

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ БІОРЕСУРСІВ
І ПРИРОДОКОРИСТУВАННЯ УКРАЇНИ

НУБІП України

Факультет інформаційних технологій

НУБІП України

УДК 004.9:631.544.4

«ПОГОДЖЕНО»

Декан факультету

«ДОПУСКАЄТЬСЯ ДО ЗАХИСТУ»

Завідувач кафедри комп'ютерних наук

інформаційних технологій

НУБІП України

Глазунова О.Г., д.п.н., професор

Голуб Б.Л., к.т.н., доцент

2022 р.

2022 р.

МАГІСТЕРСЬКА КВАЛІФІКАЦІЙНА РОБОТА

НУБІП України

на тему «Система підтримки прийняття рішень керівництвом тепличного господарства промислового типу»

Спеціальність 122 «Комп'ютерні науки»

Освітня програма «Інформаційні управляючі системи та технології»

(назва)

НУБІП України

Орієнтація освітньої програми освітньо-професійна

Гарант освітньої програми

к.е.н., доцент

Густера О.М.

Керівник магістерської кваліфікаційної роботи

к.т.н., доцент

Голуб Б.Л.

НУБІП України

Виконала

Лендел М.І.

КИЇВ-2022

НУБІП України

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ БІОРЕСУРСІВ
І ПРИРОДОКОРИСТУВАННЯ УКРАЇНИ
Факультет інформаційних технологій

ЗАТВЕРДЖУЮ
Завідувач кафедри комп'ютерних наук

к.т.н., доцент
(науковий ступінь, вчене звання) (підпис) (ПШБ)
Голуб Б.Л.
"01" листопада 2021 року

ЗАВДАННЯ
ДО ВИКОНАННЯ МАГІСТЕРСЬКОЇ КВАЛІФІКАЦІЙНОЇ РОБОТИ СТУДЕНТУ

Лендел Марині Іванівні
(прізвище, ім'я, по батькові)

Спеціальність 122 «Комп'ютерні науки»
Освітня програма «Інформаційні управляючі системи та технології»
Орієнтація освітньої програми: освітньо-професійна
Тема магістерської кваліфікаційної роботи «Система підтримки прийняття рішень керівництвом тепличного господарства промислового типу»
затверджена наказом ректора НУБіП України від "01" листопада 2021 р. №1862 «С»

Термін подання завершеної роботи на кафедру 27 жовтня 2022 р.

Вихідні дані до магістерської кваліфікаційної роботи:

1. дані про врожайність с/г культур;
2. дані про мікроклімат у процесі вирощування с/г культур.

Перелік питань, що підлягають дослідженню:

1. Аналіз предметної області
2. Дослідження інструментів OLAP
3. Проектування системи
4. Дослідження інструментів Data Mining
5. Розробка алгоритмів аналізу даних
6. Дослідження отриманих результатів

Дата видачі завдання "01" листопада 2021 р.

Керівник магістерської кваліфікаційної роботи

(підпис)

Голуб Б.Л.

(прізвище та ініціали)

Завдання прийняв до виконання

(підпис)

Лендел М.І.

(прізвище та ініціали)

ЗМІСТ	ЗМІСТ	3
ЗМІСТ.....		3
ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ		5
ВСТУП		6
1 СИСТЕМНИЙ АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ.....	1 СИСТЕМНИЙ АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ.....	10
1.1. Аналіз процесу керування тепличним господарством.....	1.1. Аналіз процесу керування тепличним господарством.....	10
1.2. Технічне завдання.....	1.2. Технічне завдання.....	14
1.3. Аналіз наявних рішень.....	1.3. Аналіз наявних рішень.....	15
1.4. Моделювання процесу керування тепличним господарством.....	1.4. Моделювання процесу керування тепличним господарством.....	18
2 ВИКОРИСТАНІ ТЕХНОЛОГІЇ	2 ВИКОРИСТАНІ ТЕХНОЛОГІЇ	26
2.1. Визначення поняття СППР	2.1. Визначення поняття СППР	26
2.2. Багатовимірні структури для збереження даних.....	2.2. Багатовимірні структури для збереження даних.....	28
2.3. Data Mining	2.3. Data Mining	29
2.4. OLAP	2.4. OLAP	30
2.5. Використані інструменти.....	2.5. Використані інструменти.....	32
3 РОЗРОБКА СИСТЕМИ	3 РОЗРОБКА СИСТЕМИ	33
3.1. Архітектура системи.....	3.1. Архітектура системи.....	33
3.2. Оперативні джерела інформації.....	3.2. Оперативні джерела інформації.....	36
3.3. Сховище даних.....	3.3. Сховище даних.....	38
3.4. Побудова розгорнутого кубу підсистеми аналізу.....	3.4. Побудова розгорнутого кубу підсистеми аналізу.....	41
3.5. Реалізація отримання даних за допомогою Data Flow.....	3.5. Реалізація отримання даних за допомогою Data Flow.....	43
3.6. Побудова звітності.....	3.6. Побудова звітності.....	45
3.7. Розрахунок KPI.....	3.7. Розрахунок KPI.....	47
3.8. Розробка алгоритмів інтелектуального аналізу даних.....	3.8. Розробка алгоритмів інтелектуального аналізу даних.....	49

4 РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ 55

4.1. Оперативний аналіз даних 55

4.2. Інтелектуальний аналіз даних 59

ВИСНОВКИ..... 66

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ..... 70

НУБІП України

НУБІП України

НУБІП України

НУБІП України

НУБІП України

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

НУБІП України

DM – data mining

KPI – key performance indicator (ключовий показник ефективності)

НУБІП України

OLAP – online analytical processing

SQL – structured query language

SSAS – SQL Server Analysis Services

SSDT-BI – SQL Server Data Tools-Business Intelligence

SSIS – SQL Server Integration Services

НУБІП України

SSRS – SQL Server Reporting Services

UML – unified modeling language

БД – база даних

ВВП – валовий внутрішній продукт

НУБІП України

СД – сховище даних

СППР – система підтримки прийняття рішень

НУБІП України

НУБІП України

НУБІП України

ВСТУП

НУБІП України

Актуальність

Сучасний світ повністю залежить від технологій, які стрімко розвиваються. Вони займають ключове місце майже у всіх напрямках нашої діяльності та допомагають у повсякденному житті майже на кожному кроці. Уся промисловість зараз автоматизована з використанням роботів, дронів, контролерів, датчиків та програмного забезпечення для більш швидкого та якісного виготовлення продукції чи надання послуг.

Саме промисловість і аграрний сектор в цілому займає значну частку у розвитку економіки України. Він забезпечує можливість експорту продукції, що у свою чергу збільшує надходження до держбюджету країни, її показник ВВП та відповідно покращує рівень життя людей.

У цій галузі важливим виступає вирощування продукції у спорудах закритого ґрунту. Із впливом сучасності для промислового вирощування овочів та фруктів у теплицях також використовуються інтелектуальні системи для моніторингу, керування та аналізу показників, які відіграють ключову роль під час вирощування рослин.

Під час вирощування важливим етапом є саме вивчення та аналіз всіх умов, які необхідні для нормального росту й розвитку рослини. У процесі росту рослин важливо враховувати оптимальні показники мікроклімату теплиці для підвищення ефективності використання ресурсів для вирощування врожаю.

Дивлячись на згадані особливості корисним є використання програмних засобів для моніторингу, збереження та аналізу показників, які важливі для підвищення ефективності вирощування. Такі розроблені системи забезпечують постійний контроль показників у спорудах закритого ґрунту, звітність про поточний стан у режимі реального часу та проведення аналізу на основі наявних даних.

На основі всіх наявних даних у системі виробник може проводити аналіз всіх ключових показників, їхні зміни та вплив у розрізі часу та приймати відповідні рішення для свого підприємства. Проте створені системи з часом

розширюються, а відповідно інформація у них теж, тому стає незручно та неефективно аналізувати попередньо внесені дані. У такому випадку існує необхідність у створенні системи, яка буде проводити аналіз показників на основі накопичених даних. Тому корисним буде проводити аналіз з використанням технологій OLAP та Data Mining.

Об'єкт та предмет дослідження

Об'єктом дослідження є вирощування овочів у спорудах закритого ґрунту. Предметом дослідження виступає процес прийняття рішень, які допоможуть керівництву теплиці підвищити ефективність у вирощуванні продукції шляхом аналізу показників, які впливають на цей процес.

Мета дослідження

Метою дослідження є визначення доцільності застосування технологій OLAP і Data Mining для підвищення ефективності вирощування овочів та фруктів у спорудах закритого ґрунту.

Зміст поставлених завдань

Дослідження та розробка системи, яка виступає предметом дослідження даної роботи, поділені на декілька етапів:

1. Дослідження предметної області: визначення актуальності, наявних рішень в обраному напрямку, проблематики та мети дослідження.
2. Дослідження та вибір інструментів, які необхідні для створення СПІР.
3. Проектування архітектури системи, її ключових вузлів.
4. Проектування та створення СД, яке зберігатиме інформацію, необхідну для проведення дослідження.
5. Ознайомлення та розробка алгоритмів СПІР для аналізу даних, результати яких допоможуть в управлінні тепличним господарством.
6. Аналіз результатів роботи розроблених алгоритмів.
7. Формулювання висновків щодо доцільності застосування обраних технологій для керування тепличним господарством промислового типу.

Методи дослідження

У процесі дослідження були використані такі методи:

1. Технології OLAP для накопичення даних у більш зручній структурі: багатомірний куб, який ефективний для проведення аналізу у розрізі різних параметрів, формування звітної інформації тощо.
2. Технології Data Mining для знаходження зв'язків між наповненими даними, які відображають результати роботи тепличного господарства.

Наукова новизна

1. Було досліджено технології OLAP і Data Mining.
2. Розроблена архітектура системи, що складається з трьох підсистем: оперативного джерела даних, сховища даних і підсистеми аналізу.
3. Як складові підсистеми аналізу були застосовані технології OLAP і Data Mining.

Апробація результатів дослідження

Основні положення та матеріали наукової роботи доповідалися та обговорювалися на науково-технічних конференціях:

1. Лендел М.І., Голуб Б.Л.: Система підтримки прийняття рішень керівництвом тепличного господарства промислового типу Збірник матеріалів XII Міжнародної науково-практичної конференції молодих вчених “Інформаційні технології: економіка, техніка, освіта”. 11-12 листопада 2021 року, НУБіП України, Київ. – С. 99. Режим доступу:

https://drive.google.com/file/d/1LhyVBHBVpMKiV3gIJFsF3EKv_n6MfY2L/view

2. Лендел М.І., Голуб Б.Л.: Система підтримки прийняття рішень керівництвом тепличного господарства промислового типу. Збірник матеріалів XIII Міжнародної науково-практичної конференції молодих вчених “Інформаційні технології: економіка, техніка, освіта”. 26-27 жовтня 2022 року, НУБіП України, Київ. – С. ? Режим доступу: ?

Структура магістерської роботи

Магістерська кваліфікаційна робота складається з 4-х розділів.

Розділ 1. Проведений аналіз та моделювання предметної області, формулювання технічного завдання та проаналізовано наявні рішення.

Розділ 2. У даному розділі було описано особливості та передумови створення СППР, обрані технології та інструменти, які використовувались для створення системи, а також описано архітектуру та особливості роботи розробленої системи.

Розділ 3. Цей розділ демонструє розробку архітектури СППР, структури СД, розгортання багатомірного кубу на основі сховища, реалізацію алгоритмів для аналізу наявних даних.

Розділ 4. Розділ містить опис щодо впровадження розробленої системи, результатів проведеного аналізу з використанням алгоритмів та сформувані висновки їхніх результатів.

Викладений матеріал наукової роботи містить:

- кількість ілюстрацій – 49
- кількість таблиць – 2
- кількість використаних джерел – 26

1 СИСТЕМНИЙ АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ

1.1. Аналіз процесу керування тепличним господарством

Однією з провідних галузей української економіки є сільське господарство. За результатами 2021-го року ця галузь вкладає більше 10 % до ВВП України, що є найбільшим відсотком серед усіх секторів. Також агропродовольча продукція складає близько 41 % за рік, що є найбільшим відсотком серед загального експорту України [1].

Для ознайомлення на рис. 1 наведено порівняльну характеристику економічної діяльності підприємств за кожним видом діяльності. Незважаючи на незначне зниження частки сільського господарства у 2021 році, воно займає 24-ге місце у ринковій економіці в Україні [2].

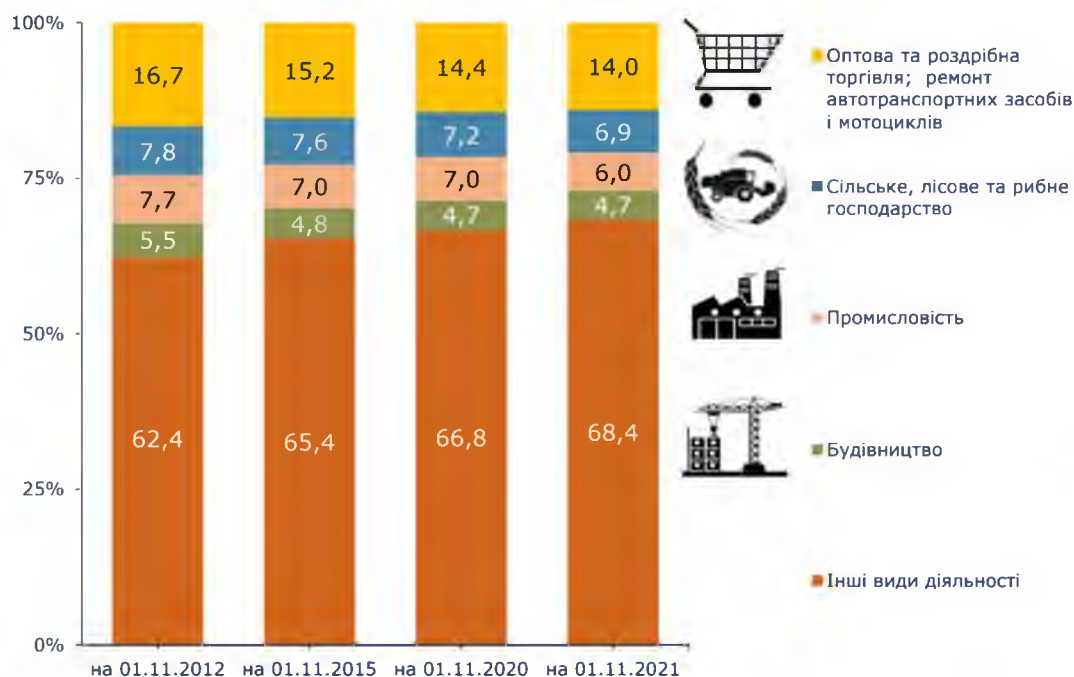


Рис. 1 Частка активних підприємств за видами економічної діяльності [2]

З огляду на розвиток сільського господарства в Україні більшу частку складає саме продукція рослинництва, врожайність якої з кожним роком збільшується у зв'язку зі збільшенням використання посівних площ та відповідно попиту на виготовлену продукцію. На рис. 2 показано структуру продукції у 2021 році у відсотках.



Рис. 2 Структура продукції сільського господарства у 2021 році

Сьогодні вирощування сільськогосподарських культур у теплицях є дуже вигідною сферою бізнесу. Такий вид діяльності характеризується хорошою окупністю, а також високою рентабельністю. В Україні тепличні господарства охоплюють близько 300 га промислових овочевих теплиць та 100 га квіткових теплиць [3].

Теплиці одні з найбільш досконалих видів культивування рослин, як можна обладнувати різними технологічними засобами для моніторингу та керування умовами для нормального росту й розвитку рослин. Споруди такого типу можна використовувати протягом всього року та за умови належного керування та догляду за рослинами отримувати велику врожайність та як наслідок вигідний прибуток.

Відповідно до агротехнічних вимог теплиці поділяють:

- за призначенням: овочеві, розсадні, квіткові, розсадно-овочеві;
- періодом експлуатації: зимові та весняні;
- способом вирощування продукції: ґрунтові та субстратні.

Теплиці також класифікують за:

- конструкційними властивостями: одноланкові (ангарні) та багатоланкові (блочні);
- матеріалом покриття: плівкові, скляні й склопластикові як із прозорим огороженням, так і непрозорим;

кількістю поверхів: одно- чи багатопверхові (баштовий тип),
 конструкцією даху: поділяються на аркові (рис. 3 (а)), двоскатні (рис. 3 (б)), односкатні (рис. 3 (в)) та нерівноскатні (рис. 3 (г)) [4].

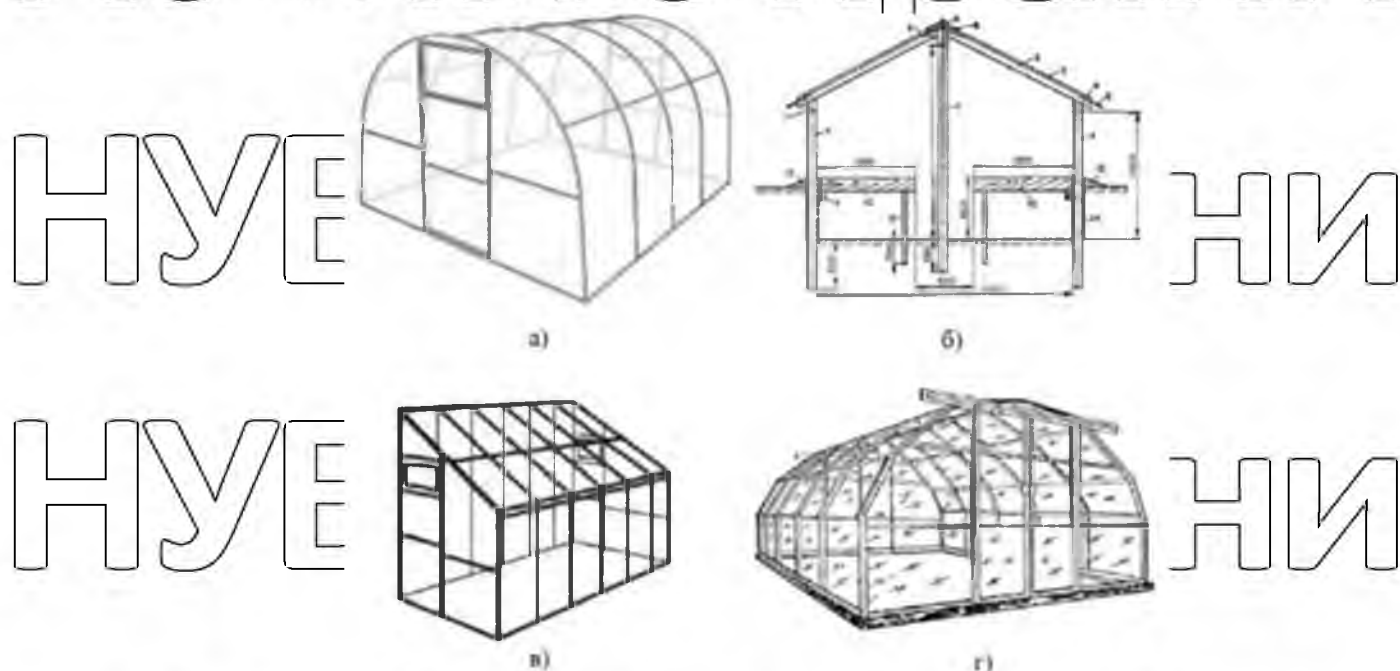


Рис. 3 Види конструкцій теплиці

Як об'єкт керування мікрокліматом, тепличні споруди відносяться до складних, враховуючи їхні умови функціонування, архітектури. Для підтримання нормального мікроклімату найбільш важливими процесами є моніторинг та управління показниками.

У процесі моніторингу важливим етапом є збереження всіх необхідних показників. Отримані значення допомагають у проведенні аналізу та прогнозування. Оброблена інформація допоможе підприємству дані про перебіг технологічних процесів, можливі відхилення тощо.

Під час вирощування продукції у спорудах закритого ґрунту важливо проводити моніторинг показників температури, вологості, освітлення, вмісту CO_2 (вуглекислий газ) та інших.

Температура виступає одним з найважливіших параметрів розвитку якісного врожаю. Проте розрізняють різні типи температури, яка впливає на ріст рослини. Розрізняють температуру рослини, повітря, а також субстрату.

Температура рослини впливає на темпи її росту та розвитку. Наприклад, температура тканини впливає на розгортання листків, розвиток бруньки та ріст стебла. Проте температуру рослини складно виміряти без інфрачервоного (ІЧ) термометру. У процесі вимірювання температури повітря можна отримати приблизну температуру рослини, а тому температура повітря є найбільш поширеним показником, який досліджується у теплиці.

Також крім температурного показника на ріст рослини та рівень врожайності впливає кількість світла. Саме кількість світла це показник, який необхідно контролювати, так як він може значно змінюватися протягом дня, місяця і року.

Існує миттєве та кумулятивне вимірювання кількості освітлення. Миттєвий рівень освітлення впливає на швидкість процесу фотосинтезу у конкретний період часу і цей параметр впливає на ріст і розвиток рослини.

Миттєве освітлення важливо контролювати для прийняття таких керуючих рішень як затінення тепличної площі чи увімкнення додаткових ламп для збільшення кількості світла.

Проте у процесі вимірювання миттєвого рівня світла виникає обмеження: існує сильна варіація впродовж дня. Спостерігаючи за результатами дослідження можна сказати, що світло, яке накопичується протягом 24 годин, важливе так само, як і миттєве світло щоб провести контроль за ростом рослини [6].

Під час вирощування рослини не менш важливим є нормальний вміст вуглекислого газу. У процесі фотосинтезу він виступає як ключовий елемент перебігу цього явища, так як за допомогою нього рослини виробляють поживні речовини, які корисні у процесі їхньої життєдіяльності. У результаті дії фотосинтезу рослини отримують вуглеводи (до прикладу цукри) та, як побічний продукт, виділяють кисень.

Важливим параметром у мікрокліматі є рівень вологості. Як занадто низька, так і занадто висока вологість всередині теплиці негативно впливає на стан рослини. Висока вологість утворює конденсат, бризки якого стікають зі стелі є джерелом поширення хвороб рослин. Також калюжі на підлозі у теплиці

сприяють росту водоростей, розмноження комах та є небезпечним фактором для роботи її працівників.

Низький рівень вологості також впливає на розвиток рослини так як через недостатню кількість води рослина буде сповільнюватись у рості та як наслідок може загинути [6].

1.2. Технічне завдання

Тепличне господарство – це підприємство, яке вирощує овочі, фрукти або розсаду для відкритого ґрунту у закритому ґрунті.

Управління тепличним господарством – це необхідні рішення та дії, які впливають на стан процесів під час виробництва, зміни роботи деяких об'єктів всередині підприємства. Під час роботи це може бути керуюче рішення щодо вирощування певного сорту, який приносить недостатній або вигідний рівень врожайності.

Для проведення аналізу процесів вирощування овочів та фруктів у тепличному господарстві потрібно розробити СІПР, яка використовуватиме СД. У процесі проектування схеми СД важливо враховувати параметри, характеристики, описані у пункті 1.1.

Для проведення аналізу тепличного господарства необхідно забезпечити зберігання таких даних:

- інформація про теплиці, а саме:
 - назва теплиці;
 - площа;
- інформація про сорт:
 - назва сорту;
 - культура, до якої даний сорт відноситься;
- оптимальні умови:
 - діапазон оптимальної температури;
 - діапазон оптимальної вологості;
- інформація про врожайність:
 - врожай.

- інформація про зафіксовану температуру;
- інформація про зафіксовану вологість.

Для ефективної роботи системи необхідно забезпечити внесення в оперативну БД вищезазначеної інформації, а також розробити структуру СД для подальшої передачі даних у неї.

Під час аналізу інструменти OLAP та Data Mining можуть надати відповідь на такий перелік питань:

- 1) Який сорт має найбільшу врожайність у розрізі вказаного часового періоду?
- 2) Яка врожайність сорту за вказаний часовий період у розрізі температури та вологості?
- 3) Наскільки часовий проміжок вирощування впливає на результуючу врожайність?
- 4) Яка тенденція зміни врожайності у теплицях?

1.3. Аналіз наявних рішень

Під час вивчення предметної області також було проаналізовано наукові роботи, які пов'язані з вирішенням питання аналізу роботи тепличного господарства. Далі наведено результати ознайомлення з наявними рішеннями.

Автори Ченг Ванг, Лілі Венг, Лінг Донг, Ксяо Кяо у своїй роботі «Дослідження системи підтримки прийняття рішень (DSS) для парників на основі видобутку даних» розглянули систему підтримки прийняття рішень (DSS)

для теплиці, побудованої зі сховища даних та технології видобутку дат. У

системі було створено сховище даних з метою диверсифікації пам'яті, використання онлайн аналітичної обробки та видобутку дат збагачує базу знань новою інформацією про сільське господарство. Впровадження системно

прийнятого аналізу SQL Server, як наслідок, зв'язок герметичності сховища

даних, видобутку дат та застосування, підвищення ефективності видобутку дат.

Комбіноване сховище даних з онлайн аналітичною обробкою та видобуванням дат для створення нового DSS.

Дана робота показує структуру системи підтримки прийняття рішень загалом: використання технології аналітичної обробки OLAP, яка допомагає аналітично обробляти дані, які зберігаються у базі знань. Також пояснює структуру модуля видобутку даних, а саме побудову окремих модулів, кожен з яких відповідає за моделювання та прогнозування, пошуку схожості, побудови асоціативних правил тощо [7].

В іншій роботі «Експертна система управління тепличним виробництвом» автори Йонгуанг Ху, Джиганг Ванг та Пінгпінг Лі розповіли про важливість використання теплиці та правильного управління всіма внутрішніми процесами і це підштовхує до розробки експертних систем.

У роботі описано та зображено функціональну структуру системи, яка складається з декількох підсистем. Важливу частину займає база знань, яка містить словник правил. Також було описано систему збору даних у режимі реального часу.

У результаті аналізу було описано короткі функціональні характеристики системи:

- Моніторинг парникового середовища в режимі реального часу;
- Оновлення інформаційної бази довідки для експертної системи;
- Виведення та відображення інформації з текстом, графічною таблицею або роздрукованою;
- Тривожний сигнал у разі екстремальних випадків або стресу врожаю.

Крім програмного у роботі було розглянуто апаратне забезпечення та впровадження системи та її результати роботи, де показано приклад діагностики захворювання на основі введених ознак, а також аналізу оптимальних умов мікроклімату та поливу.

Дана робота дала можливість розглянути цілком працездатну систему, яка допомагає керувати теплицею. Розглянуто програмну та апаратну структуру, використані технології, що допоможе у розробці власної системи [8].

Також у такому напрямленні працювали науковці Власенко Л. О.; Ладанюк А. П.; Іващук В. В. над патентом «Підсистема підтримки прийняття рішень для

керування технологічними комплексами». У цьому патенті описано СППР, у якій міститься модуль прийняття рішень, сформована база знань, модуль експертних оцінок, прогнозування та діагностики. У режимі реального часу із використанням модулів прогнозування і діагностики аналізують у якому стані перебуває система, роблять прогнозування можливих відхилень від вказаного режиму і шукають схожі факти використовуючи базу знань. Якщо схожі ситуації відсутні то за допомогою методу експертних оцінок надаються пропозиції з можливими станами системи. Далі використовуючи результати блок прийняття рішень генерує корегуючі рішення надалі заносивши отримані дані у базу знань.

Дана робота є корисною для процесу пошуку та аналізу вхідної інформації і пошуку можливих рішень та процесу збереження у базу знань для подальшого використання у процесі роботи [9].

Серед розглянутих робіт також представлена корисна модель авторів Дудник А. О.; Гачковська М. А.; Лендел Т. І.; Заєць Н. А.; Якименко І. Ю.; Комарчук Д. С. «Спосіб регулювання мікроклімату в теплиці з використанням алгоритму синтезу нейромережевого регулятора на основі заданого синергетичного закону керування». У її основі стоїть задача вдосконалити способи регулювання мікроклімату у теплиці, використовуючи алгоритм синтезу нейромережевого регулятора із заданим синергетичним законом управління. Алгоритм заснований на поведінці синергетичного контролера, який моделює показники температури й вологості за допомогою штучної нейронної мережі у результаті навчання. Особливістю створеного підходу до синтезу інтелектуальної системи управління тепличними комплексами є комбінація принципів єдності самоорганізації і навчання нейронної мережі з попереднього етапу. Через таке поєднання надалі забезпечується стійка робота системи інтелектуального керування вирощування сільськогосподарської продукції.

Дана робота є корисною у процесі аналізу правильного регулювання параметрів мікроклімату теплиці, які досліджуються [10].

Проаналізувавши існуючі рішення було зроблено висновок що багатьом з них не вистачає використання технологій аналізу да допомогою OLAP та Data

Mining. Даний вид аналізу допоможе підвищити якість керуючих рішень для управління підприємством такого напрямлення.

1.4. Моделювання процесу керування тепличним господарством

На етапі проектування системи важливо з'ясувати які вимоги поставлені, щоб наступний етап розробки проходив результативно та коректно. Для більш якісного аналізу предметної області та вимог до створюваної системи у процесі проектування розроблюються діаграми, які дають зрозуміти загальну картину плану роботи.

1.4.1. Діаграма прецедентів. У нотатції UML діаграма прецедентів дозволяє показати поведінку системи, що демонструє перелік вимог до неї. Дана діаграма демонструє зв'язки між акторами і прецедентами, наявними у системі, яка встановлює свої обмеження. Ключовими елементами діаграми прецедентів є прецеденти й актори.

Актором у діаграмі прецедентів виступає користувач або об'єкт системи, який взаємодіє з нею.

Прецедент — це дія, яку у системі виконує конкретний актор. У діаграмі варіантів використання певний прецедент може бути пов'язаний із акторами напряду або з іншими прецедентами, від дії яких залежить виконання конкретного прецеденту.

У нотатції UML діаграма прецедентів може мати такі типи взаємодії:

- розширення;
- включення;
- використання;
- вимога;
- рівнозначність.
- схожість;

Так як актор може бути напряду пов'язаний з прецедентом або через інший прецедент, то у випадку відсутності такого прямого зв'язку такий прецедент стає абстрактним.

При побудові діаграми прецедентів для системи спочатку необхідно з'ясувати, які актори взаємодіють з нею. На таблиці 1 вказаний перелік акторів, які беруть участь у взаємодії з розроблюваною системою.

Таблиця 1

Актори, які взаємодіють з системою

Актор 1	Опис 2
Оператор	Спостерігає за показниками температури та вологості у теплиці, якщо необхідно, регулює рівень вручну, опрацьовує отримані значення з датчиків, вносить інформацію про результуючу врожайність
Аналітик	Проводить аналіз роботи підприємства та формує аналітичну звітність
Директор підприємства Блок керування	Приймає рішення щодо управління господарством Зчитує дані з датчиків і управляє роботою всіх виконавчих механізмів
Датчик температури	Проводить вимірювання значень температури у теплиці
Датчик вологості	Проводить вимірювання значень вологості у теплиці

З'ясувавши які актори діють у розроблюваній системі, можна перейти до опису прецедентів, які характеризують функціонал системи (таблиця 2). На основі всіх обов'язків акторів та необхідних прецедентів побудована діаграма прецедентів, з якою можна ознайомитися на рис. 4.

У зв'язку з тим, що мета розроблюваної системи зосереджена на аналізі зафіксованих даних головним актором виступає аналітик. Саме він відповідає за ведення аналізу всіх показників та взаємозв'язків між ними. На основі аналітичної звітності директор підприємства впроваджує необхідні зміни в

зазначених бізнес-процесах. Також у системі акторами виступають оператор, блок керування, датчики вологості та температури. Блок керування та датчики, апаратне забезпечення системи виконують роль моніторингу мікроклімату в тепличному господарстві. Працівник теплиці бере участь у ручному управлінні та регулярному внесенні необхідних даних.

Таблиця 2

Перелік прецедентів у системі

Актор	Найменування	Формулювання
1	2	3
Датчики температури, вологості	Вимірювання показників мікроклімату теплиці	Моніторинг показників температури і вологості у теплиці
Вентилятор, нагрівач	Регулювання показників мікроклімату теплиці	Можливість змінення значень у температури та вологості
Блок керування	Отримання даних з датчиків	Зчитує дані з датчиків температури та вологості
Блок керування	Управління роботою виконавчих механізмів	Можливість керування станом вентилятора і нагрівача
Оператор	Формування щоденних звітів	Можливість створення звітів використовуючи відфільтровані значення за необхідними критеріями
Оператор	Спостереження за роботою виконавчих механізмів	Спостерігає за станом та роботою автоматизованої системи керування

Таблиця 2 (завершення)

1	2	3
Оператор	Ручне управління пристроями	Можливість зміни стану роботи виконавчих механізмів у вручну
Оператор	Внесення врожайності	Можливість внесення врожайності за певний період
Аналітик	Формування аналітичних запитів	Створення запитів для проведення аналізу даних
Аналітик	Обчислення показника ефективності (KPI)	Проведення обчислень ключових показників ефективності
Аналітик	Формування аналітичної звітності	Можливість створення звітів з результатами проведеного аналізу даних
Аналітик	Перегляд аналітичної звітності	Опрацювання створених звітів з проаналізованими даними
Директор підприємства	Прийняття рішень щодо управління господарством	Процес прийняття необхідних керуючих рішень у роботі тепличного господарства на основі проведеного аналізу

Опис прецеденту «Прийняття рішень щодо управління господарством»

Прецедент: прийняття рішень щодо управління господарством.

Головна дійова особа: директор підприємства.

Передумова: Аналітик повинен сформулювати звіт, переглянути та надати директору підприємства.

Основний успішний сценарій:

1 Аналітик проводить аналіз роботи тепличного господарства.

2 Формування припущень щодо залежності вхідних параметрів у роботі господарства (Альтернативний потік E-1).

3 Аналітик передає пропозиції щодо можливих керуючих рішень директору підприємства.

Альтернативні потоки:

E-1 – Прийняття рішення про тестове впровадження змін у роботі теплиці для перевірки припущення.

1.1 Вибір об'єкту дослідження, часового проміжку, умов та узгодження обраної стратегії з директором підприємства.

1.2 Проведення тестового дослідження.

1.3 Аналіз результатів дослідження.

1.4 Формування висновків щодо отриманих результатів дослідження.

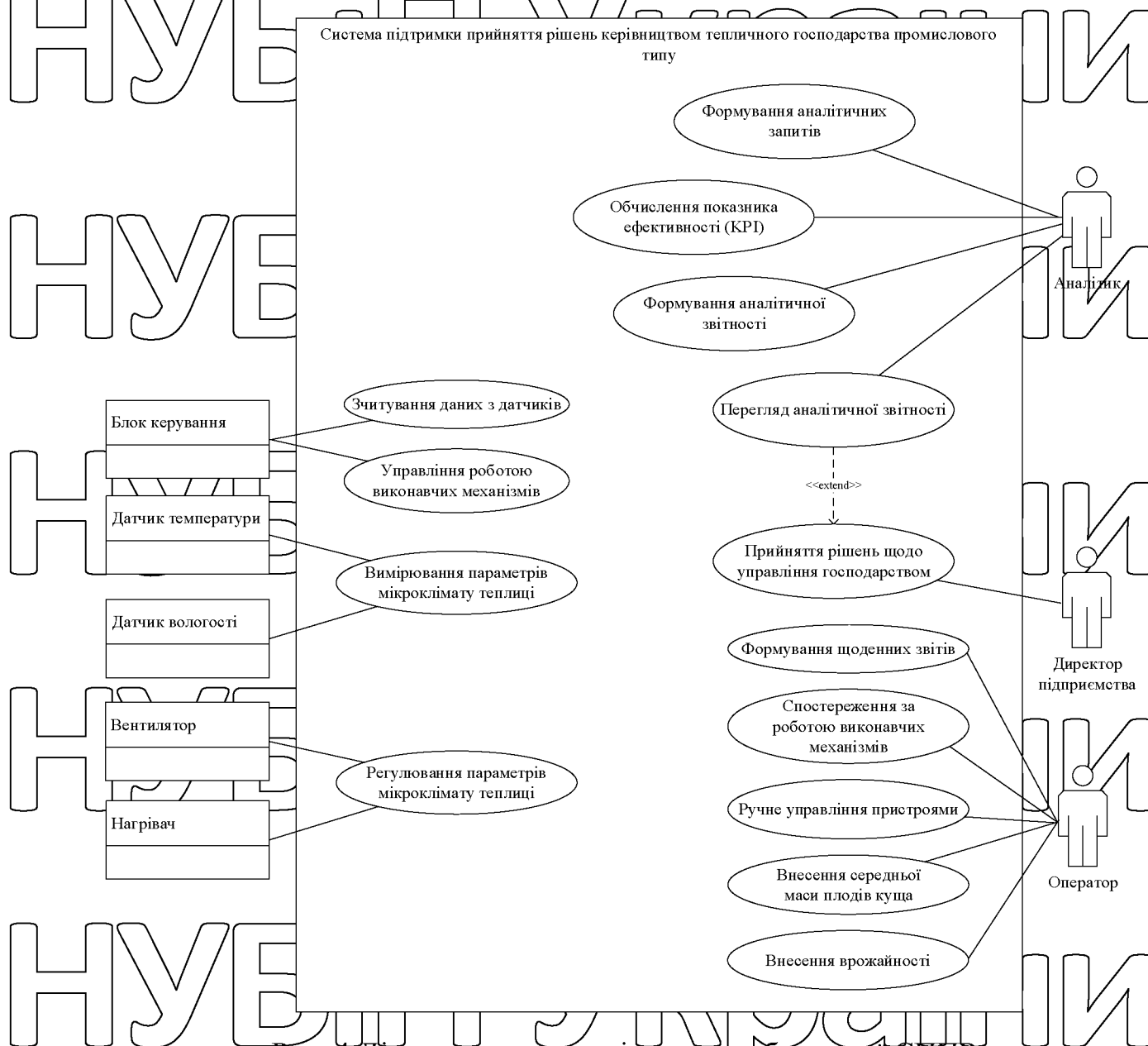


Рис. 4 Діаграма прецедентів до розроблюваної СППР

1.4.2 *Діаграма діяльності.* Для кращого розуміння принципів роботи розробленої системи будують діаграму діяльності. Дана діаграма має схожість з блок-схемою алгоритму. Вона застосовується, щоб описати поведінку і дії актору, у якого є своя доріжка з активністю. Кожен стан, показаний на діаграмі, демонструє виконання конкретної дії, а перехід до кожної наступної активності відбувається тільки за умови завершення попередньої.

Основними елементами діаграми діяльності є (рис. 5):

- прямокутник з округленими кутами відображає певну дію;
- ромб показує умову;
- вузький прямокутник означає розпаралелювання або об'єднання паралельних дій;
- чорний круг вказує на початок виконання дії;
- чорний круг у колі демонструє завершення певної діяльності;
- стрілка показує перехід однієї діяльності до іншої.



Рис. 5 Ключові елементи діаграми діяльності

На рис. 6 представлена діаграма діяльності всієї системи. На даній діаграмі можна побачити, що у системі є автоматизоване вимірювання необхідних показників, робота з вимірними і збереженими оперативними даними, чим займається оператор теплиці.

На основі збережених оперативних даних, які демонструють зміни показників у розрізі часу, відбувається аналіз директором підприємства та передача інформації до аналітика. Далі використовуючи накопичені знання аналітик формує запити для пошуку залежностей показників із результативною врожайністю на підприємстві. Від результатів аналізу залежить подальше

прийняття необхідних керуючих рішень для вектору роботи та розвитку тепличного господарства.

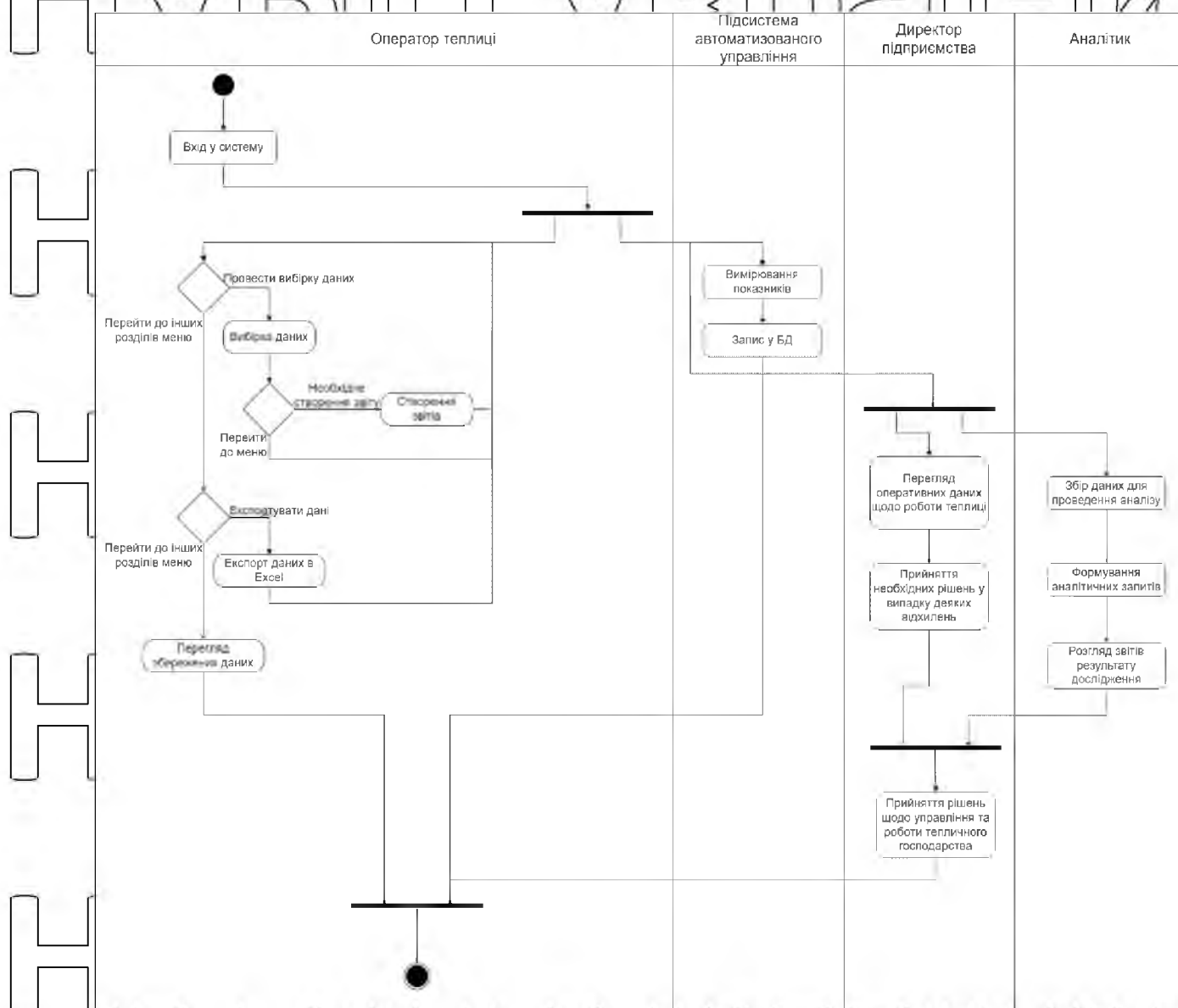


Рис.6 Діаграма діяльності

1.4.3. *Діаграма послідовності.* Діаграма послідовності використовується для опису алгоритмів дій та можливостей діючих акторів у системі, але у процесі невідомо які саме дані використовуються взаємодії з системою. Для розуміння кроків у роботі системи на етапі аналізу і проектування створюється діаграма послідовності.

Діаграма послідовності демонструє обмін даними, повідомленнями між акторами, які взаємодіють з системою. Такий спосіб відображення дозволяє

зрозуміти всі деталі процесу, яке повідомлення отримує чи надсилає конкретний актор, а також якою є його роль у дочірніх процесах.

Кожен актор, який взаємодіє з дочірньою системою, має свою лінію життя, яка вказує період часу, протягом якого він діє у ній для отримання та передачі

повідомлення. Якщо конкретний актор активний у системі протягом всього періоду роботи, то лінія життя починається від самого верху (початок діаграми) і закінчується у самому низу (кінець діаграми).

Протягом «життя» у системі актор може перебувати в активному та пасивному стані, що може бути відображено використовуючи фокус управління

вузький прямокутник, у якому верхня сторона показує на початок діяльності, а нижня – на кінець. Актор може мати декілька таких фокусів управління, але якщо він активний у системі на постійній основі, то його лінія життя може бути повністю показана використовуючи фокус управління.

На рис. 7 показано діаграму послідовності для процесу аналізу. На основі даних, експортованих з системи моніторингу, відбувається наповнення СД. Даті аналітик формує запити для аналізу фактів та формує звітну інформацію, на основі якого керівництво тепличного підприємства приймає необхідні рішення.

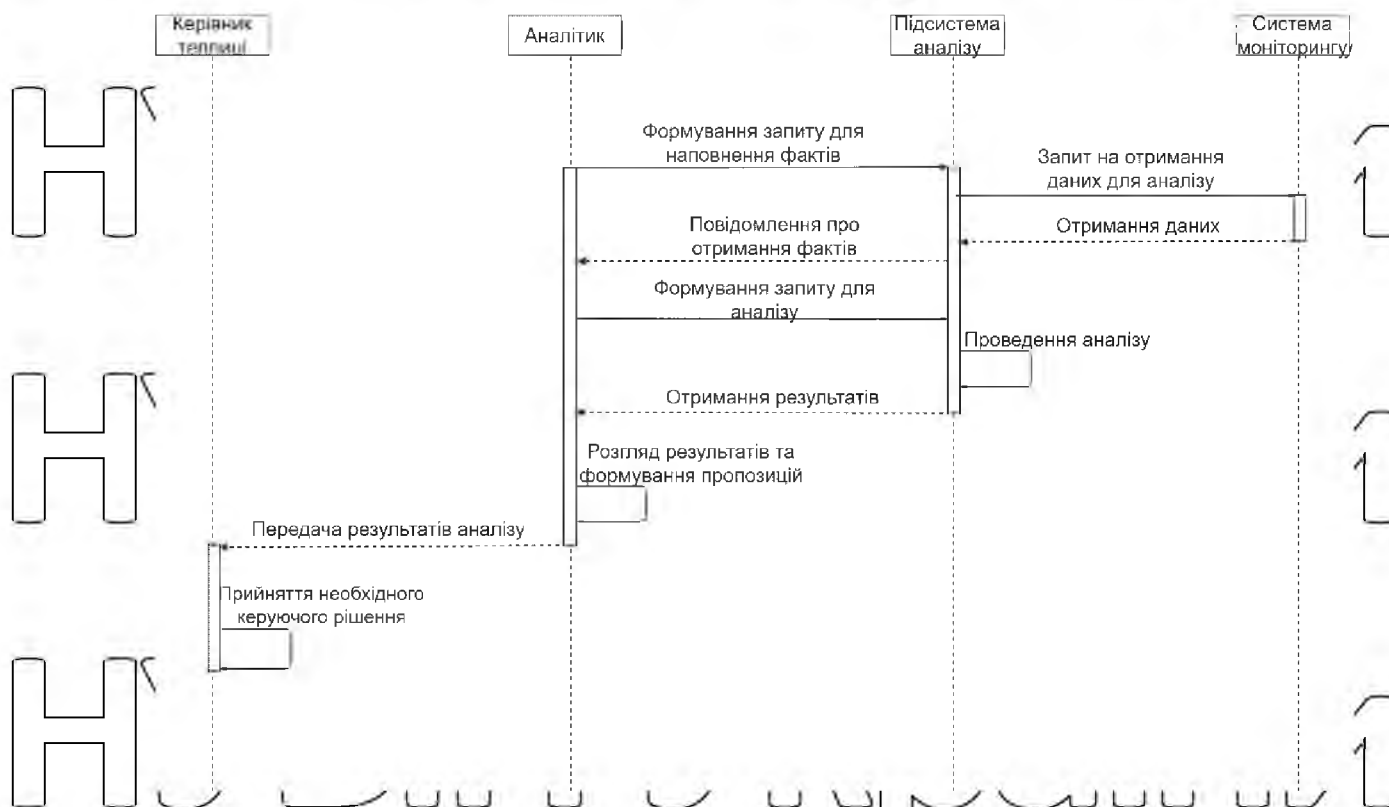


Рис. 7 Діаграма послідовності

2. ВИКОРИСТАНІ ТЕХНОЛОГІЇ

НУБІП УКРАЇНИ

2.1. Визначення поняття СППР

Система підтримки прийняття рішень (СППР) — це інформаційна система, яка допомагає підприємствам приймати рішення, які вимагають оцінки, рішучості та послідовності дій. Така інформаційна система допомагає керівництву середнього і високого рівня організації, аналізуючи великі обсяги неструктурованих даних і накопичуючи дані, які можуть допомогти у вирішенні певної проблеми та допомогти у прийнятті рішення. СППР може бути автоматизованою, керованою людиною або у поєднанні обох властивостей [11].

Система підтримки прийняття рішень створює детальні інформаційні звіти шляхом збору й аналізу даних. Таким чином, СППР відрізняється від звичайної оперативної програми, так як її метою є не збір даних, а саме проведення аналізу.

На підприємствах СППР використовується керівником і топ-менеджментом підприємства, які збирають дані та створюють звіти, які можуть використовуватися менеджерами для прийняття рішень. В основному СППР використовується для прогнозування необхідних явищ, даних про наявні та можливі факти, а також для представлення інформації користувачам у зрозумілій формі.

Теоретично СППР можна використовувати в різних сферах знань від комерційної організації до управління у таких специфічній галузі як медицина.

Одним із основних застосувань СППР в організації є звітність у реальному часі.

Це може бути дуже корисним для підприємств, які беруть участь у управлінні запасами, використовуючи підхід «саме вчасно».

У системі інвентаризації «саме вчасно» організації потрібні дані в режимі реального часу про рівень запасів, щоб розмістити замовлення у необхідний момент, щоб запобігти затримкам у виробництві та викликати негативний ефект доміно. Таким чином, СППР більше пристосована до особи чи організації, яка приймає рішення, ніж традиційна система.

Типи систем підтримки прийняття рішень:

- Керований комунікацією: дозволяє компаніям підтримувати завдання, над виконанням яких потрібно працювати більше ніж одній людині. Містить такі інтегровані інструменти, як Microsoft SharePoint Workspace і Google Docs.

- На основі моделі: дозволяє отримати доступ до фінансових, організаційних і статистичних моделей і керувати ними. Дані збираються, а параметри визначаються на основі інформації, наданої користувачами. Інформація створюється в моделі прийняття рішень

для аналізу ситуацій.

- На основі знань: надає фактичні та спеціалізовані рішення для ситуацій, використовуючи збережені факти, процедури, правила або інтерактивні структури прийняття рішень, такі як блок-схеми.

- Керований документами: керує неструктурованою інформацією в різних електронних форматах.

- На основі даних: допомагає компаніям зберігати й аналізувати внутрішні та зовнішні дані.

Переваги системи підтримки прийняття рішень

- Система підтримки прийняття рішень підвищує швидкість і ефективність прийняття рішень. Це можливо, оскільки СИПР може збирати та аналізувати дані у режимі реального часу.

- Сприяє навчанню всередині підприємства, оскільки необхідно розвинути спеціальні навички для впровадження та використання СИПР в організації.

- Автоматизує монотонні управлінські процеси, а значить керівник може використовувати більше часу на прийняття рішень.

Недоліки системи підтримки прийняття рішень:

- Вартість розробки та впровадження СИПР це величезні капіталовкладення, що робить його менш доступним для невеликих організацій.

- Компанія може залежати від СППР, оскільки вона інтегрована в щоденні процеси прийняття рішень для підвищення ефективності та швидкості. Однак менеджери, як правило, надто покладаються на систему, що позбавляє суб'єктивного аспекту прийняття рішень.

- СППР може призвести до перевантаження інформацією, оскільки інформаційна система має тенденцію розглядати всі аспекти проблеми [11].

На рис. 8 зображено ключові елементи необхідні для розробки СППР. З використанням сучасних інформаційних технологій саме підсистема аналізу допомагає аналізувати отримані дані. Вона містить дослідження використання технологій OLAP та Data Mining з можливістю інтелектуального аналізу даних.

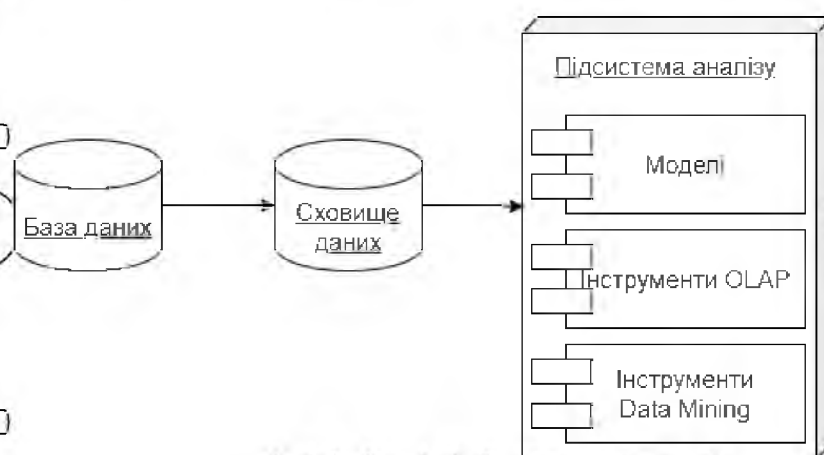


Рис. 8 Архітектура СППР

2.2. Багатовимірні структури для збереження даних

Сховище даних – це тип системи управління даними, яка призначена для забезпечення та підтримки проведення бізнес-аналітики. Сховища даних призначені виключно для виконання запитів та аналізу та часто містять великі обсяги інформації. Дані в сховищі даних зазвичай отримують з великої кількості джерел, таких як файли журналів програм і програми транзакцій, інформація з БД тощо.

Сховище даних централізує та консолідує великі обсяги даних із декількох джерел. Його аналітичні можливості дозволяють організаціям отримувати

бізнес-ідею на основі своїх даних для покращення процесу прийняття рішень [12].

Основні характеристики сховища даних:

предметно-орієнтованість: сховище даних надає інформацію

безпосередньо про конкретне явище, а не за загальними процесами бізнесу;

інтеграція: ЄД створюється шляхом об'єднання даних з різних джерел в єдиний формат. Дані повинні зберігатися в сховищі в узгодженому і загальноприйнятому вигляді з точки зору іменування, формату і кодування. Це

сприяє ефективному аналізу даних;

енергонезалежність: дані, внесені до сховища даних, повинні залишатися незмінними. Всі дані доступні лише для читання. Попередні дані не стираються при введенні поточних даних;

часова варіантність: дані, що зберігаються в сховищі даних,

документуються з елементом часу, явно або неявно. Прикладом часової варіації у сховищі даних є первинний ключ, який повинен мати елемент часу, такий як день, тиждень або місяць [13].

2.3. Data Mining

Інтелектуальний аналіз даних можна розглядати як природний розвиток інформаційних технологій. Його можна просто визначити як процедуру пошуку, збору, фільтрації та аналізу даних. Це метод вилучення корисних знань із величезних обсягів даних, що зберігаються в базах даних, центрах обробки даних або інших сховищах даних.

DM займається вторинним вивченням масивних баз даних з метою виявлення раніше невідомих зв'язків, які представляють інтерес або користь. З точки зору статистики, це можна розглядати як комп'ютерний дослідницький аналіз масивних, складних наборів даних [14].

Процес аналізу є циклічним, тобто створення інтелектуальної моделі даних є динамічним й ітеративним процесом. Після вивчення даних можливо виявити, що даних недостатньо для створення відповідних моделей або з'ясувати, що моделі не розв'язують визначену проблему належним [15].

Найбільш широко використовувані алгоритми поділяються на дві частини:

- традиційні методи: є статистика, пошук сусідів та кластеризація.
- методи нового покоління: дерева, мережі та правила.

Нижче наведено методи аналізу даних, які використовуються для проведення аналізу.

1. Класифікаційний аналіз використовується для класифікації різних даних у різних класах. Класифікація подібна до кластеризації, оскільки вона також сегментує записи даних на різні сегменти, які називаються класами, але на відміну від кластеризації класифікація показує як саме нові дані мають бути класифіковані.

2. Правила асоціації визначає зв'язки (моделювання залежності) між різними змінними у великих базах даних. Метод допомагає дослідити закономірності в даних, які можна використовувати для ідентифікації змінних у даних і повторення різних змінних, які дуже часто з'являються в наборі даних.

3. Виявлення аномалій або викидів досліджує набір даних і з'ясовує, які не відповідають очікуваному шаблону або очікуваній поведінці. Аномалії також відомі як викиди, новинки, шум, відхилення та винятки.

4. Кластеризаційний аналіз – це процес виявлення груп і кластерів у даних таким чином, що ступінь зв'язку між двома об'єктами є найвищою, якщо вони належать до однієї групи, і найменшою в іншому випадку.

5. Регресійний аналіз – це процес виявлення та аналізу зв'язку між змінними. Використовується для визначення змін залежної змінної, якщо будь-яка з незалежних змінних змінюється. Це означає, що одна змінна залежить від іншої, але не навпаки. Зазвичай він використовується для передбачення та прогнозування [16].

2.4. OLAP

OLAP — це технологія бази даних, оптимізована для запитів і звітів, а не для обробки транзакцій. Вихідними даними для OLAP є бази даних з онлайновою обробкою транзакцій (OLTP), які зазвичай зберігаються в сховищах даних. Дані OLAP отримуються з цих історичних даних і об'єднуються в структури, які

дозволяють проводити складний аналіз. Дані OLAP також організовані ієрархічно і зберігаються в кубах замість таблиць. Це складна технологія, яка використовує багатовимірні структури для забезпечення швидкого доступу до даних для аналізу. Така організація дозволяє легко відображати у зведених таблицях або діаграмах високорівневі зведення, такі як загальні обсяги продажів по всій країні або регіону, а також відображати додаткову інформацію для локацій, де продажі особливо високі або низькі.

Технологія OLAP призначена для прискорення пошуку даних. Оскільки сервер OLAP, а не Microsoft Office Excel, обчислює підсумовані значення, під час створення або зміни звіту в Excel потрібно надсилати менше даних. Такий підхід дозволяє працювати з набагато більшими обсягами вихідних даних, ніж це було б можливо, якби дані були організовані в традиційній базі даних, де Excel отримує всі окремі записи, а потім обчислює узагальнені значення.

Технологія OLAP містить два основних типи даних: міри, які є числовими даними, величини та середні значення, які використовуються для прийняття обґрунтованих бізнес-рішень, і виміри, які є категоріями, які використовуються для організації цих показників. Технологія OLAP допомагає організувати дані за багатьма рівнями деталізації, використовуючи для аналізу даних необхідні критерії [17].

В основному існує три типи OLAP:

- MOLAP (багатовимірний OLAP) – це OLAP для багатовимірних баз даних.
- ROLAP (Relational OLAP) – це OLAP, який здійснює динамічний багатовимірний аналіз на збережених даних реляційних баз даних.
- HOLAP (Hybrid OLAP) – це різноманітна інтеграція ROLAP та MOLAP. Він використовується для розробки ємності даних ROLAP, а MOLAP – найкраща можливість обробки для виконання вимог обробки [18].

2.5. Використані інструменти

У процесі розробки СППР були застосовані такі інструменти:

- СУБД SQL Server Management Studio для створення БД та СД;
- середовище розробки Visual Studio;
- служба SQL Server Analysis Services (SSAS) для розгортання кубу та проведення інтелектуального аналізу даних;
- служба SQL Server Integration Services (SSIS), яка містить інструмент Data Flow для реалізації процесу передачі даних з БД до СД;
- служба SQL Server Reporting Services (SSRS) для формування звітної інформації;
- таблицний редактор MS Excel, який надає підключення до СД для перегляду даних та можливість створювати звітну інформацію, сформовану у вигляді таблиць, графіків, діаграм тощо.

3 РОЗРОБКА СИСТЕМИ

3.1. Архітектура системи

Діаграма розгортання - це тип діаграми UML, яка показує архітектуру виконання системи, включаючи вузли, такі як апаратне або програмне середовище виконання, і проміжне програмне забезпечення, що з'єднує їх.

Діаграми розгортання зазвичай використовуються для візуалізації фізичного обладнання та програмного забезпечення системи. З її допомогою можна зрозуміти, як система буде фізично розгорнута на апаратному забезпеченні. Діаграма такого виду допомагає моделювати апаратну топологію системи в порівнянні з іншими типами діаграм UML, які в основному описують логічні компоненти системи.

Нотації діаграми розгортання

- Вузол, який представлений у вигляді куба, є фізичною сутністю, яка виконує один або декілька компонентів, підсистем або виконуваних файлів. Вузол може бути апаратним або програмним елементом (рис. 9 (a)).
- Артефакти – це конкретні елементи, які викликані процесом розробки. Прикладами артефактів є бібліотеки, архіви, конфігураційні файли, виконуваний файли тощо (рис. 9 (b)).
- Комунікаційна асоціація, яка зображується суцільною лінією між двома вузлами, показує шлях зв'язку між вузлами (рис. 9 (c)).
- Пристрій – це вузол, який використовується для представлення фізичного обчислювального ресурсу в системі. Прикладом пристрою є сервер додатків (рис. 9 (d)).
- Специфікація розгортання – це конфігураційний файл, наприклад, текстовий файл або XML-документ. Він описує, як артефакт розгортається на вузлі (рис. 9 (e)).

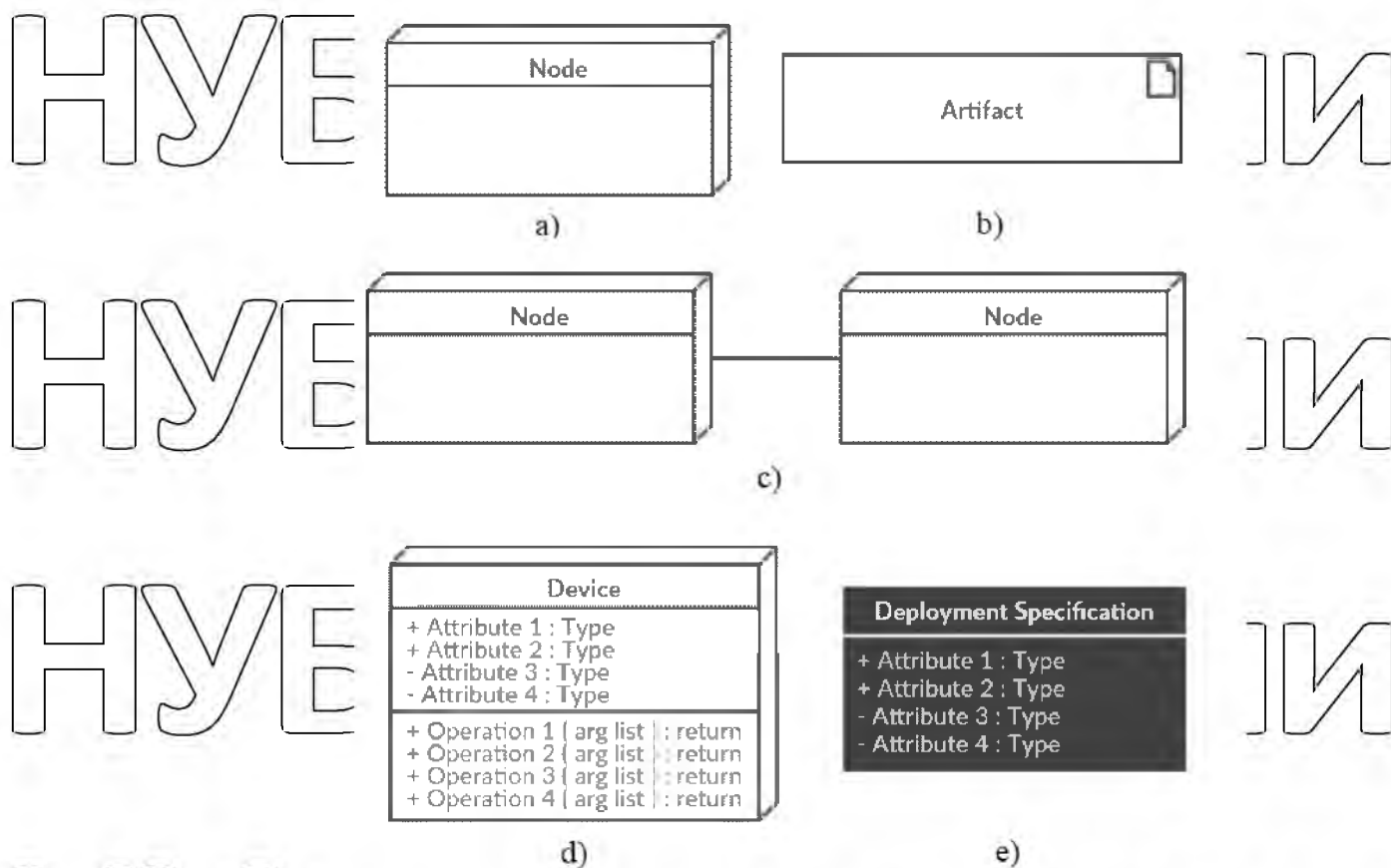


Рис. 9 Основні елементи діаграми розгортання [18]

На рис. 10 зображено архітектуру СИПР для управління тепличним господарством. Основними фізичними вузлами є:

- робоча станція теплиці:
 - модуль внесення даних в оперативну БД;
 - оперативна БД теплиці;
 - блок керування;
 - датчики температури та вологості;
- сервер сховища даних;
- робоча станція аналітика.

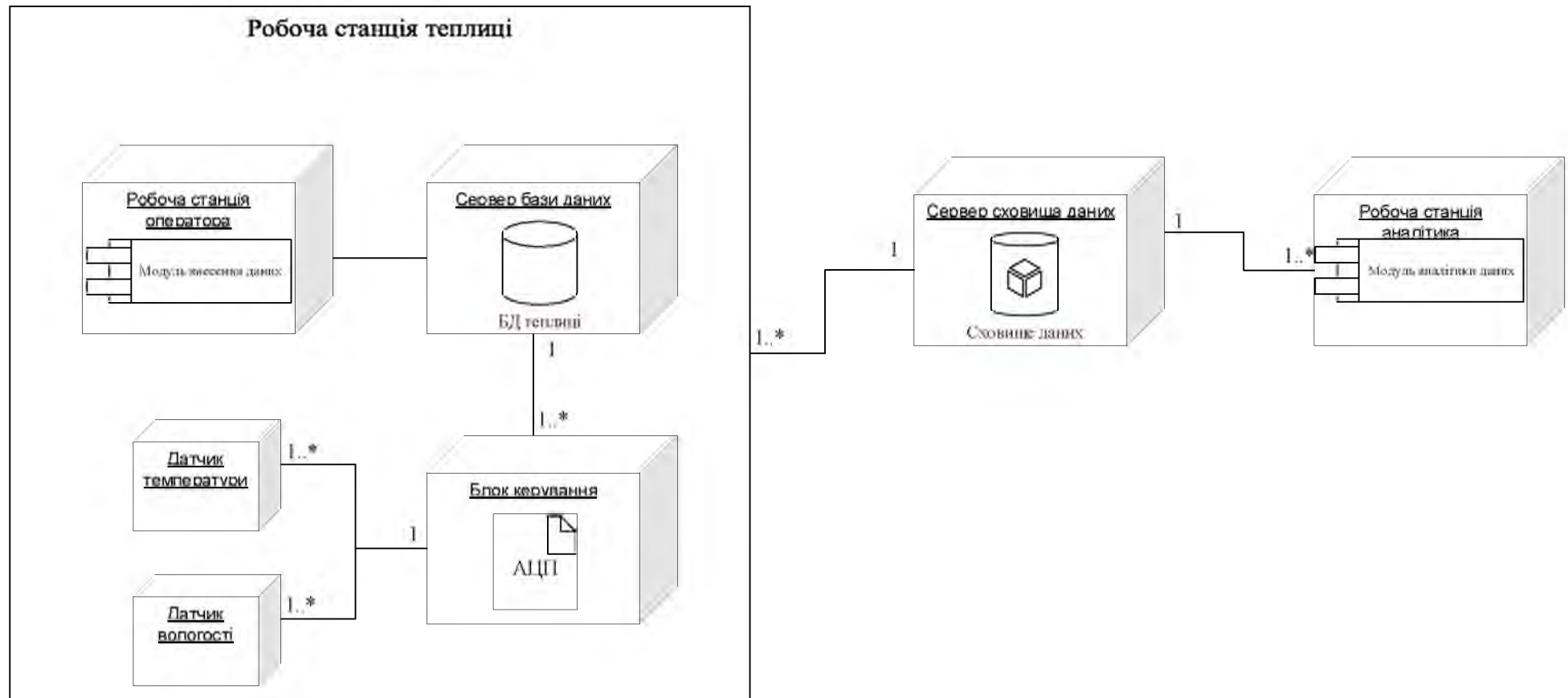


Рис. 10 Архітектура системи

3.2. Оперативні джерела інформації

Для збереження оперативних даних система моніторингу використовує БД. Таких оперативних джерел в архітектурі розроблюваної системи може бути нескінченна кількість. Надалі наявна інформація в оперативних БД можуть використовуватись для проведення аналізу. Тому на основі цієї розроблюваної підсистеми моніторингу автоматизованої системи управління мікрокліматом в теплиці система підтримки прийняття рішень формує факти, які використовуються для проведення аналізу.

Проаналізувавши вхідні та вихідні дані, які використовуються у роботі системи було розроблено оперативну базу даних, яка містить 9 таблиць:

- «Sensor» – дані про датчик: містить код датчика (`id_sensor`) як ключовий атрибут і два неключові, серед яких назва датчика (`name`) й одиниця вимірювання (`un_measurement`);
- «Greenhouse» – інформація про теплицю: містить код теплиці (`id_greenhouse`) як ключовий атрибут, а також три неключові: назва (`name`), адреса (`address`) та площа (`square`);
- «Temperature» – інформація про зафіксовану температуру в теплиці: містить зафіксований час (`datetime`), код датчику (`id_sensor`) та код теплиці (`id_greenhouse`) як ключові атрибути, а також показник (`value`) як неключовий;
- «Humidity» – інформація про зафіксовану вологість у теплиці: містить зафіксований час (`datetime`), код датчику (`id_sensor`) та код теплиці (`id_greenhouse`) як ключові атрибути, а також значення (`value`) як неключовий;
- «Harvest» – ключова таблиця, яка містить інформацію про результуючу врожайність у теплиці: наявні такі ключові атрибути як код врожаю (`id_harvest`), код сорту (`id_sort`), код культури (`id_culture`) та код теплиці (`id_greenhouse`) і неключові: врожайність (`harvest`), день посадки (`start_datetime`) та день зібрання врожаю (`finish_datetime`);

- «Culture» – дані про культуру: містить ключовий атрибут – код культури (id_culture) і неключовий – назва культури (name);
- «Sort» – дані про сорт конкретної культури: містить ключовий атрибут: код сорту (id_sort) та неключові: назва сорту (name) і код культури (id_culture), до якої належить вказаний сорт;
- «Opt_temperature» – діапазон оптимальних показників температури для сорту: містить ключовий атрибут: код сорту (id_sort), код культури (id_culture) і неключові: нижній (min_temperature) і верхній (max_temperature) поріг оптимального значення температури;
- «Opt_humidity» – діапазон оптимальних показників вологості для сорту: містить ключовий атрибут: код сорту (id_sort), код культури (id_culture) і неключові: нижній (min_humidity) і верхній (max_humidity) поріг оптимального значення вологості.

БД була розроблена з використанням СУБД MS SQL Server за допомогою запитів. Приклад створення таблиці наведений на рис. 11

```
CREATE TABLE Culture
(
    id_culture          integer IDENTITY ( 1, 1 ) ,
    name_culture       nvarchar(50) NOT NULL
)
go

ALTER TABLE Culture
    ADD CONSTRAINT XPKCulture PRIMARY KEY CLUSTERED (id_culture ASC)
go
```

Рис. 11 Створення таблиці «Culture»

Запити створення всіх таблиць БД наведено у додатку А. На рис. 12 представлена схема оперативної бази даних, яка створена у результаті виконання всіх SQL-запитів

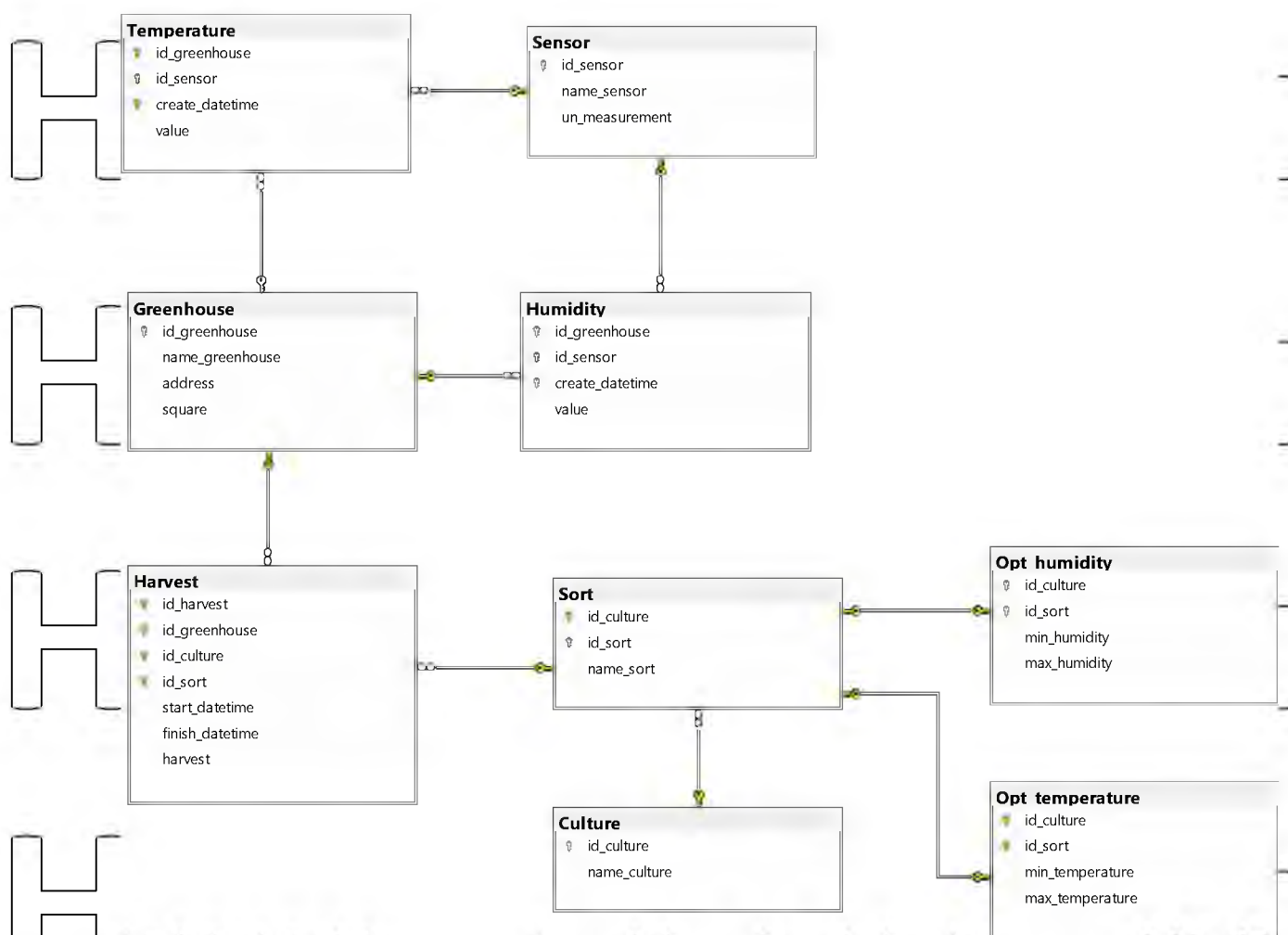


Рис. 12 Логічна модель БД тепличного господарства

3.3. Сховище даних

Для проведення аналізу даних у розроблюваній СППР використовується сховище даних, яке дозволить проводити аналіз у різних розрізах. Це забезпечується наявністю декількох вимірів, які являють собою сукупність довідкової інформації про вимірювану подію. У контексті сховищ даних подіями являються факти, які описують результати певного бізнес-процесу. Структура сховища даних зображена на рис. 13.

Для збереження необхідних даних у СД були розроблені такі таблиці:

- SortDim – вимір, що містить інформацію про сорт з такими полями:
 - id_culture – ідентифікатор культури;
 - id_sort – ідентифікатор сорту;
 - name_culture – назва культури;
 - name_sort – назва сорту;

- OptConditionDim – вимір, що зберігає оптимальні показники для сорту
 - id_condition – ідентифікатор умови;
 - id_culture – ідентифікатор культури;
 - id_sort – ідентифікатор сорту;
 - min_temperature – мінімальна оптимальна температура для сорту;
 - max_temperature – максимальна оптимальна температура для сорту;
 - min_humidity – мінімальна оптимальна вологість для сорту;
 - max_humidity – максимальна оптимальна вологість для сорту;

- GreenhouseDim – вимір, що містить інформацію про теплицю:

- id_greenhouse – ідентифікатор теплиці;
- name_greenhouse – назва теплиці;
- address – адреса, де знаходиться теплиця;
- square – площа теплиці;

- DateDim – вимір, що зберігає часовий проміжок:

- id_date – ідентифікатор дати;
- year – рік;
- month – місяць;
- day – день.

- HarvestFact – таблиця фактів про зафіксовану врожайність з такими полями:

- id_date – ідентифікатор дати;
- id_greenhouse – ідентифікатор теплиці;
- id_culture – ідентифікатор культури;
- id_sort – ідентифікатор сорту;
- avg_temperature – середня зафіксована температура за періоду вирощування врожаю;
- avg_humidity – середня зафіксована вологість за періоду вирощування врожаю;
- min_temperature – мінімальна зафіксована температура за періоду вирощування врожаю;

- max_temperature – максимальна зафіксована температура за періоду вирощування врожаю;
- min_humidity – мінімальна зафіксована вологість за періоду вирощування врожаю;
- max_humidity – максимальна зафіксована вологість за періоду вирощування врожаю;
- harvest – кількість врожаю за період вирощування.

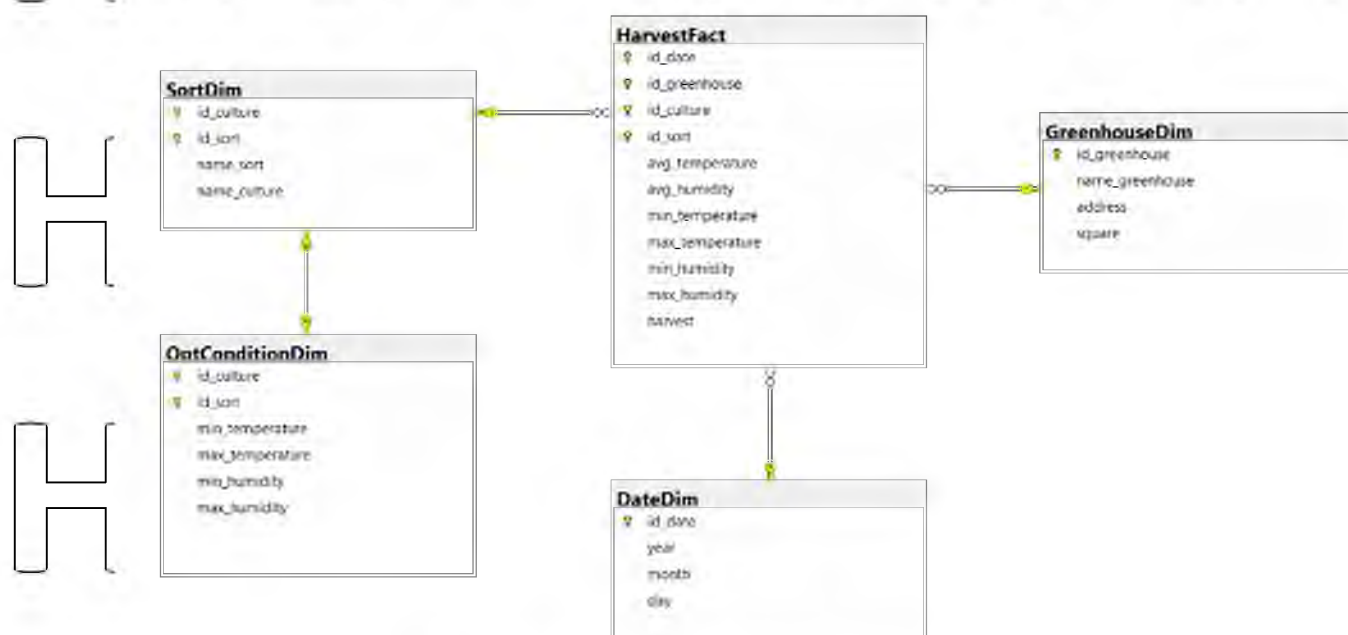


Рис. 13 Вітрина тепличного господарства

Структура СД була створена за допомогою сформованих SQL-запитів. Скрипти запитів створення всіх таблиць СД наведено у додатку Б. Так як інформація у таблицях-вимірах є відносно постійною, тому дані для цих таблиць заповнювались за допомогою SQL запитів, один з яких наведений на рис. 14.

```
BEGIN TRANSACTION
INSERT INTO GreenhouseDim ([name_greenhouse], [address], [square]) VALUES
('Цех #1', 'Київська обл., Фастівський район, м. Боярка', 5.0),
('Цех #2', 'Київська обл., Фастівський район, м. Боярка', 3.5),
('Цех #3', 'Київська обл., Фастівський район, м. Боярка', 4.0);
GO
COMMIT
GO
```

Рис. 14 Внесення умовно-постійної інформації про теплиці

3.4. Побудова розгорнутого кубу підсистеми аналізу

SQL Server Analysis Services (SSAS) – це багатовимірний OLAP-сервер, а також механізм для аналітики, який дозволяє розбивати великі обсяги даних. Ця служба входить до складу Microsoft SQL Server і допомагає виконувати аналіз з використанням різних вимірів. Має 2 варіанти: багатовимірний і табличний [19].

Для розробки кубу було використано програмне середовище Visual Studio з розширенням SSAS. На першому етапі необхідно визначити джерело даних – база даних OLAP або сховище даних. На основі визначеного джерела даних буде проводитись імпорт необхідних даних. На рис. 15 зображено формування підключення джерела даних за допомогою модулю Data Source Wizard, де ми обираємо попередньо створене сховище даних.

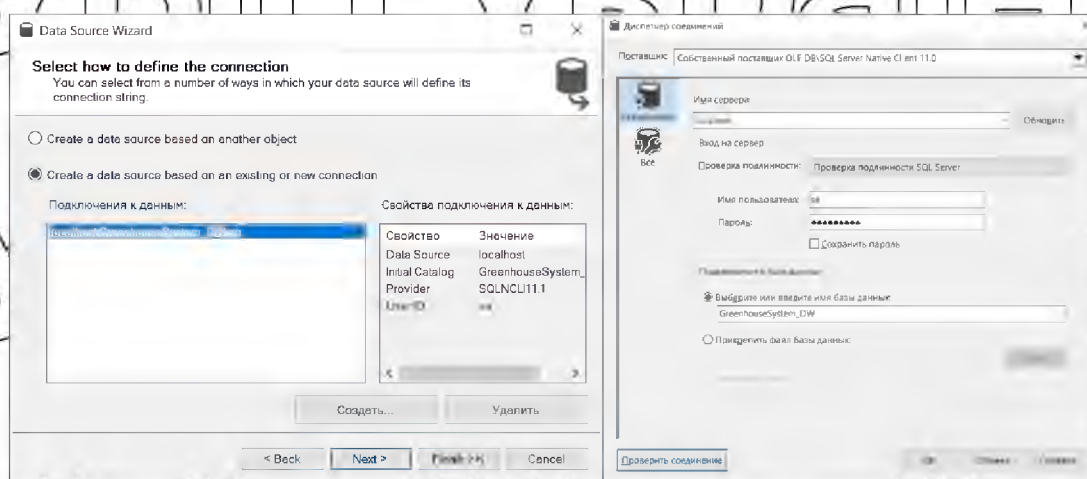


Рис. 15 Формування джерела даних

Наступним етапом у розробці куба є створення уявлення джерела даних, яке являє собою абстракцію реляційного джерела даних, яка стає основою кубів і вимірів, які створюються в багатовимірному проекті. Метою Data Source Wizard є надання контролю над структурами даних, які використовуються у проекті, і працювати незалежно від базових джерел даних (наприклад, можливість перейменувати або об'єднати стовпці без безпосередньої зміни вихідного джерела даних). У процесі розробки можна створити кілька представлень джерел даних у проекті або базі даних Analysis Services на одному або кількох джерелах даних і створити кожне з них, щоб задовольнити вимоги до іншого рішення [20].

На основі створеного джерела та уявлення джерелах даних необхідно створити виміри для OLAP куба. Дивлячись на структуру створеного сховища у пункті 3.3 куб матиме чотири виміри.

На основі створених вимірів наступним кроком є створення кубу OLAP за допомогою майстра (Cube Wizard). На рис. 16 зображено створення кубу. На першому етапі необхідно обрати таблицю фактів, яка виступає основою для створення кубу. Далі іде вибір полів та вимірів, які формуватимуть куб.

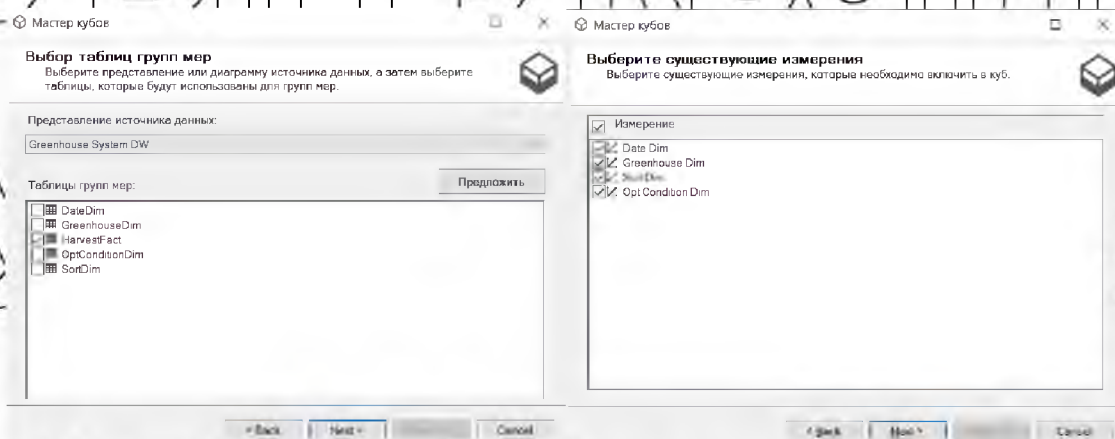


Рис. 16 Вибір таблиць вимірів та фактів для створення кубу

Після виконання всіх попередніх етапів результатом є створений Greenhouse System DW, що зображений на рис. 17.

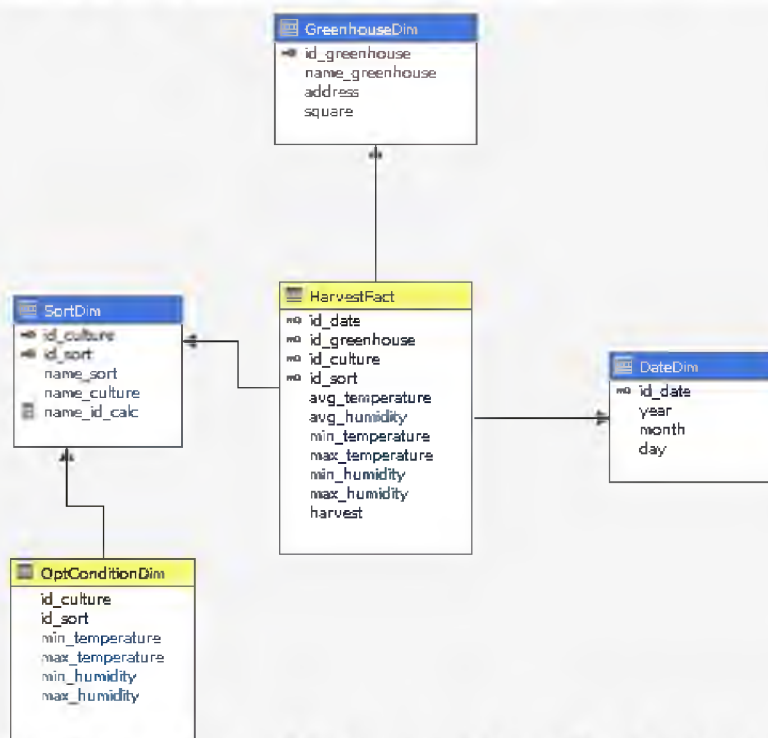


Рис. 17 Створений куб тепличного господарства

3.5. Реалізація отримання даних за допомогою Data Flow

Процес передачі даних було реалізовано за допомогою служби SQL Server Integration Services. SSIS – це інструмент, який дозволяє реалізувати інтеграцію у зручному вигляді, а саме реалізувати процес перенесення даних з одного джерела до іншого. Цей процес іноді називають ETL (від англ. Extract, Transform, Load – дослівно «вистучення, перетворення, завантаження») [21].

В інструменті SSIS є служба Data Flow, за допомогою якої було проведено заповнення таблиць вимірів та фактів.

Наповнення сховища даних (СД) відбувається на основі оперативної БД, описаної у пункті 3.2 і поділено на 2 етапи (рис. 18).

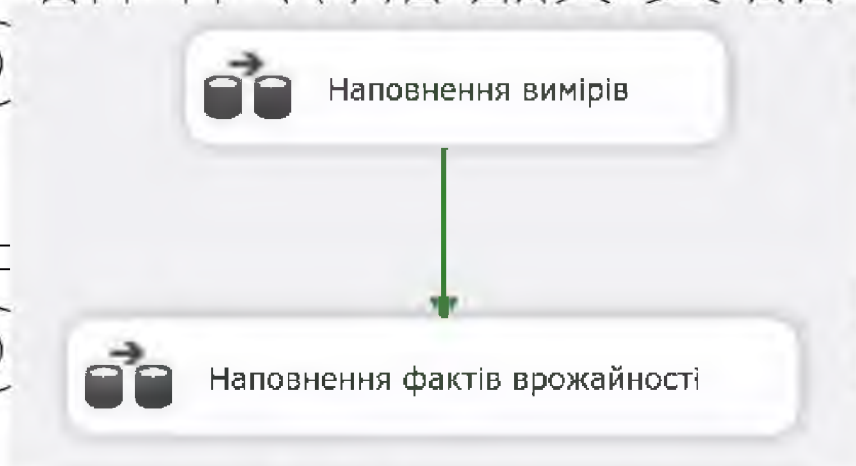


Рис. 18 Потіки даних для наповнення СД

На першому етапі ми заповнюємо таблиці вимірів, дані для яких беруться, як було зазначено вище, з оперативної БД. На рис. 19 зображено потіки, які реалізують передачу даних з БД у СД.



Рис. 19 Перший крок наповнення (таблиць-вимірів)

До прикладу далі вказано процес заповнення часового виміру DateDim.

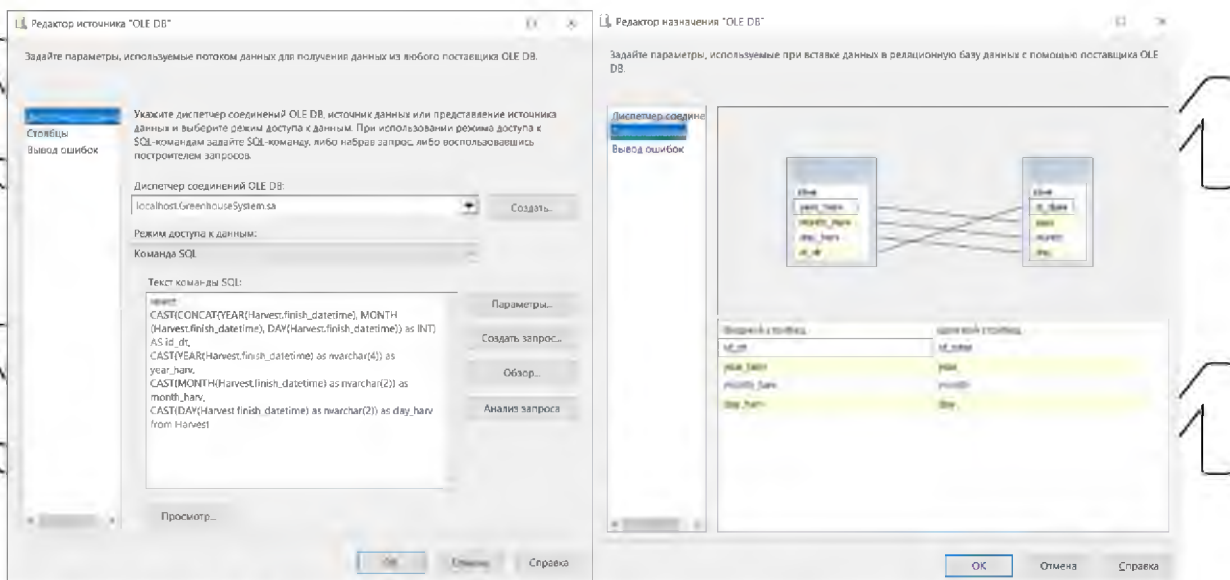


Рис. 20 Формування запиту на основі джерела оперативної БД та процес передачі даних на основі сформованої вибірки

На другому етапі проводиться заповнення таблиці фактів врожайностей.

Запит, наведений нижче, проводить вибірку врожайності з оперативної БД та дати, занесеної у СД на першому етапі

```

SELECT GreenhouseSystem.dbo.Harvest.id_greenhouse,
GreenhouseSystem.dbo.Harvest.id_culture,
GreenhouseSystem.dbo.Harvest.id_sort,
GreenhouseSystem.dbo.Harvest.harvest,
GreenhouseSystem_DW.dbo.DateDim.id_date,
MIN(Temperature.[value]) AS min_temp,
MAX(Temperature.[value]) AS max_temp,
AVG(Temperature.[value]) AS avg_temp,
MIN(Humidity.[value]) AS min_humid,
MAX(Humidity.[value]) AS max_humid,
AVG(Humidity.[value]) AS avg_humid
FROM GreenhouseSystem.dbo.Harvest
JOIN GreenhouseSystem.dbo.Greenhouse ON Harvest.id_greenhouse =
Greenhouse.id_greenhouse
JOIN GreenhouseSystem.dbo.Temperature ON Greenhouse.id_greenhouse =
Temperature.id_greenhouse
JOIN GreenhouseSystem.dbo.Humidity ON Greenhouse.id_greenhouse =
Humidity.id_greenhouse
JOIN GreenhouseSystem_DW.dbo.DateDim
ON DateDim.[year] = CAST(YEAR(GreenhouseSystem.dbo.Harvest.finish_datetime) AS
nvarchar(4))
AND DateDim.[month] =
CAST(MONTH(GreenhouseSystem.dbo.Harvest.finish_datetime) AS NVARCHAR(2))
AND DateDim.[day] = CAST(DAY(GreenhouseSystem.dbo.Harvest.finish_datetime)
AS NVARCHAR(2))
WHERE Temperature.create_datetime >= Harvest.start_datetime
AND Temperature.create_datetime <= Harvest.finish_datetime
AND Humidity.create_datetime >= Harvest.start_datetime
AND Humidity.create_datetime <= Harvest.finish_datetime
GROUP BY Harvest.id_greenhouse, Harvest.id_culture, Harvest.id_sort,
Harvest.finish_datetime, Harvest.harvest,
DateDim.id_date

```

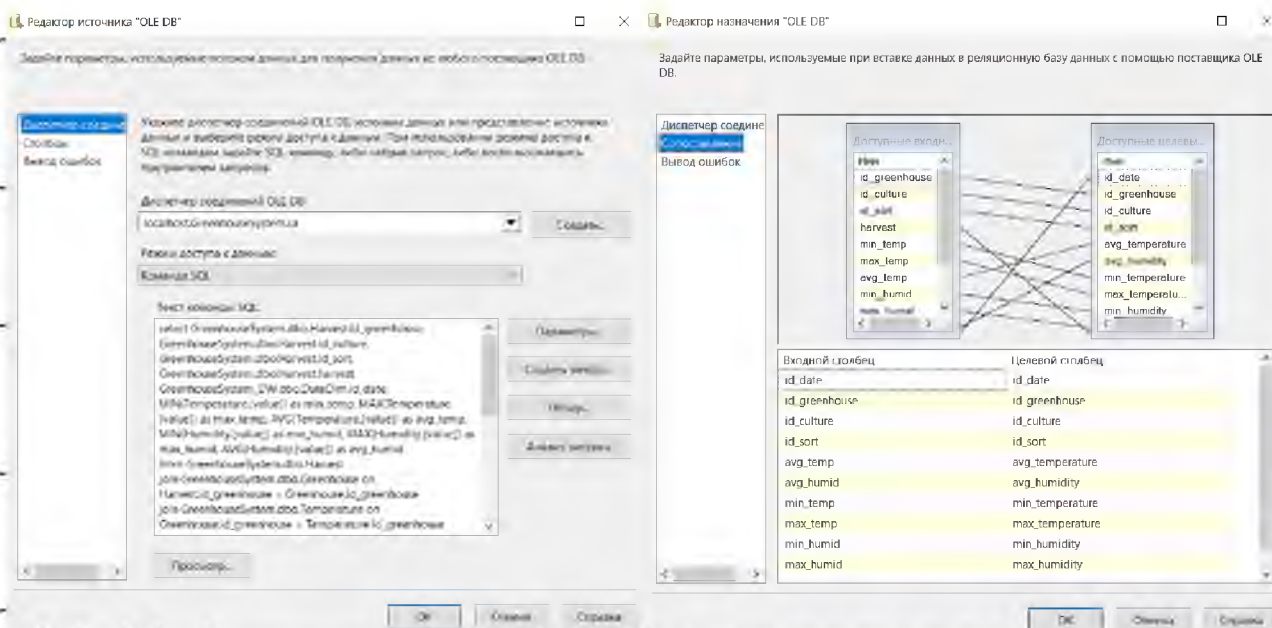



Рис. 21-22 Формування вибірки для фактів врожайності та процес передачі даних на основі сформованої вибірки

У результаті виконання всіх потоків у таблицях СД наявні дані, які перенесені з оперативної БД.

3.6 Побудова звітності

Під час проведення аналізу даних одним з найважливіших етапів є формування звітності. Саме за допомогою звітів аналітик та підприємство можуть побачити цілісну картину розвитку організації.

Для створення звітної інформації була використана служба SQL Server Reporting Services (SSRS). Служби SQL Server Reporting Services (SSRS) надають набір локальних засобів та служб для створення та розгортання звітів різного вигляду, а також управління ними [22].

Процес створення звітів починається з підключення джерела даних. Далі за допомогою майстра створюємо звіт, використовуючи запити, які формуватимуть вибірку (рис. 23). Наступним етапом є формування структури звіту: сторінки, стовпці, рядки та деталі (рис. 24). Після створення у режимі конструктора можна редагувати зовнішній вигляд звіту та додати необхідні діаграми, які будуть наочно демонструвати результати вибірки.

Запити, які використовувались для формування звітної інформації, наведені у додатку Г.

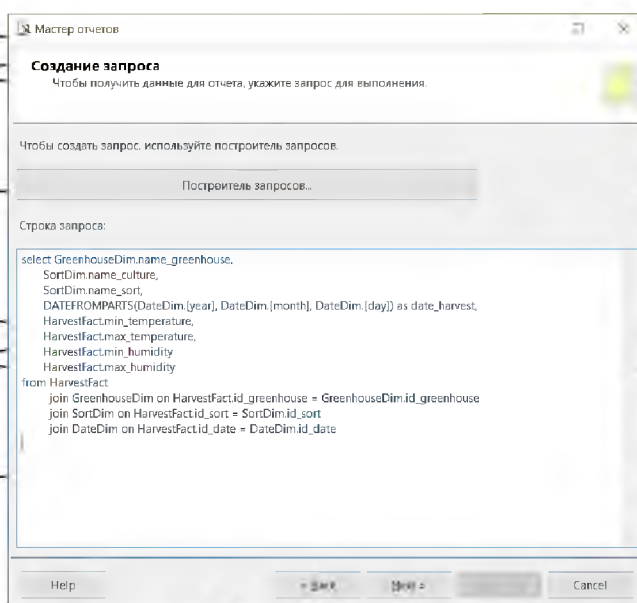


Рис. 23 Формування запиту для вибірки

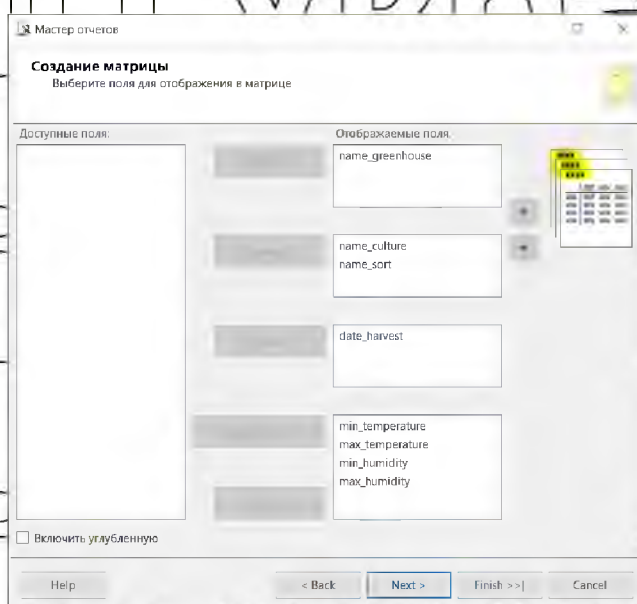


Рис. 24 Налаштування структури подання звітної інформації

Щоб зрозуміти алгоритм формування звітної інформації на рис. 25 зображено блок-схему, яка допомагає краще зрозуміти послідовність дій. Всі дії відбуваються використовуючи створені шаблони, які користувач створює натискаючи на обраний звіт. Обраний звіт розгортається у в.в.н., де користувач має можливість переглядати та зберігати у бажаних форматах *.pdf, *.xlsx, *.docx.

НУБІП України

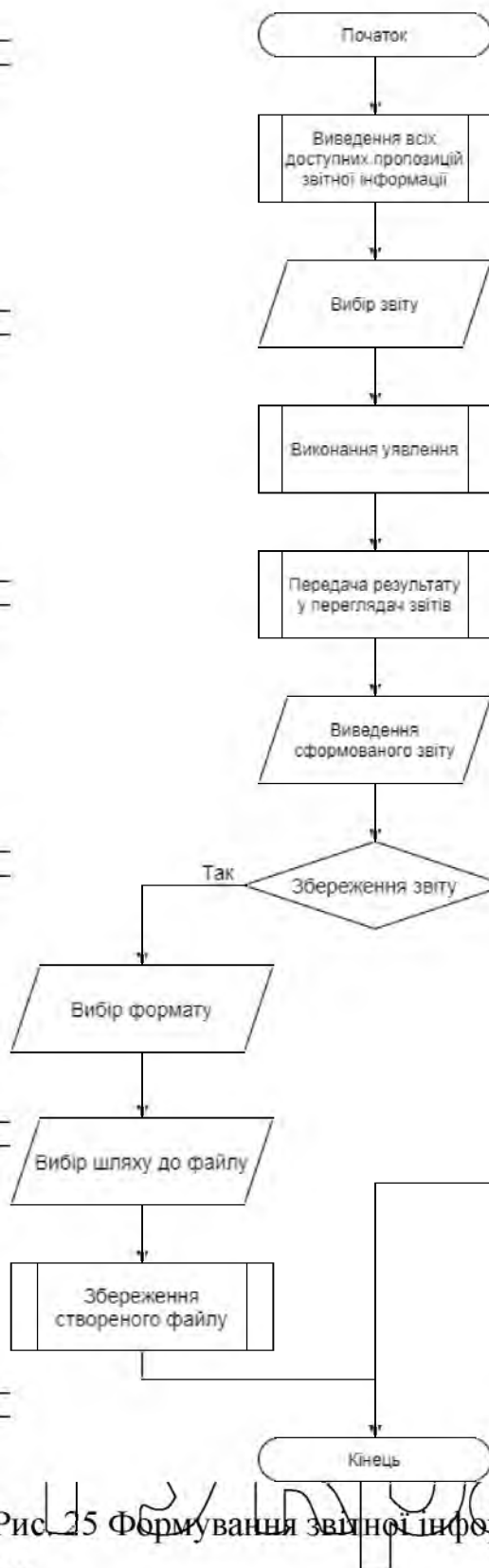


Рис. 25 Формування звітної інформації

3.7. Розрахунок КРІ

Для визначення успішності роботи підприємства аналітики обчислюють коефіцієнт ефективності. КРІ — це набір обчислень, які пов'язані з групою показників у кубі, які використовуються для оцінки успіху бізнесу. Зазвичай ці обчислення є комбінацією виразів багатовимірних виразів (MDX) або

обчислюваних елементів. KPI також мають додаткові метадані, які надають інформацію про те, як клієнтські програми повинні відображати результати обчислень KPI.

KPI обробляє інформацію про поставлену ціль, фактичну формулу продуктивності, записану в кубі, і вимірювання, щоб показати тенденцію та статус ефективності.

Компоненти для розрахунку KPI:

- Ціль – числовий вираз MDX або обчислення, яке повертає цільове значення KPI.

- Значення – числовий вираз MDX, який повертає фактичне значення KPI.

- Статус – вираз MDX, який представляє стан KPI на певний момент часу. Вираз MDX стану має повертати нормалізоване значення від -1 до 1. Значення, рівні або менші за -1, будуть інтерпретуватися як "погані" або "низькі". Значення нуль (0) інтерпретуються як «прийнятне» або «середнє». Значення, що дорівнюють або перевищують 1, будуть інтерпретуватися як «добре» або «високе» [23].

Для аналізу фактів, занесених у куб, було визначено такі KPI:

1. KPI_Sort1_Opt_Avg_Temp – визначення ефективності середньої температури за сортом «Волове серце» культури помідор.
2. KPI_Sort1_Opt_Avg_Humidity – визначення ефективності середньої вологості за сортом «Волове серце» культури помідор.
3. KPI_Sort4_Avg_Temp – визначення ефективності середньої температури за сортом «Фенікс» культури огірок.
4. KPI_Sort4_Avg_Humidity – визначення ефективності середньої температури за сортом «Фенікс» культури огірок.

Кожен з показників показує ефективність температури та вологості у розрізі сорту.

Для розрахунку КРІ необхідно розраховане фактичне значення, цільове значення та зазначені умови для визначення статусу (рис. 26-28).

* Выражение значения

```
(([Measures].[Avg Temperature], [Sort Dim].[Id Sort].&[1],[Date Dim].[Year].&[2021]) /
([Measures].[Count Harvest Fact], [Sort Dim].[Id Sort].&[1],[Date Dim].[Year].&[2021]))
```

Проблемы не найдены. Стр: 1 Симв: 173 Столб: 87 Пробелы CRLF

Рис. 26 Розрахунок фактичного значення

* Целевое выражение

```
(([Measures].[Min Temperature - Opt Condition Dim], [Sort Dim].[Id Sort].&[1]) +
([Measures].[Max Temperature - Opt Condition Dim], [Sort Dim].[Id Sort].&[1])) / 2
```

Проблемы не найдены. Стр: 1 Симв: 166 Столб: 84 Пробелы CRLF

Рис. 27 Розрахунок цільового значення

* Состояние

Признак состояния:

Шкала

Выражение состояния:

```
CASE
WHEN KPIVALUE( "KPI_Sort1_Opt_Avg_Temp" ) = KPIGOAL( "KPI_Sort1_Opt_Avg_Temp" ) THEN 1
WHEN KPIVALUE( "KPI_Sort1_Opt_Avg_Temp" ) >= ([Measures].[Min Temperature - Opt Condition
Dim], [Sort Dim].[Id Sort].&[1]) AND
KPIVALUE( "KPI_Sort1_Opt_Avg_Temp" ) <= ([Measures].[Max Temperature - Opt Condition
```

Проблемы не найдены. Стр: 6 Симв: 4 Пробелы CRLF

Рис. 28 Визначення статусу за зазначеними умовами

Також для порівняння даних за попередній часовий проміжок можна розраховувати тренд, який показує збільшення чи зменшення показника відносно попередніх даних (рис. 29).

* Тренд

Признак тренда:

Стандартная стрелка

Выражение тренда:

```
CASE
WHEN ([Measures].[Avg Temperature], [Sort Dim].[Id Sort].&[1],[Date Dim].[Year].&
[2021]) / ([Measures].[Count Harvest Fact], [Sort Dim].[Id Sort].&[1],[Date Dim].
[Year].&[2021]) >
([Measures].[Avg Temperature], [Sort Dim].[Id Sort].&[1],[Date Dim].[Year].&[2020]) /
([Measures].[Count Harvest Fact], [Sort Dim].[Id Sort].&[1],[Date Dim].[Year].&
[2020])
```

Рис. 29 Розрахунок тренду

3.8. Розробка алгоритмів інтелектуального аналізу даних

Для проведення інтелектуального аналізу було використано інструмент SQL Server Data Tools-Business Intelligence (SSDT-BI), який містить технології для бізнес аналізу: створення моделей даних Analysis Services (AS), пакетів Integration Services (IS) та звітів Reporting Services (RS).

Для вирішення задач Data Mining Analysis Services пропонує наступні методи для проведення аналізу:

- алгоритм часових рядів;
- алгоритм дерева прийняття рішень;
- алгоритм кластеризації;
- алгоритм лінійної регресії;
- алгоритм логістичної регресії;
- алгоритм нейронної мережі;
- правила взаємозв'язків;
- спрощений алгоритм Байєса.

Для демонстрації застосування завдань Data Mining нижче наведена реалізація декількох алгоритмів, для яких використовується попередньо розгорнутий та наповнений фактами куб.

3.8.1. Алгоритм часових рядів. Для створення структури інтелектуального аналізу використовується інструмент Data Mining Wizard. Початок реалізації задачі починається з вибору джерела, яке буде використовуватись для проведення аналізу. Майстер інтелектуального аналізу надає два варіанти: на основі існуючої реляційної БД чи СД з використанням створеного кубу. Для створення алгоритму використовуємо розгорнутий куб.

Наступним кроком є вибір методу для аналізу, де обираємо алгоритм часових рядів (29).

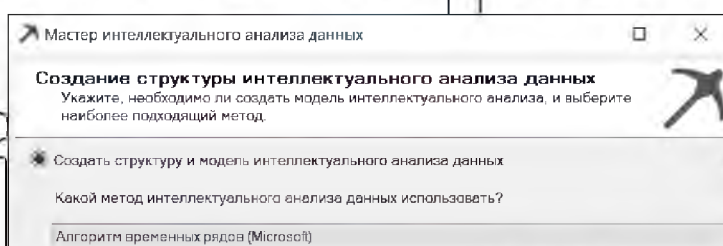


Рис. 30 Вибір алгоритму часових рядів як методу інтелектуального аналізу на основі сформованого кубу

Наступним кроком є вибір виміру, на основі якого буде проводитись інтелектуальний аналіз (рис. 31).

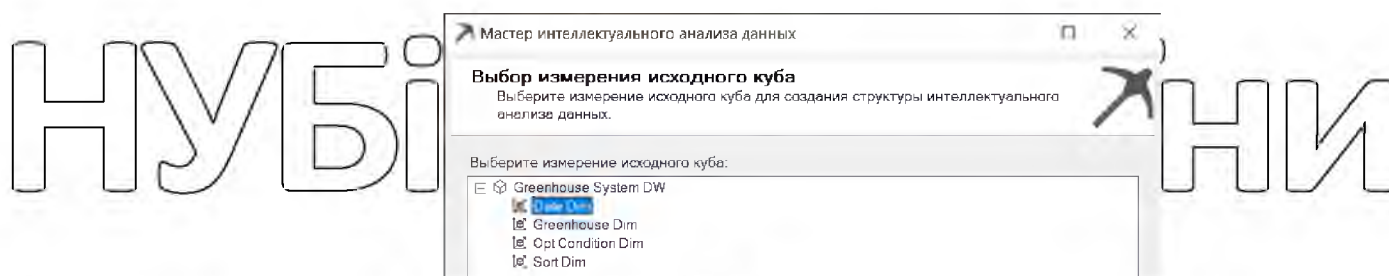


Рис. 31 Вибір виміру кубу

Далі необхідно обрати ключ структури аналізу, вхідні параметри та прогнозовані (рис. 31), де вхідний параметр впливає на процесу, який вивчається, а прогнозованим є параметр, значення якого будуть передбачуватися.

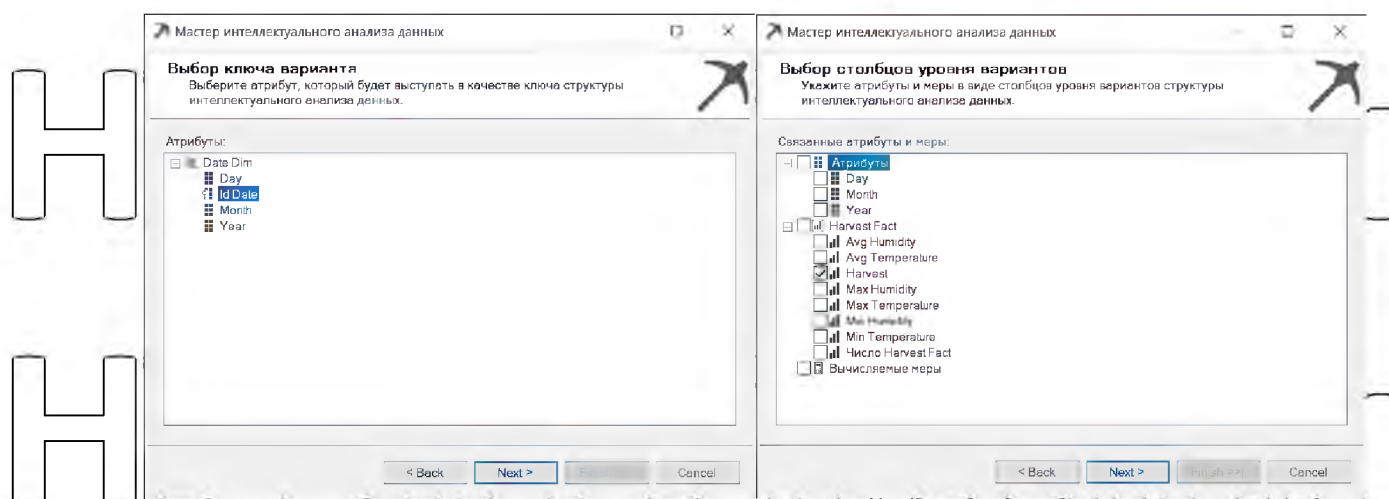


Рис. 32 Налаштування структури аналізу часового ряду

У результаті виконаних налаштувань створюється модель аналізу даних, яка наведена на рис. 33

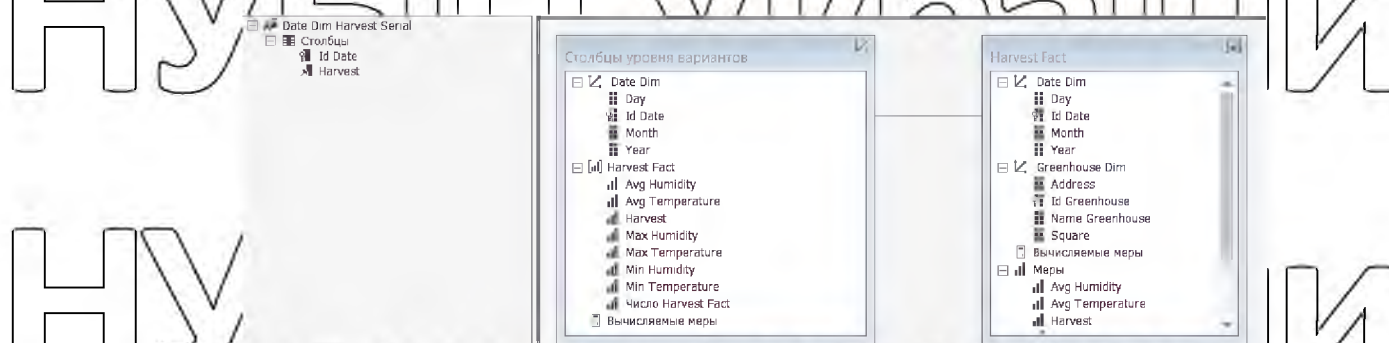


Рис. 33 Структура створеної моделі

3.8.2. Алгоритм асоціативних правил. Правила асоціації – це твердження

"якщо-тоді", які допомагають показати ймовірність взаємозв'язків між елементами даних у великих наборах даних у різних типах баз даних. Видобуток асоціативних правил має ряд застосувань і широко використовується для

виявлення кореляцій у наборах медичних даних.

У науці про дані правила асоціації використовуються для пояснення закономірностей в даних з, здавалося б, незалежних інформаційних сквищ, таких як реляційні бази даних і транзакційні бази даних. Процес використання правил асоціацій іноді називають "видобуванням правил асоціацій" або "видобуванням асоціацій" [24].

На основі розгорнутого кубу була створена нова структура інтелектуального аналізу даних, використовуючи метод асоціативних правил (рис. 34) з налаштуванням передбачення назви культури та сорту.



Рис. 34 Модель структури інтелектуального аналізу методом асоціативних правил

3.8.3. Алгоритм кластеризації. Кластеризація – це некерований алгоритм, який об'єднує групу точок даних у кластери таким чином, щоб об'єкти належали до однієї групи.

Кластеризація допомагає розбити дані на декілька підмножин. Кожна з цих підмножин містить дані, схожі між собою, і ці підмножини називаються кластерами [25].

На основі розгорнутого кубу була створена нова структура інтелектуального аналізу даних про сорти та культури, використовуючи алгоритм кластерного аналізу (рис. 35).

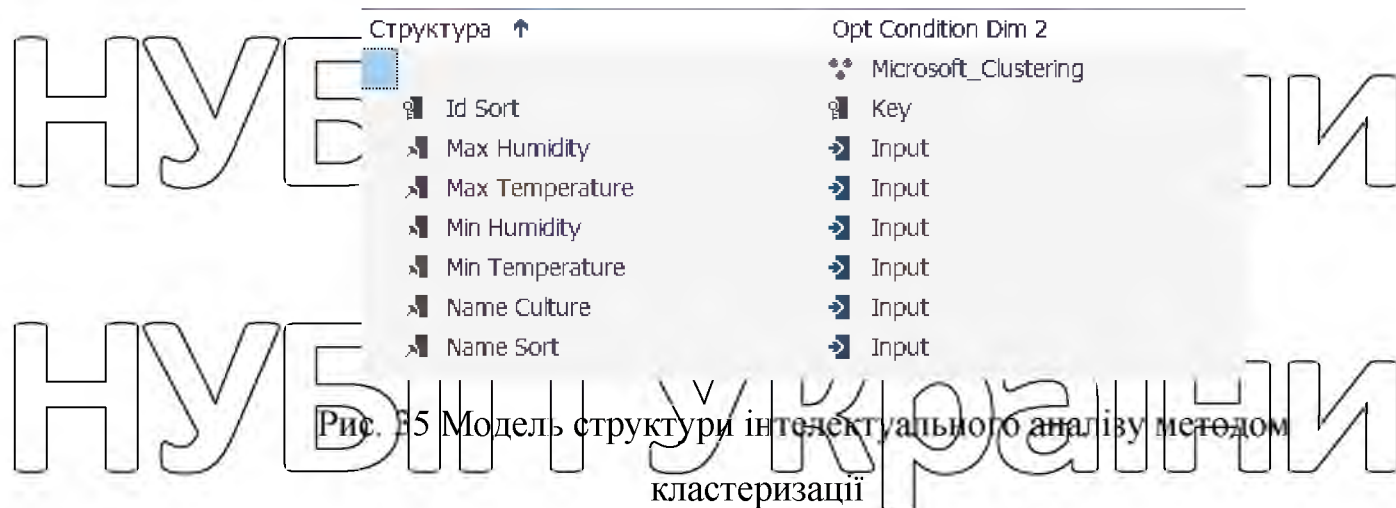


Рис. 35 Модель структури інтелектуального аналізу методом кластеризації

Разом з використанням інструментів SSIS було проведено розробку

окремих програмних модулів, які проводять інтелектуальний аналіз даних тепличного господарства.

3.8.4. Алгоритм класифікації OneR. One Rule – це простий, але точний

алгоритм класифікації, який генерує одне правило для кожного предиктора в

даних, а потім вибирає правило з найменшою сумарною помилкою в якості

"єдиного правила". Щоб створити правило для предиктора, створюється таблиця

частот для кожного предиктора проти цілі. Було показано, що OneR створює

правила лише трохи менш точні, ніж найсучасніші алгоритми класифікації, при

цьому створюючи правила, які є простими для інтерпретації людиною.

Алгоритм OneR:

Для кожного предиктора,

Для кожного значення цього предиктора скласти правило наступним

чином;

Підрахувати, як часто з'являється кожне значення цілі (класу)

Знайти найчастіший клас

Скласти правило, яке присвоює цей клас даному значенню

предиктора

Підрахувати сумарну похибку правил кожного предиктора

Вибір предиктора з найменшою сумарною похибкою. [26]

Реалізація алгоритму 1-Rule

Для реалізації алгоритму 1-Rule було виділено 2 класи високої ("HIGH") та низької ("LOW") врожайності, показник яких визначається у порівнянні з середнім показником врожайності за весь період дослідження.

За допомогою SQL визначені результати алгоритму. Нижче наведено реалізацію та результату вигляді таблиці.

Спочатку проведено вибірку даних, де визначено середню врожайність та клас врожайності за кожним сортом (рис. 36).

```

DECLARE @avgHarvest FLOAT;
SELECT @avgHarvest = AVG(HarvestFact.harvest)
FROM HarvestFact;

SELECT SortDim.name_culture,
SortDim.name_sort,
GreenhouseDim.name_greenhouse,
HarvestFact.harvest AS harvest,
@avgHarvest AS avgHarvest,
CASE
    WHEN @avgHarvest > HarvestFact.harvest
    THEN 'LOW'
    ELSE 'HIGH'
END AS harvestClassification
FROM HarvestFact
JOIN SortDim ON HarvestFact.id_sort = SortDim.id_sort
JOIN GreenhouseDim ON HarvestFact.id_greenhouse = GreenhouseDim.id_greenhouse
GROUP BY SortDim.name_culture, GreenhouseDim.name_greenhouse, HarvestFact.harvest

```

Рис. 36 Сформований запит для класифікації

Наступним етапом було визначення класифікації врожайності та її ймовірність відносно середнього показника серед заданих фактів відповідно до культури та теплиці. Сформовані запити наведені у додатку В.

4 РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

НУБІП України

4.1. Оперативний аналіз даних

4.1.1. *Результати побудови звітної інформації.* У результаті використання інструменту SSRS було сформовано декілька варіантів відображення звітної інформації.

На рис. 37 представлено звіт, який відображає зафіксовані показники мікроклімату у розрізі часу, теплиці та оптимальних показників відносно конкретного сорту.

Аналіз температури та вологості у розрізі оптимальних показників

Назва теплиці: Цех №1

Культура: Огірок
Назва сорту: "Фенікс"

	Мінімальна зафіксована температура	Максимальна зафіксована температура	Середня оптимальна температура	Мінімальна зафіксована вологість	Максимальна зафіксована вологість	Середня оптимальна вологість	Врожайність, т
16.02.2019	18,00	27,00	22,25	75,00	80,00	41,45	204,00
24.06.2019	18,00	27,00	23,05	75,00	80,00	77,80	265,00
17.10.2019	18,00	27,00	22,65	75,00	80,00	77,65	302,00
26.02.2020	18,00	27,00	20,95	75,00	80,00	77,75	212,00
29.06.2020	18,00	27,00	22,00	75,00	80,00	76,65	265,00
20.10.2020	18,00	27,00	22,20	75,00	80,00	73,25	298,00
15.02.2021	18,00	27,00	22,33	75,00	80,00	67,67	255,00
23.06.2021	18,00	27,00	24,11	75,00	80,00	65,22	290,00
15.10.2021	18,00	27,00	22,60	75,00	80,00	67,30	236,00

■ Мінімальна зафіксована температура
■ Середня оптимальна температура
■ Максимальна зафіксована температура

■ Мінімальна зафіксована вологість
■ Середня оптимальна вологість
■ Максимальна зафіксована вологість

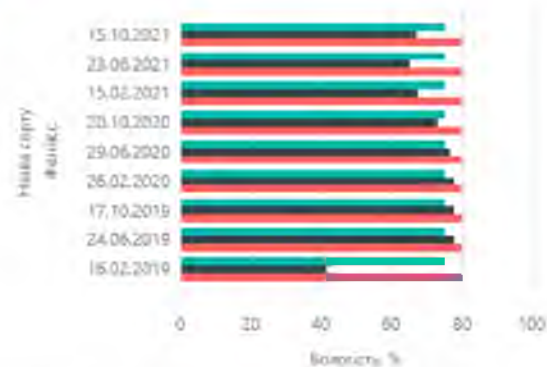
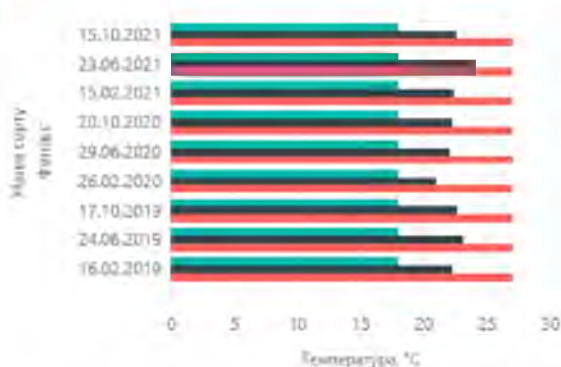


Рис. 37 Аналіз температури та вологості у розрізі оптимальних показників

НУБІП України

Висновки

Сформований звіт відображає як мікроклімат, зафіксований у процесі вирощування, впливає на результуючі показники врожайності. Судячи з наведених таблиць та графіків можна зробити висновок, що такі параметри як

температура та вологість у теплиці напряму впливають на майбутню врожайність. На прикладі врожайності, зафіксованої 16.02.2019, коли було помічено підвищену вологість та понижену температуру, можна додати що безпосередньо врожайність була найнижчою серед усіх зазначених. Такі виходи

за межі оптимальних показників негативно впливають на ріст рослин у теплиці,

що доводить проведений аналіз предметної області, описаний у пункті 1.1, де зазначено про вплив кожного параметра, який є вагомим у розвитку рослини.

Також було створено звіт, який зображено рис. 38. Він відображає зміну врожайності у розрізі теплиці та сорту, який вирощувався у ній у вказаний часовий проміжок.

Зміна врожайності у розрізі сорту

Назва теплиці: Цех №1

Культура: Огірок

Назва сорту: "Фенікс"

Дата	Врожайність, т
16.02.2019	204,00
24.06.2019	265,00
17.10.2019	302,00
26.02.2020	212,00
29.06.2020	265,00
20.10.2020	298,00
15.02.2021	255,00
23.06.2021	290,00
15.10.2021	236,00

Тенденція зміни: зростаюча



Рис. 38 Звітна інформація зрізу врожайності у розрізі сорту

Висновки

Результуюча звітна інформація демонструє тренд зміни врожайності на прикладі сорту «Фенікс» культури огірок, який вирощувався у теплиці «Цех №1». Сформована таблиця та графік показують що кінцева врожайність завжди вищується, проте відносно 2019-го року стабільно вища, незважаючи на значні

перепади. Наприклад між 17.10.2019 та 26.02.2020 різниця у врожайності сягає 90 тон, що є найбільшою різницею серед всіх зафіксованих значень. Судачи з того що низька врожайність зафіксована здебільшого у зимовий період можна зробити висновок, що вирощування у такий період року важчий так як необхідно витратити більше ресурсів на обігрів теплиць, щоб забезпечити нормальний ріст та розвиток рослини.

Наступний звіт зображено на рис. 39. Дана звітна інформація демонструє екстремуми показників мікроклімату у теплиці у розрізі врожайності сорту певної культури.

Екстремуми температури у розрізі врожайності сорту

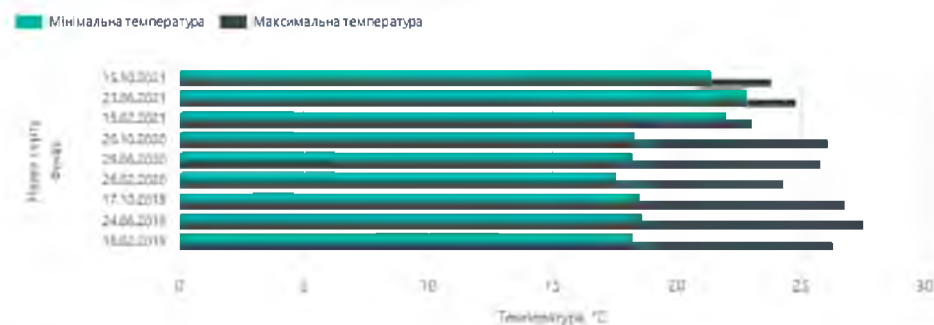
Назва теплиці: Цех №1

Культура: Огірок

Назва сорту: "Фенікс"

	Мінімальна температура	Максимальна температура	Мінімальна вологість	Максимальна вологість
фев 2019	18,20	26,30	74,50	8,40
июн 2019	18,60	27,50	74,40	81,20
окт 2019	18,50	26,80	75,80	79,50
фев 2020	17,60	24,30	71,10	84,40
июн 2020	18,20	25,80	69,00	84,30
окт 2020	18,30	26,10	70,10	76,40
фев 2021	22,00	23,00	66,00	70,00
июн 2021	22,80	24,80	61,00	69,50
окт 2021	21,40	23,80	65,20	69,40

Змінення температури в розрізі дати врожайності сорту



Змінення вологості в розрізі дати врожайності сорту

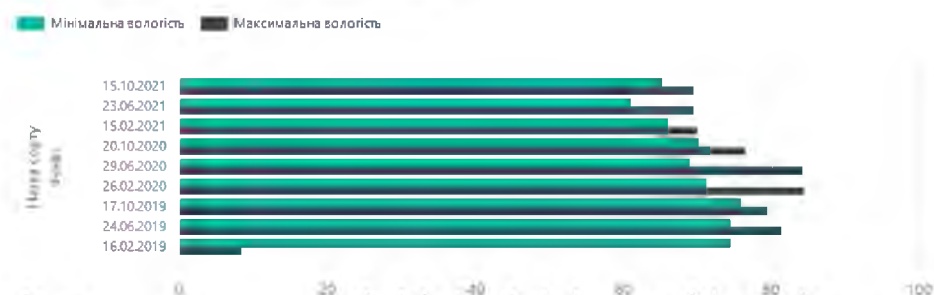


Рис. 39 Екстремуми показників мікроклімату у теплиці у розрізі врожайності сорту певної культури

Висновки

Сформований звіт відображає зафіксовані показники мікроклімату у теплиці у розрізі часу. Звітність такого типу дозволяє проаналізувати коректність роботи всіх виконавчих механізмів, справність функціонування всієї системи загалом. Звернувши увагу на стовпчикову діаграму з показниками вологості можна помітити досить високу вологість у перших двох кварталах 2020-го року, що дає причину проаналізувати якість мікроклімату, який створюється для вирощування рослин.

Також було сформовано звіт, який демонструє порівняння врожайності у розрізі року. Результат сформованого звіту на прикладі 2019-го року зображено на рис. 40.

Зміна врожайності у розрізі року

Рік: 2019

Назва культури	Назва сорту	Врожайність, т
Огірок	Амур	280,00
	Герцог	235,00
	Фенікс	253,50
Перець	Богдан	248,33
	Геркулес	276,20
	Хабанеро	315,00
Помідор	Волове серце	285,00
	Корнівський	269,67
	Черрі	281,50

Порівняння врожайності різних сортів у розрізі часу

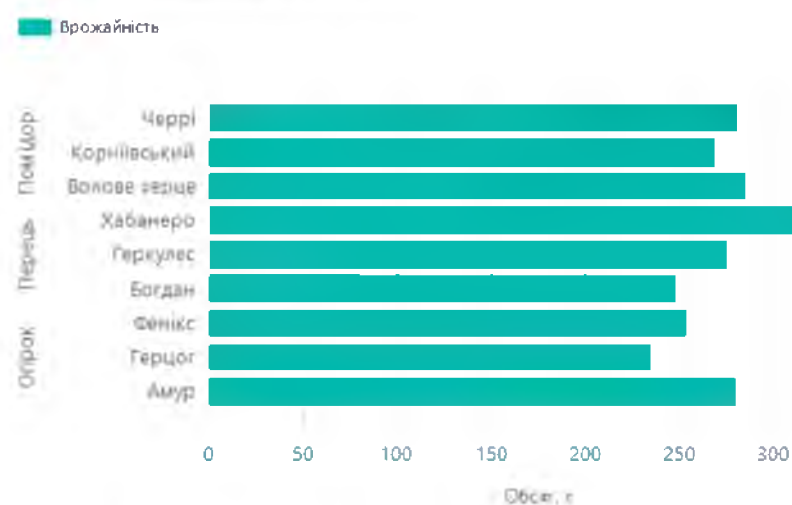


Рис. 40 Зміна врожайності у розрізі року

Висновки

Відповідно до отриманих результатів можна помітити, що середня врожайність має досить високі показники у культур помідор та перець. Лідером серед усіх сортів є перець сорту «Хабанеро». Звіти такої структури дозволяють проаналізувати який сорт чи культуру найбільш вигідно вирощувати для отримання більшої врожайності, а відповідно й більшого прибутку.

4.1.2. Результати розрахунку KPI

рис. 41.

Образить структуру	Значение	Цель	Состояние	Тренд
KPI_Sort1_Opt_Avg_Humidity	65.9999966666667	67.5		↓
KPI_Sort1_Opt_Avg_Temp	21.11111	22		↓
KPI_Sort4_Opt_Avg_Humidity	75.2222166666667	77.5		↓
KPI_Sort4_Opt_Avg_Temp	19.8888866666667	22.5		↓

Рис. 41. Результати розрахунку KPI

Висновки

На основі розрахованих даних можна зробити висновок, що середня температура та вологість за сортами знаходиться у межах норми, проте знизилась порівняно з минулим роком. Такі результати дають привід перевірити систему підтримки мікроклімату у теплицях та переглянути процес регулювання мікроклімату. У таких випадках можлива некоректна робота деяких виконавчих механізмів або ж незадовільний стан теплиць. Якщо надалі така тенденція буде продовжуватись, то це вплине на подальшу врожайність, враховуючи що зафіксовані показники зможуть виходити за межі оптимальних. Такий стан вплине на ріст рослин, які вирощуються на господарстві.

4.2. Інтелектуальний аналіз даних

4.2.1. Алгоритм часових рядів. Після обробки структури з використанням алгоритму часових рядів було сформовано графік, який зображено на рис. 42.

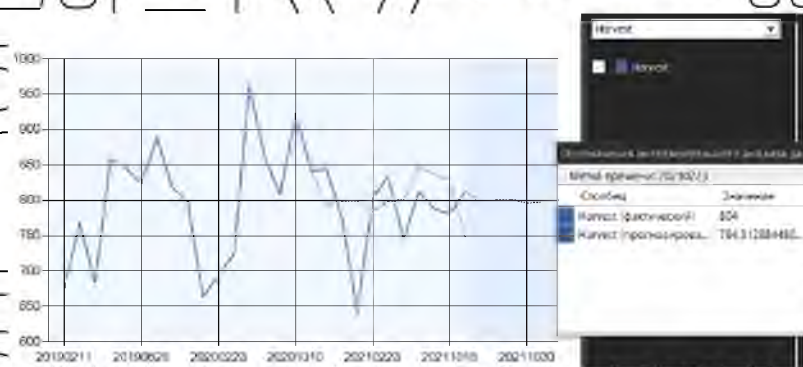


Рис. 42. Сформований графік

Висновки

На основі отриманого графіку можна побачити, що починаючи з початку 2021 року сумарна врожайність у теплицях зменшилась. Наприклад взимку 2021 врожайність досягла найнижчого значення, проте прогнозоване значення значно відрізняється від фактичного. Таку ж картину можна помітити при аналізі за

допомогою інструменту MS Excel (рис. 43). Значення R^2 дорівнює 0,11, що показує низьку залежність даних. Як висновок можна сказати прогнозування за допомогою служби SSAS дає точніший результат ніж в інструментом MS Excel.



Рис. 43 Аналіз врожайності застосовуючи поліноміальну регресійну модель

4.2.2. Алгоритм класифікації OneR. Реалізуючи алгоритм буна

проведена класифікація відповідно до двох класів:

- високий якщо врожайність вище загального середнього значення;
- низький якщо врожайність нижче загального середнього значення.

На рис. 44 показано класифікацію всіх наявних фактів.

Назва культури	Назва сорту	Назва теплиці	Врожайність, т	Середня врожайність, т	Класифікація
Огірок	Амур	Цех №4	275,00000	265,098765	HIGH
Огірок	Амур	Цех №7	253,00000	265,098765	LOW
Огірок	Амур	Цех №7	285,00000	265,098765	HIGH
Огірок	Герцог	Цех №4	220,00000	265,098765	LOW
Огірок	Герцог	Цех №4	280,00000	265,098765	HIGH
Огірок	Герцог	Цех №7	235,00000	265,098765	LOW
Огірок	Фенікс	Цех №1	204,00000	265,098765	LOW
Огірок	Фенікс	Цех №1	212,00000	265,098765	LOW
Огірок	Фенікс	Цех №1	236,00000	265,098765	LOW
Огірок	Фенікс	Цех №1	255,00000	265,098765	LOW
Огірок	Фенікс	Цех №1	265,00000	265,098765	LOW
Огірок	Фенікс	Цех №1	290,00000	265,098765	HIGH
Огірок	Фенікс	Цех №1	298,00000	265,098765	HIGH

Рис. 44 Вибірка та класифікація усіх фактів

Далі було проаналізовано дані відносно незалежної змінної «Культура». На рис. 45 наведено результат класифікації.

Назва культури	Менше ніж серед	Більше ніж середнє	Вся врожайність	Середня врожайні	Загальна середня	Класифікація	Менше очікування	Більше очікування
Помідор	5	13	18	270,888888	265,098765	HIGH	27,78	72,22
Опрок	13	14	27	263,555555	265,098765	LOW	48,15	51,85
Перець	21	15	36	263,361111	265,098765	LOW	58,33	41,67

Рис. 45 Класифікація відповідно до культури

Також було проведено класифікацію за змінною «Теплиця». На рис. 46 показано результат класифікації даних.

Назва теплиці	Менше ніж середнє	Більше ніж середнє	Вся врожайність	Середня врожайніс	Класифікація	Менше очікування	Більше очікування
Цех №1	6	3	9	258,555555	LOW	66,67	33,33
Цех №2	2	7	9	276,555555	HIGH	22,22	77,78
Цех №3	3	6	9	265,222222	HIGH	33,33	66,67
Цех №4	4	5	9	263,666666	LOW	44,44	55,56
Цех №5	7	2	9	260,222222	LOW	77,78	22,22
Цех №6	6	3	9	256,222222	LOW	66,67	33,33
Цех №7	3	6	9	268,444444	HIGH	33,33	66,67
Цех №8	2	7	9	281,555555	HIGH	22,22	77,78
Цех №9	6	3	9	255,444444	LOW	66,67	33,33

Рис. 46 Класифікація відповідно до теплиці

Висновки

За результатами проведеного аналізу можна виділити наступні правила:

– Якщо культура «Помідор», то клас врожайності високий ("High") з імовірністю 72,22%

– Якщо теплиця «Цех №2», то клас врожайності високий ("High") з імовірністю 77,78%.

– Якщо теплиця «Цех №5», то клас врожайності низький ("Low") з імовірністю 77,78%.

– Якщо теплиця «Цех №8», то клас врожайності високий ("High") з імовірністю 77,78%.

Дивлячись на сформовані правила можна зробити висновок, що на класифікацію і рівень врожайності значною мірою впливає теплиця, у якій вирощується продукція. Проте даний алгоритм дає можливість зробити лише

приблизні припущення і у процесі порівняння проаналізувати вплив кожної незалежної змінної, в дній якій проводиться розрахунок.

4.2.3. Алгоритм асоціативних правил. Використовуючи розгорнуту структуру було побудовано мережу залежностей назви культури та сорту від заданих діапазонів температури та вологості (рис. 47).

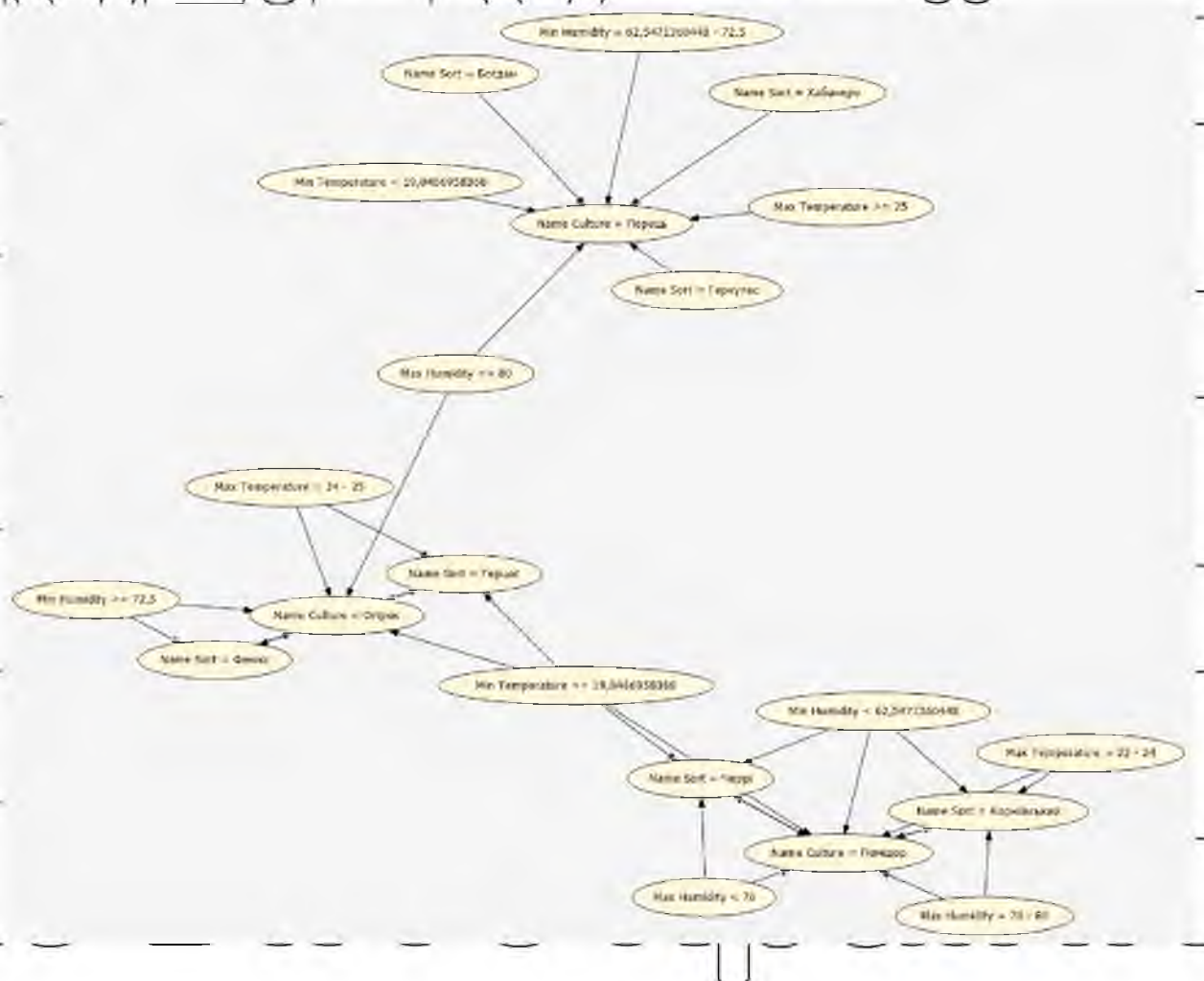


Рис. 47 Мережа залежностей

Висновки

З наведеної схеми асоціативних правил можна визначити найбільш сприятливі умови для вирощування як окремих сортів та культур, так і про можливість комбінування сортів різних культур для вирощування в єдиному просторі. Виокремимо значущі особливості:

- сорти однієї культури мають доволі схожі вимоги до середовища, що робить процес вирощування більш зрозумілим, адже дає підстави для їх вирощування в єдиному просторі;
- перці та огірки мають однакові вимоги до максимальних та мінімальних показників вологості та температури, що дає можливість комбінувати їх сорти під час вирощування в єдиному просторі;
- щодо помідорів, то їх мінімальне значення температури співпадає із показником для огірків, але через більшу чутливість сортів цієї культури, найкращим рішенням буде вирощування їх в окремому просторі.

4.2.4. Алгоритм кластеризації. Розгорнувши структуру аналізу алгоритмом кластеризації було отримано модель, наведену на рис. 48.



Рис. 48 Кластеризація наявних фактів у СД

Аналізуючи діаграму на рис. 48 можна зробити висновок, що найбільше даних зосереджено у кластері 3 ("Cluster 3"). Далі йдуть кластер 1 ("Cluster 1") та кластер 2 ("Cluster 2").

Також можна ознайомитись з більш детальним описом кластерів. На рис.

49 наведена детальна характеристика усіх кластерів..

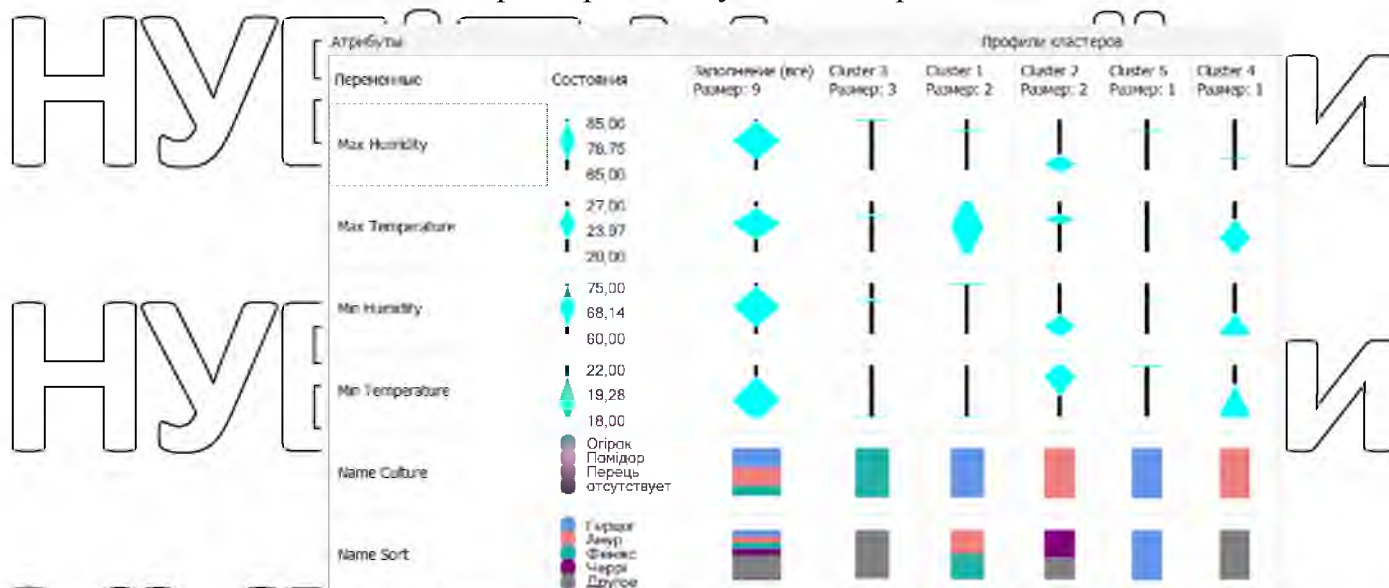


Рис. 49 Профілі кластерів

Для кращого розуміння розглянемо та проаналізуємо кожен кластер.

Кластер 1 містить дані про культуру «Огірок» (99,493 %) та містить у собі сорти «Амур» (імовірність 49,927 %) та «Фенікс» (імовірність 49,566 %).

- Максимальна вологість складає 76,8 - 81,8 % (імовірність: 99,983%)
- Мінімальна вологість складає 72,2 - 75,0 % (імовірність: 52,013 %)
- Мінімальна температура складає 18,0 - 19,3 °C (імовірність: 39,136 %)
- Максимальна температура складає 20,0 - 22,5 °C (імовірність: 18,012%)

Кластер 2 містить дані про культуру «Помідор» (100 %) та містить у собі сорти «Черри» (54,465 %) та «Волове серце» (45,523 %).

- Максимальна температура складає 24,0 - 25,4 °C (імовірність: 66,867%)
- Максимальна вологість складає 65,0 - 71,7 % (імовірність: 61,387 %)
- Мінімальна вологість складає 60,0 - 64,0 % (імовірність: 41,399 %)
- Мінімальна температура складає 20,5 - 22,0 °C (імовірність: 37,641 %)
- Мінімальна вологість складає 64,0 - 68,1 % (імовірність: 25,998 %)

- Мінімальна температура складає 19,3 - 20,5 °C (імовірність: 24,352 %)

Кластер 3 містить дані про культуру «Перець» (99,588 %) та містить у собі сорти «Геркулес» (33,196 %) та «Хабанеро» (33,196 %).

- Мінімальна вологість складає 68,1 - 72,2 % (імовірність: 99,994 %)
- Максимальна температура складає 24,0 - 25,4 °C (імовірність: 98,778%)

- Максимальна вологість складає 81,8 - 85,0 % (імовірність: 51,680 %)

- Мінімальна температура складає 18,0 - 19,3 °C (імовірність: 39,136 %)

Кластер 4 містить дані про культуру «Помідор» (100 %) та містить у собі сорт «Корніївський» (96,448 %).

- Максимальна вологість складає 65,0 - 71,7 % (імовірність: 95,387 %)

- Максимальна температура складає 20,0 - 22,5 °C (імовірність: 37,273%)

- Мінімальна вологість складає 60,0 - 64,0 % (імовірність: 24,981 %)

Кластер 5 містить дані про культуру «Огірок» (100 %) та містить у собі сорти «Герцог» (100 %).

- Максимальна вологість складає 76,8 - 81,8 % (імовірність: 96,517 %)

- Мінімальна вологість складає 68,1 - 72,2 % (імовірність: 95,571 %)

- Максимальна температура складає 24,0 - 25,4 °C (імовірність: 43,750%)

- Мінімальна температура складає 20,5 - 22,0 °C (імовірність: 42,832 %)

- Максимальна температура складає 22,5 - 24,0 °C (імовірність: 42,234%)

В результаті розгорнутої структури можна зробити висновок, що дані є однорідними та впорядкованими. Особливо це помітно на схемі мережі зв'язків

кластери майже не пов'язані один з одним. Так як більшість даних зосереджена

у кластері 3 середнє значення максимальної вологості знаходиться в межах 85%, максимальна температура близько 24,5°C. Мінімальна вологість складає близько 70% та мінімальна температура близько 18°C, що найкраще підходить для

виращування для культури перець.

ВИСНОВКИ

НУБІП України

У ході виконання магістерської кваліфікаційної роботи на тему «Система підтримки прийняття рішень керівництвом тепличного господарства

промислового типу» були розроблені модулі введення, збереження та аналізу даних. Під час аналізу було запропоновано застосування технологій OLAP і Data Mining для проведення інтелектуального аналізу великих обсягів інформації.

НУБІП України

Початковим етапом було проведено аналіз предметної області, а саме процесу керування тепличним господарством, формування постановки технічного завдання, пошук та аналіз існуючих рішень, які реалізують досліджуваний напрямок. Також було визначено акторів системи та розроблено діаграми, які описують предметну область та допомагають зрозуміти важливі бізнес-процеси.

НУБІП України

На етапі роботи над другим розділом було з'ясовані вимоги до СППР, її архітектуру, ключові вузли, проаналізовано мету та алгоритми інтелектуального аналізу даних та обрано інструменти, які допоможуть у реалізації поставлених задач використання технологій OLAP та Data Mining.

НУБІП України

У процесі розробки системи було представлено структуру оперативної БД. Наступним кроком було створення СД, яке разом з інструментом SSAS було використано для розгортання багатовимірного кубу. Створений куб був наповнений за допомогою розроблених потоків даних з використанням інструменту SSIS. Надалі імпортовані дані використані для проведення аналізу.

НУБІП України

Використовуючи вищезгаданий інструмент SSAS було створено структури інтелектуального аналізу даних, розраховано ключовий показник ефективності, а також розроблено звітну інформацію застосовуючи службу SSRS. Такі звіти забезпечили структуровану та зрозумілу подачу інформації.

НУБІП України

Використовуючи раніше створені алгоритми, у четвертому розділі було проаналізовано результати оперативного й інтелектуального аналізу даних та зроблено висновки щодо отриманих даних, їхній вплив на процес вирощування рослин у спорудах закритого ґрунту.

У результаті аналізу з використанням служби SSRS були отримані наступні дані. На основі створеного звіту було отримано таблицю та графік, які відображають зафіксовані показники мікроклімату у розрізі часу, теплиці та оптимальних показників відносно конкретного сорту. Судячи з результуючих таблиць та графіків можна помітити, що температура та вологість у теплиці напряду впливають на майбутню врожайність. На прикладі врожайності, зафіксованої 16.02.2019, коли було помічено підвищену вологість та понижену температуру, можна побачити що безпосередньо врожайність була найнижчою серед усіх зазначених.

Наступний графік відображає зміну врожайності у розрізі теплиці та сорту, який вирощувався у ній у вказаний часовий проміжок. Результуюча звітна інформація демонструє тренд зміни врожайності на прикладі сорту «Фейкс» культури огірок, який вирощувався у теплиці «Цех №1». Сформована таблиця та графік показують що кінцева врожайність завжди варіюється, проте відносно 2019-го року стабільно вища, незважаючи на значні перепади.

Також було визначено екстремуми показників мікроклімату у теплиці у розрізі врожайності сорту певної культури. Звітність такого типу дозволяє проаналізувати коректність роботи всіх виконавчих механізмів, справність функціонування всієї системи загалом. Звернувши увагу на стовпчикову діаграму з показниками вологості можна помітити досить високу вологість у перших двох кварталах 2020-го року, що дає причину проаналізувати якість мікроклімату, який створюється для вирощування рослин.

Наступний звіт демонструє порівняння врожайності у розрізі року. Відповідно до отриманих результатів можна помітити що середня врожайність має досить високі показники у культур помідор та перець. Лідером серед усіх сортів є перець сорту «Хабанеро». Звіти такої структури дозволяють проаналізувати який сорт чи культуру найбільш вигідно вирощувати для отримання більшої врожайності, а відповідно й більшого прибутку.

Також у процесі аналізу було розраховано ключовий показник ефективності, який демонструє, що середня температура та вологість за сортами

знаходиться у межах норми, проте знизилась порівняно з минулим роком. Такі результати дають привід перевірити систему підтримки мікроклімату у теплицях та переглянути процес регулювання мікроклімату.

Розглянемо результати, отримані у процесі аналізу з використанням засобів Data Mining.

У процесі аналізу за допомогою часових рядів було проаналізовано зміну врожайності. Часові ряди дозволяють прогнозувати як буде далі розвиватися ситуація. Тому на основі отриманого графіку можна побачити, що починаючи з початку 2021 року сумарна врожайність у теплицях зменшилась. Наприклад

взимку 2021 врожайність досягла найнижчого значення, прогнозоване значення відрізняється від фактичного.

У процесі аналізу методом 1-Rule було проведено аналіз в залежності від одного незалежного параметру У результаті було з'ясовано, що на класифікацію і рівень врожайності значною мірою впливає теплиця, у якій вирощується продукція.

Під час проведення аналізу з використанням асоціативних правил були сформовані моделі зв'язків між умовами вирощування і культурами. З отриманої схеми асоціативних правил можна визначити найбільш сприятливі умови для

вирощування як окремих сортів та культур, так і про можливість комбінування сортів різних культур для вирощування в єдиному просторі. Наприклад, перці та огірки мають однакові вимоги до максимальних та мінімальних показників вологості та температури, що дає можливість комбінувати їх сорти під час

вирощування в єдиному просторі. Щодо помідорів, то їх мінімальне значення температури співпадає із показником для огірків, але через більшу чутливість сортів цієї культури, найкращим рішенням буде вирощування їх в окремому просторі.

За допомогою методу кластеризації були отримані зв'язки між різними кластерами, схожі за певними ознаками, культура і умови вирощування. В результаті розгорнутої структури можна зробити висновок, що дані є однорідними та впорядкованими. Особливо це помітно на схемі мережі зв'язків

кластери майже не пов'язані один з одним. Так як більшість даних зосереджена у кластері 3 середнє значення максимальної вологості знаходиться в межах 85%, максимальна температура близько 24,5°C. Мінімальна вологість складає близько 70% та мінімальна температура близько 18°C, що найбільше підходить для вирощування для культури перець.

Тому дивлячись на отримані результати можна підтвердити доцільність використання технологій OLAP та Data Mining для аналізу процесів тепличного господарства. Обидві технології доповнюють одна одну і в той час як DM знаходить закономірності на основі відомих знань, то OLAP аналізує дані в режимі реального часу та може стверджувати або спростовувати гіпотези, наведені у процесі інтелектуального аналізу. Отримані знання корисні у процесі формування керуючих рішень тепличного господарства, що дозволить спрямувати стратегію щодо управління окремими бізнес-процесами таким чином, щоб підвищити врожайність у теплицях та відповідно прибутковості господарства загалом.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. За 2021 рік Україна збільшила зовнішньоторговельний обіг с/п продукції та продовольчих товарів, – Роман Лещенко [Електронний ресурс] – Режим доступу до сайту: <https://minagro.gov.ua/news/za-2021-rik-ukravina-zbilshila-zovnishnotor-govelnii-obig-se-produkciyi-ta-prodovolchih-tovariv-roman-leshchenko>
2. Україна у цифрах [Електронний ресурс] – Режим доступу до сайту: https://ukrstat.gov.ua/druk/publicat/kat_u/2022/zb/08/zb_Ukraine%20in%20figures21u.pdf
3. Тепличні господарства України [Електронний ресурс] – Режим доступу до сайту: <https://tripoli.land/ua/farmers/teplitsy/kievskaya>
4. Види споруд закритого ґрунту [Електронний ресурс] – Режим доступу до сайту: <https://buki.b.net/books/35505/>
5. Monitoring The Greenhouse Environment [Електронний ресурс] – Режим доступу до сайту: <https://www.greenhousegrower.com/production/plant-culture/special-series/monitoring-the-greenhouse-environment/>
6. Managing Humidity in the Greenhouse [Електронний ресурс] – Режим доступу до сайту: <https://gpnmag.com/article/managing-humidity-in-the-greenhouse/>
7. Research of Decision Support System (DSS) for Greenhouse Based on Data Mining / Wang, C., Wang, L., Dong, P., & Qiao, X. – CCTA, 2007. URL: https://link.springer.com/content/pdf/10.1007%2F978-0-387-77253-0_40.pdf
8. Expert System for Greenhouse Production Management / Yongguang Hu, Jizhang Wang and Pingping Li. DOI: 10.5772/7080, 2010. – Взята з екрану. URL: <https://www.intechopen.com/chapters/6706>
9. Підсистема підтримки прийняття рішень для керування технологічними комплексами: пат. 30556 Україна: МПК G05B 13/04 (2006.01). № u2007 3857; заявл. 10.12.2007; опубл. 25.02.2008, бюл. № 4 2 с. (<https://base.uipv.org/search/IV/search.php?action=viewdetails&IdClaim=118597&chapter=biblio>)

10. Спосіб регулювання мікроклімату в теплиці з використанням алгоритму синтезу нейромережевого регулятора на основі заданого синергетичного закону керування : пат. 142707 Україна: МПК A01G 9/24 (2006.01). № u201911720; заявл. 09.12.2019; опубл. 25.06.2020, бюл. № 12. 6 с.

(<https://base.uipv.org/searchINV/search.php?action=viewdetails&IdClaim=269207>)

11. Decision Support System (DSS) [Електронний ресурс] – Режим доступу до сайту:

<https://corporatefinanceinstitute.com/resources/knowledge/other/decision-support-system-dss/>

12. What Is a Data Warehouse? [Електронний ресурс] – Режим доступу до сайту: <https://www.oracle.com/database/what-is-a-data-warehouse/>

13. What Is a Data Warehouse: Overview, Concepts and How It Works [Електронний ресурс] – Режим доступу до сайту:

<https://www.simplilearn.com/data-warehouse-article>

14. The Concept of Data Mining [Електронний ресурс] – Режим доступу до сайту: <https://www.intechopen.com/chapters/78106>

15. Data Mining Concepts [Електронний ресурс] – Режим доступу до сайту: <https://docs.microsoft.com/en-us/analysis-services/data-mining/data-mining-concepts?view=asallproducts-allversions>

16. Data Mining Techniques [Електронний ресурс] – Режим доступу до сайту: <https://www.tutorialspoint.com/what-are-the-basic-concepts-of-data-mining>

17. Overview of Online Analytical Processing (OLAP) [Електронний ресурс] – Режим доступу до сайту: <https://support.microsoft.com/en-us/office/overview-of-online-analytical-processing-olap-15d2cdde-f70b-4277-b009-e1732b75fd6>

18. The Easy Guide to UML Deployment Diagrams [Електронний ресурс] – Режим доступу до сайту: <https://creately.com/blog/diagrams/deployment-diagram-tutorial/>

19. SSAS Tutorial: What is SSAS Cube, Architecture & Types

[Електронний ресурс] – Режим доступу до сайту: <https://www.guru99.com/ssas-tutorial/html>

20. Data Source Views in Multidimensional Models [Електронний ресурс]

– Режим доступу до сайту: <https://docs.microsoft.com/en-us/analysis-services/multidimensional-models/data-source-views-in-multidimensional-models?view=asallproducts-allversions>

21. SQL Server Integration Services (SSIS) для починаючих – частина 1

[Електронний ресурс] – Режим доступу до сайту: <https://habr.com/ru/post/330618/>

22. What is SQL Server Reporting Services (SSRS)? [Електронний ресурс]

– Режим доступу до сайту: <https://learn.microsoft.com/en-us/sql/reporting-services/create-deploy-and-manage-mobile-and-paginated-reports?view=sql-server-ver15>

23. Key Performance Indicators (KPIs) in Multidimensional Models

[Електронний ресурс] – Режим доступу до сайту: <https://docs.microsoft.com/en-us/analysis-services/multidimensional-models/key-performance-indicators-kpis-in-multidimensional-models?view=asallproducts-allversions>

24. What are Association Rules in Data Mining (Association Rule Mining)?

[Електронний ресурс] – Режим доступу до сайту: <https://www.techtarget.com/searchbusinessanalytics/definition/association-rules-in-data-mining>

25. Clustering in Data Mining [Електронний ресурс] – Режим доступу до

сайту: <https://www.javatpoint.com/data-mining-cluster-analysis>

26. OneR [Електронний ресурс] – Режим доступу до сайту

<https://www.saeedshah.com/oner.htm>

НУБІП України