

НЕЙРОМЕРЕЖЕВИЙ ПІДХІД ДО ІДЕНТИФІКАЦІЇ ЗАПОВНЮВАНOSTІ ПРИМІЩЕНЬ ЗА ПАРАМЕТРАМИ ПОВІТРЯ

Анотація. У роботі запропоновано підхід до визначення кількості людей у приміщенні на основі даних спостережень за параметрами повітря із застосуванням багатoshарової нейронної мережі. Досліджено питання вибору архітектури та параметрів нейронної мережі для розв'язання задачі прогнозування. Надано рекомендації для підвищення продуктивності моделі.

Ключові слова: концентрація CO₂, нейронна мережа, точність прогнозування, часовий ряд.

Постановка проблеми. Останнім часом питання дослідження параметрів повітря в шкільних класах, лікарнях, офісних та промислових приміщеннях стають поширеними через суттєвий вплив якості повітря на здоров'я та комфорт людей. Наявність у повітрі значної концентрації вуглекислого газу (CO₂), оксиду азоту (NO₂), дрібних часток пилу (PM_{2.5}, PM₁₀), летючих органічних сполук (VOC), бактерій та грибків призводить до проблем зі здоров'ям людини, тому моніторинг цих параметрів, можливість прогнозувати їх значення та керувати системами вентиляції є важливими для забезпечення здорового та комфортного середовища у приміщеннях. Таким чином, постає наукова проблема розроблення математичних моделей та методів прогнозування параметрів повітря за результатами моніторингу в реальному часі та суміжні задачі, зокрема керування системами вентиляції на основі спостережень, ідентифікація наповненості приміщень тощо.

Аналіз останніх досліджень. Підвищена увага до якості повітря у приміщеннях під час пандемії COVID-19, промислового забруднення, лісових пожеж та інших екологічних катастроф сприяли розробці та застосуванню моделей на основі даних для прогнозування рівнів забруднюючих речовин в приміщенні і їх впливу на організм людини. У фундаментальному дослідженні [1] наводяться відомості про вплив параметрів повітря на комфорт людини, мето-

ди вимірювання параметрів повітря, відомості про розподіл джерел забруднення і їх вплив на якість повітря.

Останнім часом багато досліджень присвячується застосуванню штучних нейронних мереж (ШНМ) для прогнозування рівня забруднюючих речовин в приміщенні. В роботі [2] штучна нейронна мережа використовується для прогнозування щогодинних концентрацій $PM_{2.5}$, PM_{10} . Для навчання нейронної мережі в якості вхідних даних обираються значення концентрацій $PM_{2.5}$, PM_{10} , час та метеорологічні параметри, а значення відповідних концентрацій, що вимірюються наступної години, використовуються як вихідні дані. Розглядаються різні архітектури нейронних мереж та досліджується вплив вибору тренувальної функції пакету прикладних програм MatLab на результат прогнозування. Порівнюються результати прогнозування з використанням 30 тренувальних функцій, серед яких байесова регуляризація зворотнього поширення (Trainbr) та метод Левенберга-Марквардта (Trainlm) виявились найкращими за значеннями середньоквадратичної та середньої абсолютної похибок. Доводиться, що обидві функції мають здатність до узагальнення та задовольняють потреби щогодинного прогнозування наявності часточок пилу.

В [3] досліджуються три моделі ШНМ, а саме ШНМ зворотного поширення, багат шарова ШНМ, а також ШНМ довготривалої та короткочасної пам'яті (LSTM). За результатами масштабного обчислювального експерименту встановлено, що найкращою для прогнозування концентрації $PM_{2.5}$, PM_{10} виявилась багат шарова ШНМ та ансамблевий метод машинного навчання – випадковий ліс, який працює за допомогою побудови численних дерев прийняття рішень під час тренування моделі й здійснює усереднений прогноз.

Автори [4] наводять огляд типів та джерел забруднювачів повітря в приміщенні, досліджують відомі в літературі прогностичні моделі ШНМ, зокрема процеси моделювання та алгоритми їх навчання. За результатами аналізу встановлено, що найчастіше використовуються традиційні ШНМ, тоді як моделі рекурентних нейронних мереж (RNN) краще розв'язують проблему прогнозування часових рядів.

Поточні дослідження найчастіше використовують значення концентрацій $PM_{2.5}$, PM_{10} , CO_2 та не враховують вплив характеристик самого приміщення, параметрів його заповнюваності на результати прогнозування якості повітря через складність включення цих показників в прогностичні моделі. Також актуальною залишається задача визначення кількості людей у приміщенні виходячи

зі значень параметрів повітря для подальшого розрахунку повітрообміну у приміщеннях та керування системами вентиляції.

Мета дослідження. Метою дослідження є створення моделі для прогнозування кількості людей у приміщенні на основі поточних даних про якість повітря, зокрема концентрації CO₂.

Викладення основного матеріалу досліджень. Одним з підходів, який можна використовувати для визначення кількості людей у приміщенні, є оцінка концентрації CO₂ в повітрі. Вважаємо, що люди видихають CO₂, і його рівень у повітрі збільшується зі збільшенням кількості людей у приміщенні. Таким чином, вимірюючи рівень CO₂, можна зробити приблизну оцінку кількості присутніх людей. Однак, визначення кількості людей у приміщенні за допомогою параметрів повітря є складним завданням, оскільки вимагає врахування різноманітних факторів, зокрема, призначення приміщення, умов вентиляції, активності людей, наявності додаткових чинників, таких як відкривання вікон, дверей, наявності інших джерел надходження CO₂, наприклад, джерел відкритого вогню тощо. Характер та кількісні показники деяких з перелічених факторів не піддаються обліку та не можуть бути додані до параметрів моделі. Математичні моделі, що визначають параметри повітрообміну приміщень відповідно до кількості людей, також не дозволяють врахувати перелічені фактори. Використовується лише незначна кількість параметрів в якості коефіцієнтів моделі, а їх значення обираються з деякого інтервалу.

Тому у дослідженні пропонується використання ШНМ для прогнозування кількості людей за результатами спостереження за параметрами повітря у приміщенні. Для побудови моделі пропонується використовувати дані вмісту CO₂ у повітрі, що одержуються з датчиків аналізаторів якості повітря та збираються в офісних приміщеннях різного призначення – конференц-залах, стандартних офісних кімнатах, які обладнано сучасними меблями та оргтехнікою. Данні про вміст CO₂ разом з даними про кількість людей у приміщенні, що також одержуються за допомогою лічильників або камер відеоспостереження, пропонується використати для навчання ШНМ так, щоб навчену на таких даних нейронну мережу в подальшому можна було використовувати для прогнозування кількості людей у приміщенні лише за даними аналізаторів якості повітря.

Для спостереження за параметрами повітря та формування навчальної вибірки у приміщеннях було розташовано сенсори, що вимірюють концентрації CO₂ щогодини протягом року. Вибір зазначеного часового інтервалу пояс-

нюється тим, що модель має враховувати зміни обумовлені порою року. Також в приміщеннях було розташовано лічильники, які вимірюють кількість людей в приміщенні в ті ж самі проміжки часу. Експериментально створений набір даних, який складався з близько 7000 записів у вигляді часових рядів, використовувався для навчання та тестування ШНМ. Дані, зібрані в приміщеннях А та В було використано для навчання та тестування моделі. Тестування відбувалось на випадково відібраної вибірці, яка складалася з 20 % записів з числа загальних даних для навчання по приміщенням А, В. Дані, що було зібрано в приміщенні С, не брали участь у налаштуванні та тестуванні моделі, їх було використано виключно для перевірки здатності моделі до узагальнення.

Попередньо, для мінімізації обчислювальних витрат набір даних було нормалізовано в такий спосіб:

$$P = (a - a_{\min}) / (a_{\max} - a_{\min}),$$

де P – нормалізоване значення параметра; a – значення параметра з вибірки; a_{\min} , a_{\max} – максимальне та мінімальне значення параметра a з вибірки.

Модель використовує виміряні концентрації CO₂ та відомі значення про кількість людей в моменти часу від $(t - 1)$ до $(t - i)$, де $i = \overline{2, m}$, щоб спрогнозувати кількість людей в приміщенні на момент t . Надалі модель бере прогнозовані значення концентрації CO₂ у приміщенні та кількість людей у момент часу t і використовує їх як вхідні дані для прогнозування кількості людей у момент часу $t + 1$. В такий спосіб можна прогнозувати кількість людей у приміщенні в часі.

Для оцінювання якості прогнозування моделі застосовувались метрики середньоквадратичної помилки прогнозу (RMSE) та середньої абсолютної помилки прогнозу (MAE):

$$RMSE = \sqrt{\frac{1}{N-1} \sum_{t=1}^N (y(t) - \hat{y}(t))^2}; MAE = \frac{1}{N} \sum_{t=1}^N |y(t) - \hat{y}(t)|,$$

де N – загальна кількість використаних даних; $y(t)$ – експериментально виміряне значення, а $\hat{y}(t)$ – прогнозоване значення параметра.

Для програмної реалізації ШНМ було використано бібліотеку програмного забезпечення з відкритим кодом для числових розрахунків з використанням графів потоку даних TensorFlow. Було побудовано багатошарову нейронну мережу, яка складається з вхідного, вихідного та кількох прихованих шарів. Багатошарові ШНМ з прихованими шарами здатні до апроксимації складних нелінійних залежностей, якими є часові ряди. Модель ШНМ оновлює вагову матри-

цю прихованого шару за допомогою алгоритму зворотнього поширення помилки і використовує вихід попереднього прихованого шару в якості входу поточного прихованого шару, щоб поступово змінювати внутрішні параметри мережі. Слід зазначити, що використання значної кількості прихованих шарів призводить до ускладнення навчання та зниження якості прогнозування. Вибір архітектури мережі, зокрема визначення кількості шарів, кількості нейронів у шарі помітно впливає на точність прогнозування кількості людей у приміщенні. Варіанти конфігурацій, що розглядалися, та якість результатів прогнозування на тестових вибірках наведено в Табл. 1, оптимальні результати та відповідний варіант конфігурації мережі виділено жирним шрифтом. Для обох приміщень оптимальним виявився один і той самий варіант конфігурації ШНМ. Оцінювання якості прогнозування, що здійснюється мережею, відбувалось за критерієм Краскала – Уолліса, що дозволило перевірити рівність медіан кількох навчальних вибірок.

Таблиця 1

Вибір конфігурації ШНМ

Приміщення	Варіант конфігурації ШНМ	Кількість нейронів у шарах мережі	MAE	RMSE
приміщення А	1	[20, 10, 1]	0.2597475	0.1976358
	2	[20, 10, 5, 1]	0.0923175	0.1259638
	3	[20, 10, 5, 5, 1]	0.0349712	0.0463273
	4	[20, 10, 5, 5, 5, 1]	0.1158294	0.1078490
	5	[20, 10, 5, 5, 5, 5, 1]	0.1435618	0.1096385
приміщення В	1	[20, 10, 1]	0.2749025	0.2112699
	2	[20, 10, 5, 1]	0.1479261	0.1611635
	3	[20, 10, 5, 5, 1]	0.0477911	0.0586487
	4	[20, 10, 5, 5, 5, 1]	0.1247318	0.1414523
	5	[20, 10, 5, 5, 5, 5, 1]	0.1636910	0.1866812

На рис. 1 наводяться результати прогнозування кількості людей у приміщенні А по днях, прогнозні значення позначено пунктирною лінією, сполушеною лінією позначено дійсну кількість людей (дані одержано з всієї тестової вибірки).

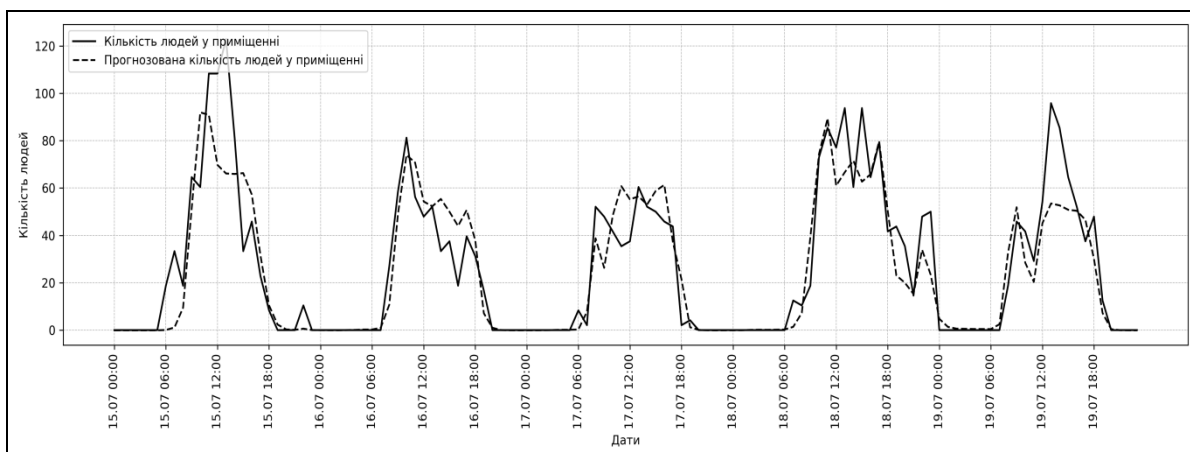


Рисунок 1 – Порівняння прогнозованих значень кількості людей у приміщенні А із спостережуваними значеннями для всієї тестової вибірки

На рис. 2 наведено усереднені значення прогнозованої кількості людей у приміщенні С (відображено пунктирною лінією) та порівняно із усередненими спостережуваними значеннями для всієї навчальної вибірки. Слід зазначити, що модель успішно відтворює характер розподілу прогнозованого параметру, хоча спостережувані дані з приміщення С не брали участь у навчанні моделі. За результатами навчання модель виявилась здатною охопити складні нелінійні залежності між різними факторами, що впливають на результат прогнозування.

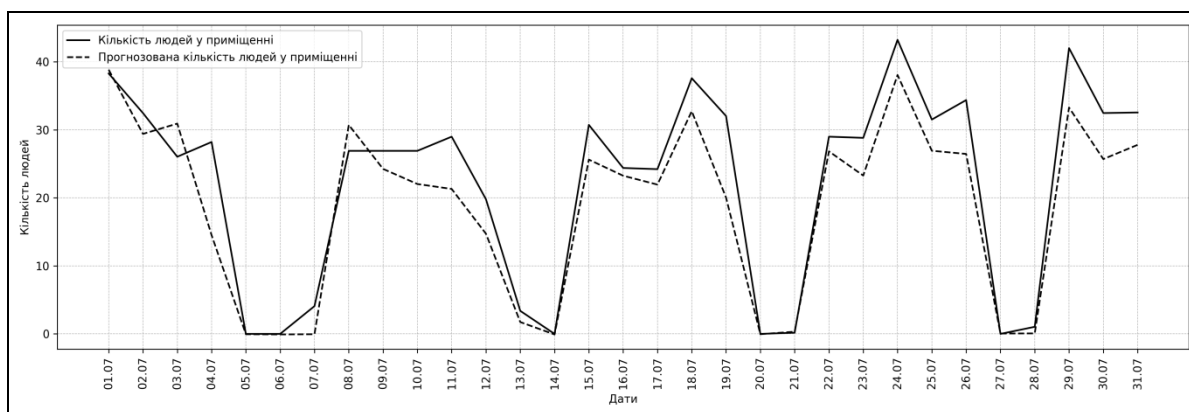


Рисунок 2 – Порівняння усереднених прогнозованих значень кількості людей у приміщенні С із усередненими спостережуваними значеннями для всієї навчальної вибірки

Незначні похибки прогнозованих значень, що залишаються в результатах прогнозу, можна пояснити обмеженістю початкових даних, похибкою самої моделі та наявністю факторів невизначеності. Слід зазначити, що продуктив-

ність моделі можна значно покращити, якщо до набору даних додати параметри впливу системи механічної вентиляції, інтенсивність людської діяльності, частоту відкривання та закривання дверей, вікон тощо.

Висновки. У роботі досліджено можливості багат шарових нейронних мереж для розв'язання задачі прогнозування. За результатами аналізу літературних джерел досліджено стан проблеми, визначено основні підходи до розв'язання задач, пов'язаних з аналізом якості повітря у приміщеннях. Побудовано математичну модель багат шарової нейронної мережі для визначення кількості людей у приміщенні за результатами спостережень за параметрами повітря та її програмну реалізацію. За результатами проведених обчислювальних експериментів із застосуванням статистичного аналізу сформульовано рекомендації щодо вибору архітектури мережі та її параметрів. Надано рекомендації для підвищення продуктивності моделі.

ЛІТЕРАТУРА

1. Pipal A.S. Measurements of Indoor Air Quality // A.S. Pipal, A. Taneja. – Handbook of Metrology and Applications. Springer, Singapore. – 2023. – Pp 1-35. DOI:10.1007/978-981-19-1550-5_90-1
2. Guo Q. Prediction of Hourly PM_{2.5} and PM₁₀ Concentrations in Chongqing City in China Based on Artificial Neural Network // Q. Guo, Z. He, Z. Wang // Aerosol and Air Quality Research. – 2023.– Volume 23, issue 6. DOI: 10.4209/aaqr.220448
3. Gao-wa S. Using Artificial Neural Networks to Predict Indoor Particulate Matter and Tvoc Concentration in an Office Building: Model Selection and Method Development // S. Gao-wa, Z. Zhen, N. Jianchun, L. Linxiao, A. Han, Y. Zhili // Energy and Built Environment. 2024. DOI: 10.1016/j.enbenv.2024.03.001.
4. Dong J. A Review of Artificial Neural Network Models Applied to Predict Indoor Air Quality in Schools // J. Dong, N. Goodman, P. Rajagopalan // Int. J. Environ. Res. Public Health. – 2023. – Vol. 20. – Pp. 1-18. DOI: 10.3390/ijerph20156441

REFERNCES

1. Pipal A.S. Measurements of Indoor Air Quality // A.S. Pipal, A. Taneja. – Handbook of Metrology and Applications. Springer, Singapore. – 2023. – Pp 1-35. DOI:10.1007/978-981-19-1550-5_90-1
2. Guo Q. Prediction of Hourly PM_{2.5} and PM₁₀ Concentrations in Chongqing City in China Based on Artificial Neural Network // Q. Guo, Z. He, Z. Wang // Aerosol and Air Quality Research. – 2023.– Volume 23, issue 6. DOI: 10.4209/aaqr.220448
3. Gao-wa S. Using Artificial Neural Networks to Predict Indoor Particulate Matter and Tvoc Concentration in an Office Building: Model Selection and Method

Development // S. Gao-wa, Z. Zhen, N. Jianchun, L. Linxiao, A. Han, Y. Zhili // Energy and Built Environment. 2024. DOI: 10.1016/j.enbenv.2024.03.001.

Dong J. A Review of Artificial Neural Network Models Applied to Predict Indoor Air Quality in Schools // J. Dong, N. Goodman, P. Rajagopalan // Int. J. Environ. Res. Public Health. – 2023. – Vol. 20. – Pp. 1-18. DOI: 10.3390/ijerph20156441.

Received 13.03.2024.

Accepted 18.03.2024.

A neural network approach to the identification of room occupantiness according to air parameters

The paper introduces an approach to determining the number of people in a room based on data from observations of air parameters using a multilayer neural network. Monitoring of air parameters, the ability to predict their values and manage ventilation systems are important to ensure a healthy and comfortable indoor environment. The purpose of the research is to develop mathematical models and methods of forecasting air parameters based on the results of real-time monitoring. Different approaches to predicting air parameters and the number of people in rooms using mathematical models in the form of equations and artificial neural networks with different architectures and types of training functions are considered. The paper proposes an approach to forecasting with the help of a multilayer neural network, which allows taking into account various factors, the nature and quantitative values of which cannot be taken into account and cannot be added to the model parameters. The CO₂ data together with the indoor occupancy data from the meters are used to train the neural network. In the future, a neural network trained on such data can be used to predict the number of people in a room based only on data from air quality analyzers.

The issue of choosing the architecture of a multilayer neural network and its parameters for solving the forecasting problem has been investigated. Neural network training is carried out by the method of error back propagation. To evaluate the forecasting quality of the model, the metrics of mean square error of forecast and mean absolute error of forecast are used. The Kruskal-Wallis criterion is used to take into account the results of forecasting on several samples. Based on the results of the computational experiment, the optimal network architecture is determined. The model successfully reproduces the nature of the distribution of the predicted parameter, as it captures the complex nonlinear dependencies between the various factors of the model. Recommendations are given to improve the performance of the model.

Гук Костянтин Григорович – магістрант кафедри ракетно-космічних та інноваційних технологій, Дніпровський національний університет імені Олеся Гончара, Дніпро.

Шевельова Алла Євгенівна – д-р фіз.-мат. наук, професор, професор кафедри обчислювальної математики та математичної кібернетики, Дніпровський національний університет імені Олеся Гончара, Дніпро.

Huk Kostyantyn – master's student of the Department of Rocket-space and Innovative Technologies, Oles Honchar Dnipro National University, Dnipro.

Sheveleva Alla – Doctor of Physical and Mathematical Sciences, Professor, Professor of the Department of Computational Mathematics and Mathematical Cybernetics, Oles Honchar Dnipro National University, Dnipro.