

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ БІОРЕСУРСІВ
І ПРИРОДОКОРИСТУВАННЯ УКРАЇНИ

НУБІП України

Факультет інформаційних технологій

УДК 004.9:005.93
«ПОГОДЖЕНО» «ДОПУСКАЄТЬСЯ ДО ЗАХИСТУ»

Декан факультету

інформаційних технологій

Завідувач кафедри комп'ютерних наук

Глазунова О.Г., д.п.н., професор Голуб Б.Л., к.т.н., доцент
НУБІП України 2022 р. НУБІП України 2022 р.

МАГІСТЕРСЬКА КВАЛІФІКАЦІЙНА РОБОТА

на тему «Інтелектуальна система аналізу результатів самоцінювання рівня цифровізації компаній та цифрових навичок їх співробітників»
Спеціальність 122 «Комп'ютерні науки»

Освітня програма Інформаційні управляючі системи та технології

Орієнтація освітньої програми освітньо-професійна

Гарант освітньої програми
К.Е.Н. ДОЦЕНТ Гуслера О.М.
(науковий ступінь та вчене звання) (підпис) (ПІБ)

Керівник магістерської кваліфікаційної роботи

старший викладач Ящук Дар'я Юріївна
(науковий ступінь та вчене звання) (підпис) (ПІБ)
доц. к.т.н. Дудник Алла Олексіївна
(науковий ступінь та вчене звання) (підпис) (ПІБ)

Виконав

Скус Ю.Ю

НУБІП України КІЇВ-2022

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ БІОРЕСУРСІВ
І ПРИРОДОКОРИСТУВАННЯ УКРАЇНИ

Факультет інформаційних технологій

ЗАТВЕРДЖУЮ

Завідувач кафедри комп'ютерних наук

Доц., к.т.н.

Голуб Б.Л.

(науковий ступінь, вчене звання)

(підпис)

(ІПБ)

“11” жовтня 2022 року

ЗАВДАННЯ

ДО ВИКОНАННЯ МАГІСТЕРСЬКОЇ КВАЛІФІКАЦІЙНОЇ РОБОТИ СТУДЕНТУ

Скус Юрій Юрійович

(прізвище, ім'я, по батькові)

Спеціальність 122 «Комп'ютерні науки»

Освітня програма «Інформаційні управляючі системи та технології»

Орієнтація освітньої програми: освітньо-професійна

Тема магістерської кваліфікаційної роботи інтелектуальна система аналізу результатів самооцінювання рівня цифровізації компаній та цифрових навичок їх співробітників

затверджена наказом ректора НУБіП України від “11” жовтня 2022р. №1431 «С»

Термін подання завершеної роботи на кафедру 27 жовтня 2022 р.

Вихідні дані до магістерської кваліфікаційної роботи:

1. Дані про анкетування працівників
2. Дані про оцінку компанії

Перелік питань, що підлягають дослідженню:

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

Дата видачі завдання “01” листопада 2021 р.

Керівник магістерської кваліфікаційної роботи

(підпис)

Яшук Д.Ю.

(прізвище та ініціали)

Дудник А.О.

(підпис)

(прізвище та ініціали)

Завдання прийняв до виконання

(підпис)

Скус Ю.Ю.

(прізвище та ініціали)

	Зміст
НУБІП України	5
Перелік умовних позначень.....	5
Вступ.....	6
1 Системний аналіз предметної області.....	9
НУБІП України	9
1.1 Опис предметної області.....	9
1.2 Огляд існуючих рішень.....	11
1.3 Постановка завдання.....	17
2 Моделювання системи.....	19
НУБІП України	19
2.1 Загальні відомості про моделювання засобами UML.....	19
2.2 Діаграма прецедентів.....	20
2.3 Діаграма послідовності.....	24
2.4 Діаграма активності.....	25
НУБІП України	27
3 Розробка системи.....	27
3.1 Логічна модель даних.....	27
3.2 Проектування сховище даних.....	29
3.3 Побудова розгорнутого куба.....	32
3.4 Реалізація міграції даних за допомогою Data Flow.....	37
НУБІП України	40
4 Результати дослідження.....	40
4.1 Побудова звітності засобами SSRS.....	40
НУБІП України	42
4.2 Розрахунок KPI.....	42
4.3 Інтелектуальний аналіз даних.....	44
4.3.1 Алгоритм 1R.....	47
4.3.2 Спрощений алгоритм Байєса.....	49
НУБІП України	52
4.3.3 Пошук асоціативних правил.....	52
4.4 Тестування системи.....	56

НУБІП України

НУБІП України

НУБІП України

НУБІП України

НУБІП України

НУБІП України

НУБІП України

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

НУБІП України

DM – Data Mining.

БД – база даних.

СД – сховище даних.

СУБД – система управління базами даних.

SQL – SQL Server Analysis Services (SSAS).

SSAS – SQL Server Integration Services.

SSIS – SQL Server Integration Services.

OLAP – On-Line Analytical Processing.

DF – Data Flow.

НУБІП України

НУБІП України

НУБІП України

НУБІП України

НУБІП України

ВСТУП

НУБІП України

Актуальність. Особливості розвитку сучасного глобального світу

спричиняють перехід до економіки знань, що зумовлює потребу цифрової трансформації задля підвищення рівня розвитку економіки та суспільства.

НУБІП України

Сучасні цифрові технології є каталізатором трансформації світу. Цифрова

трансформація має величезний вплив на бізнес і суспільне життя, відкриваючи нові можливості для розблокування економічних і соціальних переваг. Сьогодні

НУБІП України

розвиток цифрового середовища охоплює практично всі сфери життя. Зокрема, у прийнятті на роботу керівники вимагають володіння людиною цифровими навичками, які дають можливість швидко та ефективно виконувати поставлені завдання, бути успішною та використовувати потенційні можливості.

Для визначення рівня цифрових компетентностей розроблено кілька рамок, зокрема, European e-Competence Framework for ICT Professionals , European Computer Driving Licence, ICT Literacy Competencies, Global Media and Information Literacy Assessment Framework. В даному дослідженні ми взяли за

НУБІП України

основу європейську систему цифрової компетентності громадян, відому також як DigComp , що дозволяє запровадити спільний підхід до визначення і опису основних сфер цифрової компетентності громадян, і є загальним орієнтиром на європейському рівні [1].

НУБІП України

Разом з тим, в умовах цифрової трансформації важливо не лише формувати цифрову компетентність громадян, але й здійснювати цифровізацію бізнесу та соціальної сфери. Відповідно, актуалізується потреба співвіднесення рівня цифровізації установи та готовності до ефективної роботи її працівників, що визначається рівнем їх цифрової компетентності.

НУБІП України

Таким чином аналіз навичок цифрової трансформації допоможе компаніям та їх співробітникам вказати на їх сильні та слабкі сторони, на які в подальшому вони зможуть приділяти більше уваги щоб їх покращити , що безумовно покращить продуктивність компанії.

Мета дослідження: використання сучасних технологій OLAP в роботі інтелектуальної системи аналізу результатів самооцінювання рівня цифровізації компаній та цифрових навичок їх співробітників.

Завдання Для досягнення поставленої мети необхідно:

- провести системний аналіз об'єкта;
- провести самоопитування кількох підприємств
- Провести аналіз отриманих даних опитаних співробітників
- Реалізувати алгоритми, які дозволяють аналізувати та шукати певні залежності між параметрами

Методи дослідження для проведення аналізу даних, розрахунку ключових показників ефективності використовується технологія OLAP, алгоритми Data Mining для глибинного аналізу даних.

Наукова новизна вперше було впроваджено в Україні самооцінювання навичок компаній в цілому, а не окремих людей як це зроблено в проєкті «Цифрограм».

Апробація результатів дослідження. В процесі розробки магістерської роботи на тему «Інтелектуальна система аналізу результатів самооцінювання рівня цифровізації компаній та цифрових навичок їх співробітників» були проведені виступи на Міжнародній науково-практичній конференції молодих вчених «ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ: ЕКОНОМІКА, ТЕХНІКА, ОСВІТА '2022» (26-27 жовтня 2022 року, НУБіП України, Київ, 2022).

Структура роботи Кількість сторінок в роботі – 63, кількість використаних джерел – 25. Робота складається з анотації, переліку умовних позначень, вступу, чотирьох розділів:

У *першому розділі* розглядається предметна область, сформульована постановка задачі дослідження та аналіз системи самооцінювання.

Другий розділ складається з моделювання системи. Були розроблені діаграми прецедентів, послідовності, активності, які дозволяють отримати передумови для аналізу поведінки системи.

У третьому розділі розповідається про розробку та реалізацію інформаційного забезпечення системи. Було продумано та побудовано логічну модель даних, реляційну БД та спроектовано сховище даних та створено розгорнутий куб системи.

У *четвертому розділі* відображення результатів розробки програмного забезпечення системи. Були розроблені алгоритми роботи системи.

НУБІП України

НУБІП України

НУБІП України

НУБІП України

1 СИСТЕМНИЙ АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ

НУБІП УКРАЇНИ

1.1 Опис предметної області

Заклик на цифровізацію та відповідну освіту зумовлений постійним розвитком технологій. Так як, майже все, що ми робимо, потребує певного рівня цифрових навичок. До того ж, ця тенденція не буде затухати, лише набиратиме свої оберти, адже технології не стоять на місці, усе більше проникають у наш побут, стаючи невід'ємною частиною робочого й повсякденного життя. Можна сказати, що усе зводиться до постійно зростаючої потреби у правильних, цифрових навичках [2].

Хоча використання комп'ютерів, мобільних телефонів та Інтернету зростає майже серед усіх груп людей, це не обов'язково означає, що вони розвивають навички та можуть отримати від них користь у багатьох різних аспектах життя. Дослідження показали, що великі обсяги використання комп'ютерів, мобільних пристроїв та Інтернету сприяють лише цифровим навичкам на операційному рівні. Вища пізнавальна здатність до критичного пошуку та відбору інформації не є наслідком більшого споживання. Користувачі можуть просто залишатися на одному рівні і використовувати лише деякі конкретні програми. Тому велике споживання технології як такої не слід розглядати як доказ цифрової компетентності [3].

Цифрова компетентність – це поєднання знань, навичок та ставлення до використання технологій для виконання завдань, вирішення проблем, спілкування, управління інформацією, співпраці, а також для створення та обміну вмістом ефективно, належним чином, безпечно, критично, творчо, самостійно та етично.

З огляду на це розуміння, одне з найкращих визначень і найменш вразливе до випробування часом було розроблено великим дослідницьким проєктом ЄС,

після того як Європейська Комісія визнала цифрову компетентність однією з восьми ключових компетентностей для навчання впродовж життя [4].

Разом з тим, в умовах цифрової трансформації важливо не лише формувати цифрову компетентність громадян, але й здійснювати цифровізацію бізнесу та соціальної сфери. Відповідно, актуалізується потреба співвіднесення рівня цифровізації установи та готовності до ефективної роботи її працівників, що визначається рівнем їх цифрової компетентності.

Розуміння поняття цифрової компетентності настільки різноманітне, що не існує загальноприйнятого або загальновизнаного визначення. Те саме траплялося і трапляється практично з усіма концепціями стосовно цифрових інструментів та процесів. Це спричинене, серед іншого, постійним і швидким розвитком технологій, що дають можливість і створюють нові види діяльності та цілі. Прикладами можуть бути ІТ-грамотність, цифрова грамотність, медіа грамотність, інформаційна грамотність, Інтернет-грамотність тощо. Вони з'явилися одночасно з технологічним розвитком і, оскільки суспільство визнало потребу в нових компетенціях [5].

Той факт, що існує так багато різноманітних визначень цього терміна, відображає його важливість. Загальним для всіх них є те, що мова вже йде не про доступ до технологій та використання технологій, а про можливість скористатися цими значущими способами - на все життя, працю та навчання.

Для визначення рівня цифрових компетентностей розроблено кілька рамок, зокрема, European e-Competence Framework for ICT Professionals, European Computer Driving Licence, ICT Literacy Competencies, Global Media and Information Literacy Assessment Framework. В даному дослідженні ми взяли за основу європейську систему цифрової компетентності громадян, відому також як DigComp, що дозволяє запровадити спільний підхід до визначення і опису основних сфер цифрової компетентності громадян, і є загальним орієнтиром на європейському рівні. DigComp узгоджується із іншими рамками і має досвід впровадження у практику європейських країн.

Підтвердженням цьому є визначення ЄС цифрової грамотності (компетентності) як однієї з ключових компетенцій для навчання упродовж життя [1].

Аналіз Рамкової програми оновлених ключових компетентностей для навчання протягом життя показує, що цифрова компетентність передбачає впевнене, критичне та відповідальне використання і взаємодію з цифровими технологіями для навчання, професійної діяльності (роботи) та участі в житті суспільства. Вона включає цифрову й інформаційну грамотність, комунікацію і співпрацю, створення цифрового контенту (зокрема програмування), кібербезпеку та розв'язання проблем [6].

1.2 Огляд існуючих рішень

Цифрова трансформація, технології та гаджети зумовили необхідність у цифрових навичках, які вже давно вийшли за межі ІТ-індустрії та розповсюджуються на всі сфери нашого життя. Вже сьогодні ледь не кожна компанія при запрошенні на роботу вимагає цифрової грамотності, а в найближчому майбутньому вона матиме життєво важливе значення [7].

Керівників підприємств варто розуміти, що усі робочі місця у майбутньому десятилітті будуть нероздільно пов'язані з технологіями, і їм потрібно буде оцінити рівень цифрових навичок своїх працівників для розуміння стану діджиталізації компанії. Для оцінки цифрових навичок є не так багато сервісів в світі [7].

Розглянемо деякі існуючі рішення.
THE DIGITAL COMPETENCE WHEEL – це англійський сервіс для визначення цифрової компетентності людей [4]. Цей сервіс показано на рис.1.



Рис. 1 Сервіс THE DIGITAL COMPETENCE WHEEL
Опишемо далі, як працює даний сервіс.

1. Ви реєструєтесь на digital-competence.eu і підтверджуєте акаунт через пошту.

2. На головній сторінці через кнопку «START» починаємо анкетування.

3. Проходимо анкетування яке складається з 60 запитань (анкета для проходження статична).

4. Отримуємо результат у вигляді цифрового колеса ,на якому бачимо огляд того, які цифрові компетенції існують, а які з них потрібно вдосконалити.

До його головних переваг належать:

- *Швидка реєстрація.* Реєстрація відбувається тільки через підтвердження електронної пошти без внесення інших даних,

- *Детальний результат.* Кожна колонка в діаграмі представляє компетенцію з можливим балом від 0 до 100%. Чим вищий бал, тим сильніша компетентність. По центру колеса відображається оцінка вашої загальної цифрової компетентності.

Його головні недоліки:

- *Відсутність Української мови.* Даний сервіс на даний момент не підтримує українську мову тому оцінка українських компаній не можлива;

- *Відсутність платного коментарю.* Доступ до редагування питань, індивідуальних анкет, діалогів, аналіз звітів можна тільки придбавши преміум

підписку. Підписка дає керівнику швидку та просту можливість отримати огляд загального рівня цифрової компетентності організації чи підрозділу та одночасно отримати доступ до результатів окремого працівника. Це відкриває нові можливості для націлювання та диференціації розвитку цифрових компетентностей, а також відстеження процесу розвитку співробітників з часом;

- *Кількість питань.* Система аналізує пройдану анкету з 60 питань, питання однотипні та вимагають багато терпіння і часу для проходження.

Test ikanos – це іспанський сервіс для визначення цифрової компетентності людей всіх спеціальностей та студентів [8]. Цей сервіс показано на рис.2



El test ikanos te va a proporcionar tu perfil digital basado en el marco europeo de competencias digitales DigComp. El test ikanos es sencillo, funciona en cualquier dispositivo y ha sido completado por mas de 70.000 ciudadanos, profesionales, profesores o estudiantes hasta 2020.

Completarlo te llevara unos 25 minutos y al finalizar, podras descargar tus resultados en un informe personalizado en formato PDF.

test ikanos individual

Рис. 2 Сервіс Test ikanos

Опишемо далі, як працює даний сервіс.

1. На головній сторінці потрібно обрати анкетування що підходить.
2. Після вибору анкети ввести свої дані, після чого розпочнеться анкетування яке в середньому складається з 20 -30 питань.

3. Отримуємо результат у вигляді списку ключових компетентностей та їх оцінки від 0 до 10

До його головних переваг належать:

• *Індивідуальні анкети під ключові сфери.* Даний сервіс має вже підготовленні спеціальні анкети для громадян, студентів, освітніх закладів, будівельників, передових галузь, економістів, маркетологів, банківських фахівців, лікарів;

• *Запитання.* Велика кількість запитань які зроблені в різних формах , що допоможе користувачу при проходженні давати відповіді коректно та не нудьгувати .

Його головні недоліки:

• *Відсутність Української мови.* Даний сервіс на даний момент також не підтримує українську мову;

• *Інтерфейс.* Інтерфейс незрозумілий ,не завжди зрозуміло де що знаходиться та куди тобі переходити, легко заплутати,

• *Результати.* Можливо сформувати свої результати в PDF та переглядати онлайн.

SELFIE – це європейський, безкоштовний, простий у використанні інструмент, який допомагає школам оцінити рівень цифровізації у цифрову епоху [9]. Цей сервіс показано на рис.3.

НУБІП України

НУБІП України

НУБІП України

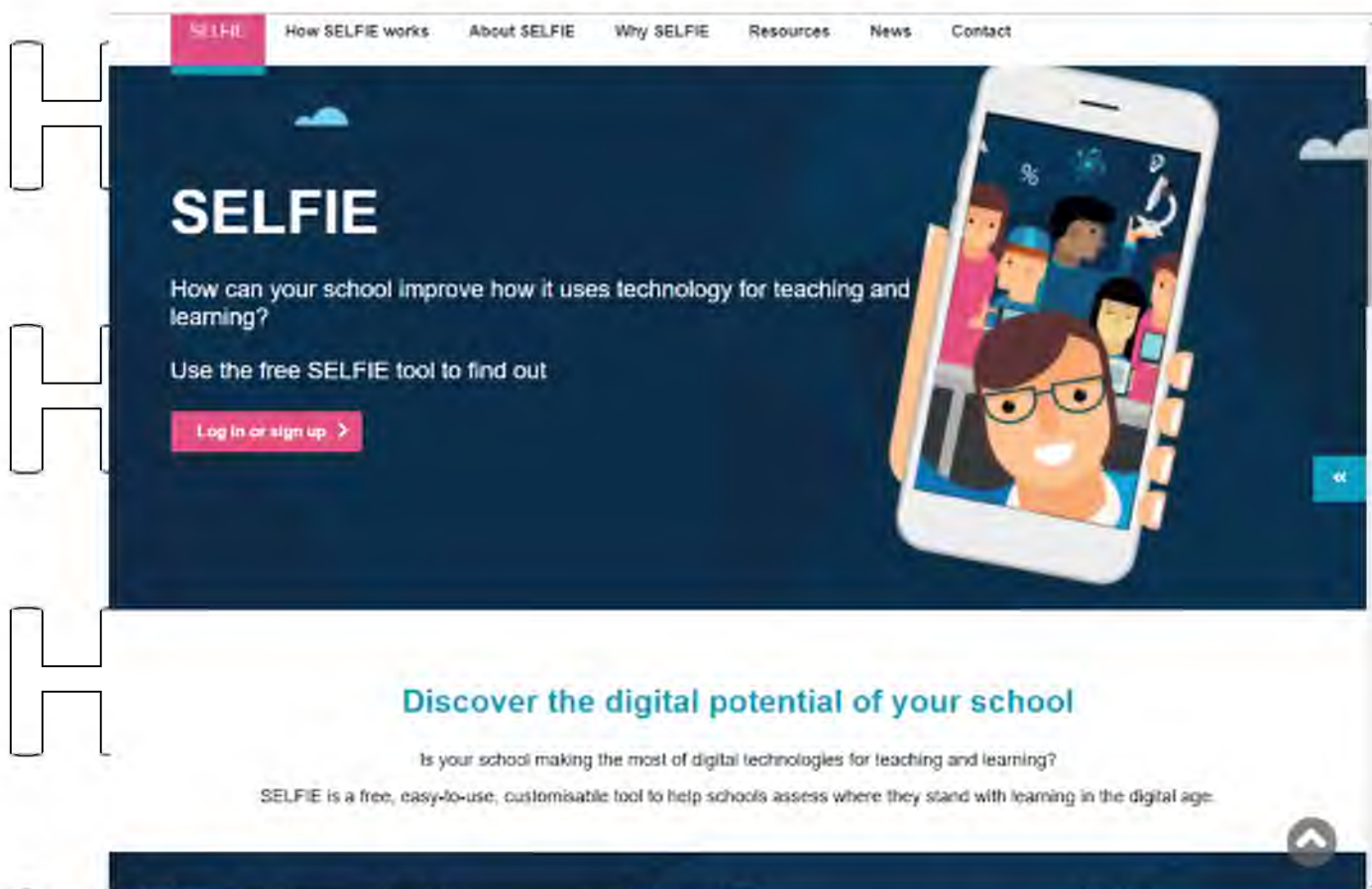


Рис. 3 Сервіс SELFIE

Опишемо далі, як працює даний сервіс.

1. Потрібно зареєструвати школу на платформі, надаючи базову інформацію, таку як тип школи (наприклад, початкова чи середня), розмір та місцезнаходження.

2. Вибрати, хто буде брати участь в опитування.

3. Обрати запитання та анкетування.

4. Та почати проходження анкетування яке триває близько 30 хвилин.

5. Після того, як учасники пройдуть анкетування, сервіс створює звіт для школи.

До його головних переваг належать:

- *Індивідуальність.* Оскільки кожна школа є унікальною, інструменти можна налаштовувати. Школа може обирати та додавати запитання та твердження відповідно до заданих потреб;

- *Анонімність.* Всі відповіді анонімні, а дані захищені. Звіти формуються з анонімних результатів;

Звіти. Після завершення SELFIE кожна школа отримує індивідуальний інтерактивний звіт, який надає як поглиблені дані, так і швидкий огляд сильних та слабких сторін;

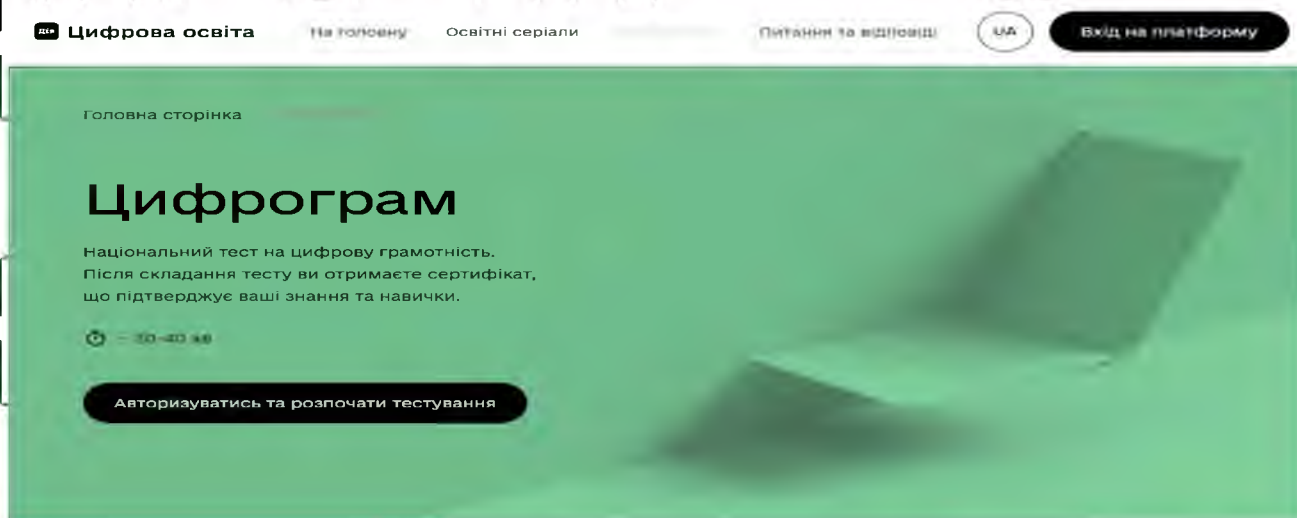
- **Оновлення.** Постій оновлення запитань, можливість проходження в нових країнах, адаптація під нових користувачів.

Його головні недоліки:

- **Відсутність Української мови.** Хоч сервіс є європейський, але української мови чи російської не має;

- **Напрямок.** Сервіс спрямований тільки на освітні заклади, і перевірити цифрову грамотність можна тільки учням, вчителям, керівникам шкіл та студентам.

Цифрограм – це український тест на цифрову грамотність. Завдання тесту систематизовані за сферами знань європейської рамки цифрових компетентностей для громадян [10]. Цей сервіс показано на рис 4.



Що це таке?

Цифрограм – це можливість перевірити цифрову грамотність будь-якому громадянину. Бета-версія

Рис. 4 Сервіс Цифрограм

Опишемо далі, як працює даний сервіс.

1. Спочатку потрібно авторизуватися на платформі. Це потрібно, щоб ви змогли зберегти ваші результати та отримати сертифікат.

2. Обрати анкетування.

3. Перед початком тестування потрібно ввести інформацію про себе.

4. Дані проходження анкетування яке складається з 90 питань.

5. Після чого сервіс покаже сформований результат.

6. І автоматично відбувається генерація сертифікату рівня володіння цифровими компетентностями.

До його головних переваг належать:

- *Першість.* Перша інформаційна система такого типу в Україні, яка не має аналогів;

- *Сертифікат.* Дає можливість використовувати сформований результат як офіційний документ на сайтах пошуку роботи. Підтвердження високого рівня цифрової грамотності підсилить резюме та підвищить шанси на працевлаштування;

- *Звіти.* Після завершення SELFIE кожна школа отримує індивідуальний інтерактивний звіт, який надає як поглиблені дані, так і швидкий огляд сильних та слабких сторін ;

- *Адаптація.* Постійні оновлення та удосконалення , створення нових анкетувань.

Його головні недоліки:

- *Запитання.* Велика кількість запитань, що забирають багато часу.

1.3 Постановка завдання

В результаті аналізу предметної області та огляду існуючих рішень, сформоване завдання, що описує набір функціональності та вимоги яким інтелектуальна система аналізу результатів самооцінювання рівня цифровізації компаній та цифрових навичок їх співробітників повинна відповідати.

Для розробки системи необхідно буде створити сховище даних , що міститиме необхідну інформацію для аналізу цифрових навичок компаній та їх співробітників на базі існуючої оперативної баз даних, до якої надходять дані через самоопитування працівників.

Сховище даних повинно зберігати такі дані стосовно опитувань як:

- Інформація про назву анкети
- Інформація про співробітника
 - Інформація про компанію якій належить співробітник
 - Інформація про результат самоопитування співробітника по певній анкеті
- Інформація про результат компаній по певній анкеті
 - Інформація про період проходження
 - Інформація про середній результат самоопитування компанії
- Інформація про дані компанії

Більш того, за допомогою інструментів OLAP та Data Mining система має відповідати на запитання які представлені нижче.

1. Скільки нових самоопитувань було пройдено за певний відрізок часу,

певної компанії?

2. Які результати самоопитування компаній по певній анкеті?

3. Яку середні результати самоопитування компанії в цілому?

4. Які кількість високих та низьких результатів самоопитування

працівників компанії?

5. Які анкети були пройдені компанією?

2. МОДЕЛЮВАННЯ СИСТЕМИ

2.1 Загальні відомості про моделювання засобами UML

Для моделювання предметної області було обрано мову UML.

UML — це аббревіатура, що позначає Unified Modeling Language. Фактично, це один з найпопулярніших методів моделювання бізнес-процесів, що є міжнародної стандартної нотації для вказівки, візуалізації і документування розробки ПО. Певний групою управління об'єктами, з'явився, як результат декількох додаткових систем нотації UML і тепер став стандартом де-факто для візуального моделювання. Основний принцип будь-якого об'єктно-орієнтованого програмування починається з побудови моделі [11].

В основному, UML - це мова моделювання для загального призначення в області розробки програмного забезпечення. Тип UML-діаграм зазвичай використовують для заміни блок-схем. Вони задовольняють як більш стандартизований метод моделювання робочих процесів, так і широкий спектр функцій для підвищення зручності читання та ефективності.

UML застосовується:

- керівниками проектів під час управління та розподілом задач і контролю за їх виконанням;
- проєктувальниками при розробці технічних завдань для програмістів;
- бізнес-аналітиками, які досліджують реальну систему;
- програмістами, які реалізують модулі інформаційної системи;

UML діаграми поділяються на структурні, поведінкові ,взаємодії.

Структурні UML Діаграми:

- Діаграма класів (Class Diagram);
- Діаграма об'єктів (Object Diagram);

- Діаграма компонентів (Component Diagram);
- Діаграма композитів (Composite Structural Diagram);
- Діаграма кооперацій (Collaboration);
- Діаграма розгортання (Deployment Diagram);
- Діаграма пакетів (Package Diagram).

Поведінкові UML Діаграми:

- Діаграма діяльності (Activity Diagram);
- Діаграма прецедентів (Use case diagram);
- Діаграма станів (State Machine diagram).

Діаграми UML взаємодії у свою чергу поділяються на:

- Діаграма послідовності (Sequence Diagram);
- Діаграма комунікації (Communication Diagram);
- Діаграма огляду взаємодії (Interaction Overview Diagram);
- Діаграма синхронізації (Timing Diagram).

2.2 Діаграма прецедентів

Діаграма прецедентів представлена на рис. 5. На ній зображено множини послідовностей дій (включаючи варіанти), що виконуються системою для того, щоб актор міг отримати певний результат.

Діаграма прецедентів в UML – це діаграма, на якій зображено відношення між акторами та прецедентами в системі. Також перекладається як діаграма варіантів використання. Такі діаграми описують функціональність, яка буде надаватись користувачам системи, котра проєктується. Представляються шляхом використання прецедентів та акторів, а також відношень між ними.

Набір усіх прецедентів діаграми фактично визначає функціональні вимоги, за допомогою яких може бути сформульоване технічне завдання.

Прецеденти є основним засобом визначення необхідної поведінки системи. Як правило, вони використовуються для опису вимог до системи, тобто,

що має робити система. Основними поняттями, пов'язаними з прецедентами є актори, прецеденти (варіанти використання), та суб'єкт [12].

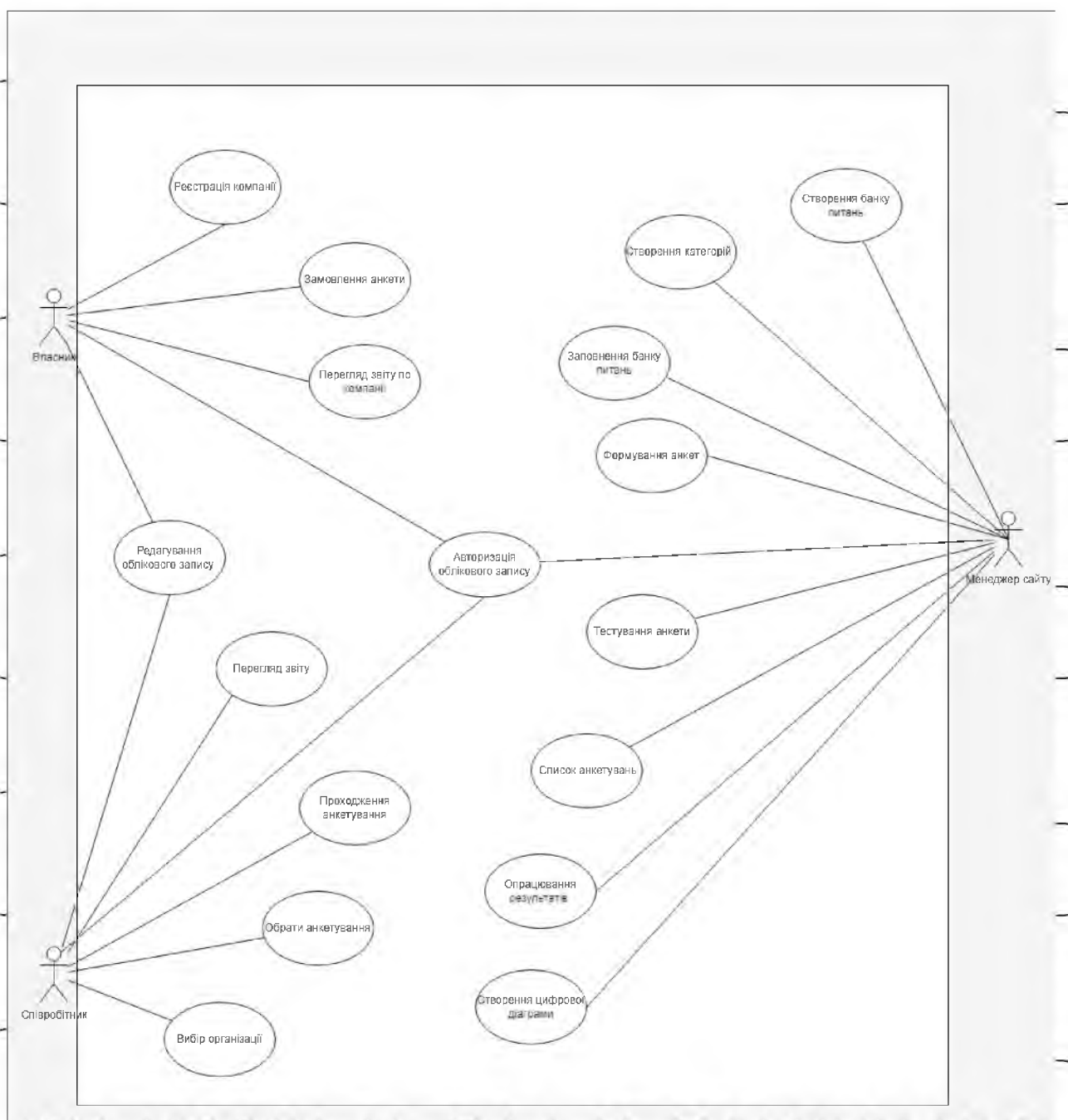


Рис. 5 Діаграма прецедентів

Опис спроектованої діаграми прецедентів представлено нижче.

Актор «Власник компанії» містить такі об'єкти як:

- Регістрація компанії;
- Замовлення анкети;

НУБІП України

- Перегляд звіту по компанії;
- Редагування облікового запису;
- Авторизація облікового запису.

Актор «Співробітник» містить такі об'єкти як:

НУБІП України

- Авторизація облікового запису;
- Редагування облікового запису;
- Проходження анкетування;

- Обрати анкетування;

НУБІП України

- Перегляд звіту;
- Вибір організації.

Актор «Менеджер сайту» містить такі об'єкти як:

- Створення банку питань;

- Створення категорій ;

НУБІП України

- Заповнення банку питань;
- Формування анкет;

- Тестування анкети;

- Список анкетувань;

НУБІП України

- Опрацювання результатів ;
- Створення цифрової діаграми.

В таблиці 1 опис акторів діаграми прецедентів.

Таблиця 1

Опис акторів діаграми прецедентів

Швидкість обробки	
Менеджер сайту	Займається наповненням контенту, платними послугами відповідає за створення ,наповнення, видалення та тестування анкет

Співробітник	Співробітник, клієнт який має доступ до анкетування та отримання індивідуальних результатів та порівняння своїх результатів з іншим членом компанії.
--------------	--

Власник компанії	Створює компанію, має доступ до порівняння результатів будь яких співробітників та загальний результат компанії
------------------	---

Для детального розуміння роботи системи нижче представлено потік подій «Створення анкетування»

Передумови:

Перед початком виконання цього прецеденту, наповнювач контенту має надійти завдання створення нової анкети. Наповнювач контенту має бути зареєстрованим та авторизованим.

Головний потік:

Запускається прецедент формування анкети. Для початку сайт перевіряє чи даний Наповнювач контенту пройшов процедуру авторизації, якщо все успішно буде запропоновано ввести тему для анкети, після введення теми, запит на введення назви анкети, після чого запропонує такі варіанти дій; Створити нове запитання чи обрати питання з банку питань.

Якщо вибрана операція «створити питання», то виконується потік S-1;

Якщо вибрана операція «Обрати питання з банку питань», то виконується потік S-2.

Створена анкета має бути протестована.

Анкета тестується.

Якщо все виконано коректно, то виконується потік S-3;

Якщо анкета не коректна, то виконується потік S-4.

Підпотоки:

S-1: Система дає запит на вибір теми для питання, його тип, та введення самої умови питання. Після цього вводять варіанти відповідей, та правильна відповідь.

Потім питання зберігається в банк питань, та добавляється в анкету.

S-2 Система дає запит на вибір запитання з банку питань.

S-3 Анкета залишається на сайті

S-4 Анкета видаляється

2.3 Діаграма послідовності

Діаграма послідовностей, відповідно до UML, являє собою співпрацю об'єктів і використовується для визначення послідовностей подій між об'єктами для певного результату. Діаграму послідовностей ми можемо побачити на Рис.6.

Діаграма послідовностей є важливим компонентом, який використовується в процесах, пов'язаних з аналізом, проєктуванням та документацією. Її використовують для уточнення діаграм прецедентів — описує поведінкові аспекти системи. Діаграма послідовності відображає взаємодію об'єктів в динаміці, в часі. При цьому інформація набуває вигляду повідомлень, а взаємодія об'єктів передбачає обмін цими повідомленнями в рамках сценарію [13].

Діаграма послідовностей, що зображена на Рис.6, містить такі об'єкти:

- «Система»;
- «Наповнювач контенту»;
- «Менеджер сайту».

Об'єкт «Наповнювач контенту» надсилає завдання на створення анкети об'єкту «Система». Після чого отримує запит на введення назви анкети .

Надсилається назва та отримується запит на введення опису анкети «Наповнювачу контенту». «Система» отримує опис анкети та відправляє запит на створення або вибору тесту. «Наповнювач контенту» заповнює анкету запитаннями та відправляє запит на тестування «Менеджер сайту». Тестування проходить «Менеджер сайту» та має вибір, залишити анкету якщо все добре, та видалити якщо йому щось не сподобається.

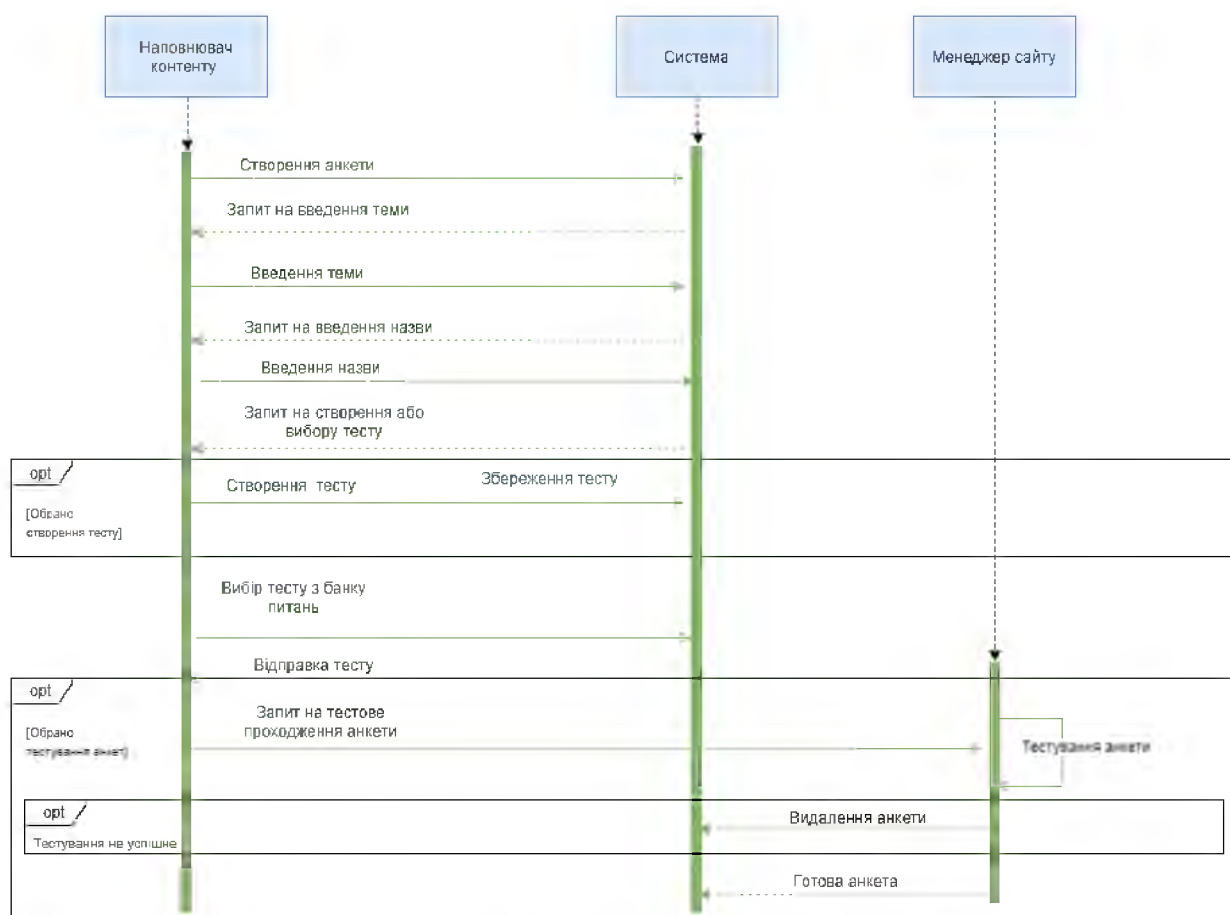


Рис. 6 Діаграма послідовностей

2.4 Діаграма активності

Діаграма активності, представлена на Рис. 7.

Діаграма діяльності використовується для моделювання послідовного робочого потоку великої діяльності, орієнтуючись на послідовності дій та відповідні умови ініціювання дії. Стам діяльності стосується виконання кожного кроку робочого процесу.

Діаграма діяльності представлена фігурами, з'єднаними стрілками.

Стрілки проходять від початку діяльності до завершення і представляють послідовний порядок виконуваних дій. Чорні кола представляють початковий стан робочого процесу. Обведений чорним колом позначає кінцевий стан.

Закруглені прямокутники представляють виконані дії, які описуються текстом всередині кожного прямокутника.

Діамантова форма використовується для представлення рішення, що є ключовою концепцією діаграми діяльності. Після завершення діяльності повинен бути обраний перехід (або набір послідовних дій) з набору альтернативних переходів для всіх випадків використання [14]

Головна сторінка сайту

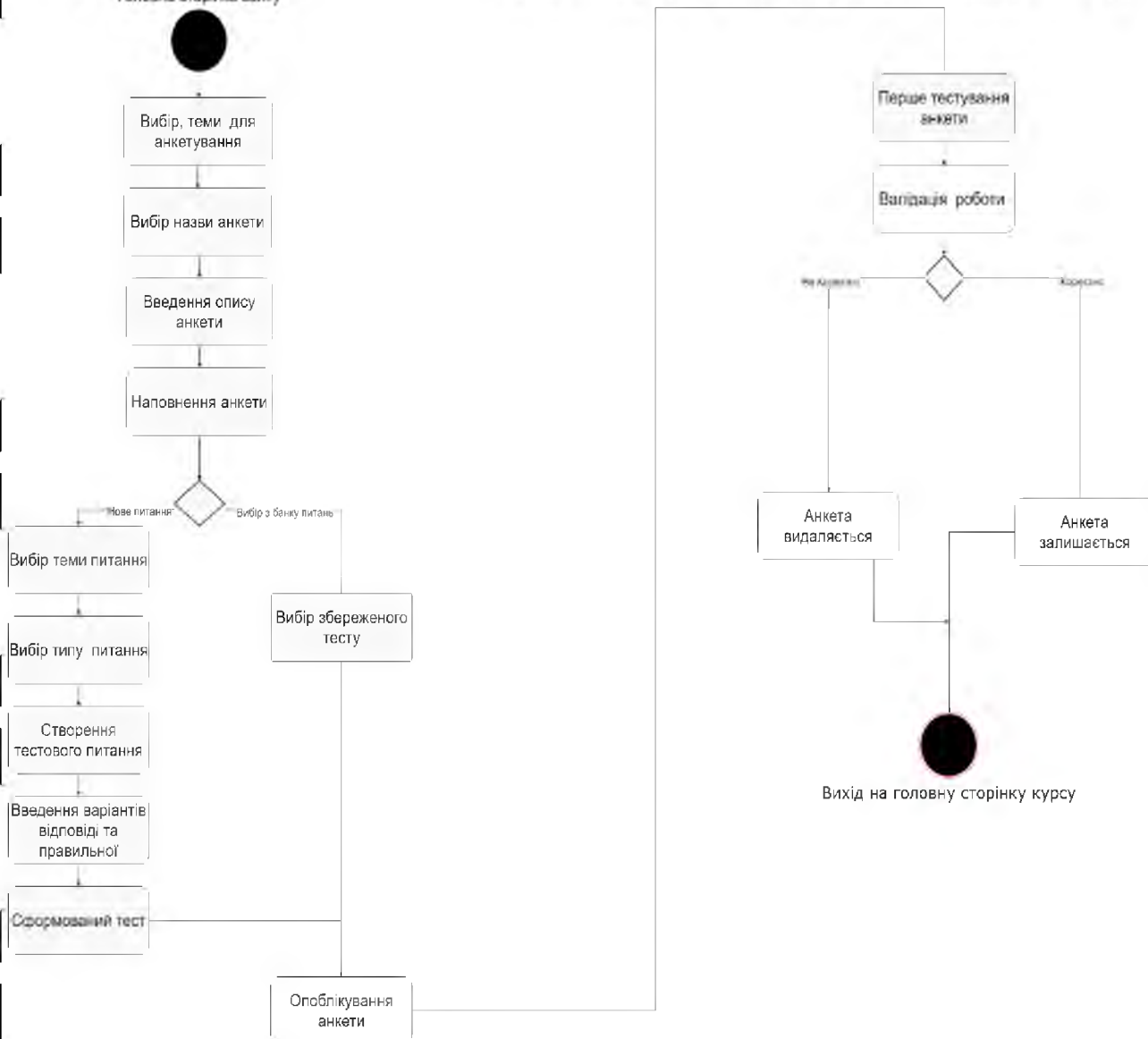


Рис. 7. Діаграма активності

3 РОЗРОБКА СИСТЕМИ

НУБІП України

3.1 Логічна модель даних

ERwin є одним з CASE засобів, що дозволяють моделювати бізнес процеси. Він відноситься до категорії I - CASE. ERwin забезпечує інтеграцію моделей верхнього рівня з моделями нижнього рівня. Моделі верхнього рівня розробляються на початкових стадіях проектування інформаційних систем. Моделі нижнього рівня розробляються на етапі створення програмного коду і тестування.

Основне призначення ERwin data modeler це моделювання даних, тому з точки зору моделювання бізнес процесів його застосування обмежене. Він дозволяє створювати моделі даних високого рівня, виконувати автоматичне перетворення цих моделей в моделі низького рівня, генерувати схеми бази даних і опис даних на рівні програмного коду.

В якості стандартів моделювання в ERwin data modeler застосовуються стандарти IDEF1X і IE (Information Engineering), засновані на діаграмах «сутність-зв'язок». Основні види моделей бізнес процесів, які можна створити за допомогою ERwin data modeler це - діаграма потоку даних, модель розподілу, модель подія/стан.

Моделі даних, які створюються допомогою ERwin data modeler, поділяються на дві категорії:

- *логічні моделі*. Ці моделі призначені для представлення елементів в термінах бізнес процесів. Для моделювання застосовуються діаграма "сутність-зв'язок", КВ модель (модель, заснована на ключах) і FA модель (повністю певна модель);

- *фізичні моделі*. Ці моделі призначені для створення бази даних інформаційної системи. Вони забезпечують переклад понять бізнес процесів на рівень понять ІТ. Для моделювання застосовуються два види моделей - трансформаційна модель і схематична модель бази даних (DBMS модель) [15].

Для побудови логічної моделі була використана ER-модель. Відповідно цій моделі інформаційне забезпечення представлено у вигляді сутностей, кожна з яких має атрибути. Крім, того сутності пов'язані між собою.

Створена логічна модель бази даних системи рис.8.

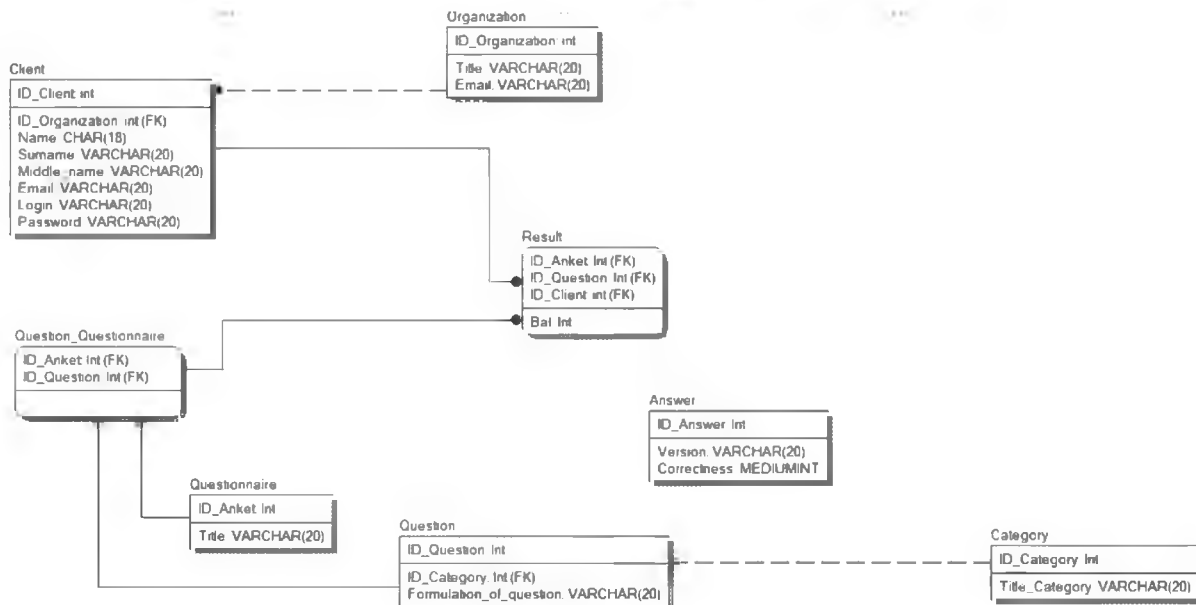


Рис.8 Логічна модель

• Клієнт «Client» описує інформацію про клієнта. Основні атрибути: ПІБ, email, логін, пароль, унікальний ідентифікатор користувача;

- Організація «Organization» описує інформацію про організацію.

Основні атрибути: унікальний ідентифікатор компанії, назва компанії, email;

• Результат «Result» описує спільну інформацію про анкету, клієнта та кількість балів зароблених користувачем під час проходження анкетування;

Основні атрибути: унікальний ідентифікатор анкети, унікальний ідентифікатор користувача, унікальний ідентифікатор питання.

• Запитання-Анкета «Question Questionnaire» проміжна сутність що зберігає ідентифікатор анкети та питання. Основні атрибути: унікальний ідентифікатор анкети, унікальний ідентифікатор питання;

• Анкета «Questionnaire» описує інформацію про анкету. Основні атрибути: унікальний ідентифікатор анкети, назва анкети;

• Запитання «Question» описує інформацію про запитання в анкеті.
Основні атрибути: унікальний ідентифікатор питання, унікальний ідентифікатор категорії, опис запитання;

- Відповіді «Answer» описує інформацію кількість балів за відповідь.

Основні атрибути: унікальний ідентифікатор відповіді, варіанти відповідей, кількість балів за відповідь;

• Категорія «Category» описує інформацію про категорії питань.
Основні атрибути: унікальний ідентифікатор

- категорії, назва категорії.

3.2 Проектування сховище даних

Проблеми розрізненості зберігання даних в рамках одного підприємства, необхідність залучення технічних фахівців для вилучення з баз даних потрібної для прийняття рішень інформації привели в 1980-і рр. до ідеї централізованого зберігання даних, необхідних для подальшого аналізу. Виник термін «сховище даних»[16].

Сховища даних є спеціалізовані бази даних, що володіють такими властивостями:

• *предметна орієнтованість*. У сховищі містяться дані, всебічно описують певну предметну область;

• *інтегрованість*. Дані збираються з безлічі різних джерел, узагальнюються і зберігаються в єдиному корпоративному сховищі;

• *забезпечення несуперечливості даних*. Дані з різних джерел можуть містити дублюючі, суперечливі відомості, тому перед їх завантаженням в сховище вони проходять процедури перевірки, узгодження, доповнення, узагальнення;

• *незмінюваність*. На відміну від баз даних транзакційних систем, в яких оперативні дані можуть редагуватися користувачами, дані в сховищі використовуються виключно в режимі читання і недоступні для коригування;

• *підтримка хронології.* Оскільки для цілей аналізу і прогнозування розвитку предметної області необхідно бачити її показники в динаміці, дані зберігаються в прив'язці в дату і за максимально можливий часовий період;

• *оптимізація під виконання складних аналітичних запитів.* Сховище проектується таким чином, щоб мінімізувати час на формування аналітичної звітності, необхідної для підтримки прийняття рішень для керівників і менеджерів.

У сховищі міститься інформація з різних джерел, яка може мати різну періодичність оновлення, різну структуру, ступінь достовірності, власників даних - відомості про ці характеристики інформації називаються метаданими і зберігаються в репозиторії сховища. У репозиторії можуть також зберігатися бізнес-терміни, правила та алгоритми обчислення показників, які визначені для даного бізнесу. Фізично репозиторій є окремою базою даних або набір таблиць в рамках бази даних сховища.

Сховище може бути реалізовано у вигляді віртуального сховища даних, вітрин даних та глобального сховища даних [16].

Для збереження необхідних даних для подальшого аналізу було спроектоване сховище даних, яке показано на рис. 9

Сд містить таблиці вимірів:

- Question_dim – містить дані про запитання для анкети;
- Anket_dim (є консольною таблицею для таблиці FeedTypeDim) – містить інформацію про анкету;
- Organization_dim – містить інформацію про організацію
- Location_dim – місце знаходження конкретної компанії; (Країна-ОбластьМісто)
- Date_dim – часовий вимір (Рік – місяць).

Факт – це величина (зазвичай числова), яка є предметом аналізу. Таблиця фактів представленого сховища: Analysis_fact, містить інформацію у розрізі господарства, породи, типу корму інформацію про:

- Проходження нових анкет

- Скільки всього анкет

- Середня кількість проходження. Інформація в таблицях вимірів є

відносно постійною, тож дані цих таблиць заповнювалися за допомогою SQL запитів.

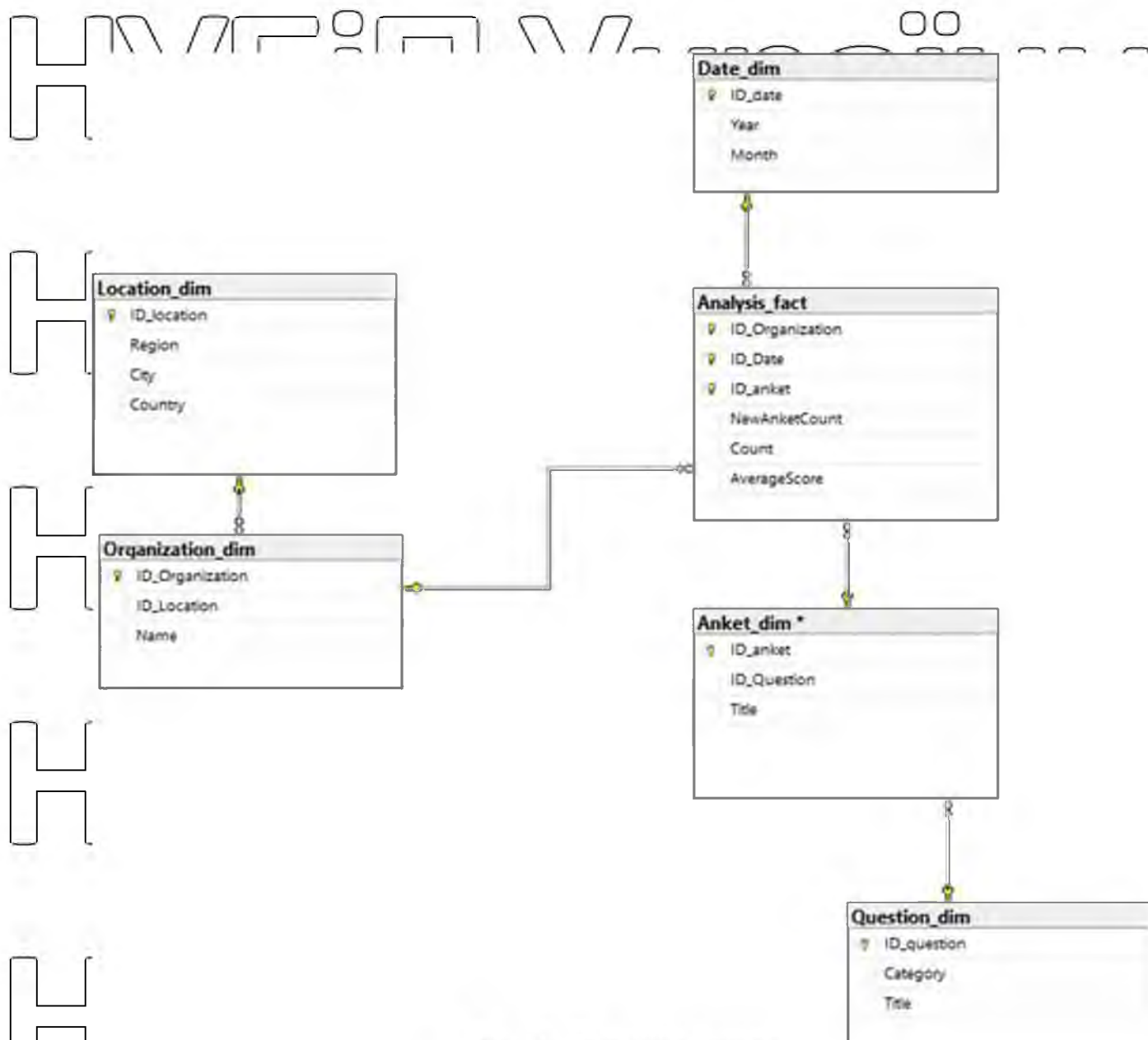


Рис. 9 Сховище даних

НУБІП України

3.3 Побудова розгорнутого куба

Оперативні дані збираються з БД, очищаються, трансформуються і «складаються» в сховище даних.

Сховище – це предметно-орієнтований, інтегрований, незмінний, що підтримує хронологію набору даних, організований для цілей підтримки прийняття рішень. Основне завдання сховища даних полягає в організації єдиного місця збору і зберігання даних, дані в більш зрозумілій для аналізу структурі [17].

Користувач (аналітик) отримує інтуїтивно зрозумілу модель даних, у вигляді багатовимірних кубів. Куб є структурою даних, яка забезпечує можливість швидкого аналізу даних за рамками обмежень реляційних баз даних.

Куби здатні відображати і підсумовувати великі обсяги даних, також надаючи користувачам доступ до будь-яких точках даних з можливістю пошуку. Таким чином, дані можуть бути зведені, фрагментовані і оброблені в міру необхідності для вирішення найбільш широкого спектра питань, що відносяться до сфери використання системи. Аналітик може отримувати зведені (наприклад, по роках) або, навпаки, детальні (по тижнях) відомості та здійснювати інші маніпуляції в процесі аналізу. Інструментом, який забезпечує необхідні для аналізу маніпуляції над даними, є OLAP (Online Analytical Processing, оперативний аналіз даних).

Оперативна аналітична обробка багатьох років використовувалася переважно бізнес-аналітиками й іншими експертами в області обробки даних. Але з недавньої появою Web-OLAP систем, що спрощують упровадження програмного забезпечення і пропонують користувачеві знайомий інтерфейс браузера, компанії стали впроваджувати складні аналітичні можливості в масштабах усього підприємства [18].

Побудова розгорнутого куба підсистеми аналізу роботи підприємства. Розгорнутий куб системи був розроблений за допомогою Visual Studio Analysis

Services

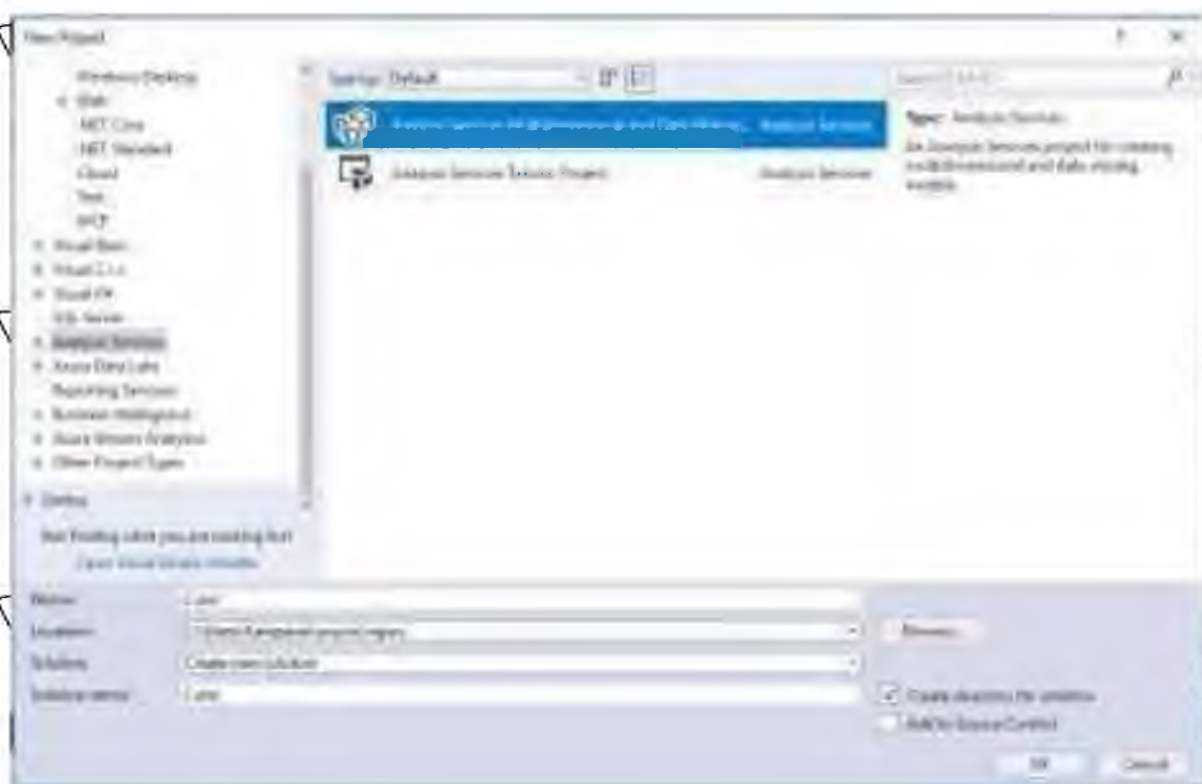
Analysis Services – це засіб аналітичних даних, що використовується в службі підтримки прийняття рішень і бізнес-аналітики.

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

14.

Крок 1. Створення проекту: необхідно обрати тип проекту Analysis Services, та папку збереження проекту (рис.10).

Крок 2. Встановлення зв'язку з джерелом даних: необхідно обрати сервер та спроектоване СД, встановити параметри автентифікації (рис.11).



НУБІП України

НУБІП України

Рис. 10 Створення проекту

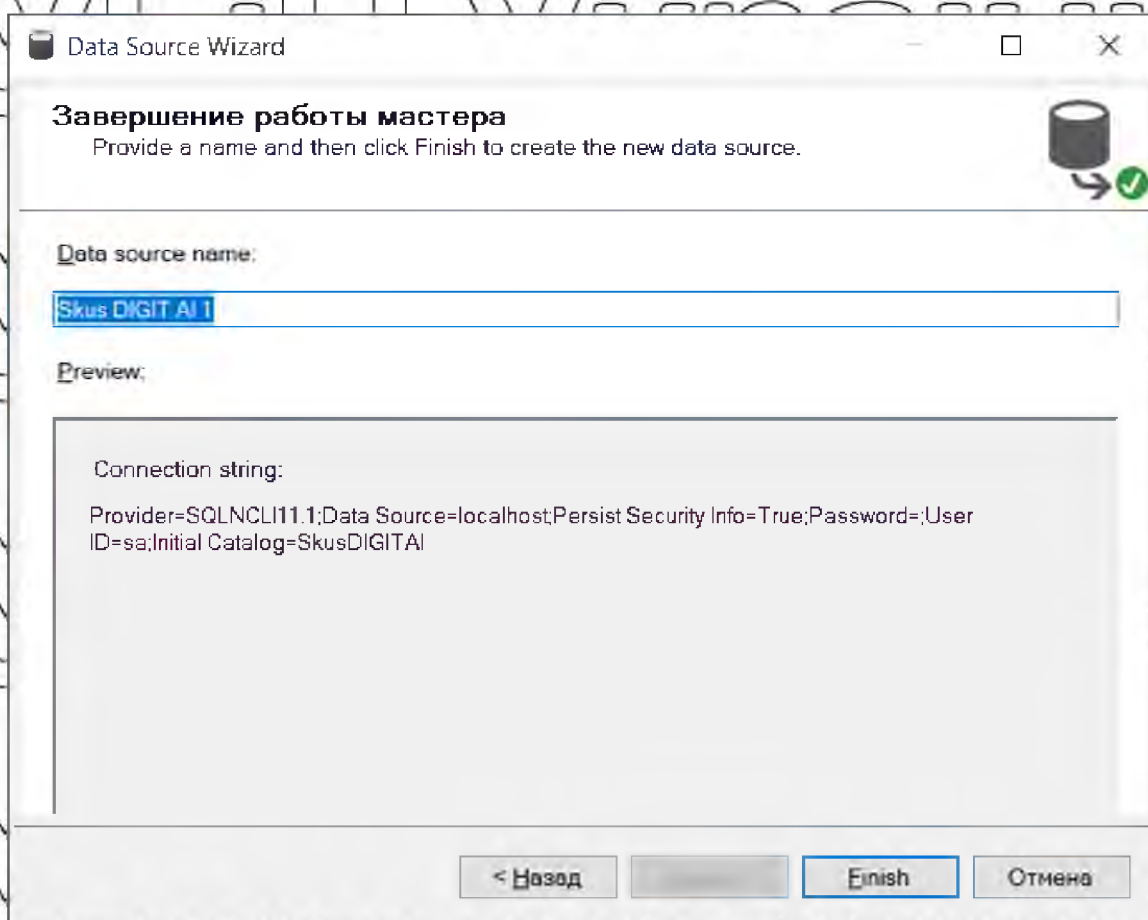


Рис. 11 Результат встановлення зв'язку з джерелом даних

Крок 3. Для створення зрізу джерела даних необхідно обрати джерело даних та необхідні таблиці, що будуть відображатися у зрізі (рис.12).

НУБІП України

НУБІП України

НУБІП України

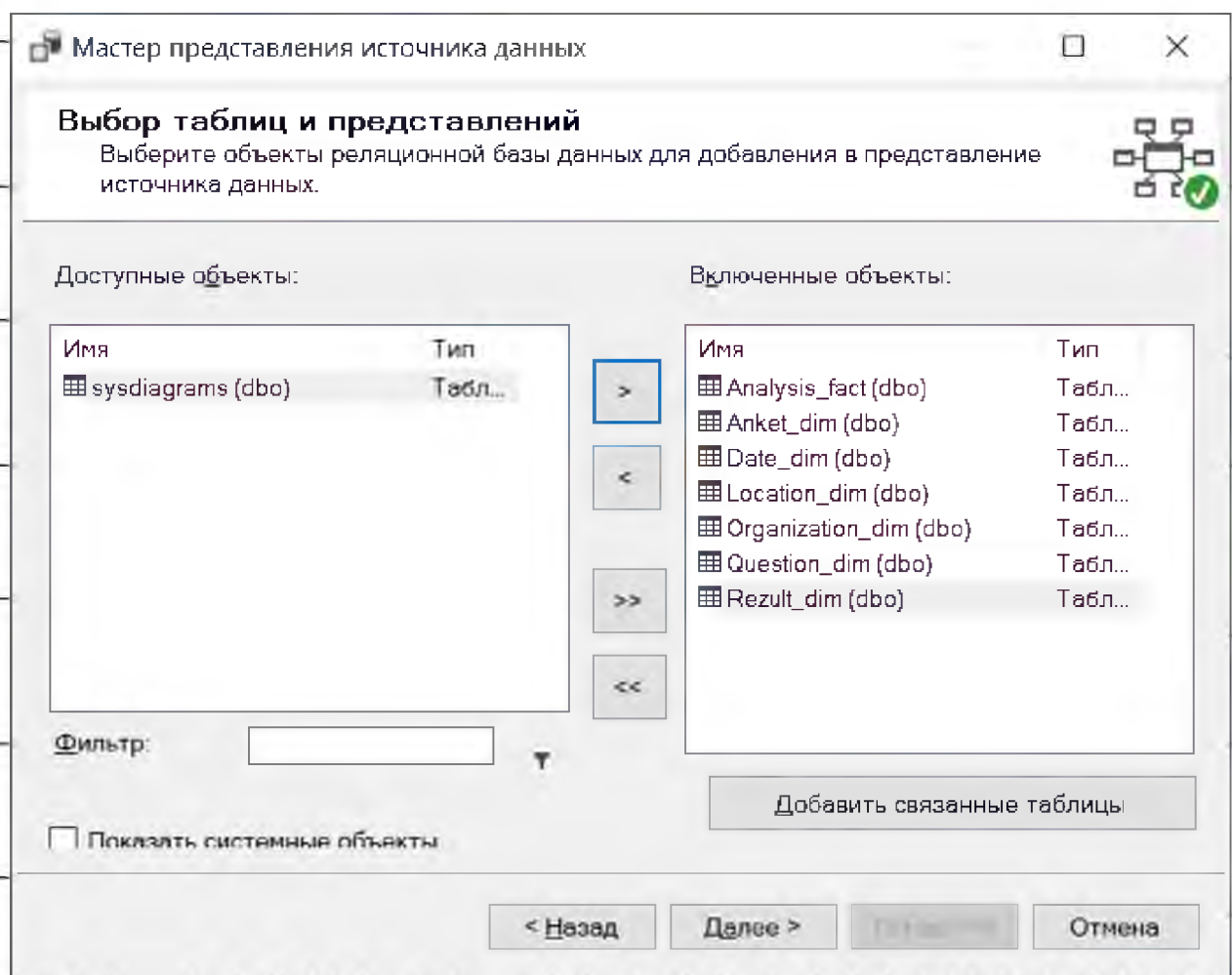


Рис. 12 Створення зрізу джерела та вибір необхідних таблиць

Крок 4. Створення кубу. Для розгортання кубу необхідно обрати таблиці

вимірів та фактів (рис 13)

НУБІП України

НУБІП України

НУБІП України

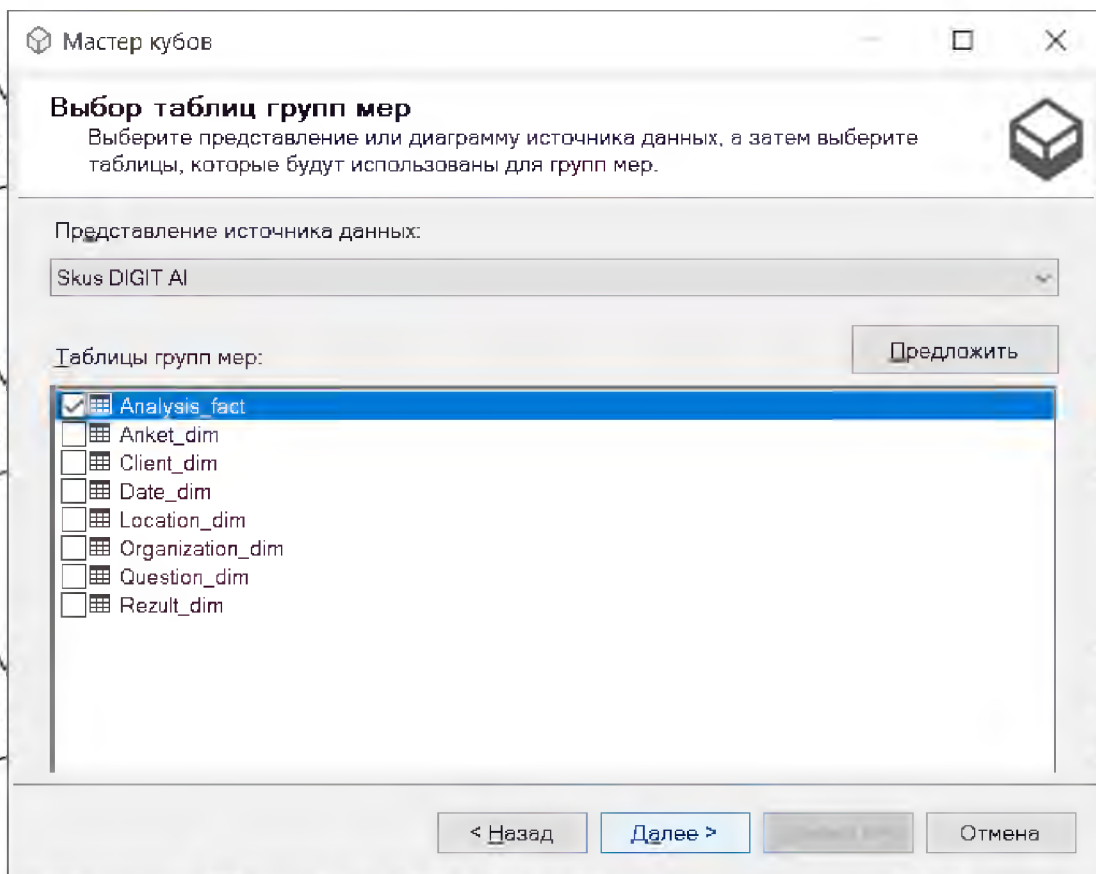


Рис. 13. Вибір таблиць

Результат формування розгорнутого кубу представлено на рис. 14.

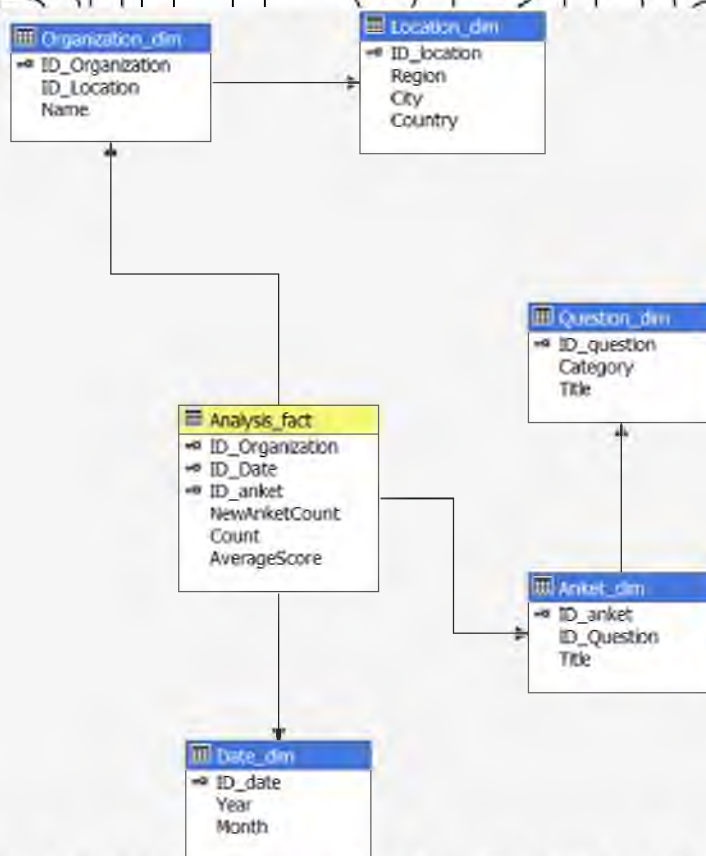


Рис. 14. Сформований куб

3.4 Реалізація міграції даних за допомогою Data Flow

Реалізація отримання даних за допомогою Data Flow. SSIS - це інструмент, який дозволяє в зручному вигляді реалізувати інтеграцію, тобто реалізувати процес перенесення даних з одного джерела в інший.

Для реалізації потоків даних за допомогою Visual Studio BI необхідно створити проєкт Integration Server (рис.15)

Создание проекта

Последние шаблоны проектов

- Integration Services Project
- Мастер проекта сервера отчетов

Поиск шаблонов (ALT+“B”)

Все языки Все платформы Все типы проектов

- Веб-приложение ASP.NET (.NET Framework)**
 Шаблоны проекта для создания приложений ASP.NET. Можно создавать приложения ASP.NET Web Forms, MVC или веб-API ASP.NET, а также добавлять множество других функций ASP.NET.
 Visual Basic Windows Облако Веб
- Облачная служба Azure (классическая)**
 Проект для создания масштабируемой службы, выполняющейся в Microsoft Azure.
 C# Azure Облако
- Новое решение**
 Создание пустого решения, не содержащего проектов
 Другое
- Консольное приложение (.NET Framework)**
 Проект приложения для командной строки
 C# Windows Консоль
- Веб-приложение ASP.NET (.NET Framework)**
 Шаблоны проекта для создания приложений ASP.NET. Можно создавать

Далее

Рис. 15 Створення проєкту

Для організації потоків даних необхідно підключитися до джерела даних та приймача даних. Налаштування підключення до джерела та приймача показано на рис.16.

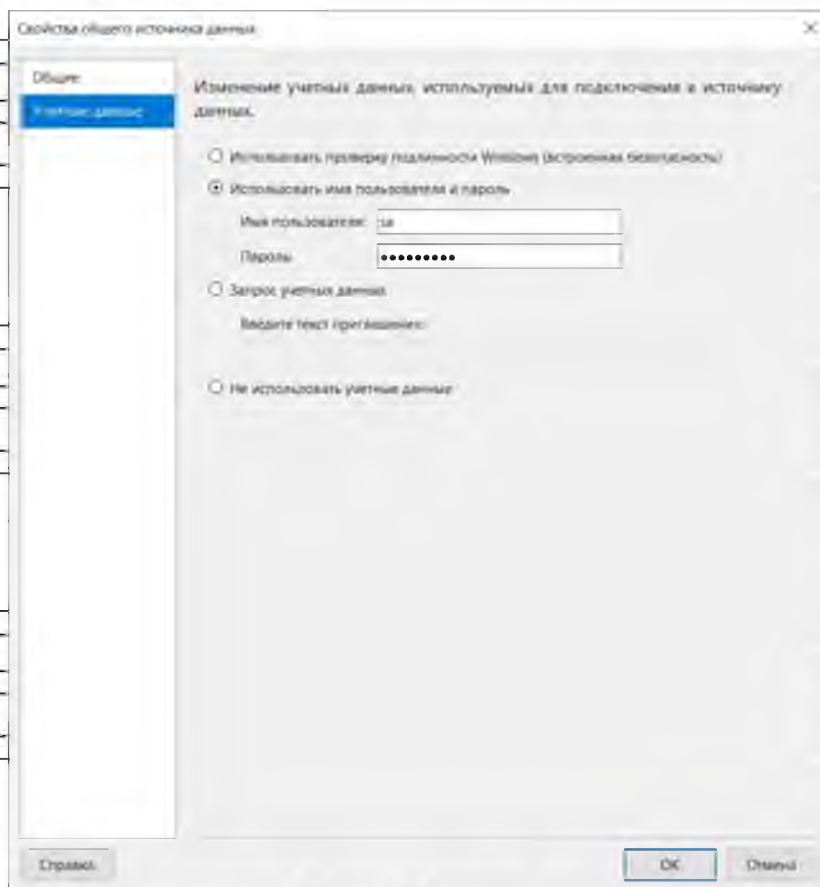


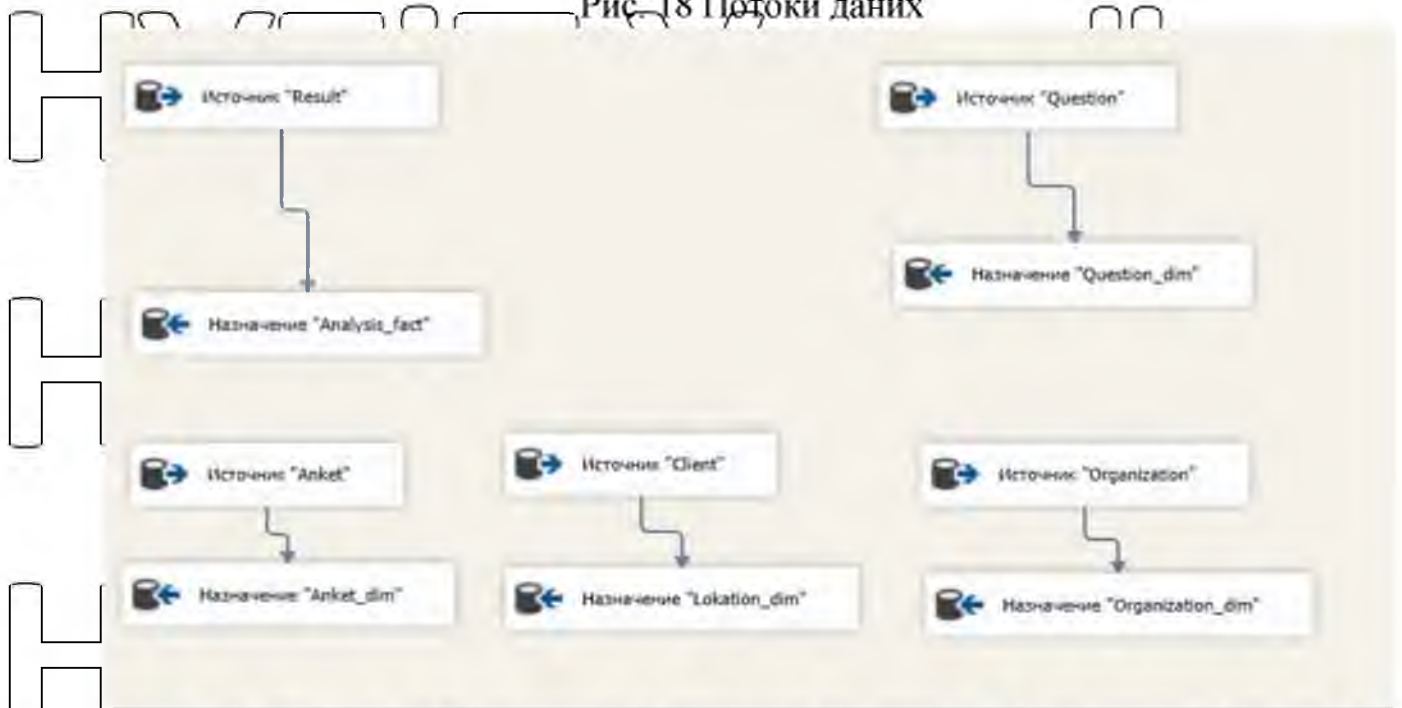
Рис. 16 Підключення до джерела та приймача даних

Щоб уникнути помилок при транспортуванні даних, які пов'язані зі зв'язками між таблицями потоки даних організовані в декілька рівнів (рис. 17-19).



Рис. 17 Потік задач

Рис. 18 Подоки даних



ID_Organiz...	ID_Date	ID_anket	NewAnket...	Count	AverageSc...
1	1	1	8	10	12
2	2	2	7	20	9
3	3	3	16	30	12
4	5	4	4	12	10
4	4	4	12	25	13
4	6	4	10	20	15
4	7	4	13	14	14
4	8	4	7	18	2
4	9	4	1	15	9
4	10	4	2	17	8
4	11	4	8	19	12
3	6	3	1	15	16
3	4	3	6	4	3

Рис. 19 Результат транспортування даних

НУБІП України

НУБІП України

4 РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

4.1 Побудова звітності засобами SSRS

Інформація для керівників може передаватися фрагментами, які можуть бути розрізнені і неповні. Організація аналітичної звітності дозволяє компаніям отримувати цінну інформацію, яка дозволить бізнесу ефективно розвиватися та уникнути різноманітних ризиків.

Бізнес аналітика – можливість приймати рішення, спираючись на інформацію отриману за даними бізнес-процесів та обліку.

Переваги автоматизації звітності описані нижче.

Значно скорочується час на збирання необхідної інформації, аналізу та прийняття рішень. Це дуже важливо, оскільки зволікання може призвести до зниження прибутку чи збитків.

Збільшення ефективності у бізнесі. Персональні звіти – це найкраща мотивація у роботі, керівник може за результатами звіту оцінити ефективність працівника.

Візуалізація бізнес-процесів. Створення графіків та діаграм, які спрощують сприйняття інформації.

Зручний інструмент для створення різноманітних звітів. Попередня підготовка звіту, що настроюється під запити користувача. Можливість пошуку помилок, які з'являються під час ручного введення даних.

Відповіді на будь-які питання. Ми готові надавати технічну підтримку програмного забезпечення весь період його використання. Вам не потрібно залучати співробітників до штату [19].

Для аналізу роботи системи використовуємо Power BI, за допомогою якою будуємо звіти та візуалізовані дані.

Побудований звіт по середньому проходженні анкет в певній організації за 2021 рік представлений на рис.20.

Опитана кількість співробітників у розрізі компанії

Компанія : **АТВ Васильків**За Анкетуванням: **Пошук**

Рік	2021
Місяць :02	15
Місяць :03	10
Місяць :04	12
Місяць :05	8
Місяць :06	13
Місяць :09	9
Місяць :11	2
Місяць :12	14

Закордонні діаграми

■ АТВ Васильків



Рис. 20 Звіт по середньому проходженні анкет в певній організації за 2021 рік

Будуємо наступний звіт щодо проходження нових анкетувань по компаніях. Вигляд звіту в режимі дизайнера представлений на рис. 21.

Кількість опитаних співробітників

Компанії : [Name]

Рік : [Year]

Місяць : [Month] [Sum(Expr1)]

Закордонні діаграми

■ Name A ■ Name B

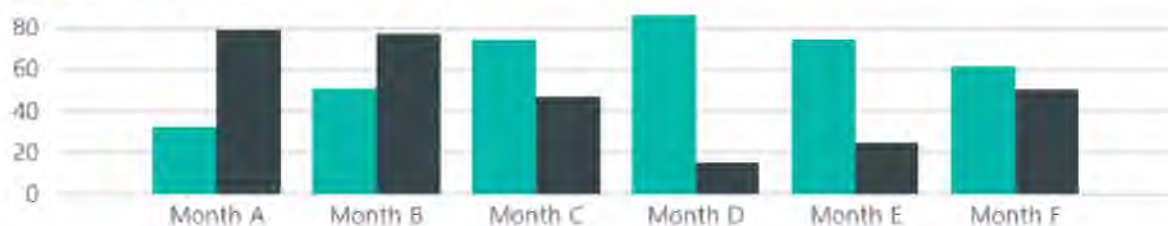


Рис. 21 Режим дизайнера звіту

Звіт, щодо проходження нових анкетувань по компаніях представлений на рис. 22.

Кількість опитаних співробітників

Компанії : АТВ Васильків

Рік : 2021

Місяць: 02	10
Місяць: 03	4
Місяць: 04	8
Місяць: 05	2
Місяць: 06	12
Місяць: 09	1
Місяць: 11	7
Місяць: 12	13

Рис. 22 Звіт по проходження нових анкетувань по компанії АТВ

Спроектована гістограма проходження нових анкетувань по компаніях представлена на рис. 23.

■ АТВ Васильків ■ Comfy Умань ■ EpicentK Київ №1 ■ Ашан Київ №1



Рис. 23 Гістограма проходження нових анкетувань по компаніях

4.2 Розрахунок КРІ

За допомогою Visual Studio BI були розраховані ключові показники ефективності (КРІ).

КРІ застосовується для обробки інформації про задану мету, записаної в кубі фактичної формули продуктивності (результати вимірювань, які показують

тренд і стан продуктивності.

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

Для визначення ефективності середньої кількості проходжень були здійснені наступні кроки:

- 1) Задано ім'я ключового показника ефективності: KPI_AVG_SCORE.
- 2) Визначення значення (рис. 24), отримуємо його з таблиці фактів.
- 3) Визначення цілі виразу.
- 4) Визначення ознаки та виразу стану, що представлені на рис. 25.
- 5) Обрахунок KPI (рис. 26).

Выражение значения

```
[Measures].[Count] / [Measures].[Число Analysis Fact]
```

Проблемы не найдены.

Стр: 1 Симв: 54 Пробелы CRLF

Рис. 24 Фактичне значення продуктивності

Состояние

Признак состояния:

Светофор

Выражение состояния:

```
CASE
WHEN KPIVALUE( "KPI_COUNT" ) > KPIGOAL("KPI_COUNT") THEN 1
WHEN KPIVALUE( "KPI_COUNT" ) <= KPIGOAL("KPI_COUNT") AND
KPIVALUE( "KPI_COUNT" ) > 5 THEN 0
```

Проблемы не найдены.

Стр: 6 Симв: 5 Пробелы CRLF

Рис. 25 Значення ознаки та виразу стану

Отобразить структуру	Значение	Цель	Состояние
KPI_AVG_SCORE	11,5	10	
KPI_COUNT	10,75	15	
KPI_NEW_COUNT	10,75	10	

Рис. 26 Результати обчислення KPI

Для визначення ефективності проходжень нових анкет були здійснені наступні кроки:

- 1) Задано ім'я ключового показника ефективності: KPI_NEW_COUNT.
- 2) Визначення значення (рис. 27), отримуємо його з таблиці фактів.
- 3) Визначення цілі виразу: 10 (рис. 28).
- 4) Визначення ознаки та виразу стану, що представлені на рис. 29.

5) Обрахунок KPI (рис. 26).

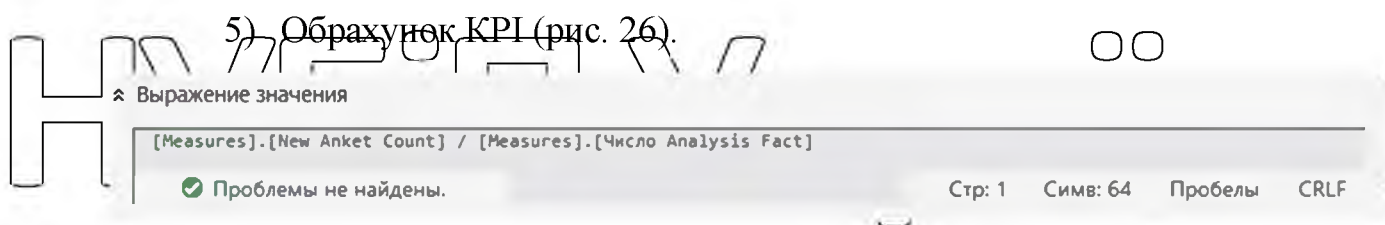


Рис. 27 Фактичне значення продуктивності

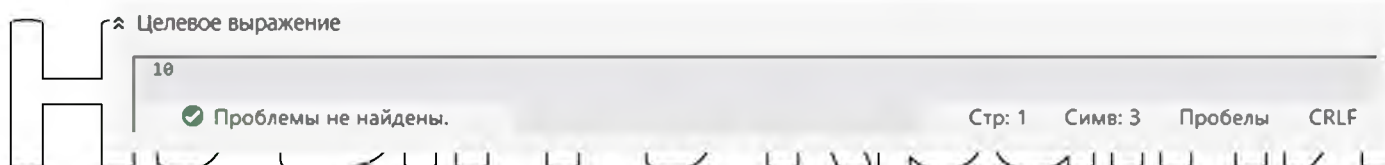


Рис. 28 Цільове значення продуктивності

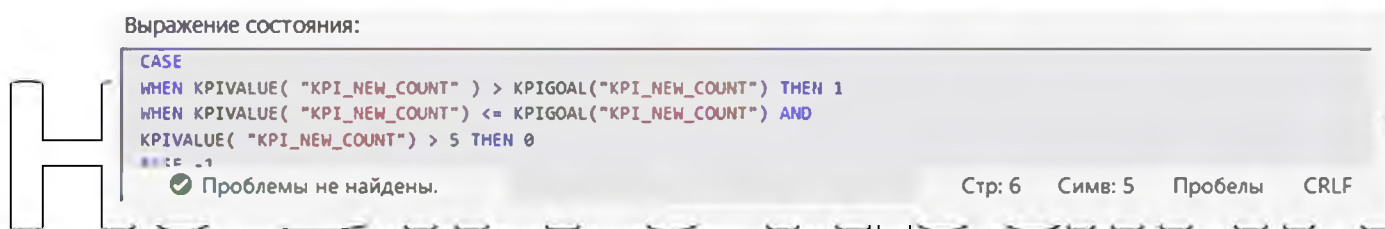


Рис. 29 Значення стану

В результаті обрахунків KPI видно, що в дослідженнях показників досягнуті поставлені цілі, про це свідчить зелене забарвлення світлофорів.

4.3 Інтелектуальний аналіз даних

Data Mining — це сукупність багатьох різних методів здобування знань. Вибір методу часто залежить від типу наявних даних і від того, яку інформацію потрібно дістати [20].

До найпоширеніших методів можна віднести такі:

об'єднання (association; іноді вживають термін *affinity*, що означає подібність, структурну близькість) — високремлення структур, що повторюються в часовій послідовності. Цей метод визначає правила, за якими можна встановити, що один набір елементів корелює з іншим. Користуючись ним, аналізують ринковий кошик пакетів продуктів, розробляють каталоги, здійснюють перехресний маркетинг тощо;

аналіз часових рядів (sequence-based analysis, або sequential association) дає змогу відшукувати часові закономірності між даними (транзакціями).

Наприклад, можна відповісти на запитання: купівля яких товарів передуює купівлі даного виду продукції? Метод застосовується, коли йдеться про аналіз цільових ринків, керування гнучкістю цін або циклом роботи із замовником (Customer Lifecycle Management);

кластеризація (clustering) — групування записів, що мають однакові характеристики, наприклад за близькістю значень полів у БД. Використовується для сегментування ринку та замовників. Можуть залучатися статистичні методи або нейромережі. Кластеризація часто розглядається як перший необхідний крок для подальшого аналізу даних;

класифікація (classification) — віднесення запису до одного із заздалегідь визначених класів, наприклад під час оцінювання ризиків, пов'язаних із видачею кредиту;

оцінювання (estimation);

нечітка логіка (fuzzy logic);

статистичні методи, що дають змогу знаходити криву, найближче розміщену до набору точок даних;

генетичні алгоритми (genetic algorithms);

фрактальні перетворення (fractal-based transforms);

нейронні мережі (neural networks) — дані пропускаються через шари вузлів, «навчених» розпізнавати ті чи інші структури — використовуються для аналізу переваг і цільових ринків, а також для приваблювання замовників [20].

До DM можна віднести ще візуалізацію даних — побудову графічного образу даних, що допомагає у процесі загального аналізу даних вбачати аномалії, структури, тренди. Частково до DM примикають дерева рішень і паралельні бази даних [20].

DM тісно пов'язана (інтегрована) зі сховищами даних (Data Warehousing, DW), які, можна сказати, забезпечують роботу Data Mining [20].

Data Mining — міждисциплінарна технологія, що виникла й розвивається на базі досягнень прикладної статистики, розпізнавання образів, методів штучного інтелекту, теорії баз даних тощо (рис. 30). Звідси й численні методи та

алгоритми, реалізовані в різних дійових системах Data Mining. Багато з таких систем інтегрують у собі відразу кілька підходів. Проте, як правило, у кожній системі присутній певний ключовий компонент, на який робиться головна ставка [20].

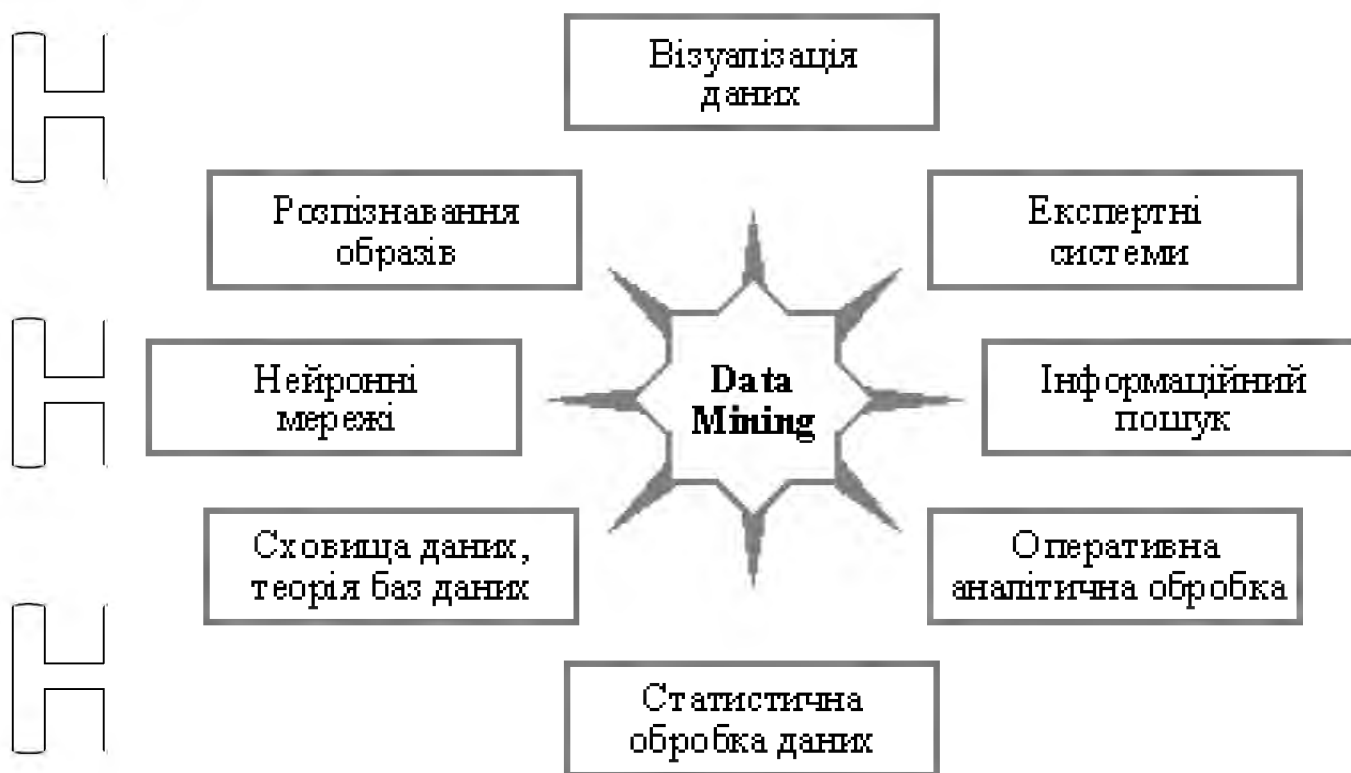


Рис. 30. Взаємозв'язок Data Mining з дисциплінами

Для побудови розглянутих моделей використовуються різні методи, й алгоритми Data Mining [20].

Більшість аналітичних методів, що використовуються у технології Data Mining – це математичні алгоритми і методи. Зокрема до методів і алгоритмів Data Mining відносять штучні нейронні мережі

деревця рішень

символьні правила;

метод найближчого сусіда і k-найближчого сусіда;

метод опорних векторів;

байєсовські мережі;

лінійну регресію;

кореляційно-регресійний аналіз;
ієрархічні методи кластерного аналізу;
неієрархічні методи кластерного аналізу;
зокрема алгоритми /с-середніх і к-медіа;

методи пошуку асоціативних правил;
метод обмеженого перебору;
еволюційне програмування і генетичні алгоритми;
різноманітні методи візуалізації даних [21].

До базових методів Data Mining належать також підходи, що використовують елементи теорії статистики. Основна їх ідея зводиться до кореляційного, регресійного та інших видів статистичного аналізу. Основним недоліком їх є усереднювання значень, що призводить до втрати інформативності даних. Це у свою чергу спричинює зменшення кількості знань, що здобуваються [21].

Основним способом дослідження задач аналізу даних є їх відображення формалізованою мовою та їх подальший аналіз за допомогою моделі [21].

Алгоритм 1R

Алгоритм 1R (OneR) — простий алгоритм формування правил для класифікації об'єкта, цей алгоритм будує правила за значенням тільки однієї незалежної змінної. Для кожного значення кожної незалежної змінної будуються правила, для даного правила обчислюється помилка — кількість об'єктів з тим самим значенням незалежної змінної, але не відповідних тому значенням залежної змінної, яка найчастіше зустрічається для даного значення незалежної змінної. У підсумку вибирається змінна по якій можна з найбільшою точністю класифікувати об'єкти. Цей алгоритм вважається найпростішим серед алгоритмів кластерного аналізу [22].

Реалізація алгоритму 1-Rule

Для реалізації алгоритму 1-Rule було виділено 2 класи високого ("HIGH") та низького ("LOW") рівня проходження тестування, показник яких визначається у порівнянні з середнім показником за весь період дослідження.

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

Спочатку проведено вибірку даних, де визначено середню кількість у розмірі 9.2 та клас проходжень за кожним підприємством (рис. 31).

	Name	Title	countFact	avgCount	resultClassification
1	АТВ Васильків	Безпека	2	9,2	LOW
2	АТВ Васильків	Безпека	8	9,2	LOW
3	АТВ Васильків	Безпека	10	9,2	HIGH
4	АТВ Васильків	Безпека	15	9,2	HIGH
5	АТВ Васильків	Безпека	19	9,2	HIGH
6	АТВ Васильків	Безпека	24	9,2	HIGH
7	АТВ Васильків	Безпека	27	9,2	HIGH
8	АТВ Васильків	Безпека	28	9,2	HIGH
9	АТВ Васильків	Безпека	35	9,2	HIGH
10	АТВ Васильків	Безпека	41	9,2	HIGH
11	АТВ Васильків	Безпека	44	9,2	HIGH
12	АТВ Васильків	Безпека	50	9,2	HIGH
13	АТВ Васильків	Вироб...	11	9,2	HIGH
14	АТВ Васильків	Вироб...	20	9,2	HIGH
15	АТВ Васильків	Вироб...	24	9,2	HIGH
16	АТВ Васильків	Вироб...	25	9,2	HIGH
17	АТВ Васильків	Вироб...	27	9,2	HIGH
18	АТВ Васильків	Вироб...	28	9,2	HIGH
19	АТВ Васильків	Вироб...	30	9,2	HIGH
20	АТВ Васильків	Вироб...	35	9,2	HIGH
21	АТВ Васильків	Вироб...	38	9,2	HIGH
22	АТВ Васильків	Вироб...	40	9,2	HIGH
23	АТВ Васильків	Вироб...	41	9,2	HIGH
24	АТВ Васильків	Вироб...	47	9,2	HIGH
25	АТВ Васильків	Пошук	3	9,2	LOW
26	АТВ Васильків	Пошук	6	9,2	LOW

Рис. 31/Вибірка фактів

Наступним етапом було визначення рівня проходження тестування відповідно до тематики опитування відносно середнього показника (рис.32).

	anketTitle	lessThanAvg	biggerThanAvg	allTestCount	avgOrganisCount	avgCount	resultClassification	lessProbability	biggerProbability
1	Спілкування	10	38	48	27	9,2	HIGH	20,83	79,17
2	Безпека	11	37	48	23	9,2	HIGH	22,92	77,08
3	Виробництво	4	44	48	29	9,2	HIGH	8,33	91,67
4	Пошук	11	37	48	23	9,2	HIGH	22,92	77,08

Рис. 32 Класифікація відповідно до тематики опитування

Аналізуючи отримані дані можна виділити наступні правила:

– Якщо тема тестування «Спілкування», то рівень проходження високий ("High") з імовірністю 79,17%

– Якщо тема тестування «Безпека», то рівень проходження високий ("High") з імовірністю 77,08%

– Якщо тема тестування «Виробництво», то рівень проходження високий ("High") з імовірністю 91,67%

– Якщо назва підприємства «Ашан Київ №1», то рівень проходження високий ("High") з імовірністю 81,25%

Можна зробити висновок, що на класифікацію і рівень проходження тестування впливає як підприємство, так і тематика тестування, проте все ж підприємство має більший вплив.

4.3.1 Спрощений алгоритм Байєса

У теорії ймовірностей та статистиці Теорема Байєса (або ж Закон Байєса, чи Правило Байєса) описує ймовірність події, спираючись на обставини, що могли би бути пов'язані з цією подією. Наприклад, припустимо, що хтось цікавиться, чи має рак певна особа, і знає вік цієї особи. Якщо рак пов'язаний з віком, то, застосовуючи теорему Байєса, інформацію про вік осіб можливо використати для точнішої оцінки ймовірності того, що вони мають рак.

При застосуванні, задіяні у теоремі Байєса ймовірності можуть мати різні інтерпретації. В одній із цих інтерпретацій теорема Байєса використовується безпосередньо у певному підході до статистичного висновування. При байєсовій інтерпретації ймовірності ця теорема виражає, як повинна раціонально змінюватися суб'єктивна міра впевненості при врахуванні свідчення. Це є байєсовим висновуванням, що є фундаментальним для байєсової статистики. Тим не менш, теорема Байєса має численні застосування у широкому спектрі обчислень із залученням ймовірностей, а не лише у байєсовому висновуванні [23].

Застосування найвних алгоритмів Байєса

- Прогнозування в режимі реального часу: Naive Bayes – класифікатор, який прагне до навчання, і він, безперечно, швидкий. Таким чином, його можна використовувати для прогнозування в режимі реального часу.

- Багатокласове передбачення: цей алгоритм також добре відомий завдяки функції багатокласового передбачення. Тут ми можемо передбачити ймовірність кількох класів цільової змінної.

- Класифікація тексту/ Фільтрація спаму/ Аналіз настроїв: найвні байєсівські класифікатори, які в основному використовуються в класифікації тексту (через кращий результат у проблемах із кількома класами та правилом незалежності), мають вищий показник успіху в порівнянні з іншими алгоритмами. Як результат, він широко використовується у фільтрації спаму (визначення спаму електронної пошти) та аналізі настроїв (в аналізі соціальних мереж, для визначення позитивних і негативних настроїв клієнтів).

- Система рекомендацій: найвний класифікатор Байєса та спільна фільтрація разом створюють систему рекомендацій, яка використовує методи машинного навчання та інтелекту даних для фільтрації невидимої інформації та прогнозування того, чи потрібний користувач даний ресурс чи ні [24].

На основі розгорнутого кубу створюємо нову структуру інтелектуального аналізу даних. Обираємо алгоритм Байєса.

На основі створеної структури для аналізу було сформовано мережу залежностей (рис. 33).

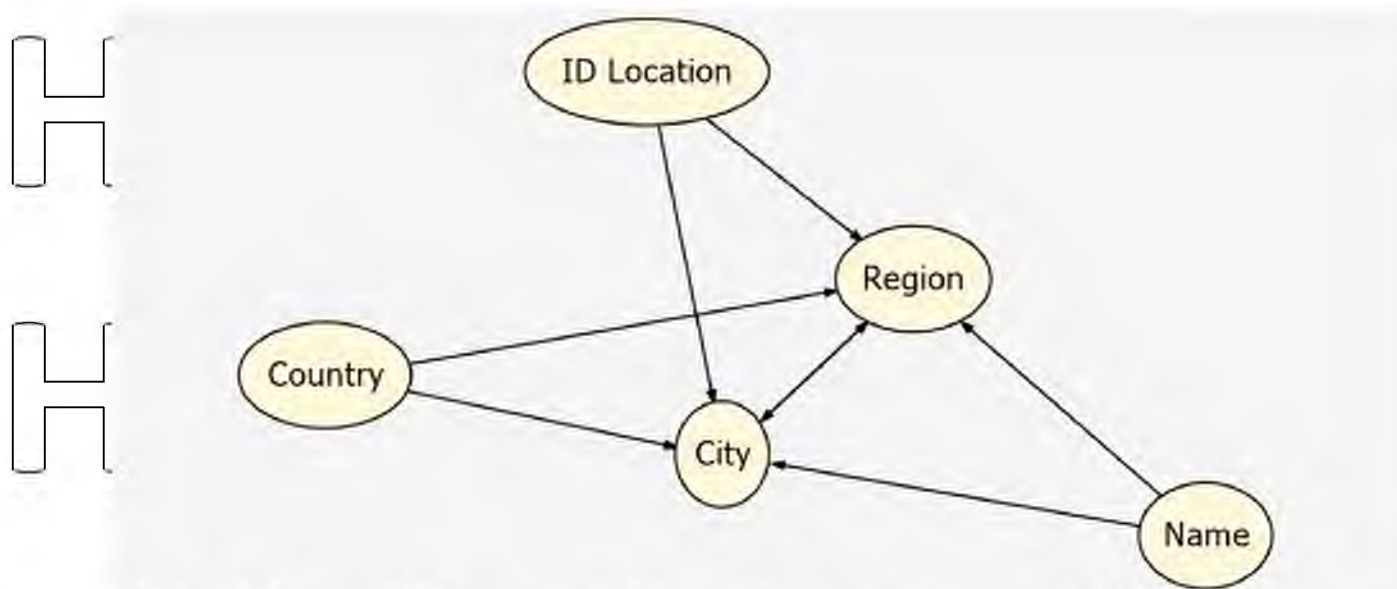


Рис. 33 Розгорнута мережа залежностей

На основі побудованої мережі можна сказати, що регіон та місто залежать від того в якій країні розташовані.

Також можна розглянути більш детальний опис за кожним прогнозованим

значенням.

Нижче на рис. 34 наведено детальний опис профілю атрибуту «Region». На основі проаналізованих даних можна зробити висновок що місто та країна мають достатній вплив.

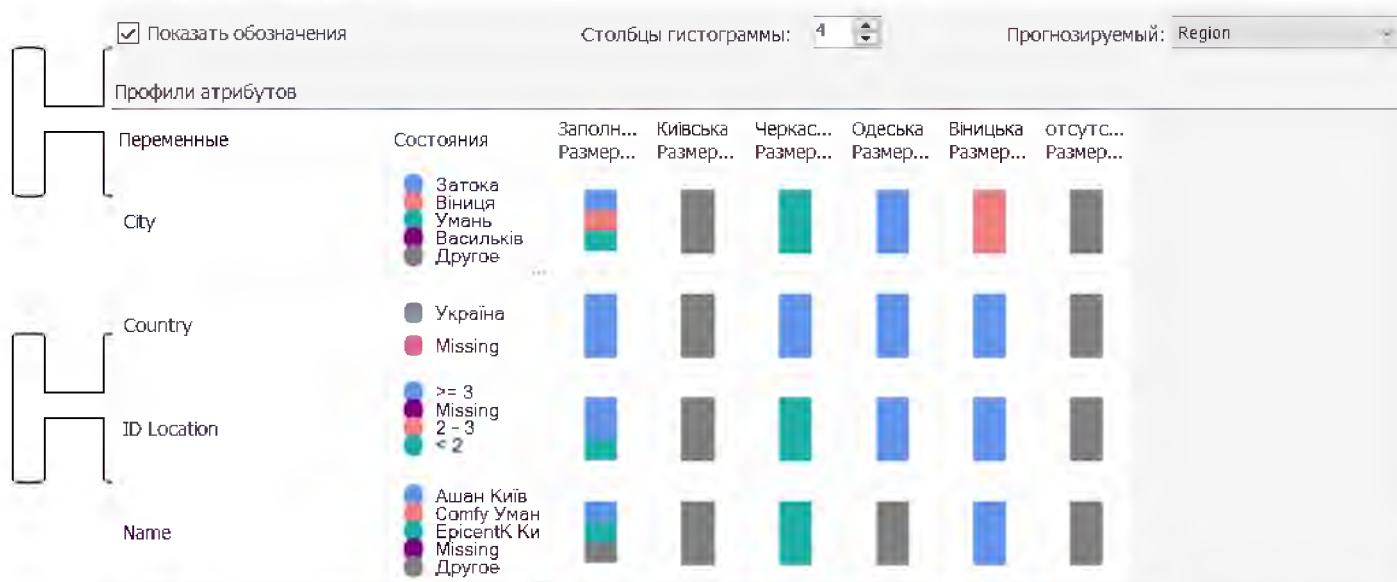


Рис. 34 Профіль атрибуту «Name Culture»

НУБІП України

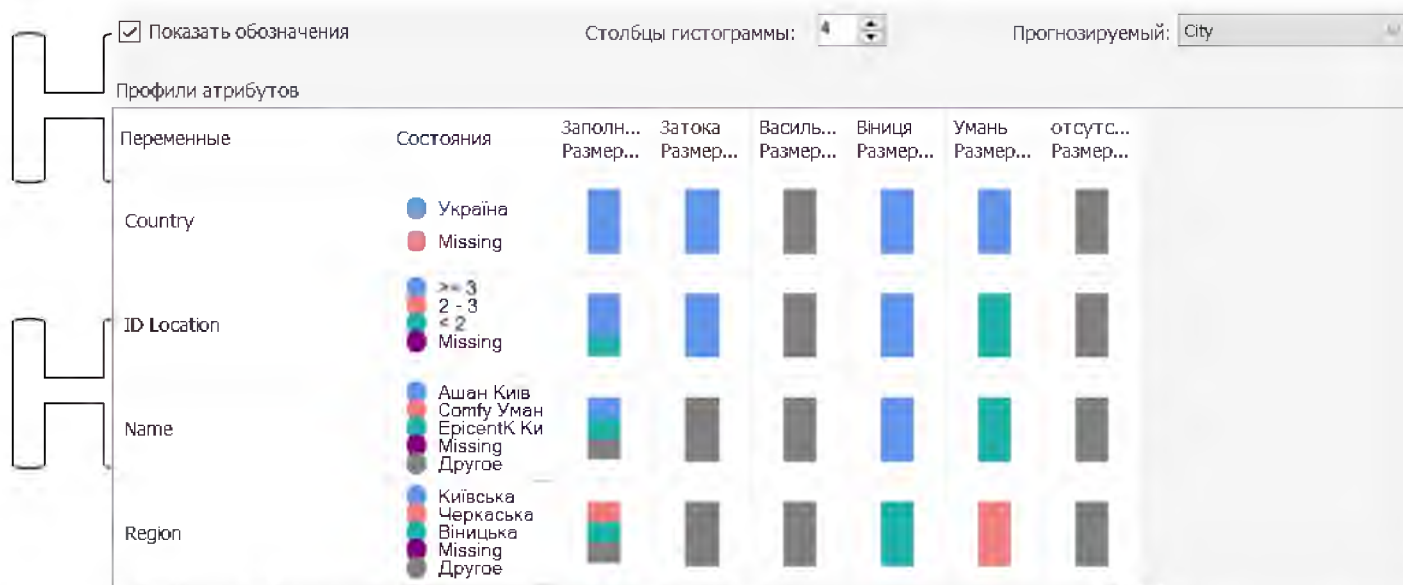


Рис. 35 Профіль атрибуту «City»

4.3.2 Пошук асоціативних правил

Пошук асоціативних правил (Association Rules) – ключова тема в

інтелектуальному аналізі даних. Асоціація має місце в тому випадку, якщо

декілька подій зв'язані одна з одною [25].

Пошук виявляє приховані зв'язки в, на перший погляд, ніяк незв'язаних

даних. Ці зв'язки - правила. Ті, які перевищують певний поріг, вважаються

цікавими. Такі правила дають можливість виконувати дії ґрунтуючись на певних

шаблонів. Вони так само допомагають в прийнятті і поясненні рішень. Як і

більшість методів DATA MINING, даний метод дозволяє перетворити

потенційно величезну кількість інформації в невеликий і зрозумілий набір

статистичних показників. Одним з найчастіше цитованих прикладів пошуку

асоціативних правил служить проблема пошуку стійких зв'язків в корзині

покупця (Market Basket Problem). Проблема пошуку стійких зв'язків в корзині

покупця полягає в тому, аби визначити які товари отримуються покупцями

разом, так, щоб фахівці з маркетингу могли відповідним чином розмістити ці

товари в магазині для підвищення об'єму продажів, і так само прийняти інші

рішення, сприяючі продажам. Деякі правила, що виявляються, можуть бути

тривіальними, наприклад, «покупці, які купують хліб, так само купують і масло».

Інші - цікаві і екстраординарні, наприклад «покупці, які купують дитячі товари, так само купують і пиво» [25].

Пошук асоціативних правил зовсім не тривіальна задача, як може здатися на перший погляд. Одна з проблем - алгоритмічна складність при знаходженні часто зустрічаючих наборів елементів, оскільки із зростанням числа елементів експоненціально зростає число потенційних наборів елементів [25].

Існують різні типи асоціативних правил. У простій формі асоціативні правила повідомляють лише про наявність або відсутність асоціації. Логічна природа таких правил озвучена в їх назві булеві асоціативні правила (Boolean Association Rule).

Правила, які збирають декілька асоціативних правил разом, називаються мультирівневі або узагальнені асоціативні правила (Multilevel or Generalized Association Rules). При побудові таких правил елементи зазвичай групуються згідно ієрархії і пошук ведеться на найвищому концептуальному рівні.

Складнішим типом правил є кількісні асоціативні правила (Quantitative Association Rules). Цей тип правил шукається із застосуванням кількісних (наприклад, ціна) або категоріальних (наприклад, стать) атрибутів

Окрім описаних вище асоціативних правил існують непрямі асоціативні правила, асоціативні правила із запереченням та інші. Не дивлячись на різні типи правил, алгоритм для пошуку асоціативних правил може бути в загальному вигляді розділений на два етапи: пошук найбільш часто зустрічаючихся наборів елементів (large (frequent) itemsets). Набір, що часто зустрічається, - це набір, в якого підтримка перевищує мінімальне значення; генерація правил на основі часто зустрічаючихся наборів [25].

На основі розгорнутого кубу була створена нова структура інтелектуального аналізу даних, використовуючи метод асоціативних правил (рис. 36).

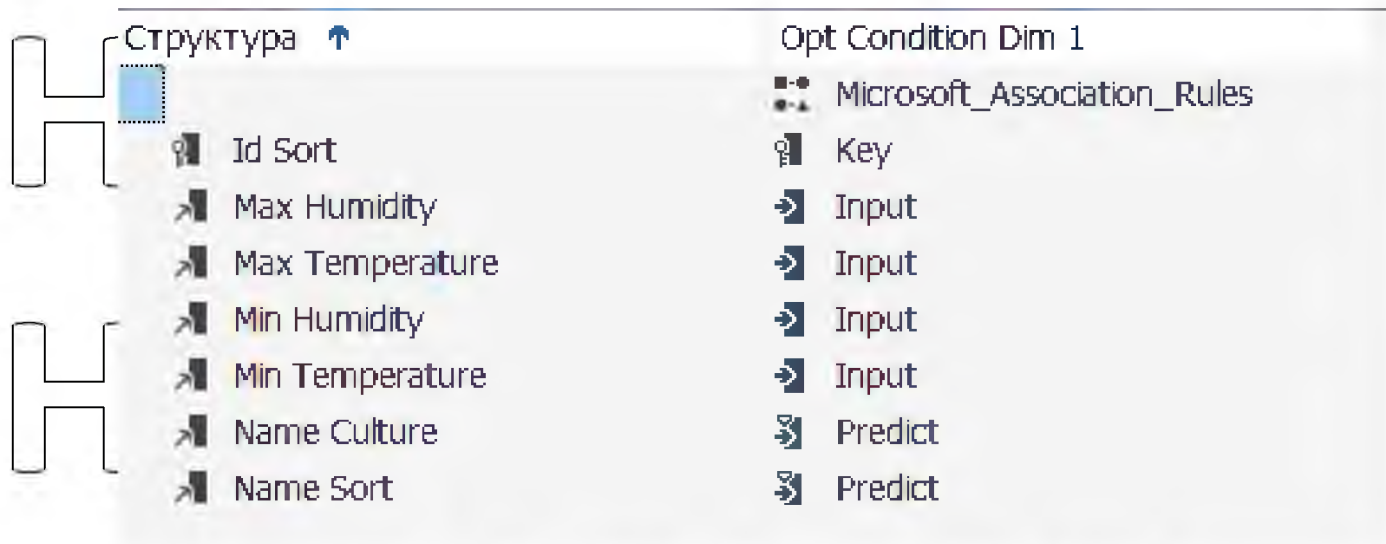


Рис. 36 Модель структури інтелектуального аналізу

Були побудовані правила з вказаною імовірністю та важливістю. Деякі з них наведені на таблиці 2.

НУБІП України

НУБІП України

НУБІП України

НУБІП України

Сформовані правила на основі знайдених взаємозалежностей

Імовірність	Важливість	Правило
1,000	0,425968732272281	Region = Київська, Country = Україна -> Name = Comfy Умань
1,000	0,425968732272281	Region = Вінницька -> Name = Ашан Київ №1
1,000	0,425968732272281	Region = Вінницька, City = Вінниця -> Name = Ашан Київ №1
1,000	0,425968732272281	Region = Вінницька, ID Location >= 3 -> Name = Ашан Київ №1
1,000	0,425968732272281	Region = Вінницька, Country = Україна -> Name = Ашан Київ №1
1,000	0,425968732272281	Region = Одеська -> Name = АТВ Васильків
1,000	0,425968732272281	City = Васильків -> Name = Comfy Умань
1,000	0,425968732272281	City = Вінниця, ID Location >= 3 -> Name = Ашан Київ №1
1,000	0,425968732272281	City = Вінниця, Country = Україна -> Name = Ашан Київ №1
1,000	0,425968732272281	ID Location = 2 - 3 -> Name = Comfy Умань
1,000	0,425968732272281	City = Васильків, Country = Україна -> Name = Comfy Умань
1,000	0,425968732272281	ID Location = 2 - 3, City = Васильків -> Name = Comfy Умань
1,000	0,425968732272281	ID Location = 2 - 3, Country = Україна -> Name = Comfy Умань
1,000	0,425968732272281	City = Затока -> Name = АТВ Васильків
1,000	0,425968732272281	City = Затока, Country = Україна -> Name = АТВ Васильків
1,000	0,425968732272281	City = Затока, ID Location >= 3 -> Name = АТВ Васильків
0,500	0,176091259055681	ID Location >= 3 -> Name = Ашан Київ №1
0,500	0,176091259055681	ID Location >= 3 -> Name = АТВ Васильків
0,500	0,176091259055681	ID Location >= 3, Country = Україна -> Name = Ашан Київ №1
0,500	0,176091259055681	ID Location >= 3, Country = Україна -> Name = АТВ Васильків

4.4 Тестування системи

Щоб користуватись системою користувачу достатньо мати будь-який обчислювальний прилад, такий як:

- комп'ютер;
- смартфон;
- планшет;
- мобільний телефон;
- ноутбук.

Головною вимогою для користування є лише наявність доступу до інтернету та встановлений будь-який браузер. Ця вимога пов'язана з тим, що система є web-орієнтованою, не потребує попереднього встановлення на клієнтський браузер, і починає працювати при переході за посиланням у браузері.

На стороні клієнта система не потребує додаткових встановлень, лише наявність будь-якого браузера.

Адмін також повинен зайти на сайт, який знаходиться на сервері і зайти з свого облікового запису.

Система була реалізована на двох підприємствах ТОВ «Епіцентрерк Калинівка Завод керамічної плитки» та компанія «IT Express» компанія яка займається логістичними перевезенням в США(офіс в Україні).

Опитування проводилося в «Епіцентрерк Калинівка Завод керамічної плитки» для цього були використані створені анкети за Європейськими стандартами. Загалом було опитано 63 людини.

Головна сторінка описує почергові кроки як потрібно працювати в системі., її вигляд представлено на рис.37.

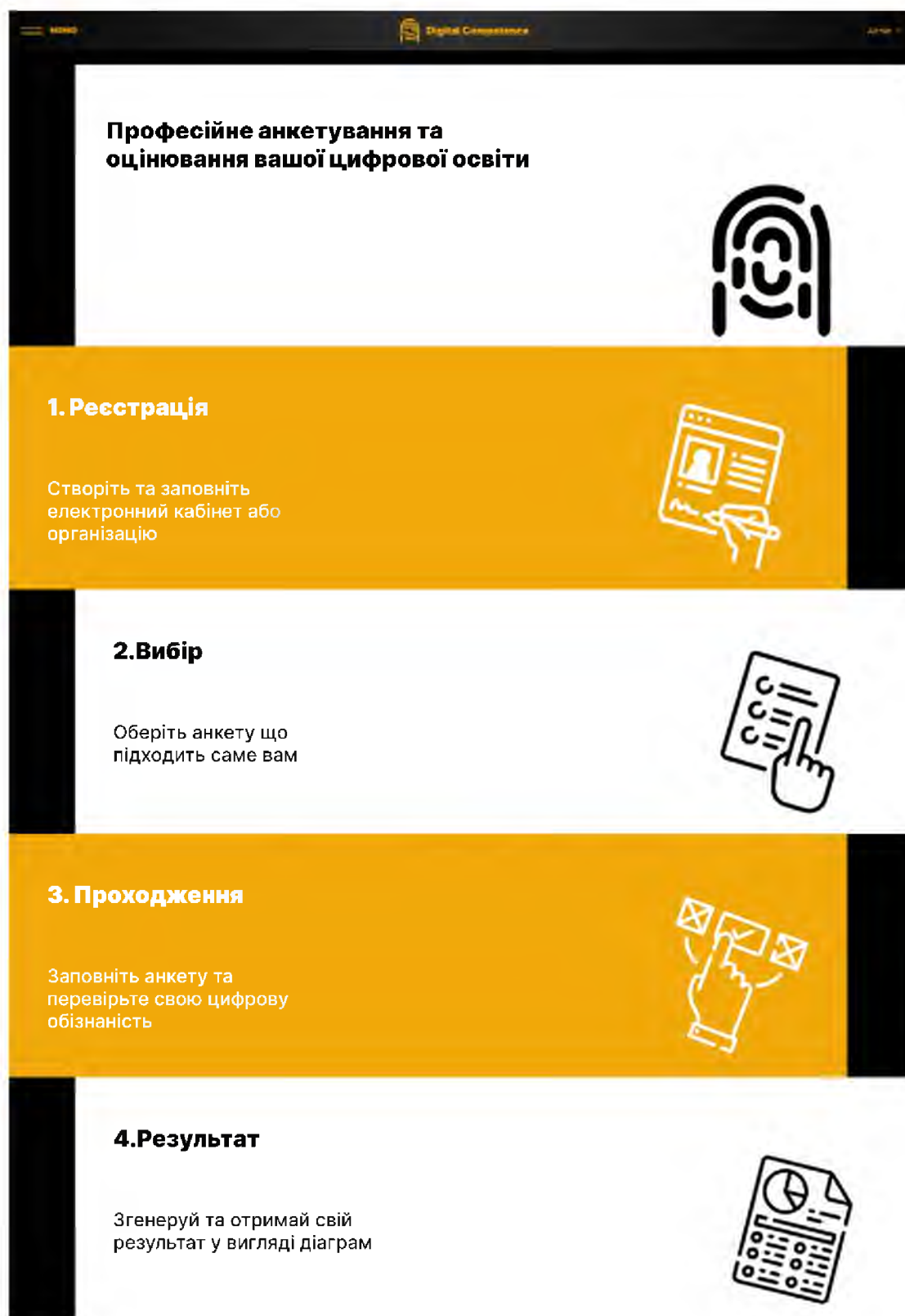


Рис.37 Головна сторінка «Як це працює?»

Після того як працівники були опитані, разом з керівником були переглянуті результати та обрані кілька графіків, щоб відобразити найактуальніші дані, які продемонстровані далі звітами за допомогою діаграм.

На діаграмі №1 (рис.38) представлено:

Скільки людей було опитано;

Результат самооцінювання за анкетною «Безпека»;

Середній бал опитаних працівників.

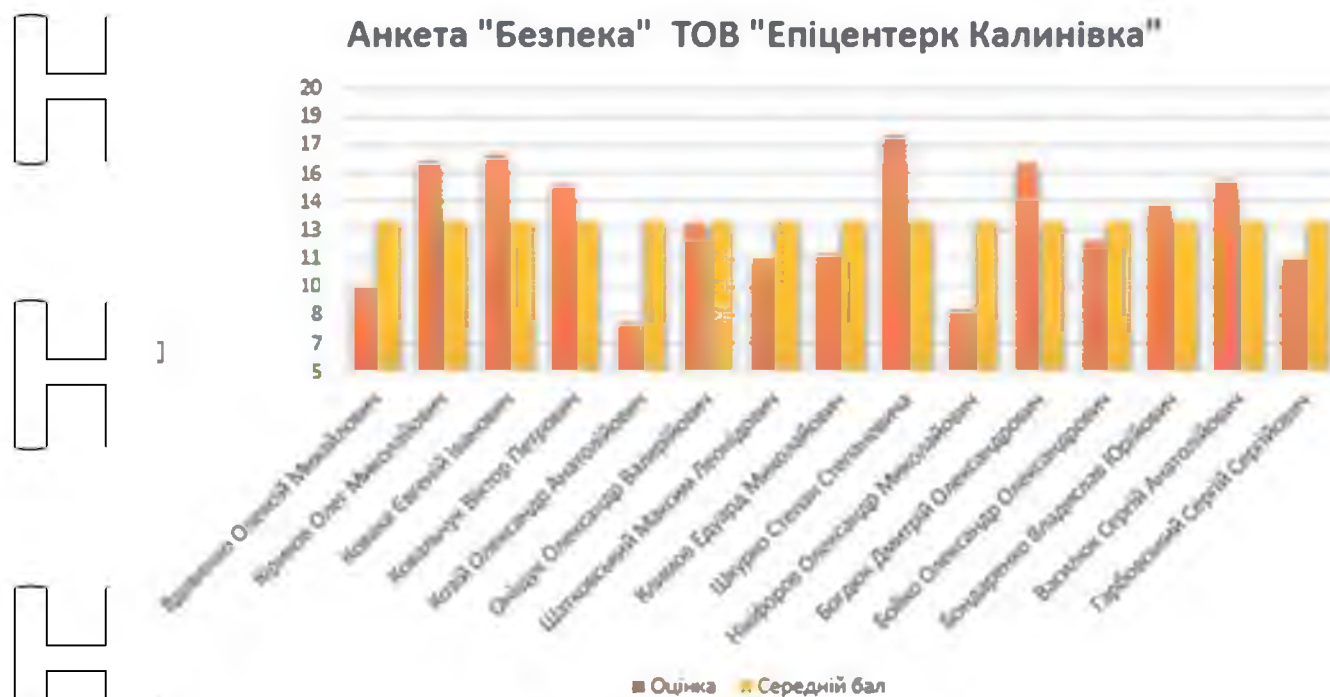


Рис. 38 Діаграма №1

Далі було проведено аналіз по більш глобальних даних.

Самоопитування відбувалося в добровільному порядку, працівники обирали анкету, хто на якій темі більше знався (до кожної анкети додавався короткий опис), можна було пройти тільки 1 анкетування.

Нижче представлено зміст анкет.

Анкета «Пошук» - здатність ідентифікувати, знаходити, отримувати, зберігати, упорядковувати й аналізувати цифрову інформацію, а також оцінювати відповідність і призначення.

Анкета «Спілкування» - здатність спілкуватися, співпрацювати, взаємодіяти та брати участь у віртуальних командах і мережах, а також використовувати відповідні медіа, тон і поведінку.

Анкета «Безпека» - здатність контролювати та захищати вашу особисту інформацію в інтернеті та розуміти наслідки особистих цифрових слідів.

Анкета «Виробництво» - можливість налаштовувати програми та пристрої відповідно до власних особистих уподобань, а також вирішувати технічні проблеми чи завдання.

Після опитування було сформовано діаграму №2 (рис. 39), на якій відображено результати відповідей 63 працівника, які були опитані, щоб проаналізувати, яких анкет було використано більше.

"Епіцентр" Калинівка види пройдених анкет

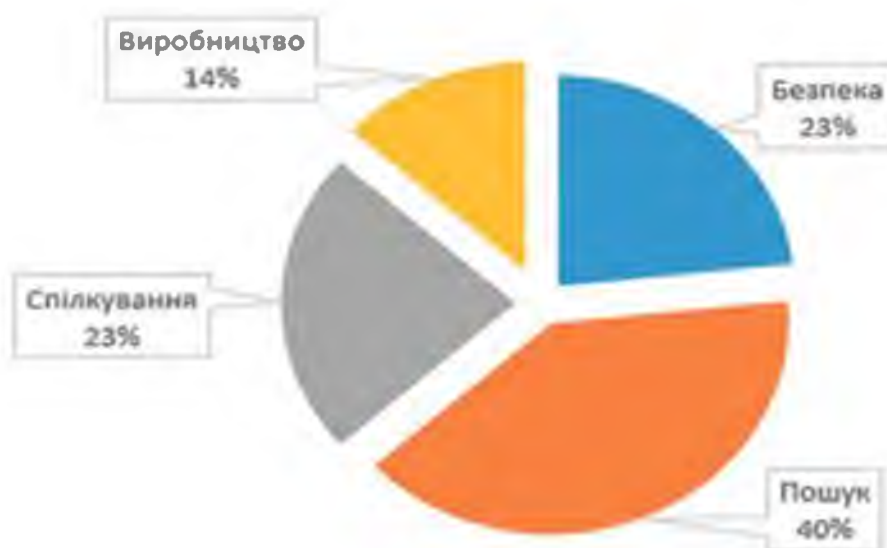


Рис. 39 Діаграма №2

Тож на діаграмі можемо побачити, що найбільше використовували анкету з назвою «Пошук».

Протестувавши систему з реальними людьми для чистоти її аналізу потрібні були результати хоча б ще однієї компанії.

Для цього було проведено дослідження в компанії «IT Express», з якою були проведені ті самі маніпуляції. Для аналізу було взято вибірку за певний період по певній анкеті рис 40.



Рис. 40 Діаграма №3

НУБІП України

НУБІП України

НУБІП України

НУБІП України

ВИСНОВКИ

НУБІП України

В магістерській роботі на тему «Інтелектуальна система аналізу

результатів самооцінювання рівня цифровізації компаній та цифрових навичок їх співробітників» були вирішені такі задачі.

НУБІП України

1. Створено систему для анкетування, тобто перевірки навичок цифровізації працівників та компаній.

2. Створено сховище даних, для цього:

- були отримали дані шляхом опитування працівників компаній;
- розроблено ПЗ передачі інформації з реляційної БД до СД.

НУБІП України

3. Аналіз даних та дослідження системи.

Завдяки отриманій інформації, накопиченій в процесі моніторингу, провели аналіз даних самоопитування працівників та оцінку їх цифрових знань.

НУБІП України

Отже, в результаті розробки магістерської роботи отримали систему, яка відображає поточні навички цифровізації компаній, що в свою чергу показує на слабкі та сильні сторони співробітників, покращивши які, компанія збільшить свою ефективність.

НУБІП України

НУБІП України

НУБІП України

СНИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. A common European framework for ICT - розміщено на платформі ecompetences.eu – [Електронний ресурс] – режим доступу:

<http://www.ecompetences.eu/>.

2. Україна – перша з розвитку цифрових навичок – розміщено на платформі Ukrinform – [Електронний ресурс] – режим доступу:

<https://www.ukrinform.ua/rubric-technology/3228076-cifrovizacia-poukrainski-u-rozvitku-cifrovih-navicok-mi-na-persomu-misci-v-evropi.html>.

3. Безтелесна Л. І. ОРГАНІЗАЦІЙНО-ПЕДАГОГІЧНИЙ МЕХАНІЗМ ФОРМУВАННЯ ЦИФРОВОЇ КОМПЕТЕНТНОСТІ СТУДЕНТІВ УНІВЕРСИТЕТІВ [Електронний ресурс] / Л. І. Безтелесна // Вісник НУВГП. –

15. – Режим доступу до ресурсу: <https://ep3.nuwm.edu.ua/21804/1/Ve931%201.pdf>

4. THE DIGITAL COMPETENCE WHEEL – розміщено на платформі Digital-competence – [Електронний ресурс] – режим доступу: <https://digital-competence.eu/dc/front/start/>.

5. Генсерук Г. Р. ЦИФРОВІ РЕСУРСИ ВИЗНАЧЕННЯ РІВНЯ ЦИФРОВОЇ КОМПЕТЕНТНОСТІ ФАХІВЦІВ [Електронний ресурс] / Галина Романівна Генсерук // Тернопільський національний педагогічний університет імені Володимира Гнатюка. –

8. Режим доступу до ресурсу: http://dspace.tnpu.edu.ua/bitstream/123456789/18847/1/45_Henseruk_Boiko.pdf.

6. Самко А. М. ЦИФРОВА КОМПЕТЕНТНІСТЬ ПЕДАГОГІЧНОГО ПЕРСОНАЛУ В СИСТЕМІ ПІСЛЯДИПЛОМНОЇ ПЕДАГОГІЧНОЇ ОСВИТИ [Електронний ресурс] / А. М. Самко // “Освітня аналітика України”. – 2021. –

Режим доступу до ресурсу: https://science.iea.gov.ua/wp-content/uploads/2021/07/3_Samko_213_2021_33_43.pdf.

7. Юхим Ю. Україна – перша з розвитку цифрових навичок серед країн Центральної та Східної Європи [Електронний ресурс] / Юлія Юхим // УКРИНФОРМ. – 2022. –

Режим доступу до ресурсу: <https://www.ukrinform.ua/rubric-technology/3228076-cifrovizacia-poukrainski-u->

rozvitku-cifrovih-navicok-mi-na-personu-misci-v-evropi.html.

8. Test ikanos individual - розміщено на платформі test.ikanos - [Електронний ресурс] - режим доступу:

<https://test.ikanos.eus/index.php/566697?lang=es&encode=> .

9. SELFIE - розміщено на платформі es.europa - [Електронний ресурс] - режим доступу: <https://es.europa.eu/education/schools-go-digital/en> .

10. Цифрограм - розміщено на платформі osvita.diia - [Електронний ресурс] - режим доступу: <https://osvita.diia.gov.ua/digigram> .

11. UML - розміщено на платформі ukr.kagutech - [Електронний ресурс] - режим доступу: <https://ukr.kagutech.com/4151830-uml-is-...-description-features-and-capabilities-tips> .

12. UML - use case - розміщено на платформі antibotan.com [Електронний ресурс] - режим доступу: http://antibotan.com/file.html?work_id=523323.

13. UML - розміщено на платформі evergreens.com [Електронний ресурс] - режим доступу: <https://evergreens.com.ua/ua/articles/uml-diagrams.html>

14. Що таке діаграма діяльності? - розміщено на платформі uk.icyscience - [Електронний ресурс] - режим доступу: <https://uk.icyscience.com/activity-diagram> .

15. Erwin - розміщено на платформі kpmg.ru - [Електронний ресурс] - режим доступу: <https://www.kpmg.ru/Automatization/ERwin.htm>

16. СХОВИЩА ДАНИХ [Електронний ресурс] // Stud.com.ua. - 2015. - Режим доступу до ресурсу: https://stud.com.ua/121123/informatika/shovischa_danih.

17. Плєскач В. Л. Спїр на оснївї сховищ даних [Електронний ресурс] / В. Л. Плєскач // studfile. - 2016. - Режим доступу до ресурсу: <https://studfile.net/preview/5429919/page:5/>.

18/ Застосування технологїї OLAP у діяльностї підприємства [Електронний ресурс] // Kursoviks. - 2016. - Режим доступу до ресурсу: <https://ua.kurovixs.com.ua/metodychki/265-lektsiya-4-zastosuvannya-tekhnologii->

olar-u-diyalnosti-pidприємства.

19. Системи звітів на базі SSRS, Cube, BI [Електронний ресурс] // syssoft. – 2014. – Режим доступу до ресурсу: <https://syssoft.pp.ua/services/sistemi-zvityi-na-bazi-ssrs-bi>.

20. Огляд алгоритмів та IC Data Mining [Електронний ресурс] // Студентська бібліотека. – 2017. – Режим доступу до ресурсу: <https://buklib.net/books/24617/>.

21. Моделі Data Mining [Електронний ресурс] // pidru4niki. – 2018. – Режим доступу до ресурсу: https://pidru4niki.com/1553080347788/informatika/modeli_data_mining.

22. 1R-алгоритм [Електронний ресурс] // nina.az. – 2021. – Режим доступу до ресурсу: <https://www.wiki.uk-ua.nina.az/1R-%D0%B0%D0%BB%D0%B3%D0%BE%D1%80%D0%B8%D1%82%D0%BC.htm>

1.

23. Теорема Баєса [Електронний ресурс] // Вікіпедія. – 2022. – Режим доступу до ресурсу: https://uk.wikipedia.org/wiki/%D0%A2%D0%B5%D0%BE%D1%80%D0%B5%D0%BC%D0%B0_%D0%91%D0%B0%D1%94%D1%81%D0%B0.

24. Easy Steps to Learn Naive Bayes Algorithm with codes in Python and R. [Електронний ресурс]. – Режим доступу: <https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/>

25. АСОЦЯТИВНІ ПРАВИЛА ТА ДЕРЕВА РІШЕНЬ [Електронний ресурс]. – 2020. – Режим доступу до ресурсу: https://msh.khnu.km.ua/pluginfile.php/573838/mod_folder/content/0/%D0%A28.pdf?forcedownload=1.

НУБІП УКРАЇНИ