

Анотація навчальної дисципліни вільного вибору здобувача вищої освіти

Дисципліна:	«Інтелектуальний аналіз даних»
Викладач:	Сяський Володимир Андрійович, к.т.н., доцент
E-mail:	syasky_v@ukr.net
Кількість кредитів:	3
Мова викладання:	українська
Вид контролю:	залік
Місце у структурно-логічній схемі:	вивчається в 6 семестрі першого (бакалаврського) рівня вищої освіти за спеціальностями: 113 Прикладна математика 121 Іженерія програмного забезпечення

Мета та завдання дисципліни

Метою викладання дисципліни **Інтелектуальний аналіз даних (Data Mining)** є формування теоретичних та практичних знань про визначальні принципи, методи, технології Data Mining та вміння застосовувати їх при вирішенні прикладних інтелектуальних задач (класифікації, кластеризації, розпізнавання образів, прогнозування, прийняття рішень тощо).

Основними **завданнями** вивчення дисципліни **Інтелектуальний аналіз даних (Data Mining)** є:

- засвоєння основних теоретичних положень та стандартів інтелектуального аналізу даних;
- опрацювання передового досвіду використання технологій Data Mining у галузях корпоративного планування, менеджменту, маркетингу, прогнозування, моделювання процесів і систем тощо;
- набуття здатності та навичок програмної реалізації основних алгоритмів інтелектуального аналізу даних та конструювання компонентів систем Data Mining різного призначення і різної проблемної орієнтації із дотриманням принципів, методів та технологій інженерії програмного забезпечення на різних стадіях життєвого циклу інформаційної системи.

Зміст навчальної дисципліни

- **Основні поняття інтелектуального аналізу даних.** Цілі та завдання інтелектуального аналізу даних. Поняття даних. Набір даних та їх атрибутів. Вимірювання. Особливості обробки даних. Модель **MapReduce** – модель паралельної обробки великих обсягів даних. Задачі Data Mining. Етапи Data Mining. Методи Data Mining.

- **Стадії і методи інтелектуального аналізу даних.** Класифікація стадій Data Mining. Вільний пошук (Discovery). Прогностичне моделювання (Predictive Modeling). Аналіз винятків (Exception Analysis). Класифікація методів Data Mining. Класифікація методів за принципом роботи з навчальними даними. Класифікація методів за принципами навчання математичних моделей. Класифікація методів за задачами Data Mining. Властивості та особливості застосування методів Data Mining.

- **Задачі інтелектуального аналізу даних.** Завдання Data Mining. Класифікація задач Data Mining. Рівні аналізу в Data Mining. Дані, інформація й знання.

- **Задача класифікації. Основні поняття та визначення класифікації.** Постановка задачі класифікації. Рівні складності задачі класифікації. Точність класифікації: оцінка рівня помилок.

- **Методи класифікації.** Алгоритм **1-rule**. Наївний байєсівський класифікатор. Метод еталонів. Класифікація образів, класи яких частково перетинаються. Метод еталонів, що дробляться. Метод дерев прийняття рішень. Дерева прийняття рішень і булеві функції.

Алгоритм конструювання дерев прийняття рішень. Метод найближчого сусіда або системи міркувань «за аналогією».

- **Застосування штучних нейронних мереж для класифікації.** Перцептрони. Лінійні розділяючі відношення. Класифікація за бінарною ознакою. Метод опорних векторів. Прошарки перцептронів. Класифікація за K -арною ознакою.

- **Задача кластеризації. Кластерні технології.** Точки, простори та відстані. Стратегії кластеризації. Багатовимірні евклідові простори та «прокляттям вимірності». Постановка задачі кластеризації.

- **Ієрархічні алгоритми кластеризації.** Базовий алгоритм ієрархічної кластеризації. Алгоритми *Single-link* та *Complete-link*. Альтернативні правила для управління ієрархічною кластеризацією. Ієрархічна кластеризація у неевклідових просторах.

- **Неієрархічні алгоритми кластеризації.** Алгоритм *k-Means (Hard-c-means)*. Алгоритм Бредлі, Файяда та Рейна (БФР). Метод найближчого сусіда. Алгоритм *CURE*. Неієрархічна кластеризація у неевклідових просторах.

- **Застосування штучних нейронних мереж для кластеризації.** Мережі векторного квантування. Самоорганізуюча карта Кохонена.

- **Адаптивні методи кластеризації.** Нечіткі алгоритми кластеризації. Генетичні алгоритми кластеризації.

- **Асоціативні правила.** Задача пошуку асоціативних правил. Асоціативні правила. Узагальнені асоціативні правила. Чисельні асоціативні правила. Алгоритм *Apriori* та його різновиди.