


КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ТАРАСА ШЕВЧЕНКА

ННІ «Інститут геології»

Кафедра *геоінформатики*

«ЗАТВЕРДЖУЮ»

Заступник директора інституту
з навчальної роботи


«26» 08 2022 року

РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

ПРОСТОРОВИЙ АНАЛІЗ ТА СТАТИСТИЧНЕ МОДЕЛЮВАННЯ В ГЕОЛОГІЇ

(повна назва навчальної дисципліни)

для студентів

галузь знань
напрямок підготовки
освітній рівень
освітня програма
блоки спеціальних дисциплін

10 Природничі науки
103 «Науки про Землю»
Бакалавр
Геологія та менеджмент надрокористування
Аналіз великих масивів даних ("Big data") в
науках про Землю
Вибіркова

вид дисципліни

Форма навчання	денна
Навчальний рік	2022/2023
Семестр	5, 6
Кількість кредитів ECTS	12
Мова викладання, навчання та оцінювання	українська
Форма заключного контролю	іспит, курсова

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

Демидов Всеволод Кирилович, кандидат фізико-математичних наук, доцент, доцент
кафедри геоінформатики

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

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

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

© Віршило І.В., Демидов В.К., 2022 рік

КИЇВ – 2022

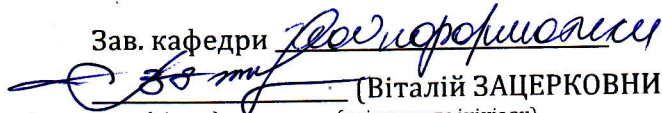
Розробники: Віршило Іван Вікторович, кандидат геологічних наук, доцент кафедри геоінформатики

Демидов Всеволод Кирилович, кандидат фізико-математичних наук, доцент, доцент кафедри геоінформатики

Затверджено

Протокол №1 від «26» серпня 2022 р.

Зав. кафедри



(Віталій ЗАЦЕРКОВНИЙ)

(підпис)

(прізвище та ініціали)

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

Протокол від «26» серпня 2022 року №1

Голова науково-методичної комісії  (Всеволод ДЕМИДОВ)

(підпис)

(прізвище та ініціали)

Мета дисципліни – ознайомлення здобувачів із основами просторового аналізу геоданих, принциповими підходами та методичними особливостями роботи з стратифікованими просторовими даними в геології, що включає в себе тривимірний простір під денною поверхнею та компоненту геологічного часу, головними методами статистичного аналізу в умовах значної неоднорідності даних та багатофакторних залежностей між параметрами стохастичного геологічного середовища.

Вимоги до вибору навчальної дисципліни:

1. Успішне опанування дисциплін з основ інформаційних технологій, статистики або аналогічних за змістом.
2. Володіти елементарними навичками роботи з персональним комп'ютером.
3. Вміти використовувати мову програмування Python для вирішення базових задач.

Анотація навчальної дисципліни / референс:

Відбувається ознайомлення з основами дослідження геоданих, основний фокус зміщений на дані геологічного середовища та геодані, що передбачає вміння аналізувати інформацію у тривимірному просторі, знання систем координат і проєкцій, особливості переходу між ними і їх вплив на аналіз даних. Розглядаються методи просторового статистичного аналізу, в тому числі варіограмного, методи інтерполяційного та апроксимаційного відновлення стохастичних функцій, в тому числі в просторово неоднорідному середовищі, методи класифікаційного, кластерного та регресійного аналізу. Здобувачі знайомляться з етапами підготовки даних, аналізу даних, моделей даних, основними методами та засобами для комп'ютерного аналізу великих масивів різномірних даних, набувають навичок використання спеціалізованого програмного забезпечення та візуалізації результатів.

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 аналізі та моделюванні даних.

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	Основні поняття науки про дані, види та типи даних	лекція, практичне заняття	Письмова робота	до 5%
1.2	Методи просторового аналізу даних	лекція, практичне заняття	Письмова робота	до 5%
1.3	Статистичні властивості просторових даних та основи варіограмного аналізу	лекція, практичне заняття	Письмова робота	до 5%
1.4	Методи апроксимації та інтерполяції даних	лекція, практичне заняття	Письмова робота	до 5%
1.5	Методи кригінгу та їх застосування в геології	лекція, практичне заняття	Письмова робота	до 10%

1.6	Основи описової статистики та регресійного аналізу для застосування до геоданих	лекція, практичне заняття	Письмова робота	до 10%
1.7	Основні методи класифікації та кластерного аналізу даних	лекція, практичне заняття	Письмова робота	до 10%
2.1	Проводити статистичний та просторовий аналіз геоданих засобами комп'ютерної обробки	практичне заняття, курсова робота, самостійне навчання	Письмова робота, програмне представлення	до 10%
2.2	Створювати просторові апроксимаційні моделі геоданих обґрунтованими методами	практичне заняття, курсова робота, самостійне навчання	Письмова робота, програмне представлення	до 10%
2.3	Програмувати прикладні задачі при аналізі та моделюванні даних на мові Python, визначати тестові приклади	практичне заняття, курсова робота, самостійне навчання	Письмова робота, програмне представлення	до 10%
3.1	Організувати подання інформації геоданих в зрозумілому для інших користувачів вигляді для ефективного вирішення поставленої задачі	практичне заняття	програмне представлення	до 5%
3.2	Формулювати електронні звіти про створені моделі даних, алгоритми та команди, ілюструвати приклади роботи розроблених програмних засобів. Поширювати їх для спільних розробок	практичне заняття, самостійне навчання	Письмова робота	До 10%
4.1	Розуміння особистої/персональної відповідальності за особисте рішення частини спільної задачі	лекція, практичне заняття, самостійне навчання	Письмова робота	до 5%

Структура курсу: лекційні і практичні заняття, курсова робота, самостійне навчання.

Схема формування оцінки:

Форми оцінювання здобувачів

1. Семестрове оцінювання:

- 1) Контрольна робота із основ просторового аналізу геоданих – 7 балів (рубіжна оцінка 4 бали)
- 2) Контрольна робота із методів статистичного аналізу геоданих – 7 балів (рубіжна оцінка 4 бали)
- 3) Курсова робота за дисципліною – 16 балів (рубіжна оцінка 10 балів). Оцінка виставляється відповідно базових критеріїв наведених в таблиці нижче в 100 балів та нормується до 16 балів за формулою:
нормована оцінка за курсову = оцінка за курсову здобувача*16/100
- 4) Оцінка за роботу на практичних заняттях – 30 балів (рубіжна оцінка 18 балів)

Базові критерії, які використовуються при оцінюванні курсових робіт

№ п/п	Вид роботи, який оцінюється	Відсоток у підсумковій оцінці за результатами захисту курсової роботи (%)
1	Теоретична та практична значимість основних результатів роботи	15
2	Відповідність отриманих результатів задачам, які ставились при написанні роботи	20
3	Власний внесок студента в основні результати роботи, які представлені до захисту	15
4	Відповідність до вимог оформлення роботи	5
5	Дотримання основних граматичних та стилістичних правил мови на якій представлена робота (як правило, української мови)	5
6	Доповідь здобувача, в якій необхідно представити основні результати власного дослідження у стислій формі	15
7	Якість презентаційного матеріалу доповіді (презентація, графічні додатки, лістинги коду)	10
8	Ґрунтовність відповідей на запитання за результатами доповіді (чіткість, повнота) та дотримання вимог академічної ввічливості	15
	Загальна кількість балів	100

2. Підсумкове оцінювання у формі іспиту: *максимальна оцінка 40 балів, рубіжна оцінка 24 бали. Під час іспиту здобувач виконує тестове завдання на комплексне знання моделей та методів просторового та статистичного аналізу геологічних даних. Підсумкове оцінювання у формі іспиту є обов'язковим.*

Результати навчальної діяльності здобувачів оцінюються за 100 бальною шкалою.

Загальна оцінка виставляється за результатами роботи студента впродовж двох семестрів, оцінювання за підготовку курсової роботи та підсумкового оцінювання, як сума (проста або зважена) балів за систематичну роботу впродовж семестрів, захисту курсової роботи та балів отриманих в результаті підсумкового оцінювання у формі іспиту.

	Семестрова кількість балів			ПКР (підсумкова контрольна робота) чи/або залік	Підсумкова оцінка
	5 семестр Контрольні та практичні роботи	6 семестр Контрольні та практичні роботи	6 семестр Курсова робота		
<i>Мінімум</i>	13	13	10	24	60
Максимум	22	22	16	40	100

Студент не допускається до підсумкового оцінювання у формі іспиту, якщо під час семестру набрав менше 36 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 14 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. Оцінювання в 5 семестрі передбачає проведення проміжного контролю, який виставляється за результатами роботи студента впродовж усього семестру, як сума (проста або зважена) балів за систематичну роботу впродовж 5 семестру. Окремої форми контролю наприкінці 5 семестру не передбачено. За дисципліною передбачено підготовка та захист курсової роботи. Теми курсових робіт визначаються студентом за погодженням з керівником та затверджуються на кафедрі. 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>5 семестр Основи просторового аналізу геоданих</i>				
1	Вступ. Тема 1 Наука про дані. Види та типи даних. Просторові дані та їх представлення.	4	4	18
2	Тема 2. Основи статистичного та просторового аналізу даних. Геостатистика.	8	6	12
3	Тема 3. Просторова інтерполяція та апроксимація. Детерміністичні методи.	6	8	30
4	Тема 4. Геостатистична апроксимація. Варіограмний аналіз. Метод кригінгу. Види кригінгу.	10	20	20
	<i>Контрольна робота 1</i>		2	
	Всього за 5 семестр	28	40	80
<i>6 семестр Статистичний аналіз даних в геології</i>				
5	Тема 5. Використання описової статистики для аналізу геоданих	10	8	20
6	Тема 6. Основи регресійного аналізу	16	20	24
7	Тема 7. Основи класифікації та кластерного аналізу геоданих	22	18	36
	<i>Контрольна робота 2</i>		2	
	<i>Курсова робота</i>			32
	Всього за 6 семестр	48	48	112
	ВСЬОГО	76	88	192

Примітка: слід зазначити теми, винесені на самостійне вивчення

Загальний обсяг: 360 год., в тому числі:

Лекцій – **76 год.**

Практичні заняття - **88 год.**

Консультації - **4 год.**

Самостійна робота – **192 год.**

РЕКОМЕНДОВАНІ ДЖЕРЕЛА:

Основні:

1. Сеньо П. С. Теорія ймовірностей та математична статистика: Підручник. – Київ, Центр навчальної літератури, 2004. – 448 с.
2. Донченко В. С., Сидоров М. В.-С., Шарапов М. М. Теорія ймовірностей та математична статистика: Навчальний посібник. – Київ, ВЦ Академія, 2009. – 288 с.
3. Chiles, J.P., Delfiner, P. (2012). Geostatistics: Modeling Spatial Uncertainty. 2-nd ed. - New York, Toronto, - John Wiley & Sons, Inc., 695 p.
4. Computational and Inferential Thinking: The Foundations of Data Science
<https://inferentialthinking.com/chapters/intro.html>
5. В.Пасічник, В.Резніченко Організація баз даних та знань. – К.: BHV, 2006.- 383 с.
6. Д.Ладичук, В.Пічура Бази геоінформаційних даних. – Херсон: ХДУ, 2007.- 103 с.
7. Інтелектуальний аналіз даних : лабораторний практикум : І-навч. посібник / О.Ю. Вінничук, І.С. Вінничук. – Чернівці : Чернівецький нац. ун-т, 2014. – 80 с.
8. Марченко О.О., Россада Т.В. Актуальні проблеми Data Mining: Навчальний посібник. — Київ. — 2017. — 150 с.
9. Davis John C. (2002) Statistics and Data Analysis in Geology 3-nd ed. - New York, Toronto, - John Wiley & Sons, Inc., 656 p.
10. Pal A., Pal S.K. Pattern recognition and big data. World Scientific Publishing, 2017. – 862 p.
11. Plaksina T. Modern Data Analytics. Applied AI and Machine Learning for Oil and Gas Industry. 2019 – 77p.
12. Дюбрул О. Использование геостатистики для включения в геологическую модель сейсмических данных. – SEG, EAGE 2002. – 296 с.
13. Документація по NumPy та Scipy <https://docs.scipy.org/doc/>
14. Документація по Pandas <http://pandas.pydata.org/pandas-docs/stable/>
15. Документація по Git <https://git-scm.com/book/uk/v2>
16. Документація по DataScience's <http://www.data8.org/datascience/index.html>

Додаткові:

17. Big Data 101 – онлайн-курс (англійською) – www.bdu.intela-edu.com/courses/course-v1:BigDataUniversity+BD0101EN+v2/about
18. Data Science 101 - онлайн-курс (англійською) – www.bdu.intela-edu.com/courses/course-v1:BigDataUniversity+DS0101EN+v1/about
19. EMC: Data Science and Big Data Analytics. In: EMC Education Services, pp. 1–508 (2012)
20. Russom, P.: Big Data Analytics. In: TDWI Best Practices Report, pp. 1–40 (2011)
21. Каждан А.Б., Гуськов О.И. 3. Математические методы в геологии. – М.: Недра, 1990. – 251 с
22. Справочник по математическим методам в геологии / Родионов Д.А., Коган Р.И., Голубева В.А. и др. – М.: Недра, 1987. – 335 с
23. Хом'як М.М. Геостатистика. Курс лекцій. Ч.1. Базові поняття статистики. – Львів: Видавничий центр ЛНУ імені Івана Франка, 2004. – 51 с.
24. Хом'як М.М. Геостатистика. Курс лекцій. Ч.2. Вступ до дисперсійного, кореляційного та регресійного аналізу. – Львів: Видавничий центр ЛНУ імені Івана Франка, 2004. – 47 с.
25. Vyzhva, Z.O. (2011). The Statistical Simulation of Random Processes and Fields. Kyiv: Obrii, 388 p.

26. Vyzhva, Z.O., Demidov V.K., Vyzhva A.S. (2018b) About statistical simulation methods of random fields on the sphere by the aircraft magnetometry data. *Visn. Kyiv University. Geology*, 82, 107-113.
27. Vyzhva, Z.O., Demidov, V.K., Vyzhva, A.S. (2020). The statistical simulation of dataset in 3-D area with spherical correlation function on Rivne NPP example. *Visn. Kyiv University. Geology*, 91, 85-93.