


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

Методи класифікації та кластерного аналізу в розв'язанні геологічних
задач

(повна назва навчальної дисципліни)

для студентів

галузь знань	10 Природничі науки
спеціальність	103 Науки про Землю
освітній рівень	Бакалавр
освітня програма	Геологія та менеджмент надрокористування
блок дисципліни	
вид дисципліни	Вибіркова

Форма навчання	денна
Навчальний рік	2023/2024
Семестр	5
Кількість кредитів ECTS	3
Мова викладання, навчання та оцінювання	українська
Форма заключного контролю	залік

Викладач: *Зацерковний Віталій Іванович, доктор технічних наук, професор, завідувач
кафедри геоінформатики*

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

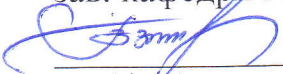
на 20__/20__ н.р. _____ (_____) «__» __ 20__ р.
(підпис, ПІБ, дата)

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

Розробники: **Зацерковний Віталій Іванович, доктор технічних наук, професор, завідувач кафедри геоінформатики**
Зєлик Ярема Ігорович, доктор технічних наук, старший науковий співробітник, доцент кафедри геоінформатик


Затверджено на засіданні кафедри
геоінформатики
Протокол № 1 від 26 серпня 2022 р.

зав. кафедри геоінформатики

 (Віталій ЗАЦЕРКОВНИЙ)
(підпис)

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

Протокол №1 від 26 серпня 2022 року

Голова науково-методичної комісії  (Всеволод ДЕМИДОВ)
(підпис)

Мета викладання дисципліни формування фундаментальних теоретичних знань з дисципліни «Методи класифікації та кластерного аналізу в розв'язанні геологічних задач» і практичних навиків з реалізації принципів аналізу даних у сучасних системах підтримки прийняття рішень, організації сховищ даних, виявлення закономірностей у даних шляхом розв'язання задач кластеризації, класифікації, регресії, прогнозування часових рядів.

Вимоги до вибору навчальної дисципліни:

З метою кращого освоєння навчального матеріалу дисципліни студентам бажано володіти знаннями та навичками програмування.

Анотація навчальної дисципліни / референс:

У програмі навчальної дисципліни розкриваються основні аспекти застосування методів класифікації та кластерного аналізу в розв'язанні геологічних задач у сучасних системах підтримки прийняття рішень, організації сховищ даних, виявлення закономірностей у даних шляхом розв'язання задач класифікації, регресії, прогнозування часових рядів.

Завдання:

- ознайомлення з методами класифікації та кластерного аналізу для розв'язання геологічних задач для аналізу закономірностей у великих обсягах даних;
- навчитися ефективно використовувати методи класифікації та кластерного аналізу для розв'язання геологічних задач;
- вивчити основні принципи обробки великих масивів даних, способи їх зберігання, візуалізації;
- оволодіння практичними навичками з реалізації вказаних моделей та методів класифікації та кластерного аналізу, зокрема:
 - ставити геологічні задачі та вибирати найбільш відповідні методи класифікації та кластерного аналізу для їх розв'язання;
 - створювати алгоритми побудови дерев рішень в розв'язанні геологічних задач;
 - використовувати найбільш відповідні методи класифікації та кластерного аналізу для розв'язання геологічних задач.

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	Місце та роль дисципліни методи класифікації та кластерного аналізу в розв'язанні геологічних задач у системі знань, основні завдання та галузі застосування	лекція, лабораторне заняття, самостійна робота	письмова робота	до 10%
1.2	Визначення призначення методів дерева рішень, опорних векторів та «найближчого сусіда» для розв'язання геологічних задач	лекція, лабораторне заняття, самостійна робота	письмова робота	до 10%
1.3	Геометрична інтерпретація задачі класифікації в розв'язанні геологічних задач	лекції, лабораторні заняття, самостійна робота	письмова робота	до 10%
1.4	Рішення задачі класифікації методом Байєса	лекції, лабораторні заняття, самостійна робота	письмова робота	до 10%
1.5	Базові алгоритми класифікації в розв'язанні геологічних задач	лекції, лабораторні заняття, самостійна робота	письмова робота	до 10%
1.6	Оптимізація параметрів моделей класифікації	лекції, лабораторні заняття, самостійна робота	письмова робота	до 10%
2.1	Використовувати методи дерева рішень, опорних векторів та «найближчого сусіда» у розв'язанні геологічних задач	лекція, лабораторне заняття, самостійна робота	письмова робота	до 5%
2.2	Створювати алгоритми побудови та аналізу моделей класифікації та кластерного аналізу	лекції, лабораторне заняття, самостійна робота	письмова робота	до 5%
2.3	Засвоїти на практиці основні методи класифікації та кластерного аналізу в розв'язанні геологічних задач та обробки даних	лекції, лабораторні заняття, самостійна робота	письмова робота	до 5%
2.4	Створювати алгоритми побудови дерев рішень в розв'язанні геологічних задач	лекції, лабораторні заняття, самостійна робота	письмова робота	до 5%
2.5	Ставити геологічні задачі та вибирати найбільш відповідні методи класифікації та кластерного аналізу для їх розв'язання	лекції, лабораторне заняття, самостійна робота	письмова робота	до 5%
2.6	Використовувати найбільш відповідні методи класифікації та кластерного аналізу для розв'язання геологічних задач	лекція, лабораторне заняття, самостійна робота	письмова робота	до 5%
3.1	Ефективне використання методи	лекції, лабораторне	письмова	до 5%

	класифікації та кластерного аналізу для розв'язання геологічних задач на основі	заняття, самостійна робота	робота	
4.1	Здатність самостійно ставити геологічні задачі та пропонувати і відповідально реалізувати в інформаційних середовищах методи їх розв'язання (здатність працювати автономно та відповідально)	лекції, лабораторні заняття, самостійна робота	письмова робота	до 5%

Структура курсу: лекційні і лабораторні заняття.

Схема формування оцінки:

Форми оцінювання студентів

1. Семестрове оцінювання:

- 1) Контрольна модульна робота «Теоретичні засади методів класифікації та кластерного аналізу в розв'язанні геологічних задач» – 10 балів (рубіжна оцінка 6 балів)
- 2) Контрольна модульна робота «Методи класифікації та кластерного аналізу в розв'язанні геологічних задач» – 10 балів (рубіжна оцінка 6 балів)
- 3) Оцінка за роботу на лабораторних заняттях – 60 балів (рубіжна оцінка 36 бали)

2. Підсумкове оцінювання у формі заліку: максимальна оцінка 20 балів, рубіжна оцінка 12 балів. Під час заліку студент в письмово-усній формі дає відповідь на три питання. Підсумкове оцінювання у формі заліку не є обов'язковим.

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	48	12	60
Максимум	80	20	100

Студент не допускається до підсумкового оцінювання у формі заліку, якщо під час семестру набрав менше 40 балів.

Організація оцінювання: Контроль здійснюється за модульно-рейтинговою системою та передбачає: виконання лабораторних робіт (де студенти мають продемонструвати якість засвоєних знань та вирішити поставлені задачі використовуючи окреслені викладачем методи та засоби), а також проведення письмової модульної контрольної роботи. Залік проводиться у письмово-усній формі.

Шкала відповідності

Зараховано / Passed	60-100
Не зараховано / Fail	0-59

СТРУКТУРА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ ТЕМАТИЧНИЙ ПЛАН ЛЕКЦІЙ І ЛАБОРАТОРНИХ ЗАНЯТЬ

№ п/п	Назва теми	Кількість годин		
		лекції	лабораторні	Самостійна робота
<i>Розділ 1 Теоретичні засади методів класифікації та кластерного аналізу в розв'язанні геологічних задач</i>				
1	Тема 1 Постановка задачі класифікації та кластерного аналізу даних. Правила класифікації та кластерного аналізу даних	4	4	6
2	Тема 2. Методи для розв'язання задач класифікації та кластерного аналізу даних	4	2	10
	<i>Модульна контрольна робота 1</i>	2		
<i>Розділ 2 Застосування методів класифікації та кластерного аналізу в розв'язанні геологічних задач</i>				
3	Тема 3. Задача прогнозування. Точність прогнозу. Види прогнозів. Прогнозування та часові ряди. Методи прогнозування. Задача візуалізації	4	4	10
4	Тема 4. Точність класифікації: оцінка рівня похибок. Оцінювання класифікаційних методів.	4	2	10
5	Тема 5. Ієрархічні методи. Ієрархічне дерево. Міри відстані. Правила об'єднання або ж зв'язку. Ітеративні методи. Метод (алгоритм) k-середніх. Інтерпретація результатів в розв'язанні геологічних задач	6	2	10
	<i>Модульна контрольна робота 2</i>	2		
	<i>Залік</i>	2		
	ВСЬОГО	28	14	46

Примітка: слід зазначити теми, винесені на самостійне вивчення

Загальний обсяг 90 год., у тому числі:

Лекцій – 28 год.

Лабораторні заняття – 14 год.

Консультації – 3 год.

Самостійна робота – 46 год.

РЕКОМЕНДОВАНІ ДЖЕРЕЛА:

1. Гладун А. Я., Рогушина Ю. В. Data Mining: пошук знань в даних / Анатолій Гладун. – К. : ТОВ «ВД «АДЕФ- Україна», 2016. – 452 с.
2. Бахрушин В. Є. Методи аналізу даних : навч. посіб. для студентів / В. Є. Бахрушин. – Запоріжжя: КПУ, 2011. – 268 с.

3. Bodyanskiy, Y., Shafronenko, A., & Mashtalir, S. (2019, May). Online Robust Fuzzy Clustering of Data with Omissions Using Similarity Measure of Special Type. In International Scientific Conference «Intellectual Systems of Decision Making and Problem of Computational Intelligence» (pp. 637-646).
4. Eisenberger J., Posner E.C. Systematic statistics used for data compression in space telemetry. J. Amer. Statist. Ass. 60, 1965. – Pp. 97–133.
5. Gautam, C., & Ravi, V. (2015). Data imputation via evolutionary computation, clustering and a neural network. Neurocomputing, 156, 134-142
6. Gupta S.S., Gnanadesikan M. Estimation of the parameters of the logistic distribution. Biometrika, 53, 1966. – pp. 565-570.
7. Halkidi M., Batistakis Y., Vazirgiannis M. On Clustering Validation Techniques //Journal of Intelligent Information Systems. — 2011. — 17, Issue 2–3. — p. 107–145
8. Hu, Z., Mashtalir, S. V., Tyshchenko, O. K., & Stolbovyi, M. I. (2018). Clustering matrix sequences based on the iterative dynamic time deformation procedure. International Journal of Intelligent Systems and Applications, 10(7), 66-73.
9. Kate, R. J. (2016). Using dynamic time warping distances as features for improved time series classification. Data Mining and Knowledge Discovery, 30(2), 283-312.
10. Mashtalir, S. V., Stolbovyi, M. I., & Yakovlev, S. V. (2019). Clustering Video Sequences by the Method of Harmonic k-Means. Cybernetics and Systems Analysis, 55(2), 200-206.
11. Sarumathi S., Shanthi N., Santhiya G. A Survey of Cluster Ensemble // International Journal of Computer Applications. — 2013. — 65, No.9. — p. 8–11
12. Springer, Cham.
13. Tiwari, K. K., Raguvanshi, V., & Jain, A. (2016). DBSCAN: An Assessment of Density Based Clustering and It's Approaches.
14. Wang, D., Lu, X., & Rinaldo, A. (2017). DBSCAN: Optimal Rates For Density Based Clustering. arXiv preprint arXiv:1706.03113.