

НУБІП України

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ БІОРЕСУРСІВ
І ПРИРОДОКОРИСТУВАННЯ УКРАЇНИ

Факультет інформаційних технологій

НУБІП України

УДК 004.9:631.1

«ПОГОДЖЕНО»

«ДОПУСКАЄТЬСЯ ДО ЗАХИСТУ»

Декан факультету

Завідувач кафедри комп'ютерних наук

інформаційних технологій

НУБІП України

Глазунова О.Г., д.п.н., професор

Голуб Б.Л., к.т.н., доцент

« ___ » _____ 2021 р.

«30» листопада 2021 р.

НУБІП України

МАГІСТЕРСЬКА КВАЛІФІКАЦІЙНА РОБОТА

на тему «Система підтримки прийняття рішень при управлінні сільськогосподарським об'єктом»

Спеціальність 122 «Комп'ютерні науки»

НУБІП України

Освітня програма «Інформаційні управляючі системи та технології»

Орієнтація освітньої програми освітньо-професійна

Гарант освітньої програми «Інформаційні управляючі системи та технології»

НУБІП України

д.т.н., доцент

Бондаренко В.Є.

Керівник магістерської кваліфікаційної роботи

к.т.н., доцент

Голуб Б.Л.,

НУБІП України

Виконала

Пронішина К.О.

КИЇВ-2021

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ БІОРЕСУРСІВ І ПРИРОДОКОРИСТУВАННЯ УКРАЇНИ

Факультет інформаційних технологій

ЗАТВЕРДЖУЮ

Завідувач кафедри комп'ютерних наук

Голуб Б.Л., к.т.н., доцент

НУБІП України

“29” жовтня 2020 року

ЗАВДАННЯ

ДО ВИКОНАННЯ МАГІСТЕРСЬКОЇ КВАЛІФІКАЦІЙНОЇ РОБОТИ СТУДЕНТЦІ

Пронішеній Катерині Олександрівні

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 при управлінні сільськогосподарським об'єктом»

затверджена наказом ректора НУБІП України від “29” жовтня 2020р. №1634 «С»

Термін подання завершеної роботи на кафедру “30” листопада 2021р.

Вихідні дані до магістерської кваліфікаційної роботи:

- 1) дані про продуктивність перепелів;
- 2) дані про врожайність, виробництво та засівну площу с/г культур, зібрані Державною Службою Статистики України.

Перелік питань, що підлягають дослідженню:

№ з/п	Питання, що підлягає дослідженню	Строк виконання	Примітка
1.	Аналіз предметної області	21.09.2020 - 24.10.2020	
2.	Дослідження інструментів OLAP	30.11.2020 - 31.10.2020	
3.	Проектування системи	01.02.2021-06.03.2021	
4.	Дослідження інструментів Data Mining	11.03.2021-10.04.2021	
5.	Розробка алгоритмів аналізу даних	22.04.2021-22.05.2021	
6.	Дослідження отриманих результатів	01.09.2021-13.11.2021	
7.	Попередній захист	30.11.2021	
8.	Захист	14.12.2021	

Дата видачі завдання “29” жовтня 2020 р.

Керівник магістерської кваліфікаційної роботи

Завдання прийнята до виконання

(підпис)

(підпис)

Голуб Б.Л.

(прізвище та ініціали)

Пронішина К.О.

(прізвище та ініціали)

НУБІП України

НУБІП України

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

5

ВСТУП

НУБІП України

6

1.0 10
1.1 10

1.2 14

НУБІП України

1.3 16
1.4 18

2 24

НУБІП України

2.1 24
2.2 26

2.3 27

НУБІП України

2.4 28
2.5 30

3 31

3.1 31

НУБІП України

3.2 32
3.3 33

3.4 35

НУБІП України

3.5 38
3.6 41

3.7 44

3.8 45
3.9 47
3.10 49

НУБІП України

3.11 52
4 59
4.1 59

НУБІП України

4.2 66

НУБІП України

ВИСНОВКИ

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

Додаток А

Додаток Б
Додаток В
Додаток Г

НУБІП України

Додаток Д

НУБІП України

НУБІП України

НУБІП України

68

71

75

78

80

83

86

НУБІП України

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

С-Г – сільське господарство

ВВП – внутрішньо валовий продукт

АПК – аграрно промисловий комплекс

UML – уніфікована мова моделювання

СД – сховище даних

БД – база даних

ІС – інформаційна система

СППР – система підтримки прийняття рішень

OLAP – Online Analytical Processing

SSIS – SQL Server Integration Services.

SSAS – SQL Server Analysis Services.

SSRS – SQL Server Reporting Services.

НУБІП України

НУБІП України

НУБІП України

НУБІП України

ВСТУП

Актуальність

Сільське господарство займає одну з найважливіших позицій в економіці України. Воно забезпечує значні частки ВВП, експорту та надходжень до Бюджету України, велику кількість робочих місць.

Окрім того, що сільське господарство є одним з найважливіших галузей економіки, воно є ще й провідною галуззю по темпам розвитку. За останні роки спостерігається тенденція на збільшення кількості підприємств АПК, а існуючі прагнуть до розширення територій та сфер діяльності.

Але існує проблема в автоматизації управління с/г об'єктами та аналізу їх роботи.

Питання автоматизації управління роботою об'єкта вирішується за допомогою встановлення різних систем: датчиків, сенсорів, мікроконтролерів, віддаленого керування, тощо. Проблеми ж автоматизації аналізу роботи об'єкта приділяють менше уваги, хоча це є одним із важливіших процесів, оскільки коректно проведений аналіз отриманих даних може значно покращити якість прийнятих рішень, щодо корегування роботи с/г об'єкту. Такий аналіз складно проводити вручну, без застосування спеціалізованого програмного забезпечення, оскільки зростає ймовірність припущення помилки через людський фактор.

Особливо це стосується великих підприємств, через величезний об'єм інформації, яку необхідно обробити.

Управління с/г об'єктом має бути спрямованим на розвиток цього об'єкту та покращення результатів його роботи. Управління об'єктом не може бути дієвим без аналізу накопичених показників його роботи. Саме аналіз внесених даних і є метою розробки системи підтримки прийняття рішень. На часі для багатьох керівників підприємств у сфері аграрного бізнесу існує необхідність створення інноваційної системи, що дозволяла б на основі накопичених даних

приймати правильні рішення. Тож, в цій роботі пропонується використання сучасних інформаційних технологій OLAP та Data Mining.

Предмет та об'єкт дослідження

Предметом дослідження є процес прийняття рішень у керуванні роботи с/г об'єкту, корегуванні його роботи з метою збільшення виробництва певного продукту, доходу, тощо.

Об'єктом дослідження є СППР при управлінні с/г підприємством з різними напрямками роботи.

Мета дослідження

Необхідно дослідити доцільність використання інформаційних технологій таких, як OLAP та Data Mining, при прийнятті рішень щодо регулювання роботи с/г підприємства з різними напрямками роботи.

Зміст поставлених завдань

Процес розробки та дослідження системи, яка є об'єктом дослідження цієї наукової роботи, був розбитий на такі етапи:

1. Дослідити предметну область: визначити проблему, її актуальність та існуючі методи її вирішення.
2. Визначити інструменти, які будуть застосовані в створенні СППР.
3. Побудувати архітектуру системи, спроектувати СД, в якому буде зберігатися вся необхідна для дослідження інформація.
4. Розробити алгоритми роботи СППР при управлінні с/г підприємством.
5. Проаналізувати отримані результати.
6. Сформулювати висновки, щодо доцільності використання обраних технологій у керуванні с/г підприємством.

Методи дослідження

Технології OLAP були використанні для представлення даних в більш

зрозумілій для аналізу структури, у вигляді багатомірного куба, що забезпечує можливість швидкого аналізу зведених або детальних даних, формування звітів тощо.

Технологія Data Mining була використана для знаходження нетривіальних

зв'язків між даними, які описують роботу с/г підприємства.

Наукова новизна

Вперше було досліджено використання технологій Data Mining та OLAP для аналізу роботи с/г підприємства, яке займається розведенням перепелів або вирощуванням с/г культур.

Позитивні результати дослідження дозволяють зробити висновок про доцільність використання таких технологій щодо побудови СППР при управлінні сільськогосподарським об'єктом.

Апробація результатів дослідження

1. Пронішина К.О. Аналіз процесу управління сільськогосподарським об'єктом // Збірник матеріалів XI Міжнародної науково-практичної конференції молодих вчених «Інформаційні технології: економіка, техніка, освіта». – Київ. –

2020 – ст. 126-127 (тези доступні за посиланням: <https://drive.google.com/file/d/1-Y7005n4Gu7ч6YlqtIP6YZuzZEсO4y97/view>)

2. Пронішина К.О. Дослідження використання методу NAIVE BAYES для вирішення задачі класифікації в розрізі СППР керівництва перепелиного господарства // Збірник матеріалів конференції «Теоретичні та прикладні аспекти розробки комп'ютерних систем». – Київ. – 2021. – ст. 49-50 (тези доступні за

посиланням: <https://drive.google.com/file/d/1e77WfQOL5T86sbVmPKbcerO676kOvBV4/view>)

3. Пронішина К.О. СППР при управлінні с/г об'єктом // Збірник матеріалів XII Міжнародної науково-практичної конференції молодих вчених «Інформаційні технології: економіка, техніка, освіта». – Київ. – 2020. – 115-116

(тези доступні за посиланням:
<https://drive.google.com/file/d/1RO94MA9GJoVfRJVYxg1nzolHPAnRyAgd/view>)

Структура магістерської роботи

Магістерська робота складається з 4 розділів.

Перший розділ присвячений аналізу та моделюванню процесу керування с/г об'єктом, формування технічного завдання та аналізу існуючих джерел.

У другому розділі описані теоретичні передумови створення СНІР, обґрунтування вибору інструментів, які були використані в цій роботі.

Під час роботи над третім розділом проекту були розроблені архітектура системи, структура СД, алгоритми аналізу даних, створені відповідні діаграми.

У четвертому розділі викладений опис впровадження системи, проведений аналіз отриманих результатів та сформувані висновки.

Текстовий матеріал доповнюють діаграми, графіки таблиці та інше. В роботі містяться:

- кількість ілюстрацій – 53;
- кількість таблиць – 4;
- кількість джерел – 28.

НУБІП України

НУБІП України

НУБІП України

1 СИСТЕМНИЙ АНАЛІЗ ПРОЦЕСУ КЕРУВАННЯ С/Г ПІДПРИЄМСТВОМ

1.1 Аналіз процесу керування с/г підприємством

Аграрний сектор відіграє визначну, стратегічну роль для економіки України. Він забезпечує близько 10% ВВП України, значну частку експорту та надходжень до державного бюджету[1]. Для порівняння у табл.1 наведені найбільш та найменш вагомі частки ВВП, відповідно до даних зібраних Державної служби статистики України [2, 3]. Бере участь у формуванні продовольчої безпеки країни, тобто здатності держави задовольняти потреби населення в харчуванні на рівні не нижче обґрунтованих норм[4].

Структура ВВП України у 2020 році

Таблиця 1

Галузь економічної діяльності	Частка ВВП
Оптова та роздрібна торгівля, ремонт автотранспортних засобів і мотоциклів	13,96
Переробна промисловість	10,13
Сільське господарство, лісове господарство та рибне господарство	9,27
Державне управління й оборона, обов'язкове соціальне страхування	7,23
Операції з нерухомим майном	6,38
...	...
Мистецтво, спорт, розваги та відпочинок	0,54
Водопостачання, каналізація, поволження з відходами	0,39

Аграрна сфера економіки забезпечує населення великою кількістю робочих місць. В Україні налічується приблизно 50 тисяч с-г підприємств. Кількість діючих підприємств у порівнянні з іншими економічними сферами наведено у табл.2[5].

Кількість діючих підприємств за видами економічної діяльності з розподілом на великі, середні, малі та мікропідприємства у 2020 році

Таблиця 2

Вид економічної діяльності	Усього, одиниць	У тому числі			
		великі підприємства, одиниць	середні підприємства, одиниць	малі підприємства, одиниць	з них мікропідприємства, одиниць
оптова та роздрібна торгівля; ремонт автотранспортних засобів і мотоциклів	98369	160	3120	95089	84327
сільське, лісове та рибне господарство	49452	36	2134	47282	42042
промисловість	47806	243	4918	42645	33079
операції з нерухомим майном	36798	2	372	36424	33741
надання інших видів послуг	3206	-	39	3167	2862
освіта	2741	-	48	2693	2299
мистецтво, спорт, розваги та відпочинок	2335	2	100	2233	1928

Для дослідження доцільності використання інформаційних технологій при прийнятті рішень, щодо керування с/г підприємством було обрано два напрямки роботи підприємства: розведення перепелів та вирощування с/г культур.

Вирощування перепелів

Останнім часом у світі визначилася чітка тенденція до збільшення виробництва продукції птахівництва. Цей інтерес обумовлюється не тільки

характеристиками продукції, а й порівняно коротким терміном відтворення, що робить цей напрямок привабливим для підприємців різного рангу.

Вирощування перепелів в Україні не таке розповсюджене, як вирощування курей, але перепела мають ряд переваг.

По-перше, перевагою перепелиного м'яса є високий вміст жиро- та водорозчинних вітамінів, мікро- та макроелементів, підвищена кількість лізоциму, сухих речовин, жиру, білка. Цінується м'ясо перепелів насамперед за

його дієтичні властивості. Воно є одним із джерел повноцінного білка в харчуванні людини має приємний аромат, ніжну консистенцію, володіє добрими смаковими якостями, високою соковитістю.

По-друге, перепелині яйця – це продукт, що містить набір всіх поживних речовин. За вмістом вітамінів та мінеральних речовин не поступаються

курячим, містять у 2-3 рази більше міді, багатьох амінокислот, заліза, кобальту, вітамінів групи В, Р, К, А. Яйця перепелів – цінний продукт харчування, які застосовують для лікування та профілактики, різних захворювань дорослих та

дітей. Вони мають антибактеріальним, протипухлинним та імуномодельючим властивостями, нормалізують жировий обмін, тиск, вміст цукру, підвищують працездатність та імунітет[6].

Робота господарства такого формату описується такими даними:

- інформація про породи перепелів;
- інформація про корм, яким їх годують, а саме:
 - вміст заліза;
 - вміст кальцію;
 - вміст фосфору;
 - енергетична цінність;
 - частка протеїну.
- інформація про розміщення господарства;

НУБІП України

- регіон;
- кліматична зона.
- інформація про умови утримання:

- температура повітря;

НУБІП України

- вологість;
- освітленість.
- інформація про продуктивність перепелів:

- вага особини;

НУБІП України

- вага яйця;
- кількість знесених яєць.

Вирощування с/г культур

Протягом років Україна тримає лідируючі позиції в світі по експорту

зернових культур. Тож можна вважати напрямок вирощування с/г культур одним з найвагоміших в сфері сільського господарства.

НУБІП України

- інформація про с/г культуру;
- інформація про розміщення господарства:

НУБІП України

- регіон;
- кліматична зона.
- інформація про продуктивність с/г культури:

- урожайність культури;

Урожайність сільськогосподарських культур – середній розмір певної продукції рослинництва з одиниці фактично зібраної площі цієї культури, який обчислено як співвідношення валового збору з основних, повторних і міжрядних посівів та фактично зібраної площі цієї культури (для багаторічних насаджень –

як співвідношення валового збору з усієї площі насаджень та площі культури у плодоносному віці).

НУБІП України

- виробництво культури;

Виробництво (валовий збір) сільськогосподарських культур – загальний розмір продукції, зібраної з основних, повторних і міжрядних посівів, у натуральному вираженні. Виробництво культур зернових і зернобобових,

соняшнику наведено у масі після доробки, буряка цукрового фабричного – у масі

без землі та домішок, культур овочевих – включаючи продукцію закритого ґрунту. Виробництво (валовий збір) культур плодоягідних, винограду

загальний розмір продукції, яка зібрана з багаторічних насаджень, незалежно від їх віку, у натуральному вираженні.

площа, з якої зібрано врожай.

Площа, з якої зібрано врожай (зібрана площа) сільськогосподарських культур – площа, на якій фактично проведені роботи зі збирання врожаю сільськогосподарських культур[7].

1.2 Технічне завдання

С/г об'єкт – це одиниця підприємства, яка займається будь-яким видом с/г діяльності: вирощування с/г культур, розведення свійських тварин, тощо.

Управління с/г об'єкта – сукупність рішень, дій уповноважених осіб, щодо керування параметрами роботи окремого об'єкта, наприклад: зміна годування перепелів, для збільшення їх продуктивності; зміна культури для вирощування, відповідно до оцінки її продуктивності в різних кліматичних зонах.

Для аналізу роботи с-г підприємства, напрямками роботи якого є розведення перепелів, та вирощування різних с-г культур необхідно розробити СППР, основою якої є СД. При розробці СД необхідно взяти до уваги те, що воно має забезпечити збереження всіх необхідних даних, які були перелічені в підрозділі

1.1.

Для збереження інформації щодо роботи об'єкта перепелиного господарства необхідно розробити оперативну БД та програмний інтерфейс для введення цієї інформації у систему. Дані, які характеризують роботу такого

об'єкта:

- інформація про розміщення господарства;
- інформація про об'єкт, де вирощують перепелів;

- інформація про породи перепелів;
- інформація про окрему особину;
- інформація про корм, яким їх годують;
- інформація про умови утримання;
- інформація про продуктивність перепелів.

Інформація щодо роботи господарства по вирощенню с-г культур зберігається у Excel файл, приклад такого файлу наведено на рис. 1. Інформація взята з відкритих даних, які наочно має Державна служба статистики України.

Посилання на набори даних подані у розділі «Список використаних джерел» [8-10].

Необхідно забезпечити інтеграцію даних з оперативним БД об'єктів по вирощенню перепелів та файлів, які зберігають інформацію про роботу об'єктів з вирощування с-г культур, у СД.

атрибути	період	Урожайність культур зернових та зернобобових	Виробництво (валовий збір) культур зернових та зернобобових	Площа з якої зібрано врожай культур зернових та зернобобових
Вінницька	2015	46	3768,3	819,3
Волинська	2015	39,4	1062,3	269,7
Дніпропетровська	2015	32,6	3866,2	1186,9
Донецька	2015	28,5	1536,1	538,6
Житомирська	2015	41,8	1459,6	349,2
Закарпатська	2015	37,5	332,4	88,5
Запорізька	2015	29	2728,1	941,7
Івано-Франківська	2015	45,1	688,5	152,6
Київська	2015	51,4	2820	548,4
Кіровоградська	2015	41	3313,9	807,9
Луганська	2015	25	992,8	396,5
Львівська	2015	45,5	1366,3	300,3
Миколаївська	2015	30,5	2896,4	948,2
Одеська	2015	29,2	3489	1193,1

Рис. 1 Фрагмент файлу, який зберігає інформацію щодо роботи підприємства. За допомогою інструментів Data Mining та OLAP знайти відповіді на такі

питання:

- a. У якої породи спостерігається найбільша несучість в розрізі сезонності?
 б. Який склад комбікорму сприяє найкращому розвитку особин в розрізі сезонності?

с. У якої культури спостерігається найбільша родючість у розрізі кліматичних зон?

д. Яка частка земель засіяна с/г культурами в розрізі кліматичних зон?

Необхідно обчислити ключові показники ефективності (KPI) підприємства по двом напрямкам його роботи за рік:

- показник ефективності перепелиного господарства – кількість знесених яєць;

- показник ефективності господарства, яке вирощує с-г культури – врожайність с-г культур.

1.3 Аналіз наявних рішень

Був проведений аналіз статей, тез, патентів та інших наукових робіт, пов'язаних з проблемою автоматизації аналізу роботи с/г підприємством. Нижче наведено опис деяких з них.

Доктор економічних наук Ясінецька І.А. та кандидат економічних наук Мушеник І.М. в своїй роботі «Інформаційні системи і технології в управлінні

діяльністю підприємства» пишуть про те, що сучасні інформаційні технології

здатні не тільки цілком задовольнити вимоги виробничих систем, але й виступають важливою передумовою їхнього розвитку. Застосування

інформаційних технологій при прийнятті рішень, щодо керування роботи підприємства, сприяє підвищенню продуктивності роботи, зменшенню витрат на

роботу підприємства, збільшення дохідності роботи, за рахунок прийняття більш якісних управлінських рішень. Після впровадження ІТ на досліджувані

підприємства, керівникам варто проводити оцінку ефективності та доцільності

такого впровадження. Виходячи з того, що ключовим та опорним показником ефективності роботи підприємства є прибуток, що тісно пов'язаний з обсягом виробництва та реалізації продукції, доцільність та ефективність впровадження

ІТ в керування підприємством прямо пропорційна тенденції зміни прибутку підприємства [11].

Також Мушеник І.М. в спільній роботі з доктором економічних наук Місюк М.В. пише про те, що впровадження інноваційних технологій є одним із ключових факторів підвищення конкурентоспроможності АПК України і

економіки країни загалом. Інноваційні зміни в роботі підприємства, в результаті освоєння нових знань, забезпечують постійне вдосконалення виробництва у межах нових підходів у керуванні підприємством. Інноваційні зміни аграрних

підприємств характеризуються перенесенням акценту на використання принципово нових прогресивних технологій, переходом до випуску високотехнологічної продукції, прогресивними організаційними і

управлінськими рішеннями в інноваційній діяльності. Незважаючи на очевидні переваги у запровадження нових технологій в Україні лише невелика кількість великих підприємств роблять значні вкладення в науково-дослідні розробки [12].

З цих двох статей можна зробити висновок про доцільність внесення інноваційних змін у різні процеси роботи с-г підприємств, щоб досягти оптимізації роботи підприємства та збільшення його прибутку.

Ахтар Сабріна в своєму патенті «Integrated IoT (Internet of Things) system solution for smart agriculture management» описує інтегровану апаратну і програмну платформу, оснащену штучним інтелектом, де дані збираються і

відслідковуються в режимі реального часу віддалено, що забезпечує прогностичний аналіз даних для автоматичного запуску превентивних дій. Описана в патенті

консолідована платформа надає дані в режимі реального часу про зростання сільськогосподарських культур, стан ґрунту, контроль над пестицидами, вибір

добрив, вибір культур, врожайність, дані роботи теплиці, дані про клімат і погоду. Ця платформа включає в себе: апаратну частину, систему моніторингу, систему управління і систему зв'язку. Дані збираються за допомогою супутника, різних датчиків та дронів. Дані проглядаються на віддалених комп'ютерах, ноутбуках або портативних пристроях, таких як планшет або смартфон, де до даних можна отримати доступ як на місці, так і віддалено [13].

До такої системи було б доречно додати технології аналізу даних (Data Mining, OLAP) для різнобічного аналізу зібраних даних та підвищення якості стратегічних рішень відносно управління таким підприємством.

1.4 Моделювання процесу керування с/г підприємством

1.4.1 *Діаграма прецедентів.* В UML діаграма прецедентів (діаграма варіантів використання, use-case) моделюють поведінку системи (предметної області) та допомагають охонити вимоги системи (предметної області).

Діаграми варіантів використання описують функції високого рівня та область застосування системи. Ці діаграми також показують взаємодії між системою та її акторами. Варто зазначити, що варіанти використання та дійові особи на таких діаграмах описують, що робить система і як учасники її використовують, але не те, як система працює всередині. Основні елементи діаграми:

- система – відповідає на питання «Що описується?»;
- актор – відповідає на питання «Хто користується системою?»;
- прецедент (приклад використання) – відповідає на питання «Що роблять актори?».

Призначення діаграми:

- ідентифікація функцій та як з ними взаємодіють актори;
- представлення системи на високому рівні – виділення акторів, які взаємодіють із системою, та функціональних можливостей, що надаються

системою, не заглиблюючись у внутрішню роботу системи;

- ідентифікація внутрішніх та зовнішніх факторів [14].

В Додатку А на рис.1 зображена діаграма прецедентів, яка візуалізує процес керування с/г підприємством.

Основними учасниками процесу є:

- оператор об'єкту (для перемісного господарства: зоотехнік, оператор пташника, тощо; для господарства, що вирощує с-г культури: агроном, меліоратор, тощо);

- керівник об'єкту підприємства (або декількох об'єктів, які мають спільні характеристики);

- менеджер (керівник) всього підприємства.

Опис прецедента «Прийняття стратегічних рішення щодо управління підприємством».

Короткий опис: Менеджер підприємства на основі отриманої інформації щодо роботи об'єктів підприємства, приймає рішення щодо регулювання їх роботи.

Передумови: Менеджер підприємства отримав звітну інформацію, щодо роботи всього підприємства.

Головний потік:

1. Менеджер аналізує ефективність роботи об'єктів підприємства.
2. Формує припущення, щодо залежності результатів роботи певного об'єкта від параметрів його роботи
//E1 – Прийнято рішення про необхідність тестового впровадження змін у роботу об'єкта для перевірки припущень.

3. Менеджер передає розпорядження, щодо змін параметрів роботи об'єкта директору цього об'єкта.

Альтернативні потоки:

//E1 – прийняття рішення про необхідність тестового впровадження змін у роботу об'єкта для перевірки припущень.

1.1. Менеджер визначає об'єкт, в якому будуть впроваджені зміни та тестовий період часу, протягом якого ці зміни будуть впроваджені.

1.2. Директору об'єкта надають розпорядження про зміни параметрів роботи.

1.3. Після закінчення тестового терміну проводиться ретельний аналіз отриманих результатів роботи об'єкта та порівняння отриманих результатів з попередніми.

1.4. Менеджер робить висновок, щодо підтвердження або спростування припущення.

Кінець опису прецедента.

1.4.2 Діаграма активності. В UML діаграма активності надає уявлення про поведінку системи, описуючи послідовність дій у процесі.

Діаграми активності схожі на блок-схеми, оскільки вони показують перебіг між діями; однак такі діаграми також можуть показувати паралельні, одночасні та альтернативні потоки.

Варіанти використання діаграм активності:

1) створення діаграм активності для моделювання найважливіших процесів предметної області;

2) створення діаграм активності, щоб покроково проілюструвати перебіг подій, які описує діаграма прецедентів;

3) створення діаграм активності на етапі проектування, щоб допомогти визначити поведінку операцій розроблюваної системи[15].

На рис.2 зображено діаграму активності, яка ілюструє процес керування с/г господарством. Діаграма розділена кількома доріжками, які виокремлюють дії учасників цього процесу, які були виділені в підпункті 1/4.1.



Рис. 2 Діаграма активності

1.4.3 *Діаграма послідовності.* Діаграма послідовності — це діаграма UML, яка ілюструє послідовність повідомлень, що передаються між об'єктами взаємодії. Діаграма послідовності складається з групи об'єктів, які представлені лініями життя, і повідомлень, якими вони обмінюються протягом тривалого часу під час взаємодії.

Діаграми послідовності можуть бути використані на декількох етапах створення системи. На етапі аналізу діаграми послідовності можуть допомогти визначити класи, які потрібні системі, і те, як об'єкти класів взаємодіють. На етапі проектування діаграми послідовності пояснюють, як система працює для здійснення взаємодій між об'єктами. Під час побудови архітектури системи можна використовувати діаграми послідовності, щоб показати поведінку шаблонів проектування та механізмів, які використовує система[16].

Діаграма послідовності на рис.3 ілюструє процес керування с/г підприємством (його окремими об'єктами).

Лінії життя на діаграмі послідовності представляють оператора об'єкта, директора об'єкта(або в його підпорядкуванні можуть бути декілька об'єктів) та менеджера(керівника) всього підприємства або якоїсь його структурної частини. Взаємодія між оператором, директором і менеджером представлено повідомленнями, що передаються між ними.

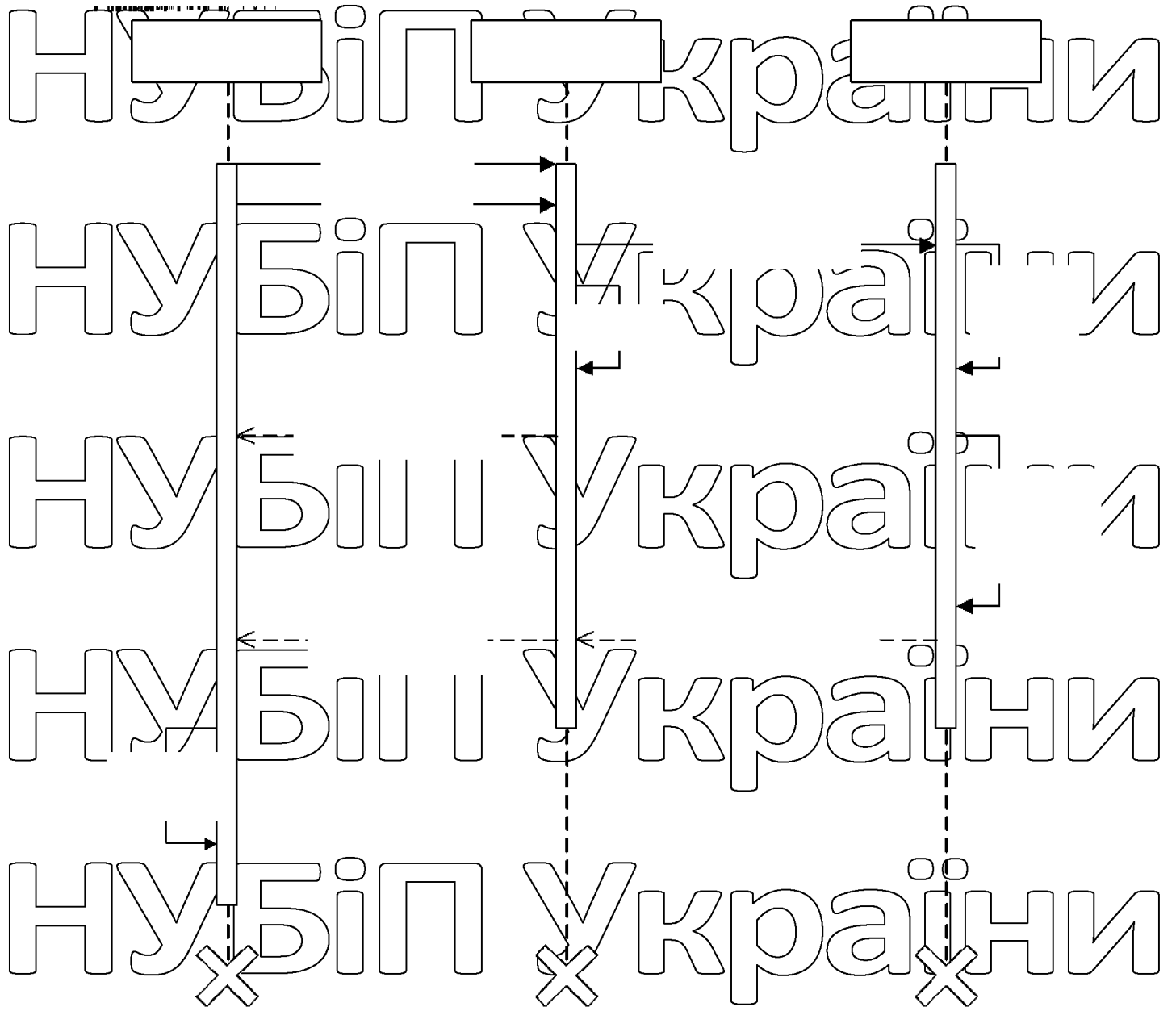


Рис. 3 Діаграма послідовності

НУБІП України

НУБІП України

2 ТЕОРЕТИЧНІ ПЕРЕДУМОВИ

2.1 Визначення СППР

СППР виникли на початку 70-х років завдяки розвитку управлінських ІС,

але досі дослідники не прийшли до єдиного визначення, що таке СППР.

Розглянемо декілька існуючих визначень:

1) основана на використанні моделей сукупність процедур для обробки даних та суджень, яка допомагає керівнику у прийнятті рішень;

2) комп'ютерна інформаційна система, що використовується для підтримки різних видів діяльності у прийнятті рішень в ситуаціях, де неможливо або не бажано мати автоматичну систему, яка повністю виконує весь процес рішення;

3) інтерактивні автоматичні системи, які допомагають особам, які приймають рішення, використовувати дані та моделі, щоб знайти рішення неструктурованих проблем[17].

Варто зазначити, що така система не замінює людину, оскільки не автоматизує повністю процес прийняття рішень, а лише допомагає людині в прийнятті певного рішення. Більшість дослідників сходяться в думці, що СППР призначені для вирішення слабо структурованих проблем. За визначенням Г.Саймона до слабо структурованих проблем відносяться ті, які містять не тільки кількісні, а і якісні параметри[18].

Умовно історію розвитку СППР можна розділити на 3 покоління.

Перше покоління – період 1970-1980 років

Такі системи зберігали великі обсяги інформації у внутрішніх і зовнішніх банках даних, моделі для аналізу даних та алгоритми обчислення показників роботи розроблялися окремими фахівцями в галузі інформатики для вирішення окремих, спеціальних проблем; в користувацькому інтерфейсі великих ЕОМ використовувалися мови програмування, які могли бути зрозумілі виключно

програмістам.

Тож, можна зробити висновок, що такі системи були мало придатні для розуміння пересічних користувачів (тобто тих, які не були спеціалістами в інформаційних технологіях).

Друге покоління – період 1980-1990 років

Такі системи зберігали достатню та необхідну інформацію для аналізу процесу; застосовувалися гнучкі моделі аналізу даних, які можна було налаштувати під вирішення багатьох різнопланових проблем; користувацький інтерфейс став «дружнім» до кінцевого користувача, через використання звичайної мови;

Третє покоління – з 90-х років і дотепер

Системи цього покоління мають ті ж характеристики, що і системи попереднього покоління, але з додатковими можливостями, які стали доступними за рахунок впровадження таких інноваційних засобів інформаційних технологій:

1) сховища, вітрини даних, які дозволяють накопичувати великі обсяги детальних та агрегованих даних;

2) OLAP технології, які надають користувачам можливість швидко та зручно маніпулювати великими обсягами даних, представляти їх у розрізі різних параметрів;

3) Data Mining – методи інтелектуального аналізу, які виявляють не очевидні та нетривіальні зв'язки між даними;

4) засоби телекомунікації, геоінформатичні системи і інші засоби, які можуть бути використані в розробці СППР відповідно до предметної області [19].

Ключові елементи СППР показані на рис. 4.

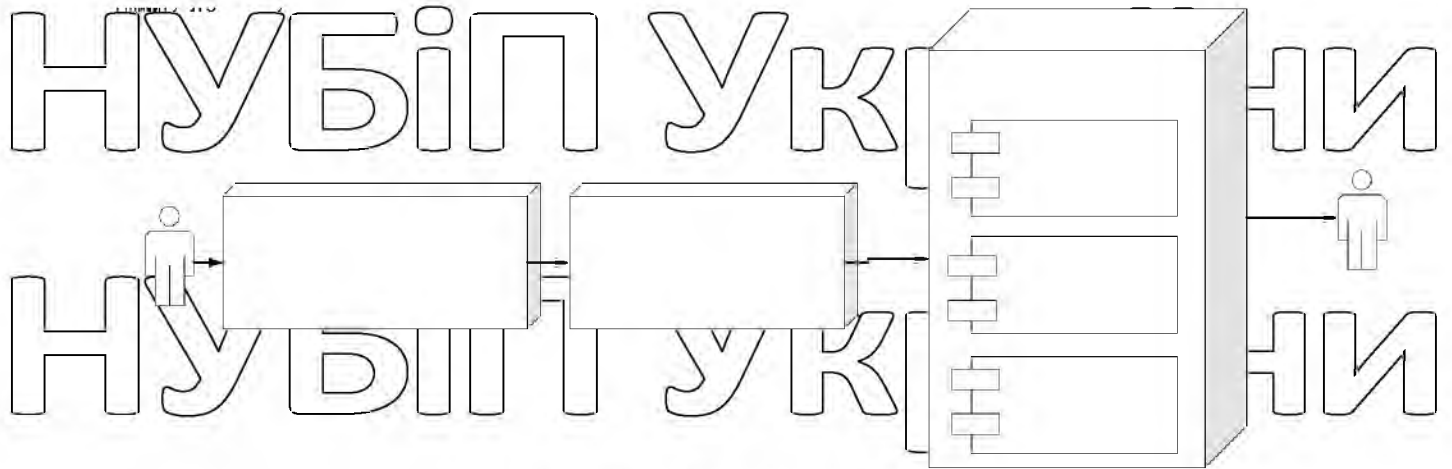


Рис. 4 Архітектура СППР

СППР має реалізовувати ряд концепцій: інтерактивність, інтегрованість, потужність, доступність, гнучкість, надійність, керованість.

2.2 Сховище даних

Сховище даних або корпоративне сховище даних (EDW) – це система, яка об'єднує дані з різних джерел в єдине, центральне, узгоджене сховище даних для підтримки аналізу даних, штучного інтелекту та машинного навчання. Система сховища даних дає змогу організації виконувати потужну аналітику на величезних обсягах історичних даних способом, яким стандартна база даних забезпечити не може.

Системи сховищ даних були частиною рішень бізнес-аналітики (BI) понад три десятиліття, але нещодавно вони отримали значний поштовх у розвитку, через появу нових типів даних і методів розміщення даних. Традиційно сховище даних розміщувалося локально і його функціональні можливості були зосереджені на вилученні даних з інших джерел, очищенні та підготовці даних, а також завантаженні та підтримці даних у реляційній базі даних. Зовсім нещодавно стало можливим таке, щоб сховище даних могло бути розміщене на спеціальному пристрої або в хмарі, також більшість сховищ даних мають деякі інструменти аналітики, візуалізації та презентації даних.

Переваги сховищ даних.

• **краща якість даних:** сховище даних централізує дані з різних джерел даних, таких як оперативні бази даних, файли тощо. Потім файли проходять етап очищення, усуваються дублікати. Дані стандартизуються, щоб створити єдине за структурою джерело даних;

• **швидше приймаються рішення за рахунок бізнес-інсайтів:** те, що дані знаходяться в різних джерелах обмежує здатність осіб, які приймають рішення, заважає визначенню коректної бізнес-стратегії. Сховища даних забезпечують інтеграцію даних з різних джерел, що дозволяє бізнес-користувачам

використовувати, аналізувати всі дані компанії при прийнятті кожного бізнес-рішення;

• **якісніше прийняття рішень:** сховище даних підтримує широкомасштабні BI-функції, такі як інтелектуальний аналіз даних (пошук

невидимих закономірностей і взаємозв'язків у даних), штучний інтелект і машинне навчання – інструменти, які аналітики і керівники бізнесу можуть використовувати, щоб отримати вагомі докази доцільності прийняття рішення практично в кожній сфері організації, від бізнес-процесів до управління

фінансами та управління запасами;

• **отримання та зростання конкурентних переваг:** усе вищезазначене в поєднанні допомагає організації знайти нові знання у накоплених даних, швидше, ніж це можливо з розрізнених сховищ даних[20].

2.3 Data Mining

Існує кілька визначень для інтелектуального аналізу даних як для бізнесу, так і для науковців. Інтелектуальний аналіз даних – це методи, алгоритми, які призначені для пошуку у великому обсягу даних закономірностей, залежностей

та тенденцій, які неможливо виявити за допомогою простого аналізу. Такі

технології дозволяють керівникам підприємств приймати якісні рішення на основі отриманих знань.

Нижче наведено список задач, які вирішується за допомогою методів інтелектуального аналізу.

Класифікація. Об'єкти класифікуються залежно від різних атрибутів.

Характеристики класів відповідно до яких будуть класифікуватися об'єкти

заздалегідь відомі. Наприклад визначенні за допомогою наборів навчальних

даних. Прикладом класифікації у сфері бізнесу є визначення, чи є особа

потенційним клієнтом компанії, залежно від його характеристик, таких як вік,

стать, сімейний статус, професія, освітня кваліфікація тощо.

Кластеризація (сегментація). Залежно від різних атрибутів здійснюється

природне групування об'єктів. На відміну від класифікації кількість груп, та їх

характеристики заздалегідь невідомі. Сегментація клієнтів – класичний приклад

кластеризації у сфері бізнесу.

Прогноз. Прогнозування майбутніх значень показників, на основі аналізу

ретроспективних даних. Прикладом застосування прогнозу в бізнесі є

прогнозування обсягу продажів на найближчі кілька місяців, років.

Асоціація. Пошук зв'язку між об'єктами в одній транзакції. Приклад транзакцій:

продаж у супермаркеті, онлайн-продаж.

Послідовність. Передбачення послідовності подій [21]

2.4 OLAP

OLAP – це сукупність технологій для виконання багатовимірного аналізу з

високою швидкістю на великих обсягах даних, які містяться в сховищі даних,

вітрині даних або іншому уніфікованому централізованому місці збереження

даних.

Більшість даних, які описують певний процес, мають кілька вимірів –

кілька категорій, на які дані розбиваються для представлення, відстеження або

аналізу. Наприклад, показники продажів можуть мати кілька вимірів, пов'язаних

з місцем розташування (регіон, країна, місто, район), часом (рік, місяць, тиждень,

день), продуктом (тип, виробник, ціна), тощо.

В базах даних набори даних зберігаються в таблицях, кожна з яких може організувати дані лише в двох вимірах одночасно. OLAP витягує дані з кількох реляційних наборів даних і реорганізує їх у багатовимірний формат, що забезпечує дуже швидку обробку та дуже глибокий аналіз.

OLAP куб – це багатовимірна база даних на основі масивів, яка дає змогу обробляти й аналізувати кілька вимірів даних набагато швидше й ефективніше, ніж традиційна реляційна база даних.

Таблиця реляційної бази даних структурована так, що зберігає окремі записи у двовимірному форматі «рядок за стовпцем». Кожен «факт» даних у базі даних знаходиться на перетині двох вимірів – рядка і стовпця.

Інструменти звітності SQL (запити, агрегатні функції) та реляційних баз даних, безумовно, можуть надати користувачеві можливість формувати звіти та проводити аналіз багатовимірних даних, що зберігаються в таблицях, але продуктивність сповільнюється зі збільшенням обсягів даних. Також це вимагає багато часу на реорганізування результатів, щоб зосередитися на різних вимірах.

Ось тут на допомогу приходить куб OLAP. Куб OLAP розширює одну таблицю додатковими шарами, кожен з яких додає додаткові виміри, формує ієрархію вимірів. Наприклад, верхній шар куба може організувати продажі за регіонами; додатковими шарами можуть бути країна, штат/провінція, місто і навіть конкретний магазин.

Теоретично куб може містити нескінченну кількість шарів. (Куб OLAP, що представляє більше трьох вимірів, іноді називають гіперкубом.) На практиці аналітики даних створюють OLAP-куби, що містять тільки потрібні їм шари для оптимального аналізу та продуктивності[22].

2.5 Використані інструменти

Для розробки СППР були використані такі програмні інструменти:

- SQL Server Analysis Services (SSIS) – інструмент, який забезпечує технології онлайн аналізу та інтелектуального аналізу даних. Встановлюється на основі SQL Server;
- Excel – табличний редактор, який надає користувачу можливості підключатися до сховища даних та формувати звіти (таблиці, графіки, прогнозування) по даним, які там зберігаються;
- Visual Studio, за допомогою якого були розроблені модулі введення та аналізу інформації. Певні алгоритми аналізу даних були реалізовані за допомогою мови C#, та розроблені в середовищі Visual Studio.

НУБІП України

3 РОЗРОБКА СИСТЕМИ

3.1 Проектування СППР при керуванні с/г об'єктом

В розрізі поставленої задачі було виділено декілька ключових користувачів

СППР при управлінні сільськогосподарським об'єктом. Їх перелік та основні зобов'язання наведені в табл.3.

Для ілюстрації взаємодії учасників з СППР було розроблено діаграму прецедентів(див. Додаток А рис.2).

НУБІП України

Користувачі СППР

Таблиця 3

Учасник процесу	Обов'язки
Оператор с/г об'єкту	Слідкує за роботою об'єкта. Формує журнал роботи об'єкта, тобто вносить дані про роботу об'єкта в оперативну БД, або файли, які зберігають оперативну інформацію.
Директор с/г об'єкта	Контролює роботу об'єкта, або декількох об'єктів. Переглядає звітну інформацію про роботу підпорядкованих об'єктів. Приймає оперативні рішення по управлінню об'єктами господарства. Приймає рішення по ліквідації критичних ситуацій.
Аналітик підприємства	Переглядає результати роботи аналітичної системи. Формує звітність по роботі об'єктів всього підприємства, ціллю якої є допомога у прийнятті рішень по управлінню підприємством.
Топ менеджер підприємства	Приймає стратегічні рішення по управлінню підприємством, окремими об'єктами підприємства.

НУБІП України

3.2 Архітектура системи

Діаграма розгортання UML — це діаграма, яка показує конфігурацію вузлів системи та компонентів, які на них працюють. Діаграми розгортання — це різновид структурної діаграми, яка використовується для моделювання фізичних аспектів об'єктно-орієнтованої системи.

Призначення діаграми розгортання:

- для ілюстрації структури системи;
- для визначення апаратного забезпечення, яке буде використовуватися для реалізації системи, і зв'язків між різними елементами обладнання;
- для моделювання фізичних апаратних елементів та шляхів зв'язку між ними;

- для планування архітектури системи;
- також корисні для документування розгортання програмних компонентів або вузлів [23].

Діаграма розгортання СППР щодо управління с/г об'єктом зображена в Додатку

Б. СППР умовно можна розділити на 3 частини:

- джерела оперативної інформації:
 - БД, яка містить дані про роботу об'єкта – реалізовано для перепелиного господарства (див. Додаток А);
 - файл, який містить дані про роботу об'єктів – джерело інформації, щодо вирощування с/г культур;
- підсистема збереження – СД ;
- підсистема аналізу:
 - підсистема оперативного аналізу – OLAP;
 - підсистема інтелектуального аналізу – Data Mining.

3.3 Оперативні джерела інформації

Дані щодо роботи с/г підприємства, яке виробляє с/г культури зберігаються у файлах відповідної структури (див. підрозділ 1.2).

Для збереження оперативної інформації щодо роботи перепелиного господарства було розроблено. Логічна схема такої БД зображена на рис. 5.

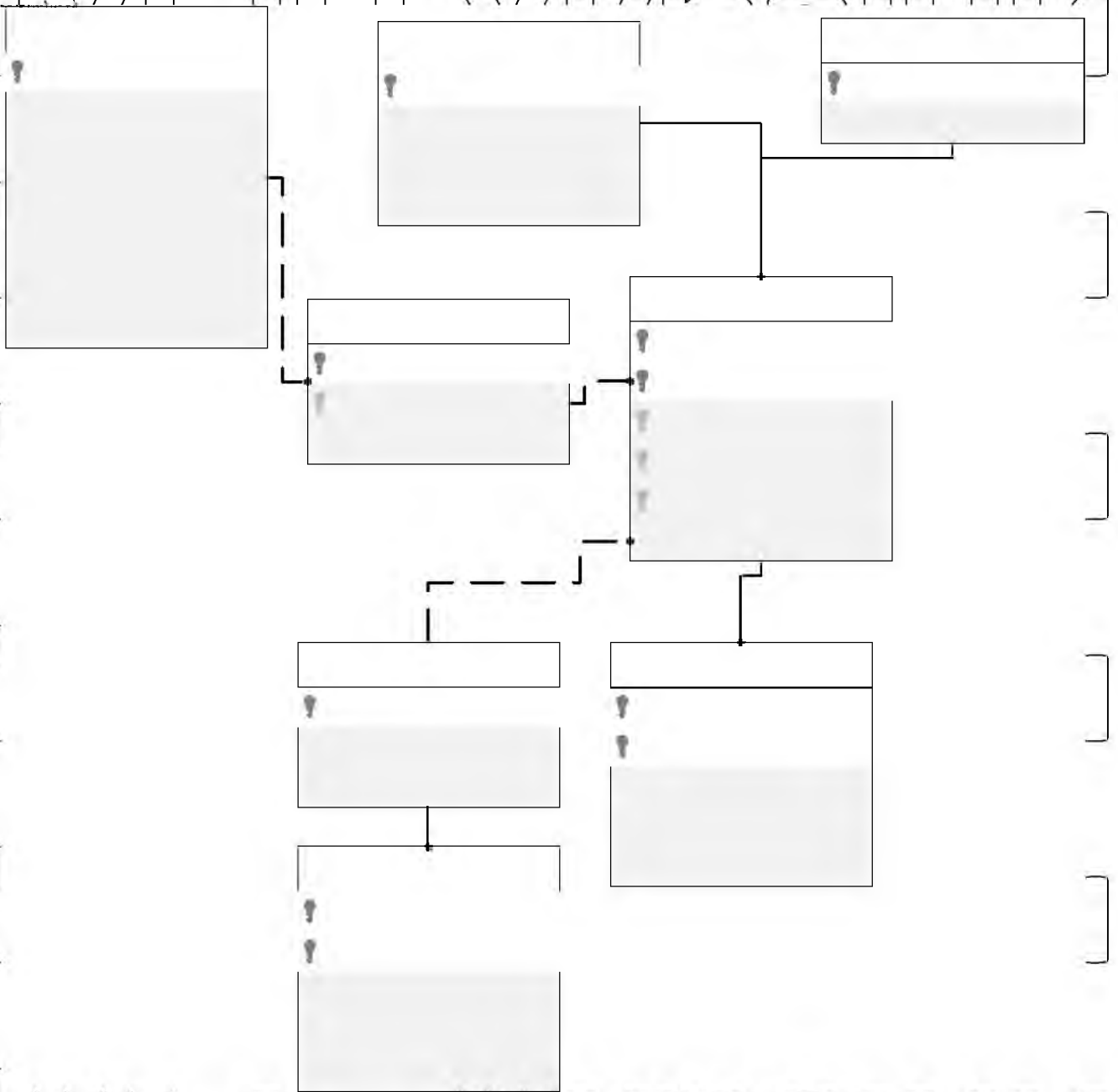


Рис. 5 Оперативна БД

В оперативній БД об'єкта перепелиного господарства зберігається інформація про:

- вид корму, яким годують групу перепелів, та його склад:
 - вміст заліза, фосфору, кальцію, натрію;

● частка протеїну;
● енергетична цінність.

- пташник, де утримують перепелів;
- мікроклімат цього об'єкту (заміри проводяться кожен день);

- порода;
- вікову групу:
 - назва вікової групи;
 - вік в якому перепілка належить до цієї групи.

- продуктивність перепілки (заміри проводяться кожен день):
 - вага перепілки;
 - кількість знесених яєць;
 - максимальна та мінімальна вага яйця зі знесених перепілкою в

цей день.

БД була реалізована в осередку SQL Server за допомогою SQL запитів, приклад запиту на створення таблиці зображений на рис. 6.

```
CREATE TABLE dbo.individual
(
  id_individual char(5) primary key,
  id_breed char(5) NOT NULL,
  id_age_group char(5) NOT NULL,
  id_feed_group char(5) NOT NULL,
  id_facility char(5) NOT NULL,
  weight money NOT NULL
)
```

Рис. 6 Приклад запиту на створення таблиці БД.

Тексти запитів створення всіх таблиць оперативної БД наведено в Додатку Е. Для заповнення БД робочою інформацією був розроблений модуль для

оператора об'єкту господарства.

3.4 Вітрини даних

Збереження даних по роботі всього с/г підприємства для подальшого їх аналізу було забезпечено сховище даних. Вітрина даних є спеціалізованим

сховищем даних, яке містить дані по одному напрямку роботи підприємства. В даній роботі розглядаються два напрямки роботи с/г підприємства, тож були розроблені дві вітрини даних.

В багатовимірному представленні дані ділять на дві групи: виміри та факти.

Вимір – це категоріальний атрибут, властивість, організований у вигляді ієрархічної структури, забезпечує інформаційний контекст числового показника.

Факт – зазвичай кількісний показник, який описує бізнес-процес, його результати. Саме факт є предметом аналізу [24].

Перепелине господарство

На рис. 7 зображена логічна схема вітрини даних, яка відповідає за збереження даних, щодо роботи об'єктів перепелиного господарства. Тип схеми

– крижинка, оскільки наявна деталізована таблиця вимірів.

Виміри:

- Область_Вимір – містить дані про область, де знаходиться господарство, а саме назва регіону та кліматичної зони;

- Господарство_Вимір (є консольною таблицею для таблиці

Область_Вимір) – містить інформацію про господарство;

- Вид_корму_Вимір – містить інформацію про назву та склад корму, яким годують перепелів;

- Порода_Вимір – містить інформацію про породу;

- Дата_Вимір – часовий вимір (Рік – місяць – день).

Факт:

- Продуктивність_Факт, містить інформацію у розрізі господарства.

породи, типу корму про:

НУБІП України

- середня вага перепелів;
- сумарна кількість знесених яєць;

- середнє значення максимальної та мінімальної маси знесених яєць;

НУБІП України

- середнє значення показників мікроклімату:
 - температури;
 - вологості;
 - освітлення.

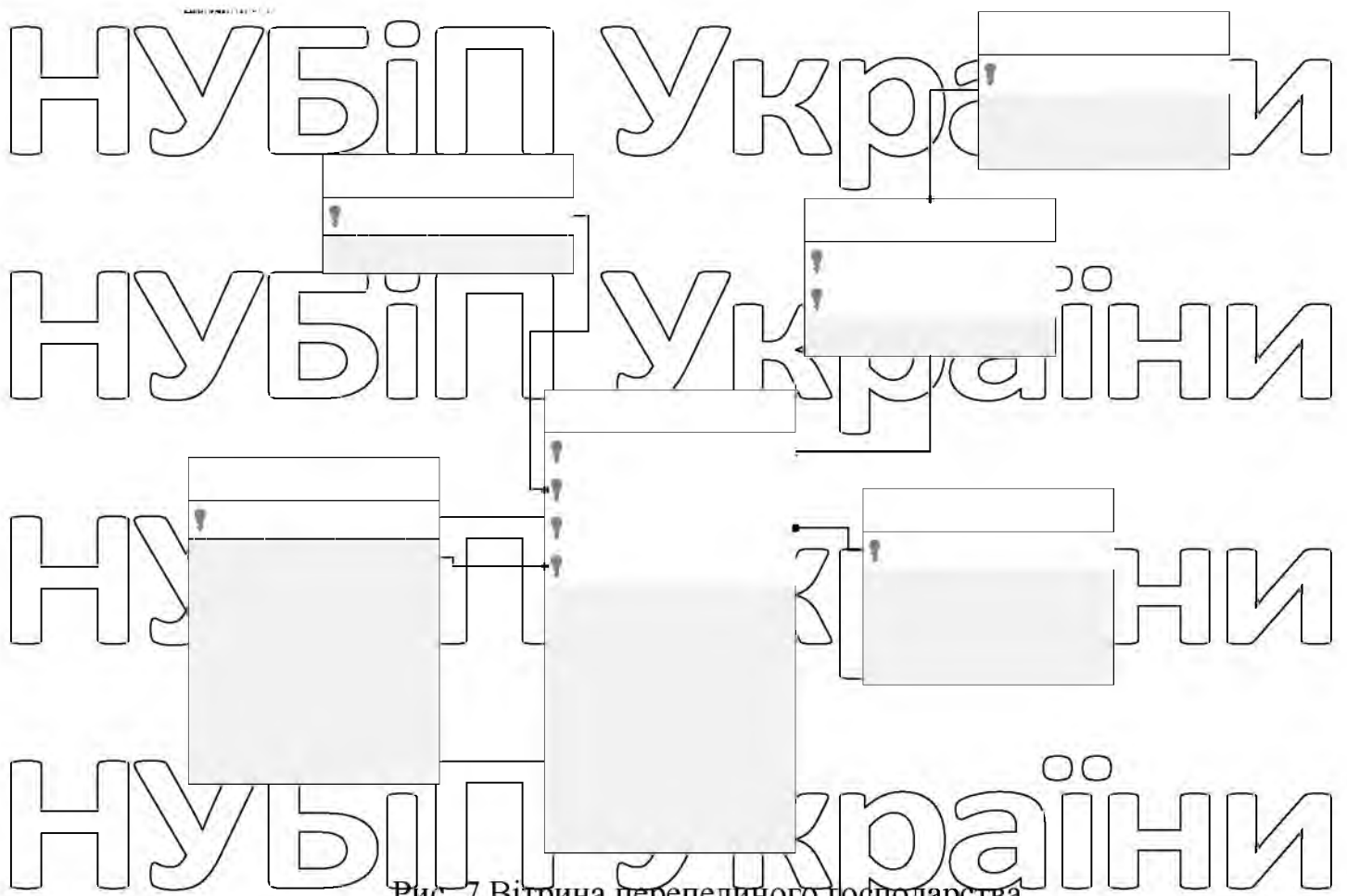


Рис. 7 Вітрина перепелиного господарства

Господарство, яке вирощує с/г культури

На рис.8 зображена логічна схема вітрини даних, яка відповідає за збереження даних, щодо роботи господарства по вирощенню с/г культур. Тип схеми – зірка, оскільки відсутня ієрархія вимірів, таблиці вимірів

НУБІП України

нормалізовані.
 Виміри.
 ● Область_Вимір – містить дані про область, де знаходиться

господарство, а саме назва регіону та кліматичної зони;

● Культура_Вимір – містить інформацію про культуру,
 ● Дата_Вимір – часовий вимір (Рік – місяць – день).
 Факт:

- Урожайність_Факт, містить інформацію в розрізі області, культури

та дати про:
 ○ урожайність культури;
 ○ виробництво культури;
 ○ засівну площу.

НУБІП України

НУБІП України

НУБІП України

НУБІП України

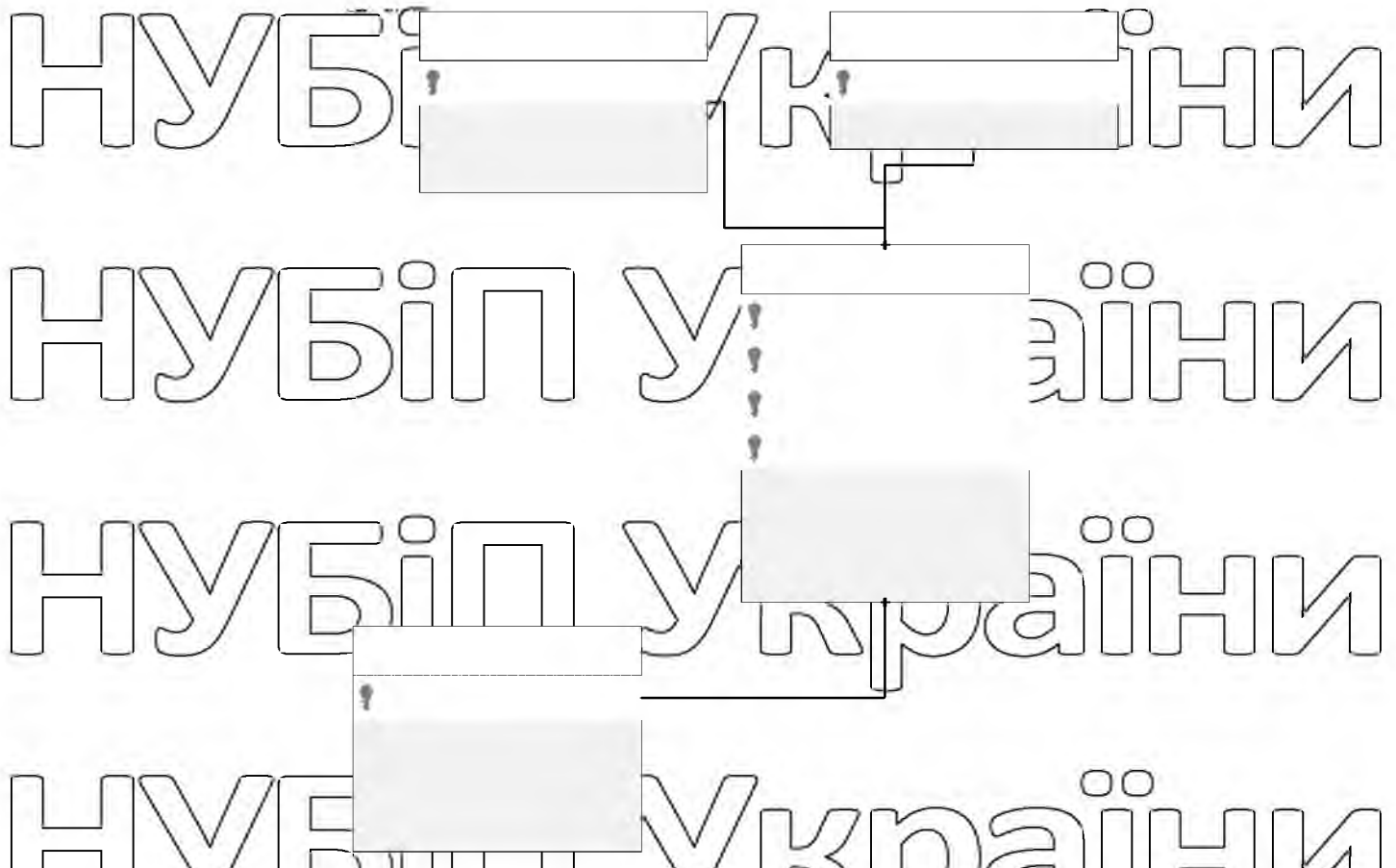


Рис. 8 Вітрина господарства, яке займається вирощуванням сіт культур

В осередку SQL Server вітрини даних розроблялися за допомогою за SQL запитів. Код запитів на створення таблиць наведений в Додатку Г. Інформація в таблицях вимірів є відносно постійною, тож дані цих таблиць заповнювалися за допомогою SQL запитів, приклад якого наведено на рис. 9. Запити на внесення інших вимірів наведені в Додатку Г.

```
INSERT INTO [dbo].[RegionDim]
([id_region]
,[name_region]
,[climate_zone])
VALUES
('1','Івано-Франківська','Полісся'),
('2','Чернігівська','Полісся'),
('3','Львівська','Полісся'),
('4','Рівненська','Полісся'),
('5','Волинська','Полісся'),
('6','Житомирська','Полісся'),
('7','Закарпатська','Полісся'),
```

Рис. 9 Внесення умовно постійної інформації

3.5 Побудова розгорнутого куба підсистеми аналізу, SSAS

Вітрини системи у форматі кубу були розроблені за допомогою Visual

Studio Analysis Services. Analysis Services – це засіб аналітичних даних, що використовується в службі підтримки прийняття рішень і бізнес-аналітики.

Процес розгортання куба в середовищі наведено по-кроково на рис.10-12.

Крок 1. Створення проекту: Analysis Services.

Крок 2. Встановлення зв'язку з джерелом даних: необхідно обрати сервер та спроектоване СД, встановити параметри автентифікації (рис. 10).

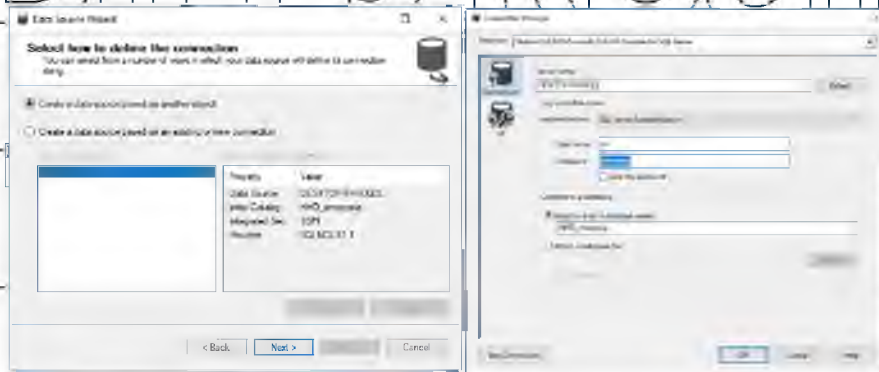


Рис. 10 Додавання джерела даних

Крок 3. Для створення зрізу джерела даних необхідно обрати джерело даних та необхідні таблиці, що будуть відображатися у зрізі (рис.11).

Крок 4. Створення кубу. Для розгортання кубу необхідно обрати таблиці вимірів та фактів (рис. 12).

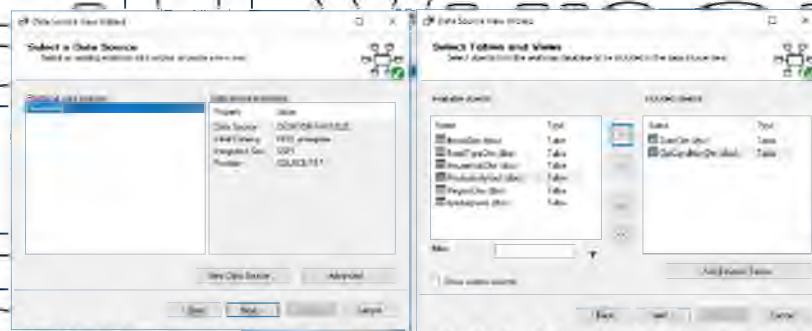
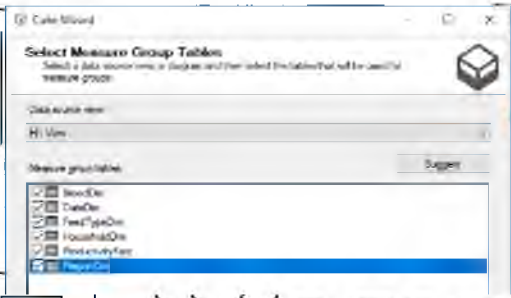


Рис. 11 Створення зрізу джерела та вибір необхідних таблиць

НУБІП України



НУБІП України

Рис. 12 Вибір таблиць

Результати формування кубів для різних напрямків роботи господарств показано на рис. 13-14.

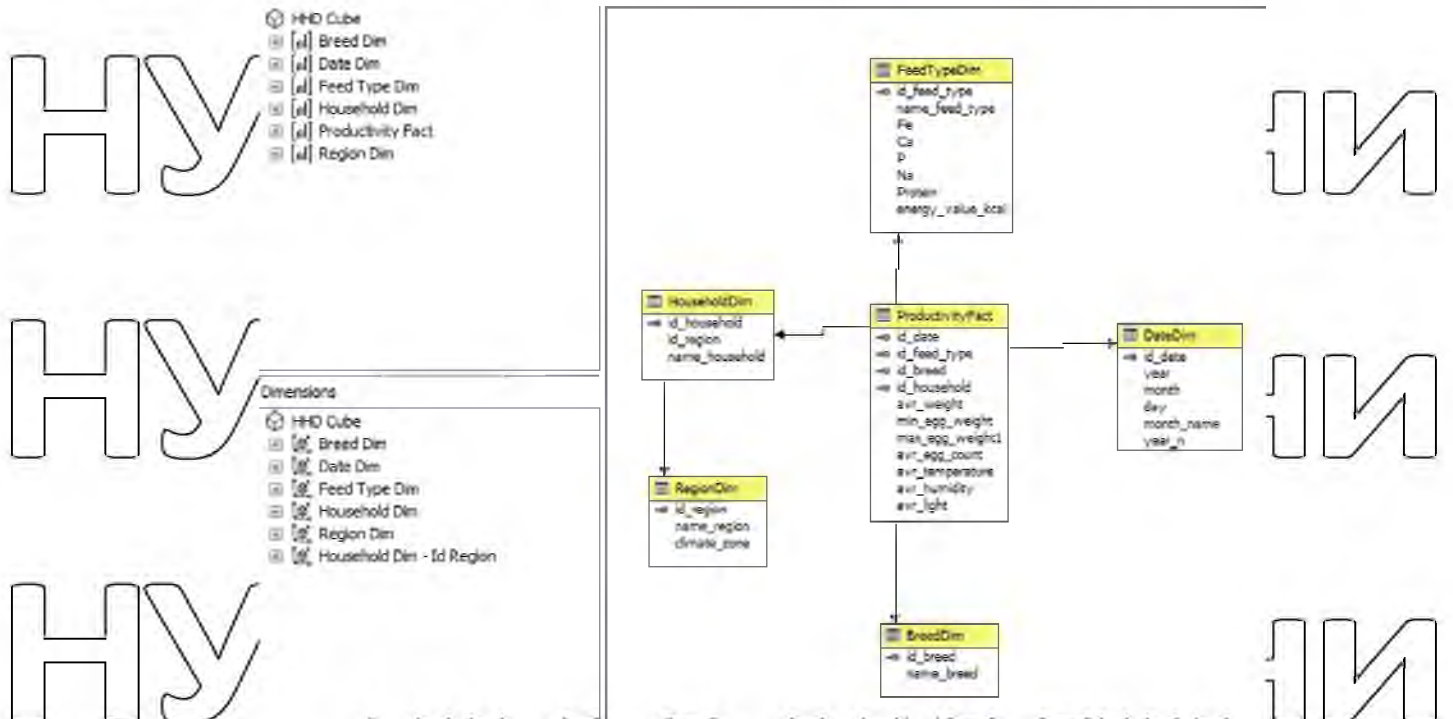


Рис. 13 Сформований куб перепелиного господарства

НУБІП України

НУБІП України

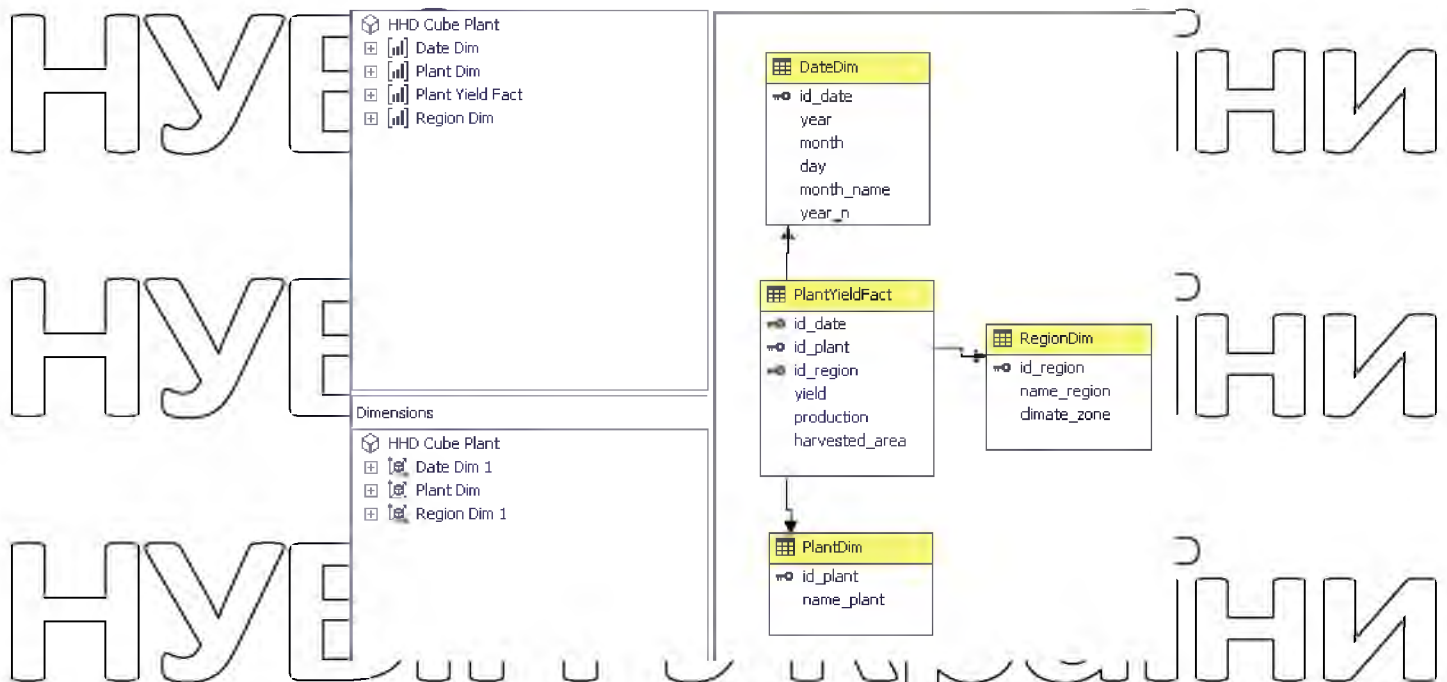


Рис. 14 Сформований куб господарства, яке займається вирощуванням с/г

культур

3.6 Реалізація отримання даних за допомогою Data Flow, SSIS

SSIS - це інструмент, який дозволяє в зручному вигляді реалізувати інтеграцію, тобто реалізувати процес перенесення даних з одного джерела в інший. Цей процес іноді називають ETL (від англ. Extract, Transform, Load - дослівно «витяг, перетворення, завантаження») [25].

Для реалізації потоків даних за допомогою Visual Studio BI необхідно створити проект Integration Server.

Для організації потоків даних необхідно підключитися до джерела даних та приймача даних. Налаштування підключення до джерела та приймача показано на рис.15.

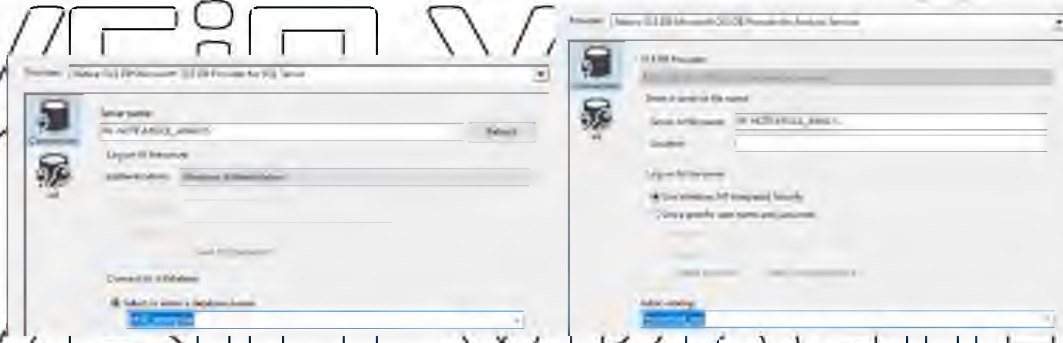


Рис. 15 Підключення до джерела та приймача даних

Щоб уникнути помилок при транспортуванні даних, які пов'язані зі зв'язками між таблицями потоки даних організовані в декілька рівнів (рис.16).



Рис. 16 Потік задач

- 1й рівень – дані транспортуються в таблиці вимірів RegionDim (консольна таблиця до HouseholdDim), DateDim, FeedTypeDim, BreedDim (рис.17);

- 2й рівень – дані транспортуються в таблицю вимірів HouseholdDim;

- 3й рівень – транспортуються детальні дані (по дням) в таблицю фактів ProductivityDim;

- 4й рівень – транспортуються агреговані дані (по місяцям) в таблицю фактів ProductivityDim. Запит на обчислення агрегованих даних зображено на рис.18.

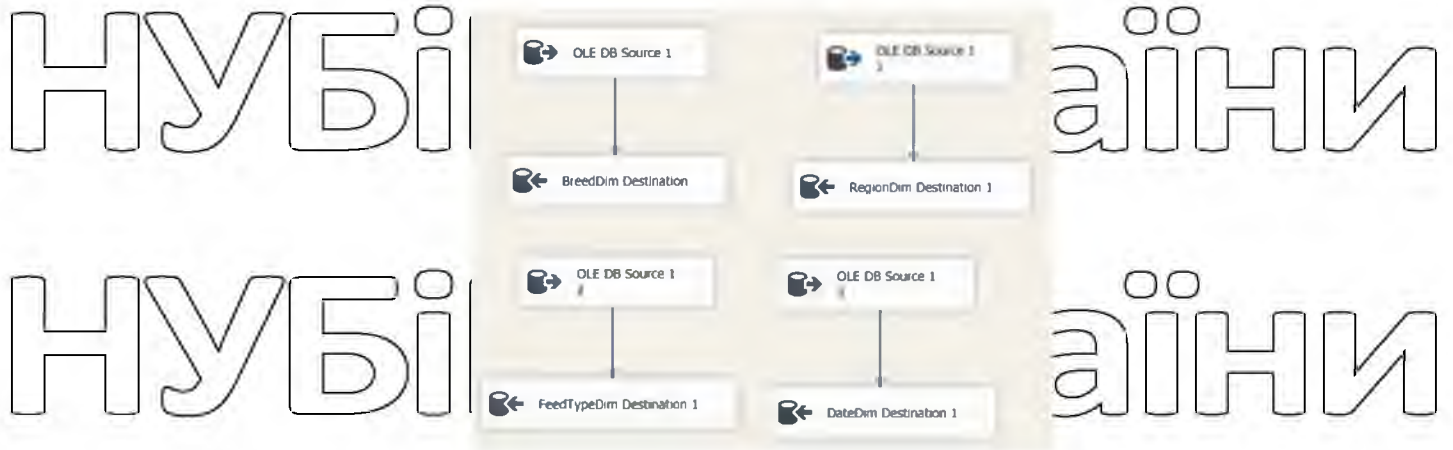


Рис. 17 Поток даних на 1-ому рівні

```

Data access mode:
SQL command
SQL command text:
select DateDim.id_date,
select table.id_feed_type, select table.id_breed, select table.id_household,
select table.avg_avr_weight, select table.avg_min_egg_weight, select table.avg_max_egg_weight1,
select table.sum_avr_egg_count,
select table.avg_avr_temperature, select table.avg_avr_humidity, select table.avg_avr_light
from DateDim, (select DateDim.year, DateDim.month, ProductivityFact.id_household, ProductivityFact.id_feed_type,
ProductivityFact.id_breed,
AVG(avr_weight) as avg_avr_weight, AVG(min_egg_weight) as avg_min_egg_weight, AVG(max_egg_weight1) as
avg_max_egg_weight1,
SUM(avr_egg_count) as sum_avr_egg_count,
AVG(avr_temperature) as avg_avr_temperature, AVG(avr_humidity) as avg_avr_humidity, AVG(avr_light) as avg_avr_light
from ProductivityFact, DateDim
where ProductivityFact.id_date = DateDim.id_date
group by DateDim.year, DateDim.month, ProductivityFact.id_household, ProductivityFact.id_feed_type,
ProductivityFact.id_breed) as select_table
where DateDim.year = select_table.year and select_table.month = DateDim.month and DateDim.day = 0

```

Рис. 18 Налаштування потоку даних на 4-ому рівні

Поява зелених позначок після запуску потоку даних сигналізує про позитивне, тобто без помилок, завершення транспортування даних (рис. 19).



Рис. 19 Індикатори позитивного завершення роботи потоку даних

За допомогою SSIS також було організовано перенесення даних з файлу, який зберігає дані про роботу с/г об'єктів по вирощуванню культур. Алгоритм організації Data Flow схожий на вищеописаний, тільки за джерело даних обирається не БД, а файл. Схема такого потоку даних зображена на рис.20.



Рис. 20 Схема потоку даних

3.7 Побудова звітності, SSRS

SQL Server Reporting Services (Служби звітності SQL Server) - програмна система створення звітів, розроблена корпорацією Microsoft. Вона може бути використана для підготовки безлічі інтерактивних і друкованих звітів.

Для створення звіту на основі даних сховища необхідно підключитися до БД, як до джерела даних. Варіанти налаштування підключення показано на рис.21.

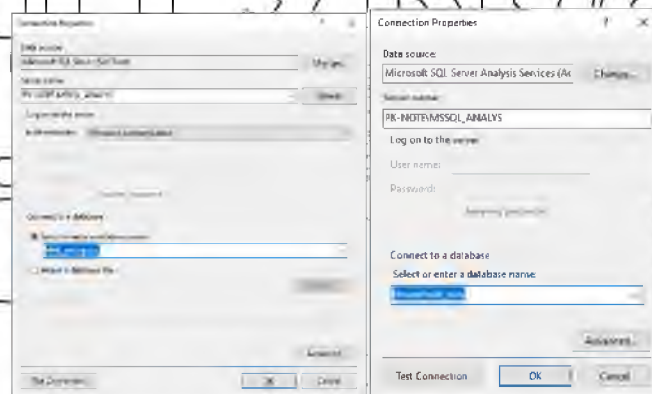


Рис. 21 Підключення до джерела даних

Відбір необхідних даних відбувається за допомогою команди SQL, а саме SELECT. Приклад такої команди для вибору необхідних для формування звіту

даних наведено на рис.22.

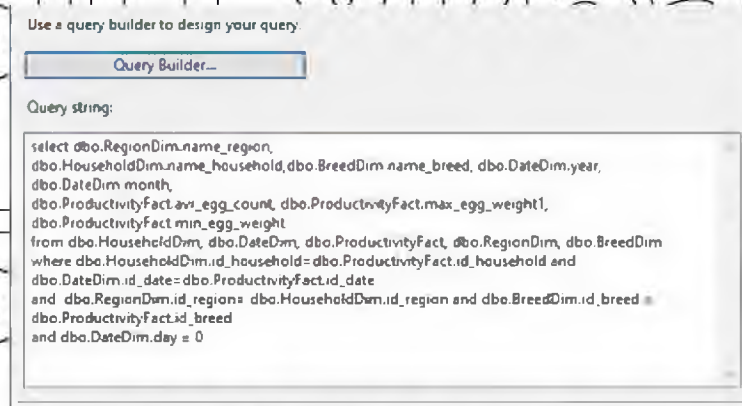


Рис. 22 Команда вибору даних

На наступному кроці налаштовуються поля звіту: рядки, стовпчики, сторінки (рис.23).

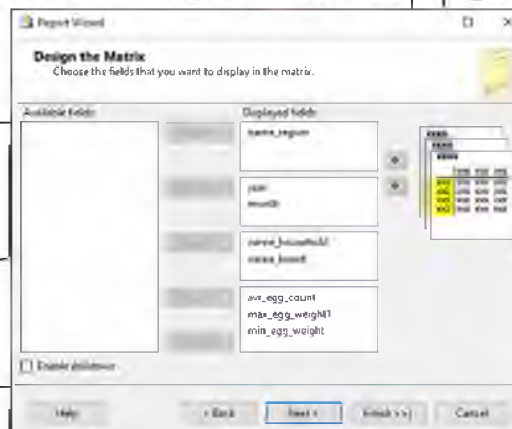


Рис. 23 Налаштування подання звіту

Подальше налаштування подання звіту відбувається через дизайнер звітів.

В цьому режимі є можливість обрати таке форматування звіту, щоб він був зрозумілим всім користувачам підсистеми аналізу, з різним рівнем обізнаності в предметній області.

3.8 Побудова звітності за допомогою EXCEL

Для побудови звіту на основі даних зі сховища необхідно встановити підключення до серверу зі створеного файлу Excel (рис.24). Потім треба обрати СД, та таблиці, де зберігаються необхідні для формування звіту дані (рис.25).

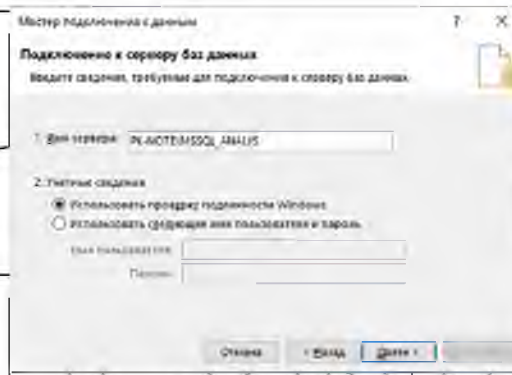


Рис. 24 Встановлення підключення

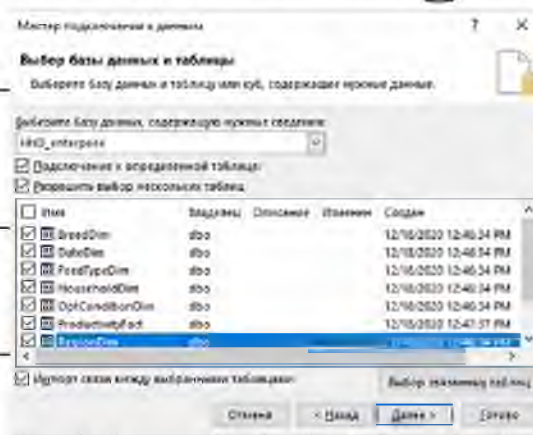


Рис. 25 Вибір необхідних таблиць

Після створення зведеної таблиці або діаграми користувач може налаштувати подання створеного об'єкта відповідно до цілі звіту (рис.26).

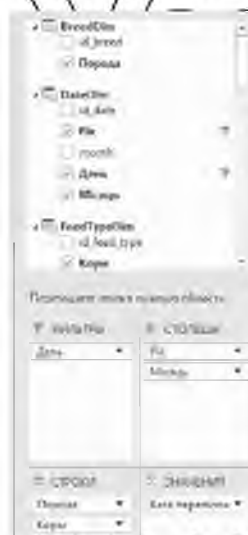


Рис. 26 Налаштування зведеної таблиці

3.9 Розрахунок KPI

KPI, ключовий індикатор ефективності, застосовується для обробки інформації про задану мету, записаного в кубі фактичного значення продуктивності і результати вимірювань, які показують тренд і стан продуктивності.

Далі наведено етапи розрахунку KPI для перепелиного господарства.

Простий об'єкт KPI складається з цілі, фактичного значення, тренду, стану.

Ціллю перепелиного господарства є кількість знесених перепілками яєць за рік.

Цільові значення виробництва наведені в табл. 4.

Цільові значення

Таблиця 4

Порода	Цільове значення продуктивності, к-сть яєць за рік
Сстонська	7000
Фараон	7000
Англійська біла	7000
Смокінгова	7000
Загалом	35000

Формування KPI, тобто визначення фактичного значення, встановлення цільового значення, формування статусу і тренду, у розрізі однієї породи наведено на рис.27-30.

* Value Expression
 (([Measures].[Avg Egg Count],[Breed Dim].[Id Breed].[1],[Date Dim].[Year].[2020]))

Рис. 27 Фактичне значення продуктивності

* Goal Expression

5000

Рис. 28 Цільове значення продуктивності

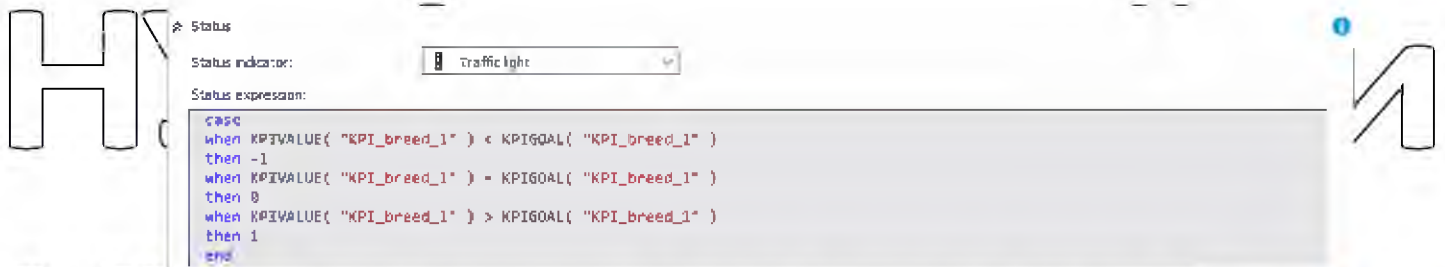


Рис. 29 Значення статусу



Рис. 30 Значення тренду

Батьківський ключовий індикатор – існуючий ключовий індикатор продуктивності, який використовує значення дочірнього індикатора при обчисленні батьківського. Формування батьківського KPI наведено на рис.31.

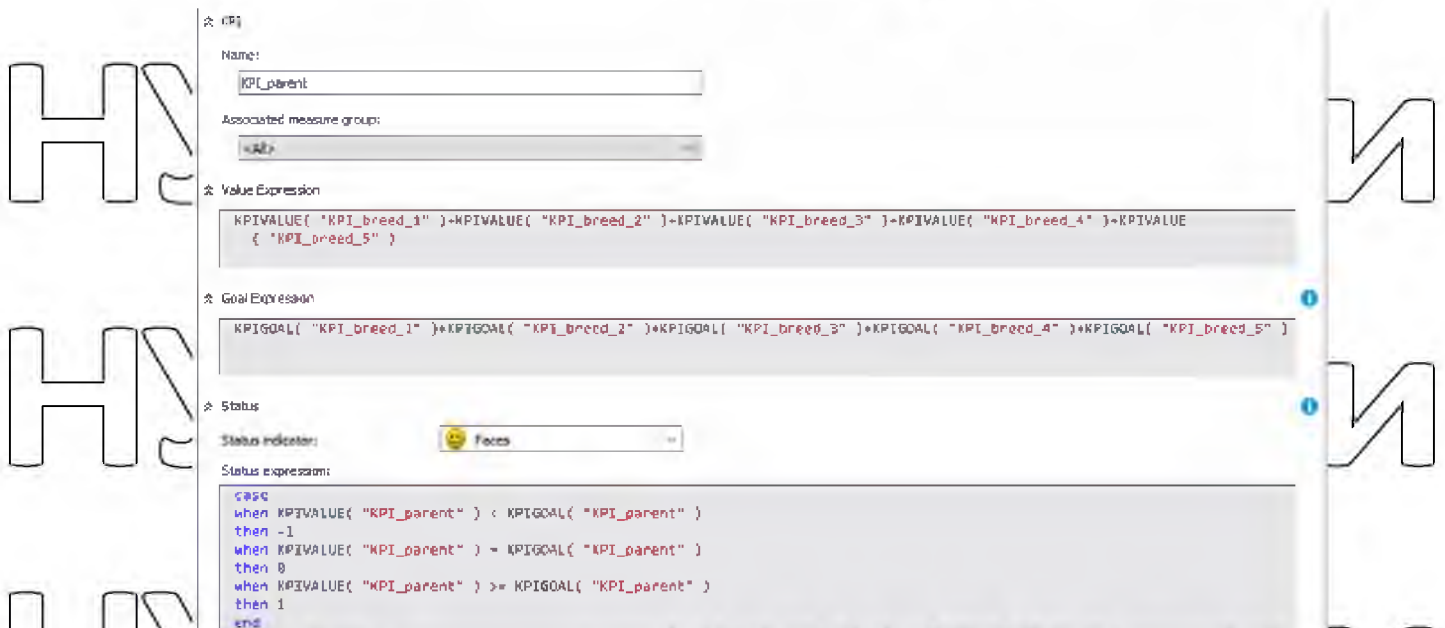


Рис. 31 Формування батьківського KPI

3.10 Проектування моделей інтелектуального аналізу, Data Mining

Microsoft SQL Server Business Intelligence – це сучасна платформа для вирішення задач бізнес-аналітики, яка надає компаніям можливість створювати

масштабовані BI рішення. Задачі, які вирішуються за допомогою Data Mining

описані в підрозділі 2.3. Microsoft BI пропонує вирішення таких задач Data Mining:

- часові ряди;
- дерева рішень;
- метод найвимогливішого Байеса;
- лінійна регресія;
- кластеризація;
- асоціативні правила.

Використання BI MS SQL Server показано на прикладі реалізації вирішення задачі часових рядів у середовищі.

Нижче наведено покроковий алгоритм побудови часового ряду, який буде показувати зміну середньої ваги перепелів в залежності від вмісту протеїну в кормі.

Додаємо в структуру сковища обчислювальну міру «Середня вага/кількість записів» (рис.15).

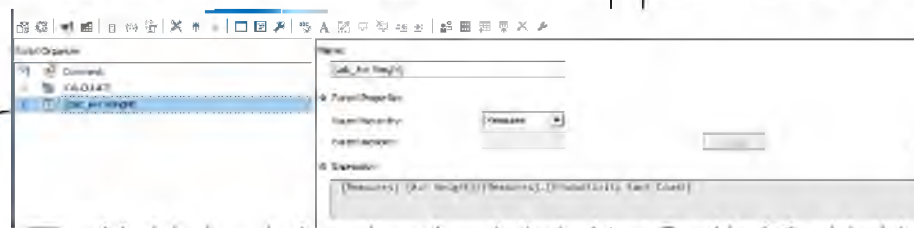


Рис. 32 Додавання обчислювальної міри

Створюємо структуру інтелектуального аналізу за допомогою Data Mining Wizard.

Спочатку необхідно обрати джерело, на основі якого буде будуватися структура інтелектуального аналізу. Wizard пропонує два варіанти: куб, який

сформований в проєкті, будь-яка інша реляційна БД або ОД. Для роботи обираємо куб, який сформований в проєкті.

Далі обираємо технологію інтелектуального аналізу – часові ряди (рис.17).

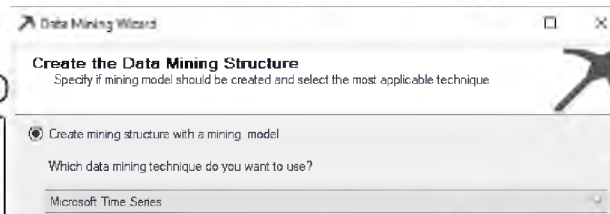


Рис. 33 Вибір технології аналізу даних

На наступному кроці обираємо таблицю вимірів, відповідно до яких буде показано та передбачено зміну фактичних показників. В задачі, яка вирішується, такою таблицею буде FeedTypeDim, що зберігає інформацію про тип корму (рис. 34). Далі необхідно обрати ключове поле з полів обраної таблиці.

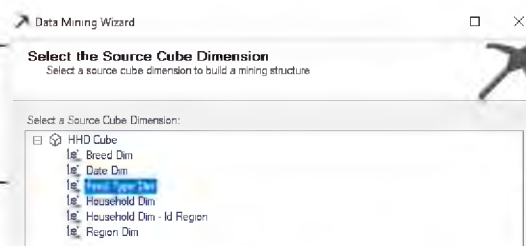


Рис. 34 Вибір таблиці вимірів

Ві пропонує обрати атрибути, які будуть відображатися в моделі інтелектуального аналізу окрім ключового. Оскільки умовсю задачі є знаходження залежності ваги перепілки від вмісту протеїну додаємо обчислювальну міру, яка позначає середню вагу перепелів (рис.35).

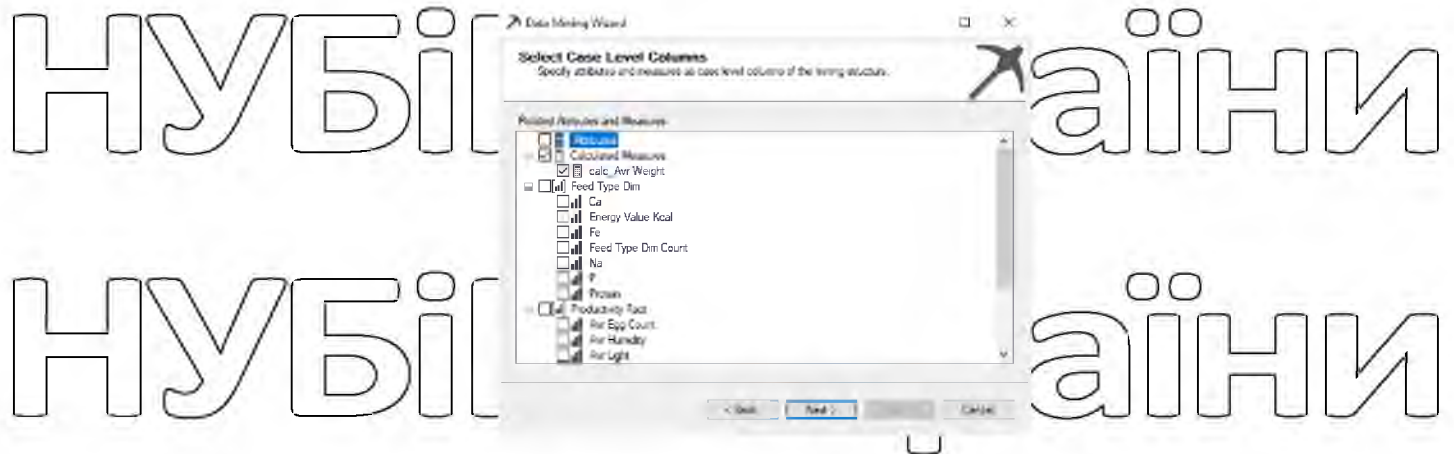


Рис. 35 Вибір атрибутів

Для вибору атрибутів необхідно визначити налаштування.

- Target – вхідна змінна, яка значною мірою впливає на перебіг процесу, який досліджується;
- Predict – змінна, значення якої будуть передбачуватися.

Також в цьому вікні можна додати додаткову таблицю вимірів (Add Nested Tables).

Ві надає можливість додаткових налаштувань моделі: фільтр даних, обмеження в кількості даних. В задачі, що розглядається, цей крок було пропущено за відсутності необхідності.

В результаті виконаних кроків формується модель аналізу даних. Її структура наведена на рис. 36

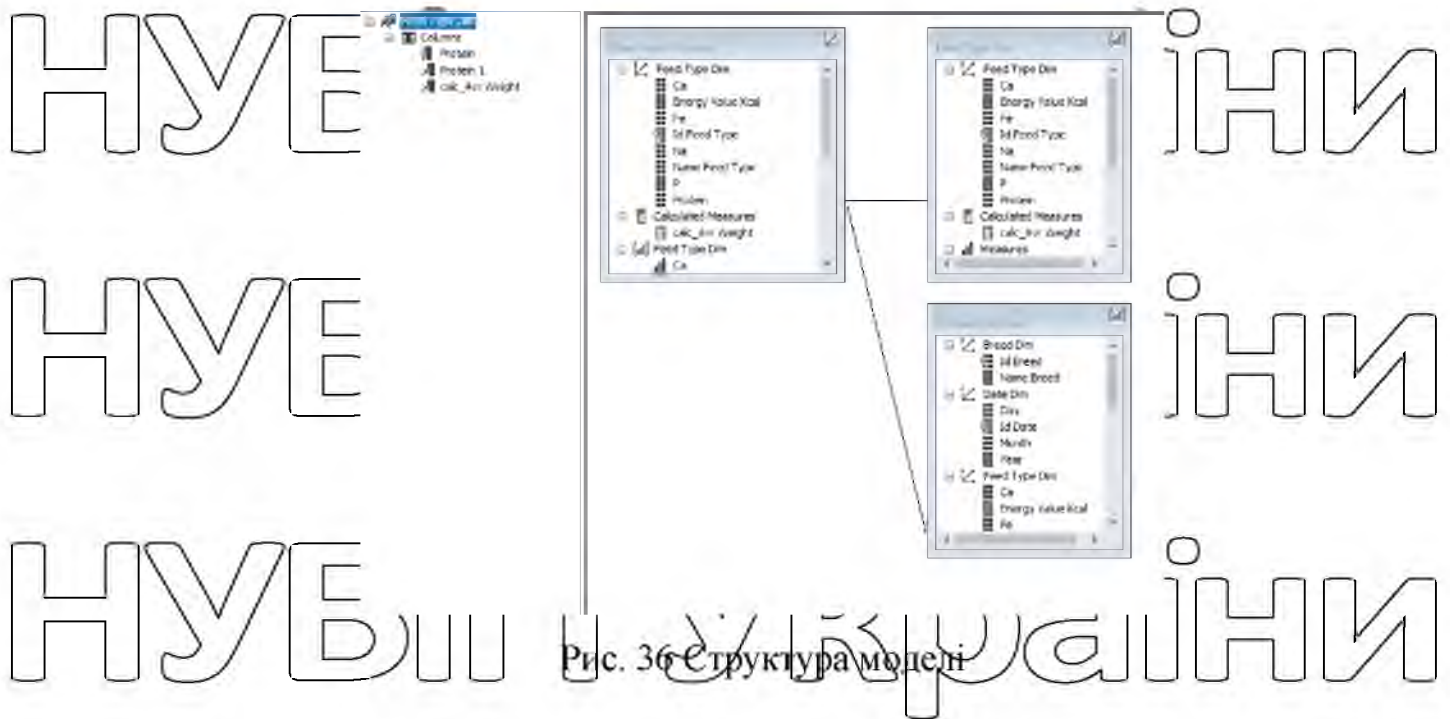


Рис. 36 Структура моделі

3.11 Розроблені алгоритми аналізу даних, Data Mining

Окрім реалізації аналізу даних інструментами SSIS, були розроблені окремі програмні модулі, які здійснювали інтелектуальний аналіз даних щодо об'єктів перепелиного господарства. Більш детально про задачі, які вирішуються за допомогою інтелектуального аналізу даних описано в підрозділі 2.3.

Було реалізовано рішення двох задач інтелектуального аналізу: класифікація, пошук асоціативних правил

3.11.1 Класифікація. Класифікація – розподіл об'єктів по визначених групах (класам) відповідно до їх ознак.

В результаті аналізу задачі було виділено два класи, відповідно до яких буде проведена класифікація:

- Клас Н – «Висока продуктивність»
- Клас L – «Низька продуктивність»

Для класифікації об'єктів необхідно визначити залежні та незалежні змінні.

Залежна змінна: клас.

Незалежні змінні: порода, корм.

Реалізовані методи класифікації: 1 rule, Naive Bayes.

1 rule (1R), 1-правило

Метод 1R найпростіший алгоритм класифікації об'єктів, адже класифікація здійснюється лише по одному параметру.

Алгоритм D-R

Визначення залежної (має бути дискретною величиною) та незалежної змінної (характеристика об'єкта);

1. Розраховуємо ймовірності потрапляння об'єктів з певної характеристикою (незалежна змінна) в певний клас (залежна змінна). Тобто, визначаємо, скільки разів зустрічається залежна змінна у об'єктів із вибраним значенням незалежної змінної;

2. Визначаємо клас, попадання в який об'єкта з певною характеристикою є більш ймовірним;

3. Обчислюємо помилку – помилкою є частота потрапляння об'єктів, які мають таку саму характеристику (значення незалежної змінної), до іншого класу відмінного від визначеного.

4. Відповідно до отриманих результатів формуємо правила: якщо «незалежна змінна», то «залежна змінна» [26].

Графічне зображення алгоритму представлено на рис.37.

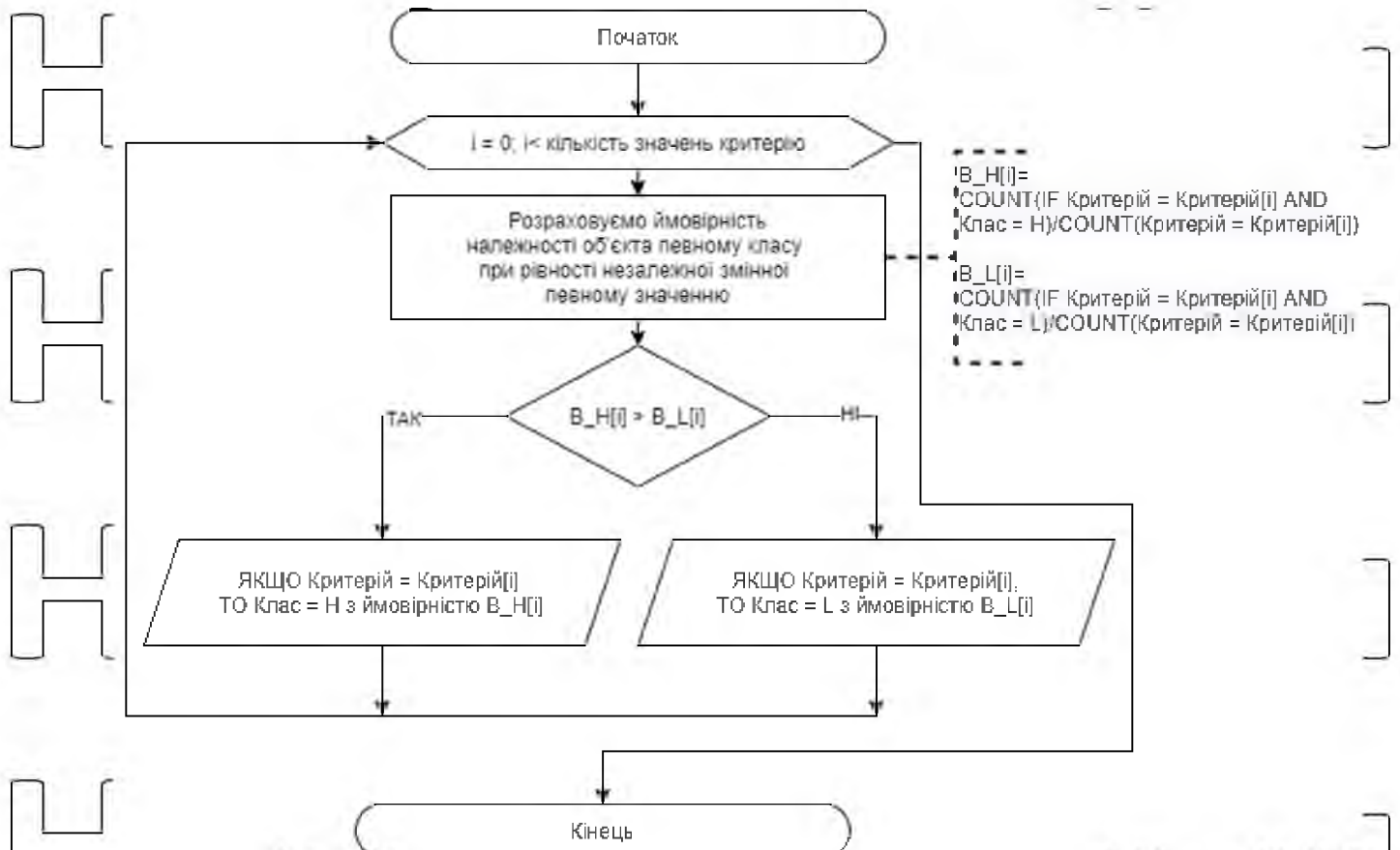


Рис. 30 Алгоритм 1 Rule

Оскільки було визначено дві незалежні змінні, відповідно до яких необхідно провести аналіз, алгоритм необхідно реалізувати окремо для кожної змінної.

Naive Bayes

Вищезгаданий метод класифікує об'єкти лише по одній ознаці, якщо ж необхідно провести класифікацію по декільком критеріям застосовують метод Naive Bayes. При цьому вважається, що всі змінні не мають залежності одна від іншої.

Ймовірність належності об'єкта відповідному класу розраховується за формулою (1).

$$P(c|x) = \frac{P(c)P(x)}{P(x)}, \quad (1)$$

де

$P(c|x)$ – умовна ймовірність класу c при настанні події x , тобто при рівності

незалежної змінної певному значенню;

$P(x|c)$ – ймовірність рівності незалежної змінної певному значенню, при належності об'єкта до класу c ;

$P(c)$ – ймовірність настання класу c ;

$P(x)$ – ймовірність рівності критерія певному значенню [27].

Трансформувавши формулу (1) маємо:

$$P(x) = \frac{P(c)P(c) \dots P(c)P(c)}{P(x)}, \quad (2)$$

де

$P(x_1, \dots, x_n | c)$ – ймовірність рівності критеріїв певним значенням при належності об'єкту класу c ;

n – кількість критеріїв.

Графічне зображення алгоритму представлено на рис.38.

3/11.2 Пошук асоціативних правил. Асоціативні правила – метод який допомагає показати ймовірність зв'язків між елементами даних у великих наборах даних у різних типах баз даних. Асоціативні правила створюються

шляхом пошуку в наборах даних шаблонів «якщо-то». Для перевірки знайдених

правил розраховують значення підтримки – частота появи певного шаблону в наборах даних [28].

Алгоритм пошуку асоціативних правил наведено на рис. 39.

НУБІП України

НУБІП України

НУБІП України

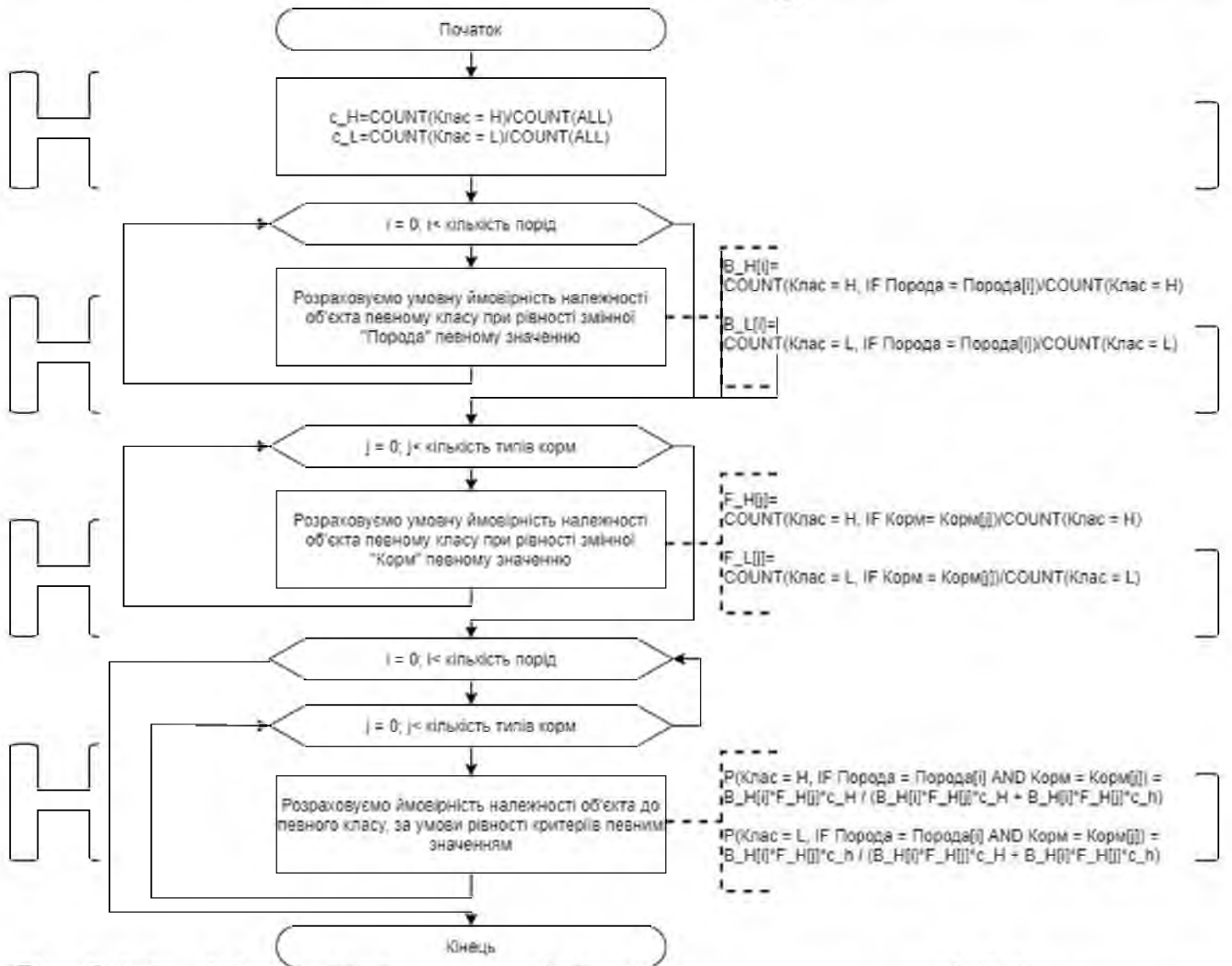


Рис. 38 Метод Naive Bayes

НУБІП України

НУБІП України

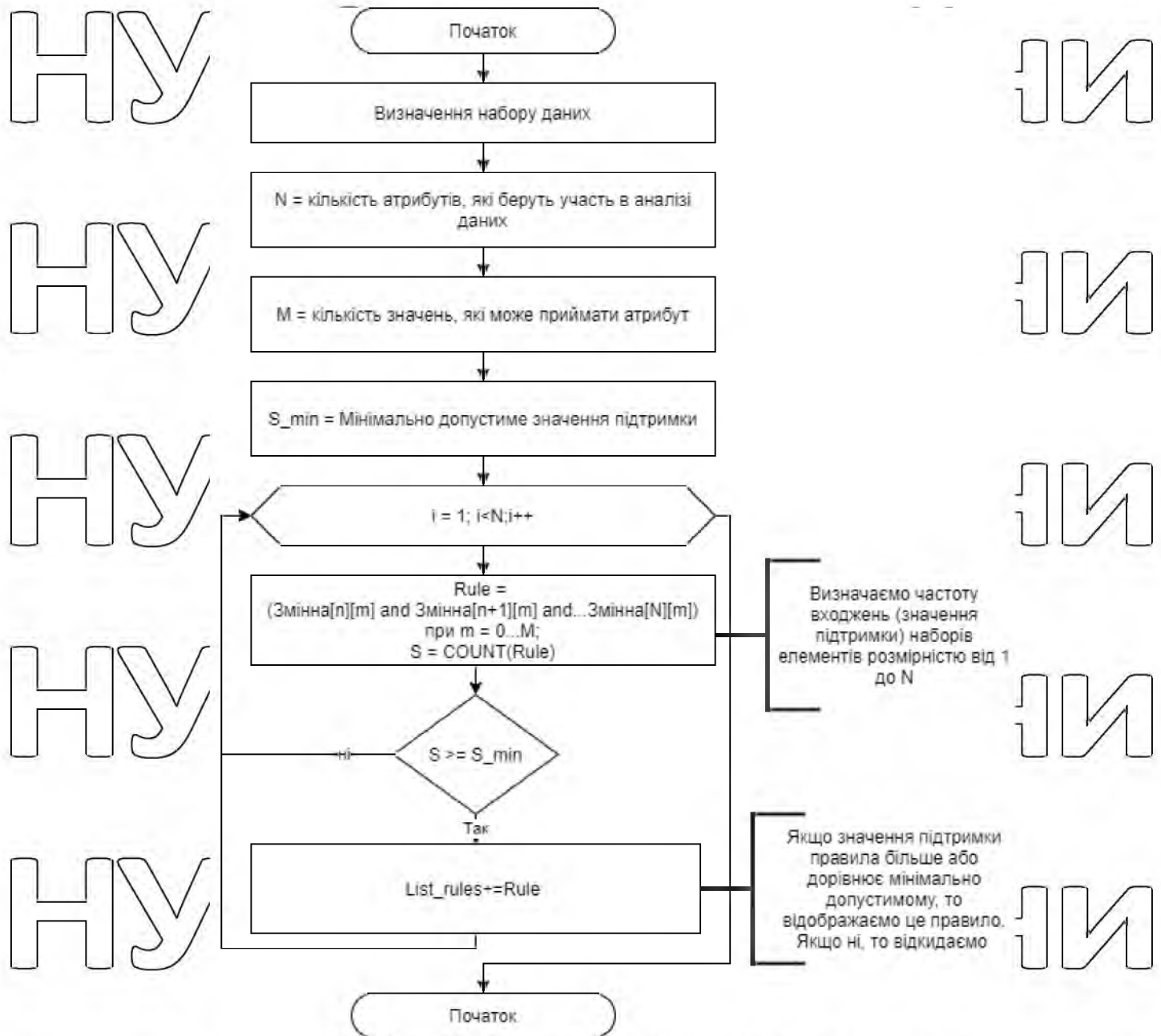


Рис. 39 Пошук асоціативних правил

Метод пошуку асоціативних правил застосовують в:

- роздрібній торгівлі, інтернет торгівлі – визначення набору товарів які покупці найчастіше купують разом, для формування рекомендацій, вибору розміщення полиць в магазинах, тощо. Наприклад: «Якщо покупець купив хліб, то купити і масло»
- медицині – визначення хвороб, за наборами симптомів; визначення

хвороб, які частіше зустрічаються разом. Наприклад «Якщо втрачено смак та нюх, висока температура, кашель, то у хворого коронавірус»

● стрімінгових платформах – визначення списку переглядів серед користувач, для формування рекомендацій для нових користувачів. Наприклад:

«Якщо користувач подивився фільм «The Social Network», то зацікавиться фільмом «Jobs»».

В розрізі задачі, що розглядається, за допомогою пошуку асоціативних правил необхідно знайти такі набори «Порода» і «Корм», які сприяють продуктивності,

як більше за середнє значення.

Код, який реалізує вищевказані алгоритми наведений в Додатку Д.

НУБІП Україні

НУБІП Україні

НУБІП Україні

НУБІП Україні

РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

4.1 Розведення перепелів

4.1.1 Інтелектуальний аналіз даних. Для аналізу даних щодо роботи

перепелиного господарства був розроблений окремий програмний додаток. В розрізі якого було реалізовано декілька методів інтелектуального аналізу.

Результати класифікації

1 rule. Метод реалізує класифікацію за одним параметром.

Класи, відповідно до яких буда проведена класифікація:

- клас «Висока продуктивність» – кількість знесених яєць більше за середнє значення за весь період спостереження;

- клас «Низька продуктивність» – кількість знесених яєць менше за середнє значення за весь період спостереження.

Залежна змінна: клас.

Незалежні змінні: порода, корм.

На рис.40 представлено результати класифікацію за параметром «Порода».

Аналіз продуктивності

Середнє значення продуктивності з "1 січень, 2020" по "31 травень, 2020" по всім породам перепілок та господарствам дорівнює: 10,547,724
 Дані поділені на два класи:
 -H – висока продуктивність;
 -L – низька продуктивність.

Продуктивність: Порода Корм Докриптеравний аналіз Асоціативна графіка Звіт

Порода	Місяць середнього	Всього середнього	Всього	Клас	Ймовірність
Співочий	303	424	727	H	59.32
Чарок	387	345	732	L	52.87
Англіська біла	473	259	732	L	64.62
Смарагдова	295	437	732	H	59.70

Рис. 40 Результат методу ІР по параметру "Порода"

Висновки

За результатами аналізу були знайдені декілька правил. Правила з найбільшою ймовірністю:

- якщо порода «Смокінгова», то клас «Висока продуктивність» – з ймовірністю ~60%;

- якщо порода «Англійська біла», то клас «Низька продуктивність» – з ймовірністю ~65%.

На рис.41 представлено результати класифікації за параметром «Корм».

Корм	Месяць середнього	Обсяг середнього	Всього	Клас	Ймовірність
Пурна	367	370	727	H	50.89
Актив	317	415	732	H	56.69
Фогель	389	343	732	L	53.14
Супер	395	337	732	L	53.96

Рис. 41 Результат методу ІК по параметру "Корм"

Класифікація за кормом є менш точною, ніж за породою, оскільки правило з найбільшою точністю «якщо корм «Актив», то клас «Висок продуктивність» має ймовірність <math><60\%</math>.

Naive Bayes. Метод реалізує класифікацію за декількома параметрами.

Порода	Корм	Ймовірність H	Ймовірність L
Англійська біла	Пурна	35.177243	63.822756
Англійська біла	Фогель	32.482123	67.517876
Естонська	Пурна	59.044550	40.955449
Естонська	Актив	64.547679	35.452326
Смокінгова	Фогель	58.505149	41.494850
Смокінгова	Супер	55.714679	44.285320
Фараон	Супер	43.060147	56.939852
Фараон	Актив	53.747420	46.252579

Рис. 42 Результати двох-критеріального аналізу

Висновки

За результатами аналізу були визначені правила з найбільшою ймовірністю (>60%):

- якщо порода «Англійська біла» і корм «Фогель», то «Низька продуктивність» з ймовірністю ~68%.

- Якщо порода «Єстонська» і корм «Актив», то «Висока продуктивність» з ймовірністю ~65%.

- якщо порода «Англійська біла» і корм «Пуріна», то «Низька продуктивність» з ймовірністю ~64%.

Висновки за результати реалізації двох методів

З отриманих результатів можна зробити такі висновки:

- 1) більший вплив на продуктивність має саме критерій «Порода»;
- 2) порода «Англійська біла» при вживанні кормів «Пуріна» і «Фогель» дає низьку продуктивність – що є приводом зміни кормів на інші, або повної відмови від вирощування перепелів цієї породи;
- 3) при вживанні корму «Актив» перепела породи «Фараон» дають в більшості випадків «Високу продуктивність», на відміну від перепелів цієї породи, які вживають корм «Супер» – що є приводом відмовитися від корму «Супер» в об'єктах, які вирощують перепелів породи «Фараон», та замінити його на «Актив».

Пошук асоціативних правил.

На рис. 43 зображено результат реалізації алгоритму пошуку асоціативних правил.

Найбільше значення підтримки, отже найбільшу кількість разів зустрічаються в наборах даних, мають правила:

- якщо порода «Англійська біла» і корм «Пуріна» то продуктивність є

меншою за середнє значення;
 • якщо порода «Смокінгова» і корм «Фогель» то продуктивність є більшою за середнє значення;

- якщо порода «Фараон» і корм «Супер» то продуктивність є меншою

за середнє значення
Висновки
 Результати реалізації пошуку асоціативних правил підтвердили отримані результати класифікації та доповнюють їх:

- перепели породи «Смокінгова» частіше продукують більше за середнє значення при вживанні корму «Фогель» – тож варто розглянути рішення, щодо закупівлі цього корму для всіх об'єктів, які вирощують перепелів породи «Смокінгова»;

- для вирощування перепелів породи «Естонська» більш вигідною є годівля кормом «Пуріна» – адже при вживанні цього корму перепела дають більшу за середнє значення продуктивність.

Аналіз продуктивності

Середнє значення продуктивності з "1 січня 2020" по "31 грудень 2020" по всім породам перепілок та господарствам дорівнює: 10,547724.
 Дані поділені на два класи:
 -Н - висока продуктивність,
 -L - низька продуктивність.

Продуктивність | Порода | Корм | Двох-критеріальний аналіз | Асоціативні правила | Звіт

Розмірність вибірки: 2923 | Параметри лів: 1495 | Параметри прав: 785

Кількість елементів в небі: | Параметри лів: 196

Параметри	Розмірність набору	Правило
259	3	Порода = Ангійська Біла, Корм = Фогель, Середня кількість яєць < 10,547724
237	3	Порода = Смокінгова, Корм = Фогель, Середня кількість яєць >= 10,547724
228	3	Порода = Фараон, Корм = Супер, Середня кількість яєць < 10,547724
217	3	Порода = Естонська, Корм = Пуріна, Середня кількість яєць >= 10,547724
213	3	Порода = Ангійська Біла, Корм = Пуріна, Середня кількість яєць < 10,547724
208	3	Порода = Фараон, Корм = Аліна, Середня кількість яєць >= 10,547724
207	3	Порода = Естонська, Корм = Аліна, Середня кількість яєць >= 10,547724

Рис. 43 Пошук асоціативних правил

4.1.2 Оперативний аналіз даних

Результат формування звітності, BI

НУБІП України

В результаті роботи з SSRS було сформовано декілька звітів.
Звіт, щодо продуктивності перепелів в розрізі породи та норму
представлений на рис. 44.

Продуктивність перепілок за 2020 рік

Середня кількість яєць, шт	Корм Актив	Корм Пуріна	Корм Супер	Корм Фогель
Порода Англійська біла		9.57		8.83
Порода Естонська	11.02	11.28		
Порода Смокінгова			10.96	11.96
Порода Фараон	10.24		10.54	



Рис. 44 Продуктивність перепелів

Висновки

Звіт підтверджує знайдені в процесі інтелектуального аналізу закономірності відносно того, що:

- перепела породи «Англійська біла» мають низьку продуктивність;
- найкраще значення продуктивності у перепелів породи «Смокінгова»

при вживанні корму «Фогель»;

- порода «Фараон» має трохи кращі показники продуктивності при вживанні корму «Супер», в порівнянні з кормом «Актив».

Звіт, щодо продуктивності перепелів, за різних умов утримання на рис.45.

НУБІП України

Продуктивність перепілок
в області: Житомирська

Господарство ТОВ "Веселувага"

Рік	Місяць	Порода Єстонська			Продуктивність Кількість яєць	Порода Фарсан			Продуктивність Кількість яєць
		Умови утримання				Умови утримання			
		Температура, С°	Вологість, %	Освітлення, годин/день		Температура, С°	Вологість, %	Освітлення, годин/день	
2020	січень	22,12	67,03	16,97	10,74	22,51	66,39	16,81	11,23
	лютий	22,03	66,78	16,76	10,21	22,17	67,90	16,82	10,14
	березень	22,56	66,39	17,06	11,10	22,57	67,13	17,13	9,94
	квітень	22,66	66,93	17,10	10,90	22,37	67,07	17,17	9,60
	травень	22,38	66,23	17,16	10,42	22,59	66,94	16,90	10,77
	червень	22,08	66,37	17,47	10,97	22,28	67,07	16,97	10,93
	липень	22,25	67,16	16,61	10,74	22,45	66,42	17,29	10,32
	серпень	22,50	66,13	17,10	9,84	22,03	67,61	17,35	10,18
	вересень	22,61	66,57	17,37	11,37	21,97	67,03	17,00	10,03
	жовтень	22,34	67,65	16,90	10,10	22,59	67,37	16,90	11,00
	листопад	22,57	67,03	17,00	10,35	22,72	67,10	16,47	10,70
	грудень	22,35	66,19	17,26	10,48	22,94	67,35	17,03	10,48

Рис. 45 Продуктивність перепелів за різних умов утримання

Висновки

Результатом аналізу цього звіту є визначення оптимальних умов утримання перепелів для збільшення їх продуктивності. Такий звіт буде корисним для директора об'єкта, для та оперативного керування параметрами роботи підпорядкованого йому пташника(пташників).

Результат формування звітності, Excel

За допомогою офісного додатку Excel були побудовані звіти для аналізу набору маси перепелів.

Звіт, що демонструє середні значення маси в залежності від різних типів корму представлений на рис. 46-47.

Середнє значення ваги перепілок		Роки													
		2020													
Порода		січень	лютий	березень	квітень	травень	червень	липень	серпень	вересень	жовтень	листопад	грудень	2020 Итого	Общий итог
Автоматична біла		116,63	116,45	115,69	115,37	116,45	116,77	116,69	115,57	115,29	116,60	116,55	116,33	115,84	115,84
Пуріна Фогель		115,86	115,95	116,00	115,32	115,63	116,40	116,97	116,10	116,32	116,46	116,00	116,73	115,89	115,89
Супер Фогель		116,40	116,97	115,91	115,42	117,26	116,18	116,09	116,05	116,24	116,78	115,11	116,19	116,79	116,79
Автоматична біла		116,87	116,93	116,11	115,39	117,13	116,50	114,78	116,25	116,21	116,01	115,11	116,42	115,81	115,81
Пуріна		114,39	116,62	116,72	116,03	116,43	114,92	114,90	116,64	116,81	116,23	116,06	116,05	116,73	116,73
Пуріна		116,45	116,26	115,52	114,75	116,79	116,08	114,65	116,03	116,59	116,44	115,31	116,80	115,87	116,87
Супер Фогель		115,46	116,00	115,49	115,89	115,73	116,39	115,20	115,61	116,19	115,87	116,19	115,90	115,90	
Супер Фогель		116,12	115,83	115,28	116,15	115,26	115,33	115,48	116,20	115,82	116,10	116,07	116,76	115,87	115,87
Супер Фогель		115,61	116,17	115,67	115,17	116,26	116,65	115,10	115,03	115,44	116,27	115,66	115,60	115,72	115,72
Автоматична біла		116,26	116,63	116,60	116,31	116,33	116,26	116,40	115,75	116,01	116,01	116,40	116,33	116,33	
Пуріна		116,39	117,26	116,66	116,22	116,26	116,74	117,21	116,04	116,85	116,31	116,33	116,38	116,39	116,39
Супер Фогель		116,73	116,14	116,45	116,33	116,95	116,05	116,60	115,45	115,30	116,40	116,33	115,40	115,84	115,84
Общій итог		115,80	116,27	116,01	115,68	116,38	116,76	115,63	115,59	115,80	116,03	115,63	116,24	115,89	115,89

Рис. 46 Звіт по середній вазі особин різних порід з різним кормом, таблиця



Рис. 47 Звіт по середній вазі особин різних порід з різним кормом, графік

Висновки

З даного звіту можна зробити висновок про оптимальний корм для особин кожної породи для набору маси. Як ми бачимо результати годування різними кормами мають досить малу розбіжність в результатах. Але чітко видно, що для набору маси у породи «Смокінгова» корм «Супер» є більш ефективним.

Результат розрахунку KPI

Результат обчислення показників ефективності зображено на рис.48.

Display Structure	Value	Goal	Status
KPI_parent	30831	28000	😊
KPI_breed_1	8105	7000	😊
KPI_breed_2	7603	7000	😊
KPI_breed_3	6734	7000	😊
KPI_breed_4	8389	7000	😊

Рис. 48 Результати обчислення KPI

Висновки

З результатів можна зробити висновок, що сумарно фактична продуктивність по всім породам перевищила цільове значення. Значення продуктивності порід «Смокінгова», «Фараон» та «Естонська» перевищили цільові, але показники породи «Англійська біла» менші за заплановані, що

підтверджує результати інтелектуального аналізу даних.

По результатам аналізу пропонується декілька рекомендацій:

1) розглянути зміну цільових значень продуктивності для порід: для

порід «Смокінгова», «Фараон» та «Естонська» у сторону збільшення, для

«Англійська біла» у сторону зменшення.

2) розглянути варіант відмови від вирощування перепелів мало

продуктивної породи – «Англійська біла», або більш глибоко проаналізувати

роботу підприємства, щоб знайти причину маленької продуктивності перепелів

цієї породи.

4.2 Вирощування с/г культур

4.2.1 Інтелектуальний аналіз даних. За допомогою SSAS були

спроєктовані моделі Data Mining для аналізу роботи щодо вирощування с/г

культур.

Результат пошуку асоціативних в розрізі кліматичної зони, культури та

врожайності показано на рис.49.

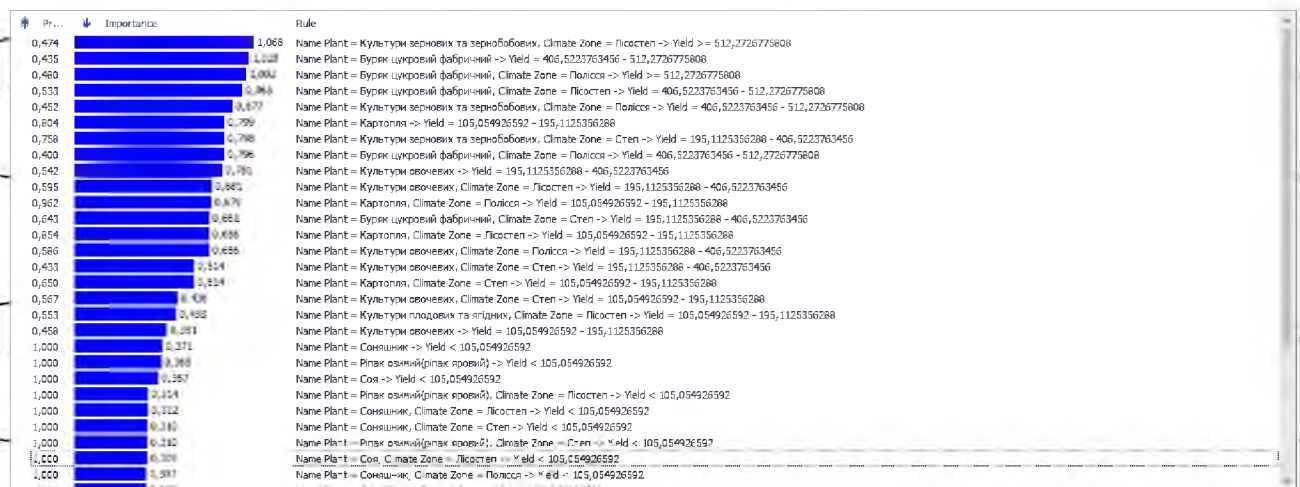


Рис. 49 Пошук асоціативних правил, врожайність

Висновки

З отриманих результатів можна зробити висновок, щодо того в якій кліматичній зоні яка культура дає більшу врожайність.

Наприклад «культури зернових та зернобобових»:

- в кліматичній зоні «Лісостеп» дають врожайність > 512,27 ц/га;
- в кліматичній зоні «Подієся» дають врожайність в межах 406,52 - 512,27 ц/га;

- в кліматичній зоні Степ дають врожайність в межах 195,11 - 406,52 ц/га.

Таку модель аналізу можна перелаштувати на аналіз щодо виробництва культур (рис.50)

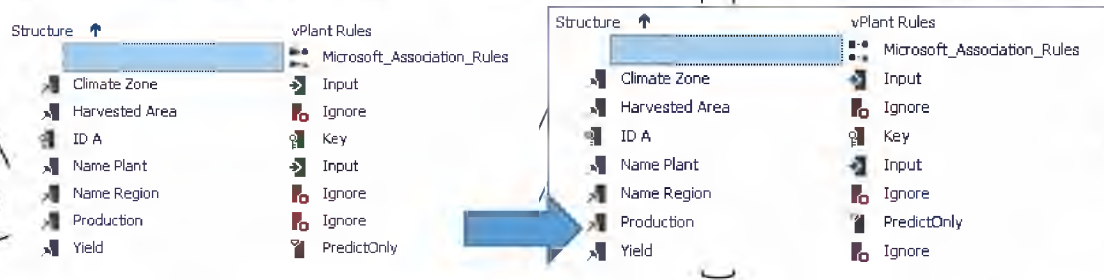


Рис. 50 Налаштування моделі

На рис. 51 показано результат пошуку асоціативних правил в розрізі кліматичної зони, культури та виробництва культури.

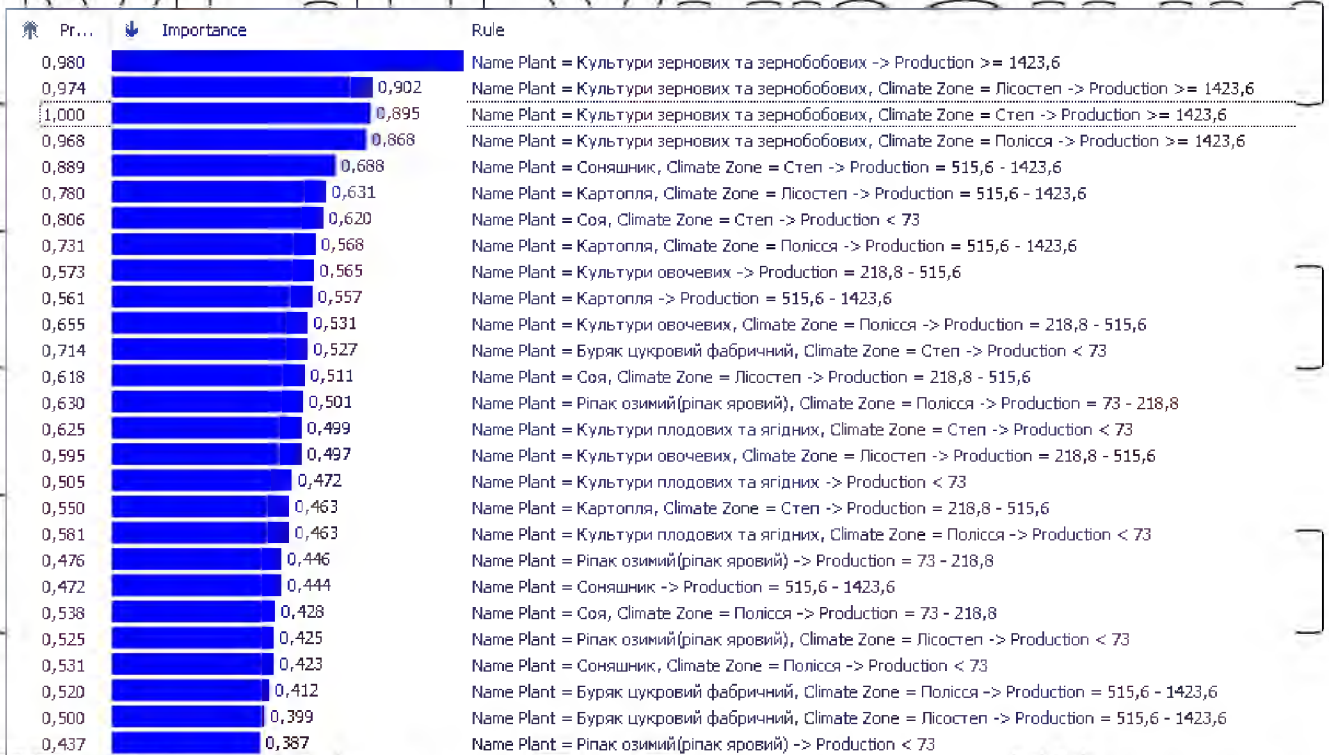


Рис. 51 Пошук асоціативних правил, виробництво

Висновки

За результатами можна визначити культуру, яка дає найбільший валовий збір в рік в кліматичній зоні. Наприклад: культури зернових та зернобобових в кліматичній зоні Лісостеп дають валовий збір >= 1423,6 тис.т.

4.2.2 Оперативний аналіз даних. Були сформовані декілька звітів для оперативного аналізу даних.

На рис. 52 показано графік врожайності в розрізі культур та кліматичних зон.

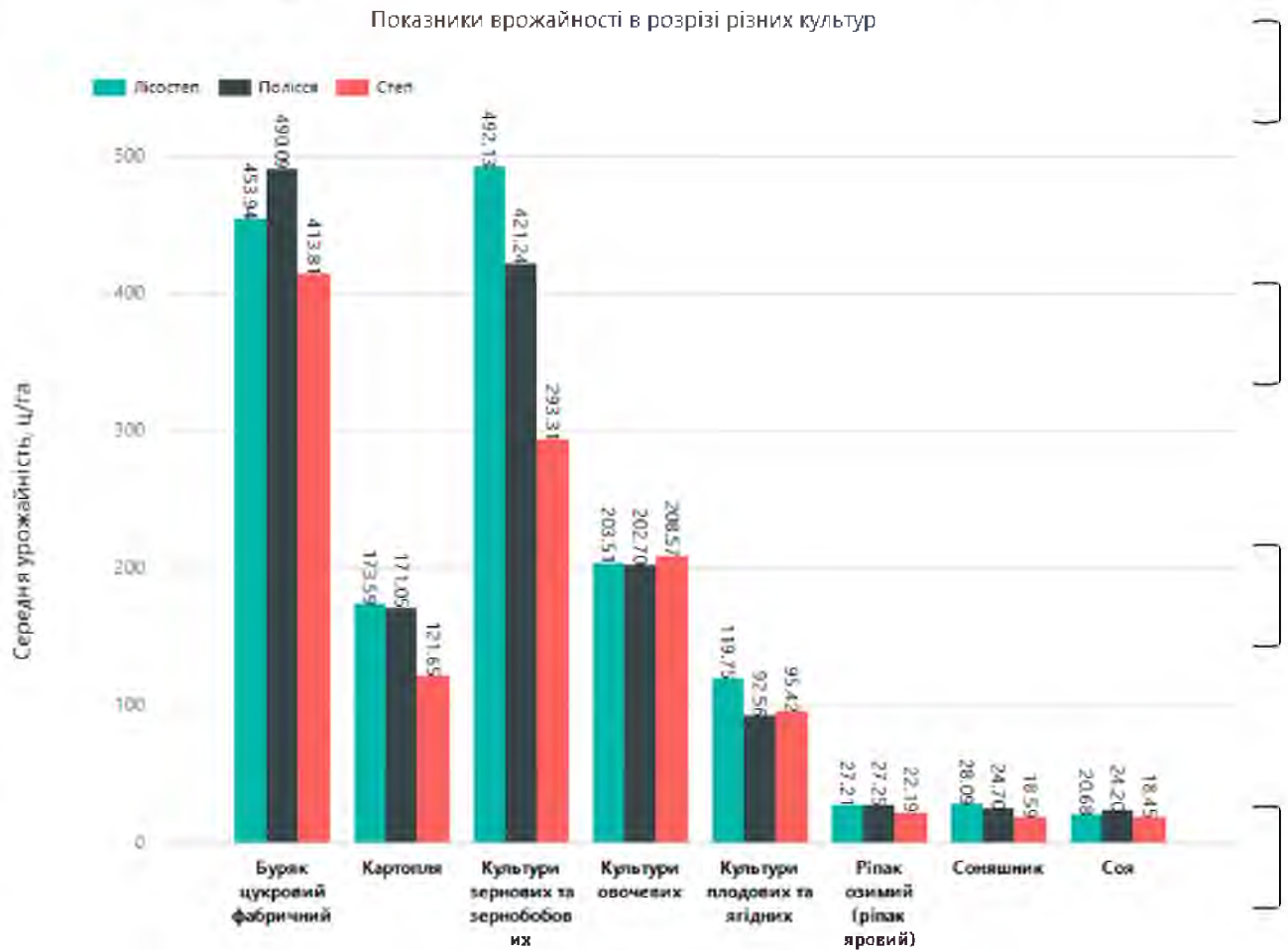


Рис. 52 Врожайність культур

Висновки

Отриманий графік підтверджує знайдені інтелектуальним аналізом закономірності щодо врожайності культур в розрізі кліматичних зон.

- найбільша врожайність зернових культур відзначається в кліматичній зоні «Лісостеп»;
- найбільша врожайність цукрового буряку спостерігається в Поліссі;
- врожайність картоплі в Лісостепу та Поліссі майже однакова, та значно менша в зоні Степу.

Також були створені звіти по інших параметрах роботи господарства. На рис.53 зображено фрагмент графіку, який ілюструє частку посівної площі в розрізі культур та кліматичних зон.

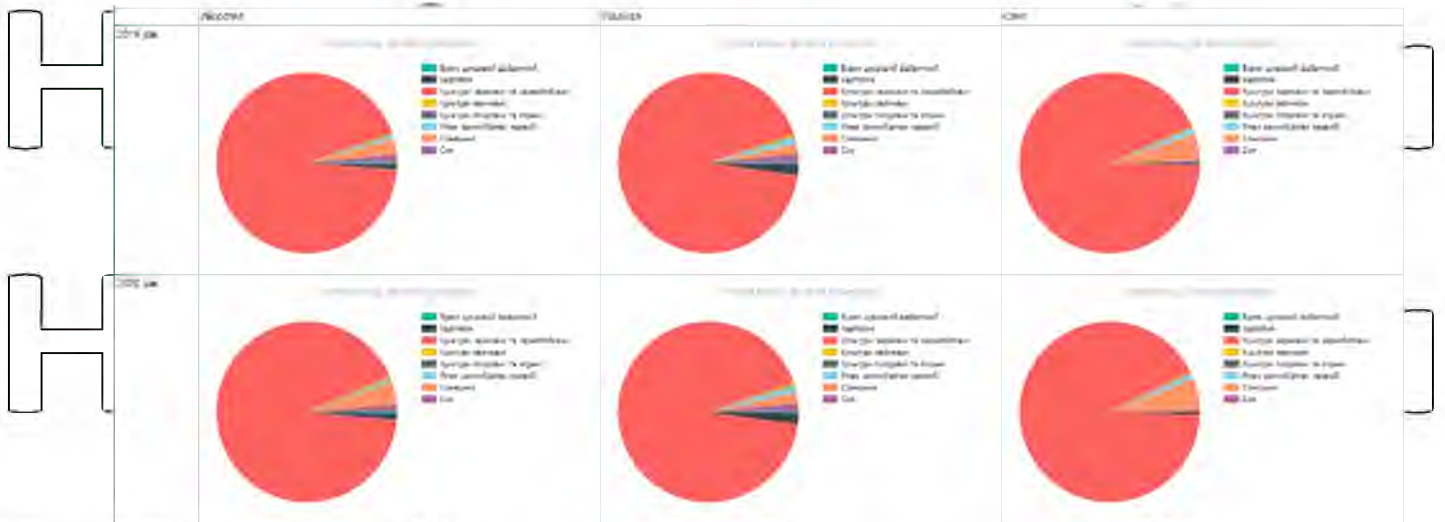


Рис. 53 Частка посівної площі

Висновки

З продемонстрованих графіків чітко видно, що найбільшу частку площі займають насадження зернових культур. Інші культури сумарно займають значно меншу площу.

В лісостепу та степу значні площі займають посіви соняшника та сої. А в Поліссі лідируючі позиції займають плодові та ягідні культури.

НУБІП України

НУБІП України

НУБІП України

ВИСНОВКИ

В ході роботи з метою створення ССПР при управлінні діяльністю с/г об'єкта були розроблені підсистеми введення, збереження інформації та аналізу.

Запропоновано використання OLAP та Data Mining технологій для аналізу великого обсягу даних.

На першому етапі було проведено аналіз предметної області, тобто процесу керування с/г об'єктом, формування технічного завдання, були розглянуті наявні наукові праці щодо предметної області. Був розроблений ряд діаграм, які описують предметну область, а саме діаграму прецедентів, послідовності та активності.

Під час роботи над другим розділом були визначені ключові елементи ССПР, розглянуті інструменти, методи, алгоритми аналізу даних, а саме інтелектуальний аналіз даних (Data Mining) та оперативний аналіз даних (OLAP).

На етапі розробки системи була спроектована та створена оперативна БД об'єкту перепелиного господарства та розроблений програмний модуль, за допомогою якого оператор об'єкта має можливість вносити дані. Для збереження інформації щодо роботи всього підприємства розроблено сховище даних, а саме дві вітрини:

- для збереження даних, щодо розведення перепелів;
- збереження даних, щодо вирощування с/г культур.

За допомогою інструментів SSAS було сформовано куб даних, організовано потоки даних від оперативних джерел інформації до сховища даних, сформовані моделі інтелектуального аналізу даних, створені звіти, щодо роботи підприємства та розраховано KPI по продуктивності передільних господарств.

Було досліджено використання декількох інструментів для формування звітності, а саме Visual Studio BI та Excel. Результати роботи з Visual Studio BI виявилися більш структурованими, інформативними та зрозумілими для

користувачів, з різним рівнем обізнаності в предметній області.

Окремо для аналізу роботи перепелиного господарства були спроектовані та реалізовані алгоритми класифікації та пошуку асоціативних правил.

В розрізі четвертого розділу було проаналізовано отримані результати оперативного та інтелектуального аналізу. На основі цих результатів були сформовані декілька рекомендацій щодо управління с/г підприємством.

Щодо роботи об'єктів по вирощуванню перепелів були знайдені такі закономірності та сформовані відповідні рекомендації, щодо регулювання роботи підприємства:

- порода «Англійська біла» при вживанні кормів «Пуріна» і «Фогель» дає низьку продуктивність – що є приводом зміни кормів на інші, або повної відмови від вирощування перепелів цієї породи;

- при вживанні корму «Актив» перепела породи «Фараон» дають в більшості випадків «Високу продуктивність», на відміну від перепелів цієї породи, які вживають корм «Супер» – що є приводом відмовитися від корму «Супер» в об'єктах, які вирощують перепелів породи «Фараон», та замінити його на «Актив».

Щодо вирощування с/г культур були знайдені такі закономірності:

- найбільша врожайність зернових культур відзначається в кліматичній зоні «Лісостеп» – що робить зону «Лісостеп» пріоритетною при виборі територій посіву зернових культур;

- найбільша врожайність цукрового буряку спостерігається в Поліссі – що робить зону «Полісся» пріоритетною при виборі територій посіву цукрового буряку;

- врожайність картоплі в Лісостепу та Поліссі майже однакова, та значно менша в зоні Степу – тож, можна зробити висновок, що кліматична зона мало впливає на продуктивність картоплі.

Знайдені залежності та закономірності в даних є підставою для формування стратегічних управлінських рішень, щодо регулювання параметрів роботи об'єктів та їх варто взяти до уваги при створенні нових об'єктів. Такі рішення призведуть до підвищення продуктивності господарств, що в свою чергу призведе до підвищення прибутковості всього підприємства.

Відповідно до отриманих результатів можна зробити висновок про доцільність використання технологій Data Mining та OLAP при формуванні управлінських рішень щодо керування об'єктами с/г господарства.

НУБІП України

НУБІП України

НУБІП України

НУБІП України

НУБІП України

Список використаних джерел

1. Єгорова О. В. ДИНАМІКА КІЛЬКОСТІ АГРАРНИХ ПІДПРИЄМСТВ УКРАЇНИ / О. В. Єгорова, А. П. Дорошенко, Є. О.

Кончаковський. // III Всеукраїнська науково-практична інтернет-конференція «Актуальні проблеми та перспективи розвитку обліку, аналізу та контролю в соціально-орієнтованій системі управління підприємством». – 2020. – С. 385–387.

(тези доступні за посиланням -

<http://dspace.pdaa.edu.ua:8080/handle/123456789/7808>)

2. Виробництво та розподіл валового внутрішнього продукту за видами економічної діяльності. Інші податки за виключенням інших субсидій пов'язаних з виробництвом (у фактичних цінах; млн грн) [Електронний ресурс]. – 2021. –

Режим доступу до ресурсу: <https://data.gov.ua/dataset/c0f1189-141b-4492-a359-32b941384030>.

3. Сільське господарство [Електронний ресурс] // НАБУ – Режим доступу до ресурсу: <https://nabu.ua/ua/silске-gospodarstvo-2.html>.

4. Проект Закону України від 08.04.2011 № 8370, «Про основи продовольчої безпеки України» [Електронний ресурс]. – 2011. – Режим доступу до ресурсу: <https://ips.ligazakon.net/document/JF6GI00A>.

5. Кількість активних підприємств за регіонами України та видами економічної діяльності / Державна служба статистики України [Електронний ресурс]. – 2021. – Режим доступу до ресурсу: <http://www.ukrstat.gov.ua>

6. Бидеев Б. А. ПРОДУКТИВНОСТЬ И БИОЛОГИЧЕСКИЕ ОСОБЕННОСТИ ПЕРЕПЕЛОВ РАЗНЫХ ПОРОД: дис. канд. с.-г. наук: 06.02.10 / Бидеев Бексолтан Александрович – Владикавказ, 2016. – 120 с. (робота доступна за посиланням - <https://cutt.ly/FRBAn7m>)

7. Методологічні пояснення [Електронний ресурс] // Ukrstat.org - публікація документів Державної Служби Статистики України – Режим доступу

до ресурсу: https://ukrstat.org/uk/operativ/operativ2005/sg/sg_rik/sg_u/sg_met.html.

8. Виробництво культур сільськогосподарських за регіонами [Електронний ресурс]. – 2021. – Режим доступу до ресурсу:

<https://data.gov.ua/dataset/c22eb3a3-d753-43ce-aadc-4562d8fbfa68>.

9. Урожайність культур сільськогосподарських за регіонами [Електронний ресурс] – Режим доступу до ресурсу:

<https://data.gov.ua/dataset/188450fa-3ab3-43b5-a783-503b52b3f847>.

10. Посівні площі основних сільськогосподарських культур

[Електронний ресурс] – Режим доступу до ресурсу:

<https://data.gov.ua/dataset/fbef4abf-8ba3-4d45-85db-c63c95e4fec7>.

11. Інформаційні системи і технології в управлінні діяльністю підприємства // Public communication in science: philosophical, cultural, political,

economic and IT context. –Volume 1 – May 15, 2020 – Houston, USA – С. 66–67.

(тези доступні за посиланням:

<https://ojs.ukrlogos.in.ua/index.php/logos/article/download/2664/2551>)

12. Місюк М.В. Вплив інноваційної діяльності на розвиток ресурсного

потенціалу аграрних підприємств / Місюк М.В., Мушеник І.М.. //

ІННОВАЦІЙНА ЕКОНОМІКА. – 2018. – №7. – С. 50–55. (стаття доступна за

посиланням: <http://inneco.org/index.php/innecoua/article/view/313/375>)

13. US 20180262571 A1, Int. Cl. H04L 29/08 ; H04W 4/38; G05B 13/02;

Integrated IoT (Internet of Things) system solution for smart agriculture management

// Akhtar; Sabrina [US] – Appl. No.: 15/451,420 – Заявлено: March 7, 2017 2020-03-

21; Опубліковано: Sep 13, 2018

14. Use-case diagrams [Електронний ресурс] // IBM – Режим доступу до

ресурсу: <https://www.ibm.com/docs/en/rational-soft-arch/9.6.1?topic=diagrams-use->

case.

15. Activity diagrams [Електронний ресурс] // IBM – Режим доступу до

ресурсы: <https://www.ibm.com/docs/en/rational-soft-arch/9.6.1?topic=diagrams-activity/>

16. Sequence diagrams [Електронний ресурс] // IBM – Режим доступу до

ресурсы: <https://www.ibm.com/docs/en/rational-soft-arch/9.6.1?topic=diagrams-sequence>

17. Ларичев О.И., Петровский А.Б. Системы поддержки принятия решений: современное состояние и перспективы развития // Итоги науки и техники. - М.: ВИНТИ - 1987. - Т. 21 - с. 131-164 (тези доступні за посиланням:

http://www.raai.org/library/papers/Larichev/Larichev_Petrovsky_1987.pdf)

18. Simon H. The new science of management decision / Herbert Simon. – New York: Harper & Brothers Publishers, 1960. – 72 с.

19. Шевчук І.Б. Конспект лекцій з навчальної дисципліни «Системи підтримки прийняття рішень» [Електронний ресурс] / Шевчук І.Б. // Львівський Національний Університет ім. Івана Франка. – 2020. – Режим доступу до ресурсу: <https://financial.lnu.edu.ua/wp-content/uploads/2015/12/konspekt-lektsiy-SPPR.pdf>.

20. Data Warehouse [Електронний ресурс] // IBM. – 2020. – Режим доступу до ресурсу: <https://www.ibm.com/cloud/learn/data-warehouse>.

21. Introduction to SQL Server Data Mining [Електронний ресурс]. – 2019. – Режим доступу до ресурсу: <https://www.sqlshack.com/introduction-to-sql-server-data-mining/>.

22. OLAP [Електронний ресурс] // IBM. – 2020. – Режим доступу до ресурсу: <https://www.ibm.com/cloud/learn/olap/>.

23. What is Deployment Diagram? [Електронний ресурс] – Режим доступу до ресурсу: <https://www.visual-paradigm.com/guide/uml-unified-modeling-language/what-is-deployment-diagram/>.

24. Навчальний посібник до вивчення дисципліни “Організація сховищ даних” / Голуб Б.Л., Ящук Д.Ю. - К:2018 р. - 165 с.

25. SQL Server Integration Services (SSIS) для начинающих – часть 1 [Електронний ресурс] – Режим доступу до ресурсу: <https://habr.com/ru/post/330618/>.

26. ЕНМК Технологія Data Mining [Електронний ресурс] / Голуб Б.Л. – Режим доступу до ресурсу: <https://elearn.nubip.edu.ua/course/view.php?id=919>

27. 6 Easy Steps to Learn Naive Bayes Algorithm [Електронний ресурс]. – 2017. – Режим доступу до ресурсу:

<https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/>.

28. Association rules [Електронний ресурс]. – 2020. – Режим доступу до ресурсу: <https://searchbusinessanalytics.techtarget.com/definition/association-rules-in-data-mining>.

НУБІП України

НУБІП України

НУБІП України

НУБІП України

НУБІП України **Додаток А**

НУБІП України

НУБІП **ДІАГРАМИ ПРЕЦЕДЕНТІВ** України

НУБІП України

НУБІП України

НУБІП України

НУБІП України

НУБІП·Є україни

НУБІП·Є україни

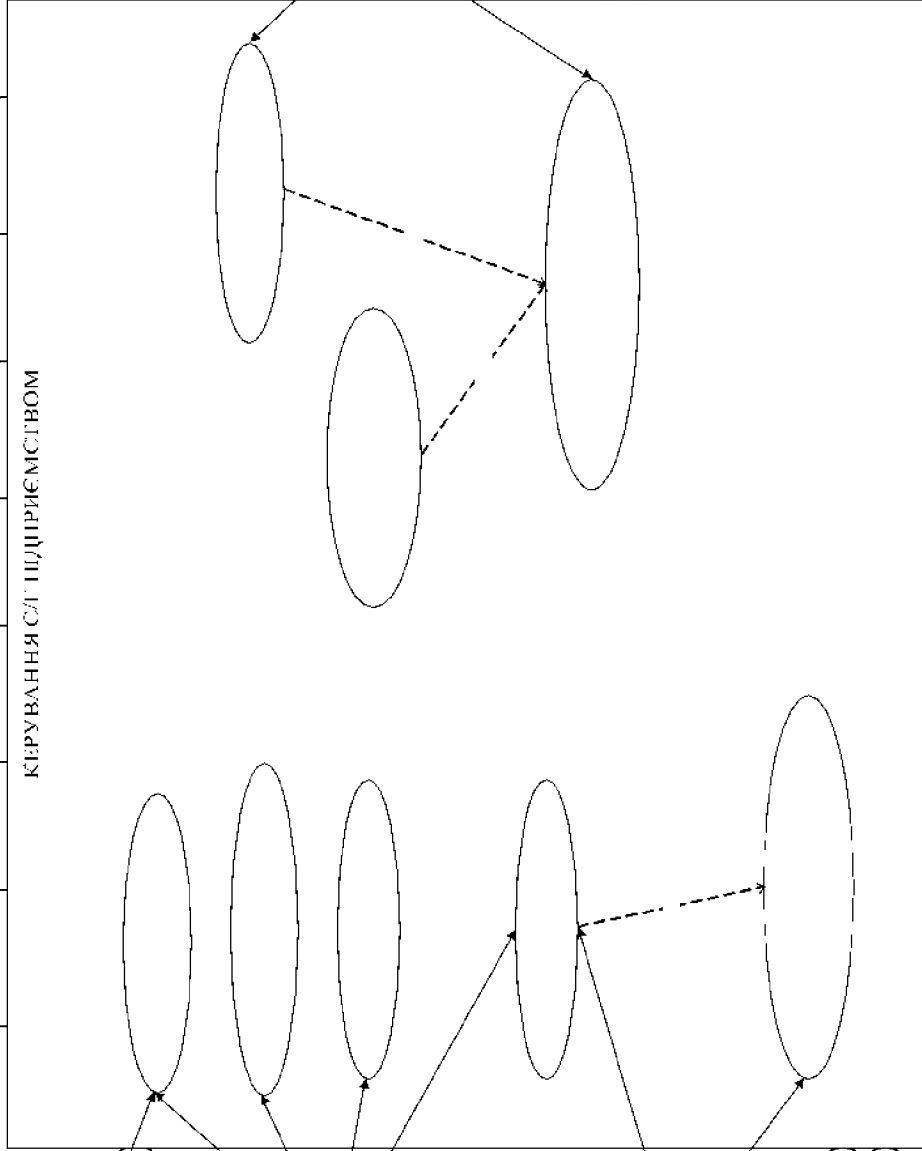
НУБІП·Є україни

НУБІП·Є україни

НУБІП·Є україни

НУБІП·Є україни

НУБІП·Є україни



КЕРУВАННЯ СЛ ПІДПРИЄМСТВОМ

© 2010-2011

Рис. 1 Діаграма прецедентів процесу керування с/г об'єктом

НУБІП України

НУБІП України

НУБІП України

НУБІП України

НУБІП України

НУБІП України

НУБІП України

НУБІПЄУКРАЇНИ

НУ

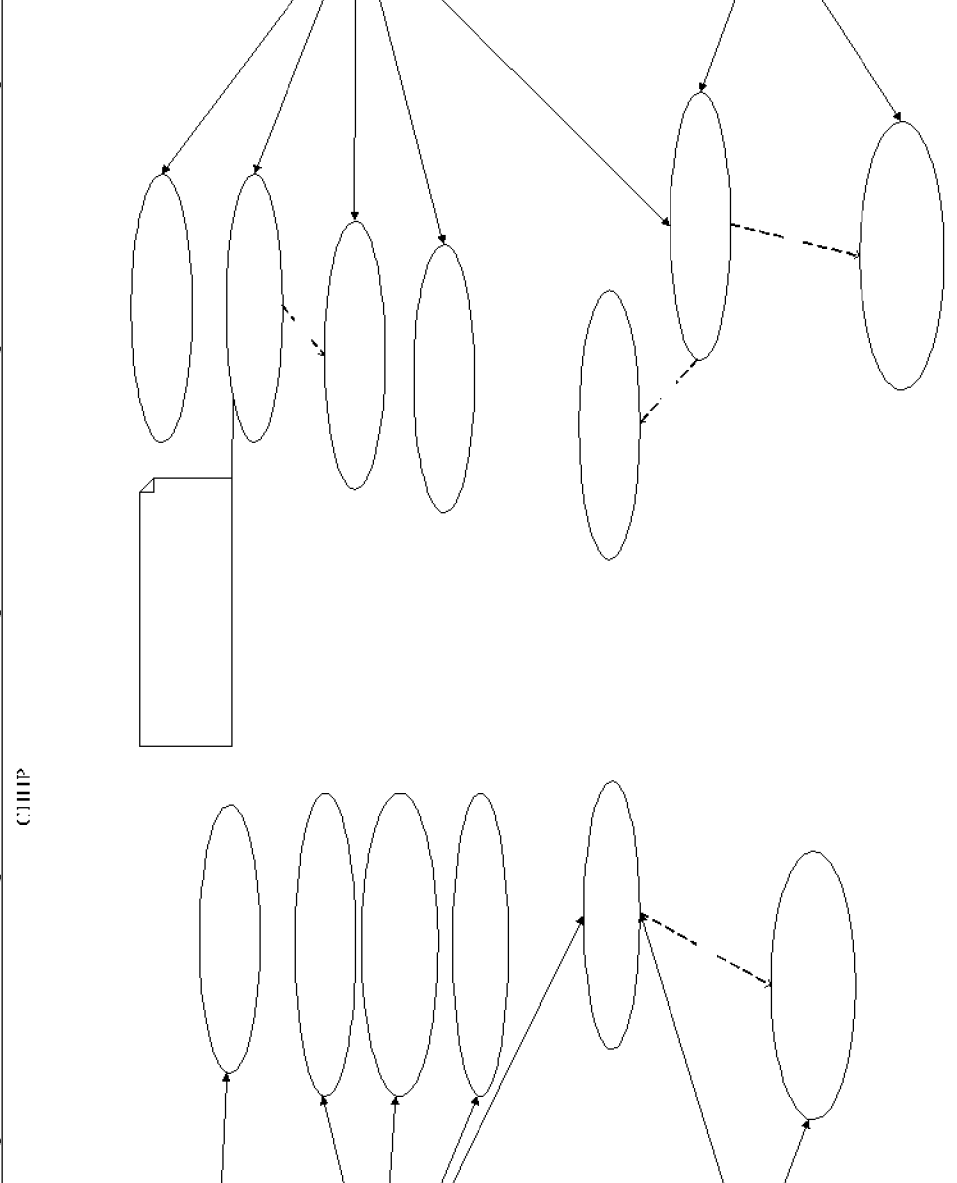
НУ

НУ

НУ

НУ

НУ



БІ

ПЄ

УК

РА

СІПР

ІПЄУКРАЇНИ

БІПЄУКРАЇНИ

Рис. 2 Діаграма прецедентів СППР

НУБІП України

НУБІП України

НУБІП України

НУБІП України

НУБІП України

НУБІП України

НУБІП України

Додаток Б

ДІАГРАМА РОЗМІЩЕННЯ

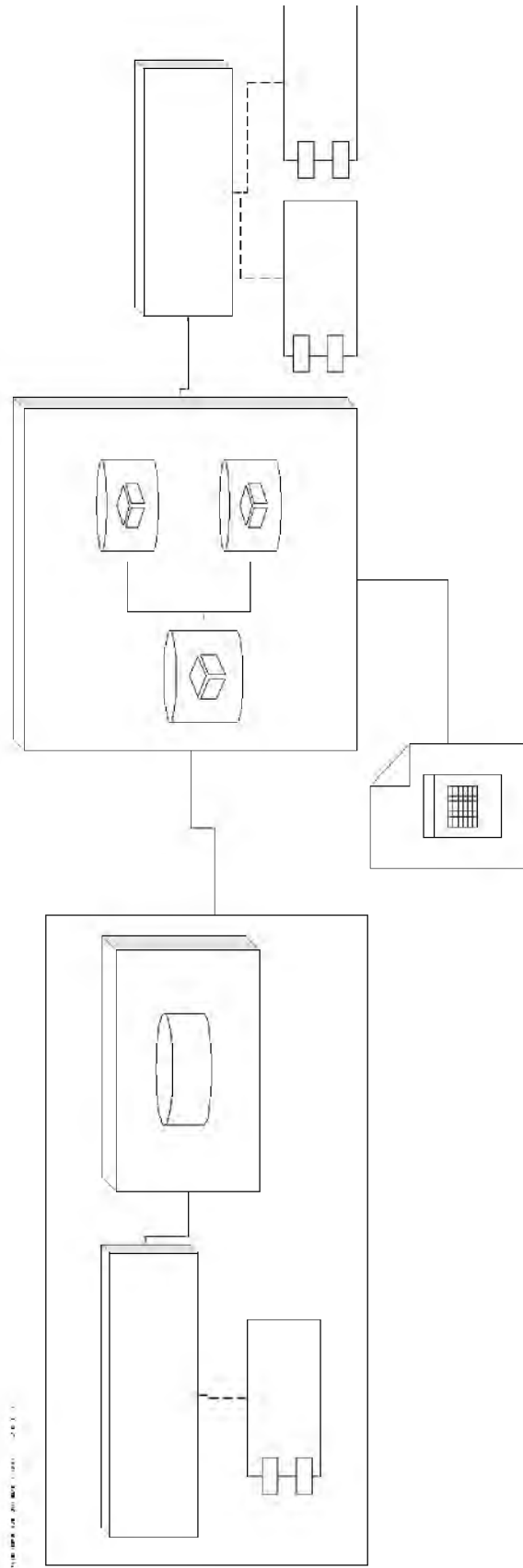


Рис. 1 Архітектура системи

СКРИПТИ ФОРМУВАННЯ ОПЕРАТИВНОЇ БД ПЕРЕПЕЛИНОГО
ГОСПОДАРСТВА

```

CREATE TABLE dbo.breed_group
(
    id_breed char(5) primary key,
    name_breed nvarchar(150) NOT NULL
)
CREATE TABLE dbo.farm_facility
(
    id_facility char(5) primary key,
    name_facility nvarchar(200) NOT NULL,
    square money NULL
)
CREATE TABLE dbo.condition_detention
(
    date_record char(10) NOT NULL,
    id_facility char(5) NOT NULL,
    temperature decimal(5,3) NOT NULL,
    humidity decimal(6,3) NOT NULL,
    light decimal(10,0) NOT NULL,
    primary key
    (date_record,
    id_facility)
)
ALTER TABLE dbo.condition_detention ADD CONSTRAINT
FK_condition_detention_farm_facility FOREIGN KEY (id_facility)
REFERENCES dbo.farm_facility (id_facility)
CREATE TABLE dbo.feed_type
(
    id_feed_type char(5) primary key,
    name_feed_type nvarchar(50) NOT NULL,
    Fe decimal(6,3) NOT NULL,
    Ca decimal(6,3) NOT NULL,
    P decimal(6,3) NOT NULL,
    Na decimal(6,3) NOT NULL,
    Protein decimal(6,3) NOT NULL,
    energy_value_kcal decimal(8,3) NOT NULL
)
CREATE TABLE dbo.feed_group
(
    id_feed_group char(5) primary key,
    id_feed_type char(5) NOT NULL,
    count_feed nchar(10) NOT NULL
)
ALTER TABLE dbo.feed_group ADD CONSTRAINT
FK_feed_group_feed_type FOREIGN KEY (id_feed_type)
REFERENCES dbo.feed_type (id_feed_type)
CREATE TABLE dbo.age_group
(
    id_age_group char(5) primary key,
    name_age_group nvarchar(50) NOT NULL,
    from_age_day int NOT NULL,
    to_age_day int NOT NULL
)
CREATE TABLE dbo.individual
(
    id_individual char(5) primary key,
    id_breed char(5) NOT NULL,
    id_age_group char(5) NOT NULL,
    id_feed_group char(5) NOT NULL,
    id_facility char(5) NOT NULL,

```

```

        weight money NOT NULL
    )
ALTER TABLE dbo.individual ADD CONSTRAINT
    FK_individual_breed_group FOREIGN KEY (id_breed)
    REFERENCES dbo.breed_group (id_breed)
ALTER TABLE dbo.individual ADD CONSTRAINT
    FK_individual_age_group FOREIGN KEY (id_age_group)
    REFERENCES dbo.age_group (id_age_group)
ALTER TABLE dbo.individual ADD CONSTRAINT
    FK_individual_farm_facility FOREIGN KEY ( id_facility)
    REFERENCES dbo.farm_facility (id_facility )
ALTER TABLE dbo.individual ADD CONSTRAINT
    FK_individual_feed_group FOREIGN KEY (id_feed_group)
    REFERENCES dbo.feed_group (id_feed_group )
CREATE TABLE dbo.productivity
(
    id_individual char(5) NOT NULL,
    date_record char(10) NOT NULL,
    count_eggs smallint NOT NULL,
    max_weight_egg decimal(6,3) NOT NULL,
    min_weight_egg decimal(6,3) NOT NULL,
    primary key (id_individual,
    date_record)
ALTER TABLE dbo.productivity ADD CONSTRAINT
    FK_productivity_individual FOREIGN KEY (id_individual)
    REFERENCES dbo.individual (id_individual)

```

СКРИПТИ СТВОРЕННЯ ТАБЛИЦЬ СД

Перепелине господарство

```

CREATE TABLE dbo.RegionDim
(
    id_region int primary key,
    name_region nvarchar(50) NOT NULL,
    climate_zone nvarchar(100) NOT NULL
)
CREATE TABLE dbo.HouseholdDim
(
    id_household int primary key,
    id_region int NOT NULL,
    name_household nvarchar(200) NOT NULL,
    FOREIGN KEY (id_region) REFERENCES dbo.RegionDim (id_region)
)
CREATE TABLE dbo.BreedDim
(
    id_breed int primary key,
    name_breed nvarchar(100) NOT NULL
)
CREATE TABLE dbo.FeedTypeDim
(
    id_feed_type int primary key,
    name_feed_type nvarchar(50) NOT NULL,
    Fe decimal(6,3) NOT NULL,
    Ca decimal(6,3) NOT NULL,
    P decimal(6,3) NOT NULL,
    Na decimal(6,3) NOT NULL,
    Protein decimal(6,3) NOT NULL,
    energy_value_kcal decimal(8,3) NOT NULL
)
CREATE TABLE dbo.DateDim
(
    id_date int primary key,
    year_char(4) NOT NULL,
    month_nvarchar(15) NOT NULL,
    day_nvarchar(15) NOT NULL
)
CREATE TABLE dbo.ProductivityFact
(
    id_date int NOT NULL,
    id_feed_type int NOT NULL,
    id_breed int NOT NULL,
    id_household int NOT NULL,
    avr_weight decimal(6,3) NOT NULL,
    min_egg_weight decimal(6,3) NOT NULL,
    max_egg_weight1 decimal(6,3) NOT NULL,
    avr_egg_count decimal(6,3) NOT NULL,
    avr_temperature decimal(5,3) NOT NULL,
    avr_humidity decimal(6,3) NOT NULL,
    avr_light decimal(10,0) NOT NULL,
    primary key (id_date,id_feed_type,id_breed,id_household),
    FOREIGN KEY (id_date) REFERENCES dbo.DateDim(id_date),
    FOREIGN KEY (id_feed_type) REFERENCES dbo.FeedTypeDim (id_feed_type),
    FOREIGN KEY (id_breed) REFERENCES dbo.BreedDim (id_breed),
    FOREIGN KEY (id_household) REFERENCES dbo.HouseholdDim (id_household)
)

```

Господарство, яке вирощує с/г культури

```

CREATE TABLE dbo.RegionDim

```

```

(
  id_region int primary key,
  name_region nvarchar(50) NOT NULL,
  climate_zone nvarchar(100) NOT NULL
CREATE TABLE dbo.PlantDim
(
  id_plant int primary key,
  name_plant nvarchar(100) NOT NULL
)
CREATE TABLE dbo.DateDim
(
  id_date int primary key,
  year_ char(4) NOT NULL,
  month_ nvarchar(15) NOT NULL,
  day_ nvarchar(15) NOT NULL
)
CREATE TABLE dbo.PlantYieldFact
(
  id_date int NOT NULL,
  id_plant int NOT NULL,
  id_region int NOT NULL,
  yield float NOT NULL,
  production float NOT NULL,
  harvested_area float NOT NULL,
  primary key (id_date, id_plant, id_region),
  FOREIGN KEY (id_date) REFERENCES dbo.DateDim(id_date),
  FOREIGN KEY (id_plant) REFERENCES dbo.PlantDim(id_plant),
  FOREIGN KEY (id_region) REFERENCES dbo.RegionDim(id_region)
)

```

Скрипти внесення постійної інформації

```

INSERT INTO [dbo].[RegionDim] ([id_region],[name_region],[climate_zone])
VALUES ('0','Україна','Україна'),
       ('26','Івано-Франківська','Полісся'),
       ('74','Чернігівська','Полісся'),
       ('46','Львівська','Полісся'),
       ('56','Рівненська','Полісся'),
       ('7','Волинська','Полісся'),
       ('18','Житомирська','Полісся'),
       ('21','Закарпатська','Полісся'),
       ('71','Черкаська','Лісостеп'),
       ('73','Чернівецька','Лісостеп'),
       ('63','Харківська','Лісостеп'),
       ('68','Хмельницька','Лісостеп'),
       ('32','Київська','Лісостеп'),
       ('80','м.Київ','Лісостеп'),
       ('53','Полтавська','Лісостеп'),
       ('59','Сумська','Лісостеп'),
       ('61','Тернопільська','Лісостеп'),
       ('5','Вінницька','Лісостеп'),
       ('1','АРК','Степ'),
       ('12','Дніпропетровська','Степ'),
       ('14','Донецька','Степ'),
       ('65','Херсонська','Степ'),
       ('35','Кіровоградська','Степ'),
       ('44','Луганська','Степ'),
       ('48','Миколаївська','Степ'),
       ('51','Одеська','Степ'),
       ('23','Запорізька','Степ')
INSERT INTO [dbo].[BreedDim] ([id_breed],[name_breed])
VALUES ('1','Єстонська'), ('2','Фараон'), ('3','Англійська біла'), ('4','Смокінгова'),
       ('5','Маньчжурська')

```

```
INSERT INTO [dbo].[PlantDim]([id_plant],[name_plant])
VALUES ('1','Культури зернових та зернобобових'), ('2','Буряк цукровий фабричний'),
('3','Соняшник'), ('4','Соя'), ('5','Ріпак озимий(ріпак яровий)'), ('6','Картопля'),
('7','Культури овочевих'), ('8','Культури плодових та ягідних')
```

РЕАЛІЗАЦІЯ АЛГОРИТМІВ

1 Rule

Код уявлень, які розраховують кількість записів, що задовільняють умову:

```

create view v_Breed_an
as
select BreedDim.name_breed, (select COUNT(ProductivityFact.avr_egg_count)
from ProductivityFact
where ProductivityFact.id_breed=BreedDim.id_breed and
ProductivityFact.avr_egg_count < (select AVG(ProductivityFact.avr_egg_count) from
ProductivityFact)) as avg_low,
(select COUNT(ProductivityFact.avr_egg_count)
from ProductivityFact
where ProductivityFact.id_breed=BreedDim.id_breed and
ProductivityFact.avr_egg_count >= (select AVG(ProductivityFact.avr_egg_count)
from ProductivityFact)) as avg_higher
from BreedDim

```

```

create view v_class_Breed
as
select v_Breed_an.name_breed as 'Порода', v_Breed_an.avg_low as 'Менше середнього',
v_Breed_an.avg_higher as 'Більше середнього', (avg_low+avg_higher) as 'Всього',
CASE
WHEN v_Breed_an.avg_low > v_Breed_an.avg_higher THEN 'L'
ELSE 'H'
END AS 'Клас',
CASE
WHEN v_Breed_an.avg_low > v_Breed_an.avg_higher
THEN ROUND((100.0 * avg_low/(1.0 * avg_low+1.0 * avg_higher)),2)
ELSE ROUND((100.0 * avg_higher/(1.0 * avg_low+1.0 * avg_higher)),2)
END AS 'Ймовірність'
from v_Breed_an

```

```

create view v_FeedType_an
as
select FeedTypeDim.name_feed_type, (select COUNT(ProductivityFact.avr_egg_count)
from ProductivityFact
where ProductivityFact.id_feed_type=FeedTypeDim.id_feed_type and
ProductivityFact.avr_egg_count < (select AVG(ProductivityFact.avr_egg_count) from
ProductivityFact)) as avg_low,
(select COUNT(ProductivityFact.avr_egg_count)
from ProductivityFact
where ProductivityFact.id_feed_type=FeedTypeDim.id_feed_type and
ProductivityFact.avr_egg_count >= (select AVG(ProductivityFact.avr_egg_count)
from ProductivityFact)) as avg_higher
from FeedTypeDim

```

```

create view v_class_FeedType
as
select v_FeedType_an.name_feed_type as 'Корм', v_FeedType_an.avg_low as 'Менше середнього',
v_FeedType_an.avg_higher as 'Більше середнього', (avg_low+avg_higher) as 'Всього',
CASE
WHEN v_FeedType_an.avg_low > v_FeedType_an.avg_higher THEN 'L'
ELSE 'H'
END AS 'Клас',
CASE
WHEN v_FeedType_an.avg_low > v_FeedType_an.avg_higher
THEN ROUND((100.0 * avg_low/(1.0 * avg_low+1.0 * avg_higher)),2)
ELSE ROUND((100.0 * avg_higher/(1.0 * avg_low+1.0 * avg_higher)),2)
END AS 'Ймовірність'
from v_FeedType_an

```


Naïve Bayes

Код уявленнь, які розраховують кількість записів, що задовільняють умову:

```

create view Productivity_all
as
SELECT      dbo.BreedDim.name_breed AS Порода,      dbo.FeedTypeDim.name_feed_type AS Корм,
dbo.FeedTypeDim.Fe, dbo.FeedTypeDim.Ca,
dbo.FeedTypeDim.P,      dbo.FeedTypeDim.Na,      dbo.FeedTypeDim.Protein AS Протеїн,
dbo.ProductivityFact.avr_egg_count AS [Кількість яєць],
CASE
    WHEN dbo.ProductivityFact.avr_egg_count >= (select AVG(ProductivityFact.avr_egg_count) from
ProductivityFact) THEN 'H'
    ELSE 'L'
END AS 'Клас'
from
dbo.BreedDim INNER JOIN      dbo.ProductivityFact ON      dbo.BreedDim.id_breed =
dbo.ProductivityFact.id_breed
INNER JOIN dbo.FeedTypeDim ON      dbo.ProductivityFact.id_feed_type = dbo.FeedTypeDim.id_feed_type
INNER JOIN dbo.HouseholdDim ON      dbo.ProductivityFact.id_household = dbo.HouseholdDim id_household
INNER JOIN dbo.DateDim ON      dbo.ProductivityFact.id_date = dbo.DateDim.id_date
INNER JOIN dbo.RegionDim ON      dbo.HouseholdDim.id_region = dbo.RegionDim.id_region
GROUP BY      dbo.BreedDim.name_breed ,      dbo.FeedTypeDim.name_feed_type ,      dbo.FeedTypeDim.Fe,
dbo.FeedTypeDim.Ca,      dbo.FeedTypeDim.P,      dbo.FeedTypeDim.Na,      dbo.FeedTypeDim.Protein

```

```

create view v_bay_help
as
select
(1.0*(select COUNT(*) from Productivity_all where Productivity_all.Клас = 'H')) as [Кількість H],
(1.0*(select COUNT(*) from Productivity_all where Productivity_all.Клас = 'L')) as [Кількість L],
(1.0*(select COUNT(*) from Productivity_all)) as [Кількість]

```

```

create view v_bay_breed
as
SELECT DISTINCT All1.Порода,
    ROUND((1.0*(select COUNT(*) from Productivity_all as All2 where All2.Клас = 'H' and All1.Порода =
All2.Порода))/(1.0*(select COUNT(*) from Productivity_all where Productivity_all.Клас = 'H')),3)
as [Ймовірність H],
    ROUND((1.0*(select COUNT(*) from Productivity_all as All2 where All2.Клас = 'L' and All1.Порода =
All2.Порода))/(1.0*(select COUNT(*) from Productivity_all where Productivity_all.Клас = 'L')),3)
as [Ймовірність L]
from Productivity_all as All1

```

```

create view v_bay_feed
as
SELECT DISTINCT All1.Корм,
    ROUND((1.0*(select COUNT(*) from Productivity_all as All2 where All2.Клас = 'H' and All1.Корм =
All2.Корм))/(1.0*(select COUNT(*) from Productivity_all where Productivity_all.Клас = 'H')),3) as
[Ймовірність H],
    ROUND((1.0*(select COUNT(*) from Productivity_all as All2 where All2.Клас = 'L' and All1.Корм =
All2.Корм))/(1.0*(select COUNT(*) from Productivity_all where Productivity_all.Клас = 'L')),3) as
[Ймовірність L]
from Productivity_all as All1

```

```

create view v_bay_class
as
select view_in.Порода,view_in.Корм,
view_in.[Ймовірність H]*100/(view_in.[Ймовірність H]+view_in.[Ймовірність L]) as [Ймовірність H],
view_in.[Ймовірність L]*100/(view_in.[Ймовірність H]+view_in.[Ймовірність L]) [Ймовірність L]
from
(select Distinct Productivity_all.Порода, Productivity_all.Корм,
v_bay_feed.[Ймовірність H]*v_bay_breed.[Ймовірність H]*(v_bay_help.[Кількість
H]/v_bay_help.[Кількість] as [Ймовірність H],
v_bay_feed.[Ймовірність L]*v_bay_breed.[Ймовірність L]*(v_bay_help.[Кількість
L]/v_bay_help.[Кількість] as [Ймовірність L]
from Productivity_all, v_bay_breed, v_bay_feed, v_bay_help
where Productivity_all.Порода = v_bay_breed.Порода and Productivity_all.Корм = v_bay_feed.Корм) as

```

```
view_in
```

Пошук асоціативних правил

```

SqlParameter bParam, fParam, avrParam;
List<Breed> list_b = new List<Breed>();
List<Feed> list_f = new List<Feed>();
ass_Rules.Clear();
int support;
string rule, t_rule;
string sql;
SqlCommand cmd;
sql = "select Count(*) from ProductivityFact";
cmd = new SqlCommand(sql, conn);
int sup_D = Convert.ToInt32(cmd.ExecuteScalar());
text_sup_D.Text = sup_D.ToString();
for (int i = 0; i < list_b.Count; i++)
{
    int id_breed = list_b[i].getID();
    sql = "select Count(*) from ProductivityFact where id_breed = @breed";
    cmd = new SqlCommand(sql, conn);
    bParam = new SqlParameter("@breed", id_breed);
    cmd.Parameters.Add(bParam);
    t_rule = "Порода = " + list_b[i].getBreed();
    support = Convert.ToInt32(cmd.ExecuteScalar());
    if (support != 0) ass_Rules.Add(new Ass_Rule(support, 1, t_rule));
    sql = "select Count(*) from ProductivityFact where id_breed = @breed and avr_egg_count > @avr";
    cmd = new SqlCommand(sql, conn);
    bParam = new SqlParameter("@breed", id_breed);
    avrParam = new SqlParameter("@avr", aver);
    cmd.Parameters.Add(bParam);
    cmd.Parameters.Add(avrParam);
    support = Convert.ToInt32(cmd.ExecuteScalar());
    rule = t_rule + ", Середня кількість яєць >= " + aver;
    if (support != 0) ass_Rules.Add(new Ass_Rule(support, 2, rule));
    sql = "select Count(*) from ProductivityFact where id_breed = @breed and avr_egg_count < @avr";
    cmd = new SqlCommand(sql, conn);
    bParam = new SqlParameter("@breed", id_breed);
    avrParam = new SqlParameter("@avr", aver);
    cmd.Parameters.Add(bParam);
    cmd.Parameters.Add(avrParam);
    support = Convert.ToInt32(cmd.ExecuteScalar());
    rule = t_rule + ", Середня кількість яєць < " + aver;
    if (support != 0) ass_Rules.Add(new Ass_Rule(support, 2, rule));
}
for (int i = 0; i < list_f.Count; i++)
{
    int id_feed = list_f[i].getID();
    sql = "select Count(*) from ProductivityFact where id_feed_type = @feed";
    cmd = new SqlCommand(sql, conn);
    fParam = new SqlParameter("@feed", id_feed);
    cmd.Parameters.Add(fParam);
    support = Convert.ToInt32(cmd.ExecuteScalar());
    t_rule = "Корм = " + list_f[i].getFeed();
    if (support != 0) ass_Rules.Add(new Ass_Rule(support, 1, t_rule));
    for (int i = 0; i < list_b.Count; i++)
    {
        int id_breed = list_b[i].getID();
        t_rule = "Порода = " + list_b[i].getBreed();
        for (int j = 0; j < list_f.Count; j++)
        {
            int id_feed = list_f[j].getID();
            sql = "select Count(*) from ProductivityFact where id_breed = @breed and id_feed_type = @feed";
            cmd = new SqlCommand(sql, conn);
            bParam = new SqlParameter("@breed", id_breed);

```

```

fParam = new SqlParameter("@feed", id_feed);
cmd.Parameters.Add(bParam);
cmd.Parameters.Add(fParam);
support = Convert.ToInt32(cmd.ExecuteScalar());
rule = t_rule + ", Корм = " + list_f[j].getFeed();
if (support != 0) ass_Rules.Add(new Ass_Rule(support, 2, rule));
    }
}
for (int i = 0; i < list_b.Count; i++)
{
    int id_breed = list_b[i].getID();
    t_rule = "Порода = " + list_b[i].getBreed();
    for (int j = 0; j < list_f.Count; j++)
    {
        int id_feed = list_f[j].getID();
        sql = "select Count(*) from ProductivityFact where id_breed = @breed and id_feed_type = @feed and avr_egg_count > @avr";
        cmd = new SqlCommand(sql, conn);
        bParam = new SqlParameter("@breed", id_breed);
        fParam = new SqlParameter("@feed", id_feed);
        avrParam = new SqlParameter("@avr", aver);
        cmd.Parameters.Add(bParam);
        cmd.Parameters.Add(fParam);
        cmd.Parameters.Add(avrParam);
        support = Convert.ToInt32(cmd.ExecuteScalar());
        rule = t_rule + ", Корм = " + list_f[j].getFeed() + ", Середня кількість яєць >= " + aver;
        if (support != 0) ass_Rules.Add(new Ass_Rule(support, 3, rule));
        sql = "select Count(*) from ProductivityFact where id_breed = @breed and id_feed_type = @feed and avr_egg_count < @avr";
        cmd = new SqlCommand(sql, conn);
        bParam = new SqlParameter("@breed", id_breed);
        fParam = new SqlParameter("@feed", id_feed);
        avrParam = new SqlParameter("@avr", aver);
        cmd.Parameters.Add(bParam);
        cmd.Parameters.Add(fParam);
        cmd.Parameters.Add(avrParam);
        support = Convert.ToInt32(cmd.ExecuteScalar());
        rule = t_rule + ", Корм = " + list_f[j].getFeed() + ", Середня кількість яєць < " + aver;
        if (support != 0) ass_Rules.Add(new Ass_Rule(support, 3, rule));
    }
}
}

```