

## АНАЛІЗ ТА ОБРОБКА ДАНИХ

**Обсяг:** 4 кредити, 120 годин (*лекційних занять - 20 годин, семінарських занять - 12 годин, аудиторних консультацій - 4 годин, самостійна робота 84 години*).

**Лектор:** професор Волков Юрій Іванович, доктор фізико-математичних наук, професор кафедри математики.

**Мета:** показати, що управління підприємством, банком, різними сферами бізнесу, у тому числі електронного, неможливе без процесів накопичення, аналізу, виявлення певних закономірностей і залежностей, прогнозування і тенденцій ризику. Метою курсу є розгляд найбільш перспективних напрямків аналізу даних: зберігання інформації, оперативний і інтелектуальний аналіз. Докладно розглянути методи та алгоритми інтелектуального аналізу. Розкрити технологію Data Mining (добування даних), тобто дослідити процес знаходження нових, дійсних і потенційно корисних знань в базах даних.

### **Завдання дисципліни:**

- Розглянути області застосування Data Mining з великими обсягами даних, а саме: у науці: - астрономії, біології, біоінформатиці, медицині, фізиці; у бізнесі – торгівлі, телекомунікаціях, банківській справі, промислового виробництва.
- Розглянути інтелектуальний аналіз даних, зокрема моделі та методи Data Mining та процес виявлення знань.
- Розглянути задачі класифікації і регресії, адаптивні методи кластеризації.

### **Вивчення спрямоване на формування компетенцій:**

#### **Компетенції соціально-особистісні:**

1. Здатність до обґрунтування доцільності застосування різних методів обробки та аналізу даних у різних сферах науки.

*Уміння:*

- проводити інтелектуальний аналіз в одній із сфер застосування Data Mining, яка відноситься до галузі власного наукового дослідження;
- обґрунтувати доцільність вибраних моделей та методів Data Mining у галузі власного наукового дослідження.

#### **Загальнонаукова компетенція:**

2. Здатність до інтелектуального аналізу даних.

*Уміння:*

- аналізувати сучасну проблематику збору та обробки інформації;
- вміння використовувати моделі та методи Data Mining;
- вміння використовувати методи побудови правил класифікації, дерев рішень, математичних функцій та вміння прогнозувати часові ряди

#### **Інструментальна компетенція:**

3. Комплексність у використанні аналітичних, інформатичних та комунікаційних технологій при обробці та аналізі даних.

*Уміння:*

- реалізовувати багатомірну класифікацію, регресію, прогнозування числових рядів засобами сучасних статистичних пакетів;
- добувати знання із Web – Web Mining;
- огляд бібліотеки Xelopes, яка містить важливі алгоритми Data Mining;
- налаштовувати прикладні програмні засоби для аналізу та обробки даних.

### **Загально-професійні компетенції:**

4. Здатність до вибору та застосування методів Data Mining для якісного інтелектуального аналізу даних.

*Уміння:*

- проводити аналіз умов застосовності методів Data Mining;
- правильно вибирати та використовувати засоби аналізу текстових документів;
- виконувати візуальний аналіз даних засобами візуалізації.

5. Здатність до практичної реалізації основних прийомів аналізу та обробки даних з використанням комп'ютерних технологій.

*Уміння:*

- реалізовувати основні методи Data Mining;
- використовувати комп'ютерні технології для обробки та візуалізації представлення даних.

**Предмет:** методи, алгоритмічні моделі і засоби їх реалізації для здійснення аналізу даних.

### **Зміст навчальної дисципліни**

1. Розглянути системи підтримки прийняття рішень та концепцію сховища даних.
2. Багатомірна модель даних. Визначення OLAP-системи.
3. Інтелектуальний аналіз даних: добування даних - Data Mining; задачі, практичне застосування, моделі та методи Data Mining.
4. Класифікація і регресія: представлення результатів, методи побудови правил класифікації, дерев рішень, математичних функцій.
5. Постановка задачі кластеризації. Базові алгоритми кластеризації та адаптивні методи кластеризації.
6. Візуальний аналіз даних – Visual Mining. Характеристики засобів візуалізації та методи візуалізації.
7. Аналіз текстової інформації – Text Mining. Витягування основних понять із тексту. Методи кластеризації текстових документів. Засоби аналізу текстової інформації.
8. Стандарти Data Mining. Бібліотека Xelopes. Приклади використання бібліотеки.
9. Розподілений аналіз даних. Використання мобільних агентів для аналізу даних.
10. Добування знань із Web – Web Mining.

**Місце у структурно-логічній схемі:** читається на другому курсі; для вивчення необхідні базові знання з математичної статистики, дискретної математики, теорії ймовірностей, баз даних, а також навички роботи з основними прикладними програмними середовищами.