

**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. № 1 від « 11 » 2020 р.)

Ректор

А. А. Мазаракі



**ТЕХНОЛОГІЇ АНАЛІЗУ ДАНИХ /
DATA ANALYSIS TECHNOLOGIES**

**РОБОЧА ПРОГРАМА /
COURSE OUTLINE**

освітній ступінь	бакалавр	/	Bachelor
галузь знань	<u>12 Інформаційні технології</u>	/	<u>Information Technology</u>
спеціальність	<u>122 Комп'ютерні науки</u>	/	<u>Computer Science</u>
спеціалізація	<u>Комп'ютерні науки</u>	/	<u>Computer Science</u>

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
цифрової економіки та системного аналізу 02.11.2020 р., протокол № 4

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

**ТЕХНОЛОГІЇ АНАЛІЗУ ДАНИХ/
DATA ANALYSIS TECHNOLOGIES**

**РОБОЧА ПРОГРАМА /
COURSE OUTLINE**

освітній ступінь	бакалавр	/	Bachelor
галузь знань	<u>12 Інформаційні технології</u>	/	<u>Information Technology</u>
спеціальність	<u>122 Комп'ютерні науки</u>	/	<u>Computer Science</u>
спеціалізація	<u>Комп'ютерні науки</u>	/	<u>Computer Science</u>

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

Для вивчення дисципліни відводиться 180 год. / 6 кредитів ЄКТС (лекцій – 28 год., лабораторних занять – 28 год., самостійної роботи – 124 год.), підсумковий контроль – екзамен

Назва теми	Кількість годин				Форма контролю
	Усього годин / кредитів	з них:			
		лекцій	лабораторні заняття / МК	самостійна робота аспірантів	
Тема 1. Передобробка даних	42	4	8	30	Т, ІДЗ, Пр, П
Тема 2. Асоціація даних	16	4	6	6	Т, ІДЗ, Пр, П
Тема 3. Кластеризація даних	16	4	6	6	Т, ІДЗ, Пр, П
Тема 4. Класифікація та регресія даних	16	2	6	8	Т, ІДЗ, Пр, П
Тема 5. Технології інтелектуальної обробки даних	22	8	0	14	Т
Тема 6. Інструментальні засоби аналізу даних	24	4	4	16	Т, ІДЗ, Пр, П
Тема 7. Створення моделі даних	12	4	0	8	Т, ІДЗ, Пр, П
Тема 8. Побудова аналітичних звітів	32	4	4	24	Т, ІДЗ, Пр, П, МК
Підсумковий контроль – екзамен					
Разом	180/6	34	34	112	x

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

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

Результати навчання	Навчальна діяльність ¹	Робочий час здобувача, год
1	2	3
<p>Знати: основні поняття <i>DataScience</i>, <i>ArtificialIntelligence</i>, <i>BusinessIntelligence</i>, етапи розв'язування задач аналізу даних, технології <i>DataMining</i>.</p> <p>Вміти: оцінювати якість даних, здійснювати консолідацію, очищення та збагачення даних.</p>	<p>Тема 1. Передобробка даних.</p> <p>Лекція 1. Вступ у DataScience. Поняття про дані План лекції</p> <ol style="list-style-type: none"> 1. Поняття про дані. 2. Типи та види даних. Форми представлення даних. 3. Вимірювання і шкали в аналізі даних. Реляційні та багатовимірні дані. Метадані. 4. Формати зберігання даних. 5. Типи наборів даних. Транзакційні дані. <p>Лекція 2. Перевірка якості даних. Методи та алгоритми очищення даних. План лекції</p> <ol style="list-style-type: none"> 1. Типовий набір інструментів оцінки якості даних в аналітичній платформі. 2. Очищення від шумів і згладжування рядів даних. 3. Фільтрація даних. Відновлення пропущених значень. 4. Редагування аномальних значень. Методи корегування аномальних значень. 5. Проблема репрезентативності. Обробка дублікатів і протиріч. 6. Семплінг. Усунення незначущих факторів. <p style="text-align: center;">Список рекомендованих джерел:</p> <p><i>Основний:</i> 1-4. <i>Додатковий:</i> 6, 8-10, 15, 26, 27. <i>Інтернет-ресурси:</i> 28-30.</p>	2 2
	<p>Лабораторне заняття 1</p> <p>Імпорт з різних джерел та первинна обробка даних у системі <i>LogiNot</i></p> <p>Мета: набути практичних навичок імпорту даних у аналітичній платформі <i>LogiNot</i> з використанням текстових файлів, файлів із роздільниками формату csv та файлів <i>Excel</i>.</p> <p>Завдання: Виконати завдання імпорту на налаштування текстового файлу в аналітичній платформі <i>LogiNot</i>. Зберегти пакет під іменем <i>Surname_txt.lgp</i></p> <p>Виконати завдання імпорту та налаштування файлу із роздільниками в аналітичній платформі <i>LogiNot</i>. Зберегти пакет під іменем <i>Surname_csv.lgp</i></p> <p>Імпортувати доданий до завдання файл <i>Excel</i>. Налаштувати типи полів та інші параметри коректного імпорту.</p>	2

¹+20% інтерактивних методів навчання виділено курсивом

Побудувати візуалізатори (таблицю, статистику, графіки, діаграми тощо) для імпортованого файлу. Зберегти пакет під іменем *Surname.xls.lgp*
Додати до пакету в *Logiplot* декілька обробників для імпортованих наборів даних із попередніх завдань.

Лабораторне заняття 2

Оцінка якості даних

Мета: ознайомитися з основними методами передобробки даних: фільтрація, оцінка якості даних, заповнення пропусків, редагування викидів з використанням вбудованих інструментів *Logiplot*.

Завдання: У файлі *Food_data.xlsx* є дані про вартість умовного середнього набору продуктів харчування в м. Києві. Для прогнозування даного показника на перспективу необхідно заповнити пропуски, які візуально спостерігаються в таблиці.

Перетворити файл *Excel* у текстовий файл формату **.txt*. Імпортувати файл у систему *Logiplot* з вірним форматкуванням типів полів (рядковий, цілий). За допомогою налаштування візуалізації обрати спосіб відображення даних у вигляді таблиці та діаграми. Для ліквідації проблем із даними використати майстер обробки «*Imputation*» (Заповнення пропусків), встановивши при налаштуванні опцію обробки як упорядкованого набору. Переконавшись, що після обробки дані приймуть вигляд, придатний для подальшого аналізу.

Переіменувати гілку з імпортом файлу *Food_data.txt*, щоб її назва відображала зміст виконаної переобробки даних.

2

Лабораторне заняття 3

Очищення даних

Мета: ознайомитися з основними методами очищення даних: парціальна передобробка, спектральна передобробка, факторний аналіз, кореляційний аналіз, усунення дублікатів і протиріч з використанням вбудованих інструментів *Logiplot*.

Завдання: Виконати спектральну обробку з даними файлу «*Trade.txt*». В якості візуалізації обрати діаграму з неперервними лініями зеленого кольору. Виконати факторний аналіз даних з файлу «*Anketa1.txt*». Виконати імпорт файлу та встановити вірні типи полів (рядковий, цілий, логічний). Залишити три фактори з чотирьох для подальшої обробки. Визначити інтервал порогу значущості, при якому з множини чотирьох факторів залишається тільки три. В таблиці залишити для відображення тільки значення трьох факторів впливу, імені, прізвища та ймовірного результату щодо повернення кредиту.

Виконати кореляційний аналіз даних з файлу «*region.txt*». Виконати імпорт файлу та встановити вірні типи полів (рядковий, цілий, дійсний). Встановити ручний вибір незначущих факторів і обрати фактори «Середня за місяць нарахована заробітна плата», «Основні фонди в економіці», «Об'єм промислової продукції», «Інвестиції в основний капітал». Порог значущості встановлюємо на рівні 0,3. Виявити дублікати на прикладі даних файлу «*Anketa.txt*».

2

	<p style="text-align: center;">Лабораторне заняття 4 Трансформація даних</p> <p>Мета: ознайомитися з основними методами трансформації даних: розбиття дати, квантування, налаштування набору даних, фільтрація, заміна значень, злиття з використанням вбудованих інструментів <i>LogiNot</i>.</p> <p>Завдання: Здійснити перетворення дати. У файлі <i>Credit.csv</i> є інформація за кредитами, отриманими у банку. Завантажити його в аналітичну платформу з вірним форматуванням типів полів (рядковий, цілий, дата/час). Отримати дані за сумами взятих кредитів у розрізі тижнів. У якості візуалізатора обрати «Куб». Отримати крос-діаграму суми взятих кредитів за тижнями в розрізі цілей кредитування</p> <p>Здійснити квантування даних. У якості початкових даних обрати дані, отримані при розбитті дати файлу «<i>Credit.csv</i>». У якості візуалізатора обрати куб (факт – «Сума кредиту», виміри – «Вік» та «Дата кредитування (Рік+Тиждень)»).</p> <p>Виконати налаштування набору даних. У якості початкових даних обрати дані, отримані при квантуванні віку кредиторів з файлу «<i>Credit.csv</i>». Визначити на яку загальну суму і на який термін було надано кредити за тижнями року; визначити на яку загальну суму і на який термін було відмовлено у наданні кредитів за тижнями року.</p> <p>Виконати фільтрацію даних. Початкові дані взяти із файлу <i>Credit.csv</i>. Здійснити вибірку позичальників, які отримали позику на суму менше 10000 грн. В якості візуалізації результату обрати статистику і переконалися у правильному застосуванні фільтру за знайденими мінімальним та максимальним значенням суми кредиту.</p> <p>Провести заміну значень. Створити звіт про суми кредитів на різні цілі у розрізі статі позичальників. В якості візуалізації результату обрати куб (факт – «Сума кредиту», виміри – «Стать» та «Мета кредитування»)</p> <p>Виконати злиття даних. Початкові дані взяти із файлів <i>TradeSales.txt</i> (обсяги продажів) та <i>TradeRest.txt</i> (обсяги залишків). Додати до даних про продажі дані про відповідні залишки. Для злиття даних вибрати майстер обробки «Join» (Злиття) і в ньому – вузол із залишками, з яким необхідно з'єднати дані за типом «Внутрішнє з'єднання».</p>	2
	<p>Самостійна робота: опрацювання матеріалу лекцій, підготовка до лабораторних занять з теми «Передобробка даних», ознайомлення з основною та додатковою літературою, інтернет-ресурсами, виконання індивідуальних завдань з використанням комп'ютера.</p> <p>Індивідуальна робота:</p> <p>Розробити шаблон пакету в системі <i>LogiNot</i>. В розробленому пакеті передбачити процедури імпорту даних із текстового файлу, файлу формату <i>csv</i>, файлу <i>Excel</i>. У шаблоні передбачити приєднання різних обробників подій. Використати всі можливі візуалізатори передобробки даних.</p> <p>Використати розроблений шаблон для виконання лабораторних робіт за темою «Передобробка даних»</p>	30

	<p style="text-align: center;">Список рекомендованих джерел: Додатковий: 12, 29. Інтернет-ресурси: 30-32, 37-38.</p>	
<p>Знати: основи афінитивного аналізу, принципи генерації популярних наборів даних та асоціативних правил.</p> <p>Вміти: використовувати алгоритм <i>apriori</i> для побудови та аналізу асоціативних правил.</p>	<p style="text-align: center;">Тема 2. Асоціація даних.</p> <p style="text-align: center;">Лекція 3. Асоціація даних. Асоціативні правила.</p> <p style="text-align: center;">План лекції</p> <ol style="list-style-type: none"> 1. Афінитивний 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. Секвенціальний аналіз. <p style="text-align: center;">Список рекомендованих джерел: Основний: 2. Додатковий: 16. Інтернет-ресурси: 29.</p>	4
	<p style="text-align: center;">Лабораторне заняття 5</p> <p style="text-align: center;">Асоціативні правила у стимулюванні оптових покупок</p> <p>Мета: навчитися генерувати асоціативні правила та на основі алгоритму <i>apriori</i>.</p> <p>Завдання: Дано базу транзакцій, утворену за результатами аналізу діяльності міжнародної компанії <i>Dapom</i> з продажу продукції власного виробництва. З метою збільшення прибутку було вирішено створити вигідні для оптових споживачів пакетні пропозиції з придбання декількох видів продукції одночасно, а потім пропонувати споживачам придбання інших видів продукції, яка найчастіше закуповується разом. З цією метою керівництво компанії поставило задачу перед аналітичним відділом визначити зв'язок між видами продукції, які найчастіше закуповують оптові споживачі. Представлено аналіз закупівель 20 оптових споживачів. Використовуючи алгоритм <i>a priori</i>, визначити популярні набори видів продукції, які купляють мінімум 5 споживачів. Побудувати асоціативні правила з рівнем достовірності не менше 70%. Впорядкувати правила за мірою добутку підтримки і достовірності і визначити 5 найбільш сильних правил. Дати письмову інтерпретацію знайденим правилам (визначити який вид продукції слід додатково запропонувати покупцям).</p> <p style="text-align: center;">Лабораторне заняття 6</p> <p style="text-align: center;">Асоціативні правила в стимулюванні роздрібних продажів</p> <p>Мета: Навчитися виділяти правила з максимальною достовірністю з метою розробки більш досконалої маркетингової стратегії для оптимізації закупок товарів. Оволодіти навичками роботи з обробником <i>AssociationRules</i> в <i>Logitot</i> для реалізації алгоритму <i>apriori</i>.</p>	4
	<p style="text-align: center;">Лабораторне заняття 6</p> <p style="text-align: center;">Асоціативні правила в стимулюванні роздрібних продажів</p> <p>Мета: Навчитися виділяти правила з максимальною достовірністю з метою розробки більш досконалої маркетингової стратегії для оптимізації закупок товарів. Оволодіти навичками роботи з обробником <i>AssociationRules</i> в <i>Logitot</i> для реалізації алгоритму <i>apriori</i>.</p>	2

	<p>Завдання: 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. 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> <p>Використовуючи алгоритм <i>a priori</i>, побудувати асоціативні правила і зробити їх інтерпретацію. Сформувати власну базу даних для подібної задачі, провести генерацію популярних наборів і визначити систему асоціативних правил. Серед початкових даних повинно бути: не менше 20 різних видів товару; кожен чек повинен містити від 2 до 4 різних видів товару; кількість чеків не менше 50.</p>	
	<p>Самостійна робота: опрацювання матеріалу лекцій, підготовка до лабораторних занять з теми «Асоціація даних», ознайомлення з основною та додатковою літературою, інтернет-ресурсами, виконання індивідуальних завдань з використанням комп'ютера, підготовка до модульного контролю.</p> <p><i>Індивідуальна робота:</i></p> <p>Розробити шаблон пакету в системі <i>LogiNot</i>. В розробленому пакеті передбачити процедури імпорту даних із файлу <i>Excel3</i> ідентифікатором транзакції та найменуванням товару або послуги. У шаблоні передбачити використання обробника «<i>AssociationRules</i>». У якості візуалізаторів використати аналіз «<i>What/If</i>», популярні набори та асоціативні правила.</p> <p>Використати розроблений шаблон для виконання лабораторних робіт за темою «Асоціація даних»</p> <p style="text-align: center;">Список рекомендованих джерел:</p> <p><i>Основний:</i> 2. <i>Додатковий:</i> 16. <i>Інтернет-ресурси:</i> 29.</p>	6
<p>Знати: поняття кластеризації та сегментації даних</p> <p>Вміти: застосовувати на практиці алгоритми <i>k-means, g-means</i>, проводити кластеризацію даних, використовувати аналітичні платформи для задач кластеризації та сегментації даних</p>	<p style="text-align: center;">Тема 3. Кластеризація даних</p> <p style="text-align: center;">Лекція 4. Алгоритми кластеризації. Кластерний аналіз.</p> <p style="text-align: center;">План лекції</p> <ol style="list-style-type: none"> 1. Формальна постановка задачі кластеризації. 2. Задачі кластеризації: вивчення даних, полегшення аналізу, стиснення даних, прогнозування, виявлення аномалій. 3. Приклади кластеризації в різних областях знань. 4. Базові алгоритми кластеризації. 5. Ієрархічні та ітеративні методи кластерного аналізу. 6. Карти та мережі Кохонена. <p style="text-align: center;">Список рекомендованих джерел:</p> <p><i>Основний:</i> 1- 3. <i>Додатковий:</i> 5, 7, 20. <i>Інтернет-ресурси:</i> 28.</p>	4

	<p style="text-align: center;">Лабораторне заняття 7</p> <p style="text-align: center;">Алгоритм k-means. Загальні принципи кластеризації.</p> <p>Мета: набуття практичних навичок роботи з алгоритмом кластеризації <i>k-means</i>.</p> <p>Завдання: Дано 10 об'єктів (A, B, C, D, E, F, G, H, I, K), які представлено координатами точок на площині. Використовуючи алгоритм <i>k-means</i>, провести кластерний аналіз даних, розбивши початкову множину точок на 2 кластери. За початкові центри кластерів прийняти точки E_i. 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_i. Порівняти отримані результати і обрати оптимальну кількість кластерів. На кожному етапі кластерного аналізу провести візуалізацію результатів розбиття на кластери. Для визначення належності об'єкта до певного кластеру використати евклідову відстань та відстань Манхеттена.</p>	4
	<p style="text-align: center;">Лабораторне заняття 8</p> <p style="text-align: center;">Сегментація даних на основі карт Кохонена</p> <p>Мета: набуття навичок роботи з обробником і візуалізатором <i>Self-Organizing Networks</i> в процесі розв'язування бізнес-задачі сегментації в <i>LogiNot</i>.</p> <p>Завдання: 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"> - 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. <p>Ця інформація може в подальшому використовуватися для:</p> <ul style="list-style-type: none"> - 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. <p>Вихідні дані представлені у вигляді файлу <i>mobile.txt</i>.</p> <p>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>	2
	<p>Самостійна робота: опрацювання матеріалу лекцій, підготовка до лабораторних занять з теми «Кластеризація даних», ознайомлення з основною та додатковою літературою, інтернет-ресурсами, виконання індивідуальних завдань з використанням комп'ютера, підготовка до модульного контролю.</p> <p>Індивідуальна робота: Розробити шаблон пакету в системі <i>LogiNot</i> для проведення кластеризації даних. В розробленому пакеті передбачити процедури імпорту даних із файлу <i>Excel</i> з характеристиками об'єктів, що підлягають кластеризації. У шаблоні передбачити використання обробників «<i>Clustering</i>», «<i>Transaction Clustering</i>»,</p>	6

	«EMClustering». Використати розроблений шаблон для виконання лабораторних робіт за темою «Кластеризація даних»	
	Список рекомендованих джерел: <i>Основний:</i> 2. <i>Додатковий:</i> 10, 16. <i>Інтернет-ресурси:</i> 29.	
Знати: основні методи класифікації даних, регресії та прогнозування даних, логістичної регресії Вміти: проводити класифікацію даних з використанням дерева рішень, здійснювати регресійний аналіз та прогнозування даних з використанням лінійної, нелінійної, динамічної та логістичної регресії.	Тема 4. Класифікація та регресія даних Лекція 5. Регресійний аналіз даних. Логістична регресія. План лекції 1. Кореляційно-регресійний аналіз. Статистичні методи аналізу. 2. Лінійна регресія. Регресія з категоріальними вхідними змінними. 3. Логістична регресія. Оцінки максимальної правдоподібності. 4. Тест Чоу. ROC-аналіз. 5. Множинна логістична регресія. Простий байесівській класифікатор. Список рекомендованих джерел: <i>Основний:</i> 1, 3. <i>Додатковий:</i> 6, 7, 11, 12, 14. <i>Інтернет-ресурси:</i> 28.	2
	Лабораторне заняття 9 Регресійний аналіз даних. Лінійна і квадратична регресія Мета: обробка даних та побудова прогнозу з використанням обробника <i>LinearRegression</i> . Завдання: <i>1. Лінійна регресія. Є набір даних щодо вартості квартир в залежності від кількості кімнат, загальної площі, житлової площі, площі кухні, типу будинку тощо (усього 11 характеристик житла). Необхідно побудувати регресійну модель, яка дозволить оцінити вартість одно- та двокімнатних квартир. Імпортуйте в <i>LogiNot</i> набір даних з файлу <i>Flats.xlsx</i>. За допомогою вузла <i>RowFilter</i> здійсніть відбір одно- та двокімнатних квартир. На основі відфільтрованих даних побудуйте лінійну модель з використанням обробника <i>LinearRegression</i>. При налаштуванні параметрів лінійної регресії:</i> - поділ на навчальну та тестову вибірку не здійснювати; - обмеження на діапазон вихідних значень не задавати; - в якості методу відбору змінних використати прямий відбір (<i>forward</i>). Імовірності включення та видалення змінної залишити за замовченням; - в якості результату аналізу обрати <i>RegressionReport</i> (Звіт за регресією). <i>На основі проведеного аналізу: оцініть якість регресійної моделі за допомогою коефіцієнта детермінації; побудуйте регресійну модель множинної лінійної регресії виду</i> $y = a_0 + a_1x_1 + a_2x_2 \dots + a_nx_n$ <i>2. Квадратична регресія. Є дані щодо заробітної плати</i>	4

робітників, їх вік та стаж. Необхідно визначити залежність заробітної плати від віку робітників.

Імпортуйте в LogiNot набір даних з файлу Salary.xlsx. Налаштуйте типи полів згідно із даними у файлі імпорту. При виборі засобів відображення результату оберіть таблицю та діаграму залежності заробітної плати від стажу робітника.

За допомогою вузла Sort здійсніть упорядкування записів вибірки за зростанням заробітної плати. У якості засобів відображення результату оберіть таблицю та діаграму залежності заробітної плати від віку робітника

Зобразити залежність заробітної плати від віку робітників за допомогою моделі квадратичної регресії:

$$y = a_0 + a_1x + a_2x^2$$

Звести модель до лінійного вигляду шляхом введення додаткового стовпця «Вік у квадраті». Для цього використати метод обробки Calculator. У якості засобів відображення результату оберіть таблицю.

Побудувати модель регресії з використанням майстра обробки LinearRegression. При налаштуванні параметрів лінійної регресії:

- поділ на навчальну та тестову вибірку не здійснювати;
- обмеження на діапазон вихідних значень не задавати;
- в якості метода відбору змінних використати повне включення (Enter). Імовірності включення та видалення змінної залишити за замовченням;
- в якості результату аналізу обрати RegressionReport (Звіт за регресією) та Diagram (Діаграма).

На основі проведеного аналізу оцініть якість регресійної моделі за допомогою коефіцієнта детермінації.

2

Лабораторне заняття 10

Логістична регресія. Оцінка кредитоспроможності позичальників.

Мета: побудова моделі бінарної логістичної регресії та визначення рівня її якості.

Завдання: Є набір даних про кредитні історії 700 клієнтів банку і 150 клієнтів, які мають намір звернутися до банку за кредитом. Інформація про клієнтів банку містить такі дані: вік, освіта, стаж комерційної діяльності, кількість повних років постійного місця проживання, річний дохід (у тисячах доларах США), борги (% від річного доходу), борги за кредитною картою (тисяч доларів США), інші борги (тисяч доларів США). Для клієнтів банку зазначається, чи є він боржником банку.

Необхідно, використовуючи 70% випадкову вибірку клієнтів, які вже брали кредит у банку, побудувати регресійну логістичну модель, яка дозволить передбачити неповернення кредиту клієнтом. Використовуючи 30% клієнтів, що залишилися, необхідно з'ясувати ступінь придатності побудованої моделі для передбачення випадків неповернення кредиту клієнтами банку. Визначити, які змінні можуть бути виключені з моделі без втрати її якості.

Імпортувати в LogiNot набір даних з файлу BankCredit.xlsx.

	<p><i>Налаштувати типи полів згідно із даними у файлі імпорту (цілий тип, дійсний тип). Обрати спосіб відображення даних імпорту у вигляді таблиці. Перевірте якість даних. У звіті наявні пропуски, викиди і екстремальні значення. Методи редагування викидів та екстремальних значень залишити без змін. На основі вузла оцінки якості даних заповнити пропуски в базі даних, використавши майстер обробки Imputation. На основі бази даних, що не містить пропусків, провести редагування викидів, використавши майстер обробки EliminateOutliers.</i></p> <p><i>Для прогнозування ймовірності відгуку на подію сформувати цільову бінарну змінну «Відгук» на основі змінної «Чи є клієнт боржником банку?» В якості методу обробки використати вузол Calculator. Розбити множину клієнтів на дві вибірки: навчальну і тестову. У якості методу обробки використати вузол Partitioning. При налаштуванні розбиття:</i></p> <ul style="list-style-type: none"> - визначити обсяг навчальної множини у 70%; - задати метод вибірки даних – Stratified, який забезпечує однакову частку відгуків в обох вибірках; - для здійснення стратифікації вказати поле Відгук; - метод відображення даних – Partitioning. <p><i>Побудувати модель прогнозування ймовірності перспектив повернення боргу кредитором за допомогою логістичної регресії, обравши в якості метода обробки вузол LogisticRegression.</i></p> <p><i>При налаштуванні параметрів логістичної регресії:</i></p> <ul style="list-style-type: none"> - обсяг навчальної множини встановити на рівні 70%; - спосіб поділу початкової множини ByColumn; - метод відбору змінних в регресійній моделі – пряий відбір (Forward); - істинні частоти «подій» та «не-подій» визначити із навчальної множини; - параметри перетворення шансів у лінійну шкалу: референсна точка – 600, відношення шансів у референсній точці – 50, крок подвоєння відношення шансів – 20, обрати тільки додатні бали з округленням до цілого значення; - в якості способів відображення даних обрати коефіцієнти регресії, звіт за регресією, якість класифікації та «What/If». <p><i>Провести аналіз отриманих візуалізаторів логістичної регресії. Провести аналіз візуалізатора коефіцієнтів регресії. Провести аналіз результатів за візуалізатором якості класифікації, побудувавши ROC-криву та Lift-криву, а також розрахувати основні метрики бінарних класифікаторів: AUC, KS, Gini.</i></p> <p><i>Провести аналіз візуалізатора «What/If».</i></p>	
	<p>Самостійна робота: опрацювання матеріалу лекцій, підготовка до лабораторних занять з теми «Класифікація та регресія даних», ознайомлення з основною та додатковою літературою, інтернет-ресурсами, виконання індивідуальних завдань з використанням комп'ютера, підготовка до модульного контролю.</p> <p><i>Індивідуальна робота:</i></p> <p>Розробити шаблон пакету в системі Logit для проведення класифікації та регресії даних і побудови прогнозової моделі. В розробленому пакеті передбачити процедури імпорту даних із</p>	8

	<p>файлу <i>Excelz</i> даними для прогнозу. У шаблоні передбачити використання обробників «<i>LinearRegression</i>», «<i>LogisticRegression</i>», «<i>ARIMAX</i>».</p> <p>Використати розроблений шаблон для виконання лабораторних робіт за темою «Класифікація та регресія даних»</p> <p>Список рекомендованих джерел:</p> <p><i>Основний:</i> 2. <i>Додатковий:</i> 10, 16. <i>Інтернет-ресурси:</i> 29.</p>	
<p>Знати: принципи візуального аналізу даних (<i>Visual Mining</i>), аналізу текстової інформації (<i>Text Mining</i>), аналізу даних в режимі реального часу (<i>Real-Time Mining</i>), аналізу даних з мережі Інтернет (<i>Web Mining</i>).</p>	<p>Тема 5. Технології інтелектуальної обробки даних</p> <p>Лекція 6. Візуальний аналіз даних - <i>Visual Mining</i>. План лекції</p> <ol style="list-style-type: none"> 1. Виконання візуального аналізу даних. 2. Характеристики засобів візуалізації даних. Методи візуалізації. 3. Методи геометричних перетворень. 4. Методи, орієнтовані на пікселі. Ієрархічні образи. <p>Лекція 7. Аналіз текстових даних - <i>Text Mining</i>. План лекції</p> <ol style="list-style-type: none"> 1. Методи класифікації текстових документів. 2. Попередня обробка тексту. Видалення стоп-слів. Стемінг. <i>N</i>-грами. Зведення реєстра. 3. Класифікація текстів. Методи кластеризації текстових документів: ієрархічні, бінарні. 4. Задача анутовання текстів. Пошук асоціацій. Первинний витяг ключових понять. Навігація текстом. Засоби аналізу текстової інформації. <p>Лекція 8. Аналіз даних у реальному часі - <i>Real Time Data Mining</i>. План лекції</p> <ol style="list-style-type: none"> 1. Ідея <i>Data Mining</i> у реальному часі. Адаптація системи до загальної концепції. Адаптивне добування даних. 2. Статичний <i>Data Mining</i> і <i>Data Mining</i> в реальному часі 3. Рекомендаційні машини. Класифікація рекомендаційних машин. 4. Аналіз ринкового кошика і секвенційний аналіз. 5. Інструменти <i>Data Mining</i> у реальному часі. <p>Лекція 9. Аналіз даних мережі Інтернет - <i>Web Mining</i>. План лекції</p> <ol style="list-style-type: none"> 1. Складності аналізу даних з мережі Інтернет. 2. Етапи <i>Web Mining</i>. Категорії <i>Web Mining</i>. Аналіз використання веб-ресурсів. 3. Методи <i>Web Mining</i> з точки зору вирішуваних завдань і реалізованих підходів. 4. Пошук інформації. Аналіз структури сегмента мережі. Використання знань з веб-ресурсів. 5. Персоналізація інформації. Пошук шаблонів в поведінці користувачів. <p>Список рекомендованих джерел:</p> <p><i>Основний:</i> 1-3. <i>Додатковий:</i> 5-7, 9, 11, 12, 14-18, 26. <i>Інтернет-ресурси:</i> 28.</p>	<p>2</p> <p>2</p> <p>2</p> <p>2</p>

	<p>застосовуються в кожному рядку стовпців. Створіть кругову діаграму, яка показує обсяг продажів за рік.</p> <p>Змініть налаштування деталізації, щоб після натискання виконавця на графіку стовпців першої діаграми, відбувалася відповідна фільтрація даних кругової діаграми.</p> <p>Завантажте файл Меню.xlsx, який містить розбивку калорій для різних типів піци Pizza Express. Створіть діаграму для порівняння середніх калорій для 3 різних типів піци. Додайте легенду до кругової діаграми (щоб вона з'явилася внизу) і налаштуйте мітки даних, щоб показати значення категорії та деталей для кожного фрагмента.</p> <p>Додайте кільцеву діаграму, щоб показати розподіл калорій за кожним видом піци. Діаграма повинна відображати середнє значення кількості калорій для кожної піци. Налаштуйте візуальну взаємодію так, щоб при натисканні на тип піци на першій діаграмі отримувати фільтровані дані на другій.</p> <p style="text-align: center;">Лабораторне заняття 12</p> <p style="text-align: center;">Імпорт даних із баз даних та веб-ресурсів</p> <p>Мета: отримання практичних навиків даних із веб-ресурсів.</p> <p>Завдання: Завантажте дані про переможців Чемпіонату Європи з футболу УЄФА, які можна знайти в таблиці результатів на сторінці Вікіпедії за адресою http://en.wikipedia.org/wiki/UEFA_European_Football_Championship Не використовуйте україномовну або російськомовну версії цієї сторінки, оскільки вони відрізняються від англійської версії, яка є найбільш повною.</p> <p>Зробити попередню очистку даних перед завантаженням їх у платформу. Відкрийте редактор PowerQuery та видаліть всі непотрібні стовпці, залишивши тільки Year (Рік) і FinalWinners (Переможці фіналу). З використанням меню Replace виконайте заміну значень, видаливши слово «Details» з усіх рядків стовпця Year. За допомогою пункту Rename залишити у стовпці FinalWinner тільки переможців фіналів, яким можна присвоїти ім'я Country (Країна).</p> <p>Створити візуалізацію звіту з використанням інструменту Map. Змінити величину маркера країни на карті, щоб розмір маркера відображав кількість перемог для кожної країни. Розфарбувати точку даних Germany різні кольори, щоб розділити переможців з ФРН та об'єднаної Німеччини. Змінити колір та стиль заголовку карти. Побудуйте кругову діаграму, що відображає внесок кожної країни у загальну кількість перемог. Проведіть сортування секторів діаграми так, щоб країни найбільшою кількістю перемог відображалися на круговій діаграмі поруч.</p>	2
	<p>Самостійна робота: опрацювання матеріалу лекцій, підготовка до лабораторних занять з теми «Інструментальні засоби аналізу даних», ознайомлення з основною та додатковою літературою, інтернет-ресурсами, виконання індивідуальних завдань з використанням комп'ютера, підготовка до модульного контролю.</p> <p>Індивідуальна робота:</p> <p>Знайдіть веб-ресурси з довідковою інформацією про импорт даних інших категорій до аналітичної платформи Microsoft Power BI. Здійсніть практичну реалізацію імпорту з двох</p>	16

	<p>із перелічених джерел даних, які не були розглянуті на лабораторних роботах з даної теми: <i>XML, MicrosoftAccessDatabases, JSON, SQL, OracleDatabases, ODBC, GoogleAnalytics, AzureSQLDatabase</i>.</p> <p>Список рекомендованих джерел:</p> <p><i>Основний:</i> 1-3 <i>Додатковий:</i> 11, 12, 14. <i>Інтернет-ресурси:</i> 28.</p>	
<p>Знати: алгоритми трансформації даних у <i>PowerQuery</i>, функції мови <i>DAx</i> для роботи з даними</p> <p>Вміти: створювати схеми даних типу «зірка» та «сніжинка», встановлювати та модифікувати зв'язки в моделі даних, використовувати мову <i>DAx</i> для створення мір та обчислювальних стовпців у моделі даних</p>	<p>Тема 7. Створення моделі даних</p> <p>Лекція 12. Створення моделі даних</p> <p>План лекції</p> <ol style="list-style-type: none"> 1. Створення зв'язків таблиці. 2. Схеми зірки та сніжинки у <i>Power BI Desktop</i>. 3. Денормалізація даних у моделі. 4. Модифікація зв'язків моделі. 5. Створення зручної моделі. <p>Список рекомендованих джерел:</p> <p><i>Основний:</i> 1-3 <i>Додатковий:</i> 8, 9, 11-14. <i>Інтернет-ресурси:</i> 28.</p>	4
	<p>Самостійна робота: опрацювання матеріалу лекцій, підготовка до лабораторних занять з теми «Передобробка даних в системі <i>Power BI Desktop</i>», ознайомлення з основною та додатковою літературою, інтернет-ресурсами, виконання індивідуальних завдань з використанням комп'ютера, підготовка до модульної контрольної роботи.</p> <p><i>Індивідуальна робота:</i> Знайти серед веб-ресурсів порталу відкритих даних https://data.gov.ua набір даних за однією з груп: будівництво, економіка та бізнес, освіта і культура, сільське господарство, транспорт, держава, земля, охорона здоров'я, соціальний захист, фінанси, екологія, молодь і спорт, податки, стандарти, юстиція. Завантажити набір даних до платформи <i>PowerBI</i> за прямим доступом до відповідної веб-сторінки. Провести трансформацію даних, залишивши тільки стовпці, які будуть використані для дослідження. Побудувати модель завантажених даних. Додати до існуючих таблиць в моделі принаймні один додатковий обчислювальний стовпець і одну міру. Побудувати звіт з використанням базових елементів візуалізації даних.</p> <p>Список рекомендованих джерел:</p> <p><i>Основний:</i> 1-3 <i>Додатковий:</i> 11-14. <i>Інтернет-ресурси:</i> 28.</p>	8

<p>Знати: основні інструменти візуалізації аналітичних звітів, засоби захисту й публікації звітів, оновлення даних у побудованих звітах.</p> <p>Вміти: створювати та формувати інтерактивні візуальні компоненти <i>PowerBIDesktop</i>, публікувати аналітичні звіти на платформі <i>PowerBIService</i>, створювати вбудований код звіту на веб-ресурсах</p>	<p align="center">Тема 8. Побудова аналітичних звітів</p> <p align="center">Лекція 13. Візуалізація звітів у Power BI Desktop</p> <p align="center">План лекції</p> <ol style="list-style-type: none"> 1. Панелі інструментів <i>PowerBIDesktop</i>. 2. Робота з таблицями. 3. Створення матриці. 4. Одномірні та багатомірні картки. 5. Створення різних типів діаграм. <p align="center">Лекція 14. Публікація та поширення аналітичних звітів</p> <p align="center">План лекції</p> <ol style="list-style-type: none"> 1. Робоча область користувача. 2. Додавання плиток на інформаційну панель. 3. Спільне використання звітів. 4. Публікація та обмін звітами. 5. Налаштування оновлення даних в опублікованих звітах. <p align="center">Список рекомендованих джерел:</p> <p><i>Основний:</i> 1-3 <i>Додатковий:</i> 11, 12, 14. <i>Інтернет-ресурси:</i> 28.</p>	<p align="center">2</p> <p align="center">2</p>
	<p align="center">Лабораторне заняття 13</p> <p align="center">Створення базових візуалізацій</p> <p>Мета: створення стандартних візуальних зображень; створення візуалізацій на основі карти; вивчення візуальних взаємодій.</p> <p>Завдання: <i>Завантажити до аналітичної платформи дані колектору міста Сан-Франциско. Використовуючи інструмент <i>Tablevisualization</i>, створити таблицю з полями <i>Typeofrequest</i>, <i>AveDaysOpen CY</i> та <i>AveDaysOpen PY</i>. На панелі <i>Filters</i> відфільтрувати таблицю так, щоб відображались лише ті випадки, коли <i>Status</i> закритий. Відформатувати таблицю рядками, що чергуються (смужками) і змінити кегль шрифту на 10.</i></p> <p><i>Створити 100-відсоткову гістограму з панелі інструментів <i>Visualizations</i>, додавши поле <i>NumberofCases</i> до області <i>Values</i>, <i>Type</i> – до області <i>Axis</i>, а <i>Year</i> – до області <i>Legend</i>.</i></p> <p><i>Змінити кольори даних на золотий та оранжевий та увімкнути мітки даних. Змініть мітки даних на чорні.</i></p> <p><i>Створити візуальний елемент <i>Treemap</i> у панелі інструментів. Додати <i>NumberofCases</i> до <i>Values</i> та <i>Neighborhood</i> до <i>Group</i>, <i>AveDaysOpen</i> до <i>Tooltips</i>.</i></p> <p><i>На новій сторінці додати лінійну діаграму, що показує дату закінчення тижня (<i>WeekEndingDate</i>) на осі та кількість випадків (<i>NumberofCases</i>) як значення.</i></p> <p><i>Провести візуалізацію з використанням елементу <i>Slicer</i> із панелі інструментів <i>Visualizations</i>. Додати <i>WeekEndingDate</i> в область <i>Fields</i>. Використати <i>Slicer</i> для регулювання діапазону дат на графіку.</i></p>	<p align="center">2</p>

	<p style="text-align: center;">Лабораторне заняття 14 Публікація аналітичних звітів</p> <p>Мета: публікація звіту на порталі <i>PowerBIService</i>, створення інформаційних панелей на порталі та встановлення графіка оновлення даних.</p> <p>Завдання: Завантажити файл, що містить дані про продажі від торгових компаній <i>Nordwind</i>. Увійти до служби <i>PowerBIService</i>. Обрати параметр <i>Publish</i> і вибрати в якості пункту призначення <i>MyWorkspace</i>. У вікні «<i>Pin</i>to dashboard» вибрати нову інформаційну панель і назвати її «<i>NorthwindSales</i>». Додати кількість замовлень за категоріями/товарами на інформаційну панель «<i>NorthwindSales</i>».</p> <p>Розгорнути панель навігації та вибрати інформаційну панель «<i>NorthwindSales</i>». Задати декілька питань в категорії «<i>Ask a question</i>». Закріпити цю таблицю на приладовій панелі.</p> <p>Вибрати «<i>Edit details</i>». Змінити заголовок і поставити відмітку в «<i>Display last refresh time</i>». Додати плитку з текстовим полем. Перестроювати і змінити розміри плиток.</p> <p>Налаштувати графік оновлення з використанням параметру «<i>SCHEDULE REFRESH</i>» в меню.</p> <p>У вікні, що з'явилося, розгорнути вузол облікових даних джерела даних та натиснути посилання «<i>Edit credentials</i>». Переконайтеся, що для методу аутентифікації встановлено значення «<i>Anonymous</i>», а рівень конфіденційності встановлено на «<i>Public</i>» перед входом. Створіть графік оновлення, який працює щодня о 7:00 год.</p>	2
	<p>Самостійна робота: опрацювання матеріалу лекцій, підготовка до лабораторних занять з теми «Візуалізація аналітичних звітів», ознайомлення з основною та додатковою літературою, інтернет-ресурсами, виконання індивідуальних завдань з використанням комп'ютера, підготовка до модульної контрольної роботи.</p> <p><i>Індивідуальна робота:</i></p> <p>Обрати веб-ресурси порталу відкритих даних https://tradingeconomics.com набір економічних показників України: https://tradingeconomics.com/ukraine/indicators.</p> <p>Завантажити набір даних до платформи <i>PowerBI</i> за прямим доступом до відповідної веб-сторінки. Побудувати звіт з використанням різних типів стовпчикових та лінійних діаграм, секторних, кільцевих та бульбашкових діаграм, карток та багаторядкових карток, карт. Здійснити візуальну фільтрацію даних за трьома обраними показниками. Опублікувати аналітичний звіт на платформі <i>PowerBIService</i>.</p> <p style="text-align: center;">Список рекомендованих джерел:</p> <p><i>Основний:</i> 1-3 <i>Додатковий:</i> 11, 12, 14. <i>Інтернет-ресурси:</i> 28.</p>	24
Разом		180

3. СПИСОК РЕКОМЕНДОВАНИХ ДЖЕРЕЛ

Основний

1. Cuesta H., Kumar S. *PracticalDataAnalysis*. Birmingham : PacktPublishingLtd, 2016. 316 p.
2. *DataScience&BigDataAnalytics: Discovering, Analyzing, VisualizingandPresentingData* /EMC EducationServices. Indianapolis : JohnWiley&Sons, Inc, 2015. 432 p.
3. Microsoft *Power BI Cookbook: CreatingBusinessIntelligenceSolutionsofAnalyticalDataModels, Reports, andDashboards*. Birmingham : PacktPublishingLtd, 2017. 802 p.
4. *Roskladka A., Ivanova O., Kulazhenko V. Data Scientist: a glanceintothefuture // Зовнішня торгівля: економіка, фінанси, право*. 2019. № 3. С. 109-120

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. *Adamson C. MasteringDataWarehouseAggregates: SolutionsforStarSchemaPerformance*. WileyPublishingInc., 2006. 318 p.
9. *Albright S. C., Winston W., Zappe C. DataAnalysisandDecisionMaking*. Boston : CengageLearning, 2016. 948 p.
10. *Cao L., Yu P. S., ZhangC., Zhang H. DataMiningforBusinessApplications*. SpringerScience; BusinessMedia, 2008. 402 p.
11. *Coodley M. O. Introductionto Microsoft Power BI: bringyourdatatolife! CreateSpaceIndependentPublishingPlatform*, 2016. 128 p.
12. *Etaati L. AdvanceAnalyticswithPower BI and R*. Auckland : RadacadSystemsLimited, 2017. 179 p.
13. *Fabrice G., Hamilton N. J. QualityMeasuresinDataMining*. 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. DataMining: ConceptsandTechniques*. MorganKaufmannPublishers, 2006. 800 p.
16. *LinoffG. S. DataAnalysisUsing SQL and Excel*. Indianapolis: Wiley, 2015. 792 p.
17. *LinoffG. S.,Berry M. J. A. DataMiningTechniques: ForMarketing, Sales, andCustomerRelationshipManagement*. Indianapolis: Wiley, 2011. 888 p.
18. *Milton M. HeadFirstDataAnalysis: A learner'sguidetobignumbers, statistics, andgooddecisions*. Sebastopol: O'ReillyMedia, 2009. 435 p.
19. *RapidMiner: DataMiningUseCasesandBusinessAnalyticsApplications* / EditedbyMarkusHofmann&RalfKlinkenberg. 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. Ітеративні, еволюційні та мультиагентні методи синтезу нечіткологічних і нейромережних моделей : монографія. Запоріжжя : ЗНТУ, 2009. 375 с.
24. Рідкокаша А. А., Голдер К. К. Основы систем штучного інтелекту : навч. посібн. Черкаси : ВІДЛУННЯ-ПЛЮС, 2002. 240 с.
25. Руденко О. Г., Бодяньський 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.

Інтернет-ресурси

28. Microsoft Power BI Guided Learning URL: <https://docs.microsoft.com/uk-ua/power-bi/guided-learning> (дата звернення 02.10.2020).
29. Офіційний сайт платформи Loginom. URL: <https://loginom.com/>
30. Rapid Miner Academy: [learn data science and Rapid Miner from leading industry experts](https://academy.rapidminer.com/). URL: <https://academy.rapidminer.com/> (дата звернення 02.10.2020).

* Курсивом виділені джерела, що є у бібліотеці КНТЕУ