



ТЕХНОЛОГІЇ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ДАНИХ

Робоча програма навчальної дисципліни (Силабус)

Реквізити навчальної дисципліни

Рівень вищої освіти	<i>Другий (магістерський)</i>
Галузь знань	<i>12 Інформаційні технології</i>
Спеціальність	<i>ОНП "Комп'ютерні науки"</i>
Освітня програма	<i>Комп'ютерні науки</i>
Статус дисципліни	<i>Нормативна</i>
Форма навчання	<i>дистанційна</i>
Рік підготовки, семестр	<i>1 курс, осінній семестр</i>
Обсяг дисципліни	<i>4 кредити ЄКТС, лекції — 30 годин, лабораторні роботи - 14 годин</i>
Семестровий контроль/ контрольні заходи	<i>МКР, РГР, залік</i>
Розклад занять	<i>https://schedule.kpi.ua/ (2 год. лекційних та 1 год. практичних занять на тиждень)</i>
Мова викладання	<i>Українська</i>
Інформація про керівника курсу / викладачів	Лектор: д.ф.-м.н., професор, Купенко Ольга Петрівна, kupenko.olga@gmail.com Практичні: д.ф.-м.н., професор, Купенко Ольга Петрівна, kupenko.olga@gmail.com
Розміщення курсу	<i>Платформа дистанційного навчання "Сікорський", https://classroom.google.com/c/Njl4MjE2MDEzNTAz?cjc=odhjncf</i>

Програма навчальної дисципліни

1. Опис навчальної дисципліни, її мета, предмет вивчення та результати навчання

Цілі дисципліни: сформувати навички щодо дослідження і виявлення алгоритмами, засобами штучного інтелекту в «сирих даних» прихованих структур, шаблонів (паттернів) або залежностей, які раніше не були відомі, нетривіальні, практично корисні та доступні для інтерпретації людиною і необхідні для ухвалення рішень в різних сферах діяльності.

Завданням дисципліни «Технології інтелектуального аналізу даних» є надання студентам ґрунтовних знань в області аналітичних досліджень інформаційного простору, вивчення методів створення, добування, консолідації, переробки, трансформації та аналізу даних.

Предметом вивчення дисципліни є основні положення, методи та алгоритми інтелектуального аналізу даних, зокрема із залученням машинного навчання та штучного інтелекту, та їх комп'ютерна реалізація з використанням мови програмування Python, зокрема спеціальних бібліотек, таких як Pandas, Numpy, sklearn, Keras.

Компетентності:

ЗК 1 Здатність до абстрактного мислення, аналізу та синтезу.

ЗК 2 Здатність застосовувати знання у практичних ситуаціях.

ЗК 5 Здатність вчитися й оволодівати сучасними знаннями.

ЗК 7 Здатність генерувати нові ідеї (креативність).

ФК 1 Усвідомлення теоретичних засад комп'ютерних наук.

ФК 5 Здатність розробляти, описувати, аналізувати та оптимізувати архітектурні рішення інформаційних та комп'ютерних систем різного призначення.

ФК 7 Здатність розробляти програмне забезпечення відповідно до сформульованих вимог з урахуванням наявних ресурсів та обмежень.

ФК 17 Здатність вибирати адекватні методи і технології обчислювального інтелекту та машинного навчання, включаючи методи глибокого навчання, еволюційного моделювання, генетичні алгоритми, та використовувати їх для вирішення задач прогнозування, керування, прийняття рішень, класифікації та інтелектуального аналізу даних в умовах невизначеності та неповної інформації.

Програмні результати навчання:

ПРН 1 Мати спеціалізовані концептуальні знання, що включають сучасні наукові здобутки у сфері комп'ютерних наук і є основою для оригінального мислення та проведення досліджень, критичне осмислення проблем у сфері комп'ютерних наук та на межі галузей знань.

ПРН 2 Мати спеціалізовані уміння/навички розв'язання проблем комп'ютерних наук, необхідні для проведення досліджень та/або провадження інноваційної діяльності з метою розвитку нових знань та процедур.

ПРН 6 Розробляти концептуальну модель інформаційної або комп'ютерної системи.

ПРН 7 Розробляти та застосовувати математичні методи для аналізу інформаційних моделей.

ПРН 9 Розробляти алгоритмічне та програмне забезпечення для аналізу даних (включно з великими).

ПРН 11 Створювати нові алгоритми розв'язування задач у сфері комп'ютерних наук, оцінювати їх ефективність та обмеження на їх застосування.

ПРН 12 Проектувати та супроводжувати бази даних та знань.

ПРН 14 Тестувати програмне забезпечення.

ПРН 18 Збирати, формалізувати, систематизувати і аналізувати потреби та вимоги до інформаційної або комп'ютерної системи, що розробляється, експлуатується чи супроводжується.

ПРН 19 Аналізувати сучасний стан і світові тенденції розвитку комп'ютерних наук та інформаційних технологій.

ПРН 20 Створювати та досліджувати інформаційні та математичні моделі систем і процесів, що досліджуються, зокрема об'єктів автоматизації.

ПРН 26 Застосовувати технології обчислювального інтелекту та інтелектуального аналізу даних, зокрема, нейронні мережі, нечіткі нейронні мережі, нейронні мережі глибокого навчання, методи машинного навчання для проектування та адаптації інтелектуальних систем прийняття рішень в різних предметних сферах.

2. Пререквізити та постреквізити дисципліни (місце в структурно-логічній схемі навчання за відповідною освітньою програмою)

Курс "Технології Інтелектуального аналізу даних" є одним із основних курсів підготовки магістрів науковців спеціальності "Штучний інтелект".

Цей курс підсумовує раніше прочитані спеціальні дисципліни в напрямку теорії та систем прийняття рішень і дає систематизоване детальне викладання основ теорії, методів та засобів побудови систем обчислювального інтелекту та їх застосування в системах

прийняття рішень в економіці, бізнесі та фінансовій сфері. Тому ця дисципліна має глибокі логічні зв'язки з попередніми дисциплінами навчального плану підготовки, зокрема з курсами "Дослідження операцій", "Теорія прийняття рішень в складних системах", "Моделювання економічних систем", "Статистичний аналіз економічних процесів".

Матеріали курсу широко використовуються при підготовці магістерської дисертації.

3. Зміст навчальної дисципліни

Освітній компонент охоплює такі теми:

Тема 1. Основні задачі Data Mining
Тема 2. Інформативність факторів. Способи підвищення інформативності
Тема 3. Алгоритм Фаррара-Глобера
Тема 4. Формування множини значимих факторів
Тема 5. Підходи до машинного навчання з вчителем та без нього
Тема 6. Навчання нейронних мереж прямого поширення.
Тема 7. Послідовні шаблони
Тема 8. Методи класифікації
Тема 9. Кластеризація. Гіпотеза компактності. Еволюційний пошук
Тема 10. Дерева рішень для класифікації та регресії
Тема 11. Класифікація з незбалансованими даними
Тема 12. Байєсові мережі
Тема 13. Відновлення інформації в наборах даних

ДИДАКТИЧНІ МЕТОДИ:

На лекційних заняттях: Лекція, пояснення, мозковий штурм, проблемні завдання.

На лабораторних заняттях: Завдання до виконання, виконання розрахунково-графічної роботи, опитування та тестування студентів

4. Навчальні матеріали та ресурси

1. Базова література

1. M. Zgurovsky, Yu. Zaychenko. Fundamentals of computational intelligence- System approach. Springer, 2016.-275 p. (НТБ ім. Г.І. Денисенка) — обов'язкова до прочитання
2. Ланде Д.В. Основи теорії і практики інтелектуального аналізу даних у сфері кібербезпеки: навчальний посібник / Д. В. Ланде, І. Ю. Субач, Ю. Є. Бояринова ; КПІ ім. Ігоря Сікорського. — Київ : КПІ ім. Ігоря Сікорського, 2018. — 300 с. (НТБ ім. Г.І. Денисенка) — ознайомитись
3. Гороховатський В. О. Методи інтелектуального аналізу та оброблення даних : навч. посіб. / В. О. Гороховатський, І. С. Творошенко ; М-во освіти і науки України, Харків. нац. ун-т радіоелектроніки. — Харків : ХНУРЕ, 2021. — 92 с. — <https://openarchive.nure.ua/handle/document/15868> — обов'язкова до прочитання
4. Інтелектуальний аналіз даних : практикум / М. Т. Фісун, І. О. Кравець, П. П. Казмірчук, С. Г. Ніколенко; М-во освіти і науки України, Чорноморський державний університет ім. Петра Могили. — Львів: Новий Світ–2000, 2023. — 160 с. (НТБ ім. Г.І. Денисенка) — ознайомитись.
5. Москаленко В. В. Моделі і методи інтелектуального аналізу багатовимірних даних за умов апіорної невизначеності. — Суми : Сумський державний університет, 2020. — 184 с. — https://essuir.sumdu.edu.ua/bitstream-download/123456789/77692/3/Moskalenko_analiz_danykh.pdf — ознайомитись

2. Допоміжна література

1. Zgurovsky M., Zaychenko Yu. Big Data: Conceptual Analysis and Applications. Springer Nature Switzerland AG. — 2020. — 275 p. (НТБ ім. Г.І. Денисенка) - факультативна
2. Землянський О.М., Пашинська Н.М., Снитюк В.Є. Інформаційно-аналітичні технології прогнозування наслідків хімічних аварій, Київ, ВПЦ "Київський університет", 2017, 167 с. (за запитом викладачу) - факультативна

3. *Huyen C. Designing Machine Learning Systems. - O'Reilly Media. USA – 2022. – 463 p. (за запитом викладачу) - факультативна*
4. *Kantarzic M. Data Mining. Concepts, Models, Methods and Algorithms / M. Kantarzic, 3rd Ed. – Publisher : Wiley, 2019. – 672 p (за запитом викладачу) - факультативна*

Навчальний контент

5. Методика опанування навчальної дисципліни (освітнього компонента)

Навчальна дисципліна охоплює 30 годин лекцій та 14 годин лабораторних робіт, виконання розрахунково-графічної роботи, а також виконання модульної контрольної роботи.

У межах дисципліни заплановано такі види занять:

- *лекції;*
- *лабораторні роботи;*
- *самостійна робота.*

Теми дисципліни взаємопов'язані, матеріал вивчається в логічній послідовності. На лекціях розкриваються найсуттєвіші теоретичні питання, які дають змогу забезпечити студентам можливість глибокого самостійного вивчення всього програмного матеріалу. Теми та порядок виконання практичних завдань сформовано в логічній послідовності і повністю узгоджується і підтримується лекційним матеріалом. Теоретичні та практичні знання та навички поглиблюються шляхом самостійної роботи з використанням рекомендованої літератури та веб-джерел.

Термін виконання (тиждень)	Назви розділів і тем
Тема 1. Основні задачі Data Mining	
1	Лекція 1. Задачі Data Mining. Ключові відмінності між засобами Data Mining і OLAP. Концепції сховищ даних
Тема 2. Інформативність факторів	
2	Лекція 2. Ентропія та інформативність факторів Способи підвищення інформативності Лабораторна робота 1. Препроцесинг інформації та побудова множинної лінійної регресії
Тема 3. Алгоритм Фаррара-Глобера	
3	Лекція 3. Алгоритм Фаррара-Глобера перевірки вхідних даних
Тема 4. Множина значимих факторів	
4	Лекція 4. Формування множини значимих факторів Лабораторна робота 2. Визначення множини значимих факторів та побудова множинної нелінійної регресії
Тема 5. Підходи до машинного навчання з вчителем та без нього	
5	Лекція 5. Підходи до машинного навчання з вчителем. Підходи до машинного навчання без вчителя
Тема 6. Навчання нейронних мереж прямого поширення	
6	Лекція 6. Методи навчання нейронних мереж прямого поширення
Тема 7. Асоціативні правила та пошук послідовних шаблонів	
7	Лекція 7. Асоціативні правила та пошук послідовних шаблонів Лабораторна робота 3. Пошук асоціативних правил та послідовних шаблонів
Тема 8. Методи класифікації	
8	Лекція 8. Класичні, ітераційні та евристичні методи класифікації Лабораторна робота 4. Класифікація об'єктів у змішаному просторі.

Термін виконання (тиждень)	Назви розділів і тем
	Тема 9. Кластеризація. Гіпотеза компактності. Еволюційний пошук
9	Лекція 9. Кластеризація
10	Лекція 10. Гіпотеза компактності. Еволюційний пошук Лабораторна робота 5. Кластеризація на множині об'єктів дійсного простору
	Тема 10. Дерева рішень для класифікації та регресії
11	Лекція 11. Алгоритми побудови дерев рішень для класифікації та регресії
	Тема 11. Класифікація з незбалансованими даними
12	Лекція 12. Підходи та методи до класифікації з незбалансованими даними. Активне навчання. Метрики для активного навчання Розрахунково-графічна робота. Повний цикл передобробки реальних даних, feature engineering, добування знань з даних, порівняння алгоритмів
	Тема 12. Байєсові мережі
13	Лекція 13. Байєсові мережі
	Тема 13. Відновлення інформації в наборах даних
14	Лекція 14. Підходи та методи до відновлення інформації в наборах даних
15	Лабораторна робота 6. Створення бази даних з MySQL та побудова аналітичних репортів в Tableau. Модульна контрольна робота

ДИДАКТИЧНІ МЕТОДИ:

На лекційних заняттях: Лекція, пояснення, мозковий штурм, проблемні завдання.

На лабораторних заняттях: Завдання до виконання, опитування та тестування студентів

6. Самостійна робота здобувача вищої освіти

Важливим аспектом для якісного засвоєння матеріалу, відпрацювання прийомів і алгоритмів вирішення основних завдань дисципліни є самостійна робота, що дає змогу перетворити отримані знання в об'єкт власної діяльності. Самостійна робота включає в себе читання літератури, огляд літератури по темі, виконання звітів з комп'ютерних практикумів, підготовку до їх захисту, виконання розрахунково-графічної роботи та підготовку до заліку.

Політика та контроль

7. Політика навчальної дисципліни (освітнього компонента)

- Правила відвідування занять — відвідування лекційних занять є обов'язковим, студенту рекомендується вести докладний конспект лекцій. Відвідування лабораторних занять вільне - для захисту результатів виконання лабораторних та розрахункових робіт та консультування з викладачем з приводу ходу виконання робіт;
- правила захисту лабораторних та розрахункових робіт - студент демонструє викладачу номер варіанту лабораторної роботи, вхідні дані, код та результати розрахунків, надає роз'яснення використаних методів та алгоритмів, аналізує власні результати;
- Штрафні бали призначаються студентам, які несвоєчасно захистили лабораторні роботи, не виконали роботи в повному обсязі, не виконали вимоги щодо якісного оформлення результатів робіт. Відповідно до Положення про систему оцінювання результатів навчання сума всіх штрафних балів не може перевищувати 10% рейтингової шкали оцінювання. Зокрема, здача лабораторної роботи пізніше встановленого терміну — -1 бал (за кожен тиждень пізніше встановленого терміну); участь у модернізації лабораторних робіт, удосконалення дидактичних матеріалів з дисципліни надається +5-10 заохочувальних балів.

- політика дедлайнів та перескладань — лабораторні та розрахункові роботи можна здавати до останнього дня перед захистом за попереднім узгодженням часу з викладачем. Перескладання заліку допустиме в період сесії та час призначений для перескладань;
- політика щодо академічної доброчесності — лабораторні або розрахункові роботи, що зроблені не самостійно, не приймаються до оцінювання;
- Кожен студент має право відпрацювати пропущені з поважної причини (лікарняний, мобільність тощо) заняття за рахунок самостійної роботи. Детальніше за посиланням: <https://kpi.ua/files/n3277.pdf>.
- Студент може підняти будь-яке питання, яке стосується процедури контрольних заходів та очікувати, що воно буде розглянуто згідно із наперед визначеними процедурами. Студенти мають право аргументовано оскаржити результати контрольних заходів, пояснивши з яким критерієм не погоджуються відповідно до оціночного.
- Підготування до лабораторних робіт та контрольних заходів здійснюється під час самостійної роботи студентів з можливістю консультування з викладачем у визначений час консультацій або за допомогою електронного листування (електронна пошта, месенджери).
- Календарний контроль проводиться з метою підвищення якості навчання студентів та моніторингу виконання студентом вимог syllabusу.

Критерій		Перший календарний контроль	Другий календарний контроль
Термін календарного контролю		Тиждень 8	Тиждень 14
Умови отримання позитивної оцінки	Поточний рейтинг	≥ 10 балів	≥ 20 балів

8. Види контролю та рейтингова система оцінювання результатів навчання (PCO)

Поточний контроль: фронтальний (усний, письмовий), лабораторні роботи.

Календарний контроль: проводиться двічі на семестр як моніторинг поточного стану виконання вимог syllabusу.

Семестровий контроль: виконання розрахунково-графічної роботи, залік

Рейтингова система оцінювання охоплює всі види тестування: контрольні роботи, активність на лабораторних заняттях та якість захисту лабораторних робіт. Кожен студент отримує свій підсумковий рейтинг за дисципліною.

Рейтинг студента з кредитного модуля у семестрі складається з балів, які він отримує за:

- написання модульної контрольної роботи;
- виконання розрахунково графічної роботи;
- робота на лабораторних заняттях та захист лабораторних робіт;
- відповіді на заліку.

Рейтингова оцінка rP з дисципліни формується як сума балів стартової оцінки rC та залікової оцінки rE :

$$rP = rC + rE$$

Стартова система оцінювання rC , максимальна величина якої становить 90 балів, складається з двох оцінок – оцінки за виконання лабораторного практикуму $rЛАБ$ (максимальна величина

якої становить 60 балів), оцінки за розрахунково-графічну роботу (максимальна величина якої становить 20 балів) та оцінки з модульної контрольної роботи rMKP (максимальна величина якої становить 10 балів):

$$rC = r\text{ЛАБ} + r\text{РГР} + r\text{МКР}$$

Система рейтингових (вагових) балів та критеріїв оцінювання:

Метод оцінювання	Кількість	Мінімальна оцінка в балах	Максимальна оцінка в балах
Лабораторні роботи	6	6	10
Розрахунково графічна робота	1	10	20
Модульна контрольна робота	1	4	10
Стартовий рейтинг		50	90
Залік	1		10
Підсумковий рейтинг		60	100

Модульна контрольна робота містить 40 питань тестового типу, правильна відповідь на кожні 4 оцінюються в один бал. За правильну відповідь на кожні 4 питання студент отримує 1 бал, неправильну – 0 балів.

Кожна з шести виконаних та захищених лабораторних робіт оцінюється за 10-бальною шкалою: максимальна оцінка – 10, мінімальна позитивна оцінка – 4. Таким чином, сумарна максимальна оцінка за виконання всього циклу лабораторних робіт складає 60 балів, а сумарна мінімальна позитивна оцінка – 50 балів.

Критерії оцінювання лабораторних робіт:

10 балів — завдання лабораторної роботи виконано на 100%, продемонстровано результати розрахунків, графічний матеріал та проведено оцінку якості роботи алгоритму, студень надає правильні відповіді на теоретичні запитання щодо залучених методів та алгоритмів.

9 балів — завдання лабораторної роботи виконано, виявлено певні недоліки.

8 балів — завдання лабораторної роботи виконано неповністю, студент орієнтується в теоретичних засадах залучених методів та алгоритмів.

7 балів — завдання лабораторної роботи виконано неповністю, присутні помилки та недоліки, неправильні відповіді на запитання.

6 балів — зроблено 60% завдання лабораторної роботи, неправильні відповіді на запитання або присутні грубі помилки, які безпосередньо впливають на отримані результати

Критерії оцінювання розрахунково-графічної роботи:

20-17 балів — завдання виконано на 100%, продемонстровано результати розрахунків, графічний матеріал, продемонстровано творчий підхід до побудови описових характеристик, очищення даних, проведено оцінку якості та порівняння результатів роботи залучених алгоритмів, студень надає правильні відповіді на теоретичні запитання.

16-14 балів — завдання виконано неповністю або виявлено певні недоліки, студент орієнтується в теоретичних засадах залучених методів та алгоритмів.

13-10 балів — зроблено 60% завдання, надано неправильні відповіді на запитання або присутні грубі помилки, які безпосередньо впливають на отримані результати

Стартова система оцінювання (максимум 60 балів) призначена для оцінювання якості і своєчасності виконання циклу лабораторних занять впродовж навчального семестру, а також якості і своєчасності виконання завдань розрахункової роботи (максимум 20 балів) та модульної контрольної роботи (максимум 10 балів).

Залікову систему оцінювання (максимум 10 балів) побудовано на основі повноти відповідей студента на запитання заліку, а також точності та аргументованості відповідей на додаткові запитання викладача. Під час оцінки знань студентів на заліку враховується розуміння питань та повнота відповідей на них, розуміння студентом зв'язку поставлених запитань з іншими темами дисципліни.

Умови допуску до екзамену: захищені усі лабораторні роботи та мінімальна позитивна стартова оцінка (сумарна оцінка з лабораторного практикуму та модульної контрольної роботи) 50 балів.

Таблиця відповідності рейтингових балів оцінкам за університетською шкалою:

<i>Кількість балів</i>	<i>Оцінка</i>
<i>100-95</i>	<i>Відмінно</i>
<i>94-85</i>	<i>Дуже добре</i>
<i>84-75</i>	<i>Добре</i>
<i>74-65</i>	<i>Задовільно</i>
<i>64-60</i>	<i>Достатньо</i>
<i>Менше 60</i>	<i>Незадовільно</i>
<i><36 або не виконано інші умови допуску до заліку</i>	<i>Не допущено</i>

9. Додаткова інформація з дисципліни (освітнього компонента)

- Опанування освітнього компонента не передбачає зарахування результатів, набутих у неформальній освіті;*
- У курсі використовується безкоштовне програмне забезпечення (Python, Keras, Pandas, Numpy), Tableau (студентам платформа надає можливість безоплатного користування протягом року).*

Робочу програму навчальної дисципліни (силабус):

Складено професором кафедри ШІ, д.ф.-м.н., професором, Купенко Ольгою Петрівною

Ухвалено кафедрою ШІ (протокол № 11 від 24.06.2025)

Погоджено Методичною комісією факультету (протокол № 7 від 25.06.2025)