

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ОДЕСЬКИЙ ДЕРЖАВНИЙ ЕКОЛОГІЧНИЙ УНІВЕРСИТЕТ

МЕТОДИЧНІ ВКАЗІВКИ
до лабораторних робіт з навчальної дисципліни
«Методи і системи підтримки рішень»
для студентів денної та заочної форми навчання
спеціальності _____ «_____»

Затверджено
на засіданні групи забезпечення спеціальності
Протокол № _____ від «___» _____ 2020р.
Голова групи _____ П.І.Б.

Затверджено
на засіданні кафедри інформатики
Протокол № _____ від «___» _____ 2020р.
Завідувач кафедри _____ Мещеряков В.І.

Одеса 2020

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ОДЕСЬКИЙ ДЕРЖАВНИЙ ЕКОЛОГІЧНИЙ УНІВЕРСИТЕТ

МЕТОДИЧНІ ВКАЗІВКИ
до лабораторних робіт з навчальної дисципліни
„Методи і системи підтримки рішень”
для магістрів денної та заочної форми навчання
спеціальності _____ «_____»

Затверджено
на засіданні групи
забезпечення спеціальності
Протокол № _____
від «_____» _____ 2020р.

Одеса – 2020

Методичні вказівки до лабораторних робіт з дисципліни „Методи і системи підтримки рішень” для магістрів першого року навчання денної та заочної форми за спеціальністю 122 «Комп’ютерні науки», рівень вищої освіти _____./

Укладач: к.ф.-м.н., Настасюк В.А. – Кафедра: Інформатики Одеського державного екологічного університету, Одеса, ОДЕКУ, 2020.

ВСТУП

Методичні вказівки до лабораторних робіт з дисципліни „Методи і системи підтримки рішень” призначені для магістрів першого року навчання факультету комп’ютерних наук ОДЕКУ. Їхньою метою є організація роботи студента за комп’ютером в процесі розв’язання практичних задач з навчальної дисципліни.

Потрібний для роботи теоретичний матеріал наведено у преамбулі до кожної лабораторної роботи. (У повному обсязі його викладено на лекціях). Студентові належить, попередньо ознайомившись з ним, виконати пропонувану послідовність дій для досягнення потрібного результату і відповісти на контрольні запитання.

Загальні поради до виконання лабораторних робіт:

- зміст кожної теми дисципліни вивчати до початку виконання лабораторної роботи;
- захищати кожну роботу окремо і відразу після завершення.

З кожної лабораторної роботи студент повинен скласти звіт, включивши в нього:

- назву роботи; мету;
- умову завдання;
- хід виконання роботи;
- відповіді на контрольні питання.

Оформлений звіт студент захищає усно.

ПРАВИЛА ТЕХНІКИ БЕЗПЕКИ ТА ОХОРОНА ПРАЦІ

Згідно з «Правилами техніки безпеки в лабораторіях інформатики» студентам заборонено:

- з'являтися та знаходитись приміщенні в нетверезому стані;
- ставити поруч з клавіатурою ємності з рідиною;
- перебувати в приміщенні в верхньому одязі та захаращувати робочі столи та стільці;
- працювати в лабораторії більше 6-ти годин на день (для вагітних жінок – більше 4-х годин);
- із власної ініціативи змінювати закріплені за ними робочі місця та перебувати в приміщенні під час роботи іншої навчальної групи;
- самостійно виконувати вмикання електроживлення лабораторії та заміну складових частин ПК, що вийшли із ладу.
-

У разі виявлення несправностей обчислювальної техніки студент повинен сповістити про це викладача або будь-кого з навчально-допоміжного персоналу лабораторії.

ДОСЛІДЖЕННЯ ПОЧАТКОВОГО ЕТАПУ ПРОЦЕСУ ВИРІШУВАННЯ

Мета: отримання навичок формування та оцінки альтернативних рішень поставленої проблеми; знайомство з програмою *Decision Explorer*: побудовою і аналізом мапи проблеми

Порядок виконання роботи

1. У середовищі *Decision Explorer* на основі запропонованих даних побудувати мапу проблеми з основними зв'язками причин і наслідків. Створити власний стиль оформлення моделі. Отриману мапу занести до звіту.
2. Виконати аналіз проблемної мапи, у тому числі:
 - виявити 2-3 ключові елементи моделі;
 - побудувати ієрархію причин для одного з цих елементів;
 - виявити петлі позитивного (негативного?) зворотного зв'язку;
 - зробити висновки про те, як може бути використана отримана інформація
 - розробити варіанти розв'язання проблеми, тобто сформулювати альтернативи
3. Занести результати до звіту

Загальні відомості

Серед систем підтримки рішень існують такі, що застосовні на початкових етапах розробки керівницького рішення. Однією з таких систем є *Decision Explorer* (Banxia Software Ltd) використовується на стадії первісного дослідження проблеми та формування альтернативних рішень. *Decision Explorer* являє собою інструмент побудови та аналізу дескриптивних (описових) нематематичних моделей, які включають у себе набір елементів досліджуваної системи (проблеми) та зв'язки між ними. Прикладами дескриптивних моделей можуть бути, зокрема, структурна модель предметної галузі, діаграма причинно-наслідкових зв'язків, дерево цілей.

Основні функції *Decision Explorer*: обробка вербальних відомостей про предметну галузь, аналіз структури досліджуваної системи, а також активізація та організація мислення особи яка опрацьовує рішення.

Програма використовує методику концептуальних мап (Mind Maps) – спосіб наочного висвітлення наявних проблем і створення нових ідей. Творець цього методу Т.Бузан описує його такими словами: "...уявна мапа будується на підставі центрального слова або концепції, навколо яких розташовуються від 5 до 10 головних ідей, що мають стосунок до нього. Кожне з цих дочірніх слів знову-таки оточують 5-10 головними ідеями". Отримана діаграма складається з вузлів (nodes) або комірок (cells), які містять поняття й питання (concept), й стрілок-посилань. Напрямок стрілки показує причинно-

наслідковий зв'язок і може бути прочитаний як пропозиція. Такий спосіб зображення дозволяє по-новому розглядати центральну концепцію, наново пов'язувати її з наявним досвідом. Діаграмою зручно користуватися, оскільки всі відомості структуровані по кластерах. Це робить мапу корисною в сесіях мережного мозкового штурму і, як вважає розробник, породжує асоціації.

1 Побудова проблемних мап у програмі Decision Explorer



Робота з *Decision Explorer* включає дві основні стадії: *синтез* та *аналіз*.

На першій стадії формується візуальна модель, мапа проблемної ситуації, яка відображає сукупність елементів досліджуваної системи та зв'язки між ними. На другій стадії отриману мапу піддають аналізу за допомогою спеціалізованого інструментарію *Decision Explorer*. Надалі фахівці обговорюють отриманий результат і коректують початкову модель.

Елементи моделі – причини та наслідки проблеми – зображуються прямокутниками або іншими фігурами з відповідними написами. Взаємозв'язки між елементами моделі передаються стрілками.

Подвійне клацання миші по порожньому полю створює елемент моделі; операцією Drag & Drop наводяться зв'язки (стрілки) між елементами.




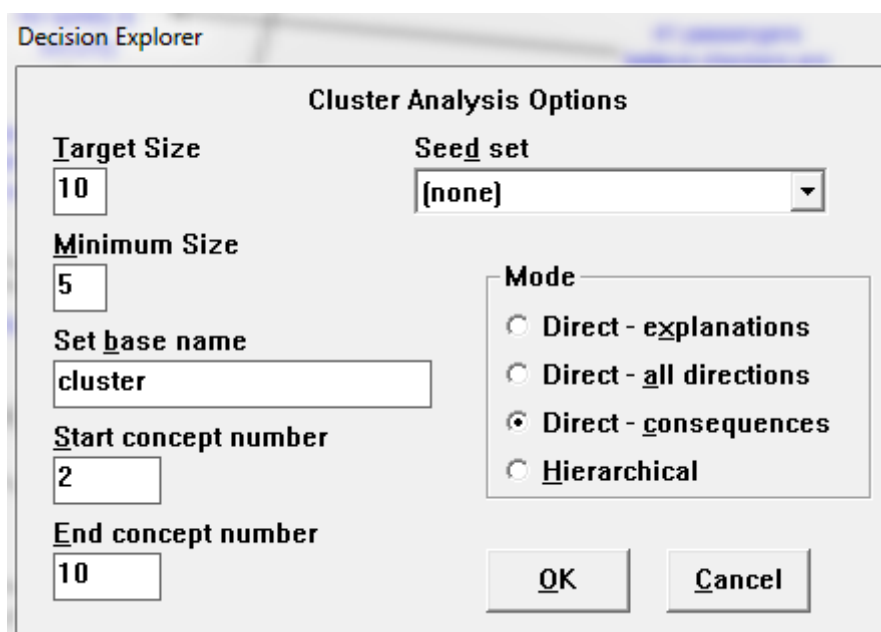
Для вибору стилю оформлення елементів моделі слугує команда **Property** → **Show style selector**, а також кнопка  на панелі інструментів. Створити свій стиль оформлення можна за допомогою команди **Property** → **Concept style properties** і відповідної кнопки **Define concept style properties**  на панелі інструментів. Ця команда відкриває діалогове вікно для визначення власного стилю оформлення. Працюючи зі стилями, слід приділити увагу можливостям кнопки “FONT” діалогового вікна **Concept style properties**. Ця кнопка дозволяє встановити не лише шрифт та його оформлення, але й той чи інший набір символів. Для правильного відображення написів на об'єктах моделі потрібно встановити набір символів кирилиці.

Масштаб зображення на екрані задається командою **View** → **Display scale** (існують і “гарячі” клавіші). Вибрати кілька елементів легко, обводячи їх мишею або клацаючи по них за натиснутої клавіші <Shift>.

2 Аналіз

На стадії аналізу отриману раніше візуальну модель досліджують за допомогою таких процедур:

- командою **Central** і **Domain** пункта меню **Analysis** визначаються “ключові” елементи моделі – тобто елементи, які мають найбільшу кількість зв’язків. Отримана інформація корисна тим, що надає можливість сформулювати основні завдання щодо виявлених елементів. (Повернення до попереднього вигляду командою **View** → **Map display** або кнопкою ); застосування команди **Central n** (вводиться з клавіатури), де $n = 2, 3, \dots$ - кількість ключових елементів дозволяє виявити довільно визначену кількість елементів;
- команда **Analysis** → **Cluster** виконує кластерний аналіз, тобто виявляє групи елементів (без перекриття), пов’язаних одне з одним більшою кількістю зв’язків ніж з елементами інших груп. Команда **Analysis** → **Cluster Options** породжує діалогове вікно



яке дозволяє сконфігурувати процедуру кластеризації. Вибір опції **Seed set – (none)** забезпечує автоматичний вибір кількості кластерів. У віконцях **Start concept number**, **End concept number** слід вказати діапазон (пронумерованих програмою) елементів, над якими проводиться кластерний аналіз. **Target size** дозволяє встановити максимальний розмір кластера 10 – 99 елементів (за замовчуванням 45) та мінімальний (8 за замовчуванням).

Вибір типів кластеризації доступний у віконці **Mode**. Наприклад, ієрархічна (**Hierarchical**) кластеризація використовує заголовки моделі та точки розгалуження як насіння для кластера, і намагається створити мапу ієрархічно пов’язаних кластерів відповідно до заданого розміру. При цьому деякі елементи можуть випасти з розгляду в такому режимі.

Кластеризація на основі подібності зв’язків між елементами (з автоматичним вибором параметрів) може також бути проведена командою **CLUSTER** з клавіатури.

- команда **Map concept** будує ієрархію “підлеглих” елементів для виділеної ланки моделі;
- команда **Loop** виявляє петлі позитивного та негативного зворотних зв’язків між елементами моделі.

Ці та інші команди вводяться з клавіатури в командний рядок або вибираються в списку меню **Analysis**. Командний рядок **DE>** з’являється внизу екрана автоматично за спроби ввести текст з клавіатури.

Постановка задачі

У пропонованій лабораторній роботі потрібно з’ясувати можливості *Decision Explorer* у розв’язанні задачі виведення уявного підприємства з кризи.

Об’єктом дослідження буде підприємство “Stylus” – компанія середнього масштабу, яка виготовляє електродвигуни та помпи, а також споживчі товари. Останнім часом компанія мала високу репутацію завдяки високій якості продукції та сервісу. Проте ситуація змінилася.

По-перше, різко скоротилася кількість замовлень на основну продукцію підприємства – електродвигуни для помпових станцій. При цьому зменшення платоспроможного попиту супроводжувалося підвищенням вимог замовників до характеристик продукції.

По-друге, підприємство змінило свій статус з державного на приватне. З цього часу розвиток виробництва потрібно було здійснювати вже не за рахунок централізованих джерел, а за рахунок самостійно зароблених коштів. Як наслідок – упродовж останніх років компанія звітує про збитки. Про стан підприємства свідчать такі цифри: половина основних фондів експлуатується понад 25 років; номенклатура продукції заводу оновлюється щорічно тільки на 1-2 %; рівень витрат на 1 грн. продукції становить 90 коп. при плані 75 коп.

У поточному році керівництво підприємства розпочало формування антикризової програми розвитку виробництва та збуту. Була скликана екстрена нарада з таким питанням на порядку денному: “Які чинники призвели до наявної ситуації?” Іншими словами, керівництво розпочало перший етап процесу розробки рішення – дослідження проблемної ситуації, виявлення її причин та наслідків.

Експерти та аналітики визначили пари взаємопов’язаних чинників (табл. далі).

Чинник, який впливає	Чинник, на який впливає
Великі витрати матеріалів та сировини	Висока собівартість продукції
Висока собівартість продукції	Недостатність власних коштів
Висока собівартість продукції	Низький попит на продукцію
Значна частка пасивної частини ОФ	Висока собівартість продукції
Мала частка ТНВ в обсязі продукції	Низький попит на продукцію
Недостатній аналіз ринку	Низький попит на продукцію
Недостатній аналіз ринку	Мала частка ТНВ в обсязі продукції
Недостатність власних коштів	Низька оновлюваність продукції
Недостатність власних коштів	Старіння оборотних фондів (ОФ)
Низька виконавча дисципліна	Висока собівартість продукції
Низька оновлюваність продукції	Низький попит на продукцію
Низький обсяг виробництва	Недостатність власних коштів
Низький попит на продукцію	Низький обсяг виробництва
Низький рівень стратегіч. планування	Низька виконавча дисципліна
Низький рівень стратегіч. планування	Слабке просування продукції на ринку
Низький рівень стратегіч. планування	Мала частка ТНВ в обсязі продукції
Низький рівень стратегіч/ планування	Недостатній аналіз ринку
Слабке просування продукції на ринку	Низький попит на продукцію
Старіння оборотних фондів (ОФ)	Висока собівартість продукції
Старіння оборотних фондів (ОФ)	Великі витрати матеріалів і сировини
Старіння оборотних фондів (ОФ)	Низька оновлюваність продукції

Хід роботи

Дія	Як зробити
-----	------------

За наведеною таблицею
будуємо мапу проблеми

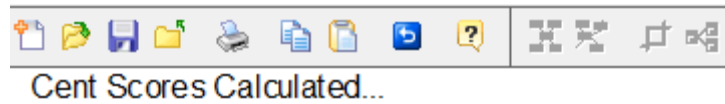
Аркуш з візуальною моделлю захищаємо від змін **View → ProtectView**

Дія	Як зробити	Примітка
Виявити та занести до звіту 3 ключові ланки (вузли) системи	> Central 3	Команда супроводжується оголошенням, що View 1 (створена мапа)

причинно-наслідкових зв'язків і на їх основі сформулювати 3 основні завдання антикризового керування ВАТ “Stylus”

захищений від змін, у діалоговому вікні слід вибрати View 2.

На екран буде виведено:



Top 3 concepts in descending order.

10 низький попит

9 from 13 concepts

4 велика собівартість

9 from 13 concepts

9 погане стратегіч. планування

7 from 13 concepts

На основі цієї інформації формуємо основні задачі антикризового керування і занесимо до звіту з лабораторної роботи.

Якщо інтуїтивна уява дослідника проблеми не співпала з результатом застосування команди **Central**, це означає, що побудована мапа потребує поточнення – доповнення новими причинними зв'язками.

Дія	Як зробити	Примітка
Визначити і занести до звіту 2 відносно незалежні підсистеми чинників та їхні елементи, які потурають кризовому становищу підприємства	У командному рядку набрати “Cluster”	Результат зберігається окремо від основної мапи; його можна вивести на екран командою “MAP Cluster1” в командному рядку. Утворені кластери - не єдино можливі, а аналіз - це лише пропозиція про те, як можна поділити модель

На аркушах **View 3**, **View 4** з'являться схеми - кластери, які складаються з 7 та 6 елементів. Вони відповідають двом системам антикризового керування, які можна назвати:

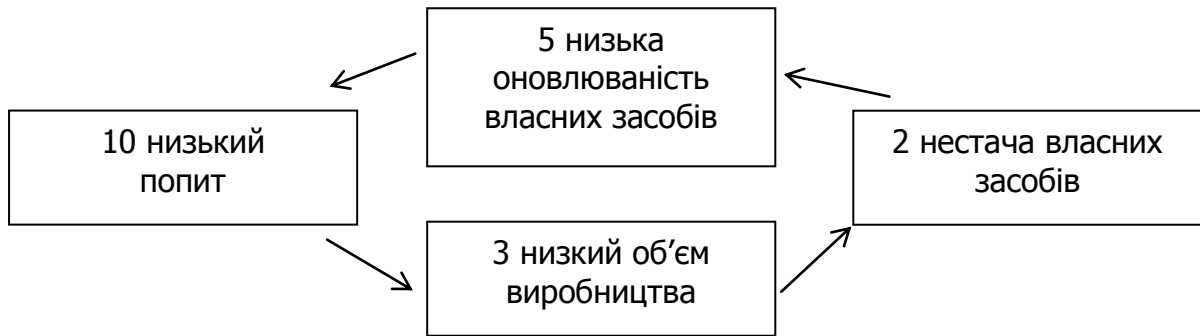
- 1) підсистема вдосконалення керування;
- 2) підсистема вдосконалення техніки та технологій.

Дія	Як зробити	Примітка
Виявити петлі	> Loop	

зворотнього зв'язку
Перевірити наявність
більшої кількості
петель

> **map loop1**,
> **map loop2** ... тощо
в залежності від
кількості знайдених
петель

Результат виконання операції **Loop**:



Дія	Як зробити	Примітка
З'ясувати причини нестачі власних фінансових коштів - побудувати ієрархію причин для відповідного елемента	<p>Застосувати команду Analysis → Map concept до елемента “Нестача власних коштів” (виділити його попередньо мишкою);</p> <p>У виниклому діалоговому вікні Map Options задати такі параметри ієрархії :</p> <ul style="list-style-type: none"> - на вкладці Map type задати тип ієрархії – Tree; Selected method - Use selected concept <p>У Tree options задати: inward links (вхідні зв'язки)</p>	<p>Якщо вікно Map Options не з'явилося, його можна викликати за допомогою команди Analysis → Mapping options</p>

Занести послідовно отримані результати до звіту. Сформулювати висновки

Деякі корисні команди Decision Explorer

У командному рядку:

cls – очистити аркуш

mapall – отображення на листе всей визуальной модели (карты)

map ... – отображення на листе части визуальной модели, в т.ч.:

map loop1 – отображення первой петли (команда доступна после выполнения команды **loop**)

Команды меню:

View → **Protect View** – защита листа от изменений.

View → **Display scale** → **Select Zoom** – выбор масштаба изображения

Питання до захисту роботи

1. Яке призначення систем підтримки рішень?
2. У чому відмінність СПР від решти інформаційних систем?
3. Які класи задач можливо розв'язати за допомогою СПР?
4. Визначити основні критерії класифікації СПР.
5. Назвіть основні концептуальні засади класифікації.

ЛАБОРАТОРНА РОБОТА №2

Ухвалення рішень при багатьох критеріях за допомогою методу аналізу ієрархій

Мета: Набути навичок розв'язання багатокритеріальних задач із використанням пакету MS Excel.

Порядок виконання роботи

Завдання 1: визначити найкращу альтернативу при багатьох критеріях за допомогою МАІ для наведеного прикладу

Завдання 2 – індивідуальне. Послідовність роботи аналогічна завданню 1.

Скласти **звіту** з лабораторної роботи з

- формулюванням індивідуального завдання;
- ієрархією "цілі-критерії-альтернативи";
- світлинами екрану монітора, матрицями порівнянь критеріїв і альтернатив, що містять, обчислення векторів пріоритетів, перевірку узгодженості і визначення найкращої альтернативи;
- висновками із завдання.

Теоретичні відомості

Метод аналітичної ієрархії (МАІ, Analytic Hierarchy Process – АНП) використовують у задачах пошуку найкращого рішення, де альтернативи і критерії (атрибути) відомі заздалегідь. Він зручніший за аксіоматичні методи коли кількість альтернатив мала.

Підхід МАІ складається з послідовності кроків.

1. Перший етап – зображення задачі у вигляді ієрархії: мета (цілі) – критерії – альтернативи.

2. На другому етапі виконують попарні порівняння елементів кожного рівня. Результати порівнянь переводять в числа за допомогою спеціальної таблиці

Рейтинг	Опис
1	Однакова перевага
3	Помірна перевага
5	Явна перевага
7	Очевидна перевага
9	Виняткова перевага

Також можна присвоювати значення рейтингу 2, 4, 6 і 8, як середні від найближчих рейтингів.

3. Розраховують коефіцієнти важливості для коефіцієнтів кожного рівня. При цьому перевіряється узгодженість суджень ОВ.

4. Підраховують кількісний індикатор якості кожної з альтернатив і визначають найкращу.

Математичні основи МАІ ґрунтуються на теорії матриць. Матриці попарних порівнянь, що використовуються в методі, є додатними обернено-симетричними незвідними.

Додатними обернено-симетричними називаються квадратні матриці $A = (a_{ij})$, для яких $a_{ij} > 0$ і $a_{ij} = 1/a_{ji}$. Квадратна матриця незвідна, якщо її не можна при-

вести до вигляду $\begin{bmatrix} A_1 & 0 \\ A_2 & A_3 \end{bmatrix}$, де A_1 і A_3 – квадратні матриці, 0 – нульова матриця.

У МАІ оцінками (пріоритетами) альтернатив, отриманих з матриці їхніх парних порівнянь, слугують нормалізовані значення *головного власного вектора* матриці. Вектори, власні для матриці A , отримують з рівняння $Ax = \lambda x$. Відповідні їм числа λ називають *характеристичними* або *власними* до матриці A , вони є коренями *характеристичного рівняння* $|A - \lambda I| = 0$ (I – одинична матриця).

За теоремою Фробеніуса-Перрона незвідна додатна матриця завжди має дійсне додатне некрратне власне число λ_{max} , більше від решти власних чисел. *Головному власному числу* λ_{max} і відповідає *головний власний вектор*.

Таким чином для отримання оцінок альтернатив, тобто головного вектора матриці порівнянь, потрібно розв'язати характеристичне рівняння відносно λ ; виділити з коренів найбільший, λ_{max} ; з матричного рівняння знайти відповідний йому власний вектор; значення вектора нормалізувати.

Наведений алгоритм є доволі громіздким. У МАІ використовується спрощена процедура пошуку власного вектора, яка ґрунтується на наслідковій теоремі Фробеніуса-Перрона про *примітивні* матриці і не потребує розв'язання характеристичного рівняння.

Матриця з означеними вище характеристиками називається *примітивною*, якщо її головне власне число *єдине* (в іншому разі існують характеристичні числа матриці, що співпадають за модулем з головним числом).

Вважаючи матриці попарних порівнянь примітивними, застосовують спрощений алгоритм. Як показано в наступному [прикладі](#), значення елементів матриці спочатку нормалізують на суми стовпців, тоді елементами головного вектора, а, отже, і оцінками альтернатив, будуть середні значення по рядках.

Описаний підхід застосовується до всіх елементів ієрархії. Наприкінці процедури оцінюють узгодженість усієї моделі вподобань.

Постановка задачі

Компанії Sleepwell Hotels потрібно вибрати найкращий пакет бухгалтерського програмного забезпечення із запропонованих кількома постачальниками. Це завдання було доручене завідувачу відділом. Той виділив трьох постачальників, пропонуване програмне забезпечення яких зможе задовольнити основні потреби компанії: Revenue Technology Corporation (RTC), PRAISE Strategic Solutions (PSS) і El Cheapo (EC) і означив **критерії**, важливі на його думку, у виборі програмного забезпечення: 1) загальна вартість програмної системи, 2) забезпечення обслуговування протягом наступного року, 3) складність і надійність основних математичних процедур і 4) можливість адаптації системи під умови Sleepwell.

Хід роботи

1. Створюємо аркуш “**Вартість**” у робочій книжці Excel **ПЗ.xlsx** для аналізу першого критерію (**загальна вартість**) і заносимо на аркуш дані як на **світлині 1**. Таблицю слід читати таким чином: вказаний в рядку постачальник порівнюється з постачальником із стовпця. Якщо перший з них прийнятніший, то відповідне число від 3 до 9 записується в комірку на перетині рядка і стовпця. У симетричну відносно діагоналі матриці комірку вписується обернене число. Очевидно, що оскільки будь-який постачальник рівнозначний самому собі, в усіх діагональних комірках стоятимуть значення 1.
2. За показником загальної вартости першому постачальникові надано середню між помірною і явною перевагою порівняно з другим постачальником, тому в комірку другого стовпця першого рядка (**C4**) заноситься число 4. Постачальникові 3 (EC) надано перевагу від однакового до помірного перед постачальником 1 (RTC), тому в комірку 3-го стовпця першого рядка записано число 1/2 (комірка **D4**).

	A	B	C	D
2				
3		RTC	PSS	EC
4	RTC		1	4
5	PSS		0,25	1
6	EC		2	7

	A	B	C	D
2				
3		RTC	PSS	EC
4	RTC	1	4	0,5
5	PSS	=1/C4	1	=1/7
6	EC	=1/D4	=1/D5	1

Св. 1. Попарне порівняння за показником вартости

3. Нормалізацію матриці виконують **сумуванням чисел в кожному стовпці** і подальшим **діленням кожного елемента стовпця на свою суму**. Результати цієї операції показано в комірках **B12:D14** на **світлині 2**.

	A	B	C	D	E
3		RTC	PSS	EC	
4	RTC	1	4	0,5	
5	PSS	0,25	1	0,143	
6	EC	2	7	1	
7					
8	СУМА	3,25	12	1,643	
9					
10	НОРМАЛІЗАЦІЯ				
11		RTC	PSS	EC	Середнє
12	RTC	0,308	0,333	0,304	0,315
13	PSS	0,077	0,083	0,087	0,082
14	EC	0,615	0,583	0,609	0,602

	A	B	C	D	E
3		RTC	PSS	EC	
4	RTC	1	4	=1/B6	
5	PSS	=1/C4	1	=1/C6	
6	EC	2	7	1	
7					
8	СУМА	=СУММ(B4:B7)	=СУММ(C4:C7)	=СУММ(D4:D7)	
9					
10	НОРМАЛІЗАЦІЯ				
11		RTC	PSS	EC	Середнє
12	RTC	=B4/B8	=C4/C8	=D4/D8	=СРЗНАЧ(B12:D12)
13	PSS	=B5/B8	=C5/C8	=D5/D8	=СРЗНАЧ(B13:D13)
14	EC	=B6/B8	=C6/C8	=D6/D8	=СРЗНАЧ(B14:D14)

Св. 2. Нормалізована матриця для критерію загальної вартости

4. Наступний крок полягає в обчисленні **балу для кожного продавця** за критерієм загальної вартости. Ці значення показані на **св. 2** в стовпці **Е**. Видно, що найвищий середній бал за цим критерієм має постачальник **ЕС**.

5. Тепер потрібно обчислити **коефіцієнт узгоджености** і перевірити його значення. Цю операцію виконують, щоб пересвідчитися в узгоджености заданих вподобань з початкової таблиці. Наприклад, якщо за критерієм загальної вартости задана явна перевага постачальника 1 перед постачальником 2 і помірна перевага постачальника 2 в порівнянні з постачальником 3, то постачальників 1 і 3 уже не можна вважати рівнозначними, бо це призведе до неузгоджености. Ще більша неузгодженість виникне, якщо задати, аби 3 було краще.

Обчислення коефіцієнта узгоджености складається з трьох етапів.

1. Обчислюють міру узгоджености для кожного постачальника.
2. Визначають індекс узгоджености ІС.

3. Обчислюють коефіцієнт узгоджености як відношення IS/IP , де IP - індекс рандомізації.

Як показано на **св. 3**, для першого постачальника (RTC) кількість балів у рядку **B4:D4** множиться на середній рейтинг кожного постачальника (комірки **E12:E14**) і результат ділиться на середній рейтинг першого постачальника (комірка **E12**).

Аналогічні обчислення здійснюються для 2 і 3 постачальників. У ідеальному випадку міри узгоджености мають дорівнювати числу можливих альтернативних рішень (у нашому випадку є 3 рішення, тобто 3 постачальники). Для обчислення індексу узгоджености визначають середню міру узгоджености всіх трьох постачальників, з неї віднімається кількість можливих варіантів рішення n і результат ділиться на $n - 1$. Індекс узгоджености IS показаний на **св. 3** в комірці **F16**, його значення = 0,001.

Останній етап визначення **коефіцієнта узгоджености** полягає в діленні IS на індекс рандомізації IP , значення якого для різних значень n обчислюються в методі МАІ спеціальним чином і наведені в таблиці нижче

n	Індекс рандомізації
2	0,00
3	0,58
4	0,90
5	1,12
6	1,24
7	1,32
8	1,41
9	1,45
10	1,51

Коефіцієнт узгоджености записаний в комірці **F20** і дорівнює 0,002.

10	НОРМАЛІЗАЦІЯ					
11		RTC	PSS	EC	<i>Середнє</i>	<i>Міра узгоджености</i>
12	RTC	0,308	0,333	0,304	0,315	3,002
13	PSS	0,077	0,083	0,087	0,082	3,000
14	EC	0,615	0,583	0,609	0,602	3,004
15						
16	IU =					0,001
17						
18	IP =					0,58
19						
20	Коефіцієнт узгоджености=					0,002

10	НОРМАЛІЗАЦІЯ					
11		RTC	PSS	EC	Середнє	Міра узгоджености
12	RTC	=B4/B8	=C4/C8	=D4/D8	=CPЗНАЧ(B12:D12)	=МУМНОЖ(B4:D4;E12:E14)/E12
13	PSS	=B5/B8	=C5/C8	=D5/D8	=CPЗНАЧ(B13:D13)	=МУМНОЖ(B5:D5;E12:E14)/E13
14	EC	=B6/B8	=C6/C8	=D6/D8	=CPЗНАЧ(B14:D14)	=МУМНОЖ(B6:D6;E12:E14)/E14
15						
16	IY =					=(CPЗНАЧ(F12:F14)-3)/2
17						
18	IP =					0,58
19						
20	Коефіцієнт узгоджености=					=F16/F18

Св. 3. Коефіцієнт узгоджености для критерію загальної вартости

У разі абсолютної узгоджености переваг міра узгоджености дорівнюватиме 3, отже, ІС дорівнюватимуть нулю, і коефіцієнт узгоджености також дорівнюватиме нулю. Дуже великий ($> 0,10$) коефіцієнт, означає, що менеджер був недостатньо послідовний в своїх оцінках, тому слід повернутися назад і переглянути результати попарних порівнянь (в більшості випадків виявляється елементарна помилка, і коефіцієнт узгоджености сигналізує про її наявність).

6. Тепер потрібно виконати те ж саме для решти критеріїв. Для цього слід тричі скопіювати робочий аркуш "Вартість", створивши тим самим три нових робочих аркуші "Обслуговування", "Складність" і "Адаптація", а потім треба просто змінити параметри попарних порівнянь. Результати цих дій показані на св. 4-6.

	A	B	C	D	E	F
3		RTC	PSS	EC		
4	RTC	1	0,5	6		
5	PSS	2	1	8		
6	EC	0,167	0,125	1		
7						
8	СУМА	3,167	1,625	15		
9						
10	НОРМАЛІЗАЦІЯ					
11		RTC	PSS	EC	Середнє	Міра узгоджености
12	RTC	0,316	0,308	0,400	0,341	3,020
13	PSS	0,632	0,615	0,533	0,593	3,032
14	EC	0,053	0,077	0,067	0,065	3,003
15						
16	IY =					0,009
17						
18	IP =					0,58
19						
20	Коефіцієнт узгоджености=					0,016

Св. 4. Коэффициент узгоджености для критерію обслуговування

	A	B	C	D	E	F
3		RTC	PSS	EC		
4	RTC	1	1	5		
5	PSS	1	1	5		
6	EC	0,2	0,2	1		
7						
8	СУМА	2,200	2,2	11		
9						
10	НОРМАЛІЗАЦІЯ					
11		RTC	PSS	EC	Середнє	Міра узгоджености
12	RTC	0,455	0,455	0,455	0,455	3,000
13	PSS	0,455	0,455	0,455	0,455	3,000
14	EC	0,091	0,091	0,091	0,091	3,000
15						
16					IY =	0,000
17						
18					IP =	0,58
19						
20					Коефіцієнт узгоджености=	0,000

Св. 5. Коефіцієнт узгоджености за критерієм складности

	A	B	C	D	E	F
3		RTC	PSS	EC		
4	RTC	1	0,25	3		
5	PSS	4	1	6		
6	EC	0,3	0,167	1		
7						
8	СУМА	5,333	1,417	10		
9						
10	НОРМАЛІЗАЦІЯ					
11		RTC	PSS	EC	Середнє	Міра узгоджености
12	RTC	0,188	0,176	0,300	0,221	3,040
13	PSS	0,750	0,706	0,600	0,685	3,109
14	EC	0,063	0,118	0,100	0,093	3,013
15						
16					IY =	0,027
17						
18					IP =	0,58
19						
20					Коефіцієнт узгоджености=	0,047

Св. 6. Коефіцієнт узгоджености для критерію адаптації

У всіх випадках значення коефіцієнта узгоджености перебувають у межах від 0 до 0,047, це означає, що менеджер був досить послідовний в своїх оцінках. Крім того, можна помітити, що компанія PSS виявилася кращою за критерієм

обслуговування, RTC і PSS — кращі за критерієм складності, а PSS — більш адаптативна.

7. На цьому перший етап роботи закінчується. На другому етапі здійснюють попарні порівняння для визначення ваги критеріїв. Процес аналогічний попередньому в тому, що знову виконується порівняння, проте тепер *порівнюються не постачальники, а критерії*. Ці дії виконують на робочому аркуші “Вага” (св. 7).

	A	B	C	D	E	F	G
3		Вартість обслуг.	Обслуговув.	Складність	Адаптація		
4	Вартість	1	6	0,5	3		
5	Обслуговування	0,167	1	0,125	0,333		
6	Складність	2	8	1	5		
7	Адаптація	0,333	3	0,2	1		
8							
9	СУМА	3,500	18,000	1,825	9,333		
10							
11	НОРМАЛІЗАЦІЯ						
12		Вартість обслуг.	Обслуговув.	Складність	Адаптація	Середнє	Міра узгоджености
13	Вартість	0,286	0,333	0,274	0,321	0,304	4,071
14	Обслуговування	0,048	0,056	0,068	0,036	0,052	4,011
15	Складність	0,571	0,444	0,548	0,536	0,525	4,087
16	Адаптація	0,095	0,167	0,110	0,107	0,120	4,023
17							
18						IY =	0,016
19						IP =	0,9
20							
21						Коефіцієнт узгоджености=	0,018

Св. 7. Коефіцієнт узгоджености для ваги критеріїв

Виявилось, що показник складності і надійности математичних алгоритмів має найбільшу вагу (52,5% у вічку F14), за ним йде вартість (30,4% у вічку F12). Добре, що міри узгоджености виявилися близькими до 4, тому індекс узгоджености і коефіцієнт узгоджености близькі до нуля.

8. Останній крок полягає в обчисленні зважених середніх оцінок для кожного варіанту рішення і застосуванні отриманих результатів для ухвалення рішення про те, у якого постачальника буде куплено нове програмне забезпечення. Завершальні обчислення зроблені на аркуші “Порівняння” (св. 8).

	A	B	C	D	E	F
1						
2	Критерії	Ваги	RTC	PSS	EC	
3	Вартість	0,304	0,315	0,082	0,602	
4	Обслуговування	0,052	0,341	0,593	0,065	
5	Складність	0,525	0,455	0,455	0,091	
6	Адаптація	0,120	0,221	0,685	0,093	
7						
8	Зважені середні ретинги		0,379	0,377	0,245	
9						

	B	C	D	E
1				
2	Ваги	RTC	PSS	EC
3	=Ваги!F13	=Вартість!E12	=Вартість!E13	=Вартість!E14
4	=Ваги!F14	=Обслуговування!E12	=Обслуговування!E13	=Обслуговування!E14
5	=Ваги!F15	=Складність!E12	=Складність!E13	=Складність!E14
6	=Ваги!F16	=Адаптація!E12	=Адаптація!E13	=Адаптація!E14
7				
8		=СУММПРОИЗВ(\$B3:\$B6;C3:C6)	=СУММПРОИЗВ(\$B3:\$B6;D3:D6)	=СУММПРОИЗВ(\$B3:\$B6;E3:E6)

Св. 8. Зважене середнє рейтингів з використанням ваг

На підставі отриманих результатів можна зробити висновок, що компанія RTC (показник 0,378 в комірці **C8**) дещо перевершує компанію PSS (0,376 в комірці **D8**) а компанія EC від них помітно відстала.

Варіанти завдань 1

Варіант 1

Джек обирає університет, в якому б він хотів отримати вищу освіту. Він зупинився на двох з них: Гарварді і Стенфорді і визначив такі критерії вибору університету: плата за навчання, престиж університету, вартість життя і достоїнства місцини, де знаходиться університет.

Оплата в Гарварді трохи вища, ніж в Стенфорді. Престиж університетів приблизно однаковий. Вартість життя в Гарварді помітно дешевше, та зате достоїнства міста, де розташований Стенфорд, помітно вище. Достоїнства міста, де розташований університет для Джека трохи більше важливі, ніж вартість життя в ньому. У свою чергу, престиж університету трохи важливіший, в порівнянні з міською красою. А от розмір оплати важить значно більше навіть в порівнянні з престижем.

В який університет краще поступити Джеку? Чому дорівнюють середні рейтинги університетів за критерієм престижу? Чому дорівнюють середні ваги критеріїв?

Варіант 2

Потрібно вибрати один з двох варіантів, А і Б програмного забезпечення (ПЗ) для створення інтернет-магазину. Застосовують критерії відбору ПЗ:

1. Вартість.
2. Супровід розробників.
3. Зручність інтерфейсу.
4. Надані функції.

Супровід розробників (наприклад, безкоштовна тех. підтримка, навчання персоналу) при виборі ПЗ оцінюються як помітно важливіші в порівнянні з характеристиками користувацького інтерфейсу. Ще важливішим критерієм є функції (можливості) ПЗ. Але основним при ухваленні рішення все ж є вартість.

Припустимо, А - це дорога система з широким набором користувацьких функцій, зручним інтерфейсом, супроводжувана розробниками а система Б - проста і недорога розробка.

Купівля якого ПЗ буде прийнятнішою відповідно до вказаних критеріїв?

Варіант 3

Вирішивши купити автомобіль, людина звузила свій вибір до трьох моделей: Mercedes, Mitsubishi і Honda. Чинниками, що впливають на його рішення, є: вартість автомобіля (С), вартість обслуговування (О), вартість поїздки по місту (Г) і сільській місцевості (М). Наступна таблиця містить потрібні відомості з експлуатації автомобіля, що відповідають трирічному терміну.

Модель автомобіля	С (дол.)	О (дол.)	Г (дол.)	М (дол.)
<i>Mercedes</i>	62000	1800	4500	1500
<i>Mitsubishi</i>	35000	1200	2250	750
<i>Honda</i>	40000	600	1125	600

Найсуттєвішими критеріями при ухваленні рішення є вартість автомобіля і вартість його обслуговування. Поїздки по сільській місцевості здійснюються рідко порівняно з поїздками по місту.

Використовуйте вказані вартості для побудови матриць порівнянь. Оцініть узгодженість матриць і визначіть модель автомобіля, яку слід вибрати.

Варіант 4

Gert's Sports - мережа спортивних магазинів, що швидко розвивається на Східному узбережжі США. Власник мережі Боб Гертц накопичив чималий капітал, щоб відкрити нові крамниці в районі Чикаго. Для забезпечення крамниць потрібно буде розширити склади. За підтримкою Гертц може звернутися до послуг однієї з трьох фінансових компаній, у кожній з них є свої переваги щодо надання кредиту і обслуговування клієнтів. Боб оцінив рейтинги цих компаній тако:

Рейтинги за умовами кредиту

	Big Bank	Little Bank	US Bucks
Big Bank	1	2	0,143
Little Bank	0,5	1	6
US Bucks	7	0,167	1

Рейтинги з обслуговування клієнтів

	Big Bank	Little Bank	US Bucks
Big Bank	1	0,25	1
Little Bank	4	1	0,5
US Bucks	1	2	1

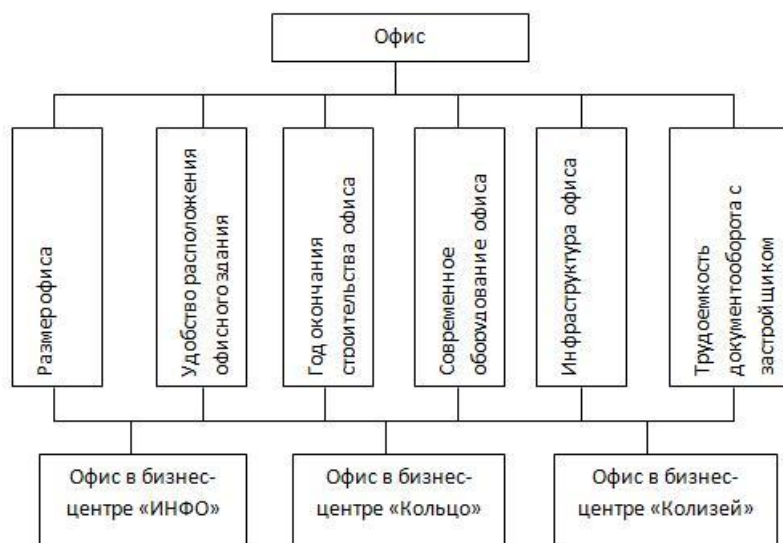
З допомогою МАІ визначить єдине джерело фінансування для компанії Гертца. Визначить узгодженість рейтингів.

Варіант 5

Індивідуальний підприємець вирішив купити (або побудувати за договором з будівельною компанією) нежитлове приміщення (офіс) для подальшого отримання доходу від здачі його в оренду. На ринку існують три альтернативи з приблизно однаковою вартістю (очевидної переваги жодна не має). У кожній альтернативи є інші переваги і недоліки за критеріями, які встановив підприємець:

- розміри офісу (площа);
- зручність розташування офісної будівлі;
- рік закінчення будівництва офісу;
- сучасне устаткування офісу (цифрова телефонна лінія, високошвидкісний Інтернет та ін.);
- інфраструктура офісу (парковка, охорона, місце харчування, фітнеса тощо);
- трудомісткість документообігу із забудовником.

Завдання полягає у виборі одного з трьох варіантів офісу, який якнайповніше задовольняє потреби підприємця.



Найважливішими оцінками нерухомості підприємець вважає зручність розташування офісної будівлі й інфраструктуру офісу. Менш важливими він вважає розміри офісу і його устаткування. Проте розміри офісу помітно важливіші при ухваленні рішення в порівнянні з трудомісткістю документообігу і роком закінчення будівництва офісу.

При парній оцінці трьох варіантів за кожним з критеріїв отримано матриці порівняння, як показано в таблиці.

Чи є отримані оцінки узгодженими?

Розмір офісу				Інфраструктура офісу			
	ІНФО	Кільце	Колізей		ІНФО	Кільце	Колізей
ІНФО	1	2	1/3	ІНФО	1	1/2	3
Кільце	1/2	1	1/5	Кільце	2	1	3
Колізей	3	5	1	Колізей	1/3	1/3	1
Зручність розташування				Сучасне устаткування офісу			
	ІНФО	Кільце	Колізей		ІНФО	Кільце	Колізей
ІНФО	1	1	2	ІНФО	1	4	5
Кільце	1	1	3	Кільце	1/4	1	5
Колізей	1/2	1/3	1	Колізей	1/5	1/5	1
Трудомісткість документообігу				Рік закінчення будівництва			
	ІНФО	Кільце	Колізей		ІНФО	Кільце	Колізей
ІНФО	1	5	3	ІНФО	1	1/5	1/7
Кільце	1/5	1	3	Кільце	5	1	1/3
Колізей	1/3	1/3	1	Колізей	7	3	1

Варіант 6

Gert's Sports - мережа спортивних магазинів, які швидко розвивається, на Східному узбережжі США. Власник мережі Боб Гертц накопичив капітал, щоб відкрити нові магазини. Він може побудувати магазини трьох типів: супермаркети (вартість \$3,5 млн.), торгові центри (\$1, 7 млн) і інтернет-крамницю (\$1 млн). У супермаркеті працює 150 чоловік, у торговому центрі – 65 чоловік, в інтернет-крамниці - 50 чоловік. Гертц може вкласти у відкриття магазинів до \$10 млн., при цьому він хоче досягти кількох цілей, зокрема, максимізувати дохід разом з максимізацією кількості зайнятих. Очікуваний дохід для супермаркету, торгового магазину і інтернет-магазину складає \$1, \$0,5 і \$1 млн. відповідно. Кількість магазинів кожного типу обмежена демографічними чинниками краю: інтернет-крамниць може бути не більше однієї, супермаркетів - не більше трьох, а торгових центрів - не більше семи. Оцініть ситуацію з допомогою МАІ. Передбачено врахування двох критеріїв - доходу і кількості зайнятих. Боб оцінив рейтинги за обома критеріями:

Рейтинг прибутковості

	Супермаркет	Торговий центр	інтернет-магазин
Супермаркет	1	4	7
Торговий центр	0,25	1	5

Інтернет-магазин	0,142857	0,2	1
------------------	----------	-----	---

Рейтинг за кількістю зайнятих

	Супермаркет	Торговий центр	інтернет-магазин
Супермаркет	1	0.25	0.333
Торговий центр	4	1	0.5
Інтернет-магазин	3	0.2	1

На підставі цих рейтингів знайдіть найкращі рішення з будівництва магазинів. Оцініть узгодженість рейтингів.

Варіант 7

Gert's Sports - мережа спортивних магазинів, які швидко розвивається, на Східному узбережжі США. Власник мережі Боб Гертц шукає постачальників хокейного спорядження. Він чекає різкого підвищення рівня продажів у зв'язку з незвично холодною зимою. Він дійшов висновку, що при виборі постачальника треба спиратися на його спроможність забезпечити своєчасну доставку замовлення. Рейтинги чотирьох можливих постачальників Боб оцінив так

Рейтинги постачальників

	Sticks Supply	Puck's House	Rinks Inc.	Goal Tenders
Sticks Supply	1	3	1	0,5
Puck's House	0,33333	1	0,5	0,25
Rinks Inc.	1	2	1	1
Goal Tenders	2	4	1	1

Методом МАІ виберіть двох кращих постачальників. Чи був Боб послідовний при складанні рейтингів?

Варіант 8

Відділ кадрів фірми звужив пошук майбутнього співробітника до трьох кандидатур: Стіва (S), Джейн (J) і Майлса (M). Остаточний відбір ґрунтується на трьох критеріях: співбесіда (C) досвід роботи (O) і рекомендації (P).

Відділ кадрів має найважливішим критерієм при прийомі на роботу рекомендації з попередніх місць роботи. Трохи поступаються йому за важливістю результати співбесіди з претендентом. Досвід роботи порівняно з рекомендаціями має істотно меншу важливість.

Після проведеної співбесіди з трьома претендентами, збору даних, що стосуються до досвіду їх роботи і рекомендацій, побудовані матриці АС, АО

і АР. Якого з трьох кандидатів слід узяти на роботу? Оцініть узгодженість даних.

$$A_C = \begin{matrix} & \begin{matrix} S & J & M \end{matrix} \\ \begin{matrix} S \\ J \\ M \end{matrix} & \begin{pmatrix} 1 & 3 & 4 \\ \frac{1}{3} & 1 & \frac{1}{5} \\ \frac{1}{4} & 5 & 1 \end{pmatrix} \end{matrix}, \quad A_S = \begin{matrix} & \begin{matrix} S & J & M \end{matrix} \\ \begin{matrix} S \\ J \\ M \end{matrix} & \begin{pmatrix} 1 & 3 & 4 \\ \frac{1}{3} & 1 & \frac{1}{5} \\ \frac{1}{4} & 5 & 1 \end{pmatrix} \end{matrix}, \quad A_P = \begin{matrix} & \begin{matrix} S & J & M \end{matrix} \\ \begin{matrix} S \\ J \\ M \end{matrix} & \begin{pmatrix} 1 & \frac{1}{2} & 1 \\ 2 & 1 & \frac{1}{2} \\ 1 & 2 & 1 \end{pmatrix} \end{matrix}.$$

Варіант 9

Допоможіть Марлен Віт обрати роботу після закінчення коледжу. Вона отримала три пропозиції від працедавців у Бакерсфилде, Фресно і Ойлдейле і визначила три важливі для неї критерії вибору: заробітна плата, стабільність роботи і привабливість міста.

1. Чому дорівнюють середні рейтинги за критерієм зарплати?
2. Чи узгоджені оцінки Марлен? Як можна змінити попарні оцінки, щоб узгодити їх?
3. Чому дорівнюють середні ваги критеріїв?
4. Яку роботу слід обрати?

Варіант 10

Кевин і Джун (*К* і *Д*) купують новий будинок. Наявними є три варіанти - А, Б і С. Кевин і Джун погоджували два критерії для вибору будинку: площа зелених луків (*Л*) і близькість до місця роботи (*Б*), а також розробили матриці порівнянь, наведені нижче. Потрібно оцінити три будинки в порядку пріоритетності і розрахувати коефіцієнт узгодженості кожної матриці.

$$A_{KD} = \begin{matrix} & \begin{matrix} K & D \end{matrix} \\ \begin{matrix} K \\ D \end{matrix} & \begin{pmatrix} 1 & 2 \\ \frac{1}{2} & 1 \end{pmatrix} \end{matrix}, \quad A_{KB} = \begin{matrix} & \begin{matrix} Л & Б \end{matrix} \\ \begin{matrix} Л \\ Б \end{matrix} & \begin{pmatrix} 1 & \frac{1}{3} \\ 3 & 1 \end{pmatrix} \end{matrix}, \quad A_{DB} = \begin{matrix} & \begin{matrix} Л & Б \end{matrix} \\ \begin{matrix} Л \\ Б \end{matrix} & \begin{pmatrix} 1 & 4 \\ \frac{1}{4} & 1 \end{pmatrix} \end{matrix},$$

$$A_{KL} = \begin{matrix} & \begin{matrix} A & B & C \end{matrix} \\ \begin{matrix} A \\ B \\ C \end{matrix} & \begin{pmatrix} 1 & 2 & 3 \\ \frac{1}{2} & 1 & 2 \\ \frac{1}{3} & \frac{1}{2} & 1 \end{pmatrix} \end{matrix}, \quad A_{KB} = \begin{matrix} & \begin{matrix} A & B & C \end{matrix} \\ \begin{matrix} A \\ B \\ C \end{matrix} & \begin{pmatrix} 1 & 2 & \frac{1}{2} \\ \frac{1}{2} & 1 & \frac{1}{3} \\ 2 & 3 & 1 \end{pmatrix} \end{matrix},$$

$$A_{DL} = \begin{matrix} & \begin{matrix} A & B & C \end{matrix} \\ \begin{matrix} A \\ B \\ C \end{matrix} & \begin{pmatrix} 1 & 4 & 2 \\ \frac{1}{4} & 1 & 3 \\ \frac{1}{2} & \frac{1}{3} & 1 \end{pmatrix} \end{matrix}, \quad A_{DB} = \begin{matrix} & \begin{matrix} A & B & C \end{matrix} \\ \begin{matrix} A \\ B \\ C \end{matrix} & \begin{pmatrix} 1 & \frac{1}{2} & 4 \\ 2 & 1 & 3 \\ \frac{1}{4} & \frac{1}{3} & 1 \end{pmatrix} \end{matrix}.$$

Варіант 11

Допоможіть Гордону Шамвею вибрати нове авто. Він зупинив свій вибір на трьох моделях, Buick Regal, Toyota Camry і Honda Accord, і вказав три основні для нього критерії ціна, надійність (по відгуках покупців), швидкість.

- а) чому дорівнюють середні рейтинги за критерієм швидкості?
- б) чому дорівнюють середні ваги критеріїв?

- в) чи Був Чарльз послідовний при завданні вагів?
- г) який автомобіль ви рекомендуєте купити?

Завдання 2

Запропоновано вибрати купівлю за допомогою МАІ. Для цього потрібно:

1. Сформулювати критерії вибору.
2. Визначити і обґрунтувати рейтинги альтернатив за кожним критерієм. Оцінити їх узгодженість.
3. Визначити рейтинг критеріїв. Оцінити їх узгодженість.
4. Вибрати найбільш прийнятне рішення.

Кількість критеріїв має бути не менше трьох. Кількість варіантів рішень (товарів) - не менше чотирьох (вказати марки). Схема пошуку рішення аналогічна завданню 1.

Варіанти завдань - приклади товарів для вибору:

1. Пральна машина
2. Телевізор
3. mp3 -плеер
4. Мобільний телефон
5. Автомобіль
6. Кухонна плита
7. Кухонний комбайн
8. Ноутбук
9. Холодильник
10. Мікрохвильова піч
11. Цифровий фотоапарат
12. DVD -плеер
13. Лазерний принтер
14. Сканер
15. Монітор

Контрольні питання

1. Навести приклади ухвалення рішень з урахуванням багатьох цілей і критеріїв.
2. Яке математичне підґрунтя має МАІ?
3. Алгоритм методу МАІ.
4. Як визначають вагові коефіцієнти критеріїв?
5. Як обчислюють і перевіряють коефіцієнти узгодженості?
6. Як виглядає стандартна шкала порівняння?
7. Що таке індекс рандомізації IP?
8. Наведіть схему розв'язання задачі в процесорі Excel.

ОЦІНКА БАГАТОКРИТЕРІАЛЬНИХ АЛЬТЕРНАТИВ ЗА ДОПОМОГОЮ СПР PRIME Decisions

Мета: знайомство та використання СПР PRIME Decisions для підтримки керівницьких рішень

Порядок виконання роботи

За наведеним прикладом поставити і розв'язати задачу вибору з кількох альтернатив. Проілюстрований копіями екрану монітора результат занести до звіту.

Загальні відомості

Системи підтримки рішень (СПР) засновані на формалізації методів одержання вихідних і проміжних оцінок та алгоритмізації самого процесу вироблення рішення. Людино-машинна процедура вирішування за допомогою СПР являє собою циклічний процес взаємодії людини і комп'ютера. Однією з СПР- програм є доволі поширена PRIME Decisions, яка використовується як у навчальною метою так і на практиці. У PRIME Decisions реалізовано методику PRIME (Preference Ratios In Multiattribute Evaluation), розроблену в Гельсінському університеті. У PRIME інформація про вподобання (переваги) альтернатив надається *інтервально*, бо ми часто не знаємо точного значення нашого вподобання, але можемо виставити йому оцінку на кшталт: “це добре”, “це погано”, “це трохи краще, ніж те” тощо). Інтервали допомагають у моделюванні думок як математичних обмежень, напр. “добрий” можливо визначити як такий, що в 2 - 3 рази кращий, ніж “поганий”.

PRIME Decisions дотримується встановлених процедурних кроків (табл. нижче) у розв'язанні аналітичних проблем вирішування.

Процедурний крок	PRIME Decisions еквівалент
Обрати мету	-
Визначити атрибути (критерії)	Вікно Value Tree
Визначити альтернативи	Вікно альтернатив
Оцінити вподобання	Вікно “Інформація про вподобання”
Визначити найкращу альтернативу	Інтервали оцінок, ваги, домінування, вікна Decision Rules

Постановка задачі

Як приклад застосування PRIME Decisions у цій роботі запропоновано розв'язати задачу вибору найкращої комплектації комп'ютера, оцінюючи за вибраними шкалами всі елементи устаткування. Хід роботи описано нижче.

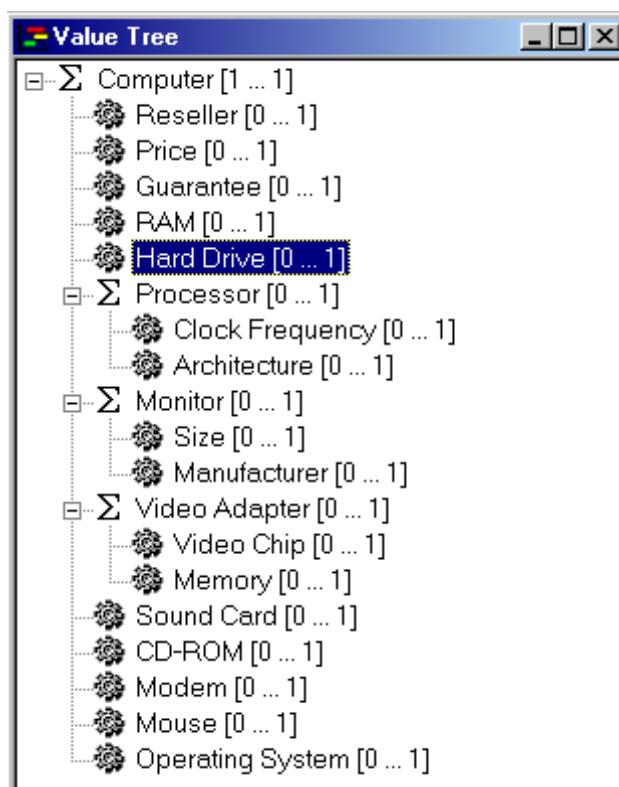
Хід роботи

1 Дерево оцінок

1. Після запуску програми на екрані з'являється основне робоче вікно та три додаткові вікна: **Value Tree**; **Alternatives**; **Preference Information**.

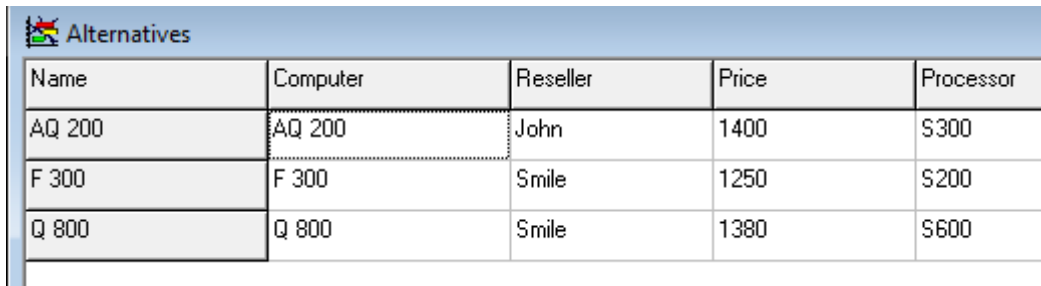
Клацання правою клав'яшею миші у вікні **Value Tree** виводить меню з доступними командами. Виберіть команду **New Main Goal** і введіть назву головної мети: “Комп'ютер”.

2. Тепер потрібно додати атрибути комп'ютера: “Reseller” (продавець), “Price” (ціна), “Guarantee” (гарантійний строк), “RAM” (об'єм оперативної пам'яті), “Hard drive” (об'єм жорсткого диску). Потім додати нову (“проміжкову”) мету “Processor”, до якої прив'язати атрибути “Clock Frequency” і “Architecture”. У цьому прикладі ми створюємо дерево значень, на кшталт того, що на світлинці нижче. Цифри в дужках вказують на інтервал ваги, пов'язаний з відповідним елементом. Оскільки модель ще не розрахована, то кожній меті (окрім головної) й атрибутам формально присвоєно значення від 0 до 1.



2 Визначення альтернатив

3. Коли атрибути означено, будуємо сітку (матрицю) можливих альтернатив у вікні Alternatives:



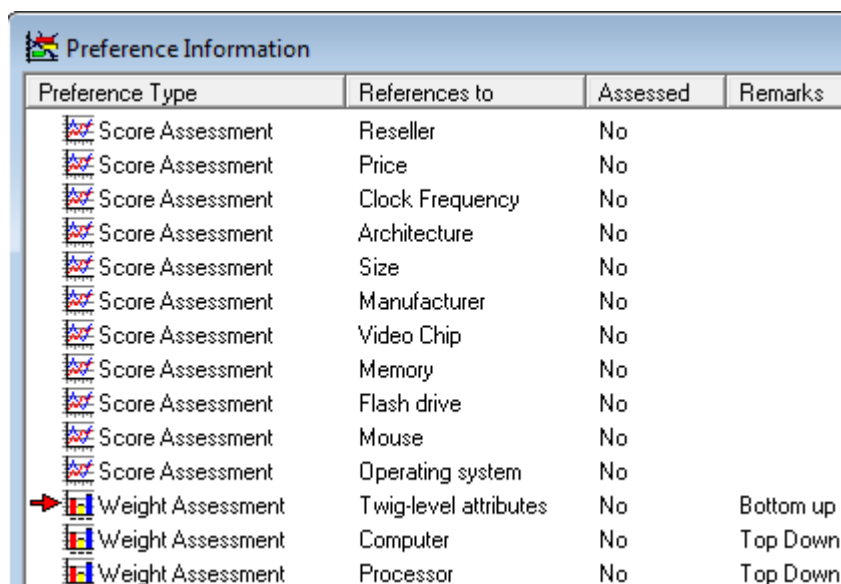
Name	Computer	Reseller	Price	Processor
AQ 200	AQ 200	John	1400	S300
F 300	F 300	Smile	1250	S200
Q 800	Q 800	Smile	1380	S600

Назви атрибутів уже є у заголовках колонок таблиці; перший стовпець – це головна мета. Командою **New Alternative** контекстного меню цього вікна заповнюємо рядки таблиці відповідними значеннями атрибутів.

Введіть альтернативи (5-6) для мети та значення кожного атрибута. (У деяких випадках комірки допускається залишати незаповненими).

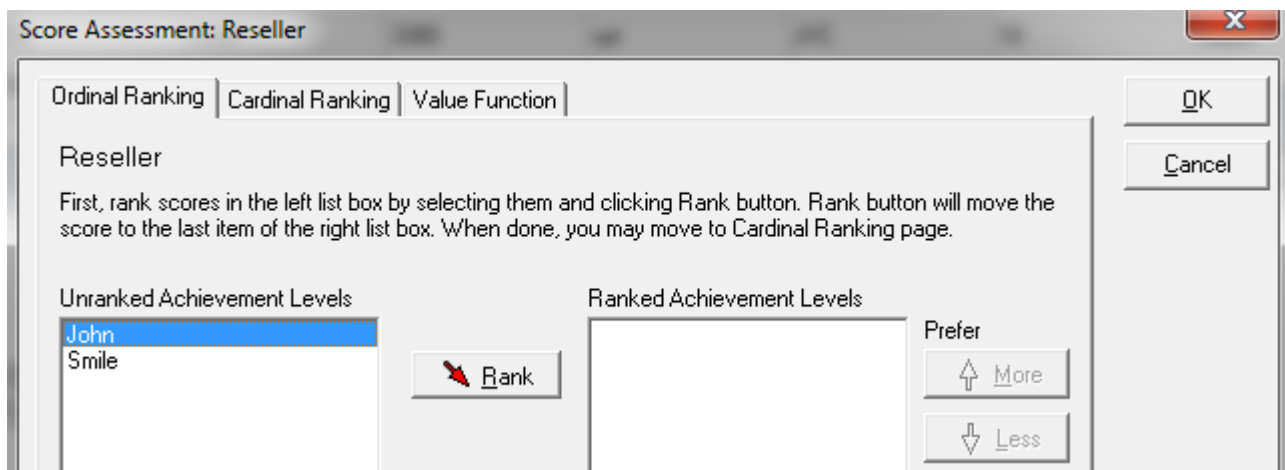
3 Відомості про вподобання та оцінка альтернатив

Атрибути у PRIME Decisions відіграють роль *критеріїв* ухвалення остаточного рішення і мусять бути оцінені за тією чи іншою шкалою вподобань. Уподобання (переваги) визначають у вікні **Preference Information**. Видно (світлина нижче), що атрибути (значення критеріїв) запропоновано оцінювати як **Score Assessment** (у балах), а додатковим цілям призначати вагові бали **Weight Assessment**. (Червоні стрілки вказують на стиль оцінки ваги, яка була вибрана).



Preference Type	References to	Assessed	Remarks
<input checked="" type="checkbox"/> Score Assessment	Reseller	No	
<input checked="" type="checkbox"/> Score Assessment	Price	No	
<input checked="" type="checkbox"/> Score Assessment	Clock Frequency	No	
<input checked="" type="checkbox"/> Score Assessment	Architecture	No	
<input checked="" type="checkbox"/> Score Assessment	Size	No	
<input checked="" type="checkbox"/> Score Assessment	Manufacturer	No	
<input checked="" type="checkbox"/> Score Assessment	Video Chip	No	
<input checked="" type="checkbox"/> Score Assessment	Memory	No	
<input checked="" type="checkbox"/> Score Assessment	Flash drive	No	
<input checked="" type="checkbox"/> Score Assessment	Mouse	No	
<input checked="" type="checkbox"/> Score Assessment	Operating system	No	
<input checked="" type="checkbox"/> Weight Assessment	Twig-level attributes	No	Bottom up
<input checked="" type="checkbox"/> Weight Assessment	Computer	No	Top Down
<input checked="" type="checkbox"/> Weight Assessment	Processor	No	Top Down

4. Проранжуємо (виставимо по порядку) атрибути з точки зору очікуваної корисності. Виберемо котрийсь з елементів (атрибут) і двічі клацнемо мишкою в колонці **Preference Type** по полю, що відповідає цьому елементові – з’явиться панель **Score Assessment**. Вибираємо закладку **Ordinal Ranking** (порядковий рейтинг). (Клацання, наприклад, по першому рядку в табл. **Preference Information** виведе **Score Assessment** для атрибута “Reseller” (наступна світлина). Кнопка **Rank** перекидає елементи з віконця Unranked до Ranked, де кнопками ↑ і ↓ можливо їх ранжувати в порядку вподобання (на порядок вказує цифра у дужках біль кожного значення критерію). Потрібно провести таку процедуру з усіма елементами панелі **Preference Information**, які мають тип вподобання **Score Assessment**. Відповідно в колонці **Assessed** для кожного оціненого атрибута позначка “No” змінюється на “Yes”.

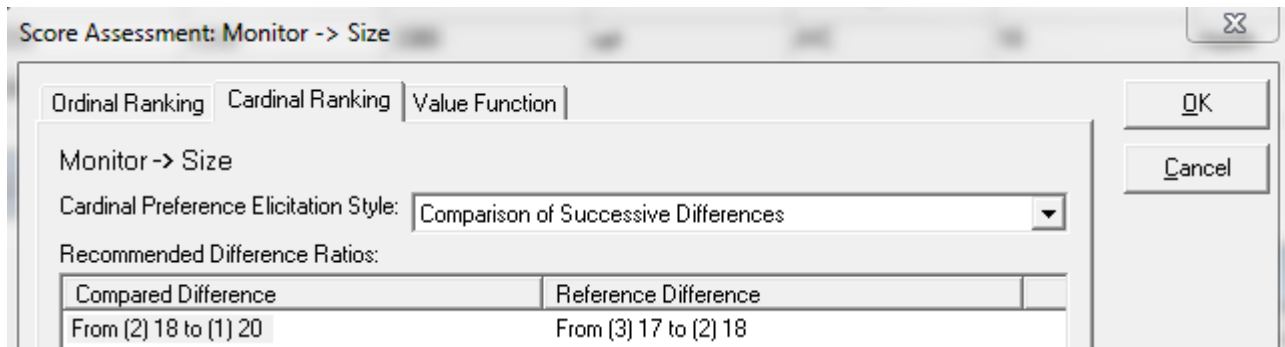


Замість описаної процедури можна обрати **Elicitation Tour** у меню **Model**, який проведе через усі етапи ранжування.

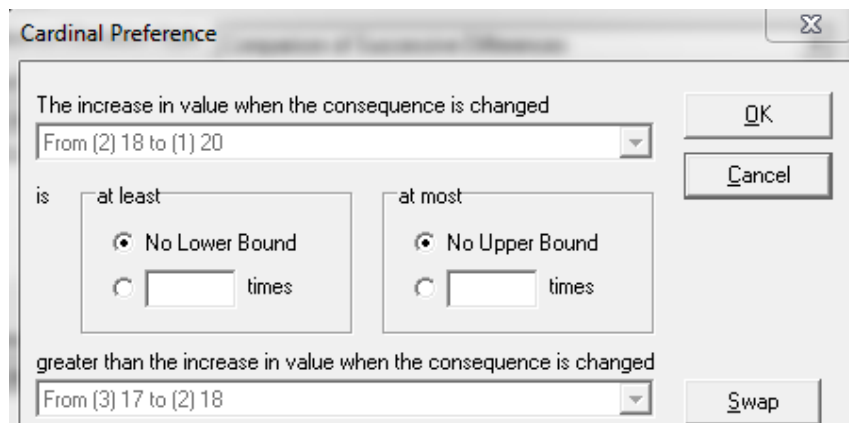
Таким чином ми розташували значення кожного атрибуту на уявних лінійках. Наступний крок – нанесення шкал на ці лінійки.

5. Інформації про силу (ступінь) переваг, якої не було в порядковому рейтингу, надає т.зв. кардинальний рейтинг (**Cardinal ranking**). Числа шкали оцінок за цим рейтингом не повинні бути від’ємними, хоча можуть лишатися невизначеними.

Знову послідовно клацаємо на кожному пунктові вікна “Інформація про вподобання”.



У показаному на світлинні прикладі порівнюється відмінність між (2) та (1) місцем у рейтингу (мониторів за розміром) з відмінністю між (3) і (2) місцями. Чи перша відстань на рейтинговій шкалі більша/менша від другої? Чи вони однакові? Форма рекомендацій **Cardinal ranking** залежить від обраного стилю оцінювання (“Cardinal Preference Elicitation Style”). Якщо вибрати “Порівняння послідовних відмінностей (Comparison Successive Differences)”, то червона стрілка **Add** виведе наступне вікно:



Пропозиція у ньому звучить так: “Збільшення оцінки при зміні наслідку з (2) на (1) принаймні (і щонайбільше) в ... разів більше, ніж збільшення оцінки, коли наслідок змінюється з (3) на (2)”. Іншими словами, якщо у віконці “at least” поставити цифру 2, то вийде, що при виборі монітора різниця між 20-дюймовим пристроєм та 18-дюймовим для нас у 2 рази суттєвіша ніж між 17- і 18-дюймовим. “No lower bound” і “no upper bound” означає необмежені оцінки. З міркувань, наведених двома абзацами нижче, візьмемо “at least” число, яке не дуже відрізняється від одиниці, напр. 1,2.

Іншими варіантами “Cardinal Preference Elicitation Style” є

- ✓ порівняння двох відмінностей від найнижчого рівня (“Comparison of Two Differences from Lowest Level”),
- ✓ порівняння різниці від найнижчого та найвищого рівня (“Comparison of Difference from Lowest and to Highest Level”)
- ✓ пряма оцінка за шкалою [0, 1] (“Direct Rating in [0, 1]-scale”).

Увага! Кардинальний рейтинг може породжувати викривлення, оскільки числові межі бувають важкими для визначення. Тому (перший тип викривлення) для порівняння важливо, щоб різниці були приблизно однакового розміру, принаймні, однакового порядку величини. Коли десь різниця в порівнянні набагато більша (менша) ніж еталонна відмінність, неможливо достовірно визначити порядок величин меж. До того ж (другий тип викривлення) після оцінювання відмінностей у бодай навіть цілих числах, надалі доведеться мати справу з оберненими до них дрібними величинами. Тому було б добре, аби відмінність різниць була меншою або такої ж величини, як інша різниця в одному порівнянні. Правило “зі стелі”: межі легко зрозуміти, якщо вони варіюються від 1 до 10. Можна міняти відмінності в порівнянні (за допомогою кнопки **Swap**), яка обертає межі і тим самим видаляє викривлення другого типу

У “порівнянні послідовних відмінностей” (наприклад, порівняння різниць між 3-м і 2-м рівнями з різницею між 4-м до 3-м рівнем) можуть виникати обидва типи викривлення, якщо різниці дуже відрізняються.

У “порівнянні двох різниць” (наприклад, якщо є п'ять рівнів досягнень, порівняйте різницю від 5-го до 2-го рівня із різницею між 5-м і 3-м рівнями) останній тип викривлення виключено, оскільки кожна різниця більша, ніж та, з якою його порівнюють. Проте перший тип викривлення може з'явитися, якщо відміна найнижчого рівня від другого нижнього рівня дуже мала.

У “порівнянні від найнижчого та найвищого рівня” порівнюються відстань від найнижчого до якогось рівня та від цього рівня до найвищого. Оскільки перша з цих відстаней змінюється із переходом до нового рівня, можуть виникати обидва типи викривлення.

У **прямому оцінюванні за шкалою** різниця між найгіршим та найкращим рівнем (= 1) - це еталонна відмінність. З нею порівнюють інші різниці від найнижчого рівня до обраного рівня. Оскільки еталон лишається незмінним, перший тип зміщення не виникає, за винятком випадків, коли відмінність найнижчого рівня і другого нижнього рівня незначна. Більш того, оцінки менші за одиницю, позаяк жодна різниця не перевищує еталонну, і тому прямий рейтинг підпадає під зміщення другого типу.

У нашому прикладі прямий рейтинг найбільш застосовний до порівняння цін. Для решти порівнянь "порівняння двох відмінностей" було б найнадійнішим способом виявити кардинальні переваги для особи, яка вирішує.

4 Вагова оцінка

Другий етап у визначенні переваг полягає в оцінці *ваги* критеріїв (атрибутів). PRIME визначає вагу атрибута як виграш у загальній оцінці завдяки зміні від найгіршого наслідку цього критерію (атрибуту) до найкращого. PRIME вико-

ристовує інтервальне зважування: найбільшу корисність оцінено у 100 балів; ваги інших атрибутів порівнюються з цим значенням і задаються в інтервалі [0, 100].

PRIME має два стилі зважування:

знизу вгору -	потрібно зважити атрибути (критерії) моделі один відносно одного
згори донизу -	потрібно порівняти ваги суб-атрибутів (та суб-цілей) з з вагою обраного “еталонного” атрибуту

Ваги можна визначати обома способами; щоб вибрати, який саме використовувати, потрібно поставити/зняти прапорець **Top Down Weights** у меню **Options**. Обидва способи оцінки ваги мають свої переваги. Спадна оцінка (Top Down) забезпечує структурований простий для розуміння спосіб визначення ваги; спосіб “знизу вгору” (**Bottom Up**) показує всі атрибути в одному вікні й не потребує зважування цілей. Як правило, метод “bottom-up” годиться для невеличких моделей, тоді як “top Down” краще застосовувати для багатоцільових моделей.

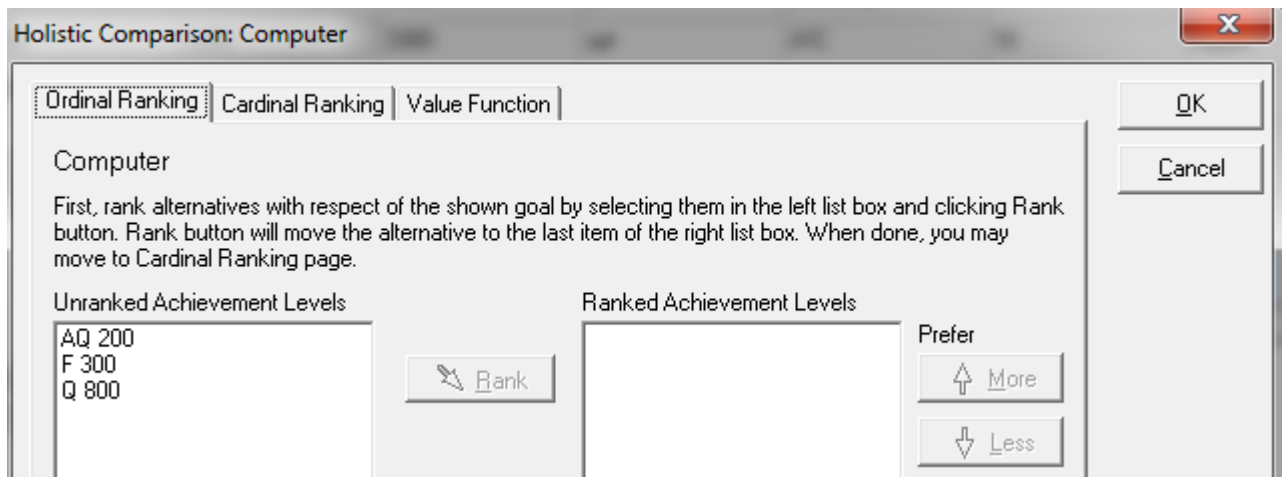
6. Проведемо процес зважування для кожної мети в моделі. Виберемо **Options** → **Top Down Weights** і, двічі клацаючи на панелі вподобань по позиціях “Weight Assessment” навпроти червоних стрілок, введемо оцінки у цілих числах від 0 до 100 для “Low bond” і “Upper bond” для кожного атрибуту. Один (найважливіший для нас) атрибут має бути оцінений і згори й знизу у 100.

	Worst Conseq.	Best Conseq.	Lower bound	Upper bound
Reseller	Ipsum	Smile	8	9
Price	1410	1250	100	100
Processor	Implicit (Goal)	Implicit (Goal)	25	35
Monitor	samsung	Sony	17	20
Video Adapter	Implicit (Goal)	Implicit (Goal)	5	6
Flash drive	2.1	3.0	0	2
Mouse	f	a	0	2

Атрибути, які мають єдине значення (тобто без вибору), за визначенням мають обидві межі 0 (бо немає можливості отримати краще значення).

5 Цілісні порівняння

У цілісному порівнянні оцінюють порядкові та кардинальні рейтинги стосовно цілей. Засоби для цілісного порівняння схожі на засоби бальної оцінки. Вікно **Holistic Comparison** можна викликати, двічі клацнувши на потрібному рядку вікна **Preference Information**:



7. Проводимо бальне і кардинальне ранжування для “Holistic Comparison” аналогічно п. 4.

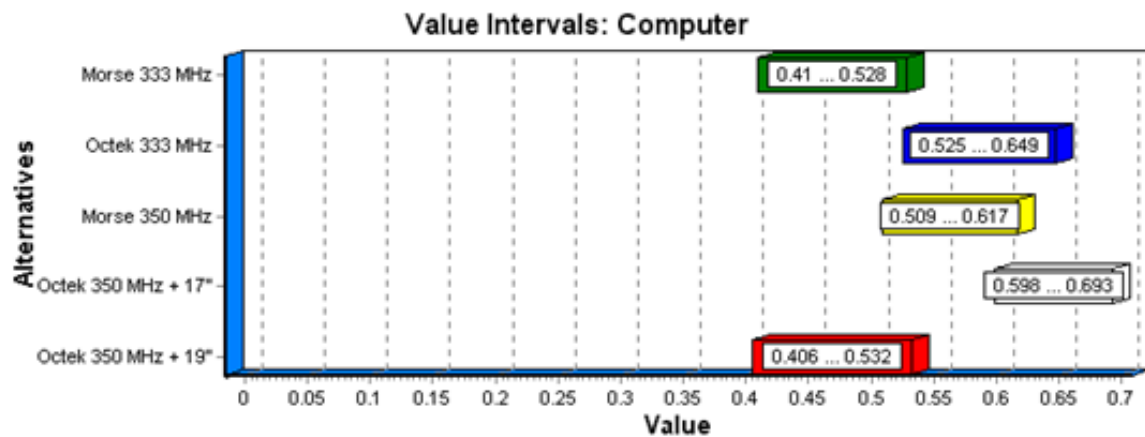
6 Розрахунок моделі

8. Спочатку для швидкого перегляду результатів запускаємо **Calculate Model Partially** в меню **Model** – ця процедура розв’язує лише найважливіші частини моделі: інтервали значень основної мети, структури домінування та правила вирішування. Потім даємо команду **Calculate Model**, яка запускає повний процес обчислення моделі.

Методом розв’язання розглядуваної задачі є лінійне програмування. Результати обчислень виводяться у вікна **Value Intervals**, **Weights**, **Dominance**, **Decision Rules**, які можна викликати з пункту **Windows** головного меню.

Value Intervals показує діапазон можливих значень. Кожна альтернатива має інтервал значень для кожного атрибута й мети – діапазон його можливих значень. Поняття інтервалу значень альтернативи стосується інтервалу значення альтернативи основної мети, оскільки містить загальне значення альтернативи.

На графіку **Value Intervals** прямокутник, який зображує модель комп’ютера, тим правіший, чим краща модель. Бачимо (на світлинці далі), що модель комп’ютера Ostek 350 MHz+17” домінує абсолютно на моделями Morse 333 MHz та Ostek 350 MHz + 19”, бо найгірша оцінка першого виявилася більшою від найкращих оцінок останніх (проекції відповідних прямокутників на горизонтальну вісь не перетинаються)..

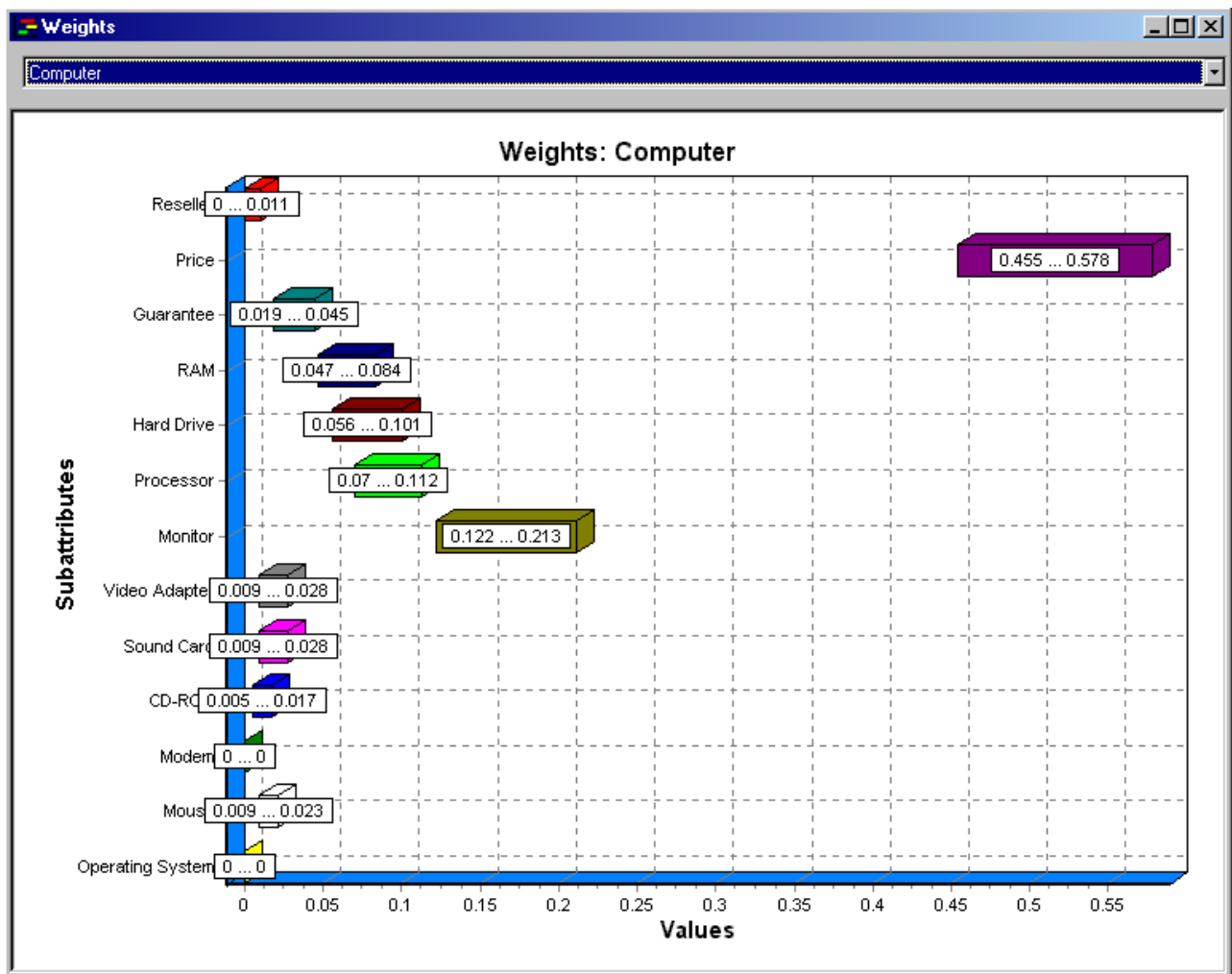


Це вікно не надає можливості зробити остаточний висновок про домінування. Аби бути впевненим у цьому питанні, слід розглянути структури домінування у вікні **Dominance** (див. нижче).

Інтервали значень можна розглядати як ненормовані або нормовані. Інтервали ненормованих значень репрезентують частину значення альтернативи від загальної оцінки за атрибутом або метою. Оскільки лише атрибути рівня гілки (twig-level) роблять внесок в оцінку моделі, сума найвищих балів (ваг) кожного атрибута дорівнює 1. З іншого боку, інтервали можна нормалізувати на $[0, 1]$ -масштаб. Значення масштабуються так, що найбільшій верхній межі задається значення 1, а нижня верхня межа - значення 0. Нижні межі масштабуються аналогічно. Нормалізовані інтервали значень пропонують загальний погляд на значення, що є корисним при порівнянні альтернатив між собою. Однак інформація про вплив на загальну величину втрачається.

9. Застосувати нормалізацію, вибравши **Normalized Intervals** в меню **Options** та перерахуйте модель.

Вікно **Weights** показує важливість атрибуту (критерію) або мети порівняно з іншими атрибутами. Видно лише під атрибути мети.



10. Виведемо на екран вікно домінування (**Dominance**), яке містить матрицю переваг. Червона крапка в матриці вказує на те, що альтернатива цього стовпця переважає альтернативу з відповідного рядка. Якщо крапка зелена, ситуація зворотна.

	Morse 333 MHz	Octek 333 MHz	Morse 350 MHz	Octek 350 MHz + 17"	Octek 350 MHz + 19"
Morse 333 MHz	○	●	●	●	
Octek 333 MHz	→ ●	○			→ ●
Morse 350 MHz	→ ●		○	●	→ ●
Octek 350 MHz + 17"	→ ●		→ ●	○	→ ●
Octek 350 MHz + 19"		●	●	●	○

З прикладу на світліні можна зробити висновок що альтернативи Octek 333 та Octek 350+17 найкращі, оскільки нічим не домінуються. Проте це само по собі не уможливорює остаточний висновок, який лишається за людиною.

Питання самоконтролю

1. Що таке точкові критеріальні оцінки і що таке інтервальні?
1. Яка потреба вводити інтервальні критеріальні оцінки замість точкових?
2. Чим відрізняються методики оцінки критеріїв (атрибутів) “знизу вгору” та “згори донизу”?
3. Для чого застосовують нормалізацію інтервальних оцінок?

СТВОРЕННЯ ПРОЕКТУ DEDUCTOR Studio

Мета: початкове знайомство з пакетом Deductor, пробне створення проекту аналізу даних; первісна обробка даних

Порядок виконання роботи

За наведеним прикладом здійснити обробку даних із навчального файлу. Результати (роздруки копій екрану з коментарями) занести до звіту.

Загальні визначення і відомості про платформу Deductor

Керівницьке рішення ухвалюється часто на основі аналізу великих обсягів даних. Вхідні дані мають бути репрезентативними, точними й достовірними. Однак вони, як правило, поступають із різних джерел і не завжди якісні. Дані можуть містити орфографічні помилки, фіктивні та логічно неправильні значення, порушувати формат запису, нарешті просто бути відсутніми. Тож, **попередня обробка даних** – конче потрібний крок для забезпечення задовільного результату аналізу. Вона включає в себе **очищення даних**: заповнення пропусків, редагування аномалій та помилок, згладжування даних, виявлення дублікатів і протиріч тощо, а також (за потреби) **попереднє перетворення даних**: зниження розмірності, усунення незначущих чинників тощо. Функції виправлення даних мість у собі програмний пакет Deductor.

Аналітична платформа Deductor складається з двох компонентів:

- Deductor Warehouse – сховище консолідованих даних з різних джерел;
- Deductor Studio – програма аналізу даних, робоче місце аналітика.

Робота в Deductor Studio побудована на створенні сценарію обробки даних. **Сценарій** є послідовність кроків для отримання потрібного результату. Їх створюють за допомогою трьох майстрів – імпорту, експорту та обробки даних. **Мастер імпорта** потрібен, бо Deductor Studio не дозволяє вводити дані вручну. (Версія програми Academic дозволяє імпорт лише з текстових файлів, тож, наприклад, дані Excel доводиться попередньо переводити у текстовий формат .txt). Створені за допомогою майстрів сценарії сукупно утворюють **проект**.

У Deductor дані очищує обробник **Парціальная обработка**. Спочатку виконується **заповнення пропусків**. Для цього обробник пропонує два способи: **Аппроксимация** й **Максимальна правдоподібність**. Апроксимація заповнює пропущене значення, усереднюючи сусідні. Максимальна правдоподібність підставляє значення найімовірніше значення зі статистичної точки зору – цей метод рекомендовано для неупорядкованих даних.

Редагуючи **аномалії**, **Парціальная обробка** із різною точністю вишукує в даних випадкові або рідкі події (різкі відхилення), які не вкладаються в загальну картину досліджуваного явища (процеса).

Ще однією функцією обробника є **очищення від шумів**. Шумами називають швидкі випадкові зміни значень. Вони, як правило, не бажані, бо приховують загальні тенденції та утруднюють побудову моделі прогнозу. Операція варта до застосування, коли шум має нормальний розподіл із малою дисперсією.

Згладжування даних потрібно, коли ряд даних нерівномірний – містить багато дрібних структур, що утруднює пошук загальних закономірностей. Згладжування дозволяє вказати смугу пропускання частот – відсікти коливання числових даних з частотою вище деякого порогу. Вужча смуга відповідає вищій мірі згладжування; проте надто вузька смуга може призвести до втрати корисної інформації.

Дублікати й протиріччя в даних можливо знайти за допомогою одноіменного обробника.

1 Експорт, імпорт та візуалізація даних

Стандартний запуск програми відкриває робоче вікно Deductor Studio. Меню **Файл** містить окрім звичних команд виклик демоприкладів аналізу даних і сховищем даних; **Вид** містить команди підключення даних та підготовки звітів, а також підготовки сценаріїв.

1. Створіть теку “DS_імпорт_експорт_чистка” для результатів роботи. Скопіюйте до неї файл початкових даних Credit.txt, який входить в поставку версії програми (тека C:\Program Files\BaseGroup\Deductor\Samples).

2. Аби завантажити ці дані запусіть **Мастер** кнопкою **F6** або через контекстне меню вкладки **Сценарии**.

- Перевірте, чи завантажувється перший рядок таблиці (data frame) як рядок заголовків. (Можливо, напр., розпочати імпорт з потрібного рядка, без заголовків).
- Перевірте роздільник цілої та дробової частин чисел. Якщо у початкових даних була десяткова точка, то слід вказати саме точку як роздільник, інакше числа не будуть розпізнані.
- Налагодьте параметри стовпців: спробувати поміняти назву; перевірити правильність автоматичного визначення типів даних для кожного поля (ліва панель, поле **Тип даних**).
- Розпочніть імпорт кнопкою **Пуск**. Вказати **Таблиця**, як засіб виведення. Поставити свою мітку – назву гілки сценарію, напр. “Імпорт Credit.txt”. У правій панелі відобразяться отримані дані.

- Зробіть зміни у вузол сценарію: **Настроити** → **Таблиця и Статистика**; ознайомтеся із вкладкою **Статистика** в панелі візуалізації.
 - Налагодьте лише візуалізацію даних для вузла сценарію: в його контекстному меню виберіть **Мастер визуализации**; задайте виведення **Таблиця, Статистика, Діаграма**; налагодьте діаграму як точкову по полю “Сумма кредита”, підписи осі X – поле “Срок кредита”; ознайомтеся із вкладкою **Діаграма**; за допомогою її панелі інструментів або контекстного меню змініть тип діаграми, поля та ін.
 - За допомогою майстра візуалізації, додайте відображення типу **Гистограмма**. Налагодьте її для виведення сум кредитів.
- Запротоколюйте всі зміни, роблячи копії екрану.
- Використовуючи майстер експорту, вивантажте дані у вигляді текстового файлу “Export Credit.txt” у свою теку. Задайте експорт лише частини стовпців таблиці (довільно).
 - Збережіть у своїй теці створений проект імпорту-експорту для наступного повторення: “<Своє прізвище>_Import-Export”. Перевірте ще раз, чи проект відкривається в програмі.

2 Поновлення пропущених даних

- Запустивши новий проект, збережіть його у файлі “Очищення даних”.
- Підготуйте в Excel таблицю значень функції $\sin(x)$ або $\cos(x)$, де аргумент змінюється від 0 до 10 з кроком 0,05. Видаліть декілька значень функції (не аргументу), аби з’явилися пропущені дані. Збережіть отримане в своїй теці у текстовому форматі.
- Завантажте створені дані в Deductor. Назвіть гілку “Поновлення $\sin(x)$ ”. Задайте відображення **Таблиця и диаграмма**.
- У розділі **Очистка даних** відкритого вікна виберіть пункт **Парциальная обработка**, перемикач **Аппроксимация**. Надалі не задавайте ніякої обробки даних. Відобразіть **Таблиця и Диаграмма**.
- Порівняйте оброблені дані з необробленими.

3 Редагування аномалій

- Скопіюйте у свою теку демонстраційний файл BuKuDu.txt. Завантажте його у той же проект за допомогою майстра імпорту. Назвіть його “Редаг_аномалій_ BuKuDu”. Відобразіть таблицю і точкову діаграму для поля “у”.
- Для створеної гілки проведіть парціальну обробку. При цьому

- в її вікні **Восстановление пропущенных данных** установіть премикач **Отключить**;
- у вікні редагування аномальних значень виберіть поле **Количество**, встановіть прапорець **Редактирование аномальных значений**, виберіть малу **Степень подавления**;
- згладжування не задавайте. Виконайте редагування аномалій (кнопки **Пуск** і **Далее**). Відобразіть знову таблицю і точкову діаграму для поля “Сумма кредита”.

3. Порівняйте оброблені й необроблені дані. Виведіть у правій нижній частині вікна панель деталізації. З’ясуйте за її допомогою, на які значення були замінені аномальні.

4. Для тієї ж гілки “Редаг_аномалій_VuKuDu” проведіть ще одне редагування аномалій, вибравши більшу міру їх придушення. Надайте вузлам сценарію зрозумілі назви. Порівняйте нові величини для аномальних значень. Зробіть і запротоколюйте висновки.

4 Згладжування та очищення від шумів

1. Скопіюйте у свою теку демонстраційний файл VuKuDu.txt. У тому ж проєкті створіть ще одну гілку сценарію “Згладжування шумів_”, імпортувавши дані цього файлу. Відобразіть при цьому лінійчасту діаграму.

2. Проведіть для цього вузла згладжування та очищення від шумів двома способами (по черзі) обробника **Парциальная обработка** (крок **Спектральная обработка**). Для кожного перетворення виведіть лінійчасту діаграму. Тут потрібні налаштування:

- для опції **Сглаживание** залишити смугу пропускання 50;
- для видалення шуму встановити *середній* ступінь віднімання.

3. Надайте отриманим вузлам сценарію зрозумілі назви. Порівняйте за допомогою діаграм початкові дані й отримані після кожного з перетворень. Занесіть висновки у звіт із лабораторної роботи.

Каталог *C:\Program Files\BaseGroup\Deductor\Manual* містить файли:

Руководство аналитика.pdf, Импорт и экспорт данных.pdf, Deductor.pdf.

Питання самоконтролю

1. Які компоненти має Deductor Studio?
2. Як створити проект у Deductor Studio?
3. Що зберігає файл проекту? Чи є в ньому імпортовані дані?
4. В якому форматі можливий експорт даних з Deductor Studio?.
5. Для чого обробляють дані?
6. Для чого призначений обробник “Парціальна обробка”?
7. Для чого поновлювати пропущені дані? Для яких даних годиться кожен з способів заповнення пропусків Deductor?
8. Які дані вважають зашумленими? Які варіанти згладжування і віднімання шуму пропонує Deductor?
9. Що таке аномалії і як вони можуть вплинути на результат аналізу даних?

ІНТЕЛЕКТУАЛЬНИЙ АНАЛІЗ ДАНИХ У СПР

Мета: *виконати кластерний аналіз даних за методом k -середніх*

Порядок виконання роботи

Виконати приклад, наведений в інструкції. Інтерпретувати отриманий результат розбиття. Якщо величина силуету виявиться замалою, повторити розрахунки для іншого числа кластерів.

Складаючи звіт з роботи, надати аналіз отриманих результатів, відповіді на питання, виділені в інструкції блакитним кольором.

Теоретичні відомості

Для аналізу та вироблення пропозицій у СПР задіяні різні методи, зокрема, інтелектуальний аналіз даних. Метою такого аналізу є виявлення прихованих закономірностей і тенденцій у великих (і часто неструктурованих) даних. Його використовують поряд із статистичним аналізом з метою побудувати модель досліджуваного явища (складної системи). Створена модель потім закладається у майбутню СПР.

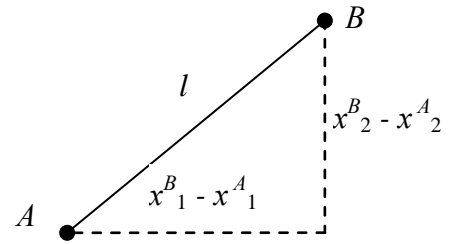
До методів інтелектуального аналізу даних належить **кластерний аналіз**, спосіб розбиття різних об'єктів на гурти подібних собі. Працюючи з цими гуртами – визначаючи, що в їхніх членів спільного, а що відрізняє одне від одного – часто можливо виявити корисну інформацію. Кластеризація на відміну від звичайної класифікації не задає заздалегідь класифікаційні гурти, вони виявляються самі по собі в процесі обробки даних.

У лабораторній роботі розбиття на кластери проводиться методом **k -середніх**. Для його використання кількість кластерів k має бути визначеною заздалегідь. Гіпотеза про кількість кластерів може ґрунтуватися на теоретичних міркуваннях, результатах попередніх досліджень або здогадці. Алгоритм призначає випадковим чином у багатовимірному просторі характеристик центри майбутніх кластерів (k штук), потім розраховує відстані між ними та кожним об'єктом, приписуючи їх до того кластера, до якого вони найближче. Закінчивши приписування, алгоритм розраховує середні значення (центри мас) для кожного кластера. набір середніх являють собою координати нового (скоректованого) положення кластерних центрів. Далі повторюється обчислення відстаней і приписування об'єктів до того чи іншого кластера. Ці операції тривають допоки центри не припинять “мігрувати” у просторі.

Відстань між точками береться евклідова:

$$l = \sqrt{\sum_{i=1}^n (\alpha_i - c_i)^2}. \quad (1)$$

Тут набір x_i – координати об’єкта, а c_i – координати центру конкретного кластера в n -вимірному просторі. Найпростіше задача оцінки відстані виглядає на площині (2D-просторі), де справедлива звичайна теорема Пітагора (див. рис.).



Коли розбиття на m кластерів не дозволяє однозначно інтерпретувати результат, процедуру повторюють для іншого k . Об’єктивним показником якості кластеризації є т.зв. **силует**. Він розраховується як різниця середньої відстані до членів сусіднього кластера x' і середньої відстані до членів свого кластера x , поділена на найбільше з цих двох величин:

$$s = \frac{x' - x}{\max(x', x)} \quad (2)$$

Задовільне розбиття дає $s \rightarrow 1$. Коли $s \leq 0$ розбиття незадовільне, слід збільшити (зменшити?) кількість кластерів.

Постановка задачі

Деяка компанія із продажу винних виробів, маючи клієнтську базу даних, хоче розподілити список покупців на гурти за споживацькими вподобаннями (знати, хто що *завжди* замовляє) аби успішно проводити свою маркетингову політику. Відомо, що в минулому році було 32 пропозиції угод (*4_Lab_CIP_start-table.xlsx*, аркуш *DB*); є також перелік з 324-х замовлень (аркуш *Offer*).

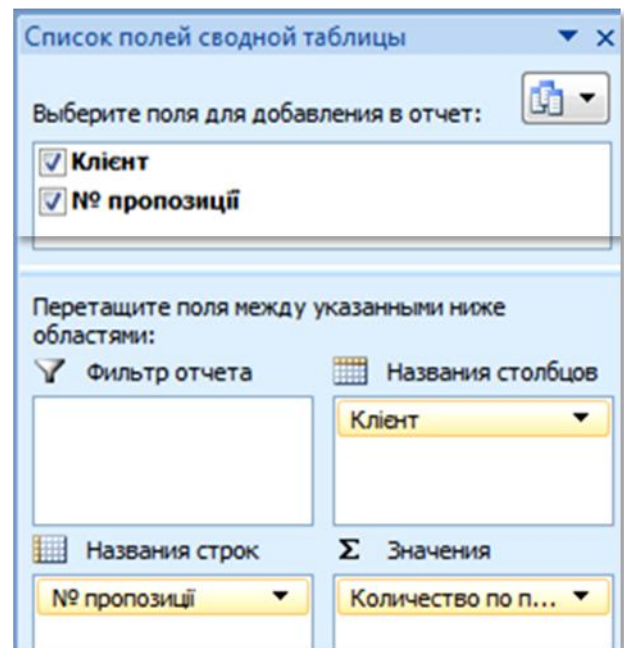
Завдання: методом k -середніх провести аналіз клієнтської бази даних з метою знайти найкраще розбиття по споживацьких гуртах.

1. З орієнтованої по рядках табл. *Offer* робимо зведену таблицю, з якої буде видно, які клієнти брали участь у тих чи інших оборудках. Виконуємо команду **Вставка** → **Сводная таблица** і у виниклому діалоговому вікні вказуємо всю табл. *Offer* як джерело даних і вибираємо опцію “Створення таблиці на новому аркуші”. (Якщо на аркуші приховано частину рядків, їх слід попередньо розкрити).

У діалоговому вікні майстра зведених таблиць (**світлина 1**) формуємо, як показано, зведену таблицю. У клітинках навпроти кожного клієнта з’явиться 1, якщо пара “клієнт – оборудка” існує, і 0 (порожня клітина) якщо ні.

Тепер скопіюємо табл. *Offer* на новий аркуш *Matrix* і туди ж вставимо значення зі зведеної табл. без №№ оборуд, бо вони вже є в інформаціях про замовлення. Вийде розширений варіант матриці *Offer* (світлина 2). Оскільки всі клієнтські вектори бінарні, координати центрів кластерів матимуть значення від 0 до 1.

2. Кількість кластерів встановимо 4, орієнтуючись, скажімо, на купівельну спроможність клієнтів: надактивні, активні, середні, низької купівельної спроможности.



Св. 1

Скопіюємо вміст *Matrix* на новий аркуш *4MC* і вставимо 4 стовпці після цінового максимуму у колонки від **Н** до **К** (світлина 3). Назвемо стовпці *Cluster 1 ... Cluster 4*.

№ предложения	Период	Сорт	Минимальное количество, кг	Скидка, %	Происхождение	После пика сезона	Adams	Allen	Anderson	Bailey	Baker	Barnes	Bell	Be
1	January	Malbec	72	56	France	ЛОЖЬ								
2	January	Pinot Noir	72	17	France	ЛОЖЬ								1
3	February	Espumante	144	32	Oregon	ИСТИНА								
4	February	Champagne	72	48	France	ИСТИНА								
5	February	Cabernet Sauvignon	144	44	New Zealand	ИСТИНА								
6	March	Prosecco	144	86	Chile	ЛОЖЬ								
7	March	Prosecco	6	40	Australia	ИСТИНА				1	1			
8	March	Espumante	6	45	South Africa	ЛОЖЬ								
9	April	Chardonnay	144	57	Chile	ЛОЖЬ	1							
10	April	Prosecco	72	52	California	ЛОЖЬ					1	1		
11	May	Champagne	72	85	France	ЛОЖЬ								
12	May	Prosecco	72	83	Australia	ЛОЖЬ								
13	May	Merlot	6	43	Chile	ЛОЖЬ								
14	June	Merlot	72	64	Chile	ЛОЖЬ								
15	June	Cabernet Sauvignon	144	19	Italy	ЛОЖЬ								
16	June	Merlot	72	88	California	ЛОЖЬ								
17	July	Pinot Noir	12	47	Germany	ЛОЖЬ								1
18	July	Espumante	6	50	Oregon	ЛОЖЬ	1							
19	July	Champagne	12	66	Germany	ЛОЖЬ					1			
20	August	Cabernet Sauvignon	72	82	Italy	ЛОЖЬ								
21	August	Champagne	12	50	California	ЛОЖЬ								1
22	August	Champagne	72	63	France	ЛОЖЬ								1
23	September	Chardonnay	144	39	South Africa	ЛОЖЬ								
24	September	Pinot Noir	6	34	Italy	ЛОЖЬ				1				1
25	October	Cabernet Sauvignon	72	59	Oregon	ИСТИНА								
26	October	Pinot Noir	144	83	Australia	ЛОЖЬ				1				1

Св. 2

G	H	I	J	K	L	M	N	O	P
Після пика	Кластер 1	Кластер 2	Кластер 3	Кластер 4	Adams Allen Anderson Bailey Baker				
зона									
ЛОЖЬ									
ЛОЖЬ									
ИСТИНА									
ИСТИНА									
ИСТИНА									
ЛОЖЬ									
ИСТИНА								1	1
ЛОЖЬ									
ЛОЖЬ						1			
ЛОЖЬ									1
ЛОЖЬ									

Св. 3

У них мають з'явитися відстані обладунків від кластерних центрів. Поки що всі кластерні центри - нулі. Евклідова відстань об'єкта "Адамс" від кластера в інтерпретації Excel виглядає тако:

$\{= \text{КОРЕНЬ}(\text{СУММ}((\text{L}\$2:\text{L}\$33-\text{H}2:\text{H}\$33)^2))\}$

Її введемо у клітину L34 під замовленнями Адамса. Це буде відстань об'єкта "Адамс" до центру першого кластера. Результат – число 1,732. Воно має такий зміст: Адамс уклав 3 угоди, але позаяк кластерні центри – нулі, відповідь буде $\sqrt{3} = 1,732$.

Додамо Адамсові розрахунок відстаней до решти кластерних центрів, перетягнувши комірку L34 до низу на L37, а потім змінивши вручну посилання на кластерний центр зі стовпця H на стовпець I, J і K в комірках нижче. В результаті, наприклад, у L36 вийде формула

$\{= \text{КОРЕНЬ}(\text{СУММ}((\text{L}\$2:\text{L}\$33-\text{J}\$2:\text{J}\$33)^2))\}$

Дужки {} з'являються самі після натискання Ctrl+Shift+Enter. Це потрібно робити, щоб частина формули L\$2:L\$33-H2:H\$33 «знала», куди звертатися для розрахунку різниць крок за кроком

3. Розрахуємо відстань решти покупців до кластерних центрів заповнюючи L34:L36 → DG34:DG37 і введемо заголовки рядків "Відстань до кластера ..." у стовпці G:

	A	B	C	D	E	F	G	H	
34								Відстань до кластера 1	
35	Сумарна відстань від покупців до їхніх найближчих кластерних центрів								Відстань до кластера 2
36	174,6							Відстань до кластера 3	
37								Відстань до кластера 4	
38								Мінімальна відстань	
39								Найближчий кластер	

Тепер відома відстань кожного клієнта до всіх 4-х кластерних центрів.

4. Розпочнемо розподіл клієнтів по кластерах.

Розрахуємо мінімальну відстань до кластерів для Адамса у клітинці **L38**: **=МИН(L34:L37)**. Знайдемо (комірка **L39**) відносно розташування елемента з таким (мінімальним) значенням у переліку відстаней **L34:L37**:

=ПОИСКПОЗ(L38,L34:L37,0)

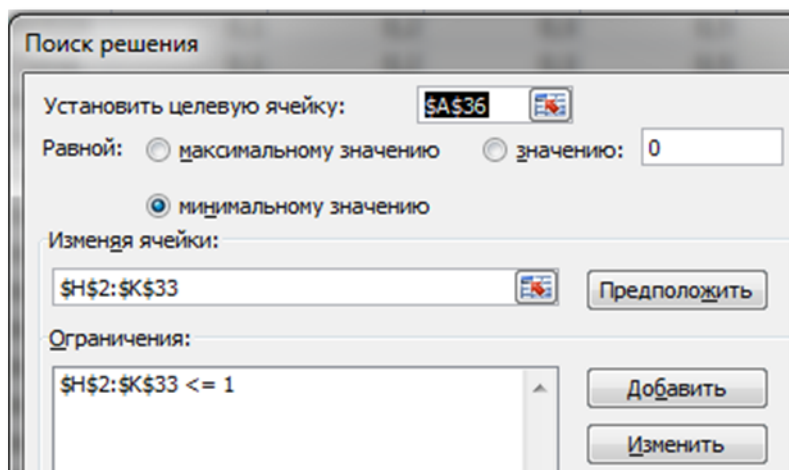
Ця функція показуватиме № найближчого до клієнта центру кластера. На разі відстані до всіх кластерів однакові, тож формула вибирає перший (**L34**) і повертає 1. Перетягнемо дві останні формули до → **DG38:DG39**. Для кожного клієнта відстані до всіх кластерів очікувано виявляються однаковими.

5. Встановимо найкраще розташування кластерних центрів. Це означає знайти такі значення в стовпцях **H:K**, які *мінімізують* загальну відстань між покупцями і центрами прив'язки, вказаних у рядку **L39**. У комірці **A36** просумуємо всі відстані:

=СУММ(L38:DG38)

Ця сума в рамках розглянутого підходу підлягає мінімізації, тобто є *цільовою* для операції Excel **Поиск решения**.

6. Шукаючи розв'язок, скористаємось **еволюційним алгоритмом**, який використовує комбінацію випадкового пошуку і “схрещування” (подібно біологічній еволюції). **Змінні задачі**: кластерні центри; **обмеження**: кластерні центри повинні мати значення від 0 до 1 (бо координати всіх обладунків не перевищують 1). Задача “Пошукові розв'язку” – мінімізувати **A36** через зміну **H2:K33** з умовою **H2:K33 ≤ 1**.



Потрібно вибрати належні опції еволюційного алгоритму, натиснувши **Параметры**. Варто встановити максимальний час пошуку більше ніж за замовчуванням, скажімо 600 секунд; вибрати опцію “зробити необмежені змінні не негативними”; розмір популяції 100; рівень мутацій 0,15; збіжність 0,00001.

Можливо перервати пошук розв'язку, натиснувши Esc, і вийти з найкращим числом, яке на той час він зміг знайти

7. По завершенню роботи алгоритму маємо у стовпцях

Н:К координати кластерних центрів та у рядку **39** – розбиття клієнтів по кластерах.

Чим темніший зелений колір комірки у рядку обладнання (координати), тим далі відстоїть кластерний центр уздовж відповідної координатної осі від початку координат. Кластер 1 у деяких рядках має майже нульові ($\leq 0,01$) координати. У наведеному прикладі це обладнання 17, 22, 24, 26 і 30. Щоб з'ясувати, яких покупців він гуртує, слід роздивитися дані про обладнання. Аналіз показує, що вказані обладнання в основному укладені на вино сорту *Pino Noir*. (Винятки: обладнання 22 на шампанське і 30 на Мальбек). Отже можна вважати, що до кластеру 1 входять покупці, які цікавляться шампанським.

Отримані кластерні центри можуть відрізнятися від поданих тут, бо еволюційний алгоритм використовує випадкові числа і відповідь виходить різною

Яких покупців об'єднує кластер 2?

Останні 2 кластери важко інтерпретувати. У такому разі радять окремо дослідити самих покупців у кластері й визначити які угоди їм до вподоби.

8. Перевіримо, як розподілені обладнання по кластерах.

Продублюємо аркуш *4МС* → *ByCluster*, пронумеруємо на ньому колонки від **Н** до **К** від 1 до 4 (фото 4) і очистимо на ньому блок **Н2:К33** з кластерними координатами.

	Е	Ф	Г	Н	І	Ј	К
1	Снижка, %	Походження	Після піку сезона	1	2	3	4
2	56	France	ЛОЖЬ				
2	17	France	ЛОЖЬ				
4	32	Oregon	ИСТИНА				
2	48	France	ИСТИНА				
4	44	New Zealand	ИСТИНА				
4	86	Chile	ЛОЖЬ				

Св. 4

Тепер тут міститиметься кількість покупців, які взяли участь у тій чи іншій обладнанні. Наприклад, у **Н2** буде кількість покупців кластера 1, що погодилися на пропозицію №1, а саме січневий Мальбек. Для цього потрібно скласти числа **L2:DG2** на аркуші *4МС*, але лише покупців з 1 кластера. Який покупець у якому кластері, видно у рядку 39.

Підрахунок за умовою проводиться функцією **СУММЕСЛИ()**. У **Н2** вводимо формулу

=СУММЕСЛИ('4МС'!\$L\$39:\$DG\$39;Н\$1;'4МС'!\$L2:\$DG2)

Заповнюємо цією формулою блок комірок **Н2:К33**. Виниклі числа є кількість обладнань у конкретному кластері.

Застосуємо **Автофільтр**. Відсортувавши колонки **Н:К**, бачимо які оборудки найпопулярніші в кожному з кластерів. Для кластеру 1 популярність **Pino Noir** підтверджується. Сортування кластеру 2 виводить на гору дрібногуртових покупців (закупка мінімальної кількості у 6 одиниць). Інтерпретація кластера 3 не однозначна. Покупцям з кластеру 4 з якоїсь причини не сподобалася серпнева пропозиція на шампанське. Також 5 з 6 найбільших оборудок – на французьке вино, а 9 з 10 перших за величиною – на великий об’єм краму.

Значки \$ вставляються вручну після появи формули для того, щоб можна було поширити її без змін на прилеглий діапазон

Ще одна особливість розподілу: кластери 3 і 4 перетинаються.

	F	G	H	I	J	K
Походження	Після піка сезону	1	2	3	4	
France	ЛОЖЬ	3	0	6	1	
France	ЛОЖЬ	2	0	2	6	
Oregon	ИСТИНА	2	0	4	0	
France	ИСТИНА	7	0	5	0	
New Zealand	ИСТИНА	2	0	2	0	
Chile	ЛОЖЬ	5	0	7	0	
Australia	ИСТИНА	4	12	3	0	
South Africa	ЛОЖЬ	9	8	3	0	
Chile	ЛОЖЬ	7	0	3	0	

9. Оскільки отриманий розподіл по 4 кластерах не може бути однозначно інтерпретований, варто спробувати метод для *іншого числа* кластерів. Аби визначитися з кількістю кластерів, перевіримо спочатку ступінь приналежності покупців до “своїх” кластерів, тобто порівняємо відстані між усіма покупцями.

Створимо аркуш *Distance*. На ньому по горизонталі та вертикалі розташуємо покупців (фото 5), скопіювавши їх з аркуша *Matrix*. (При перетворенні рядка на стовпець користуємось опцією “Транспонировать” меню “Специальная вставка”). На перетині рядка p і колонки q має з’явитися евклідова відстань між клієнтами p і q ,

$$l_{pq} = \sqrt{\sum_{i=1}^{n^2} (x_i^{(p)} - x_i^{(q)})^2}.$$

де $x_i^{(p)}$ – i -координата p -клієнта з таблиці *Matrix*.

Для заповнення таблиці *Distance* користуємось функцією **СМЕЩ(<блок комірок>, n , m)**, яка бере означений набір комірок і переміщує його на n рядків та m колонок донизу і вправо. Так, у комірці **C3** буде формула (фото 5), яка розраховує відстань l_{11} клієнта *Adams* до самого себе.

A	B	C	D	E	F	G	H	I	J	K
		0	1	2	3	4	5	6	7	
		Adams	Allen	Anderson	Bailey	Baker	Barnes	Bell	Bennett	Brooks
0	Adams	=КОРЕНЬ(СУММ((СМЕЩ(Matrix!\$H\$2:\$H\$33;0;Distance!\$C\$1)-СМЕЩ(Matrix!\$H\$2:\$H\$33;0;Distance!\$A3))^2))								
1	Allen	2,236	0,000	2,000	2,000	2,449	2,449	2,449	2,000	2,449
2	Anderson	2,236	2,000	0,000	2,000	2,449	2,449	1,414	2,000	2,449
3	Bailey	1,732	2,000	2,000	0,000	2,000	2,449	2,449	2,000	2,449
4	Baker	2,646	2,449	2,449	2,000	0,000	2,000	2,828	2,449	2,828
5	Barnes	2,646	2,449	2,449	2,449	2,000	0,000	2,828	2,449	2,449
6	Bell	2,646	2,449	1,414	2,449	2,828	2,828	0,000	2,449	2,828
7	Bennett	1,732	2,000	2,000	2,000	2,449	2,449	2,449	0,000	2,000
8	Brooks	2,646	2,449	2,449	2,449	2,828	2,449	2,828	2,000	0,000
9	Brown	1,414	2,236	2,236	1,000	2,236	2,646	2,646	1,732	2,646
10	Butler	2,449	2,646	2,646	2,236	3,000	2,646	3,000	2,646	2,646

Св. 5

10. Скопіюємо покупців та їхні гуртові прив'язки з аркуша 4MC на новий аркуш *Silhouette* до стовпців A і B відповідно. Надамо стовпцям C:F назви “Відстань до клієнтів в кластері 1, 2, 3, 4”. У комірці C2 розрахуємо середню відстань клієнта “Адамс” (він у кластері 2) до клієнтів у кластері 1. Приналежність покупця до кластера 1 визначаємо з 39-го рядка аркуша 4MC), і а відстань до Адамса беремо з аркуша *Distances*:

=СРЗНАЧЕСЛИ('4MC'!\$L\$39:\$DG\$39;1;Distance!\$C3:\$CX3).

Для стовпців D:F формула та ж, тільки № кластера замінюється на 2, 3, 4. Копіюємо ці 4 формули на всіх покупців (св. 6).

У стовпці G розрахуємо найближчий гурт покупців. Для Адамса це

=МИН(C2:F2).

A	B	C	D	E	F	G
Клиент	Кластер	Відстань до клієнтів у кластері 1	Відстань до клієнтів у кластері 2	Відстань до клієнтів у кластері 3	Відстань до клієнтів у кластері 4	Найменше
Adams	2	2,285	1,461	2,688	2,374	1,461
Allen	1	1,974	2,224	2,476	2,127	1,974
Anderson	4	2,143	2,224	2,558	1,130	1,130
Bailey	2	2,085	1,477	2,462	2,150	1,477
Baker	1	2,337	2,418	2,703	2,556	2,337
Barnes	3	2,417	2,639	2,345	2,556	2,345
Bell	4	2,551	2,639	2,897	1,296	1,296

Св. 6

У стовпці H розрахуємо другий за близькістю гурт покупців.

=НАИМЕНЬШИЙ(C2:F2;2)

Відстань до членів свого кластера у стовпці I:

=ИНДЕКС(C2:F2, B2).

Силует (2) розраховуємо у стовпці **J**, замінивши в нього функцію

=ЕСЛИ(I2=G2, H2, G2)

яка порівнює власну кластерну відстань у **I** з найближчим кластером у **G**, і якщо вони співпадуть, то відповідь – **H**. У протилежному випадку це **G**.

Можливий результат наведено в табл. нижче.

A	B	C	D	E	F	G	H	I	J
Клієнт	Кластер	Відстань до клієнтів у кластері 1	Відстань до клієнтів у кластері 2	Відстань до клієнтів у кластері 3	Відстань до клієнтів у кластері 4	Найменше	Наступне найменше	Мій кластер	Найближчий кластер
Adams	2	2,285	1,461	2,688	2,374	1,461	2,285	1,461	2,285
Allen	1	1,974	2,224	2,476	2,127	1,974	2,127	1,974	2,127
Anderson	4	2,143	2,224	2,558	1,130	1,130	2,143	1,130	2,143
Bailey	2	2,085	1,477	2,462	2,150	1,477	2,085	1,477	2,085
Baker	1	2,337	2,418	2,703	2,556	2,337	2,418	2,337	2,418
Barnes	3	2,417	2,639	2,345	2,556	2,345	2,417	2,345	2,417
Bell	4	2,551	2,639	2,897	1,296	1,296	2,551	1,296	2,551
Bennett	1	1,923	1,730	2,534	2,150	1,730	1,923	1,923	1,730

Св. 7

У колонці **K** розраховуємо силует (2) кожного покупця:

=(J2-I2)/МАКС(J2,I2)

Аналізуючи числа в **K**, встановіть, які клієнти перебувають не в своєму кластері.

Після цього усереднимо значення в **K**, щоб отримати підсумковий силует. У розглядуваному прикладі це 0,149 (Св. 7), що набагато ближче до 0 ніж до 1. Це вказує на те, що слід вибрати інше число кластерів k .

Повторіть розрахунки для іншого числа кластерів.

Примітка. Повністю переробляти розрахунки для іншого числа кластерів (наприклад 5) не потрібно. Слід зробити копію аркуша 4МС → 5МС і додати туди стовпець для 5-го кластера. Після цього переглянути на наступних аркушах посилання на комірки.

Питання самоконтролю

1. З якою метою проводять кластерний аналіз ?
2. Аналіз методом k-середніх для деякої множини об'єктів дає такі значення силуету: $s = -0,15; 0,55; 0,78$ при розбитті на 3, 4 або 5 кластерів, відповідно. На якому числі розбиття слід зупинитися?
3. Яких покупців об'єднує кластер 2 у п. 7?

ЛІТЕРАТУРА

Основна

1. Волошин О.Ф., Мащенко С.О. Моделі та методи прийняття рішень. Навч. посібник. Видав.-поліграф. центр "Київський університет", 2010.
2. Герасимов Б.М., Дивизинюк М.М., Субач І.Ю. Системы поддержки принятия решений: проектирование, применение, оценка эффективности. Севастополь, 2004.
3. Данилевич С.Б., Дьячкова О.В. Современные информационные технологии в экономике. Бизнес-анализ данных средствами аналитической платформы Deductor. – Харьков: Изд-во НУА, 2013.

Додаткова

1. Настасюк В.А. Теорія рішень. Конспект лекцій. Одеса, 2010.
2. F. Burstein, C. Holsapple. (Editors) Handbook on Decision Support Systems 1, 2. Springer, 2008.