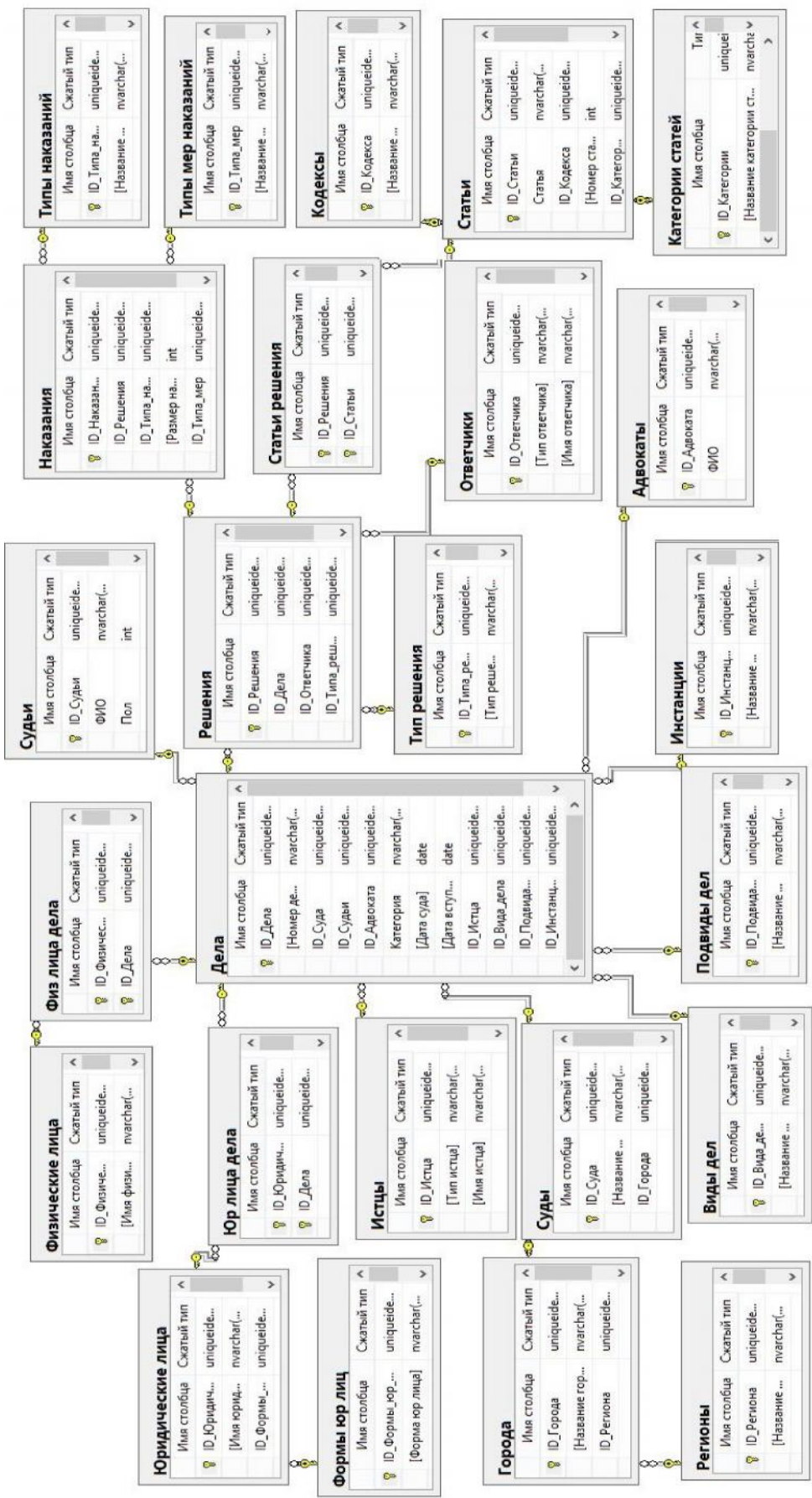


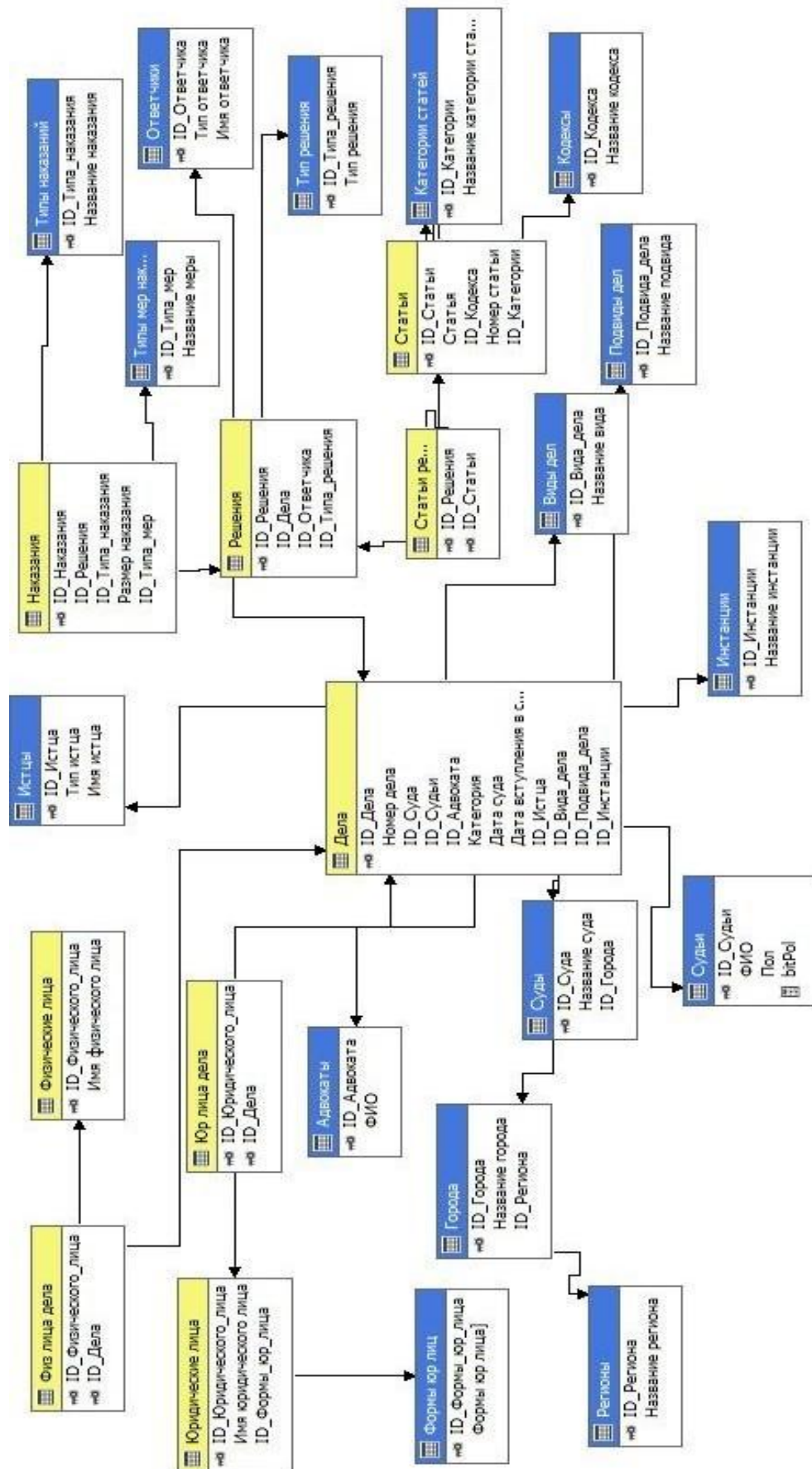
ДОДАТОК А

Фізична схема бази даних



ДОДАТОК Б

Схема куба даних судочинства



ДОДАТОК В

Графічний матеріал атестаційної роботи

Харківський національний університет радіоелектроніки
Кафедра "Електронних обчислювальних машин"

Кваліфікаційна робота на тему:
Вилучення та аналіз даних в судочинстві
за допомогою OLAP та Data Mining

Виконав: студент групи СПм-20-1 Полонець К.С.
Керівник: доц. Ільїна І.В.

Мета кваліфікаційної роботи

Метою роботи є аналіз записів, отриманих в результаті обробки документів і рішень судів за допомогою методів і засобів технології Business Intelligence.

Об'єктом дослідження є система діловодства

Задачі:

- Аналіз предметної області;
- Методи покращення існуючих систем;
- Розробка і реалізування системи аналізу даних для подальшого використання

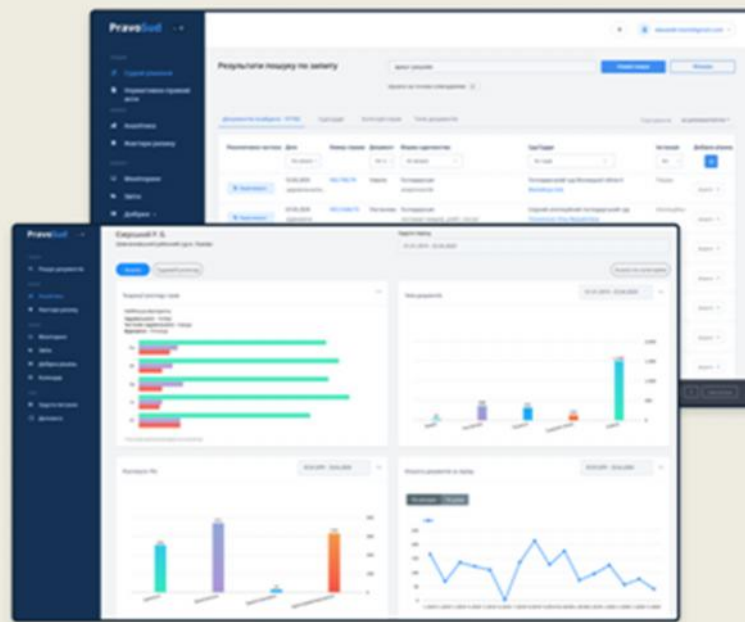
Постановка задачі

- У наш час відбувається інтенсивне накопичення величезних обсягів даних різного типу в різних предметних областях, що вимірюються в петабайтах, це в свою чергу дає можливість вирішувати завдання отримання нових фактів, залежностей і прихованих кореляцій, а також дозволяє вирішувати деякі аналітичні завдання, такі як прогнозування, перевірка статистичних гіпотез, розрахунок агрегатних показників тощо.

Актуальність задачі

- Аналіз судочинства досить актуальний, оскільки результати аналізу будуть цікавими як великим організаціям, таким як правоохоронні органи та правозахисні організації, так і окремим особам, таким як незалежні юристи та люди, які перебувають під слідством. Для великих організацій буде більш цікавий глибокий аналіз для пошуку прихованих залежностей і відхилень, а окремим особам аналіз за конкретними суддями та адвокатами, який може допомогти передбачити, яке рішення може бути винесене певним суддею або вибрати найбільш ефективного адвоката.

Огляд аналогів

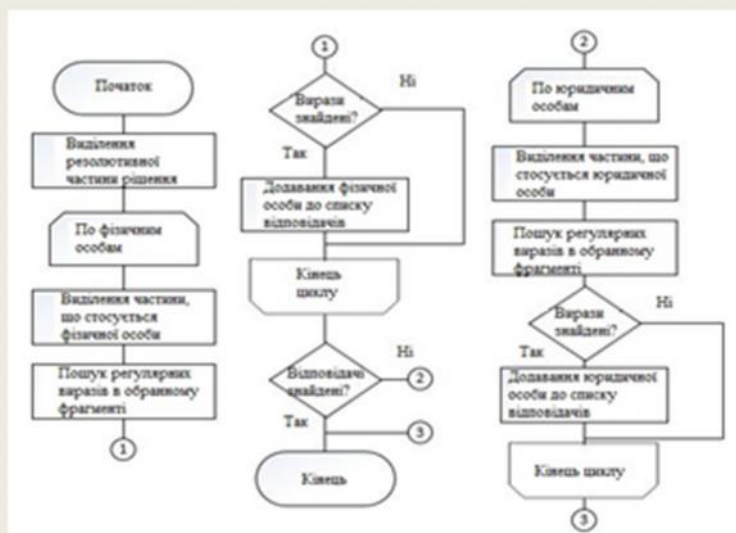


Технології, використані для розробки

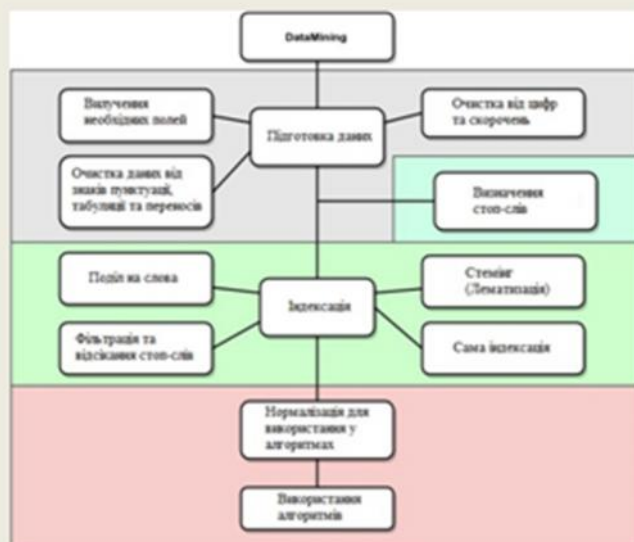
Для реалізації проекту було використано:

- Для вилучення даних – середовище PyCharm, мова програмування - Python;
- Створення БД – середовище Microsoft SQL Server Management Studio;
- Запис даних у БД – середовище Microsoft Visual Studio, мова програмування – C#.

Схема алгоритму для вилучення відповідей



Деревоподібна схема послідовності етапів у використанні технології Data Mining



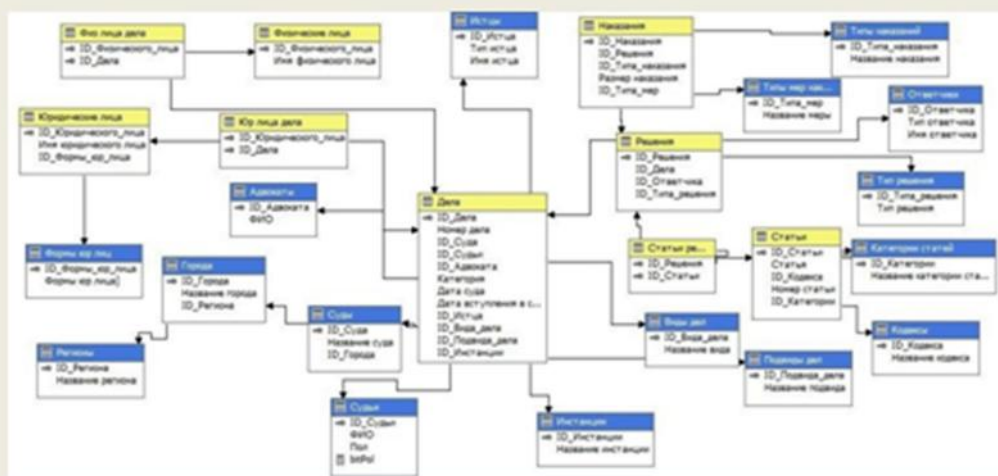
Результати типів аналізу

- **MDS** – 4000 рішень, час виконання 25 хвилин;
- **LTSA** – 4000 рішень, час виконання 19 хвилин;
- **LLE** – 4000 рішень, час виконання 7 хвилин;
- **t-SNE** – 4000 рішень, час виконання 3 хвилини.

Фізична схема бази даних



Схема куба даних судочинства (сніжинка)



Висновки

У ході виконання роботи було освоєно весь цикл технологій, що використовується в Data Mining. Зокрема, підготовка даних, очищення даних, індексування даних, формування необхідної структури даних успішного застосування алгоритмів Data Mining.

Було обрано алгоритми для обробки багатовимірних об'єктів і проведено порівняльний аналіз вибраних алгоритмів. Виявлено явні групи рішень, пов'язані між собою різними атрибутами, причому дані групи формуються повністю автоматично з допомогою технологій і алгоритмів Data Mining.