

Б. П. Книш, канд. техн. наук, доц.; Я. А. Кулик, канд. техн. наук, доц.

РОЗРОБКА МОДЕЛІ З ВИКОРИСТАННЯМ РЕГРЕСІЇ ОПОРНИХ ВЕКТОРІВ ДЛЯ СИСТЕМИ ВИМІРЮВАННЯ КОНЦЕНТРАЦІЇ ВУГЛЕКИСЛОГО ГАЗУ НА ОСНОВІ БЕЗДРОТОВОГО ОПТИКО-ЕЛЕКТРОННОГО СЕНСОРА

В роботі запропоновано модель з використанням регресії опорних векторів для системи вимірювання концентрації вуглекислого газу на основі бездротового оптико-електронного сенсора. В середовищі, що може бути небезпечним для людини, розташовується сенсор, який дистанційно передає дані про концентрацію газу на комп'ютер, де за допомогою машинного навчання проводиться обробка отриманої інформації. В якості такого сенсора запропоновано використати оптико-електронний сенсор концентрації вуглекислого газу завдяки високій точності, чутливості, стабільності та надійності, а також швидкій реакції, стійкості до завад, безконтактності вимірювань, мінімальному технічному обслуговуванню та можливості дистанційного моніторингу. Також в роботі запропоновано використовувати модель регресії опорних векторів для обробки даних бездротового оптико-електронного сенсора концентрації вуглекислого газу. Регресія опорних векторів добре працює з нелінійними даними та, порівняно з іншими методами, демонструє високу точність при малій та середній кількості вхідних даних. Такі системи можуть виявляти вуглекислий газ завдяки вимірюванню його концентрації з високою точністю внаслідок збору даних в режимі реального часу оптико-електронним сенсором, що в поєднанні з обробкою зібраних даних за допомогою машинного навчання демонструє великий потенціал у вимірюванні концентрації вуглекислого газу, а, отже, і раннього попередження про його появу. Основними показниками ефективності запропонованої моделі вимірювання концентрації вуглекислого газу з використанням регресії опорних векторів, що визначалися в процесі дослідження, було обрано такі характеристики як середньоквадратична похибка (MSE), середня абсолютна похибка (MAE) та коефіцієнт детермінації (R^2). Для аналізу цих показників була проведена оцінка зміни MSE, MAE та точності за метрикою R^2 під час навчання та перевірки моделі, порівняння даних з оптичного сенсора та з моделі, а також оцінки зміни MSE та MAE в процесі вимірювання.

Ключові слова: вуглекислий газ, оптико-електронний сенсор, концентрація, машинне навчання, регресія опорних векторів.

Вступ

На сьогодні в різних галузях техніки присутні різноманітні методи дослідження та аналізу газів, що дозволяють вимірювати та контролювати їх фізичні параметри, зокрема концентрацію, густину, кількісний склад, вміст домішок тощо. Важливим параметром серед них є концентрація.

На основі різних методів контролю концентрації газу розроблено широку класифікацію різноманітних сенсорів, датчиків, газоаналізаторів і систем, які дозволяють проводити вимірювання концентрації тих чи інших газів та їх сполук в різних середовищах.

Одним з таких газів, виявлення та контроль якого має надзвичайно важливе значення, є вуглекислий газ. Цей газ не є ні токсичним, ні вибухонебезпечним, але його основною небезпекою залишаються його задушливі властивості. Оскільки він важчий за повітря, то замінює кисень у замкнутих або погано вентильованих приміщеннях. Крім того, його неможливо виявити без відповідного обладнання, що додатково підвищує його небезпеку. За концентрації від 7 % до 10 % може настати смерть від асфіксії, навіть якщо присутня достатня кількість кисню [1].

Таким чином, підвищення точності контролю та швидкість реагування на появу вуглекислого газу є надзвичайно важливою задачею. Проблема її вирішення сприяла

розвитку ефективних систем вимірювання концентрації та раннього попередження про появу вуглекислого газу. Одним з варіантів вирішення цієї задачі є системи, що поєднують сенсори з нейронними мережами або алгоритмами машинного навчання. Це дає змогу здійснювати інтелектуальне та надійне вимірювання концентрації вуглекислого газу шляхом вимірювання його концентрації в динамічних умовах.

Нейронні мережі працюють краще тоді, коли є набір спектрограм або масивів даних з декількох сенсорів. Але для аналізу фізичних процесів вони є досить складними через чутливість до різноманітних шумів і перешкод, зміщення точки нуля, сенсорну деградацію, нелінійну залежність тощо. На противагу нейронним мережам, алгоритми машинного навчання потребують менше часу на навчання, менш схильні до перенавчання, швидше та краще працюють за обмежених обсягів тренувальних даних (до кількох тисяч точок). Наприклад, найкращим рішенням для вимірювань з одного сенсора будуть методи машинного навчання, такі як RF або SVM [2, 3].

Існують різноманітні сенсори концентрації газу, зокрема каталітичні, електрохімічні, оптико-електронні тощо. Серед них виділяють оптико-електронні сенсори завдяки високій точності, чутливості, стабільності та надійності, а також швидкій реакції, стійкості до завад, безконтактності вимірювань, мінімальному технічному обслуговуванню та можливості дистанційного моніторингу. Ці характеристики корисні у разі застосування цих сенсорів для контролю отруйних газів, наприклад, вуглекислого газу, оскільки забезпечують мінімальний вплив людини протягом довгого часу. Це дає змогу вирішувати різноманітні задачі виявлення, моніторингу та раннього попередження про появу вуглекислого газу.

Таким чином, в середовищі, що може бути небезпечним для людини, розташовується сенсор, який дистанційно передає дані про концентрацію газу на комп'ютер, де за допомогою машинного навчання проводиться обробка отриманої інформації. Такі системи можуть виявляти вуглекислий газ завдяки вимірюванню його концентрації з високою точністю внаслідок збору даних в режимі реального часу оптико-електронним сенсором. В подальшому ці дані передаються на комп'ютер, де завдяки алгоритмам машинного навчання здійснюється їх обробка. Це демонструє великий потенціал у вимірюванні концентрації вуглекислого газу, вимірюванні його концентрації і, отже, раннього попередження про його появу.

Все це створює перспективну основу у випадку розгортання в реальних середовищах інтелектуальних систем вимірювання концентрації вуглекислого газу шляхом поєднання оптико-електронних сенсорів концентрації та методів машинного навчання.

Таким чином, **метою роботи** є розробка моделі з використанням машинного навчання для системи вимірювання концентрації вуглекислого газу на основі бездротового оптико-електронного сенсора та оцінка її точності.

Бездротовий оптико-електронний сенсор концентрації вуглекислого газу

Система вимірювання концентрації вуглекислого газу містить оптико-електронний сенсор концентрації вуглекислого газу в робочому середовищі, що передає дані про концентрацію вуглекислого газу на комп'ютер.

Було обрано оптичний сенсор концентрації газу [4], який завдяки конструктивним особливостям вирізняється високою точністю та чутливістю, оскільки дозволяє усунути впливи деяких фізичних параметрів, а саме молярних мас, питомих показників поглинання газів, густин, тисків, температур, товщини шару газу та розсіювальних властивостей матеріалів. Крім того, він не залежить від можливої нестабільності роботи й старіння фотоприймачів і світлодіодів. На рис. 1 наведено схему оптичного сенсора концентрації газу, що складається з трьох кювет – вимірювальної 1, в якій міститься газова суміш, що аналізується, кювети з чистим повітрям 2 та кювети з іншим чистим газом 3, наявність якого досліджується в газовій суміші.

Аналізована газова суміш прокачується через вхідний патрубок на вимірювальній кюветі 1. Інфрачервоне випромінювання від джерел інфрачервоного випромінювання 5, 6, 7, які керуються блоком керування 11, проходить спочатку через вхідні оптичні системи 17 для розсіювання світлових потоків, а потім через вимірювальну кювету 1, де частково поглинаються аналізованою газовою сумішшю. Вона тисне на діафрагми 4 кювети з чистим повітрям 2 та кювети з іншим чистим газом 3, пропорційно його концентрації. Крім того, інфрачервоне випромінювання проходить через кювету з чистим повітрям 2 і кювету з іншим чистим газом 3 та потрапляє на вихідні оптичні системи 17', які збирають світлові потоки. Після цього інфрачервоне випромінювання, яке пройшло через кювети 1, 2, 3, потрапляє на приймачі інфрачервоного випромінювання 8, 9, 10, на виході яких з'являються електричні сигнали, які пропорційні концентрації аналізованої газової суміші у вимірювальній кюветі 1, чистого повітря в кюветі 2, та іншого чистого газу в кюветі 3. Сигнали з приймачів інфрачервоного випромінювання 8, 9, 10 підсилюються за допомогою підсилювачів 12, об'єднуються за допомогою комутатора 13, надходять до аналого-цифрового перетворювача 14, де аналоговий сигнал перетворюється в цифровий, потрапляють через шину на вхід мікропроцесорного пристрою 15, який має обернений зв'язок з комутатором 13. Цифровий код, який пропорційний концентрації кожного з газів, обробляється в мікропроцесорному пристрої 15 і виводиться через шину на цифрове індикаторне табло 16 [5].

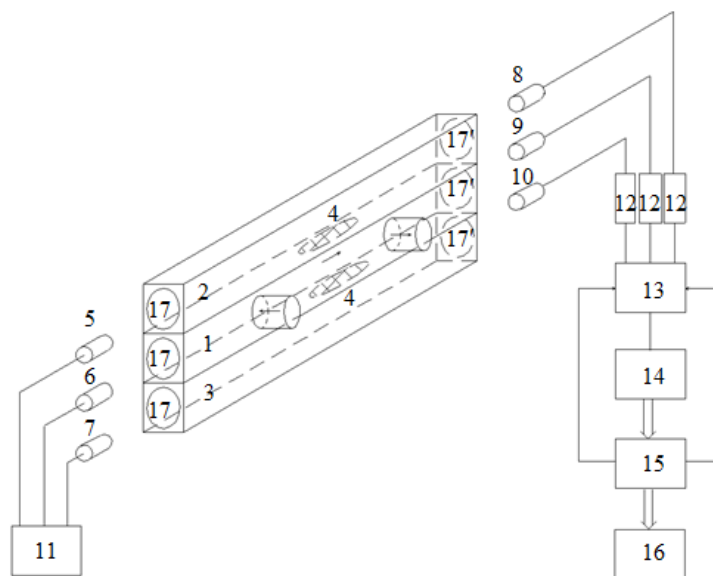


Рис. 1. Схема оптичного сенсора концентрації газу: 1 – вимірювальна кювета; 2 – кювета з чистим повітрям; 3 – кювета з іншим чистим газом; 4 – діафрагма; 5, 6, 7 – джерела інфрачервоного випромінювання; 8, 9, 10 – приймачі інфрачервоного випромінювання; 11 – блок керування; 12 – підсилювач; 13 – комутатор; 14 – аналого-цифровий перетворювач; 15 – мікропроцесорний пристрій; 16 – індикаторне табло; 17, 17' – вхідні і вихідні оптичні системи, відповідно

Аналізована газова суміш, що прокачується через вимірювальну кювету, порівнюється з чистим повітрям та з іншим чистим газом, а саме з вуглекислим газом, які знаходяться у відповідних кюветах. Причому вуглекислий газ стандартизований у відповідності до ДСТУ 4817:2007 [6].

Описаний вище оптичний сенсор концентрації газу можна модифікувати шляхом заміни комутатора, аналого-цифрового перетворювача, мікропроцесора та цифрового індикаторного табло на модуль бездротового зв'язку. Він передаватиме сигнали з приймачів на комп'ютер, де за рахунок використання алгоритмів машинного навчання буде відбуватися обробка цих даних, що дозволить здійснювати вимірювання концентрації вуглекислого газу в аналізованій газовій суміші.

Методи машинного навчання у системах вимірювання концентрації газів

На сьогодні методи машинного навчання демонструють високу ефективність у вимірюванні концентрації газу на основі сенсорних даних. Методи машинного навчання можуть обробляти та інтерпретувати дані сенсорів, забезпечуючи покращену точність та чутливість, що дозволяє швидко виявляти гази та вимірювати концентрації на основі сенсорних систем [7].

В роботі [8] розробили «віртуальний» сенсор чадного газу. Це модель машинного навчання, що прогнозує масу викидів чадного газу з вихлопу бензинового двигуна за даними наявних сенсорів (кисневий сенсор HEGO). Навчена модель демонструє високу точність, оскільки дані з сенсорів можна конвертувати в достовірні оцінки чадного газу без додаткового обладнання. Проте цей метод не є універсальним, оскільки досліди проводилися лише на одному автомобілі і невідомо як модель працюватиме за інших умов.

В дослідженні [9] використали ймовірнісний градієнтний бустинг для калібрування показів недорогих сенсорів вуглекислого газу в промислових умовах (важке машинобудування). За даними 8-місячного польового розгортання було встановлено, що модель може навчитися конвертувати необроблені сигнали сенсорів у точні значення концентрації, тобто усувати необхідність дорогого лабораторного калібрування. Також вона дозволяє будувати ймовірнісні карти небезпеки розподілу вуглекислого газу на виробництві на основі отриманих прогнозів. Проте створення карти небезпеки – складний процес, оскільки побудова ймовірнісних моделей та ієрархічної обробки вимагає значних обчислювальних ресурсів при масштабуванні мережі.

У роботі [10] використовується регресія Lasso для апроксимації концентрації вимірюваного газу, яка показує високу точність. Метод застосовує L1-регуляризацию до лінійної регресії, дозволяючи деяким коефіцієнтам ознак зменшуватися до нуля. Це корисно для вибору ознак, а також запобігає перенавчанню, зменшує мультиколінеарність. Проте цей метод використовує лінійну регресію, яка гірше працює для апроксимації інших типів залежності, наприклад, експоненціального.

Залежність одного фізичного параметра від іншого у системах вимірювання концентрації газу часто має нелінійний характер. Оскільки, залежність інтенсивності випромінювання від концентрації газу є експоненціальною, то методи на основі лінійної регресії не дозволяють оцінити точність запропонованої системи вимірювання концентрації вуглекислого газу. Найоптимальнішим методом машинного навчання для такої задачі є регресія опорних векторів. Він використовує концепцію маржі для регресії, добре працює з нелінійними даними та, порівняно з іншими методами, демонструє високу точність за малої та середньої кількості вхідних даних.

У роботі [11] розроблено гібридну регресію опорних векторів методом регресії на основі опорних векторів GA-SVR, який додатково використовує генетичні алгоритми та дозволяє точно прогнозувати зміну концентрації вуглекислого газу в соляних резервуарах. Поєднання з генетичним алгоритмом покращує точність, але ускладнює модель і підвищує обчислювальні витрати.

У дослідженні [12] надано оцінку розчинності вуглекислого газу в іонних рідинах за допомогою методів машинного навчання. Запропоновано оптимізовані моделі регресії опорних векторів регресії (PSO-SVR, GWO-SVR, SSA-SVR), які забезпечують високу точність. Проте навіть вони потребують ретельної оптимізації гіперпараметрів для досягнення найкращих результатів.

В роботі [13] виконано обчислювальне моделювання з використанням регресії опорних векторів для прогнозування поглинання вуглекислого газу в природних умовах. Метод продемонстрував хорошу точність, але все ж таки недостатньо високу. Також метод виявився чутливим до вибору гіперпараметрів.

Таким чином, можна стверджувати, що наявні моделі на основі регресії опорних векторів

показують високу точність, проте мають суттєві недоліки, що робить їх непридатними для використання у нашій системі. Тому виникає необхідність у розробці власної моделі вимірювання концентрації вуглекислого газу з використанням регресії опорних векторів.

Запропонована модель вимірювання концентрації вуглекислого газу з використанням регресії опорних векторів

В роботі запропоновано модель вимірювання концентрації вуглекислого газу з використанням регресії опорних векторів, яка обробляє дані з вищенаведеного модифікованого оптичного сенсора концентрації вуглекислого газу. Для навчання цієї моделі запропоновано ознаки:

- I_{air} – вхідна ознака, яка є інтенсивністю світла на виході кювети з чистим повітрям;
- I_{ref} – вхідна ознака, яка є інтенсивністю світла на виході кювети з вуглекислим газом відомої концентрації;
- I_{sample} – вхідна ознака, яка є інтенсивністю світла на виході вимірюваної кювети з невідомою концентрацією вуглекислого газу;
- CO_{2ppm} – цільова змінна, яка є фактичною концентрацією вуглекислого газу в повітрі (ppm).

Модель отримує на вході I_{air} , I_{ref} , I_{sample} , які показують як аналізований газ поглинає світло порівняно з газами відомих концентрацій. Ці значення дозволяють виявити функціональну залежність між поглинанням світла і концентрацією вуглекислого газу, базуючись на фізичному законі Бугера-Ламберта-Бера. Навчання моделі проводилось на 2500 зразках. Оскільки модель застосовується в контрольованих умовах, цього достатньо для невеликої кількості ознак.

При розробці моделі та виборі її гіперпараметрів необхідно врахувати, що:

- вхідні ознаки мають нелінійний зв'язок з цільовою змінною, оскільки інтенсивність світла має експоненціальну залежність від концентрації вуглекислого газу;
- вхідні дані нормалізовані і мають приблизно однаковий масштаб, що характерно для оптичних сигналів;
- потрібно балансувати між узагальнюючою здатністю моделі та точністю.

Значення гіперпараметрів показано в таблиці 1.

Таблиця 1

Гіперпараметри запропонованої моделі регресії опорних векторів

Гіперпараметр	Позначення	Значення
Ядро функції	kernel	RBF
Параметр регуляризації	C	50
Ширина ε -зони нечутливості	epsilon	0.1
Параметр gamma ядра RBF	gamma	scale
Евристика	shrinking	True
Допуск для умов зупинки оптимізації	tol	0.001
Максимальна кількість ітерацій	max_iter	-1 (без обмежень)

Для ядра моделі вибрано радіально-базисну функцію (RBF), що добре підходить для нелінійних залежностей. Формула RBF-ядра має вигляд:

$$K(x_i, x_j) = e^{-\gamma \|x_i - x_j\|^2}, \quad (1)$$

де γ – коефіцієнт, який визначає ширину гаусіанського ядра; $\|x_i - x_j\|^2$ – відстань між двома точками у просторовому представленні ознак.

Параметр gamma ядра RBF вибрано як scale, тому ширина ядра визначається за формулою:

$$\gamma = \frac{1}{nf \cdot \text{Var}(X)}, \quad (2)$$

де nf – кількість ознак, яка для цієї системи дорівнює 3; $\text{Var}(X)$ – середня дисперсія ознак, що дорівнює середньому значенню дисперсій усіх колонок вхідних даних X по всім ознакам [14].

Параметр регуляризації C визначає компроміс між помилкою на перевіірочній вибірці та складністю моделі. Оскільки для запропонованої системи рівень шумів є невеликим, значення C обрано як 50. Це зроблено для того, щоб модель отримувала більші значення маржі за похибки. Це призводить до того, що модель наближається до вхідних даних, але й створює ризик перенавчання. Ширину ε -зони нечутливості обрано як 0,1. Це значення задає достатню кількість підтримувальних векторів для балансу між точністю та стійкістю моделі. Допуск для умов зупинки оптимізації складає 0,001, що забезпечує достатню точність моделі за не дуже довгий час обчислення. Для пришвидшення отримання розв'язку оптимізації використовується евристика типу «згортання». Максимальна кількість ітерацій не обмежується, що допомагає контролювати час навчання.

Підготовка даних для запропонованої моделі починається зі збору сигналів з приймачів інфрачервоного випромінювання. Дані нормалізуються, що необхідно для стабільної роботи моделі та компенсації відмінностей між інтенсивностями. Після формування повного набору даних вони розподіляються на навчальну та перевіірочну вибірки, що дозволяє оцінити узагальнюючу здатність моделі. Частка для навчальної та перевіірочної вибірки становить 80 % та 20 %, відповідно. Модель регресії опорних векторів на кожному кроці виконує 5-кратну кросс-валідацію. Навчальні дані діляться на п'ять частин, і кожна з них по черзі використовується як перевіірочна вибірка, а решта – для навчальної. Це дозволяє отримати надійну оцінку якості моделі без залежності від конкретного розбиття.

Дослідження запропонованої моделі вимірювання концентрації вуглекислого газу з використанням регресії опорних векторів

Основними показниками ефективності запропонованої моделі вимірювання концентрації вуглекислого газу з використанням регресії опорних векторів, що визначалися в процесі дослідження, було обрано:

- середньоквадратичну похибку (MSE):

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2, \quad (3)$$

- середню абсолютну похибку (MAE):

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|, \quad (4)$$

- коефіцієнт детермінації (R^2):

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \quad (5)$$

де y_i – виміряні значення концентрації вуглекислого газу з сенсора; $\hat{y}_i$ – прогнозовані моделлю значення концентрації вуглекислого газу; $\bar{y}$ – усереднене значення концентрації вуглекислого газу з сенсора; $i = 1, 2, \dots, n$ – кількість вимірювань; n – загальна кількість вимірювань.

Обрані показники ефективності є типовими для оцінювання якості моделей машинного навчання. Вони є стандартизованими статистичними метриками для оцінювання відповідності між фактичними вимірюваннями та прогнозованими результатами [15].

Дослідження запропонованої моделі було проведено шляхом моделювання зміни MSE, MAE та точності за метрикою R^2 під час навчання та перевірки моделі, порівняння даних з оптичного сенсора та з моделі, а також оцінки зміни MSE та MAE в процесі вимірювання.

На рис. 2 показано графіки зміни MSE, MAE та точності за метрикою R^2 під час навчання та перевірки запропонованої моделі. Криві навчання побудовані для оцінки поведінки моделі під час навчання, залежно від обсягу доступних навчальних даних. Вони показують як змінюється точність (або похибка) моделі для навчального та перевірного наборів із зростанням кількості даних, що використовуються для навчання.

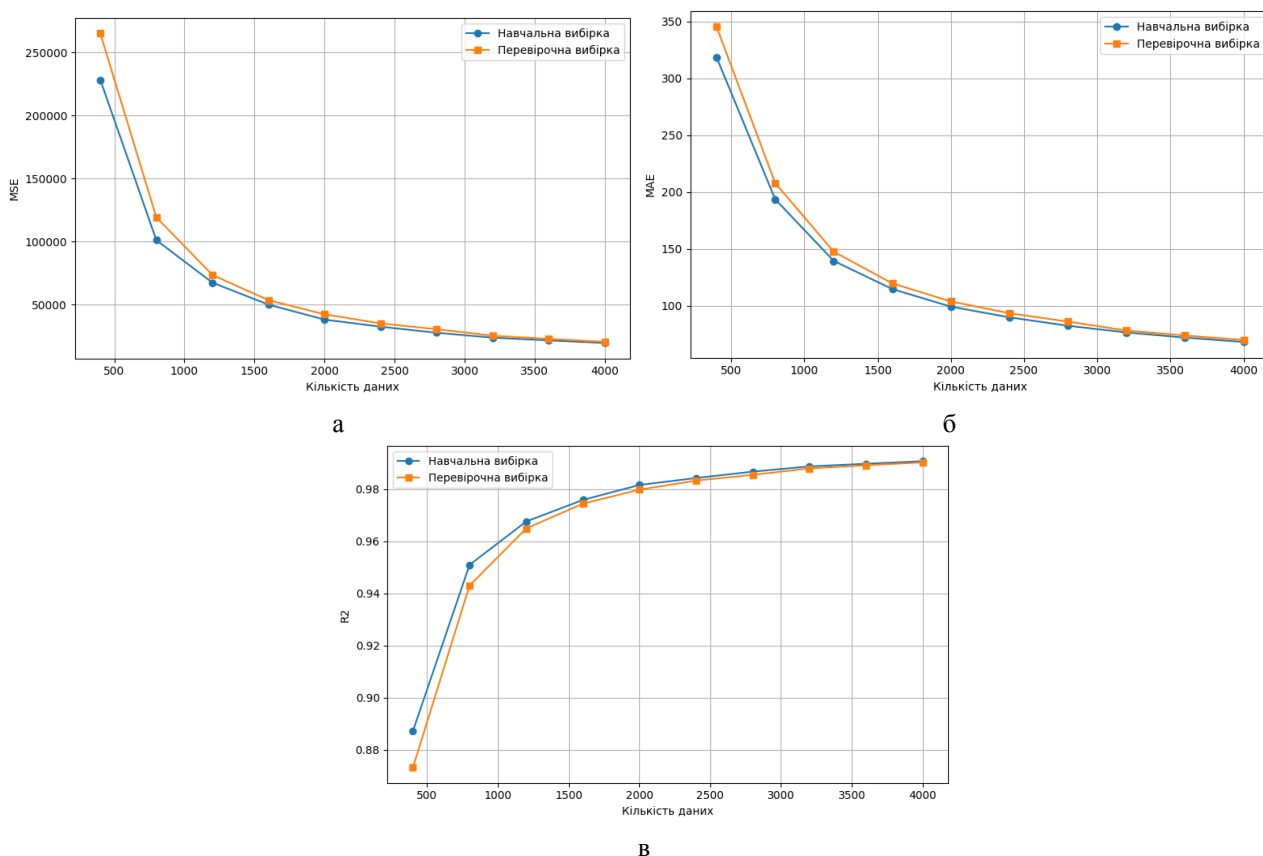


Рис. 2. Графіки зміни MSE (а), MAE (б) та точності за метрикою R^2 (в) під час навчання та перевірки запропонованої моделі

Як видно з рис. 2, модель поступово знижує похибку на навчальних та перевірочних даних. Також для навчальної і перевірочної вибірки спостерігається зростання значення R^2 зі збільшенням кількості вимірювань. Це свідчить про те, що модель успішно навчається та поліпшує свою точність прогнозування за більшої кількості даних. Після 3000–4000 вимірювань обидві криві майже вирівнюються, і значення R^2 наближається до 0,99. Це ознака доброї узагальнюючої здатності моделі, тобто вона не просто запам'ятовує дані, а справді навчається їх закономірностям. Значення $R^2 > 0,96$ на великій кількості даних. Запропонована модель демонструє стійке підвищення точності із збільшенням кількості навчальних прикладів без ознак перенавчання. Крива навчання свідчить, що додаткові дані продовжують покращувати її продуктивність, хоча ефект від цього поступово зменшується. Модель збалансована та добре адаптована до поточного завдання. Крім того, обидві криві розташовані близько одна до одної, що також свідчить про те, що модель добре навчається і узагальнює.

На рис. 3 показано графік порівняння даних з оптичного сенсора та з моделі регресії опорних векторів.

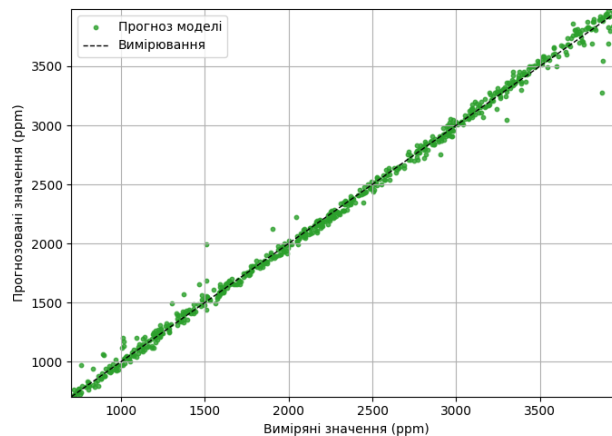


Рис. 3. Графік порівняння даних з оптичного сенсора та з моделі регресії опорних векторів

Як видно з рис. 3, було обрано діапазон концентрації вуглекислого газу від 500 до 4000 ppm, оскільки це відповідає небезпечному рівню вуглекислого газу в будинках. Проведено 5000 вимірювань. Коефіцієнт детермінації $R^2 = 0,992$. Це означає, що понад 99 % варіацій у вимірних значеннях концентрації вуглекислого газу пояснюється моделлю. Це досить високий показник, який свідчить про добру узгодженість прогнозу з реальними даними. Цей графік дозволяє візуально швидко оцінити наскільки добре модель апроксимує залежність між вимірними та прогнозованими значеннями концентрації вуглекислого газу.

На рис. 4 показано графіки зміни похибок MSE та MAE в процесі вимірювання.

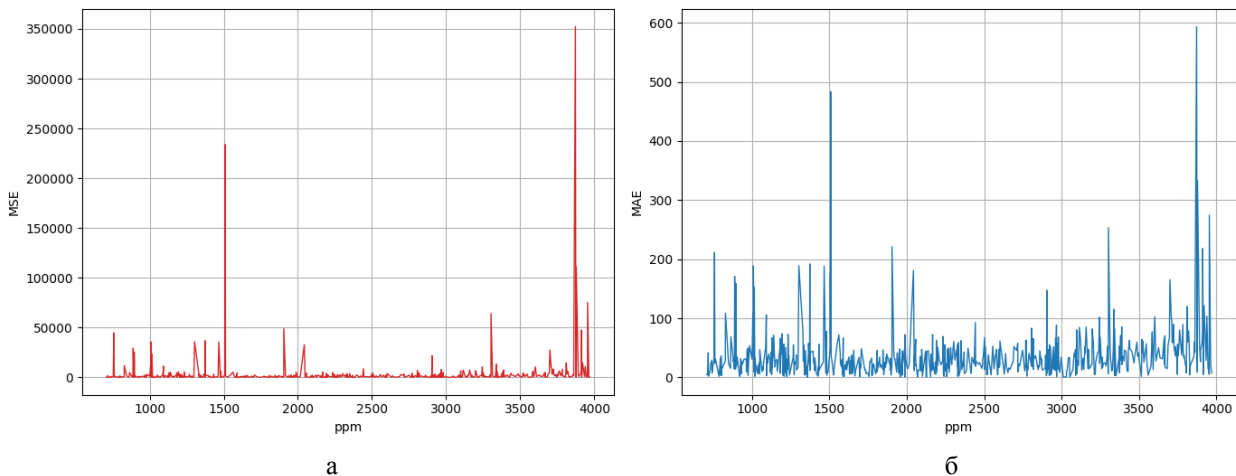


Рис. 4. Графіки зміни похибок в процесі вимірювання: а – MSE; б – MAE

Як видно з рис. 4, розподіл похибок MSE та MAE дозволяє оцінити наскільки рівномірно модель здійснює прогнозування значень концентрації вуглекислого газу та чи є значні відхилення. Встановлено, що загальне значення MSE складає 16710,472 ppm², а MAE – 65,561 ppm. Таким чином, відхилення прогнозованих значень концентрації перебуває в межах малих похибок, що вказує на високу точність і надійну роботу моделі та її потенційне застосування в системах моніторингу вуглекислого газу в режимі реального часу.

Висновки

В роботі запропоновано модель з використанням регресії опорних векторів для системи вимірювання концентрації вуглекислого газу на основі бездротового оптико-електронного сенсора. Ця модель обробляє дані з бездротового оптико-електронного сенсора концентрації

вуглекислого газу. Порівняно з наявними моделями вона є досить точною та простою, адже не вимагає значних обчислювальних ресурсів. Показниками ефективності запропонованої моделі регресії опорних векторів, що визначалися в процесі дослідження, було обрано такі характеристики як середньоквадратична похибка (MSE), середня абсолютна похибка (MAE) та коефіцієнт детермінації (R^2). Для оцінки якості навчання моделі розраховувались MSE, MAE та точності за метрикою R^2 для навчальної та перевірконої вибірки. Встановлено, що модель поступово знижує похибку на навчальних та перевіркочних даних. Також для навчальної і перевірконої вибірки спостерігається зростання R^2 зі збільшенням кількості вимірювань, причому значення R^2 стабілізується після 0,96 на великій кількості даних. Це свідчить про те, що модель успішно навчається та поліпшує свою точність прогнозування за більшої кількості даних. Після 1500–2000 вимірювань обидві криві майже вирівнюються, і значення R^2 наближається до 0,99. Це є ознакою доброї узагальнюючої здатності моделі, яка також демонструє стійке підвищення точності із ростом кількості навчальних прикладів без ознак перенавчання. Виконувалось порівняння даних з оптичного сенсора та з моделі, а також оцінки зміни MSE та MAE в процесі вимірювання. Було обрано діапазон концентрації вуглекислого газу від 500 до 4000 ppm, оскільки це відповідає небезпечному рівню вуглекислого газу в будинках. Проведено 5000 вимірювань. Коефіцієнт детермінації $R^2 = 0,992$. Це означає, що понад 99 % варіацій у вимірних значеннях концентрації вуглекислого газу пояснюється моделлю. Це досить високий показник, який свідчить про добру узгодженість прогнозу з реальними даними. Розподіл похибок MSE та MAE дозволив встановити, що загальне значення MSE складає 16710,472 ppm², а MAE – 65,561 ppm, що вказує на високу точність запропонованої моделі. Таким чином, відхилення прогнозованих значень концентрації перебуває в межах малих похибок, це вказує на високу точність і надійну роботу моделі та її потенційне застосування в системах моніторингу вуглекислого газу в режимі реального часу.

СПИСОК ЛІТЕРАТУРИ

1. Вуглекислий газ. URL: <https://gastech.com.ua/UA/vuglekyslyj-gaz-co2>. (дата звернення: 28.07.2025).
2. Qilong P., Fouzi H., Ying S. A comparison of machine learning methods for ozone pollution prediction. *Journal of Big Data*. 2023. Volume 10. P. 1–31. <https://doi.org/10.1186/s40537-023-00748-x>.
3. Quantitative and Qualitative Analysis of Multicomponent Gas Using Sensor Array / F. Shurui et al. *Sensors*. 2019. Volume 19. P. 1–12. <https://doi.org/10.3390/s19183917>.
4. Оптичний сенсор концентрації газу : пат. 68725 Україна : МПК G01N 21/01. № u201110898 ; заявл. 12.09.2011 ; опубл. 10.04.2012, Бюл. № 7.
5. Білінський Й. Й., Книш Б. П., Іоніна К. Ю. Інфрачервоний триканальний сенсор концентрації газу. *Вісник Хмельницького національного університету*. 2011. №6. С. 41–49.
6. ДСТУ 4817: 2007. Діоксид вуглецю газоподібний і скраплений. Технічні умови. [Чинний від 2009-01-01]. Київ : УкрНДНЦ, 2007. 39 с.
7. Machine Learning in Wireless Sensor Networks: Algorithms, Strategies, and Applications / M. A. Alsheikh et al. *IEEE Communications Surveys & Tutorials*. 2014. №16. P. 1–23. DOI: 10.1109/COMST.2014.2320099.
8. Machine-Learning-Based Emission Models in Gasoline Powertrains – Part 2: Virtual Carbon Monoxide / N. Kempema et al. *SAE International Journal of Engines*. 2023. Volume 16. P. 799–807. <https://doi.org/10.4271/03-16-06-0045>.
9. Probabilistic Machine Learning with Low-Cost Sensor Networks for Occupational Exposure Assessment and Industrial Hygiene Decision Making / A. Patton et al. *Annals of Work Exposures and Health*. 2022. Volume 66. P. 580–590. <https://doi.org/10.1093/annweh/wxab105>.
10. Chowdhury M. A. Z., Oehlschlaeger M. A. Artificial Intelligence in Gas Sensing: A Review. *ACS Sensors*. 2025. Volume 10. P. 1538–1563. <https://doi.org/10.1021/acssensors.4c02272>.
11. Predictive Modelling of Carbon Dioxide Injectivity Using SVR-Hybrid / M. Mardhatillah et al. *Offshore Technology Conference Asia*. 2022. <https://doi.org/10.4043/31472-ms>.
12. Estimating CO₂ solubility in ionic liquids by using machine learning methods / Z. Liu et al. *Journal of Molecular Liquids*. 2023. Volume 391. P. 1–20. <https://doi.org/10.1016/j.molliq.2023.123308>.
13. Computational simulation using machine learning models in prediction of CO₂ absorption in environmental applications / H. Jin et al. *Journal of Molecular Liquids*. 2022. Volume 358. P. 1–12. <https://doi.org/10.1016/j.molliq.2022.119159>.

14. Ebden M. Gaussian Processes: A Quick Introduction. *Mathematics*. 2015. №5. P. 1–13. <https://doi.org/10.48550/arXiv.1505.02965>.
15. Predicting Methane Concentrations in Underground Coal Mining Using a Multi-Layer Perceptron Neural Network Based on Mine Gas Monitoring Data / M. Tutak et al. *Sustainability*. 2024. №16. P. 1–23. <https://doi.org/10.3390/su16198388>.

Стаття надійшла до редакції 12.10.2025.

Стаття пройшла рецензування 28.10.2025.

Книш Богдан Петрович – канд. техн. наук, доцент кафедри загальної фізики, e-mail: tutmos-3@i.ua.

Кулик Ярослав Анатолійович – канд. техн. наук, доцент кафедри автоматизації та інтелектуальних інформаційних технологій.

Вінницький національний технічний університет.