

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ БІОРЕСУРСІВ  
І ПРИРОДОКОРИСТУВАННЯ УКРАЇНИ  
Факультет інформаційних технологій

УДК 004.9:631.544.4

«ПОГОДЖЕНО»

Декан факультету  
інформаційних технологій

«ДОПУСКАЄТЬСЯ ДО ЗАХИСТУ»

Завідувач кафедри комп'ютерних наук

Глазунова О.Г., д.п.н., професор

Голуб Б.Л., к.т.н., доцент

2022 р.

2022 р.

## МАГІСТЕРСЬКА КВАЛІФІКАЦІЙНА РОБОТА

на тему «Дослідження впливу електронних освітніх ресурсів на формування  
освітньої компетентності учнів»

Спеціальність 122 «Комп'ютерні науки»

Освітня програма «Інформаційні управляючі системи та технології»

(назва)

Орієнтація освітньої програми освітньо-професійна

Гарант освітньої програми

к.е.н., доцент

Густера О.М.

Керівник магістерської кваліфікаційної роботи

Старший викладач

Яциук Д.Ю.

Виконав

Бардей Д.Р.

КИЇВ-2022

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ БІОРЕСУРСІВ  
І ПРИРОДОКОРИСТУВАННЯ УКРАЇНИ

Факультет інформаційних технологій

**ЗАТВЕРДЖУЮ**

Завідувач кафедри комп'ютерних наук

к.т.н., доцент, Голуб Б.Л.

(науковий ступінь, вчене звання) (підпис)

(ПІБ)

“01” листопада 2021 року

**ЗАВДАННЯ**

**ДО ВИКОНАННЯ МАГІСТЕРСЬКОЇ КВАЛІФІКАЦІЙНОЇ РОБОТИ СТУДЕНТУ**

**Бардей Денису Руслановичу**

(прізвище, ім'я, по батькові)

Спеціальність 122 «Комп'ютерні науки»

Освітня програма «Інформаційні управлючі системи та технології»

Орієнтація освітньої програми: освітньо-професійна

Тема магістерської кваліфікаційної роботи «Дослідження впливу електронних освітніх ресурсів на формування освітньої компетентності учнів»

затверджена наказом ректора НУБіП України від “01” листопада 2021 р. №1862 «С»

Термін подання завершеної роботи на кафедру 27 жовтня 2022 р.

Вихідні дані до магістерської кваліфікаційної роботи:

1. Отримання необхідних даних для формування коефіцієнта компетентності
2. Аналіз даних, отриманих при дослідженні
3. Формування звітів та рекомендацій щодо покращення КК

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)

Завдання прийняв до виконання

(підпис)

Бардей Д.Р.

(прізвище та ініціали)

## КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів дипломної роботи	Строк виконання етапів роботи	Примітка
1	Аналіз проблемної області	03.09.2022 26.09.2022	Опис та аналіз актуальності та мети теми, об'єкту, предмету та методів дослідження, а також формування поставлених задач.
2		26.09.2022 29.10.2022	Збір даних по існуючим системам та огляд літературних джерел. Опис предметної області. Аналіз показників які впливають на коефіцієнт освітньої компетентності.
3	Моделювання системи	02.11.2022 13.12.2022	Розробка діаграм прецедентів використання, послідовності діаграму класів.
4	Інформаційне забезпечення системи	13.12.2022 21.01.2022	Проектування алгоритму підрахунку коефіцієнту компетентності
5		20.01.2022 16.02.2022	Розробка БД
6		17.02.2022 11.03.2022	Дослідження OLAP технологій та методів Data Mining. Проектування СД.
7		12.03.2022 06.04.2022	Розробка куба інтерактивної аналітичної обробки даних (OLAP) і моделі інтелектуального аналізу даних.
8		07.04.2022 29.04.2022	Вибір та встановлення програмних засобів для розробки ПЗ. Розробка алгоритмів роботи системи.
9		05.05.2022 19.06.2022	Розробка сайту, як основу для збору необхідної інформації
10		01.09.2022 28.09.2022	Інтеграція алгоритму до сайту підключення система аналізу
11	29.09.2022 12.10.2022	Впровадження та експлуатація системи	
13	Написання пояснювальної записки	02.11.2022 27.11.2022	Оформлення пояснювальної записки магістерської роботи, а також інших супровідних документів (презентації, доповіді)

Студент

Бардей Д.Р.

(підпис)

Керівник роботи

Ящук Д.Ю.

(підпис)

<b>ЗМІСТ</b>	<b>ЗМІСТ</b>	<b>4</b>
<b>ВСТУП</b>		<b>5</b>
<b>1 АНАЛІЗ ПРОБЛЕМНОЇ ОБЛАСТІ</b>		<b>9</b>
1.1 Загальна характеристика		9
1.2 Аналіз існуючих рішень		10
1.3 Постановка завдання		11
1.4 Процес отримання необхідних даних.		15
<b>2 МОДЕЛЮВАННЯ СИСТЕМИ</b>		<b>17</b>
2.1 Загальні відомості		17
2.2 Об'єктне та функціональне моделювання		19
2.3 Абстракція предметної області		27
<b>3 РОЗРОБКА СИСТЕМИ АНАЛІЗУ</b>		<b>30</b>
3.1 OLAP-технології		30
3.2 Архітектура системи		32
3.3 Опис бази даних		32
3.4 Структура сховища даних		38
3.5 Створення OLAP куба		41
3.6 Створення потоків даних за допомогою Data Flow		46
3.7 Використані технології		50
3.8 Поєднання технологій		52
<b>4 ВПРОВАДЖЕННЯ ТА ЕКСПЛУАТАЦІЯ СИСТЕМИ</b>		<b>55</b>
4.1 Загальні поняття технології Data Mining		55
4.2 Decision Trees Algorithm		58
4.3 Naive Bayes Algorithm		62
4.4 Дослідження використання методу асоціативних правил		66
4.5 Дослідження використання алгоритмів кластеризації		71
4.6 Розрахунок KPI		75
4.7 Експлуатація		76
<b>ВИСНОВКИ</b>		<b>85</b>
<b>СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ</b>		<b>86</b>

# ВСТУП

# НУБІП України

Електронні освітні ресурси – це веб сайти які дозволяють отримувати інформацію незалежно від місця перебування, єдина вимоги, це доступ в мережу Internet та безпосередньо пристрій, на якому можна зайти в мережу.

# НУБІП України

На зорі мережі Internet подібні ресурси мало чим відрізнялись від звичних нам бібліотек, про те з часом вони набували все більшої складності. Так з часом з'явилися ресурси, які базуються на одній конкретній темі, далі з часом з'явилися форуми, і нововведенням стала можливість обговорити ті чи інші теми з іншими користувачами ресурсу.

# НУБІП України

В сучасному вигляді збереглись ще ті перші подібні ресурси, про те при розробці ресурсів, які б відповідали всім сучасним вимогам цього буде замало.

Для повноцінного аналізу потрібний ресурс який міг би при певних умовах повністю покрити всі потреби в здобуванні освіти. Тобто розробити веб сайт, який був би нічим іншим як онлайн школою.

# НУБІП України

За допомогою подібного ресурсу ми в повній мірі зможемо оцінити яким чином електронні освітні ресурси впливають на формування освітньої компетенції учнів.

# НУБІП України

В привичному всім очному форматі навчання учитель перевіряє та оцінює успішність учнів, керуючись директивними документами, педагогічною теорією та практикою. Найголовнішими з цих вимог є індивідуальний характер перевірки та оцінювання знань, систематичність, диференційованість, об'єктивність, умотивованість оцінок, вимогливість учителя, єдність вимог.

# НУБІП України

Реформування загальної середньої освіти передбачає втілення у життя принципу гуманізації освіти, методологічну переорієнтацію процесу навчання з інформативної форми на розвиток особистості людини, індивідуально-диференційований, особистісно-орієнтований підхід до навчання оцінювання навчальних досягнень кожного учня.

# НУБІП України

Система оцінювання покликана визначати на кожному етапі навчання рівень успішності учнів відповідно до вимог державного стандарту загальної середньої освіти, рівень компетентності учнів, їх готовності до застосування засвоєних знань на практиці. Адже навчальна діяльність повинна не просто дати людині суму знань, а сформувати комплекс компетенцій.

### Актуальність теми

Сфера освіти є невід'ємною частиною держави. В тій чи іншій мірі це невід'ємна частина життя кожного громадянина. Її вплив на життя людини важко переоцінити, а важливість стає в один ряд з сферою медицини чи економіки. Проте сферу освіти виокремлює дуже важливий фактор – з освіти починається все. Від якості освіти залежить загальний рівень всіх сфер як в житті людини, так і на загальнодержавному рівні. Тому жодна держава в світі не повинна нехтувати якістю освіти яка надається на її теренах.

Задля забезпечення якості створюються безліч інститутів, які контролюють, аналізують та вносять нові рекомендації, які повинні позитивно впливати на розвиток. Сфера освіти не є виключенням, і вона потребує постійного оновлення, щоб йти в ногу з сучасними тенденціями сфера повинна пристосовуватись до викликів.

Протягом сотень років людство зіштовхувалося з викликами, які не завжди виходило легко подолати. Задля збільшення шансів подолати той чи інший виклик людству доводилось пристосовуватись. І з ускладненням системи виклики стають все складнішими, тому і шляхи подолання потрібно теж ускладнювати, але разом з цим для сприйняття людини вони повинні бути досить простими.

З 2020 року в світі актуальна тенденція переходу від традиційного очного формату до дистанційного, який являє собою новим нетрадиційним способом навчання, якому ще потрібно подолати великий шлях до вдосконалення системи, щоб впоратись з всіма сучасними викликами.

Задля цього сфера потребує перегляду системи оцінювання яка б оцінювала також і неформальне оцінювання, яке також формує компетентність учнів. Освітня компетентність учня вбирає в себе всі показники його процесу навчання, і дає повну картину того, наскільки якісно буде навчений учень.

### **Об'єкт і предмет дослідження**

Об'єктом дослідження є процес навчання в українській школі. Предметом дослідження – алгоритм визначення коефіцієнта компетентності та розробка системи для отримання, зберігання, обробки, аналізу та формування звітної інформації.

### **Мета роботи**

Мета роботи полягає в розробці системи, яка б могла оцінити та відобразити всю повноту освітньої компетентності учня, система повинна бути незалежною від формату навчання, зберігати та обробляти великі об'єми даних, бути достатньо простою в керуванні, задля можливості розгортання в будь-якій області сфери освіти і доступною для сприйняття всіма категоріями населення.

### **Зміст поставлених завдань**

Для досягнення поставленої мети необхідно:

- провести системний аналіз об'єкта;
- сформулювати вимоги та структуру до системи оцінювання рівня компетентності;
- розробити структуру інформаційного забезпечення системи;
- розробити алгоритми аналізу даних, провести дослідження системи і випробування в реальних умовах.

### **Методи дослідження**

При проведенні досліджень були використані певні стандартні програмні продукти, такі як: ERwin Data Modeler, SQL Server та методи Data mining.

### **Наукова новизна**

В процесі розробки були визначені всі можливі чинники, які впливають на якість освіти. Створено алгоритм розрахунку коефіцієнта компетентності, який зображує рівень якості освіти.

### **Апробація результатів дослідження**

В процесі розробки магістерської роботи на тему «Дослідження впливу електронних освітніх ресурсів на формування освітньої компетентності учнів» були проведені виступи на XIII Міжнародній науково-практичній конференції молодих вчених «Інформаційні технології: економіка, техніка, освіта»

### **Структура роботи**

Структурні елементи роботи розміщені в такій послідовності:

У першому розділі розглядається предметна область, сформульована постановка задачі дослідження та аналіз стану аналізу впливу електронних освітніх ресурсів на освітню компетентність учнів

Другий розділ присвячено моделюванню системи. Були розроблені діаграми прецедентів використання, пакетів, класів та розгортання, які дозволяють отримати передумови для аналізу поведінки системи.

Третій розділ присвячено розробці та реалізації інформаційного забезпечення системи. Було побудовано логічну модель даних, реляційну БД та спроектовано сховище даних та створено розгорнутий куб системи.

У четвертому розділі розглянуто розробка програмного забезпечення системи. Були розроблені алгоритми роботи системи, програмне забезпечення моніторингу процесів формування освітньої компетенції учнів та їх аналізу.



# 1 АНАЛІЗ ПРОБЛЕМНОЇ ОБЛАСТІ

## НУВБІП УКРАЇНИ

### 1.1 Загальна характеристика

Компетенція — загальна здатність, що базується на знаннях, досвіді, цінностях, здібностях, набутих завдяки навчанню. Основними компетенціями, яких потребує сучасне життя, є:

- **соціальні** — здатність брати на себе відповідальність, брати участь у спільному прийнятті рішень, врегулюванні конфліктів ненасильницьким шляхом, у функціонуванні та розвитку демократичних інститутів суспільства;
- **полікультурні** — пов'язані з життям у полікультурному суспільстві, передбачають розуміння несхожості людей, взаємоповагу до їх мови, релігії, культури тощо;
- **комунікативні** — володіння усною і писемною мовою та іншими мовами;
- **інформаційні** — зумовлені зростанням ролі інформації у сучасному суспільстві й передбачають оволодіння інформаційними технологіями, вміння здобувати, критично осмислювати і використовувати різноманітну інформацію;
- **саморозвитку та самоосвіти** — потреба і готовність постійно вдосконалюватись як у професійному, так і в особистому та суспільному плані.

Тобто потрібно розуміти, що компетентність учня не можна описати лиш за допомогою оцінки за предмет, чи в цілому за середній бал по всім предметам. Потрібно врахувати інші фактори, які хоч якось можуть вплинути на формування освітньої компетентності учня. З цією метою було визначено 2 можливих види оцінювання учнів — формальне та неформальне.

До формального оцінювання ми віднесли класичну систему оцінювання, яка залежить від таких факторів:

- Загальна оцінка за семестр;
- Середня оцінка за домашнє завдання;
- Середня оцінка під часу уроку;

- Середня оцінка за самостійні та контрольні роботи.

Також, щоб дослідження надало більш точні дані в програму були введені щоденні тести, які проводяться після кожного уроку, в залежності від уроку, в усні чи письмовій форма:

Для формування неформальної оцінки враховуються наступні фактори:

- Час витрачений на самостійно опрацювання матеріалів уроку;
- Час витрачений на самостійне опрацювання матеріалів для виконання домашнього завдання;
- Середній час виконання домашнього завдання;
- Середнє значення показника самоорганізації учня, яке базується на само оцінюванні учня, та з точки зору вчителя;
- Середній час вивчення додаткових матеріалів
- Активність під час уроку;
- Кількість не виконаного домашнього завдання;
- Кількість пропущених занять.

Тому потрібно в першу чергу розробити сайт, за допомогою якого учень зможе, не залежно від поточного місця перебування, повноцінно навчатись.

Розроблений сайт буде чудовою платформою для отримання широкого аспекту даних, які потрібні для аналізу, і відповідно створенню висновків про те, як змінився рівень компетентності учня, якщо порівнювати з класичним форматом навчання.

### 1.2 Аналіз існуючих рішень

За напрямком дослідження були проаналізовані останні дослідження та публікації.

Так в статті «Формування освітньої компетентності по курсу предмету «Я досліджую світ»» [24] було отримано вичерпну інформацію про те, як можна

визначити формат оцінювання, були сформовані терміни **Формальне** та **Неформальне** навчання. Визначені формати отримання необхідної інформації, для формування коефіцієнту освітньої компетентності саме по гуманітарним предметам. На прикладі було показано яким чином можна зібрати необхідну інформацію в усній формі.

В статті Рекомендації від УЦОЯО щодо формування читацької та математичної компетентності учнів на рівні початкової освіти[25] було отримано вичерпну інформацію про формування освітньої компетентності на прикладі предмету «Алгебра». Було отримано інформацію про методи отримання необхідних даних по технічним предметам.

Так використовуючи досвід попередників було сформовано метод, за допомогою якого можна одночасно підрахувати загальний коефіцієнт освітньої компетентності учнів.

### 1.3 Постановка завдання

Для отримання інформації, яка необхідна для формування ІК, потрібно створити електронний освітній ресурс, який відповідав би сучасним вимогам.

Він повинен повністю задовільнити потреби учнів та вчителів. Також він повинен бути простим в адмініструванні, та з широким доступом до потрібної інформації для аналітиків.

Функціонал сайту повинен містити:

- Створення та видалення особистих кабінетів;
- Створення академічних груп ( далі класів);
- Надання прав користувачам;
- Завантаження навчальних матеріалів та завдань;
- Завантаження виконаних учнем завдань;
- Створення, видалення та редагування новин;
- Створення та присвоєння викладачу предметів;
- Виставлення оцінок за виконані роботи;

НУБІП України

- Перегляд предметів, які викладаються в класі, що відносяться до конкретного класу;
- Перегляд викладачів;

НУБІП України

- Проходження тестів;
- Проходження опитування;
- Відображення інформації про КК відповідно до заданого масштабу

Web – сайт повинен містити на головній сторінці форму авторизації, і у разі невдалої авторизації повинен бути закритим для перегляду неавторизованому користувачу. Після успішної авторизації, перед очима

НУБІП України

користувача буде стрічка новин, та ключові кнопки керування:

- Аналітика;
- Моя успішність;

НУБІП України

- Мої уроки;
- Мій розклад;
- Класний журнал;
- Класна кімната;

НУБІП України

- Домашнє завдання;
- Кнопка виходу з профіля;
- Особистий кабінет;
- Головна;
- Новини;

НУБІП України

- Галерея;
- Контакти;
- Кнопка зміни мови на сайті.

НУБІП України

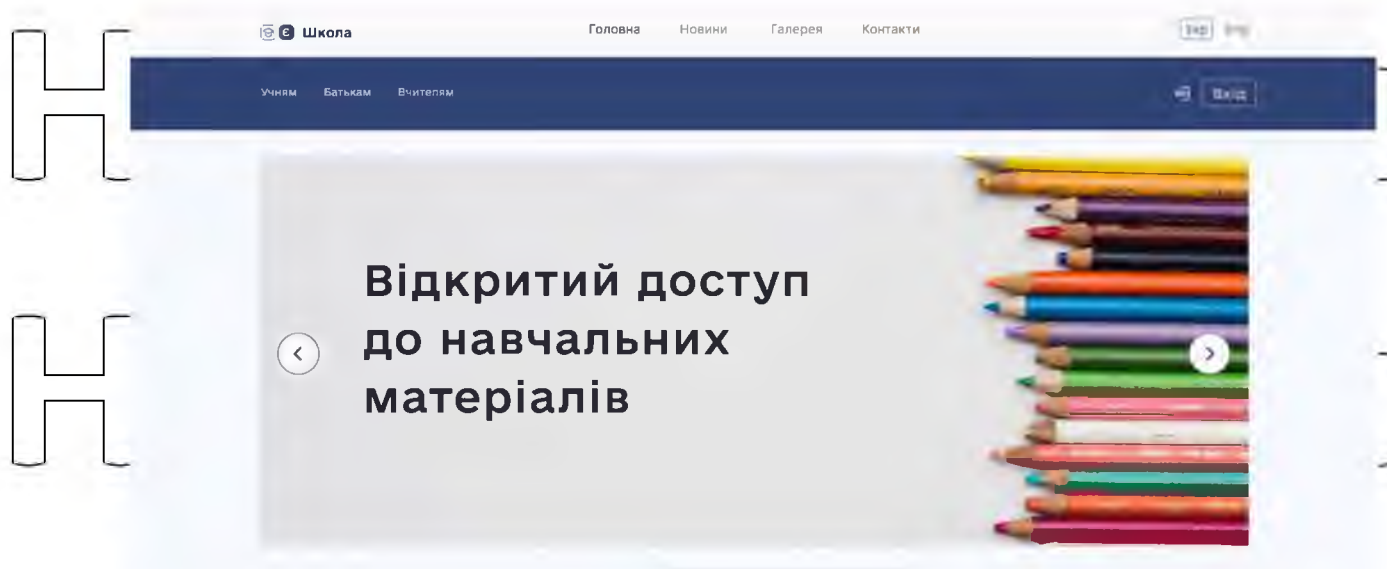


Рис. 1 Головна сторінку сайту

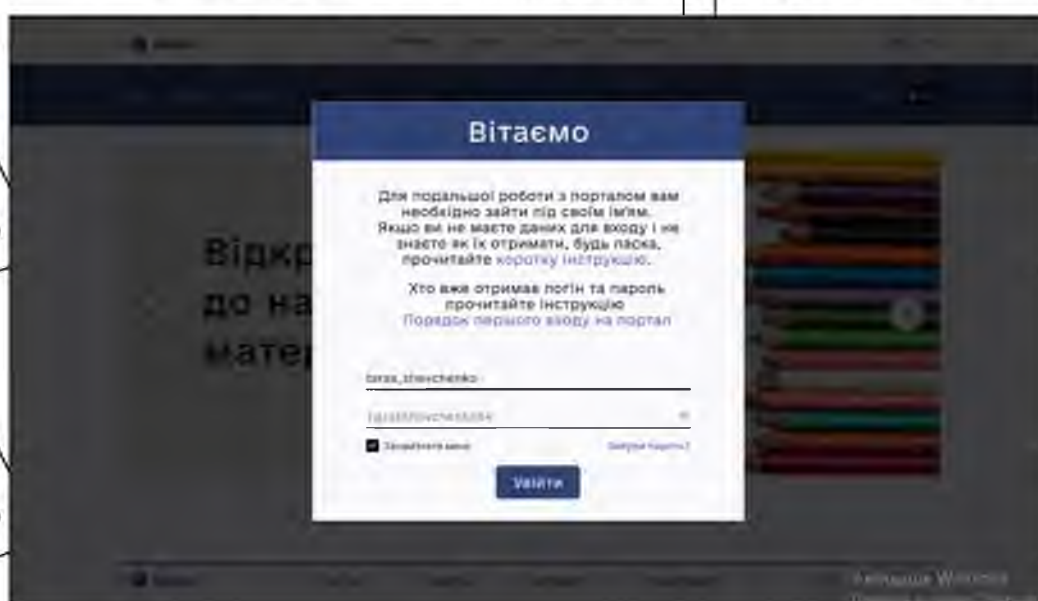


Рис. 2 Сторінка авторизації

Створений сайт має інтуїтивно зрозумілий інтерфейс, який має мінімальну кількість елементів, за для повного зосередження на навчальному процесі та роботі. Також додатково було додано опцію «Запам'ятай мене», яка дає можливість в подальшому користувачам не витрачати зайвий час на вхід.

Отже, розробка даного сайту дозволить в повній мірі надати учням весь необхідний функціонал для повноцінного навчання, що в свою чергу дасть змогу отримати всі необхідні дані.

Маючи у розпорядженні подібний сайт є змога розгорнути систему аналізу, розгорнута система повинна не впливати на роботу сайту, лиш доповнювати його.

Для розгортання системи потрібно лише додати необхідні елементи системи. Вся система являє собою вкладка «Аналітика» в яку користувач може зайти натиснувши кнопку «Аналітика», в залежності від ролі користувач побачить перед собою відповідну інформацію:

- Користувач з роллю «Вчитель» буде бачити перед собою загальні результати навчання кожного учня по відповідному предмету, який викладає вчитель, буде сформовано коефіцієнт освітньої компетентності класу чи конкретного учня, також буде бачити рекомендації щодо покращення викладання для підвищення коефіцієнта освітньої компетентності;

- Користувач з роллю «Учень» буде бачити результати свого навчання, він зможе переглянути як загальний коефіцієнт освітньої компетентності так і конкретно по кожному предмету. Також йому буде відображатись загальні рекомендації по всім предметам, та результати по конкретному предмету;

- Користувач з роллю «Аналітик» натиснувши кнопку «Аналітика» матиме змогу бачити загальний коефіцієнт класу по всім предметам, коефіцієнт по конкретному предмету, коефіцієнт по конкретному учню, та коефіцієнт по конкретному предмету та конкретному учню. Як особливість ролі він буде мати змогу встановлювати еталонні значення, або використовувати уже зафіксовані результати навчання по минулим семестрам для того, щоб системі було з чим порівнювати та формувати висновки.

Для кожної ролі буде доступним перегляд результатів по двом проміжках.

Проміжки обираються за побажанням користувача, але зазвичай будуть порівнюватись між собою два семестри, в нашому випадку це буде очне навчання і дистанційне. В такому випадку користувач отримає всю звітну



7. Учень проходить тест;  
8. Система отримує інформацію про результат тесту;  
9. Учень самостійну контрольну роботу;  
10. Вчитель оцінює роботу;

11. Система отримує оцінку про результат самостійної контрольної роботи;  
12. Учень проходить опитування;  
13. Вчитель проходить опитування;

14. Аналітик формує критерії утворення КК;  
15. Учень та вчитель переглядає КК.

НУБІП України

НУБІП України

НУБІП України

НУБІП України



## 2. МОДЕЛЮВАННЯ СИСТЕМИ

# НУБІП України

### 2.1 Загальні відомості

2.1.1 Системний аналіз. Системний аналіз – це сукупність методологічних

засобів, які використовуються для підвищення ступеня обґрунтованості рішень у складних (слабко-структурованих) проблемах політичного, військового, наукового, соціального і економічного характеру. Системний аналіз передбачає розгляд об'єктів як систем, переважно цілеспрямованих. Основні методологічні засади системного аналізу базуються на принципах системного підходу.

Системний аналіз відрізняється від інших методів дослідження тим, що:

- враховує принципову складність об'єкта, що досліджується; бере до уваги розгалужені та стійкі взаємні зв'язки його з оточенням; враховує неможливість спостереження ряду властивостей об'єкта та оточуючого середовища;
- реальні явища, їх властивості та зв'язки з оточенням переводяться далі в абстрактні категорії теорії систем;
- ґрунтуючись на відомих властивостях складних систем дозволяє виявити нові конкретні властивості та взаємні зв'язки конкретного об'єкта дослідження;
- на відміну від інших методів, в яких точно визначені об'єкти, включає як один з важливих етапів визначення об'єкта, його знаходження чи конструювання;
- орієнтується не на розв'язання «правильно сформульованих» задач, а на створення правильної постановки задачі, вибір відповідних методів для її розв'язання;
- основне в СА — знайти шлях, яким можна перетворити складну проблему в простішу, яким чином не лише складну до розв'язання, але й для розуміння, проблему перетворити в послідовність задач, для яких існують методи їх розв'язання;

# НУБІП України

• SA завжди конкретний — завжди має справу з конкретною проблемою, конкретним об'єктом дослідження, є продуктивним тоді, коли застосовується до розв'язання завдань певного типу [1].

Тож якщо підсумувати все з визначень системного аналізу — можна коротко сказати, що він створений для полегшення роботи людини в речах, які раніше були рутинними чи взагалі неможливими.

2.1.2 UML. Універсальна мова моделювання (Unified Modelling Language або UML) — це мова позначень або побудови діаграм, призначена для визначення, візуалізації і документування моделей зорієнтованих на об'єкти

систем програмного забезпечення. UML не є методом розробки, іншими словами, у конструкціях цієї мови не повідомляється про те, що робити першим, а що останнім, і не надається інструкції щодо побудови вашої системи, але ця мова

допомагає вам наочно переглядати компонування системи і полегшує співпрацю з іншими її розробниками. Розробкою UML керує Object Management Group (OMG). Ця мова є загальноприйнятим стандартом графічного опису програмного забезпечення.

UML розроблено для розробки структури зорієнтованого на об'єкти програмного забезпечення, ця мова має дуже обмежену користь для програмування на основі інших парадигм.

Конструкції UML створюються з багатьох модельних елементів, які позначають різні частини системи програмного забезпечення.

Елементи UML використовуються для побудови діаграм, які відповідають певній частині системи або точці зору на систему.

У Umbrello UML Modeller реалізовано підтримку таких типів діаграм:

- Діаграма випадків використання показує дієвих осіб (людей або інших користувачів системи), випадки використання (сценарії використання системи) та їх взаємодію;
- Діаграми класів, на яких буде показано класи та зв'язки між ними;
- Діаграми послідовності, на яких показано об'єкти і послідовність методів, якими ці об'єкти викликають інші об'єкти;

- Діаграми співпраці, на яких буде показано об'єкти та їх взаємозв'язок з наголосом на об'єкти, які беруть участь у обміні повідомленнями;
- Діаграми стану, на яких буде показано стани, зміну станів і події у об'єкті або частині системи;
- Діаграми діяльності, на яких буде показано дії та зміни однієї дії іншою, які є наслідком подій, що сталися у певній частині системи;
- Діаграми компонентів, на яких буде показано програмні компоненти високого рівня (на зразок KParts або Java Beans);
- Діаграми впровадження, на яких буде показано екземпляри компонентів та їх взаємодію;
- Діаграми взаємозв'язку сутностей, на яких буде показано дані, взаємозв'язки і умови обмеження зв'язків між даними[2].

## 2.2 Об'єктне та функціональне моделювання

2.2.1 *Діаграма прецедентів.* Діаграма варіантів використання (use case diagram) – Візуальне моделювання в UML можна представити як деякий процес спуску від найбільш загальної і абстрактної концептуальної моделі вихідної системи до логічної, а потім і до фізичної моделі відповідної програмної системи.

Для досягнення цих цілей спочатку будується модель у формі так званої діаграми варіантів використання (use case diagram), яка описує функціональне призначення системи або, іншими словами, те, що система буде робити в процесі свого функціонування. Діаграма варіантів використання є вихідним концептуальним уявленням або концептуальною моделлю системи в процесі її проектування і розробки[3].

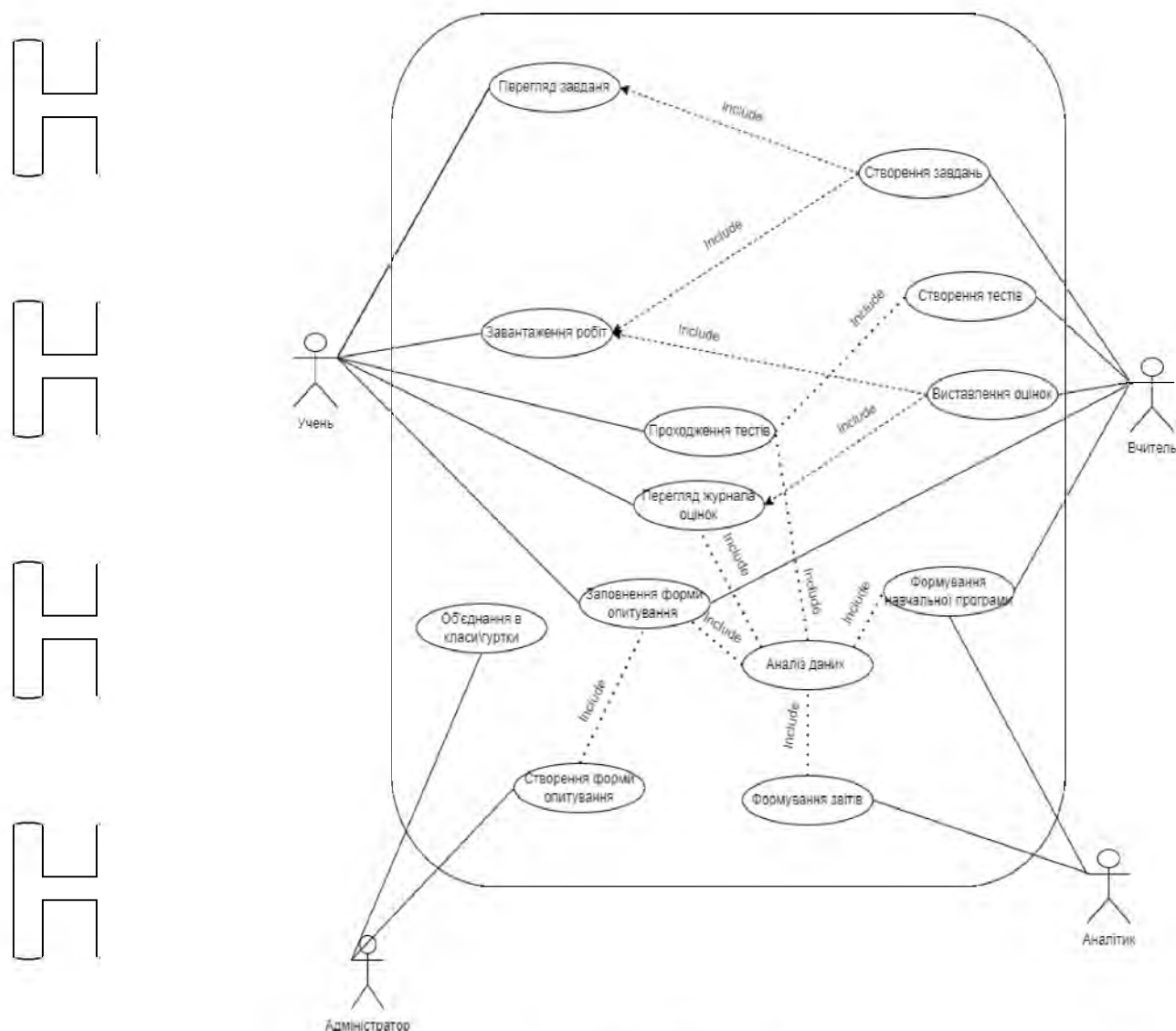


Рис. 4 Діаграма Прецедентів

Діаграма до даної системи зображена на рис. 4. На діаграмі зображено взаємодія між акторами, які взаємодіють з системою, та прецеденти в самій системі. Актори знаходяться поза системою. Сама діаграма являє собою візуалізацію того, що може чи повинен робити актор в системі, але не дає відповіді на те яким чином буде досягнуто тої чи іншої мети, які для цього потрібні ресурси та інше.

Змодельована система містить таких акторів:

- Адміністратор;
- Вчитель;
- Студент(учень);
- Аналітик.

Актор «Адміністратор» включає такі прецеденти:

- Об'єднання в класи/гуртки;
- Створення форми опитування.

Актор «Викладач» включає такі прецеденти:

- Створення завдань;
- Виставлення оцінок;
- Створення тестів;
- Заповнення форми опитування;
- Формування навчальної програми

Актор «Учень» включає такі прецеденти:

- Проходження тестів;
- Завантаження робіт;

- Перегляд завдання;
- Заповнення форми опитування;
- Перегляд журналу оцінок.

Актор «Аналітик» включає такі прецеденти:

- Формування звітів;
- Формування навчальної програми;
- Аналіз даних.

Розглянемо приклад на конкретному сценарії - Учень хоче виконати завдання, завантажити результат та перевірити, який відгук залишив вчитель.

*Оптимістичний сценарій:*

1. Учень заходить в навчальну програму, та обирає завдання, яке він хоче виконати.
2. Завантажує роботу, виконує її та завантажує виконане завдання.
3. Вчитель перевіряє, залишає оцінку в журналі.
4. Учень та вчитель проходить опитування.
5. Учень проходить тестування.
6. Система отримує данні та формує КК.

Результат: Учень завантажив роботу, яку перевірили і залишили відгук.

*Задовільний сценарій:*

1. Учень заходить в навчальну програму, обирає завдання.
2. Завантажує роботу виконує завдання та завантажує на сайт.
3. Вчитель перевіряє завдання та залишає оцінку.
4. Учень та вчитель проходить опитування.
5. Учень не проходить тестування.
6. Система формує КК використовуючи не повну вибірку даних.

Результат: Система формує КК, проте він не є точним, так як система не отримала всі необхідні данні.

*Негативний сценарій:*

1. Учень заходить в навчальну програму, обирає завдання.
2. Завантажує роботу виконує завдання та завантажує на сайт.
3. Вчитель оцінює роботу.
4. Учень та вчитель не проходить опитування.
5. Учень не проходить тест.
6. Система формує КК який не відповідає стандартам

Результат: Система сформувала КК, проте він не відповідає стандартам, так як система не отримала інформації у повній мірі

2.2.2 *Діаграма послідовності.* Діаграма послідовності — відображає взаємодії об'єктів впорядкованих за часом. Зокрема, такі діаграми відображають задіяні об'єкти та послідовність відправлених повідомлень.

Іншими словами, діаграма послідовностей відображає часові особливості передачі і прийому повідомлень об'єктами.

Діаграми послідовностей можна використовувати для уточнення діаграм прецедентів, більш детального опису логіки сценаріїв використання. Це відмінний засіб документування проекту з точки зору сценаріїв використання.

Діаграми послідовностей зазвичай містять об'єкти, які взаємодіють у рамках сценарію, повідомлення, якими вони обмінюються, і які повертаються результати, які пов'язані з повідомленнями.

На діаграмі послідовності відображаються тільки ті об'єкти, які безпосередньо беруть участь у взаємодії.

Лінія життя об'єкта зображується пунктирною вертикальною лінією, асоційованою з єдиним об'єктом на діаграмі послідовності. Лінія життя служить для позначення періоду часу, протягом якого об'єкт існує в системі і, отже, може потенційно брати участь у всіх її взаємодіях. Якщо об'єкт існує в системі постійно, то і його лінія життя повинна продовжуватися по всій площині діаграми послідовності від самої верхньої її частини до самої нижньої.

У процесі функціонування об'єктно-орієнтованих систем одні об'єкти можуть перебувати в активному стані, безпосередньо виконуючи певні дії, або стані пасивного очікування повідомлень від інших об'єктів. Щоб явно виділити подібну активність об'єктів, в мові UML застосовується спеціальне поняття, що отримало назву фокуса управління. Фокус управління зображується у формі витягнутого вузького прямокутника, верхня сторона якого позначає початок отримання фокусу управління об'єкта, а його нижня сторона - закінчення фокусу управління. Прямокутник розташовується нижче позначення відповідного об'єкта і може замінювати його лінію життя, якщо на всьому її протязі він є активним[4].

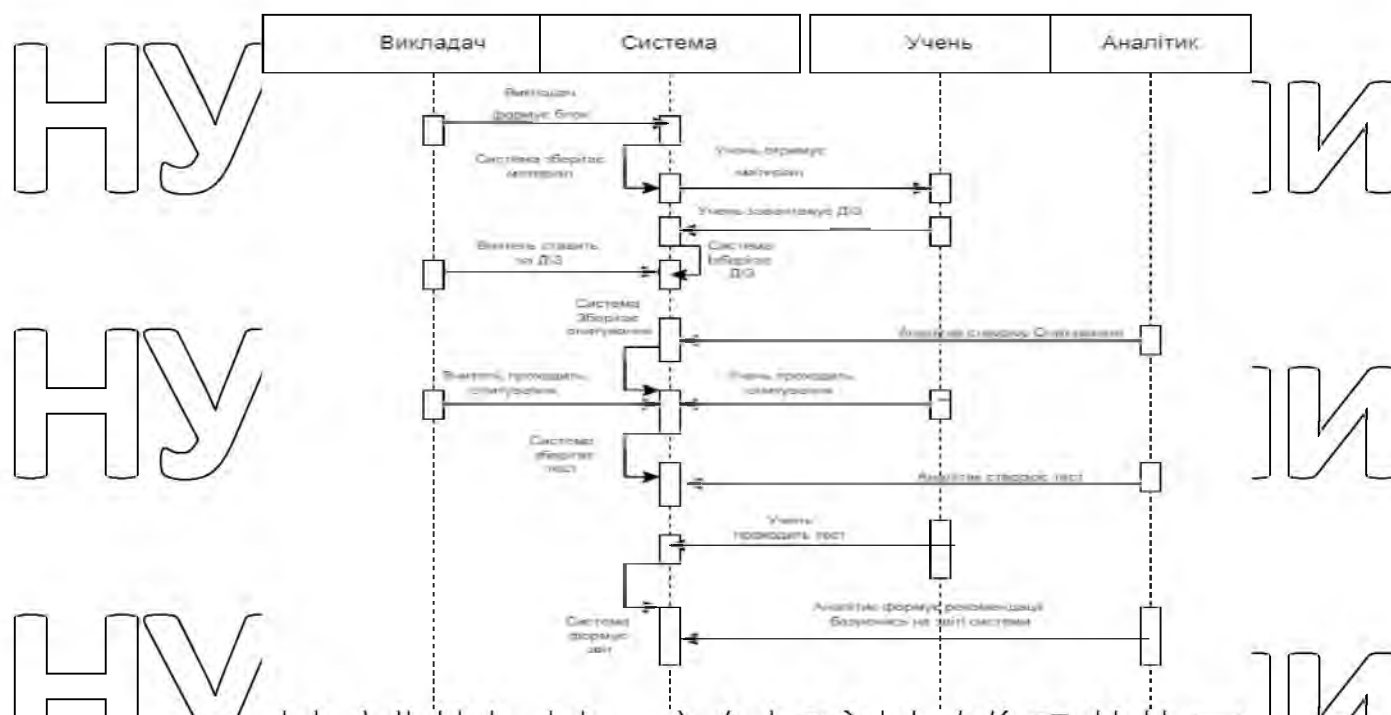


Рис. 5 Діаграма послідовності

Діаграма послідовності на рисунку 5 має такі об'єкти:

- Аналітик;
- Система;
- Викладач;
- Учень;

Часова послідовність дій може відрізнитись за термінами, але по крокам при звичайному розвитку подій завжди буде мати наступні кроки.

1. Об'єкт «Викладач» взаємодіє з системою, формує блок, перевіряє завдання, які виконує учень. Та проходить опитування, де оцінює учня по неформальному оцінюванню.
2. Об'єкт «Система» взаємодіє з рештою об'єктів, та є корінним об'єктом. Вона надає змогу об'єктам взаємодіяти між собою. Зберігає та аналізує інформацію, формує звіти.
3. Об'єкт «Учень» отримує від системи завдання вчителя, виконує їх та завантажує в систему. Проходить опитування та тести.
4. Об'єкт «Аналітик» взаємодіє з системою створюючи тести та опитування. На базі звітів створює рекомендації.



### 2.2.3 Діаграма компонентів

На рисунку 6 зображено діаграму компонентів розробленої системи, видно всі елементи системи, їх кількість та як вони між собою пов'язані.

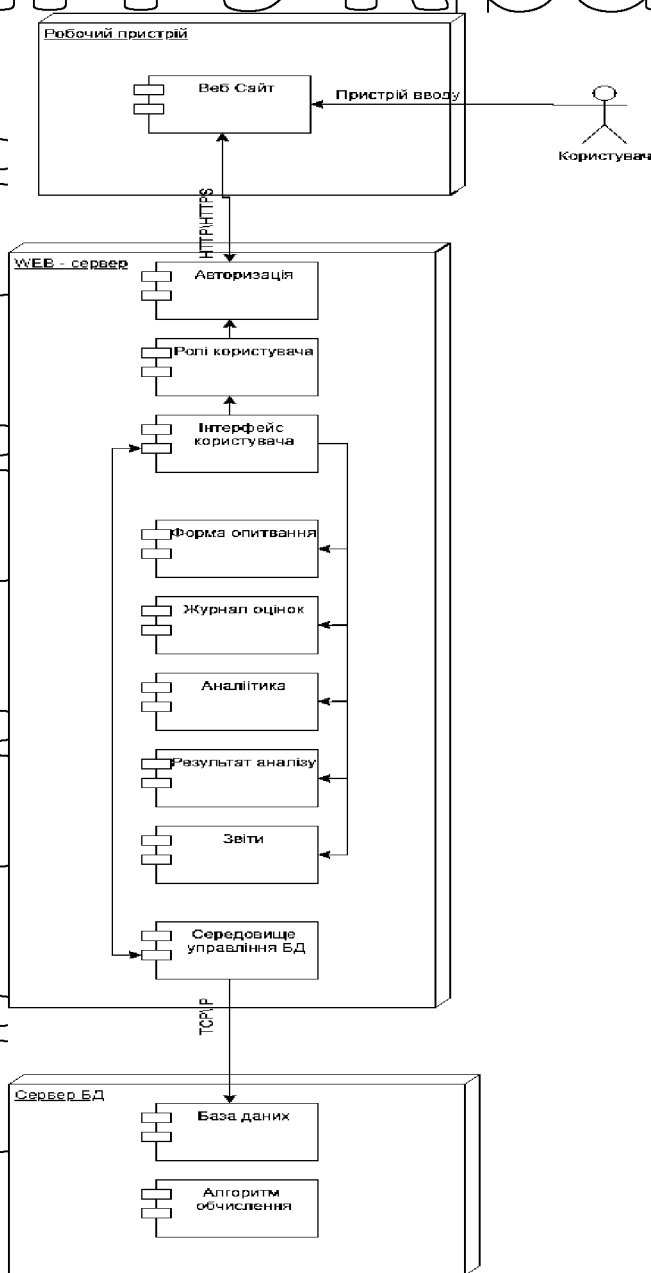


Рис. 6 Діаграма компонентів

### 2.2.4 DFD діаграма

Діаграми потоків даних (Data Flow Diagrams - DFD) є ієрархію функціональних процесів, пов'язаних потоками даних. Мета такого подання - продемонструвати, як кожен процес перетворить свої вхідні дані у вихідні, а також виявити відносини між цими процесами. [26]

Для побудови DFD традиційно використовуються дві різні нотації, які відповідають методам Йордона - ДеМарко і Гейне-Серсона. Ці нотації незначно відрізняються один від одного графічним зображенням символів. Відповідно до даних методами модель системи визначається як ієрархія діаграм потоків даних, описують асинхронний процес перетворення інформації від її введення в систему до видачі споживачеві. [26]

Практично будь-який клас систем успішно моделюється за допомогою DFD-орієнтованих методів. Вони з самого початку створювалися як засіб проектування інформаційних систем (тоді як SADT - як засіб моделювання систем взагалі) і мають більш багатий набір елементів, які адекватно відображають специфіку таких систем. [26]

Метод моделювання IDEF3, який є частиною сімейства стандартів IDEF, був розроблений в кінці 1980-х років для закритого проекту BBC США. Цей метод призначений для таких моделей процесів, в яких важливо зрозуміти послідовність виконання дій і взаємозалежності між ними. [26]

Хоча IDEF3 і не досяг статусу федерального стандарту США, він набув широкого поширення серед системних аналітиків як доповнення до методу функціонального моделювання IDEF0 (моделі IDEF3 можуть використовуватися для деталізації функціональних блоків IDEF0, що не мають діаграм декомпозиції). Основою моделі IDEF3 служить так званий сценарій процесу, який виділяє послідовність дій і підпроцесів аналізованої системи. [26]

На рисунку 7 можна переглянути які дані використовує алгоритм для аналізу і формування звітів.

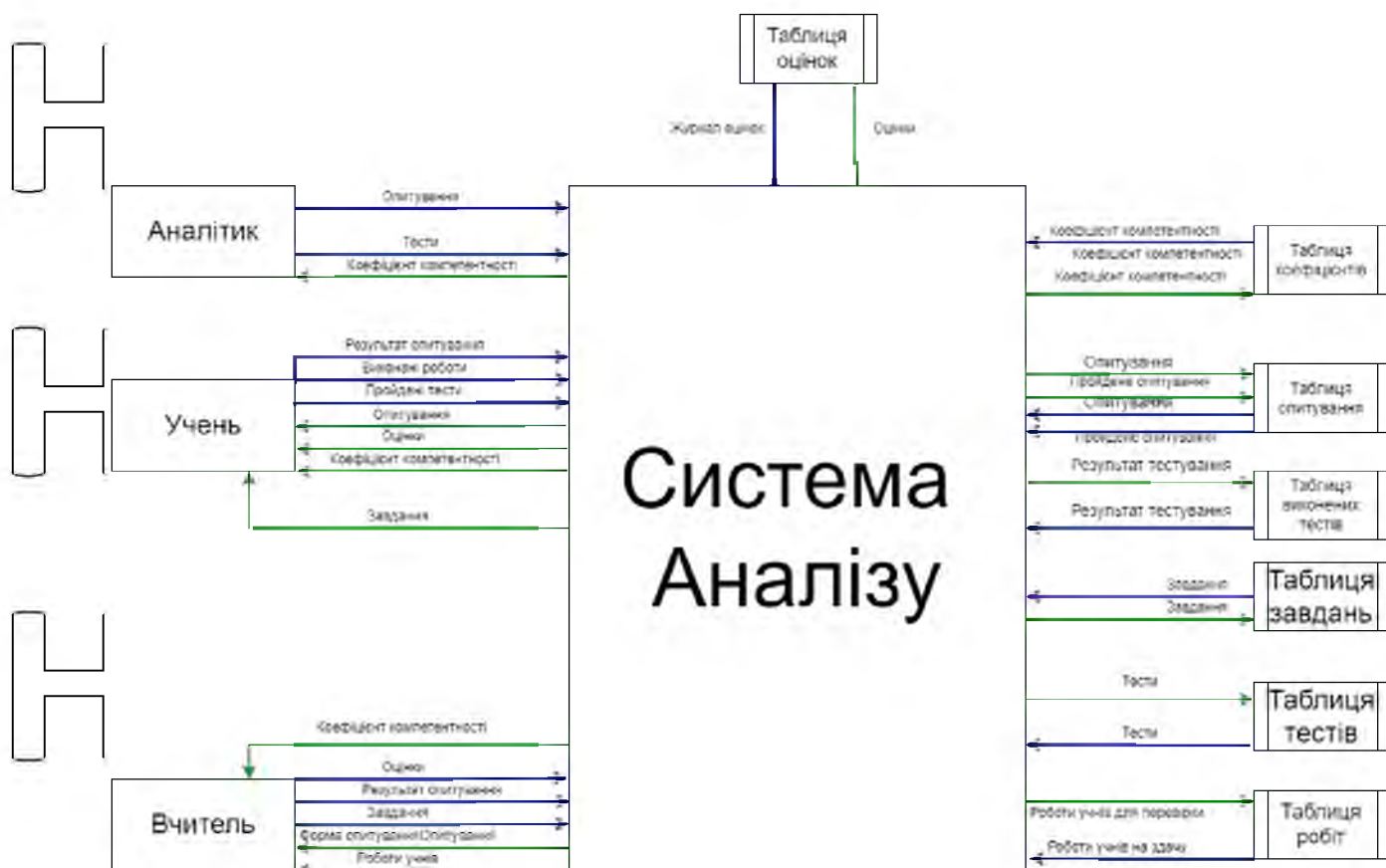


Рис. 7 DFD-діаграма потоків даних в системі аналізу

### 2.3 Абстракція предметної області

Абстракція виділяє істотні характеристики деякого об'єкта, що відрізняють його від всіх інших видів об'єктів, таким чином, чітко описує його концептуальні межі з точки зору спостерігача [23].

Для розробленої системи були спроектовані абстракції, що представлені на рисунку 8-11.

Абстракція «Учень»

Важливі властивості:

Логін

Пароль

Роль

Обов'язки:

Перегляд навчального плану

Завантаження робіт

Перегляд оцінок

Рис. 8 Абстракція «Учень»

Абстракція «Аналітик»

Важливі властивості:

Логін

Пароль

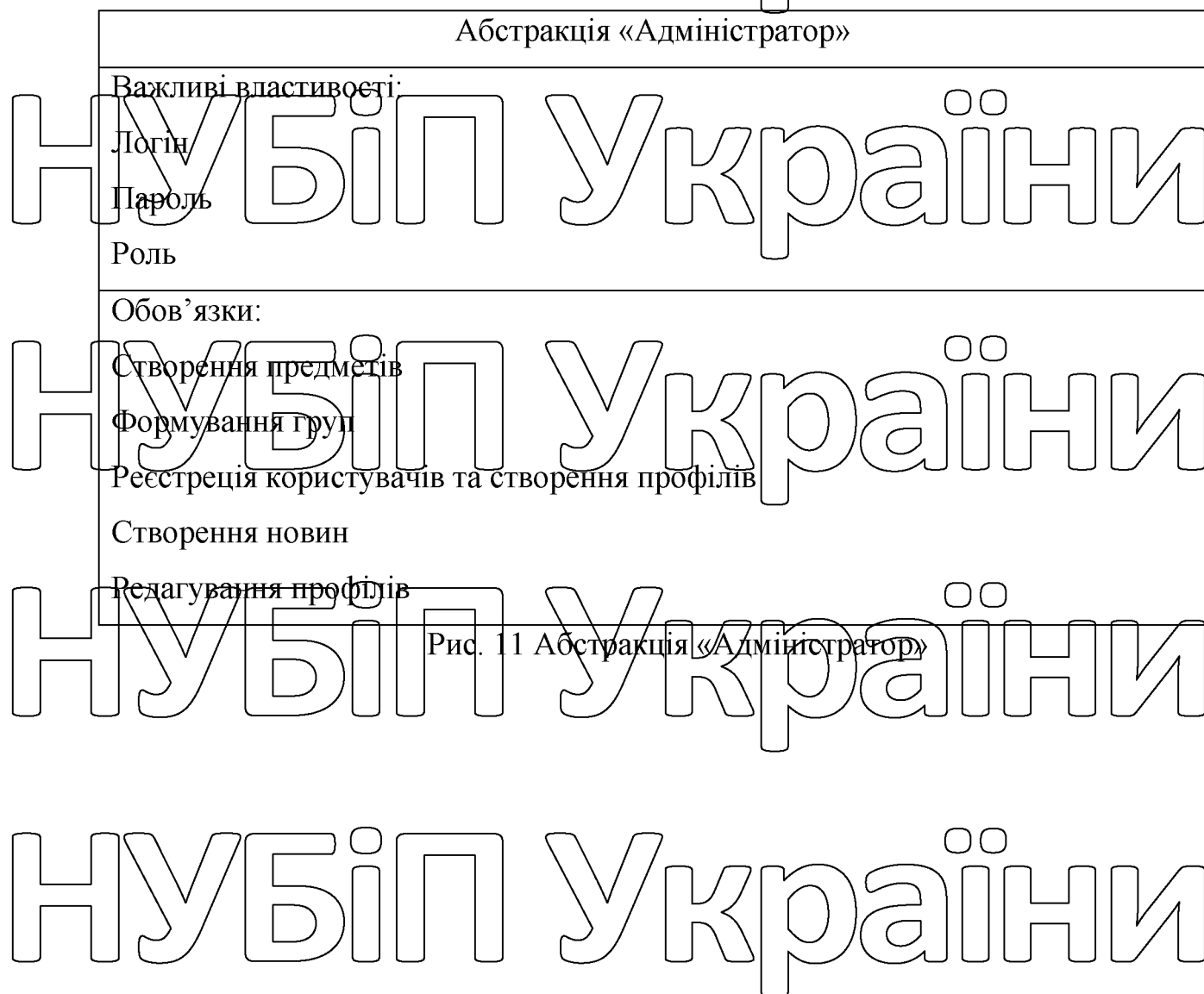
Роль

Обов'язки:

Аналіз отриманої інформації

Створення стандартів оцінювання освітньої компетентності

Рис. 9 Абстракція «Аналітик»



## 3 РОЗРОБКА СИСТЕМИ АНАЛІЗУ

### 3.1. OLAP-технології

OLAP-системи (On-Line Analytical Processing) представляють інструмент для аналізу великих обсягів даних в режимі реального часу на основі гнучкого перегляду, отримання довільних зрізів даних і виконання аналітичних операцій деталізації, згортки, порівняння в часі і ін. Якщо системи регламентованої звітності дозволяють відповісти на питання: "Який дохід 17 підприємства за останній квартал?", то OLAP-системи дають відповідь на питання: "Якого постачальника вигідно вибрати і чому?"

Вимоги до OLAP-систем сформульовані в тесті FASMI (Fast Analysis of Shared Multidimensional Information - швидкий аналіз розділяється багатовимірної інформації):

- Fast (швидкий) - надання результатів аналізу за прийнятний час (зазвичай не більше п'яти секунд), нехай навіть ціною менш детального аналізу
- Analysis (аналіз) - можливість здійснення будь-якого логічного і статистичного аналізу, характерного для цього додатка, і його збереження в доступному для кінцевого користувача вигляді;
- Shared (розділяється) - розрахований на багато користувачів доступ до даних з підтримкою відповідних механізмів блокувань і засобів авторизованого доступу;
- Multidimensional (багатовимірний) - багатовимірне концептуальне представлення даних, включаючи повну підтримку для ієрархій та множинних ієрархій;
- Information (інформація) - можливість звертатися до будь-якої потрібної інформації незалежно від її обсягу і місця зберігання

В багатовимірних СД поряд з вихідними даними зберігаються агреговані дані, з метою скорочення часу виконання складних запитів, що призводять до великих обсягів обчислень по детальним даними.

Всі OLAP-системи діляться на три класи за типом вихідної БД;

- MOLAP (Multidimensional OLAP) - багатовимірний OLAP (використовуються багатовимірні БД). Застосування доцільно за умови не дуже високих обсягів вихідних даних для аналізу, стабільному наборі інформаційних вимірювань, а також в тих випадках, коли час відповіді системи на нерегламентовані запити є найбільш критичним параметром.

- ROLAP (Relational OLAP) - реляційний OLAP (використовуються реляційні БД). Набули поширення дві основні схеми реалізації: схема 18 "зірка" і схема "сніжинка". Переваги: в більшості корпоративних ХД використовуються реляційні СУБД і інструменти ROLAP працюють безпосередньо над ними; наявність великих можливостей захисту даних і розмежування прав доступу користувачів. Недолік - менша в порівнянні з MOLAP продуктивність.

- HOLAP (Hybrid OLAP) - гібридний OLAP. Використовуються як багатовимірні, так і реляційні бази даних: найбільш затребувані агреговані бізнес-показники зберігаються в багатовимірному просторі, а ресурсомісткі детальні дані - в реляційному. Як недоліки можна виділити: складність синхронізації регламентів завантаження, модифікації алгоритмів узгодження і верифікації моделей консолідації, розподілу прав доступу до інформаційних ресурсів системи.

Перевагами OLAP-систем є:

- простота використання і сприйняття зведених таблиць;
- повнота аналітичних даних;
- повна і легка настройка звіту користувачем без залучення програміста;

- можливість деталізувати звіт в процесі аналізу даних;
- швидке формування звітів;
- несутеречливість даних між звітами;
- консолідація інформації з різних БД;
- підвищений захист даних [14].

### 3.2. Архітектура системи

Архітектура системи дослідження впливу електронних освітніх ресурсів на формування освітньої компетентності учнів представлена на рисунку 12. На ньому представлено три основні вузли на яких розміщена система, а саме:

- Робоча станція вчителя, що призначена для роботи викладача та містить модуль роботи з введеними даними.
- Робоча станція аналітика, що призначена для роботи аналітика з наявною інформацією в системі та містить модуль введення даних.
- Сервер має обробляти всі запити, що надходять з робочих станцій та містить базу даних, сховище даних та модуль обробки запитів

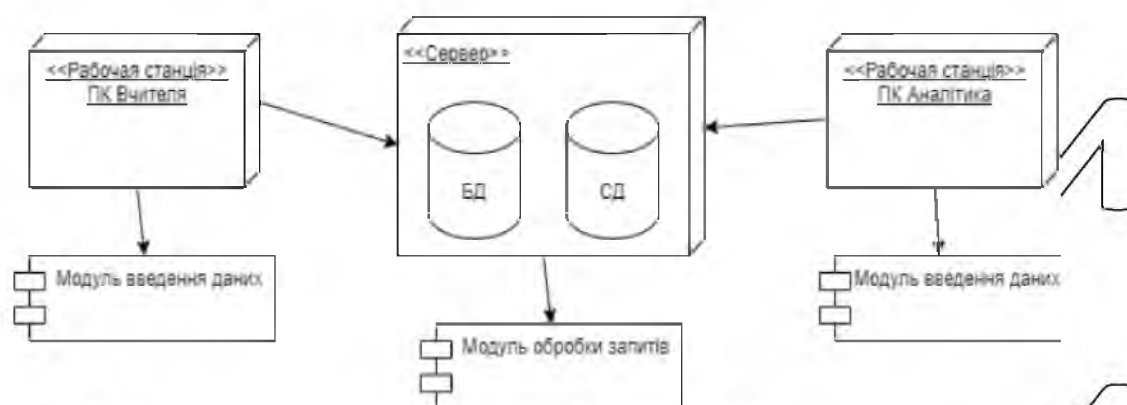


Рис. 12 Топологія системи

### 3.3. Опис бази даних

Система дослідження впливу електронних освітніх ресурсів на формування освітньої компетентності учнів на основі раніше розробленого програмного забезпечення системи з використанням веб-технологій. Отже



джерелом даних системи виступає оперативна база даних, схема якої представлена на рисунку 13. БД має такі сутності:

1. Аналітик – сутність призначена зберігати дані про користувача-аналітика та має такі поля:

- ID\_Аналітик (int) – поле, що виступає первинним ключем;
- Прізвище (String) – поле, що містить дані про прізвище аналітика;
- Ім'я (String) – поле, що містить дані про ім'я аналітика;
- По-батькові (String) – поле, що містить дані про по-батькові аналітика; Логін (String) – поле, що містить дані про логін

аналітика;

- Ел. пошта (String) – поле, що містить дані про електронну пошту аналітика;

- Пароль (String) – шифроване поле, що містить дані про пароль

аналітика

2. Статистика – сутність призначена зберігати дані про статистику навчально закладу та має такі поля:

- ID\_Статистика (int) – поле, що виступає первинним ключем;
- ID\_Учень (int) – поле, що виступає зовнішнім ключем та пов'язує сутність з іншою сутністю Учень;
- Кількість відвіданих занять (int) – поле, що містить дані про кількість відвіданих занять учнем;
- Середня оцінка (int) – поле, що містить дані про середнє 12

значення отриманих учнями оцінок;

- Ефективність (int) – поле, що містить дані про розраховану ефективності;

- Час на виконання завдання (int) – поле, що містить дані про час, що

учень витратив на завдання;

- Бали за пройдені опитування (int) – поле, що містить дані про оцінку учня отриману під час опитування;

- Час на вивчення матеріалу (int) – поле, що містить дані про час, що був витрачений учнем на закріплення матеріалу;
- Коефіцієнт успішності (int) – поле, що містить дані про розрахований коефіцієнт успішності.

3. Просунутий журнал оцінок – сутність призначена зберігати дані про журнал оцінок та має такі поля:

- ID\_Учень (int) – поле, що виступає зовнішнім ключем та пов'язує сутність з іншою сутністю Учень;
- ID\_Оцінки (int) – поле, що виступає зовнішнім ключем та пов'язує сутність з іншою сутністю Оцінка;
- ID\_Журнала (int) – поле, що виступає первинним ключем;
- Прізвище (String) – поле, що містить дані про прізвище учня;
- Ім'я (String) – поле, що містить дані про ім'я учня;
- По\_батькові (String) – поле, що містить дані про по\_батькові учня;
- Час на виконання завдання (int) – поле, що містить дані про час, що учень витратив на завдання;
- Час на вивчення матеріалу (int) – поле, що містить дані про час, що був витрачений учнем на закріплення матеріалу;
- Коефіцієнт успішності (int) – поле, що містить дані про розрахований коефіцієнт успішності;
- Коефіцієнт розуміння матеріалу (int) – поле, що містить дані про розрахований коефіцієнт розуміння матеріалу.

4. Оцінка – сутність призначена зберігати дані про оцінку учня та має такі поля:

- ID\_Оцінки (int) – поле, що виступає первинним ключем;
- Кількість балів (int) – поле, що містить дані про оцінку.

5. Робота учня – сутність призначена зберігати дані про виконання завдання учня та має такі поля:

- ID\_Роботи (int) – поле, що виступає первинним ключем;

• Об'єм файлу (String) – поле, що містить дані про шлях, де зберігається файл з роботою учня.

6. Завдання – сутність призначена зберігати дані про завдання та має такі поля:

- ID\_Завдання (int) – поле, що виступає первинним ключем;

- Постанова (String) – поле, що містить дані про саме завдання.

7. Вчитель – сутність призначена зберігати дані про користувача вчителя та має такі поля:

- ID\_Вчитель (int) – поле, що виступає первинним ключем;

- Прізвище (String) – поле, що містить дані про прізвище вчителя;

- Ім'я (String) – поле, що містить дані про ім'я вчителя;

- По-батькові (String) – поле, що містить дані про по-батькові

вчителя; • Логін (String) – поле, що містить дані про логін вчителя;

• Ел. пошта (String) – поле, що містить дані про електронну пошту вчителя;

- Пароль (String) – шифроване поле, що містить дані про пароль вчителя.

8. Предмет викладання – сутність призначена зберігати дані про користувача-аналітика та має такі поля:

- ID\_Вчитель (int) – поле, що виступає зовнішнім ключем та пов'язує сутність з іншою сутністю Вчитель;

- ID\_Предмет (int) – поле, що виступає первинним ключем;

- Назва (String) – поле, що містить дані про назву предмету.

9. Учень – сутність призначена зберігати дані про користувача-учня та має такі поля:

- ID\_Учень (int) – поле, що виступає первинним ключем;

- ID\_Клас (int) – поле, що виступає зовнішнім ключем та пов'язує сутність з іншою сутністю Клас;

- Прізвище (String) – поле, що містить дані про прізвище учня;

• Ім'я (String) – поле, що містить дані про ім'я учня;

• По-батькові (String) – поле, що містить дані про по-батькові учня;

- Ім'я (String) – поле, що містить дані про ім'я учня;
- По-батькові (String) – поле, що містить дані про по-батькові;
- Логін (String) – поле, що містить дані про логін учня;
- Ел. пошта (String) – поле, що містить дані про електронну пошту

учня.

10. Клас – сутність призначена зберігати дані про клас та має такі поля:

- ID\_Клас (int) – поле, що виступає первинним ключем;
- Номер класу (int) – поле, що містить дані про номер класу;
- Кількість учнів (int) – поле, що містить дані про кількість учнів що

навчаються в класі.

11. Форма опитування – сутність призначена зберігати дані про опитування та має такі поля:

- ID\_Форми (int) – поле, що виступає первинним ключем;
- ID\_Учень (int) – поле, що виступає зовнішнім ключем та пов'язує сутність з іншою сутністю Учень;
- ID\_Предмет (int) – поле, що виступає зовнішнім ключем та пов'язує сутність з іншою сутністю Предмет;

- ID\_Вчитель (int) – поле, що виступає зовнішнім ключем та пов'язує сутність з іншою сутністю Вчитель;

- Питання (String) – поле, що містить дані про питання; 15
- Час на самостійне опрацювання матеріалів (int) – поле, що містить дані про час витрачений на опрацювання матеріалів;

- Назва (String) – поле, що містить дані про назву опитування.

12. Навчальна програма – сутність призначена зберігати дані про навчальну програму та має такі поля:

- ID\_Навчальна програма (int) – поле, що виступає первинним ключем;

- ID\_Вчителя (int) – поле, що виступає зовнішнім ключем та пов'язує сутність з іншою сутністю Вчитель;

• Кількість уроків (int) – поле, що містить дані про кількість уроків в навчальному плані;

• Кількість завдань (int) – поле, що містить дані про кількість завдань в навчальному плані;

• Навчальні матеріали (String) – поле, що містить дані про навчальні матеріали.

13. Адміністратор – сутність призначена зберігати дані про користувача-адміністратора та має такі поля:

• ID\_Адміністратор (int) – поле, що виступає первинним ключем;

• Прізвище (String) – поле, що містить дані про прізвище адміністратора;

• Ім'я (String) – поле, що містить дані про ім'я адміністратора;

• По-батькові (String) – поле, що містить дані про по-батькові адміністратора;

• Логін (String) – поле, що містить дані про логін адміністратора;

• Ел. пошта (String) – поле, що містить дані про електронну пошту адміністратора;

• Пароль (String) – шифроване поле, що містить дані про пароль адміністратора.

НУБІП України

НУБІП України

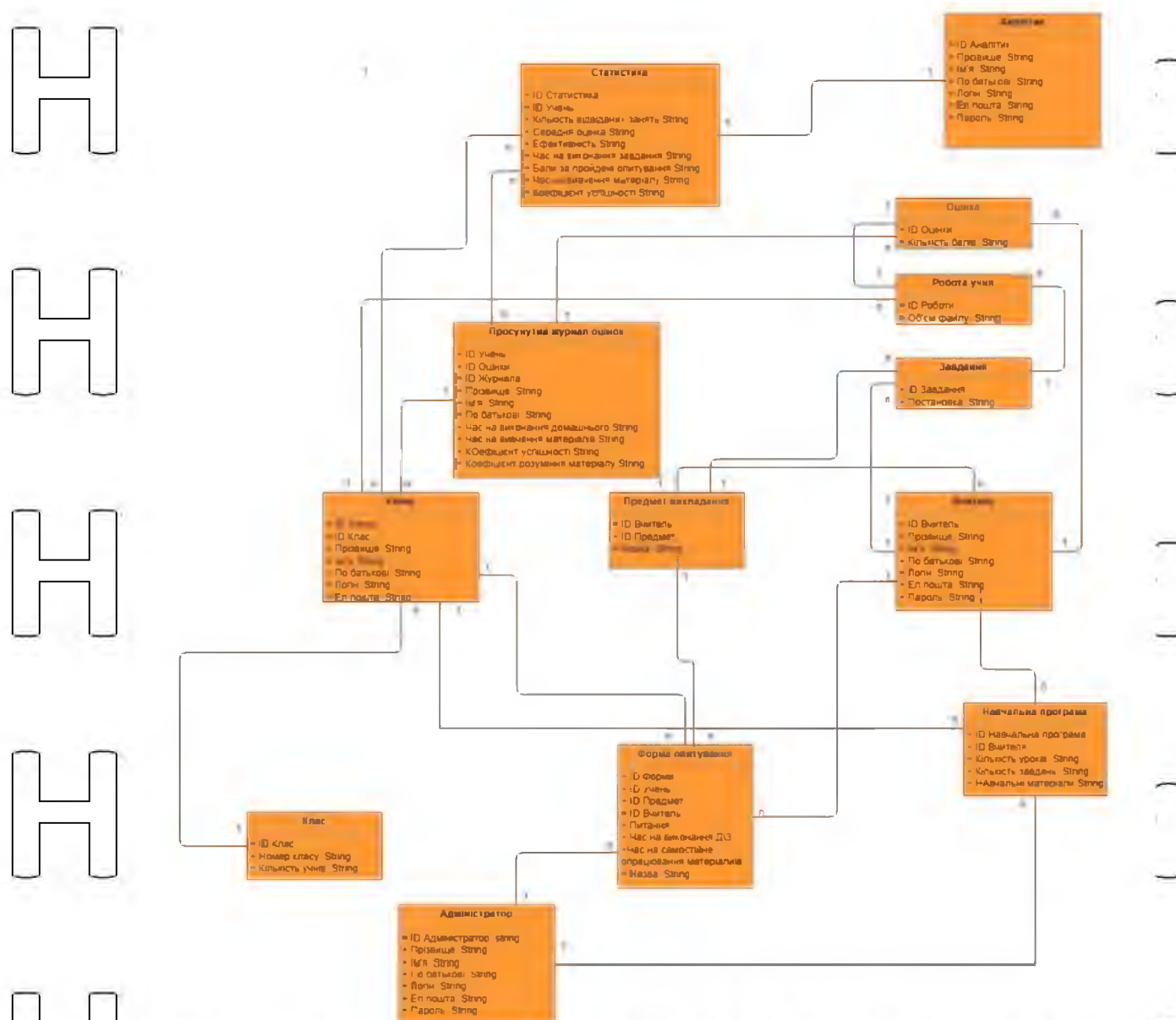


Рис. 13 Схеми БД

### 3.4. Структура сховища даних

Головним етапом створення системи є створення сховища даних (СД), наповнення якого буде відбуватись за рахунок оперативної БД. На рисунку 14 представлено створене СД.

1. Task\_Dim – вимір, що зберігає дані про результати оцінювання виконаних завдань учнем та має такі поля:

- Id task – унікальне значення даного виміру;
- Time – час, що був витрачений на виконання завдання;

- Mark – оцінка за виконані завдання;
- Id\_student – зовнішній ключ, що належить виміру Student\_Dim;
- Date – дата оцінювання завдання;
- Id\_subject – зовнішній ключ, що належить виміру Subject\_Dim;
- Amount\_task – загальна кількість завдань.

2. Poll\_Dim – вимір, що зберігає дані про результати опитування учня та має такі поля:

- Id\_poll – унікальне значення даного виміру;
- Time – час витрачений на засвоєння матеріалу;

- Mark – оцінка опитування;
- Date – дата опитування;
- Id\_student – зовнішній ключ, що належить виміру Student\_Dim;
- Id\_subject – зовнішній ключ, що належить виміру Subject\_Dim.

3. Student\_Dim – вимір, що зберігає дані про учня та має такі поля:

- Id\_student – унікальне значення даного виміру;
- Name\_student – ім'я студента;
- Id\_class – зовнішній ключ, що належить виміру Class\_Dim.

4. Class\_Dim – вимір, що зберігає дані про клас та має такі поля:

- Id\_class – унікальне значення даного виміру;
- Name\_class – назва класу.

5. Date\_Dim – вимір, що зберігає дані про дати звіту та має такі поля:

- Id\_date – унікальне значення даного виміру;
- Date\_quarter – дати звітів.

6. Subject\_Dim – вимір, що зберігає дані про предмет та має такі поля:

- Id\_subject – унікальне значення даного виміру;
- Name\_subject – назва предмету.

7. Progress\_Fact – таблиця-фактів, що має такі поля:

- Id\_date – унікальне значення даного виміру;
- Id\_subject – унікальне значення даного виміру;

- Id\_class – унікальне значення даного виміру;
- Progress\_by\_poll – успішність опитування;
- Progress\_by\_task – успішність оцінок завдань;
- Overall\_progress – загальна успішність.



Рис. 14 Структура сховища даних

Особливістю СД в порівнянні з оперативною БД полягає в тому, що таблиці-виміри оновлюються не часто. Операцію по наповненню таблиці-вимір Subject Dim представлено на рисунку 15

```

INSERT INTO [dbo].[Subject_Dim]
    ([Id_subject]
    ,[Name_subject])
VALUES
    ('1','Математика'),
    ('2','Біологія'),
    ('3','Англійська мова'),
    ('4','Українська мова'),
    ('5','Фізика')
GO
  
```

Рис. 15 SQL запит наповнення таблиці-вимір Subject Dim.

Наповнення інших таблиць-вимірів продемонстровано на рисунках 8-11.



Id_student	Name_student	Id_class
101	Закревський Подолян	1
102	Царук Шарль	1
103	Орлик Вишата	1
104	Парубок Яків	1
105	Кушніренко Немир	1
106	Яворський Гатусил	1
107	Ситенко Марко	1
108	Андрушко Акилина	1
109	Павличенко Мальвіна	1
110	Балог Незабудка	1
201	Носковська Квітослава	2
202	Ганкевич Магадара	2
203	Качура Бажана	2
204	Карпека Євгенія	2
205	Снігур Анастасія	2
206	Ніколайчук Йосип	2
207	Горняткевич Русудан	2
208	Киричук Лад	2
209	Сенькович Боян	2
210	Горбаренко Олесь	2

Рис. 16 Наповнення таблиці-вимір Student\_Dim

Id_date	Date_quarter	Id_class	Name_class
1	2020-10-25	1	206-a
2	2020-12-25	2	207-a
3	2021-03-25	3	208-a
4	2021-05-25	4	206-b
5	2021-05-25	5	207-b

Рис. 17 Наповнення таблиць-вимірів Date\_Dim та Class\_Dim відповідно.

Id_poll	Time	Mark	Date	Id_student	Id_subject	Id_task	Time	Mark	Id_student	Date	Id_subject	Amount_task
1	45	10	2020-09-02	101	1	1	45	9	201	2020-09-02	1	10
2	45	11	2020-09-09	101	1	2	45	9	201	2020-09-09	1	10
3	45	10	2020-09-16	101	1	3	50	10	201	2020-09-16	1	10
4	45	10	2020-09-23	101	1	4	45	10	201	2020-09-23	1	10
5	45	10	2020-09-30	101	1	5	55	9	201	2020-09-30	1	10
6	45	12	2020-10-14	101	1	6	50	9	201	2020-10-14	1	10
7	45	11	2020-10-21	101	1	7	40	10	201	2020-10-21	1	10
8	45	11	2020-10-28	101	1	8	35	8	201	2020-10-28	1	10
9	45	8	2020-11-04	101	1	9	45	8	201	2020-11-04	1	10
10	45	9	2020-11-11	101	1	10	40	9	201	2020-11-11	1	6
11	45	10	2020-11-18	101	1	11	40	9	201	2020-11-18	1	6
12	45	11	2020-11-25	101	1	12	45	9	201	2020-11-25	1	6
13	45	9	2020-12-02	101	1	13	35	9	201	2020-12-02	1	6
14	45	10	2020-12-09	101	1	14	45	10	201	2020-12-09	1	6
15	45	12	2020-12-16	101	1	15	50	9	201	2020-12-16	1	6
16	45	12	2020-09-03	101	2	16	45	9	201	2020-09-03	2	10
17	45	11	2020-09-10	101	2	17	40	11	201	2020-09-10	2	10
18	45	10	2020-09-17	101	2	18	40	8	201	2020-09-17	2	10
19	45	10	2020-09-24	101	2	19	45	10	201	2020-09-24	2	10
20	45	10	2020-10-01	101	2	20	40	9	201	2020-10-01	2	10

Рис. 18 Наповнення таблиць-вимірів Poll\_Dim та Task\_Dim відповідно.

### 3.5. Створення OLAP куба

OLAP куб – багатомірна структура, що значно прискорює процес аналізу за рахунок раніше агрегованих даних. Для розкортання куба скористаємось



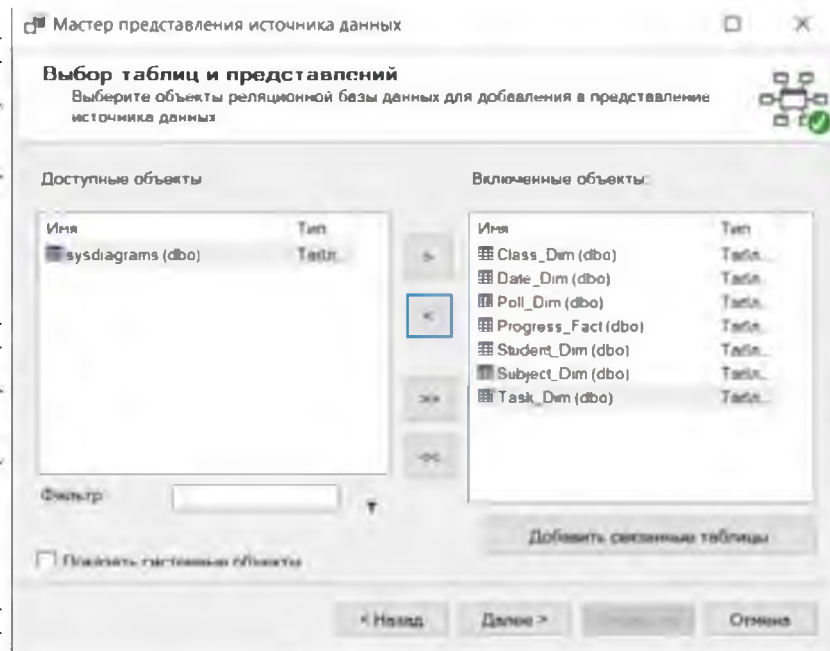


Рис. 20 Уявлення джерела даних

Важливим аспектом розгортання є правильно створені виміри. На рисунку 21. представлено перший вимір – Date\_Dim, що містить дату чверті. Він дає можливість проводити аналіз у розрізі різних часових періодах.

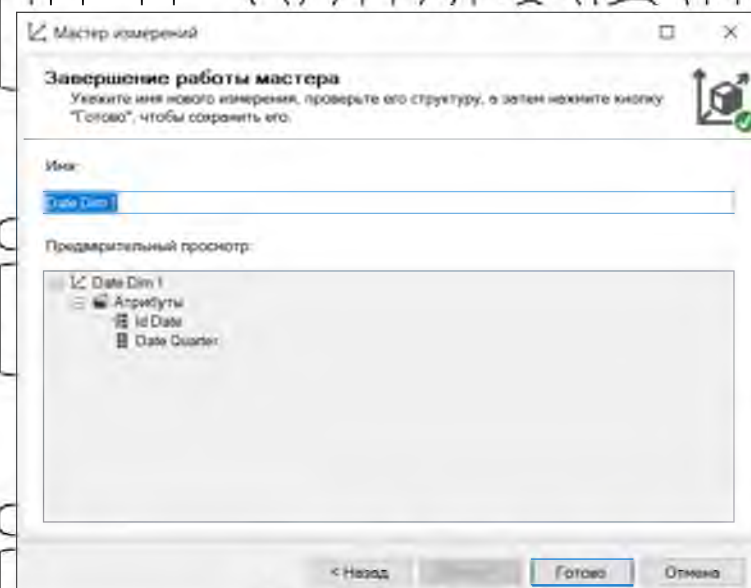


Рис. 21 Вимір Date\_Dim

Наступним виміром створюємо Subject\_Dim. Структура виміру представлено на рисунку 22



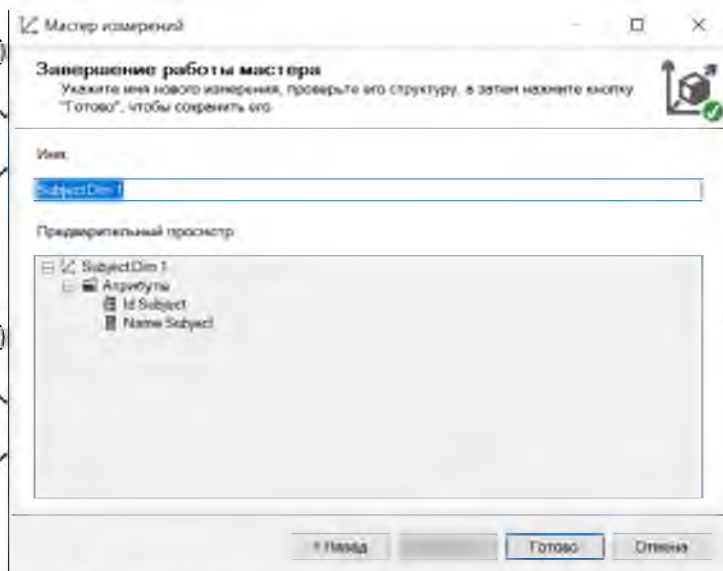


Рис. 22 Структура виміру Subject\_Dim

На рисунку 23 представлено структуру Poll\_Dim, що підтягує виміри Student\_Dim та Class\_Dim.

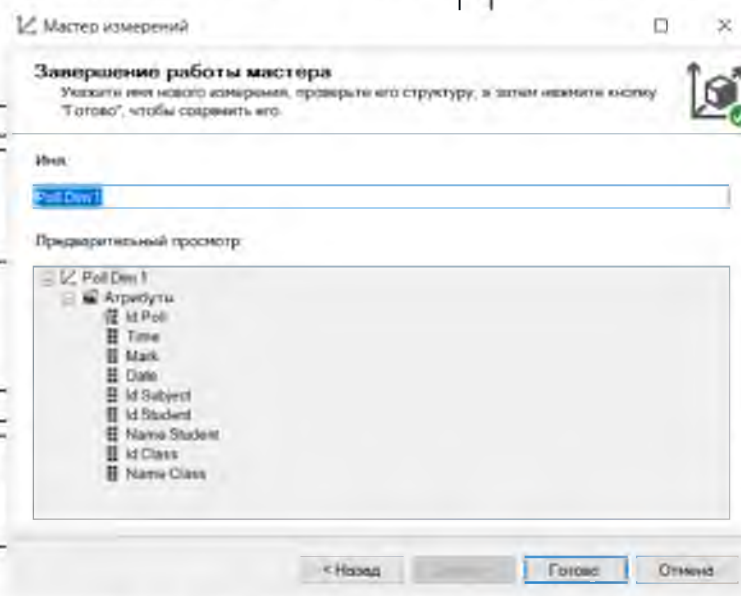


Рис. 23 Структура виміру Poll\_Dim

Останнім виміром створюємо Task\_Dim, що підтягує виміри Student\_Dim та Class\_Dim. Структура виміру наведена на рисунку 24.

НУБІП Україна

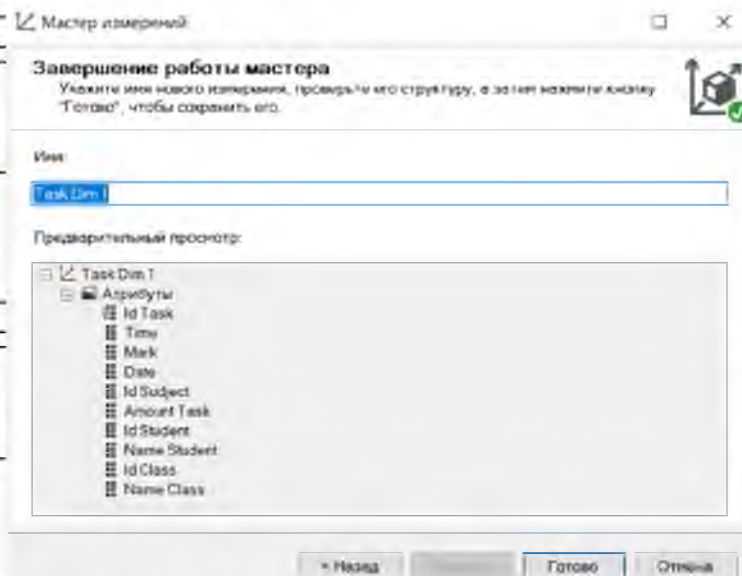


Рис. 24 Структура виміру Task Dim.

Наступним кроком є створення самого куба, для цього потрібно визначити міри, як представлено на рисунку 25.

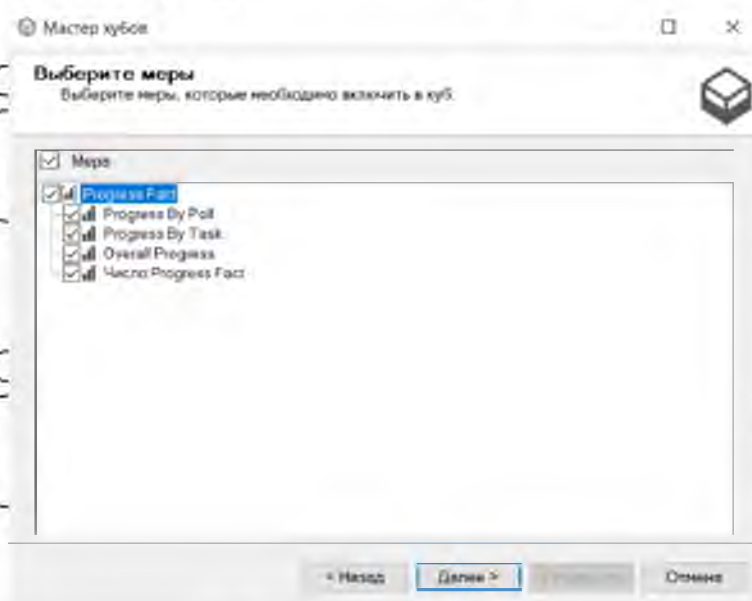


Рис. 25. Відбір мір куба.

В результаті підключення всіх компонентів вимірів та мір, що представлено на рисунку 26., можна розгорнути куб.

НУБІП України

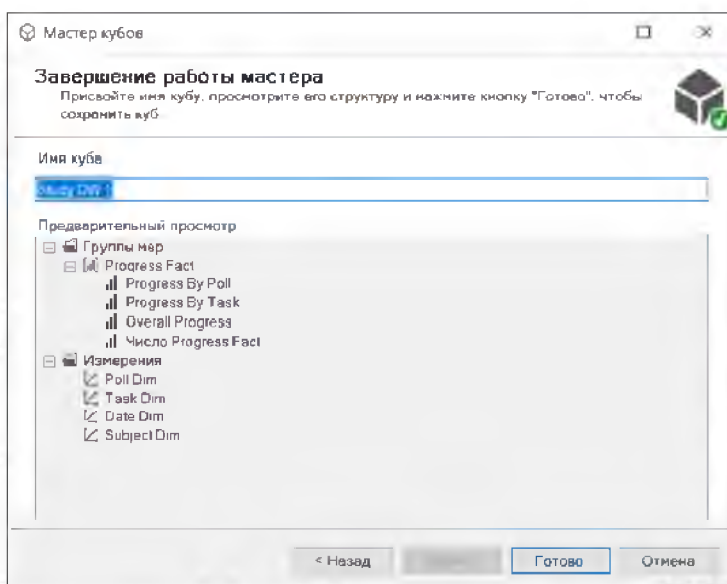


Рис. 27 Структура куба.

В результат успішного обробки та розгортання куба було згенерована багатомірна структура куба, що представлена на рисунку 28.

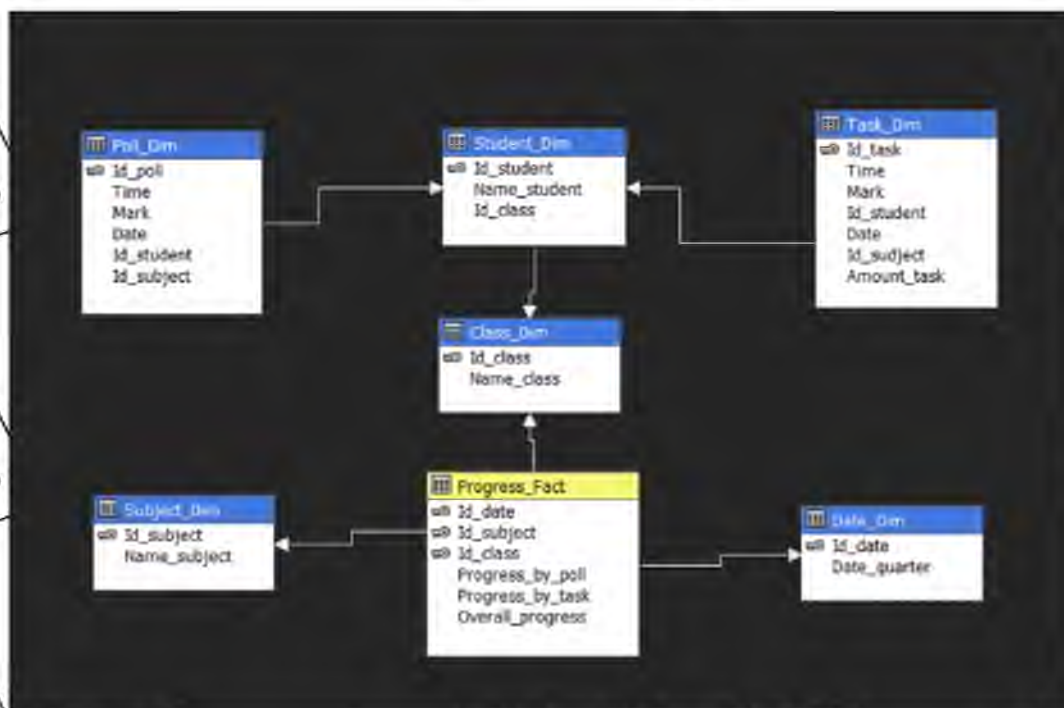


Рис. 28. Розгорнутий куб

### 3.6. Створення потоків даних за допомогою Data Flow

Важливою особливістю багатомірних структур є агреговані дані. Для створення та наповнення агрегованими даними скористаємось службою SSIS - платформа для побудови рішень для інтеграції даних на рівні підприємства та перетворення даних. Служба призначена для вирішення складних бізнесзавдань



шляхом копіювання або завантаження файлів, складених даних, очищення даних, а також керування об'єктами та даними SQL Server [16].

Для початку створимо проєкт, знову скориставшись середовищем Visual Studio для роботи із службою SSIS. Визначившись з функціоналом, що буде покладено на різні потоки даних, розташуємо їх в потрібному порядку, як представлено на рисунку 29.



Рис. 29. Потоки даних проєкту

Далі встановлюємо потрібний порядок функцій в кожному потоці. На рисунку 30 представлено блоки першого потоку.

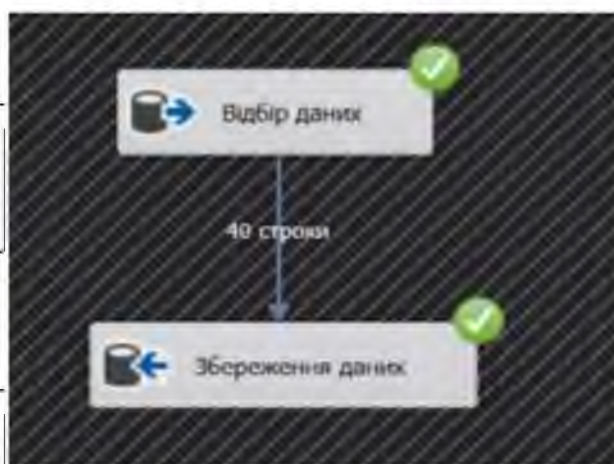


Рис. 30. Перший потік даних

Розглянемо блоки з рисунку 30. більш детально. Їх функціонал представлено на рисунку 31:

- Відбір даних – блок призначений сформувати набір даних, над якими будуть проводитись операції.

- **Збереження даних** – блок призначений записування даних у вказані таблиці.

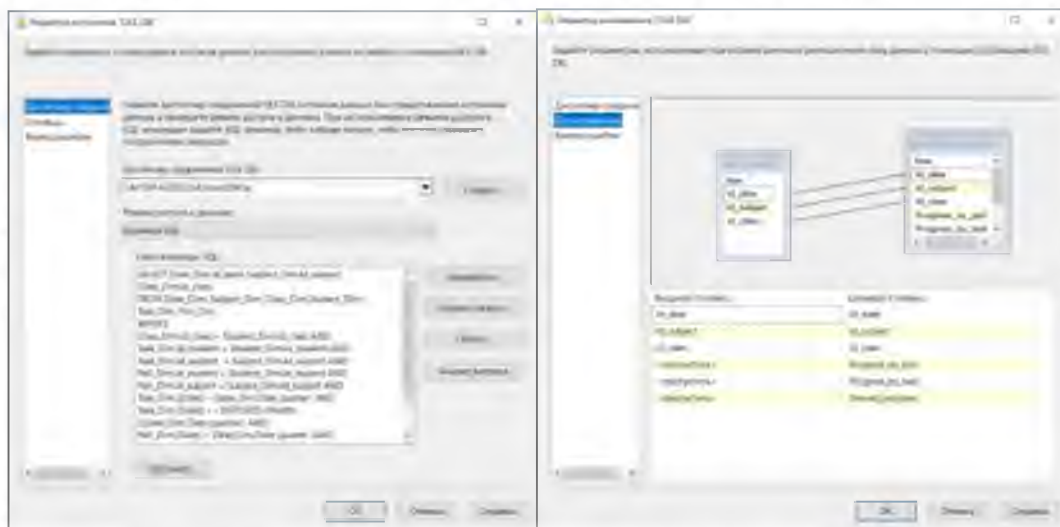


Рис. 31. Блоки «Відбір даних» та «Збереження даних».

На рисунку 32. представлено 2 потік даних, а на рисунку 33. представлено 3 потік що містить блок з розрахунками.



Рис 32. Структура 2 потоку даних.

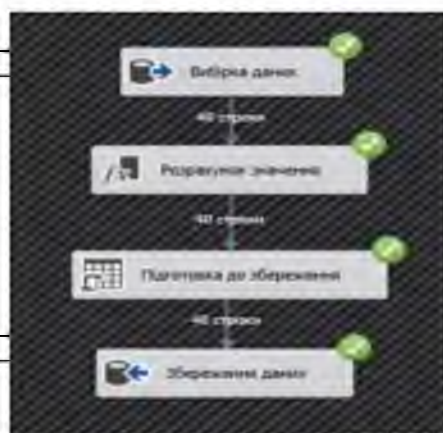


Рис 33. Структура 3 потоку даних.



Блоки 3 потоку даних детальніше розглянуто на рисунку 34.

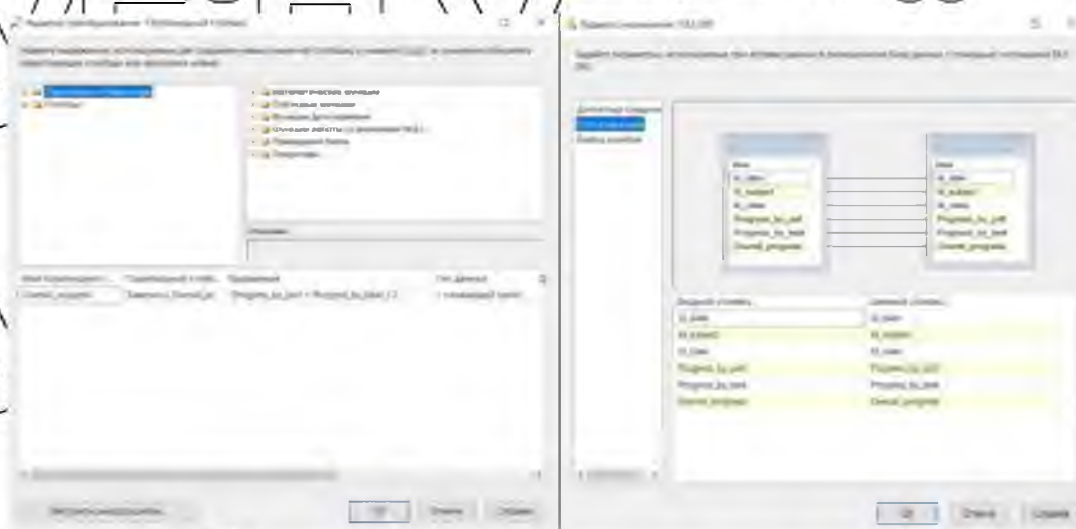


Рис. 34. Блоки «Створення набору даних» та «Збереження даних» другого потоку даних

Результат виконання всіх потоків даних приведено на рисунку 35.

	Id_date	Id_subject	Id_class	Progress_by_poll	Progress_by_task	Overall_progress
1	1	1	1	0.190476190476192	0.237551020408164	0.214013605442178
2	1	1	2	0.189841269841271	0.202904968047826	0.196373118944548
3	1	2	1	0.19904761904762	0.245555555555557	0.222301587301588
4	1	2	2	0.197777777777778	0.218097299525872	0.207937538651825
5	1	3	1	0.18825396825397	0.235294578437437	0.211774273345703
6	1	3	2	0.193015873015874	0.219559266130694	0.206287569573284
7	1	4	1	0.191111111111112	0.24092970521542	0.216020408163266
8	1	4	2	0.19015873015873	0.206903834260977	0.198531282209854
9	1	5	1	0.200952380952382	0.244734693877551	0.222843537414967
10	1	5	2	0.202857142857143	0.220140383426098	0.211498763141621
11	2	1	1	0.191388888888889	0.23029761904762	0.210843253968255
12	2	1	2	0.191111111111112	0.21947068903319	0.205290900072151
13	2	2	1	0.197222222222223	0.236531746031746	0.216876984126985
14	2	2	2	0.197777777777779	0.217068709330502	0.207423243554141
15	2	3	1	0.187222222222223	0.228293650793652	0.207757936507937
16	2	3	2	0.195277777777779	0.221647637085138	0.208462707431459
17	2	4	1	0.190277777777778	0.228524531024532	0.209401154401155
18	2	4	2	0.189166666666668	0.219515602453103	0.204341134559885
19	2	5	1	0.197222222222223	0.239464285714286	0.218343253968255
20	2	5	2	0.204722222222223	0.233919552669553	0.219320887445888
21	3	1	1	0.203611111111111	0.250927128427128	0.227269119769119
22	3	1	2	0.200277777777779	0.21485524891775	0.207566513347764
23	3	2	1	0.201388888888889	0.246854978354978	0.224121933621934
24	3	2	2	0.204722222222223	0.218452922077924	0.211587572150073
25	3	3	1	0.213009845288327	0.262087420315269	0.237548632801798

Рис. 35. Наповнення Progress Fact.

### 3.7. Використані технології

1. HTML - скорочення від "HyperText Mark-up Language" - перекладається як "Мова розмітка гіпертексту" (Гіпертекст - це текст, що не послідовно зв'язаний з іншими документами, тобто у вас є змога з першої сторінки документу перейти на останню). Іншими словами HTML - це мова розмітки, або ще один спосіб зберігання інформації. За допомогою HTML ти позначаєш текст, вказуючи своєму веб-переглядачу, як він має розуміти позначений текст, так само як і на жорсткому диску інформація зберігається в блоках, кластерах, секторах, доріжках і тільки за допомогою, такої, визначеної структури твій комп'ютер розуміє, що треба, а що не треба зчитувати[18];

2. CSS (аббревіатура від Cascading Style Sheets, що в перекладі означає каскадні таблиці стилів) - це спеціальна мова (мова стилів), за допомогою якої описують вигляду документів (як і де відобразити елементи веб-сторінки), написаних мовами розмітки даних. Найчастіше CSS використовується для документів, котрі розмічені мовою HTML, XHTML та XML[18];

3. JavaScript - це об'єктно-орієнтована мова скриптів, яка використовується для розробки вбудованих додатків, які виконуються як на стороні клієнта, так і на стороні сервера. Частіше за все вони виконуються в формі JavaScript в клієнтській частині і реалізуються в додатку у вигляді інтегруємого веб-браузером компонента, що дозволяє розробляти покращені інтерфейси і динамічні веб-сайти. JavaScript використовує можливості середовища, в якому виконуються написані на ньому сценарії. JavaScript являється діалектом стандарту ECMAScript і охарактеризований як динамічна мова скриптів. На JavaScript вплинули багато мов. Його не рідко порівнювали з іншими мовами програмування, але в підсумку прийшли до висновку, що JavaScript це в дійсності особлива мова[19].

4. SQL - це непроцедурного мова БД, орієнтований на великий обсяг оброблюваної інформації. Непроцедурного значить, що в ньому, перш за все, приділяється увага які дані викликати, видаляти або вставляти, а не якимось чином це робити. Кількісно-орієнтований ж означає, що за допомогою цієї мови

можна обробляти значні обсяги інформації в групах. [20]

5. PHP – це популярна мова програмування, особливо серед веб-розробників. Автором початкової версії є Рasmus Лердорф. Ідея його полягала у розробці набору інструментів спрощення процесу створення динамічних веб-сторінок. Незважаючи на те, що сучасний PHP є мовою загального призначення, найчастіше його використовують як серверний інструмент для генерації HTML-коду, який потім інтерпретується веб-браузером. Що таке PHP? Це мова програмування з відкритими вихідними кодами, над розвитком якого працюють програмісти-ентузіасти з усього світу. Він має простий синтаксис, частково схожий на Java та C++. Це проект, що постійно розвивається, на даний момент актуальною є 7-ма версія мови. За статистикою, кожен шостий програмний продукт створено на PHP.

#### *Переваги PHP*

- Простий та інтуїтивно зрозумілий синтаксис. PHP швидко освоюють навіть програмісти-новачки. Він увіобраз усі найкращі особливості таких популярних мов, як C, Java та Perl. PHP-код легко читається незалежно від способу використання (для написання невеликих скриптів або створення потужних додатків з використанням об'єктно-орієнтованого підходу до реалізації програми).
- Кросплатформність та гнучкість. PHP сумісний із усіма популярними платформами (Linux, Windows, Mac OS). Написані на ньому програми успішно працюють на різному серверному ПЗ (IIS, Nginx, Apache та багатьох інших).
- Відмінна масштабованість. PHP дозволяє досягти максимальної продуктивності додатків, написаних на ньому, з зростанням апаратних ресурсів. Веб-програми, розподілені на кілька серверів, здатні справлятися із суттєвими навантаженнями (великим трафіком).
- Вбудовуваність у HTML-документи. На просту HTML-сторінку можна легко додати контент, що динамічно змінюється шляхом вставки блоків коду



PHP. Вони додаються подібно до HTML-тегів, не порушують структуру документа.

- Активний розвиток та вдосконалення. Спільнота розробників постійно працює над впровадженням додаткового функціоналу, що розширює можливості мови, спрощенням синтаксису та покращенням захисту від можливих атак.

- Детальна документація. На офіційному сайті проекту представлені повні відомості про кожну функціональну одиницю мови з прикладами використання.

- Простий пошук рішень проблем, що виникають. В інтернеті існує безліч форумів, присвячених програмуванню на PHP.

- Широкі перспективи її подальшого розвитку. Більшість CMS були створені на чистому PHP та фреймворках. Цим зумовлені популярність та затребуваність PHP програмістів.

### 3.8. Послання технологій

Для реалізації даного проекту використовувалось наступне програмне забезпечення:

- Visual Studio Code – використовувався для розробки Front-end частини проекту. Достатньо простий та функціональний інструмент, за допомогою якого можна створювати продукти абсолютно різного масштабу.
- Microsoft Visual Studio – інструмент використовувався для розгортання OLAP – кубу, для створення системи розрахунку КК.
- PHPStorm – використався для написання коду мовою PHP, за допомогою останнього БД вдалось підключити до Front-end частини проекту.
- MySQL – використовувалась для створення БД.

За допомогою розробленого скрипта вдалось через мову PHP на пряму підключити OLAP технологію до БД проекту. Нижче буде наведені кроки, для досягнення цієї цілі.

1. Для встановлення потрібної бібліотеки прописуємо команду в консолі (рис. 36.)

```
composer require kabachello/phpolapi
```

Рис. 36. Встановлення необхідної бібліотеки

2. Підключення

```
<?php
require_once 'vendor/autoload.php';
use phpOLAPI\Xmla\Connection\Connection;
use phpOLAPI\Xmla\Connection\Adaptator\SoapAdaptator;

$conection = new Connection(
    new SoapAdaptator('http://localhost:8080/mondrian/xmla')
    [
        'BdeSchocl' => 'Provider=Mondrian;DataSource=MondrianFoodMart;',
        'Kkko' => 'FoodMart',
        'Statistic' => 'FoodMart'
    ]
);

$conection = new Connection(
    new SoapAdaptator('http://localhost/olap/msmdpump.dll', 'username',
        'password'),
    [
        'DataSourceInfo' => null,
        'CatalogName' => 'Adventure Works DW 2008R2 SA'
    ]
);
```

3. Побудова MDX за допомогою API

```
use phpOLAPI\Mdx\Query;

$query = new Query("[Sales]");
$query->addElement("[Measures].[Unit Sales]", "COL");
$query->addElement("[Measures].[Store Cost]", "COL");
$query->addElement("[Measures].[Store Sales]", "COL");
$query->addElement("[Gender].[All Gender].Children", "COL");
$query->addElement("[Promotion Media].[All Media]", "ROW");
$query->addElement("[Product].[All Products].[Drink].[Alcoholic Beverages]",
    "ROW");
$query->addElement("[Promotion Media].[All Media].Children", "ROW");
$query->addElement("[Product].[All Products]", "ROW");
$query->addElement("[Time].[1997]", "FILTER");

$conection = ...

$resultSet = $conection->statement(
    $query->toMdx()
);
```

4. Формування вибірки

```
use phpOLAPI\Renderer\Table\HtmlTableRenderer;
use phpOLAPI\Renderer\Table\CsvTableRenderer;
```



```

use phpOLAP\Renderer\AssocArrayRenderer
$connection = ...
$resultSet = $connection->statement("
SELECT
    {
        [Statistics].[AvarageState],
        [Positive].[PositiveRating]
    } ON COLUMNS,
    {
        [Statistics].[AvarageState].[AvarageState].[Avarage],
        [Negative].[NegativeRating].[AvarageState].[Avarage]
    } ON ROWS
FROM
    [KKkof]
WHERE
    ([Customer].[Customer Geography].[Country].[Ukraine])
");
$array = (new AssocArrayRenderer($resultSet))->generate();
var_dump($array);
// HTML table
$tableRenderer = new HtmlTableRenderer($resultSet);
echo $tableRenderer->generate();

// CSV
header("Content-type: application/vnd.ms-excel");
header("Content-disposition: attachment; filename='export.csv'");
$csv = new CsvTableLayout($resultSet);
print($csv->generate());

```

За допомогою цього скрипта ми можемо на пряму використовувати всі переваги OLAP технології, не вдаючись до прямого підключення Visual Studio.

## 4 ВПРОВАДЖЕННЯ ТА ЕКСПЛУАТАЦІЯ СИСТЕМИ

### 4.1. Загальні поняття технології Data Mining

Інтелектуальний аналіз даних (Data Mining), або розвідка даних - термін, що застосовується для опису здобуття знань у базах даних, дослідження даних, обробки зразків даних, очищення і збору даних. Це процес виявлення кореляції, тенденцій, шаблонів, зв'язків і категорій.

Data Mining - предметна область, що виникла і розвивається на базі таких наук, як прикладна статистика, розпізнавання образів, штучний інтелект, теорія баз даних тощо.

Виникнення і розвиток Data Mining зумовлені різними факторами, серед яких вирізняємо основні: вдосконалення програмно-апаратного забезпечення, вдосконалення технологій зберігання і запису даних; накопичення великої кількості ретроспективних даних; вдосконалення алгоритмів обробки інформації.

Сутність і мету технології Data Mining можна описати так: це технологія, призначена для пошуку у великих інформаційних масивах даних неочевидних, об'єктивних, корисних на практиці закономірностей. ІАД здійснюється за допомогою використання технологій розпізнавання шаблонів, а також статистичних і математичних методів.

При розвідці даних багаторазово виконуються операції і перетворення над "сирими" даними (відбір ознак, стратифікація, кластеризація, візуалізація і регресія), що призначені для знаходження:

- структур, які інтуїтивно зрозумілі для людей і краще розкривають суть бізнес-процесів, що лежать в основі їх протікання;
- моделей, які можуть передбачити результат або значення певних ситуацій, використовуючи історичні або суб'єктивні дані.

Інтелектуальний аналіз даних - процес автоматичного пошуку прихованих закономірностей або взаємозв'язків між змінними у великих масивах

необроблених даних, що поділяється на задачі класифікації, моделювання і прогнозування. Класичне визначення цього терміна дав у 1996 р. один із засновників цього напрямку Г. П'ятецький-Шапіро.

Data Mining - це процес виявлення у необроблених даних раніше невідомих нетривіальних, практично корисних і доступних інтерпретацій знань, необхідних для прийняття рішень у різних сферах діяльності.

За визначенням SAS Institute, Data Mining - це процес виділення, дослідження і моделювання великих обсягів даних для виявлення невідомих до цього структур (patterns) з метою досягнення переваг у бізнесі.

За визначенням Gartner Group, Data Mining - це процес, мета якого - виявляти нові кореляції, зразки і тенденції у результаті просіювання великого обсягу даних з використанням методик розпізнавання зразків і статистичних та математичних методів.

В основу технології Data Mining покладено концепцію шаблонів (patterns), що є закономірностями, які властиві вибіркам даних і можуть бути подані у формі, зрозумілій людині.

Задачі Data Mining:

1. Класифікація (Classification) - виявляються ознаки, які характеризують групи об'єктів досліджуваного набору даних - класи; за цими ознаками новий об'єкт можна віднести до того або іншого класу. Для вирішення задач класифікації можуть використовуватися методи: найближчий сусід (Nearest Neighbor); k-найближчий сусід (k-Nearest Neighbor); байє-совські мережі (Bayesian Networks); індукція дерев рішень; нейронні мережі (neural networks).

2. Кластеризація (Clustering) - результатом її є поділ об'єктів на групи.

3. Асоціація (Associations) - знаходять закономірності між пов'язаними подіями у наборі даних. Найбільш відомий алгоритм рішення задачі пошуку асоціативних правил - алгоритм Аргіогі.

4. Послідовність (Sequence), або послідовна асоціація (sequential association), - дає можливість знайти часові закономірності між транзакціями. Завдання послідовності подібне до асоціації, але її метою є встановлення



закономірностей між подіями, пов'язаними за часом, тобто послідовність визначається високою ймовірністю ланцюжка пов'язаних за часом подій.

5. Прогнозування (Forecasting) - на основі особливостей історичних даних оцінюються майбутні значення показників. Застосовуються методи математичної статистики, нейронні мережі тощо.

6. Визначення відхилень (Deviation Detection), аналіз відхилень або викидів - виявлення й аналіз даних, що найбільше відрізняються від загальної чисельності даних, виявлення нехарактерних шаблонів.

7. Оцінювання (Estimation) - зводиться до прогнозу безперервних значень ознак.

8. Аналіз зв'язків (Link Analysis) - задача знаходження залежностей у наборі даних.

9. Візуалізація (Visualization, Graph Mining) - створюється графічний образ аналізованих даних. Для вирішення задач візуалізації використовуються графічні методи, що показують наявність закономірностей в даних.

10. Підбивання підсумків (Summarization) - опис конкретних груп об'єктів за допомогою аналізованого набору даних.

Зазначені вище задачі поділяються за призначенням на описові і предиктивні.

Описові, або дескриптивні (descriptive), задачі пов'язані з поліпшенням розуміння аналізованих даних. Ключовий момент у таких моделях - простота і прозорість результатів для сприйняття людиною. До такого типу задач належать кластеризація і пошук асоціативних правил.

Рішення предиктивних (predictive), або прогнозуючих, задач поділяється на два етапи. На першому етапі на підставі набору даних з відомими результатами будується модель. На другому етапі вона використовується для прогнозу результатів на підставі нових наборів даних. Вимагається, щоб побудовані моделі працювали максимально точно. До цього типу задач відносять задачі класифікації і регресії. Сюди можна віднести і задачу пошуку

асоціативних правил, якщо результати її рішення можуть бути використані для прогнозу появи деяких подій.

За способами рішення задачі поділяють на такі, що вирішують за допомогою вчителя і без його допомоги. Категорія навчання з учителем представлена такими задачами Data Mining: класифікація, оцінка, прогнозування, категорія навчання, без учителя - задачею кластеризації.

У випадку рішення з допомогою вчителя задача аналізу даних розв'язується у кілька етапів. Спочатку за допомогою конкретного алгоритму

Data Mining будується модель аналізованих даних - класифікатор. Потім

класифікатор піддається навчанню. Іншими словами, перевіряється якість його роботи і, якщо вона незадовільна, відбувається додаткове навчання класифікатора. Так продовжується доти, доки не буде досягнуто необхідного

рівня якості або не стане зрозуміло, що обраний алгоритм не працює коректно з даними, або дані не мають структури, здатної проявитися. До цього типу задач відносять задачі класифікації і регресії.

Рішення без допомоги вчителя об'єднує задачі, що виявляють описові моделі, наприклад, закономірності в часових рядах макропоказників. Очевидно, якщо ці закономірності існують, то модель має їх проявити. Перевагою цих задач

є можливість їх рішення без будь-яких попередніх знань про дані аналізу. До них належать кластеризація і пошук асоціативних правил [5].

#### 4.2. Decision Trees Algorithm

Дерева рішень — це непараметричний метод навчання з керівництвом, який використовується для класифікації та регресії. Мета полягає в тому, щоб створити модель, яка прогнозує значення цільової змінної, вивчаючи прості правила прийняття рішень, виведені з характеристик даних. Дерево можна розглядати як кусково-постійне наближення.

Деякі переваги дерев рішень:

- Простий для розуміння та інтерпретації. Дерева можна візуалізувати.
- Вимагає невеликої підготовки даних. Інші методи часто вимагають нормалізації даних, створення фіктивних змінних і видалення пустих значень.

Однак зауважте, що цей модуль не підтримує пропущені значення.

- Вартість використання дерева (тобто прогнозування даних) є логарифмічною кількістю точок даних, які використовуються для навчання дерева.

- Здатний обробляти як числові, так і категоріальні дані. Однак реалізація наукового навчання наразі не підтримує категоріальні змінні. Інші методи зазвичай спеціалізуються на аналізі наборів даних, які мають лише один тип змінної. Для отримання додаткової інформації див. алгоритми.

- Здатний вирішувати проблеми з кількома виходами.

- Використовується модель білої коробки. Якщо дана ситуація спостерігається в моделі, пояснення умови легко пояснюється булевою логікою. Навпаки, у моделі чорного ящика (наприклад, у штучній нейронній мережі) результати можуть бути складнішими для інтерпретації.

- Можливість перевірити модель за допомогою статистичних тестів. Це дозволяє пояснити надійність моделі.

- Працює добре, навіть якщо його припущення дещо порушуються справжньою моделлю, на основі якої були згенеровані дані.

До недоліків дерев рішень можна віднести:

- Учні з дерева рішень можуть створювати надскладні дерева, які погано узагальнюють дані. Це називається переобладнанням. Щоб уникнути цієї проблеми, необхідні такі механізми, як обрізка, встановлення мінімальної кількості зразків, необхідних у вузлі листя, або встановлення максимальної глибини дерева.

- Дерева рішень можуть бути нестабільними, оскільки невеликі варіації в даних можуть призвести до створення зовсім іншого дерева. Ця проблема пом'якшується використанням дерев рішень в ансамблі.

- Прогнози дерев рішень не є ні гладкими, ні безперервними, а кусково-постійними наближеннями, як показано на малюнку вище. Тому вони погано вміють екстраполювати.

• Відомо, що проблема вивчення оптимального дерева рішень є NP-повною за кількох аспектів оптимальності і навіть для простих концепцій. Отже, практичні алгоритми навчання дерева рішень засновані на евристичних алгоритмах, таких як жадібний алгоритм, де в кожному вузлі приймаються локально оптимальні рішення. Такі алгоритми не можуть гарантувати повернення глобально оптимального дерева рішень. Це можна пом'якшити шляхом навчання кількох дерев у ученні ансамблю, де функції та зразки випадково відбираються із заміною.

• Існують поняття, які важко вивчити, оскільки дерева рішень не легко виражають їх, наприклад, XOR, проблеми парності чи мультиплектора.

• Учні дерева рішень створюють упереджені дерева, якщо деякі класи домінують. Тому рекомендується збалансувати набір даних перед поєднанням з деревом рішень [6].

На основі створеного раніше куба Study\_DW створюємо структуру інтелектуального аналізу на основі алгоритму дерева прийняття рішення, що представлено на рисунку 37.

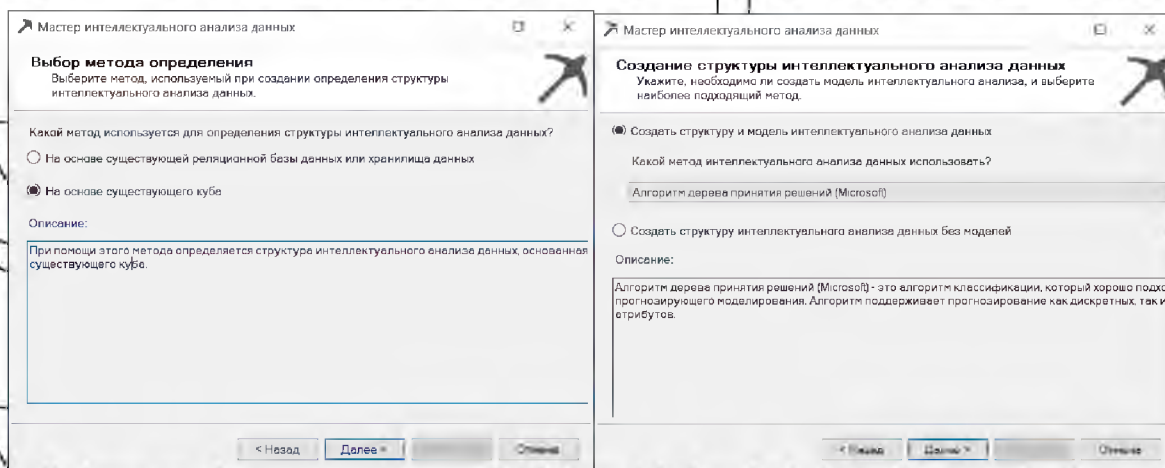


Рис. 37 Створення нової структури на основі алгоритму дерева рішень

Далі визначаємо часовий вимір кубу – Poll\_Dim та визначаємо ключове поле часу – Id\_Poll, що представлено на рисунку 38.

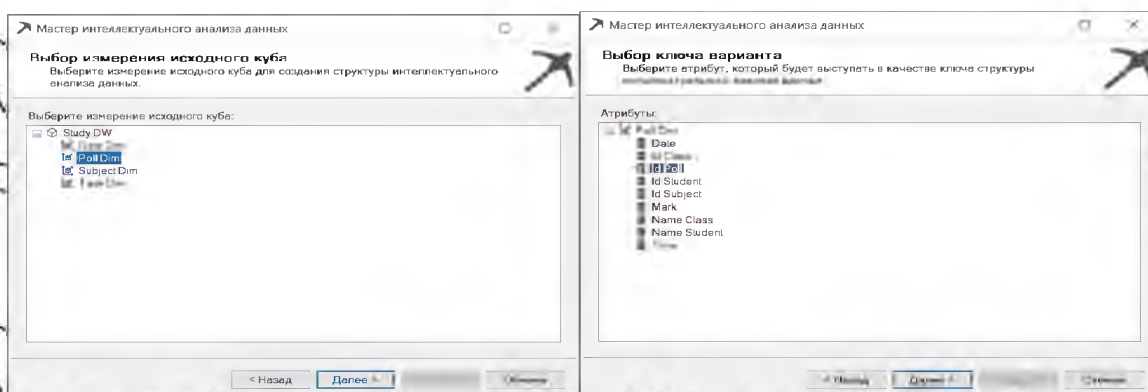


Рис. 38 Вибір ключового виміру куба та ключа структури

Наступним кроком є вибір досліджуваних даних, а саме Date, Id\_Subject, Mark, Name Class, Name Student, Time, та прогнозування значення Name Class, як представлено на рисунку 39.

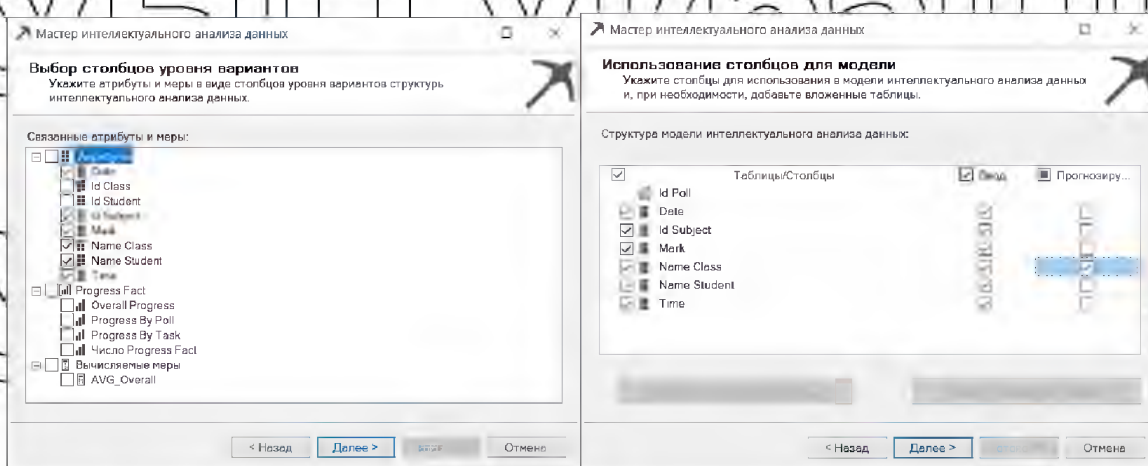


Рис. 39 Вибір полів, що аналізуємо та прогнозуємо для створеної структури  
Визначення типу полів структури та її збереження, що представлено на  
рисунку 40.

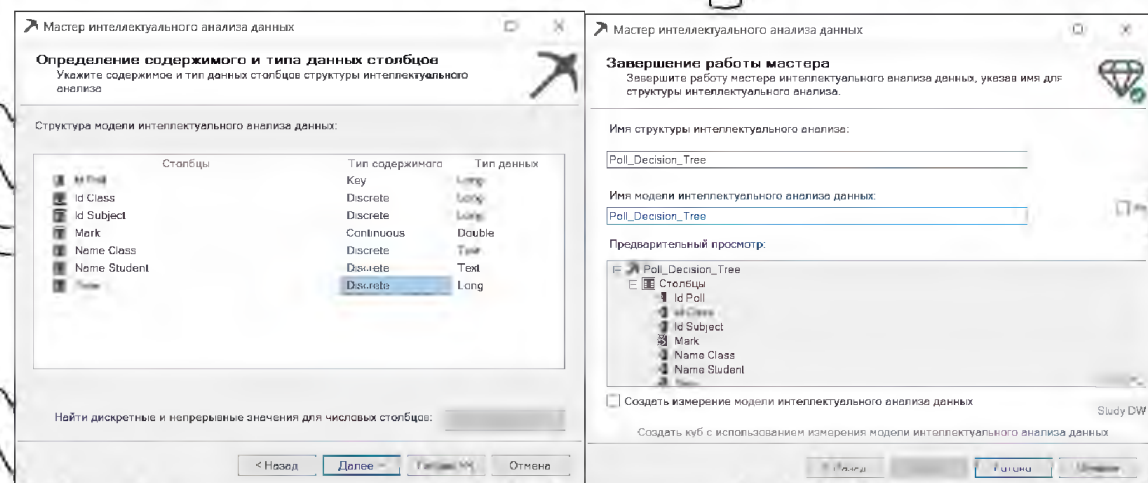


Рис. 40 Визначення типів полів та збереження структури



Результат розгортання структури інтелектуального аналізу показано на  
 рисунку 41.

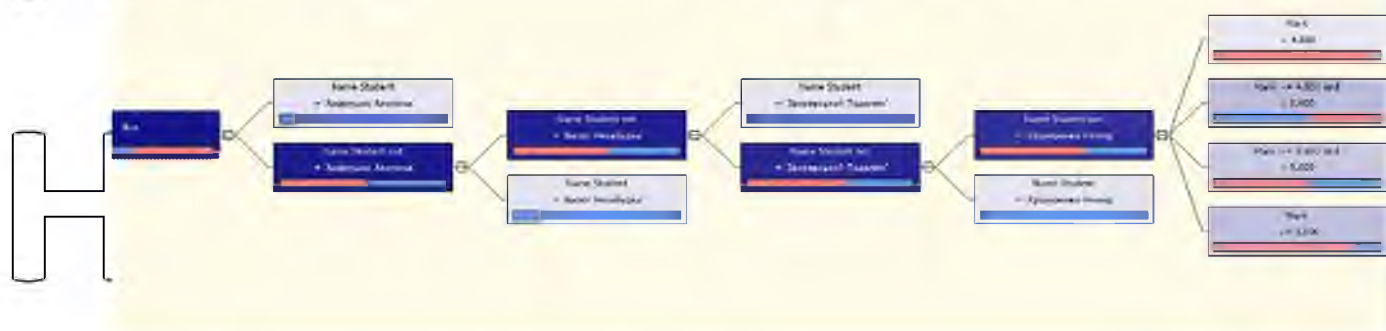


Рис. 41 Дерево рішень

Представлене на рисунку 41, дерево рішень досить велике, і основна маса вузлів зосереджена саме на менах учнів, адже кожен учень належить тільки до одного класу, але досить цікавим виступає 6 рівень – прогнозування класу в залежності від оцінок. Розглянемо цей рівень більш детально:

- Оцінка менша за 4,8, то 94,29% - ймовірність, що це клас 207-а.
- Оцінка знаходиться в проміжку 4,8-8,4, то 55,69% з клас 206-а, 44,22% - клас 207-а.
- Оцінка має значення в проміжку 8,4-9,6 то 43,81% - 206-а клас, 56,15% - клас 207-а.
- Оцінка вище за 9,6, то 15,08% - клас 206-а, та 84,86% - 207-а.

#### 4.3. Naive Bayes Algorithm

Наївні методи Байеса — це набір алгоритмів навчання з наглядом, заснованих на застосуванні теореми Байеса з «наївним» припущенням про умовну незалежність між кожною парою ознак, заданою значенням змінної класу. Теорема Байеса стверджує наступний зв'язок із заданою змінною класу  $y$ , залежний вектор ознак  $x_1$  через  $x_n$ : [7]

$$P(y | x_1, \dots, x_n) = \frac{P(y)P(x_1, \dots, x_n | y)}{P(x_1, \dots, x_n)}$$

Використовуючи наївне припущення умовної незалежності, що [7]

$$P(x_i | y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) = P(x_i | y),$$

для усіх  $i$ , це співвідношення спрощується до [7]

$$P(y | x_1, \dots, x_n) = \frac{P(y) \prod_{i=1}^n P(x_i | y)}{P(x_1, \dots, x_n)}$$

Оскільки  $P(x_1, \dots, x_n)$  є константою з урахуванням вхідних даних, ми можемо використовувати таке правило класифікації [7]

$$P(y | x_1, \dots, x_n) \propto P(y) \prod_{i=1}^n P(x_i | y)$$

$$\hat{y} = \arg \max_y P(y) \prod_{i=1}^n P(x_i | y),$$

Його можна використовувати для оцінки максимальну апостеріорну (MAP) оцінку  $P(y) \prod_{i=1}^n P(x_i | y)$ , перша - це відносна частота класу  $y$  в тренувальний комплект. [7]

Різні наївні байєсівські класифікатори відрізняються головним чином за припущеннями, які вони роблять щодо розподілу  $P(x_i | y)$ .

Незважаючи на їх, на перший погляд, занадто спрощені припущення, наївні байєсівські класифікатори досить добре працювали в багатьох реальних ситуаціях, як відомо, класифікація документів і фільтрація спаму. Вони вимагають невеликої кількості навчальних даних для оцінки необхідних параметрів. (Щодо теоретичних причин, чому наївний Байєс добре працює, і на які типи даних він працює, див. посилання нижче).

Наївні байєсівські учні та класифікатори можуть бути надзвичайно швидкими порівняно з більш складними методами. Відокремлення розподілів



умовних ознак класу означає, що кожен розподіл може бути незалежно оцінений як одновимірний розподіл. Це, у свою чергу, допомагає полегшити проблеми, пов'язані з прокляттям розмірності [7].

На основі створеного раніше куба Study\_DW створюємо структуру інтелектуального аналізу на основі спрощеного алгоритму Байеса, що представлено на рисунку 41.

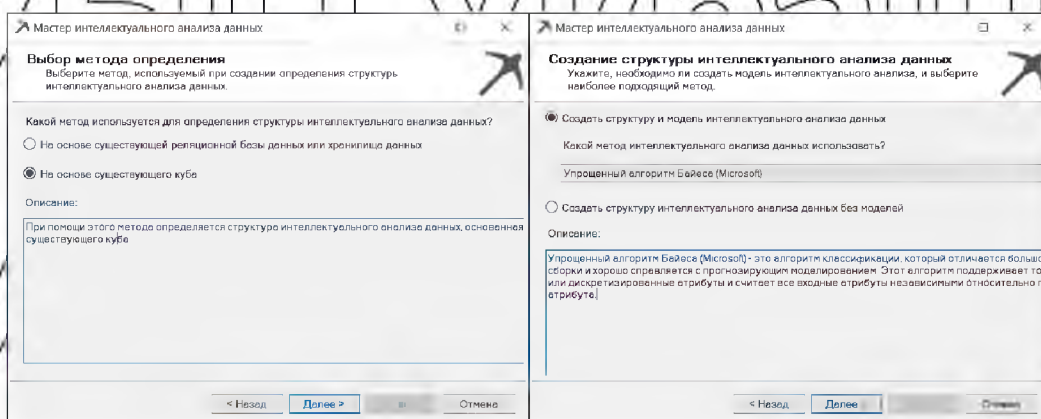


Рис. 41 Створення нової структури на основі спрощеного алгоритму Байеса.

Далі визначаємо часовий вимір/кубу – Task\_Dim та визначаємо ключове поле часу – Id\_Task, що представлено на рисунку 42.

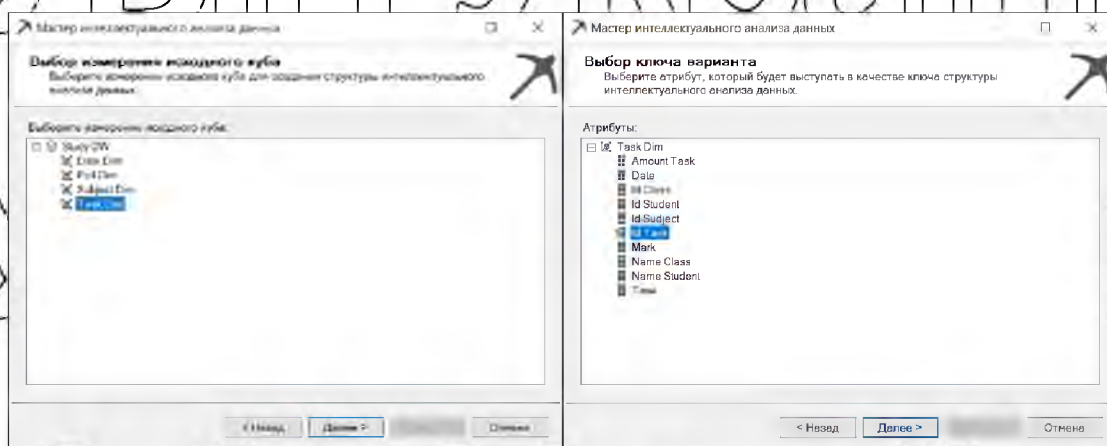


Рис. 42 Вибір ключового виміру куба та ключа структури

Наступним кроком є вибір досліджуваних даних, а саме Date, Id\_Subject, Mark, Name Class, Name Student, Time, та прогнозування значення Name Class, як представлено на рисунку 43.



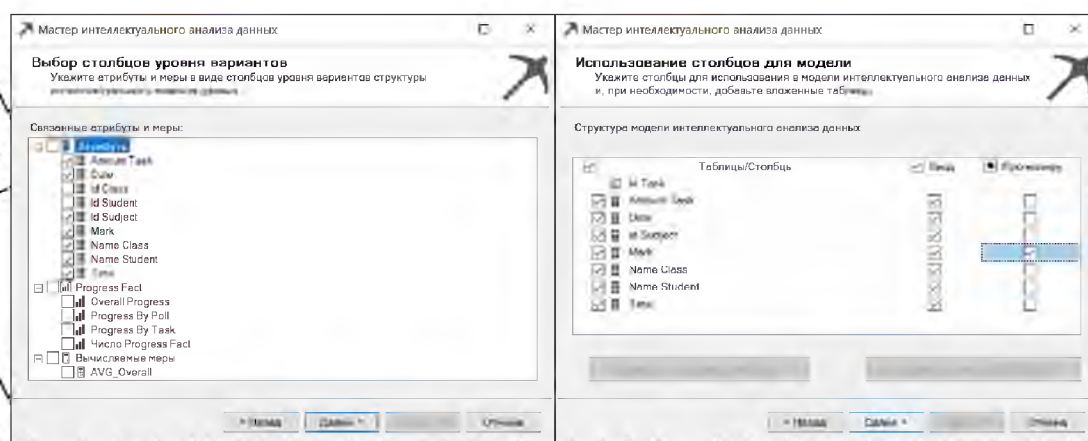


Рис. 43. Вибір полів, що аналізуємо та прогнозуємо для створюваної структури

Визначення типу полів структури та її збереження, що представлено на

рисунок 44.

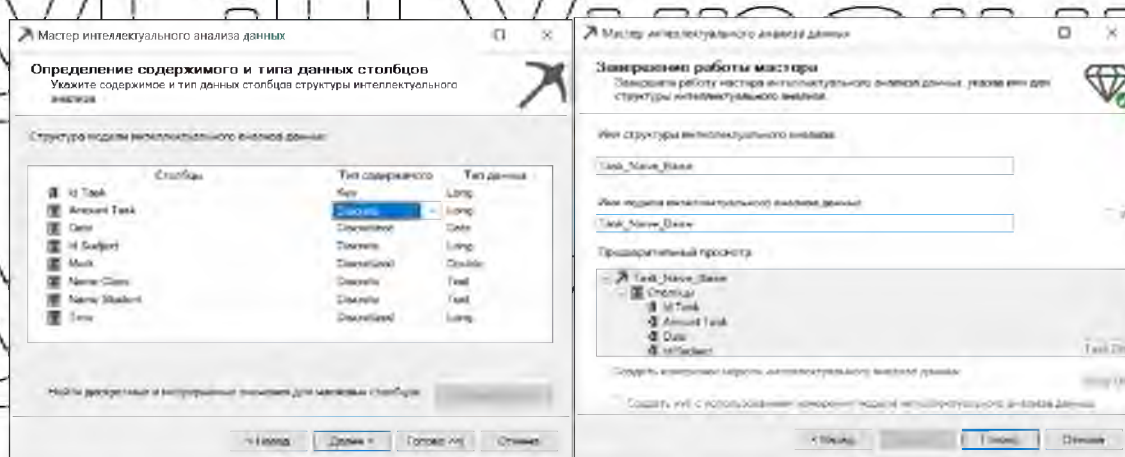


Рис. 44. Визначення типів полів та збереження структури.

Результат розгорнення структури інтелектуального аналізу показано на  
рисунок 45.

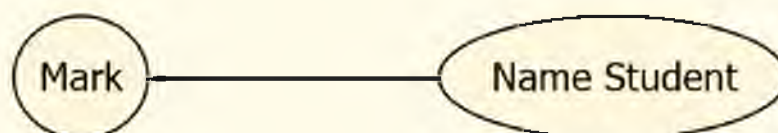


Рис. 45. Мережа зв'язків

Представлена на рисунку 45, мережа дозволяє зрозуміти, що значення оцінки напряму залежить саме від учня. Проаналізуємо атрибути:

- Оцінку, менше за 2,7 найімовірніше отримає Ганкевич Магара з вірогідністю 90% або Сенькович Боян – 10%.

- Оцінку в проміжку 2,7-5,5 найімовірніше отримає Ганкевич Магара з вірогідністю 68% чи Ситенко Марко – 27%.

- В проміжку від 5,5 до 7,5, оцінки отримують Качура Бажана (18,34%), Орлик Вишата (16,72%), Ситенко Марко (16,34%), Ганкевич Магара (15,63%).

- Оцінку в межах 7,5-10,2 отримують близько 13 чоловік з імовірністю 6-5% кожен.

- Оцінка більше за 10,2 отримає Снігур Анастасія (25%), Кушніренко Немир (20%), Закревський Подольян (13%), Андрушко Акилина (12%).

#### 4.4. Дослідження використання методу асоціативних правил

Правила асоціації — це заснований на правилах метод машинного навчання для виявлення цікавих зв'язків між змінними у великих базах даних. Він призначений для визначення сильних правил, виявлених у базах даних, використовуючи певні показники цікавості. У будь-якій транзакції з різноманітними елементами правила асоціації призначені для виявлення правил, які визначають, які і чому певні елементи пов'язані.

Правила асоціації використовуються сьогодні в багатьох областях застосування, включаючи аналіз використання Інтернету, виявлення вторгнень, безперервне виробництво та біоінформатику. На відміну від аналізу послідовності, навчання правил асоціації зазвичай не враховує порядок елементів у транзакції або між транзакціями [8].

Алгоритм правила асоціації сам по собі складається з різних параметрів, які можуть ускладнити виконання тим, хто не має певного досвіду в аналізі даних, з багатьма правилами, які важко зрозуміти.

Незважаючи на це, навчання правил асоціації є чудовою системою для прогнозування поведінки у взаємозв'язках даних. Це робить його гідною уваги

технікою класифікації або виявлення закономірностей у даних під час реалізації методів машинного навчання.

Правила асоціації створюються шляхом пошуку даних для поширених моделей «якщо-то», а також за допомогою певного критерію в розділі «Підтримка та впевненість» для визначення найважливіших зв'язків. Підтримка

це доказ того, як часто той чи інший елемент з'являється у наведених даних, оскільки впевненість визначається тим, скільки разів твердження якщо-то визнаються істинними. Однак є третій критерій, який можна використовувати,

він називається Lift, і його можна використовувати для порівняння очікуваної впевненості та фактичної впевненості. Підвищення покаже, скільки разів очікується, що твердження якщо-то буде істинним.

Правила асоціації створені для обчислення з наборів елементів, які створені двома або більше елементами. Якби правила були побудовані на основі аналізу всіх можливих наборів елементів із даних, то правил було б стільки, що вони не мали б жодного значення. Ось чому правила асоціації зазвичай складаються з правил, які добре представлені даними. [9]

Існує багато різних методів аналізу даних, які можна використовувати для пошуку певної аналітики та результатів, наприклад, є класифікаційний аналіз,

кластеризаційний аналіз та регресійний аналіз. Залежно від того, що ви шукаєте зі своїми даними, залежить від того, яку техніку ви повинні використовувати.

Правила асоціації в основному використовуються для пошуку аналітики та прогнозування поведінки клієнтів. Для класифікаційного аналізу, швидше за все,

він буде використаний для запитань, прийняття рішень і прогнозування поведінки. Кластеризаційний аналіз в основному використовується, коли немає

припущень щодо ймовірних зв'язків у даних. Регресійний аналіз використовується, коли потрібно передбачити значення безперервної залежності від ряду незалежних змінних [10].

*Переваги.* Використання правил асоціації має багато переваг, як-от пошук шаблону, який допомагає зрозуміти кореляції та спільні випадки між наборами даних. Дуже хорошим прикладом із реального світу, який використовує правила

асоціації, може бути медицина. Медицина використовує правила асоціації, щоб допомогти діагностувати пацієнтів. Під час діагностики пацієнтів необхідно враховувати багато змінних, оскільки багато захворювань мають схожі симптоми. Використовуючи правила Асоціації, лікарі можуть визначити умовну ймовірність захворювання, порівнюючи зв'язки симптомів з минулих випадків

[11].  
Недоліки. Однак правила асоціації також призводять до багатьох різних проблем, таких як пошук відповідних параметрів і порогових налаштувань для алгоритму майнінгу. Але є також недолік великої кількості виявлених правил.

Причина в тому, що це не гарантує, що правила будуть визнані відповідними, але це також може призвести до низької продуктивності алгоритму. Іноді реалізовані алгоритми будуть містити занадто багато змінних і параметрів. Для тих, хто не має гарної концепції інтелекту даних, це може спричинити проблеми з його розумінням [12].

На основі створеного раніше куба Study DW створюємо структуру інтелектуального аналізу на основі алгоритму правил взаємозв'язків, що представлено на рисунку 46.

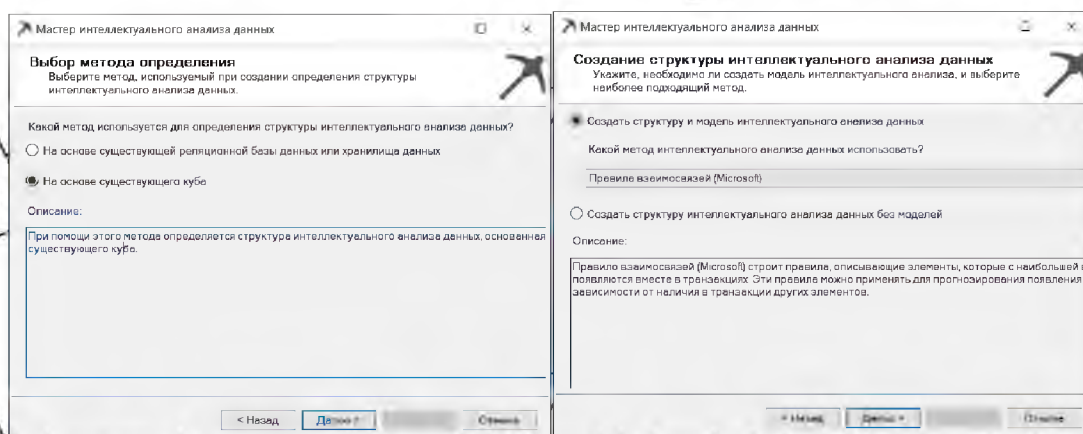


Рис. 46 Створення нової структури на основі алгоритму правил взаємозв'язків

Далі визначаємо часовий вимір кубу – Poll\_Dim та визначаємо ключове поле часу – Id\_Poll, що представлено на рисунку 47.



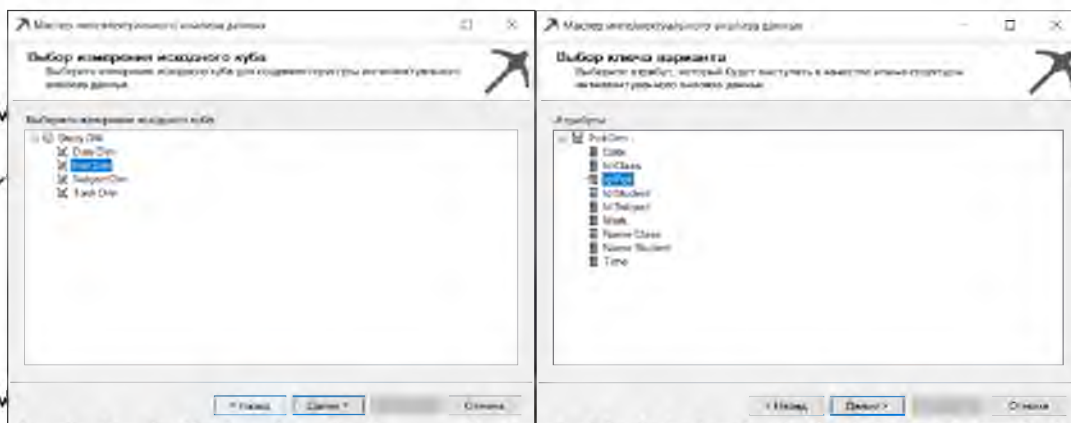


Рис. 47 Вибір ключового виміру куба та ключа структури

Наступним кроком є вибір досліджуваних даних, а саме Id\_Subject, Mark, Name Class, Name Student, Time та прогнозування значення Name Class, як представлено на рисунку 48.

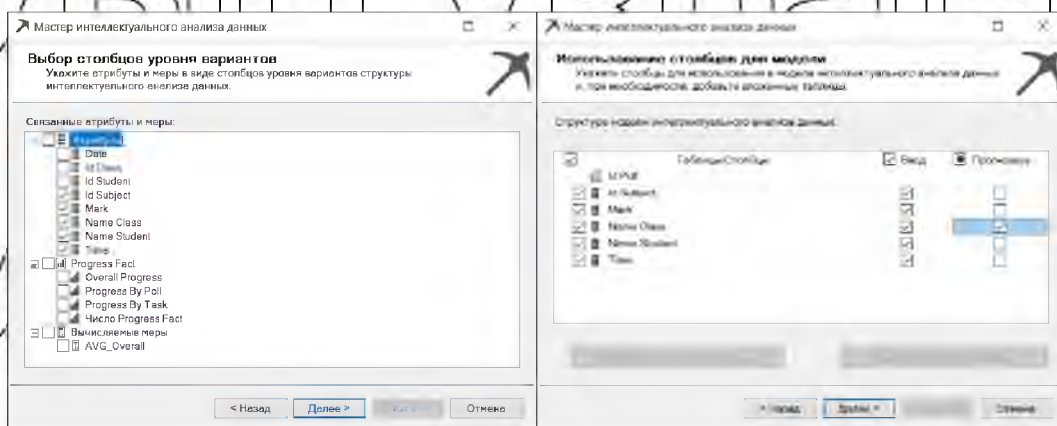


Рис. 48 Вибір полів, що аналізуємо та прогнозуємо для структури

Визначення типу полів структури та її збереження, що представлено на рисунку 49.

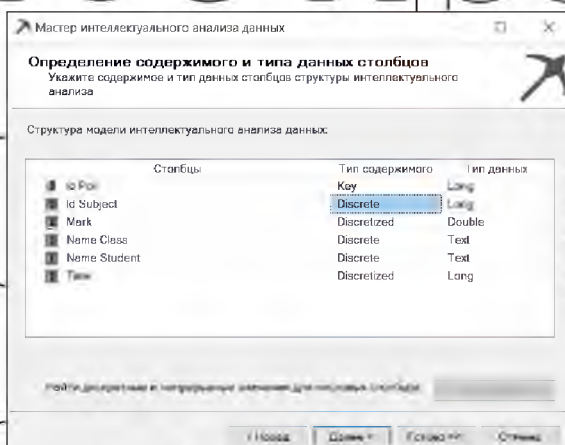


Рис. 49 Визначення типів полів та збереження структури.

Результат розгортання структури інтелектуального аналізу представлено в таблиці 1.

# НУБІП УКРАЇНИ

Таблиця 1

Ймовірність	Важливість	Правило	Оцінка
0,9	0,84	Name Student = Снігур Анастасія, Id Subject = 2	>= 10,27
0,767	0,77	Name Student = Снігур Анастасія, Id Subject = 1	>= 10,27
0,733	0,75	Name Student = Кушніренко Немир, Id Subject = 5	>= 10,27
0,700	0,73	Name Student = Снігур Анастасія, Id Subject = 5	>= 10,27
0,667	0,79	Name Student = Снігур Анастасія	>= 10,27
0,667	0,79	Name Student = Снігур Анастасія, Name Class = 207-а	>= 10,27
0,667	0,79	Name Student = Снігур Анастасія, Time = 45	>= 10,27
0,633	0,68	Name Student = Закревський Подолян, Id Subject = 5	>= 10,27
0,633	0,68	Name Student = Кушніренко Немир, Id Subject = 2	>= 10,27
0,567	0,64	Name Student = Кушніренко Немир, Id Subject = 3	>= 10,27
0,567	0,64	Name Student = Снігур Анастасія, Id Subject = 3	>= 10,27
0,567	1,23	Name Student = Ганкевич Магара, Id Subject = 1	2,72-5,5
0,553	0,69	Name Student = Кушніренко Немир, Time = 45	>= 10,27

Загалом можна підсумувати, що більшість правила стосуються конкретних учнів, що тримаються свого діапазону успішності.

#### 4.5. Дослідження використання алгоритмів кластеризації

Кластеризація — це завдання поділу сукупності або точок даних на кілька груп таким чином, щоб точки даних в тих самих групах були більш схожими на інші точки даних у тій же групі, ніж в інших групах. Простіше кажучи, мета полягає в тому, щоб відокремити групи зі схожими ознаками та об'єднати їх у кластери.

Загалом, кластеризацію можна розділити на дві підгрупи:

- Жорстка кластеризація: у жорсткій кластеризації кожна точка даних або повністю належить до кластера, або ні.
- М'яка кластеризація: у м'якій кластеризації замість того, щоб помістити кожену точку даних в окремий кластер, призначається ймовірність або ймовірність того, що ця точка даних буде в цих кластерах.

Оскільки завдання кластеризації є суб'єктивним, то засобів, які можна використати для досягнення цієї мети, достатньо. Кожна методологія дотримується різного набору правил для визначення « подібності » між точками даних. Насправді відомо більше 100 алгоритмів кластеризації. Але деякі алгоритми широко використовуються, давайте розглянемо їх докладніше:

- Моделі підключення: як випливає з назви, ці моделі засновані на уявленні про те, що точки даних, розташовані ближче в просторі даних, демонструють більше схожості одна з одною, ніж точки даних, розташовані далі.

Ці моделі можуть мати два підходи. У першому підході вони починають з класифікації всіх точок даних в окремі кластери, а потім їх агрегування в міру зменшення відстані. У другому підході всі точки даних класифікуються як один кластер, а потім розділяються у міру збільшення відстані. Крім того, вибір функції відстані є суб'єктивним. Ці моделі дуже легко інтерпретувати, але їм бракує масштабованості для обробки великих наборів даних. Прикладами цих моделей є алгоритм ієрархічної кластеризації та його варіанти.

• **Центроїдні моделі:** це ітеративні алгоритми кластеризації, в яких поняття подібності впливає з близькості точки даних до центрoїда кластерів. Алгоритм кластеризації K-Means є популярним алгоритмом, який відноситься до цієї категорії. У цих моделях немає кластерів, необхідних в кінці, необхідно вказати заздалегідь, тому важливо мати попередні знання про набір даних. Ці моделі виконуються ітеративно, щоб знайти локальний оптимум.

• **Моделі розподілу:** ці моделі кластеризації засновані на уявленні про те, наскільки ймовірно, що всі точки даних в кластері належать до одного розподілу (наприклад, нормальний, гаусів). Ці моделі часто страждають від переобтяжки.

Популярним прикладом цих моделей є алгоритм максимізації очікування, який використовує багатствимірні нормальні розподіли.

• **Моделі щільності:** ці моделі здійснюють пошук у просторі даних областей із різною щільністю точок даних у просторі даних. Він ізолює різні регіони з різною щільністю та призначає точки даних у межах цих регіонів в одному кластері. Популярними прикладами моделей щільності є DBSCAN і OPTICS [13].

На основі створеного раніше куба Study\_DW створюємо структуру інтелектуального аналізу на основі алгоритму кластеризації, що представлено на **рисунку 50**.

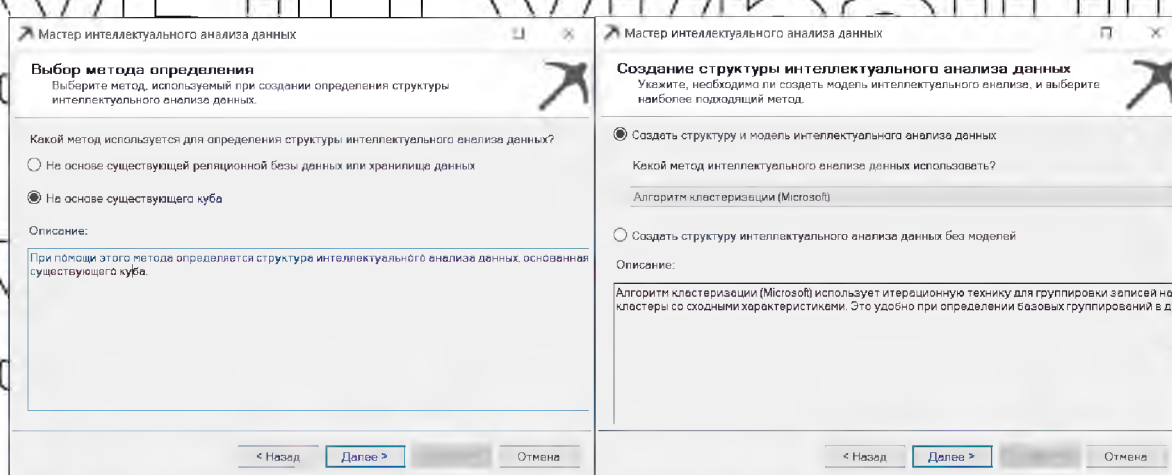


Рис. 50 Створення нової структури на основі алгоритму кластеризації

Далі визначаємо часовий вимір кубу – Task\_Dim та визначаємо ключове поле часу – Id\_Task, що представлено на **рисунку 51**.



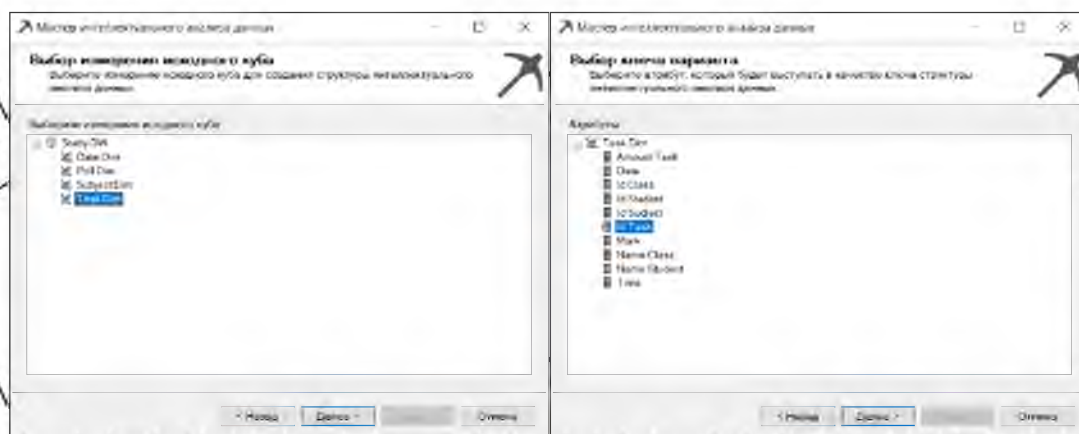


Рис. 51 Выбор ключевого измерения куба та ключа структуры

Наступним кроком є вибір досліджуваних даних, а саме Date, Id\_Subject, Mark, Name\_Class, Name\_Student, Time, та прогнозування значення Name Class, як представлено на рисунку 52.

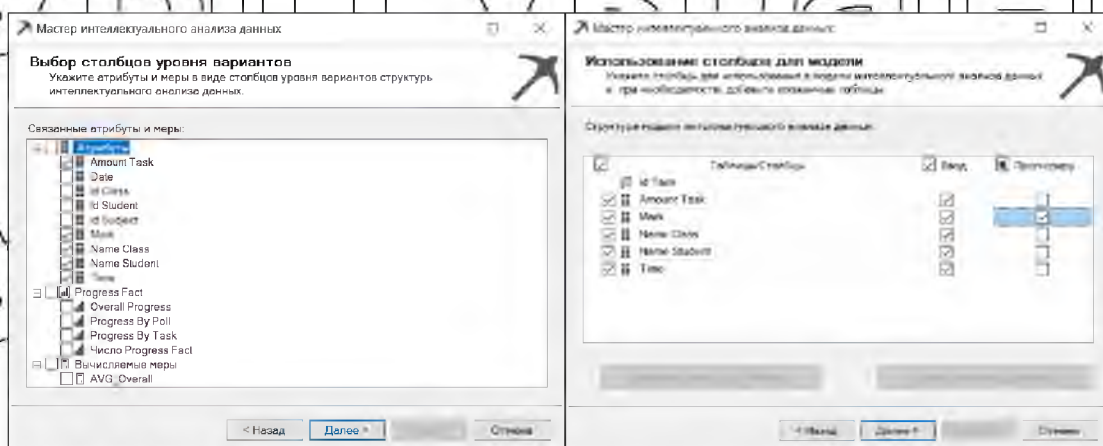


Рис. 52 Выбор полів, що аналізуємо та прогнозуємо для структури  
Визначення типу полів структури та її збереження, що представлено на рисунку 53.

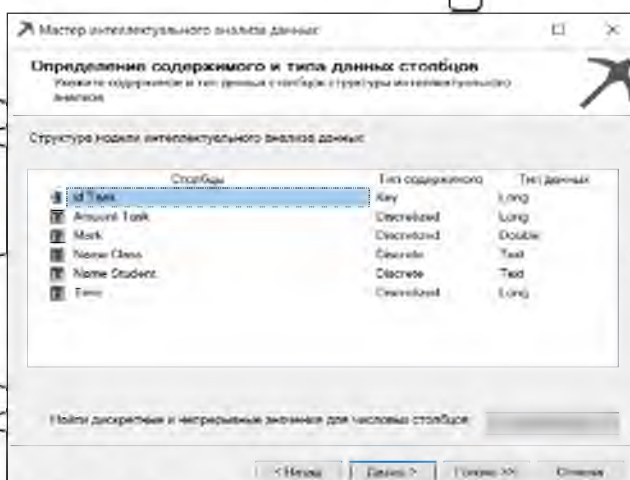


Рис. 53 Визначення типів полів та збереження структури

Результат розгортання структури інтелектуального аналізу представлено на рисунку 54.

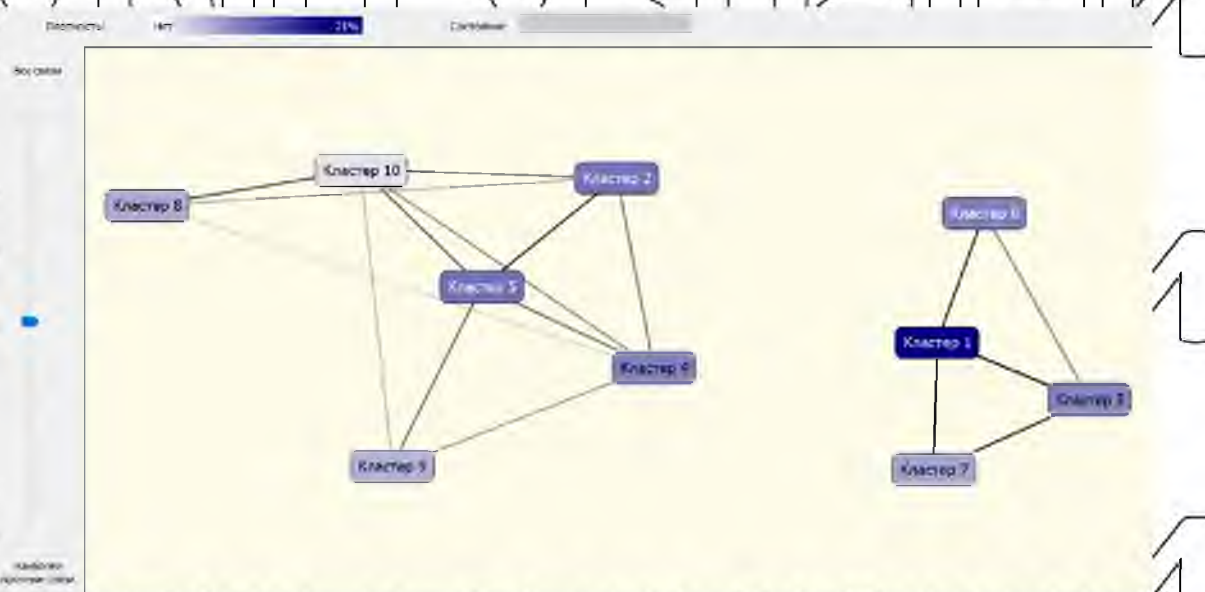


Рис. 54 Кластерна діаграма.

#### Аналіз кластерів:

- Кластер 1: учні 206-а класу, що витратили на виконання завдання 41-48 хвилин на 6 (52,97%) чи 10 (47,02%) завдань отримали оцінки в діапазоні: 9-10,1 (27,38%); 7,9-9 (27,22%); 4-7,9(22,41%); 10,1-12(20,74%).

- Кластер 2: учні 207-а класу, що витратили на виконання завдання 41-48 (74,7%) хвилин на 6 (37,8%) чи 10 (62,16%) завдань отримали оцінки в діапазоні: 9-10,1(69,8%); 7,9-9 (6,5%); 10,1-12(23,63%).

- Кластер 3: учні 206-а класу, що витратили на виконання завдання 32-40 (68,87%) хвилин на 6 (52,56%) чи 10 (47,43%) завдань отримали оцінки в діапазоні: 9-10,1 (36%); 7,9-9 (53,41%); 4-7,9(8,2%).

- Кластер 4: учні 207-а класу, що витратили на виконання завдання 32-40 (72,98%) хвилин на 6 (69,14%) чи 10 (30,86%) завдань отримали оцінки в діапазоні: 9-10,1 (40,74%); 7,9-9 (34,07%); 4-7,9(9,11%); 10,1-12(15,93%).

- Кластер 5: учні 207-а класу, що витратили на виконання завдання 41-48 (78,45%) хвилин на 6 (54,76%) чи 10 (45,23%) завдань отримали оцінки в діапазоні: 9-10,1 (21,48%); 7,9-9 (55,92%); 4-7,9(21,50%).

- Кластер 6: учні 206-а класу, що витратили на виконання завдання 32-40



(65,58%) хвилини на 6 (50,55%) чи 10 (49,44%) завдань отримали оцінки в діапазоні: 9-10,1 (31,59%); 10,1-12 (61,89%).

- Кластери 7-10: містять дані що мають певні відхилення.

#### 4.6. Розрахунок КРІ

КРІ – показник ефективності підрозділу (підприємства), який допомагає організації в досягненні стратегічних і тактичних (оперативних) цілей. Використання ключових показників ефективності дає компанії можливість оцінити її поточний стан і допомогти в оцінці успішності реалізації обраної стратегії. [17]

Для розрахунку КРІ скористаємось проектом раніше створеним за допомогою служби SSAS. На рисунку 45 подано формули, за допомогою яких розраховується КРІ успішності учнів за опитуванням та оцінюванням завданнями.



Рис. 55. КК успішності учнів за опитуванням та оцінюванням завданнями

На рисунку 56. показано КК досягнення цілей успішності учнів.

Отобразить структуру	Значение	Цель	Состояние
KPI_Progress	0,22	0,25	
KPI_Progress_Poll	0,2	0,25	
KPI_Progress_Task	0,23	0,25	

Рис. 56. КК досягнення цілі.

На рисунку 55 зображена вибірка з якої береться необхідна інформація для розрахунку коефіцієнту освітньої компетентності учня та результат обчислення і формування КК.

Використовуючи отримані дані система може сформувати висновки базуючись на обчисленнях які базуються на алгоритмах кластеризації, в результаті чого ми маємо КК досягнення цілі, який використаємо в подальшому для формування рекомендацій щодо покращення обчислювальних показників. А

також використовуючи вищезгадані дані можна зробити висновок, як сильно та якими показниками відрізняються результати навчання по обраній вибірці часу.

Подібна процедура застосовується для визначення як коефіцієнта компетентності по конкретному предмету в конкретного учня, так і для

визначення загального КК учня. Оперуючи цими даними можна сформувати

рекомендаційні листи, де будуть сформовані компетентними особами рекомендації щодо покращення рівня освіти як в конкретного учня, так і в класі

в цілому, в подальшому коли система буде зіштовхуватись з подібними показниками вона буде знати які саме рекомендаційні листи потрібно

використати, так як при певній кількості повторень(залежить від налаштувань) проблема стає типовою, і до неї застосовуються типові рішення.

Трішки видозмінені обчислення використовуються для обчислення загального коефіцієнта класу, потім класів чи навіть школи.

#### 4.7. Експлуатація

Систему підрахунку досить легко ввести в експлуатацію, та налаштувати під потреби. Система має продуманий функціонал, за допомогою якого є змога

масштабувати систему під абсолютно любі потреби. Наприклад систему можна

розвернути до конкретного класу, школи. Також система може об'єднати всі школи в районі, області та навіть в країні, обмежується вона лише технічними

характеристиками апаратного забезпечення, і потребує невеликого пере налаштування системи в різних масштабах.

Реалізована система надає змогу наглядно демонструвати прогрес конкретної вибірки учнів, конкретного учня чи класу. За допомогою системи є

можливість зрівняти показники по учням, по класам в різних школах, Чи зрівняти між собою 2 семестри з різним форматом навчання.

Отримана інформація дозволяє зробити висновки про ефективність навчального процесу і за допомогою цієї інформації можна зрозуміти в яких місцях навчальний процес не розкривається на повну. І виправивши відкриті системою недоліки можна покращити загальний рівень якості навчання, та відповідно, рівень освітньої компетентності учня.

Для отримання результату, в першу чергу відповідні компетентні спеціалісти повинні розробити певні стандарти, відповідно до яких рівень освітньої компетентності учнів буде найбільш високим. Далі потрібно застосувати ці стандарти в еталонному семестрі та ввести ці данні в систему.

Після цих дій в системі буде достатньо інформації, і вона почне свою роботу.

Система оперує таким поняттям як Навчальний блок. Це умовний відрізок часу, за який система вбираю ту кількість інформації, як необхідна їй для створення перших коефіцієнтів, відповідно до яких і буду створювати КК – коефіцієнт компетентності.

Здебільшого навчальний блок являє собою одну з тем навчальної програми, і, в залежності від складності тем, проте мінімально в нього входить:

- Класна робота;
- Домашня робота;
- Самостійна/Контрольна робота;
- Спеціалізований(наданий системою) тест;
- Результати опитувань.

Як виглядає журнал оцінок з сторони учня, який базується на подібних блоках, можна спостерігати на рисунку 57



Тема	Коефіцієнт компетентності	Оцінювання	Оцінка
Числові нерівності	100%	Класна робота	10
		Домашня робота	10
		Самостійна робота	10
		Тест	10
Квадратична функція	100%	Класна робота	10
		Домашня робота	11
		Самостійна робота	9
		Тест	9
Числові послідовності	100%	Класна робота	10
		Домашня робота	11
		Самостійна робота	9
		Тест	9

Рис. 57 Журнал оцінок з сторони учня

В даному журналі ми можемо спостерігати формальне оцінювання (Класна, домашня, самостійна робота та тест). Для формування повноцінного КК учню з своєї, а вчителю з своєї, сторони потрібно пройти опитування і якщо система отримає всю необхідну інформацію то уже по пройденому навчальному блоку в учня буде можливість спостерігати його успішність, яка буде характеризуватись коефіцієнтом компетентності.

Ролі «Вчитель» система надає більше прав, так вчитель може оцінювати учня, ставити та видаляти оцінки. Також для визначення КК вчитель повинен пройти опитування.

НУБІП України

НУБІП України

Рис. 58 Опитування учня

Рис. 59. Опитування вчителя

На рисунку 58 та 59 є можливість спостерігати опитування учня та вчителя відповідно, ціль якого зібрати всі атрибути неформального оцінювання. Подібні опитування видозмінюються відповідно до предмету та глибини розрахунку КК. Також, для більш точніших та суб'єктивних оцінок є можливість додавання в



систему опитування з сторони батьків, для підтвердження чи спростування тих чи інших показників учня.

Після того, як система отримає необхідну кількість інформації, вона формує звіт, де учень може побачити:

- Загальний коефіцієнт освітньої компетентності;
- Конкретний, по предмету, коефіцієнт освітньої компетентності;
- Рекомендації щодо покращення загального КК
- Рекомендації щодо покращення КК по конкретному предметі;
- Результат формального оцінювання;
- Результат неформального оцінювання.

На рисунку 60 можна спостерігати результат роботи алгоритму, який порахував КК по предмету «Алгебра», створив рекомендації щодо покращення та передав їх в систему, яка в свою чергу створила позитивний для сприйняття, осучаснений стиль візуалізації інформації.

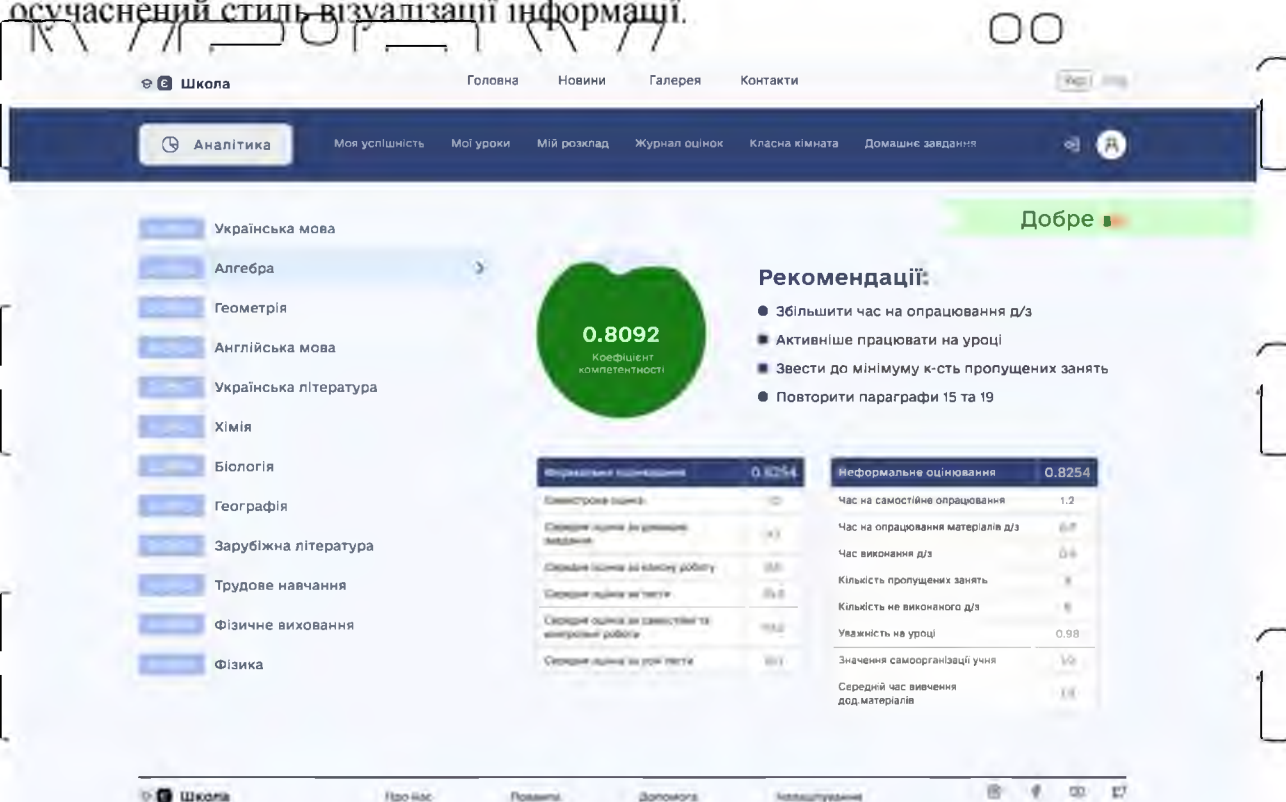


Рис. 60 Звіт по предмету «Алгебра»

На рисунку 61 можна спостерігати звіт з загальною успішністю учня по всім предметам програми. В звіті вказується:



Рис 61 Загальна успішність конкретного учня.

Як додатковий функціонал, який можна реалізувати за потреби, учню буде рекомендовано, базуючись на отриманих даних, ряд спеціальностей, в який відповідно до знань, навичок та позитивних сторін особистості, найбільше підійшли б учню, також було б запропоновано ряд вузів, в яких найкраще навчають по вищезгаданих спеціальностях.

Для актуальності результатів підрахунку КК у аналітику та відповідного компетентного спеціаліста є доступ до вибірки інформації, за допомогою якої він може зробити висновки, базуючись на яких можна покращити освітній процес.

Втручання людини в систему необхідне лише для внесення нових стандартів, або утворенню нової адаптивної системи оцінок, після чого система уже сама зможе надавати рекомендації щодо покращення загального рівня освітньої компетентності вибірки учнів.



На рисунку 62 зображено як виглядає система, коли вибірка даних це конкретний клас. В даному звіті є можливість спостерігати:

- Загальний коефіцієнт класу;
- Рекомендації щодо його покращення;
- Спільні проблеми учнів з низьким рівнем КК;
- Список предметів з найнижчим загальним рівнем КК.

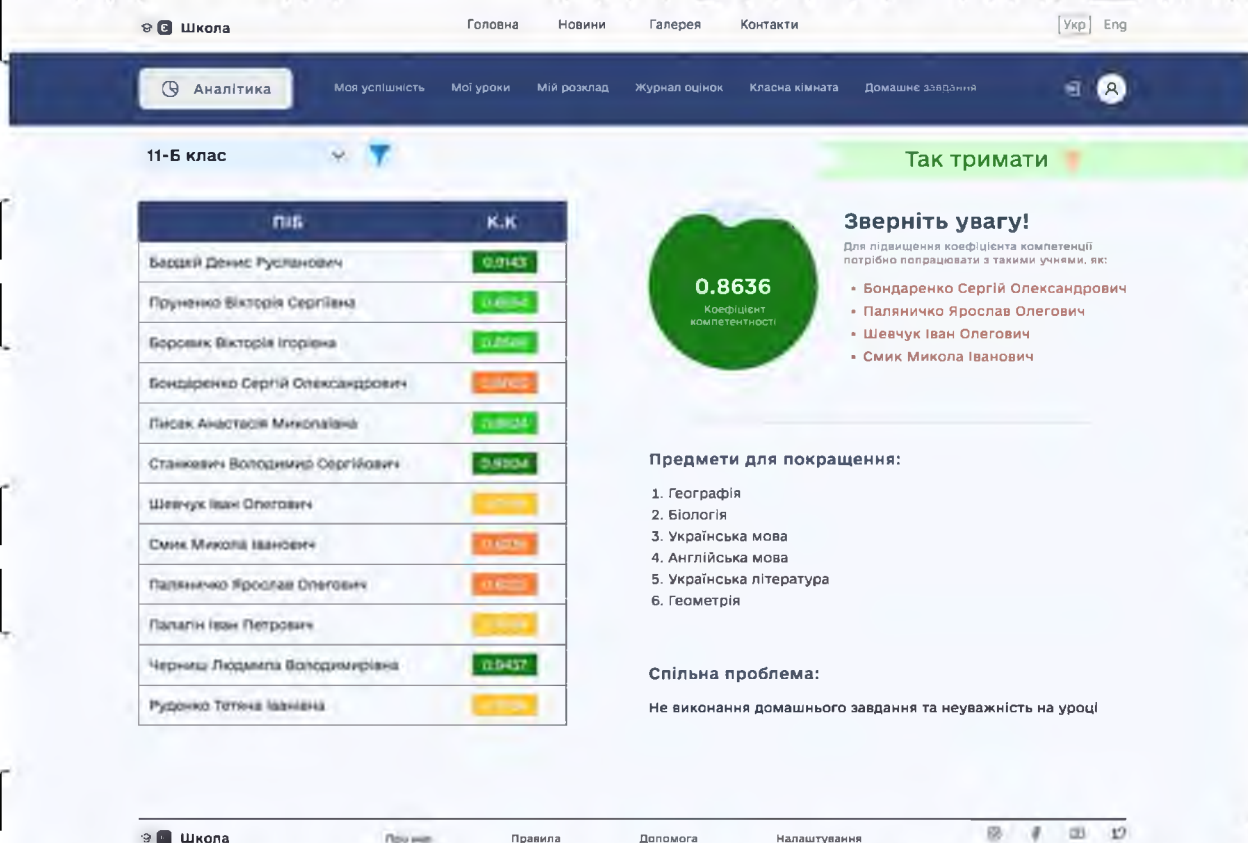


Рис. 62 Загальний КК 11 б класу, на конкретних учнів

На рисунку 63, зображено можливість вибору конкретного учня в обраному класі, для перегляду КК та формування висновків, щодо покращення КК конкретного учня

На зображенні ми бачимо:

- Рекомендації, які відносяться до обраного учня;
- Предмети, в яких найнижчий КК.





Рис. 63. З вибірки «116 Клас» обрано учня з найнижчим КК

Як основну ціль система повинна показувати різницю між 2 часовими вибірками, зрівнювати еталонне очне навчання з дистанційним. Як позитивні електронні освітні ресурси впливають на освітню компетенцію учня, та які показники погіршуються при такому форматі навчання



Рис. 64. Порівняння двох семестрів за результативністю

На рисунку 64 зображено результат порівняння двох семестрів між собою, судячи з звіту ми можемо побачити, що при дистанційному форматі навчання погіршуються показники гуманітарних предметів в цілому, та неформальні параметри решти предметів. Система аналізуючи отриманні дані вносить рекомендації щодо покращення показників.

НУБІП України

НУБІП України

НУБІП України

НУБІП України

НУБІП України

НУБІП України

## ВИСНОВКИ

НУБІП України

Під час розробки системи, яка б змогла розрахувати як впливають

електронні освітні ресурси на формування освітньої компетентності учнів, було розроблено алгоритм, який використовує для обчислення всі показники, які так чи інакше могли б вплинути на формування освітньої компетентності. Було визначено які саме показники впливають на освітню компетентність учнів.

Були визначені та використані технології обробки, зберігання, аналізу та обчислення інформації які як найкраще проявили себе в заданому середовищі, а розроблений алгоритм є універсальним методом визначення освітньої компетентності учня.

Під час реалізації системи було розроблено сайт, який може впоратись з всіма викликами сучасного світу, і допоможе забезпечити якість освітніх послуг на найвищому рівні. На сайті реалізовано весь необхідний функціонал, який допоможе зібрати всю необхідну інформацію, яка необхідна для того щоб сформувати коефіцієнт компетентності, який в свою чергу в зображує на скільки якісно надають освітні послуги в певному часовому відрізу.

Отже результатом роботи в першу чергу стала система, яка дозволить не тільки підтримувати необхідний рівень якості освітніх послуг, а й забезпечить всією необхідною інформацією, яка необхідна для розвитку і покращення сфери освіти в цілому.

НУБІП України

НУБІП України



## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Системний аналіз - [Електронний ресурс] - Режим доступу:  
<http://dspace.wunu.edu.ua/jspui/bitstream/316497/627/1/%D0%A1%D0%B8%D1%81%D1%82%D0%B5%D0%BC%D0%BD%D0%B8%D0%B9%20%D0%B0%D0%B1%D0%B0%D0%B8%D1%96%D0%B7.pdf>
2. Основи UML - [Електронний ресурс] - Режим доступу:  
<https://docs.kde.org/trunk5/uk/umbrello/umbrello/uml-basics.html>
3. Діаграм прецедентів - [Електронний ресурс] - Режим доступу:  
<https://ukrbuka.net/88640-Diagrammy-precedentov-variantov-ispolzovaniya.html>
4. Діаграма послідовності - [Електронний ресурс] - Режим доступу:  
<http://flash.retejo.info/exefpagxo/uml/diagrama-poslidovnosti>
5. Інтелектуальні технології Data Mining і Text Mining. - [Електронний ресурс]  
[https://pidru4niki.com/1623021247786/informatika/intelektualni\\_tehnologivi\\_data\\_mining\\_text\\_mining](https://pidru4niki.com/1623021247786/informatika/intelektualni_tehnologivi_data_mining_text_mining)
6. Decision Trees. URL - [Електронний ресурс] - : <https://scikit-learn.org/stable/modules/tree.html>
7. Naive Bayes. URL - [Електронний ресурс] : [https://scikit-learn.org/stable/modules/naive\\_bayes.html](https://scikit-learn.org/stable/modules/naive_bayes.html)
8. Piatetsky-Shapiro, Gregory (1991), Discovery, analysis, and presentation of strong rules, in Piatetsky-Shapiro, Gregory; and Frawley, William J.; eds., Knowledge Discovery in Databases, AAAI/MIT Press, Cambridge, MA - [Друкований ресурс]
9. Garcia, Enrique (2007). "Drawbacks and solutions of applying association rule mining in learning management systems" (PDF). - [Електронний ресурс] - : <https://sci2s.ugr.es/keel/pdf/specific/congreso/3-associationrules-Final.pdf>
10. Техніки аналізу даних: 5 найкращих для погляду. - [Електронний ресурс]



<https://www.precisely.com/blog/datagovernance/top-5-data-mining-techniques>

11. "What are Association Rules in Data Mining (Association Rule Mining)". SearchBusinessAnalytics. – [Електронний ресурс] - :

<https://searchbusinessanalytics.techtarget.com/definition/association-rules-in-data-mining>

12. "Drawbacks and solutions of applying association rule mining in learning management systems". ResearchGate. – [Електронний ресурс] - :

<https://www.researchgate.net/publication/289657906>

13. An Introduction to Clustering and different methods of clustering. – [Електронний ресурс] - : <https://www.analyticsvidhya.com/blog/2016/11/an-introduction-to-clustering-and-different-methods-of-clustering/#:~:text=Soft%20Clustering%3A%20In%20soft%20clustering,clusters%20of%20the%20retail%20store.>

14. OLAP-технології та звітність [Електронний ресурс] - [https://stud.com.ua/62441/menedzhment/olap\\_tehnologiyi\\_zvitnist](https://stud.com.ua/62441/menedzhment/olap_tehnologiyi_zvitnist)

15. What is Analysis Services? – [Електронний ресурс] - : <https://docs.microsoft.com/enus/analysis-services/analysis-services-overview?view=asallproductsallversions>

16. SQL Server Integration Service [Електронний ресурс] - <https://learn.microsoft.com/en-us/sql/integration-services/sql-server-integration-services?view=sql-server-ver15>

17. What is SSRS? – [Електронний ресурс] - : [SSRS Tutorial: What is SQL Server Reporting Services? \(guru99.com\)](https://www.guru99.com/ssrs-tutorial-what-is-sql-server-reporting-services.html)

18. Підручник HTML [Електронний ресурс] - Режим доступу: [https://css.in.ua/article/shcho-take-css\\_3](https://css.in.ua/article/shcho-take-css_3)

19. Підручник по JS - [Електронний ресурс] - Режим доступу: <http://ruszura.in.ua/uncategorized/shcho-take-javascript-vak-funktsionuyu-javascript.html>

20. SQL база даних Для чого призначена база даних? [Електронний ресурс]



<https://www.ukraine.com.ua/uk/blog/programming/sql-baza-dannih-dlya-chego-prednaznachena-baza-dannih.html>

21. Що таке PHP? - [Електронний ресурс]

<https://freehost.com.ua/ukr/faq/wiki/chto-takee->

[php/?gclid=Cj0KCQjwwfiaBhC7ARIsAGvcPe6lXoJQP3GrJTdVdb0REthfY](http://php/?gclid=Cj0KCQjwwfiaBhC7ARIsAGvcPe6lXoJQP3GrJTdVdb0REthfY)

[Uyb72UyM7iVX80Kxоnc6-88F-tiXY-YaAnleEALw\\_wcB](http://Uyb72UyM7iVX80Kxоnc6-88F-tiXY-YaAnleEALw_wcB)

22. Діаграма діяльності - [Електронний ресурс] - Режим доступу:

<https://uk.icvscience.com/activity-diagram>

23. Олександр Леоненков. Самовчитель UML, 2007. – [Електронний ресурс]

– Режим доступу: <http://kniga.scienceontheweb.net/samovchitel-uml-2.html>.

24. Формування освітньої компетентності по предмету «Я пізнаю світ» -

[Електронний ресурс]

[http://jds.multycourse.com.ua/ua/print\\_page/module/27](http://jds.multycourse.com.ua/ua/print_page/module/27)

25. Рекомендації від УЦОЯО щодо формування читацької та математичної компетенції учнів на рівні початкової освіти – [Електронний ресурс] -

<https://mon.gov.ua/ua/news/ucyoao-pidgotuvav-rekomendacivi-shodo-formuvannya-chitackoyi-ta-matematichnovi-kompetentnostej-uchniv-na-rivni-pochatkovoyi-osviti>

26. DFD діаграма потоків даних - [Електронний ресурс]

[https://studwood.net/1048448/informatika/model\\_potokiv\\_danih](https://studwood.net/1048448/informatika/model_potokiv_danih)