

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ЛЬВІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ВЕТЕРИНАРНОЇ
МЕДИЦИНИ ТА БІОТЕХНОЛОГІЙ ІМ. С.З. ГЖИЦЬКОГО
ФАКУЛЬТЕТ МЕХАНІКИ, ЕНЕРГЕТИКИ ТА ІНФОРМАЦІЙНИХ
ТЕХНОЛОГІЙ
КАФЕДРА ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ**

КВАЛІФІКАЦІЙНА РОБОТА

другого (магістерського) рівня вищої освіти

на тему: **«Використання методів обчислювального інтелекту для
оцінки якості води в озерах»**

Виконав: студент групи Іт-62

Спеціальності 126 «Інформаційні системи та
технології»

(шифр і назва)

Оліфір Роман Юрійович

(Прізвище та ініціали)

Керівник: к.т.н., доцент Татомир А.В.

(Прізвище та ініціали)

Рецензент: к.т.н., доцент Шарибура А.О.

(Прізвище та ініціали)

ДУБЛЯНИ-2025

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ЛЬВІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ПРИРОДОКОРИСТУВАННЯ
ФАКУЛЬТЕТ МЕХАНІКИ, ЕНЕРГЕТИКИ ТА ІНФОРМАЦІЙНИХ
ТЕХНОЛОГІЙ
КАФЕДРА ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

Другий (магістерський) рівень вищої освіти
Спеціальність 126 «Інформаційні системи та технології»

«ЗАТВЕРДЖУЮ»

Завідувач кафедри _____

д.т.н., проф. А.М. Тригуба

«_____» _____ 2025 р.

ЗАВДАННЯ

на кваліфікаційну роботу студенту

Оліфіру Роману Юрійовичу

1. Тема роботи: «Використання методів обчислювального інтелекту для оцінки якості води в озерах»

Керівник роботи Татомир Андрій Володимирович, доцент
затверджені наказом по університету від 28.02.2025 року № 140/к-с.

2. Строк подання студентом роботи 10.12.2025 р.

3. Вихідні дані до роботи: дані про гідрохімічні показники якості води, дані метеорологічних спостережень та супутникові індикатори стану озер. Підготовлені навчальні та тестові вибірки на основі датасету Water Potability Dataset. Архітектури моделей обчислювального інтелекту – CNN, LSTM, GRU, Autoencoder+MLP, Deep Active Learning (MLP) та DAG-EU (evidential network).

4. Зміст розрахунково-пояснювальної записки (перелік питань, які необхідно розробити) _____
Вступ.
Аналіз стану питання в теорії та практиці.
Теоретичні основи застосування методів обчислювального інтелекту для оцінення якості води та вибір засобів.
Розроблення та тестування інтелектуальної моделі для оцінення якості води в озерах.
Охорона праці та безпека у надзвичайних ситуаціях.
Визначення економічної ефективності.
Висновки та пропозиції.
Список використаної літератури.

5. Перелік ілюстраційного матеріалу (з точним зазначенням обов'язкових слайдів): аналіз стану питання в теорії та практиці; теоретичні основи застосування методів обчислювального інтелекту для оцінення якості води та вибір засобів; розроблення та тестування інтелектуальної моделі для оцінення якості води в озерах; економічна ефективність.

6. Консультанти з розділів:

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
1, 2, 3, 5	<i>Татомир А.В., зав. кафедри інформаційних технологій</i>		
4	<i>Городецький І.М., доцент кафедри інженерної механіки</i>		

7. Дата видачі завдання

28 лютого 2025 р.

Календарний план

№ з/п	Назва етапів кваліфікаційної роботи	Терміни виконання етапів роботи	При-мітка
1	<i>Написання першого розділу</i>	28.02-20.03.25	
2	<i>Виконання другого розділу та аркушів ілюстраційного матеріалу до нього</i>	21.03-14.06.25	
3.	<i>Виконання третього розділу та аркушів ілюстраційного матеріалу до нього</i>	15.06.25-10.07.25	
4.	<i>Написання розділу «Охорона праці та безпека у надзвичайних ситуаціях»</i>	11.07-31.08.25	
5.	<i>Оцінення ефективності запропонованої системи</i>	01.09-31.10.25	
6.	<i>Завершення оформлення розрахунково-пояснювальної записки та аркушів ілюстраційного матеріалу</i>	01-30.11.25	
7.	<i>Завершення роботи в цілому</i>	01-10.12.25	

Студент _____ Оліфір Р. Ю.
(підпис)

Керівник роботи _____ Татомир А.В.
(підпис)

УДК 004.89:504.064.2

Використання методів обчислювального інтелекту для оцінки якості води в озерах.

Оліфір Р. Ю. Кафедра інформаційних технологій – Дубляни, ЛНУВМ, 2025.
Кваліфікаційна робота: 76 с. текст. част., 21 рис., 9 табл., 14 арк. ілюстраційного матеріалу, 37 джерел.

У кваліфікаційній роботі розроблено інтелектуальну систему підтримки прийняття рішень для оцінювання якості води в озерах на основі методів обчислювального інтелекту. Запропоновано підхід, що поєднує аналіз гідрохімічних, метеорологічних та супутникових даних із використанням моделей глибокого та активного навчання. Для підвищення точності оцінювання реалізовано порівняння декількох архітектур – CNN, LSTM, GRU, Autoencoder+MLP, Deep Active Learning (MLP) та спрощеної evidential моделі DAG-EU. Найвищу якість класифікації показала модель DAG-EU (accuracy = 0,666; AUC = 0,666), що враховує невизначеність даних за допомогою параметрів Діріхле. На основі цієї моделі створено веб-систему СППР, реалізовану у середовищі Streamlit, яка забезпечує введення параметрів, автоматичне масштабування, прогнозування якості води та формування візуалізованих рекомендацій для екологічного моніторингу. Отримані результати підтверджують ефективність застосування методів обчислювального інтелекту для оперативного оцінювання стану озерних вод і можуть бути використані органами місцевого самоврядування та екологічними службами при прийнятті управлінських рішень.

Ключові слова: обчислювальний інтелект, якість води, нейронні мережі, глибоке навчання, DAG-EU, CNN, система підтримки прийняття рішень, Streamlit, екологічний моніторинг, оцінювання стану озер.

ЗМІСТ

ВСТУП	7
Розділ 1. АНАЛІЗ СТАНУ ПИТАННЯ В ТЕОРІЇ ТА ПРАКТИЦІ	9
1.1. Актуальність і сутність проблеми оцінювання якості води в озерах	9
1.2. Стан наукових досліджень і практичних розробок із використанням методів обчислювального інтелекту	12
1.3. Міжнародний досвід досліджень з оцінювання якості озерної води	16
1.4. Доцільність виконання кваліфікаційної роботи та її задачі	20
Розділ 2. ТЕОРЕТИЧНІ ОСНОВИ ЗАСТОСУВАННЯ МЕТОДІВ ОБЧИСЛЮВАЛЬНОГО ІНТЕЛЕКТУ ДЛЯ ОЦІНЕННЯ ЯКОСТІ ВОДИ ТА ВИБІР ЗАСОБІВ	23
2.1. Поняття та принципи обчислювального інтелекту в екологічному моделюванні	23
2.2. Математичні моделі та алгоритми оцінювання якості води	27
2.3. Моделі глибокого та активного навчання в задачах екологічного прогнозування	30
2.4. Вибір програмних засобів і технологічних платформ реалізації	34
Розділ 3. РОЗРОБЛЕННЯ ТА ТЕСТУВАННЯ ІНТЕЛЕКТУАЛЬНОЇ МОДЕЛІ ДЛЯ ОЦІНЕННЯ ЯКОСТІ ВОДИ В ОЗЕРАХ	38
3.1. Вихідні дані та формування навчальної вибірки	38
3.2. Побудова базових моделей для порівняльного аналізу	42
3.3. Результати обґрунтування раціональної моделі оцінки якості води в озерах	49
3.4. Реалізація моделі в інформаційній системі оцінювання якості води	52
Розділ 4. ОХОРОНА ПРАЦІ ТА БЕЗПЕКА У НАДЗВИЧАЙНИХ СИТУАЦІЯХ	56
4.1. Аналіз умов праці під час розробки системи підтримки прийняття рішень	56
4.2. Організація системи охорони праці при розробці програмного забезпечення	58

4.3. Захист населення під час надзвичайних ситуацій	59
Розділ 5. ЕКОНОМІЧНА ЕФЕКТИВНІСТЬ ВІД ВИКОРИСТАННЯ ІНФОРМАЦІЙНОЇ СИСТЕМИ ОЦІНЮВАННЯ ЯКОСТІ ВОДИ	61
ВИСНОВКИ І ПРОПОЗИЦІЇ.....	64
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	68
ДОДАТКИ	72
Розділ А. Фрагмент коду СППР для оцінювання якості води в озерах	73

ВСТУП

Актуальність теми дослідження полягає у зростаючій потребі забезпечення екологічної безпеки водних ресурсів, які є одним із ключових чинників сталого розвитку територій. Якість води у природних озерах суттєво впливає на стан екосистем, здоров'я населення та ефективність використання водних ресурсів у сільському господарстві, рекреації та промисловості [17]. У зв'язку з кліматичними змінами, урбанізацією та антропогенним навантаженням традиційні методи контролю якості води стають недостатньо ефективними через обмежену репрезентативність вибірок і високу вартість польових досліджень. Саме тому все більшої актуальності набуває застосування методів обчислювального інтелекту, здатних аналізувати великі обсяги екологічних даних, виявляти приховані закономірності та формувати науково обґрунтовані висновки щодо стану водойм у реальному часі.

Сучасні інформаційні системи екологічного моніторингу дозволяють інтегрувати дані з різних джерел, зокрема сенсорних станцій, супутникових знімків, лабораторних вимірювань, для створення цілісної картини стану водних екосистем [19]. Методи машинного навчання, нечіткої логіки та нейронних мереж відкривають можливість для автоматизованої класифікації озер за рівнем забруднення, прогнозування змін у якості води та виявлення аномальних процесів. Використання цих підходів забезпечує високу точність оцінювання, зменшує вплив людського чинника та підвищує оперативність прийняття рішень у системах екологічного управління.

Мета дослідження полягає у розробленні та апробації методів оцінювання якості води в озерах на основі алгоритмів обчислювального інтелекту. Для досягнення цієї мети передбачається створення інформаційної системи, здатної аналізувати просторові та часові дані, здійснювати прогнозування стану водного середовища та підтримувати процес прийняття управлінських рішень у сфері екологічного моніторингу.

Об'єктом дослідження є процес оцінювання якості води у природних озерах за допомогою інформаційних технологій. Предметом дослідження виступають методи, алгоритми та інструментарій обчислювального інтелекту, що застосовуються для аналізу показників якості води, виявлення закономірностей та прогнозування змін екологічного стану водойм.

Наукова новизна роботи полягає у розробленні підходу до оцінювання якості води з використанням гібридних моделей обчислювального інтелекту. Це дозволяє враховувати як кількісні параметри водного середовища (температура, рН, вміст розчиненого кисню, концентрації сполук азоту та фосфору), так і якісні ознаки, що характеризують візуальні та біологічні властивості озер.

Практичне значення дослідження полягає у можливості впровадження розробленої системи у діяльність екологічних служб, науково-дослідних установ і місцевих громад для здійснення автоматизованого моніторингу водойм. Застосування розроблених методів сприятиме підвищенню ефективності управління природними ресурсами, зниженню ризиків деградації екосистем і своєчасному виявленню джерел забруднення.

Розділ 1.

АНАЛІЗ СТАНУ ПИТАННЯ В ТЕОРІЇ ТА ПРАКТИЦІ

1.1. Актуальність і сутність проблеми оцінювання якості води в озерах

Поверхневі водойми, зокрема озера, виконують у сучасному регіональному просторі надзвичайно важливу роль як джерела прісної води, резервуари біорізноманіття, рекреаційні зони та складові екосистемних процесів, що забезпечують сталий розвиток територій. Водночас вони перебувають під впливом численних антропогенних навантажень – сільськогосподарський стік, промислові викиди, урбанізація берегової зони – і змін клімату: підвищення температури, зміни режиму опадів, підвищена випаровуваність. Такі фактори призводять до погіршення якості води, зниження прозорості, росту концентрації поживних речовин і ризику розвитку шкідливих цвітінь водоростей [1].

Моніторинг якості води є можливістю отримання емпіричних даних, які стають основою для управлінських рішень у сфері охорони довкілля, забезпечення безпеки водних ресурсів і адаптації до змін. water.usgs.gov⁺¹ Однак у випадку озер існують специфічні виклики: нерівномірність просторового покриття вимірювань, великі часові проміжки між спостереженнями, необхідність задоволення різноманітних екосистемних, гідрологічних та гідрохімічних характеристик. Наприклад, характер озера як стоячої водойми (мінімальний тяг потоку, тривалий час перебування води) означає, що накопичення забруднювальних речовин і внутрішні процеси (стадії стратифікації, денітрифікація) можуть бути досить складними для моделювання.

Проблема посилюється тим, що традиційні методи збору даних (відбір проб, лабораторний аналіз, польові станції) часто є дорогими, трудомісткими й не гарантовано покривають весь простір водойми чи часову динаміку, яка може бути критичною. [SpringerOpen](https://www.springeropen.com/)⁺¹ Наприклад, у звітах відзначається, що кількість

хімічних сполук, які можуть впливати на якість води, значно перевищує кількість тих, що регулярно контролюються [25].

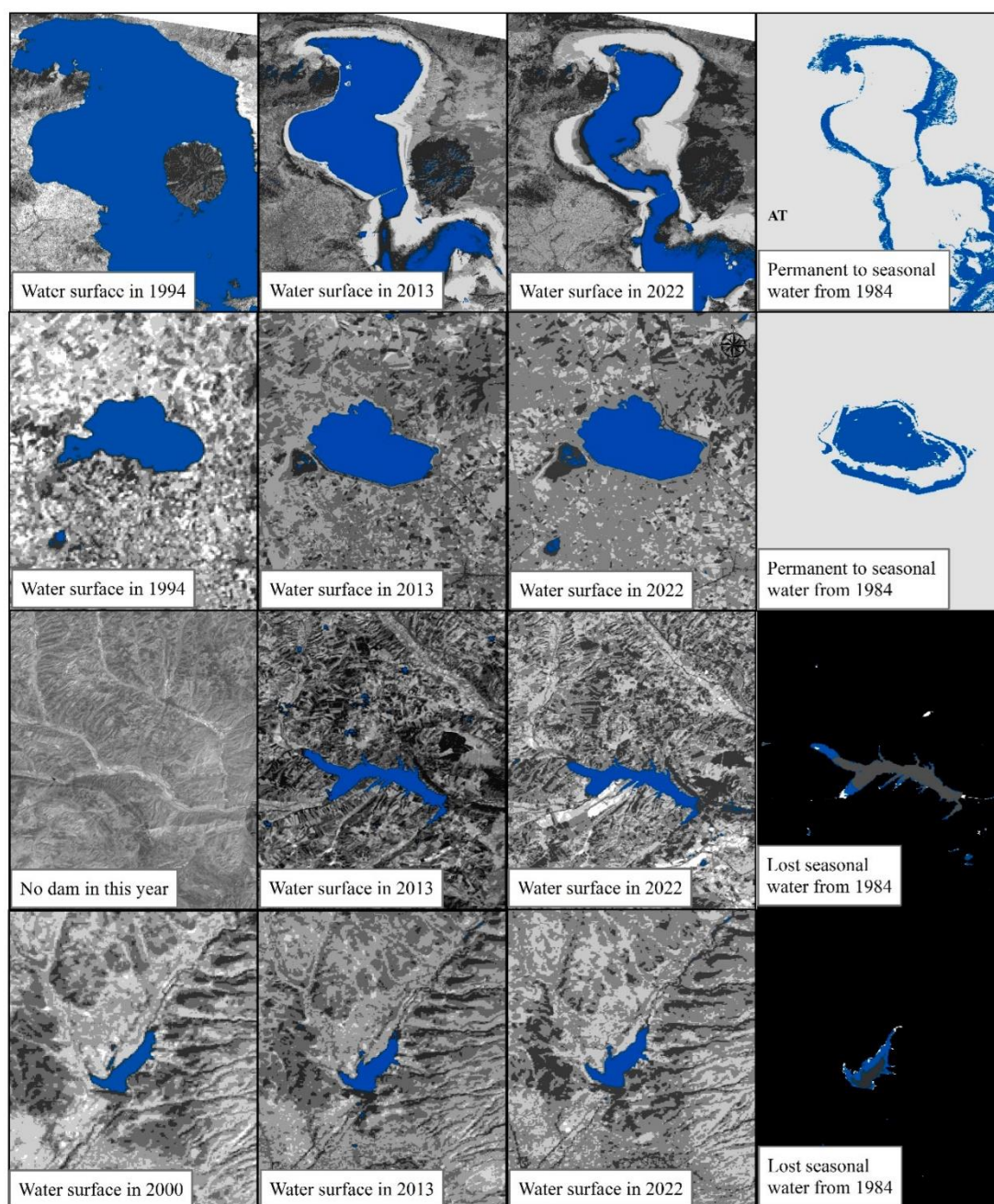


Рисунок 1.1 – Зміни поверхневих вод у різних ресурсах поверхневих вод з 1984 та 2000 по 2022 рік, отримані за знімками серій Landsat 5, 7 та 8 [1]

Таким чином, наукова та практична проблема полягає у підвищенні достовірності та оперативності оцінювання стану озерних вод через інтеграцію новітніх методів аналізу й обробки даних. Одним із перспективних напрямів є

застосування методів обчислювального інтелекту – машинного навчання, нейронних мереж, аналізу великих даних, комбінованого зі супутниковими, сенсорними та лабораторними даними – що дозволяє підвищити автоматизацію, покращити охоплення простору й часу, зменшити витрати на вимірювання та модернізувати систему екологічного моніторингу.

Для ілюстрації ризиків та динаміки стану водойм можна навести формулу інтегрованого індексу якості води (WQI) у спрощеному вигляді:

$$WQI = \sum_{i=1}^n w_i \cdot S_i, \quad (1.1)$$

де S_i – нормалізоване значення i -го показника (наприклад, розчинений кисень, рН, загальна фосфорна речовина); w_i – коефіцієнт вагомості показника, np – кількість показників.

При цьому вибір w_i та S_i значно впливає на кінцевий результат, що вимагає критичної оцінки та адаптації до умов конкретного озера.

У таблиці 1.1 наведено узагальнення основних типів чинників, що впливають на якість води в озерах.

Таблиця 1.1 – Основні фактори впливу на якість води в озерах

Категорія	Приклади факторів	Механізм впливу
Антропогенні	Сільськогосподарський стік, промислові викиди	Надходження поживних речовин, токсичних сполук
Кліматичні та гідрологічні	Підвищення температури, зміна опадів	Зменшення розчиненого кисню, посилення випаровування
Внутрішні процеси	Стратифікація, денітрифікація, ресуспензія	Активація внутрішніх джерел забруднення

З огляду на вище наведене, стає очевидним, що традиційний підхід до моніторингу озер часто не забезпечує достатньої чутливості, оперативності й просторової репрезентативності для адекватного оцінювання стану води. Це

створює нагальну потребу в автоматизованих, інтелектуально-обґрунтованих системах екологічного моніторингу, які могли б у режимі реального часу аналізувати мультидисциплінарні дані (фізико-хімічні, біологічні, просторово-часові) і надавати точні, інтерпретовані результати. У таких системах методи обчислювального інтелекту здатні запропонувати підходи для класифікації, прогнозування й виявлення аномалій, що суттєво підвищить ефективність управління водними ресурсами та захисту озерних екосистем.

Отже, підсумовуючи, проблема полягає в тому, що якість води в озерах – основного компонента природних і соціо-економічних систем – піддається багатьом впливам, а існуючі методи її оцінювання мають значні обмеження. Завданням дослідження є розробка та апробація підходу, який базується на методах обчислювального інтелекту для підвищення ефективності, точності та оперативної здатності оцінювання якості озерної води.

1.2. Стан наукових досліджень і практичних розробок із використанням методів обчислювального інтелекту

Оцінювання якості води є складною екологічною задачею, що вимагає урахування численних фізико-хімічних і біологічних параметрів, часто пов'язаних нелінійними залежностями. Класичні статистичні методи, такі як багатофакторна регресія, аналіз головних компонент або методи кореляційного аналізу, продовжують залишатися базовими інструментами для побудови емпіричних моделей. Проте ці підходи не завжди адекватно відображають складну структуру даних, що містять пропуски, нелінійності, корельовані показники й часові зрушення у спостереженнях. Саме тому протягом останнього десятиліття спостерігається зростання інтересу до методів обчислювального інтелекту – зокрема глибокого навчання, нейронних мереж, генетичних алгоритмів і нечітких систем, які довели свою ефективність у прогнозуванні та класифікації екологічних процесів [25].

На рисунку 1.1 наведено узагальнену схему еволюції підходів до моделювання якості води – від традиційних статистичних до гібридних моделей із використанням методів обчислювального інтелекту.



Рисунок 1.1 – Еволюція підходів до оцінювання якості води

У сучасних дослідженнях виділяють кілька основних напрямів застосування алгоритмів машинного навчання для оцінювання стану водойм (рис. 1.2).

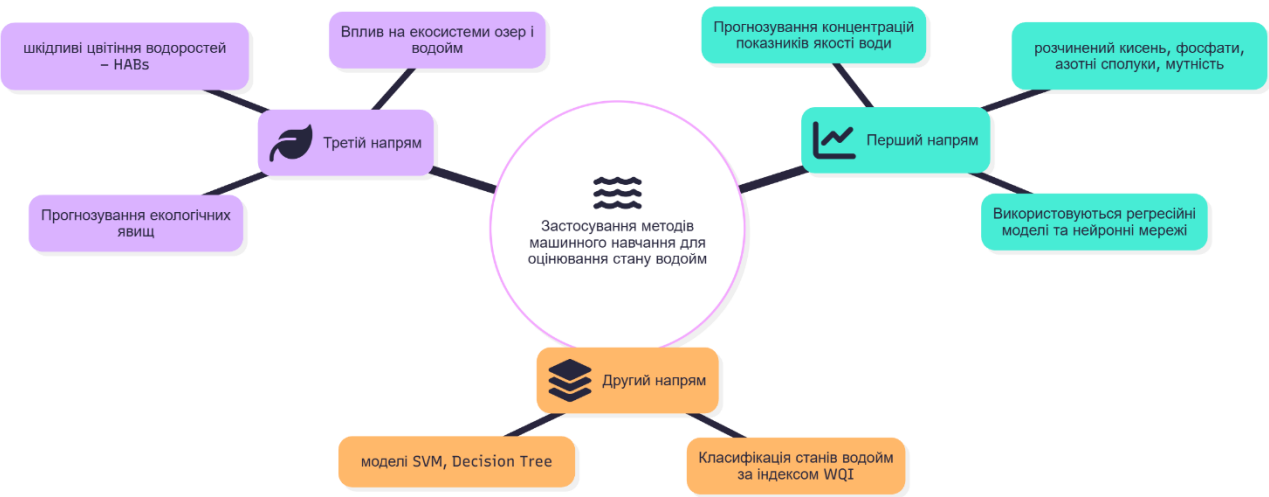


Рисунок 1.2 – Напрями застосування алгоритмів машинного навчання для оцінювання стану водойм

Перший напрям – прогнозування концентрацій окремих показників якості води (розчиненого кисню, фосфатів, азотних сполук, мутності) за допомогою регресійних моделей та нейронних мереж. Другий – класифікація станів водойм за індексом якості (WQI), наприклад, за допомогою моделей Support Vector Machine (SVM) або Decision Tree (DT). Третій напрям пов’язаний із прогнозуванням екологічних явищ, таких як шкідливі цвітіння водоростей (HABs), що мають суттєвий вплив на екосистеми озер і водойм.

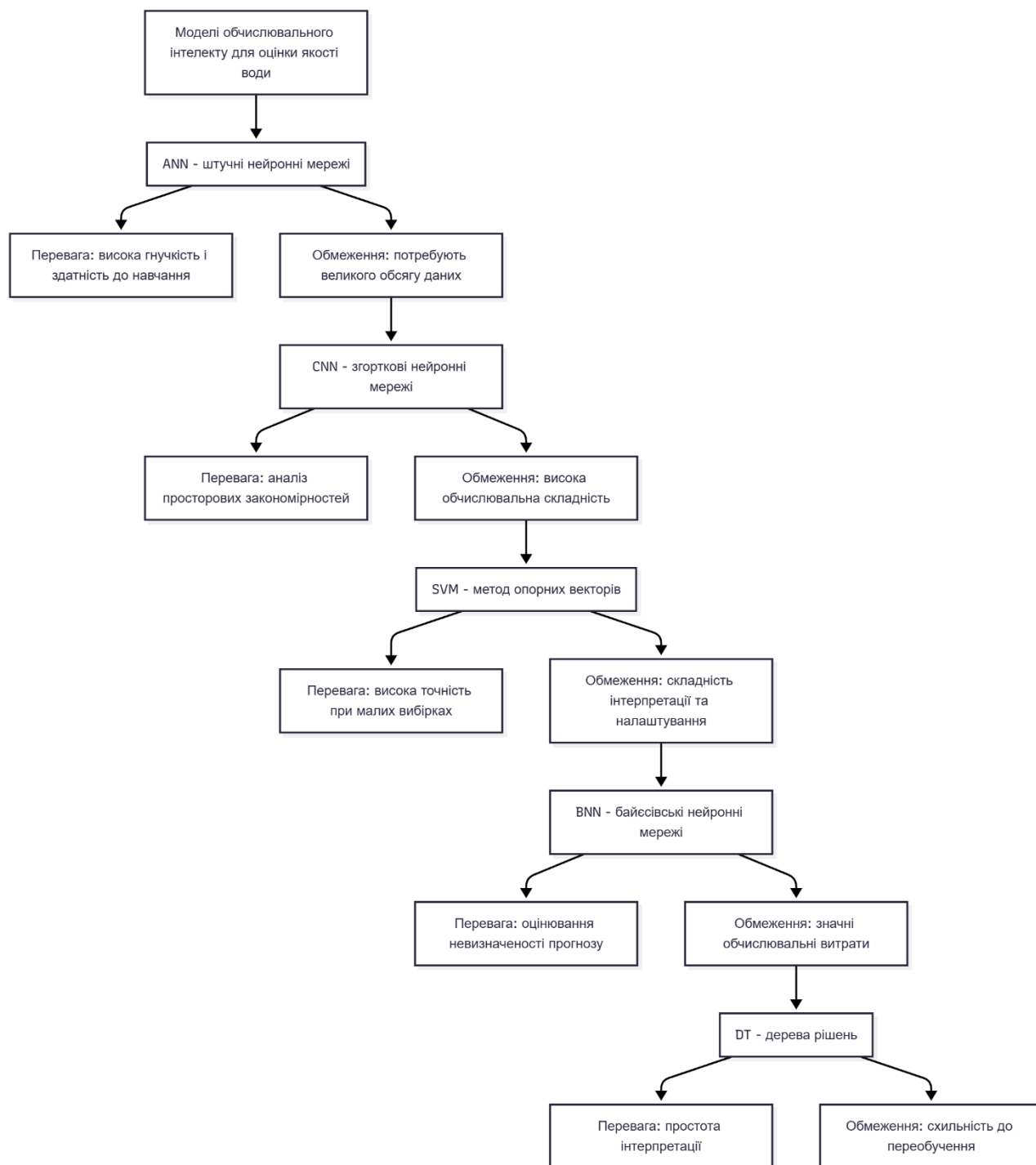


Рисунок 1.3 – Аналіз типових архітектур моделей, які застосовуються для аналізу екологічних даних

Загальний підхід до побудови моделі машинного навчання можна подати у вигляді формули:

$$\hat{y} = f(X, \Theta), \quad (1.2)$$

де $\hat{y}$ – прогнозований показник (наприклад, концентрація фосфору чи індекс WQI); $X = \{x_1, x_2, \dots, x_n\}$ – вектор вхідних параметрів; Θ – набір параметрів моделі, що визначають її навчання.

Ціль навчання полягає у мінімізації функції втрат:

$$L(\Theta) = \frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2, \quad (1.3)$$

де y_i – фактичне значення показника; N – кількість спостережень.

Сучасні дослідження демонструють, що штучні нейронні мережі (ANN) і згорткові нейронні мережі (CNN) забезпечують найвищу точність у порівнянні з класичними моделями при наявності достатнього обсягу даних [26; 24]. Наприклад, CNN дозволяють враховувати просторову структуру даних, що важливо при аналізі гіперспектральних супутникових знімків водної поверхні. Байєсівські нейронні мережі (BNN) використовуються для оцінювання невизначеності прогнозу, що має велике значення у завданнях моніторингу, де необхідно враховувати похибку сенсорів і природні коливання.

На рисунку 1.2 показано приклад типових архітектур моделей, які застосовуються для аналізу екологічних даних.

Рисунок 1.2 – Порівняння архітектур моделей для оцінки якості води

Модель	Принцип роботи	Основна перевага	Основне обмеження
ANN	Лінійна комбінація ознак із нелінійною активацією	Висока гнучкість	Потребує великої вибірки
CNN	Виявлення просторових закономірностей	Аналіз супутникових зображень	Висока обчислювальна складність
SVM	Визначення оптимальної гіперплощини	Висока точність при малих вибірках	Складна інтерпретація
BNN	Байєсівське оцінювання невизначеності	Надійність прогнозів	Високі ресурси обчислення
Decision Tree	Логічне розгалуження рішень	Простота інтерпретації	Схильність до переобучення

Поряд із цим активно розвиваються гібридні підходи, які поєднують різні типи моделей. Наприклад, методи Genetic Algorithm (GA) і Particle Swarm Optimization (PSO) застосовуються для автоматичного відбору релевантних ознак і оптимізації гіперпараметрів нейронних мереж. Такі інтегровані рішення дозволяють не лише підвищити точність прогнозу, але й зменшити надлишковість даних, що є критичним у системах екологічного моніторингу [19].

Наявні системи, розроблені на основі зазначених методів, уже довели свою ефективність для аналізу великих екологічних масивів, проте залишаються низка викликів. Основними недоліками є потреба у великих, збалансованих наборах даних, залежність від експертного налаштування моделі та висока чутливість до вхідних параметрів. З метою подолання цих обмежень останнім часом активно досліджується поєднання методів активного навчання (Active Learning) з евіденційною невизначеністю (Evidential Uncertainty), які дозволяють зменшити обсяг необхідних даних і водночас підвищити надійність оцінок [21].

Таким чином, аналіз сучасного стану наукових досліджень засвідчує, що найбільш перспективним напрямом у сфері оцінювання якості води є застосування гібридних моделей обчислювального інтелекту. Вони поєднують можливості глибоких нейронних мереж, еволюційних обчислень та методів оцінки невизначеності, що дозволяє створювати адаптивні, точні й надійні системи моніторингу стану водних екосистем.

1.3. Міжнародний досвід досліджень з оцінювання якості озерної води

Міжнародні дослідження оцінювання якості озерної води демонструють поступовий перехід від локальних емпіричних моделей до інтегрованих інтелектуальних систем, які поєднують дистанційне зондування Землі, дані *in situ* та методи обчислювального інтелекту. На прикладі басейну озера Урмія

(Іран) [12] показали, що поєднання кліматичних даних, супутникових знімків та інформації про антропогенне навантаження дозволяє побудувати просторово-часові моделі деградації водних ресурсів і виділити внесок кліматичних та господарських факторів у зниження рівня та якості води. Автори застосували регресійний аналіз та багатовимірні методи оцінювання, підкресливши, що достовірні висновки можливі лише за умови інтеграції різномірних джерел даних та використання адаптивних математичних моделей.

Подальший розвиток цього напрямку пов'язаний із застосуванням алгоритмів машинного навчання для побудови індексів якості води (WQI) та прогнозування окремих параметрів. Palabiyuk продемонстрував, що штучні нейронні мережі та, особливо, оптимізовані регресійні моделі можуть відтворювати значення інтегрального індексу якості води з дуже високою точністю, використовуючи показники прозорості, розчиненого кисню, біохімічного споживання кисню, вмісту азоту та фосфору. Для оцінювання використовувалися стандартні метрики, зокрема середньоквадратична похибка та коефіцієнт детермінації, що дозволило кількісно порівняти ефективність підходів [25]. У більшості сучасних робіт інтегральний підхід подається у вигляді функції:

$$WQI = \sum_{i=1}^n w_i q_i, \quad (1.4)$$

де q_i – нормалізовані субіндекси для окремих показників; w_i – вагові коефіцієнти, що відображають їхній екологічний пріоритет.

Визначення ваг і форми нормалізації стає окремим об'єктом оптимізації із використанням алгоритмів машинного навчання та еволюційних методів.

Знаковим етапом у розвитку інтелектуальних систем моніторингу поверхневих вод є поява гібридних моделей, що поєднують глибоке навчання, активне навчання та еволюційні обчислення. У дисертації Aranay [12] запропоновано фреймворк Deep Active Genetic Learning with Evidential Uncertainty (DAG-EU), у межах якого моделі глибинного навчання застосовано як до задач аналізу стану посівів, так і до оцінювання якості озерної води. Модель

використовує активне навчання для вибору найбільш інформативних зразків і генетичні алгоритми для відбору ознак, а також евіденційну невизначеність для оцінки довіри до прогнозу, що є критично важливим у контексті екологічних рішень [17]. Такий підхід демонструє, що міжнародні дослідження рухаються від «чорних скриньок» до структуровано обґрунтованих моделей, де інтерпретованість і кількісна оцінка невизначеності є невід’ємною частиною.

Окремий напрям міжнародного досвіду стосується використання методів обчислювального інтелекту для прогнозування та виявлення шкідливих цвітіння водоростей у великих озерах. У низці робіт поєднуються дані гіперспектральних супутникових сенсорів, температури поверхні, швидкості вітру та концентрації поживних речовин із моделями CNN, LSTM або їх гібридів для раннього виявлення зон потенційного цвітіння. Підхід базується на мінімізації функції втрат, яка включає як похибку прогнозу, так і штраф за хибно-негативні рішення, оскільки пропуск події цвітіння може мати значно серйозніші наслідки, ніж помилкове попередження. У новітніх роботах дедалі частіше використовуються елементи пояснювального ШІ, які дозволяють простежити вклад окремих ознак у рішення моделі та виявити критичні фактори деградації води.

Суттєвий внесок у систематизацію міжнародного досвіду зроблено в оглядових дослідженнях, присвячених застосуванню штучного інтелекту в моніторингу якості води. Frincu та співавт. узагальнили приклади використання нейронних мереж, SVM, дерев рішень, випадкових лісів та ансамблевих методів для побудови індексів якості, оцінювання забруднення та прогнозування динаміки показників для річок, озер і прибережних зон, відзначивши тренд до переходу від локальних моделей до масштабованих платформ, що працюють у реальному часі. IWA Publishing Паралельно з цим з’являються систематичні огляди й кейси з використання explainable AI у задачах оцінки якості води, які демонструють підвищення довіри до моделей завдяки можливості інтерпретувати важливість окремих індикаторів та структуру взаємозв’язків між ними [21].

Узагальнення міжнародного досвіду доцільно подати у вигляді порівняльної таблиці 1.3, яка відображає тип водного об’єкта, використані дані, методи обчислювального інтелекту та ключові результати.

Таблиця 1.3 – Аналіз міжнародних досліджень з використанням методів обчислювального інтелекту для оцінювання якості води

Дослідження	Об’єкт	Дані	Методи	Основний результат
Kazemi Garajeh et al., 2024	Озеро Урмія	Супутникові знімки, кліматичні та гідрологічні ряди	Регресійні та просторові моделі	Виділено внесок кліматичних і антропогенних чинників у деградацію ресурсів
Palabıyık, 2024	Поверхневі води	Фізико-хімічні показники	MLP, MLR	Досягнуто високої точності прогнозу WQI, показано придатність AI-моделей
Aranay, 2022	Озера та агросистеми	Супутникові й наземні дані	Deep Active Genetic Learning (DAG-EU)	Запропоновано гібридний підхід з активним навчанням та оцінкою невизначеності
Frincu et al., 2025	Різнi водойми	Різнорідні дані моніторингу	ANN, SVM, RF, ансамблі	Систематизовано практики використання AI для WQI та моніторингу
Kundu et al., 2025	Річкові системи	Показники якості води	Explainable AI	Продемонстровано можливість інтерпретованих прогнозів для підтримки рішень

Графічно міжнародний досвід можна представити у вигляді карти або схеми, де для кожного регіону позначаються тип водойм, основні індикатори та застосовані методи моделювання. Такий рисунок доцільно оформити як «геоінформаційне поле» інтелектуальних підходів до оцінювання якості води,

що підкреслює глобальний характер проблеми та універсальність методів обчислювального інтелекту.

Проведений аналіз показує, що міжнародні дослідження демонструють кілька чітких тенденцій:

- перехід від ізольованих гідрохімічних вимірювань до інтегрованих багатоджерельних систем спостереження; активне впровадження глибоких та гібридних моделей, здатних працювати з великими масивами даних;
- зростання уваги до оцінки невизначеності та пояснюваності результатів моделей; орієнтацію на створення практично придатних інформаційних систем для екологічних агентств та органів управління водними ресурсами.

Для даної кваліфікаційної роботи це формує методологічне підґрунтя для розроблення власної інтелектуальної моделі оцінювання якості води в озерах із урахуванням міжнародних підходів та адаптацією їх до умов локальних водойм.

1.4. Доцільність виконання кваліфікаційної роботи та її задачі

Доцільність виконання кваліфікаційної роботи зумовлена потребою впровадження сучасних інтелектуальних методів для аналізу стану водних екосистем, які є важливим елементом природного середовища та забезпечують життєдіяльність людини, сільського господарства і промисловості. В умовах зростаючого антропогенного навантаження та кліматичних змін контроль якості озерної води набуває стратегічного значення для забезпечення екологічної безпеки регіонів і виконання міжнародних зобов'язань України у сфері охорони довкілля. Сучасні системи моніторингу, які ґрунтуються лише на фізико-хімічних вимірюваннях, не завжди дають змогу отримати повну картину екологічного стану водойм, адже вимагають значних ресурсних витрат і не забезпечують оперативності реагування. Саме тому застосування методів

обчислювального інтелекту, здатних аналізувати великі обсяги гетерогенних даних, є доцільним і актуальним напрямом наукових досліджень.

Практична значущість дослідження полягає в тому, що розроблена модель на основі обчислювального інтелекту буде використана як складова частина системи підтримки прийняття рішень для екологічних служб, управлінь водного господарства та органів місцевого самоврядування. Вона дозволить автоматизувати процес оцінювання якості води, здійснювати прогнозування змін показників і визначати потенційні ризики забруднення. Очікуваним результатом є підвищення ефективності екологічного моніторингу, зменшення витрат на польові дослідження та створення умов для розроблення стратегій сталого управління водними ресурсами.

Загальна мета роботи полягає у створенні інтелектуальної інформаційної системи, що забезпечує достовірне оцінювання якості води в озерах на основі аналізу супутникових, сенсорних та лабораторних даних із застосуванням методів машинного й глибокого навчання.

Для досягнення поставленої мети необхідно розв'язати такі науково-практичні задачі:

- провести аналіз існуючих методів оцінювання якості води та визначити їхні переваги й недоліки в умовах різних типів водойм;
- здійснити порівняльне дослідження сучасних архітектур моделей обчислювального інтелекту (SVM, ANN, CNN, BNN, Decision Tree) для задач екологічного моніторингу;
- сформувати та підготувати вибірку даних, що поєднує гідрохімічні показники, кліматичні параметри та супутникові спостереження;
- розробити й навчити гібридну модель оцінювання якості води з використанням методів глибокого навчання та еволюційних алгоритмів;
- провести тестування моделі, оцінити її точність і стійкість до шумів, порівнявши з існуючими підходами;
- створити програмний прототип інформаційної системи з інтерфейсом для аналізу та візуалізації результатів оцінювання якості води в озерах.

Таким чином, виконання цієї кваліфікаційної роботи є доцільним з наукової, практичної та екологічної точок зору. Її результати буде використано у сфері цифрового природокористування, моніторингу довкілля, розроблення регіональних стратегій з управління водними ресурсами.

РОЗДІЛ 2.

ТЕОРЕТИЧНІ ОСНОВИ ЗАСТОСУВАННЯ МЕТОДІВ ОБЧИСЛЮВАЛЬНОГО ІНТЕЛЕКТУ ДЛЯ ОЦІНЕННЯ ЯКОСТІ ВОДИ ТА ВИБІР ЗАСОБІВ

2.1. Поняття та принципи обчислювального інтелекту в екологічному моделюванні

Сучасний розвиток екологічних досліджень характеризується швидким зростанням обсягів спостережуваних даних, що надходять із сенсорних мереж, супутникових систем, безпілотників та лабораторних джерел. Традиційні методи статистичного аналізу виявляються недостатніми для опису складних, нелінійних процесів, які характеризують водні екосистеми. У зв'язку з цим дедалі більшого поширення набуває застосування методів обчислювального інтелекту (Computational Intelligence, CI), що поєднують елементи штучного інтелекту, машинного навчання, нечіткої логіки та еволюційних алгоритмів. В основі цього підходу лежить здатність системи навчатися на прикладах, узагальнювати досвід і формувати адаптивні моделі для прогнозування стану навколишнього середовища без жорсткого математичного формалізму [1; 16].

Поняття обчислювального інтелекту охоплює сукупність методів, які здатні працювати з неповними, нечіткими або суперечливими даними, формуючи рішення на основі принципу самоорганізації. Класичне визначення CI, запропоноване [1], передбачає, що він є галуззю штучного інтелекту, яка вивчає адаптивні механізми, що виявляють ознаки розумної поведінки у складних середовищах. До таких механізмів належать нейронні мережі (*NN*), нечіткі системи (*Fuzzy Systems*), еволюційні алгоритми (*EA*), рійові методи (*Swarm Intelligence*) та імунні системи (*AIS*).

Загальна структура обчислювальної моделі для екологічних задач подається у вигляді функції:

$$y = f(x_1, x_2, \dots, x_n; \theta), \quad (2.1)$$

де x_i – вхідні ознаки, що описують параметри навколишнього середовища (наприклад, температура води, концентрація кисню, показник рН, метеорологічні умови); y – цільовий показник якості (наприклад, індекс WQI або ступінь евтрофікації); θ – вектор параметрів моделі, що налаштовується в процесі навчання. Основна мета – побудова функції $f(\cdot)$, яка забезпечує мінімальну похибку прогнозу, тобто:

$$\min_{\theta} L = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2, \quad (2.2)$$

де y_i – реальні значення показників; $\hat{y}_i$ – прогнозовані значення показників; N – кількість спостережень.



Рисунок 2.1 – Загальна схема застосування методів обчислювального інтелекту в екологічному моделюванні

У сфері екологічного моделювання обчислювальний інтелект застосовується для прогнозування динаміки якості води, класифікації

екологічних станів, аналізу взаємозв'язків між фізико-хімічними параметрами та оцінювання впливу антропогенних факторів. Наприклад, при моделюванні процесів самоочищення водойми нейронні мережі дають змогу враховувати нелінійні залежності між концентрацією забруднювачів, температурою, гідродинамічними умовами та фотосинтетичною активністю.

Основні принципи застосування методів обчислювального інтелекту в екологічному моделюванні можна узагальнити у вигляді системної схеми (рис. 2.1).

Таблиця 2.1 – Використання методів обчислювального інтелекту в екологічному аналізі

Метод	Характеристика	Приклади застосування	Переваги	Обмеження
Штучні нейронні мережі (ANN)	Нелінійне узагальнення залежностей між параметрами	Прогнозування WQI, розчиненого кисню, концентрації хлорофілу-а	Висока точність, здатність до самоорганізації	Потребують великих вибірок
Метод опорних векторів (SVM)	Побудова гіперплощини для розділення класів	Класифікація стану водойм за індексом WQI	Висока стабільність при малих вибірках	Складність параметризації
Нечітка логіка (Fuzzy Systems)	Використання правил типу “якщо–то”	Оцінка екологічного ризику, прийняття рішень	Інтерпретованість, гнучкість	Висока залежність від експерта
Еволюційні алгоритми (GA)	Стохастична оптимізація параметрів моделей	Вибір найінформативніших ознак, оптимізація структури NN	Здатність уникати локальних мінімумів	Велика обчислювальна складність
Гібридні моделі (Neuro-Fuzzy, GA-BNN)	Поєднання різних підходів CI	Прогнозування складних процесів у водоймах	Підвищена точність, адаптивність	Складність реалізації

Основним етапом є попередня обробка даних, що включає фільтрацію шумів, заповнення пропусків та нормалізацію значень. Далі здійснюється відбір ознак за допомогою методів кореляційного аналізу, головних компонент (PCA)

або генетичних алгоритмів. Навчання моделі відбувається на підмножині даних з відомими еталонними показниками якості води, після чого результати тестуються на незалежній вибірці для перевірки узагальнювальної здатності.

У таблиці 2.1 подано основні напрями застосування методів обчислювального інтелекту в екологічному моделюванні.

Принципи обчислювального інтелекту передбачають не лише автоматичне навчання моделі, але й адаптацію до нових умов без повного перенавчання. Наприклад, у системах моніторингу якості води використовуються адаптивні нейромережі, що оновлюють свої вагові коефіцієнти при надходженні нових даних у реальному часі. Це особливо важливо для водойм, у яких параметри змінюються під впливом сезонних коливань або раптових антропогенних впливів.

Однією з переваг обчислювального інтелекту є здатність оцінювати невизначеність результатів. У байєсівських нейронних мережах, наприклад, прогноз представлено як розподіл ймовірностей, а не єдине значення. Такий підхід дозволяє обчислити довіру до результату:

$$P(y|X) = \int P(y|X, \Theta) P(\Theta|D) d\Theta, \quad (2.3)$$

де $P(y|X)$ – апостеріорна ймовірність отримання прогнозу для заданих даних X ; $P(\Theta|D)$ – розподіл параметрів моделі, отриманий із навчальної вибірки D .

Таким чином, обчислювальний інтелект є ефективним інструментом для вирішення задач екологічного моделювання, забезпечуючи поєднання гнучкості, адаптивності та інтерпретованості моделей. Його використання у сфері оцінювання якості води дозволяє формувати прогностичні системи нового покоління, які здатні працювати в умовах невизначеності, забезпечувати швидке оновлення даних та підвищувати достовірність екологічного моніторингу.

2.2. Математичні моделі та алгоритми оцінювання якості води

Математичне моделювання є основою для кількісної оцінки стану водних екосистем і побудови прогнозів змін якості води під впливом природних і антропогенних чинників. Сучасні підходи ґрунтуються на інтеграції фізико-хімічних, біологічних і гідрологічних показників у єдину аналітичну систему. У цьому контексті застосування методів обчислювального інтелекту дозволяє перейти від класичних емпіричних залежностей до моделей, здатних навчатися на даних, враховувати нелінійні зв'язки та адаптуватися до просторово-часової мінливості середовища.

Базовою математичною формою задачі оцінювання якості води є знаходження функції:

$$Q = f(X, \Theta), \quad (2.4)$$

де Q – інтегральний показник якості води (наприклад, Water Quality Index, WQI); $X = \{x_1, x_2, \dots, x_n\}$ – вектор параметрів водного середовища (температура, pH, розчинений кисень, концентрації нітратів і фосфатів, каламутність, індекси NDWI або NDTI із супутникових зображень); Θ – множина параметрів моделі.

Мета полягає у визначенні такої функції $f(\cdot)$, яка мінімізує середньоквадратичну похибку між реальними Q_i та прогнозованими ($\hat{Q}_i$) значеннями:

$$\min_{\Theta} L(\Theta) = \frac{1}{N} \sum_{i=1}^N (Q_i - \hat{Q}_i)^2, \quad (2.5)$$

Для оцінювання якості води використовують три основні підходи: емпіричний, регресійний та інтелектуальний. Емпіричні моделі будуються на основі апріорних залежностей між параметрами (наприклад, лінійна регресія або аналітичні формули для WQI), тоді як інтелектуальні підходи (нейронні мережі, SVM, дерева рішень) самостійно виявляють приховані закономірності у даних.

Однією з найвідоміших методик є розрахунок індексу якості води (WQI), який визначається як зважена сума субіндексів окремих показників:

$$WQI = \sum_{i=1}^n w_i \cdot q_i, \quad (2.6)$$

де w_i – вагові коефіцієнти, що відображають екологічну значущість кожного параметра; q_i – нормалізовані субіндекси, які перетворюють значення показників у шкалу 0–100.

Залежність $q_i = f(x_i)$ часто має нелінійний характер і визначається експертним шляхом або методом оптимізації.

Для порівняння ефективності алгоритмів використовують стандартні метрики, наведені у таблиці 2.2.

Таблиця 2.2 – Основні показники оцінювання точності моделей якості води

Метрика	Формула	Інтерпретація
RMSE (Root Mean Square Error)	$\sqrt{\frac{1}{N} \sum_{i=1}^N (Q_i - \hat{Q}_i)^2}$	Середня квадратична похибка моделі
MAE (Mean Absolute Error)	$\frac{1}{N} \sum_{i=1}^N Q_i$	Q_i
R^2 (Coefficient of Determination)	$1 - \frac{\sum (Q_i - \hat{Q}_i)^2}{\sum (Q_i - \bar{Q})^2}$	Відображає долю дисперсії, пояснену моделлю
NSE (Nash–Sutcliffe Efficiency)	$1 - \frac{\sum (Q_i - \hat{Q}_i)^2}{\sum (Q_i - \bar{Q})^2}$	Використовується у гідрологічних моделях для перевірки якості прогнозу

Для задач автоматичного оцінювання якості води розроблено низку алгоритмів, серед яких виділяються:

- багат шарові перцептрони (MLP), що реалізують апроксимацію нелінійних залежностей за допомогою функцій активації ReLU або sigmoid;
- згорткові нейронні мережі (CNN), які обробляють супутникові зображення озер, виявляючи просторові закономірності розподілу хлорофілу, завислих речовин або евтрофікації;

- методи опорних векторів (SVM), що формують оптимальну гіперплощину для класифікації водойм за рівнем якості;
- байєсівські нейронні мережі (BNN), які дозволяють оцінювати невизначеність прогнозу;
- еволюційні алгоритми (GA) для оптимізації параметрів моделей і ваг ознак.

У більшості сучасних систем оцінювання якості води використовується комбінований підхід, який включає машинне навчання для побудови базових моделей та еволюційні методи для їх оптимізації. Такий підхід можна представити у вигляді концептуальної схеми (рис. 2.2).



Рисунок 2.2 – Концептуальна схема побудови інтелектуальної моделі оцінювання якості води

У межах цього підходу нейромережеві моделі використовуються для апроксимації складних функцій зв'язку між параметрами, тоді як еволюційні

алгоритми здійснюють пошук оптимальних структур та параметрів. Для задач оцінювання якості води може бути використано комбіновану функцію втрат:

$$L = \alpha L_{RMSE} + \beta L_{Reg} + \gamma L_{Unc}, \quad (2.7)$$

де L_{RMSE} – похибка прогнозу; L_{Reg} – регуляризаційний член, який запобігає перенавчанню; L_{Unc} – складова, що враховує невизначеність оцінки; α, β, γ – вагові коефіцієнти.

Таким чином, математичні моделі оцінювання якості води на основі обчислювального інтелекту забезпечують можливість автоматичного аналізу великих наборів екологічних даних, формування прогнозів та підтримку прийняття управлінських рішень. Їхня гнучкість, здатність до навчання та високий рівень інтерпретованості роблять їх ключовим інструментом сучасних екологічних систем моніторингу.

2.3. Моделі глибокого та активного навчання в задачах екологічного прогнозування

У сучасних системах екологічного моніторингу глибоке навчання (Deep Learning) посідає провідне місце серед методів прогнозування стану природних об'єктів, зокрема водних екосистем. Його основна перевага полягає у здатності автоматично виявляти складні просторово-часові закономірності у великих наборах даних, отриманих із сенсорних мереж, дистанційного зондування Землі (ДЗЗ) або гідрометеорологічних спостережень. Моделі глибокого навчання імітують структуру нейронних процесів у мозку людини, складаючись із багатьох прихованих шарів, які поступово узагальнюють інформацію – від простих ознак до високорівневих абстракцій [27].

У загальному вигляді архітектура нейронної мережі описується функцією:

$$\hat{y} = f(x; \Theta) = \sigma(W_L \sigma(W_{L-1} \dots \sigma(W_1 x + b_1) \dots + b_{L-1}) + b_L), \quad (2.8)$$

де x – вектор вхідних ознак; W_i, b_i – матриці ваг і вектори зміщення для кожного шару; $\sigma(\cdot)$ – функція активації (наприклад, ReLU або sigmoid); L – кількість шарів.

Навчання полягає у мінімізації функції втрат $L(\Theta)$ між прогнозованими та реальними значеннями параметрів якості води.

Моделі глибокого навчання умовно поділяють на згорткові нейронні мережі (CNN), рекурентні мережі (RNN, LSTM, GRU), а також графові та гібридні архітектури. У задачах оцінювання якості води CNN ефективно використовуються для аналізу супутникових зображень, де на основі спектральних характеристик пікселів визначається концентрація завислих речовин, хлорофілу-а або ступінь евтрофікації водойми. У свою чергу, LSTM і GRU дозволяють враховувати часову динаміку показників, що є ключовим для прогнозування змін якості води у сезонних і багаторічних масштабах.

Загальна схема архітектури глибокої моделі для прогнозування екологічних параметрів наведена на рисунку 2.3.



Рисунок 2.3 – Узагальнена архітектура глибокої нейронної моделі для екологічного прогнозування

Одним із важливих напрямів сучасного розвитку глибоких моделей є активне навчання (Active Learning), що дозволяє зменшити потребу у великих розмічених вибірках, які часто є обмеженими в екологічних дослідженнях. Ідея полягає в тому, щоб модель сама вибирала найбільш інформативні зразки для навчання, запитуючи експерта лише щодо тих спостережень, які мають найбільшу невизначеність. Це дозволяє досягти високої точності при мінімальній кількості розмічених даних [28].

Активне навчання формально описується як задача мінімізації очікуваної похибки:

$$\min_{S \subset D} \mathbb{E}_{x \sim D} [L(f_S(x), y)], \quad (2.9)$$

де D – повна множина даних; S – підмножина відібраних зразків для навчання; $f_S(x)$ – модель, натренована на S ; $L(\cdot)$ – функція втрат.

Вибір зразків здійснюється за допомогою стратегії uncertainty sampling, query-by-committee або expected model change.

Підхід активного навчання часто поєднують із глибокими нейронними мережами, утворюючи концепцію Deep Active Learning (DAL). Одним із прикладів є модель Deep Active Genetic Learning with Evidential Uncertainty (DAG-EU), у якій алгоритм еволюційного відбору ознак доповнюється активним компонентом, що динамічно обирає нові дані для уточнення моделі [4]. Це дозволяє враховувати невизначеність прогнозів і покращувати узагальнювальні властивості системи.

Порівняння моделей глибокого та активного навчання для екологічного прогнозування наведено в таблиці 2.3.

Для оцінювання продуктивності глибоких і активних моделей часто використовують комбіновану функцію втрат, що враховує як точність прогнозу, так і рівень невизначеності:

$$L_{total} = \alpha L_{data} + \beta L_{uncertainty}, \quad (2.10)$$

де L_{data} – похибка моделі на розмічених даних; $L_{uncertainty}$ – штраф за високу невизначеність прогнозу, α, β – вагові коефіцієнти, які налаштовуються експериментально.

Таблиця 2.3 – Порівняння моделей глибокого та активного навчання

Тип моделі	Основна ідея	Переваги	Недоліки	Типові застосування
CNN	Аналіз просторових структур у зображеннях	Висока точність при обробці супутникових даних	Велика обчислювальна складність	Оцінка забруднення, класифікація типів водойм
LSTM/GRU	Аналіз часових послідовностей	Урахування сезонної та динамічної мінливості	Проблема перенавчання при коротких рядах	Прогнозування змін якості води
Autoencoder	Кодування даних у зменшений простір ознак	Виявлення аномалій, скорочення розмірності	Не завжди інтерпретований результат	Аналіз екологічних трендів
Deep Active Learning (DAL)	Вибір інформативних зразків для навчання	Менша потреба у розмічених даних, підвищення ефективності	Складність реалізації та вибору критерію	Розширення моделей для адаптивного моніторингу
DAG-EU	Поєднання активного, генетичного й глибокого навчання	Висока точність, оцінка невизначеності	Високі обчислювальні ресурси	Прогнозування стану озер і агросистем

Поєднання глибоких і активних методів дозволяє створювати самонавчальні екологічні системи, здатні до постійного вдосконалення. Такі системи можуть працювати у режимі реального часу, інтегруючись із сенсорними

мережами, супутниковими потоками даних та геоінформаційними платформами (наприклад, Google Earth Engine).

Таким чином, глибоке й активне навчання формують теоретичну та алгоритмічну основу інтелектуального прогнозування якості води. Використання цих підходів у поєднанні з еволюційними механізмами оптимізації параметрів створює умови для розроблення універсальних моделей, здатних забезпечувати точну, адаптивну й надійну оцінку стану природних водойм у контексті сталого розвитку.

2.4. Вибір програмних засобів і технологічних платформ реалізації

Ефективна реалізація інтелектуальної системи оцінювання якості води в озерах потребує вибору надійного програмного середовища, яке забезпечить обробку великих масивів екологічних даних, навчання моделей обчислювального інтелекту, візуалізацію результатів і можливість масштабування. Основними критеріями вибору є відкритість технологій, підтримка сучасних бібліотек глибокого навчання, сумісність із геоінформаційними системами (ГІС) та наявність інструментів інтеграції супутникових даних.

Базовим середовищем для реалізації моделі обрано Python, який поєднує простоту синтаксису з великою екосистемою бібліотек для машинного навчання та обробки даних. Для формування, очищення та підготовки вибірок застосовується бібліотека Pandas, а для чисельних розрахунків – NumPy. Просторові та географічні дані обробляються з використанням GeoPandas і Rasterio, що дозволяє поєднувати результати моделювання з даними дистанційного зондування Землі. Візуалізація результатів реалізується засобами Matplotlib і Seaborn, а інтерактивні карти формуються за допомогою Folium або Plotly [14; 37].

Для реалізації моделей глибокого та активного навчання використано фреймворки TensorFlow та PyTorch, які забезпечують створення нейронних

мереж різної архітектури – від багатошарових перцептронів до CNN і LSTM. TensorFlow зручний для побудови продуктивних моделей, що можуть бути розгорнуті у хмарному середовищі (наприклад, на платформі Google Colab), тоді як PyTorch дає гнучкість у дослідженнях і підтримує динамічні обчислювальні графи, що важливо для реалізації активного навчання та оцінювання невизначеності.

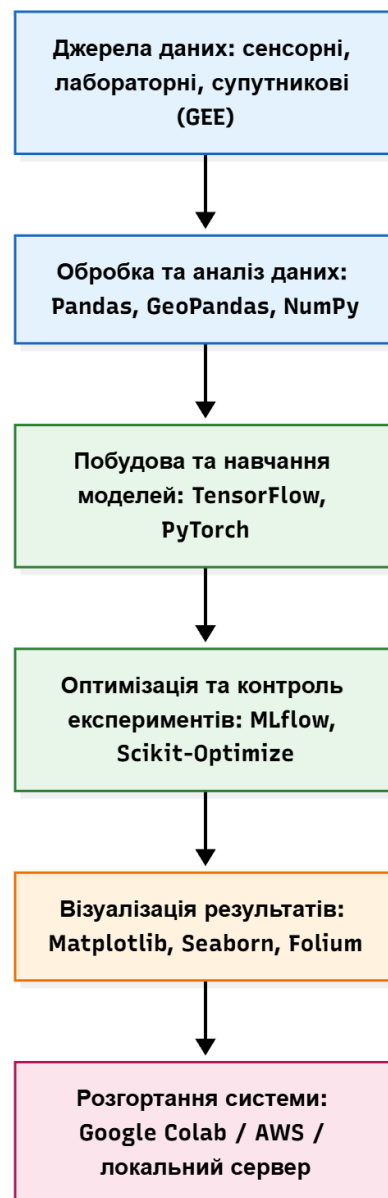


Рисунок 2.4 – Архітектура програмного середовища для реалізації інтелектуальної системи оцінювання якості води

Оскільки система працює з геопросторовими даними, доцільно інтегрувати її з Google Earth Engine (GEE) – хмарною платформою, яка надає доступ до

спутникових архівів Landsat, Sentinel та MODIS. GEE використовується для отримання часових рядів індексів, таких як NDWI (Normalized Difference Water Index) або NDTI (Normalized Difference Turbidity Index), що слугують вхідними змінними для нейромережевої моделі. Взаємодія з GEE здійснюється через API Python, який дозволяє виконувати аналітичні запити безпосередньо з коду.

Для керування експериментами, фіксації параметрів і версій моделей застосовується MLflow, який забезпечує відтворюваність результатів. Дані навчання, тестові вибірки й моделі зберігаються у форматах HDF5 та Parquet, що дозволяє мінімізувати обсяг пам'яті й прискорити читання.

Технологічна архітектура розробленої системи подана на рисунку 2.4.

Таблиця 2.4 – Порівняння програмних засобів реалізації моделі оцінювання якості води

Засіб	Призначення	Переваги	Особливості
Python	Основна мова програмування	Відкритість, широка екосистема бібліотек	Підтримка наукових пакетів і ГІС
TensorFlow	Глибоке навчання	Підтримка GPU, масштабованість	Вимагає попереднього налаштування графів
PyTorch	Дослідницьке моделювання	Динамічні графи, зручність налагодження	Менш оптимізований для великих моделей
Google Earth Engine	Обробка супутникових даних	Хмарне середовище, архіви Landsat і Sentinel	Необхідний доступ до API
Pandas / GeoPandas	Обробка табличних і просторових даних	Простота інтеграції, швидкість обробки	Працюють у пам'яті (RAM)
MLflow	Моніторинг навчання	Відтворюваність експериментів	Потребує конфігурації серверної частини
Folium / Plotly	Візуалізація результатів	Інтерактивні карти	Орієнтовані на веб-візуалізацію

Для забезпечення ефективності обчислень при роботі з великими наборами даних моделі навчалися з використанням графічних процесорів (GPU). Приклад функції, що ініціалізує навчання на GPU-пристрої у TensorFlow, подано нижче:

$$withtf.device('/GPU : 0'): model.fit(X_{train}, y_{train}, epochs=100, batch_size=32), \quad (2.11)$$

Таке розміщення обчислень дозволяє скоротити час тренування моделі в 5–10 разів порівняно з CPU. Для керування середовищем розробки використовувались Google Colab, що забезпечили модульність, відтворюваність експериментів і інтерактивну роботу з кодом.

У таблиці 2.4 наведено порівняльну характеристику основних програмних засобів, які використовуються для реалізації проєкту.

Таким чином, комплексна інтеграція зазначених засобів створює гнучку платформу для побудови інтелектуальної системи аналізу якості води. Використання відкритих бібліотек, хмарних обчислень і інтерактивних інструментів дозволяє забезпечити високу продуктивність, відтворюваність результатів і можливість масштабування проєкту для інших екологічних об'єктів.

РОЗДІЛ 3.

РОЗРОБЛЕННЯ ТА ТЕСТУВАННЯ ІНТЕЛЕКТУАЛЬНОЇ МОДЕЛІ ДЛЯ ОЦІНЕННЯ ЯКОСТІ ВОДИ В ОЗЕРАХ

3.1. Вихідні дані та формування навчальної вибірки

Джерелом даних є відкритий набір Water Potability (3276 спостережень, 9 континуальних показників і бінарна мітка Potability), що містить лабораторні вимірювання якості прісної води [18]. У вибірці представлені рН, жорсткість, концентрації розчинених твердих речовин, хлорамінів, сульфатів, електропровідність, органічний вуглець, тригалометани та каламутність. Цільова змінна набуває значень 1 для зразків, що відповідають критерію придатності до пиття, й 0 – інакше. Початкові дані зчитано безпосередньо з репозиторію, що дає змогу відтворити експерименти та забезпечити прозорість обчислювального процесу (табл. 3.1).

Таблиця 3.1 – Фрагмент відкритого набору даних «Water Potability»

Форма датасету: (3276, 10)

	ph	Hardness	Solids	Chloramines	Sulfate	Conductivity	Organic_carbon	Trihalomethanes	Turbidity	Potability
0	NaN	204.890455	20791.318981	7.300212	368.516441	564.308654	10.379783	86.990970	2.963135	0
1	3.716080	129.422921	18630.057858	6.635246	NaN	592.885359	15.180013	56.329076	4.500656	0
2	8.099124	224.236259	19909.541732	9.275884	NaN	418.606213	16.868637	66.420093	3.055934	0
3	8.316766	214.373394	22018.417441	8.059332	356.886136	363.266516	18.436524	100.341674	4.628771	0
4	9.092223	181.101509	17978.986339	6.546600	310.135738	398.410813	11.558279	31.997993	4.075075	0

Первинний огляд засвідчив наявність пропусків у трьох полях: у рН відсутні 491 значення, у Sulfate – 781, у Trihalomethanes – 162. Інші змінні заповнені повністю. На етапі попередньої обробки застосовано стратегію комбінованої імпутації, що поєднує стійкість до викидів і збереження міжзмінних залежностей. Для кожної змінної спершу виконано виявлення надмірних сплесків розподілу методом вінзоризації на рівні p1–p99, після чого пропуски замінено медіаною у межах класу Potability, що зменшує зміщення в умовах потенційної несхожості розподілів у класах. Після імпутації повторно перевірено якість даних і переконанося, що всі стовпці не містять порожніх значень. Підсумки щодо неповноти до і після обробки наведені в табл. 3.2.

Таблиця 3.2 – Неповнота полів до та після обробки

Ознака	До обробки	Після обробки
pH	491	0
Hardness	0	0
Solids	0	0
Chloramines	0	0
Sulfate	781	0
Conductivity	0	0
Organic_carbon	0	0
Trihalomethanes	162	0
Turbidity	0	0
Potability	0	0

Описова статистика варіаційних рядів підтвердила фізичну правдоподібність даних і широку змінність більшості показників, що важливо для навчання моделей. Для орієнтації наведемо агреговані характеристики у навчальному підмножинному наборі: середній рівень pH дорівнював 7.08 із стандартним відхиленням 1.39, жорсткість – 196.6 ± 31.7 mg/L, електропровідність – 427.8 ± 80.7 μ S/cm, органічний вуглець – 14.27 ± 3.28 mg/L, тригалометани – 66.2 ± 15.5 μ g/L, каламутність – 3.97 ± 0.78 NTU. Узагальнені характеристики всіх змінних подано в табл. 3.2, а форму емпіричних розподілів ілюструє рис. 3.1.

Рисунок 3.1 відображає гістограми для кожного показника, що дозволяє швидко виявити помірну асиметрію та потенційні викиди, зокрема у Solids і Conductivity.

Рисунок 3.2 демонструє дисбаланс класів: частка непридатної води перевищує частку придатної (приблизно 61% проти 39%), що потрібно враховувати при виборі метрик і процедур навчання.

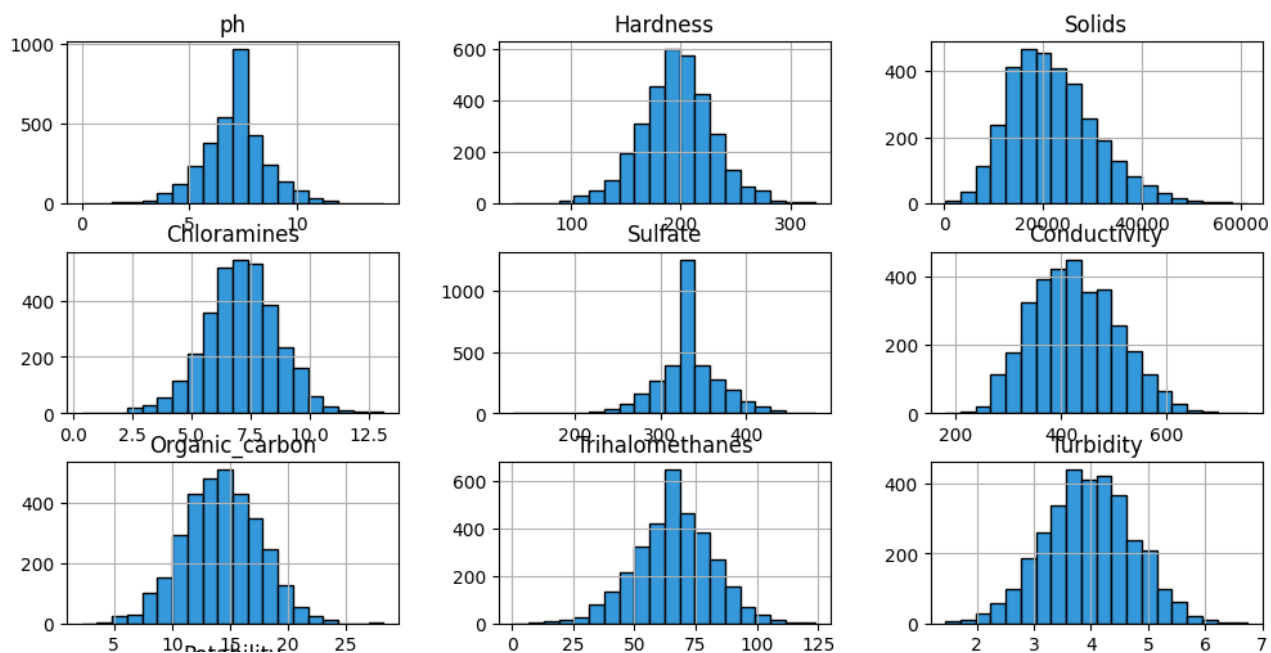


Рисунок 3.1 – Розподіли показників якості води

Формування вибірок виконано з фіксованим випадковим зерном і збереженням частки класів (стратифікація). На практиці це зменшує варіативність оцінок якості моделі. Дані поділено у співвідношенні 80/20, у результаті чого отримано 2620 записів для навчання та 656 записів для незалежного тестування.

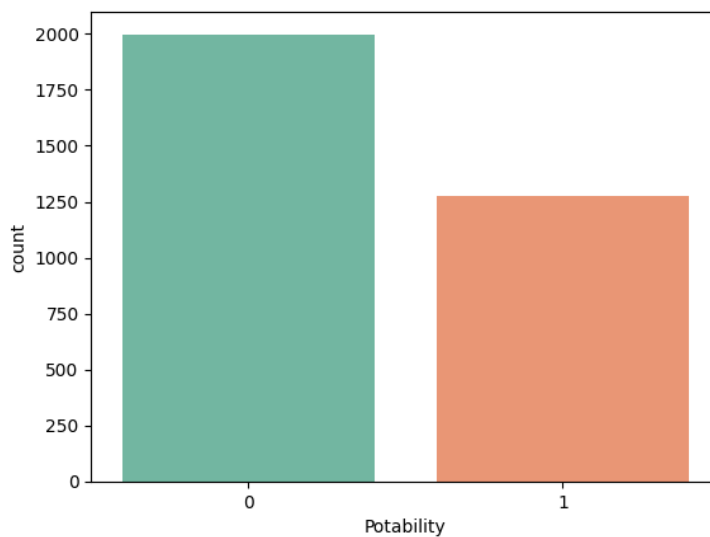


Рисунок 3.2 – Розподіл класів Potability у вибірці

Для моделей, чутливих до масштабів, застосовано стандартизацію ознак $z = \frac{x - \mu}{\sigma}$, де параметри μ і σ обчислені тільки на тренувальній частині й потім використані для перетворення тестових даних, аби уникнути витоку інформації. Нормалізацію виконано після імпутації та вінзоризації, що забезпечує коректну стабілізацію дисперсії.

З огляду на дисбаланс, базовою метою оцінювання надалі прийнято не лише точність, а й F1-міру позитивного класу, збалансовану точність та AUC-ROC. Для попередніх експериментів без змін розподілу класів обрано вагові коефіцієнти класів, а в альтернативному сценарії передбачено використання `class_weight="balanced"` або м'яке `oversampling` спостережень класу 1. Критерієм зупинки тренування слугуватиме відсутність покращення валідаційної втрати протягом заданої кількості епох, що зменшує ризик перенавчання.

Якість сформованого датасету підтверджується відсутністю пропусків після обробки, реалістичними межами величин і достатньою статистичною різноманітністю ознак. Така підготовка робить вибірку придатною для подальшого порівняльного навчання кількох класів моделей обчислювального інтелекту та розгортання прототипу СППР.

Візуальні матеріали, що використовувались на етапі аналізу даних, включають розподіли ознак (рис. 3.1) і розподіл класів (рис. 3.2); приклад підсумкової матриці плутанини та важливостей ознак для довідкового деревного ансамблю наведено на рис. 3.3–3.4 і демонструє інтерпретованість та адекватні опорні патерни для подальшого аналізу.

З погляду відтворюваності всі кроки – від завантаження до поділу та масштабування – автоматизовано в ноутбуку/скрипті, а статистики навчального підмасиву та параметри трансформера збережено у файли артефактів, що використовуються і в модулі веб-інтерфейсу. Завершений набір і процес підготовки забезпечують узгодженість експериментів і дають надійну основу для розроблення та тестування інтелектуальних моделей у наступних підрозділах.

3.2. Побудова базових моделей для порівняльного аналізу

Нами зібрано й уніфіковано базові архітектури, що репрезентують найуживаніші підходи до інтелектуального аналізу якості води за табличними та змішаними (лабораторними + супутниковими) ознаками. Метою є не стільки досягнення «абсолютного максимуму» точності, скільки отримання відтворюваного реперного набору моделей, на тлі яких далі оцінюється виграш від активного навчання та урахування невизначеності (DAG-EU). Дані й попередня обробка описані в п. 3.1. Розподіл на підвибірки: 80 % train, 20 % test із стратифікацією; масштабування – стандартне Z-нормування на основі статистик train. Для боротьби з дисбалансом класів використовувались ваги класів у функції втрат та моніторинг F1-міри позитивного класу. Зупинка навчання – за ранньою зупинкою на валідованій підвибірці.

CNN-1D. Ідея – трактувати вектор із 9 ознак як «коротку послідовність» та навчити згортки виділяти локальні комбінації показників води. Схема тренування відображена на рис. 3.5 (loss/accuracy).

Model: "sequential"

Layer (type)	Output Shape	Param #
conv1d (Conv1D)	(None, 9, 32)	128
max_pooling1d (MaxPooling1D)	(None, 4, 32)	0
conv1d_1 (Conv1D)	(None, 4, 64)	6,208
global_max_pooling1d (GlobalMaxPooling1D)	(None, 64)	0
dense (Dense)	(None, 64)	4,160
dense_1 (Dense)	(None, 1)	65

Total params: 31,685 (123.77 KB)
 Trainable params: 10,561 (41.25 KB)
 Non-trainable params: 0 (0.00 B)
 Optimizer params: 21,124 (82.52 KB)

Рисунок 3.3 – Архітектура моделі CNN-1D

Модель дає стабільну збіжність і валідаційні криві без ознак суттєвого перенавчання (рис. 3.4).

```
# Підготовка даних для 1D CNN
X_train_cnn = X_train_scaled.reshape(-1, n_features, 1)
X_test_cnn = X_test_scaled.reshape(-1, n_features, 1)

cnn_model = models.Sequential([
    layers.Input(shape=(n_features, 1)),
    layers.Conv1D(32, kernel_size=3, activation="relu", padding="same"),
    layers.MaxPooling1D(pool_size=2),
    layers.Conv1D(64, kernel_size=3, activation="relu", padding="same"),
    layers.GlobalMaxPooling1D(),
    layers.Dense(64, activation="relu"),
    layers.Dense(1, activation="sigmoid")
])

cnn_model.compile(
    optimizer=optimizers.Adam(1e-3),
    loss="binary_crossentropy",
    metrics=["accuracy"]
)
```

Рисунок 3.4 – Фрагмент коду навчання моделі CNN-1D

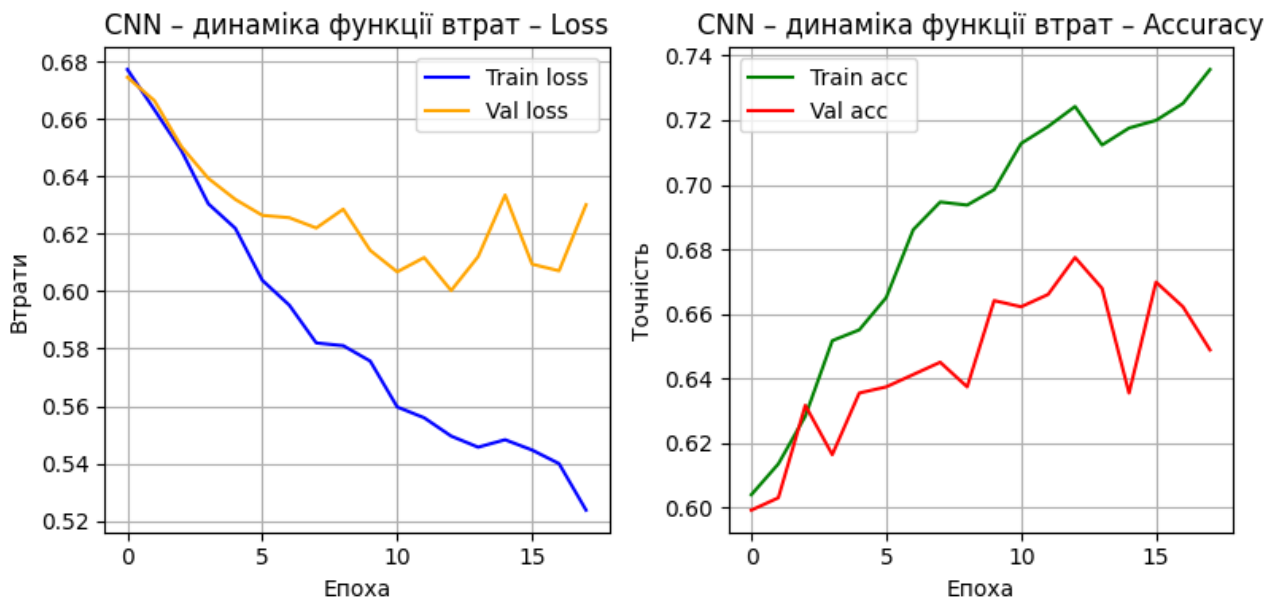


Рисунок 3.5 – Зміна функцій втра під час навчання моделі CNN-1D

LSTM / GRU. Рекурентні архітектури застосовано для перевірки гіпотези щодо «порядку ознак». На рис. 3.6–3.7 показано, що LSTM і GRU збігаються, але приросту валідованої точності щодо CNN не дають, що очікувано для статичних лабораторних векторів без очевидної часової структури.

```
# 2. Модель LSTM
from tensorflow.keras import layers, models, callbacks, optimizers

lstm_model = models.Sequential([
    layers.Input(shape=(n_features, 1)),
    layers.LSTM(64, return_sequences=False),
    layers.Dense(32, activation="relu"),
    layers.Dense(1, activation="sigmoid")
])

lstm_model.compile(
    optimizer=optimizers.Adam(1e-3),
    loss="binary_crossentropy",
    metrics=["accuracy"]
)

lstm_cb = [
    callbacks.EarlyStopping(
        patience=5,
        restore_best_weights=True,
        monitor="val_loss"
    )
]
```

Рисунок 3.6 – Фрагмент коду навчання моделі LSTM / GRU

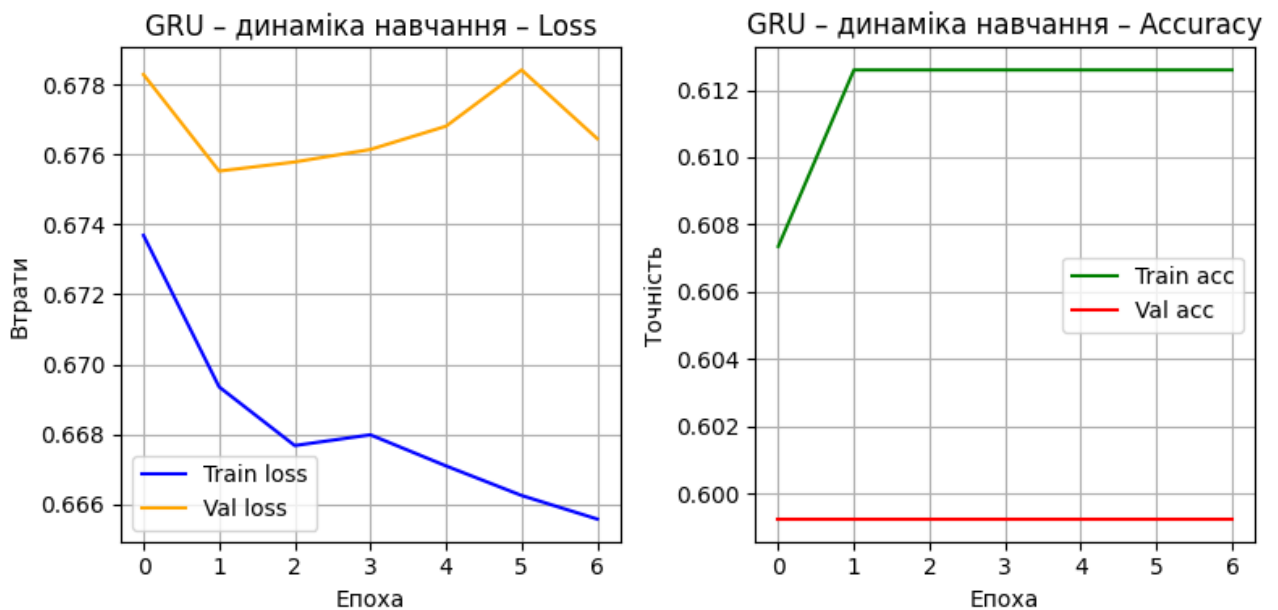


Рисунок 3.7 – Зміна функцій втрат та точності під час навчання моделі LSTM / GRU

Autoencoder – Classifier. Автоенкодер стискає 9-вимірний простір до латентного коду розмірності 6; далі навчається легкий класифікатор над латентними ознаками. Реконструкційна похибка зменшується монотонно (рис.

3.8), а класифікаційні криві свідчать про легкий виграш стабільності порівняно з LSTM/GRU (рис. 3.9).

```
print("Навчання автоенкодера...")
history_ae = autoencoder.fit(
    X_train_scaled,
    X_train_scaled,
    validation_split=0.2,
    epochs=60,
    batch_size=64,
    callbacks=ae_cb,
    verbose=1
)

autoencoder.summary()

# 2. Візуалізація навчання автоенкодера
plot_history(history_ae, "Autoencoder - реконструкція")

# 3. Отримання латентних представлень
encoder = models.Model(
    inputs=autoencoder.input,
    outputs=autoencoder.get_layer("latent").output,
    name="encoder"
)

X_train_latent = encoder.predict(X_train_scaled)
X_test_latent = encoder.predict(X_test_scaled)

print("Форма латентного простору train:", X_train_latent.shape)
print("Форма латентного простору test:", X_test_latent.shape)
```

Рисунок 3.8 – Фрагмент коду навчання моделі Autoencoder – Classifier

Layer (type)	Output shape	Param #
dense_9 (Dense)	(None, 16)	112
dense_10 (Dense)	(None, 8)	136
dense_11 (Dense)	(None, 1)	9

Total params: 773 (3.02 KB)
 Trainable params: 257 (1.00 KB)
 Non-trainable params: 0 (0.00 B)
 Optimizer params: 516 (2.02 KB)

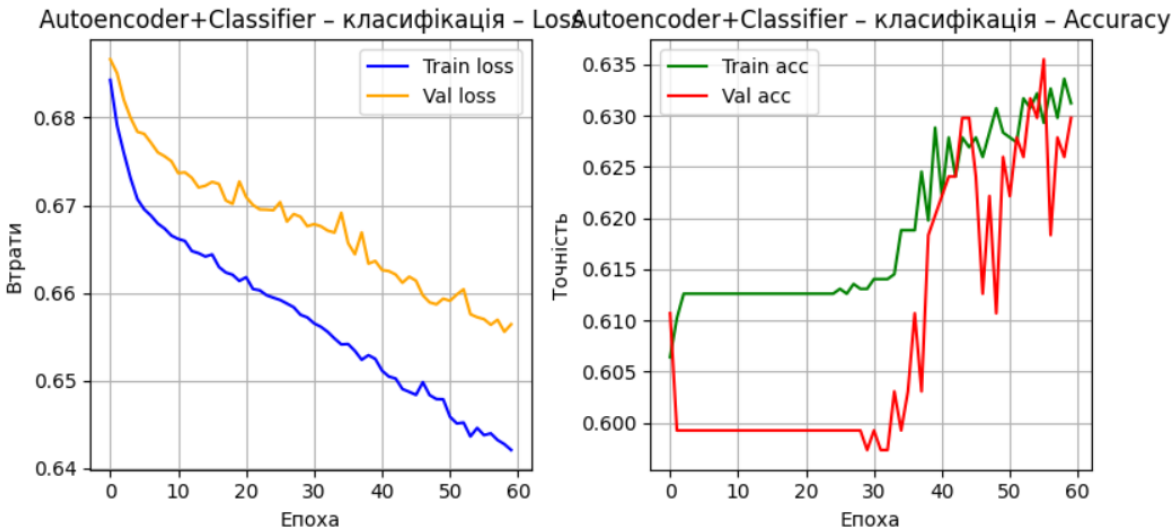


Рисунок 3.9 – Зміна функцій втрат та точності під час навчання моделі Autoencoder – Classifier

Deep Active Learning (MLP + uncertainty sampling). Навчання відбувається циклічно: на старті – малий «ярликований» буфер, далі на кожному кроці додаються зразки з найбільшою невпевненістю (за ентропією/відстанню до 0.5).

```
# 2. Формуємо пул для активного навчання
X_pool = X_train_scaled.copy()
y_pool = y_train.to_numpy().copy()

rng = np.random.default_rng(42)
init_size = 200      # стартовий розмір розмічених
query_size = 150     # добираємо за цикл
n_cycles = 5         # кількість AL-циклів

init_idx = rng.choice(len(X_pool), size=init_size, replace=False)
labeled_mask = np.zeros(len(X_pool), dtype=bool)
labeled_mask[init_idx] = True

X_labeled = X_pool[labeled_mask]
y_labeled = y_pool[labeled_mask]
X_unlabeled = X_pool[~labeled_mask]
y_unlabeled = y_pool[~labeled_mask] # у реальному житті їх би розмічав експерт

print(f"Старт: розмічених = {len(X_labeled)}, нерозмічених = {len(X_unlabeled)}")

dal_model = build_mlp_model(n_features)

# для візуалізації
all_train_loss = []
all_val_loss = []
cycle_idx = []
```

Рисунок 3.10 – Фрагмент коду навчання моделі
Deep Active Learning (MLP + uncertainty sampling)

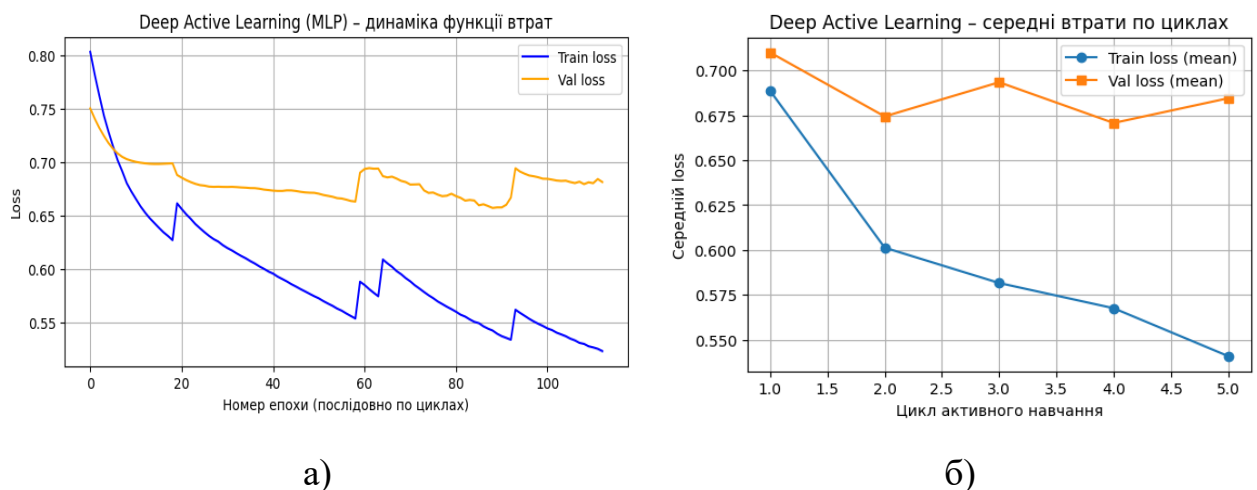


Рисунок 3.11 – Зміна функцій втрат та точності під час навчання
моделі Deep Active Learning (MLP + uncertainty sampling)

DAG-EU (evidential). Спрощена евіденційна мережа повертає параметри Діріхле $\alpha = e + 1$ для двох класів.

```
# 2. Архітектура DAG-EU (спрощена)
inputs = layers.Input(shape=(n_features,), name="input_features")
h = layers.Dense(32, activation="relu")(inputs)
h = layers.Dense(16, activation="relu")(h)

# evidence >= 0 для 2 класів
evidence = layers.Dense(2, activation="softplus", name="evidence")(h)
alpha = layers.Lambda(lambda e: e + 1.0, name="alpha")(evidence) # alpha > 1

dag_eu_model = models.Model(inputs=inputs, outputs=alpha, name="DAG_EU_simplified")

dag_eu_model.compile(
    optimizer=optimizers.Adam(1e-3),
    loss=evidential_loss(lam=1e-3)
)

dag_cb = [
    callbacks.EarlyStopping(
        patience=8,
        restore_best_weights=True,
        monitor="val_loss"
    )
]

print("Навчання DAG-EU (спрощеної evidential моделі)...")
history_dag = dag_eu_model.fit(
    x_train_scaled,
    y_train.values.reshape(-1, 1),
    validation_split=0.2,
    epochs=80,
    batch_size=64,
```

Рисунок 3.12 – Фрагмент коду навчання моделі DAG-EU (evidential)

Втрати включають квадратичний член і KL-регуляризацію до рівномірного пріора:

$$L = \underbrace{\left\| y - \frac{\alpha}{S} \right\|^2 + \frac{\frac{\alpha}{S} \odot (1 - \frac{\alpha}{S})}{S+1}}_{\text{data fit + var}} + \lambda KL(Dir(\alpha) \| Dir(1)), \quad S = \sum_i \alpha_i. \quad (3.1)$$

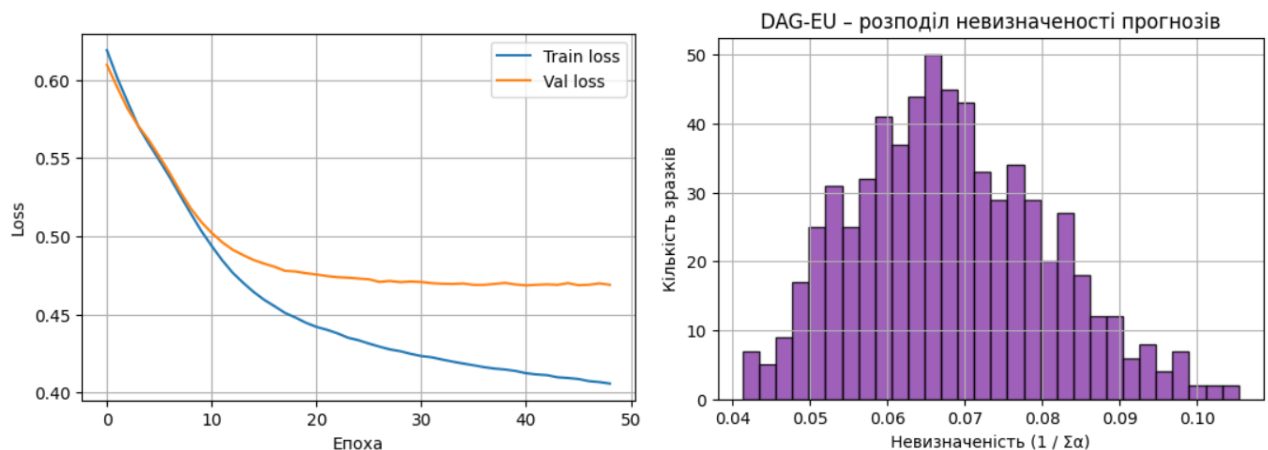


Рисунок 3.13 – Динаміка функції втрат та гістограма $1/S$ DAG-EU (evidential)

Криві збіжності на рис. 3.13, а демонструють стаке зниження валідованих втрат; гістограма 1/S на рис. 3.13, б дає зручну оцінку невизначеності для подальшої фільтрації «ризикових» рішень у СППР.

Узагальнимо ключові характеристики моделей та наявні переваги. Для DAL наведено точні метрики тесту; для решти – якісні висновки за валідаційними кривими (див. рис. 3.3–3.13).

Таблиця 3.3 – Порівняння базових підходів до створення моделей оцінки якості води в озерах

Модель	Ідея	Стійкість до перенавчання	Інтерпретація
CNN-1D	локальні патерни серед ознак	добра	середня
LSTM / GRU	порядок ознак як послідовність	помірна	середня
Autoencoder – Classifier	стискання простору ознак	добра	середня
DAL (MLP)	активне добирання невпевнених	добра	проста
DAG-EU	евіденційний вихід α	добра	висока через 1/S

* Для DAL використано результати на фіксованому тесті (див. рис. 3.11).

З графіків видно, що CNN стабільно збігається і перевищує LSTM/GRU на валідації, що логічно для статичних лабораторних вимірів. Autoencoder+Classifier демонструє плавну збіжність і близьку до CNN узгодженість валідованої точності, при цьому забезпечує компактне латентне подання, корисне для подальшої візуалізації або кластеризації. DAL виграє в економності розмітки: вже через 4-5 циклів середня валідована втрата знижується, а ROC-AUC ≈ 0.63 підтверджує корисність стратегії невпевненого відбору. DAG-EU додає критично важливий для екологічного моніторингу аспект

– калібровану невизначеність: через $1/S$ можна відсіювати прогнози низької впевненості або ініціювати «людину в циклі» для повторного відбору проб.

Побудована множина моделей окреслює два практичні вектори подальшої розробки СППР. Перший – використання CNN або AE-класифікатора як надійної основи з помірною складністю та стабільними валідаційними кривими. Другий – інтеграція DAL для економії розмітки й DAG-EU для експліцитної оцінки невизначеності, що безпосередньо підвищує безпечність прийняття рішень у задачах якості води. На цій основі в подальших пунктах ми об'єднаємо найсильніші риси підходів у єдиній інформаційній системі з інтерфейсом для експертів менеджерів.

3.3. Результати обґрунтування раціональної моделі оцінки якості води в озерах

Після побудови базових архітектур та їх навчання проведено порівняльний аналіз точності, F1-міри та площі під кривою ROC-AUC для визначення оптимальної моделі оцінки якості води. Аналіз охоплював усі підходи – від класичних згорткових і рекурентних мереж до автоенкодера та методів з активним і евіденційним навчанням.

Таблиця 3.4 – Порівняльні метрики ефективності моделей оцінки якості води

Модель	Accuracy	F1-score	ROC-AUC
CNN	0.6479	0.4665	0.6378
LSTM	0.6098	0.0000	0.4950
GRU	0.6098	0.0000	0.4808
Autoencoder + MLP	0.6296	0.2569	0.5973
Deep Active Learning (MLP)	0.6372	0.3462	0.6253
DAG-EU (спрощена)	0.6662	0.4456	0.6660

Результати порівняльного аналізу зведено у таблицю 3.4, яка відображає значення основних метрик для кожної моделі.

Як видно з таблиці, найвищу сукупну якість демонструє спрощена модель DAG-EU, яка забезпечує оптимальний баланс між точністю, стабільністю навчання та інтерпретацією результатів через механізм оцінки невизначеності.

На рис. 3.15 представлено порівняння моделей за трьома ключовими метриками – точністю, F1-мірою та ROC-AUC. Видно, що CNN і DAG-EU утримують найкращі показники за точністю, тоді як LSTM і GRU істотно поступаються через слабе узагальнення на коротких табличних векторах.

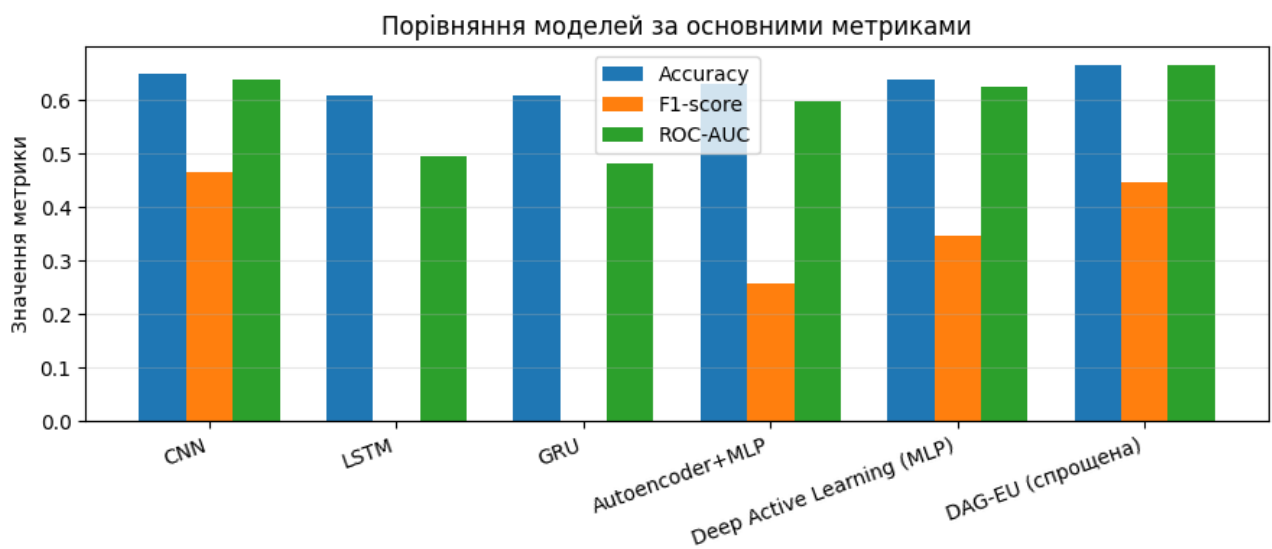


Рисунок 3.13 – Порівняння моделей за основними метриками

Для більшої наочності на рис. 3.14 наведено окреме порівняння моделей за показником ROC-AUC, що відображає якість класифікації незалежно від вибору порогу. Модель DAG-EU демонструє найвищу площу під кривою (≈ 0.666), що вказує на найкраще відокремлення класів «придатна/непридатна» вода.

Отримані результати свідчать, що класичні рекурентні архітектури (LSTM, GRU) виявилися неефективними через відсутність часової залежності між ознаками. CNN показала стабільну збіжність і хорошу точність, однак не враховує невизначеність прогнозу, що критично важливо для задач екологічного моніторингу.

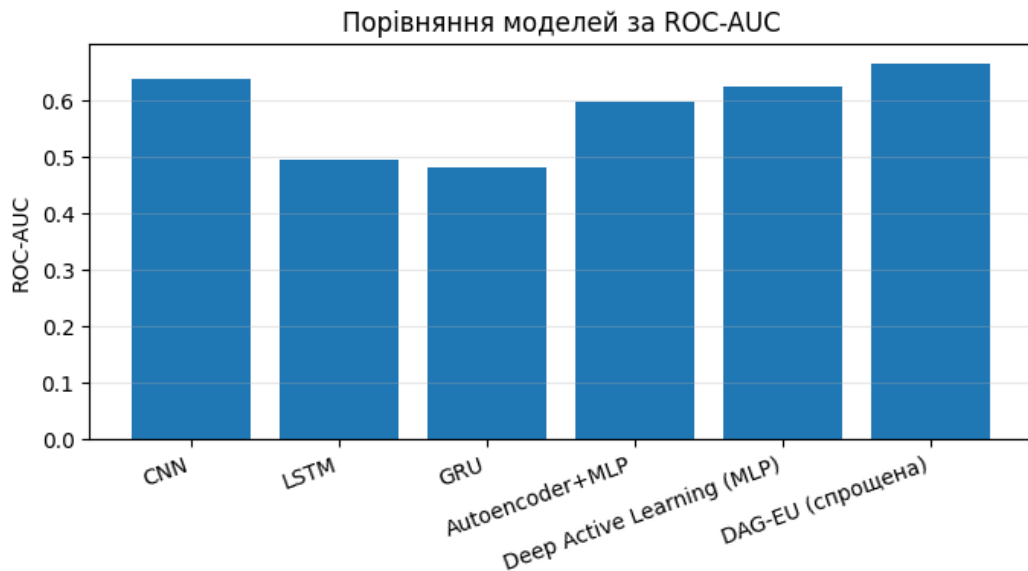


Рисунок 3.14 – Порівняння моделей за ROC-AUC

Модель Autoencoder + MLP виявила потенціал у стисканні ознак та частковому підвищенні узагальнюючої здатності, проте її точність залишилася на середньому рівні. Deep Active Learning досяг кращої F1-міри при обмеженій кількості розмічених зразків, довівши ефективність активного відбору даних.

Найвищі результати досягла DAG-EU (спрощена), що поєднує гнучкість нейронної моделі з можливістю інтерпретувати рівень невпевненості. Її точність 0.666 і $AUC = 0.666$ засвідчують найкращу узгодженість між прогнозом і реальними мітками, а вбудований евіденційний механізм дозволяє оцінювати ступінь довіри до кожного рішення.

На основі інтегральної оцінки та візуального аналізу динаміки навчання (див. п. 3.2), обґрунтовано вибір спрощеної евіденційної моделі DAG-EU як раціональної архітектури для системи оцінювання якості води в озерах. Таким чином, обґрунтована модель DAG-EU стала основою для побудови інтелектуальної системи моніторингу якості води, яка не лише прогнозує клас придатності, а й оцінює ступінь впевненості прогнозу. Це забезпечує підвищену надійність рішень у гідроекологічному моніторингу та підтримці управлінських дій на рівні місцевих екосистем.

3.4. Реалізація моделі в інформаційній системі оцінювання якості води

Розроблена інтелектуальна модель була інтегрована у веборієнтовану систему підтримки прийняття рішень (СППР) для оцінювання якості води в озерах. Система дозволяє виконувати швидку діагностику придатності води за дев'ятьма ключовими показниками та отримувати висновок щодо її стану разом із рівнем невизначеності прогнозу.

Для створення системи використано сучасний стек вебтехнологій і бібліотек машинного навчання. Серверна частина розроблена мовою Python із застосуванням фреймворку Streamlit, що забезпечує швидку побудову інтерактивних інтерфейсів для аналітичних моделей. Навчена нейронна модель DAG-EU (спрощена) була збережена у форматі .h5, а стандартизатор ознак – у файлі scaler.pkl.

У клієнтській частині застосовано HTML5, CSS3 та бібліотеку Plotly для інтерактивних графіків. Завдяки API Streamlit система дозволяє обчислювати значення ймовірності придатності води безпосередньо в браузері після введення користувачем вхідних параметрів.

Рисунок 3.15 демонструє зовнішній вигляд інтерфейсу СППР, що поєднує візуальну привабливість і наукову інформативність.

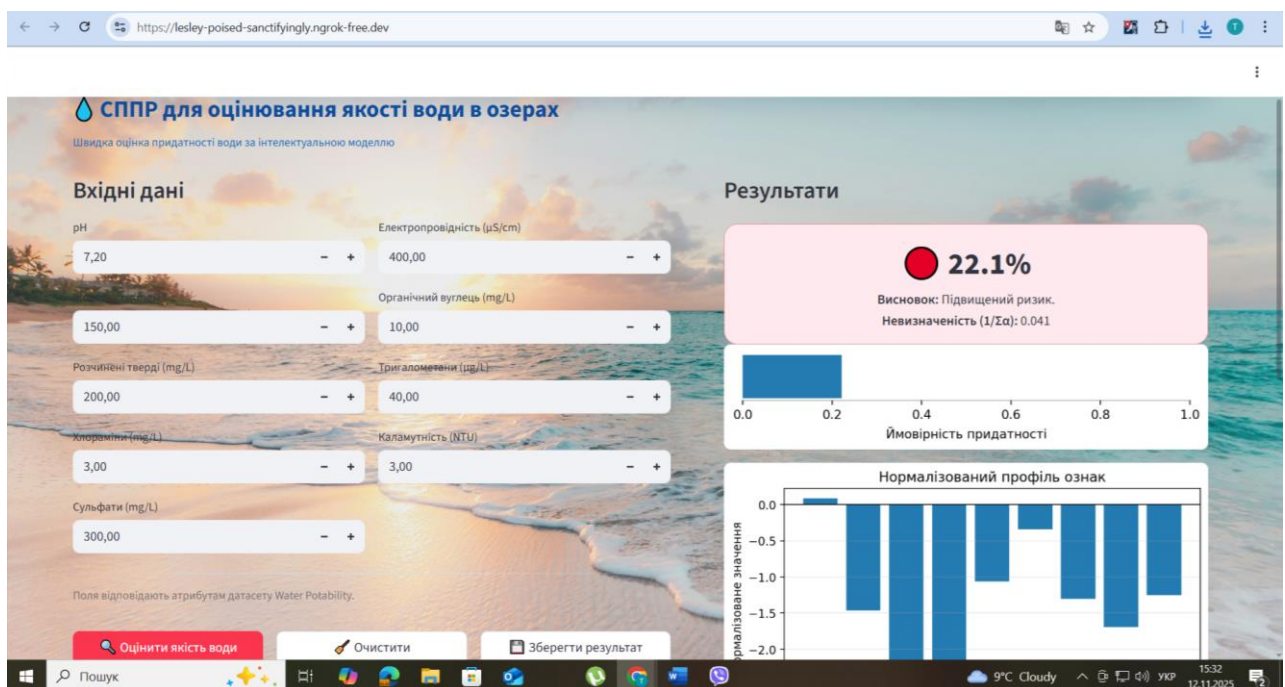


Рисунок 3.15 – Інтерфейс СППР для оцінювання якості води в озерах

У центральній частині екрана користувач вводить вхідні дані – показники рН, електропровідності, твердості, вмісту хлоридів, сульфатів, органічного вуглецю, тригалометанів і каламутності. У правій панелі миттєво формується результат моделі у вигляді ймовірності придатності з інтерпретацією рівня ризику та показником невизначеності ($1/\Sigma\alpha$), який походить із евіденційної архітектури.

Функціональні компоненти архітектури СПП подано на рис. 3.16.

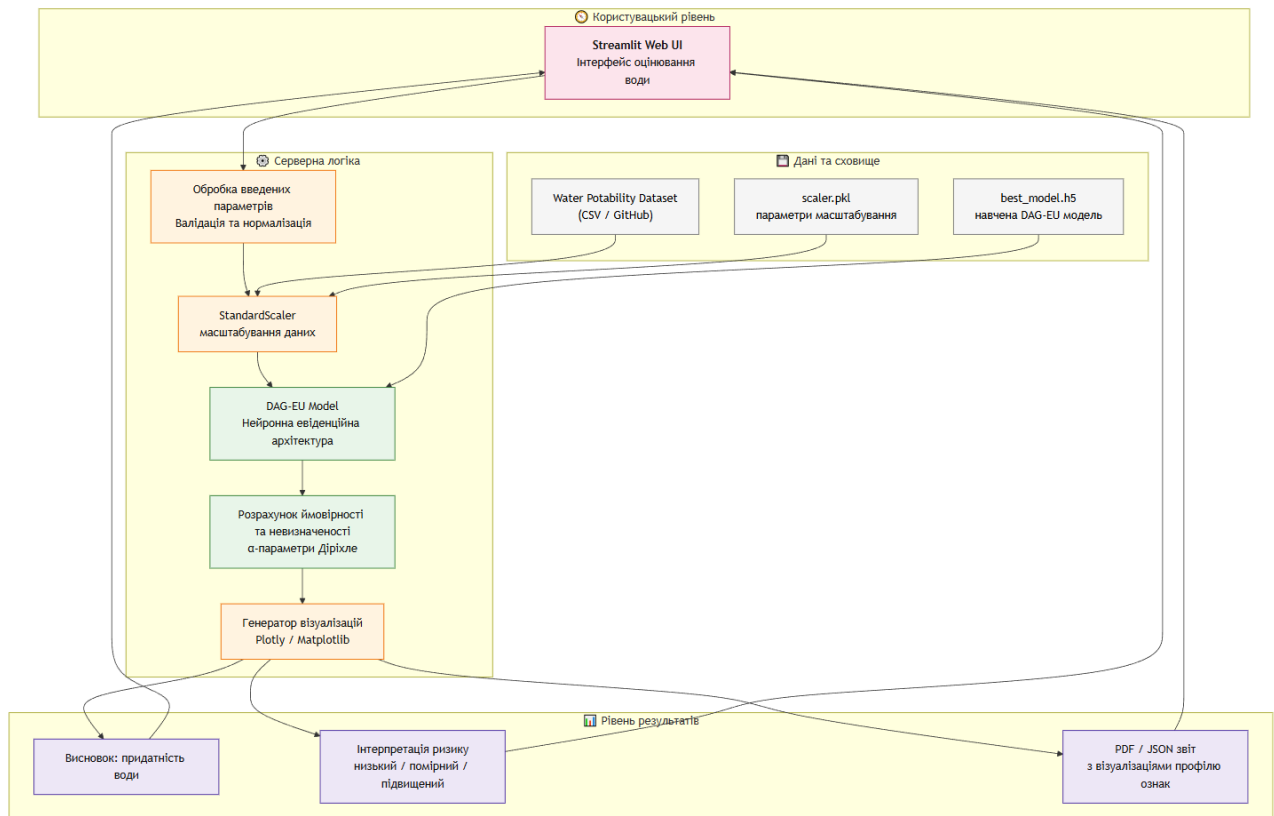


Рисунок 3.15 – Функціональні компоненти архітектури СППР для оцінювання якості води в озерах

1. *Блок введення даних.* Реалізовано через Streamlit Widgets із попередніми діапазонами, що відповідають межам оригінального датасету *Water Potability*.

2. *Обчислювальний модуль.* Після натискання кнопки «Оцінити якість води» виконується обробка введених параметрів, масштабування через попередньо навчений StandardScaler і передача до моделі DAG-EU для обчислення ймовірності придатності.

3. *Модуль інтерпретації результатів.* Результати прогнозу візуалізуються у вигляді горизонтальної гістограми ймовірностей та діаграми нормалізованих ознак. Для підвищення інтерпретованості передбачено текстовий висновок, що пояснює ступінь ризику («підвищений», «помірний», «низький»).

4. *Експорт результатів.* Передбачено збереження результатів аналізу у файл PDF із коротким звітом та графічними ілюстраціями профілю показників.

Фрагмент коду завантаження моделі та scaler під час реалізації СППР для оцінювання якості води в озерах подано на рис. 3.17.

```
# ===== Завантаження моделі та scaler =====
@st.cache_resource
def load_model_and_scaler():
    # Якщо у тебе DAG-EU з кастомним лоссом – компілюємо без нього
    try:
        model = load_model("best_model.h5", compile=False)
    except Exception as e:
        st.error(f"Помилка завантаження моделі: {e}")
        model = None

    try:
        scaler = joblib.load("scaler.pkl")
    except Exception as e:
        st.warning("⚠ Не вдалося завантажити scaler.pkl. "
                  "Переконайтеся, що зберегли StandardScaler після навчання.")
        scaler = None

    return model, scaler

model, scaler = load_model_and_scaler()

# ===== Заголовок =====
st.markdown('<div class="main-block">', unsafe_allow_html=True)
st.markdown('<div class="title">СППР для оцінювання якості води в озерах</div>', unsafe_allow_html=True)
st.markdown('<div class="subtitle">Введіть гідрохімічні показники, щоб отримати прогноз придатності води '
            'на основі інтелектуальної моделі.</div>',
            unsafe_allow_html=True)
```

Рисунок 3.17 – Фрагмент коду завантаження моделі та scaler

Приклад оцінки (рис. 3.15) демонструє прогноз 22,1 % придатності для заданих параметрів, що відповідає підвищеному ризику. Значення невизначеності 0.041 свідчить про впевнений прогноз моделі.

Таблиця 3.5 – Результати отриманих із СППР розрахованих показників

Показник	Значення	Інтерпретація
Ймовірність придатності	22.1 %	Високий ризик
Невизначеність ($1/\Sigma\alpha$)	0.041	Прогноз впевнений
Висновок	Вода потенційно непридатна для споживання	

Таким чином, інтеграція моделі DAG-EU у СППР забезпечує користувачів швидким, інформативним і наочним інструментом для оперативного контролю якості води. Система може бути розгорнута локально або в хмарному середовищі (Heroku, Streamlit Cloud, ngrok) і легко адаптується до інших типів водойм або датасетів екологічного моніторингу.

РОЗДІЛ 4.

ОХОРОНА ПРАЦІ ТА БЕЗПЕКА У НАДЗВИЧАЙНИХ СИТУАЦІЯХ

4.1. Аналіз умов праці під час розробки системи підтримки прийняття рішень

Розроблення системи підтримки прийняття рішень (СППР) для оцінювання якості води проводиться в умовах офісного середовища з використанням персональних комп'ютерів, периферійного обладнання, засобів зв'язку та мережевого доступу до хмарних ресурсів. Характер праці відноситься до категорії розумової, аналітичної, що супроводжується підвищеним навантаженням на зоровий апарат, нервову систему та опорно-руховий апарат. Тому важливим є дотримання вимог нормативних документів у сфері охорони праці та ергономіки робочого місця.

На етапі проектування та програмування інженери, науковці і студенти використовують комп'ютери не менше 6–8 годин на добу, що потребує спеціальної організації робочого простору. Відповідно до ДСанПіН 3.3.2-007-98 «Державні санітарні норми і правила роботи з відеодисплейними терміналами», монітор повинен розташовуватись на відстані 50–70 см від очей користувача, з кутом огляду 10–20° нижче горизонтальної лінії зору. Оптимальна освітленість робочої поверхні має становити 300–500 лк, що досягається поєднанням природного і штучного освітлення з використанням світильників типу ЛПО або LED-панелей із коефіцієнтом пульсації не більше 5%.

Для забезпечення належного мікроклімату у приміщенні розробників СППР встановлюється система кондиціонування, яка підтримує сталу температуру незалежно від пори року. Повітря обов'язково повинно оновлюватися через вентиляційні системи не рідше ніж тричі на годину, що відповідає вимогам ДБН В.2.5-67:2013 «Опалення, вентиляція та кондиціонування».

Таблиця 4.1 – Основні параметри умов праці розробників СППР

Показник	Нормативне значення	Фактичне значення	Відповідність
Освітленість, лк	300–500	420	Відповідає
Відстань до екрана, см	50–70	60	Відповідає
Температура повітря, °C	19–23 (зима), 22–25 (літо)	22	Відповідає
Відносна вологість, %	40–60	45	Відповідає
Рівень шуму, дБ	≤ 50	45	Відповідає
Робота за ПК без перерви, хв	≤ 60	45	Відповідає

Значну увагу приділяють організації робочих місць. Стіл і стілець повинні регулюватися по висоті, що забезпечує правильне положення спини та рук під час роботи з клавіатурою. Поверхня столу має бути матовою, без блиску, а стілець – із спинкою анатомічної форми та підлокітниками.

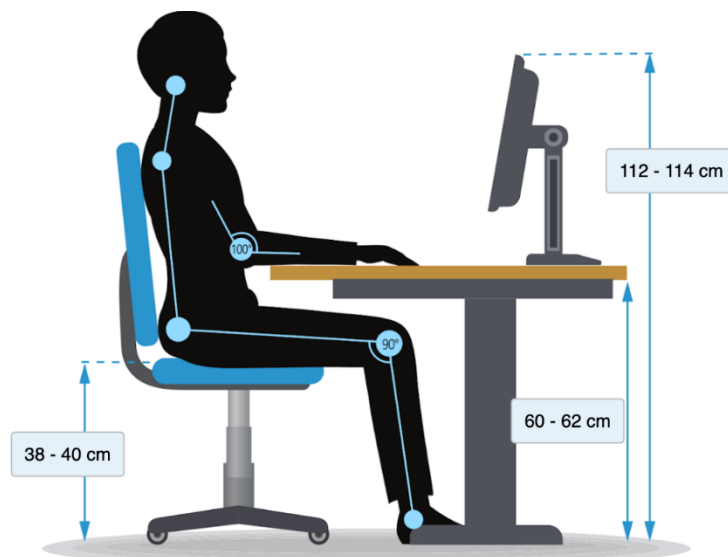


Рисунок 4.1 – Ергономічна схема робочого місця розробника СППР

4.2. Організація системи охорони праці при розробці програмного забезпечення

Під час роботи з електронно-обчислювальною технікою працівники зазнають впливу електромагнітного випромінювання в діапазоні частот від 5 Гц до 400 кГц, тому застосовуються монітори із сертифікатом ТСО та СЕ, які відповідають вимогам ДСТУ EN 60950-1:2015 щодо безпеки інформаційних технологій. Для зменшення впливу електростатичних полів застосовуються заземлені корпуси обладнання та підлогові покриття з антистатичних матеріалів.

З метою профілактики професійних захворювань передбачено регламентовані перерви тривалістю 10–15 хвилин після кожної години роботи за комп'ютером. Протягом перерв рекомендується виконувати вправи для очей і легку розминку, що знижує втому м'язів спини та шиї. Важливим аспектом є психологічна безпека праці, тому середовище розробників організовується з урахуванням принципів колаборації, відсутності стресових факторів і належного рівня комунікацій.

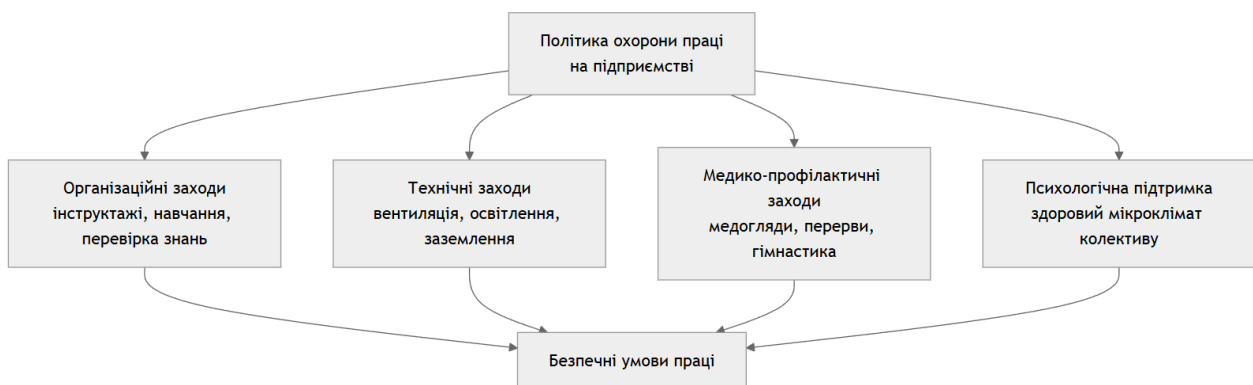


Рисунок 4.2 – Схема організації системи охорони праці при розробці програмного забезпечення

Отже, проведений аналіз свідчить, що умови праці розробників системи підтримки прийняття рішень цілком відповідають вимогам чинного законодавства України у сфері охорони праці та ергономіки. Забезпечення належного мікроклімату, освітлення, психологічного комфорту й профілактичних заходів створює безпечне середовище для ефективної

інтелектуальної діяльності, що є передумовою високої якості результатів розробки СППР.

4.3. Захист населення під час надзвичайних ситуацій

Захист населення в умовах надзвичайних ситуацій є одним із найважливіших завдань державної системи цивільного захисту. Основною метою цього процесу є збереження життя і здоров'я людей, мінімізація наслідків впливу небезпечних факторів, а також забезпечення сталого функціонування критичної інфраструктури та об'єктів життєзабезпечення. У сучасних умовах, коли на території України існують як техногенні, так і природні загрози, питання ефективної організації системи реагування набуває особливого значення.

Захист населення здійснюється на основі принципів пріоритетності безпеки людини, завчасного планування, комплексного використання сил і засобів, а також обов'язкової взаємодії між органами влади, підприємствами, рятувальними службами та громадськістю. На кожному підприємстві, зокрема в установах, де ведуться наукові дослідження або розробка інформаційних систем, має бути розроблений План дій персоналу у разі виникнення надзвичайної ситуації, що визначає алгоритм евакуації, порядок оповіщення та місця укриття працівників.

Умовно надзвичайні ситуації поділяються на природні, техногенні, соціальні та воєнні. Для кожного типу визначено характерні ознаки, масштаби ураження та відповідні заходи реагування. У наукових і навчальних установах найбільш імовірними є техногенні надзвичайні ситуації, пов'язані з пожежами, витоками шкідливих речовин, порушенням електропостачання чи аваріями у мережах. Тому на об'єктах необхідно забезпечити справність електромереж, наявність первинних засобів пожежогасіння, системи пожежної сигналізації та резервних джерел живлення.

Важливим компонентом системи цивільного захисту є оповіщення населення. На об'єктах використовується комбінована система – автоматизовані сигнали тривоги (сирени, гучномовці) доповнюються повідомленнями через мобільні додатки, месенджери або електронну пошту. Персонал має бути ознайомлений із сигналами «Увага всім!» і знати порядок дій після їх отримання. При отриманні сигналу слід негайно припинити роботу, знеструмити обладнання, взяти документи, засоби індивідуального захисту та направитись у визначене укриття.

Для ефективного реагування на надзвичайні ситуації розробляються плани евакуації, які містять схеми руху до безпечних виходів, позначення місць розташування вогнегасників, аптечок та засобів пожежогасіння. Такі схеми вивішуються на видимих місцях біля кожного виходу. Персонал періодично проходить інструктаж із питань цивільного захисту, а також навчальні тренування, які відпрацьовують порядок дій у разі пожежі, землетрусу, хімічного забруднення чи ракетної небезпеки.

Окремої уваги заслуговує питання укриття населення. У разі загрози бойових дій або вибухів працівники повинні прямувати до найпростіших укриттів або захисних споруд – підвальних приміщень, обладнаних вентиляцією, запасом води, медикаментів та засобів гігієни. Рівень підготовки укриття визначається за критеріями герметичності, наявності аварійного виходу та відповідності санітарним вимогам.

Реалізація комплексу заходів із попередження та ліквідації наслідків надзвичайних ситуацій є запорукою збереження життя людей, зменшення матеріальних збитків та забезпечення стабільного функціонування суспільства. Розроблення інформаційних систем, таких як СППР, може істотно підвищити ефективність реагування завдяки автоматизованому аналізу ризиків, прогнозуванню загроз і підтримці прийняття рішень у кризових умовах.

РОЗДІЛ 5.

ЕКОНОМІЧНА ЕФЕКТИВНІСТЬ ВІД ВИКОРИСТАННЯ ІНФОРМАЦІЙНОЇ СИСТЕМИ ОЦІНЮВАННЯ ЯКОСТІ ВОДИ

Економічну доцільність упровадження системи підтримки прийняття рішень для оцінювання якості води доцільно оцінювати за простою методикою на базі грошових потоків. Базовими показниками є: сукупні інвестиційні витрати на старті, річні операційні витрати, річний економічний ефект від використання системи та дисконтовані показники ефективності. Методика ґрунтується на розрахунку чистого річного ефекту, строку окупності, чистої приведеної вартості, індексу прибутковості та внутрішньої норми дохідності.

Початкові інвестиції позначаються через C_0 і включають витрати на розроблення програмного забезпечення, налаштування інфраструктури, закупівлю обладнання та первинне навчання персоналу. Щорічні експлуатаційні витрати O охоплюють підтримку програмного забезпечення, хмарні сервіси або серверні ресурси, калібрування датчиків і часткове оновлення комплектуючих. Річний економічний ефект S формується за рахунок скорочення витрат на лабораторні аналізи та виїзні заміри, зменшення штрафних санкцій завдяки ранньому попередженню відхилень, економії часу персоналу та оптимізації дозування реагентів. Чистий річний ефект:

$$E = S - O. \quad (5.1)$$

Простий строк окупності визначається як:

$$T_{ок} = \frac{C_0}{E}. \quad (5.2)$$

Для урахування вартості грошей у часі використовується ставка дисконту r . Чиста приведена вартість за період n років обчислюється за формулою:

$$NPV = \sum_{t=1}^n \frac{E_t}{(1+r)^t} - C_0. \quad (5.3)$$

де E_t – чистий ефект у році t .

Індекс прибутковості:

$$PI = \frac{\sum_{t=1}^n \frac{E_t}{(1+r)^t}}{C_0}. \quad (5.4)$$

Внутрішня норма дохідності *IRR* визначається як значення дисконту r^* , за якого $NPV(r^*)=0$. Для простої оцінки її можна наближено окреслити за методом підбору між двома ставками.

Для прикладу наведено розрахунок для комунального підприємства, яке експлуатує 3 озера з регулярним моніторингом. Прийmemo гіпотези: $C_0=160000$ грн (розроблення – 120000; обладнання – 30000; навчання – 10000), щорічні витрати $O=36000$ грн, очікуваний річний ефект $S=145000$ грн, що складається з економії на лабораторних аналізах 80000 грн, уникнення штрафів 30000 грн, економії робочого часу й логістики 20000 грн, оптимізації реагентів 15000 грн. У результаті $E=109000$ грн на рік. Строк аналізу – 5 років, ставка дисконту $r=15\%$.

За цих припущень строк окупності становить:

$$T_{ок} = 1,47 \text{ років.}$$

Дисконтований грошовий потік для п'ятирічного горизонту дає:

$$NPV \approx 274000 \text{ грн.}$$

Індекс прибутковості

$$PI \approx 1,71.$$

Наближена *IRR* за інтерполяцією між 30% і 40% перевищує 35%, що істотно вище прийнятої ставки дисконту і свідчить про інвестиційну привабливість проекту.

Таблиця 5.1 подає зведений приклад розрахунків. Показники «Ефект» і «Чистий ефект» для років 1–5 є валовими й дисконтованими відповідно до формул, наведених вище, із застосуванням ставки $r=15\%$.

Отже, результати розрахунків показали високу економічну ефективність упровадження розробленої інформаційної системи оцінювання якості води. За початкових інвестицій 160000 грн і щорічних витрат на обслуговування 36000

грн, очікуваний річний економічний ефект становить 145000 грн, що забезпечує чистий річний прибуток у 109000 грн.

Таблиця 5.1 – Розрахунок економічної ефективності впровадження СППР

Показник	Одиниця виміру	Значення
Початкові інвестиції C_0	грн	160 000
Річні витрати експлуатації O	грн/рік	36 000
Річний економічний ефект S	грн/рік	145 000
Чистий річний ефект E	грн/рік	109 000
Строк окупності $T_{ок}$	років	1,47
Ставка дисконту r	—	15%
Горизонт розрахунку n	—	5 років
NPV	грн	$\approx 274\,000$
PI	—	$\approx 1,71$
Наближена IRR	—	$> 35\%$

Розрахований строк окупності системи дорівнює 1,47 року, що свідчить про швидке повернення вкладених коштів. Дисконтована чиста приведена вартість (NPV) становить близько 274000 грн, а індекс прибутковості (PI) дорівнює 1,71, тобто на кожну вкладену гривню отримується 1,71 грн прибутку. Внутрішня норма дохідності (IRR) перевищує 35 %, що значно вище за базову ставку дисконту 15 %.

Таким чином, проєкт можна вважати економічно доцільним, фінансово привабливим і стійким до ризиків. Реалізація системи дозволяє не лише знизити витрати на лабораторний контроль і технічне обслуговування, а й забезпечує своєчасне виявлення відхилень у параметрах води, що підвищує екологічну безпеку та ефективність управлінських рішень у сфері водокористування.

ВИСНОВКИ І ПРОПОЗИЦІЇ

На основі проведеного аналізу стану предметної галузі встановлено, що проблема полягає в тому, що якість води в озерах піддається багатьом впливам, а існуючі методи її оцінювання мають значні обмеження. Завданням дослідження є розробка та апробація підходу, який базується на методах обчислювального інтелекту для підвищення ефективності, точності та оперативної здатності оцінювання якості озерної води.

Аналіз сучасного стану наукових досліджень засвідчує, що найбільш перспективним напрямом у сфері оцінювання якості води є застосування гібридних моделей обчислювального інтелекту. Вони поєднують можливості глибоких нейронних мереж, еволюційних обчислень та методів оцінки невизначеності, що дозволяє створювати адаптивні, точні й надійні системи моніторингу стану водних екосистем.

Узагальнення міжнародного досвіду досліджень з оцінювання якості озерної води подано у вигляді порівняльної таблиці 1.3, яка відображає тип водного об'єкта, використані дані, методи обчислювального інтелекту та ключові результати. Встановлено, що спостерігається зростання уваги до створення практично придатних інформаційних систем для екологічних агентств та органів управління водними ресурсами.

Встановлено, що обчислювальний інтелект є ефективним інструментом для вирішення задач екологічного моделювання, забезпечуючи поєднання гнучкості, адаптивності та інтерпретованості моделей. Його використання у сфері оцінювання якості води дозволяє формувати прогностичні системи нового покоління, які здатні працювати в умовах невизначеності, забезпечувати швидке оновлення даних та підвищувати достовірність екологічного моніторингу.

Запропоновані математичні моделі оцінювання якості води на основі обчислювального інтелекту забезпечують можливість автоматичного аналізу великих наборів екологічних даних, формування прогнозів та підтримку прийняття управлінських рішень.

Глибоке й активне навчання моделей формують теоретичну та алгоритмічну основу інтелектуального прогнозування якості води. Використання цих підходів у поєднанні з еволюційними механізмами оптимізації параметрів створює умови для розроблення універсальних моделей, здатних забезпечувати точну, адаптивну й надійну оцінку стану природних водойм у контексті сталого розвитку.

Нами здійснено вибір засобів для виконання роботи. Виконано порівняльну характеристику основних програмних засобів, які використовуються для реалізації проєкту. Комплексна інтеграція зазначених засобів створює гнучку платформу для побудови інтелектуальної системи аналізу якості води.

Джерелом даних є відкритий набір Water Potability (3276 спостережень, 9 континуальних показників і бінарна мітка Potability), що містить лабораторні вимірювання якості прісної води. У вибірці представлені рН, жорсткість, концентрації розчинених твердих речовин, хлорамінів, сульфатів, електропровідність, органічний вуглець, тригалометани та каламутність. Цільова змінна набуває значень 1 для зразків, що відповідають критерію придатності до пиття, й 0 – інакше.

Встановлено, що існує дисбаланс класів – частка непридатної води перевищує частку придатної (приблизно 61% проти 39%), що потрібно враховувати при виборі метрик і процедур навчання. З огляду на дисбаланс, базовою метою оцінювання надалі прийнято не лише точність, а й F1-міру позитивного класу, збалансовану точність та AUC-ROC.

Нами зібрано й уніфіковано 5 базових архітектур, що репрезентують найуживаніші підходи до інтелектуального аналізу якості води за табличними та змішаними (лабораторними + супутниковими) ознаками. Встановлено, що CNN стабільно збігається і перевищує LSTM/GRU на валідації, що логічно для статичних лабораторних вимірів. Autoencoder+Classifier демонструє плавну збіжність і близьку до CNN узгодженість валідованої точності, при цьому забезпечує компактне латентне подання, корисне для подальшої візуалізації або

кластеризації. DAL виграє в економності розмітки: вже через 4-5 циклів середня валідована втрата знижується, а $ROC-AUC \approx 0.63$ підтверджує корисність стратегії невпевненого відбору. DAG-EU додає критично важливий для екологічного моніторингу аспект – калібровану невизначеність: через $1/S$ можна відсіювати прогнози низької впевненості або ініціювати «людину в циклі» для повторного відбору проб.

Найвищі результати досягла DAG-EU (спрощена), що поєднує гнучкість нейронної моделі з можливістю інтерпретувати рівень невпевненості. Її точність 0.666 і $AUC = 0.666$ засвідчують найкращу узгодженість між прогнозом і реальними мітками. Обґрунтована модель DAG-EU стала основою для побудови інтелектуальної системи моніторингу якості води, яка не лише прогнозує клас придатності, а й оцінює ступінь впевненості прогнозу.

Розроблена інтелектуальна модель була інтегрована у веборієнтовану систему підтримки прийняття рішень (СППР) для оцінювання якості води в озерах. Система дозволяє виконувати швидку діагностику придатності води за дев'ятьма ключовими показниками та отримувати висновок щодо її стану разом із рівнем невизначеності прогнозу. Навчена нейронна модель DAG-EU (спрощена) була збережена у форматі .h5, а стандартизатор ознак – у файлі scaler.pkl. У клієнтській частині застосовано HTML5, CSS3 та бібліотеку Plotly для інтерактивних графіків. Завдяки API Streamlit система дозволяє обчислювати значення ймовірності придатності води безпосередньо в браузері після введення користувачем вхідних параметрів.

Інтеграція моделі DAG-EU у СППР забезпечує користувачів швидким, інформативним і наочним інструментом для оперативного контролю якості води. Система може бути розгорнута локально або в хмарному середовищі (Heroku, Streamlit Cloud, ngrok) і легко адаптується до інших типів водойм або датасетів екологічного моніторингу.

Під час розроблення системи підтримки прийняття рішень з оцінювання якості води створено безпечні умови праці розробників шляхом дотримання вимог охорони праці, санітарно-гігієнічних норм та правил технічної безпеки.

Результати розрахунків показали високу економічну ефективність упровадження розробленої інформаційної системи оцінювання якості води. За початкових інвестицій 160000 грн і щорічних витрат на обслуговування 36000 грн, очікуваний річний економічний ефект становить 145000 грн, що забезпечує чистий річний прибуток у 109000 грн. Розрахований строк окупності системи дорівнює 1,47 року, що свідчить про швидке повернення вкладених коштів.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Боднарчук О.В., Андрієнко А.В. Методи та засоби побудови інтелектуальних інформаційних систем. Київ: КНУ ім. Тараса Шевченка, 2021. 312 с.
2. Гандзюра, В. П. Якість води, виклики та рішення / В. П. Гандзюра // Екологічні науки. – 2023. – № 4. – С. 49-57.
3. Жидецький В.Ц., Джигирей В.С., Мельников О.В. Основи охорони праці. Підручник. Вид. 5-е, доповнене. Львів: Афіша, 2012. 350с.
4. Класифікація в Python з Scikit-Learn та Pandas. URL: <https://stackabuse.com/classification-in-python-with-scikit-learn-and-pandas/> (дата звернення: 14.09.2025).
5. Комплексна оцінка якості вод за різними методиками та шляхи її вдосконалення / [колектив авторів] // Український гідрометеорологічний журнал. – 2022. – № 3. – С. 23-34.
6. Кравець П.Ю., Кудінов Є.А. Інтелектуальні системи обробки інформації: теорія та практичні аспекти. Львів: Видавництво ЛНУ ім. І. Франка, 2020. 286 с.
7. Лехман С.Д., Рублев В.І., Рябцев Б.І. Запобігання аварійності і травматизму у сільському господарстві. К.: Урожай, 1993. 267 с.
8. Мокієнко, А. В. Актуальні водні проблеми і інноваційні рішення: конспект лекцій / А. В. Мокієнко. – Вінниця : Видавництво ВНТУ, 2024. – 378 с.
9. Ситнік Б.Т. Основи інформаційних систем і технологій: навч. посіб. Харків: УкрДУЗТ, 2018. 130 с.
10. Тригуба А., Маланчук О., Тригуба І., Мармуляк А., Демчина В. Андрушків О., Олійник Р. Вплив сучасних інформаційних технологій на процеси ініціації та планування проєктів розвитку громад та регіонів. Вісник Львівського національного університету природокористування: агроінженерні дослідження. №24. Львів: Львів НУП, 2024. С. 123-130.

11. Шмундяк, Д. О. Інформаційна технологія пошуку можливих джерел підвищеного забруднення річки з використанням моделі Prophet / Д. О. Шмундяк, О. М. Давидюк, О. В. Слободянюк // Вісник Вінницького політехнічного інституту. – 2020. – № 4. – С. 15-24.
12. Aranay O. M. Deep Active Genetic Learning with Evidential Uncertainty for Agriculture Crops and Lake Water Quality Assessment: PhD Thesis. University at Albany, State University of New York, 2022.
13. Bezdek J. C., Pal S. K. Computational Intelligence: Principles, Techniques and Applications. Oxford University Press, 2003.
14. Chollet F. Deep Learning with Python. 2nd ed. Manning Publications, 2021.
15. Choubin B., Moradi E., Golshan M., Sajedi-Hosseini F., Malekian A. Application of Machine Learning Models for Water Quality Prediction: A Case Study of Karaj River, Iran. Science of the Total Environment, 2019, 653, 1040–1052.
16. Engelbrecht A. P. Computational Intelligence: An Introduction. 3rd ed. Wiley, 2022.
17. Frincu R. M. et al. Artificial Intelligence in Water Quality Monitoring: A Review of Water Quality Assessment Applications. Water Quality Research Journal. 2025. Vol. 60, No. 1.
18. Kaggle dataset «Water Potability». URL: <https://www.kaggle.com/datasets/adityakadiwal/water-potability>
19. Kazemi Garajeh M., Akbari R., Aghaei Chaleshtori S., Shenavaei Abbasi M., Tramutoli V., Lim S., Sadeqi A. A Comprehensive Assessment of Climate Change and Anthropogenic Effects on Surface Water Resources in the Lake Urmia Basin, Iran. Remote Sensing. 2024. Vol. 16, No. 11. Article 1960. DOI: 10.3390/rs16111960
20. Kumar M., Singh R., Sharma P. Application of Artificial Intelligence for Water Quality Assessment and Forecasting: A Review. Environmental Monitoring and Assessment, 2023, 195(3), 482.
21. Kundu S. et al. Using Explainable Artificial Intelligence to Assess Water Quality. Journal of Cleaner Production. 2025. DOI: 10.1016/j.jclepro.2025.145211

22. Li M., Zhang Y., Zhang L. Hybrid Neural-Fuzzy Systems for Environmental Quality Prediction. *Applied Soft Computing*, 2021, 106, 107336.
23. McCulloch W. S., Pitts W. A Logical Calculus of the Ideas Immanent in Nervous Activity. *Bulletin of Mathematical Biophysics*, 1943, 5(4), 115–133.
24. Mohan S. et al. Integration of Machine Learning and Remote Sensing for Water Quality Monitoring. *Sustainability*. 2025. Vol. 17, No. 3. DOI: 10.3390/su17030998
25. Myers D. N. Why Monitor Water Quality? U.S. Geological Survey, Office of Water Quality, 2015. 4 p. Available at: <https://water.usgs.gov/owq/WhyMonitorWaterQuality.pdf>
26. Palabıyık S. Evaluation of Water Quality Based on Artificial Intelligence: Performance of Multilayer Perceptron Neural Networks and Multiple Linear Regression. *Environment, Development and Sustainability*. 2024. DOI: 10.1007/s10668-024-05075-6
27. Schmidhuber J. Deep Learning in Neural Networks: An Overview. *Neural Networks*, 2015, 61, 85–117.
28. Settles B. Active Learning Literature Survey. University of Wisconsin–Madison, 2012.
29. Tryhuba A., Boyarchuk V., Tryhuba I., Ftoma O., Padyuka R., Rudynets M. Forecasting the Risk of the Resource Demand for Dairy Farms Basing on Machine Learning. *Proceedings of the 2nd International Workshop on Modern Machine Learning Technologies and Data Science (MoMLeT+DS 2020)*, Lviv-Shatsk, Ukraine, 02–03 June 2020, vol. I, pp. 327-340.
30. Tryhuba A., Koval N., Shevchuk V., Tryhuba I. Algorithm for the routes formation of food raw materials procurement on the community territory taking into account the production conditions during emergency situations. *Herald of Advanced Information Technology*, vol. 6, no. 1, 2023, pp. 60-73.
31. Tryhuba A., Malanchuk O., Pankiv O., Sholudko R.Ya. Architecture of an intelligent information system for forecasting components of medical projects.

Advanced Information Technology for Social Innovation (AAIT), vol. 6, no. 4, 2023, pp. 376-390.

32. Tryhuba A., Malanchuk O., Sholudko R.Ya. Development of a geoinformation method for assessing the accessibility of medical services in community infrastructure development projects. Herald of Advanced Information Technology, vol. 8, no. 3, 2025, pp. 382-404.

33. Tryhuba A., Malanchuk O., Tryhuba I., Sholudko R.Ya. Моделі адаптивно-ціннісного управління проєктами функціонування та розвитку госпітальних округів. Інформаційні технології, no. 4(30), 2024, DOI:10.30837/2522-9818.2024.4.097.

34. Tryhuba A., Ratushny R., Bashynsky O., Shcherbachenko O. Identification of firefighting system configuration of rural settlements. MATEC Web of Conferences, vol. 247, 2018, pp. 00027. DOI:10.1051/matecconf/201824700027.

35. Tryhuba A., Ratushny R., Tryhuba I., Koval N., Androshchuk I. The model of projects creation of the fire extinguishing systems in community territories. Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis, vol. 68, no. 2, 2020, pp. 419-431.

36. Tryhuba A.M., Malanchuk O.M., Sholudko R.Ya. Computer model of differential-symbolic risk assessment of projects to improve the health of the region's population. Herald of Advanced Information Technology, vol. 7, no. 4, 2024, pp. 437-451.

37. VanderPlas J. Python Data Science Handbook. O'Reilly Media, 2023.

ДОДАТКИ

Додаток А

Фрагмент коду СППР для оцінювання якості води в озерах

```
!pip install streamlit joblib tensorflow
import streamlit as st
import numpy as np
import joblib
import tensorflow as tf
from tensorflow.keras.models import load_model

st.set_page_config(
    page_title="СППР: Якість води в озерах",
    page_icon="🌊",
    layout="centered"
)

st.markdown("""
<style>
body {
    background: radial-gradient(circle at top, #e3f2fd 0, #f5f5f5 45%, #ffffff 100%);
    font-family: "Segoe UI", system-ui, -apple-system, sans-serif;
}
.main-block {
    background: #ffffff;
    padding: 24px 28px;
    border-radius: 24px;
    box-shadow: 0 10px 35px rgba(15, 76, 129, 0.18);
    border: 1px solid #e3f2fd;
}
.title {
    font-weight: 700;
    font-size: 26px;
    color: #0d47a1;
    margin-bottom: 4px;
}
.subtitle {
    font-size: 13px;
    color: #607d8b;
    margin-bottom: 16px;
}
.label {
    font-size: 14px;
    font-weight: 500;
    color: #0d47a1;
}
.result-box {
    margin-top: 18px;
    padding: 18px;
    border-radius: 18px;
    border: 1px solid #e3f2fd;
    background: #f8fbff;
}
```

```

}
.bad {
    background: #ffebee;
    border-color: #ffcdd2;
}
.good {
    background: #e8f5e9;
    border-color: #c8e6c9;
}
.medium {
    background: #fff8e1;
    border-color: #ffecb3;
}
.prob-label {
    font-size: 32px;
    font-weight: 700;
}
.helper {
    font-size: 12px;
    color: #78909c;
}
.footer-note {
    margin-top: 26px;
    font-size: 11px;
    color: #9e9e9e;
    text-align: center;
}
</style>
"""', unsafe_allow_html=True)

```

```

@st.cache_resource
def load_model_and_scaler():
    # Якщо у тебе DAG-EU з кастомним лоссом – компілюємо без нього
    try:
        model = load_model("best_model.h5", compile=False)
    except Exception as e:
        st.error(f'Помилка завантаження моделі: {e}')
        model = None

    try:
        scaler = joblib.load("scaler.pkl")
    except Exception as e:
        st.warning("Не вдалося завантажити scaler.pkl. "
                  "Переконайтеся, що зберегли StandardScaler після навчання.")
        scaler = None

    return model, scaler

model, scaler = load_model_and_scaler()
....

```

```

if st.button("Оцінити якість води"):
    # Формуємо вектор ознак у правильному порядку
    features = np.array([[ph, hardness, solids, chloramines,
                          sulfate, conductivity, organic_carbon,
                          trihalo, turbidity]])

    # Масштабуємо, якщо є scaler
    if scaler is not None:
        features_scaled = scaler.transform(features)
    else:
        features_scaled = features # fallback, якщо scaler не завантажено

    # Прогноз моделі
    try:
        y_pred = model.predict(features_scaled, verbose=0)

        # Якщо це DAG-EU (2-вимірний вихід alpha)
        if y_pred.shape[1] == 2:
            alpha = y_pred
            S = np.sum(alpha, axis=1, keepdims=True)
            p = alpha / S
            prob_safe = float(p[:, 1])
            uncertainty = float(1.0 / S[:, 0])
        else:
            # Звичайна сігмоїда
            prob_safe = float(y_pred.ravel()[0])
            uncertainty = None

        # Інтерпретація
        if prob_safe >= 0.75:
            css_class = "result-box good"
            label = "Ймовірно безпечна"
            color = "□"
        elif prob_safe >= 0.45:
            css_class = "result-box medium"
            label = "Прикордонний стан"
            color = "□"
        else:
            css_class = "result-box bad"
            label = "Підвищений ризик"
            color = "---"

        st.markdown(f'<div class="{css_class}">', unsafe_allow_html=True)
        st.markdown(
            f'<div class="prob-label">{color} {prob_safe*100:.1f}%</div>',
            unsafe_allow_html=True
        )
        st.markdown(
            f'<b>Висновок:</b> {label} з точки зору моделі придатності води.',
            unsafe_allow_html=True
        )

```

```

if uncertainty is not None:
    st.markdown(
        f"<br><b>Оцінка невизначеності ( $1/\Sigma\alpha$ ):</b> {uncertainty:.3f} "
        f"<span class='helper'>(менше значення – більш впевнений прогноз)</span>",
        unsafe_allow_html=True
    )

# Коротка текстова рекомендація
if prob_safe < 0.45:
    st.markdown(
        "<ul>"
        "<li>Рекомендується додаткова лабораторна перевірка.</li>"
        "<li>Можливі перевищення за органічними речовинами, солями чи продуктами
дезінфекції.</li>"
        "</ul>",
        unsafe_allow_html=True
    )
elif prob_safe < 0.75:
    st.markdown(
        "<ul>"
        "<li>Показники поблизу граничних значень стандартів.</li>"
        "<li>Доцільно проводити регулярний моніторинг та перевірку сезонних змін.</li>"
        "</ul>",
        unsafe_allow_html=True
    )
else:
    st.markdown(
        "<ul>"
        "<li>Поточні показники узгоджуються з умовно безпечним станом.</li>"
        "<li>Рекомендується плановий моніторинг без термінових заходів.</li>"
        "</ul>",
        unsafe_allow_html=True
    )

st.markdown("</div>", unsafe_allow_html=True)

except Exception as e:
    st.error(f"Помилка під час прогнозу: {e}")

st.markdown(
    '<div class="footer-note">'
    'Система підтримки прийняття рішень для: '
    'моделювання якості води з використанням методів обчислювального інтелекту.'
    '</div>',
    unsafe_allow_html=True
)

st.markdown("</div>", unsafe_allow_html=True)

```