

**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 цифрової економіки та системного аналізу

ЗАТВЕРДЖЕНО

вченою радою
(пост. п. 6 від «27» 02 2019 р.)

Ректор



А. А. Мазаракі

**ТЕХНОЛОГІЇ АНАЛІЗУ ДАНИХ /
DATA ANALYSIS TECHNOLOGIES**

**РОБОЧА ПРОГРАМА
COURSE OUTLINE**

| | |
|-------------------------|--|
| освітній ступінь | PhD |
| галузь знань | 12 Інформаційні технології / Information Technology |
| спеціальність | 122 Комп'ютерні науки / Computer Sciences |
| спеціалізація | Комп'ютерні науки / Computer Sciences |

Київ 2020

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

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

Робочу програму розглянуто і затверджено на засіданні кафедри кібернетики та системного аналізу 14.02.2020 р., протокол № 13

Рецензенти: В. В. Кулаженко, кандидат економічних наук,
М. Г. Шарафутдінов, директор з розвитку компанії «Center Research & Development», бізнес-аналітик

ТЕХНОЛОГІЇ АНАЛІЗУ ДАНИХ/ DATA ANALYSIS TECHNOLOGIES

РОБОЧА ПРОГРАМА COURSE OUTLINE

| | |
|-------------------------|--|
| освітній ступінь | PhD |
| галузь знань | 12 Інформаційні технології / Information Technology |
| спеціальність | 122 Комп'ютерні науки / Computer Sciences |
| спеціалізація | Комп'ютерні науки / Computer Sciences |

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

Для вивчення дисципліни відводиться 90 год. / 3 кредити ЄКТС (лекцій – 16 год., лабораторних занять – 12 год., самостійної роботи – 62 год.), підсумковий контроль – залік

| Назва теми | Кількість годин | | | | Форма контролю |
|--|-------------------------|-----------|--------------------------|------------------------------|----------------|
| | Усього годин / кредитів | з них: | | | |
| | | лекцій | лабораторні заняття / МК | самостійна робота аспірантів | |
| Тема 1. Наука про дані – <i>Data Science</i> | 4 | 2 | 0 | 2 | Т, ІДЗ, П, Пр |
| Тема 2. Консолідація та передобробка даних | 16 | 4 | 2 | 10 | Т, ІДЗ, П, Пр |
| Тема 3. Технології інтелектуальної обробки даних | 50 | 8 | 6 | 36 | Т, ІДЗ, П, Пр |
| Тема 4. Інструментальні засоби аналізу даних | 20 | 2 | 4 | 14 | Т, ІДЗ, Пр, П |
| Підсумковий контроль – залік | | | | | |
| Разом | 90/3 | 16 | 12 | 62 | х |

Примітка: Т – тестування; ІДЗ – виконання індивідуальних домашніх завдань; П – перевірка індивідуальних завдань; Пр. – презентація індивідуального завдання.

2. ТЕМАТИКА ТА ЗМІСТ ЛЕКЦІЙНИХ, ПРАКТИЧНИХ (СЕМІНАРСЬКИХ) ЗАНЯТЬ І САМОСТІЙНОЇ РОБОТИ ЗДОБУВАЧІВ

| Результати навчання | Навчальна діяльність ¹ | Робочий час здобувача, год |
|--|---|----------------------------|
| 1 | 2 | 3 |
| <p>Знати: основні задачі та етапи проведення аналізу даних.</p> <p>Вміти: працювати з інтерфейсом та основними вузлами аналітичної платформи <i>Deductor Studio</i>.</p> | <p>Тема 1. Наука про дані – <i>Data Science</i>.</p> <p style="text-align: center;">Лекція 1. Вступ у <i>Data Science</i>. Поняття про дані</p> <p style="text-align: center;">План лекції</p> <ol style="list-style-type: none"> 1. Поняття про дані. Типи та види даних. Форми представлення даних. 2. Вимірювання і шкали в аналізі даних. Реляційні та багатовимірні дані. Метадані. 3. Технологія <i>Knowledge Discovery in Databases</i>. Формування вибірки даних. 4. Технологія <i>Data Mining</i>. 5. Збір і систематизація даних. <p style="text-align: center;">Список рекомендованих джерел:</p> <p><i>Основний:</i> 2, 4. <i>Додатковий:</i> 13, 14, 19. <i>Інтернет-джерела:</i> 25.</p> | 2 |
| | <p>Самостійна робота: опрацювання матеріалу лекцій, підготовка до лабораторних занять з теми, ознайомлення з основною та додатковою літературою, інтернет-ресурсами, виконання індивідуальних завдань з використанням комп'ютера.</p> <p><i>Індивідуальна робота:</i> За темою дисертації підготувати дані для аналітичного дослідження. За технологією <i>Knowledge Discovery in Databases</i> відокремити та описати етапи:</p> <ul style="list-style-type: none"> - вибору даних для аналізу; - очищення даних; - трансформації даних; - підготовки даних до моделювання засобами <i>Data Mining</i>. | 2 |
| <p>Знати: основні задачі консолідації даних; типи та структури сховищ даних і <i>OLAP</i>-системи; методи та алгоритми передобробки даних</p> | <p>Тема 2. Консолідація та передобробка даних</p> <p style="text-align: center;">Лекція 2. Задачі консолідації даних</p> <p style="text-align: center;">План лекції</p> <ol style="list-style-type: none"> 1. Поняття консолідації. Джерела даних. 2. Основні задачі консолідації даних. Схема процесу консолідації. 3. Очищення даних. Збагачення даних. Спеціалізовані сховища даних. 4. ETL-процес. 5. Сховища даних у системах підтримки прийняття рішень (СППР). 6. <i>OLAP</i>-системи. | 2 |

¹+20% інтерактивних методів навчання виділено курсивом

| | | |
|--|--|----|
| <p>Вміти: здійснювати імпорт у реляційне сховище даних та використовувати його у процесі консолідації даних; проводити квантування, сортування, групування, очищення та фільтрацію даних.</p> | <p>Лекція 3. Методи та алгоритми передобробки даних План лекції</p> <ol style="list-style-type: none"> 1. Типовий набір інструментів передобробки в аналітичній платформі. 2. Очищення від шумів і згладжування рядів даних. 3. Фільтрація даних. Відновлення пропущених значень. 4. Редагування аномальних значень. Обробка дублікатів і протиріч. 5. Семплінг. Усунення незначущих факторів. 6. Перетворення часових рядів. Квантування даних. 7. Сортування даних. Злиття даних. Об'єднання даних. 8. Нормалізація даних. 9. Групування та розгрупування даних. 10. Трансформація даних. <p>Список рекомендованих джерел: <i>Основний:</i> 1-3. <i>Додатковий:</i> 6, 8-12. <i>Інтернет-джерела:</i> 25, 26.</p> | 2 |
| | <p>Самостійна робота: опрацювання матеріалу лекцій, підготовка до лабораторних занять з теми, ознайомлення з основною та додатковою літературою, інтернет-ресурсами, виконання індивідуальних завдань з використанням комп'ютера, підготовка до тестування за контрольними питаннями, наведеними у методичних рекомендаціях до виконання лабораторних завдань.</p> <p><i>Індивідуальна робота:</i> За визначеною аналітикою дисертаційного дослідження здійснити процес консолідації даних та виконати їх передобробку:</p> <ul style="list-style-type: none"> - обґрунтувати вибір типу сховища даних; - здійснити опис ETL-процесу; - розробити методику усунення дублікатів, протиріч та аномальних значень; - провести візуальний аналіз даних; - обґрунтувати необхідність нормалізації та квантування даних; - здійснити трансформацію даних. | 10 |
| | <p>Лабораторне заняття 1 Система Deductor. Базові навички роботи у системі Deductor Мета: ознайомитися з головним вікном <i>Deductor Studio</i>, навчитися виконувати маніпуляції з вкладками; ознайомитися зі сценарієм – базовим поняттям в <i>Deductor Studio</i>, навчитися оперувати основними вузлами сценарію. Завдання: 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> | 2 |

| | | |
|--|---|----------|
| <p>Знати: алгоритми побудови асоціативних правил та кластеризації даних; методів класифікації даних; методи та алгоритми візуального аналізу даних, аналізу текстових даних, даних, що отримано у реальному часі, даних мережі Інтернет</p> | <p>Тема 3. Технології інтелектуальної обробки даних</p> <p>Лекція 4. Асоціація даних. Асоціативні правила</p> <p>План лекції</p> <ol style="list-style-type: none"> 1. Афінітивний аналіз. Поняття типової транзакції. Предметний набір. 2. Основні поняття <i>Rules Mining</i>. Асоціативні правила. Умова та наслідок асоціативного правила. 3. Підтримка та достовірність правил. Значущість асоціативних правил. Міри корисності правил. Ліфт, левередж та покращення асоціативних правил. 4. Алгоритм <i>a priori</i>. Генерація асоціативних правил. 5. Ієрархічні асоціативні правила. Методи пошуку ієрархічних асоціативних правил. 6. Секвенціальний аналіз. | <p>2</p> |
| <p>Вміти: використовувати асоціативні правила у прикладних задачах; застосовувати на практиці алгоритми кластеризації та сегментації даних.</p> | <p>Лекція 5. Алгоритми кластеризації. Кластерний аналіз.</p> <p>План лекції</p> <ol style="list-style-type: none"> 1. Формальна постановка задачі кластеризації. 2. Базові алгоритми кластеризації. 3. Алгоритми кластеризації k-means та g-means. 4. Критерій збіжності алгоритму. Міри відстаней у кластеризації. Міри Евкліда і Манхеттена. 5. Програмні засоби кластеризації. Мережі Кохонена. Самоорганізуючі карти Кохонена. | <p>2</p> |
| | <p>Лекція 6. Класифікація даних. Логістична регресія</p> <p>План лекції</p> <ol style="list-style-type: none"> 1. Огляд методів класифікації. Процес класифікації. 2. Точність класифікації. Оцінювання класифікаційних методів. 3. Кореляційно-регресійний аналіз. Статистичні методи аналізу. 4. Лінійна регресія. Регресія з категоріальними вхідними змінними. 5. Логістична регресія. Оцінки максимальної правдоподібності. 6. Тест Чоу. ROC-аналіз. 7. 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. | <p>2</p> |
| | <p>Лекція 7. Алгоритми інтелектуального аналізу даних</p> <p>План лекції</p> <ol style="list-style-type: none"> 1. Візуальний аналіз даних - <i>Visual Mining</i> 2. Аналіз текстових даних – <i>Text Mining</i> 3. Аналіз даних у реальному часі - <i>Real Time Data Mining</i> 4. Аналіз даних мережі Інтернет – <i>Web Mining</i> <p>Список рекомендованих джерел:</p> <p><i>Основний:</i> 1-3. <i>Додатковий:</i> 5-7, 15-24. <i>Інтернет-джерела:</i> 25, 26.</p> | <p>2</p> |

| | | |
|--|--|----|
| | <p>Самостійна робота: опрацювання матеріалу лекцій, підготовка до лабораторних занять з теми, ознайомлення з основною та додатковою літературою, інтернет-ресурсами, виконання індивідуальних завдань з використанням комп'ютера, підготовка до тестування за контрольними питаннями, наведеними у методичних рекомендаціях до виконання лабораторних завдань.</p> <p><i>Індивідуальна робота:</i></p> <p>У відповідності до цілей дисертаційної роботи визначити основні технології інтелектуальної обробки даних дисертаційного дослідження (асоціативний аналіз, кластерний аналіз, класифікаційний аналіз). За визначеною технологією здійснити інтелектуальний аналіз даних (визначити множину асоціативних правил та їх характеристики, здійснити кластеризацію та (або) класифікацію даних). Визначити тип основного джерела даних для аналізу (текстові документи, графіки й діаграми, інтернет-ресурси) та описати процес і особливості обробки даних з цих джерел (<i>Text Mining, Visual Mining, Web Mining</i> тощо).</p> | 36 |
| | <p style="text-align: center;">Лабораторне заняття 2</p> <p style="text-align: center;">Асоціативні правила в системі <i>Deductor</i></p> <p>Мета: 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>Завдання:</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.</p> <p><i>Необхідно:</i></p> <ul style="list-style-type: none"> - передбачити, які товари покупці можуть вибрати в залежності від того, що вже є в їхніх кошиках; - виявити найбільш популярні товарні набори, які складаються з більше ніж одного предмету; - запропонувати рекламні акції типу: «Кожному, хто купив А і В – товар С в подарунок». <p>Вихідні дані. Представлені у файлі <i>Чеки.txt</i> двома полями – Номер транзакції і Товар. Оскільки номенклатура товарів побутової хімії надзвичайно різноманітна, то товари представлені в узагальненій формі без торгових марок: порошки, миючі засоби і т.д. (всього 37 найменувань).</p> <ol style="list-style-type: none"> 1. Використовуючи алгоритм <i>a priori</i>, побудувати асоціативні правила і зробити їх інтерпретацію. 2. Сформувати власну базу даних для подібної задачі і виконати п.1. Серед початкових даних повинно бути: <ul style="list-style-type: none"> ○ не менше 20 різних видів товару; ○ кожен чек повинен містити від 2 до 4 різних видів товару; ○ кількість чеків не менше 50. | 2 |

Лабораторне заняття №3

2

Алгоритм *k-means*. Загальні принципи кластеризації.

Мета: набуття практичних навичок роботи з алгоритмом кластеризації *k-means*.

Завдання:

Дано 10 об'єктів (*A, B, C, D, E, F, G, H, I, K*), які представлено координатами точок на площині.

1. Використовуючи алгоритм *k-means*, провести кластерний аналіз даних, розбивши початкову множину точок на 2 кластери. За початкові центри кластерів прийняти точки *E* і *J*.
2. Використовуючи алгоритм *k-means*, провести кластерний аналіз даних, розбивши початкову множину точок на 3 кластери. За початкові центри кластерів прийняти точки *A, F* і *J*.
3. Порівняти отримані результати і обрати оптимальну кількість кластерів.
4. На кожному етапі кластерного аналізу провести візуалізацію результатів розбиття на кластери.
5. Для визначення належності об'єкта до певного кластеру використати евклідову відстань та відстань Манхеттена.

Лабораторне заняття 4

2

Сегментація клієнтів з використанням карт Кохонена

Мета: набуття навичок роботи з обробником і візуалізатором Карта Кохонена в процесі розв'язування бізнес-задачі сегментації в *Deductor Studio*.

Завдання:

Постановка задачі. Керівництво філії регіональної телекомунікаційної компанії, яка надає послуги мобільного зв'язку, поставило задачу сегментації абонентської бази. Її цілі полягають в наступному:

- побудова профілів абонентів шляхом виявлення їх схожої поведінки в плані частоти, тривалості та часу дзвінків, а також щомісячних витрат;
- оцінка найбільш і найменш прибуткових сегментів.

Ця інформація може в подальшому використовуватися для:

- розробки маркетингових акцій, направлених на визначення групи клієнтів;
- розробки нових тарифних планів;
- оптимізації витрат на адресну SMS-розсилку про нові послуги і тарифи;
- попередження відтоку клієнтів у інші компанії.

Вихідні дані. Представляються у вигляді файлу *mobile.txt*.

Розв'язати бізнес-задачу сегментації абонентів за допомогою підходу, який базується на алгоритмі Кохонена і складається з двох кроків:

- кластеризація об'єктів алгоритмом Кохонена;
- побудова та інтерпретація карти Кохонена.

| | | |
|--|--|--|
| <p>Знати: принципи функціонування сучасних аналітичних платформ</p> <p>Вміти: використовувати програмні продукти <i>Microsoft Power Pivot, Microsoft Power View, Microsoft Power Query, Deductor Studio</i> для розв'язування основних задач аналізу даних; створювати звіти в аналітичній платформі <i>Microsoft Power BI</i> з використанням сучасних методів візуалізації даних</p> | <p align="center">Тема 4. Інструментальні засоби аналізу даних</p> <p align="center">Лекція 8. Аналітичні платформи</p> <p align="center">План лекції</p> <ol style="list-style-type: none"> 1. Програмне забезпечення в області аналізу даних. Аналітичні платформи. 2. Технології аналізу даних у продуктах <i>Microsoft Corporation. Microsoft Power Pivot. Microsoft Power View. Microsoft Power Query.</i> 3. Хмарні технології <i>Microsoft</i> для аналізу та візуалізації даних. 4. <i>Microsoft Power BI.</i> 5. Аналітична платформа <i>Deductor Studio.</i> <p align="center">Список рекомендованих джерел:</p> <p><i>Основний:</i> 1-3. <i>Додатковий:</i> 6, 26, 25-8, 11, 13, 15, 16, 18-22. <i>Інтернет-джерела:</i> 25, 26.</p> <p>Самостійна робота: опрацювання матеріалу лекцій, підготовка до лабораторних занять з теми, ознайомлення з основною та додатковою літературою, інтернет-ресурсами, виконання індивідуальних завдань з використанням комп'ютера, підготовка до тестування за контрольними питаннями, наведеними у методичних рекомендаціях до виконання лабораторних завдань.</p> <p><i>Індивідуальна робота:</i></p> <ol style="list-style-type: none"> 1. Використати аналітичну платформу <i>Deductor</i> для реалізації основних технологій аналізу даних дисертаційного дослідження (асоціація, кластеризація, класифікація, регресія). 2. Використати аналітичну платформу <i>Microsoft Power BI</i> для створення звіту за результатами аналізу даних дисертаційного дослідження із застосуванням візуальних елементів. <p align="center">Лабораторне заняття 5</p> <p align="center">Створення та обробка даних в системі Power BI</p> <p>Мета: набуття базових навичок роботи в аналітичній платформі <i>Power BI.</i></p> <p>Завдання: Використати дані Державної служби статистики України http://www.ukrstat.gov.ua/operativ/operativ2018/ct/sctp/sctp_2019_u.xlsx щодо цін за оренду однокімнатної квартири у м. Києві впродовж 2018-2019 рр.</p> <ol style="list-style-type: none"> 1. Здійснити реєстрацію у програмі <i>Microsoft Power BI Desktop.</i> 2. Безпосередньо в програмі ввести дані Державної служби статистики України (таблиця 1). 3. Налаштувати необхідний формат даних таблиці (ціле число, дата, текстове поле, грошові одиниці). 4. Провести візуалізацію даних, використовуючи такі інструменти: <ol style="list-style-type: none"> 1) графік; 2) стовпчикова діаграма; 3) кругова діаграма. | <p align="center">2</p> <p align="center">14</p> <p align="center">2</p> |
|--|--|--|

5. Змінити кольори елементів візуалізацій, шрифт і стиль надписів, встановлених за замовченням.
6. Продемонструвати інтерактивність візуалізацій, по черзі виділивши їх окремі компоненти.
7. Побудувати агреговані таблиці, здійснивши ієрархію дат до рівня кварталу (таблиця 2) та року (таблиця 3). Таблиці розташувати на окремих аркушах в Power BI.
8. Використавши панель аналітики, за даними таблиці 1:
 - 1) встановити на графіку лінії максимуму, мінімуму та середнього значення;
 - 2) побудувати прогноз ціни оренди однокімнатної квартири на перше півріччя 2019 року.

2

Лабораторне заняття 6

Імпорт даних та базові операції в Power BI

Мета: набуття базових навичок візуалізації даних в аналітичній платформі *Power BI*.

Завдання:

Завдання 1.

1. Створіть новий файл Power BI Desktop і завантажте таблицю синглів, що продаються мільйонами примірників, на веб-сайті *The Guardian*.
2. Завантажте файл *million-sellers.csv*.
3. Створіть стовпчикову діаграму мільйонних продажів відомих виконавців
4. Створіть кругову діаграму, яка показує обсяг продажів за рік.
5. Змініть налаштування деталізації, щоб після натискання виконавця на графіку стовпців першої діаграми, відбувалася відповідна фільтрація даних кругової діаграми.

Завдання 2.

1. Відкрийте звіт Power BI *Menu.pbix*. Ви повинні побачити, що він містить розбивку калорій для різних типів піци *Pizza Express*.
2. Створіть діаграму для порівняння середніх калорій для 3 різних типів піци.
3. Додайте легенду до кругової діаграми (щоб вона з'явилася внизу) і налаштуйте мітки даних, щоб показати значення категорії та деталей для кожного фрагмента.
4. Додайте кільцеву діаграму, щоб показати розподіл калорій за кожним видом піци. Діаграма повинна відображати середнє значення кількості калорій для кожної піци.
5. Налаштуйте візуальну взаємодію так, щоб при натисканні на тип піци на першій діаграмі отримувати фільтровані дані на другій. Коли ви натискаєте на тип даних *Legera* на першому графіку, це має показати другий.

Разом

90

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*
* Курсивом виділені джерела, що є у бібліотеці КНТЕУ

Додатковий

5. *Гнатієнко Г. М., Снитюк В. Є. Експертні технології прийняття рішень: монографія. 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.
8. Панфилов А. Н., Скоба А. Н., Кузнецова А. В., Зуев В. А. Интеллектуальный анализ данных. Новочеркасск: Лик, 2016. – 76 с.
9. Ситник В. Ф., Краснюк М. Т. Інтелектуальний аналіз даних (дейтамайнінг) : навч. посібн. Київ : КНЕУ, 2007. 376 с.
10. Субботін С. О. Подання й обробка знань у системах штучного інтелекту та підтримки прийняття рішень : навч. посібн. Запоріжжя : ЗНТУ, 2008. 341 с.
11. Яковлев В. Б. Финансовый анализ данных в *Deductor Studio*. М.: ОнтоПринт, 2018. – 168 с.
12. Adamson C. Mastering Data Warehouse Aggregates: Solutions for Star Schema Performance. Wiley Publishing Inc., 2006. 318 p.
13. Albright S. C., Winston W., Zappe C. Data Analysis and Decision Making. Boston : Cengage Learning, 2016. 948 p.
14. Cao L., Yu P. S., Zhang C., Zhang H. Data Mining for Business Applications. Springer Science; Business Media, 2008. 402 p.
15. Coodley M. O. Introduction to Microsoft Power BI: bring your data to life! CreateSpace Independent Publishing Platform, 2016. 128 p.
16. Etaati L. Advance Analytics with Power BI and R. Auckland : Radacad Systems Limited, 2017. 179 p.
17. Fabrice G., Hamilton N. J. Quality Measures in Data Mining. Berlin; Heidelberg: Springer-Verlag, 2007. 361 p.
18. Ferrari A., Russo M. Introducing Microsoft Power BI. Redmond : Microsoft Press, 2016. 407 p.
19. Han J., Kamber M. Data Mining: Concepts and Techniques. Morgan Kaufmann Publishers, 2006. 800 p.

20. Linoff G. S. Data Analysis Using SQL and Excel. Indianapolis: Wiley, 2015. 792 p.
21. Linoff G. S., Berry M. J. A. Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management. Indianapolis: Wiley, 2011. 888 p.
22. Milton M. Head First Data Analysis: A learner's guide to big numbers, statistics, and good decisions. Sebastopol: O'Reilly Media, 2009. 435 p.
23. RapidMiner: Data Mining Use Cases and Business Analytics Applications / Edited by Markus Hofmann & Ralf Klinkenberg. Minneapolis : CRC Press, 2004. 518 p.
24. Upton G. Categorical data analysis by example. New Jersey: John Wiley & Sons Inc, 2017. 198 p.

Internet-ресурсы

25. Microsoft Power BI Guided Learning URL: <https://docs.microsoft.com/uk-ua/power-bi/guided-learning> (дата звернення 18.03.2020).
26. *Deductor Studio* URL: <https://basegroup.ru/deductor/description> (дата звернення 18.03.2020).