

**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ ТОРГОВЕЛЬНО-ЕКОНОМІЧНИЙ  
УНІВЕРСИТЕТ**  
**СИСТЕМА УПРАВЛІННЯ ЯКІСТЮ**  
**Система забезпечення якості освітньої діяльності та якості вищої освіти**  
*сертифікована на відповідність ДСТУ ISO 9001:2015 / ISO 9001:2015*  
**Кафедра кібернетики та системного аналізу**

**ЗАТВЕРДЖЕНО**

вченою радою

(пост. 10 п. 10 від "21" червня 2018 р.)

Ректор



А.А. Мазаракі

**ТЕХНОЛОГІЇ АНАЛІЗУ ДАНИХ /  
DATA ANALYSIS TECHNOLOGIES**

**ПРОГРАМА ТА РОБОЧА ПРОГРАМА /  
CURRICULUM AND SYLLABUS**

<b>освітній ступінь</b>	<b>магістр / master</b>
<b>галузь знань</b>	<b>12 Інформаційні технології / Information Technology</b>
<b>спеціальність</b>	<b>121 Інженерія програмного забезпечення / Software Engineering</b>
<b>спеціалізація</b>	<b>Інженерія програмного забезпечення / Software Engineering</b>

**Київ 2018**

**Розповсюдження і тиражування без офіційного дозволу КНТЕУ  
заборонено**

Автор: А. А. Роскладка, доктор економічних наук, професор

Програму розглянуто і затверджено на засіданні кафедри кібернетики та системного аналізу 14.05.2018р., протокол № 23

Рецензенти: В. В. Кулаженко, кандидат економічних наук,  
І. В. Фабрика, кандидат економічних наук, заступник керівника  
Центру управління стратегічними змінами АТ «Ощадбанк».

**ТЕХНОЛОГІЇ АНАЛІЗУ ДАНИХ /  
DATA ANALYSIS TECHNOLOGIES**

**ПРОГРАМА ТА РОБОЧА ПРОГРАМА /  
CURRICULUM AND SYLLABUS**

<b>освітній ступінь</b>	<b>магістр / master</b>
<b>галузь знань</b>	<b>12 Інформаційні технології / Information Technology</b>
<b>спеціальність</b>	<b>121 Інженерія програмного забезпечення / Software Engineering</b>
<b>спеціалізація</b>	<b>Інженерія програмного забезпечення / Software Engineering</b>

Редактор  
Комп'ютерна верстка

Підп. до друку . Формат 60x84/16. Папір письм.  
Ризографія. Ум. друк. арк. \_\_\_\_\_. Ум. фарбо-відб. \_\_\_\_\_.  
Обл.-вид. арк. \_\_\_\_\_. Тираж \_\_\_\_\_ пр. Зам. .

---

Центр підготовки навчально-методичних видань КНТЕУ.  
02156, Київ-156, вул. Кіото, 19.

## ВСТУП

Програма та робочі програми дисципліни «Технології аналізу даних» призначена для студентів магістратури КНТЕУ денної форми навчання галузі знань галузі знань 12 «Інформаційні технології», спеціальності 121 «Інженерія програмного забезпечення», спеціалізації «Інженерія програмного забезпечення».

Програму підготовлено відповідно до освітньо-професійних програм підготовки бакалаврів та магістрів із зазначених галузей знань, спеціальностей та спеціалізацій.

Програма та робоча програма складається з таких розділів:

1. Мета, завдання та результати вивчення дисципліни (компетентності), її місце в освітньому процесі.
2. Зміст дисципліни.
3. Структура дисципліни та розподіл годин за темами (тематичний план).
4. Тематика та зміст лекційних, практичних (семінарських) занять і самостійної роботи студентів.
5. Список рекомендованих джерел.

### ***1. МЕТА, ЗАВДАННЯ ТА РЕЗУЛЬТАТИ ВИВЧЕННЯ ДИСЦИПЛІНИ (КОМПЕТЕНТНОСТІ), ЇЇ МІСЦЕ В ОСВІТНЬОМУ ПРОЦЕСІ***

*Метою* вивчення дисципліни «Технології аналізу даних» є надання фундаментальних теоретичних знань і набуття практичних навичок з питань формування, дослідження та всебічного аналізу інформації в різних галузях науки.

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

*Предметом* вивчення дисципліни є основні положення та методи аналізу даних та їх комп'ютерна реалізація за допомогою аналітичних платформ.

*Зміст компетенції* – здатність здійснювати дослідження, видобування та інтелектуальний аналіз даних з різноманітних інформаційних ресурсів на основі математичних моделей і методів науки про дані (*Data Science*) для процесів підтримки прийняття рішень; здатність використовувати програмне забезпечення для аналізу даних з метою проведення математичних та методологічних досліджень; здатність до комп'ютерної реалізації математичних моделей реальних систем і процесів; проектувати, застосовувати і супроводжувати програмні засоби моделювання, бізнес-аналітики, прийняття рішень, оптимізації обробки інформації, інтелектуального аналізу даних.

У результаті опанування програми дисципліни студент повинен *знати*:

- аналітичні засоби інформаційного аналізу;
- основні поняття та визначення інтелектуального аналізу даних;
- моделі та методи побудови моделей та аналізу залежностей у даних;
- критерії порівняння моделей і методів інтелектуального аналізу даних.
- методи реалізації *OLAP* та *Data Mining* технологій;
- типи сховищ даних;
- поняття трансформації, консолідації, очищення даних;

- принципи візуалізації даних;
- основні складові частини нейронної мережі;
- принципи застосування нейронно-мережових технологій;
- сучасні програмні засоби для проектування і розробки систем інтелектуального аналізу даних.

Після вивчення дисципліни студент повинен

вміти:

- обґрунтовувати й аналізувати вибір конкретного типу моделі та методу інтелектуального аналізу даних при вирішенні практичних задач;
- застосовувати алгоритми *Data Mining* при обробці даних;
- розробляти та будувати моделі сховищ даних;
- проводити аналіз даних для виявлення знань;
- використовувати *OLAP*-систему при обробці баз даних;
- використовувати сучасні програмні засоби для проектування та дослідження систем інтелектуального аналізу даних;
- створювати програми для інтелектуального аналізу даних при розв'язку конкретних практичних задач;
- аналізувати результати побудови та використання систем інтелектуального аналізу даних при вирішенні прикладних задач.

Дисципліна «Технології аналізу даних» базується на знаннях, отриманих студентами під час вивчення таких навчальних дисциплін:

- «Комп'ютерна дискретна математика»;
- «Об'єктно-орієнтовне програмування»;
- «Бази даних»;
- «Методи і засоби передачі даних»;
- «Експертні системи»;
- «Основи штучного інтелекту».
- 

## 2. ЗМІСТ ДИСЦИПЛІНИ

### Тема 1. Наука про дані – *Data Science*

Глосарій *Data Science*. Історія розвитку *Artificial Intelligence* і *Business Intelligence*. Зв'язок понять «дані», «інформація» та «знання». Характеристика фахівця з аналізу даних. *Soft skills* та *hard skills* аналітика даних. Приклади застосування *Data Science* у різних галузях людської діяльності.

Типи та види даних. Форми представлення даних. Вимірювання і шкали в аналізі даних. Реляційні та багатовимірні дані. Змінні, постійні та умовно-постійні дані. Довідкові, оперативні та архівні дані. Точкові дані та дані за період. Первинні і вторинні дані. Метадані.

Формати зберігання даних. Типи наборів даних. Транзакційні дані.

Етапи розв'язування задач аналізу даних: висунення гіпотез, збір і систематизація даних, побудова моделі, яка пояснює факти, тестування моделі та інтерпретація результатів, застосування отриманої моделі.

Технологія *Knowledge Discovery in Databases*. Формування вибірки даних. Очищення даних. Трансформація даних.

Технологія *Data Mining*. Задачі *Data Mining*: класифікація, регресія, кластеризація, асоціація, послідовність даних. Поняття про аналітичні системи. Актуальні бізнес-задачі.

### Список рекомендованих джерел

*Основний*: 2, 4.

*Додатковий*: 9, 10, 15.

*Інтернет-джерела*: 28, 30.

### Тема 2. Консолідація даних

Поняття консолідації. Основні критерії оптимальності консолідації даних. Джерела даних. Основні задачі консолідації даних. Схема процесу консолідації. Розробка стратегії консолідації. Поняття транзакції. Очищення даних. Збагачення даних. Деталізовані та агреговані дані.

Спеціалізовані сховища даних. *ETL*-процес. Сховища даних у системах підтримки прийняття рішень (СППР). Відмінності СППР та *OLTP*-систем. Основні вимоги до сховищ даних. Основні задачі розробки сховища даних. Архітектури сховищ даних. Нерегламентовані запити. Регулярні звіти. Реляційні сховища даних. Гібридні сховища даних. Вітрини даних. Віртуальні сховища даних. Нечіткі зрізи. Одноплатформенні та крос-платформенні сховища даних.

*OLAP*-системи. Схеми побудови сховищ даних: «зірка» та «сніжинка». Виміри, атрибути і факти. *OLAP*-куб. Структура багатовимірного куба. Операції *OLAP*-аналізу: переріз, транспонування, згортка (групування), деталізація (декомпозиція) *OLAP*-кубів.

### Список рекомендованих джерел

*Основний*: 1, 3.

*Додатковий*: 8.

*Інтернет-джерела*: 28, 30.

### Тема 3. Передобробка даних.

Трансформація даних. Основні методи трансформації даних. Перетворення часових рядів. Квантування даних. Сортування даних. Злиття даних. Об'єднання даних. Налаштування набору даних. Таблична підстановка значень. Обчислювальні значення. Нормалізація даних. Трансформація впорядкованих даних. Групування та розгрупування даних. Внутрішнє та зовнішнє з'єднання даних. Нормалізація та кодування даних. Методи нормалізації даних.

Система *Enterprise Data Quality Management (EDQM)*. Місце проведення оцінки якості даних Рівні якості даних: технічний, аналітичний та концептуальний рівень. Оцінка придатності даних до аналізу. Технології та методи оцінки якості даних. Профайлінг. Візуальна оцінка якості даних. Причини надходження в систему «брудних даних».

Передобробка даних. Типовий набір інструментів передобробки в аналітичній платформі. Очищення від шумів і згладжування рядів даних. Фільтрація даних. Відновлення пропущених значень. Редагування аномальних значень. Методи корегування аномальних значень. Проблема репрезентативності.

Обробка дублікатів і протиріч. Зниження вимірності вхідних даних. Скорочення кількості ознак. Скорочення кількості значень ознак і записів. Семплінг. Усунення незначущих факторів.

#### Список рекомендованих джерел

Основний: 1-3, 6.

Додатковий: 26, 27.

Інтернет-джерела: 28, 30.

#### Тема 4. Пошук асоціативних правил – *Rules Mining*

Афінітивний аналіз. Поняття типової транзакції. Предметний набір. Основні поняття *Rules Mining*. Асоціативні правила. Умова та наслідок асоціативного правила.

Підтримка та достовірність правил. Значущість асоціативних правил. Міри корисності правил. Ліфт, левередж та покращення асоціативних правил.

Алгоритм *a priori*. Пошук предметних наборів. Генерація асоціативних правил. Ієрархічні асоціативні правила. Методи пошуку ієрархічних асоціативних правил. Послідовні шаблони. Програмні засоби пошуку асоціативних правил. Практичний аспект застосування технології асоціативних правил. Секвенціальний аналіз.

#### Список рекомендованих джерел

Основний: 1, 5, 7.

Додатковий: 19, 20.

Інтернет-джерела: 28, 30.

#### Тема 5. Кластеризація даних.

Формальна постановка задачі кластеризації. Задачі кластеризації: вивчення даних, полегшення аналізу, стиснення даних, прогнозування, виявлення аномалій. Приклади кластеризації в різних областях знань. Представлення результатів кластеризації даних.

Базові алгоритми кластеризації. Алгоритм кластеризації *k-means*. Критерій збіжності алгоритму. Міри відстаней у кластеризації. Міри Евкліда і Манхеттена. Алгоритм *g-means*. Кластеризація за Гюстафсоном-Кесселем.

Програмні засоби кластеризації. Мережі Кохонена. Самоорганізуючі карти Кохонена. Методика побудови карти Кохонена. Вибір кількості нейронів карти. Алгоритм навчання мережі Кохонена.

Ієрархічні алгоритми кластерного аналізу. Міри схожості. Методи об'єднання та зв'язку.

Ітеративні алгоритми кластерного аналізу. Алгоритм PAM ( Partitioning Around Methods).

Адаптивні методи кластеризації. Проблеми алгоритмів кластеризації. Невизначеність у виборі критерія якості кластеризації. Проблеми машинних ресурсів. Задача вибору кількості кластерів.

#### **Список рекомендованих джерел**

*Основний:* 1- 3, 5, 7.

*Додатковий:* 20.

*Інтернет-джерела:* 28, 30.

#### **Тема 6. Класифікація та регресія даних.**

Огляд методів класифікації. Процес класифікації. Точність класифікації. Оцінювання класифікаційних методів.

Кореляційно-регресійний аналіз. Статистичні методи аналізу. Байєсівська класифікація. Лінійна регресія. Регресія з категоріальними вхідними змінними. Методи відбору змінних до регресійної моделі. Обмеження у застосуванні регресійних моделей. Використання фіктивних змінних.

Логістична регресія. Оцінки максимальної правдоподібності. Значущість вхідних змінних. Використання логістичної регресії для розв'язування задач класифікації. Тест Чоу. *ROC*-аналіз.

Інтерпретація моделі логістичної регресії. Інтерпретація моделі для дихотомічної змінної. Інтерпретація моделі для поліхотомічної змінної. Інтерпретація логістичної моделі для неперервної змінної. Множинна логістична регресія. Простий байєсівській класифікатор.

Методи прогнозування даних. Часовий ряд та його компоненти. Моделі прогнозування часових рядів. Прогнозування в торгівлі та логістиці.

#### **Список рекомендованих джерел**

*Основний:* 1, 3, 7.

*Додатковий:* 11, 12, 14.

*Інтернет-джерела:* 28, 30

#### **Тема 7. Аналіз даних на основі машинного навчання**

Дерева рішень. Структура дерева рішень. Процес конструювання дерев рішень. Вибір атрибуту розбиття у вузлі дерева. Принцип «розділяй і володарюй». Алгоритми побудови дерев рішень. Повне дерево рішень. Міри ефективності дерев рішень. Критерії вибору найкращих атрибутів розгалуження. Індекс Джині. Зменшення ентропії. Відношення приросту інформації. Алгоритми *ID3* та *C4.5*. Проблема перенавчання. Невідомі значення атрибутів. Алгоритм *CART*. Регресійне дерево рішень. Спрощення дерев рішень. Перенавчання і складність моделей. Критерії оптимізації дерев рішень. Відсікання гілок. Регуляризаційні мережі (Regularization Networks)

#### **Список рекомендованих джерел**

*Основний:* 1, 6.

*Додатковий:* 11, 12, 14.

*Інтернет-джерела:* 28, 30.

## **Тема 8. Нейронні мережі та генетичні алгоритми аналізу даних**

Задачі аналізу даних на основі штучних нейронних мереж. Поняття нейрона та перцептрона. Біологічний та штучний нейрони. Структура нервової системи. Перцептрон Розенблатта. Структура нейронної мережі. Чутливі нейрони (*S*-елементи). Асоціюючі нейрони (*A*-елементи). Реагуючі нейрони (*R*-елементи). Модель зорової системи.

Формальні нейрони штучних нейронних мереж. Бінарна та біполярна функції активації нейронів. Поріг сприйняття подразнення. Розв'язування задач розпізнавання образів на основі окремих нейронів. Правило Хебба. Модель Хебба. Алгоритм Хебба налаштування вагів нейронів. Нейронна мережа Хебба.

Міра Хеммінга. Модель Хеммінга. Функція активації нейронів у моделі Хеммінга. Нейронна мережа Хеммінга. Підмережа *Maxnet*. Навчання, самонавчання та адаптація нейронної мережі Хеммінга. Алгоритми навчання нейронних мереж.

Еволюційні технології та генетичні алгоритми. Концептуальні засади еволюційної теорії. Основні положення теорії генетичних алгоритмів.

Моделі генетичних алгоритмів. Програмне забезпечення та сфери застосування генетичних алгоритмів. Мурашині алгоритми та генетичне програмування.

Нечіткі методи інтелектуального аналізу даних. Концепція нечітких обчислень. Нечітка логіка в системах *Data Mining*. Програмне забезпечення нечітких методів. Сучасна практика застосування нечітких методів в системах *Data Mining*.

### **Список рекомендованих джерел**

*Основний:* 1-3, 5-7.

*Додатковий:* 9, 11, 12, 14-18.

*Інтернет-джерела:* 28, 30.

## **Тема 9. Технології інтелектуальної обробки даних**

Виконання візуального аналізу даних. Характеристики засобів візуалізації даних. Методи візуалізації. Методи геометричних перетворень. Відображення іконок. Методи, орієнтовані на пікселі. Ієрархічні образи.

Задача та етапи аналізу текстів *Text Mining*. Задачі аналізу текстової інформації. Загальний опис процесу отримання понять з тексту. Етапи *Text Mining*: пошук інформації, попередня обробка документів, витяг інформації, застосування методів *Text Mining*, інтерпретація результатів. Методи класифікації текстових документів.

Попередня обробка тексту. Видалення стоп-слів. Стеммінг. *N*-грами. Зведення реєстра. Класифікація текстів. Методи кластеризації текстових документів: ієрархічні, бінарні. Представлення текстових документів. Задача анотування текстів. Пошук асоціацій. Первинний витяг ключових понять. Навігація текстом. Засоби аналізу текстової інформації.

Ідея *Data Mining* у реальному часі. Адаптація системи до загальної концепції. Адаптивне добування даних. Статичний *Data Mining* і *Data Mining* в реальному



часі. Застосування *Data Mining* в реальному часі. Рекомендаційні машини. Класифікація рекомендаційних машин. Підхід на основі змісту. Спільне фільтрування. Аналіз ринкового кошика і секвенційного аналіз. Посилення навчання і агенти. Інструменти *Data Mining* у реальному часі. Інструмент *Amazon.com* – механізм рекомендацій. Інструмент *Prudsys* – рекомендаційна машина *Prudsys*.

Додаток з відкритим кодом - *SpamAssassin*. Складнощі аналізу даних з мережі Інтернет. Етапи *Web Mining*. Категорії *Web Mining*. Аналіз використання веб-ресурсів. Використання веб-структур та веб-контенту.

Методи *Web Mining* з точки зору вирішуваних завдань і реалізованих підходів. Пошук інформації. Аналіз структури сегмента мережі. Використання знань з веб-ресурсів. Персоналізація інформації. Пошук шаблонів в поведінці користувачів.

### **Список рекомендованих джерел**

*Основний:* 1-3, 5-7.

*Додатковий:* 9, 11, 12, 14-18.

*Інтернет-джерела:* 28, 30.

### **Тема 10. Інструментальні засоби аналізу даних**

Програмне забезпечення в області аналізу даних. Аналітичні платформи. Технології аналізу даних у продуктах *Microsoft Corporation*. Технологія моделювання даних у *Microsoft Power Pivot*. Інтерактивний інструмент *Microsoft Power View* для дослідження, графічного відображення та представлення даних. Надбудова *Microsoft Power Query* в задачах бізнес-аналітики.

Хмарні технології *Microsoft* для аналізу та візуалізації даних. Організація бізнес-аналітики рівня *Business Intelligence (BI)*. Платформа *Microsoft BI*. Завантаження даних *Power BI* з різних інформаційних джерел.

Аналітична платформа *Deductor Studio*. Базові навички роботи у системі *Deductor*. Візуалізатори в *Deductor Studio*. Сортування, заміна та фільтрація інформації в *Deductor Studio*. Консолідація даних в системі *Deductor*. Побудова аналітичної звітності.

Система бізнес-аналітики *Weka*. Колекція інструментів і алгоритмів *Weka*. Прогнозування в системі *Weka*. Використання алгоритмів машинного навчання в системі бізнес-аналітики *Weka*. Візуалізація результатів аналізу даних у вигляді графової моделі. Мови візуального моделювання.

### **Список рекомендованих джерел**

*Основний:* 1-3, 5-7.

*Додатковий:* 9, 11, 12, 14-18.

*Інтернет-джерела:* 28, 30.

### 3. СТРУКТУРА ДИСЦИПЛІНИ ТА РОЗПОДІЛ ГОДИН ЗА ТЕМАМИ (ТЕМАТИЧНИЙ ПЛАН)

#### Тематичний план для спеціалізацій «Цифрова економіка», «Інженерія програмного забезпечення»

Для вивчення дисципліни відводиться 225 год. / 7,5 кредитів ECTS (лекцій – 44 год., лабораторних занять – 44 год., самостійної роботи – 137 год.), підсумковий контроль – екзамен

Назва теми	Кількість годин				Форма контролю
	Усього годин / кредитів	з них:			
		лекції	лабораторні заняття / МК	самостійна робота студентів	
Тема 1. Наука про дані – <i>Data Science</i>	28	6	4	18	Т, ІДЗ, П
Тема 2. Консолідація даних	20	4	4	12	Т, ІДЗ, П
Тема 3. Передобробка даних	14	4	0	10	Т, ІДЗ, П
Тема 4. Пошук асоціативних правил – <i>Rules Mining</i>	16	2	4	10	Т, ІДЗ, П
Тема 5. Кластеризація даних	32	4	8	20	Т, ІДЗ, П, МК
Тема 6. Класифікація та регресія даних	12	4	0	8	Т, ІДЗ, П,
Тема 7. Аналіз даних на основі машинного навчання	36	4	12	20	Т, ІДЗ, Пр, П
Тема 8. Нейронні мережі та генетичні алгоритми аналізу даних	14	6	0	8	Т, ІДЗ, П
Тема 9. Технології інтелектуальної обробки даних	48	8	12	28	Т, ІДЗ, П
Тема 10. Інструментальні засоби аналізу даних	5	2	0	3	Т, ІДЗ, Пр, П, МК
Підсумковий контроль – екзамен					
<b>Разом</b>	<b>225/7,5</b>	<b>44</b>	<b>44</b>	<b>137</b>	<b>x</b>

Результати навчання	Навчальна діяльність <sup>1</sup>	Робочий час студента, год <sup>2</sup>	Оцінювання в балах
1	2	3	4
<p><b>Розуміння:</b> теоретичних засад науки про дані (<i>Data Science</i>).</p> <p><b>Знання:</b> основних задач та етапів проведення аналізу даних.</p> <p><b>Вміння:</b> працювати з інтерфейсом та основними вузлами аналітичної платформи <i>Deductor Studio</i>.</p>	<p><b>Тема 1. Наука про дані – <i>Data Science</i>.</b></p> <p><b>Лекція 1. Вступ у <i>Data Science</i>. Поняття про дані</b> <b>План лекції</b></p> <ol style="list-style-type: none"> <li>1. Поняття про дані.</li> <li>2. Типи та види даних. Форми представлення даних.</li> <li>3. Вимірювання і шкали в аналізі даних. Реляційні та багатовимірні дані. Метадані.</li> <li>4. Формати зберігання даних.</li> <li>5. Типи наборів даних. Транзакційні дані.</li> </ol> <p><b>Лекція 2. Основні задачі аналізу даних. Практичне застосування <i>Data Science</i></b> <b>План лекції</b></p> <ol style="list-style-type: none"> <li>1. Характеристика фахівця з аналізу даних. <i>Soft skills</i> та <i>hard skills</i> аналітика даних.</li> <li>2. Технологія <i>Knowledge Discovery in Databases</i>. Формування вибірки даних.</li> <li>3. Технологія <i>Data Mining</i>. Задачі <i>Data Mining</i>: класифікація, регресія, кластеризація, асоціація, послідовність даних.</li> <li>4. Поняття про аналітичні системи.</li> <li>5. Актуальні бізнес-задачі.</li> </ol> <p><b>Лекція 3. Етапи аналізу даних</b> <b>План лекції</b></p> <ol style="list-style-type: none"> <li>1. Висунення гіпотез.</li> <li>2. Збір і систематизація даних.</li> <li>3. Побудова моделі, яка пояснює факти.</li> <li>4. Тестування моделі та інтерпретація результатів.</li> <li>5. Застосування отриманої моделі.</li> </ol> <p><b>Лабораторне заняття 1</b> <b>Система <i>Deductor</i>. Базові навички роботи у системі <i>Deductor</i></b> <b>Мета:</b> ознайомитися з головним вікном <i>Deductor Studio</i>, навчитися виконувати маніпуляції з вкладками; ознайомитися зі</p>	<p>2</p> <p>2</p> <p>2</p> <p>2</p> <p>2</p>	<p>1</p>

<sup>1</sup>+20% інтерактивних методів навчання відісно курсивом

<sup>2</sup> А – спеціалізації «Цифрова економіка», «Інженерія програмного забезпечення»; В – спеціалізації «Комп'ютерні науки», «Логістика та управління ланцюгами постачання», «Економіка і організація туризму».

1	2	3	4
	<p>сценарієм – базовим поняттям в <i>Deductor Studio</i>, навчитися оперувати основними вузлами сценарію.</p> <p><b>Завдання:</b> Створити новий проект в системі <i>Deductor</i>. Вивчити призначення вкладок «Сценарии» та «Подключения». Здійснити імпорт даних з текстових файлів, створених у програмах Блокнот та MS Excel (csv-файл). Вивчити основні операції вузла «Настройка набора данных». Використовуючи Мастер обработки, створити декілька вузлів (Корреляционный анализ, Фильтрация, Сортировка, Замена данных, Группировка тощо). Представити інформацію у вигляді таблиць, діаграм, гістограм. Виконати експорт окремих результатів у текстові файли.</p> <p style="text-align: center;"><b>Лабораторне заняття 2</b> <b>Базові вузли та візуалізатори в Deductor Studio</b></p> <p><b>Мета:</b> ознайомитися з базовими візуалізаторами в <i>Deductor Studio</i>, навчитися застосовувати їх під час створення проекту; ознайомитися з вузлами Сортвання, Заміна і Фільтрація в <i>Deductor Studio</i>, навчитися застосовувати їх під час створення проекту.</p> <p><b>Завдання:</b> Налаштувати візуалізатори до вузлів імпорту: Таблица, Статистика. Створити новий проект. Імпортувати в нього текстовий файл <i>CreditSample.txt</i>, що йде в поставці <i>Deductor</i> у каталозі <i>Samples</i>. Відсортувати цей набір даних за визначеними полями в порядку зростання. Зробіть вказану заміну даних полів за методом підстановки. Провести фільтрацію набору даних за визначеними критеріями.</p> <p style="text-align: center;"><b>Список рекомендованих джерел:</b> Основний: 2, 4. Додатковий: 9, 10, 15. Інтернет-джерела: 28, 30.</p> <p><b>Самостійна робота:</b> підготовка до практичних занять з теми, ознайомлення з основною та додатковою літературою, інтернет-ресурсами, виконання індивідуальних завдань з використанням комп'ютера.</p>	<p style="text-align: center;">2</p> <p style="text-align: center;">18</p>	<p style="text-align: center;">1</p>
<p><b>Розуміння:</b> сутності консолідації даних.</p>	<p style="text-align: center;"><b>Тема 2. Консолідація даних</b> <b>Лекція 4. Задачі консолідації даних</b></p>	<p style="text-align: center;">2</p>	

1	2	3	4
<p><b>Знання:</b> основних задач консолідації даних, типів та структур сховищ даних і OLAP-систем.</p> <p><b>Вміння:</b> створювати реляційне сховище даних та використовувати його у процесі консолідації даних</p>	<p align="center"><b>План лекції</b></p> <ol style="list-style-type: none"> <li>1. Поняття консолідації.</li> <li>2. Основні критерії оптимальності консолідації даних. Джерела даних.</li> <li>3. Основні задачі консолідації даних. Схема процесу консолідації.</li> <li>4. Розробка стратегії консолідації. Поняття транзакції.</li> <li>5. Очищення даних. Збагачення даних. Деталізовані та агреговані дані.</li> </ol> <p align="center"><b>Лекція 5. Сховища даних та OLAP-системи</b></p> <p align="center"><b>План лекції</b></p> <ol style="list-style-type: none"> <li>1. Спеціалізовані сховища даних.</li> <li>2. ETL-процес. Сховища даних у системах підтримки прийняття рішень (СППР).</li> <li>3. Архітектури сховищ даних.</li> <li>4. Реляційні сховища даних. Гібридні сховища даних. Вітрини даних. Віртуальні сховища даних.</li> <li>5. OLAP-системи. Схеми побудови сховищ даних.</li> <li>6. Виміри, атрибути і факти. OLAP-куб.</li> </ol> <p align="center"><b>Лабораторні заняття 3-4</b></p> <p align="center"><b>Deductor Studio. Консолідація даних на прикладі аптечної мережі</b></p> <p><b>Мета:</b> оволодіти навичками створення реляційного сховища та консолідації даних аптечної мережі в <i>Deductor Studio</i>.</p> <p><b>Завдання:</b></p> <p><i>Постановка задачі. Компанія, яка володіє невеликою аптечною мережею, займається роздрібним продажем лікувальних препаратів. Керівництво компанії прийняло рішення про впровадження системи аналітичної OLAP-звітності, в якій цікавить інформація про динаміку продажу, завантаженість торговельних точок, найбільш популярних товарів у роздрібних розрізах. Оскільки існуюча звітна система потерпає навантаження (компанія постійно розширює свою мережу), було вирішено створити одне консолідоване джерело сховище даних, яке служило б базою для OLAP-звітності.</i></p> <p><i>Попередньо програмісти створили процедуру вивантаження даних із звітної системи в структуровані текстові файли (у якості проби сформована «пачка» даних за</i></p>	<p align="center">2</p> <p align="center">4</p>	<p align="center">2</p> <p align="center">2</p>

1	2	3	4
	<p>декілька місяців).</p> <p><b>Вихідні дані</b> представлені у текстових файлах: <i>Товари.txt</i>, <i>Групи товарів.txt</i>, <i>Відділи.txt</i>, <i>Продаж.txt</i>.</p> <p><b>Необхідно:</b></p> <ol style="list-style-type: none"> <li>1. Спроекувати структуру реляційного сховища даних (СД).</li> <li>2. Наповнити СД первинною інформацією.</li> <li>3. Розробити процедури поповнення СД і контроль несуперечливості даних, які в ньому містяться.</li> <li>4. Запропонувати набір OLAP-звітів.</li> <li>5. Сформувавши власну базу даних для подібної задачі і виконати пункти 1-4.</li> <li>6. База даних повинна складатися з 4 файлів: <ul style="list-style-type: none"> <li><i>Відділи.txt</i> – повинен містити <u>коди</u> і <u>назви</u> трьох відділів економічної системи.</li> <li><i>Групи товарів.txt</i> повинен містити <u>коди</u> і <u>назви</u> п'яти груп товарів.</li> <li><i>Товари.txt</i> повинен містити <u>коди</u>, <u>назви</u> і <u>групи товарів</u> (по 5-10 найменувань товару в кожній групі).</li> <li><i>Продаж.txt</i> – повинен містити не менше 100 записів з такими полями: <ul style="list-style-type: none"> <li>• дата покупки;</li> <li>• код товару;</li> <li>• код відділу;</li> <li>• час придбання товару (в годинах);</li> <li>• кількість одиниць товару;</li> <li>• сума покупки.</li> </ul> </li> </ul> </li> </ol> <p><b>Список рекомендованих джерел:</b>  <b>Основний:</b> 1, 3.  <b>Додатковий:</b> 8.  <b>Інтернет-джерела:</b> 28, 30.</p> <p><b>Самостійна робота:</b> підготовка до практичних занять з теми, ознайомлення з основною та додатковою літературою, інтернет-ресурсами, виконання індивідуальних завдань з використанням комп'ютера.</p>	12	
<p><b>Розуміння:</b> теоретичних основ підготовки даних для аналізу</p> <p><b>Знання:</b> методів та алгоритмів</p>	<p><b>Тема 3. Передобробка даних</b></p> <p><b>Лекція 6. Трансформація даних</b></p> <p><b>План лекції</b></p> <ol style="list-style-type: none"> <li>1. Основні методи трансформації даних.</li> <li>2. Перетворення часових рядів. Квантування даних.</li> <li>3. Сортування даних. Злиття даних.</li> </ol>	2	

1	2	3	4
<p>передоброби даних</p> <p><b>Вміння:</b> проводити квантування, сортування, групування, очищення та фільтрацію даних.</p>	<p>Об'єднання даних.</p> <p>4. Нормалізація даних.</p> <p>5. Трансформація впорядкованих даних.</p> <p>6. Групування та розгрупування даних.</p> <p><b>Лекція 7. Методи та алгоритми передоброби даних</b></p> <p><b>План лекції</b></p> <p>1. Типовий набір інструментів передоброби в аналітичній платформі.</p> <p>2. Очищення від шумів і згладжування рядів даних.</p> <p>3. Фільтрація даних. Відновлення пропущених значень.</p> <p>4. Редагування аномальних значень. Методи корегування аномальних значень.</p> <p>5. Проблема репрезентативності. Обробка дублікатів і протиріч.</p> <p>6. Семплінг. Усунення незначущих факторів.</p> <p><b>Список рекомендованих джерел:</b> <i>Основний:</i> 1-3, 6. <i>Додатковий:</i> 26, 27. <i>Інтернет-джерела:</i> 28, 30.</p> <p><b>Самостійна робота:</b> підготовка до практичних занять з теми, ознайомлення з основною та додатковою літературою, інтернет-ресурсами, виконання індивідуальних завдань з використанням комп'ютера.</p>	<p>2</p> <p>10</p>	
<p><b>Розуміння:</b> сутності задачі асоціації даних.</p> <p><b>Знання:</b> алгоритмів побудови асоціативних правил.</p> <p><b>Вміння:</b> використовувати асоціативні правила у прикладних економічних задачах.</p>	<p><b>Тема 4. Пошук асоціативних правил – <i>Rules Mining</i></b></p> <p><b>Лекція 8. Асоціативні правила</b></p> <p><b>План лекції</b></p> <p>1. Афінітивний аналіз. Поняття типової транзакції. Предметний набір.</p> <p>2. Основні поняття <i>Rules Mining</i>. Асоціативні правила. Умова та наслідок асоціативного правила.</p> <p>3. Підтримка та достовірність правил. Значущість асоціативних правил. Міри корисності правил. Ліфт, левередж та покращення асоціативних правил.</p> <p>4. Алгоритм <i>a priori</i>. Генерація асоціативних правил.</p> <p>5. Ієрархічні асоціативні правила. Методи пошуку ієрархічних асоціативних правил.</p> <p>6. Секвенціальний аналіз.</p>	<p>2</p> <p>4</p>	<p>2</p>

1	2	3	4
	<p style="text-align: center;"><b>Лабораторні заняття 5-6</b></p> <p style="text-align: center;"><b>Асоціативні правила в стимулюванні роздрібних продаж</b></p> <p><b>Мета:</b> Навчитися генерувати асоціативні правила та на основі алгоритму <i>a priori</i> виділяти правила з максимальною достовірністю з метою розробки більш досконалої маркетингової стратегії для оптимізації закупок товарів. Оволодіти навичками роботи з обробником <i>Асоціативні правила в Deductor Studio</i> для реалізації алгоритму <i>a priori</i>.</p> <p><b>Завдання:</b></p> <p style="text-align: center;"><i>Постановка задачі.</i> Роздрібна мережа по продажу товарів побутової хімії поставила задачу аналізу купівельних кошиків для оптимізації розміщення товарів на вітринах і проведення крос-продажів. Відділ маркетингу надав 5000 чеків, у яких відображені покупки, зроблені клієнтами магазинів. Необхідно:</p> <ul style="list-style-type: none"> <li>- передбачити, які товари покупки можуть вибрати в залежності від того, що вже є в їхніх кошиках;</li> <li>- виявити найбільш популярні товарні набори, які складаються з більше ніж одного предмету;</li> <li>- запропонувати рекламні акції типу: «Кожному, хто купив А і В – товар С в подарунок».</li> </ul> <p style="text-align: center;"><i>Вихідні дані.</i> Представлені у файлі Чеки.txt двома полями Номер транзакції і Товар. Оскільки номенклатура товарів побутової хімії надзвичайно різноманітна, то товари представлені в узагальненій формі без торгових марок: порошки, мючі засоби і т.д.(всього 37 найменувань).</p> <ol style="list-style-type: none"> <li>1. Використовуючи алгоритм <i>a priori</i>, побудувати асоціативні правила і зробити їх інтерпретацію.</li> <li>2. Сформувати власну базу даних для подібної задачі і виконати п.1. Серед початкових даних повинно бути: <ul style="list-style-type: none"> <li>○ не менше 20 різних видів товару;</li> <li>○ кожен чек повинен містити від 2 до 4 різних видів товару;</li> <li>○ кількість чеків не менше 50.</li> </ul> </li> </ol> <p style="text-align: center;"><b>Список рекомендованих джерел:</b>  Основний: 1, 5, 7.  Додатковий: 19, 20.  Інтернет-джерела: 28, 30.</p>	10	



1	2	3	4
	<p><b>Самостійна робота:</b> підготовка до практичних занять з теми, ознайомлення з основною та додатковою літературою, інтернет-ресурсами, виконання індивідуальних завдань з використанням комп'ютера.</p>		
<p><b>Розуміння:</b> теоретичних основ кластерного аналізу.</p> <p><b>Знання:</b> алгоритмів кластеризації даних.</p> <p><b>Вміння:</b> застосовувати на практиці алгоритми кластеризації та сегментації даних.</p>	<p><b>Тема 5. Кластеризація даних</b> <b>Лекція 9. Алгоритми кластеризації</b> <b>План лекції</b></p> <ol style="list-style-type: none"> <li>1. Формальна постановка задачі кластеризації.</li> <li>2. Задачі кластеризації: вивчення даних, полегшення аналізу, стиснення даних, прогнозування, виявлення аномалій.</li> <li>3. Приклади кластеризації в різних областях знань. Представлення результатів кластеризації даних.</li> <li>4. Базові алгоритми кластеризації.</li> </ol> <p><b>Лекція 10. Кластерний аналіз. Ієрархічні та ітеративні методи</b> <b>План лекції</b></p> <ol style="list-style-type: none"> <li>1. Ієрархічні алгоритми кластерного аналізу. Міри схожості. Методи об'єднання та зв'язку.</li> <li>2. Ітеративні алгоритми кластерного аналізу.</li> <li>3. Алгоритм кластеризації k-means. Алгоритм g-means.</li> <li>4. Критерій збіжності алгоритму. Міри відстаней у кластеризації. Міри Евкліда і Манхеттена.</li> <li>5. Алгоритм PAM (Partitioning Around Methods).</li> <li>6. Адаптивні методи кластеризації.</li> <li>7. Кластеризація за Гюстафсоном-Кесселем.</li> <li>8. Програмні засоби кластеризації. Мережі Кохонена. Самоорганізуючі карти Кохонена.</li> </ol> <p><b>Лабораторні заняття №7-8</b> <b>Алгоритм k-means. Загальні принципи кластеризації.</b></p> <p><b>Мета:</b> набуття практичних навичок роботи з алгоритмом кластеризації <i>k-means</i>.</p> <p><b>Завдання:</b> <i>Дано 10 об'єктів (A, B, C, D, E, F, G, H, I, K), які представлено координатами точок на площині.</i></p> <ol style="list-style-type: none"> <li>1. Використовуючи алгоритм <i>k-means</i>, провести кластерний аналіз даних,</li> </ol>	<p>2</p> <p>2</p> <p>4</p> <p>2</p>	<p></p> <p>2</p> <p>1</p>

1	2	3	4
	<p>розбивши початкову множину точок на 2 кластери. За початкові центри кластерів прийняти точки E і J.</p> <p>2. Використовуючи алгоритм <i>k-means</i>, провести кластерний аналіз даних, розбивши початкову множину точок на 3 кластери. За початкові центри кластерів прийняти точки A, F і J.</p> <p>3. Порівняти отримані результати і обрати оптимальну кількість кластерів.</p> <p>4. На кожному етапі кластерного аналізу провести візуалізацію результатів розбиття на кластери.</p> <p>5. Для визначення належності об'єкта до певного кластеру використати евклідову відстань та відстань Манхеттена.</p> <p style="text-align: center;"><b>Лабораторне заняття 9</b> <b>Сегментація клієнтів з використанням карт Кохонена</b></p> <p><b>Мета:</b> набуття навичок роботи з обробником і візуалізатором <i>Карта Кохонена</i> в процесі розв'язування бізнес-задачі сегментації в <i>Deductor Studio</i>.</p> <p><b>Завдання:</b></p> <p><i>Постановка задачі.</i> Керівництво філії регіональної телекомунікаційної компанії, яка надає послуги мобільного зв'язку, поставило задачу сегментації абонентської бази. Її цілі полягають в наступному:</p> <ul style="list-style-type: none"> <li>- побудова профілів абонентів шляхом виявлення їх схожої поведінки в плані частоти, тривалості та часу дзвінків, а також щомісячних витрат;</li> <li>- оцінка найбільш і найменш прибуткових сегментів.</li> </ul> <p>Ця інформація може в подальшому використовуватися для:</p> <ul style="list-style-type: none"> <li>- розробки маркетингових акцій, направлених на визначення групи клієнтів;</li> <li>- розробки нових тарифних планів;</li> <li>- оптимізації витрат на адресну SMS-розсилку про нові послуги і тарифи;</li> <li>- попередження відтоку клієнтів у інші компанії.</li> </ul> <p><b>Вихідні дані.</b> Представляються у вигляді файлу <i>mobile.txt</i>.</p> <p>Розв'язати бізнес-задачу сегментації абонентів за допомогою підходу, який</p>	<p style="text-align: center;">2</p> <p style="text-align: center;">20</p>	<p style="text-align: center;">40</p>



1	2	3	4
	<p>3. Моделі прогнозування часових рядів. 4. Прогнозування в торгівлі та логістиці.</p> <p><b>Список рекомендованих джерел:</b> <i>Основний:</i> 1, 3, 7. <i>Додатковий:</i> 11, 12, 14. <i>Інтернет-джерела:</i> 28, 30</p> <p><b>Самостійна робота:</b> підготовка до практичних занять з теми, ознайомлення з основною та додатковою літературою, інтернет-ресурсами, виконання індивідуальних завдань з використанням комп'ютера.</p>		
<p><b>Розуміння:</b> сутності машинного навчання як інструменту аналізу даних.</p> <p><b>Знання:</b> методів й алгоритмів побудови дерев рішень та машинної класифікації.</p> <p><b>Вміння:</b> працювати в аналітичній платформі <i>Microsoft Power BI</i>.</p>	<p><b>Тема 7. Аналіз даних на основі машинного навчання.</b></p> <p><b>Лекція 13. Древа рішень</b> <b>План лекції</b></p> <ol style="list-style-type: none"> <li>1. Структура дерева рішень. Процес конструювання дерев рішень.</li> <li>2. Вибір атрибуту розбиття у вузлі дерева.</li> <li>3. Принцип «розділяй і володарюй». Алгоритми побудови дерев рішень.</li> <li>4. Міри ефективності дерев рішень. Критерії вибору найкращих атрибутів розгалуження.</li> <li>5. Регресійне дерево рішень. Спрощення дерев рішень.</li> <li>6. Критерії оптимізації дерев рішень.</li> </ol> <p><b>Лекція 14. Алгоритми машинної класифікації</b> <b>План лекції</b></p> <ol style="list-style-type: none"> <li>1. Індекс Джині.</li> <li>2. Зменшення ентропії. Відношення приросту інформації.</li> <li>3. Алгоритми ID3 та C4.5. Проблема перенавчання.</li> <li>4. Алгоритм CART.</li> <li>5. Регуляризаційні мережі (Regularization Networks)</li> </ol> <p><b>Лабораторне заняття 11</b> <b><i>Power BI</i>: завантаження, реєстрація та основи роботи. <i>Power BI Desktop, Power BI Service, Power BI Mobile</i></b></p> <p><b>Мета:</b> ознайомлення зі структурою аналітичної платформи <i>Microsoft Power BI</i>; інсталяція, реєстрація та початок роботи із системою.</p>	<p>2</p> <p>2</p> <p>2</p> <p>2</p>	<p>1</p> <p>1</p>

1	2	3	4
	<p><b>Завдання:</b> Завантажити систему Microsoft Power BI з офіційного сайту розробника microsoft.com. Провести реєстрацію програми з використанням корпоративного акаунту Office365. Ознайомитися із функціональними можливостями та основними принципами роботи в середовищах Power BI Desktop, Power BI Service, Power BI Mobile. Встановити та налаштувати DAX Studio.</p> <p style="text-align: center;"><b>Лабораторне заняття 12</b></p> <p style="text-align: center;"><b>Power BI: створення даних у програмі та проста візуалізація</b></p> <p><b>Мета:</b> створення таблиць даних безпосередньо в аналітичній платформі Power BI Desktop та їх тривіальний аналіз.</p> <p><b>Завдання:</b> Знайти актуальні дані щодо зміни курсу основних валют до гривні за останні 10 років. Внести дані до таблиці редактора Power BI. Встановити необхідні формати даних. Виконати просту візуалізацію даних з використанням графіків і діаграм аналітичної платформи. Здійснити фільтрацію даних за різні періоди і прогнозування курсу основних валют на найближчі три місяці.</p> <p style="text-align: center;"><b>Лабораторне заняття 13</b></p> <p style="text-align: center;"><b>Імпорт даних та базові операції в Power BI</b></p> <p><b>Мета:</b> вивчення засобів імпорту даних з різних локальних джерел та веб-ресурсів.</p> <p><b>Завдання:</b> Провести аналіз системи управління доступом до джерел даних. Здійснити налаштування конфіденційності джерела даних. Завантажити дані з декількох джерел, встановити зв'язки між таблицями та здійснити проектування структури моделі. Використати різні джерела імпорту даних: текстовий файл, MS Excel, MS Access, Azure, SQL, веб-ресурси.</p> <p style="text-align: center;"><b>Лабораторне заняття 14</b></p> <p style="text-align: center;"><b>Очищення даних. Трансформація даних. Параметри запиту</b></p> <p><b>Мета:</b> формування та налаштування запитів в системі Power BI. Підготовка даних до моделювання.</p> <p><b>Завдання:</b> Побудувати початковий запит до моделей DirectQuery. Переіменувати стовпці наборів даних. Усунути дублікати даних. Здійснити інтеграцію декількох</p>	<p style="text-align: center;">2</p> <p style="text-align: center;">2</p> <p style="text-align: center;">2</p> <p style="text-align: center;">2</p>	<p style="text-align: center;">1</p> <p style="text-align: center;">1</p> <p style="text-align: center;">1</p> <p style="text-align: center;">1</p>

1	2	3	4
	<p><i>захитів. Встановити потрібний формат початкових даних.</i></p> <p style="text-align: center;"><b>Лабораторне заняття 15</b>  <b>Побудова моделей даних у Power BI.</b>  <b>Аналіз діяльності торговельної мережі</b>  <b>Мета:</b> побудова розширеного звіту аналізу діяльності торговельної мережі.  <b>Вихідні дані.</b> <i>Результати економічної діяльності торговельної мережі.</i>  <b>Завдання:</b> <i>Завантажити початкові дані для моделі, використовуючи локальні дані щодо діяльності торговельної мережі та різні інтернет-джерела. Побудувати структуру моделі діяльності торговельної мережі, встановити необхідні зв'язки між різними джерелами. Здійснити візуалізацію звіту з використанням діаграм, графіків, гістограм та інтерактивних карт. Провести аналіз діяльності торговельної мережі за різними регіонами виготовлення та споживання товарів, датою реалізації продукції, постачальниками сировини. Визначити тенденції та побудувати прогноз розвитку діяльності торговельної мережі.</i></p> <p style="text-align: center;"><b>Лабораторне заняття 16</b>  <b>Побудова моделей даних у Power BI.</b>  <b>Аналіз діяльності туристичної агенції</b>  <b>Мета:</b> побудова розширеного звіту аналізу діяльності туристичної агенції.  <b>Вихідні дані.</b> <i>Результати діяльності туристичної агенції.</i>  <b>Завдання:</b> <i>Завантажити початкові дані для моделі, використовуючи локальні дані щодо діяльності туристичної агенції та різні інтернет-джерела. Побудувати структуру моделі діяльності туристичної агенції, встановити необхідні зв'язки між різними джерелами. Здійснити візуалізацію звіту з використанням діаграм, графіків, гістограм та інтерактивних карт. Провести аналіз діяльності туристичної агенції за різними регіонами замовлень, туристичними DESTINATIONAMI та напрямками туристичних маршрутів, датою реалізації туристичних замовлень. Визначити тенденції та побудувати прогноз розвитку діяльності туристичної агенції.</i></p> <p style="text-align: center;"><b>Список рекомендованих джерел:</b></p>	20	

1	2	3	4
	<p><i>Основний:</i> 1, 6.  <i>Додатковий:</i> 11, 12, 14.  <i>Інтернет-джерела:</i> 28, 30.</p> <p><b>Самостійна робота:</b> підготовка до практичних занять з теми, ознайомлення з основною та додатковою літературою, інтернет-ресурсами, виконання індивідуальних завдань з використанням комп'ютера.</p>		
<p><b>Розуміння:</b>  сутності штучного інтелекту, структури та призначення нейронних мереж</p> <p><b>Знання:</b>  нейронно-мережових та генетичних алгоритмів аналізу даних</p> <p><b>Вміння:</b>  використовувати алгоритми Хебба і Хеммінга для побудови й аналізу нейронних мереж</p>	<p><b>Тема 8. Нейронні мережі та генетичні алгоритми аналізу даних</b></p> <p><b>Лекція 15. Штучні нейрони та нейронні мережі</b>  <b>План лекції</b></p> <ol style="list-style-type: none"> <li>1. Задачі аналізу даних на основі штучних нейронних мереж.</li> <li>2. Поняття нейрона та перцептрона. Біологічний та штучний нейрони. Структура нервової системи.</li> <li>3. Перцептрон Розенблатта.</li> <li>4. Структура нейронної мережі.</li> <li>5. Модель зорової системи.</li> </ol> <p><b>Лекція 16. Нейронні мережі Хебба та Хеммінга</b>  <b>План лекції</b></p> <ol style="list-style-type: none"> <li>1. Формальні нейрони штучних нейронних мереж.</li> <li>2. Бінарна та біполярна функції активації нейронів.</li> <li>3. Розв'язування задач розпізнавання образів на основі окремих нейронів.</li> <li>4. Правило Хебба. Модель Хебба. Алгоритм Хебба налаштування вагів нейронів. Нейронна мережа Хебба.</li> <li>5. Міра Хеммінга. Модель Хеммінга. Функція активації нейронів у моделі Хеммінга. Нейронна мережа Хеммінга. Підмережа Maxnet.</li> </ol> <p><b>Лекція 17. Нейронно-мережові та генетичні алгоритми</b>  <b>План лекції</b></p> <ol style="list-style-type: none"> <li>1. Навчання, самонавчання та адаптація нейронної мережі Хеммінга. Алгоритми навчання нейронних мереж.</li> <li>2. Моделі генетичних алгоритмів.</li> <li>3. Мурашині алгоритми та генетичне програмування.</li> </ol>	<p>2</p> <p>2</p> <p>2</p> <p>8</p>	

1	2	3	4
	<p>4. Нечіткі методи інтелектуального аналізу даних.</p> <p>5. Сучасна практика застосування нечітких методів в системах <i>Data Mining</i>.</p> <p><b>Список рекомендованих джерел:</b>  <i>Основний:</i> 6, 7.  <i>Додатковий:</i> 11, 12, 14, 21-23, 25.  <i>Інтернет-джерела:</i> 28, 30.</p> <p><b>Самостійна робота:</b> підготовка до практичних занять з теми, ознайомлення з основною та додатковою літературою, інтернет-ресурсами, виконання індивідуальних завдань з використанням комп'ютера.</p>		
<p><b>Розуміння:</b> призначення інтелектуальних методів аналізу даних.</p> <p><b>Знання:</b> методів та алгоритмів візуального аналізу даних, аналізу текстових даних, даних, що отримано у реальному часі, даних мережі Інтернет.</p> <p><b>Вміння:</b> створювати звіти в аналітичній платформі <i>Microsoft Power BI</i> з використанням сучасних методів візуалізації даних.</p>	<p><b>Тема 9. Технології інтелектуальної обробки даних</b></p> <p><b>Лекція 18. Візуальний аналіз даних - <i>Visual Mining</i></b>  <b>План лекції</b></p> <ol style="list-style-type: none"> <li>1. Виконання візуального аналізу даних.</li> <li>2. Характеристики засобів візуалізації даних. Методи візуалізації.</li> <li>3. Методи геометричних перетворень.</li> <li>4. Методи, орієнтовані на пікселі. Ієрархічні образи.</li> </ol> <p><b>Лекція 19. Аналіз текстових даних – <i>Text Mining</i></b>  <b>План лекції</b></p> <ol style="list-style-type: none"> <li>1. Задача та етапи аналізу текстів <i>Text Mining</i>.</li> <li>2. Методи класифікації текстових документів.</li> <li>3. Попередня обробка тексту. Видалення стоп-слів. Стеммінг. <i>N</i>-грами. Зведення реєстра.</li> <li>4. Класифікація текстів. Методи кластеризації текстових документів: ієрархічні, бінарні.</li> <li>5. Задача анотування текстів. Пошук асоціацій. Первинний витяг ключових понять. Навігація текстом. Засоби аналізу текстової інформації.</li> </ol> <p><b>Лекція 20. Аналіз даних у реальному часі - <i>Real Time Data Mining</i></b>  <b>План лекції</b></p> <ol style="list-style-type: none"> <li>1. Ідея <i>Data Mining</i> у реальному часі.</li> </ol>	<p>2</p> <p>2</p> <p>2</p> <p>2</p>	



1	2	3	4
	<p>Адаптація системи до загальної концепції. Адаптивне добування даних.</p> <p>2. Статичний <i>Data Mining</i> і <i>Data Mining</i> в реальному часі</p> <p>3. Рекомендаційні машини. Класифікація рекомендаційних машин.</p> <p>4. Аналіз ринкового кошика і секвенційного аналіз.</p> <p>5. Інструменти <i>Data Mining</i> у реальному часі.</p> <p><b>Лекція 21. Аналіз даних мережі Інтернет – <i>Web Mining</i></b> <b>План лекції</b></p> <p>1. Складнощі аналізу даних з мережі Інтернет.</p> <p>2. Етапи <i>Web Mining</i>. Категорії <i>Web Mining</i>. Аналіз використання веб-ресурсів.</p> <p>3. Методи <i>Web Mining</i> з точки зору вирішуваних завдань і реалізованих підходів.</p> <p>4. Пошук інформації. Аналіз структури сегмента мережі. Використання знань з веб-ресурсів.</p> <p>5. Персоналізація інформації. Пошук шаблонів в поведінці користувачів.</p> <p><b>Лабораторні заняття 17-18</b> <b>Візуалізація даних у <i>Power BI</i>. Побудова звіту за результатами вступної кампанії</b> <b>Мета:</b> Набуття базових навичок візуалізації даних в аналітичній платформі <i>Power BI</i>. Побудова розширеного звіту аналізу результатів вступної кампанії до закладів вищої освіти України. <b>Вихідні дані.</b> <i>Офіційні інтернет-джерела, що містять результати вступної кампанії до закладів вищої освіти України у поточному навчальному році (vstup.info, abit-poisk.org.ua, vstup.edebo.gov.ua).</i> <b>Завдання:</b> <i>Завантажити початкові дані для моделі, використовуючи офіційні інтернет-джерела щодо результатів вступної кампанії до закладів вищої освіти України. Побудувати структуру моделі вступної кампанії, встановити необхідні зв'язки між різними джерелами. Здійснити візуалізацію звіту з використанням діаграм, графіків, гістограм та інтерактивних карт. Провести аналіз вступної кампанії за різними закладами вищої освіти, регіонами розташування ЗВО, обсягом ліцензії за</i></p>	<p>4</p> <p>4</p> <p>2</p>	<p>2</p> <p>2</p> <p>1</p>

1	2	3	4
	<p><i>певною освітньою програмою, кількістю місць підготовки фахівців за державним замовленням, середнім балом абітурієнтів, прохідним балом на місця державного замовлення. Визначити тенденції та побудувати прогноз розвитку діяльності визначених закладів вищої освіти.</i></p> <p style="text-align: center;"><b>Лабораторні заняття 19-20</b></p> <p style="text-align: center;"><b>Використання інтерактивних карт.</b></p> <p style="text-align: center;"><b>Побудова звіту про вибори в Україні</b></p> <p><b>Мета:</b> побудова розширеного звіту аналізу результатів виборів до Верховної Ради України.</p> <p><b>Вихідні дані.</b> <i>Офіційні інтернет-джерела, що містять результати виборчої кампанії до Верховної Ради України, зокрема <a href="http://svk.gov.ua">svk.gov.ua</a>.</i></p> <p><b>Завдання:</b> <i>Завантажити початкові дані для моделі, використовуючи офіційні інтернет-джерела щодо результатів виборчої кампанії до Верховної ради України. Побудувати структуру моделі виборчої кампанії, встановити необхідні зв'язки між різними джерелами. Здійснити візуалізацію звіту з використанням діаграм, графіків, гістограм та інтерактивних карт. Провести аналіз виборчої кампанії за різними партіями, мажоритарними округами, регіонами голосування. Визначити лідерів та аутсайдерів голосування. Партії, що подолали 5% прохідний бар'єр. Визначити тенденції у порівнянні з минулими виборами та побудувати прогноз результатів наступних виборів.</i></p> <p style="text-align: center;"><b>Лабораторне заняття 21</b></p> <p style="text-align: center;"><b>Web-аналітика в Power BI. Аналіз даних з web-сторінок. Чемпіонати світу і Європи з футболу</b></p> <p><b>Мета:</b> Набуття базових навичок аналізу даних з web-сторінок в аналітичній платформі Power BI.</p> <p><b>Вихідні дані.</b> <i>Дані про переможців Чемпіонату Європи з футболу УЄФА можна знайти в таблиці результатів на сторінці Вікіпедії за адресою <a href="http://en.wikipedia.org/wiki/UEFA_European_Football_Championship">http://en.wikipedia.org/wiki/UEFA_European_Football_Championship</a></i></p> <p><b>Завдання:</b> <i>Здійснити підключення до джерела даних в Інтернеті і перехід за доступними в ньому</i></p>	<p style="text-align: center;">2</p> <p style="text-align: center;">28</p>	<p style="text-align: center;">40</p>

1	2	3	4
	<p>таблицями. Виконати збір і перетворення даних в редакторі Power Query. Створити та налаштувати візуалізації карти і кругової діаграми переможців футбольних чемпіонатів відповідно до кількості перемог. Побудувати рейтингову систему найкращих бомбардирів чемпіонатів.</p> <p><b>Лабораторне заняття 22</b>  <b>Модульний контроль №2.</b>  <b>«Інтелектуальний аналіз даних»</b></p> <p><b>Список рекомендованих джерел:</b>  <i>Основний:</i> 1-3, 5-7.  <i>Додатковий:</i> 9, 11, 12, 14-18.  <i>Інтернет-джерела:</i> 28, 30.</p> <p><b>Самостійна робота:</b> підготовка до практичних занять з теми, ознайомлення з основною та додатковою літературою, інтернет-ресурсами, виконання індивідуальних завдань з використанням комп'ютера.</p>		
<p><b>Розуміння:</b>  принципів функціонування програмного забезпечення інтелектуального аналізу даних.</p> <p><b>Знання:</b>  сучасних аналітичних платформ</p> <p><b>Вміння:</b>  використовувати програмні продукти Microsoft Power Pivot, Microsoft Power View, Microsoft Power Query, Weka для розв'язування основних задач аналізу даних.</p>	<p><b>Тема 10. Інструментальні засоби аналізу даних</b></p> <p><b>Лекція 21. Аналітичні платформи</b>  <b>План лекції</b></p> <ol style="list-style-type: none"> <li>1. Програмне забезпечення в області аналізу даних. Аналітичні платформи.</li> <li>2. Технології аналізу даних у продуктах Microsoft Corporation. Microsoft Power Pivot. Microsoft Power View. Microsoft Power Query.</li> <li>3. Хмарні технології Microsoft для аналізу та візуалізації даних.</li> <li>4. Microsoft Power BI.</li> <li>5. Аналітична платформа Deductor Studio.</li> <li>6. Система бізнес-аналітики Weka.</li> </ol> <p><b>Список рекомендованих джерел:</b>  <i>Основний:</i> 3.  <i>Додатковий:</i> 9, 10, 13-19.  <i>Інтернет-джерела:</i> 28-30</p> <p><b>Самостійна робота:</b> підготовка до практичних занять з теми, ознайомлення з основною та додатковою літературою, інтернет-ресурсами, виконання індивідуальних завдань з використанням комп'ютера.</p>	<p style="text-align: center;"><b>2</b></p> <p style="text-align: center;"><b>3</b></p>	
<b>Підсумковий контроль – екзамен</b>			
<b>Разом</b>	<b>225</b>	<b>100</b>	

## 5. СПИСОК РЕКОМЕНДОВАНИХ ДЖЕРЕЛ

### *Основний*

1. Cuesta H., Kumar S. Practical Data Analysis. Birmingham : Packt Publishing Ltd, 2016. 316 p.
  2. Data Science & Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data / EMC Education Services. Indianapolis : John Wiley & Sons, Inc, 2015. 432 p.
  3. Microsoft Power BI Cookbook: Creating Business Intelligence Solutions of Analytical Data Models, Reports, and Dashboards. Birmingham : Packt Publishing Ltd, 2017. 802 p.
  4. Roskladka A., Ivanova O., Kulazhenko V. Data Scientist: a glance into the future. *Зовнішня торгівля: економіка, фінанси, право*. 2019. № 3. С. 109-120
  5. Гнатієнко Г. М., Снитюк В. Є. *Експертні технології прийняття рішень*. Київ : Маклаут, 2008. 444 с.
  6. Матвійчук А. В. Штучний інтелект в економіці: нейронні мережі, нечітка логіка: монографія. Київ : КНЕУ, 2011. 439 с.
  7. Олійник А. О., Субботін С. О., Олійник О. О. Інтелектуальний аналіз даних: навч. посібн. Запоріжжя : ЗНТУ, 2012. 278 с.
- \* Курсивом виділені джерела, що є у бібліотеці КНТЕУ

### *Додатковий*

8. Adamson C. Mastering Data Warehouse Aggregates: Solutions for Star Schema Performance. Wiley Publishing Inc., 2006. 318 p.
9. Albright S. C., Winston W., Zappe C. Data Analysis and Decision Making. Boston : Cengage Learning, 2016. 948 p.
10. Cao L., Yu P. S., Zhang C., Zhang H. Data Mining for Business Applications. Springer Science; Business Media, 2008. 402 p.
11. Coodley M. O. Introduction to Microsoft Power BI: bring your data to life! CreateSpace Independent Publishing Platform, 2016. 128 p.
12. Etaati L. Advance Analytics with Power BI and R. Auckland : Radacad Systems Limited, 2017. 179 p.
13. Fabrice G., Hamilton N. J. Quality Measures in Data Mining. Berlin; Heidelberg: Springer-Verlag, 2007. 361 p.
14. Ferrari A., Russo M. Introducing Microsoft Power BI. Redmond : Microsoft Press, 2016. 407 p.
15. Han J., Kamber M. Data Mining: Concepts and Techniques. Morgan Kaufmann Publishers, 2006. 800 p.
16. Linoff G. S. Data Analysis Using SQL and Excel. Indianapolis: Wiley, 2015. 792 p.
17. Linoff G. S., Berry M. J. A. Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management. Indianapolis: Wiley, 2011. 888 p.
18. Milton M. Head First Data Analysis: A learner's guide to big numbers, statistics, and good decisions. Sebastopol: O'Reilly Media, 2009. 435 p.
19. RapidMiner: Data Mining Use Cases and Business Analytics Applications / Edited by Markus Hofmann & Ralf Klinkenberg. Minneapolis : CRC Press, 2004. 518 p.
20. Upton G. Categorical data analysis by example. New Jersey : John Wiley & Sons Inc, 2017. 198 p.
21. Witten I. H., Eibe F., Hall M. A. Data mining: practical machine learning tools and techniques. Morgan Kaufmann Publishers, 2011. 630 p.
22. Лысенко Ю.Г., Н. Н. Иванов Н. Н., Минц А. Ю. Нейронные сети и генетические алгоритмы : учебн. пособ. Донецк : ООО «Юго-Восток, Лтд», 2003. 265 с.
23. Олійник А. О., Субботін С. О., Олійник О. О. Неітеративні, еволюційні та мультиагентні методи синтезу нечіткологічних і нейромережних моделей : монографія. Запоріжжя : ЗНТУ, 2009. 375 с.
24. Рідкокаша А.А., Голдер К. К. Основи систем штучного інтелекту : навч. посібн. Черкаси : ВІДЛУННЯ-ПЛЮС, 2002. 240 с.

25. Руденко О. Г., Бодянський Є. В. Штучні нейронні мережі. Харків : Компанія СМІТ, 2006. 404 с.
26. Ситник В. Ф., Краснюк М. Т. Інтелектуальний аналіз даних (дейтамайнінг) : навч. посібн. Київ : КНЕУ, 2007. 376 с.
27. Субботін С. О. Подання й обробка знань у системах штучного інтелекту та підтримки прийняття рішень : навч. посібн. Запоріжжя : ЗНТУ, 2008. 341 с.

*Internet-ресурси*

28. Microsoft Power BI Guided Learning URL: <https://docs.microsoft.com/uk-ua/power-bi/guided-learning> (дата звернення 25.08.2019).
29. Weka 3: Data Mining Software in Java. URL: <http://www.cs.waikato.ac.nz/ml/weka> (дата звернення 25.08.2019).
30. Руководство аналитика Deductor 5.3 URL: <https://basegroup.ru/deductor/manual/guide-analyst-530> (дата звернення 25.08.2019).