


КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ТАРАСА ШЕВЧЕНКА

ННІ «Інститут геології»

Кафедра *геоінформатики*

«ЗАТВЕРДЖУЮ»

Заступник директора інституту
з навчальної роботи


«31» 08 2021 року

РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

**АНАЛІЗ ВЕЛИКИХ МАСИВІВ ДАНИХ І БАЗИ ДАНИХ
(BIG DATA AND DATABASES)**

(повна назва навчальної дисципліни)

для студентів

галузь знань
напрямок підготовки
освітній рівень
освітня програма
блок дисциплін
вид дисципліни

19 Архітектура та будівництво

193 «Геодезія та землеустрій»

Магістр

Оцінка землі та нерухомого майна

Інтелектуальний аналіз даних

Вибіркова

Форма навчання	денна
Навчальний рік	2022/2023
Семестр	3
Кількість кредитів ECTS	4
Мова викладання, навчання та оцінювання	українська
Форма заключного контролю	іспит

Викладачі: *Віршило Іван Вікторович, кандидат геологічних наук, доцент кафедри геоінформатики*

Пролонговано: на 20__/20__ н.р. _____ (_____) «__» __ 20__ р.
(підпис, ПІБ, дата)

© Віршило І.В., 2021 рік

КИЇВ - 2021

Розробники: Віршило Іван Вікторович, кандидат геологічних наук, доцент, доцент
кафедри геоінформатики

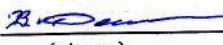
Затверджено
« 14 » 05 2021 р.

зав. кафедри геоінформатики
 (Зацерковний В.І.)
(підпис) (прізвище та ініціали)

Протокол № 11 від « 14 » травень 2021 р.

Схвалено науково - методичною комісією інституту **ННІ «Інститут геології»**

Протокол від « 31 » 08 2021 року №1

Голова науково-методичної комісії  (Демидов В.К.)
(підпис) (прізвище та ініціали)

« 31 » 08 2021 року

Мета дисципліни – ознайомлення студентів із основними методами та підходами до інтелектуального аналізу даних (Data mining) та їх взаємодією з сучасними базами даних(в тому числі і геоданих), детальне вивчення компонентів систем підтримки прийняття рішень, методів фільтрації, трансформації та очищення даних, визначення закономірностей даних. Вироблення у студентів навичок практичної роботи з великими обсягами різноманітних даних (Big Data) у сучасних системах керування базами даних, сховищах та вітринах даних у різноманітних геологічних та споріднених задачах. Навчити студентів досліджувати логічні взаємозв'язки атрибутивної інформації, проводити різні стадії підготовки та інтелектуального аналізу геоданих.

Вимоги до вибору навчальної дисципліни:

1. Успішне опанування щонайменше одного курсу з керування або розробки баз даних, основ великих масивів даних.
2. Володіти навичками програмування, роботи з електронними таблицями, базами даних та іншими джерелами даних.

Анотація навчальної дисципліни / референс:

Відбувається ознайомлення з основами проектування систем інтелектуального аналізу даних, методами роботи з великими даними, побудовою специфічних структур баз даних (сховища даних, вітрини даних). Вивчаються загальні парадигми науки про дані (Data science), їх зв'язок з проектуванням баз даних, їх місце в сучасних системах підтримки прийняття рішень та аналізу даних. Студенти набувають практичних навичок інтелектуального аналізу атрибутивної інформації.

Завдання:

- ознайомити студентів із основними елементами систем підтримки прийняття рішень;
- ознайомити студентів із основними парадигмами організації даних в базах даних, сховищах даних та вітринах даних;
- набуття студентами необхідних методичних та методологічних знань і практичних навичок підготовки даних для подальшого інтелектуального аналізу;
- засвоєння студентами базових знань з основ програмування на мові Python при роботі з великими даними.

Результати навчання:

Результат навчання (1. знати; 2. вміти; 3. комунікація; 4. автономність та відповідальність)		Форма/Методи викладання і навчання	Форма/Методи оцінювання	Відсоток у підсумковій оцінці з дисципліни
Код	Результат навчання			
1.1	Склад систем підтримки прийняття рішень	лекція, практичне заняття	Письмова робота	до 5%
1.2	Реляційні бази даних, сховища та вітрини даних	лекція, практичне заняття	Письмова робота	до 5%
1.3	Методи підготовки даних до інтелектуального аналізу	лекція, практичне заняття	Письмова робота	до 10%
1.4	Моделі Data Mining	лекція, практичне заняття	Письмова робота	до 10%
1.5	Методи Data Mining	лекція, практичне заняття	Письмова робота	до 10%
1.6	Засоби Data Mining	лекція, практичне заняття	Письмова робота	до 10%

2.1	Створювати сховища даних із існуючих реляційних баз даних та інших джерел	практичне заняття, самостійне навчання	Письмова робота, програмне представлення	до 10%
2.2	Проводити підготовку даних для аналізу у системах Data Mining	практичне заняття, самостійне навчання	Письмова робота, програмне представлення	до 10%
2.3	Писати і тестувати програми для роботи з великими даними на мові Python, визначати тестові приклади	практичне заняття, самостійне навчання	Письмова робота, програмне представлення	до 10%
3.1	Організувати подання інформації в базах даних в зрозумілому для інших користувачів вигляді для ефективного вирішення поставленої задачі	практичне заняття	програмне представлення	до 5%
3.2	Формулювати письмові звіти про створені бази даних, алгоритми та команди, ілюструвати приклади роботи розроблених програмних засобів	практичне заняття, самостійне навчання	Письмова робота	До 10%
4.1	Розуміння особистої/персональної відповідальності за особисте рішення частини спільної задачі	--/-	--/-	до 5%

Структура курсу: лекційні і практичні заняття.

Схема формування оцінки:

Форми оцінювання студентів

1. Семестрове оцінювання:

- 1) Контрольна робота із основ організації великих даних, методів підготовки даних до інтелектуального аналізу – 10 балів (рубіжна оцінка 6 балів)
- 2) Контрольна робота із методів інтелектуального аналізу – 10 балів (рубіжна оцінка 6 балів)
- 3) Оцінка за роботу на практичних заняттях – 40 балів (рубіжна оцінка 24 бали)

2. Підсумкове оцінювання у формі іспиту: максимальна оцінка 40 балів, рубіжна оцінка 24 бали. Під час іспиту студент виконує тестове завдання на комплексне знання моделей та методів Data Mining, основних алгоритмів роботи з великими даними, створення сховищ та вітрин даних.

Результати навчальної діяльності студентів оцінюються за 100 бальною шкалою.

Залік виставляється за результатами роботи студента впродовж усього семестру, як сума (проста або зважена) балів за систематичну роботу впродовж семестру.

	Семестрова кількість балів	ПКР(підсумкова контрольна робота) чи/або залік	Підсумкова оцінка
Мінімум	36	24	60
Максимум	60	40	100

Студент не допускається до підсумкового оцінювання у формі іспиту, якщо під час семестру набрав менше 20 балів.

Організація оцінювання: Контроль здійснюється за модульно-рейтинговою системою та передбачає: виконання 2 семестрових практичних проєктів (де студенти мають продемонструвати якість засвоєних знань та вирішити поставлені задачі використовуючи окреслені викладачем методи та засоби), виконання 10 самостійних практичних робіт (де студенти мають продемонструвати якість засвоєних знань та вирішити поставлені задачі без обмеження інструментарію та техніки вирішення проблеми) та проведення 2 письмових модульних контрольних робіт. Підсумкове оцінювання проводиться у формі письмового іспиту.

Шкала відповідності

Відмінно / Excellent	90-100
Добре / Good	75-89
Задовільно / Satisfactory	60-74
Незадовільно / Fail	0-59

СТРУКТУРА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ ТЕМАТИЧНИЙ ПЛАН ЛЕКЦІЙ І ПРАКТИЧНИХ ЗАНЯТЬ

№ п/п	Назва теми	Кількість годин		
		лекції	практичні	Самостійна робота
<i>Розділ 1 Основи організації великих даних</i>				
1	Вступ. Тема 1 Системи підтримки прийняття рішень. OLTP, OLAP та Data Mining системи.	6	4	20
2	Тема 2. Сховища даних та вітрини даних. Підготовка даних.	4		18
	<i>Контрольна робота 1</i>		2	2
<i>Розділ 2 Системи та методи інтелектуального аналізу</i>				
3	Тема 3. Моделі, методи та засоби Data Mining	8	6	18
4	Тема 4. Обробка геоданих засобами OLAP	6		20
	<i>Контрольна робота 2</i>		2	2
	ВСЬОГО	24	14	80

Примітка: слід зазначити теми, винесені на самостійне вивчення

Загальний обсяг 120 год., в тому числі:

Лекцій – 24 год.

Практичні заняття - 14 год.

Консультації - 2 год.

Самостійна робота – 80 год.

РЕКОМЕНДОВАНІ ДЖЕРЕЛА:

Основні:

1. В.Пасічник, В.Резніченко Організація баз даних та знань. – К.: ВНУ, 2006.- 383 с.
2. Д.Ладичук, В.Пічура Базы геоінформаційних даних. – Херсон: ХДУ, 2007.- 103 с.
3. Інтелектуальний аналіз даних : лабораторний практикум : І-навч. посібник / О.Ю. Вінничук, І.С. Вінничук. – Чернівці : Чернівецький нац. ун-т, 2014. – 80 с.
4. Марченко О.О., Россада Т.В. Актуальні проблеми Data Mining: Навчальний посібник. — Київ. — 2017. — 150 с.

Додаткові:

6. Big Data 101 – онлайн-курс (англійською) – www.bdu.intela-edu.com/courses/course-v1:BigDataUniversity+BD0101EN+v2/about
7. Data Science 101 - онлайн-курс (англійською) – www.bdu.intela-edu.com/courses/course-v1:BigDataUniversity+DS0101EN+v1/about
8. SQL and Relational Databases 101 - онлайн-курс (англійською) – <http://courses.bdu.intela-edu.com/courses/course-v1:BDU+DB0101EN+v1/about>
9. Getting Started with DB2 OLAP Server for OS/390 Seungrahn Hahn, Ashvin Amin, Klemens Dickhoefer, Gerard Laumay / SG24-5665-00 / www.redbooks.ibm.com
10. e-Business Intelligence Front-End Tool Access to OS/390 Data Warehouse // Viviane Anavi-Chaput, Kjell Hansson, Charles Lewis, Cheryl Raitakari / SG24-5688-00/ / www.redbooks.ibm.com