



# СИЛАБУС НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

## ІНТЕЛЕКТУАЛЬНИЙ АНАЛІЗ ДАНИХ

ID 4684

Шифр, назва спеціальності та освітній рівень	126 Інформаційні системи та технології (бакалавр)	Назва освітньої програми	Інформаційні системи та технології (2024)
Тип програми	Освітньо-професійна	Мова викладання	Українська
Факультет	Факультет комп'ютерно-інформаційних систем і програмної інженерії (ФІС)	Кафедра	Каф. комп'ютерних наук (КН)

### Викладач/викладачі

**Козбур Галина Володимирівна**, канд. техн. наук, доцент, доцент кафедри комп'ютерних наук, [профіль на порталі "Науковці ТНТУ"](#)

## Загальна інформація про дисципліну

Мета курсу	Формування теоретичних знань і практичних навичок застосування методів інтелектуального аналізу даних для побудови моделей, аналізу та інтерпретації даних, обґрунтування вибору алгоритмів і програмних засобів та використання отриманих результатів у задачах підтримки прийняття рішень в інформаційних системах.
Формат курсу	для денної та заочної форм навчання, з можливістю проведення в аудиторіях та/або он-лайн (в середовищі дистанційного навчання Atutor)
Компетентності ОП	<p>Інтегральна компетентність: Здатність розв'язувати складні спеціалізовані задачі та практичні проблеми в області інформаційних систем та технологій, або в процесі навчання, що характеризуються комплексністю та невизначеністю умов, які потребують застосування теорій та методів інформаційних технологій.</p> <p>Загальні компетентності (ЗК):</p> <p>ЗК01. Здатність до абстрактного мислення, аналізу та синтезу.</p> <p>ЗК02. Здатність застосовувати знання у практичних ситуаціях.</p> <p>ЗК03. Здатність до розуміння предметної області та професійної діяльності.</p> <p>Спеціальні (фахові) компетентності (СК):</p> <p>СК06. Здатність використовувати сучасні інформаційні системи та технології (виробничі, підтримки прийняття рішень, інтелектуального аналізу даних та інші), методики й техніки кібербезпеки під час виконання функціональних завдань та обов'язків.</p> <p>СК11. Здатність до аналізу, синтезу і оптимізації інформаційних систем та технологій з використанням математичних моделей і методів.</p> <p>СК13. Здатність проводити обчислювальні експерименти, порівнювати результати експериментальних даних і отриманих рішень.</p> <p>СК15. Здатність проєктувати архітектуру інформаційних систем із врахуванням вимог до обробки великих даних, інформаційної безпеки та взаємодії компонентів у розподілених середовищах.</p>
	<p>ПРО1. Знати лінійну та векторну алгебру, диференціальне та інтегральне числення, теорію функцій багатьох змінних, теорію рядів, диференціальні рівняння для функції однієї та багатьох змінних, операційне числення, теорію ймовірностей та математичну статистику в обсязі, необхідному для розробки та використання інформаційних систем, технологій та інфокомунікацій, сервісів та інфраструктури організації.</p> <p>ПРО2. Застосовувати знання фундаментальних і природничих наук, системного аналізу та технологій моделювання, стандартних алгоритмів та дискретного аналізу при розв'язанні задач проєктування і використання інформаційних систем</p>

<p>Програмні результати навчання з ОП</p>	<p>та технологій.</p> <p>ПР03. Використовувати базові знання інформатики й сучасних інформаційних систем та технологій, навички програмування, технології безпечної роботи в комп'ютерних мережах, методи створення баз даних та інтернет-ресурсів, технології розроблення алгоритмів і комп'ютерних програм мовами високого рівня із застосуванням об'єктно-орієнтованого програмування для розв'язання задач проектування і використання інформаційних систем та технологій.</p> <p>ПР05. Аргументувати вибір програмних та технічних засобів для створення інформаційних систем та технологій на основі аналізу їх властивостей, призначення і технічних характеристик з урахуванням вимог до системи і експлуатаційних умов; мати навички налагодження та тестування програмних і технічних засобів інформаційних систем та технологій.</p> <p>ПР16. Здатність інтегрувати методи інтелектуального аналізу даних у прикладні інформаційні системи для підтримки прийняття рішень, прогнозування та виявлення аномалій.</p>
<p>Обсяг курсу</p>	<p><b>Очна (денна) форма здобуття освіти:</b></p> <p>Кількість кредитів ECTS — 4,5; лекції — 18 год.; лабораторні заняття — 36 год.; самостійна робота — 81 год.;</p> <p><b>Заочна форма здобуття освіти:</b></p> <p>Кількість кредитів ECTS — 4,5; лекції — 8 год.; лабораторні заняття — 8 год.; самостійна робота — 119 год.;</p>
<p>Ознаки курсу</p>	<p>Рік навчання — 3; семестр — 6; Обов'язкова (для здобувачів інших ОП може бути вибірковою) дисципліна; кількість модулів — 2;</p>
<p>Форма контролю</p>	<p>Поточний контроль: захист лабораторних робіт; 2 модульні тести в системі Atutor.</p> <p>Підсумковий контроль: екзамен</p>
<p>Компетентності та дисципліни, що є передумовою для вивчення</p>	<p>Вища математика; Теорія імовірностей, імовірнісні процеси і математична статистика; Програмування; Теорія алгоритмів; Організація баз даних.</p>
<p>Матеріально-технічне та/або інформаційне забезпечення</p>	<p>ПК, Інтернет; MS Excel; WEKA; Python.</p>

## СТРУКТУРА КУРСУ

<b>Лекційний курс</b>	<b>Годин</b>	
	<b>ОФЗО</b>	<b>ЗФЗО</b>
Лекція 1. Вступ до ІАД. Етапи, способи, моделі, методи ІАД. Основні задачі. Дані, їх види.	2	1
Лекція 2. Пошуковий аналіз даних (EDA). Статистичні методи обробки числових даних. Види розподілів даних. Коробкові діаграми.	2	1
Лекція 3. Основи машинного навчання. Задача класифікації. Дерево рішень як класифікатор. Критерії розбиття на класи. Перенавчання алгоритму.	2	1
Лекція 4. Метричні методи класифікації. Метрики якості класифікаторів. Ансамблі дерев рішень. Вибір тренувальних та валідаційних даних для навчання.	2	1
Лекція 5. Задача кластеризації. Ітеративні та ієрархічні методи кластеризації.	2	1
Лекція 6. Вимірювання відстаней. Техніки валідації моделі кластеризації.	2	1
Лекція 7. Задача регресії. Однофакторна та множинна регресії. Метрики якості для регресійної моделі. Мультиколінеарність.	2	1
Лекція 8. Поняття про інші задачі, методи та засоби інтелектуального аналізу даних (часові ряди, асоціативні правила, ін.).	2	0,5
Лекція 9. Візуалізація даних.	2	0,5
<b>РАЗОМ:</b>	<b>18</b>	<b>8</b>

<b>Лабораторний практикум (теми)</b>	<b>Годин</b>	
	<b>ОФЗО</b>	<b>ЗФЗО</b>
ЛР 1. Пошуковий аналіз даних. Дослідження розподілу даних з допомогою електронних таблиць.	9	2
ЛР 2. Аналіз даних з допомогою класифікаційного дерева.	9	2
ЛР 3. Побудова моделі кластеризації.	9	2

Теми занять, короткий зміст

ЛР 4. Знаходження рівняння множинної регресії.

9

2

Передбачено альтернативну траєкторію виконання лабораторного практикуму шляхом опанування курсів освітньої платформи у сфері аналізу даних (Datacamp) у межах неформальної освіти з отриманням відповідних сертифікатів. Така форма роботи спрямована на формування практичних навичок застосування сучасних методів інтелектуального аналізу даних, використання інструментальних засобів і технологій, а також обґрунтованого вибору алгоритмів для розв'язання прикладних задач. Перелік рекомендованих курсів, технологій і навчальних треків визначається викладачем з урахуванням програмних результатів навчання та цілей дисципліни.

РАЗОМ: 36

8

## Курсова робота/проект

Мета виконання курсової роботи	Виконання курсової роботи забезпечує досягнення програмних результатів навчання, зокрема здатності реалізовувати алгоритми аналізу даних засобами сучасних інформаційних технологій, обґрунтовувати вибір моделей і інструментів, оцінювати їх ефективність та застосовувати отримані результати в інформаційних системах для підтримки прийняття рішень.
Завдання курсової роботи	1) аналіз проблеми та/чи методів і алгоритмів інтелектуального аналізу для її вирішення; 2) критичний аналіз існуючих методів і засобів розв'язання аналогічної проблеми; 3) обґрунтування вибору методів та/чи засобів розв'язання проблеми; 4) виконання поставленого індивідуального завдання та оформлення звітної документації.
Структура курсової роботи	Титульний лист; завдання на курсову роботу; анотація; зміст; перелік умовних позначень; вступ; основна частина; висновки; список використаних джерел; додатки.
Обсяг курсової роботи	Рекомендований обсяг - 20-30 сторінок (не включаючи додатків).
Етапи виконання	Вибір та затвердження теми курсової роботи; критичний аналіз нормативно-правової бази, спеціальної літератури з проблем, що розглядаються, пошук додаткових джерел інформації; складання плану курсової роботи; узагальнення та аналіз накопиченого матеріалу, обробка даних, обґрунтування пропозицій; написання тексту і оформлення курсової роботи; захист курсової роботи згідно з встановленим графіком.  Можлива альтернативна траєкторія виконання КР – через сертифікацію в рамках неформальної освіти.
Оцінювання курсової роботи	Зміст курсової роботи – 75 балів, захист курсової роботи – 25 балів.
Форма контролю	Захист курсової роботи передбачає: - стислу доповідь (5 хв.), в якій необхідно коротко висвітлити зміст одержаних результатів дослідження. Зробити акцент на висновках та рекомендаціях. Доповідь під час захисту має супроводжуватись презентацією результатів; - відповіді на запитання наукового керівника та членів комісії / дискусія. Курсова робота та її захист оцінюється відповідно до вимог кредитно-модульної системи.

Технічне й програмне  
забезпечення

Технічні засоби для демонстрування результатів виконання курсової роботи (ноутбук, проєктор). Пакет програмних продуктів Microsoft Office, Python.

## ІНШІ ВИДИ РОБІТ

Теми, короткий зміст

Самостійне опрацювання тем:

Генетичні алгоритми.

Система аналітичного інтерактивного оброблення OLAP.

Ансамблеві математичні моделі ІАД.

Статистична обробка часових рядів.

Проходження курсів на інтерактивній освітній платформі Datacamp з отриманням сертифікатів. Назви курсів, технологій чи треків пропонуються викладачем або вибираються студентом. За кожен отриманий сертифікат студент отримує 5 балів.

Термін проходження курсів - до завершення навчального семестру.

## Інформаційні джерела для вивчення курсу

1. Електронний курс «Інтелектуальний аналіз даних» в системі електронного навчання Atutor (ID: 4684, лектор - Козбур Г.В.), який містить:

- документацію до курсу (робоча програма, силабус);
- рекомендації до роботи з електронним курсом;
- актуальний потижневий план проходження дисципліни;
- теоретичний матеріал та слайди лекцій;
- методичні вказівки до виконання лабораторних робіт та індивідуальні завдання;
- терміни захистів лабораторних робіт та систему оцінювання;
- систему тестування, терміни проходження тренувальних та підсумкових тестів;
- усі актуальні оголошення, опитування, рекомендації;
- додаткову літературу з дисципліни.

Основні зовнішні джерела:

1. Parteek Bhatia. Data Mining and Data Warehousing. Principles and Practical Techniques. Cambridge University Press. 2009.
2. Mohammed J. Zaki, Wagner Meira Jr. Data Mining and Analysis: Fundamental Concepts and Algorithms.
3. Han, Jiawei. Data mining : concepts and techniques / Jiawei Han, Micheline Kamber, Jian Pei. – 3rd ed. 2012.
4. Інтелектуальний аналіз даних та машинне навчання. Частина 1. Базові методи та засоби аналізу даних / Я. В. Іванчук, В. І. Месюра, А. А. Яровий, О. Д. Манжілевський – Вінниця : ВНТУ, 2021. – 69 с.

Основні електронні інформаційні ресурси

1. <https://ir.lib.vntu.edu.ua/bitstream/handle/123456789/36454/120997.pdf?sequence=2&isAllowed=y>
2. <https://ru.coursera.org/specializations/data-mining>
3. <https://myweb.sabanciuniv.edu/rdehkharghani/files/2016/02/The-Morgan-Kaufmann-Series-in-Data-Management-Systems-Jiawei-Han-Micheline-Kamber-Jian-Pei-Data-Mining.-Concepts-and-Techniques-3rd-Edition-Morgan-Kaufmann-2011.pdf>
4. [https://academia.dk/BiologiskAntropologi/Epidemiologi/DataMining/Witten\\_and\\_Frank\\_DataMining\\_Weka\\_2nd\\_Ed\\_2005.pdf](https://academia.dk/BiologiskAntropologi/Epidemiologi/DataMining/Witten_and_Frank_DataMining_Weka_2nd_Ed_2005.pdf)
5. <http://www.cs.waikato.ac.nz/ml/weka/downloading.html>

## Політики курсу

Політика контролю	Використовуються такі засоби оцінювання та методи демонстрування результатів навчання: поточне опитування; тестування; виконання та захист лабораторних робіт; виконання та захист курсової роботи; бесіди та обговорення проблемних питань; дискусії; індивідуальні консультації. Можливий ректорський контроль.
Політика щодо консультування	Консультації при вивченні дисципліни проводяться згідно графіка, затвердженого на кафедрі КН. Консультування передбачено як очно, так і з використанням ресурсів електронного навчального курсу у середовищі електронного навчання університету.
Політика щодо перескладання	Студент має право на одне повторне складання модульного контролю у заліковий тиждень. Повторне складання тесту відбувається лише в присутності викладача. Повторний захист курсової роботи відбувається в терміни, визначені графіком освітнього процесу. Здобувач ВО має право на зарахування результатів навчання здобутих у неформальній чи інформальній освіті.
Політика щодо академічної доброчесності	При складанні усіх видів контролю у середовищі електронного навчання завжди активується система розпізнавання особи, що складає контроль. Усі лабораторні роботи у ЕНК перевіряються вбудованою системою Антиплагіат. При складанні усіх форм контролю забороняється списування, у тому числі з використанням сучасних інформаційних технологій.
Політика щодо відвідування	Відвідування занять є обов'язковим компонентом освітнього процесу. За наявності поважних причин (наприклад, хвороба, особливі потреби, відрядження, сімейні обставини, участь у програмах академічної мобільності тощо) навчання може здійснюватися за індивідуальним графіком, погодженим з деканом факультету.

## СИСТЕМА ОЦІНЮВАННЯ

### Розподіл балів, які отримують студенти за курс

Модуль 1			Модуль 2			Підсумковий контроль		Разом з дисципліни
Аудиторна та самостійна робота			Аудиторна та самостійна робота			Теоретичний курс	Практичне завдання	100
Теоретичний курс (тестування)	Лабораторна робота		Теоретичний курс (тестування)	Лабораторна робота				
8	20		7	20		15	10	
№ лекції	Види робіт	К-ть балів	№ лекції	Види робіт	К-ть балів			
Лекція 1	Лабораторна робота №1	15	Лекції 4,5	Лабораторна робота №3	15			
Лекція 2, 3	Лабораторна робота №2	15	Лекція 6	Лабораторна робота №4	15			
			Лекції 7, 8, 9					

### Розподіл балів, які отримують студенти за виконання та захист КР

Модуль 1		Підсумковий контроль		Разом за КР
Виконання розділу 1		Захист КР		100
75		25		
Види робіт	К-ть балів			
Викладення змісту розділів	60			
Наявність належно оформлених вступу, висновків, переліку посилань і додатків	10			
Відповідність курсової роботи вимогам ЄСКД	5			

## Розподіл оцінок

Сума балів за навчальну діяльність	Шкала ECTS	Оцінка за національною шкалою, екзамен
90-100	A	Відмінно
82-89	B	Добре
75-81	C	Добре
67-74	D	Задовільно
60-66	E	Задовільно
35-59	FX	Незадовільно з можливістю повторного складання
1-34	F	Незадовільно з обов'язковим повторним вивченням дисципліни

Затверджено рішенням кафедри КН, протокол №1 від «26» серпня 2024 року.