Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 1 (31)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

УДК 004.94
DOI: https://doi.org/10.30837/2522-9818.2025.1.073

О. Ляшенко, І. Михайліченко

МОДЕЛЬ АВТОНОМНОЇ СИСТЕМИ МОНІТОРИНГУ ТА ОПТИМІЗАЦІЇ ІТ-ІНФРАСТРУКТУРИ З ВИКОРИСТАННЯМ ТРАНСФОРМЕРІВ

Предметом дослідження є модель автономної системи моніторингу та оптимізації ІТ-інфраструктури, що використовує трансформери для аналізу багатовимірних часових рядів і виявлення аномалій у реальному часі. У статті розглянуто сучасні підходи до моніторингу ІТ-інфраструктури, зокрема використання машинного навчання та класичних методів статистичного аналізу. Огляд літератури показує, що наявні підходи мають обмеження, зокрема низьку ефективність в умовах динамічних змін параметрів системи. **Мета статті** – розроблення автономної системи, здатної здійснювати багатофакторний аналіз у реальному часі та автоматично реагувати на виявлені загрози. Запропоновано модель на основі трансформерів, що дає змогу виявляти складні аномалії та прогнозувати можливі збої. З огляду на окреслену мету необхідно розв'язати такі **завдання:** сформувати модель багатовимірного аналізу часових рядів; розробити алгоритм виявлення аномалій і прогнозування можливих проблем; реалізувати механізми автономного налаштування системи для стабілізації ІТ-інфраструктури. Упроваджено такі **методи:** математичне моделювання, машинне навчання (трансформери), статистичний аналіз (крос-валідація), алгоритми прогнозування на основі часових рядів. **Здобуті результати.** Модель досягла середньої абсолютної похибки (MAE) 4.3 % на синтетичних даних, що доводить її здатність точно виявляти аномалії. Крос-валідація підтвердила стабільність навчання без перенавчання, а гістограма залишків продемонструвала рівномірний розподіл похибок. Крім того, теплові карти кореляції показали взаємозв'язки між ключовими параметрами ІТ-інфраструктури. **Висновки.** Розроблена система дає змогу автоматично виявляти та прогнозувати проблеми в роботі ІТ-інфраструктури, забезпечуючи автономне налаштування параметрів для підтримки її стабільності. Запропонований підхід може бути інтегрований у сучасні системи управління ІТ-інфраструктурою для покращення їх ефективності.

Ключові слова: автономна система; трансформери; багатовимірні часові ряди; ІТ-інфраструктура; виявлення аномалій; прогнозування.

Вступ

У сучасному світі інформаційних технологій стабільність та безпека ІТ-інфраструктури мають критичне значення для надійної роботи організацій будь-якого масштабу. Різноманіття серверів, мережевого обладнання, систем зберігання та інших компонентів генерують значні обсяги даних, що містять цінну інформацію для аналізу й ухвалення оперативних рішень. Проте збільшення складності систем, динаміка їх змін і стрімке зростання обсягів даних знижують ефективність класичних підходів до моніторингу та діагностики інфраструктури.

Аналіз проблеми

Сучасні комерційні системи забезпечують певний рівень автоматизації та використовують методи машинного навчання. Проте вони часто мають обмеження, серед яких недостатня здатність до глибокого аналізу багатовимірних даних і виявлення багатофакторних аномалій, залежність від ручного налаштування.

Традиційні методи моніторингу зазвичай ґрунтуються на використанні фіксованих порогів для кожного окремого показника, що дає змогу виявляти прості аномалії, але є недостатньо точним для визначення комплексних проблем, що виникають у взаємодії між різними параметрами [1].

Застосування трансформерів у аналізі часових рядів продемонструвало значні переваги, зокрема здатність опрацьовувати довгострокові залежності та висока продуктивність в аналізі даних у реальному часі.

Запропонована в цій роботі система розв'язує питання, пов'язані з обмеженнями, за допомогою застосування адаптованих трансформерів, таких як *Timesformer* та *Informer*, що забезпечують ефективний аналіз часових рядів, зважаючи як на довгострокові, так і короткострокові залежності. Основними перевагами запропонованого підходу є виявлення багатофакторних аномалій – використання трансформерів дає змогу ідентифікувати складні

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 1 (31)

аномалії в режимі реального часу навіть за умов багатовимірних відхилень у різних компонентах системи.

Також система здатна використовувати історичні дані для передбачення майбутніх загроз, що допомагає запобігати потенційним збоям і порушенням у роботі інфраструктури, і застосовувати автономне налаштування: з огляду на виявлені аномалії та прогнозовані проблеми, система може автоматично запускати процеси оптимізації, щоб забезпечити стабільну та енергоефективну роботу інфраструктури.

Метою статті є розроблення та впровадження автономної системи моніторингу IT-інфраструктури, що здатна здійснювати багатофакторний аналіз у реальному часі та автономно реагувати на виявлені загрози. У роботі вперше запропоновано трансформерну модель для складного моніторингу часових рядів, модифіковано архітектуру трансформера для потреб багатовимірного аналізу IT-інфраструктури, розроблено механізм автономної оптимізації на основі прогнозування проблем. Досягнуті результати мають потенціал для подальшого розвитку сучасних підходів до управління IT-інфраструктурою в умовах динамічного середовища.

Розв’язання завдання

Розроблена система для моніторингу та аналізу IT-інфраструктури містить кілька ключових модулів, що забезпечують збір і попереднє оброблення інформації, навчання моделі на основі трансформерів для виявлення аномалій, оцінювання якості моделі за допомогою крос-валідації, а також візуалізацію результатів прогнозування та аналізу аномалій. Основні компоненти системи (рис. 1):

– модуль збору та генерації даних;

– модуль попереднього оброблення та підготовки інформації: для забезпечення стабільності моделі всі параметри нормалізуються з використанням *StandardScaler*;

– модуль моделі на основі трансформера: модель, яку застосовує система, є трансформером, налаштованим на оброблення багатовимірних часових рядів;

– модуль крос-валідації: для оцінювання стабільності та узагальненості моделі реалізовано п’ятикратну крос-валідацію;

– модуль візуалізації результатів: система генерує візуальні графіки для аналізу якості роботи моделі. Зокрема для кожного фолду крос-валідації будується графік динаміки втрат.

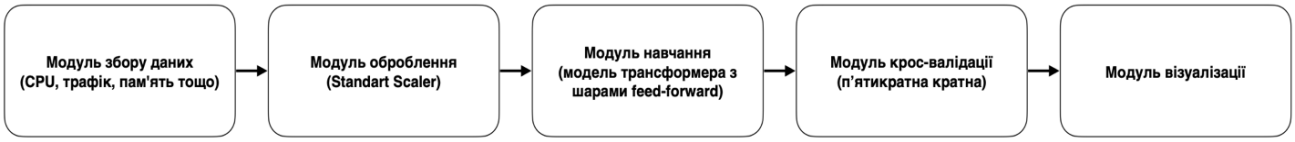

Рис. 1. Основні компоненти системи

Визначальний принцип роботи системи передбачає декілька ітерацій, перша з яких – генерація та попереднє оброблення інформації. Система генерує синтетичні дані, до яких додаються аномальні відхилення для моделювання потенційних проблем у IT-інфраструктурі. Дані масштабуються й перетворюються в послідовності для подачі на вхід трансформерної моделі. Наступний етап – навчання трансформера з використанням крос-валідації [1]. Дані проходять п’ятикратну крос-валідацію, де модель навчається і перевіряється на валідаційних наборах. Параметри моделі, зокрема кількість шарів, розмір вектора ознак і кількість голів уваги, оптимізовані для точного виявлення аномалій у послідовностях. Упровадження крос-валідації дає змогу визначити найкращі параметри для запобігання перенавчанню. Після цього починається прогнозування та оцінювання якості моделі. Після навчання модель застосовується до тестових даних для отримання прогнозів використання CPU. Порівнюючи реальні значення з прогнозами, система ідентифікує відхилення, які вказують на можливі аномалії. Якість моделі визначається за допомогою метрик MAE та RMSE, що дає змогу оцінити похибку моделі [2]. Наприкінці візуалізуються результати й аналізуються аномалії. Результати подаються у вигляді графіків, що відбивають реальні значення та прогнозовані дані, а також гістограму розподілу залишків, що дає змогу оцінити точність роботи моделі та її здатність фіксувати аномалії (рис. 2).

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 1 (31)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

Рис. 2. Узагальнена схема принципу роботи системи

Матеріали й методи

Для створення навчального й тестового наборів використовуються синтетичні дані, що моделюють поведінку ключових параметрів IT-інфраструктури, таких як:

– використання CPU (*cpu_usage*);
– використання пам'яті (*memory_usage*);
– мережевий трафік (*network_traffic*);
– температура сервера (*temperature*);
– енергоспоживання (*energy_consumption*).

Кожен із цих параметрів генерується з використанням періодичних функцій (синусоїд і косинусоїд) з додаванням нормального шуму, що дає змогу створити реалістичний часовий ряд. Наприклад, застосування CPU моделюється так: *cpu_usage=np.sin(np.linspace(0,50,num_samples))*20+ 50+np.random.randn(num_samples)*5.*

Інші параметри моделюються аналогічним способом.

Аномальні значення додаються до деяких відліків у кожному параметрі для тестування здатності моделі виявляти аномалії. Зміна значень в окремих точках призводить до значних відхилень, що допомагає оцінити здатність моделі фіксувати аномалії. Щоб забезпечити стабільне навчання моделі, дані кожного параметра нормалізуються до стандартного розподілу із середнім значенням 0 та стандартним відхиленням 1 [3]. Цей процес описується такою формулою:

$$x_i = \frac{x_i - \mu_x}{\sigma_x}\ , \qquad (1)$$

де x_i – значення показника; μ_x – середнє значення; σ_x – стандартне відхилення.

Оскільки модель працює з часовими рядами, дані організовуються у послідовності фіксованої довжини L (наприклад, 30 відліків) [4]. Цей процес створює фрагменти даних, де кожен фрагмент має L відліків, що дає змогу моделі аналізувати часову динаміку змін. Формалізовано це виглядає так:

$$X = \left\{ \left(x_i, x_{i+1}, \ldots, x_{i+L-1} \right), y_{i+L} \right\}\ , \qquad (2)$$

де X – вхідна послідовність; y – цільова змінна для прогнозування; i – початкова точка фрагменту.

Архітектура трансформера для аналізу аномалій

Модель для аналізу часових рядів побудована на основі трансформера, що містить кілька шарів з увагою, *feed-forward*-шари і позиційне кодування. Трансформери, зокрема *Timesformer* та *Informer*, демонструють високу ефективність у задачах аналізу багатовимірних часових рядів завдяки здатності моделювати довгострокові залежності [5].

Трансформер не використовує рекурентні зв'язки, тому потребує позиційного кодування для врахування порядку елементів у послідовності [6]. Позиційне кодування додається до вхідного вектора *x_i* у вигляді додаткових ознак і розраховується за формулами:

$$PE\left(pos, 2i\right) = \sin\left(\frac{pos}{10000^{2i/d}}\right) \qquad (3)$$

$$PE\left(pos, 2i+1\right) = \cos\left(\frac{pos}{10000^{2i/d}}\right), \qquad (4)$$

де pos – позиція в послідовності; i – індекс ознаки; d – розмірність вхідного вектора. Позиційне

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)
Innovative technologies and scientific solutions for industries. 2025. No. 1 (31)

кодування додається до вхідного вектора x_i, забезпечуючи необхідну інформацію про порядок елементів у послідовності. Механізм *multi-head attention* є ключовим компонентом трансформера, що дає змогу моделювати залежності між елементами вхідної послідовності [7]. Він обчислює кілька паралельних *head* уваги, кожна з яких розраховується за формулою:

$$Attention(Q,K,V) = softmax\left(QK^T/\sqrt{d_k}\right)V, \quad (5)$$

де Q,K,V – матриці запитів, ключів та значень; d_k – розмірність простору ключів. Механізм *multi-head attention* сприяє одночасному зосередженню на різних аспектах взаємозв'язків між елементами. Вихід обчислюється як конкатенація всіх голів (*head*), після чого застосовується лінійна трансформація:

$$MultiHead(Q,K,V) = Concat(head_1,\ldots,head_h)W^O, \quad (6)$$

де h – кількість голів; W^O – матриця ваг.

Кожен блок трансформера також містить два *feed-forward*-шари з нелінійністю ReLU [8]. Вхідні дані проходять крізь два лінійні шари, що дає змогу моделі навчати більш складні залежності:

$$FFN(x) = max(0, xW_1 + b_1)W_2 + b_2, \quad (7)$$

де W_1 і W_2 – матриці ваг; b_1 і b_2 – зміщення [9]. Функція втрат, яка мінімізується під час навчання, – це середньоквадратична похибка (*Mean Squared Error*, MSE), що обчислюється як

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - y_i\right)^2, \quad (8)$$

де y_i – справжнє значення; y_i – прогнозоване значення. MSE підходить для задачі прогнозування та забезпечує чутливість до значних помилок, що корисно у виявленні аномалій.

Навчання моделі та гіперпараметри

Для оцінювання узагальненості моделі застосовується п'ятикратна крос-валідація, що дає змогу визначити стабільність і точність моделі на різних підмножинах даних. Крос-валідація розділяє дані на п'ять непересічних наборів, де кожен з них використовується один раз як валідаційний набір, а чотири інші слугують для навчання [10].

Крос-валідація забезпечує більш стійкі результати, оскільки дає змогу моделі зважати на різні аспекти даних під час навчання. Кожен фолд обчислює середні значення втрат для навчального й валідаційного наборів, що допомагає оцінити рівень узагальненості моделі. Процес крос-валідації зображено на рис. 3.

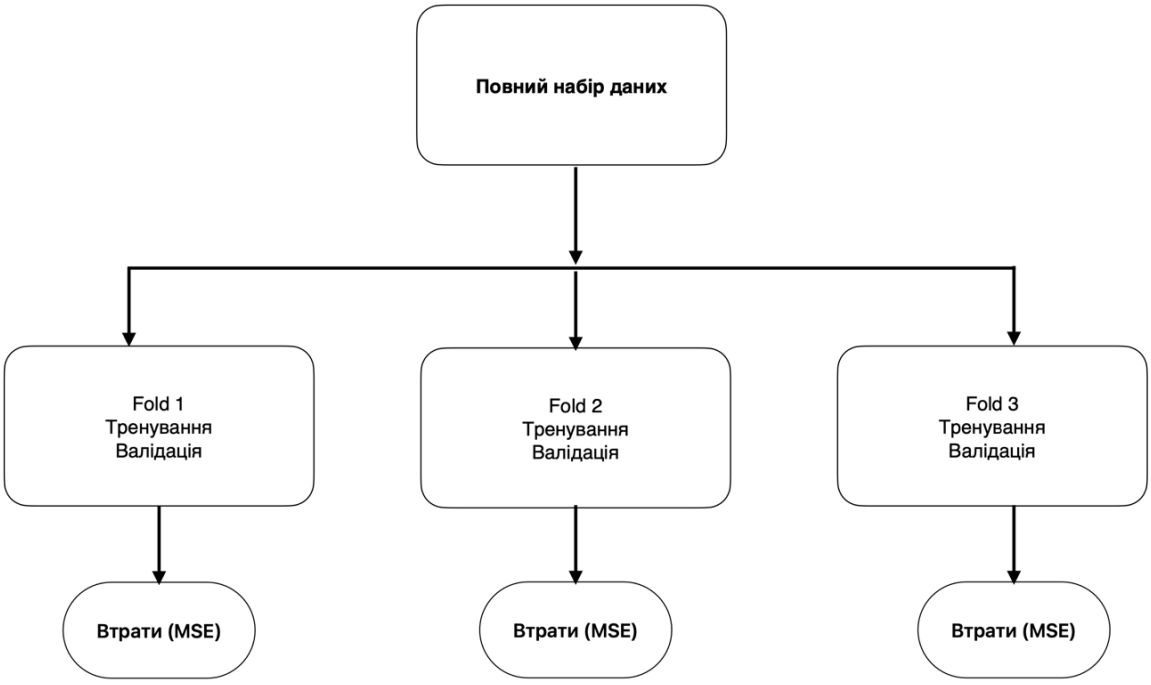

Рис. 3. Процес крос-валідації

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 1 (31)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

Щоб уникнути перенавчання, застосовують метод ранньої зупинки, який завершує навчання, якщо функція втрат на валідаційних даних не покращується протягом визначеного числа епох. У реалізації це значення встановлено на 10. Це означає, що навчання припиняється, якщо впродовж десяти епох модель не демонструє покращення на валідаційному наборі.

Додатково для підвищення стабільності навчання використовують адаптивні регулятори швидкості навчання (*Learning Rate Scheduler*) [11], які зменшують коефіцієнт навчання на 50 % у разі, якщо втрати на валідаційних даних не покращуються протягом трьох епох. Це дає змогу моделі більш повільно оновлювати ваги на пізніх етапах навчання, забезпечуючи кращу збіжність.

Функція втрат, що використовується для оптимізації, – це середньоквадратична похибка (*Mean Squared Error*, MSE). MSE забезпечує чутливість до великих помилок, що є корисним для завдань, де критично важливо виявляти значні відхилення від нормального стану.

Метрики оцінювання точності та якості

Для оцінювання ефективності моделі та її здатності виявляти аномалії використовують кілька метрик, що дають змогу провести кількісний аналіз якості прогнозів. Основними метриками є середня абсолютна похибка (*Mean Absolute Error*, MAE), корінь середньоквадратичної похибки (*Root Mean Squared Error*, RMSE). RMSE є чутливішою до великих похибок, ніж MAE, і забезпечує більше значення для завдань, де великі помилки є критичними.

Система реалізована на основі мови *Python* з використанням сучасних бібліотек для нейронних мереж, машинного навчання та оброблення даних, що допомагає ефективно працювати з часовими рядами, навчати модель, здійснювати крос-валідацію та оцінювати точність прогнозів. Основні бібліотеки та фреймворки:

– *PyTorch* – обраний як основний фреймворк для побудови трансформерної моделі. *PyTorch* надає інструменти для створення складних нейронних мереж і забезпечує високу ефективність у виконанні на GPU, що особливо важливо для навчання глибинних моделей;

– *scikit-learn* – використовується для попереднього оброблення даних і масштабування ознак, що дає змогу підготувати вхідні дані у відповідному форматі. Крім цього, *scikit-learn* забезпечує зручні методи для обчислення метрик точності (MAE, RMSE), необхідних для оцінювання продуктивності моделі;

– *NumPy* – застосовується для роботи з багатовимірними масивами даних і математичними операціями, такими як масштабування та підготовка векторів, що покращує швидкість виконання операцій;

– *Matplotlib* та *Seaborn* – інструменти для створення візуалізацій, що використовуються для аналізу втрат на навчальних і валідаційних даних, а також для візуалізації залишків та інших показників, що допомагають оцінити якість роботи моделі.

Етапи навчання моделі

Генерація синтетичних даних. Для тренування моделі генеруються синтетичні дані про роботу IT-інфраструктури, що містять показники, зокрема використання CPU, пам’ять, мережевий трафік, температура й енергоспоживання [12]. Ці дані генеруються як часові ряди, до яких додається випадковий шум, що дає змогу моделі навчитися зважати на варіативність у реальних системах.

Введення аномалій. До обраних точок даних додаються аномальні відхилення, що моделюють потенційні проблеми в IT-інфраструктурі. Це дає змогу перевірити, як модель справляється з виявленням таких аномалій і чи може вона диференціювати їх від нормальних варіацій.

Масштабування даних. Щоб запобігти впливу різниці в масштабах між параметрами, кожен показник нормалізується до стандартного розподілу з середнім значенням 0 і стандартним відхиленням 1 [13]. Це сприяє більш стабільному навчальному процесу та запобігає домінуванню будь-якої ознаки у вхідних даних.

Формування послідовностей. Для забезпечення моделі контекстом послідовності дані розбиваються на фрагменти фіксованої довжини. Кожен фрагмент – це часовий проміжок, що дає змогу моделі брати до уваги часову динаміку параметрів. Цільова змінна для кожної послідовності є значенням CPU у наступний момент часу, що моделює завдання прогнозування.

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 1 (31)

Позиційне кодування. Оскільки трансформери не використовують рекурентних зв'язків, позиційне кодування дає змогу зважати на порядок елементів у послідовності [14]. Для цього застосовуються синусоїдальні функції, які додаються до кожного вектора вхідних даних, надаючи інформацію про позицію елементів у послідовності.

Механізм багатоголової уваги. Модель використовує кілька голів уваги, що аналізують різні аспекти взаємозв'язків між елементами вхідної послідовності. Багатоголова увага допомагає моделі більш ефективно захоплювати як короткострокові, так і довгострокові залежності, що є критичним для виявлення аномалій у багатофакторних часових рядах.

Feed-forward-шари. Кожен блок трансформера містить *feed-forward*-шари, які обробляють вбудовані подання після уваги, застосовуючи нелінійну функцію активації [15]. Це дає змогу моделі навчатися більш складних залежностей між ознаками та завдяки цьому підвищувати точність прогнозування.

Для мінімізації функції втрат використовують алгоритм оптимізації *AdamW*, який бере до уваги адаптивну швидкість навчання з ваговим спадом (*weight decay*), що запобігає перенавчанню. Функцією втрат обрано середньоквадратичну похибку (*Mean Squared Error*, MSE), яка підходить для завдань прогнозування, забезпечуючи чутливість до значних помилок.

Для забезпечення стабільної продуктивності моделі застосовується п'ятикратна крос-валідація. Кожен фолд розбиває дані на навчальний та валідаційний набори, де модель навчається на більшій частині даних, а перевіряється на валідаційному наборі. Це дає змогу уникнути перенавчання та оцінити, наскільки модель здатна узагальнювати залежності на різних підмножинах даних.

Щоб уникнути перенавчання, в процесі навчання використовується рання зупинка: якщо валідаційні втрати не знижуються протягом кількох епох поспіль, навчання припиняється. Крім того, застосовується динамічне зменшення швидкості навчання, яке активується, якщо втрати на валідаційних даних не покращуються впродовж визначеного числа епох, що сприяє більш точному налаштуванню ваги на пізніх етапах навчання.

Результати досліджень та їх обговорення

Результати експериментів на синтетичних даних показують здатність моделі виявляти аномалії та прогнозувати використання CPU в IT-інфраструктурі. Модель пройшла навчання за допомогою п'ятикратної крос-валідації, що дало змогу оцінити її стабільність і узагальненість. Нижче наведено детальний аналіз досягнутих результатів. Основними метриками оцінювання моделі є *Mean Absolute Error* (MAE) та *Root Mean Squared Error* (RMSE), що вказують на середню абсолютну похибку та середньоквадратичне відхилення відповідно.

MAE для моделі становить 4.3001, що означає середню похибку прогнозу в межах 4.3%.

RMSE становить 5.5387, що свідчить про те, що модель має більшу чутливість до великих помилок, хоча загальна точність залишається на прийнятному рівні.

Такі результати вказують на здатність моделі фіксувати тенденції у використанні CPU, але також на присутність незначних відхилень, що можуть бути спричинені як особливостями генерації синтетичних даних, так і потенційними обмеженнями архітектури моделі.

Графік крос-валідації, на якому показано динаміку втрат (*Train Loss* і *Validation Loss*) для кожного з п'яти фолдів, свідчить про стабільну роботу моделі на навчальних і валідаційних даних (рис. 4).

Різниця між *Train Loss* і *Val Loss* для кожного фолду вказує на відсутність суттєвого перенавчання, а також на стабільність збіжності для кожного набору даних. Зокрема після 40–60 епох модель досягає стабільних значень втрат як на навчальних, так і на валідаційних даних. Діаграма динаміки втрат для різних фолдів підтверджує, що модель досягає збіжності внаслідок регуляризації (*dropout*) та механізму ранньої зупинки, що забезпечує стабільність навчання та запобігає перенавчанню.

Теплова карта кореляцій між параметрами демонструє взаємозв'язки між основними показниками IT-інфраструктури. Найбільш значуща кореляція спостерігається між використанням CPU та енергоспоживанням (коефіцієнт кореляції – 0.75), що вказує на взаємозалежність між навантаженням на процесор та рівнем споживання енергії. Інші параметри мають нижчий рівень кореляції, що вказує на їх незалежність. Це є важливим для побудови багатофакторної моделі (рис. 5).

Теплова карта кореляцій дає змогу також оцінити, які параметри необхідно брати до уваги в аналізі аномалій, оскільки сильні кореляції можуть впливати на точність прогнозів.

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 1 (31)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

Рис. 4. Графік крос-валідації

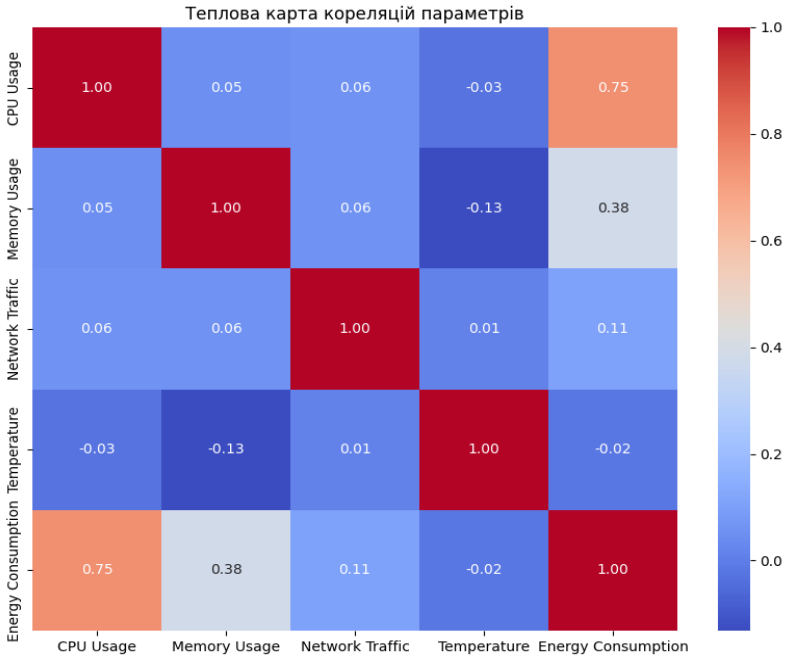

Рис. 5. Теплова карта кореляції між параметрами

Гістограма розподілу залишків показує симетричний розподіл із середнім, близьким до нуля, що є позитивною ознакою для моделі, оскільки свідчить про відсутність систематичних похибок у прогнозуванні (рис. 6). Більшість залишків зосереджені навколо нуля, що вказує на високу точність моделі в процесі прогнозування звичайних значень.

Присутність незначної кількості великих залишків вказує на відносно рідкісні великі відхилення, що можуть бути спричинені аномальними подіями або значними коливаннями в синтетичних даних. Це наголошує на важливості налаштування моделі для виявлення таких великих залишків, як потенційних аномалій.

Графік прогнозування аномалій демонструє, що модель успішно слідкує за трендом реальних значень і точно відтворює нормальні коливання у використанні CPU (рис. 7). Червоні точки на графіку позначають контрольовані аномалії, що були додані до даних під час генерації. Ці точки збігаються з підвищеними значеннями в прогнозах моделі, що вказує на її здатність фіксувати різкі зміни та відхилення.

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)
Innovative technologies and scientific solutions for industries. 2025. No. 1 (31)

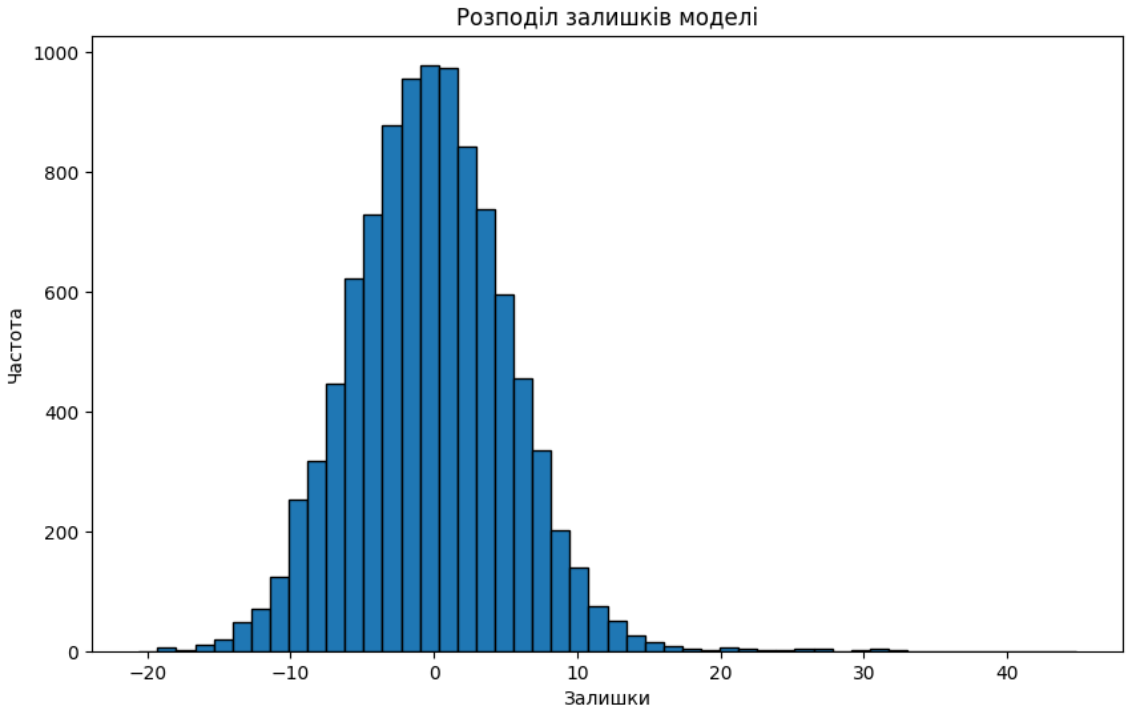

Рис. 6. Гістограма розподілу залишків

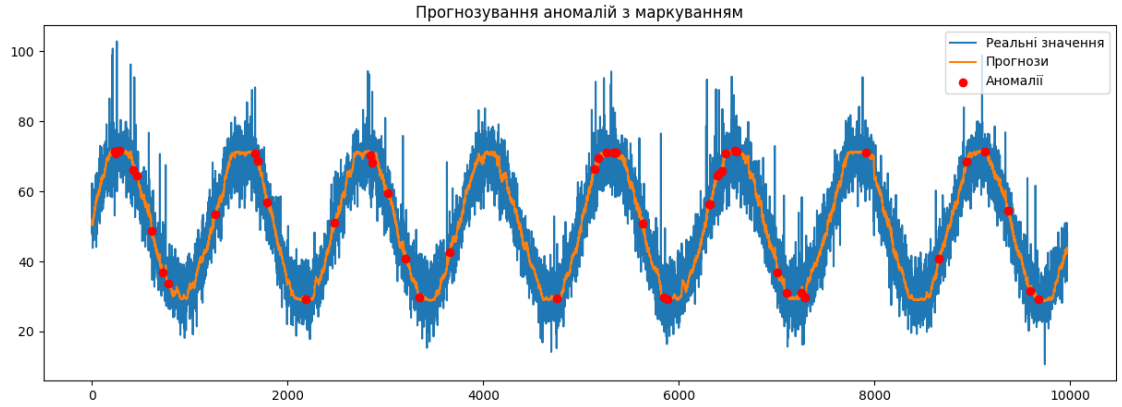

Рис. 7. Графік прогнозування аномалій

Аналіз графіка прогнозування вказує на високу чутливість моделі до різких змін у даних, що є критично важливим для вчасного виявлення аномалій в умовах реального часу. Це підтверджує можливість використання моделі для автоматичного виявлення відхилень у критичних параметрах IT-інфраструктури.

Висновки

Досягнуті результати експериментів підтверджують високу здатність моделі до аналізу багатофакторних часових рядів і виявлення аномалій. Крос-валідація демонструє стабільність навчання без значного перенавчання, тоді як розподіл залишків свідчить про відсутність систематичних похибок у прогнозах. Графіки кореляцій та прогнозів підтверджують, що модель успішно фіксує ключові взаємозв'язки між параметрами та вчасно реагує на значні відхилення.

Основні результати доводять, що розроблена система є ефективною для завдань моніторингу та прогнозування в IT-інфраструктурі та може бути використана як основа для подальшого розширення функціональності, наприклад, для виявлення більш комплексних типів аномалій та інтеграції із системами автономного налаштування.

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 1 (31)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

Список літератури

1. Mykhailichenko I., Ivashchenko H., Barkovska O., Liashenko O., Application of Deep Neural Network for Real-Time Voice Command Recognition, *2022 IEEE 3rd KhPI Week on Advanced Technology (KhPIWeek)*, Kharkiv, Ukraine, 2022, P. 1–4, DOI: 10.1109/KhPIWeek57572.2022.9916473
2. Barkovska O., Pyvovarova D., Kholiiv V., Ivashchenko H., Rosinskyi D. Model zberezhennia informatsiinykh obiektiv iz pryskorenymy metodamy obrobky tekstiv, *CEUR Workshop Proceedings.* 2021. № 2870. P. 286–299. DOI: 10.20944/preprints202412.2147.v1
3. Hunko M., Tkachov V., Liashenko O., Rabchan Y. Arkhitektura zastosunku dlia otrymannia danykh iz naukometrychnykh baz danykh // *Materialy konferentsii IEEE 3rd KhPI Week on Advanced Technology, KhPI Week 2022*. 2022. DOI: 10.1109/KhPIWeek57572.2022.9916473
4. Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A. N., Kaiser Ł., Polosukhin I. Attention is All You Need. *Advances in Neural Information Processing Systems*. Vol. 30. 2017. P. 5998–6008. DOI: 10.48550/arXiv.1706.03762
5. Wu H., Xu J., Wang L., Chen K. Long-Term Time-Series Forecasting with Triangular Matrix-Based Attention. *Proceedings of the 9th International Conference on Learning Representations (ICLR)*. 2021. 15 p. DOI: 10.48550/arXiv.2106.13008
6. Машинне навчання у виявленні аномалій / Л. Г. Кравченко та ін. *Системний аналіз і прикладна інформатика.* 2018. № 5. С. 90–102.
7. Informer: новий підхід до прогнозування / В. О. Мельник та ін. *Журнал штучного інтелекту*. 2019. № 4. С. 12–25.
8. Василенко В. О., Петров І. М. Методи виявлення аномалій у багатовимірних часових рядах. Київ, 2019. 256 с.
9. Lin T., Guo T., Wang K., Xu J. A Survey on Transformer Architectures in Time Series Applications. *arXiv*, 2021. 29 p. DOI: 10.48550/arXiv.2106.13008
10. Li S., Jin X., Xuan Y., Zhou X., Chen W., Wang Y.-X., Yan X. Enhancing the Locality and Breaking the Memory Bottleneck of Transformer on Time Series Forecasting. *Advances in Neural Information Processing Systems.* Vol. 32. 2019. P. 5243–5253. DOI: 10.48550/arXiv.1907.00235
11. Aggarwal C. C. Outlier Analysis. 2nd ed. Cham: *Springer*, 2017. 466 p. DOI: 10.1007/978-3-319-47578-3
12. Іванченко М. С., Гриценко О. В. Застосування трансформерів у аналізі часових рядів. Львів, 2021. 220 с.
13. Goodfellow I., Bengio Y., Courville A. Deep Learning. *Cambridge*, 2016. 775 p. DOI: 10.5555/3086952
14. Zhou T. et al. Fedformer: Frequency Enhanced Decomposed Transformer for Long-Term Series Forecasting. *International Conference on Machine Learning.* 2022. P. 27268–27286. URL: https://proceedings.mlr.press/v162/zhou22g.html
15. Zhou H. et al. Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting. P*roceedings of the AAAI Conference on Artificial Intelligence.* Vol. 35. No. 12. 2021. P. 11106–11115. DOI: https://doi.org/10.1609/aaai.v35i12.17325

References

1. Mykhailichenko, I., Ivashchenko, H., Barkovska, O., Liashenko, O., (20220, "Application of Deep Neural Network for Real-Time Voice Command Recognition ", *2022 IEEE 3rd KhPI Week on Advanced Technology (KhPIWeek)*, Kharkiv, Ukraine, 2022, P. 1-4, DOI: 10.1109/KhPIWeek57572.2022.9916473
2. Barkovska, O., Pyvovarova, D., Kholiiv, V., Ivashchenko, H., Rosinskyi, D. (2021), "Model for Preserving Information Objects with Accelerated Text Processing Methods" ["Model zberezhennia informatsiinykh obiektiv iz pryskorenymy metodamy obrobky tekstiv"], *CEUR Workshop Proceedings*, No. 2870, P. 286–299. DOI: 10.20944/preprints202412.2147.v1
3. Hunko, M., Tkachov, V., Liashenko, O., Rabchan, Y. (2022), "Application Architecture for Retrieving Data from Scientometric Databases" ["Arkhitektura zastosunku dlia otrymannia danykh iz naukometrychnykh baz danykh"], *IEEE 3rd KhPI Week on Advanced Technology, KhPI Week 2022*. DOI: 10.1109/KhPIWeek57572.2022.9916473
4. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., Polosukhin, I. (2017), "Attention is All You Need", *Advances in Neural Information Processing Systems*, Vol. 30, P. 5998–6008. DOI: 10.48550/arXiv.1706.03762
5. Wu, H., Xu, J., Wang, L., Chen, K. (2021), "Long-Term Time-Series Forecasting with Triangular Matrix-Based Attention", *Proceedings of the 9th International Conference on Learning Representations (ICLR)*, 15 p. DOI: 10.48550/arXiv.2106.13008
6. Kravchenko, L. H., et al. (2018), "Machine Learning in Anomaly Detection" [“Mashynne navchannia u vyiavlenni anomalii"], *System Analysis and Applied Informatics*, No. 5, P. 90–102.
7. Melnyk, V. O., et al. (2019), "Informer: A New Approach to Forecasting" [“Informer: novyi pidkhid do prohnozuvannia"], *Journal of Artificial Intelligence*, No. 4, P. 12–25.
8. Vasylenko, V. O., Petrov, I. M. (2019), "Methods for Anomaly Detection in Multidimensional Time Series" ["Metody vyiavlennia anomalii u bahatovymirnykh chasovykh riadakh"], Kyiv, 256 p.
9. Lin, T., Guo, T., Wang, K., Xu, J. (2021), "A Survey on Transformer Architectures in Time Series Applications", *arXiv*, 29 p. DOI: 10.48550/arXiv.2106.13008
10. Li, S., Jin, X., Xuan, Y., Zhou, X., Chen, W., Wang, Y.-X., Yan, X. (2019), "Enhancing the Locality and Breaking the Memory Bottleneck of Transformer on Time Series Forecasting", *Advances in Neural Information Processing Systems*, Vol. 32, P. 5243–5253. DOI: 10.48550/arXiv.1907.00235

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) *Innovative technologies and scientific solutions for industries. 2025. No. 1 (31)*

11. Aggarwal, C. C. (2017), "Outlier Analysis", 2nd ed., *Springer*, Cham, 466 p. DOI: 10.1007/978-3-319-47578-3
12. Ivanchenko, M. S., Hrytsenko, O. V. (2021), "Application of Transformers in Time Series Analysis" ["Zastosuvannia transformeriv u analizi chasovykh riadiv"], Lviv, 220 p.
13. Goodfellow, I., Bengio, Y., Courville, A. (2016), "Deep Learning", Cambridge, 775 p. DOI: 10.5555/3086952
14. Zhou, T., et al. (2022), "Fedformer: Frequency Enhanced Decomposed Transformer for Long-Term Series Forecasting", *International Conference on Machine Learning*, P. 27268–27286. available at: https://proceedings.mlr.press/v162/zhou22g.html
15. Zhou, H., et al. (2021), "Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting", *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 35, No. 12, P. 11106–11115. DOI: https://doi.org/10.1609/aaai.v35i12.17325

Надійшла (Received) 01.03.2025

Відомості про авторів / About the Authors

Ляшенко Олексій Сергійович – кандидат технічних наук, доцент, Харківський національний університет радіоелектроніки, декан факультету комп'ютерної інженерії та управління, Харків; Україна;e-mail: oleksii.liashenko@nure.ua; ORCID ID: https://orcid.org/0000-0002-0146-3934

Михайліченко Ігор Володимирович – Харківський національний університет радіоелектроніки, асистент кафедри електронних обчислювальних машин, Харків; Україна, e-mail: ihor.mykhailichenko@nure.ua; ORCID ID: https://orcid.org/0009-0005-5476-8992

Liashenko Oleksii – PhD (Computer Engineering Science), Associate Professor, Kharkiv National University of Radio Electronics, Dean at the Faculty of Computer Engineering and Management, Kharkiv, Ukraine.

Mykhailichenko Ihor – Kharkiv National University of Radio Electronics, Assistant at the Department of Electronic Computing Machines, Kharkiv, Ukraine.

AUTONOMOUS MONITORING AND OPTIMIZATION SYSTEM FOR IT INFRASTRUCTURE USING TRANSFORMERS

The **subject matter** of the article is an autonomous IT infrastructure monitoring and optimization system using transformers for analyzing multidimensional time series and detecting anomalies in real time. The study reviews current approaches to IT infrastructure monitoring, including machine learning and traditional statistical methods. A literature review reveals that existing methods often lack efficiency in dynamically changing system parameters. The **goal** of this research is to develop an autonomous system capable of performing real-time multifactor analysis and autonomously responding to detected threats. A transformer-based model is proposed, allowing for complex anomaly detection and failure prediction. The methodology includes mathematical modeling, machine learning (transformers), statistical analysis (cross-validation), and time series forecasting. The following **tasks** were solved in the article: formulation of a model for multidimensional time series analysis, development of an algorithm for anomaly detection and problem forecasting, implementation of autonomous adjustment mechanisms for stabilizing IT infrastructure. The following **methods** used are – mathematical modeling, machine learning methods (transformers), statistical analysis (cross-validation), and forecasting algorithms based on time series. The following **results** the model achieved a mean absolute error (MAE) of 4.3% on synthetic data, confirming its ability to accurately detect anomalies. Cross-validation validated the stability of training without overfitting, while a residual histogram showed a symmetrical error distribution. Additionally, correlation heatmaps highlighted interdependencies between key IT infrastructure parameters. **Conclusions:** the proposed system effectively detects and predicts IT infrastructure failures, ensuring autonomous parameter adjustment to maintain stability. The developed approach can be integrated into modern IT infrastructure management systems to enhance operational efficiency.

Keywords: autonomous system; transformers; multidimensional time series; IT infrastructure; anomaly detection; forecasting.

Бібліографічні описи / Bibliographic descriptions

Ляшенко О. С., Михайліченко І. В. Модель автономної системи моніторингу та оптимізації IT-інфраструктури з використанням трансформерів. *Сучасний стан наукових досліджень та технологій в промисловості*. 2025. № 1 (31). С. 73–82. DOI: https://doi.org/10.30837/2522-9818.2025.1.073

Liashenko, O., Mykhailichenko, I. (2025), "Autonomous monitoring and optimization system for IT infrastructure using transformers", *Innovative Technologies and Scientific Solutions for Industries*, No. 1 (31), P. 73–82. DOI: https://doi.org/10.30837/2522-9818.2025.1.073