



Силабус курсу Інтелектуальний аналіз даних

Ступінь вищої освіти – бакалавр
Галузь знань – 01 Освіта/Педагогіка
Спеціальність – 015 Професійна освіта
Спеціалізація - 015.39 Цифрові технології
Освітньо-професійна програма – Професійна освіта
(Цифрові технології)

Рік навчання: 2

Семестр: 3

Кількість кредитів: 5

Керівник курсу

к. е. н., доцент **БАШУЦЬКА Оксана Степанівна**

Контактна інформація

o.bashutska@gmail.com, +380673728589

Опис дисципліни

Мета навчальної дисципліни «Інтелектуальний аналіз даних та знань» — формування у студентів здатності аналізувати великі масиви даних, використовуючи сучасні інструменти та методи інтелектуального аналізу, зокрема Excel, MATLAB та SPSS. Курс спрямований на розвиток як теоретичних знань, так і практичних навичок для виявлення прихованих закономірностей, прогнозування та прийняття обґрунтованих рішень у сфері цифрових технологій.

В курсі вивчаються основні методи інтелектуального аналізу даних, включаючи класифікацію, кластеризацію, регресійний аналіз, асоціативний аналіз, а також візуалізацію даних. Особлива увага приділяється методам попередньої обробки даних, аналізу часових рядів, а також зменшенню розмірності даних.

Головним завданням курсу «Інтелектуальний аналіз даних» є ознайомлення студентів із сучасними методами аналізу даних та інструментальними засобами їх реалізації. Студенти мають набути навичок побудови моделей для вирішення реальних задач у сфері цифрових технологій, включаючи застосування програм Excel, MATLAB та SPSS для автоматизації процесів аналізу даних.

Оволодіння курсом повинне забезпечити студентам навички практичного використання методів інтелектуального аналізу даних для вирішення задач у різних сферах діяльності, зокрема аналізу великих даних, прогнозування та прийняття рішень на основі отриманих моделей.

Структура курсу

Години (лек. / сем.)	Тема	Результати навчання	Завдання
2/2	Тема 1. Вступ до інтелектуального аналізу даних	Володіти поняттями інтелектуального аналізу даних (Data Mining). Знати основні етапи процесу видобування знань із даних (Knowledge Discovery in Databases, KDD).	Практичне завдання
2/2	Тема 2. Види даних та попередня обробка даних.	Розрізняти типи даних: категорійні, числові, часові ряди тощо. Виявляти проблеми з даними: відсутні значення, шум, дублікати. Знати основні методи нормалізації та очищення даних.	Практичне завдання
2/2	Тема 3. Статистичні основи інтелектуального аналізу	Володіти основними поняттями статистики: середнє, медіана, дисперсія, кореляція. Знати основи теорії ймовірності. Розуміти випадкові величини та їх розподіли.	Практичне завдання
2/2	Тема 4. Візуалізація даних	Знати методи візуалізації даних в Excel, MATLAB та SPSS. Уміти будувати гістограми, діаграми розсіювання, боксплоти.	Практичне завдання
2/2	Тема 5. Класифікація як метод інтелектуального аналізу даних.	Володіти основними поняттями класифікації. Вміти будувати моделі класифікації: дерева рішень, логістична регресія. Здійснювати оцінку якості моделей класифікації.	Практичне завдання
2/2	Тема 6. Кластеризація.	Володіти основними поняттями кластеризації. Знати методи кластеризації: метод k-середніх, ієрархічна кластеризація. Володіти метриками для оцінки якості кластеризації.	Практичне завдання
2/2	Тема 7. Регресійний аналіз.	Уміти будувати лінійну регресію та поліноміальну регресію. Володіти методом найменших квадратів.	Практичне завдання
2/2	Тема 8. Основи асоціативного аналізу.	Знати правила асоціації: поняття, приклади. Володіти алгоритмом Аргіогі. Використовувати асоціативний аналіз для рекомендаційних систем.	Практичне завдання
2/2	Тема 9. Аналіз часових рядів.	Знати основні поняття часових рядів: тренд, сезонність, шум. Уміти проводити прогнозування на основі часових рядів.	Практичне завдання
2/2	Тема 10. Основи машинного навчання	Володіти основними поняттями машинного навчання. Розрізняти навчання з учителем та без учителя. Знати основні алгоритми машинного навчання.	Практичне завдання
2/2	Тема 11. Зменшення розмірності даних	Розуміти поняттями високовимірних даних. Знати метод головних компонент (PCA). Володіти техніки зменшення розмірності в Excel та MATLAB.	Практичне завдання

2/2	Тема 12. Методи оптимізації в аналізі даних	Використовувати методи оптимізації для поліпшення моделей аналізу даних. Проводити оптимізацію в MATLAB.	Практичне завдання
2/2	Тема 13. Перенавчання та узагальнення.	Володіти поняттями перенавчання (overfitting) і недонавчання (underfitting). Використовувати методи боротьби з перенавчанням. Використання крос-валідації.	Практичне завдання
2/2	Тема 14. Етика та правові аспекти аналізу даних	Володіти поняттями конфіденційності даних. Розуміти відповідальність за прийняття рішень на основі аналізу даних. Знати етичні аспекти у використанні інтелектуальних систем.	Практичне завдання
2/2	Тема 15. Підсумки курсу та перспективи інтелектуального аналізу даних.	Розуміти потенціал та розвиток інтелектуального аналізу даних. Вміти наводити та розробляти приклади реальних кейсів.	Практичне завдання

Політика оцінювання

Підсумковий бал (за 100-бальною шкалою) з дисципліни визначається як середньозважена величина, в залежності від питомої ваги кожної складової залікового кредиту:

Оцінювання

Модуль 1		Модуль 2		Модуль 3	Модуль 4	Модуль 5
10%	10%	10%	10%	5%	15%	40%
Поточне оцінювання	Модульний контроль	Поточне оцінювання	Модульний контроль	Тренінги	Самостійна робота	Екзамен
Оцінка за поточне оцінювання визначається як середнє арифметичне з оцінок, отриманих під час занять (теми 1 - 7). Модульний контроль: Модульна робота по змістовому модулю1		Оцінка за поточне оцінювання визначається як середнє арифметичне з оцінок, отриманих під час занять (теми 7 - 15). Модульний контроль: Модульна робота по змістовому модулю2		Оцінка, отримана під час тренінгу	Оцінюється за результатами виконання завдання, яке вноситься на самостійну роботу.	Два теоретичних питання по 20 балів. Дві задачі по 30 балів за кожну.

Шкала оцінювання студентів:

ECTS	Бали	Зміст
A	90-100	відмінно
B	85-89	добре
C	75-84	добре
D	65-74	задовільно
E	60-64	достатньо
FX	35-59	незадовільно з можливістю повторного складання
F	1-34	незадовільно з обов'язковим повторним курсом

Рекомендовані джерела інформації

1. Болюбаш Н.М. Інтелектуальний аналіз даних: навч. посіб. / Н. М. Болюбаш. – Миколаїв : Вид-во ЧНУ ім. Петра Могили, 2023. – 320 с.
2. Акіменко В.В. Прикладні задачі інтелектуального аналізу даних (DATA MINING). К.: КНУ ім. Тараса Шевченка, 2018. 152 с.
3. Олійник А. О., Субботін С. О., Олійник О. О. Інтелектуальний аналіз даних: навчальний посібник, Запоріжжя: ЗНТУ, 2012, 278 с.
4. Актуальні проблеми Data Mining : навч. посіб. для студентів факультету комп'ютерних наук та кіберне-тики / О. О. Марченко, Т. В. Россада. – К. : КНУ ім. Т. Шевченка, 2017. – 150 с.
5. Гороховатський В. О., Творошенко І. С. Методи інтелектуального аналізу та оброблення даних : навч. посіб. / В. О. Гороховатський, І. С. Творошенко. – Харків : ХНУРЕ, 2021. – 92 с.
6. Черняк О.І., Захарченко П.В. Інтелектуальний аналіз даних: Підручник, К., 2014, 599 с.
7. Sivanandam S. N., Sumathi S., Deepa S. N. Introduction to Fuzzy Logic using MATLAB, 2007. 430 p.
8. Сявавко М., Рибицька О. Математичне моделювання за умов невизначеності. — Львів: Українські технології, 2000. — 320 с.
9. Paluszek M., Thomas S. Practical MATLAB Deep Learning. 1st Ed.: Apress, 2020. _252p.
10. Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani. An Introduction to Statistical Learning. Springer, 2023
11. Han, Jiawei. Data mining : concepts and techniques / Jiawei Han, Micheline Kamber, Jian Pei. – 3rd ed. 740 p.

Електронні ресурси

1. Kaggle (<https://www.kaggle.com>)
2. Coursera (<https://www.coursera.org>)
3. edX (<https://www.edx.org>)
4. UCI Machine Learning Repository (<https://archive.ics.uci.edu/ml/index.php>)
5. DataCamp (<https://www.datacamp.com>)
6. Towards Data Science (<https://towardsdatascience.com>)
7. Портал відкритих даних України. URL: <https://data.gov.ua/>