

**МІНІСТЕРСТВО ОСВІТИ ТА НАУКИ УКРАЇНИ
ПВНЗ
«МІЖНАРОДНИЙ НАУКОВО-ТЕХНІЧНИЙ УНІВЕРСИТЕТ
ІМЕНІ АКАДЕМІКА Ю. БУГАЯ»**

Кафедра інформаційних та комунікаційних технологій

Освітня компонента: «Інтелектуальний аналіз даних»

Дослідницька РГР – Ескізний проєкт

Виконав:
студент 4 курсу денної форми навчання
група І26
Бондаренко О.В.

Перевірила:
доц. Гриб'юк О.О.

Київ – 2025

Тема проєкту: Аналіз поведінкових даних та розробка аналітичних моделей для прийняття рішень.

Прізвище та ім'я виконавців проєкту: Бондаренко Олег.

Концепція проєкту: посилання на [github](https://github.com).

Мета проєкту: Формування системи інтелектуального аналізу поведінкових та предметно-орієнтованих даних з метою моделювання сценаріїв розвитку бізнес-рішень, виявлення трендів та оптимізації взаємодії.

Завдання проєкту:

1. Отримати вибіркові дані поведінкових метрик користувачів веб-проєкту з різних джерел (первинна предметна область) та провести їх первинну обробку.
2. Здійснити регресійний та survival-аналіз для виявлення впливу поведінкових ознак на ймовірність покупки або конверсії.
3. Провести розкладання часових рядів поведінкових метрик на складові (тренд, сезонність).
4. Провести апробацію розробленого комплексу методів на новій предметній області – ідентифікація діамантів.
5. Сформувати висновки та рекомендації щодо оптимізації бізнес-рішень на основі виявлених закономірностей, трендів та результатів моделювання.

ЕСКІЗНИЙ ПРОЄКТ

1. Архітектура системи (Flow)

Концептуальна архітектура (зображена на Рис. 1) описує логічний потік даних у системі. Цей потік реалізує повний цикл інтелектуального аналізу даних:

1. Модуль діагностики (інформаційна система) слугує для збору первинних транзакційних даних (OLTP) та їх збереження у DB1 (базі даних діамантів).
2. Модулі аналітики (ML) (кластеризація, класифікація, апроксимація) працюють «за лаштунками», використовуючи дані з DB1 (характеристики) та DB2 (ринкові показники) для навчання моделей.
3. Результати аналізу зберігаються у DB3 (базі груп/кластерів) або повертаються користувачу через Модуль виведення інформації.

Технічна архітектура реалізує цей концептуальний потік за допомогою сучасної клієнт-серверної моделі, що складається з трьох чітко розділених рівнів:

1. Рівень представлення (Client-Side). Фронтенд, реалізований як статичний веб-додаток.
2. Рівень додатку (Server-Side). Бекенд, реалізований як API-сервіс на Python/FastAPI.
3. Рівень даних (Database). Реляційна СУБД PostgreSQL.

Поеднання архітектур відбувається наступним чином:

- «Модуль діагностики» та «Модуль підготовки, введення і перетворення інформації» реалізовані як зв'язка Фронтенд (UI) + Бекенд (API). Бекенд містить експертний компонент ReportCalculator (на основі правил IDC) для виконання розрахунків Етапу 1 (OLTP-обробка).
- «Модуль кластеризації», «Модуль класифікації» та «Модуль апроксимації» реалізовані як окремі ML-сервіси (функції) всередині того ж Бекенду (Python). Вони виконують аналітику Етапу 2 та надають результати для візуалізації на Фронтенді (Етап 3).

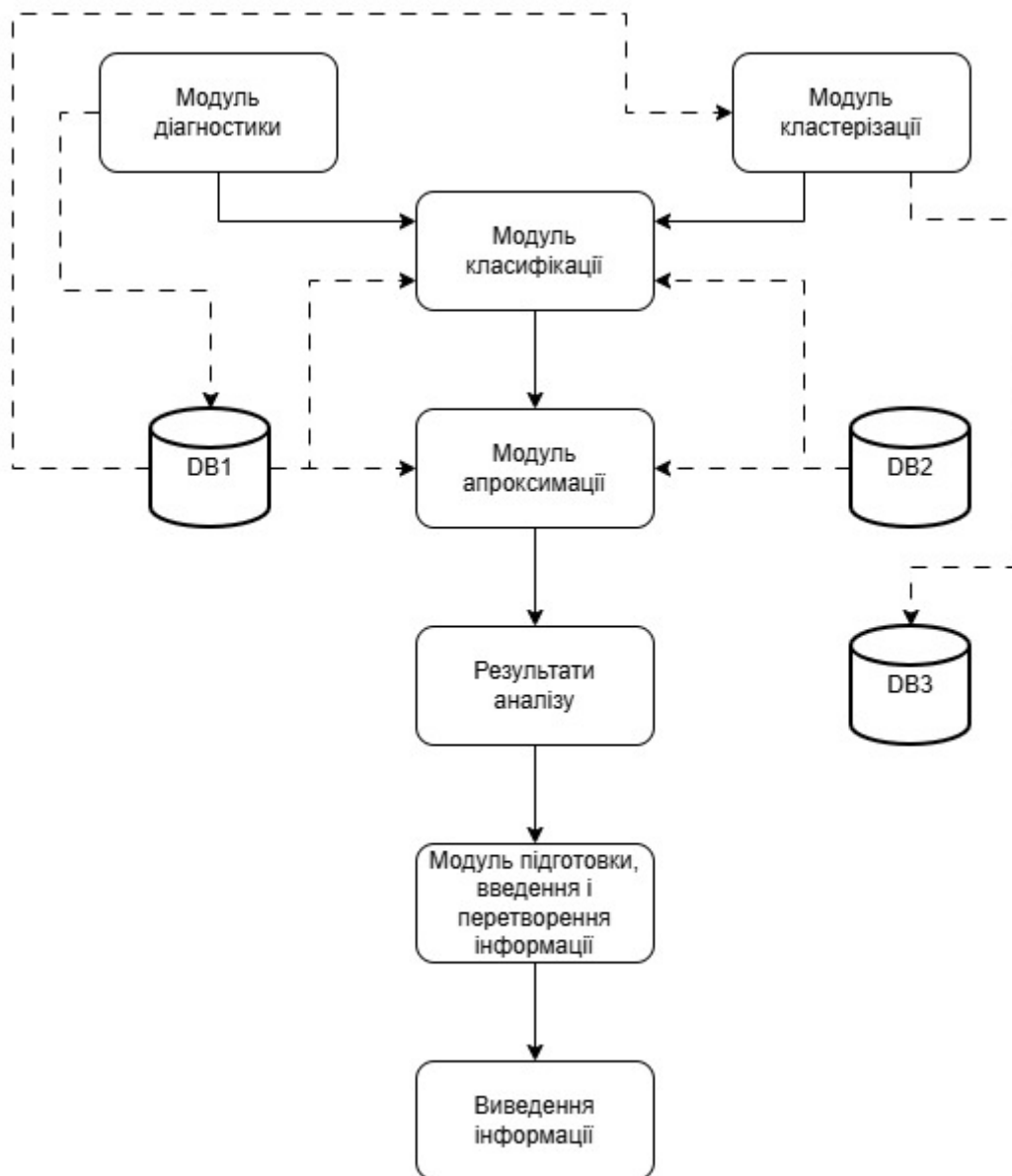


Рисунок 1 – Архітектура системи (Flow)

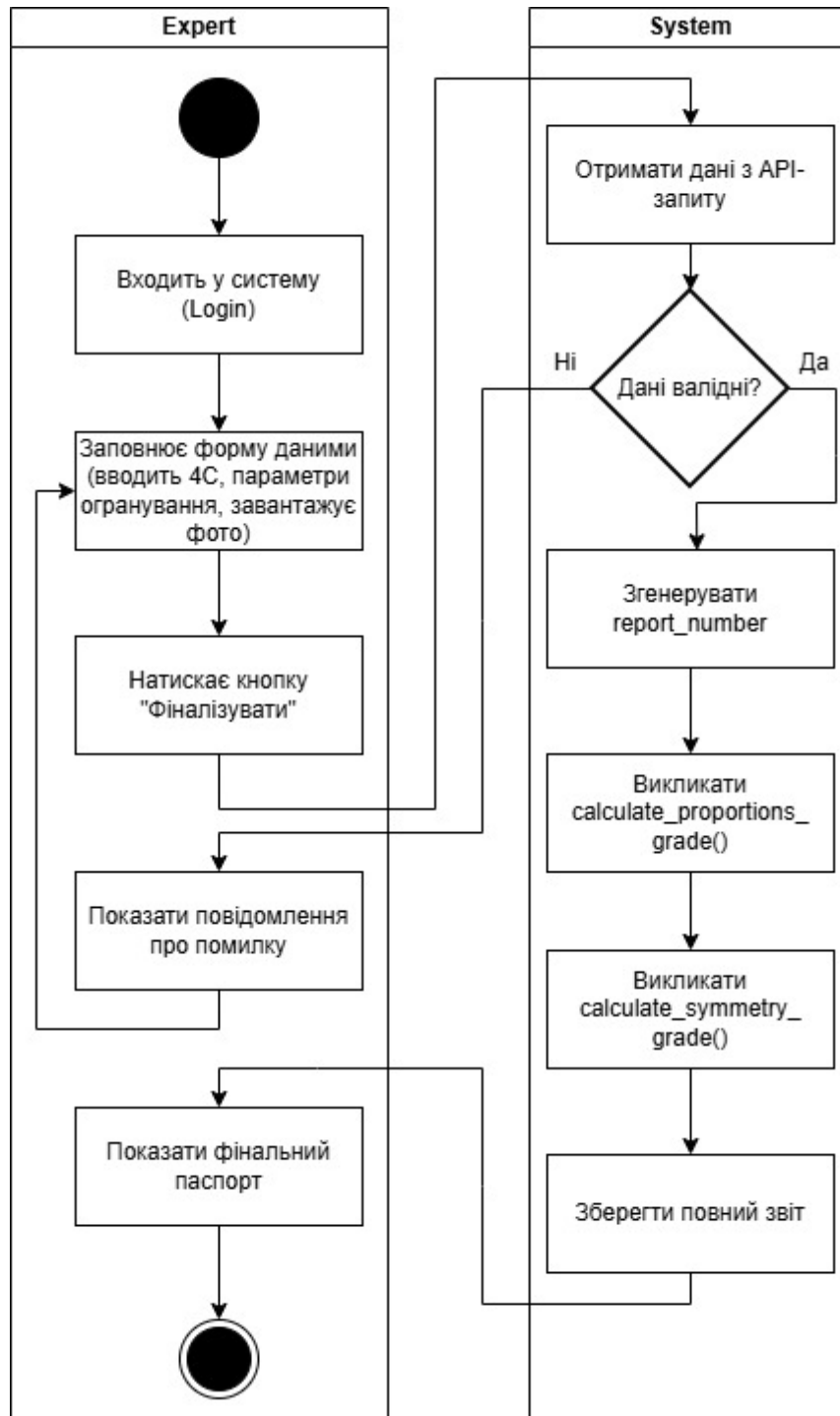
Роздивимося подрібніше Етап 1.

"Модуль діагностики" – це система на основі правил. Ми використовуємо:

- Міжнародні стандарти (IDC-Rules): Для визначення `color_grade`, `clarity_grade`.
- Розрахунковий модуль (назвемо його ReportCalculator): Логіка `if/else`, яка на основі `crown_angle`, `table_width_percent` тощо, присвоює `proportions_grade` та `symmetry_grade`.

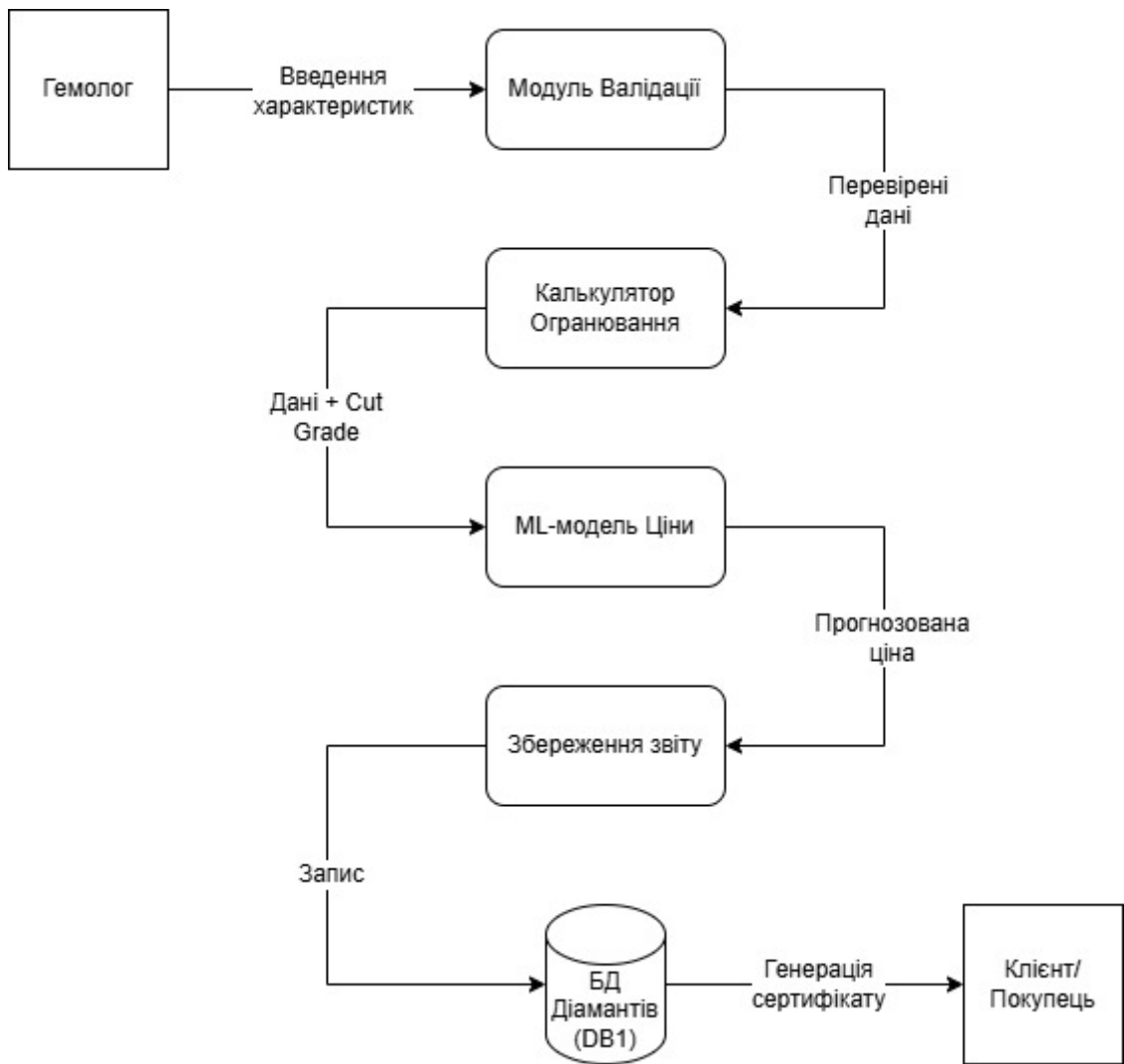
Це детерміновані алгоритми (не стохастичні), вони завжди дають однаковий результат для однакових вхідних даних.

Ось як виглядає Діаграма Активності (Activity Diagram) для Етапу 1:

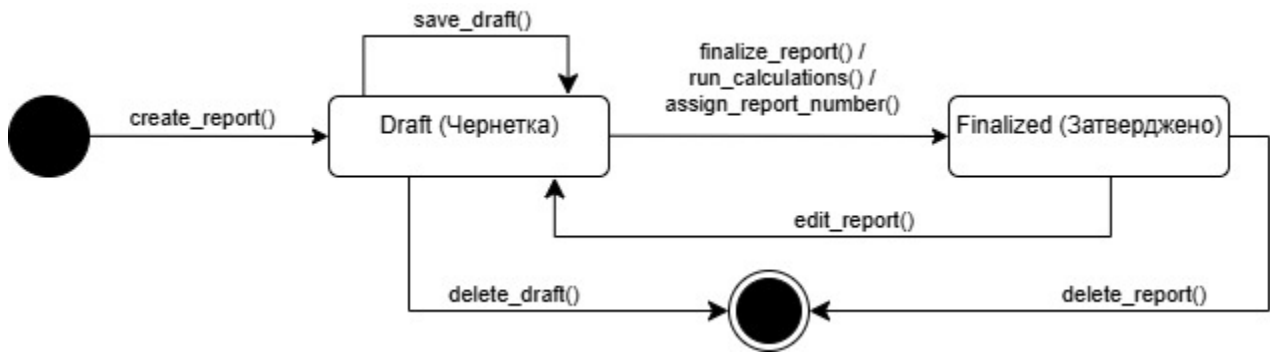


Можна чітко побачити OLTP-процес: "Заповнює форму даними ..." -> "Зберегти повний звіт".

Діаграма потоків даних (Data Flow Diagram) Етап 1:

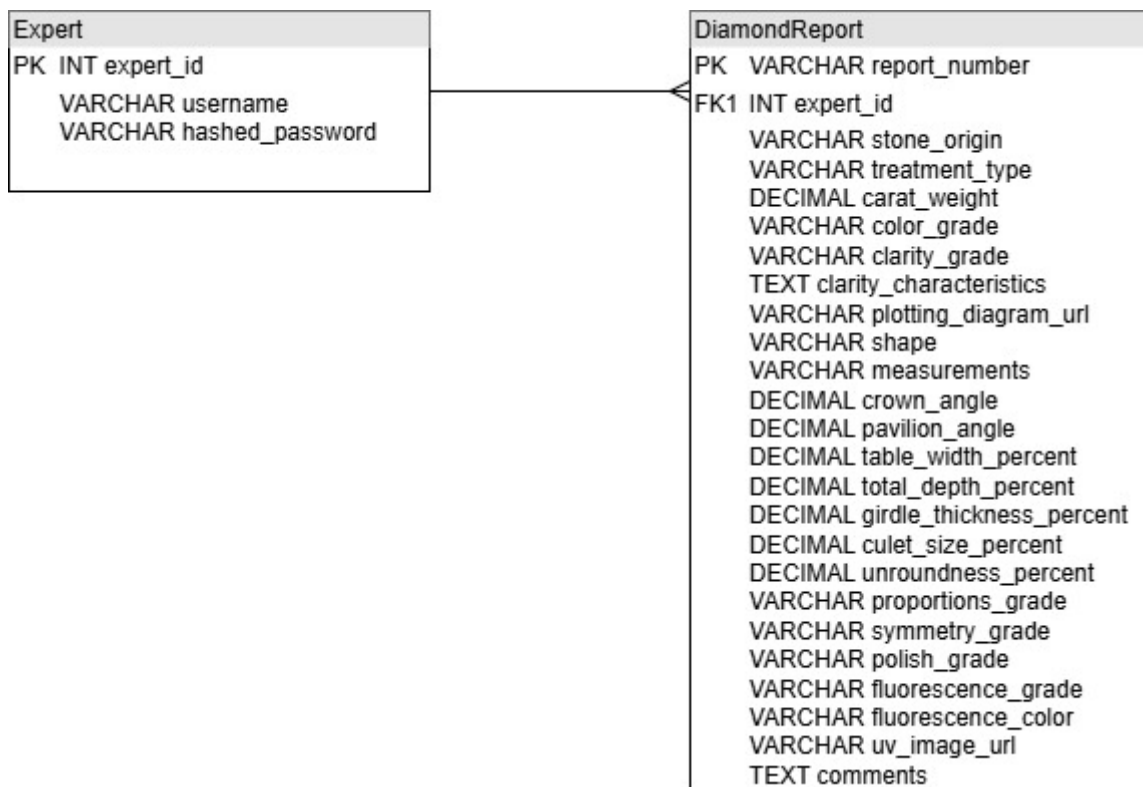


Діаграма Станів (State Machine Diagram):

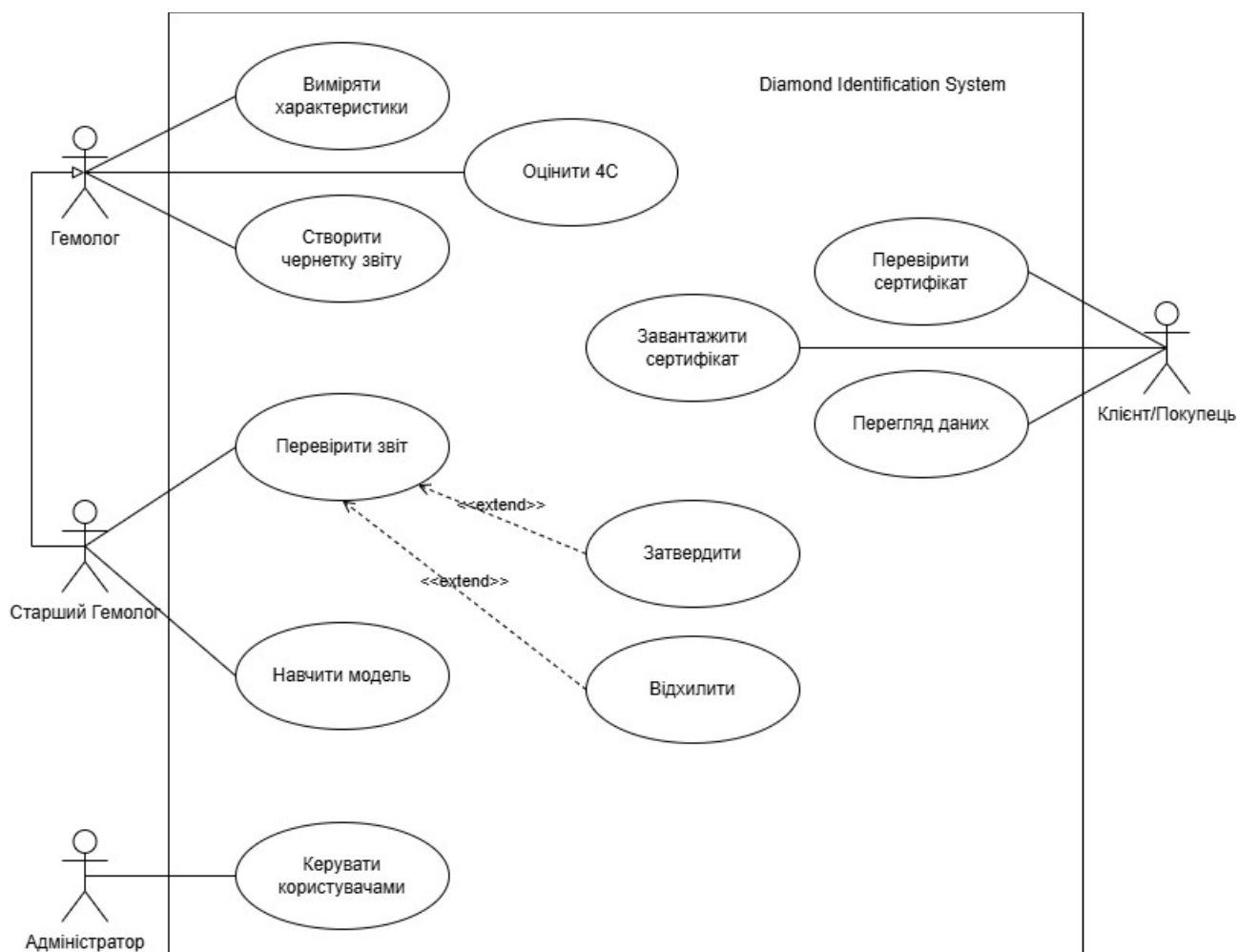
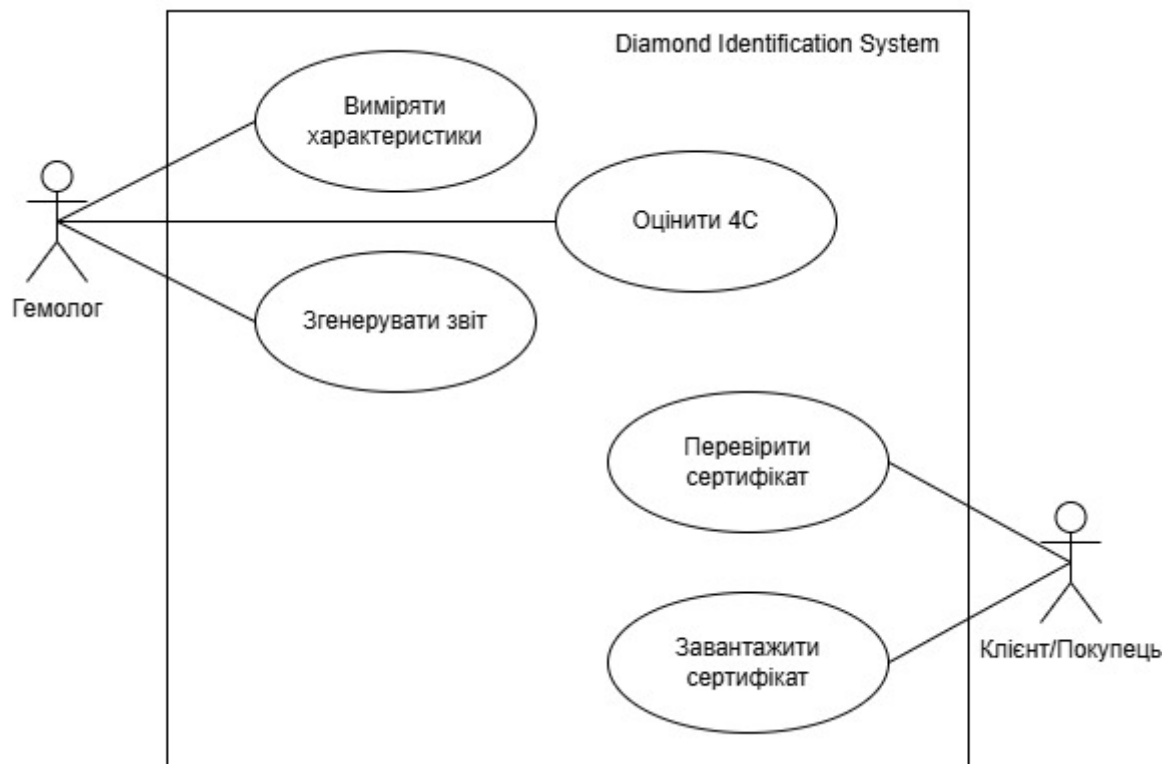


А ось так виглядає Діаграма сутностей (Entity-Relationship Diagram, ERD) для Етапу 1 (нормалізована структура OLTP):

- Сутність 1: Expert (Довідник експертів)
- Сутність 2: DiamondReport (Журнал звітів/транзакцій)
- Зв'язок: Один Expert може створити багато (*) DiamondReport (зв'язок "один-до-багатьох").



Діаграма прецедентів/акторів (Use Case Diagram):



Діаграми Взаємодії (Sequence Diagram) / Діаграма Послідовності:

