмних продуктів - Excel, R тощо, з обов'язковим орієнтуванням на тему дослідження);
- метод навчання через дослідження;
- робота з навчально-методичною літературою (із розкриттям суті при спілкуванні).

8. Форми контролю

Поточний контроль знань здобувачів вищої освіти ступеня доктора філософії проводиться шляхом опитування (за результатами опрацьованого матеріалу).

Підсумковий контроль знань проводиться у письмовій формі, з подальшою усною співбесідою. Контроль знань здобувачів виконується під час приймання результатів виконання практичних робіт за допомогою контрольних запитань. Контроль знань в кінці розділів здійснюється за допомогою тестів.

Форми та методи проведення поточного та підсумкового контролю розробляються лектором дисципліни.

9. Методичне забезпечення

Науково-методичне забезпечення навчального процесу передбачає:

- державні стандарти, навчальні плани, конспект лекцій; методичні вказівки до виконання практичних занять; електронні варіанти тестів для поточного і підсумкового контролю;

- електронний навчальний курс навчальної дисципліни (на навчальному порталі НУБіП України eLearn - <https://elearn.nubip.edu.ua/course/view.php?id=4875>).

10. Рекомендована література

Основна

1. Конспект лекцій з навчальної дисципліни Обробка інформації в комп'ютерно-інтегрованих системах автоматизації для здобувачів третього (освітньо-наукового) рівня вищої освіти за спеціальністю 151 «Автоматизація та комп'ютерно-інтегровані технології» (кваліфікація: PhD доктор філософії). К:Видавничий центр НУБіП, 2022. – 72 с.
2. Методичні вказівки до проведення практичних занять з навчальної дисципліни Обробка інформації в комп'ютерно-інтегрованих системах автоматизації для здобувачів третього (освітньо-наукового) рівня вищої освіти за спеціальністю 151 «Автоматизація та комп'ютерно-інтегровані технології» (кваліфікація: PhD доктор філософії). К:Видавничий центр НУБіП, 2022. – 43 с.
3. Machine Learning Models and Algorithms for Big Data Classification / S. Suthaharan – Integrated Series in Information Systems 36, Springer Science+Business Media New York, 2019. – 359 pp.
4. Статистический анализ и визуализация данных с помощью R / С.Э. Мاستицкий, В.К. Шитиков – Хайдельберг – Лондон – Тольятти. 2014. – 401 с. 7. R в действии. Анализ и визуализация данных в программе R / Кабаков Р. И. пер. с англ. – К.: ДМК Пресс, 2017. – 588 с.
5. R in a Nutshell / J. Adler. – Second Edition. O'Reilly Media, Inc. 2017. 722 pp.
6. Аналіз даних за допомогою програми Excel для Windows. Режим доступу: <https://support.microsoft.com/uk-ua/office/%D0%B0%D0%BD%D0%B0%D0%BB%D1%96%D0%B7-%D0%B4%D0%B0%D0%BD%D0%B8%D1%85-%D0%B7%D0%B0-%D0%B4%D0%BE%D0%BF%D0%BE%D0%BC%D0%BE%D0%B3%D0%BE%D1%8E-%D0%BF%D1%80%D0%BE%D0%B3%D1%80%D0%B0%D0%BC%D0%B8->

excel-%D0%B4%D0%BB%D1%8F-windows-9f66d632-5262-4752-8675-505b6cd4eb58

7. Бахрушин В.Є. Методи аналізу даних : навчальний посібник для студентів / В.Є. Бахрушин. – Запоріжжя : КПУ, 2011. – 268 с.
8. Роїк М. В. Огляд програмних засобів статистичного аналізу даних / М. В. Роїк, О. І. Присяжнюк, В. О. Денисюк / Ефективна економіка № 7, 2017. Режим доступу: <http://www.economy.nayka.com.ua/?op=1&z=5676>.
9. Технології оброблення великих даних: конспект лекцій з дисципліни «Технології оброблення великих даних» [Електронний ресурс] : навч. посіб. / Л.М. Олещенко; КПП ім. Ігоря Сікорського. – Електронні текстові дані (1 файл: 5,55 Мбайт). – Київ: КПП ім. Ігоря Сікорського, 2021. – 227 с. Режим доступу: https://ela.kpi.ua/bitstream/123456789/42206/1/%D0%9AonspLekts_Tekhnolohii-obroblennia-velykykh-danykh_%D0%9Eleshchenko.pdf

Допоміжна

1. Learning Data Mining with R / B. Makhabel – Packt Publishing. 2021. – 314 pp. Режим доступу: <https://smartboost.com/blog/data-mining-with-r-or-python/>
2. Bramer, M. Principles of Data Mining / M. Bramer. – Third edition. London : Springer-Verlag London Ltd. 2016. – 526 p.
3. Aggarwal, C.C. Computational Intelligence in Data Mining: / C.C. Aggarwal. – Springer International Publishing Switzerland. 2017. 734 P. Режим доступу: <https://books.google.com.ua/books?id=S3ckDwAAQBAJ&pg=PA271&lpg=PA271&dq=Aggarwal,+C.C.+Data+Mining:+The+Textbook&source=bl&ots=XMJiKFDwZU&sig=ACfU3U0LWpa-WTjbzqnWGwGBxnCQiS3IeQ&hl=uk&sa=X&ved=2ahUKEwje8aePwIP2AhWQyqQKHfMxBVQ4MhDoAXoECAsQAQw#v=onepage&q=Aggarwal%2C%20C.C.%20Data%20Mining%3A%20The%20Textbook&f=false>
4. Bishop, C. M., Pattern recognition and machine learning / C. M. Bishop. – New York: Springer, 2016. – 738 p.
5. Ланде Д.В. Оброблення надвеликих масивів даних (Big Data): навчальний посібник / Ланде Д.В., Субач І.Ю., Гладун А.Я. Київ 2021. – 168 с.

11. Інформаційні ресурси

<https://wikipedia.org>
<https://www.kpi.kharkov.ua/ukr/> – головна сторінка ХПІ
<http://www.google.com.ua> – (пошуковий сайт).
<https://nubip.edu.ua/> – головна сторінка НУБіП України.
<https://nubip.edu.ua/departament/naukova-biblioteka>
<http://elibrary.nubip.edu.ua> – електронна наукова бібліотека НУБіП України.
<http://www.nbuv.gov.ua/> – Національна бібліотека України імені В.І. Вернадського, Київ.