

Міністерство освіти і науки України
Харківський національний університет імені В.Н. Каразіна
Кафедра теоретичної та прикладної системотехніки

“ЗАТВЕРДЖУЮ”

В.о. декана факультету комп'ютерних наук


Світлани КОЛОВАНОВА
“ 30 ” червня 2023 р.



Робоча програма навчальної дисципліни
Сучасний числовий наліз

рівень вищої освіти третій (освітньо-науковий)

галузь знань 12 «Інформаційні технології»

спеціальність 122 «Комп'ютерні науки»

освітньо-наукова програма Комп'ютерні науки

вид дисципліни обов'язкова

факультет комп'ютерних наук

2023 / 2024 навчальний рік

Програму обговорено та рекомендовано до затвердження вченою радою факультету комп'ютерних наук

«29» червня 2023 року, протокол № 14

РОЗРОБНИК ПРОГРАМИ:

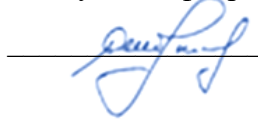
кандидат технічних наук, доцент з во кафедри теоретичної та прикладної системотехніки

Стрілець Вікторія Євгенівна .

Програму схвалено на засіданні кафедри теоретичної та прикладної системотехніки

«08» червня 2023 року, протокол № 13

Завідувач кафедри теоретичної та прикладної системотехніки



Сергій ШМАТКОВ.

Програму погоджено з гарантом освітньої програми «Комп'ютерні науки»

Гарант освітньої програми «Комп'ютерні науки»



Микола ТКАЧУК

Програму погоджено методичною комісією факультету комп'ютерних наук

«21» червня 2023 року, протокол № 12

Голова методичної комісії факультету комп'ютерних наук



Лариса ВАСИЛЬЄВА

ВСТУП

Програма навчальної дисципліни «Сучасний числовий аналіз» розроблена відповідно до освітньо-наукової програми підготовки докторів філософії на третьому рівні вищої освіти спеціальності 122 Комп'ютерні науки.

1. Опис навчальної дисципліни

1.1. Мета викладання навчальної дисципліни

Мета викладання навчальної дисципліни – надання здобувачам ступеня PhD сучасних знань та фахових компетентностей, які є необхідними для засвоєння основ інтелектуального аналізу даних, методології машинного навчання для дослідження складних ієрархічних багаторівневих систем (СІБС) та процесів, управління складними комп'ютерними системами.

1.2. Основні завдання вивчення дисципліни

- отримання теоретичних і практичних знань і умінь в області числового аналізу і машинного навчання;
- формування умінь і навичок створення і використання моделей і методів числового аналізу для проведення наукових досліджень;
- формування умінь і навичок набуття навичок з адаптації стандартних алгоритмів до нових – чисельних рішень складних прикладних задач;
- також придбання знань про пакети прикладних програм спеціалізованого призначення, які забезпечують розв'язання тих чи інших задач дисертаційних досліджень.

В ході вивчення дисципліни у аспіранта повинні формуватися такі компетентності.

Загальні компетентності

ЗК01. Здатність до абстрактного мислення, аналізу та синтезу.

Спеціальні (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 застосовувати сучасні інформаційні технології, бази даних та інші електронні ресурси, спеціалізоване програмне забезпечення у науковій та навчальній діяльності.

1.3. Кількість кредитів 4

1.4. Загальна кількість годин 120

1.5. Характеристика навчальної дисципліни	
Обов'язкова	
Денна форма навчання	Заочна (вечірня) форма навчання
Рік підготовки	
2-й	
Семестр	
4-й	4-й
Лекції	
30 год.	30 год.
Практичні, семінарські заняття	
- год.	- год.
Самостійна робота	
90 год.	90 год.
Індивідуальні завдання	
- год.	

1.6. Заплановані результати навчання

Згідно з вимогами освітньо-професійної програми, студенти мають досягти таких результатів навчання:

знати:

- основні задачі аналізу, інтелектуальної обробки даних і синтезу СІБС та процесів. управління складними комп'ютерними системами;
- методи аналізу та обробки великих масивів даних;
- методи машинного навчання для розв'язання задач аналізу, інтелектуальної обробки даних і синтезу СІБС, управлінню складними комп'ютерними системами;
- методи аналізу стохастичних процесів з використанням сучасних моделей;
- методи планування і виконання комп'ютерних експериментів та управління ними.

вміти:

- здійснювати вибір методів машинного навчання для розв'язання задач аналізу, інтелектуальної обробки даних і синтезу СІБС, управління складними комп'ютерними системами;
- проводити верифікацію математичних методів, оцінку якості математичних методів на основі існуючих критеріїв;

- вирішувати задачі чисельного характеру аналізу, інтелектуальної обробки даних і синтезу СІБС, управління складними комп'ютерними системами з застосуванням спеціалізованих пакетів;
- вміти створювати дослідницькі групи для проведення аналізу та обробки великих масивів даних;
- вміти створювати та використовувати нове програмне забезпечення для аналізу та обробки великих масивів даних;
- пояснювати, кількісно та якісно оцінювати, корегувати отримані результати.

В результаті вивчення дисципліни у студента повинні формуватися такі *програмні результати навчання (РН)*.

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.

РН03. Формулювати і перевіряти гіпотези; використовувати для обґрунтування висновків належні докази, зокрема, результати теоретичного аналізу, експериментальних досліджень (опитувань, спостережень,...) і математичного та/або комп'ютерного моделювання, наявні літературні дані.

РН04. 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'ютерній науці та дотичних міждисциплінарних напрямах.

РН05. Планувати і виконувати експериментальні та/або теоретичні дослідження з комп'ютерних наук та дотичних міждисциплінарних напрямів з використанням сучасних інструментів, критично аналізувати результати власних досліджень і результати інших дослідників у контексті усього комплексу сучасних знань щодо досліджуваної проблеми.

РН06. 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.

РН07. Розробляти та реалізовувати наукові та/або інноваційні інженерні проекти, які дають можливість переосмислити наявне та створити нове цілісне знання та/або професійну практику і розв'язувати значущі наукові та технологічні проблеми комп'ютерної науки з дотриманням норм академічної етики і врахуванням соціальних, економічних, екологічних та правових аспектів.

РН08. 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.

2. Тематичний план навчальної дисципліни

Розділ 1. Моделі і методи робастного оцінювання.

Тема 1. Теорія обчислювального навчання.

Індуктивне навчання за приватними емпіричними даними (навчальній вибірці). Дедуктивне навчання на основі формалізації знань експертів. Методи робастного оцінювання.

Інформаційно-аналітичне забезпечення процесів робастного оптимального проектування та інтелектуального діагностування (РОП&ІД) на основі методів машинного навчання.

Тема 2. Методи генерації пробних (навчальних) вибірок (рідкісних сіток) на основі теорії планування експериментів.

Методи планування експериментів. Повний та дробовий факторні експерименти типу 2^n та 3^n . Центральне композиційне планування. Ортогональне центральне композиційне планування. Рототабельне центральне композиційне планування. Плани Боксу-Хантера, Рехшафнера. Некомпозиційне планування. Плани Боксу-Бенкіна. Латинські і греко-латинські квадрати Ейлера. Критерії оптимальності планування.

Тема 3. Імовірнісні методи генерації пробних (навчальних) вибірок (рідкісних сіток).

Генеральні і репрезентативні вибірки, вибірки напрямків. Метод істотних вибірок. Послідовності Соболя (ЛПт послідовності), Нідеррайтера. Різновиди методу Монте-Карло; методи латинського гіперкуба, максимуму ентропії, \max і \min , поліноміального хаосу розширення.

Тема 4. Робастні методи ідентифікації математичних моделей систем та процесів.

Типи та види робастного оцінювання. Метод максимальної правдоподібності (М-оцінювання).

Некоректно поставлені задачі. Алгоритми регуляризації (робастні алгоритми): адаптивні, інваріантні. Методи регуляризації в задачах ідентифікації, апроксимації даних та прогнозування часових рядів.

Робастні штучні нейронні мережі (ШНМ). Гіперпараметри. Регуляризація у глибокому навчанні (заснованому на навчанні уявленням). Ядерне згладжування.

Тема 5. Методи оцінювання інформативності (значущості) змінних стану при невизначеності даних.

Методи оцінювання диференціальної інформативності з врахуванням точності вимірювання змінних стану и наявність парної кореляції між ними: кореляційного аналізу, дисперсійного аналізу і методи розпізнавання образів. Методи розпізнавання образів: детерміністські (дискримінантного аналізу, багатовимірного шкалювання і логічні), ймовірно-статистичні (методи Байеса, послідовного аналізу і оцінювання на основі теорії інформації). Стохастичний аналіз інформативності: індекси Соболя, Taguchi S / N Ratio.

Оцінювання інформативності на основі методів структурно-параметричного аналізу і синтезу регресійних моделей: факторного аналізу (головних компонент (МГК), нелінійні МГК, Грамма-Шмідта, аналізу компонентів на основі теорії інформації) і спрямованого перебору (ітеративні – на основі різних типів апроксиматорів, в тому числі

ШНМ, що навчаються), послідовного аналізу варіантів, вагові з адаптацією, локально-стохастичні на основі самоорганізації.

Тема 6. Прогнозування багатовимірних часових рядів на основі концепції тренд-аналізу. Математичні моделі контролю стану (ММКС) динамічних систем і процесів.

Багатовимірні часові ряди. Багатовимірні трендові регресійні моделі. Ранг коінтеграції (розмірність простору коінтегрованих часових рядів).

Типи статистичних ММКС: моделі стохастичною фільтрації, регресивні (структурно-параметричні моделі) і ймовірнісні моделі.

Моделі стохастичною фільтрації (фільтри Калмана-Бьюсі).

Регресивні моделі: згладжування часового ряду (змінного середнього і експоненційного згладжування), авторегресійні трендові моделі: лінійні (ARIMA, GARCH і SSM) та нелінійні (ШНМ з тимчасовими затримками, рекурентні ШНМ).

Ймовірнісні моделі: мережі Петрі, ланцюги Маркова (приховані Маркові моделі).

Розкладання часових рядів. Критерії тренду.

Розділ 2. Математичні моделі та обчислювальні методи стратифікації даних

Тема 7. Математичні моделі і методи кластеризації даних (стратифікація прецедентів на підгрупи (класи)).

Методи стратифікації прецедентів на підгрупи (класи).

Навчання з вчителем: узагальнення та перенавчання, крос-валідація, порівняння класифікаторів. Вибір методу обчислення відстані між об'єктами: метрики Мінковського, Хеммінга, Евкліда, зважена евклідова відстань, статистики Стьюдента та Романовського.

Навчання без вчителя (кластерний аналіз): детерміністські методи (дискримінантного аналізу, опорних векторів, багатовимірною шкалювання і логічні), ієрархічна кластеризація, алгоритм k-середніх, нечіткий алгоритм k-середніх. Ймовірнісні моделі кластеризації: байєсівські моделі (байєсівські мережі довіри), EM-алгоритми. Непараметричні моделі кластеризації. Графові методи кластеризації: алгоритм виділення зв'язкових компонент, алгоритм ФОРЕЛ, функціонали якості кластеризації. Ієрархічна кластеризація (таксономія): агломеративна ієрархічна кластеризація, дендрограма; властивості стиснення, розтягування і редукованості.

Тема 8. Системні математичні моделі розпізнавання станів систем і процесів на основі даних моніторингу.

Аналітичні методи класифікації (на основі алгебр): агломеративні, факторного аналізу, дискримінантного аналізу, ШНМ (у тому числі мережі Хопфілда, Хеммінга, Кохонена).

Логічні методи класифікації, на базі теоретико-множинного уявлення: мультіагентні, нечіткої логіки, лінгвістичної апроксимації (наприклад, у формі продукційних правил, заснованих на обчисленні предикатів).

Статистичні методи класифікації: EM-алгоритми, стохастичні ШНМ.

Розділ 3. Задачі та методи теорії прийняття рішень в умовах невизначеності.

Тема 9. Пряма і зворотна задачі розрахунку розмірних конструкторських ланцюгів і методи їх вирішення.

Математичні моделі та обчислювальні методи синтезу рішень задач стохастичною оптимізацією: М-, V- і Р-задачі. Багатокритеріальні задачі стохастичною оптимізацією зі змішаними умовами: MV-, МН-задачі.

Статистичні оцінки довірчих інтервалів математичного очікування функцій для нелінійних залежностей методом Монте-Карло.

Синтез квазірішень багатокритеріальних задач системної модифікації в детермінованій і стохастичній формулюваннях. Методи побудови множини Парето.

Тема 10. Застосування меметичних алгоритмів для вирішення задач робастного оцінювання.

Методи обчислювального інтелекту: ітеративні (локальні), вагові локальні (з адаптацією), локально-стохастичні на основі самоорганізації: стохастичною апроксимацією, стохастичні квазіградієнтні, мультиагентні (генетичні алгоритми, методи диференціальної еволюції, імунні), популяційні (імітації руху: зграї перелітних птахів; мурашиних, бджолиних колоній). Меметичні алгоритми як гібридні алгоритми, засновані на використанні різних стратегій (гіпереврістік).

3. Структура навчальної дисципліни

Назви розділів і тем	Кількість годин				
	Денна (заочна, вечірня) форма				
	Усього	у тому числі			
Л		п	лаб	інд	с.р.
<i>Розділ 1. Моделі і методи робастного оцінювання</i>					
Тема 1. Теорія обчислювального навчання.	10	2			9
Тема 2. Методи генерації пробних (навчальних) вибірок на основі теорії планування експериментів.	10	2			9
Тема 3. Імовірнісні методи генерації пробних (навчальних) вибірок.	10	2			9
Тема 4. Робастні методи ідентифікації математичних моделей систем та процесів.	14	4			9
Тема 5. Методи оцінювання інформативності (значущості) змінних стану при невизначеності даних.	14	4			9

Тема 6. Прогнозування багатовимірних часових рядів на основі концепції тренд-аналізу. Математичні моделі контролю стану динамічних систем і процесів.	14	4				9
<i>Розділ 2. Математичні моделі та обчислювальні методи стратифікації даних</i>						
Тема 7. Математичні моделі і методи кластеризації даних.	14	4				9
Тема 8. Системні математичні моделі розпізнавання станів систем і процесів на основі даних моніторингу.	10	2				9
<i>Розділ 3. Задачі та методи теорії прийняття рішень в умовах невизначеності</i>						
Тема 9. Пряма і зворотна задачі розрахунку розмірних конструкторських ланцюгів і методи їх вирішення.	14	4				9
Тема 10. Застосування математичних алгоритмів для вирішення задач робастного оцінювання.	10	2				9
Усього годин	120	30				90

4. Теми практичних занять

-

5. Завдання для самостійної роботи

№ п/п	Зміст	Кількість годин
1	Дослідити методи генерації пробних (навчальних) вибірок (рідкісних сіток) на основі теорії планування експериментів.	9
2	Дослідити імовірнісні методи генерації пробних (навчальних) вибірок (рідкісних сіток).	9
3	Дослідити робастні методи ідентифікації математичних моделей систем та процесів.	9
4	Дослідити методи оцінювання інформативності (значущості) змінних стану при невизначеності даних.	9
5	Дослідити методи прогнозування багатовимірних часових рядів на основі концепції тренд-аналізу.	9
6	Підготовка до підсумкової контрольної роботи	9

7	Дослідити математичні моделі і методи кластеризації даних (стратифікація прецедентів на підгрупи (класи)).	9
8	Дослідити системні математичні моделі розпізнавання станів систем і процесів на основі даних моніторингу.	9
9	Дослідити методи розв'язання прямої та зворотної задач розрахунку розмірних конструкторських ланцюгів і методи їх вирішення.	9
10	Дослідити приклади застосування меметичних алгоритмів для вирішення задач робастного оцінювання.	9
	Разом	90

6. Індивідуальні завдання

(не має)

7. Методи контролю

Контроль засвоєння навчального матеріалу здійснюється шляхом:

- звітування про самостійну роботу на лекційних заняттях;
- проведення письмового підсумкового контролю знань.

Аспіранти, які не виконали всі види робіт, що включені до навчального плану, до заліку не допускаються.

Підсумковий контроль здійснюється шляхом проведення заліку.

8. Схема нарахування балів

Бали за поточний контроль знань впродовж семестру (по темах)						Разом	Залік	Сума
T1	T2, 3	T4	T5, 6	T7, 8	T9, 10			
8	8	8	8	8	8	60	40	100

Бали за темами нараховуються при контрольному опитуванні на лекційних заняттях.

Критерії оцінювання залікових робіт студентів

Вимоги	Кількість балів
--------	-----------------

Показані всебічні систематичні знання та розуміння навчального матеріалу; безпомилково виконані завдання.	35-40
Показані повні знання навчального матеріалу; помилки, якщо вони є, не носять принципового характеру.	30-35
Показано повне знання необхідного навчального матеріалу, але допущені помилки.	20-30
Показано повне знання необхідного навчального матеріалу, але допущені суттєві помилки	10-20
Показано недосконале знання навчального матеріалу, допущені суттєві помилки.	5-10
Показано недосконале знання навчального матеріалу, допущені суттєві помилки, які носять принциповий характер; обсяг знань не дозволяє засвоїти предмет.	1-5

Шкала оцінювання

Сума балів за всі види навчальної діяльності протягом семестру	Оцінка	
	для чотирирівневої шкали оцінювання	для дворівневої шкали оцінювання
90 – 100	відмінно	зараховано
70-89	добре	
50-69	задовільно	
1-49	незадовільно	не зараховано

9. Рекомендована література

Основна література

1. MacKay David J.C. Information Theory, Inference and Learning Algorithms / David J.C. MacKay. – Cambridge University Press, 2003. – 628 p.
2. Bishop Christopher M. Pattern Recognition and Machine Learning / Christopher M. Bishop. – New York: Springer, 2006. – 738 p.
3. Deuffhard P. Numerical Analysis in Modern Scientific Computing / P. Deuffhard, A. Hohmann // Texts in Applied Mathematics, vol. 43. – 2003. – 339 p.
4. Kincaid D. Numerical Analysis: Mathematics of Scientific Computing / D. Kincaid, W. Cheney. – American Mathematical Society, 2009. – 788 p.
5. Vachharajani V. Numerical Analysis: A Programming Approach / Vinay Vachharajani. – BPB Publications, 2018. – 596 p.

6. Stewart D.E. Numerical Analysis: A Graduate Course / David E. Stewart // CMS/CAIMS Books in Mathematics, vol. 4. – 2022. – 632 p.
7. Modern numerical methods and their application in mechanical engineering / Yi Wang, Bo Yu, Filippo Berto, Weihua Cai, Kai Bao // Advances in Mechanical Engineering. – 2019. – Vol. 11 (11). – P. 1-3.
8. A Survey of Modern Numerical Analysis / David M. Young // Society for Industrial and Applied Mathematics, 1973. – Vol. 15, No. 2. – P. 503-523.
9. Методи машинного навчання у задачах системного аналізу і прийняття рішень : монографія / Стрілець В.Є., Шматков С.І., Угрюмов М.Л. та ін. – Харків : ХНУ імені В.Н. Каразіна, 2020. – 169 с.

Допоміжна

1. Rasmussen C. E., Williams C. K. I. Gaussian Processes for Machine Learning. – Cambridge, Massachusetts: MIT Press, 2006. – 248 p.
2. Bishop Christopher M. Pattern Recognition and Machine Learning. – New York: Springer, 2006. – 738 p.
3. Deep Learning Tutorial. – LISA lab, University of Montreal, 2015. – 167 p.
4. Hal Daume III. A course in machine learning. – Edited by John Mark Ockerbloom, 2015. – 193 p.
5. Coelho L. P. Building machine learning systems with Python / Luis Pedro Coelho, Willi Richert. – Packt Publishing Ltd., 2015. – 302 p.
6. Raschka S. Python machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow2 / Sebastian Raschka, Vahid Mirjalili. – Packt Publishing, 2019. – 772 p.
7. Cielen D. Introduction Data Science: Big Data, Machine Learning, and more, using Python tools / Davy Cielen, Arno D.B. Meysman, Mohamed Ali. – Manning, 2016. – 320 p.
8. Fenner M. Machine Learning with Python for Everyone / Mark Fenner. – O'Reilly Media, Inc., 2019.
9. Sutton Richard S., Barto Andrew G. Reinforcement Learning: An Introduction. – Cambridge, Massachusetts: MIT Press, 2018. – 426 p.
10. Hastie Trevor, Tibshirani Robert, Friedman Jerome. The Elements of Statistical Learning. Data Mining, Inference, and Prediction. – New York: Springer, 2009. – 745 p.

10. Посилання на інформаційні ресурси в Інтернеті, відео-лекції, інше методичне забезпечення

1. <https://www.datacamp.com/courses/deep-learning-in-python>
2. <https://www.kaggle.com/learn/machine-learning>
3. Datasets – <https://www.kaggle.com>