

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ЗАПОРІЗЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ЕКОНОМІЧНИЙ ФАКУЛЬТЕТ
КАФЕДРА ЕКОНОМІЧНОЇ КІБЕРНЕТИКИ

ЗАТВЕРДЖУЮ



Д. Череп _____ факультету

В. Череп _____

28 серпня 2023 _____

Data Mining

РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

підготовки магістра

очної (денної) та заочної (дистанційної) форм здобуття освіти

спеціальності 051 «Економіка»

освітньо-професійна програма «Економічна кібернетика»,

«Управління персоналом та економіка праці»

Укладач Очеретін Д. В., к.е.н., доцент, доцент кафедри економічної кібернетики

Обговорено та ухвалено
на засіданні кафедри економічної кібернетики

Протокол № 1 від « 24 » серпня 2023 р.
В.о. завідувача кафедри економічної
кібернетики

Н.К. Максишко

Ухвалено науково-методичною радою
економічного факультету

Протокол № 1 від « 28 » серпня 2023 р.
Голова науково-методичної ради
економічного факультету

Н.О. Дугієнко

Погоджено
Гарант освітньо-професійної програми

І.В. Козін

В.М. Гельман

2023 рік

1. Опис навчальної дисципліни

1	2	3	
Галузь знань, спеціальність, освітня програма рівень вищої освіти	Нормативні показники для планування і розподілу дисципліни на змістові модулі	Характеристика навчальної дисципліни	
		очна (денна) форма здобуття освіти	заочна (дистанційна) форма здобуття освіти
Галузь знань <u>05 Соціальні та поведінкові науки</u>	Кількість кредитів – 4	Вибіркова	
Спеціальність <u>051 Економіка</u>	Загальна кількість годин – 120	Семестр:	
		3 -й	-
Освітньо-професійна програма <u>Економічна кібернетика, Управління персоналом та економіка праці</u>	Змістових модулів – 6	Лекції	
		22 год.	-
Рівень вищої освіти: магістерський	Кількість поточних контрольних заходів – 26	Лабораторні	
		10 год.	-
		Самостійна робота	
		88 год.	-
		Вид підсумкового семестрового контролю: залік	

2. Мета та завдання навчальної дисципліни

Метою викладання навчальної дисципліни «Data mining» є формування системи фундаментальних знань щодо процесу виокремлення, дослідження та моделювання великих обсягів даних для виявлення невідомих до цього структур з застосуванням статистичних та математичних методів.

Предметом вивчення навчальної дисципліни є набуття навичок застосування методів інтелектуального аналізу даних за допомоги мови програмування для статистичної обробки даних R.

Об'єктом вивчення дисципліни є оволодіння практичними навичками проведення інтелектуального аналізу даних за допомоги мови програмування для статистичної обробки даних R.

Основними **завданнями** вивчення дисципліни «Data mining» є:

- оволодіти основними поняттями Data mining;
- вивчити класифікацію методів та стадій Data mining;
- вивчити завдання Data mining;
- оволодіти основними поняттями мови програмування для статистичної обробки даних R;
- навчитися проводити пошук у великих обсягах даних неочевидних, об'єктивних та корисних на практиці закономірностей;
- навчитися застосовувати статистичні та кібернетичні методи Data mining.

У результаті вивчення навчальної дисципліни студент повинен набути таких результатів навчання (знання, уміння тощо) та компетентностей:

Заплановані робочою програмою результати навчання та компетентності	Методи і контрольні заходи
<p>ІК Здатність визначати й розв'язувати складні економічні задачі та проблеми, приймати відповідні аналітичні та управлінські рішення у сфері економіки або у процесі навчання, що передбачає проведення досліджень та/або здійснення інновацій за невизначених умов та вимог</p>	<p>Репродуктивні методи (лекція, пояснення, робота з методичними матеріалами). Наочні методи (схеми, моделі, алгоритми).</p>
<p>ЗК-04 Здатність спілкуватися з представниками інших професійних груп різного рівня (з експертами з інших галузей знань/видів економічної діяльності)</p> <p>ЗК-08 Здатність проводити дослідження на відповідному рівні</p> <p>СК-02 Здатність до професійної комунікації в сфері економіки іноземною мовою</p> <p>СК-03 Здатність збирати, аналізувати та обробляти статистичні дані, науково-аналітичні матеріали, які необхідні для розв'язання комплексних економічних проблем, робити на їх основі обґрунтовані висновки</p> <p>СК-04 Здатність використовувати сучасні інформаційні технології, методи та прийоми дослідження економічних та соціальних процесів, адекватні встановленим потребам дослідження</p> <p>СК-06 Здатність формулювати професійні задачі в сфері економіки та розв'язувати їх, обираючи незалежні напрями і відповідні методи для їх розв'язання, беручи до уваги наявні ресурси</p> <p>СК-09 Здатність застосовувати науковий підхід до формування та виконання ефективних проєктів у соціально-економічній сфері</p> <p>СК-11 Здатність планувати і виробляти проєкти у сфері економіки, здійснювати її інформаційне, методичне, матеріальне, фінансове та кадрове забезпечення</p> <p>СК-12 Здатність знаходити, обробляти, інтерпретувати економічні дані та їх використовувати для дослідження процесів в сфері економічної діяльності на базі застосування математичних методів, моделей та комп'ютерних технологій.</p> <p>СК-13 Здатність моделювати проблеми управління та їх наслідки і пропонувати можливі шляхи вирішення із використанням методів економічної кібернетики та сучасних інформаційних технологій</p> <p>СК-14 Здатність до проведення досліджень у сфері інформатики, спрямованих на пошук інноваційного використання нових та існуючих комп'ютерних технологій в сфері економічної діяльності.</p>	<p>Евристичні методи</p>
<p>ПРН-01 Формулювати, аналізувати та синтезувати рішення науково-практичних проблем</p> <p>ПРН-05 Дотримуватися принципів академічної доброчесності</p> <p>ПРН-07 Обирати ефективні методи управління економічною діяльністю, обґрунтовувати пропонувані рішення на основі релевантних даних та наукових і прикладних досліджень</p> <p>ПРН-08 Збирати, обробляти та аналізувати статистичні дані, науково-аналітичні матеріали, необхідні для вирішення комплексних економічних завдань</p> <p>ПРН-09 Приймати ефективні рішення за невизначених умов і вимог, що потребують застосування нових підходів, методів та інструментарію соціально-економічних досліджень</p>	<p>Методи контролю і самоконтролю: усний, письмовий.</p> <p>Практичні методи (лабораторні та контрольні завдання).</p> <p>Контрольні заходи:</p> <ul style="list-style-type: none"> – лабораторні завдання; – теоретичне тестування; – залік.

ПРН-10 Застосовувати сучасні інформаційні технології та спеціалізоване програмне забезпечення у соціально-економічних дослідженнях та в управлінні соціально-економічними системами

ПРН-11 Визначати та критично оцінювати стан та тенденції соціально-економічного розвитку, формувати та аналізувати моделі економічних систем та процесів

ПРН-15 Організувати розробку та реалізацію соціально-економічних проєктів із врахуванням інформаційного, методичного, матеріального, фінансового та кадрового забезпечення

ПРН-17 Досліджувати та прогнозувати розвиток соціально-економічних систем на основі системного підходу та з використанням сучасних інформаційних технологій

Міждисциплінарні зв'язки

Вивчення дисципліни «Data Mining» базується на знаннях, вміннях та навичках, отриманих під час вивчення таких дисциплін, як «Аналіз та моделювання соціально-економічних систем», «Методологія наукових досліджень в інформаційній економіці», «Професійно-орієнтований практикум іноземною мовою».

Після вивчення курсу «Аналіз та моделювання соціально-економічних систем» студент повинен володіти сучасними математичними моделями та методами аналізу соціально-економічних систем, що становлять основу кількісного обґрунтування та сприяють підвищенню якості управлінських рішень.

Після вивчення курсу «Методологія наукових досліджень в інформаційній економіці» студент повинен володіти знаннями з основ методології, методів і понять наукового дослідження; отримати практичні навички і уміння застосування методів проведення наукового дослідження.

Після вивчення курсу «Професійно-орієнтований практикум іноземною мовою» студент повинен володіти практичними навичками спілкування іноземною мовою в науковій та професійній діяльності, вміти розуміти та інтерпретувати інформацію з міжнародних науково-метричних баз та видань.

3. Програма навчальної дисципліни

Змістовий модуль 1. Основні поняття про Data Mining

Тема 1. Data Mining як мультидисциплінарна галузь

Зміст терміна «Data Mining». Поняття статистики. Поняття машинного навчання. Поняття штучного інтелекту. Розвиток технології баз даних. Поняття Data Mining. Data Mining як частина ринку інформаційних технологій: класифікація аналітичних систем, думки експертів про Data Mining. Проблеми технології Data Mining. Відмінності Data Mining від інших методів аналізу даних. Існуючі підходи до аналізу.

Тема 2. Набір даних та їх атрибутів

Поняття про дані у широкому розумінні. Набір даних та їх атрибутів. Змінна. Значення. Генеральна сукупність. Вибірка. Параметри. Статистики. Гіпотези. Вимірювання. Числові та символічні змінні. Дискретні та неперервні дані. Шкали: номінальна, порядкова, інтервальна, відносна, діхотомічна. Типи наборів даних. Табличні дані. Графічні дані. Формати зберігання даних. Основні положення баз даних. Системи управління базами даних (СУБД). Вимоги до СУБД. Класифікація видів даних. Метадані.

Змістовий модуль 2. Методи, стадії та завдання Data Mining

Тема 3. Методи та стадії Data Mining

Класифікація стадій Data Mining. Класифікація методів Data Mining. Статистичні методи Data Mining. Кібернетичні методи Data Mining. Властивості методів Data Mining.

Тема 4. Завдання Data Mining. Інформація та знання

Завдання Data mining: класифікація, кластеризація, асоціація, послідовна асоціація, прогнозування, оцінювання, аналіз зв'язків, візуалізація. Класифікація завдань Data mining: за стратегіями: навчання з вчителем, навчання без вчителя, інші. Класифікація завдань Data mining в залежності від моделей, що використовуються: описові, прогнозуючі. Розподіл завдань Data mining: автоматичне дослідження та відкриття (вільний пошук), пояснення та опис, зв'язок понять, Поток «від даних до рішень». Поток «від завдання до додатку». Інформація. Властивості інформації. Знання та їх властивості.

Змістовий модуль 3. Метод пошуку асоціативних правил

Тема 5. Метод пошуку асоціативних правил

Сфери застосування асоціативних правил. Завдання аналізу ринкового кошика. Транзакційна база даних. Визначення та характеристики асоціативних правил. Межі підтримки та достовірності асоціативного правила. Алгоритм Apriori та його різновиди. Програмні засоби реалізації методу пошуку асоціативних правил. Реалізація методу пошуку асоціативних правил у пакеті arules середовища R.

Змістовий модуль 4. Методи класифікації та кластеризації

Тема 6. Метод кластерного аналізу

Поняття про кластерний аналіз. Завдання кластерного аналізу. Методи кластерного аналізу: ієрархічні та неієрархічні. Міри подібності: квадрат евклідової відстані, Манхеттінська відстань, відстань Чебишева, відсоток незгоди. Методи об'єднання або зв'язку. Алгоритм k-середніх. Перевірка якості кластеризації.

Тема 7. Метод дерева рішень

Застосування методу дерева рішень для завдань класифікації та прогнозування. Переваги дерев рішень. Процес конструювання дерева рішень. Критерій розщеплення. Зупинка побудови дерева. Скорочення дерева або відсікання гілок. Алгоритми, що реалізують дерева рішень.

Змістовий модуль 5. Методи штучних нейронних мереж та дискримінантний аналіз

Тема 8. Метод штучних нейронних мереж

Завдання Data mining, що вирішуються з допомогою штучних нейронних мереж. Елементи нейронних мереж. Архітектура нейронних мереж. Навчання нейронних мереж. Моделі нейронних мереж: одношаровий та багатошаровий перцептрон. Програмне забезпечення для роботи з нейронними мережами. Карти Кохонена, що самоорганізуються.

Тема 9. Метод дискримінантного аналізу

Проблеми дискримінантного аналізу. Дискримінантний факторний аналіз. Геометричний прогнозний дискримінантний аналіз. Ймовірнісний дискримінантний аналіз. Вимірювання якості моделі: лямбда Уїлкса, коефіцієнт детермінації, скорегований коефіцієнт детермінації. Дискримінантний аналіз якісних змінних (метод DISQUAL). Переваги та недоліки дискримінантного аналізу.

Змістовий модуль 6. Метод аналізу часових рядів

Тема 10. Метод аналізу часових рядів

Послідовність етапів регресійного аналізу. Завдання регресійного аналізу. Тренд, сезонність та цикл. Точність прогнозу часових рядів. Види прогнозів часових рядів. Методи

прогнозування часових рядів. Однофакторна лінійна регресія в середовищі R. Множинна лінійна регресія у середовищі R.

4. Структура навчальної дисципліни

Змістовий модуль	Усього годин	Аудиторні (контактні) години					Самостійна робота, год		Система накопичення балів		
		Усього годин	Лекційні заняття, год		Лабораторні заняття, год		о/д ф.	з/дист ф.	Теор. зав-ня, к-ть балів	Практ. зав-ня, к-ть балів	Усього балів
			о/д ф.	з/дист ф.	о/д ф.	з/дист ф.					
1	15	6	4		2		9		5	2	7
2	15	6	4		2		9		5	2	7
3	15	2	2		-		13		3	3	6
4	15	6	4		2		9		8	8	16
5	15	8	6		2		7		6	9	15
6	15	6	2		4		9		3	6	9
Усього за змістові модулі	90	32	12		20		58		30	30	60
Підсумковий семестровий контроль залік	30	-	-	-			30		20	20	40
Загалом											120
											100

5. Темі лекційних занять

№ змістового модуля	Назва теми	Кількість годин	
		о/д.ф.	з/дист.ф.
1	Data Mining як мультидисциплінарна галузь	2	
	Набір даних та їх атрибутів	2	
2	Методи та стадії Data Mining	2	
	Завдання Data Mining. Інформація та знання	2	
3	Метод пошуку асоціативних правил	2	
4	Метод кластерного аналізу	2	
5	Метод дерева рішень	2	
6	Метод штучних нейронних мереж	4	
7	Метод дискримінантного аналізу	2	
8	Метод аналізу часових рядів	2	
Разом		22	

6. Темі лабораторних занять

№ змістового модуля	Назва теми	Кількість годин	
		о/д.ф.	з/дист.ф.
1	Data Mining як мультидисциплінарна галузь	2	
	Набір даних та їх атрибутів	-	
2	Методи та стадії Data Mining	2	
	Завдання Data Mining. Інформація та знання	-	
3	Метод пошуку асоціативних правил	-	
4	Метод кластерного аналізу	2	
5	Метод дерева рішень	-	
6	Метод штучних нейронних мереж	2	
7	Метод дискримінантного аналізу	-	

8	Метод аналізу часових рядів	2	
Разом		10	

7. Види і зміст поточних контрольних заходів

№ змістового модуля	Вид поточного контрольного заходу	Зміст поточного контрольного заходу	Критерії оцінювання**	Усього балів
1	2	3	4	5
1	Теоретичне опитування	Теоретичні питання за темою 1 розділу 3 робочої програми	Теоретичні знання студента оцінюються в 1 бал, викладач задає одне теоретичне питання, правильна відповідь на яке дозволяє отримати студенту 1 бал.	1
	Практичне завдання	Виконання лабораторної роботи 1	Лабораторна робота полягає у завантаженні даних та їх візуалізації на мові R - <u>2 бали</u> : студент повністю виконав практичне завдання; - <u>1 бал</u> : студент частково виконав практичне завдання; - <u>0 балів</u> : студент не виконав практичне завдання.	2
	Тестування	Тестування за темою 1 розділу 3 робочої програми у системі Moodle	Контрольне тестування за темою складається з 2 тестових питань. За кожну правильну відповідь студент отримує 1 бал.	2
	Тестування	Тестування за темою 2 розділу 3 робочої програми у системі Moodle	Контрольне тестування за темою складається з 2 тестових питань. За кожну правильну відповідь студент отримує 1 бал.	2
Усього за ЗМ 1 контр. заходів	4			7
2	Теоретичне опитування	Теоретичні питання за темою 3 розділу 3 робочої програми	Теоретичні знання студента оцінюються в 1 бал, викладач задає одне теоретичне питання, правильна відповідь на яке дозволяє отримати студенту 1 бал.	1
	Практичне завдання	Виконання лабораторної роботи 2	Завданням лабораторної роботи є оцінка статистичних характеристик на мові R - <u>2 бали</u> : студент повністю виконав практичне завдання; - <u>1 бал</u> : студент частково виконав практичне завдання; - <u>0 балів</u> : студент не виконав практичне завдання.	2
	Тестування	Тестування за темою 3	Контрольне тестування за темою складається з 2 тестових питань. За	2

		розділу 3 робочої програми у системі Moodle	кожну правильну відповідь студент отримує 1 бал.	
	Тестування	Тестування за темою 4 розділу 3 робочої програми у системі Moodle	Контрольне тестування за темою складається з 2 тестових питань. За кожну правильну відповідь студент отримує 1 бал.	2
Усього за ЗМ 2 контр. заходів	4			7
3	Теоретичне опитування	Теоретичні питання за темою 5 розділу 3 робочої програми	Теоретичні знання студента оцінюються в 1 бал, викладач задає одне теоретичне питання, правильна відповідь на яке дозволяє отримати студенту 1 бал.	1
	Практичне завдання	Виконання лабораторної роботи 3	Завданням лабораторної роботи є побудова асоціативних правил на мові R Лабораторна робота складається з 3 завдань. Правильне виконання одного завдання оцінюється у 1 бал.	3
	Тестування	Тестування за темою 5 розділу 3 робочої програми у системі Moodle	Контрольне тестування за темою складається з 2 тестових питань. За кожну правильну відповідь студент отримує 1 бал.	2
Усього за ЗМ 3 контр. заходів	3			6
4	Теоретичне опитування	Теоретичні питання за темою 6 розділу 3 робочої програми	Теоретичні знання студента оцінюються в 1 бал, викладач задає одне теоретичне питання, правильна відповідь на яке дозволяє отримати студенту 1 бал.	1
	Практичне завдання	Виконання лабораторної роботи 4	Завданням лабораторної роботи є проведення кластерного аналізу на мові R Лабораторна робота складається з 6 завдань. Правильне виконання одного завдання оцінюється у 1 бал.	6
	Тестування	Тестування за темою 6 розділу 3 робочої програми у системі Moodle	Контрольне тестування за темою складається з 3 тестових питань. За кожну правильну відповідь студент отримує 1 бал.	3

	Теоретичне опитування	Теоретичні питання за темою 7 розділу 3 робочої програми	Теоретичні знання студента оцінюються в 1 бал, викладач задає одне теоретичне питання, правильна відповідь на яке дозволяє отримати студенту 1 бал.	1
	Практичне завдання	Виконання лабораторної роботи 4	Завданням лабораторної роботи є побудова дерева рішень на мові R - <u>2 бали</u> : студент повністю виконав практичне завдання; - <u>1 бал</u> : студент частково виконав практичне завдання; - <u>0 балів</u> : студент не виконав практичне завдання.	2
	Тестування	Тестування за темою 7 розділу 3 робочої програми у системі Moodle	Контрольне тестування за темою складається з 3 тестових питань. За кожну правильну відповідь студент отримує 1 бал.	3
Усього за ЗМ 4 контр. заходів	6			16
5	Теоретичне опитування	Теоретичні питання за темою 8 розділу 3 робочої програми	Теоретичні знання студента оцінюються в 1 бал, викладач задає одне теоретичне питання, правильна відповідь на яке дозволяє отримати студенту 1 бал.	1
	Практичне завдання	Виконання лабораторної роботи 6	Завданням лабораторної роботи є побудова штучних нейронних мереж на мові R Лабораторна робота складається з 6 завдань. Правильне виконання одного завдання оцінюється у 1 бал.	6
	Тестування	Тестування за темою 8 розділу 3 робочої програми у системі Moodle	Контрольне тестування за темою складається з 2 тестових питань. За кожну правильну відповідь студент отримує 1 бал.	2
	Теоретичне опитування	Теоретичні питання за темою 9 розділу 3 робочої програми	Теоретичні знання студента оцінюються в 1 бал, викладач задає одне теоретичне питання, правильна відповідь на яке дозволяє отримати студенту 1 бал.	1
	Практичне завдання	Виконання лабораторної роботи 7	Завданням лабораторної роботи є проведення дискримінантного аналізу на мові R Лабораторна робота складається з 3 завдань. Правильне виконання одного завдання оцінюється у 1 бал.	3
	Тестування	Тестування за темою 9 розділу 3	Контрольне тестування за темою складається з 2 тестових питань. За кожну правильну відповідь студент	2

		робочої програми у системі Moodle	отримує 1 бал.	
Усього за ЗМ 56 контр. заходів	6			15
6	Теоретичне опитування	Теоретичні питання за темою 10 розділу 3 робочої програми	Теоретичні знання студента оцінюються в 1 бал, викладач задає одне теоретичне питання, правильна відповідь на яке дозволяє отримати студенту 1 бал.	1
	Практичне завдання	Виконання лабораторної роботи 8	Завданням лабораторної роботи є аналіз часових рядів на мові R. Лабораторна робота складається з 6 завдань. Правильне виконання одного завдання оцінюється у 1 бал.	6
	Тестування	Тестування за темою 10 розділу 3 робочої програми у системі Moodle	Контрольне тестування за темою складається з 2 тестових питань. За кожну правильну відповідь студент отримує 1 бал.	2
Усього за ЗМ 6 контр. заходів	3			9
Усього за змістові модулі контр. заходів	26			60

8. Підсумковий семестровий контроль

Форма	Види підсумкових контрольних заходів	Зміст підсумкового контрольного заходу	Критерії оцінювання	Усього балів
1	2	3	4	5
Залік	Теоретичне завдання	Тестування за змістовими модулями у системі Moodle	Контрольне тестування за вивченим матеріалом курсу складається з 10 питань. За кожну правильну відповідь студент отримує 2 бали.	20
	Практичне завдання	Захист індивідуального практичного завдання залікової роботи у системі Moodle	Індивідуальне практичне завдання полягає в розв'язанні двох аналітичних завдань на проведення аналізу даних на мові R, за правильне виконання кожного з яких студент отримує 10 балів. Результат виконання завдань оцінюється за такою шкалою: - <u>10 балів</u> – завдання повністю виконано без помилок; - <u>9 балів</u> – студент в цілому	20

			<p>виконав практичне завдання, але не повно та допустивши деякі неточності;</p> <p>- <u>8 балів</u> – студент правильно визначив сутність практичного завдання, але виконав його частково й допустив при цьому одну помилку, що не впливає на загальне розуміння практичного завдання;</p> <p>- <u>7 балів</u> – студент правильно визначив сутність практичного завдання, але виконав його частково й допустив при цьому дві помилки, що не впливають на загальне розуміння практичного завдання;</p> <p>- <u>6 балів</u> – студент правильно визначив сутність практичного завдання, але виконав його частково й допустив при цьому три помилки, що не впливають на загальне розуміння практичного завдання;</p> <p>- <u>5 балів</u> – студент правильно визначив сутність практичного завдання, але виконав його недостатньо або поверхово, допустивши при цьому одну помилку, що впливає на загальне розуміння практичного завдання;</p> <p>- <u>4 бали</u> – студент правильно визначив сутність практичного завдання, але виконав його недостатньо або поверхово, допустивши при цьому дві помилки, що впливають на загальне розуміння практичного завдання;</p> <p>- <u>3 бали</u> – студент правильно визначив сутність практичного завдання, але виконав його недостатньо або поверхово, допустивши при цьому три помилки, що впливають на загальне розуміння практичного завдання;</p> <p>- <u>2 бали</u> – студент частково або поверхово виконав практичне завдання, допустивши при цьому одну помилку, що суттєво впливає на загальне розуміння практичного завдання;</p> <p>- <u>1 бал</u> – студент частково або поверхово виконав практичне завдання, допустивши при цьому дві помилки, що суттєво</p>	
--	--	--	--	--

			впливають на загальне розуміння практичного завдання; - 0 балів – студент не виконав практичне завдання.	
Усього за підсумковий семестровий контроль				40

9. Рекомендована література

Основна:

1. Ситник В. Ф., Краснюк М. Т. Інтелектуальний аналіз даних (дейтамайнінг) : навч. посіб. Київ : КНЕУ, 2007. 376 с.
2. Черняк О. І., Захарченко П. В. Інтелектуальний аналіз даних : підруч. Київ : Знання, 2014. 599 с.
3. Олійник А. О., Субботін С. О., Олійник О. О. Інтелектуальний аналіз даних : навч. посіб. Запоріжжя : ЗНТУ, 2012. 278 с.
4. Гладун А. Я., Рогушина Ю. В. Data Mining: пошук знань в даних. Київ : ТОВ «ВД «АДЕФ- Україна», 2016. 452 с.
5. Рідкокаша А. А., Голдер К. К. Основи систем штучного інтелекту. Черкаси : Відлуння-плюс, 2002. 240 с.
6. Провост Фостер, Фоусетт Том Data Science для бізнесу: Як збирати, аналізувати і використовувати дані / пер. з англ. Анастасія Дудченко. Київ : Наш формат, 2019. 400 с.

Додаткова:

1. Ciaburro Giuseppe, Venkateswaran Balaji. Neural Networks with R. Birmingham : Packt Publishing, 2017. 314 p.
2. Berry M.J.A., Linoff G. Data mining techniques: for marketing, sales, and customer relationship management. Indianapolis : Wiley Publishing, 2004. 672 p. URL: <http://ebooks.znu.edu.ua/files/Bibliobooks/Inshi4/0006100.pdf>.
3. Pyle Dorian. Business modeling and data mining. Burlington : Morgan Kaufmann Publishers, 2003. 650 p.
4. Yanchang Z., Yonghua C. Data Mining Applications with R. Waltham, Oxford, Amsterdam : Elsevier, 2014. 471 p. URL: <http://ebooks.znu.edu.ua/files/Bibliobooks/Kudin/0036204.pdf>.
5. Azzalini A., Bruno S. Data Analysis and Data Mining. An Introduction. New York : Oxford University Press, 2012. 289 p. URL: <http://ebooks.znu.edu.ua/files/Bibliobooks/Kudin/0036206.pdf>.
6. Gisele L.P., Alex A.F. Automating the Design of Data Mining Algorithms: an Evolutionary Computation Approach. Heidelberg : Springer-Verlag Berlin Heidelberg, 2010. 197 p. URL: <http://ebooks.znu.edu.ua/files/Bibliobooks/Kudin/0036216.pdf>.
7. Stephane T. Data Mining and Statistics for Decision Making. New York : John Wiley & Sons, 2011. 704 p. URL: <http://ebooks.znu.edu.ua/files/Bibliobooks/Kudin/0036219.pdf>.
8. Плєскач В. Л., Затонацька Т. Г. Інтелектуальні технології Data Mining і Text Mining. *Інформаційні системи і технології на підприємствах*. Київ : Знання, 2011. С. 540–559.
9. Kandethody M. Ramachandran, Chris P. Tsokos Mathematical Statistics With Applications in R. London, San Diego, Cambridge, Oxford : Esevier, 2021. 680 p. URL: <https://doi.org/10.1016/C2018-0-02285-9>.
10. Paolo Giordani, Maria Brigida Ferraro, Francesca Martella. An Introduction to Clustering with R. Singapore : Springer Singapore, 2020. 340 p. DOI: <https://doi.org/10.1007/978-981-13-0553-5>.

11. Laura Chihara, Tim Hesterberg. Mathematical Statistics with Resampling and R. Hoboken, New Jersey : Wiley, 2011. 434 p.

12. Peter Dalgaard. Introductory Statistics with R. Second Edition. New York : Springer, 2008. 370 p. URL: https://www.academia.dk/BiologiskAntropologi/Epidemiologi/PDF/Introductory_Statistics_with_R__2nd_ed.pdf.

13. Danielle Navarro. Learning statistics with R: A tutorial for psychology students and other beginners (Version 0.6). University of New South Wales, 613 p. URL: <http://compcogscisydney.org/learning-statistics-with-r>.

14. Vijay Kotu and Bala Deshpande. Data Science. Concept and Practice. Second Edition. Cambridge : Elsevier, 2019. 549 p. URL: <https://asolanki.co.in/wp-content/uploads/2019/04/Data-Science-Concepts-and-Practice-2nd-Edition-3.pdf>.

Інформаційні джерела:

1. Сайт Державної служби статистики України. URL: <http://ukrstat.gov.ua>.

2. Сайт Національного банку України. URL: <http://www.bank.gov.ua>.

3. Статистичний відділ Організації об'єднаних націй. URL: <https://unstats.un.org/home>.

4. Відділ статистики Продовольчої та сільськогосподарської організації Об'єднаних націй (ФАО). URL: <http://www.fao.org/economic/ess/en>.

5. Сайт Всесвітнього банку. URL: <http://www.worldbank.org>.

6. Сайт Євростату. URL : <https://ec.europa.eu/eurostat>.

7. 50 Data Mining Resources – Tutorials, Techniques and More. URL: <https://www.ngdata.com/data-mining-resources>.

8. The R Project for Statistical Computing. URL: <https://www.r-project.org>.

9. Open source and enterprise-ready professional software for R-RStudio. URL: <https://www.rstudio.com>.

10. Kaggle: Your home for dataset. URL: <https://www.kaggle.com>.

11. Data Visualization Packages for R to Consider in 2023. CiteDrive. URL: <https://citedrive.medium.com/data-visualization-packages-for-r-to-consider-in-2023-b303d91a5ffb>.

12. Imaobong Njokko. Data Analysis in R- Series(VI); Introduction to Joins using dplyr. July 2022. URL: <https://medium.com/@imanjokko/data-analysis-in-r-series-vi-joining-data-using-dplyr-fc0a83f0f064>.

13. Data Analysis and Graphics Using R. February 23, 2023 by SAROJ. PYOFLIFE. URL: <https://pyoflife.com/data-analysis-and-graphics-using-r-pdf>.