

**МІЖНАРОДНИЙ НАУКОВО-НАВЧАЛЬНИЙ ЦЕНТР
ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ ТА СИСТЕМ
НАН УКРАЇНИ ТА МОН УКРАЇНИ**

ЗАТВЕРДЖЕНО

Директор Міжнародного науково-навчального центру інформаційних технологій та систем НАН та МОН України

Олександр БОЛКОВ
«23» квітня 2024 р



РОБОЧА НАВЧАЛЬНА ПРОГРАМА

**З дисципліни «МЕТОДИ ІНТЕЛЕКТУАЛЬНОГО ОБРОБЛЕННЯ ТА
АНАЛІЗУ ДАНИХ»**

Рівень вищої освіти третій
Ступінь вищої освіти доктор філософії
Галузь знань 12 – інформаційні технології
Спеціальність 122 – комп'ютерні науки

Шифр ДВЦ 2 Дисципліна за вибором Міжнародного Центру

Форма навчання _____ денна _____ Курс 2 Семестр 3

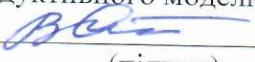
Всього годин /кредитів ЄКТС 90 /3 за навчальним планом

- лекції (Л) 22
- семінарські заняття (СЗ) 16
- практичні заняття (ПЗ) —
- індивідуальні заняття (ІЗ) 2
- самостійна робота студентів (СРС) 50
- підсумковий контроль дисципліни – іспит

м. Київ

Укладач(і) робочої навчальної програми:


доктор технічних наук, професор, завідувач відділу інформаційних технологій
індуктивного моделювання

 Володимир СТЕПАШКО

(підпис)

e-mail: shepetukha@irtc.org.ua

Робочу програму погоджено з гарантом освітньо-наукової програми

Гарант освітньої програми  / Володимир СТЕПАШКО
(підпис)

Затверджено: Вченою радою Міжнародного науково-навчального центру інформаційних
технологій та систем НАН та МОН України
Протокол № 3 від 23.04.2024 р.

Вчений секретар Вченої ради



Микола КОМАР

Ухвалено: Науково-методичною радою Міжнародного науково-навчального центру
інформаційних технологій та систем НАН та МОН України
Протокол № 2 від 15.04.2024 р.

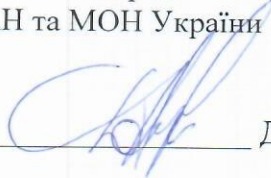
Голова Науково-методичної ради



Людмила КОЗАК

Ухвалено: Радою молодих вчених Міжнародного науково-навчального центру
інформаційних технологій та систем НАН та МОН України
Протокол № 4 від 15.04.2024 р.

Голова Ради молодих вчених



Дмитро ВОЛОШЕНЮК

Введено в дію наказом директора Міжнародного науково-навчального центру
інформаційних технологій та систем НАН та МОН України № 57 від 23.04.2024 р.

1.ЗАГАЛЬНІ ВІДОМОСТІ

Найменування показників	Характеристика дисципліни за денною формою навчання
Вид дисципліни	за вибором Міжнародного Центру
Мова викладання, навчання та оцінювання	українська
Загальний обсяг кредитів / годин	3 / 90
Курс	2-й
Семестр	3-й
Кількість змістових модулів з розподілом:	3
Обсяг кредитів	3
Обсяг годин, в тому числі:	90
Лекції	22
Семінарські заняття	16
Самостійна робота	50
Форма підсумкового контролю	іспит

Дисципліна «Методи інтелектуального оброблення та аналізу даних» належить до переліку дисциплін циклу професійної підготовки за вибором Міжнародного Центру (обов'язкова). Вона забезпечує важливий аспект професійного світогляду аспіранта і спрямована на формування вміння розробляти й використовувати в наукових дослідженнях сучасні методи, засоби та інформаційні технології оброблення і аналізу статистичних, експериментальних, дослідних даних, що породжуються економічними, екологічними, технічними та медико-біологічними об'єктами різного рівня в процесі їхнього функціонування і містять релевантну інформацію про поточний стан, внутрішні закономірності та тенденції розвитку цих об'єктів.

2. МЕТА І ЗАВДАННЯ ДИСЦИПЛІНИ

Метою викладання навчальної дисципліни «Методи інтелектуального оброблення та аналізу даних» є навчити аспіранта формулювати та розв'язувати задачі оброблення, і аналізу реальних даних, що породжуються економічними, екологічними, технічними та медико-біологічними об'єктами різного рівня в процесі їхнього функціонування і містять релевантну інформацію про поточний стан, внутрішні закономірності та тенденції розвитку цих об'єктів, ефективно працювати з інформаційними джерелами, створювати нові знання за допомогою проведення оригінальних теоретичних та експериментальних досліджень із застосуванням цих методів.

Основними завданнями є: 1) ознайомлення з основними напрямками дослідження складних динамічних системами та методологією побудови математичних моделей керування такими системами; 2) ознайомлення з принципами прийняття рішень в інтелектуалізованих системах та підходами до їх застосування в задачах керування; 3) ознайомлення з основними шляхами застосування методів штучного інтелекту в інформаційних системах та технологіях, спрямованих на вирішення проблем інтелектуального керування. Здатність продукувати нові ідеї, розв'язувати комплексні проблеми у сфері комп'ютерних наук, застосовувати методологію наукової та педагогічної діяльності, а також проводити власне наукове дослідження, результати якого мають наукову новизну, теоретичне та практичне значення.

Інтегральна компетентність

Здатність продукувати нові ідеї, розв'язувати комплексні проблеми у сфері комп'ютерних наук, застосовувати методологію наукової та педагогічної діяльності, а також проводити власне наукове дослідження, результати якого мають наукову новизну, теоретичне та практичне значення.

Загальні компетентності:

ЗК01. Здатність до абстрактного мислення, аналізу та синтезу.

ЗК02. Здатність до пошуку, оброблення та аналізу інформації з різних джерел.

ЗК04. Здатність розв'язувати комплексні проблеми комп'ютерних наук на основі системного наукового світогляду та загального культурного кругозору з дотриманням принципів професійної етики та академічної доброчесності.

ЗК05. Здатність до спілкування з колегами, широким академічним товариством та громадськістю українською та однією з іноземних мов європейського простору.

ЗК07. Здатність виявляти проблеми, формалізувати ідеї, визначати мету, ставити та розв'язувати завдання.

Фахові компетентності:

СК01. Здатність виконувати оригінальні дослідження, досягати наукових результатів, які створюють нові знання у комп'ютерних науках та дотичних до них міждисциплінарних напрямках і можуть бути опубліковані у провідних наукових виданнях з комп'ютерних наук та суміжних галузей.

СК03. Здатність виявляти, ставити та розв'язувати дослідницькі науково-прикладні завдання та/або вирішувати проблеми в сфері комп'ютерних наук, оцінювати та забезпечувати якість виконуваних досліджень.

СК06. Здатність аналізувати та оцінювати сучасний стан і тенденції розвитку комп'ютерних наук та інформаційних технологій.

СК 07. Здатність до створення та адекватного застосування інтелектуальних методів аналізу інформації та керування складними динамічними об'єктами різної природи.

СК 08. Здатність аналізувати дані та одержувати необхідні знання для розв'язання нестандартних завдань з використанням математичних методів та методів комп'ютерного моделювання.

СК09. Здатність до розвитку теорії, розроблення нових методів та інтелектуальних засобів для розв'язання прикладних завдань з різних сфер, зокрема, технічної, економічної, екологічної, медичної, біологічної тощо.

СК10. Здатність до проведення наукових досліджень з інтелектуального оброблення, аналізу та інтерпретації інформації про об'єкти різної природи.

Програмні результати навчання

РН01. Мати передові концептуальні та методологічні знання з комп'ютерних наук і на межі предметних галузей, а також дослідницькі навички, достатні для проведення

наукових і прикладних досліджень на рівні останніх світових досягнень з відповідного напрямку, отримання нових знань та/або здійснення інновацій.

PH03. Формулювати і перевіряти гіпотези; використовувати для обґрунтування висновків належні докази, зокрема, результати теоретичного аналізу, експериментальних досліджень і математичного та/або комп'ютерного моделювання, наявні літературні дані.

PH04. Розробляти та досліджувати концептуальні, математичні і комп'ютерні моделі процесів і систем, ефективно використовувати їх для отримання нових знань та/або створення інноваційних продуктів у комп'ютерних науках та дотичних міждисциплінарних напрямках.

PH06. Застосовувати сучасні інструменти і технології пошуку, оброблення та аналізу інформації, зокрема, статистичні методи аналізу даних великого обсягу та/або складної структури, спеціалізовані бази даних та інформаційні системи.

PH08. Визначати актуальні наукові та практичні проблеми у сфері комп'ютерних наук, глибоко розуміти загальні принципи та методи комп'ютерних наук, а також методологію наукових досліджень, застосувати їх у власних дослідженнях у сфері комп'ютерних наук та у викладацькій практиці.

PH13. Здійснювати розроблення нових методів та інтелектуальних засобів для розв'язання прикладних завдань з різних сфер, зокрема, технічної, економічної, екологічної, медичної, біологічної тощо.

PH15. Проводити інтелектуальний аналіз складних об'єктів за різними видами первинної інформації (зображення, складні сигнали, тексти, електронні медичні записи, відео та аудіо записи).

PH 17. Розробляти комп'ютерні системи оброблення та аналізу інформації різного виду (цифрової, текстової, зображень, відеоряду, сигналів тощо).

3. СТРУКТУРА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Номер лекції	Назва лекції	Кількість годин		
		Лекції	Семінарські /практичні заняття	Самостійна робота
ЗМ1: Проблема аналізу даних різної природи (1 кредит)				
1	<i>Дані різної природи як джерело інформації та знань</i> Природа походження даних, їх різновиди та характеристики. Задачі попереднього оброблення даних.	2	-	4
2	<i>Аналіз даних як задача моделювання складних процесів і систем</i> Моделювання як категорія пізнання світу. Дедуктивний та індуктивний підходи до побудови моделей. Основні етапи процесу моделювання. Типові задачі моделювання та розпізнавання.	4	2	6

ЗМ2: Методи інтелектуального аналізу даних, обчислювального інтелекту та м'яких обчислень (1 кредит)				
5	<i>Методи інтелектуального аналізу даних</i> Методи математичної статистики. Штучні нейронні мережі. Еволюційні та генетичні методи. Машинне та глибинне навчання.	4	4	10
6	<i>Методи обчислювального інтелекту та м'яких обчислень в задачах аналізу даних</i> Жорсткі та м'які обчислення. Методи нечіткої логіки. Нечіткі нейромережі та їх навчання. Ройові алгоритми в задачах моделювання.	4	4	10
ЗМ3: Проблема конструювання інтелектуальних інформаційних технологій оброблення, аналізу та інтерпретації даних (1 кредит)				
8	<i>Індуктивне моделювання як сучасний напрямок інтелектуального аналізу даних</i> Концепція самоорганізації моделей, постановка задачі індуктивного моделювання. Основні результати теорії, Класифікація типових алгоритмів та критеріїв. Сучасні тенденції розвитку наукового напрямку.	4	4	10
9	<i>Прикладні аспекти розроблення і застосування інформаційних технологій оброблення, аналізу та інтерпретації даних</i> Концепція інтелектуального моделювання. Формування бази знань про процес моделювання. Інтелектуальний інтерфейс користувача засобів аналізу. Агрегування та візуалізація результатів аналізу даних.	4	2	10
	ВСЬОГО	22	16	50

4. ПРОГРАМНИЙ МАТЕРІАЛ

Змістовий модуль 1. Проблема аналізу даних різної природи

Тема 1. Дані різної природи як джерело інформації та знань

Природа походження даних: вимірювання, статистика, експерименти. Різновиди даних: числові, порядкові, рангові, текстові, зображення. Характеристики даних: повнота, релевантність, обсяг, точність, зашумленість.

Задачі роботи з даними: первинне опрацювання, усунення недоліків, збереження в БД, пошук в БД, візуалізація, підготовка даних для аналізу. Сучасна концепція Великих даних.

Тема 2. Аналіз даних як задача моделювання складних процесів і систем

Моделювання як фундаментальна категорія пізнання світу. Два протилежні підходи до побудови моделей: на основі теорії (дедуктивний) та на основі даних (індуктивний). Основні етапи процесу моделювання. Типові різновиди задач моделювання як задач оптимізації.

Типові задачі моделювання в рамках індуктивного аналізу даних різної природи: розвідувальний аналіз даних; побудова (структурно-параметрична ідентифікація) моделей з даних; розпізнавання образів з учителем (класифікація) та без учителя (кластеризація).

Змістовий модуль 2. Методи інтелектуального аналізу даних, обчислювального інтелекту та м'яких обчислень

Тема 3. Методи інтелектуального аналізу даних

Методи математичної статистики: кореляційний, факторний, регресійний аналіз, перевірка гіпотез. Виведення за прецедентами та дерева рішень. Штучні нейронні мережі: основні парадигми та архітектури, методи навчання.

Еволюційні та генетичні методи моделювання на основі даних. Машинне навчання з даних як напрямок штучного інтелекту. Глибинне навчання як сучасна тенденція розвитку нейромереж.

Тема 4. Методи обчислювального інтелекту та м'яких обчислень в задачах аналізу даних

Жорсткі та м'які обчислення в задачах прийняття рішень. Співвідношення між методами обчислювального інтелекту та м'яких обчислень.

Математичні методи нечіткої логіки: різновиди введення нечіткості, фазифікація та дефазифікація даних. Типові архітектури нечітких нейромереж та методи їх навчання.

Ройові алгоритми в задачах моделювання як парадигми реалізації групового інтелекту: алгоритми бджолиних та мурашиних колоній.

Змістовий модуль 3. Проблема конструювання інтелектуальних інформаційних технологій оброблення, аналізу та інтерпретації даних

Тема 5. Індуктивне моделювання як сучасний напрямок інтелектуального аналізу даних

Історичні аспекти розвитку наукового напрямку. Концепція і принципи самоорганізації моделей за масивами даних. Означення і постановка задачі індуктивного моделювання.

Основні результати теорії методу групового врахування аргументів (МГУА). Класифікація алгоритмів МГУА перебірного та ітераційного (нейромережевого) типу. Зовнішні критерії селекції моделей та їх властивості.

Сучасні тенденції розвитку наукового напрямку.

Тема 6. Прикладні аспекти розроблення і застосування інформаційних технологій оброблення, аналізу та інтерпретації даних

Концептуальні основи інтелектуального моделювання: автономне, вбудоване, комплексне. Отримання інформації для формування бази знань про процес моделювання як основи конструювання засобів.

Проектування інтелектуального інтерфейсу користувача засобів моделювання в технологіях оброблення, аналізу та інтерпретації даних.

Засоби агрегування та візуалізації результатів аналізу даних. Засоби прогнозу аналітики для задач інформаційної підтримки прийняття управлінських рішень.

5. САМОСТІЙНА РОБОТА

Самостійна робота охоплює:

- 1) підготовку до семінарських занять,
- 2) опрацювання наукової літератури,
- 3) підготовку до іспиту.

№ п/п	Зміст самостійної роботи	Обсяг СР (годин)
1.	Підготовка до семінарських занять	10
2.	Опрацювання наукової літератури	20
3.	Підготовка до іспиту	20
Усього за навчальною дисципліною		50

6. РЕЙТИНГОВА СИСТЕМА ОЦІНЮВАННЯ РЕЗУЛЬТАТІВ НАВЧАННЯ

Контроль знань аспірантів здійснюється на підставі Положення про організацію та проведення поточного і підсумкового/семестрового контролю результатів навчання здобувачів вищої освіти третього (освітньо-наукового) рівня.

За результатами поточного і підсумкового контролю формується загальна оцінка результатів (ЗО) засвоєння кожної дисципліни за формулою:

$$ZO = k_1 * \text{ПоК} + k_2 * \text{ПідК},$$

де k_1 , k_2 - коефіцієнти переведення балів поточного (ПоК) та підсумкового контролю (ПідК) відповідно; $k_1 = 0,4$, $k_2 = 0,6$.

Максимальна кількість балів у поточному контролі встановлюється таким чином:

Форми навчальної діяльності	Максимальна сумарна оцінка в балах
усне опитування	65
активна робота на заняттях	35
Всього	100

У підсумковому контролі (іспит/залік) оцінки виставляють за 100-бальною системою з перерахунком рейтингових показників нормованої в національну шкалу та шкалу ECTS

За 100-бальною шкалою	За національною шкалою		За шкалою ECTS
	Іспит	Залік	
91 – 100	відмінно	зараховано	A (відмінно)
81-90	добре		B (дуже добре)
71-80			C (добре)
65-70	задовільно		D (задовільно)
60-65			E (достатньо)
30-59	незадовільно	не зараховано	FX (незадовільно – з можливістю повторного складання)
1-29			F (неприйнятно – з обов'язковим повторним навчанням)

7. ОРІЄНТОВНИЙ ПЕРЕЛІК ЗАПИТАНЬ НА ЕКЗАМЕН

Змістовий модуль 1 – 12 запитань

1. Можливі джерела походження даних та їх різновиди.
2. Основні характеристики первинних даних.
3. Типові задачі попереднього оброблення даних.
4. Збереження та пошук в БД, візуалізація, підготовка даних для аналізу.
5. Сутність та характеристики сучасної концепції Великих даних.
6. Два підходи до побудови моделей: дедуктивний та індуктивний.
7. Основні етапи процесу моделювання.
8. Типові різновиди задач моделювання як задач оптимізації.
9. Типові задачі розвідувального аналізу даних.
10. Задача структурно-параметричної ідентифікації моделей з даних
11. Задача розпізнавання образів з учителем.
12. Задача розпізнавання образів без учителя.

Змістовий модуль 2 – 12 запитань

1. Кореляційний та факторний методи аналізу даних.
2. Регресійний аналіз як метод побудови моделей. Перевірка гіпотез.
3. Метод головних компонент у задачах аналізу даних.
4. Методи виведення за прецедентами та побудова дерев рішень.
5. Основні парадигми штучних нейронних мереж та їх навчання.
6. Еволюційні та генетичні методи оптимізації та моделювання на основі даних.
7. Машинне навчання з даних як напрямок штучного інтелекту.
8. Глибинне навчання як сучасна тенденція розвитку нейромереж.
9. Математичні методи нечіткої логіки; різновиди введення нечіткості.

10. Типові архітектури нечітких нейромереж та методи їх навчання.
11. Ройові алгоритми бджолиних колоній в задачах оптимізації та моделювання.
12. Ройові алгоритми мурашиних колоній в задачах оптимізації та моделювання.

Змістовий модуль 3 – 12 запитань

1. Концепція і принципи самоорганізації моделей за масивами даних.
2. Означення і постановка задачі індуктивного моделювання.
3. Основні результати теорії методу групового врахування аргументів (МГУА).
4. Класифікація алгоритмів МГУА перебірного типу.
5. Класифікація алгоритмів МГУА ітераційного (нейромережевого) типу.
6. Зовнішні критерії селекції моделей та їх властивості.
7. Означення та основні характеристики інтелектуального моделювання.
8. Властивості автономної системи інтелектуального моделювання.
9. Властивості вбудованої системи інтелектуального моделювання.
10. Властивості комплексної системи інтелектуального моделювання.
11. Особливості інтелектуального інтерфейсу користувача засобів моделювання.
12. Інформаційна підтримки управлінських рішень засобами агрегування та візуалізації результатів аналізу даних.

8. ПОЛІТИКА ДОБРОЧЕСНОСТІ

Виконання навчальних завдань має відповідати вимогам Кодексу академічної доброчесності Міжнародного науково-навчального центру інформаційних технологій та систем НАН та МОН України, затвердженого вченою радою Міжнародного центру 20 січня 2022 року, протокол № 1.

9. РЕКОМЕНДОВАНА ЛІТЕРАТУРА

Основна:

1. Інтелектуальний аналіз даних: навчальний посібник / А.О. Олійник, С.О. Субботін, О.О. Олійник. Запоріжжя: ЗНТУ, 2012. 278 с. Режим доступу: http://eir.zntu.edu.ua/bitstream/123456789/2155/1/Subbotin_Intellectual_data_analysis.pdf – Назва з екрану.
2. Zgurovsky M.Z., Zaychenko Yu.P. The fundamentals of computational intelligence: System approach. Studies in Computational Intelligence Book series, vol. 652. Berlin: Springer International Publishing, 2017. 389 p.
3. Schlesinger M.I., Hlaváč V. Ten Lectures on Statistical and Structural Pattern Recognition. Computational Imaging and Vision Book series, vol. 24. Berlin: Springer Verlag, 2002, 522 p.
4. Stepashko V.S. Formation and development of self-organizing intelligent technologies of inductive modeling. *Cybernetics and Computer Engineering*. 2018. № 4 (194). P. 41-60. Режим доступу: <http://kvt-journal.org.ua/tag/stepashko-v-s/> – Назва з екрану.
5. Степашко В.С. Елементи теорії індуктивного моделювання. Стан та перспективи розвитку інформатики в Україні: монографія / Кол. авторів. Київ: Наукова думка, 2010. С. 481-496.
6. Stepashko V. On the Self-Organizing Induction-Based Intelligent Modeling / In: *Advances in Intelligent Systems and Computing III*. AISC book series, Volume 871. Cham:

Springer, 2019. P. 433-448. Режим доступу: https://link.springer.com/chapter/10.1007/978-3-030-01069-0_31 – Назва з екрану.

7. Степашко В.С., Булгакова О.С., Зосімов В.В. Ітераційні алгоритми індуктивного моделювання. К.: Наукова думка, 2018. 190 с.

8. Мороз О.Г., Степашко В.С. Перебірні алгоритми індуктивного моделювання на основі генетичних операторів. К.: Освіта України, 2021. 216 с.

9. Stepashko V.S., Savchenko-Syniakova Ye.A., Pidnebesna H.A. Problem of Constructing an Ontological Metamodel of Iterative GMDH Algorithms. *Cybernetics and Computer Engineering*. 2022. № 3 (209). P. 21-33.

10. Savchenko, M., Synytsya, K., & Savchenko-Synyakova, Y. (2023). Recommendation Methods for Information Technology Support of Lifelong Learning Situations. In: *Information Technology for Education, Science, and Technics / Faure, E., Danchenko, O., Bondarenko, M., Tryus, Y., Bazilo, C., Zaspas, G. (eds). Lecture Notes on Data Engineering and Communications Technologies*, vol 178. Cham: Springer, 2023, pp. 552-564. https://doi.org/10.1007/978-3-031-35467-0_33. . ISSN 1439-7358.

Додаткова:

1. Шаховська Н.Б., Болюбаш Ю.Я. Модель великих даних “сутність-характеристика”. // Вісник НУ «Львівська політехніка», Серія: Інформаційні технології та мережі, 2015, вип. 814. С. 186-196.

2. Колодчак О.М. Інтелектуальний аналіз даних // Вісник Національного університету "Львівська політехніка". Комп'ютерні системи та мережі. 2013. № 773. С. 49-58. – Режим доступу: http://nbuv.gov.ua/UJRN/VNULPKSM_2013_773_11 – Назва з екрану.

3. Гладун А.Я., Рогушина Ю.В. Data mining. Пошук знань в даних. К.: АДЕФ-Україна, 2016. 452 с.

4. Степашко В.С., Єфіменко С.М., Савченко Є.А. Комп'ютерний експеримент в індуктивному моделюванні. Київ: Наукова думка, 2014. 222 с. *

5. Єфіменко С.М., Степашко В.С. Прогнозна аналітика як ефективний інструмент підтримки рішень у системах цифрової економіки. *Control Systems and Computers*. 2018. № 6. С. 25-35.

6. Stepashko V., Voloschuk R. Development of a Decision Support System for the Country Economic Security Area. Proc. of the IEEE 18th International Conference on Computer Science and Information Technologies (CSIT). 19-21 October 2023, Lviv, Ukraine. Lviv: LNPU, 2023, pp. 1-4. doi: 10.1109/CSIT61576.2023.10324295.

7. Yefimenko S.M. Intelligent Information Technology for Inductive Modeling of Complex Processes on the Basis of Recurrent-and-Parallel Computations. *Control Systems and Computers*. 2023. № 1. С. 54-64. <https://doi.org/10.15407/csc.2023.01.054>.