

Силабус курсу:

ІНТЕЛЕКТУАЛЬНИЙ АНАЛІЗ ДАНИХ



Ступінь вищої освіти:	бакалавр
Спеціальність:	122 «Комп'ютерні науки»
Рік підготовки:	4
Семестр викладання:	7
Кількість кредитів ЄКТС:	2,5
Мова(-и) викладання:	українська
Вид семестрового контролю	іспит

Автор курсу та лектор:

к.т.н., доц., Шумова Лариса Олександрівна

вчений ступінь, вчене звання, прізвище, ім'я та по-батькові

доцент кафедри комп'ютерних наук та інженерії

посада

shumova@snu.edu.ua

електронна адреса

+38-050-225-75-60

телефон

Viber

месенджер

за розкладом

консультації

Цілі вивчення курсу:

Анотація навчального курсу

Наведені в курсі матеріали спрямовані на формування у здобувачів вищої освіти знань основних теоретичних положень та практичних методів інтелектуального аналізу даних і технологій Data mining, які необхідні для фахівців з комп'ютерних наук.

При вивченні дисципліни здобувачі вищої освіти отримують навички виявлення в "сирих" даних раніше невідомих нетривіальних практично корисних і доступних інтерпретації знань, необхідних для прийняття рішень в різних сферах людської діяльності. Необхідною умовою сучасного аналізу даних є ефективне використання комп'ютерних програм, від функціональної повноти і алгоритмічної продуманості яких залежить підсумкова інтерпретація результатів дослідження та надійність висновків. Такі можливості дає вільно розповсюджене програмне забезпечення для аналізу даних Weka (Waikato Environment for Knowledge Analysis).

Предметом вивчення дисципліни "Інтелектуальний аналіз даних" є технології інтелектуального аналізу даних.

Метою викладання навчальної дисципліни «Інтелектуальний аналіз даних» є теоретична та практична підготовка здобувачів

вищої освіти у напрямку розв'язання задач обробки великих масивів інформації: освоєння технологій оперативного та інтелектуального аналізу даних з візуалізацією результатів аналізу в процесі розв'язування прикладних задач у галузі комп'ютерних наук, пошуку у необроблених масивах даних раніше невідомих, практично корисних знань та закономірностей, необхідних для прийняття рішень.

Основними завданнями вивчення дисципліни «Інтелектуальний аналіз даних» є забезпечити розуміння і засвоєння здобувачами вищої освіти технологій інтелектуального аналізу даних, підготовка до їх вибору, реалізації і використання при вирішенні прикладних задач, ознайомлення зі станом та перспективами розвитку інтелектуальних методів обчислень як одного з напрямів штучного інтелекту.

Результати навчання:

Здобувач вищої освіти, що вивчив курс «Інтелектуальний аналіз даних», повинен знати: основні поняття та визначення інтелектуального аналізу даних; поширені методи та алгоритми вирішення основних завдань аналізу; критерії порівняння моделей і методів інтелектуального аналізу даних; вміти: обґрунтовувати вибір засобів і методів для розв'язку конкретних задач аналізу даних; створювати і застосовувати типові математичні моделі для дослідження явищ і процесів за результатами спостережень при розв'язку задач: класифікації, регресії, кластеризації, пошуку асоціативних правил; представляти результати аналізу в зручному для сприйняття вигляді, інтерпретувати їх у відповідності з поставленою задачею.

В результаті вивчення курсу здобувач вищої освіти повинен отримати навички ефективного використання сучасних технологій та програмних засобів інтелектуального аналізу даних та обробки інформації.

Знання і навички, отримані при вивченні дисципліни, будуть розвинуті на бакалаврському рівні при проходженні дисциплін «Методи штучного інтелекту»; на рівні підготовки магістрів спеціальності «Комп'ютерні науки» при проходженні дисциплін «Проектування інтелектуальних систем», «Теорія аналізу та обробки даних для IoT та IoE», «Системи підтримки прийняття рішень», будуть використовуватись у професійному контексті спеціалістів - аналітиків даних.

Передумови до початку вивчення:

Вивчення даного курсу базується на таких дисциплінах як «Дискретна математика», «Теорія імовірності, імовірнісні процеси та математична статистика», «Теорія прийняття рішень», «Об'єктно-орієнтоване програмування».

Мета курсу (набуті компетентності)

Мета курсу "Інтелектуальний аналіз даних" – підготовка бакалаврів за спеціальністю 122 «Комп'ютерні науки», відповідно до державних стандартів, встановлених освітньо-кваліфікаційною характеристикою та освітньо-професійною програмою, здатних здійснювати розробку, впровадження і супровід інтелектуальних систем аналізу й обробки даних організаційних, технічних, природничих і соціально-економічних систем.

В наслідок вивчення даного навчального курсу здобувач вищої освіти набуде наступних компетентностей.

Загальні компетентності:

- ЗК01. Здатність до абстрактного мислення, аналізу та синтезу;
- ЗК02. Здатність застосовувати знання у практичних ситуаціях;
- ЗК03. Знання та розуміння предметної області та розуміння професійної діяльності;
- ЗК06. Здатність вчитися й оволодівати сучасними знаннями;
- ЗК07. Здатність до пошуку, оброблення та аналізу інформації з різних джерел;
- ЗК08. Здатність генерувати нові ідеї (креативність);
- ЗК11. Здатність приймати обґрунтовані рішення.

Спеціальні (фахові, предметні) компетентності:

- СК01. Здатність до математичного формулювання та досліджування неперервних та дискретних математичних моделей, обґрунтовування вибору методів і підходів для розв'язування теоретичних і прикладних задач у галузі комп'ютерних наук, аналізу та інтерпретування;
- СК02. Здатність до виявлення статистичних закономірностей недетермінованих явищ, застосування методів обчислювального інтелекту, зокрема статистичної, нейромережевої та нечіткої обробки даних, методів машинного навчання та генетичного програмування тощо;
- СК04. Здатність використовувати сучасні методи математичного моделювання об'єктів, процесів і явищ, розробляти моделі й алгоритми чисельного розв'язування задач математичного моделювання, враховувати похибки наближеного чисельного розв'язування професійних задач;
- СК06. Здатність до системного мислення, застосування методології системного аналізу для дослідження складних проблем різної природи, методів формалізації та розв'язування системних задач, що мають суперечливі цілі, невизначеності та ризики;
- СК07. Здатність застосовувати теоретичні та практичні основи методології та технології моделювання для дослідження характеристик і поведінки складних об'єктів і систем, проводити обчислювальні експерименти з обробкою й аналізом результатів;
- СК11. Здатність до інтелектуального аналізу даних на основі методів обчислювального інтелекту включно з великими та погано структурованими даними, їхньої оперативної обробки та візуалізації результатів аналізу в процесі розв'язування прикладних задач.

Структура курсу

№	Тема	Години (Л/ПЗ)	Стислий зміст	Інструменти і завдання
1.	Введення до курсу.	2/0	Поняття інтелектуального аналізу даних (Data Mining). Поняття «дані», «інформація» і «знання». Класифікація задач Data Mining Практичне застосування Data Mining. Моделі Data Mining Методи Data Mining. Засоби інтелектуального аналізу даних у СППР. Неефективність використання OLTP систем для аналізу даних.	Участь в обговоренні Тести
2.	Сховища даних та OLAP – технології.	6/4	Технології вилучення, перетворення та завантаження даних. Концепція організації зберігання даних. Моделі даних інформаційного сховища. Концепція багатовимірного сховища даних. Типи OLAP. Реалізація сховищ і вітрин даних. ETL процес. Очистка даних.	Практична робота «Робота з OLAP-кубом у MS Excel»
3.	Інформаційна технологія великі дані (Big data).	4/0	Великі дані. Класифікація методів аналізу великих даних (Big Data Analytics).	Участь в обговоренні Тести
4.	Методи розвідувального аналізу даних.	6/4	Основні цілі розвідувального аналізу. Основні методи розвідувального аналізу. Кореляційний аналіз. Факторний аналіз. Дискримінантний аналіз. Багатовимірне шкалювання (Multidimensional scaling; MDS).	Практична робота «Механізми збору та підготовки даних для їх подальшого аналізу»
5.	Методи класифікації даних.	6/4	Постановка задачі. Методи побудови класифікаційних правил. 1R алгоритм. Метод Naive Bayes. Точність класифікації: оцінка рівня помилок. Оцінювання класифікаційних методів.	Практична робота «Дослідження методів та алгоритмів класифікації даних для вирішення практичних задач аналізу даних»
6.	Дерева рішень.	6/4	Методика «розділяй та володарюй» для побудови дерев рішень. Алгоритм максимального покриття для побудови дерев рішень.	Практична робота «Програмні засоби побудови дерева рішень»
7.	Методи кластеризації даних.	6/4	Задача кластеризації. Порівняння класифікації та кластеризації. Коротка характеристика підходів до кластеризації. Алгоритм kmeans, Ієрархічні алгоритми кластеризації. Неієрархічні алгоритми кластеризації. Оцінка якості кластеризації. Застосування кластерного аналізу.	Практична робота «Дослідження методів та алгоритмів кластеризації даних»
8.	Метод асоціативних правил (affinity analysis).	6/4	Постановка задачі. Секвенціальний аналіз. Представлення результатів. Алгоритми пошуку асоціативних правил Алгоритм Apriori та його різновиди.	Практична робота «Побудова асоціативних правил»

Рекомендована література

1. Акіменко В.В. Прикладні задачі інтелектуального аналізу даних (DATA MINING). К.: КНУ ім. Тараса Шевченка, 2018. 152 с. Щербаков Є. В., Щербакова М. Є. Мовні засоби системного програмування: навчальний посібник. – Луганськ: Вид-во СНУ ім. В. Даля, 2006. – 376 с.
2. Пасічник В.В., Шаховська Н.Б. Сховища даних: Навчальний посібник. – Львів: «Магнолія 2006», 2008. – 496 с.
3. Ситник В. Ф., Краснюк М.Т. Інтелектуальний аналіз даних (дейтамайнінг): Навч. посібник. – К: КНЕУ, 2007. – 376 с.
4. Черняк О.І. Інтелектуальний аналіз даних: Підручник / О.І. Черняк, П.В. Захарченко ; Київ. нац. ун-т ім. Т. Шевченка. — К. : Знання, 2014. — 599 с.
5. Inmon W.H. Building the Data Warehouse, 4th Edition. – Hoboken, NJ:Wiley, 2005. – 576 p.
6. Kretowski, M. Evolutionary Decision Trees in Large-Scale Data Mining /M. Kretowski , - Cham: Springer International Publishing, 2019, 180 p.
7. Ranga Suri N.N.R., Murty N., Athithan M.G. Outlier Detection: Techniques and Applications. A Data Mining Perspective / N.N.R. Ranga Suri, N. Murty, M.G. Athithan., - Cham: Springer International Publishing, 2019, 214 p.
8. Witten I. Data Mining: Practical Machine Learning Tools and Techniques / Ian Witten, Eibe Frank, Mark A. Hall. – Morgan Kaufmann, 2016. – 654p.

Методичне забезпечення

1. Конспект лекцій з дисципліни «Вступ до інтелектуального аналізу даних» (для студентів денної та заочної форми навчання спеціальності 122 "Комп'ютерні науки") Уклад: О.К. Лифар. – Сєвєродонецьк: Вид-во СНУ ім. В. Даля, 2019. – 213с.
2. Методичні вказівки до практичних занять з дисципліни «Інтелектуальний аналіз даних» для здобувачів вищої освіти денної і заочної форм навчання за спеціальністю «Комп'ютерні науки» / Укл. : Шумова Л.О. – Київ: СНУ ім. В. Даля, 2025. –40 с.
3. Методичні вказівки до виконання контрольної роботи з дисципліни «Інтелектуальний аналіз даних» для здобувачів вищої освіти заочної форми навчання за спеціальністю «Комп'ютерні науки» / Укл.: Шумова Л.О. – Київ: СНУ ім. В. Даля, 2025. - 27 с.

Оцінювання курсу

За повністю виконані завдання студент може отримати визначену кількість балів:

Інструменти і завдання	Кількість балів
Участь в обговоренні	10
Тести	30
Індивідуальні завдання	30
Іспит	30
Разом	100

Шкала оцінювання студентів

Сума балів за всі види навчальної діяльності	Оцінка ECTS	Оцінка за національною шкалою	
		для екзамену, курсового проекту (роботи), практики	для заліку
90 – 100	A	відмінно	зараховано
82-89	B	добре	
74-81	C		
64-73	D	задовільно	
60-63	E		
35-59	FX	незадовільно з можливістю повторного складання	не зараховано з можливістю повторного складання
0-34	F	незадовільно з обов’язковим повторним вивченням дисципліни	не зараховано з обов’язковим повторним вивченням дисципліни

Політика курсу

Плагіат та академічна доброчесність:

Під час виконання завдань студент має дотримуватись політики академічної доброчесності. Запозичення мають бути оформлені відповідними посиланнями. Списування є забороненим.

Дотримання академічної доброчесності студентами передбачає:

- самостійне виконання навчальних завдань, завдань поточного та підсумкового контролю результатів навчання (для осіб з особливими освітніми потребами ця вимога застосовується з урахуванням їхніх індивідуальних потреб і можливостей);
- посилання на джерела інформації у разі використання ідей, розробок, тверджень, відомостей;
- дотримання норм законодавства про авторське право і суміжні права;
- надання достовірної інформації про результати власної навчальної (наукової, творчої) діяльності, використанні методики досліджень і джерела інформації.

Порушенням академічної доброчесності вважається:

академічний плагіат - оприлюднення (частково або повністю) наукових (творчих) результатів, отриманих іншими особами, як результатів власного дослідження (творчості) та/або відтворення опублікованих текстів (оприлюднених творів мистецтва) інших авторів без зазначення авторства;

самоплагіат - оприлюднення (частково або повністю) власних раніше опублікованих наукових результатів як нових наукових результатів;

фабрикація - вигадкування даних чи фактів, що використовуються в освітньому процесі або наукових дослідженнях;

фальсифікація - свідомо зміна чи модифікація вже наявних даних, що стосуються освітнього процесу чи наукових досліджень;

списування - виконання письмових робіт із залученням зовнішніх джерел інформації, крім дозволених для використання, зокрема під час оцінювання результатів навчання.

За порушення академічної доброчесності здобувачі освіти можуть бути притягнені до такої академічної відповідальності:

- повторне проходження оцінювання (контрольна робота, іспит, залік тощо);
- повторне проходження відповідного освітнього компонента освітньої програми.

Завдання і заняття:

Всі завдання, передбачені програмою курсу мають бути виконані своєчасно і оцінені в спосіб, зазначений вище. Аудиторні заняття мають відвідуватись регулярно. Пропущені заняття (з будь-яких причин) мають бути відпрацьовані з отриманням відповідної оцінки не пізніше останнього тижня поточного семестру. За об'єктивних причин (наприклад, хвороба) навчання може відбуватись в он-лайн формі за погодженням із керівником курсу. В разі поважної причини (хвороба, академічна мобільність тощо) терміни можуть бути збільшені за письмовим дозволом декана.

Поведінка в аудиторії:

На заняття студенти вчасно приходять до аудиторії відповідно до діючого розкладу та обов'язково мають дотримуватися вимог техніки безпеки.

Під час занять студенти:

- не вживають їжу та жувальну гумку;
- не залишають аудиторію без дозволу викладача;
- не заважають викладачу проводити заняття.

Під час контролю знань студенти:

- є підготовленими відповідно до вимог даного курсу;
- розраховують тільки на власні знання (не шукають інші джерела інформації або «допомоги» інших осіб);
- не заважають іншим;
- виконують усі вимоги викладачів щодо контролю знань.