

НУБІП України

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ БІОРЕСУРСІВ
І ПРИРОДОКОРИСТУВАННЯ УКРАЇНИ

Факультет інформаційних технологій

НУБІП України

УДК 004.9:626-047.36

«ПОГОДЖЕНО»

«ДОПУСКАЄТЬСЯ ДО ЗАХИСТУ»

Декан факультету

Завідувач кафедри комп'ютерних наук

інформаційних технологій

Голуб Б.Л., к.т.н., доцент

Глазунова О.Г., д.п.н., професор

2021 р.

«30» листопада 2021 р.

МАГІСТЕРСЬКА КВАЛІФІКАЦІЙНА РОБОТА

НУБІП України

на тему Система моніторингу водних екосистем

Спеціальність

122 Комп'ютерні науки

(код і назва)

Освітня програма Комп'ютерний еколого-економічний моніторинг

(назва)

Орієнтація освітньої програми

освітньо-професійна

(освітньо-професійна або освітньо-наукова)

Гарант освітньої програми

кандидат технічних наук

Басараб Р. М.

(науковий ступінь та вчене звання)

(підпис)

(ПІБ)

НУБІП України

Керівник магістерської кваліфікаційної роботи

Бородкін Г. О.

(науковий ступінь та вчене звання)

(підпис)

(ПІБ)

Виконав

Гермаш В. О.

(підпис)

(ПІБ студента)

НУБІП України

КИЇВ-2021

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ БІОРЕСУРСІВ
І ПРИРОДОКОРИСТУВАННЯ УКРАЇНИ
Факультет інформаційних технологій

ЗАТВЕРДЖУЮ

Завідувач кафедри комп'ютерних наук

Голуб Б.Л., к.т.н., доцент

"29" жовтня 2020 року

ЗАВДАННЯ

ДО ВИКОНАННЯ МАГІСТЕРСЬКОЇ КВАЛІФІКАЦІЙНОЇ РОБОТИ СТУДЕНТУ

Гермашу Владиславу Олександровичу

(прізвище, ім'я, по батькові)

Спеціальність 122 Комп'ютерні науки

(код і назва)

Освітня програма Комп'ютерний еколого-економічний моніторинг

(назва)

Орієнтація освітньої програми Освітньо-професійна

(освітньо-професійна або освітньо-наукова)

Тема магістерської кваліфікаційної роботи Система моніторингу водних екосистем

затверджена наказом ректора НУБіП України від "29" жовтня 2020 р. №1635 «С»

Термін подання завершеної роботи на кафедру _____

(дні, місяць, рік)

Вихідні дані до магістерської кваліфікаційної роботи, впорядковані значення вимірюваних показників та сформована звітна інформація, що дозволить прийняти рішення

Перелік питань, що підлягають дослідженню:

№ з/п	Питання, що підлягає дослідженню	Строк виконання	Примітка
1.	Аналіз предметної області моніторингу водних екосистем		
2.	Моделювання предметної області		
3.	Проектування оперативної бази даних сховища даних		
4.	Розробка сховища даних		
5.	Проектування архітектури системи		
6.	Розробка алгоритмів для реалізації системи		
7.	Впровадження системи		
8.	Попередній захист		
9.	Захист		

Дата видачі завдання "29" жовтня 2020 р.

Керівник магістерської кваліфікаційної роботи _____

(підпис)

Бородкін Г. О.

(прізвище та ініціали)

Завдання прийняв до виконання _____

(підпис)

Гермаш В. О.

(прізвище та ініціали студента)

АНОТАЦІЯ

Випускна дипломна робота магістра містить 4 розділи.
Основною метою цієї роботи є оцінка якості та ефективності системи моніторингу водних екосистем для прийняття науково обґрунтованих рішень про ефективність природоохоронних заходів.

У першому розділі проведено аналіз існуючих рішень та розкритий аналіз проблемної області, представлені основні завдання, які має виконувати система, визначена мета.

Другий розділ описує підходи моделювання системи. В цьому розділі описуються побудовані діаграми, що дозволять зрозуміти роботу кожного з процесів визначення якості води.

У третьому розділі описується архітектура системи. Представлені алгоритми обробки інформації та технології

Зображені структури оперативної бази даних та сховища даних.

У четвертому розділі проводяться результати дослідження системи та вказуються рекомендації щодо впровадження інформаційної системи.

Описуються використані методи Data mining, що були застосовані для покращення процесу аналізу якості води.

ABSTRACT

The master's thesis contains 4 sections.

The main purpose of this work is to assess the quality and effectiveness of the aquatic ecosystem monitoring system for making scientifically sound decisions on the effectiveness of environmental measures.

The first section analyzes the existing solutions and analyzes the problem area, presents the main tasks to be performed by the system, the defined goal

The second section describes system modeling approaches. This section describes the diagrams that will help you understand the operation of each of the processes of determining water quality.

The third section describes the system architecture. Algorithms of information processing and technology are presented.

The structures of the operational database and data warehouse are shown.
The fourth section presents the results of the study of the system and provides recommendations for the implementation of the information system.

Describes the data mining methods used, which were used to improve the process of water quality analysis.

НУБІП України

НУБІП України

НУБІП України

НУБІП України

НУБІП України

НУБІП України

НУБІП України	ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ	6
НУБІП України	ВСТУП	7
НУБІП України	СИСТЕМНИЙ АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ	9
НУБІП України	1.1	10
НУБІП України	1.2	11
НУБІП України	1.3	Ошибка! Закладка не определена.
НУБІП України	МОДЕЛЮВАННЯ СИСТЕМИ	18
НУБІП України	2.1 Діаграма прецедентів	18
НУБІП України	2.1	19
НУБІП України	2.1	20
НУБІП України	РОЗРОБКА СИСТЕМИ	23
НУБІП України	3.1 Структура оперативної бази даних	23
НУБІП України	3.2 Структура сховища даних	24
НУБІП України	3.3 Вибір систем управління базою	25
НУБІП України	3.4 Архітектура системи	26
НУБІП України	3.5 Організаційна структура програмного забезпечення	27
НУБІП України	3.6 Вибір інструментарію для створення ППЗ	27
НУБІП України	3.7 Алгоритмізація та програмування програмних модулів	28
НУБІП України	РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ	38
НУБІП України	4.1 Вимоги до апаратного та програмного забезпечення	38
НУБІП України	4.2 Загальний вигляд системи	40
НУБІП України	4.3 Звітна інформація	43
НУБІП України	4.4 Розрахунок КРІ	45

НУБІП України

НУБІП України

НУБІП України

НУБІП України

НУБІП України

НУБІП України

НУБІП України

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

СУБД – Система управління базами даних

БД – База даних

GUI – Графічний інтерфейс користувача

HTML – Hyper Text Markup Language (мова гіпертекстової розмітки)

MVC – Model view controller (модель уявлення контролер)

ППЗ – Прикладне програмне забезпечення

ГДК – Гранично допустимі концентрації

СД – сховище даних

KPI – Key Performance Indicator

BI – Business Intelligence

НУБІП України

НУБІП України

НУБІП України

НУБІП України

НУБІП України

ВСТУП

НУБІП України

У наш час вода вважається одним із найдефіцитніших природних ресурсів на нашій планеті. Залежно від якості води, вона може бути джерелом життя та міцного здоров'я або хворобами та смертю. Тому дослідження якості води у зв'язку з кількістю забруднення є актуальним.

НУБІП України

Зростаюча деградація навколишнього середовища за останні роки, спричинена збільшенням населення та зміною клімату, збільшує потребу дослідників вивчити негативний вплив на навколишнє середовище, особливо на джерела води та її наслідки.

НУБІП України

Зростаюче забруднення води в океанах, озерах та річках у всьому світі вимагає більш досконалих методів у системах моніторингу навколишнього середовища, особливо у галузі моніторингу якості води.

НУБІП України

Традиційно визначення жості води проводилося вручну, коли були отримані зразки води та відправлені на дослідження в лабораторії, що враховує процес, витрати та людські ресурси.

В основному, існує багато параметрів, які необхідно виміряти для аналізу якості води. Проте запропонована система WQM вимірює ключові параметри води:

НУБІП України

- Значення рН води.
- Помутніння води.
- Рівень води в баку.

Температура і вологість навколишнього повітря.

НУБІП України

До теперішнього часу в багатьох організаціях накопичені значні обсяги даних, на основі яких є можливість вирішення різноманітних аналітичних і управлінських завдань.

Проблеми зберігання і обробки аналітичної інформації стають все більш актуальними і привертають увагу фахівців і фірм.

НУБІП України

В ідеалі робота аналітиків і керівників різних рівнів повинна бути організована так, щоб вони могли мати доступ до всієї інформації, що їх цікавить і

користуватися зручними і простими засобами подання та роботи з цією інформацією.

Об'єктом є оцінка якості водних екосистем.

НУБІП України

Предметом є система моніторингу.

Для дослідження та оцінки якості води потрібно опрацьовувати великі обсяги інформації, тому використання сховища даних для вирішення даної проблеми буде доцільним. Сховище даних повинно надати можливість аналізу даних в розрізі часу.

НУБІП України

Мета дослідження оцінити якість та ефективність системи моніторингу водних екосистем для прийняття науково обґрунтованих рішень про ефективність природоохоронних заходів.

НУБІП України

НУБІП України

НУБІП України

НУБІП України

НУБІП України

СИСТЕМНИЙ АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ

НУБІП України

1.1 Опис предметної області

Вода – один з найважливіших складових, що потрібні для життя. Якість води характеризується великою кількістю різних показників, саме ці показники визначають якість води для безпечного використання.

Відповідно до Водного кодексу України оцінювання якості водних ресурсів здійснюється за нормативами екологічної безпеки водокористування та екологічних норм якості водних об'єктів.

За допомогою діючих нормативів можна оцінювати якість води, що використовується для комунально-побутового, господарсько-питного і рыбогосподарського водокористування.

Для визначення якості води досліджують її хімічні та біологічні компоненти та властивості.

Одним з ключових хімічних показників це показник кислотності. Він вказує на наявність в водному середовищі іонів водню.

Рівень кислотності нижче 6.5 - 7 дає зрозуміти, що в водному середовищі підвищена кислотність.

Якщо рівень вище, то означає, що підвищена лужність води (здатність води нейтралізувати сильні кислоти).

Біологічним показником є біологічне споживання кисню. Це кількість кисню, що використовують організми у воді для свого росту та створення біомаси.

В Україні, окрім вимог щодо очищення води, також регламентують показник кислотності води. Він нормується двома стандартами:

ГОСТ2874-82 – 6.0-9.0.

ДСанПіН – 6.5-8.5.

За європейськими стандартами якості води, показник кислотності питної води повинен бути в межах – 6.5-9.5.

Гранично допустимі концентрації рівень, що встановлює допустиму концентрації різних речовин у воді, якщо значення вище певного рівня, то вода вважається непридатною для використання.

Для оцінки якості води використовують такі метод:

Метод детального аналізу;

Біологічні методи,

Хімічні методи.

Метод детального аналізу це зіставлення вимірюваних значень кожного з показників з його рівнями ГДК.

Біологічні методи в їх основі лежить використання біологічних особливостей видів та показників структур угруповань біоти водойм.

Однією з процедур біологічного методу є біотестування — це процедура за якою виконується оцінка токсичності водного середовища з використання тест-об'єктів. Тобто реакція деяких видів живих організмів на забруднення.

Ще одна процедура біологічних методів це біомоніторинг. Це безперервний процес спостереження за екологічним станом водних об'єктів.

Аналіз існуючих рішень

В наш час існує не так багато рішень, які представляють собою засоби моніторингу водних екосистем. Зазвичай це системи, що демонструють готові розрахунки, які були проведені екологічною інспекцією, що представляло собою збір інформації, аналіз вручному режимі.

Інтерактивна карта забрудненості річок України «Чиста вода». Відкрите джерело, в якому зібрано дані з понад 400 пунктів контролю річкових вод. Надає можливість переглянути до 16 показників забрудненості, а також графік змін протягом 5 років (рис.1.1).

На карті відображається інформація про найбільші річкові басейни України.

Як додаткова інформація, можна побачити кількість викидів в річки підприємств в розрізі областей України (рис. 1.2).



Рис. 1.1. Інтерактивна карта забрудненості річок України «Чиста вода»

Вибір області:

- Волинська
- Рівненська
- Житомирська
- Київська
- Чернігівська
- Сумська
- Львівська
- Тернопільська
- Хмельницька
- Вінницька
- Черкаська
- Полтавська
- Харківська
- Луганська
- Донецька
- Закарпатська
- Івано-Франківська
- Чернівецька

Ми знаємо, що ці підприємства забруднюють воду. Чому вони продовжують це робити?

Підприємства повинні очищати свою воду, яку скидають у річку. Але часом вони не можуть цього зробити, бо вони мають застарілі технології, а модернізація коштує надто дорого. Коже необхідно обладнати каналізаційні вузли. В той же час деякі підприємства одяжно встановили на своїх територіях, щоб уникнути додаткових витрат.

В деяких випадках державі може надати підприємству право скидати в річку забруднену воду. Після, коли всі річки будуть очищені, всі вони можуть скинути, обслуговувати. Вони встановили «трансакційно доступний облік», скорочено — ТДС.

ТДС потрібні для того, щоб поступово зменшувати рівень забруднення. За право забруднювати довільно підприємства платять еквівалентні податки. Якщо вони перевищують встановлені норми — то зобов'язані сплатити штраф.

Одна з функцій ТДС — зробити невідомі забруднення річок. В Україні досі державні структури еквівалентні податки. У найгіршому, що до 2020 року кількість річкових забруднень має зменшитися на 15%. Тому підприємства повинні модернізувати свої очисні споруди, щоб відповідати новим вимогам.

Оберіть свою область зі списку. Ви отримаєте список підприємств, які забруднюють річку у вашому регіоні. Ці підприємства відповідальні за значну частину забруднення в річку. Наприклад, цифра 284,3 млн ГРАТ АК «КИЇВВОДОКАНАЛ» означає, що підприємство скинуло в річку 284,3 млн кубічних метрів забрудненої води. Надаємо, що в 1 кубічному метрі міститься 1000 літрів.

ЄДРПОУ	Назва	Виток, млн куб.
00307079	ІПТ ЗАКАЛІВСЬКА ЦЕЛЮЛІЗНА ФАБРИКА	1,071
22078727	ІПТ ПЛІВАЦЬКА СІЛЬСЬКА РАДА ЗАКАЛІВСЬКОГО Р-Н	0,429
14337529	АНТ «ІНТОНОВ» М.З.М.ІСТОМБЕЛЬ	0,225
27977023	ІПТ ІВАНЕМОЇ МСЛО РАДИ ІВАНЕМОДЖИНАЛІ МУЗІН	0,176
1682644	ТОВ «ІПЕР» ВОСІК СІЛЬСЬКА РАДА ВОСІКОВСЬКОГО Р-Н	0,109
02352863	ІПТ ІФ «ВАНОВОДОКАНАЛ»	0,119
2185710	ТОВ ЧОРНОСТАРСЬКА ІВАНІВА МАМАРІОНОВА Р-Н	0,101

Рис. 1.2. Кількість викидів

Державний водний кадастр: облік повертневих водних об'єктів (рис. 1.3). Цей ресурс дає змогу побачити рівень забруднення кожного з річкових басейнів України. Окрім стану води річок, можна відобразити інформацію про пункти моніторингу та гідрологічні пости.

НУБІП України

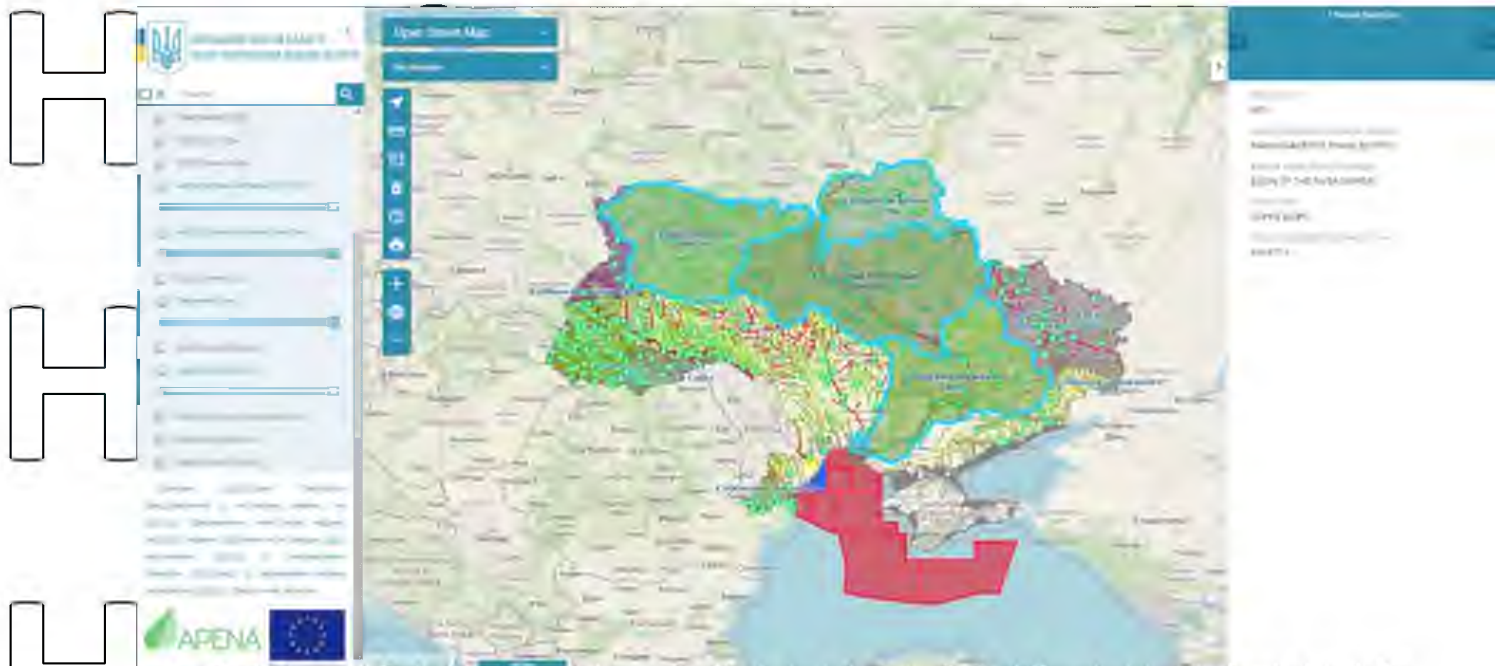


Рис. 1.3. Державний водний кадастр: облік поверхневих водних об'єктів

Державний моніторинг вод здійснюється з метою забезпечення збирання, обробки, збереження, узагальнення та аналізу інформації про стан водних об'єктів, прогнозування його змін та розроблення науково обґрунтованих рекомендацій для прийняття рішень у галузі використання, охорони вод та відтворення водних ресурсів.

Даний ресурс відображає інформацію по 10 показниках забруднення на кожному з постів моніторингу. Також можна побачити, що є і інформація про перевищення норми (Рис. 1.4).

Що найбільш важливе, це наявність таблично-графічного представлення даних, чого не було в попередніх ресурсах (Рис.1.5). Такий звіт відображає зміни кожного з показників за введений період.

Ці ресурси обмежені в можливостях та не дають змогу провести ефективний аналіз рівню забруднення води.

НУБІП України

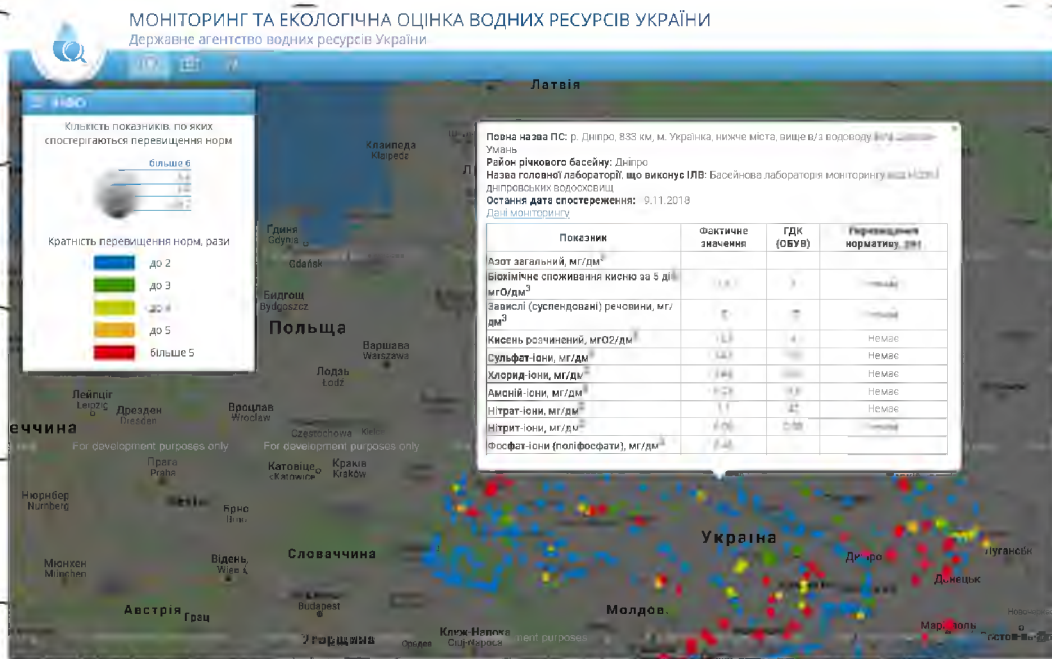


Рис 1.4. Моніторинг та екологічна оцінка водних ресурсів України

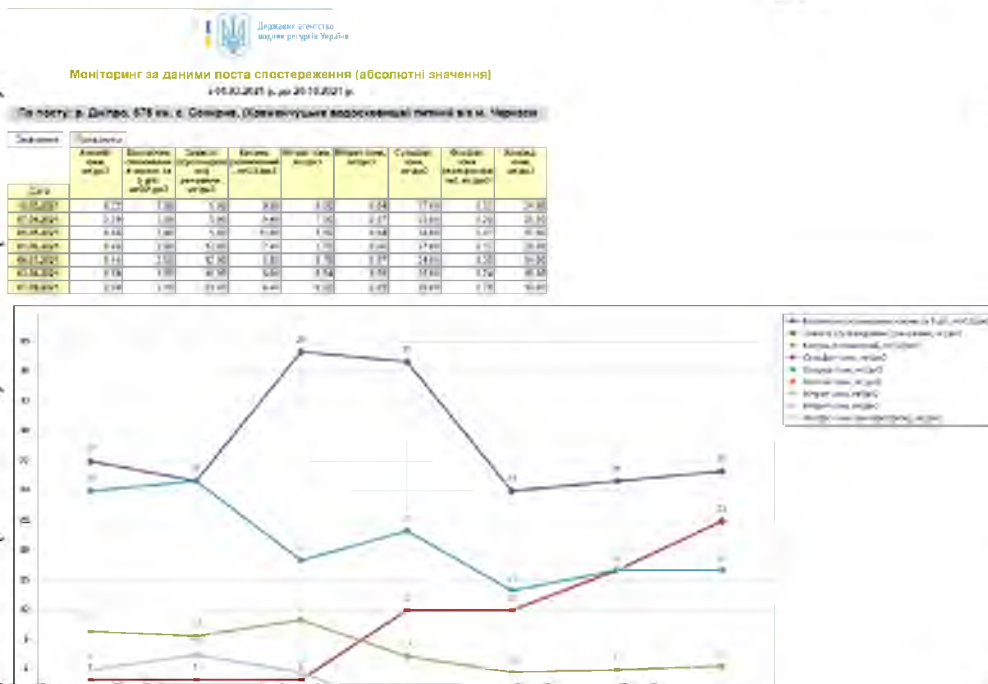


Рис 1.5. Абсолютні значення моніторингу

Найбільш наповненою платформою екологічної інформації в Україні є “Відкрите довкілля” (Рис.1.6). Даний ресурс дає більш розгорнену інформацію про стан води. На основі даних про стан води, можемо побачити відсоток забрудненості на конкретному річковому басейні та спостерігати зміни показників в часовому розрізі



Рис. 1.6. Відкрите довкілля

Дослідивши рішення моніторингу водних екосистем в Україні, можна зробити висновок, що автоматизація моніторингу водних ресурсів в Україні в секторі ІТ знаходиться на етапі розвитку. Зі статті в журналі «Нейрон» «Розумна система моніторингу якості води з економічно ефективним використанням Інтернету речей», наскільки автоматизація моніторингу води є ефективним в аналізі та прийнятті рішень, щодо запобігання та контролю рівню забруднення води.

Нейрон – це загальнонауковий журнал з відкритим доступом, який є частиною сімейства Sci Press. Будь-яка робота, що містить науково точні та цінні дослідження, що відповідає прийнятим етичним та науковим видавничим стандартам, розглядатиметься для публікації.

Система, запропонована в цій статті, є ефективним, недорогим рішенням IoT для моніторингу якості води в режимі реального часу.

Розроблена система, що має цільові плати Arduino Mega і NodeMCU, успішно взаємодіє з кількома датчиками.

Для відстеження якості води в режимі реального часу розроблено ефективний алгоритм.

Веб-додаток, тобто ThingSpeak, використовується для моніторингу таких параметрів, як значення рН, каламутність води, рівень води в резервуарі,

температура та вологість навколишнього атмосфери через веб-сервер. далі, ці вимірювані параметри також відстежуються в мобільному додатку ThingSpeak. Також цю роботу необхідно провести для аналізу кількох інших параметрів, таких як електропровідність, вільний залишковий хлор, нітрати та розчинений у воді кисень.

На відміну від вже існуючих рішень в даній системі передбачено можливість більш детального аналізу по кожному показнику, що дасть змогу скоротити час на прийняття рішень, які необхідні для запобігання забруднень водних ресурсів.

Постановка завдання

Провівши аналіз предметної області та порівнявши з вже існуючими рішеннями було вирішено створити та провести дослідження системи, що дозволить спостерігати зміни рівня забруднення в реальному часі.

Включатиме методи та алгоритми за якими буде виконуватись аналіз змін та прийняття рішень.

Так як моніторинг водних ресурсів потребує обробки великої кількості даних було прийнято рішення використати сховище даних (СД) це дасть можливість виконати збір інформації з різних, часто розрізнених джерел і дозволить представити дані в зручному форматі, відповідному потребам кінцевих користувачів і напрямками їх діяльності.

Розробка СД покликана створити ефективне інформаційне забезпечення для вирішення конкретних стратегічних завдань.

Задоволення якихось приватних потреб або автоматизація звичайних щоденних процедур, пов'язаних з обробкою інформації окремими співробітниками, в цілі та завдання створення СД не входять.

Цим системи, побудовані на технології сховища даних, і відрізняються від різного роду систем оперативної обробки даних.

Перед створенням СД потрібно визначити:

Напрямок діяльності та предметну область;

Конкретні дані, що будуть знаходитись в сховищі;

- Процедури роботи з даними.

Побудоване сховище даних повинно відповісти на такі питання:

НУБІП УКРАЇНИ

- Які території мали найбільшу забрудненість за рік?
- Середні показники якості води в розрізі сезонів.
- Визначення якості води в розрізі показників забрудненості.

За допомогою використання технологій Data Mining, потрібно дослідити

ефективність системи, як такої, що зможе виконувати аналіз великого об'єму

даних обробки та приведення цих даних в зрозумілий вигляд для прийняття рішень.

НУБІП УКРАЇНИ

НУБІП УКРАЇНИ

НУБІП УКРАЇНИ

НУБІП УКРАЇНИ

НУБІП УКРАЇНИ

НУБІП УКРАЇНИ

МОДЕЛЮВАННЯ СИСТЕМИ

НУБІП України

2.1 Діаграма прецедентів

Для візуального представлення взаємодії прецедентів з акторами було розроблено діаграму прецедентів (рис. 2.1. Діаграма прецедентів).

Акторами в нашій системі є :

Адміністратор роботи з датчиками. Його основною задачею є контроль роботи датчиків в системою, він відповідає за внесення показників вимірювань

Датчик. Основне джерело даних показників вимірювань

Еколог. Оцінює якість води на основі отриманих вимірювань та сформованих звітів. Формує звіти зі своїми висновками, щодо рівня якості водних ресурсів.

Аналітик Екологічних питань. Саме він виконує прогнозування змін якості показників. Формує та визначає ключові показники ефективності.

Керівник екологічної інспекції. Роль керівника це прийняття рішень про заходи зменшення забрудень.

В системі моніторингу водних екосистем виділено наступні прецеденти:

внесення показників вимірювань. Актором в даному випадку буде датчик. За допомогою даного прецеденту буде відбуватись внесення основного об'єму інформації;

налаштування датчиків. Адміністратор роботи з датчиками виконує налаштування датчиків;

формування поточних звітів. На основі отриманих даних еколог формує поточні звіти;

перегляд поточних звітів. З даним прецедентом взаємодію три актори системи:

еколог, аналітик екологічних питань та керівник екологічної інспекції. Всі вони мають змогу переглянути поточні звіти, що були сформовані екологом;

- формування KPI. Формуванням KPI займається аналітик екологічних питань
- прийняття рішень про заходи зменшення забруднень. Керівник екологічної інспекції на основі раніше сформованих звітів та KPI приймає рішення.



Рис. 2.1. Діаграма прецедентів

2.1 Діаграма послідовності

Для представлення взаємодії між об'єктами системи, побудовано діаграму послідовності. В діаграмі послідовностей об'єкти розташовані в порядку передачі інформації в часі.

На діаграмі послідовності на рисунку 2.2 показано, як користувач може формувати журнал та звітну інформацію. Основним результатом, що потрібен для аналізу та прийняття рішень, являються звіти, саме вони допоможуть зрозуміти ситуацію забрудненості та впливають на рішення. Послідовність дій, яка потрібна для отримання результату наступна. Потрібні дані показників вимірювань, саме вони є основним елементом в досягненні результату.

Користувач вносить всі дані вимірювань кожного показника в систему або це виконується автоматично з додаткових датчиків та пристроїв, що в реальному часі поповнюють нашу систему актуальними значеннями показників.

Вся інформація по внесеним показникам буде зберігатись в журналі, де можна побачити, на якій території та в який час були внесені показники.

На основі готового сформованого журналу показників за допомогою методів Data mining та інших алгоритмів формуються звіти та ключові показники ефективності.

І в результаті користувач отримує готові сформовані звіти, по яким може робити висновки та приймати рішення.

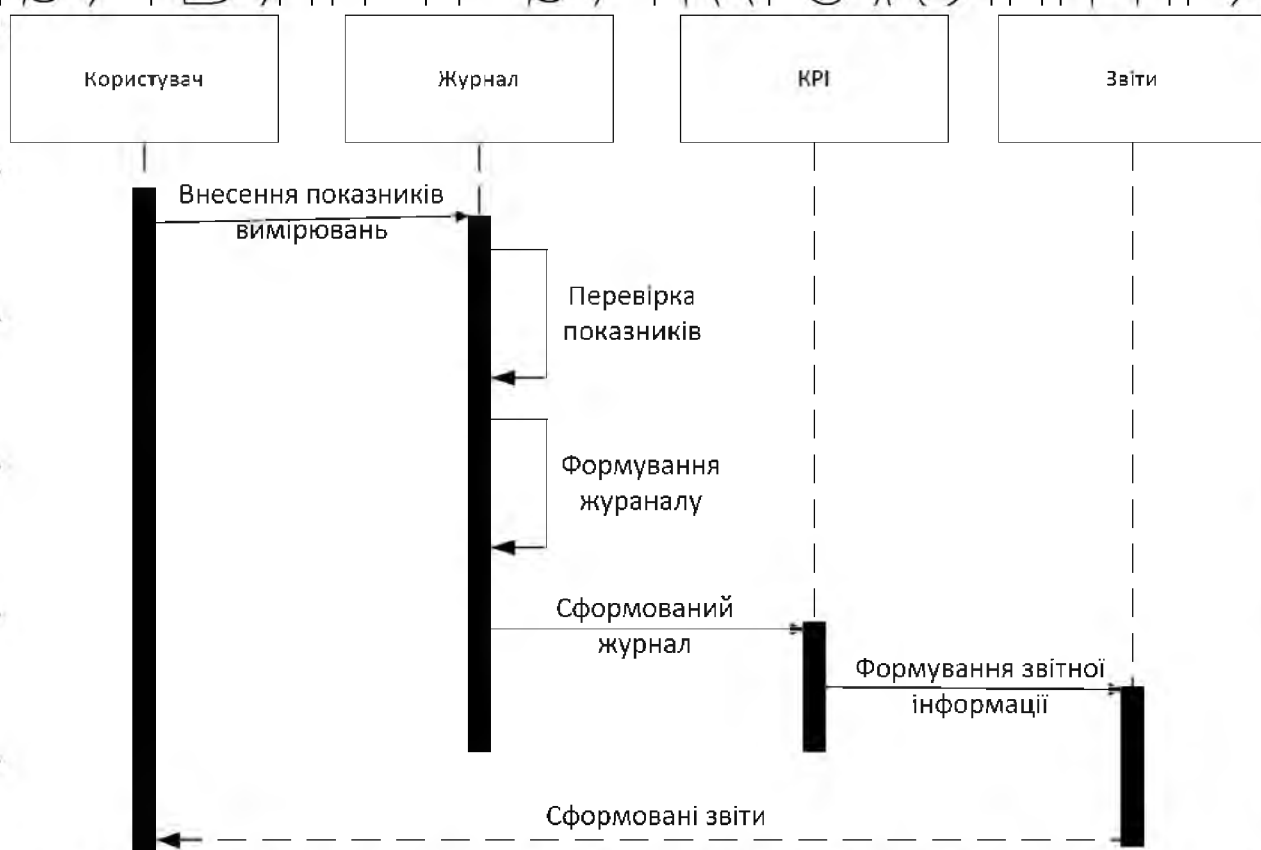


Рис. 2.2. Діаграма послідовності

2.1 Діаграма активності

Для зображення послідовного великого потоку діяльності, було спроектовано діаграму активності (Рис.2.3).
 На діаграмі описані основні аспекти поведінки системи. Зображений алгоритм дій показує, як наша система буде приймати та обробляти зібрані значення показників.

Як бачимо над внесеними показниками відбуваються такі процеси, як розрахунок КРІ та генерація звітної інформації.

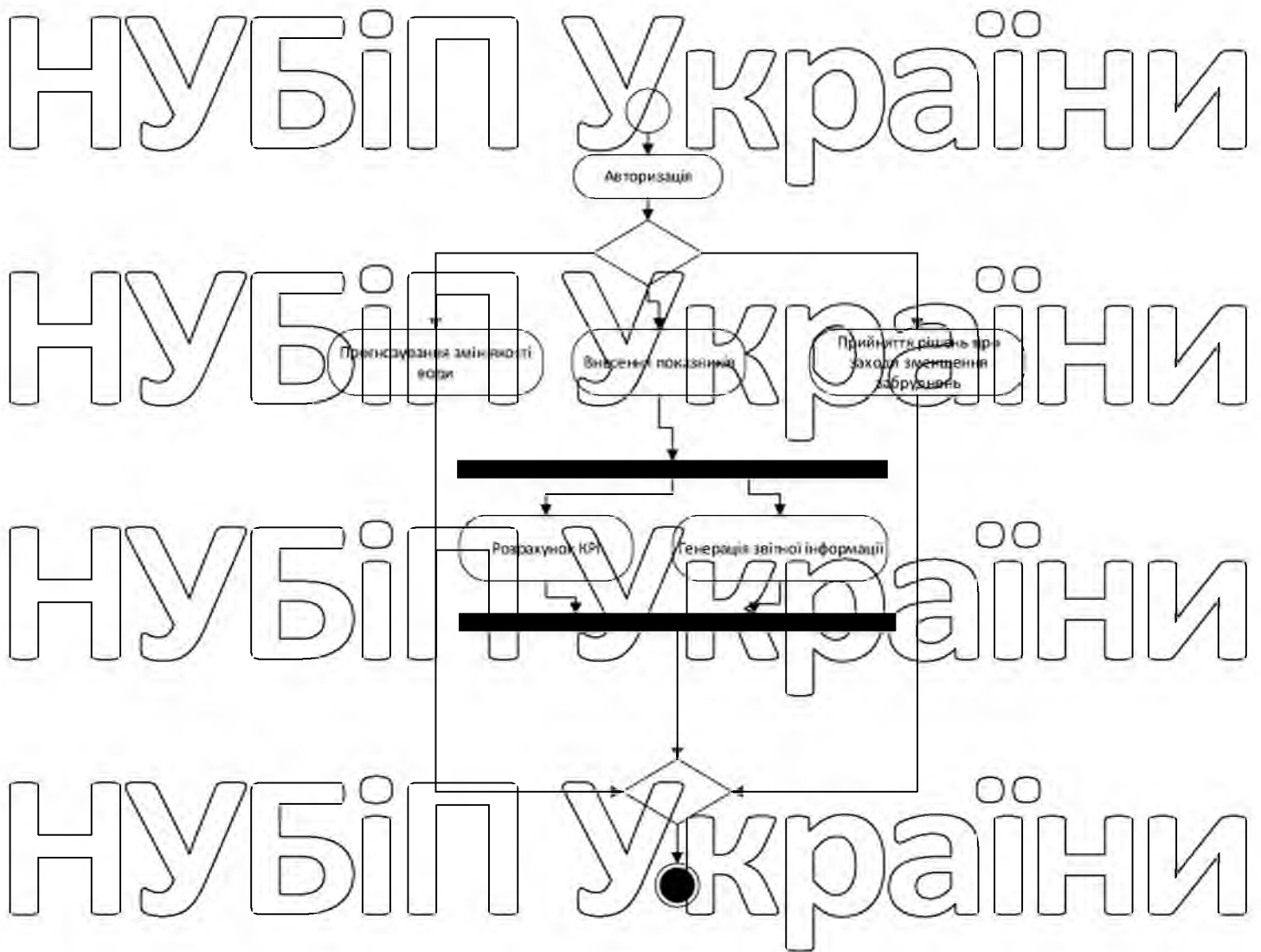


Рис. 2.3. Діаграма активності.

РОЗРОБКА СИСТЕМИ

НУБІП України

3.1 Структура оперативної бази даних

Для точного визначення та аналізу якості води потрібно великий обсяг даних. Щоб дослідити як змінюється рівень показників та зробити висновки потрібно зберігати дані за тривалий проміжок часу.

Найкращим рішенням є створення сховища даних. Заповнення сховища даними може виконуватись як за допомогою оперативної бази даних (рис. 3.1) так і з використанням зовнішніх джерел.

В оперативній базі даних заплановано зберігання даних за декілька останніх місяців.

В оперативній БД представлені такі сутності:

Indicator – інформація про датчики, які використовуються у системі

Indicator_Param – дані налаштування показників

Water_Category – довідникова інформація про типи водоєм

Place_Measurment – території, на яких ведуться виміри показників

Measurment_Log – журнал виконаних замірів показників

Зовнішніми джерелами є датчики та сенсори вимірювання показників, дані яких можуть потрапляти до окремо від БД.

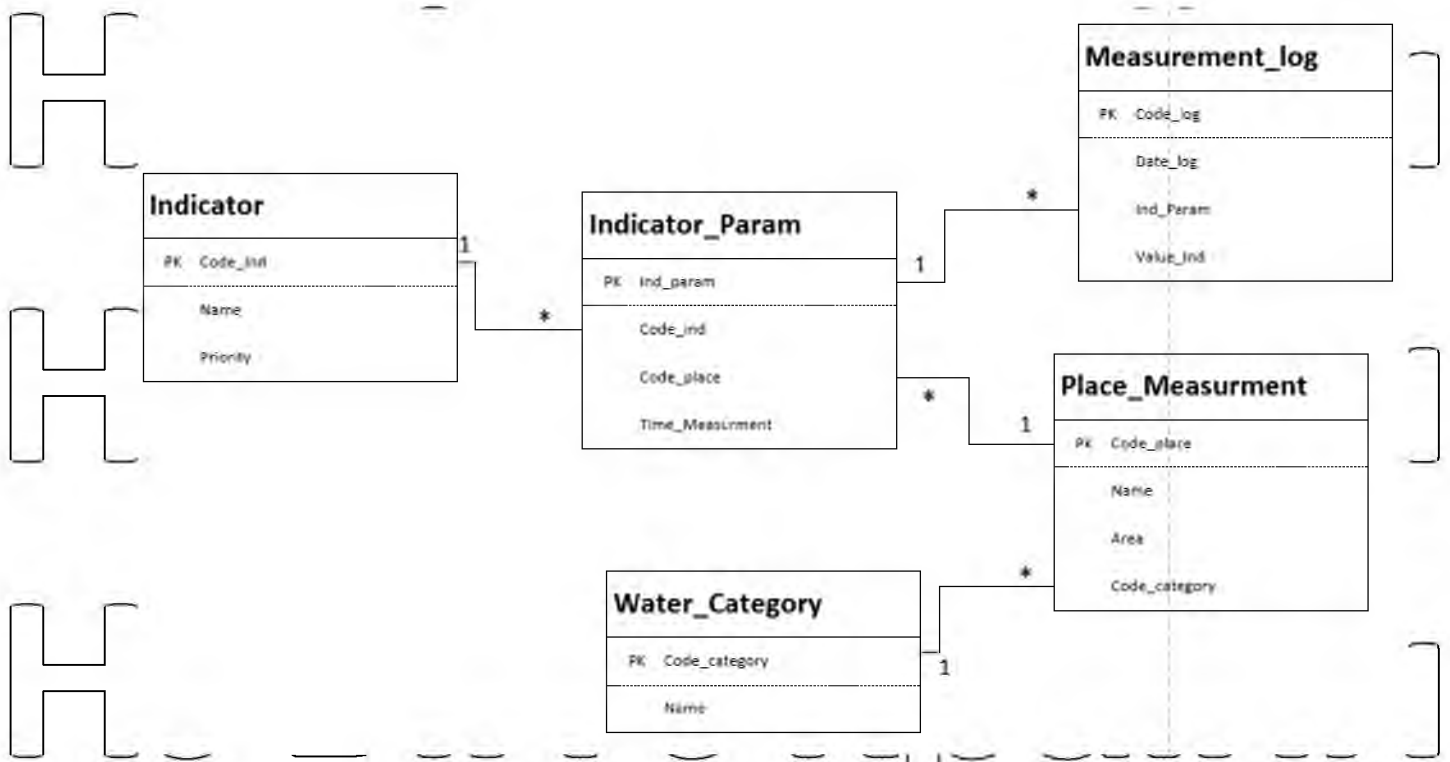


Рис. 3.1. Структура оперативної бази даних

3.2 Структура сховища даних

Основним завданням нашої системи є підтримка прийняття рішень за допомогою аналізу зібраних даних. Оскільки потрібно зберігати великі обсяги даних за тривалий проміжок часу, завдяки використанню OLAP технологій система може мати доступ до такої кількості даних. І саме головне, що OLAP допоможе у вирішенні інтелектуального аналізу даних та прийняття правильних і чітких рішень.

Сховище даних побудовано шляхом інтеграції даних з безлічі різномірних джерел, які підтримують аналітичну звітність, структуровані та або спеціальні запити та прийняття рішень. Зберігання даних передбачає очищення даних, їх інтеграцію та консолідацію.

В сховищі даних показано (рис.3.2.) схему «крижинка».

Таблицями вимірів виступають DateDim (збереження часової інформації), IndicatorDim (показники системи), Indicator_ParamDim (налаштування показників), Water_CategoryDim (дані про типи водойм), Place_MeasurmentDim (території вимірювань)

Також на сусемі зображено таблицю фактів, що зв'язана з усіма таблицями вимірів та зберігає у собі значення над якими будуть виконуватись розрахунки.

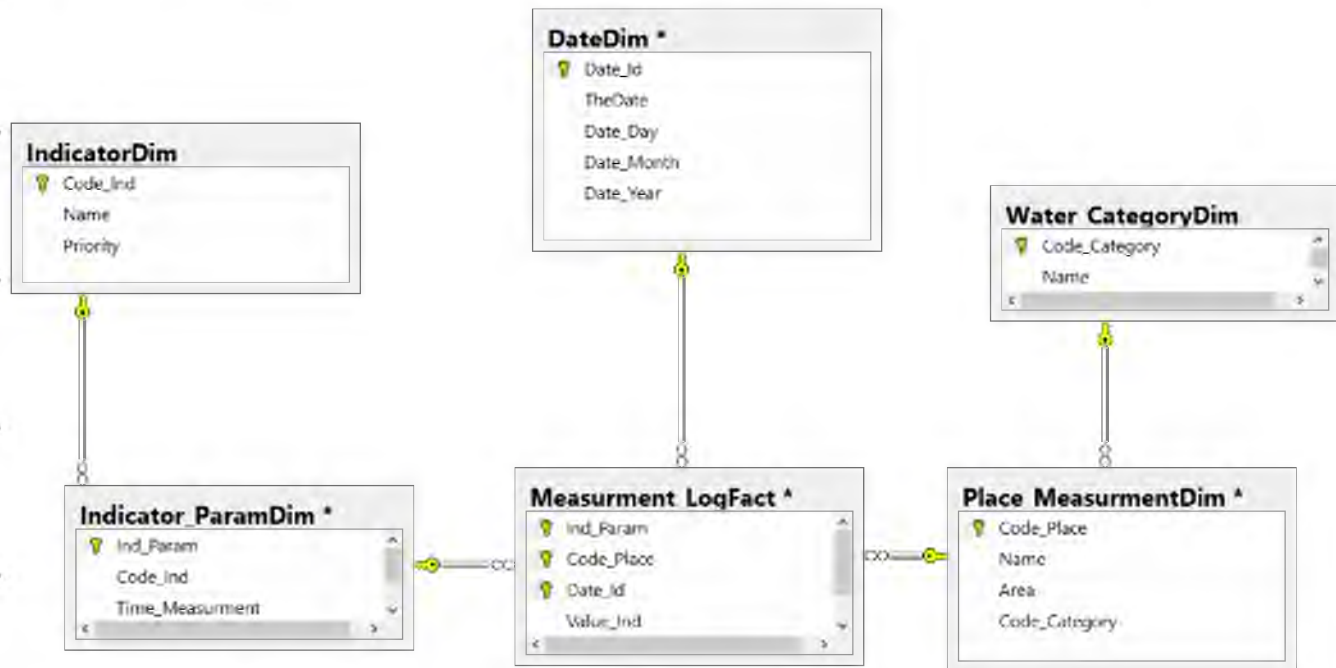


Рис. 3/2. Структура сховища даних

3.3 Вибір систем управління базою

Реляційна база даних реалізована за допомогою СУБД MS SQL SERVER. Це можна пояснити наступними можливостями даної СУБД:

однією з переваг SQL Server це простота його застосування, зокрема адміністрування. SQL Server Manager, що входить до складу всіх редакцій Microsoft SQL Server, являє собою повнофункціональний і досить простий засіб для адміністрування цієї СУБД;

надійність і безпека. SQL SERVER надає шифрування даних;

швидкість.

І так як система розробляється з використанням .net платформ, дана СУБД надасть великий вибір інструментів для роботи

3.4 Архітектура системи

Для побудови надійної системи найважливішим етапом є проектування архітектури системи. Добре спроектована організація компонентів системи їх

якість і взаємодія. Дозволить спростити її розробку та в подальшому зменшить ризики виникнення проблем в роботі системи чи в розширенні функціоналу або інших нововведень.

Архітектура системи зображена на рис. 3.3.

Вузлами на діаграмі робочі станції для еколога та адміністратора.

Також вузол робочої станції з модулем аналітик. В системі знаходиться база даних, в яку надходить інформація після змін екологом, дані дані з БД переходять в сховище даних, також виділено зовнішні джерела, якими виступають датчики та сенсори вимірювання показників.

За допомогою модуля аналітики проводяться всі розрахунки з даними такі як розрахунок КРД що дозволяють прийняти рішення.

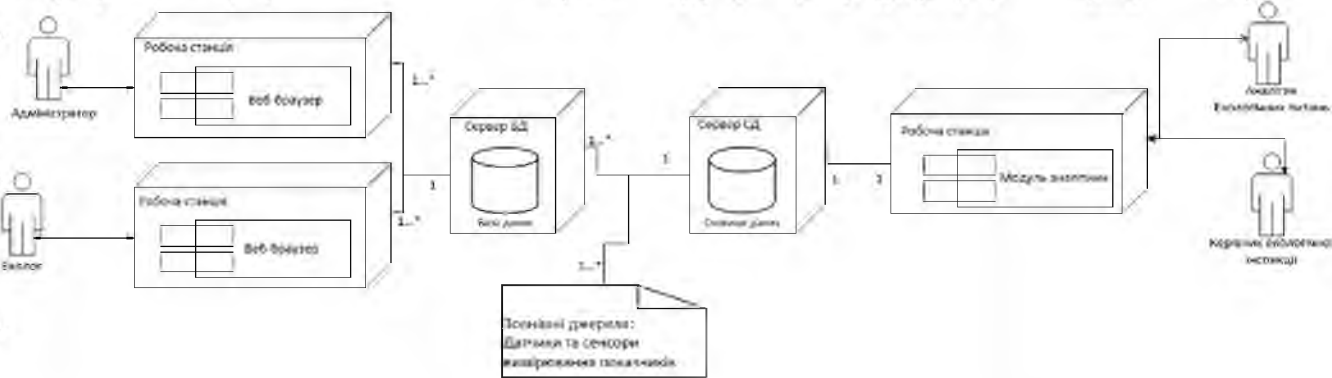


Рис. 3.3. Архітектура системи

3.5 Організаційна структура програмного забезпечення

Для стабільної та надійної роботи в інформаційній системі повинно бути мінімум три функціональні частини – модулі для зберігання даних, можливість їх оброблення та інтерфейс користувача.

Найбільш доцільним є використання клієнт-серверної архітектури. Це дозволить розподілити систему на два фізичні модулі.

Для чіткого розмежування користувацького інтерфейсу і логіки системи на окремі компоненти, було застосовано схему розподілення MVC.

Шаблон MVC призначений для розробки динамічних веб додатків, полягає в розбитті на блоки, що будуть відповідати за вирішення різних задач.

Включає в себе наступні компоненти:

модель – це компонент, що використовується для роботи з даними, які зберігаються та оброблюються на сервері;
представлення – компонент, який відповідає за взаємодію з користувачами;
контролер – відповідає за зв'язок між моделлю і представленням. Контролер обробляє дії і запити користувача.

3.6 Вибір інструментарію для створення ПІЗ

Для реалізації системи було обрано Visual Studio 2019, оскільки для розробки розробка системи було обрано мову програмування C# та технологію .NET ASP.MVC.

Visual Studio 2019 – це інтегроване середовище розробки, який надає можливість для розробки додатків для Windows, Android і iOS, а також сучасних веб-додатків.

Перевагами даного середовища розробки є:

можливість швидко і просто налаштувати взаємодію з СУБД MS SQL SERVER, який використовується для розробки даної системи;

підтримка технологій та мов програмування, які потрібні для реалізації системи C#, ASP.Net MVC, HTML / JavaScript;

підтримка систем управління версіями, що дозволяє зберігати систему та моніторити процес розробки;

зручний механізм відладки. Та внесення програмних змін без повторної компіляції проекту.

3.7 Алгоритмізація та програмування програмних модулів

3.7.1 Моделювання бізнес-логіки системи. З метою спрощення структури та організації роботи з моделю системи, розроблено діаграму пакетів (рис.3.4).

BusinessLogic містить весь код, який реалізує бізнес-логіку сервера. GUI описує інтерфейс користувача. DBAccess містить доступ та логіку для роботи з базою даних. DataBaseServer містить базу даних. В DataWarehouseServer розміщено

сховище даних. AnalyticBlock містить програмний код та алгоритми для реалізації аналізу даних.

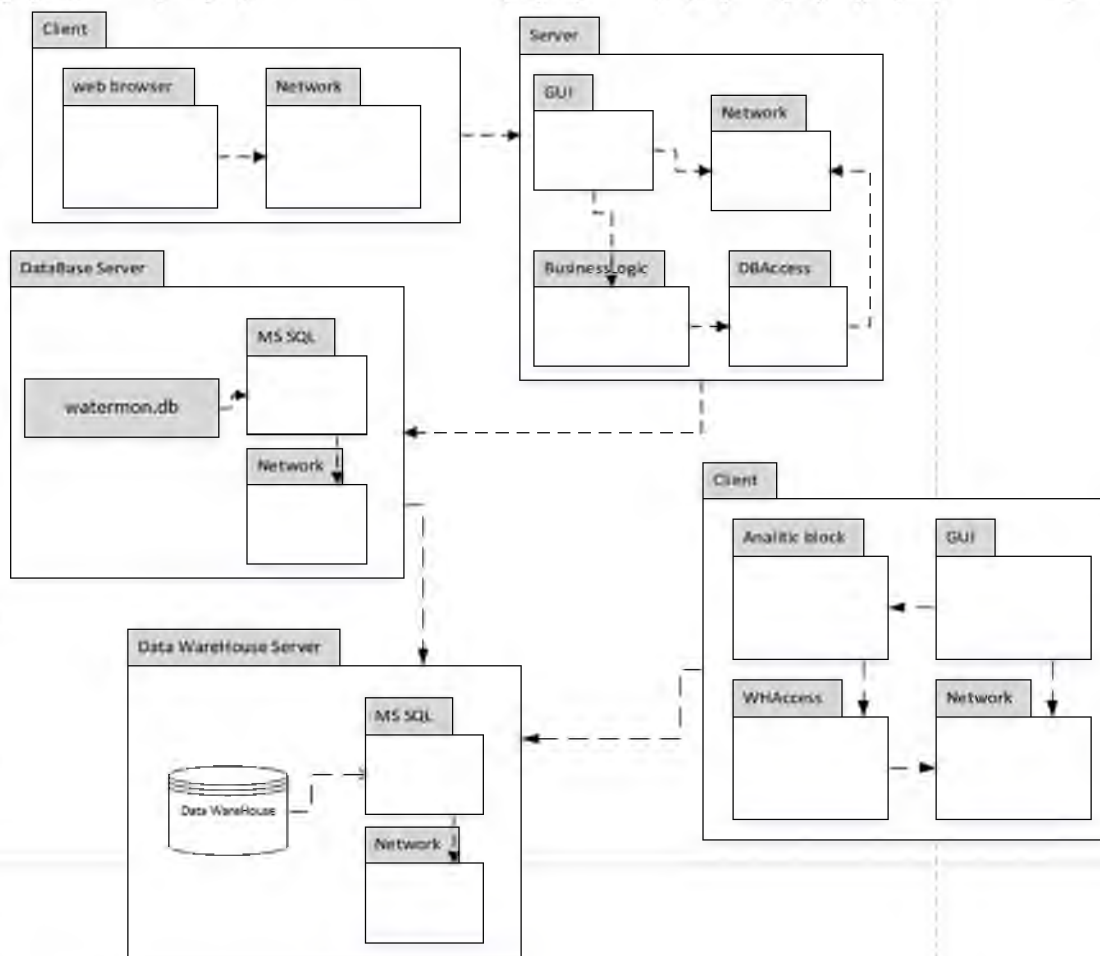


Рис. 3.4 Діаграма пакетів

Для реалізації основних механізмів, що потрібні для коректної і правильної роботи спроектовано блок-схеми. За допомогою блок-схем вдалося зобразити послідовність виконання алгоритмів та механізмів необхідних для реалізації вирішення проблем.

Процес внесення показників до системи на перший погляд не представляє нічого складного і не потребує великих зусиль для реалізації. Але це один з найважливіших етапів на шляху до виконання аналізу. Поток з великої кількості даних потрібно контролювати, тому додаткові перевірки на правильність занесених даних показників, дасть змогу уникнути помилок в майбутніх обрахунках якості води.

Перевірка коректності введених значень показників зображена на рис.3.5.

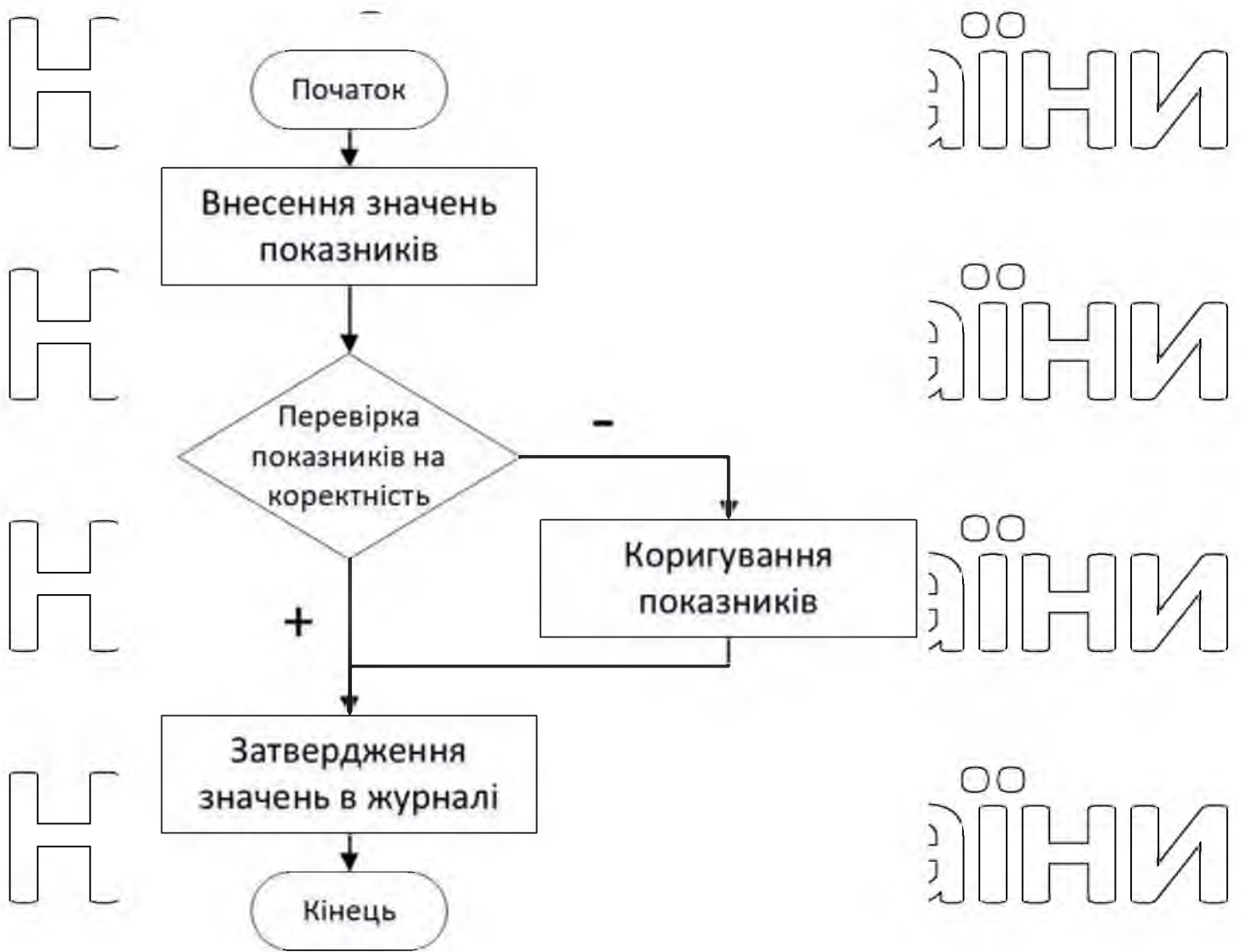


Рис. 3.5 Блок-схема Перевірка коректності введених значень показників

Формування висновків, що допоможуть прийняти рішення зображені на блок-схемі рис.3.6. Вказавши параметри для відбору даних по показникам виконується аналіз ГДК.

На основі аналізу сформованих звітів та графіків користувачі можуть дослідити отримані результати та прийняти рішення, які необхідні для зменшення забруднення.

НУБІП України



Рис. 3.6 Блок-схема формування висновків

Дослідивши предметну область були виділені загальні правила та алгоритми для аналізу якості води. Всі ці правила та алгоритми відображені в блок-схемі

(рис.3.7)

За заданими параметрами відбору, отримуємо дані для аналізу.

За допомогою занесених норм кожного показника, до нашої системи,

виконується порівняння та класифікація на дві групи: показники в межах норми та показників, які перевищують норму.

Далі виконуються додаткові класифікації по територіям та періодам.

Надається можливість формування проміжних висновків та генеруються відповіді в форматі звітності.



Рис. 3.7 Блок-схема реалізації аналізу ГДК.

3.7.2 Реалізація алгоритмів функціонування системи. Механізм аналізу показників реалізований на мові програмування C#.

Для визначення входження значення показника до норми використано наступний розрахунок: фактичне значення показника / показник базового рівня.

Для побудови класифікації було використано алгоритм побудови елементарних правил 1Rule, цей алгоритм будує правила за значенням тільки однієї незалежної змінної.

Алгоритм 1rule представляє собою пошук змінної, що дозволить максимально точно класифікувати об'єкти.

Алгоритм працює наступним чином. Необхідно перевірити кожну змінну шляхом обчислення «помилки» — кількості об'єктів, що не задовольняють правилу з тих, що мають значення цієї змінної. У підсумку обирається змінна з найменшою помилкою.

Пошуку за однією змінною не завжди буде достатньо, щоб відповісти на всі питання та прийняти рішення. Для класифікації з декількома незалежними змінними було застосовано алгоритм Наївного байеса.

Модель наївного байєсівського класифікатора приймає два припущення:

- порядок проходження ознак об'єкта не має значення;
- ймовірності ознак не залежать одне від одного при даному класі.

Ідея алгоритму полягає в розрахунку умовної ймовірності належності об'єкта до c_r при рівності його незалежних змінних певним значенням. З теорії

ймовірності відомо, що її можна визначити за формулою (рис.3.8)

$$P(y = c_r | E) = P(E | y = c_r) \cdot P(y = c_r) / P(E).$$

Рис.3.8. Формула розрахунку умовної ймовірності об'єкта.

Для кожного з правил за формулою Байеса визначається його ймовірність.

Припускаючи, що незалежні змінні приймають значення незалежно один від одного, висловимо ймовірність $P(E | y = c_r)$ через добуток ймовірностей для кожної незалежної змінної (рис.3.9.).

$$P(E | y = c_r) = P(x_1 = c_p^1 | y = c_r) \cdot P(x_2 = c_a^2 | y = c_r) \cdot \dots \cdot P(x_m = c_b^m | y = c_r).$$

Рис.3.9. Формула ймовірності через добуток ймовірностей для кожної змінної

Імовірність належності об'єкта до класу c_r за умови рівності його змінної x_h деякому значенню c_d^h визначається за формулою (рис.3.10):

$$P(x_h = c_d^h) = P(x_h = c_d^h \text{ і } y = c_r) / P(y = c_r)$$

Рис.3.10. Імовірність належності об'єкта до класу

Одним з дійсних переваг цього методу є те, що пропущені значення не створюють ніякої проблеми. При підрахунку ймовірності вони просто пропускаються для всіх правил, і це не впливає на співвідношення ймовірностей.

3.7.3 Забезпечення інтерфейсу з базою даних.

Так як система розробляється на мові програмування C# з використанням технології ASP .NET MVC, тому використано засоби .NET.

Платформа .Net Framework пропонує велику кількість способів взаємодії з базами даних.

Ado.Net технологія складається з керуваних класів, які дозволяють підключатись до джерел даних, виконувати команди і керувати автономними даними.

Основною особливістю є написання однакового коду для доступу до даних, як в веб-додатках, так і в клієнт-серверних настільних додатках.

Entity Framework – Арі-інтерфейс/сироектований для роботи з будь-якими базами даних. Також включає в себе альтернативу Linq.

Dapper представляє собою технологію мапінга результатів sql-запитів з класами C#. Забезпечує велику продуктивність і виконує запити швидше, ніж інші.

Для системи обліку показників спортивних змагань найбільше підходить Dapper, оскільки виконується швидше і підходить для розробки веб-додатків.

Для роботи з базою даних було створено репозиторій в якому зберігаються всі дані, які використовуються в процесі роботи з системою.

3.7.4 Формування звітної інформації.

Для побудови графіків було використано **java script**, оскільки являється зручним для створення графіків.

Для побудови графіків є багато засобів одним з таких є **Morris**.

Morris – засіб, що використовується у побудові графіків з використанням java script.

Morris дозволяє створювати красиві графіки в найкоротші терміни.

Створення графіків та діаграм досить просто. Засіб надає можливість створювати різні види графіків та діаграм.

Power BI Desktop – це додаток, що дозволяє підключатись до джерел даних та перетворювати і візуалізувати дані.

За допомогою **Power BI** можна підключатись, до декількох різних джерел даних, також об'єднувати їх в моделі даних.

Ця модель даних дозволяє створювати візуальні елементи та колекції візуальних елементів, якими можна ділитися з іншими людьми у вигляді звітів. Більшість

користувачів, що працюють над проектами бізнес-аналітики, застосовують

Power BI Desktop для створення звітів, а службу **Power BI** - для надання до звітів загального доступу для інших користувачів.

3.7.5 Опис BI та створення в його середовищі проекту служби SSAS (побудова розгорнутого куба)

SQL Server Business Intelligence (BI) – це низка інструментів, що дозволяє перетворювати необроблені дані в інформацію, що підприємства можуть використовувати для прийняття рішень.

Інструменти бізнес-аналітики використовуються для того, щоб дозволити розробникам виконувати операції вилучення, завантаження та перетворення над даними зі сховищ даних.

Компоненти програми включають в себе показники ефективності, аналітику, звітність підприємств.

Служби аналізу **SQL Server (SSAS)** - технологія стеку **Microsoft Business Intelligence** для розробки рішень онлайн-аналітичної обробки (**OLAP**).

Простіше кажучи, **SSAS** використовують щоб створити розгорнутий куб, з використанням даних з вітрин даних / сховища даних для глибшого та швидшого аналізу даних.

Використання розгорнутого кубу дозволить аналізувати великі обсяги даних та без зайвих зусиль розподілити за розмірами.

Для створення розгорнутого кубу потрібно обрати таблиці, що ми хочемо

включити до кубу.

Після завершення процесу створення розгорнутого кубу автоматично згенерується схема кубу (рис.3.11).

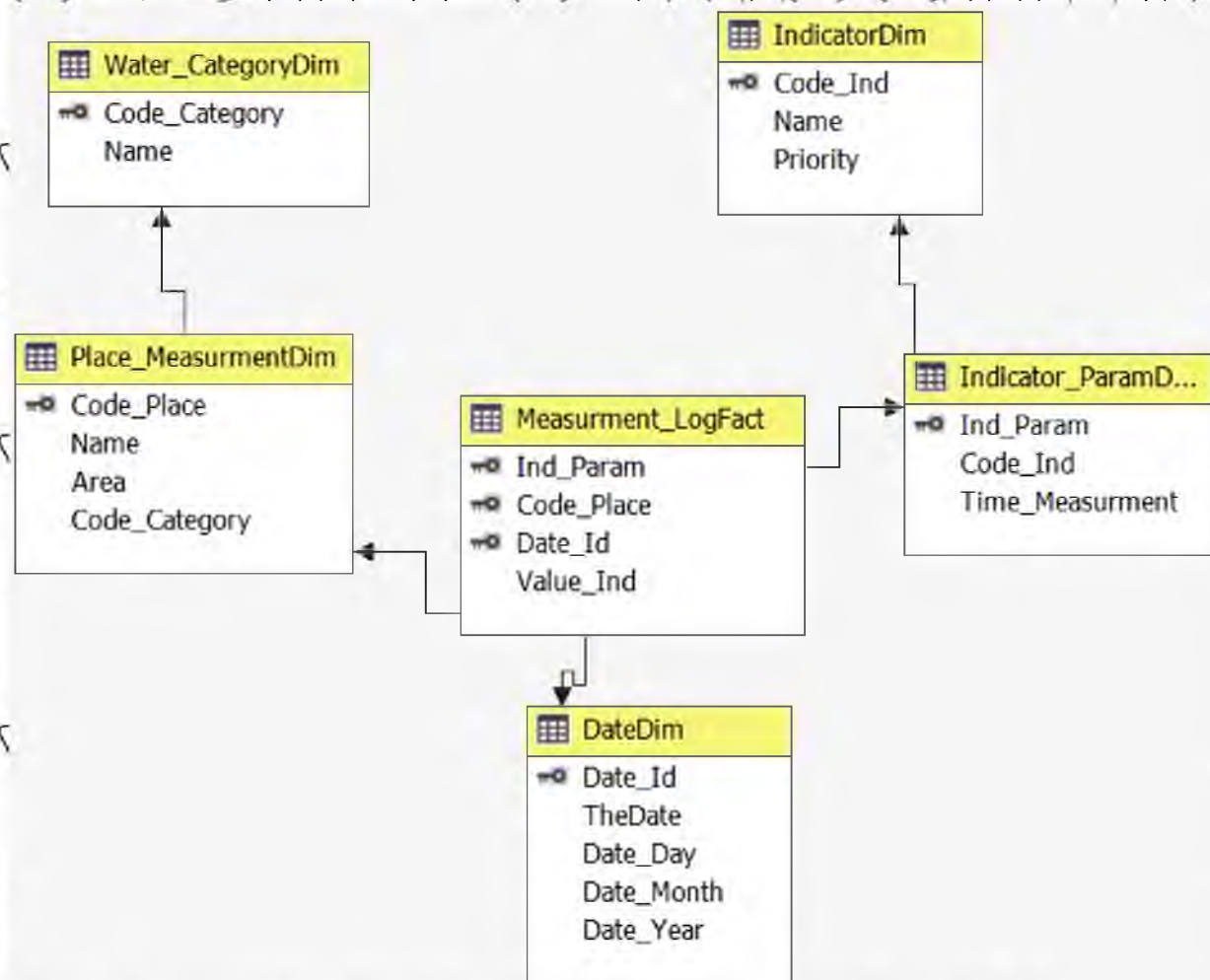


Рис.3.11. Схема розгорнутого кубу

3.7.6 Реалізація отримання даних за допомогою Data Flow

Служба інтеграції SQL Server (SSIS) - компонент програмного забезпечення бази даних Microsoft SQL Server, що використовується для виконання завдань міграції даних.

SSIS - швидкий і гнучкий інструмент зберігання даних, що використовують для вилучення, завантаження та перетворення даних, таких як агрегування, очищення, злиття даних.

Це дозволяє полегшити процес переміщення даних з однієї бази в іншу.

SSIS витягує дані з різноманітних джерел, таких як бази даних SQL Server, бази даних Oracle, файли Excel та DB2.

Структура за якою відбувається міграція даних зображена на рис. 3.12 – 3.15.



Рис.3.12. Загальна структура

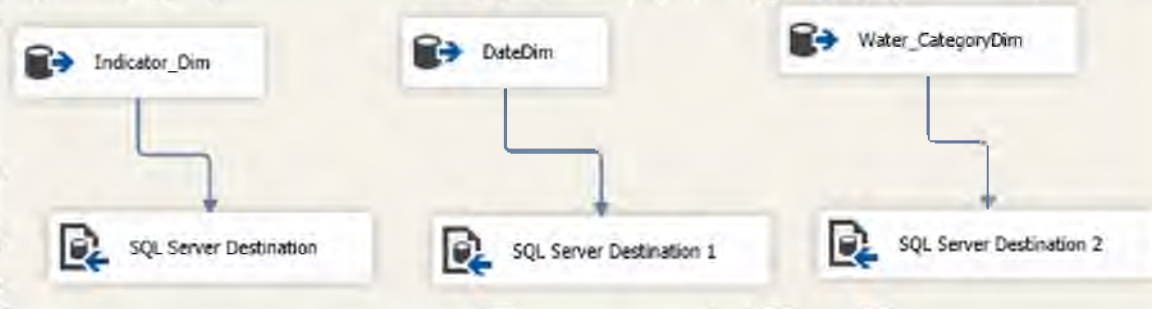


Рис.3.13. Потоки даних задач 1 рівня



Рис.3.14. Потіки даних задач 2 рівня



Рис.3.15. Потіки даних задач 3 рівня

РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

4.1 Вимоги до апаратного та програмного забезпечення

4.1.1 Діаграма розміщення. Розробка діаграми розміщення, як правило, є останнім етапом специфікації моделі програмної системи.

При розробці діаграми розміщення переслідують наступні цілі:

- визначити розподіл компонентів системи по її фізичних вузлах;
- показати фізичні зв'язки між всіма вузлами реалізації системи на етапі її виконання;
- виявити вузькі місця системи і реконфігурирувати її топологію для досягнення необхідної продуктивності.

Діаграма розміщення зображена на рисунку 4.1.

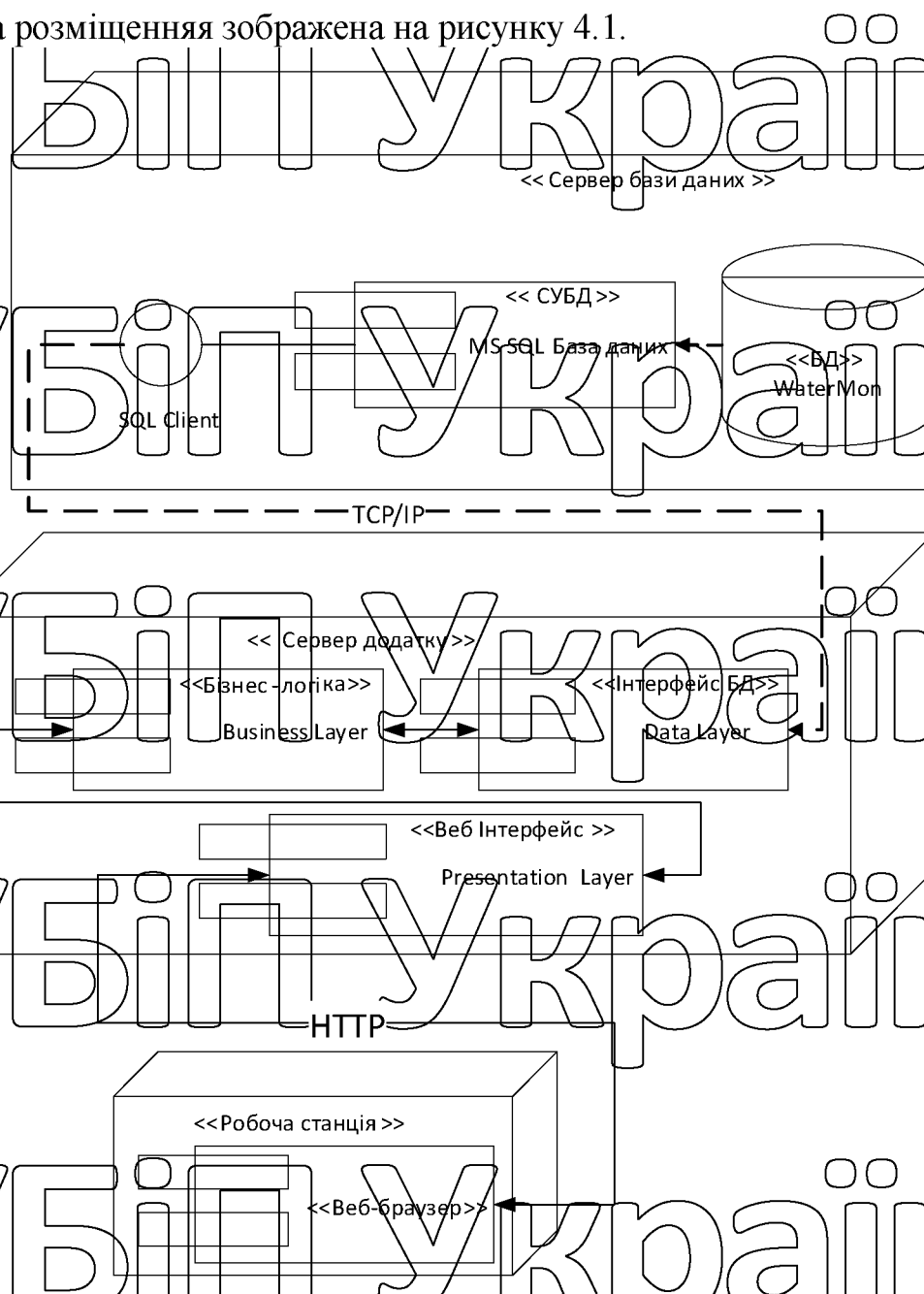


Рис. 4.1. Діаграма розгортання.

4.1.2 Вимоги до апаратного забезпечення. Для повноцінної роботи системи слід використовувати сервер баз даних.

Серверна частина

- процесор **intel pentium E2140**;
- ємність оперативної пам'яті – 8 Гігабайт;
- жорсткий диск місткістю 20 Гігабайт.

Клієнтська частина

- процесор **intel pentium E2140**;
- ємність оперативної пам'яті – 4 Гігабайт;
- жорсткий диск місткістю 5 Гігабайт.

4.1.3 Вимоги до програмного забезпечення.

Серверна частина

- **microsoft sql server 2012** і новіші;
- **windows 7** і новіші;
- **iis server express**.

Клієнтська частина

- веб-браузер;
- інтернет зв'язок;

4.2 Загальний вигляд системи

Одним з наших компонентів системи являється журнал вимірювань значень показників (рис. 4.2.).

Журнал представлений у вигляді таблиці, де відображено назву показників, територію на якій було проведено вимірювання показників та тип водного ресурсу.

Також зазначена дата проведення вимірювання показників та саме фактичне значення.

НУБІП України

Для адміністратора контролю показників передбачено можливість коригування показників та всіх відомостей про них. Також передбачено фільтр, пошук по територіях або конкретних показниках.

Пошук...

Журнал показників

Показати 10

Фільтр: Південний Ковш

Показник	Територія	Тип водного ресурсу	Дата і час	Значення
Кисень розчинений	м. Київ, Південний Ковш	Річка	8.4	2018/07/10
Нітрат-іони	м. Київ, Південний Ковш	Річка	2.7	2018/07/10
Сульфат-іони	м. Київ, Південний Ковш	Річка	10	2018/07/10
Фосфат-іони	м. Київ, Південний Ковш	Річка	0.35	2018/07/10
Хлорид-іони	м. Київ, Південний Ковш	Річка	14	2018/07/10
Кисень розчинений	м. Київ, Південний Ковш	Річка	9	2018/09/10
Нітрат-іони	м. Київ, Південний Ковш	Річка	1.4	2018/09/10
Сульфат-іони	м. Київ, Південний Ковш	Річка	47	2018/09/10
Фосфат-іони	м. Київ, Південний Ковш	Річка	0.45	2018/09/10
Хлорид-іони	м. Київ, Південний Ковш	Річка	18.2	2018/09/10

Показано 10 записів

Попередня Наступна

Рис. 4.2. Журнал Показників

НУБІП України

Для обробки занесених даних в журнал передбачено відбір за допомогою параметрів, що допоможуть відібрати потрібні значення для аналізу (Рис.4.3).

Параметри для аналізу розбито на три частини.

Блок відбору територій, для яких потрібно виконати аналіз (Рис.4.4).

Блок показників, значення яких будуть брати участь в розрахунках (Рис.4.5).

Період за який будуть відібрані значення показників (Рис.4.6).

Пошук...

Введення параметрів

Виберіть території

- м. Київ, Південний Ковш
- с.Козаровичі, Київське водосховище
- м.Вишгород, н/б Київської ГЕС, питний водозабір м.Київ
- м.Київ, гідропарк, техн в/з ТЕЦ-4
- водосховище, 500 м нижче БСА
- м. Українка, нижче міста, вище в/з водоводу Біла Церква-Умань
- м. Переяслав-Хмельницький, вище гирла р.Трубіж

Виберіть показники

- Кисень розчинений
- Датчик PH
- Амоній-іони
- Нітрат-іони
- Нітрит-іони
- Сульфат-іони
- Фосфат-іони
- Хлорид-іони

Період даних моніторингу

Початок:

Кінець:

Рис. 4.3 Введення параметрів формування звітної інформації

НУБІП України

Виберіть території

м. Київ, Південний Ковш

с.Козаровичі, Київське водосховище

м.Вишгород, н/б Київської ГЕС, питний водозабір м.Київ

м.Київ, гідропарк, техн в/з ТЕЦ-4

водосховище, 500 м нижче БСА

м. Українка, нижче міста, вище в/з водоводу Біла Церква-Умань

м. Переяслав-Хмельницький, вище гирла р.Трубіж

Рис. 4.4 Вибір територій

НУБІП України

Виберіть показники

Кисень розчинений

Датчик рН

Амоній-іони

Нітрат-іони

Нітрит-іони

Сульфат-іони

Фосфат-іони

Хлорид-іони

Рис. 4.5 Вибір показників для розрахунків.

НУБІП України

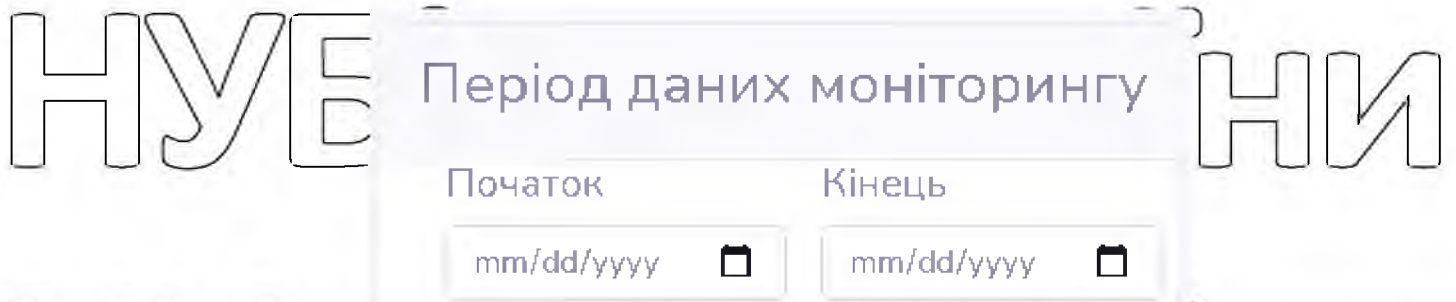


Рис. 4.6 Період даних моніторингу

4.3 Звітна інформація

Найважливішою частиною системи є блок звітної інформації. Саме за допомогою звітної інформації будуть прийматися рішення та проводитиметься аналіз виміюваних значень.

Побудований звіт Середні показники якості води територій за вказаний період (рис.4.7.) надає можливість оцінити рівень вмісту кожного показника на кожній території та за вказані роки.

Середні показники якості води по місяцям за рік (рис.4.8.) допомагає визначити, як змінюються значення показників зі зміною сезонів.

Зміна значень показників за період (рис.4.9.) даний звіт показує, як взагальному змінюється конкретний показник за період, на основі чого можна робити висновок по рівню забрудненості.

Території з найбільшою забрудненістю в розрізі показників за період (рис.4.10.) звіт по територіям показує, де найбільше забруднення кожного з показників були зафіксовані за вказаний період. Можна зробити висновок, що найбільші значення показник в було зафіксовано на території 2 показником РН

Середні показники якості води територій з 2019 по 2021 роки

Роки/Показники	Дослідна територія 1		Дослідна територія 2	
	Датчик DO, %	Датчик PH, %	Датчик DO, %	Датчик PH, %
2019	4.60	4.27	4.25	4.43
2020	4.60	4.18	4.74	5.89
2021	3.70	5.70	4.40	4.27

12/23/2020 3:38:36 PM

Рис.4.7. Середні показники якості води територій за вказаний період

Середні показники якості води по місяцям за 2021 рік

	Дослідна територія 1		Дослідна територія 2	
	Датчик DO, %	Датчик PH, %	Датчик DO, %	Датчик PH, %
January	3.41	5.63	4.33	4.33
February	3.98	5.77	4.46	4.21

12/23/2020 3:37:40 PM

Рис.4.8. Середні показники якості води по місяцям за рік

Зміна значень показників за період з 12/1/2020 по 2/3/2021

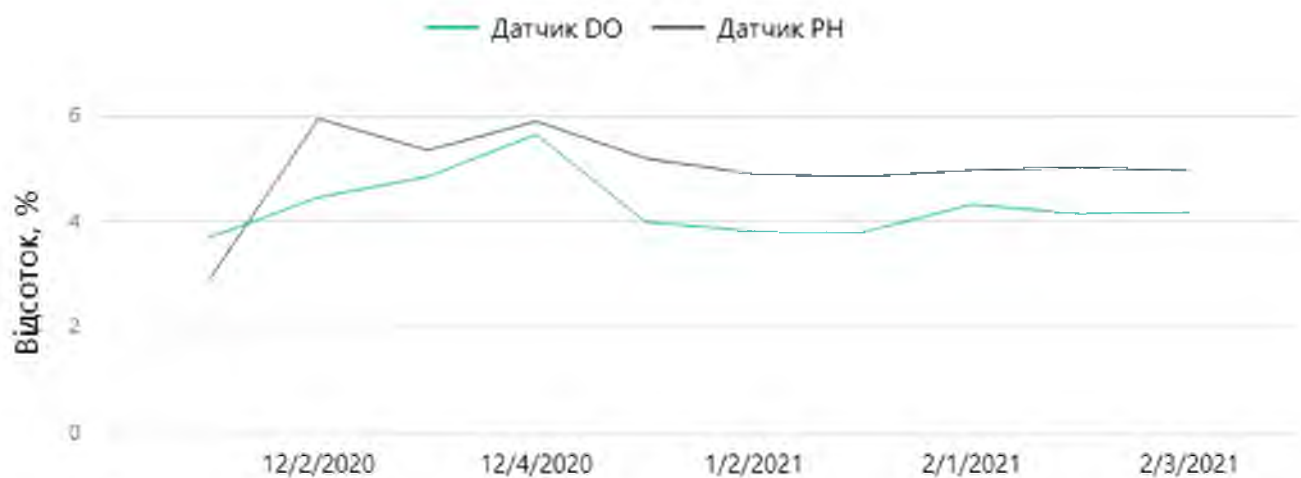


Рис.4.9. Зміна значень показників за період

НУБІП України

НУБІП України

Території з найбільшою забрудненістю в розрізі показників з 12/1/2020 по 2/3/2021



Рис.4.10. Території з найбільшою забрудненістю в розрізі показників за період

4.4 Розрахунок КРІ

Ключові показники ефективності (КРІ) - це критичні значення показників прогресу до запланованого результату.

Ключові показники ефективності, що забезпечать стратегічне та оперативне вдосконалення, створюють аналітичну базу, що в подальшому допоможе у прийнятті рішень та допоможе зосередити увагу на тому, що має найбільше значення.

Для побудови КРІ використано засоби для інтеграції та роботи в системі середовища MS EXCEL.

Побудова ключових показників відбувається програмно за допомогою шаблону EXCEL.

За створеним КРІ можемо побачити, на яких територіях рівень перевищує норму (4.11), за розрахунком можна зробити висновок, що для території 1 середнє значення показника на дату 12/1/2020 задовільняє норму.

Далі на всі наступні дати норму перевищено, на основі даної інформації можна зробити висновок, що території 1 потрібно прийняти міри для зменшення забрудень.

Також можна побачити, де були найбільші показники і чи вони задовільняють норму (4.12.).

Як можемо побачити за вказаний рік на всіх територія та по усіх показниках максимальні значення показників перевищили допустиму норму. Отже можна зробити висновок, що для кожна територія має проблеми з забрудненнями.

Рік	2021		...
	Максимальні показники	Норма показника	Статус
Дослідна територія 1	6.1	3.5	🔴
Датчик DO	4.3	3.5	🔴
Датчик PH	6.1	3.5	🔴
Дослідна територія 2	4.52	3.5	🔴
Датчик DO	4.52	3.5	🔴
Датчик PH	4.5	3.5	🔴
Максимальний показник	6.1	3.5	🔴

Рис.4.11 Відображення максимальних показників

Територія	12/1/2020		12/2/2020		12/3/2020		12/4/2020		1/1/2021		1/2/2021	
	Середні значення показників	Норма показника	Середні значення показників	Норма	Середні значення показників	Норма	Середні значення показників	Норма	Середні значення показників	Норма	Середні значення показників	Норма
Дослідна територія 1	3.475	3.5	4.365	3.5	4.3	3.5	5.4	3.5	4.8	3.5	4.3165	3.5
Датчик DO	3.45	3.5	4.33	3.5	4.1	3.5	6.5	3.5	3.5	3.5	3.333	3.5
Датчик PH	3.5	3.5	4.4	3.5	4.5	3.5	4.3	3.5	6.1	3.5	5.3	3.5
Дослідна територія 2	3.15	3.5	6.05	3.5	5.9	3.5	6.15	3.5	4.4	3.5	4.4	3.5
Датчик DO	3.55	3.5	4.6	3.5	5.6	3.5	4.8	3.5	4.5	3.5	4.3	3.5
Датчик PH	2.35	3.5	7.5	3.5	6.2	3.5	7.5	3.5	4.3	3.5	4.5	3.5
Середні показники	3.3125	3.5	5.2075	3.5	5.1	3.5	5.775	3.5	4.6	3.5	4.35825	3.5

Рис.4.12 Відображення середніх показників

4.5 Засоби Data Mining

Для дослідження та знаходження у даних прихованих знань, що раніше не були відомі, було використано методи та алгоритми технології Data Mining.

Одним з алгоритмів було обрано IR-алгоритм (Рис.4.13.). Значення було розподілено на два окремі класи:

H – High pollution (Висока забрудненість)

L – Low pollution (Низька забрудненість)

Залежна змінна – Якість води

Незалежні змінні – Середні значення показників в розрізі місяців

В роботі алгоритму було використано різні показники, їх було приведено до середньої норми якості.

В результаті роботи алгоритму можна побачити наскільки кожен з показників впливає на рівень забруднення.

Також бачимо, що найменші рівні забруднення 2.7 – 2.73 фактичної якості, були в літній період часу. А найбільші показники забруднення 2.84 – 2.85 спостерігалися в кінці зими та весною.

На основі чого можна зробити висновок, що саме відліга впливала на забруднення води.

Класифікація значень показників за 2020 рік

Місяць	I	Датчик PH	Мутність	Вміст заліза	Норма якості	Фактична якість	Клас
Березень	7.89		0.51	0.12	2.81	2.84	Висока забрудненість
Вересень	7.86		0.4	0.13	2.81	2.76	Висока забрудненість
Жовтень	7.88		0.41	0.14	2.81	2.83	Висока забрудненість
Квітень	7.9		0.46	0.12	2.81	2.82	Висока забрудненість
Липень	7.87		0.34	0.11	2.81	2.7	Низька забрудненість
Лютий	7.86		0.56	0.13	2.81	2.85	Висока забрудненість
Серпень	7.85		0.36	0.1	2.81	2.73	Низька забрудненість
Січень	7.71		0.58	0.14	2.81	2.81	Висока забрудненість
Травень	7.88		0.42	0.1	2.81	2.80	Низька забрудненість
Червень	7.87		0.35	0.1	2.81	2.78	Низька забрудненість

Рис.4.13. Класифікація значень показників

На рисунку 4.14 показано класифікацію значень показників, з використанням алгоритму, який будується на основі формули Bayes.

Саме цей метод допоможе правильно класифікувати значення, оскільки за допомогою цього методу можна розглянути декілька незалежних змінних.

За знайденою умовною ймовірністю можна побачити відсоток належності кожного з показників до класу (Рис 4.15).

В результаті можемо спостерігати відсоток ймовірності прогнозування входу до класу в цілому для визначення якості води.

Різниця в відсотках ймовірності обумовлена не достатньою кількістю навчальної вибірки

Класифікація значень показників 2020 рік

Місяць	Датчик pH	Мутність	Вільні азоти	Норма кислі	Шкідлива кислі
Березень	7,89	0,51	0,12	2,81	2,84
Вересень	7,86	0,4	0,13	2,81	2,79
Жовтень	7,88	0,41	0,14	2,81	2,83
Квітень	7,9	0,46	0,12	2,81	2,82
Липень	7,87	0,34	0,11	2,81	2,7
Лютий	7,86	0,56	0,13	2,81	2,85
Серпень	7,85	0,36	0,1	2,81	2,73
Січень	7,71	0,58	0,14	2,81	2,81
Травень	7,86	0,48	0,1	2,81	2,85
Червень	7,87	0,39	0,1	2,81	2,76

Рис.4.14. Класифікація значень з використанням формули Bayes

Метод	↑	Клас розподілу	Значення		
Середні значення показників		Низька забрудненість	7,84	0,45	0,12
		Висока забрудненість	7,88	0,54	0,13
Варіації показників		Низька забрудненість	0,0076	0,0092	0,0003
		Висока забрудненість	0,0004	0,0012	0,00005
Умовна ймовірність %		Низька забрудненість	0,045	0,039	0,20
		Висока забрудненість	0,093	0,033	0,43
Безумовні ймовірності класу %		Низька забрудненість	66,66		
		Висока забрудненість	33,33		
Ймовірність прогнозування %		Низька забрудненість	34,28		
		Висока забрудненість	65,71		

Рис. 4.15. Ймовірність входу до класу

НУБІП України

ВИСНОВКИ

Провівши аналіз предметної області, було встановлено, які показники являються найбільш задіяними для визначення якості води.

На основі цих показників були побудовані алгоритми роботи системи.

Також провівши аналіз існуючих рішень, дослідивши всі подібні ресурси, було виявлено їх недоліки та переваги.

На основі зібраної інформації сформовано постановку завдання. Це дозволить досліджувальній системі мати перевагу перед вже існуючими засобами та системами.

Для розуміння роботи всієї системи було побудовано діаграми, що дозволять більш глибоко розуміти суть кожного процесу.

НУБІП України

В спроектованій архітектурі системи було розміщено оперативну базу даних та сховище даних, що дозволить обробляти великі обсяги даних в розрізі часу.

Інформаційна система розроблена на мові програмування `C#` з використанням технології ASP.NET MVC.

За допомогою використання сучасних технологій розробки веб додатків інформацію представлено в зручному для користувача вигляді.

Використання клієнт-серверної архітектури дало змогу розвантажити клієнтську частину, так як обробка даних відбувалась повністю на стороні серверу.

За допомогою створених звітів та KPI можна відповісти на поставлені питання.

Які території мали найбільшу забрудненість за рік?

Середні показники якості води в розрізі сезонів.

Визначення якості води в розрізі показників забрудненості.

За допомогою звіту Території з найбільшою забрудненістю в розрізі показників за період можна прийняти рішення, де найбільші забруднення кожного з показників були зафіксовані та прийняти заходи для запобігання забрудненню в подальшому.

Зміна значень показників за період показує, як взагалі змінюється конкретний показник за період, на основі чого можна визначити в який період часу показник перевищує норму та прийняти рішення для запобігання збільшення в майбутньому.

З використанням методів Data mining було виконано класифікацію значень показників.

З використання алгоритму 1Rule було виявлено наскільки кожен з показників впливає на рівень забруднення.

Також знайдені найменші рівні забруднення 2.7 – 2.73 фактичної якості, були в літній період часу. А найбільші показники забруднення 2.84 – 2.85 спостерігалися в кінці зими та весною.

На основі чого можна зробити висновок, що саме відлига впливала на забруднення води в період за 2020 рік.

НУБІП України

З використанням алгоритму, який будується на основі формули Bayes. Було класифіковано значення, оскільки за допомогою цього методу можна розглянути декілька незалежних змінних.

НУБІП України

За знайденою умовною ймовірністю можна побачити відсоток належності кожного з показників до класу.

В результаті можемо спостерігати відсоток ймовірності прогнозування входу до класу в цілому для визначення якості води.

НУБІП України

Різниця в відсотках ймовірності обумовлена не достатньою кількістю навчальної вибірки

Отже, для більш точного аналізу потрібна більша кількість вимірюваних значень показників.

НУБІП України

Провівши аналіз всіх застосованих методів та технологій в системі, можна зробити висновок, що дана система може використовуватись, як окреми повноцінний продукт на ринку, що досить вдало зможе конкурувати з іншими системами.

З використанням технологій Data mining, робить систему інкальною в своєму роді.

НУБІП України

Тому використання такої системи дасть змогу автоматизувати процеси аналізу та обробки показників якості води

НУБІП України

НУБІП України

НУБІП України

НУБІП України

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. W3Schools [Електронний ресурс] Режим доступу: <https://www.w3schools.com/> (дата звернення 05.10.2021) – Назва з екрана

2. TRUMPXCEL [Електронний ресурс] Режим доступу: <https://trumpexcel.com/kpi-dashboard-in-excel-part-1/> (дата звернення 19.10.2021) –

Назва з екрана

3. OLAP [Електронний ресурс] Режим доступу: <https://olap.com/learn/bi-olap/olap-bi-definitions/business-intelligence/> (дата звернення 25.10.2021) – Назва з

екрана

4. Б.Л.Голуб, Є.М.Шукайло. Методичний посібник до вивчення дисципліни "Програмування та алгоритмічні мови". Методичний посібник. – Видавничий центр НАУ, 2003 – 64 с.

5. Козлов В.А. Открытые информационные системы. –М.:Финансы и статистика, 1999.

6. Липаев В.В. Системное проектирование сложных программных средств для информационных систем. –М.:Синтег, 1999

7. Перминов Г.И. Хранилища данных. Интернет-университет информационных технологий - ИНТУИТ.ру, 2009

8. Барабда і Данве, 2015 М. Н. Барабде , С. Р. Данве Система постійного моніторингу якості води для водних ресурсів у віддалених місцях IJRCEE , 3(2015), стор. 2320 - 9798

9. Чень і Хань, 2018. Іхен Чен, Давей Хань **Моніторинг якості води в розумному місті: пілотний проект** (2018), с. 307-316

10. Монарті та ін., 2018. Нагесвара Рао Монарті, Ч.Мужеш, П. Відья Сагар

Система моніторингу якості води за допомогою IoT

11. 4-та Міжнародна конференція з досягнень в галузі електрики, електроніки, інформації, комунікацій та біоінформатики (2018)

12. Омар Фарук та ін., 2017. MD Омар Фарук, Інджамамул Хоке Ему, MD. Назмул Хакел, Майтрі Дей, Н.К. Дас, Мрінмой Дей

13. **Розробка та впровадження економічно ефективної системи оцінки якості води** Регіон 10 конференції гуманітарних технологій IEEE, Дакка, Бангладеш (2017), стор. 860 - 863

14. Прасад та ін., 2015. А. Н. Прасад, К. А. Мамун, Ф. Р. Іслам, Х. Хаква

Розумна система моніторингу якості води Університет південної частини

Тихого океану. 2-й Азіатсько-Тихоокеанський Всесвітній конгрес з комп'ютерних наук та інженерії Конференція IEEE (2015)

15. Шафі та ін., 2018

Уфера Шафі, Рафія Мумтаз, Хірра Анвар, Алі Мустафа

Камар, Хамза Хуршид Виявлення забруднення поверхневих вод за допомогою Інтернету речей Школа електротехніки та обчислювальної техніки,

Національний університет науки і технологій, Конференція IEEE (2018), с. 92 - 96

16. Сугарія та ін., 2018

Tx Sugarriyaa, C. Rakshaya, K. Ramyadevi, M. Ramya, П. Г. Раші

17. **Розумна система моніторингу якості води для застосувань у режимі реального часу** Int. J. Pure Appl. Математика, 118 (2018), с. 1363 - 1369

18. Бейкер К. Водна безпека: виклики та можливості дослідження. Наука. 2012; 337: 914-5.

19. Поверховий М.В., ван дер Гаг Б, Бернс Б.П. Досягнення в онлайнному моніторингу якості питної води та системах раннього попередження. Вода Res. 2011; 45: 741-7.

20. Chung WY, Yoo JH. Віддалений моніторинг якості води на широкій території. Привід датчика В Хім. 2015; 217: 51–7.

21. Методическое пособие по дисциплине «Экологический мониторинг»: Учебное пособие/ Под ред. Г.И. Хараева. – Улан-Удэ: Изд-во ВСГТУ, 2004. – 77

с.
22. Руководство по программированию на C# [Электронный ресурс] Режим доступа: <https://msdn.microsoft.com> (дата звернення 15.04.2021) – Назва з екрана.

23. ASP Tutorial [Электронный ресурс] Режим доступа: <http://www.w3schools.com/sql/> (дата звернення 21.04.2021) – Назва з екрана

24. Стоян Стефанов JavaScript Шаблоны. 2011. 734 с.

25. ASP.NET MVC 3 Framework Адам Фримен, Стивен Сандерсон 2012 р. 672

с.
26. Вольфсон Б. /Гибкие методологии разработки. Версия 1.2. /Вольфсон Борис. 2009р. 112стр.

27. Хендерсон .Професійний посібник з SQL Server. Структура й реалізація/ Кен Хендерсон.

28. Технология разработки программных продуктов : учеб. Пособие для студ. проф. образования / А.В.Рудаков. 2-е изд., стер. – М.: Издательский центр «Академия», 2006. – 208 с.

29. Карвин Б.К 21 Профаммирование баз данных SQL. Типичные ошибки и ихустранение / Б. Карвин. — М.: Рид Групп, 2012. —336 с. — (Профессиональные компьютерные книги)

30. Томає Ерл Основи Big Data: концепції, алгоритми і технології ISBN: 978-0-13-429107-9 2018р. 320 с.

31. Алексеева М. Б., Балан С. Н. Основи теорії систем і системного аналізу.

32. Володимир Боровик, Володимир Гамаюн. Автоматизоване робоче місце. Проектування інформаційних систем і баз даних. Київ.«НАУ-друк».2010.128 с.

33. Об'єктно-орієнтоване проектування ІС Режим доступа: <https://studopedia.org/9-154040.html>

34. Діаграма класів – Режим доступу: https://life-prog.ru/ukr/1_4646_priznachennya-diagrami-klasiv.html

35. Діаграма компонентів – Режим доступу: <https://studfiles.net/preview/381095/page:41/>

36. Амортизаційні відрахування та методи їх розрахунку – Режим доступу: <http://ua.texreferat.com/referat-2655-3.html>

37. Use case diagrams – Режим доступу: <http://lvivqaclub.blogspot.com/2008/10/use-case-uml-diagram.html>

38. Діаграма розгортання – Режим доступу: https://studopedia.com.ua/1_280679_diagrama-rozgortannya-deployment-diagram.html

39. А. А. Барсегян Технології аналізу даних: Data Mining, OLAP 2 видавництво.

40. Інформаційні системи і технології на підприємствах – Плєскач В. Л. – OLAP системи – Режим доступу: <https://westudents.com.ua/glavy/27272-OLAP-sistemi.html>

41. OneR Algorithm Режим доступу: <https://www.saedsayad.com/oner.htm>

42. Microsoft Power BI: Data Visualization режим доступу: <https://powerbi.microsoft.com/en-us/>

43. Alberto Ferrari and Marco Russo. Analyzing Data with Power BI and Power Pivot for Excel. 2017 p. 434 c.

44. Reza RadPro. Power BI Architecture: Sharing, Security, and Deployment Options for Microsoft Power BI Solutions. 2018p. 544c.

45. Luis Torgo. Data Mining with R: Learning with Case Studies, Second Edition. 2010p. 426c.

46. Matthew A. Russell and Mikhail Klassen. Mining the Social Web: Data Mining Facebook, Twitter, LinkedIn, Instagram, GitHub, and More. 2018p. 438c.

47. Priya Pedamkar Data warehouse & data mining 2009p.

48. Margaret H. Dunham Data mining introductory and advanced topics. 2002p.

49. Russell K. Anderson. Visual Data Mining: The VisMiner Approach. 2012p.
202с.

50. Mike Barlow. Real-Time Big Data Analytics: Emerging Architecture. 2013p.
520с.

51. Марченко О.О., Россада Т.В. Актуальні проблеми Data Mining 2017р. 150с.

НУБІП України

НУБІП України

НУБІП України

НУБІП України

НУБІП України