СИЛАБУС
НАВЧАЛЬНОЇ ДИСЦИПЛІНИ
«DATA MINING»

Рівень вищої освіти	Другий (магістерський)
Назва кафедри	Статистики та математичних методів в економіці
Тип дисципліни	Основна
Керівник курсу	**Лазоренко Віталій Валерійович**, кандидат економічних наук, доцент, доцент кафедри статистики, інформаційних технологій та математичних методів в економіці
Профайл керівника курсу	*http://nasoa.edu.ua/spivrobitnyky*
Контактна інформація керівника курсу (e-mail)	E-mail: Lazorenko@nasoa.edu.ua

ОПИС ДИСЦИПЛІНИ

	Кредитів ECTS	Годин	За видами занять:			
			Лекцій	Практичних та семінарських занять	Самостійна підготовка	Вид підсумкового контролю
Обсяг:	Денна форма навчання					
	4	120	20	22	78	Залік
	Заочна форма навчання					
	4	120	2	4	114	Залік

Опис дисципліни	**Навчальна дисципліна: «Data Mining »** є вибірковою компонентою навчального плану. **Мета вивчення навчальної дисципліни** - формування у студентів теоретичних знань, практичних умінь і навичок щодо застосування сучасних методів інтелектуального аналізу даних (Data mining) в різних сферах людської діяльності. Надання поглиблених знань з теорії та практики інтелектуального аналізу даних, за допомогою якого можна виявити раніше невідомі, нетривіальні закономірності в даних. **Завдання вивчення навчальної дисципліни** полягає в оволодінні студентами теоретичними знаннями та практичними навичками з обробки, аналізу, візуалізації, прогнозування даних та застосування сучасних аналітичних платформ і технологій Data Mining для вирішення прикладних задач у професійній діяльності. **Формат проведення дисципліни:** лекції, практичні заняття, самостійна робота. Здобувачі вищої освіти мають змогу отримати індивідуальні консультації.
Професійні компетентності	**Здобувач повинен знати:** - основні методи консолідації, трансформації, візуалізації, оцінки якості, очищення та попередньої обробки даних;

	- принципи побудови та структурну організацію сховищ даних; - алгоритми пошуку асоціативних правил і кластерного аналізу; - статистичні методи та моделі класифікації і регресії; - методики аналізу та прогнозування часових рядів; - технологію побудови ансамблів і порівняння моделей; - можливості вітчизняних і зарубіжних універсальних програмих аналітичних платформ, що застосовуються для аналізу даних; - проблемні питання впровадження аналітичних програмних продуктів і технологій в професійну діяльність організацій і установ.
Програмні результати навчання	**Здобувач повинен вміти:** - практично застосовувати методи консолідації, трансформації, візуалізації, оцінки якості, очищення та попередньої обробки даних для якісної підготовки даних до аналізу; - створювати сховища даних, виконувати їх завантаження, витягувати дані зі сховищ; - застосовувати технології інтелектуального аналізу електронних масивів даних для вирішення конкретних практичних проблем; - використовувати можливості вітчизняних і зарубіжних універсальних програмних засобів і аналітичних платформ для пошуку закономірностей, зв'язків, правил, знань в електронних масивах даних; - вільно орієнтуватися на сучасному динамічному ринку аналітичних програмних продуктів; - розв'язувати засобами Data mining економічні задачі в умовах функціонування інформаційних систем; застосовувати отримані практичні навички використання методики аналізу та прогнозування часових рядів, технологію побудови ансамблів і порівняння моделей в у певній предметній галузі.
Програма навчальної дисципліни	Змістовий модуль 1: Теоретичні основи та інструменти інтелектуального аналізу даних Тема 1. Предмет, методи і завдання дисципліни *Data Mining* Тема 2. Ймовірнісний калькулятор в ПСП SPSS, STATISTICA та Microsoft Excel Тема 3. Класичні технології інтелектуального аналізу *Data Mining*. Методи класифікації і кластеризації, багатофакторного дисперсійного аналізу Тема 4. Дискримінантний аналіз в *Data Mining* засобами ПСП IBM SPSS Statistics Тема 5. Методи класифікації в *Data Mining* засобами STATISTICA Змістовий модуль 2: Прикладні технології кластеризації, факторного аналізу, прогнозування та візуалізації Тема 6. Кластерний аналіз в *Data Mining* засобами IBM SPSS Statistics Тема 7. Кластерний аналіз в *Data Mining* засобами STATISTICA Тема 8. Факторний аналіз. Когнітивний регресійний аналіз засобами ПСП SPSS, Statistica Тема 9. Задачі *Data Mining*: прогнозування та візуалізація, часові ряди Тема 10. АП DEDUCTOR. Комплексний підхід до інтелектуального аналізу даних. Традиційний процес *Data Mining*
Методи навчання	У процесі викладання дисципліни використовується кредитно-трансферна система організації навчального процесу, що відповідає вимогам ЄКТС. Для досягнення навчальних завдань, загального розвитку особистості, професійної підготовки доцільно використовувати різні методи навчання, а саме: словесний, наочний; пояснювально-ілюстративний, пошуковий, проблемний, проблемно-пошуковий, дослідницький; методи контролю і

	самоконтролю (усний, письмовий, тестування тощо).
Матеріально-технічне забезпечення навчальної дисципліни	Комп'ютери з програмним забезпеченням для виконання різних видів освітньої діяльності: Microsoft Office, Веб-браузери, IBM SPSS Statistics, STATISTICA, Deductor, Microsoft Excel та іншим офісним і аналітичним ПЗ. Мультимедійний проектор, комп'ютер або ноутбук, використання платформи MicrosoftTeems та Moodledля дистанційного навчання
Політики навчальної дисципліни	***Політика щодо академічної доброчесності.*** Дотримання академічної доброчесності здобувачами вищої освіти НАСОА є передумовою для ефективного опанування результатами навчання і отримання позитивної оцінки з поточного та підсумкового контролів. Складові політики академічної доброчесності регламентуються: Положенням про академічну доброчесність Національної академії статистики, обліку та аудиту, затвердженим рішенням Вченої ради НАСОА, протокол від 26 червня 2026 р. №15. Відповідно до Положення про внутрішню систему забезпечення якості освіти в НАСОА та Положення про академічну доброчесність НАСОА у разі виявлення викладачем порушення студентом норм та правил академічної доброчесності, викладач зобов'язаний застосувати норми п.4.17.2 Положення про академічну доброчесність НАСОА, зокрема: «повторне проходження оцінювання (тест, контрольна робота, залік, іспит тощо); повторне проходження відповідного освітнього компонента (модуля) освітньої програми». ***Політика щодо відвідування занять.*** Здобувачі вищої освіти денної форми навчання зобов'язані відвідувати заняття. Поважними причинами для відсутності на заняттях є хвороба, академічна мобільність або інші випадки відсутності, які підтверджені документально. Якщо здобувач вищої освіти відсутній на заняттях з поважної причини, він презентує виконані завдання під час самостійної підготовки та відповідно до графіку консультацій викладача. ***Політика щодо перескладання.*** Здобувачі вищої освіти повинні дотримуватися термінів виконання усіх завдань, передбачених програмою навчальної дисципліни. Ліквідація академічної заборгованості та перескладання заліку проводиться після закінчення екзаменаційної сесії за окремим розкладом, складеним навчально методичним відділом. ***Політика щодо оскарження результатів оцінювання.*** Якщо здобувач вищої освіти не згоден з оцінюванням рівня його знань він може оскаржити виставлену викладачем оцінку у встановленому порядку відповідно до Положення про організацію освітнього процесу у Національній академії статистики, обліку та аудиту, затвердженого рішенням Вченої ради НАСОА, протокол від 24 квітня 2020 р. №8 (зі змінами та доповненнями, протокол від 04 листопада 2025 р. №4).

КРИТЕРІЇ ТА МЕТОДИ ОЦІНЮВАННЯ

За результатами семестру студент отримує підсумкову оцінку за 100-бальною системою, яка розраховується як накопичена сума отриманих балів за всі види поточного, модульного та підсумкового контролю.

Умовою допуску до підсумкового контролю є набрання здобувачем вищої освіти 35 балів у сукупності за всіма темами дисципліни.

Якщо за результатами модульно-рейтингового контролю студент отримав сумарну кількість балів за два змістовні модулі, що не перевищує 34 бали, то здобувач вищої освіти вважається таким, що не виконав усі види робіт, які передбачаються навчальним планом з дисципліни «Data Mining» і направляється на повторний курс вивчення дисципліни.

Форми контролю	Види навчальної роботи	Оцінювання
Денна форма навчання		
ПОТОЧНИЙ КОНТРОЛЬ	Робота на лекціях	-
	Робота на практичних заняттях, у т.ч.:	
	Виконання практичного завдання	Від 1до 5 балів
МОДУЛЬНИЙ КОНТРОЛЬ	Модульний контроль № 1	Максимальна оцінка – 10 балів
	Модульний контроль № 2	Максимальна оцінка – 10 балів
ПІДСУМКОВИЙ КОНТРОЛЬ	Залік	Максимальна оцінка – 30 балів
Заочна форма навчання		
ПОТОЧНИЙ КОНТРОЛЬ	Робота на лекціях	-
КОНТРОЛЬНА РОБОТА ДЛЯ ЗАОЧНОЇ ФОРМИ НАВЧАННЯ	Модульний контроль № 1	Максимальна оцінка – 35 балів
	Модульний контроль № 2	Максимальна оцінка – 35 балів
ПІДСУМКОВИЙ КОНТРОЛЬ	Залік	Максимальна оцінка – 30 балів

КРИТЕРІЇ ОЦІНЮВАННЯ РІВНЯ НАВЧАЛЬНИХ ДОСЯГНЕНЬ ЗДОБУВАЧІВ ВИЩОЇ ОСВІТИ

Засоби оцінювання та методи демонстрування результатів навчання включають виконання завдань розрахунково-аналітичного характеру, розв'язання ситуаційних управлінських кейсів, підготовку та представлення презентацій за результатами виконаних завдань, аналізу управлінських ситуацій та міні-досліджень. Такі форми роботи спрямовані на перевірку здатності здобувачів застосовувати теоретичні знання для прийняття управлінських рішень, оцінювання альтернатив і формування оптимальних моделей поведінки організації.

Практичні заняття передбачають змістовне обговорення ключових питань теми, аналіз чинної нормативно-правової бази, що регулює управлінську діяльність, а також проведення дискусій щодо актуальних проблем менеджменту та сучасних підходів до організаційного розвитку. У ході практичного заняття здійснюється усне та письмове опитування для перевірки розуміння теоретичного матеріалу та уміння застосовувати його в конкретних управлінських ситуаціях. Максимальна кількість балів, яку здобувач може отримати за роботу на одному практичному занятті, становить 4 бали.

Самостійна робота здобувачів спрямована на глибше опрацювання теоретичних і прикладних аспектів менеджменту, дослідження управлінських моделей, підходів та інструментів, а також на виконання аналітичних завдань, що формують здатність до обґрунтованого прийняття управлінських рішень. Оцінювання самостійної роботи здійснюється під час практичних занять у формі усного чи письмового опитування, включно з питаннями, винесеними для індивідуального опрацювання.

Шкала оцінювання роботи здобувачів вищої освіти на практичних заняттях

Кількість балів	Критерії оцінювання
4-5	В повному обсязі володіє навчальним матеріалом, вільно самостійно та аргументовано його викладає під час усних виступів та письмових відповідей, глибоко та всебічно розкриває зміст теоретичних питань та практичних завдань, використовуючи при цьому літературу.
1-3	Не в повному обсязі володіє навчальним матеріалом. Фрагментарно, поверхово (без аргументації та обґрунтування) викладає його під час усних виступів та письмових відповідей, недостатньо розкриває зміст

	теоретичних питань та практичних завдань, допускаючи при цьому суттєві неточності.
0	Не володіє навчальним матеріалом та не в змозі його викласти, не розуміє змісту теоретичних питань та практичних завдань.

Критерії оцінювання модульного контролю

Формою модульного контролю є модульні контрольні роботи, які проводяться у формі тестових завдань закритої форми, що формуються та виконуються в системі Moodle, та оцінюються від 0 до 10 балів

Критерії оцінювання підсумкового контролю

Формою підсумкового контролю є екзамен, який проводяться у формі тестових завдань закритої форми, що формуються та виконуються в системі Moodle, та оцінюється від 0 до 30 балів.

Шкала оцінювання знань здобувачів вищої освіти

Оцінка за національною шкалою	**Оцінка в балах**	**Оцінка за шкалою ECTS**	
Залік		**Оцінка**	**Пояснення**
Зараховано	**90-100**	A	Здобувач вищої освіти демонструє високий рівень знань навчального матеріалу в обсязі, що відповідає програмі дисципліни, достатньо реалізовує теоретичні положення дисципліни в практичних розрахунках, виконав практичні завдання відмінно або з незначною кількістю помилок. За час навчання при проведенні практичних занять проявив вміння самостійно вирішувати поставлені завдання, активно включатись в дискусії, може відстоювати власну позицію в питаннях та рішеннях, що розглядаються
	82-89	B	Здобувач вищої освіти демонструє гарні знання, добре володіє матеріалом, що відповідає програмі дисципліни, робить на їх основі аналіз можливих ситуацій та вміє застосовувати теоретичні положення при вирішенні практичних задач, але допускає окремі неточності. Вміє самостійно виправляти допущені помилки, кількість яких є незначною
	74-81	C	Здобувач вищої освіти в загальному добре володіє матеріалом, знає основні положення матеріалу, що відповідає програмі дисципліни, робить на їх основі аналіз можливих ситуацій та вміє застосовувати при вирішенні типових практичних завдань, але допускає окремі неточності. Помилки у відповідях та розрахунках не є системними. Знає характеристики основних положень, що мають визначальне значення при проведенні практичних занять та поясненні прийнятих рішень, в межах дисципліни, що вивчається.
	64-73	D	Здобувач вищої освіти засвоїв основний теоретичний матеріал, передбачений програмою дисципліни. При вирішенні практичних завдань допускає значну кількість недоліків і суттєвих помилок
	60-63	E	Здобувач вищої освіти має певні знання, передбачені в програмі дисципліни, володіє основними положеннями,

			що вивчаються на рівні, який визначається як мінімально допустимий. З використанням основних теоретичних положень, здобувач вищої освіти з труднощами пояснює правила вирішення практичних завдань дисципліни, відсутнє розуміння порядку виконання завдання та взаємозв'язків з іншими дисциплінами.
Незараховано	**35-59**	FX	Здобувач вищої освіти може відтворити окремі фрагменти з курсу. Незважаючи на те, що програму дисципліни здобувач вищої освіти виконав, працював він пасивно, його відповіді під час практичних робіт в більшості є неправильними, необґрунтованими. Цілісність розуміння матеріалу з дисципліни у здобувача вищої освіти відсутні. Здобувач вищої освіти має можливість повторного складання.
	1-34	F	Здобувач вищої освіти повністю не виконав вимоги програми навчальної дисципліни. Його знання на підсумкових етапах навчання є фрагментарними. Обов'язковий повторний курс.

ПИТАННЯ ДЛЯ ПІДСУМКОВОГО КОНТРОЛЮ ЗНАНЬ

1. Сутність, предмет і завдання інтелектуального аналізу даних (Data Mining).
2. Основні етапи становлення та розвитку технологій Data Mining.
3. Місце Data Mining у системі сучасних інформаційних технологій.
4. Основні завдання інтелектуального аналізу даних та їх класифікація.
5. Основні методи та алгоритми Data Mining.
6. Етапи процесу інтелектуального аналізу даних.
7. Основні джерела даних та особливості їх використання в Data Mining.
8. Методи попередньої обробки, очищення та трансформації даних.
9. Поняття якості даних та основні підходи до її оцінювання.
10. Основні сфери практичного застосування технологій Data Mining.
11. Роль теорії ймовірностей і математичної статистики в інтелектуальному аналізі даних.
12. Поняття випадкової величини та основні характеристики її розподілу.
13. Основні закони розподілу ймовірностей, що використовуються в Data Mining.
14. Нормальний розподіл та його застосування в аналізі даних.
15. Використання ймовірнісного калькулятора в IBM SPSS Statistics.
16. Особливості обчислення ймовірностей у програмному середовищі STATISTICA.
17. Використання статистичних функцій Microsoft Excel для обчислення ймовірностей.
18. Порівняння можливостей SPSS, STATISTICA та Microsoft Excel для статистичного аналізу даних.
19. Сутність і основні характеристики класичних технологій Data Mining.
20. Класифікація як завдання інтелектуального аналізу даних.
21. Основні алгоритми класифікації та особливості їх застосування.
22. Кластеризація як метод виявлення прихованої структури даних.
23. Основні відмінності між класифікацією та кластеризацією.
24. Сутність і призначення дисперсійного аналізу.
25. Однофакторний і багатофакторний дисперсійний аналіз: особливості застосування.
26. Основні припущення та умови використання багатофакторного дисперсійного аналізу.
27. Інтерпретація результатів дисперсійного аналізу в задачах Data Mining.
28. Сутність і завдання дискримінантного аналізу.
29. Основні припущення та умови застосування дискримінантного аналізу.
30. Поняття дискримінантної функції та принципи її побудови.
31. Відбір змінних для проведення дискримінантного аналізу.
32. Етапи виконання дискримінантного аналізу в IBM SPSS Statistics.

33. Оцінювання статистичної значущості дискримінантних функцій.
34. Інтерпретація результатів дискримінантного аналізу.
35. Оцінювання точності класифікації об'єктів за результатами дискримінантного аналізу.
36. Основні методи класифікації даних у середовищі STATISTICA.
37. Дерева рішень як інструмент класифікації об'єктів.
38. Принципи побудови та інтерпретації дерев класифікації.
39. Застосування статистичних моделей для класифікації даних.
40. Особливості підготовки навчальної та тестової вибірок.
41. Основні показники оцінювання якості класифікаційних моделей.
42. Поняття перенавчання моделі та способи його запобігання.
43. Порівняння результатів застосування різних методів класифікації у STATISTICA.
44. Сутність, мета та завдання кластерного аналізу.
45. Основні етапи проведення кластерного аналізу.
46. Методи вимірювання відстаней і подібності між об'єктами.
47. Ієрархічні методи кластеризації та їх особливості.
48. Метод k-середніх: алгоритм і сфери застосування.
49. Проведення ієрархічного кластерного аналізу в IBM SPSS Statistics.
50. Використання методу k-середніх у IBM SPSS Statistics.
51. Методи визначення оптимальної кількості кластерів.
52. Інтерпретація та оцінювання якості результатів кластеризації.
53. Функціональні можливості STATISTICA для проведення кластерного аналізу.
54. Особливості підготовки даних до кластеризації у STATISTICA.
55. Використання ієрархічних методів кластеризації у STATISTICA.
56. Застосування методу k-середніх у середовищі STATISTICA.
57. Побудова та інтерпретація дендрограм.
58. Візуалізація результатів кластерного аналізу.
59. Порівняння результатів кластеризації, отриманих різними методами.
60. Практичне використання кластерного аналізу для розв'язання прикладних задач.
61. Сутність і завдання факторного аналізу.
62. Основні етапи проведення факторного аналізу.
63. Поняття латентних факторів та факторних навантажень.
64. Метод головних компонент та особливості його застосування.
65. Критерії визначення кількості факторів.
66. Методи обертання факторів та їх призначення.
67. Проведення факторного аналізу засобами IBM SPSS Statistics і STATISTICA.
68. Сутність регресійного аналізу та його використання для виявлення взаємозв'язків між змінними.
69. Основні етапи побудови регресійної моделі.
70. Оцінювання якості регресійної моделі та інтерпретація її параметрів.
71. Особливості використання когнітивного регресійного аналізу в задачах інтелектуального аналізу даних.
72. Прогнозування як завдання інтелектуального аналізу даних.
73. Основні методи прогнозування, що застосовуються в Data Mining.
74. Часові ряди: поняття, структура та основні компоненти.
75. Методи виявлення тенденцій розвитку часових рядів.
76. Методи згладжування часових рядів.
77. Використання регресійних моделей для прогнозування.
78. Основні підходи до оцінювання точності прогнозів.
79. Призначення та основні методи візуалізації даних.
80. Візуалізація результатів прогнозування засобами аналітичних програмних продуктів.
81. Практичне застосування прогнозних моделей для обґрунтування управлінських рішень.
82. Призначення та основні функціональні можливості аналітичної платформи Deductor.
83. Архітектура та основні компоненти аналітичної платформи Deductor.

84. Основні етапи реалізації процесу Data Mining у середовищі Deductor.
85. Імпорт, інтеграція та попередня обробка даних у Deductor.
86. Застосування методів класифікації та кластеризації в Deductor.
87. Використання аналітичної платформи Deductor для прогнозування.
88. Візуалізація та інтерпретація результатів інтелектуального аналізу даних у Deductor.
89. Сутність комплексного підходу до інтелектуального аналізу даних.
90. Основні етапи традиційного процесу Data Mining та їх взаємозв'язок.
91. Оцінювання ефективності застосування технологій Data Mining для розв'язання прикладних задач.

РЕКОМЕНДОВАНА ЛІТЕРАТУРА

1. Акіменко В. В. *Прикладні задачі інтелектуального аналізу даних (Data Mining)*. Київ: КНУ ім. Тараса Шевченка, 2018. 152 с.
2. Гнатюк В. *Вступ до R на прикладах*. Харків: Харківський національний економічний університет, 2010. 101 с.
3. Капітон А. М., Тищенко Д. О., Десятко А. М., Лазоренко В. В. *Evolution and distribution analysis of multimodal artificial intelligence systems = Еволюція та аналіз розвитку мультимодальних систем штучного інтелекту*, 2024. https://doi.org/10.26906/SUNZ.2024.4.075
4. Кулаженко В. В., Лазоренко В. В., Кузнецов О. Ф. Дослідження алгоритму формування фондового портфеля інвестора за допомогою теорії нечітких множин. *Інвестиції: практика та досвід*, 2021, № 2, с. 30–37. http://www.investplan.com.ua/pdf/2_2021/7.pdf
5. Майборода Р. Є. *Комп'ютерна статистика*. Київ: ВПЦ "Київський університет", 2019. 589 с.
6. Майборода Р. Є., Сугакова О. В. *Аналіз даних за допомогою пакета R: навчальний посібник*. Київ: КНУ ім. Тараса Шевченка, 2015. 65 с.
7. Притульська Н., Антюшко Д., Лазоренко В. Іродієтичні продукти: очікування та вподобання споживачів. *Міжнародні науково-практичні наші товари nd arkets*, 2023, № 45(1), с. 41–53.
8. Снитюк В. Є. *Прогнозування. Моделі. Методи. Алгоритми: навчальний посібник*. Київ: Маклаут, 2008. 364 с.
9. Чорней Р. К. *Практикум з теорії ймовірностей та математичної статистики*. Київ, 2006. 328 с.
10. Черняк О. І., Захарченко П. В. *Інтелектуальний аналіз даних: підручник*. Київ, 2014. 599 с.
11. Agresti A. *Statistical Methods for Social Sciences*. Boston: Pearson, 2018.
12. Han J., Pei J., Tong H. *Data Mining: Concepts and Techniques*. 4th ed. San Francisco: Morgan Kaufmann, 2022. 752 p.
13. Khorolska K., Bebeshko B., Desiatko A., Lazorenko V. 3D models classification with use of convolution neural network. *CEUR Workshop Proceedings*, 2021, vol. 3179, pp. 25–34. http://ceur-ws.org/Vol3179/Paper_3.pdf
14. Khorolska K., Lazorenko V., Bebeshko B., Desiatko A., Kharchenko O., Yaremych V. Usage of clustering in decision support system. In: Raj J.S., Palanisamy R., Perikos I., Shi Y. (eds). *Intelligent Sustainable Systems*. Lecture Notes in Networks and Systems, vol. 213. Singapore: Springer, 2022. https://doi.org/10.1007/978-981-16-2422-3_49
15. Tan P.-N., Steinbach M. *Introduction to Data Mining*. 2nd ed. Pearson, 2018. 864 p.
16. Zaki M. J., Meira W. Jr. *Data Mining and Machine Learning: Fundamental Concepts and Algorithms*. 2nd ed. Cambridge: Cambridge University Press, 2020.
17. Zumel N., Mount J. *Practical Data Science with R*. Manning Publications Co., 2014. 417 p.

Розглянуто та затверджено на засіданні кафедри статистики та математичних методів в економіці, протокол №1 від 27 серпня 2026 р.