КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ТАРАСА ШЕВЧЕНКА

ННІ «Інститут геології»

Кафедра ***геоінформатики***

«ЗАТВЕРДЖУЮ»

Заступник директора інституту
з навчальної роботи
_____Всеволод ДЕМИДОВ
«31» серпня 2026 року

РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

ІНТЕЛЕКТУАЛЬНИЙ АНАЛІЗ ДАНИХ І ШТУЧНИЙ ІНТЕЛЕКТ

(повна назва навчальної дисципліни)

для студентів

галузь знань	**Е Природничі науки, математика та статистика**
спеціальність	**Е4 Науки про Землю**
освітній рівень	**Магістр**
освітня програма	**Промислова геофізика і петрофізика**
блок дисциплін	
вид дисципліни	**Вибіркова**

Форма навчання	**заочна**
Навчальний рік	**2026/2027**
Семестр	**3**
Кількість кредитів ECTS	**6**
Мова викладання, навчання та оцінювання	**українська**
Форма заключного контролю	**залік**

Викладачі: *Демидов Всеволод Кирилович, кандидат фізико-математичних наук, 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)

© Демидов В.К., 2026 рік

КИЇВ – 2026

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

Затверджено
Протокол № 1 від «31» серпня 2026 р.

Зав. кафедри геоінформатики
(підпис) (Віталій ЗАЦЕРКОВНИЙ)
(прізвище та ініціали)

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

Протокол від «31» серпня 2026 року №1

Голова науково-методичної комісії (підпис) (Всеволод ДЕМИДОВ)
(прізвище та ініціали)

1. **Мета дисципліни** – *формування фундаментальних теоретичних знань з наукового напрямку та комплексу інформаційних технологій інтелектуального аналізу даних, призначеного для пошуку у великих обсягах даних неочевидних, об'єктивних та корисних на практиці закономірностей та видобутку знань із різнорідних баз даних. Сформувати теоретичні знання та практичні навички інтелектуального аналізу даних і застосування штучного інтелекту для розв'язання задач промислової геофізики та петрофізики: підготовки даних, виявлення закономірностей, класифікації, регресії, кластеризації та прогнозування.*

– набуття теоретичних і практичних навиків з реалізації принципів інтелектуального аналізу даних у сучасних системах підтримки прийняття рішень, організації сховищ даних, виконання процесів оперативної аналітичної обробки даних (OLAP), виявлення закономірностей у даних шляхом розв'язання задач кластеризації, класифікації, регресії, прогнозування часових рядів, використання знаннєорієнтованого (моделі подання знань та логічне виведення) та конекціоністського (штучні нейронні мережі) підходів у штучному інтелекті до аналізу даних.

2. Вимоги до вибору навчальної дисципліни:

Нема

3. Анотація навчальної дисципліни / референс**:**

Відбувається ознайомлення з основами реалізації принципів інтелектуального аналізу даних у сучасних системах підтримки прийняття рішень, організації сховищ даних, виконанням процесів оперативної аналітичної обробки даних (OLAP), виявлення закономірностей у даних шляхом розв'язання задач кластеризації, класифікації, регресії, прогнозування часових рядів, використання знаннєорієнтованого (моделі подання знань та логічне виведення) та конекціоністського (штучні нейронні мережі) підходів штучного інтелекту до інтелектуального аналізу даних.. Студенти отримують практичні навички командної розробки для представлення даних на прикладах даних гео-направленості. Відбувається знайомство з особливостями представлення гео-даних в базах даних. Допускається локальне або хмарне навчальне середовище з фіксованими залежностями; завдання реалізуються як з GPU так і без.

Академічна доброчесність. Виконання всіх навчальних і контрольних завдань здійснюється з дотриманням Хартії академічної доброчесності та чинних положень Київського національного університету імені Тараса Шевченка. Запозичення, використання даних, ідей, текстів, зображень, цифрових матеріалів та інших результатів мають бути належно атрибутовані. Порушення академічної доброчесності розглядаються відповідно до університетських процедур.

Опанування програмного матеріалу дисципліни, поточний та підсумковий контроль здійснюються на принципах академічної доброчесності відповідно до ст. 42 Закону України «Про освіту» та рекомендацій Міністерства освіти і науки України щодо запобігання академічному плагіату (Лист МОН № 1/9-565 від 15.08.2018 р.), а також Положення про використання штучного інтелекту при реалізації освітніх програм у Київському національному університеті імені Тараса Шевченка https://senate.knu.ua/?p=3466

Використання штучного інтелекту. Для освітнього компонента загалом встановлюється рівень 2 використання генеративного штучного інтелекту згідно «Положення про використання штучного інтелекту при реалізації освітніх програм у Київському національному університеті імені Тараса Шевченка».

Для окремих видів оцінюваних робіт у п.п. 6.1 Форми оцінювання студентів робочої програми визначені спеціальні обмеження, зокрема повна заборона використання генеративного ШІ під час контрольних заходів, що виконуються під наглядом викладача.

Визнання результатів неформальної та/або інформальної освіти. Результати навчання, здобуті в неформальній та/або інформальній освіті, можуть бути подані до визнання у встановленому Університетом порядку за умови їх відповідності результатам навчання цієї дисципліни та наявності підтвердних матеріалів.

4. Завдання:

- *розуміти призначення систем підтримки прийняття рішень, сховищ даних, OLTP та OLAP;*
- *організовувати аналітичні дані з урахуванням якості, метаданих і походження;*
- *ознайомлення з архітектурою та призначенням систем підтримки прийняття рішень, сховищ даних та систем оперативної аналітичної обробки даних;*
- *ознайомлення з методами розв 'язання задач кластеризації, класифікації, регресії для аналізу закономірностей у великих обсягах даних, методами прогнозування часових рядів даних;*
- *ознайомлення з реалізацією знаннєорієнтованого та конекціоністського (штучні нейронні мережі) підходів у штучному інтелекті до аналізу даних;*
- *оволодіння практичними навиками з реалізації вказаних моделей та методів інтелектуального аналізу даних;*
- *ставити задачі кластеризації, класифікації, регресії, пошуку асоціативних правил і прогнозування часових рядів та обирати методи відповідно до структури геофізичних даних;*
- *створення віртуальних сховищ даних та оперування з різними категоріями даних у сховищі.*

5. 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 завдання, етапи та межі застосування інтелектуального аналізу геофізичних і петрофізичних даних*	*лекція, практичне заняття*	*Письмова робота*	*до 10%*
1.2	*Архітектуру та призначення сховищ даних та систем оперативної аналітичної обробки даних (OLAP) та їх реалізацію в СППР*	*лекція, практичне заняття*	*Письмова робота*	*до 10%*
1.3	*Моделі подання знань та реалізацію логічного виведення в інтелектуальному аналізі даних у межах знаннєорієнтованого напряму штучного інтелекту*	*лекція, практичне заняття*	*Письмова робота*	*до 10%*
1.4	*Методи кластеризації, класифікації, регресії, асоціативних правил і прогнозування; принципи валідації, вибору метрик і запобігання витоку даних*	*лекція, практичне заняття*	*Письмова робота*	*до 10%*

1.5	*Основи архітектури й навчання нейронних мереж, регуляризації та контролю перенавчання*	*лекція, практичне заняття*	*Письмова робота*	*до 10%*
2.1	*Використовувати комплексні переваги інтелектуального аналізу даних від інших підходів у різних галузях застосування інтелектуальних інформаційних систем*	*практичне заняття, самостійне навчання*	*Письмова робота, програмне представлення*	*до 10%*
2.2	*Подавати прості експертні правила, застосовувати їх до даних і порівнювати результати з моделями машинного навчання*	*практичне заняття, самостійне навчання*	*Письмова робота, програмне представлення*	*до 10%*
2.3	*Реалізовувати та порівнювати моделі кластеризації, класифікації, регресії або прогнозування; обґрунтовувати поділ даних і метрики*	*практичне заняття, самостійне навчання*	*Письмова робота, програмне представлення*	*до 10%*
2.4	*Навчати базову нейронну мережу, контролювати перенавчання та порівнювати її з класичним методом на однакових вибірках*	*практичне заняття, самостійне навчання*	*Письмова робота, програмне представлення*	*до 10%*
3.1	*Вміти організувати командну розробку для ефективного вирішення поставленої задачі*	*практичне заняття*	*--/-*	*до 5%*
4.1	*Розуміння особистої/персональної відповідальності за особисте рішення частини спільної задачі*	*--/-*	*--/-*	*до 5%*

Структура курсу: *лекційні і практичні заняття, самостійна робота.*

6. Схема формування оцінки:

6.1 Форми оцінювання студентів

Семестрове оцінювання:

1) Оцінка за виконання практичних робіт – 80 балів (рубіжна оцінка 48 балів)

Дозволяється використовувати ШІ при підготовці практичних робіт як допоміжного інструменту за умови збереження авторського вкладу в підготовку програмного коду практичної роботи та розрахункових задач. Дозволено використовувати ШІ для довідкових функцій під час підготовки практичної роботи та як інструмент перевірки орфографії, технічного оформлення роботи. Декларування факту та способу використання ШІ є обов'язковим. Під час захисту практичних робіт використання ШІ заборонено.

Підсумкове оцінювання *у формі заліку*: *максимальна оцінка 20 балів, рубіжна оцінка 12 балів. Під час заліку студент виконує поставлені завдання з використанням знань та вмінь з інтелектуального аналізу даних. Під час підготовки письмової залікової роботи використання ШІ заборонено.* ***Підсумкове оцінювання у формі заліку не є обов'язковим, при відмові від участі у даній формі оцінювання студент не отримає відповідні бали до підсумкової оцінки.***

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 та заліку.*

	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.*

6.2 Організація оцінювання: *Контроль здійснюється за модульно-рейтинговою системою та передбачає: виконання 4 практичних робіт (де студенти мають продемонструвати якість засвоєних знань та вирішити поставлені задачі використовуючи окреслені викладачем методи та засоби). Підсумкове оцінювання проводиться у формі письмового заліку.*

6.3 Шкала відповідності

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

7. СТРУКТУРА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ
ТЕМАТИЧНИЙ ПЛАН ЛЕКЦІЙ І ПРАКТИЧНИХ ЗАНЯТЬ

№ п/п	**Назва теми**	**Кількість годин**		
		лекції	**практичні**	**Самостійна робота**
Розділ 1 *Теоретичні засади, архітектура та функціонування сучасних інформаційних систем інтелектуального аналізу даних*				
1	**Вступ**. **Тема 1.** Основні завдання інтелектуального аналізу даних, системи підтримки прийняття рішень та системи оперативної обробки транзакцій	1	1	38
2	**Тема 2.** Сховища даних та системи оперативної аналітичної обробки даних (OLAP) - основа сучасних високопродуктивних систем інтелектуального аналізу даних	1	1	40
Розділ 2 *Напрями та методи штучного інтелекту в інтелектуальному аналізі даних*				
3	**Тема 3.** Кластеризація, класифікація, регресія, пошук асоціативних правил - основні завдання інтелектуального аналізу даних	1	1	40
4	**Тема 4.** Штучні нейронні мережі в інтелектуальному аналізі даних у межах	1	1	50

	конекціоністського напряму штучного інтелекту			
	Залік			2
	ВСЬОГО	**4**	**4**	**170**

Примітка: слід зазначити теми, винесені на самостійне вивчення

Загальний обсяг *180 год.*, в тому числі:
Лекцій **– *4** год.*
Практичні заняття **- *4** год.*
Консультації **- *2** год.*
Самостійна робота **- *170** год.*

Теми для самостійного навчання:

1. *Життєвий цикл аналізу даних: предметна задача, цільова змінна, одиниця спостереження та критерій успіху; відмінності описового аналізу, прогнозу й причинного висновку.*
2. *Каротажні дані, лабораторні вимірювання керна та метадані: одиниці, глибини, узгодження інтервалів, пропуски, дублікати та фізично неможливі значення. Розмежування помилки вимірювання й геологічної аномалії.*
3. *Python, NumPy, pandas і візуалізація: імпорт таблиць, узгодження типів, описові статистики, розподіли та зв'язки ознак. Фіксація джерела, ліцензії та версії набору даних.*
4. *OLTP, системи підтримки прийняття рішень, експертні правила й логічне виведення: простий приклад правила для геофізичних ознак та межі його застосування.*
5. *Відмінності бази даних, аналітичного сховища й озера даних; факти та виміри, схема «зірка», деталізація запису, OLAP-агрегації й ризик некоректного усереднення вимірювань.*
6. *ETL/ELT: очищення, об'єднання джерел, контроль схеми та походження даних; CSV і Parquet; SQL-запити для групування за свердловиною, інтервалом або літологією.*
7. *Реалізація навчального сховища у SQLite або DuckDB; використання pandas для підготовки та аналізу агрегованих даних.*
8. *Локальний Jupyter або хмарний ноутбук, фіксація залежностей і випадкового стану, Git та журнал експериментів. Обмеження ресурсів; непублічні промислові дані не завантажуються до зовнішніх сервісів без дозволу власника.*
9. *Розподіл на навчальну, валідаційну й тестову вибірки; групова крос-валідація за свердловинами, просторово відокремлена перевірка або хронологічний поділ залежно від задачі. Запобігання витоку даних за допомогою Pipeline.*
10. *Регресія та класифікація: лінійні моделі, регуляризація, kNN, дерева, Random Forest і градієнтний бустинг; порівняння з простим базовим прогнозом. Підготовка ознак навчається лише на навчальній вибірці.*
11. *Метрики: MAE, RMSE і R^2; precision, recall, F1, матриця помилок, ROC-AUC та PR-AUC. Дисбаланс класів, вибір порога й помилки біля меж літологічних класів.*
12. *Кластеризація k-means і DBSCAN, масштабування ознак, PCA; оцінювання стабільності кластерів і геологічна інтерпретація. Візуальні групи не є автоматичним підтвердженням геологічних класів.*
13. *Асоціативні правила: support, confidence та lift. Часові ряди: лагові ознаки, сезонність і базовий прогноз; перевірка без використання майбутніх спостережень.*
14. *Інтерпретація моделей: permutation importance, SHAP як додатковий інструмент; корельовані ознаки, невизначеність і межі перенесення моделі на нову свердловину.*
15. *Перцептрон і багатошаровий перцептрон, функції активації, функції втрат, оптимізатори, навчання мініпакетами; базова реалізація в PyTorch на CPU.*
16. *Криві навчання, нормалізація, регуляризація, dropout і рання зупинка; вибір гіперпараметрів без використання тестової вибірки.*
17. *Порівняння MLP з деревами та ансамблями на табличних петрофізичних даних. Обґрунтування складності моделі та оцінювання на однаковому поділі даних.*

18. *Огляд CNN для зображень керна та часових/послідовнісних моделей для сигналів. Трансформери й генеративний ШІ — додаткове ознайомлення без обов'язкового навчання великих моделей.*
19. *Відтворюваність експериментів, збереження моделі та опис навчальних даних, метрик і обмежень. MLflow як додатковий засіб обліку експериментів.*
20. *Використання генеративного ШІ для допомоги з кодом: перевірка результатів, декларування використання та особисте пояснення рішення; достовірність і конфіденційність даних.*

РЕКОМЕНДОВАНІ ДЖЕРЕЛА:

Основні:

1. *Han J., Pei J., Tong H., (2022), Data Mining: Concepts and Techniques. 4th ed. Morgan Kaufmann. https://shop.elsevier.com/books/data-mining/han/978-0-12-811760-6*
2. *James G., Witten D., Hastie T., Tibshirani R., Taylor J., (2023), An Introduction to Statistical Learning: with Applications in Python. Springer. https://www.statlearning.com/*
3. *Géron A., (2025), Hands-On Machine Learning with Scikit-Learn and PyTorch. O'Reilly Media. https://www.oreilly.com/library/view/hands-on-machine-learning/9798341607972/*
4. *Reis J., Housley M., (2022), Fundamentals of Data Engineering: Plan and Build Robust Data Systems. O'Reilly Media. https://www.oreilly.com/library/view/fundamentals-of-data/9781098108298/*
5. *scikit-learn Developers. User Guide: preprocessing, pipelines, model selection, metrics and common pitfalls. URL: https://scikit-learn.org/stable/user_guide.html*
6. *I. Kamber, Micheline. II. Pei, Jian., (1012), Data Mining Concepts and Techniques, - Morgan Kaufmann Publishers is an imprint of Elsevier.*
7. *Nathan Greeneltch, (2019), Python Data Mining. Quick Start Guide, - Published by Packt Publishing Ltd.*

Додаткові:

8. *Molnar C., (2025),Interpretable Machine Learning: A Guide for Making Black Box Models Explainable. 3rd ed.. https://christophm.github.io/interpretable-ml-book/*
9. *Hyndman R. J., Athanasopoulos G., (2021), Forecasting: Principles and Practice. 3rd ed. OTexts. https://otexts.com/fpp3/*
10. *Dorman M., Graser A., Nowosad J., Lovelace R., (2025),Geocomputation with Python. Chapman & Hall/CRC. https://py.geocompx.org/*
11. *PyTorch Contributors. Tutorials. https://pytorch.org/tutorials/*
12. *DuckDB Contributors. Documentation. https://duckdb.org/docs/stable/*
13. *MLflow Contributors. ML Experiment Tracking. https://mlflow.org/docs/latest/ml/tracking*
14. *Вінничук О. Ю., (2014), Інтелектуальний аналіз даних: лабораторн. практикум: навч. посіб. М-во освіти і науки України, Чернів. нац. ун-т ім. Юрія Федьковича; уклад.: Вінничук О. Ю., Вінничук І. С. - Чернівці: ЧНУ.*
15. *Ліщина Н.М., (2016),Інтелектуальний аналіз даних [Текст] : конспект лекцій для студентів напряму підготовки 6.050101 «Комп'ютерні науки» денної форми навчання / уклад. Н.М.Ліщина. - Луцьк : Луцький НТУ.*
16. *Черняк О. І., (2014), Інтелектуальний аналіз даних: підручник. Київ. нац. ун-т ім. Тараса Шевченка. - К.: Знання.*