

ДЕРЖАВНИЙ ВИЩИЙ НАВЧАЛЬНИЙ ЗАКЛАД
«ПРИДНІПРОВСЬКА ДЕРЖАВНА АКАДЕМІЯ БУДІВНИЦТВА ТА АРХІТЕКТУРИ»

КАФЕДРА КОМП'ЮТЕРНИХ НАУК, ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ
ТА ПРИКЛАДНОЇ МАТЕМАТИКИ
(повна назва кафедри)

«ЗАТВЕРДЖУЮ»
Проректор з навчально-виховної роботи
Галина ЄВССВА
« 02 » вересня 2021 року



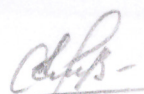
СИЛАБУС НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

«Інтелектуальний аналіз даних»

спеціальність _____ (назва навчальної дисципліни)
122 «Комп'ютерні науки»
освітньо-професійна програма _____ (шифр і назва спеціальності)
«Комп'ютерні науки»
(назва освітньої програми)
освітній ступінь _____ бакалавр
(назва освітнього ступеня)
форма навчання _____ денна
(денна, заочна, вечірня)
розробник _____ Кривенкова Людмила Юріївна
(прізвище, ім'я, по батькові)

1. АНОТАЦІЯ НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Дисципліна «Інтелектуальний аналіз даних» належить до переліку нормативних навчальних дисциплін циклу професійної підготовки, забезпечує професійний розвиток бакалавра та спрямована на формування в майбутнього фахівця основних понять, теоретичних положень і методів аналізу даних. Розглядаються етапи аналізу даних, існуючі методи Data Mining і застосовуються техніки: класифікація і регресії, кластеризація, прогнозування та візуалізація, побудова дерев рішень, пошук асоціативних правил, аналіз часових рядів. Практична значимість курсу полягає у оволодіння базовими знаннями: моделями, методами, технікою, апаратом і алгоритмами вирішення завдань в галузі аналізу даних та відпрацюванні застосування технік Data Mining.



2. ЗМІСТ НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

	Години	Кредити	Семестр	
			VI	VII
Всього годин за навчальним планом, з них:	135	4,5	135	
Аудиторні заняття, у т. ч.:	60		60	
лекції	30		30	
лабораторні роботи	14		14	
практичні заняття	16		16	
Самостійна робота, у т. ч.:	75		75	
підготовка до аудиторних занять	17		17	
підготовка до контрольних заходів				
виконання курсового проекту або роботи				
опрацювання розділів програми, які не викладаються на лекціях	28		28	
підготовка до екзамену	30	1	30	
Форма підсумкового контролю			екзамен	

3. СТИСЛИЙ ОПИС НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Мета дисципліни. Метою викладання дисципліни «Інтелектуальний аналіз даних» є формування системи знань про типи завдань, що виникають в області інтелектуального аналізу даних (Data Mining), здатність до інтелектуального багатовимірного аналізу даних та їхньої оперативної аналітичної обробки з візуалізацією результатів аналізу в процесі розв'язання прикладних задач, вивчення основних підходів та алгоритмів розв'язання задач аналізу даних, методів і моделей сучасної обробки даних, формування практичних навичок роботи з сучасними пакетами прикладних програм для рішення задач аналізу і інтерпретації даних, що виникають в процесі професійної діяльності.

Завдання дисципліни. Основними завданнями вивчення дисципліни «Інтелектуальний аналіз даних» є вивчення основних принципів та інструментарію постановки задач, побудови моделей, використання методів розв'язання задач інтелектуального аналізу даних; вироблення у майбутніх фахівців вміння самостійного вирішення завдань з вибору методів аналізу в практичних ситуаціях; вивчення технології сучасного аналізу даних на комп'ютері; інтерпретація даних реалізації.

Пререквізити дисципліни. Успішне опанування курсу «Інтелектуальний аналіз даних» передбачає знання та навички з курсів «Математичний аналіз», «Інформатика», «Теорія ймовірностей, ймовірнісні процеси та математична статистика», «Алгоритмізація та програмування».

Постреквізити дисципліни. Вивчення дисципліни забезпечує формування у фахівців знання основних понять і методів інтелектуального аналізу даних, вміння застосовувати набути знання у процесі розв'язання професійних задач теоретичного та прикладного характеру при вивченні дисциплін: «Проектування інформаційних систем», «Математичні методи прийняття рішень».

Компетентності:

- **К.** Здатність розв'язувати складні спеціалізовані задачі та практичні проблеми у галузі комп'ютерних наук або у процесі навчання, що передбачає застосування теорій та методів комп'ютерних наук, інформаційних технологій і характеризується комплексністю та невизначеністю умов.
- **ЗК-1.** Здатність до абстрактного мислення, аналізу та синтезу.
- **ЗК-2.** Здатність застосовувати знання у практичних ситуаціях.
- **ЗК-3.** Знання та розуміння предметної області та розуміння професійної діяльності.

- **ЗК-6.** Здатність вчитися і оволодівати сучасними знаннями.
- **ЗК-7.** Здатність до пошуку, оброблення та аналізу інформації з різних джерел.
- **СК-1.** Здатність до математичного та логічного мислення, формулювання та досліджування математичних моделей, зокрема дискретних математичних моделей, обґрунтування вибору методів і підходів для розв'язування теоретичних і прикладних задач в галузі комп'ютерних наук, інтерпретування отриманих результатів.
- **СК-11.** Здатність до інтелектуального багатовимірного аналізу даних та їхньої оперативної аналітичної обробки з візуалізацією результатів аналізу в процесі розв'язання прикладних задач в галузі комп'ютерних наук.

Програмні результати навчання:

- **РН-1.** Здобувати систематичні знання в галузі комп'ютерних наук, аналізувати проблеми з точки зору сучасних наукових парадигм, осмислювати і робити обґрунтовані висновки з наукової і навчальної літератури та результатів експериментів.
- **РН-2.** Реалізовувати засвоєні поняття, концепції, теорії та методи в інтелектуальній і практичній діяльності в галузі комп'ютерних наук, осмислювати зміст і послідовність застосування способів виконання дій, узагальнювати і систематизувати результати робіт.
- **РН-4.** Оцінювати предмет навчальної діяльності, визначати загальну мету і конкретні задачі, вибирати адекватні засоби їх розв'язання для досягнення результату, здійснювати необхідний самоконтроль, використовувати довідкову літературу і технічну документацію, розвивати та застосовувати у професійній діяльності свої творчі здібності, організовувати робоче місце, планувати робочий час.
- **РН-5.** Використовувати технології та інструментарії пошукових систем, методи інтелектуального аналізу даних і текстів, здійснювати опрацювання, інтерпретацію та узагальнення даних.
- **РН-23.** Використовувати технології OLAP, DataMining, TextMining, WebMining в процесі інтелектуального багатовимірного аналізу даних; розв'язувати професійні задачі з використанням методів класифікації, прогнозування, кластерного аналізу, пошуку асоціативних правил.

Заплановані результати навчання:

- знати програмний матеріал; основні поняття, технології і методи із всіх розділів курсу;
- застосовувати базові знання під час розв'язання прикладних і наукових завдань;
- застосовувати методології багатовимірного аналізу, технологій Data Mining, використання інструментальних засобів в області аналізу даних;
- використовувати технології та методи інтелектуального аналізу даних в процесі багатовимірного аналізу даних, здійснювати опрацювання, інтерпретацію та узагальнення даних.
- використовувати технології інтелектуального аналізу даних для розв'язування професійних задач з використанням методів класифікації, прогнозування, кластерного аналізу, пошуку асоціативних правил;
- виконувати необхідні розрахунки в професійній діяльності.

знати:

- базові визначення, основні етапи аналізу і інтерпретації даних;
- можливості аналізу інформації на основі методів інтелектуального аналізу даних;
- основні поняття, задачі, стадії та сучасні методи Data Mining;
- побудову і використання моделей для аналізу даних;
- використання підходів інтелектуального аналізу даних в сучасних областях економіки, науки і освіти;

вміти:

- формулювати завдання аналізу даних;
- вибирати адекватні алгоритми їх вирішення;

- вибирати необхідне програмне забезпечення;
- виконувати комплексний аналіз інформації з метою прийняття рішень на основі використання сучасних методів і систем для інтелектуального аналізу даних.
- оцінювати ефективність застосування різних моделей і методів у конкретних ситуаціях, якість отриманих рішень;
- вирішувати задачі на комп'ютері в середовищі електронних таблиць (пакет аналізу, надбудови «Data Mining» та інші додатки Excel);
- аналізувати отримані результати дослідження.

Методи навчання – практичний, наочний, словесний, робота з книгою. Практичних навичок студенти набувають на аудиторних лекційних, практичних та лабораторних заняттях.

Форми навчання – аудиторна, позааудиторна, індивідуальна.

4. СТРУКТУРА (ТЕМАТИЧНИЙ ПЛАН) ДИСЦИПЛІНИ

Назва змістових модулів і тем	Кількість годин, у тому числі				
	усього	л	п	лаб	с/р
Змістовий модуль 1. Обробка даних та прогнозування					
Основи інтелектуального аналізу даних.	4	2			2
Методи використання навчальної інформації.	16	6	2	2	6
OLAP і Data Mining.	12	2	4	2	4
Методи прогнозування	20	4	4	4	8
Разом за змістовим модулем 1	52	14	10	8	20
Змістовий модуль 2. Методи аналізу					
Методи класифікації.	16	6	2	2	6
Методи кластерного аналізу	24	6	4	4	10
Методи пошуку шаблонів даних	6	2			4
Аналіз часових рядів	7	2			5
Разом за змістовим модулем 2	53	16		6	25
Підготовка до екзамену	30				30
Усього годин	135	30	16	14	75

5. ЛЕКЦІЙНИЙ КУРС

№ зан.	Тема занять	Кількість годин
1	Основи інтелектуального аналізу даних. Уведення у аналіз даних. Предмет, мета і зміст курсу. Методи первісної обробки даних. Методи дослідження структури даних: візуалізація та автоматичне групування даних.	2
2 - 4	Методи використання навчальної інформації. Кореляційний і регресійний аналіз даних. Множинний регресійний аналіз. Лінійна множинна регресійна модель. Перевірка адекватності моделі. Нелінійне оцінювання параметрів.	6
5	OLAP і Data Mining. Поняття OLAP. Методи, стадії, задачі Data Mining. Процес Data Mining. Стандарти Data Mining.	2
6, 7	Методи прогнозування. Аналіз багатомірних угруповань.	4
8 - 10	Методи класифікації. Дерева рішень. Методи опорних векторів, «найближчого сусіда», Байеса.	6
11 - 13	Методи кластерного аналізу. Ієрархічна та секційна кластеризація. Методи кластеризації: процедура Мак-Кина,	6

	метод k-медоїдів, сітчасті методи.	
14	Методи пошуку шаблонів даних. Асоціаційні правила. Послідовне відображення шаблонів даних. Метод Аргіоті. Розробка OLAP-кубів під час аналізу багатомірних даних у БД.	2
15	Статистична обробка часових рядів і прогнозування.	2

6. ТЕМИ ПРАКТИЧНИХ ЗАНЯТЬ

№ зан.	Тема занять	Кількість годин
1 - 3	Кореляційно-регресійний аналіз. Проведення парного та багатомірного кореляційно-регресійний аналізу даних спостережень, аналіз отриманих моделей.	6
4, 5	Прогнозування очікуваних значень цільових параметрів. Прогнозування на основі моделей, отриманих при проведенні кореляційно-регресійного аналізу. Розрахунки прогнозів.	4
6	Методи класифікації. Рішення задачі класифікації методом Байеса	2
7, 8	Методи кластерного аналізу. Рішення задачі кластеризації методом k-середніх в середовищі EXCEL	4

7. ТЕМИ ЛАБОРАТОРНИХ ЗАНЯТЬ

№ зан.	Тема занять	Кількість годин
1	Кореляційний аналіз. Підготовка інформації до проведення кореляційно-регресійного аналізу. Проведення кореляційного аналізу даних спостережень в середовищі MS EXCEL.	2
2-4	Регресійний аналіз Прогнозування очікуваних значень цільових параметрів. Проведення парного та багатомірного регресійний аналізу даних спостережень. Рішення задач оптимізації.	6
5	Методи класифікації. Рішення задачі класифікації методом Байеса	2
6, 7	Методи кластерного аналізу. Рішення задачі кластеризації методом k-середніх в середовищі EXCEL	4

8. САМОСТІЙНА РОБОТА

№ п/п	Вид роботи / Назва теми	Кількість годин
1	підготовка до аудиторних занять	17
2	опрацювання розділів програми, які не викладаються на лекціях:	28
	Основи інтелектуального аналізу даних	2
	Методи використання навчальної інформації	2
	OLAP і Data Mining	6
	Методи прогнозування	2
	Методи класифікації	2
	Методи кластерного аналізу	4
	Методи пошуку шаблонів даних	4
	Статистична обробка часових рядів і прогнозування.	2
	OLAP, Data Mining, СППР	4
3	Підготовка до екзамену	30

9. МЕТОДИ КОНТРОЛЮ

Письмовий, усний, практична перевірка.

10. ПОРЯДОК ТА КРИТЕРІЇ ОЦІНЮВАННЯ

Контроль успішності студента здійснюється за допомогою 100-бальної системи оцінювання, що має відповідні оцінки в національній шкалі і шкалі ECTS.

Змістовий модуль 1. Обробка даних та прогнозування

Лабораторні роботи (максимальна кількість балів – 100 за кожну):

№1 «Кореляційно-регресійний аналіз»,

№2 «Прогнозування очікуваних значень цільових параметрів»

- виконання лабораторної роботи та її оформлення – 60 балів;
- відповідь на теоретичне питання №1, №2 при захисті лабораторної роботи (максимальна кількість балів на одне питання - 20 балів): 40 балів;
 - правильна відповідь на питання 20 балів;
 - відповідь на питання потребує деяких уточнень (були допущені неточності в формулюваннях або тлумаченні результатів) 16 – 19 балів;
 - відповідь на питання із незначними помилками (сутність розкрита, але були помилки в поясненнях або висновках) 6 -15 балів;
 - відповідь на питання зі значними помилками (не розкрита сутність питання, висновки протирічать початковим даним) 1 - 5 балів;
 - неправильна відповідь або немає відповіді 0 балів.

Підсумкова оцінка зі змістового модуля 1 визначається як середня між оцінками за лабораторні роботи змістового модуля 1.

Змістовий модуль 2. Методи аналізу

Лабораторні роботи (максимальна кількість балів – 100 за кожну):

№3 «Методи класифікації»,

№4 «Методи кластерного аналізу»

- виконання лабораторної роботи та її оформлення – 60 балів;
- відповідь на теоретичне питання №1, №2 при захисті лабораторної роботи (максимальна кількість балів на одне питання - 20 балів): 40 балів;
 - правильна відповідь на питання 20 балів;
 - відповідь на питання потребує деяких уточнень застосування визначень або формулювання понять і способів використання наведених методів 16 – 19 балів;
 - відповідь на питання із незначними помилками (сутність розкрита, але були помилки в графічній інтерпретації або застосуванні метода, що реалізується) 6 -15 балів;
 - відповідь на питання зі значними помилками (не розкрита сутність питання або використано помилковий метод розв'язання, аналіз результатів поверхневий) 1 - 5 балів;
 - неправильна відповідь або немає відповіді 0 балів.
- Підсумкова оцінка зі змістового модуля 2 визначається як середня між оцінками за лабораторні роботи змістового модуля 2.

- екзамену

Екзамен проводиться в письмовій формі у вигляді відповідей на білети, що містять два теоретичних питання і 1 практичне завдання. До складання екзамену допускаються студенти, які повністю виконали навчальний план дисципліни – отримали не менше 60 балів з кожного змістового модуля.

Екзамен (максимальна кількість балів – 100):

- відповідь на теоретичні питання (максимальна кількість балів на одне питання - 25 балів): 50 балів;
- правильна відповідь на питання 25 балів;

- відповідь на питання потребує деяких уточнень застосування визначень або формулювання понять і способів використання наведених методів 19 – 24 балів;
- відповідь на питання із незначними помилками (сутність розкрита, але були помилки в формулюваннях) 9 -18 балів;
- відповідь на питання зі значними помилками (не розкрита сутність питання) 1 – 8 балів;
- неправильна відповідь або немає відповіді 0 балів.
 - виконання практичного завдання максимальна кількість балів: 50 балів;
- правильне виконання завдання, реалізація виконана без помилок, проведено аналіз результатів 50 балів;
- робота виконана повністю, але потребує деяких уточнень (були допущені неточності в формулюваннях або тлумаченні результатів) 40 – 49 балів;
- робота виконана повністю, але аналіз результатів недостатній або відсутній 30 – 39 балів;
- виконання завдання із незначними помилками (алгоритм правильний але були помилки в реалізації або використанні помилкових операторів) 20 –29 балів;
- виконання завдання зі значними помилками (реалізована програма, яка не працює) 10 – 19 балів;
- виконання завдання з грубими помилками (неправильний алгоритм або метод) 1 – 9 балів;
- неправильне виконання завдання або немає розв'язання 0 балів.

Підсумкова оцінка з дисципліни визначається як середня між підсумковими оцінками за змістові модулі 1, 2 та оцінкою екзамену.

11. ПОЛІТИКА КУРСУ

Порядок зарахування пропущених занять:

- пропущена лекція відпрацьовується підготовкою конспекту відповідно до теми пропущеного заняття та його захистом;
- пропущені практичні заняття відпрацьовуються студентами виконанням відповідної практичної роботи самостійно та її захистом;
- пропущені лабораторні заняття відпрацьовуються студентами виконанням відповідної лабораторної роботи самостійно та її захистом.

Зміни в нарахуванні балів у випадках несвоєчасного виконання завдань не відбувається.

Дотримання академічної доброчесності студента передбачає:

- самостійне та добросовісне виконання завдань, в тому числі поточного та підсумкового контролю;
- відповідальне ставлення до своїх обов'язків;
- повага до честі й гідності інших осіб;
- посилення на джерела інформації у разі запозичення ідей, розробок, тверджень, відомостей;
- використання при виконанні завдань лише перевірених та достовірних джерел інформації.

За порушення академічної доброчесності студент може бути притягнутий до академічної відповідальності (повторне проходження оцінювання).

12. РЕКОМЕНДОВАНА ЛІТЕРАТУРА

Основна

1. Іванов С.М., Максишко Н.К., Бречко Д.О. Інтелектуальний аналіз даних: конспект лекцій Запоріжжя: ЗНУ, 2020. 156 с.
2. Олійник А. О., Субботін С. О., Олійник О. О. Інтелектуальний аналіз даних: навчальний посібник / Запоріжжя: ЗНТУ, 2012. 278 с.
3. Черняк О.І., Захарченко П.В. Інтелектуальний аналіз даних: підручник, Київ: Знання, 2010. 837 с.


4. В.Є. Бахрушин Методи аналізу даних: навчальний посібник для студентів. – Запоріжжя : КПУ, 2011. 268 с.
5. Анализ данных и процессов: учеб. пособие / А. А. Барсегян, М. С. Куприянов, И. И. Холод, М. Д. Тесс, С. И. Елизаров. - 3-е изд., перераб. и доп. — СПб.: БХВ-Петербург, 2009.
6. Чубукова И.А. Data Mining: учебное пособие, М.: Интернет-университет информационных технологий: БИНОМ: Лаборатория знаний, 2006.
7. Паклин Н.Б., Орешков В.И. Бизнес-аналитика: от данных к знаниям: Уч. пособие. 2-е изд., испр. – СПб.: Питер, 2013. 704 с.
8. В. Дюк, А. Самойленко. «Data Mining», - СПб:- Питер, 2001. - 366 с.
9. Ершова Н.М., Скрипник В.П. Экономико-математические методы и модели принятия решений в условиях определенности, неопределенности и риска. Днепропетровск: Издательство «Свидлер А.Л.», 2010. 350 с.

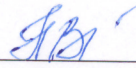
Допоміжна

1. Методичні вказівки до виконання курсової роботи з дисципліни «Інтелектуальний аналіз даних» для студентів ступеня бакалавра спеціальності 122 «Комп'ютерні науки» денної форми навчання / Укладачі: Єршова Н. М., Кривенкова Л. Ю. – Дніпро: ДВНЗ ПДАБА, 2019. - 43 с.
2. Сергеев-Горчинский О. О., Ищенко Г. В. Интеллектуальный анализ данных: компьютерный практикум. Київ: КПП ім. Ігоря Сікорського 2018. 75 с.
3. Зайцев К. С. Применение методов Data Mining для поддержки процессов управления ИТ-услугами: учебное пособие. - М.: МИФИ, 2009. - 96 с.
4. Карлберг Конрад. Бизнес-анализ с помощью Excel: пер. с англ. – К.: Диалектика, 1997. – 448 с.
5. Марманис Х., Бабенко Д. Алгоритмы интеллектуального Интернета. Передовые методики сбора, анализа и обработки данных. – СПб.: Символ-Плюс, 2011. – 480 с.
6. Мусаев А. А. Интеллектуальный анализ данных: учебное пособие. СПб.: СПбГТИ(ТУ), 2018. – 172с.
7. Шумейко А. А. Интеллектуальный анализ данных (Введение в Data Mining): учеб. пособ. / А. А. Шумейко, С. Л. Сотник. – Днепропетровск: Белая Е. А., 2012. – 212 с.

13. INTERNET-РЕСУРСИ

1. Приклади інтелектуального аналізу даних: найпоширеніші програми інтелектуального аналізу даних 2021 року (Електронний ресурс) / Режим доступу: <https://uk.myservername.com/data-mining-examples>
2. Учебники по интеллектуальному, анализу данных (Електронний ресурс) / Режим доступу: <https://msdn.microsoft.com/ru-ru/library/bb677206.aspx>
3. Курс по Data Mining (Електронний ресурс) / Режим доступу: <http://intuit.ru>
4. Луньков А. Д., Харламов А. В. Интеллектуальный анализ данных: учебно-методическое пособие. Режим доступу: http://elibrary.sgu.ru/uch_lit/1141.pdf

Розробник  (Людмила КРИВЕНКОВА)
(підпис)

Гарант освітньої програми  (Наталя ВЕЛЬМАГІНА)
(підпис)

Силабус затверджено на засіданні кафедри комп'ютерних наук, інформаційних технологій та прикладної математики